Gravitational effective field theory islands, low-spin dominance, and the four-graviton amplitude

Zvi Bern[1], Dimitrios Kosmopoulos[1] and Alexander Zhiboedov[2,*]

[1] Mani L Bhaumik Institute for Theoretical Physics, UCLA Department of Physics and Astronomy, Los Angeles, CA 90095, United States of America
[2] CERN, Theoretical Physics Department, CH-1211 Geneva 23, Switzerland

E-mail: alexander.zhiboedov@cern.ch

Abstract

We analyze constraints from perturbative unitarity and crossing on the leading contributions of higher-dimension operators to the four-graviton amplitude in four spacetime dimensions, including constraints that follow from distinct helicity configurations. We focus on the leading-order effect due to exchange by massive degrees of freedom which makes the amplitudes of interest infrared finite. In particular, we place a bound on the coefficient of the R^3 operator that corrects the graviton three-point amplitude in terms of the R^4 coefficient. To test the constraints we obtain nontrivial effective field-theory data by computing and taking the large-mass expansion of the one-loop minimally-coupled four-graviton amplitude with massive particles up to spin 2 circulating in the loop. Remarkably, we observe that the leading EFT coefficients obtained from both string and one-loop field-theory amplitudes lie in small islands. The shape and location of the islands can be derived from the dispersive representation for the Wilson coefficients using crossing and assuming that the lowest-spin spectral densities are the largest. Our analysis suggests that the Wilson coefficients of weakly-coupled gravitational physical theories are much more constrained than indicated by bounds arising from dispersive considerations of $2 \to 2$ scattering. The one-loop four-graviton amplitudes used to obtain the EFT data are computed using modern amplitude methods, including generalized unitarity, supersymmetric decompositions and the double copy.

---

*Author to whom any correspondence should be addressed.

Keywords: EFT bounds, scattering amplitudes, gravity

(Some figures may appear in colour only in the online journal)

1. Introduction

Remarkably, systematic bounds can be placed on possible corrections to Einstein gravity [1–7]. Such corrections naturally appear due to the presence of heavy particles in the theory. To leading order in Newton's constant G, such particles can be exchanged at tree-level, as in string theory, or at one-loop, as in the case of matter minimally coupled to gravity. By expanding such amplitudes at low energies and matching to a low-energy effective field theory one finds an infinite series of higher-derivative corrections to Einstein gravity. The coefficients in front of these higher-derivative operators, or Wilson coefficients, satisfy various bounds due to unitarity and causality of the underlying amplitude [1, 2]. In this paper, we focus on the leading corrections to Einstein gravity.[3] A central question, which we investigate in this paper, is to understand if there are principles that can greatly restrict the values of physically allowed Wilson coefficients.

Consistency bounds on the Wilson coefficients received a lot of attention recently in the context of $2 \rightarrow 2$ scattering, which is also the subject of our paper. The basic tool to derive such bounds is given by dispersion relations which express low-energy Wilson coefficients as weighted sums of the discontinuity of the amplitude. Unitarity constrains the form of the discontinuity of the amplitude which can be further used to derive the bounds. The simplest examples of this type constrain the sign of Wilson coefficients. More interesting bounds arise when one accommodates constraints coming from crossing symmetry. Including those leads to the two-sided bounds on the Wilson coefficients [3–6]. In this way the ultraviolet (UV) complete theories form bounded regions in the space of couplings. Reference [5] also analyzed a few examples of physical EFTs in the context of scattering of scalars and noted that they lie near the boundaries of the allowed region due to the importance of low-spin contributions to the partial-wave expansions (see section 10.3 and appendix D of reference [5]).

In the context of physical theories, especially gravitational ones, it is then natural to ask the following question:

Is it possible that the Wilson coefficients of physical theories live in much smaller regions than the bounds coming from considerations of $2 \rightarrow 2$ *scattering suggest?*

By physical theories in this paper we mean perturbatively consistent *S*-matrices that satisfy unitarity, causality, and crossing for any $n \rightarrow m$ scattering processes. Constructing such *S*-matrices is far beyond the scope of bootstrap methods that focus on $2 \rightarrow 2$ scattering, but such examples are provided to us by string theory and matter minimally coupled to gravity.[4] We can then imagine that consistency of the full *S*-matrix is reflected back on the $2 \rightarrow 2$ scattering through more stringent constraints on Wilson coefficients that one would naively have found by analyzing $2 \rightarrow 2$ scattering. In this paper we present data extracted from field-theory and string-theory $2 \rightarrow 2$ scattering amplitudes that suggest that the above assertion is indeed true and we identify a principle behind it. This principle is *low-spin dominance* (LSD), which, if fundamentally correct, might be traced back to the consistency of the full gravitational *S*-matrix, beyond $2 \rightarrow 2$ scattering. However, demonstrating this is beyond the scope of the present paper.

---

[3] In particular, higher-loop effects do not affect the discussion in this paper since by assumption gravity is weakly coupled and we are focusing on the leading-order effect.

[4] Perturbatively consistent *S*-matrices occupy a somewhat intermediate position between fully non-perturbatively consistent quantum gravities (often referred to as landscape) and consistency of $2 \rightarrow 2$ scattering studied by bootstrap methods.

The universal nature of gravity together with the strict consistency requirements that graviton scattering obeys make such an assertion plausible. It is a well-known fact that scattering of massless spinning particles is very constrained [8]. In fact, massless particles of spin larger than two do not admit a non-trivial *S*-matrix [9]. Gravitons, being massless spin-two particles, are thus expected to have an especially constrained *S*-matrix, and the assertion made in the previous paragraph is thus particularly plausible for graviton scattering which is the subject of the present paper.

Firstly, we use the techniques of reference [5] to derive bounds between low-energy couplings of the same dimensionality in gravitational scattering.[5] We focus on the first few corrections to Einstein gravity. We then ask where do the Wilson coefficients obtained from string theory and from the low-energy limit of the one-loop minimally-coupled amplitudes land in the space allowed by the general bounds. Remarkably, in all cases studied here we find that both the string and field-theory coefficients land on a small *theory island*, which to a good approximation is a thin line segment in the space of EFT coefficients. (See, for example, figures 9 and 12 in section 5).[6] The location of this island can be found by assuming that lowest-spin partial waves dominate the dispersive representation of the low-energy couplings, which is the LSD principle. See section 5.2 for the precise mathematical formulation. More generally, we show how one can combine an assumed hierarchy among the spectral densities of various spins with crossing symmetry to systematically derive stronger bounds on the Wilson coefficients. We impose crossing symmetry via the use of null constraints [3, 4].

The idea that LSD is a true property of physical theories can be traced back to causality, or the statement that the amplitude cannot grow too fast in the Regge limit. Otherwise we could have simply added a tree-level exchange by a large-spin particle which would contribute to a given spin partial wave. Due to causality we cannot do this (see e.g. reference [11]). The situation is particularly dramatic in gravity. In this case the only particle that can be exchanged at tree-level in graviton scattering without violating causality is the graviton itself. Moreover, its self-coupling has to be the one of Einstein gravity references [11–14]. Alternatively, particles of all spins have to be exchanged at tree level to preserve causality, which is the mechanism realized in string theory. It is important to emphasize that at the level of $2 \to 2$ scattering LSD does not follow from causality and we do not prove it in this paper, rather we use it as a principle to organize the known data, and suggest that it may hold more generally. It would be interesting to understand if it follows from considerations on the consistency of $n \to m$ graviton scattering.

Alternatively, it is also possible that our finding of LSD could be special for the models considered here and bears little significance for more general gravitational models. This possibility, which we cannot exclude, would be still very interesting. Indeed, as we demonstrate, any such violation is an indication for non-stringy, non-weakly-coupled-matter physics. For example, it would be very interesting to see if one can somehow violate LSD by making the matter sector strongly coupled, e.g. by considering large-*N* QCD coupled to gravity [15].

Curiously, the phenomenon of LSD generates hierarchies between different Wilson coefficients in the absence of any symmetry. We call this phenomenon *hierarchy from unitarity* and it is something that could have puzzled an unassuming low-energy physicist. We find that specific combinations of Wilson coefficients whose dispersive representation does not involve

---

[5] It would be very interesting to generalize our analysis to include bounds that relate couplings of different dimensionality along the lines of references [3, 4, 7].

[6] In reference [10] the string-theory island was interpreted in terms of unitarity constraints coupled with world-sheet monodromy constraints.

the lowest-spin partial waves can be much smaller than their counterparts that do have them in their definitions.

Secondly, we apply the dispersive sum rules [4, 16] to amplitudes with various helicity configurations of the external gravitons.[7] We derive various bounds on the inelastic scattering (the one in which the final and initial state gravitons have different helicities) in terms of the elastic one (see e.g. reference [17]). We also place a precise bound on the R^3 coefficient in terms of the R^4 coefficient (see (6.13)). Such a bound translates the problem of making the analysis of reference [11] quantitatively precise to the problem of the bounding the leading R^4 contact coefficient in terms of the gap of the theory. This has been recently done in reference [7] in a similar perturbative setting for $D = 10$ maximal supergravity; see also reference [18] for the nonperturbative analysis of the same problem. It would, of course, be very interesting to generalize these studies to more general cases of graviton scattering.

In order to provide data for checking and understanding the derived constraints, we first compute the one-loop four-graviton scattering amplitude with the gravitons minimally coupled to massive matter up to spin 2. Amplitudes corresponding to the ones discussed here, but with massless particles circulating in the loop were obtained a while ago in reference [19] and corresponding gauge-theory amplitudes with massive particles in the loop were computed in reference [20]. We use the same type of organization of the amplitude in terms of supersymmetric multiplets as applied in the earlier calculations, since they naturally group contributions according to their analytic properties.

To evaluate the amplitudes, we make use of standard tools including the unitarity method [21] and the Bern–Carrasco–Johansson (BCJ) [22, 23] double copy, which gives gravity integrands in terms of corresponding gauge-theory ones. We build on the D-dimensional version of the unitarity method of reference [20] in order to fix the rational terms in the amplitudes. At four points gauge-theory tree-level amplitudes automatically satisfy the duality between color and kinematics, so the associated double-copy relations also hold automatically on the unitarity cuts. We use this to express the cuts of the gravity loop integrands directly in terms of the corresponding gauge-theory ones. By using the double copy our computation parallels the corresponding gauge-theory one [20] allowing us to import many of the same steps into the gravitational amplitude calculations.

A complication with massive amplitudes is that there is a class of terms that depend on the mass but do not have branch cuts in any kinematic variable. This makes their construction tricky in the context of the unitarity method. Reference [24] introduced an approach to this problem. Here we instead solve the problem differently by making use of a special property of the scattering amplitudes under study that exploits their simple dependence on the mass of the particle circulating in the loop. In our case (i.e. a single mass circulating in the loop) we instead use knowledge of the UV properties of the amplitudes to fix all remaining functions in the amplitude not determined by unitarity. This procedure is greatly aided by arranging the amplitude in terms of integrals that have no mass or spacetime dependence in their coefficients. To ensure the veracity of our amplitudes we perform a number of nontrivial checks on the mass dependence, and infrared and UV properties. Related to this, we also note a simple relation between UV divergences of appropriate spacetime dimension shifts of the amplitudes and the terms in the large-mass expansion in four dimensions (see equation (2.65)).

We analyze our amplitudes in the large-mass limit and match to a low-energy effective field theory. In this way we systematically obtain corrections to Einstein gravity due to the presence of a heavy spinning particle. These corrections are organized in inverse powers of the particle's

---

[7] Flat space superconvergence considered in reference [14] is a particular example of these more general sum rules.

mass. As already noted not only are our results for the Wilson coefficients fully consistent with the general analysis of bounds on gravitational scattering, but are restricted to small islands.

Since we focus on the leading effect due to heavy particles in the weakly coupled setting neither IR divergences, nor logarithms due to the loops of massless particles make an appearance in our analysis. Taking these into consideration is an important task which we leave for the future.

Our paper naturally consists of two parts: in the first part we explain in detail the construction of the one-loop massive amplitudes used to provide theoretical data that we interpret in the second part in terms of bounds on coefficients of gravitational EFTs. Readers who are interested in the EFT constraints can skip section 2 on the construction of the one-loop amplitudes. Particularly important plots that illustrate the theory islands and the concept of LSD in the partial-wave expansion are given in figures 9–12.

In more detail, the sections are organized as follows: in section 2 we describe our construction of the one-loop four-graviton amplitude with massive matter up to spin 2 in the loop. In section 3 we compute graviton scattering in a general low-energy effective theory. By expanding our amplitudes in the low-energy limit, we extract the Wilson coefficients of the effective field theory. In section 4 we describe the general properties of the gravitational amplitudes stemming from unitarity and causality. In section 5 we derive two-sided bounds on Wilson coefficients that follow from a single helicity configuration that describes elastic scattering; comparing to known data from string theory and our computed one-loop amplitude, we show that the results fall into small islands. We trace the position of these islands using LSD of partial waves. In section 6 we obtain bounds that arise from considering multiple helicities. We bound the low-energy expansion coefficients of inelastic amplitudes in terms of elastic ones. We also derive a bound for the coefficient of the R^3 operator in terms of the R^4 coefficient. Finally, we provide our concluding remarks in section 7. We include various appendices. In appendix A we describe in some detail our definition of minimal coupling of gravity to a massive spinning particle. In appendix B we collect tree-level graviton four-point amplitudes in various string theories. In appendix C we present details on the derivation of some low-energy bounds that are not listed in the main text of the paper. In appendix D we analyze an amplitude function with an accumulation point in the spectrum that partially violates LSD, but show that the corresponding low-energy coefficients still land on the small islands. Appendix E collects the Wigner d-matrices used throughout the paper. In appendix F we present our results for the one-loop amplitudes. We give the expressions for one-loop integrals in terms of which the amplitudes are expressed in appendix G. Finally, in appendix H we expand these results to high orders in the large-mass expansion.

2. Construction of one-loop four-graviton scattering amplitudes

In this section we describe the construction of the one-loop four-graviton amplitudes with massive matter up to spin 2 in the loop. We collect the results in appendix F. We first briefly review the methods used to obtain the amplitudes. Then, following the generalized-unitarity method we build the integrand-level generalized-unitarity cuts. We describe a natural and efficient organization of the unitarity cuts and the amplitudes motivated by supersymmetry. This organization also meshes well with the double-copy construction which we use to obtain gravitational unitarity cuts from gauge-theory ones. Having obtained the unitarity cuts we describe the necessary integral reduction and cut merging into the amplitudes. This process fixes all but

a few pieces of the amplitudes, which we obtain by exploiting the known UV properties of the amplitudes. After calculating the amplitudes, we comment on some interesting UV properties we observe. Finally, we conclude this section by listing the consistency checks we performed on our calculation.

2.1. Basic methods

2.1.1. Spinor helicity.   We use the spinor-helicity method [25] to describe the external graviton states of amplitudes (for reviews see reference [26]). The natural quantities in this formalism are two component Weyl spinors

$$(\lambda_i)_\alpha \equiv [u_+(k_i)]_\alpha, \qquad (\tilde{\lambda}_i)_\alpha \equiv [u_-(k_i)]_\alpha, \tag{2.1}$$

which we write in a 'bra' and 'ket' notation as

$$|k_i^+\rangle = |i\rangle = \lambda_i, \qquad |k_i^-\rangle = |i] = \tilde{\lambda}_i, \qquad \langle k_i^-| = \langle i| = \lambda_i, \qquad \langle k_i^+| = [i| = \tilde{\lambda}_i, \tag{2.2}$$

where k_i^μ refers to the null momentum of the *i*th external particle, while the '$\pm$' superscript refers to the helicity of the corresponding state. The spinor inner products are defined using the antisymmetric tensors $\varepsilon^{\alpha\beta}$ and $\varepsilon^{\dot\alpha\dot\beta}$,

$$\langle k_i^-|k_j^+\rangle = \langle ij\rangle = \varepsilon^{\alpha\beta}(\lambda_i)_\alpha(\lambda_j)_\alpha, \qquad \langle k_i^+|k_j^-\rangle = [ij] = -\varepsilon^{\dot\alpha\dot\beta}(\tilde{\lambda}_i)_{\dot\alpha}(\tilde{\lambda}_i)_{\dot\alpha}. \tag{2.3}$$

These spinor products are antisymmetric in their arguments and we choose a convention where they satisfy $\langle ij\rangle[ij] = 2k_i \cdot k_j$.

In order to construct amplitudes with external gravitons, our starting point is the corresponding ones with external gluons. For calculations involving external gluons the helicity polarization vectors are defined as

$$\varepsilon_\mu^+(k_i; q_i) = \frac{\langle q_i|\gamma_\mu|i]}{\sqrt{2}\langle iq_i\rangle}, \qquad \varepsilon_\mu^-(k_i; q_i) = \frac{[q_i|\gamma_\mu|i\rangle}{\sqrt{2}[iq_i]}, \tag{2.4}$$

where q_i are arbitrary null 'reference momenta' which drop out of the final gauge-invariant amplitudes. Note that we do not use a shorthand notation for the spinors corresponding to the reference momenta. The polarization tensors for gravitons are simply given in terms of products of gluon polarization vectors,

$$\varepsilon_{\mu\nu}^+(k; q) = \varepsilon_\mu^+(k; q)\,\varepsilon_\nu^+(k; q), \qquad \varepsilon_{\mu\nu}^-(k; q) = \varepsilon_\mu^-(k; q)\,\varepsilon_\nu^-(k; q), \tag{2.5}$$

which automatically satisfy the graviton tracelessness condition, due to the Fierz identity. When these polarization vectors are contracted into external momenta k_i^μ or loop momenta ℓ^μ we define,

$$k_1 \cdot \varepsilon_2^+ = \frac{\langle q_2|k_1|2]}{\sqrt{2}\langle q_2 2\rangle} \equiv \frac{\langle q_2|1|2]}{\sqrt{2}\langle q_2 2\rangle}, \qquad \ell \cdot \varepsilon_2^+ = \frac{\langle q_2|\ell|2]}{\sqrt{2}\langle q_2 2\rangle} \equiv \frac{\langle q_2|\ell|2]}{\sqrt{2}\langle q_2 2\rangle}, \text{ etc}, \tag{2.6}$$

where we also use the abbreviation $\varepsilon_2^+ \equiv \varepsilon^+(k_2; q_2)$.

We note that, in general, for loop calculations some care is needed when using dimensional regularization. To take advantage of the spinor-helicity formulation in a one-loop calculation we need to choose an appropriate version of dimensional regularization. Specifically, instead of taking the external polarization tensors and momenta to be $(4 - 2\epsilon)$-dimensional as in conventional dimensional regularization [27], we use the so called four-dimensional helicity (FDH)

scheme [28, 29] where both external and loop state counts are kept in four dimensions and only the loop momentum is continued to $4 - 2\epsilon$ dimensions. Because the massive one-loop amplitudes that we obtain here are neither UV nor infrared divergent, the precise distinction between the different versions of dimensional regularization drops out from the final results for the amplitudes. We do, however, need to regularize intermediate steps because individual loop integrals are UV divergent, with the divergence canceling in final results.

2.1.2. Generalized unitarity.   In order to construct the loop integrands we use the generalized-unitarity method [21]. This method systematically builds complete loop-level integrands using as input on-shell tree-level amplitudes. A central advantage is that simplifications and features of the latter are directly imported into the former. Reviews of the generalized-unitarity methods are found in references [30, 31].

In general, the task of computing an amplitude is to reduce it to a linear combination of known scalar integrals. Using standard integral-reduction techniques (see e.g. references [32, 33]) any four-point one-loop amplitude can be written as a linear combination of box, triangle, bubble and tadpole integrals,

$$\mathcal{M}_4^{1-\text{loop}} = \left(d_{s,t}I_4(s,t) + c_sI_3(s) + b_sI_2(s) + \text{perms.}\right) + b_0I_2(0) + a_0I_1, \quad (2.7)$$

where the permutations run over distinct relabelings of the integrals. At the four-point level there are a total of 11 coefficients. These coefficients depend on polarization vectors, momenta, masses and the dimensional-regularization parameter ϵ. We define the basis integrals appearing in equation (2.7) by

$$I_4(s,t) = \int \frac{\mathrm{d}^D L}{(2\pi)^D} \frac{-i(4\pi)^{D/2}}{(L^2 - m^2)((L + k_1)^2 - m^2)((L + k_1 + k_2)^2 - m^2)((L - k_4)^2 - m^2)},$$

$$I_3(s) = \int \frac{\mathrm{d}^D L}{(2\pi)^D} \frac{i(4\pi)^{D/2}}{(L^2 - m^2)((L + k_1)^2 - m^2)((L + k_1 + k_2)^2 - m^2)},$$

$$I_2(s) = \int \frac{\mathrm{d}^D L}{(2\pi)^D} \frac{-i(4\pi)^{D/2}}{(L^2 - m^2)((L + k_1 + k_2)^2 - m^2)},$$

$$I_1 = \int \frac{\mathrm{d}^D L}{(2\pi)^D} \frac{i(4\pi)^{D/2}}{(L^2 - m^2)}, \quad (2.8)$$

where $D = 4 - 2\epsilon$, $s = (k_1 + k_2)^2$, $t = (k_2 + k_3)^2$ and $u = (k_1 + k_3)^2$. We obtain the remaining integrals in equation (2.7) by permuting the external legs. The unitarity method efficiently targets the coefficients of the integrals in equation (2.7). The integrals $I_2(0)$ and I_1 are respectively bubble on external leg and tadpole contributions, and are independent of kinematic variables. As we discuss below, because they lack dependence on kinematic variables, these latter integrals require special treatment to determine their coefficients.

Traditionally, unitarity of the scattering matrix is implemented at the integrated level via dispersion relations (see e.g. reference [34]). However, for our purposes, it is much more convenient to use an integrand-level version of unitarity [21]. This is based on the concept of a generalized-unitarity cut that reduces an integrand to a sum of products of tree-level amplitudes. For example, for the *s*-channel cut displayed in figure 1(a),

$$i\mathcal{M}_4^{1-\text{loop}}\Big|_{s-\text{cut}} = \int \frac{\mathrm{d}^{4-2\epsilon} L}{(2\pi)^{4-2\epsilon}} \frac{1}{L_1^2 - m^2} \frac{1}{L_2^2 - m^2} \sum_{\text{states}} \mathcal{M}_4^{\text{tree}}(1, 2, L_1, L_2)$$

$$\times \mathcal{M}_4^{\text{tree}}(-L_1, -L_2, 3, 4)\Big|_{s-\text{cut}}, \quad (2.9)$$
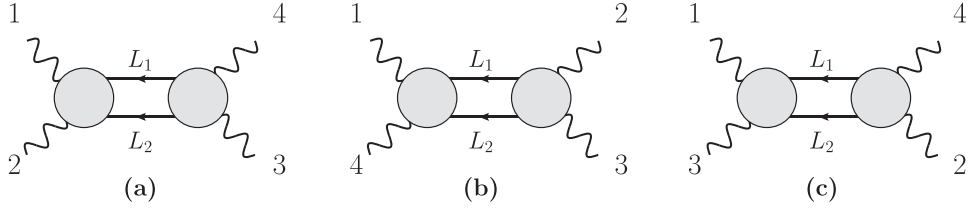
Figure 1. The (a) *s*-, (b) *t*- and (c) *u*-channel two-particle cuts of a one-loop four-point amplitude. The exposed lines are all on shell and the blobs represent tree-level amplitudes.

where $L_1 = L$ and $L_2 = -L - k_1 - k_2$ represent the two cut legs, and $\mathcal{M}_4^{\text{tree}}$ denote the tree-level amplitudes. The sum runs over all intermediate physical states that contribute for a given set of external states. The three generalized-unitarity cuts of the one-loop four-point amplitude are shown in figure 1. In this figure the exposed lines are all on shell and the blobs represent on-shell tree-level amplitudes.

To obtain the full one-loop amplitude we must combine the unitarity cuts. One possibility is to carry this out prior to integration by finding a single integrand with the correct unitarity cuts in all channels [35]. Some non-trivial examples where this approach was implemented are high-loop computations in super-Yang–Mills and supergravity (see e.g. references [36]). On the other hand, in high-multiplicity QCD calculations (see e.g. reference [37]) the cuts are usually combined after reducing to a basis of integrals. We apply the latter approach here. We do so by promoting each cut propagator to a Feynman propagator, and each cut to a Feynman integral. We then use FIRE6 [33] to reduce each Feynman integral to the scalar integrals appearing in equation (2.7). In each cut channel we only determine coefficients of basis integrals with cuts in that channel. By systematically evaluating each cut we determine all coefficients except for those of integrals without kinematic dependence, i.e. $I_2(0)$ and I_1. In the case of gauge theory, the corresponding coefficients are determined by imposing the known UV behavior of the amplitudes [20]. Below, we describe an analogous procedure for the case of gravitational amplitudes.

2.1.3. Double copy. To efficiently obtain the unitarity cuts of the four-graviton amplitude, we use the double-copy construction [22, 38] which expresses gravitational scattering amplitudes directly in terms of gauge-theory ones. Here we use the BCJ form of the double-copy relations [22, 23], which is more natural when organizing expressions in terms of diagrams.

To apply the BCJ double copy, we start by writing a four-point one-loop gauge-theory amplitude in the following form:

$$i\mathcal{A}_4^{1-\text{loop}} = g^4 \sum_i \int \frac{\mathrm{d}^{4-2\epsilon} L}{(2\pi)^{4-2\epsilon}} \frac{1}{S_i} \frac{n_i c_i}{D_i}. \tag{2.10}$$

The sum runs over all distinct four-point one-loop graphs with trivalent vertices. We denote the gauge-theory coupling constant by g. We label each graph by an integer i. The S_i are the symmetry factors of the graphs. The color factor c_i of each graph is obtained by dressing each vertex by a structure constant $\tilde{f}^{abc}$, since we take all particles to be in the adjoint representation. Our normalization of the structure constants follows that of reference [22]. The denominator D_i contains the propagators of each graph. Finally, we capture all non-trivial kinematic dependence by the numerator n_i.
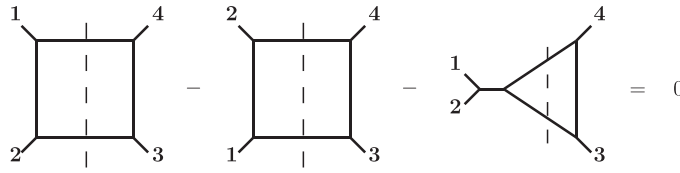
Figure 2. Example of a color or numerator relation for the one-loop four-point amplitudes. Here the diagram represent either color or kinematic numerators. We use these relations on the generalized-unitarity cuts, as indicated by the dashed lines. The relations are effectively tree-level ones except that the state and color sums on the cut legs are carried out.

The color factors obey color-algebra relations of the type

$$c_i - c_j - c_k = 0, \tag{2.11}$$

where i, j and k are some graphs. These relations follow from the Jacobi identity obeyed by the structure constants $\tilde{f}^{abc}$. For a representation obeying color-kinematics duality, the numerators satisfy the same Jacobi relations, i.e.

$$\tilde{n}_i - \tilde{n}_j - \tilde{n}_k = 0. \tag{2.12}$$

The tildes on the numerators signify that these numerators do not have to be the same as the ones appearing in equation (2.10), but as noted in reference [22] these can be kinematic numerators from a different theory. Given such a representation we may obtain the corresponding gravitational amplitude simply by replacing the color factor with the corresponding kinematic numerator,

$$c_i \to \tilde{n}_i, \tag{2.13}$$

so that

$$i\mathcal{M}_4^{1-\text{loop}} = \left(\frac{\kappa}{2}\right)^4 \sum_i \int \frac{\mathrm{d}^{4-2\epsilon}L}{(2\pi)^{4-2\epsilon}} \frac{1}{S_i} \frac{n_i \tilde{n}_i}{D_i}, \tag{2.14}$$

where κ is the gravitational coupling constant, which is given in terms of Newton's constant by,

$$\kappa^2 = 32\pi G. \tag{2.15}$$

The matter content of the resulting gravitational theory is determined by the choice of the numerators n_i and $\tilde{n}_i$. We use this to control the type of particle circulating in the loop.

While the BCJ double copy is usually formulated at the level of the full integrand, since we extract the final answer directly from the cuts it is more convenient to use it at the level of generalized-unitarity cuts. In figure 2 we depict an example of a color relation in terms of cut graphs. In this way we can ignore duality relations that contain diagrams without support on the given cut. For duality relations with support on the cut, the particles of each tree-level amplitude entering the cut remain on the same side of the cut for all three diagrams, as is the case in figure 2. Effectively this amounts to using the duality relations for the two tree-level amplitudes on each side of the cut (see e.g. figure 3). The tree-level relations are sufficient to ensure that for the cut diagram, the double-copy replacement formula (2.13) holds.
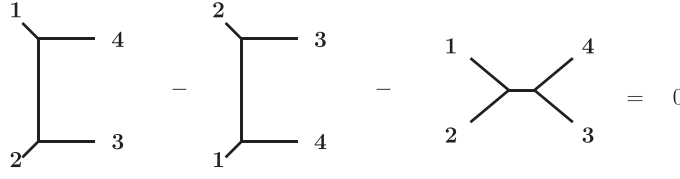
Figure 3. The tree-level four-point color or numerator Jacobi identity. This can be used to set the numerator of one of the diagrams to zero.

We also take advantage of a property of four-point tree-level amplitudes noted in the original BCJ paper [22], which states that we can effectively set one of the three numerators to zero. In this way the duality relation implies that the other two numerators must be equal (up to a possible sign). In order to achieve that, we absorb the propagator of the diagram whose numerator is set to zero into the numerators of the other two diagrams. Specifically, consider the four-point all-gluon tree-level amplitude,

$$i\mathcal{A}_4^{\text{tree}} = g^2 \left(\frac{n_s c_s}{s} + \frac{n_t c_t}{t} + \frac{n_u c_u}{u} \right). \tag{2.16}$$

Using the color Jacobi identity depicted in figure 3, $c_s = c_t - c_u$, we can eliminate c_s in favor of the other two color factors, so that

$$i\mathcal{A}_4^{\text{tree}} = g^2 \left(\frac{n_t' c_t}{t} + \frac{n_u' c_u}{u} \right), \tag{2.17}$$

where

$$n_t' = n_t + n_s \frac{t}{s}, \qquad n_u' = n_u - n_s \frac{u}{s}. \tag{2.18}$$

Demanding that the numerators n_t' and n_u' (and $n_s' = 0$) satisfy the duality relation of figure 3 then implies that

$$n_t' = n_u'. \tag{2.19}$$

The analysis is identical in the presence of matter.

2.2. Setup of the calculation

Our goal is to obtain the four-point one-loop amplitude with external gravitons and with minimally-coupled massive spinning particles up to spin 2 circulating in the loop. Following the generalized-unitarity method we first build the integrand-level generalized-unitarity cuts in equation (2.9). For the one-loop four-point amplitude there are three independent cuts, labeled by the Mandelstam invariant that can be build out of the external graviton momenta on the tree-level amplitudes, i.e. *s*-, *t*- and *u*-channel cuts. We consider all spins up to spin 2 for the massive particles and we denote their mass by *m*. While these masses can be different for each particle since only a single particle at a time circulates in the loop there is no need to put an index labeling the massive particle. We take the massive particles to be real.

We note that there is an ambiguity in the definition of the minimal coupling of a spin 2 particle to gravity. In this paper, we fix this ambiguity by demanding that we recover pure gravity in the appropriate massless limit. This choice also preserves tree-level unitarity [39] and causality [40]. We discuss this ambiguity and the choice we make in this paper in appendix A.
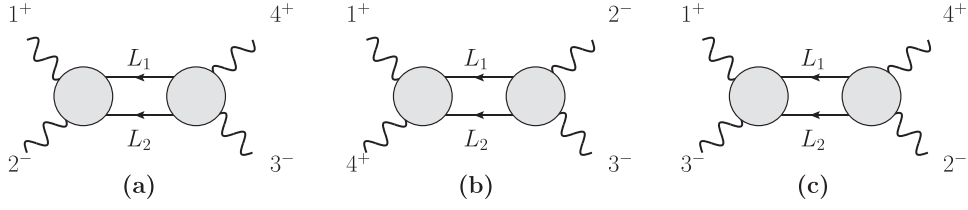
Figure 4. The (a) *s*-, (b) *t*- and (c) *u*-channel two-particle cuts of a one-loop four-point amplitude with two negative- and two positive-helicity external gluons or gravitons (double-minus configuration). The internal lines represent massive spinning particles. The exposed lines are all on shell and the blobs represent tree-level amplitudes.

For the amplitude in question there exist three independent helicity configurations. Specifically, we calculate

$$\mathcal{M}_4^{1-\text{loop}}(1^+, 2^-, 3^-, 4^+), \qquad \mathcal{M}_4^{1-\text{loop}}(1^+, 2^+, 3^-, 4^+), \qquad \mathcal{M}_4^{1-\text{loop}}(1^+, 2^+, 3^+, 4^+). \tag{2.20}$$

We refer to these configurations respectively as *double-minus*, *single-minus* and *all-plus*. All other amplitudes are related to these via relabelings and parity.

We build the generalized-unitarity cuts from four-point tree-level gravity amplitudes. The double copy implies that these tree-level amplitudes can be directly obtained from the corresponding gauge-theory ones, which can be described by the three diagrams shown in figure 3. As we discussed in the previous section we can use the color Jacobi identity in the gauge-theory case to remove one diagram, at the expense of other diagrams obtaining numerators that are nonlocal in the external kinematic invariants. The net effect is that after multiplying and dividing by appropriate propagators every contribution to the cut can be assigned to cut box diagrams. Moreover, on the generalized cuts the four-point tree-level BCJ numerator relations set the remaining numerators equal to each other, as noted in equation (2.19).

For example, for the four-point gauge-theory amplitude the *s*-channel cut in figure 4(a) is of the form,

$$i\mathcal{A}_4\big|_{s-\text{cut}} = g^4 \int \frac{\mathrm{d}^{4-2\epsilon}L}{(2\pi)^{4-2\epsilon}} \, n_{\text{g},s} \left(\frac{c_{1234}}{D_{1234}} + \frac{c_{1342}}{D_{1342}} \right)\bigg|_{s-\text{cut}}. \tag{2.21}$$

The box color factor (see figure 5) is given by

$$c_{abcd} = \tilde{f}^{e_a g_1 g_4} \tilde{f}^{e_b g_2 g_1} \tilde{f}^{e_c g_3 g_2} \tilde{f}^{e_d g_4 g_3}, \tag{2.22}$$

where *abcd* takes in the values indicated in equation (2.21). As usual, repeated indices are summed. The denominators are given by products of the usual Feynman propagators,

$$D_{abcd} = \left(L^2 - m^2\right)\left((L + k_a)^2 - m^2\right)\left((L + k_a + k_b)^2 - m^2\right)$$
$$\times \left((L + k_a + k_b + k_c)^2 - m^2\right), \tag{2.23}$$

where the k_a are external momenta. Finally, $n_{\text{g},s}$ is a gauge-theory kinematic factor common to both box diagrams.
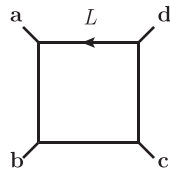
Figure 5. Generic box diagram whose color factor and denominator are given by equations (2.22) and (2.23) respectively. The external momenta are taken incoming while the direction of the loop momentum is indicated by the arrow.

The gravitational cuts are similar. For example, the *s*-channel cut is of the form,

$$i\mathcal{M}_4|_{s-\text{cut}} = \left(\frac{\kappa}{2}\right)^4 \int \frac{d^{4-2\epsilon}L}{(2\pi)^{4-2\epsilon}} \, n_{\text{GR},s} \left(\frac{1}{D_{1234}} + \frac{1}{D_{1342}}\right)\Bigg|_{s-\text{cut}}, \tag{2.24}$$

where $n_{\text{GR},s}$ is the single *s*-channel gravitational kinematic numerator. As we noted in section 2.1, the gravity amplitudes follow from replacing the color factor with a gauge-theory kinematic numerator. As we describe in more detail below, the particle content circulating in the loop is determined by the choice of the gauge-theory numerators.

As can be seen from figures 4(a) and (c), for the indicated helicity configuration, the *u*-channel cut is obtained by relabeling the momenta and spinors in the *s*-channel cut: $k_2 \leftrightarrow k_3$ and $\langle 2| \leftrightarrow \langle 3|$. For the single-minus and all-plus configurations, all three cuts are given by appropriate relabelings of a single cut.

2.3. Supersymmetric decompositions

We are interested in the problem of minimally-coupled massive matter circulating in the loop. A convenient way to organize this calculation is by following the massless case where each particle can be described as a linear combination of supersymmetric multiplets circulating in the loop. Our organization is in direct correspondence to this supersymmetric decomposition [19, 41, 42]. This allows us to recycle the results of the calculation of the lower-spin particles circulating in the loop into contributions for the higher-spin particles. It also has the advantage of grouping together terms that contain integrals of the same tensor rank. For the gauge-theory case such a decomposition has already been used to organize the contributions of massive spin 0, 1/2 and 1 particles circulating in the loops [20, 42]. The double-copy construction will allow us to import this into minimally-coupled gravitational theories with up to spin-2 massive particles.

2.3.1. Gauge theory. We start by examining the corresponding amplitudes in gauge theory. For the case of one-loop four-point amplitudes with external gluons and massive matter circulating in the loop, reference [42] showed that[8]

$$\begin{aligned}
\mathcal{A}_4^{S=0} &= \mathcal{A}_4^{\{0\}}, \\
\mathcal{A}_4^{S=1/2} &= -2\mathcal{A}_4^{\{0\}} + \mathcal{A}_4^{\{1/2\}}, \\
\mathcal{A}_4^{S=1} &= 3\mathcal{A}_4^{\{0\}} - 4\mathcal{A}_4^{\{1/2\}} + \mathcal{A}_4^{\{1\}},
\end{aligned} \tag{2.25}$$

[8] We combine the 'gauge boson' and 'scalar' contributions of reference [42] into the massive $S = 1$ result.

where the notation $\mathcal{A}_4^{\{S\}}$ denotes the new piece we need to calculate at spin-S. In this way we express the amplitudes with a given spinning particle circulating in the loop in terms of the simpler-to-calculate new pieces. Inverting the above equations, we may think of the new pieces as amplitudes with multiplets circulating in the loop. These massive multiplets have the same degrees of freedom as the corresponding massless ones. Hence, in terms of on-shell supersymmetric representation theory they satisfy the BPS condition [43]. For recent calculations involving massive supersymmetric multiplets in gauge theory see references [44, 45].

Before turning to the corresponding gravitational decomposition it is useful to first look at the massless limit. For the gauge theory case we have that as $m \to 0$,

$$
\begin{aligned}
\mathcal{A}_4^{S=0} &\to \mathcal{A}_4^{S=0,\,m=0}, \\
\mathcal{A}_4^{S=1/2} &\to \mathcal{A}_4^{S=1/2,\,m=0}, \\
\mathcal{A}_4^{S=1} &\to \mathcal{A}_4^{S=1,\,m=0} + \mathcal{A}_4^{S=0,\,m=0},
\end{aligned}
\tag{2.26}
$$

where we see that the $S = 1$ case is nontrivial, which follows from the mismatch in the number of degrees of freedom in a massive and massless vector boson.

2.3.2. Background field gauge. A nice way to understand the above supersymmetric decomposition is in terms of background field gauge [30]. While we do not use background field gauge to compute the amplitudes, it does offer useful insight into the structure of the amplitude. For the different particles circulating in the loop we can write the effective action as

$$
\begin{aligned}
\Gamma_{S=0}[A] &= \ln \det_{[0]}^{-1/2} \left(D^2 + m^2 \right), \\
\Gamma_{S=1/2}[A] &= \ln \det_{[1/2]}^{1/4} \left(D^2 \,\mathbb{I} + m^2 \,\mathbb{I} - g\frac{1}{2}\sigma^{\mu\nu}F_{\mu\nu} \right), \\
\Gamma_{S=1}[A] &= \ln \det_{[1]}^{-1/2} \left(D^2\mathbb{I} + m^2\mathbb{I} - g\Sigma^{\mu\nu}F_{\mu\nu} \right) \\
&\quad + \ln \det_{[0]} \left(D^2 + m^2 \right) + \ln \det_{[0]}^{-1/2} \left(D^2 + m^2 \right),
\end{aligned}
\tag{2.27}
$$

where $\mathbb{I}$ is the identity matrix, $\sigma_{\mu\nu}/2$ is the spin-1/2 Lorentz generator and $\Sigma_{\mu\nu}$ is the spin-1 Lorentz generator. Ignoring the $F_{\mu\nu}$ terms and focusing on the $(D^2 + m^2)$ term, each state corresponds to a power to which the determinant is raised: $-1/2$ for a bosonic state and $+1/2$ for a fermionic state. For a massive real scalar there is precisely one bosonic state corresponding the $-1/2$ power to which the first determinant is raised. For the $S = 1/2$ fermion the determinant is raised to the $1/4$ power to account for the 4×4 Dirac determinant, effectively leaving a single power that corresponds to the two states of a Majorana fermion. Similarly, for the $S = 1$ vector, the determinant is over a 4×4 Lorentz generator space so the exponent of $-1/2$ in the first term in $\Gamma_{S=1}[A]$ corresponds to four bosonic states. This is then reduced to two states by the ghost determinant corresponding to the second term and increased by one state from the scalar longitudinal degree of freedom required by a massive vector boson corresponding to the third term. This extra degree of freedom is incorporated into equation (2.25) as well. To make the supersymmetric cancellations more manifest we rewrite the Dirac determinant as a product of determinants so that the similarity to the bosonic case is clear,

$$
\begin{aligned}
\det_{[1/2]}^{1/2}(\slashed{D} + im)\det_{[1/2]}^{1/2}(\slashed{D} - im) &= \det_{[1/2]}^{1/4}(\slashed{D}^2 + m^2) \\
&= \det_{[1/2]}^{1/4}\left(D^2 + m^2 - g\frac{1}{2}\sigma^{\mu\nu}F_{\mu\nu} \right),
\end{aligned}
\tag{2.28}
$$

where $\not{D}^2 = \frac{1}{2}\{\not{D}, \not{D}\} + \frac{1}{2}[\not{D}, \not{D}]$. This corresponds to using the second order formalism for fermions described in reference [46]. The fact that the mass enters into equation (2.27) in such a simple manner can also be understood in terms of a Kaluza–Klein reduction of the massless case from five dimensions, truncated to keeping only the lowest massive state in the loop.

The effective-action determinants (2.27) can be straightforwardly applied to show that the supersymmetric decomposition organizes contributions in terms of power counting of the resulting diagrams. The terms with leading power of loop momenta come from the D^2 terms in equation (2.27), because $F_{\mu\nu}$ contains only the external gluon momenta. If we set the $F_{\mu\nu}$ terms to zero then in all supersymmetric combinations the balance between the bosons and fermions implies the leading powers of loop momentum cancel. Subleading terms in supersymmetric combinations come from using one or more factors of $F_{\mu\nu}$ when generating a graph; each $F_{\mu\nu}$ reduces the maximum power of momentum by one. Terms with a lone $F_{\mu\nu}$ vanish, thanks to $\mathrm{Tr}(\sigma^{\mu\nu}) = \mathrm{Tr}(\Sigma^{\mu\nu}) = 0$. This reduces the leading power in an m-point one-particle-irreducible diagram from ℓ^m down to ℓ^{m-2}. For $\mathcal{A}_m^{\{1\}}$, a comparison of the traces of products of two and three $\sigma^{\mu\nu}$ and $\Sigma_{\mu\nu}$ shows that further cancellations reduce the leading behavior all the way down to ℓ^{m-4}. In gauges other than Feynman background field gauge, these cancellations would be more more obscure.

2.3.3. Gravity.   Now consider the gravitational case. In the $m \to 0$ limit we again have a mismatch in the number of degrees of freedom for $S \geqslant 1$ between the massive and massless cases,

$$
\begin{aligned}
\mathcal{M}_4^{S=0} &\to \mathcal{M}_4^{S=0,\, m=0}, \\
\mathcal{M}_4^{S=1/2} &\to \mathcal{M}_4^{S=1/2,\, m=0}, \\
\mathcal{M}_4^{S=1} &\to \mathcal{M}_4^{S=1,\, m=0} + \mathcal{M}_4^{S=0,\, m=0}, \\
\mathcal{M}_4^{S=3/2} &\to \mathcal{M}_4^{S=3/2,\, m=0} + \mathcal{M}_4^{S=1/2,\, m=0}, \\
\mathcal{M}_4^{S=2} &\to \mathcal{M}_4^{S=2,\, m=0} + \mathcal{M}_4^{S=1,\, m=0} + \mathcal{M}_4^{S=0,\, m=0}.
\end{aligned}
\tag{2.29}
$$

In the massless case with spinning particles circulating in the loop we can again decompose the amplitudes in terms of ones with supermultiplets circulating in the loop [19],

$$
\begin{aligned}
\mathcal{M}_4^{S=0,\, m=0} &= \mathcal{M}_4^{\{0\},\, m=0}, \\
\mathcal{M}_4^{S=1/2,\, m=0} &= -2\mathcal{M}_4^{\{0\},\, m=0} + \mathcal{M}_4^{\{1/2\},\, m=0}, \\
\mathcal{M}_4^{S=1,\, m=0} &= 2\mathcal{M}_4^{\{0\},\, m=0} - 4\mathcal{M}_4^{\{1/2\},\, m=0} + \mathcal{M}_4^{\{1\},\, m=0}, \\
\mathcal{M}_4^{S=3/2,\, m=0} &= -2\mathcal{M}_4^{\{0\},\, m=0} + 9\mathcal{M}_4^{\{1/2\},\, m=0} - 6\mathcal{M}_4^{\{1\},\, m=0} + \mathcal{M}_4^{\{3/2\},\, m=0}, \\
\mathcal{M}_4^{S=2,\, m=0} &= 2\mathcal{M}_4^{\{0\},\, m=0} - 16\mathcal{M}_4^{\{1/2\},\, m=0} + 20\mathcal{M}_4^{\{1\},\, m=0} \\
&\quad - 8\mathcal{M}_4^{\{3/2\},\, m=0} + \mathcal{M}_4^{\{2\},\, m=0}.
\end{aligned}
\tag{2.30}
$$

The $\{S\}$ pieces in each case are in direct correspondence to the supermultiplets circulating in the loop, as defined in reference [19].[9] For example,

$$
\mathcal{M}_4^{\{1/2\},\, m=0} = \mathcal{M}_4^{\mathcal{N}=1,\, m=0},
\tag{2.31}
$$

---

[9] We use a real scalar while reference [19] used a complex one. Hence there is relative factor of 1/2 for that contribution.

where $\mathcal{N} = 1$ denotes the chiral multiplet consisting of a Weyl fermion and two real scalars.

Using the relation between the massive and massless amplitudes in equation (2.29) and the massless supersymmetric decomposition in equation (2.30), we can organize our computation in a similar way as for gauge theory in equation (2.25). Specifically, we have

$$
\begin{aligned}
\mathcal{M}_4^{S=0} &= \mathcal{M}_4^{\{0\}}, \\
\mathcal{M}_4^{S=1/2} &= -2\mathcal{M}_4^{\{0\}} + \mathcal{M}_4^{\{1/2\}}, \\
\mathcal{M}_4^{S=1} &= 3\mathcal{M}_4^{\{0\}} - 4\mathcal{M}_4^{\{1/2\}} + \mathcal{M}_4^{\{1\}}, \\
\mathcal{M}_4^{S=3/2} &= -4\mathcal{M}_4^{\{0\}} + 10\mathcal{M}_4^{\{1/2\}} - 6\mathcal{M}_4^{\{1\}} + \mathcal{M}_4^{\{3/2\}}, \\
\mathcal{M}_4^{S=2} &= 5\mathcal{M}_4^{\{0\}} - 20\mathcal{M}_4^{\{1/2\}} + 21\mathcal{M}_4^{\{1\}} - 8\mathcal{M}_4^{\{3/2\}} + \mathcal{M}_4^{\{2\}}.
\end{aligned}
\tag{2.32}
$$

The massive multiplets circulating in the loop are 'short', i.e. they have the same degrees of freedom as the corresponding massless ones. Hence, they obey the BPS condition [43]. While we have not tried directly embedding these amplitudes into supergravity theories, it is an interesting question to do so. Here we use the relation to supermultiplets for a more modest aim of reorganizing the contributions, so that as the spin increases the new pieces become simpler. Examples of supergravity calculations involving massive multiplets are given in reference [47].

We note that in general, care is required when using dimensional regularization in conjunction with helicity methods and supersymmetric decompositions. To allow for different choices of regularization scheme, we would need to correct the last line of equation (2.25) to be [42]

$$
\mathcal{A}_4^{S=1} = (3 - 2\delta_R\epsilon)\mathcal{A}_4^{\{0\}} - 4\mathcal{A}_4^{\{1/2\}} + \mathcal{A}_4^{\{1\}},
\tag{2.33}
$$

where $\delta_R = 0$ in the FDH scheme [28] and $\delta_R = 1$ in either conventional dimensional regularization [27] or the 't Hooft-Veltman scheme [48]. One may then propagate this correction to the gravitational amplitudes through the double copy. While the correction is of $\mathcal{O}(\epsilon)$, it can interfere with infrared or UV singularities to give nontrivial contributions. However, for the massive amplitudes that we are computing here, the distinction between different schemes is not important because the four-graviton amplitudes with massive matter in the loop are both UV and infrared finite (see sections 2.6 and 2.8).

We also note that the coefficient of the scalar ($\mathcal{M}_4^{\{0\}}$) counts the degrees of freedom of the particle in question, modulo a minus sign for the fermions. Recall that all particles are taken to be real. Also, given the general setup of our amplitudes (equations (2.21) and (2.24)), we get a similar decomposition for the corresponding numerators $n_{g,\alpha}$ and $n_{GR,\alpha}$.

Finally, the supersymmetric decomposition is simplified in the case of the single-minus and all-plus configurations. In these cases, supersymmetric Ward identities [49] show that only the $\mathcal{M}_4^{\{0\}}$ piece gives a nonzero contribution for each spin particle in the loop. Using this observation, equation (2.32) becomes

$$
\begin{aligned}
\mathcal{M}_4^{S=0}(1^\pm, 2^+, 3^+, 4^+) &= \mathcal{M}_4^{\{0\}}(1^\pm, 2^+, 3^+, 4^+), \\
\mathcal{M}_4^{S=1/2}(1^\pm, 2^+, 3^+, 4^+) &= -2\mathcal{M}_4^{\{0\}}(1^\pm, 2^+, 3^+, 4^+), \\
\mathcal{M}_4^{S=1}(1^\pm, 2^+, 3^+, 4^+) &= 3\mathcal{M}_4^{\{0\}}(1^\pm, 2^+, 3^+, 4^+), \\
\mathcal{M}_4^{S=3/2}(1^\pm, 2^+, 3^+, 4^+) &= -4\mathcal{M}_4^{\{0\}}(1^\pm, 2^+, 3^+, 4^+), \\
\mathcal{M}_4^{S=2}(1^\pm, 2^+, 3^+, 4^+) &= 5\mathcal{M}_4^{\{0\}}(1^\pm, 2^+, 3^+, 4^+).
\end{aligned}
\tag{2.34}
$$

Therefore for these two helicity configurations, it is sufficient to calculate the $S = 0$ amplitude where only a scalar circulates in the loop.

2.4. Kinematic numerators through the double copy

Following the double-copy construction we can directly obtain the gravitational unitarity-cut numerators from gauge-theory ones. We may express the double copy in terms of spinning particles or new pieces (equation (2.32)) circulating in the loop. In the latter case, we find an especially compact representation for the numerators.

Taking the product of gauge-theory kinematic numerators we decompose them into cut numerators of the gravitational case. In terms of spin we have

$$
\begin{aligned}
n_{g,\alpha}^{\tilde{S}} n_{g,\alpha}^{S=0} &= n_{GR,\alpha}^{\tilde{S}}, \quad \text{for } \tilde{S} = 0, 1/2, 1, \\
n_{g,\alpha}^{S=1/2} n_{g,\alpha}^{S=1/2} &= n_{GR,\alpha}^{S=1} + n_{GR,\alpha}^{S=0}, \\
n_{g,\alpha}^{S=1} n_{g,\alpha}^{S=1/2} &= n_{GR,\alpha}^{S=3/2} + n_{GR,\alpha}^{S=1/2}, \\
n_{g,\alpha}^{S=1} n_{g,\alpha}^{S=1} &= n_{GR,\alpha}^{S=2} + n_{GR,\alpha}^{S=1} + n_{GR,\alpha}^{S=0},
\end{aligned}
\tag{2.35}
$$

where α denotes the cut under consideration. These are in direct correspondence to the Clebsch–Gordan decomposition. In terms of the new pieces in equation (2.32),

$$
\begin{aligned}
n_{g,\alpha}^{\{S\}} n_{g,\alpha}^{\{0\}} &= n_{GR,\alpha}^{\{S\}}, \quad \text{for } S = 0, 1/2, 1, \\
n_{g,\alpha}^{\{1/2\}} n_{g,\alpha}^{\{1/2\}} &= n_{GR,\alpha}^{\{1\}}, \\
n_{g,\alpha}^{\{1\}} n_{g,\alpha}^{\{1/2\}} &= n_{GR,\alpha}^{\{3/2\}}, \\
n_{g,\alpha}^{\{1\}} n_{g,\alpha}^{\{1\}} &= n_{GR,\alpha}^{\{2\}}.
\end{aligned}
\tag{2.36}
$$

Observe that the gauge-theory numerator $n_{g,\alpha}^{S=0}$ along with either $n_{g,\alpha}^{S=1/2}$ or $n_{g,\alpha}^{S=1}$ are sufficient to construct all the gravitational numerators up to spin 2. We explicitly verified that both constructions yield the same result. References [20, 42] calculated the corresponding amplitudes $\mathcal{A}_4^{S=0}$, $\mathcal{A}_4^{S=1/2}$ and $\mathcal{A}_4^{S=1}$, from which we may extract the desired kinematic numerators. As a consistency check, we match $\mathcal{A}_4^{\{1\}}$, which was calculated in reference [45].

From the above construction we find a remarkably simple form for the kinematic numerators. For the double-minus gauge-theory numerators we have,

$$
n_{g,\alpha}^{\{S\}} = \left([14]\langle 23\rangle\right)^2 (\psi_\alpha)^{2-2S},
\tag{2.37}
$$

while for the corresponding gravity numerators we have,

$$
n_{GR,\alpha}^{\{S\}} = \left([14]\langle 23\rangle\right)^4 (\psi_\alpha)^{4-2S},
\tag{2.38}
$$

where

$$
\psi_s \equiv \frac{\langle 2|\ell|1] \langle 3|\ell|4]}{s[14]\langle 23\rangle}, \qquad \psi_t \equiv \frac{(m^2 + \mu^2)}{t}, \qquad \psi_u \equiv \frac{\langle 3|\ell|1] \langle 2|\ell|4]}{u[14]\langle 32\rangle}.
\tag{2.39}
$$

For the single-minus and all-plus configurations it is sufficient to calculate the numerators for a scalar circulating in the loop. For the single-minus configuration for gauge theory and gravity we find

$$n_{\text{g},s}^{S=0} = \frac{(m^2 + \mu^2)}{s} \frac{[12]}{\langle 12 \rangle} \langle 3|\ell|4]^2, \qquad n_{\text{GR},s}^{S=0} = \frac{(m^2 + \mu^2)^2}{s^2} \frac{[12]^2}{\langle 12 \rangle^2} \langle 3|\ell|4]^4, \quad (2.40)$$

while for the all-plus configuration we have

$$n_{\text{g},s}^{S=0} = (m^2 + \mu^2)^2 \frac{[12][34]}{\langle 12 \rangle \langle 34 \rangle}, \qquad n_{\text{GR},s}^{S=0} = (m^2 + \mu^2)^4 \left(\frac{[12][34]}{\langle 12 \rangle \langle 34 \rangle} \right)^2. \quad (2.41)$$

For these two configurations, we obtain the numerators for the *t*- and *u*-channel cuts by appropriate relabelings.

Following the conventions of reference [20], we break the $(4 - 2\epsilon)$-dimensional loop momentum L into a four-dimensional part ℓ and a (-2ϵ)-dimensional part μ. We write $L = (\ell, \mu)$. Using this notation we have for example

$$\ell^2 = m^2 + \mu^2, \tag{2.42}$$

when the cut condition $L^2 = m^2$ is satisfied. We take $\epsilon < 0$ so that we can break the loop momentum in this fashion. Further, whenever a four-dimensional vector $v \equiv (v, 0)$ is contracted with the $(4 - 2\epsilon)$-dimensional loop momentum, the (-2ϵ)-dimensional part is projected out,

$$v \cdot L = v \cdot \ell. \tag{2.43}$$

Reference [50] (equations (5.20) and (5.36)) calculated Compton amplitudes for a massive particle in four dimensions of up to spin 1 in gauge theory and up to spin 2 in gravity. We find similar spin dependence in our double-minus numerators as for these Compton amplitudes.

2.5. Integral reduction and cut merging

In this subsection we use standard integration-by-parts (IBP) methods to reduce the generalized-unitarity cuts we calculated in the previous section in terms of a basis of master integrals. This allows us to fix all integral coefficients in equation (2.7) other than those of the tadpole and the bubble on external leg. We discuss these remaining pieces in section 2.6. We show details for the double-minus configuration with a scalar in the loop; the remaining helicity configurations and particle-in-the-loop contributions are similar. In order to keep the expressions concise, we do not include the helicity labels.

The general strategy for constructing the full amplitudes is to evaluate the cuts one by one in terms of the master integrals appearing in equation (2.7). If a given integral has a generalized cut in the channel being evaluated then that channel fully determines its coefficient. By stepping through the three channels in figure 1 we obtain the coefficients of all master integrals except $I_2(0)$ and I_1. The box integrals each have cuts in two channels so consistency requires that they give the same coefficient.

We start with the *s*-channel cut of the $S = 0$ double-minus helicity configuration defined in equation (2.20). The discussion for the *u*-channel cut follows in the same way, since it is given

by a relabeling of the *s*-channel one. We have

$$
i\mathcal{M}_4^{S=0}\bigg|_{s-\text{cut}} = \left(\frac{\kappa}{2}\right)^4 (\langle 23\rangle[14])^4
$$

$$
\times \int \frac{\mathrm{d}^{4-2\epsilon}L}{(2\pi)^{4-2\epsilon}} \left(\frac{\langle 2|\ell|1]\,\langle 3|\ell|4]}{s[14]\langle 23\rangle}\right)^4 \left(\frac{1}{D_{1234}} + \frac{1}{D_{1342}}\right)\bigg|_{s-\text{cut}}. \tag{2.44}
$$

We define

$$
v_1^\mu = \langle 2|\gamma^\mu|1], \qquad v_2^\mu = \langle 3|\gamma^\mu|4], \tag{2.45}
$$

which live in the four-dimensional subspace so that the v_i effectively project out the $(4-2\epsilon)$-dimensional components. This implies that $v_i \cdot \ell = v_i \cdot L$, as follows from the prescription [20] that $\epsilon < 0$. Next, we lift the cut condition and regard our expression as part of the full integrand that we would have obtained by Feynman rules,

$$
i\mathcal{M}_4^{S=0,\,s-\text{channel}} = \left(\frac{\kappa}{2}\right)^4 (\langle 23\rangle[14])^4
$$

$$
\times \int \frac{\mathrm{d}^{4-2\epsilon}L}{(2\pi)^{4-2\epsilon}} \left(\frac{\langle 2|L|1]\,\langle 3|L|4]}{s[14]\langle 23\rangle}\right)^4 \left(\frac{1}{D_{1234}} + \frac{1}{D_{1342}}\right), \tag{2.46}
$$

keeping in mind that it is only valid for contributions that have an *s*-channel cut. We apply IBP identities [51] in $4-2\epsilon$ dimensions to reduce the above integrals to the master integrals in equation (2.7), using the software FIRE6 [33]. Upon reducing to master integrals we remove the ones that have no *s*-channel cut.

Next, we turn to the *t*-channel cut. We have

$$
i\mathcal{M}_4^{S=0}\bigg|_{t-\text{cut}} = \left(\frac{\kappa}{2}\right)^4 (\langle 23\rangle[14])^4
$$

$$
\times \int \frac{\mathrm{d}^4\ell}{(2\pi)^4}\frac{\mathrm{d}^{-2\epsilon}\mu}{(2\pi)^{-2\epsilon}} \left(\frac{(m^2+\mu^2)}{t}\right)^4 \left(\frac{1}{D_{1234}} + \frac{1}{D_{1423}}\right)\bigg|_{t-\text{cut}}, \tag{2.47}
$$

where we subdivide the integration into four- and (-2ϵ)-dimensional parts

$$
\int \frac{\mathrm{d}^{4-2\epsilon}L}{(2\pi)^{4-2\epsilon}} = \int \frac{\mathrm{d}^4\ell}{(2\pi)^4}\frac{\mathrm{d}^{-2\epsilon}\mu}{(2\pi)^{-2\epsilon}}. \tag{2.48}
$$

As for the *s* channel, we lift the cut conditions and regard our expression as part of the full integrand,

$$
i\mathcal{M}_4^{S=0,\,t-\text{channel}} = \left(\frac{\kappa}{2}\right)^4 (\langle 23\rangle[14])^4
$$

$$
\times \int \frac{\mathrm{d}^4\ell}{(2\pi)^4}\frac{\mathrm{d}^{-2\epsilon}\mu}{(2\pi)^{-2\epsilon}} \left(\frac{(m^2+\mu^2)}{t}\right)^4 \left(\frac{1}{D_{1234}} + \frac{1}{D_{1423}}\right). \tag{2.49}
$$

As we discuss in appendix G.2, the integrals with the (-2ϵ)-dimensional components of loop-momentum μ in the numerators can be expressed directly in terms of the master integrals
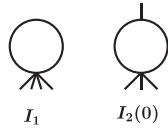
Figure 6. The tadpole and bubble-on-external-leg integrals.

defined in equation (2.8). After reducing our expression we eliminate master integrals that vanish on the *t*-channel cut.

As noted above, an important consistency condition arises from the fact that the box integrals have cuts in two channels, so we can determine their coefficients from either channel. We explicitly verified that the coefficients we obtain for the box integrals from looking at the different channels are the same. In this way we are able to extract all coefficients of equation (2.7) other than a_0 and b_0, since the corresponding integrals have no cuts in any channel; we obtain these remaining two integral coefficients in section 2.6.

2.6. Ultraviolet behavior and rational pieces

Analyzing the different generalized-unitarity cuts we obtain all the coefficients in equation (2.7) other than a_0 and b_0. Tadpoles and bubbles on external legs (see figure 6) vanish on any cut, therefore their coefficients are not accessible through generalized unitarity. In this subsection we use the known UV properties of the amplitude to obtain these coefficients. Similarly to section 2.5, we discuss the double-minus configuration as a specific example. The other configurations follow in the same manner.

First, we observe that simply neglecting these two integrals leads to an inconsistent answer. Expanding to leading order in ϵ we get

$$\mathcal{M}_4^{S=0}\big|_{a_0 \to 0, \, b_0 \to 0} = (\langle 23 \rangle [14])^4 \frac{m^2 \tilde{Q}(s,t,m)}{\epsilon} + \mathcal{O}(\epsilon^0), \tag{2.50}$$

where $\tilde{Q}(s,t,m)$ is some rational function. We note that we expect no UV divergence to appear in the four-graviton amplitude since the only counterterm that we could write to absorb it is the Gauss–Bonnet term, which is evanescent in four dimensions [52]. Also, the UV divergence not coming out local hints that we neglected to include some integrals.

We use the vanishing of the UV divergence to obtain the remaining coefficients a_0 and b_0. Note that since a_0 and b_0 are rational functions of ϵ, it seems that we do not have enough information to fully fix them. We surpass this difficulty by realizing that our problem admits a second integral basis where the integral coefficients do not depend on ϵ and m. This basis is overcomplete in that it contains $(4 - 2\epsilon)$- and higher-dimensional integrals, while the one introduced in equation (2.7) only includes $(4 - 2\epsilon)$-dimensional integrals. We discuss this basis in detail and provide an algorithmic way of switching between the two bases in appendix G.2.

We start by bringing the quantity we get from cut merging to the basis containing higher-dimensional integrals. In doing so, we remove any integrals that vanish on all cuts. We refer to this piece expressed in this basis as the 'on-shell-constructable' piece. Then, we need to figure out which integrals without unitarity cuts we need to include in our expression, since in principle one could add infinitely many higher-dimensional integrals.

We consider the pieces that would arise in a Feynman-rules calculation. For the amplitude at hand we would find integrals up to quartically divergent. After we IBP reduce them and express them in the basis that contains the higher-dimensional integrals, we would in principle

find all scalar integrals up to quartically divergent appearing. Hence we need to consider:

$$I_1, \ I_2(k^2), \ I_1^{6-2\epsilon}, \ I_2^{6-2\epsilon}(k^2), \ I_2^{8-2\epsilon}(k^2), \tag{2.51}$$

where we imagine the limiting case $k^2 \to 0$. For our definition of higher-dimensional integrals see equation (G1). These are all the integrals that have no unitarity cuts and are up to quartically divergent. The reason why we consider the above limiting case is because the coefficients of these integrals might have a $1/k^2$ pole.

Then, the unknown piece in our amplitude up to normalization takes the form

$$\lim_{k^2 \to 0} \left(\beta_1 I_1 + \beta_2 I_2(k^2) + \beta_3 I_1^{6-2\epsilon} + \beta_4 I_2^{6-2\epsilon}(k^2) + \beta_5 I_2^{8-2\epsilon}(k^2) \right)$$
$$= \gamma_1 I_1 + \gamma_2 I_2(0) + \gamma_3 I_1^{6-2\epsilon}, \tag{2.52}$$

where we use

$$I_2^{6-2\epsilon}(k^2) = I_1 + \frac{k^2}{6} I_2(0) + \mathcal{O}(k^4), \qquad I_2^{8-2\epsilon}(k^2) = I_1^{6-2\epsilon} + \frac{k^2}{6} I_1 + \mathcal{O}(k^4), \tag{2.53}$$

following appendix B of reference [20]. The unknown coefficients γ_i are rational functions of the kinematics that do not contain ϵ or m dependence. In writing this expression we assume that the coefficients of the integrals are at worse divergent as $1/k^2$ and that the expression is finite in the $k^2 \to 0$ limit when a massive particle is circulating in the loop.[10]

Upon these considerations, our amplitude takes the form

$$\mathcal{M}_4^{S=0} = \mathcal{M}_4^{S=0} \bigg|_{\text{on-shell-constr.}} + \frac{1}{(4\pi)^{2-\epsilon}} \left(\frac{\kappa}{2} \right)^2 \mathcal{M}_4^{\text{tree}}$$
$$\times \left(\gamma_1 I_1 + \gamma_2 I_2(0) + \gamma_3 I_1^{6-2\epsilon} \right), \tag{2.54}$$

where the γ_i are coefficients are determined from the requirement that UV divergences cancel, the integrals I_1 and I_2 are given in equations (G2) and (G4), and

$$\mathcal{M}_4^{\text{tree}} \equiv \mathcal{M}_4^{\text{tree}}(1^+, 2^-, 3^-, 4^+) = \left(\frac{\kappa}{2} \right)^2 \frac{(\langle 23 \rangle [14])^4}{stu}. \tag{2.55}$$

The Mandelstam variables are defined below equation (2.8). Note that the little-group scaling for the unknown terms is fixed to be $(\langle 23 \rangle [14])^4$. Since the γ_i are rational functions of the kinematics, we are free to multiply and divide by (stu) to introduce $\mathcal{M}_4^{\text{tree}}$.

Next, demanding that the amplitude has no UV divergence and that all three γ_i are independent of the mass m uniquely fixes γ_1 and γ_3 to nonzero values while setting $\gamma_2 = 0$. The results for the amplitudes collected in appendix F the values of the γ_i are all chosen to make the amplitudes UV finite.

As a simple consistency check, we repeat this analysis adding integrals that have no unitarity cuts and are divergent worse than quartically (namely $(8 - 2\epsilon)$- and higher-dimensional tadpoles). We verify that the answer is the same, i.e. the coefficients of these new integrals are set to zero.

---

[10] We drop ill-defined 'cosmological constant' tadpole diagrams with a $1/\left(\sum_{i=1}^4 k_i \right)^2$ propagator.

We now briefly comment on the $m \to 0$ limit. In this limit, both I_1 and $I_2(0)$ vanish, and the amplitude has no UV divergence, as can be seen from equation (2.50). Hence our amplitude has the correct UV behavior in this limit as well. Moreover, all integrals that have no unitarity cuts become scaleless and hence zero in dimensional regularization. Therefore, there are no unfixed coefficients and our construction based on the unitarity cuts captures the full amplitude.

Finally, we want to clarify why we chose the overcomplete basis with higher-dimensional integrals. The coefficients a_0 and b_0 are in principle arbitrary rational functions of ϵ and m, and of the kinematic variables s and t. The existence of a basis where the integral coefficients do not contain ϵ or m significantly restricts the functional dependence of a_0 and b_0 on ϵ and m. The existence of such a basis is a nontrivial fact that may be understood via analyzing the calculation in a covariant gauge, as we explain in appendix G.2. In this way the two integral bases are not completely equivalent, since the latter one is exploiting more of the specific properties of the problem at hand.

2.7. Further ultraviolet properties

2.7.1. Quadratic and quartic divergence. Next, we want to analyze the quadratic and quartic divergence of our amplitude. In dimensional regularization, the signature of these divergences are poles around $\epsilon = 1$ and $\epsilon = 2$ in the final answer. Since we compute our amplitude to all order in ϵ we may probe these poles. We tackle this question in the basis that contains only $(4 - 2\epsilon)$-dimensional integrals. In the basis that includes higher-dimensional integrals, many of the integrals are quadratically or quartically divergent, which obscures the analysis. We discuss the double-minus configuration for a spin-0 particle in the loop; the other cases are similar. As we demonstrate, our amplitude has no quadratic or quartic divergences.

We start with the quadratic divergence. In the chosen basis the only quadratically divergent integral is the tadpole. However, the coefficient of the tadpole is linear in $(\epsilon - 1)$, hence there is no contribution to the quadratic divergence from it. It then suffices to expand all coefficients around $\epsilon = 1$ and only keep the divergent piece. We get

$$\mathcal{M}_4^{S=0} = -\frac{1}{(4\pi)^{2-\epsilon}} \left(\frac{\kappa}{2}\right)^2 \mathcal{M}_4^{\text{tree}} \frac{su}{(\epsilon - 1)t^6} (3s^2u^2 + 10m^2stu + 6m^4t^2)$$
$$\times \left(\iota_2(s) + \iota_2(u)\right) + \mathcal{O}\left((\epsilon - 1)^0\right), \tag{2.56}$$

where we introduced

$$\iota_2(s) = s \left(2m^2 I_3(s) - I_2(s)\right). \tag{2.57}$$

Note that $I_n \equiv I_n^{4-2\epsilon} = I_n^{2-2(\epsilon-1)}$. It appears that our amplitude has a quadratic divergence which is non local due to the $1/t^6$ factor. However, using equation (G11) with $D = 2 - 2(\epsilon - 1)$ we get

$$I_3^{4-2(\epsilon-1)}(s) = \frac{1}{2(\epsilon - 1)} \left(-2m^2 I_3^{2-2(\epsilon-1)}(s) + I_2^{2-2(\epsilon-1)}(s)\right), \tag{2.58}$$

which gives

$$\iota_2(s) = -2s(\epsilon - 1) I_3^{4-2(\epsilon-1)}(s). \tag{2.59}$$

Since $I_3^{4-2(\epsilon-1)}(s)$ is finite as $\epsilon \to 1$, we see that there is no quadratic divergence.

The analysis of the quartic divergence follows in a similar manner. In this basis there are no quartically-divergent integrals. We expand our result around $\epsilon = 2$ to get

$$\mathcal{M}_4^{S=0} = \frac{1}{(4\pi)^{2-\epsilon}} \left(\frac{\kappa}{2}\right)^2 \mathcal{M}_4^{\text{tree}} \frac{su}{(\epsilon - 2)t^5}(5su + 6m^2t)\left(\iota_4(s) + \iota_4(u)\right)$$
$$+ \mathcal{O}\left((\epsilon - 2)^0\right), \tag{2.60}$$

where

$$\iota_4(s) = 2sI_1 + s(s - 6m^2)I_2(s) + 4m^4sI_3(s). \tag{2.61}$$

Using equation (G11) two consecutive times we get

$$\iota_4(s) = -4s(\epsilon - 2)I_3^{4-2(\epsilon-2)}(s), \tag{2.62}$$

which shows that there is no quartic divergence.

In our problem we demonstrated that the coefficients of the integrals which have no unitarity cuts in equation (2.54) contain no m dependence. This property was crucial in order to fix them using the vanishing of the (logarithmic) UV divergence. In a more general situation we expect this property to no longer be true. In such a scenario, analyzing the higher UV divergences offers an alternative method for obtaining these coefficients. For example, if we demand the vanishing of the quadratic and quartic divergences along with the logarithmic one for our problem, we may fix the coefficients γ_i to the values we found above, without needing to impose that they do not contain m dependence.

2.7.2. Ultraviolet divergences in higher dimensions. We can also inspect the UV properties in higher dimensions. This is straightforward because we obtain expressions for the amplitudes valid to all orders in ϵ. In the calculations we keep the external kinematics and helicities fixed in four dimensions. In addition, we use the FDH scheme [28] which keeps the number of physical states circulating in the loop fixed at their four-dimensional values. However, we can still analytically continue the loop momentum to any dimension and study the divergence properties. We can use this as a rather nontrivial check on our expressions and also to point out a simple relation between appropriately defined divergences in higher dimensions and coefficients of terms in the large-mass expansion. We discuss the double-minus configuration with a scalar in the loop for concreteness, however our results hold for all cases considered in this paper.

Since our expressions are valid to all orders in the dimensional-regularization parameter ϵ, we may shift the spacetime dimension of the loop momentum by 2σ via $\epsilon \to \epsilon - \sigma$. For example for $\sigma = 1$ the spacetime dimension is shifted to $D = 6 - 2\epsilon$. In the shifted dimension we define the coefficients of the UV divergences as

$$\mathcal{M}_4^{S=0}\Big|_{\epsilon\to\epsilon-\sigma} = \frac{1}{(4\pi)^\sigma} \frac{1}{(\sigma - 1)!} \frac{1}{\epsilon} \sum_{n=0}^{\sigma-1} \left(m^2\right)^n \delta_{\sigma,n}^{\text{UV}}(s,t) + \mathcal{O}\left(\epsilon^0\right), \tag{2.63}$$

where we choose the normalization in a particular way to account for the angular loop-integration in different dimensions. Note that there is no UV divergence for $\sigma = 0$, corresponding to $D = 4 - 2\epsilon$, due to the lack of a corresponding counterterm matrix element. For $\sigma = 1$ the counterterm is an R^3-type operator. In this case, up to a numerical factor the coefficient of the divergence matches the tree amplitude generated from an R^3 insertion, which we describe in section 3. For convenience we restrict ourselves to even dimensions (for $\epsilon \to 0$) because in these cases the dimension-shifting formulas (equation (G11)) bring the higher-dimensional

integrals back to four dimensions. For $2 \leqslant \sigma \leqslant 6$ we have explicitly confirmed that the divergences are local and therefore correspond to appropriate derivatives of R^4-type operators, as expected. We note that because individual integrals have nonlocal coefficients the fact that the divergences are local provides a rather nontrivial check on our expressions. Furthermore, starting at $D = 8 - 2\epsilon$ all the integrals are divergent, so their coefficients feed into this check.

Finally, we point out a relation between the UV divergences in equation (2.63) and terms in the large-mass expansion in four dimensions. Specifically, defining

$$\mathcal{M}_4^{S=0} = \sum_{n=1}^{\infty} \frac{\delta_n^{\text{IR}}(s,t)}{m^{2n}}, \tag{2.64}$$

we have explicitly checked through $\sigma = 4$ or equivalently $D = 12 - 2\epsilon$ that

$$\delta_{\sigma,0}^{\text{UV}} = \delta_\sigma^{\text{IR}}. \tag{2.65}$$

Similarly, we find that $\delta_{\sigma,n}^{\text{UV}}$ is proportional to $\delta_{\sigma-n,0}^{\text{UV}} = \delta_{\sigma-n}^{\text{IR}}$ with a common proportionality constant for all multiplets. For this correspondence to hold it is important to keep both the external and internal states at their four dimensional value; only the loop momentum is analytically continued to higher spacetime dimensions. It is quite remarkable that such simple relations exist between the coefficients of the divergences in higher dimensions and the terms in the large-mass expansion of the amplitudes in four dimensions.

2.8. Consistency checks

We have carried out a variety of checks on our amplitudes. Basic self-consistency checks are that UV and infrared or mass singularities be of the right form.

UV singularities must be local. In general this is nontrivial and happens only after the UV singularities are combined. In our case, the coefficient of the divergences vanishes. We also verified that the $1/(\epsilon - 1)$ pole cancels, consistent with the fact that the Ricci-scalar counterterm vanishes by the equation of motion. Similarly for the $1/(\epsilon - 2)$ pole, the expression is not only local but it vanishes. Further, we verified that the divergences obtained by analytically continuing the loop momentum to higher dimensions while keeping the state counts and external kinematics to their four-dimensional values are also local.

Another nontrivial check comes from looking at the $m \to 0$ limit. Since the internal lines are massive, there is no IR divergence for our amplitude. However, we may regard the mass of the internal lines as an infrared regulator for the corresponding massless amplitude and study the IR divergence as $m \to 0$. In gravity the infrared singularities are quite simple since there are only soft singularities and no collinear or mass singularities [53]. The soft singularities arise only from gravitons circulating in the loop. This implies that as $m \to 0$ all contributions must be infrared finite except for the $\mathcal{M}_4^{\{2\}}$ piece (corresponding to $\mathcal{N} = 8$ supergravity in the massless limit), since this is the only piece that has a graviton circulating in the loop in this limit.

To carry out this check we start with the on-shell-constructable piece. We start from the $(4 - 2\epsilon)$-dimensional integral basis, where only the boxes and the triangles have IR singularities. The triangles only contain simple logarithms (see equation (G6)) while the boxes contain both logarithms and dilogarithms (see equation (G7)). To simplify the check we use equation (G11) to trade the boxes for triangles and $(6 - 2\epsilon)$-dimensional boxes. Importantly, the latter have no IR singularities. In this form the infrared singularities are all pushed into triangle integrals, hence it suffices to verify that their coefficients vanish as $m \to 0$, which we confirm for all but the $\mathcal{M}_4^{\{2\}}$ piece. Next we look at the contributions with no unitary cuts. These pieces are

zero if we take $m \to 0$ before we expand in ϵ. On the other hand, divergent terms appear if we first expand in ϵ and then take $m \to 0$. However, given that the UV divergence vanishes in four-dimensions these infrared divergent $\log(m)$ pieces also cancel among themselves. For the $\mathcal{M}_4^{\{2\}}$ piece there is indeed an infrared singularity that develops as $m \to 0$. In this case, we recover the known infrared divergence of pure Einstein gravity [53].

As another check, we have explicitly verified that as $m \to 0$ our massive results match the massless ones given in reference [19], up to an overall sign in $\mathcal{M}_4^{\{1/2\}}$, as noted in reference [54]. One simple way to implement this check is to start with the expressions for the amplitudes in the $(4 - 2\epsilon)$-dimensional integral basis, set $m \to 0$ in the integral coefficients, and replace the massive integrals with massless ones.

Finally, we note that contributions from individual integrals do not decay at large mass as required by decoupling, while the amplitudes have the required property. This involves nontrivial cancellation between the pieces, providing yet another check.

3. Amplitudes in the low-energy effective field theory

In this section we study four-graviton scattering in a general parity-even low-energy EFT. Such EFTs start from the Einstein–Hilbert action and extend to systematically include higher-dimension operators. We include a massless scalar field to our analysis, corresponding to the dilaton found in string theory.

We match this EFT to the one-loop amplitudes determined in section 2 and collected in appendix F. In this context, the EFT is valid for energies significantly smaller than the mass of the spinning particle in the loop. In this way we determine the modification to the low-energy theory of gravity due to the presence of a heavy particle. We take this as a nontrivial model of UV physics feeding into low-energy physics.

For the lowest-dimension operators, we calculate the four-point tree-level amplitudes in this EFT and compare them to the expansion of our one-loop amplitudes in the large-mass limit in order to obtain their Wilson coefficients. More generally, since we later put bounds on the coefficients appearing in the amplitudes themselves, there is no need to relate these back to a Lagrangian. For comparison to the bounds derived in subsequent sections we also present the one-loop scattering amplitudes expanded in the large-mass limit. In appendix H we present the expansions to much higher orders, which should be useful for further studies of the bounds.

Finally, we obtain the Regge limits of our one-loop amplitudes. These are useful later in analyzing low-energy coefficients via dispersion relations.

3.1. Setup of the effective field theory

The first few terms of the EFT describing low-energy gravitational scattering are

$$S_{\text{EFT}} = \int \mathrm{d}^4 x \sqrt{-g} \left[-\frac{2}{\kappa^2} R + \frac{1}{2} \partial_\mu \phi \partial^\mu \phi + \frac{2\beta_\phi}{\kappa^2} \phi\, C + \frac{8}{\kappa^3} \frac{\beta_{R^3}}{3!} R^3 \right.$$

$$\left. + \frac{2\beta_{R^4}}{\kappa^4} C^2 + \frac{2\tilde{\beta}_{R^4}}{\kappa^4} \tilde{C}^2 + \cdots \right], \tag{3.1}$$

where R is the Ricci scalar, κ is given in terms of Newton's constant G via $\kappa^2 = 32\pi G$, and the metric is $g_{\mu\nu} = \eta_{\mu\nu} + \kappa h_{\mu\nu}$ in terms of the graviton field $h_{\mu\nu}$. To describe the most general parity-even theory that captures low-energy four-graviton scattering, we include the massless scalar field ϕ. The factors of $\frac{1}{\kappa}$ are chosen to normalize the kinetic term canonically and to

remove the factors of κ that would appear in the three-point tree-level amplitudes built out of a single insertion of $\phi\,C$ or R^3 and the four-graviton tree-level amplitudes built out of a single insertion of C^2 or $\tilde{C}^2$. The β_ϕ, β_{R^3}, β_{R^4} and $\tilde{\beta}_{R^4}$ are Wilson coefficients that depend on the details of the UV physics. The composite operators are defined by

$$
\begin{aligned}
R^3 &\equiv R^{\mu\nu\kappa\lambda}R_{\kappa\lambda\alpha\gamma}R^{\alpha\gamma}{}_{\mu\nu}\,, \\
C &\equiv R^{\mu\nu\kappa\lambda}R_{\mu\nu\kappa\lambda}\,, \\
\tilde{C} &\equiv \frac{1}{2}R^{\mu\nu\alpha\beta}\epsilon_{\alpha\beta}{}^{\gamma\delta}R_{\gamma\delta\mu\nu},
\end{aligned}
\tag{3.2}
$$

where $R_{\lambda\mu\nu\kappa}$ is the Riemann tensor. One can systematically add higher-dimension operators; we choose not to do so here since our later analysis is at the amplitude level, so the mapping back to Lagrangian coefficients is not necessary.

In writing the effective action (3.1) we apply the equations of motion and integrate by parts to reduce the number of terms to a minimum independent set. In particular, in constructing the higher-dimensional operators we replace instances of the Ricci scalar and tensor, R and $R_{\mu\nu}$, with appropriate contractions of the matter stress–energy tensor. We drop such terms since they give rise to higher-point matter interactions, which do not affect our analysis. For example, we do not include R^2-type terms because the squares of R and $R_{\mu\nu}$ do not contribute due to the equations of motion, while the contraction of two Riemann tensors C can be traded for the Gauss–Bonnet contribution, which is equal to a total derivative in four dimensions [52]. Furthermore, we do not include operators that our calculation is not sensitive to. Specifically, we do not consider the other possible contraction of three Riemann tensors, since it does not contribute to four-graviton scattering [55, 56]. We restrict ourselves to parity-even interactions and neglect the parity-odd operators $\phi\,\tilde{C}$ and $C\tilde{C}$ [57, 58]. The possible parity-even contractions of four Riemann tensors were obtained in reference [59]. Recent studies that use similar Lagrangians are found in references [58, 60–62].

3.2. Scattering amplitudes in the effective field theory

To describe the amplitudes it is useful to extract overall dependence on the spinors, leaving only functions of s, u with simple crossing properties. Specifically, for the independent helicity configurations we define,

$$
\begin{aligned}
\mathcal{M}_4(1^+, 2^-, 3^-, 4^+) &= (\langle 23\rangle[14])^4\, f(s, u), \\
\mathcal{M}_4(1^+, 2^+, 3^-, 4^+) &= \big([12][14]\langle 13\rangle\big)^4\, g(s, u), \\
\mathcal{M}_4(1^+, 2^+, 3^+, 4^+) &= \left(\frac{[12][34]}{\langle 12\rangle\langle 34\rangle}\right)^2 h(s, u),
\end{aligned}
\tag{3.3}
$$

corresponding to the *double-minus*, *single-minus* and *all-plus* helicity configurations. As usual we do not include the overall i that normally would appear in Feynman diagrams. All other amplitudes are given by permutations and complex conjugation,

$$
\mathcal{M}_4(1^{h_1}, 2^{h_2}, 3^{h_3}, 4^{h_4}) = \mathcal{M}_4^*(1^{-h_1}, 2^{-h_2}, 3^{-h_3}, 4^{-h_4}),
\tag{3.4}
$$

where we define the complex conjugation to not act on the $i\epsilon$ prescription.[11] In parity-preserving theories complex conjugation acts only on the spinors swapping the positive and negative helicity spinors,

$$\mathcal{M}_4(1^{h_1}, 2^{h_2}, 3^{h_3}, 4^{h_4}) = \mathcal{M}_4(1^{-h_1}, 2^{-h_2}, 3^{-h_3}, 4^{-h_4})\big|_{\lambda_i \leftrightarrow \tilde{\lambda}_i}, \tag{3.5}$$

so that the f, g and h functions are unaltered.

In general relativity the leading-order results for the amplitudes above take the form

$$f_{\mathrm{GR}}(s,u) = \left(\frac{\kappa}{2}\right)^2 \frac{1}{stu} + \cdots, \tag{3.6}$$

$$g_{\mathrm{GR}}(s,u) = \frac{1}{(4\pi)^2}\left(\frac{\kappa}{2}\right)^4 \frac{1}{s^2 t^2 u^2} \frac{s^2 + t^2 + u^2}{360} + \cdots, \tag{3.7}$$

$$h_{\mathrm{GR}}(s,u) = -\frac{1}{(4\pi)^2}\left(\frac{\kappa}{2}\right)^4 \frac{s^2 + t^2 + u^2}{120} + \cdots \tag{3.8}$$

Below we do not consider loop effects in general relativity itself (due to gravitons circulating in the loops) but focus on the properties of the higher-derivative operators in the gravitational EFT generated by integrating out massive degrees of freedom. For the same reason IR divergences are not an issue for our analysis since the corrections of interest are manifestly IR finite.

Permutation symmetry of the amplitudes plays a crucial role in our analysis and manifests itself in the following crossing relations

$$f(s,u) = f(u,s), \tag{3.9}$$

$$g(s,u) = g(u,s) = g(s,t), \tag{3.10}$$

$$h(s,u) = h(u,s) = h(s,t), \tag{3.11}$$

where $s + t + u = 0$.

With our normalizations the three-point amplitude arising from the Einstein term is[12]

$$\mathcal{M}_3^{\mathrm{GR}}(1^+, 2^+, 3^-) = \frac{\kappa}{2}\left(\frac{[12]^3}{[23][31]}\right)^2. \tag{3.12}$$

The three-points amplitudes with an insertion of the $\phi\,C$ or the R^3 operator are [56]

$$\mathcal{M}_3^{\phi\,C}(1^+, 2^+, 3^\phi) = \beta_\phi[12]^4, \qquad \mathcal{M}_3^{R^3}(1^+, 2^+, 3^+) = \beta_{R^3}([12][23][31])^2, \tag{3.13}$$

where β_ϕ and β_{R^3} are the Wilson coefficient for these operators appearing in equation (3.1).

At four points $\phi\,C$ contributes to the double-minus and all-plus configurations,

$$\mathcal{M}_4^{\phi\,C}(1^+, 2^+, 3^+, 4^+) = -3(\beta_\phi)^2 stu\left(\frac{[12][34]}{\langle 12\rangle\langle 34\rangle}\right)^2,$$

$$\mathcal{M}_4^{\phi\,C}(1^+, 2^-, 3^-, 4^+) = -(\beta_\phi)^2 \frac{(\langle 23\rangle[14])^4}{t}. \tag{3.14}$$

---

[11] More precisely, given complex conjugation $\gamma : z \mapsto z^*$, we define $f^*(z)$ as $f^* \equiv \gamma \circ f \circ \gamma : z \mapsto (f(z^*))^*$. For identical scalar particles (3.4) becomes the familiar Hermitian analyticity.

[12] We implicitly use complex momenta so that the three-point amplitude does not vanish from kinematic constraints.

On the other hand, there are two independent helicity configurations, the all-plus and the single-minus configurations, that contain a single insertion of R^3. We obtain these amplitudes following reference [56]. We find

$$\mathcal{M}_4^{R^3}(1^+, 2^+, 3^-, 4^+) = \beta_{R^3} \left(\frac{\kappa}{2}\right) \frac{1}{stu} \left([12][14]\langle 13\rangle\right)^4, \tag{3.15}$$

and

$$\mathcal{M}_4^{R^3}(1^+, 2^+, 3^+, 4^+) = 10\beta_{R^3} \left(\frac{\kappa}{2}\right) stu \left(\frac{[12][34]}{\langle 12\rangle\langle 34\rangle}\right)^2, \tag{3.16}$$

which is slightly rearranged compared to reference [56]. We build a double-minus contribution out of two insertions of R^3 [2],

$$\mathcal{M}_4^{R^3}(1^+, 2^-, 3^-, 4^+) = (\beta_{R^3})^2 \frac{su}{t} \left(\langle 23\rangle[14]\right)^4. \tag{3.17}$$

For the R^4-type operators the amplitudes are [62]

$$\begin{aligned}
\mathcal{M}_4^{R^4}(1^+, 2^+, 3^+, 4^+) &= \beta_{R^4}^- \frac{(s^2 + t^2 + u^2)^2}{2} \left(\frac{[12][34]}{\langle 12\rangle\langle 34\rangle}\right)^2, \\
\mathcal{M}_4^{R^4}(1^+, 2^-, 3^-, 4^+) &= \beta_{R^4}^+ \left(\langle 23\rangle[14]\right)^4,
\end{aligned} \tag{3.18}$$

where by $\mathcal{M}_4^{R^4}$ we refer to the amplitudes build out of both C^2 and $\tilde{C}^2$, with

$$\beta_{R^4}^\pm \equiv \beta_{R^4} \pm \tilde{\beta}_{R^4}. \tag{3.19}$$

Using these four-point amplitudes we may extract the coefficients β_{R^3}, β_{R^4} and $\tilde{\beta}_{R^4}$ by matching to our one-loop calculation in the large-mass limit. Since we did not include the massless scalar field in the construction of our one-loop amplitudes, we have $\beta_\phi = 0$ in this case.

Next, we bring our one-loop amplitudes in a form suitable to compare to the EFT amplitudes listed above. We start with the double-minus amplitude. As usual, we organize the contributions according to the supersymmetric decomposition (2.32),

$$\mathcal{M}_4^{\{S\}}(1^+, 2^-, 3^-, 4^+) = (\langle 23\rangle[14])^4 \, f^{\{S\}}(s, u). \tag{3.20}$$

In the large-mass limit, for the double-minus amplitudes we have,

$$
\begin{aligned}
f^{\{0\}} = \mathcal{K} \Bigg(&\frac{1}{6300m^4} + \frac{t}{41\,580m^6} + \frac{81(s^2+u^2)+155su}{15\,135\,120m^8} \\
&+ \frac{\left(161(s^2+u^2)+324su\right)t}{151\,351\,200m^{10}} + \cdots \Bigg),
\end{aligned}
$$

$$
\begin{aligned}
f^{\{1/2\}} = \mathcal{K} \Bigg(&\frac{1}{1120m^4} + \frac{t}{8400m^6} + \frac{15(s^2+u^2)+28su}{554\,400m^8} \\
&+ \frac{\left(153(s^2+u^2)+313su\right)t}{30\,270\,240m^{10}} + \cdots \Bigg),
\end{aligned}
$$

$$
\begin{aligned}
f^{\{1\}} = \mathcal{K} \Bigg(&\frac{1}{180m^4} + \frac{t}{1680m^6} + \frac{22(s^2+u^2)+39su}{151\,200m^8} \\
&+ \frac{\left(20(s^2+u^2)+43su\right)t}{831\,600m^{10}} + \cdots \Bigg),
\end{aligned}
\qquad (3.21)
$$

$$
\begin{aligned}
f^{\{3/2\}} = \mathcal{K} \Bigg(&\frac{1}{24m^4} + \frac{t}{360m^6} + \frac{9(s^2+u^2)+14su}{10\,080m^8} \\
&+ \frac{\left(8(s^2+u^2)+21su\right)t}{75\,600m^{10}} + \cdots \Bigg),
\end{aligned}
$$

$$
f^{\{2\}} = \mathcal{K} \left(\frac{1}{2m^4} + \frac{s^2+su+u^2}{120m^8} + \frac{stu}{504m^{10}} + \cdots \right),
$$

where

$$
\mathcal{K} = \frac{1}{(4\pi)^2} \left(\frac{\kappa}{2} \right)^4. \qquad (3.22)
$$

For the all-plus and single-minus configurations it suffices to give the result for the spin-0 contribution since we obtain the remaining amplitudes via equation (2.34). For the single-minus configuration we have

$$
\mathcal{M}_4^{S=0}(1^+, 2^+, 3^-, 4^+) = \left([12]\langle 13\rangle[14] \right)^4 g(s,u), \qquad (3.23)
$$

where

$$
g(s,u) = \mathcal{K} \left(\frac{1}{5040m^2 stu} + \frac{1}{6306\,300m^8} + \frac{(s^2+su+u^2)}{441\,080\,640m^{12}} + \cdots \right). \qquad (3.24)
$$

Finally, for the all-plus configuration we have,

$$
\mathcal{M}_4^{S=0}(1^+, 2^+, 3^+, 4^+) = \left(\frac{[12][34]}{\langle 12\rangle\langle 34\rangle} \right)^2 h(s,u), \qquad (3.25)
$$

where

$$h(s, u) = \mathcal{K} \left(\frac{stu}{504m^2} + \frac{(s^2 + su + u^2)^2}{3780m^4} + \frac{(s^2 + su + u^2)stu}{7920m^6} \right.$$

$$+ \frac{75(s^6 + u^6) + 225(s^5 u + su^5) + 559(s^4 u^2 + s^2 u^4) + 743s^3 u^3}{7207\,200m^8}$$

$$\left. + \frac{3(s^2 + su + u^2)^2 stu}{400\,400m^{10}} + \cdots \right). \tag{3.26}$$

The fact that the highest power of m appearing in the large-mass expansion is m^{-2} may be contrasted to the high powers of the m in the coefficients of the integrals in the $(4 - 2\epsilon)$-dimensional integral basis (see equations (F26), (F31) and (F32)). It is a nontrivial consistency check that our amplitudes vanish in the large-mass limit, as expected from decoupling. Indeed, the above large-mass behavior hinges on nontrivial cancellations between all pieces of the amplitude.

Now consider the matching and extraction of the Wilson coefficients β_{R^3}, β_{R^4} and $\tilde{\beta}_{R^4}$ (or, equivalently, β_{R^3} and $\beta_{R^4}^{\pm}$). Since the relation between the Wilson coefficients and the amplitudes is linear, the Wilson coefficients satisfy the supersymmetric decomposition (equation (2.32)). Hence, we may organize our results in terms of the multiplets circulating in the loop. One may then assemble the corresponding coefficients for any spinning particle circulating in the loop using equation (2.32).

Since the all-plus and single-minus amplitudes are nonzero only for the $\{0\}$ piece, we have

$$(\beta_{R^3})^{\{0\}} = \frac{1}{(4\pi)^2} \left(\frac{\kappa}{2} \right)^3 \frac{1}{m^2} \frac{1}{5040},$$

$$(\beta_{R^4}^-)^{\{0\}} = \frac{1}{(4\pi)^2} \left(\frac{\kappa}{2} \right)^4 \frac{1}{m^4} \frac{1}{7560}, \tag{3.27}$$

$$(\beta_{R^3})^{\{S\} \neq 0} = (\beta_{R^4}^-)^{\{S\} \neq 0} = 0,$$

where our notation $(\beta_X)^{\{S\}}$ means the value of β_X as determined by the data for the new piece for a given spin S. The double-minus configuration is nonzero for any multiplet circulating in the loop. We find

$$(\beta_{R^4}^+)^{\{0\}} = \frac{1}{(4\pi)^2} \left(\frac{\kappa}{2} \right)^4 \frac{1}{m^4} \frac{1}{6300},$$

$$(\beta_{R^4}^+)^{\{1/2\}} = \frac{1}{(4\pi)^2} \left(\frac{\kappa}{2} \right)^4 \frac{1}{m^4} \frac{1}{1120},$$

$$(\beta_{R^4}^+)^{\{1\}} = \frac{1}{(4\pi)^2} \left(\frac{\kappa}{2} \right)^4 \frac{1}{m^4} \frac{1}{180}, \tag{3.28}$$

$$(\beta_{R^4}^+)^{\{3/2\}} = \frac{1}{(4\pi)^2} \left(\frac{\kappa}{2} \right)^4 \frac{1}{m^4} \frac{1}{24},$$

$$(\beta_{R^4}^+)^{\{2\}} = \frac{1}{(4\pi)^2} \left(\frac{\kappa}{2} \right)^4 \frac{1}{m^4} \frac{1}{2}.$$

3.3. Regge limits of the amplitudes

For our analysis of the amplitudes with dispersion relations in the next sections, we need the behavior of the amplitudes for $t \to \infty$ and $s \in \mathbb{R}$ fixed, with $|s| < 4m^2$ and for $s \to \infty$ and $t \in \mathbb{R}$

fixed, with $|t| < 4m^2$. We extract these directly from the explicit values of the amplitudes in appendix H.

We start with the $t \to \infty$ limit. For the new-pieces in the supersymmetric decomposition we have

$$f^{\{S\}}(s, u) \sim \frac{\log(t)}{t^2}, \quad S = 0, \ 1/2, \ 1,$$

$$f^{\{3/2\}}(s, u) \sim \frac{\log^2(t)}{t^2}, \qquad f^{\{2\}}(s, u) \sim \frac{1}{t}, \tag{3.29}$$

$$g^{\{0\}}(s, u) \sim \frac{1}{t^2}, \qquad h^{\{0\}}(s, u) \sim t^2,$$

where the f, g, and h function are related to the amplitude via equation (3.3). Using equation (2.32) we assemble the contributions for each particle of a given spin,

$$f^S(s, u) \sim \frac{\log(t)}{t^2}, \quad S = 0, \ 1/2, \ 1,$$

$$f^{3/2}(s, u) \sim \frac{\log^2(t)}{t^2}, \qquad f^2(s, u) \sim \frac{1}{t}, \tag{3.30}$$

$$g^S(s, u) \sim \frac{1}{t^2}, \qquad h^S(s, u) \sim t^2, \quad S = 0, \ 1/2, \ 1, \ 3/2, \ 2.$$

For the f^S functions corresponding to the case of no helicity flips between incoming and incoming states the spin 2 contribution dominates as expected.

Next, we consider the $s \to \infty$ limit. We find

$$f^{\{S\}}(s, u) \sim \frac{1}{s}, \quad S = 0, \ 1/2, \ 1, \ 3/2, \ 2,$$

$$g^{\{0\}}(s, u) \sim \frac{1}{s^2}, \qquad h^{\{0\}}(s, u) \sim s^2, \tag{3.31}$$

which gives

$$f^S(s, u) \sim \frac{1}{s}, \qquad g^S(s, u) \sim \frac{1}{s^2}, \qquad h^S(s, u) \sim s^2, \quad S = 0, \ 1/2, \ 1, \ 3/2, \ 2. \tag{3.32}$$

Note that the limits of the functions g^S and h^S in equations (3.31) and (3.32) follow from those in equations (3.29) and (3.30), since these functions are crossing symmetric.

4. Properties of gravitational amplitudes

We now turn to the properties of the low-energy effective field theory, arising from taking the large-mass expansion of the one-loop four-graviton amplitudes calculated in section 2. Here we do not consider loop effects in general relativity itself (due to gravitons circulating in the loops) but focus on the properties of the leading-order higher-derivative operators in a weakly-

coupled gravitational EFT generated by integrating out massive degrees of freedom. For the same reason IR divergences are not an issue for our analysis since the corrections of interest are manifestly IR finite. We also note that we do not need to deal with UV divergences, since the one-loop four graviton amplitude with polarization tensors restricted to four dimensions considered here is UV finite [52].

To make the analysis more complete we also include the example of tree-level graviton scattering in string theory (see appendix B). As recently discussed in [63], in this case the scattering amplitudes have a great degree of universality; e.g. to leading order considered here they do not depend on the details of the string compactification.

A general question we ask in this paper is the following: where do physical theories land in the space of couplings that satisfy the bounds from causality, unitarity and crossing? In this section we review general properties of the gravitational amplitudes relevant for the derivation of the bounds. In subsequent sections we then proceed with the derivation of various bounds on the Wilson coefficients, following the recent developments of references [3–6, 64]. Finally, we check that the results presented in the paper satisfy the expected bounds and analyze the region in the space of couplings covered by known theories.

4.1. Low-energy expansion

The functions f, g and h defined in equation (3.3) correspond to the independent helicity configurations. We consider their low-energy expansion[13]

$$
f(s, u) = \left(\frac{\kappa}{2}\right)^2 \frac{1}{stu} + |\beta_{R^3}|^2 \frac{su}{t} - |\beta_\phi|^2 \frac{1}{t} + \sum_{i=0}^{\infty} f_{2i,i} s^i u^i
$$

$$
+ \sum_{i=1}^{\infty} \sum_{j=0}^{[\frac{i}{2}]} f_{i,j}(s^{i-j}u^j + s^j u^{i-j}),
$$

(4.1)

$$
g(s, u) = \left(\frac{\kappa}{2}\right) \frac{\beta_{R^3}}{stu} + \sum_{p,q=0}^{\infty} g_{2p+3q,q} \sigma_2^p \sigma_3^q,
$$

$$
h(s, u) = \left[10 \left(\frac{\kappa}{2}\right) \beta_{R^3} - 3\beta_\phi^2\right] stu + \sum_{p,q=0, 2p+3q \geqslant 4}^{\infty} h_{2p+3q,q} \sigma_2^p \sigma_3^q,
$$

where $[x]$ means the integer part of x, and we introduced

$$
\sigma_k \equiv (-1)^k \frac{s^k + t^k + u^k}{k}.
$$

(4.2)

In equation (4.1) we explicitly write the massless exchange poles and assume that the rest of the amplitude admits a simple low-energy expansion. This is a structure expected from integrating out massive degrees of freedom and indeed, the amplitudes analyzed in the present paper are of this type. It does not however capture correctly the structure of the amplitude once the loops involving massless particles are included. For example, consider the case of the one-loop correction due to gravitons circulating in the loop; see equation (3.6) and appendix F.4. In

---

[13] Here we allowed for parity-odd and parity-violating effects which render certain coefficients in the low-energy expansion complex.

this case, we see that for the all-plus amplitude one-loop Einstein gravity generates a non-zero $h_{2,0}$. Similarly, for the single-minus amplitude $g(s,u)$ the one-loop result in Einstein gravity has poles in each of the Mandelstam variables. Finally, the double-minus amplitude $f(s,u)$, see equation (F35), contains IR divergences, logarithms, as well as higher-order singularities in $1/t$ compared to the formulas above. In the present paper we focus on the effect produced by integrating out massive degrees of freedom and do not analyze the effects from one-loop massless exchanges. Mathematically, this is simply due to the fact that at leading order that we are interested in, the two effects lead to additive contributions to the scattering amplitude and can be analyzed separately. Moreover, the corrections to graviton scattering due to integrating out massive degrees of freedom satisfy all the basic properties that we discuss later in the section and as such the low-energy expansion generated in this way satisfies consistency bounds. Physically, integrating out massive degrees of freedom leads to effects that are localized in the impact parameter space $b \lesssim 1/m_{\text{gap}}$, which are encoded via higher-derivative operators in a gravitational EFT, whereas the one-loop effect due to graviton exchange contributes at any impact parameter, which also manifests itself through the fact that such corrections do not admit the representation (4.1). It would be very interesting to develop a systematic and unified approach to treat both effects in the context of gravitational scattering, but this is beyond the scope of the present paper and we leave it for future work.

In writing equation (4.1) we take into account the crossing relations (3.9). Note that $\sigma_3 = -stu$ for $u = -s - t$. In the expansion (4.1) $f_{i,j}$ are real. For parity-preserving theories $h_{k,j}$ and $g_{k,j}$ are real as well. In the formulas above β_{R^3} encodes the unique non-minimal correction to the three-point amplitude of gravitons as defined in equation (3.13). Similarly β_ϕ encodes the non-minimal coupling of a scalar to two gravitons. In parity-preserving theories these are real.

For completeness we take into account the possibility of non-minimal coupling to a massless scalar in the amplitude above with the three-point amplitude given in equation (3.13). Curiously the non-minimal correction to the three-point graviton amplitude and the non-minimal coupling to a massless scalar mix in the first term in the low-energy expansion of $h(s,u)$ (4.1). This point was discussed in detail in reference [56].

We list the explicit results for these functions obtained for the amplitudes considered in the present paper as an expansion in large mass in equations (3.21), (3.24) and (3.26), as well as in appendix H. The exact form of the amplitudes is found in appendix F.

4.2. Unitarity constraints

Here we consider the constraints that arise from unitarity. Since the gravitational EFTs of interest are weakly coupled, we limit ourselves to perturbative unitarity as discussed for example in references [5, 50]. It would be of course be very interesting to implement unitarity nonperturbatively along the lines of references [65, 66]. In four dimensions this would also require understanding the constraints of unitarity at the level of IR-finite observables; see e.g. references [67, 68] for a recent discussion. Luckily, for our purpose of investigating the leading-order corrections to the gravitational EFT these subtle but important issues are irrelevant. Here we primarily follow the discussion of reference [5] but modify it to account for differing helicity configurations.

To discuss the constraints coming from unitarity we note that the general incoming two-graviton state is a superposition of amplitudes with different helicities. In total there are four choices for the incoming and four choices for the outgoing states. We therefore consider the

following matrix of possible amplitudes

$$\begin{pmatrix} \mathcal{M}_{(+,-,-,+)} & \mathcal{M}_{(+,-,-,-)} & \mathcal{M}_{(-,-,-,+)} & \mathcal{M}_{(-,-,-,-)} \\ \mathcal{M}_{(+,-,+,+)} & \mathcal{M}_{(+,-,+,-)} & \mathcal{M}_{(-,-,+,+)} & \mathcal{M}_{(-,-,+,-)} \\ \mathcal{M}_{(+,+,-,+)} & \mathcal{M}_{(+,+,-,-)} & \mathcal{M}_{(-,+,-,+)} & \mathcal{M}_{(-,+,-,-)} \\ \mathcal{M}_{(+,+,+,+)} & \mathcal{M}_{(+,+,+,-)} & \mathcal{M}_{(-,+,+,+)} & \mathcal{M}_{(-,+,+,-)} \end{pmatrix}, \tag{4.3}$$

where we labeled the helicities of the gravitons using an all-incoming convention.

Consider now scattering in the physical t-channel $14 \to 23$. To describe this situation we consider the center-of-mass frame and choose helicity spinors as follows [5, 50]

$$\lambda_1 = t^{1/4} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \lambda_4 = t^{1/4} \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

$$\lambda_2 = it^{1/4} \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{pmatrix}, \qquad \lambda_3 = it^{1/4} \begin{pmatrix} \sin\frac{\theta}{2} \\ -\cos\frac{\theta}{2} \end{pmatrix}. \tag{4.4}$$

Since we consider particles 1 and 4 to be incoming we take $\tilde{\lambda}_{1,4} = \lambda_{1,4}^*$. Particles 2 and 3 are outgoing, therefore $\tilde{\lambda}_{2,3} = -\lambda_{2,3}^*$. With this choice we get $t = \langle 14 \rangle [14]$, $s = \langle 12 \rangle [12] = -t\sin^2\frac{\theta}{2}$, so that $\cos\theta = 1 + \frac{2s}{t}$. Evaluating the matrix (4.3) for this kinematics we obtain

$$\mathcal{M}(s,t) \equiv \begin{pmatrix} t^4 f(s,u) & s^2 t^2 u^2 g^*(s,u) & s^2 t^2 u^2 g^*(s,u) & h^*(s,u) \\ s^2 t^2 u^2 g(s,u) & u^4 f(s,t) & s^4 f(t,u) & s^2 t^2 u^2 g^*(s,u) \\ s^2 t^2 u^2 g(s,u) & s^4 f(t,u) & u^4 f(s,t) & s^2 t^2 u^2 g^*(s,u) \\ h(s,u) & s^2 t^2 u^2 g(s,u) & s^2 t^2 u^2 g(s,u) & t^4 f(s,u) \end{pmatrix}. \tag{4.5}$$

Unitarity restricts the form of the discontinuity of various amplitudes. We introduce the t-channel discontinuity as

$$\text{Disc}_t \, \mathcal{M}(s,t) \equiv \frac{\mathcal{M}(s, t+i\epsilon) - \mathcal{M}(s, t-i\epsilon)}{2i}. \tag{4.6}$$

Through the optical theorem or unitarity, the discontinuity (4.6) is related to the square of the $2 \to n$ amplitude where we insert a complete set of intermediate states. It is convenient to decompose intermediate states into the irreducible representation of the Poincare group, which are therefore labeled by the total energy $\sqrt{t}$, spin J and potentially other quantum numbers i. The simplest example is when we have an exchange by a single particle of mass $\sqrt{t}$ and spin J. In this case (4.6) produces the square of the corresponding three-point couplings multiplied by a kinematical polynomial.[14] A slightly more general situation is when we have an exchange by multiple particles of the same mass and spin. In this case the discontinuity produces (4.6) a sum over the products of the corresponding three-point couplings. Finally, we can also have a multi-particle state as an intermediate state of total energy $\sqrt{t}$ and spin J. It is convenient to

---

[14] In the case of external scalars these are familiar Legendre polynomials. For spinning particles the analogous polynomials in $D = 4$ are Wigner d-functions, which are written explicitly below.

think about it again as a single-particle state with the continuous label for the species (which corresponds to the distribution of the total energy among the constituent particles). The result is always the same: we can write equation (4.6) as a sum of kinematical polynomials multiplied by various *spectral densities* $\rho_J(t)$ which encode the sum over the products of couplings to the intermediate states of energy $\sqrt{t}$ and J in a given theory. What we just said is simply a restatement of the standard partial-wave expansion in a language which is perhaps slightly more intuitive. For a more complete and detailed derivation see references [65, 66]. In this way we can write the discontinuity in terms of the spectral density. For example,

$$
\text{Disc}_t \, \mathcal{M}_{(+,-,-,+)}(s,t) = \sum_{J=0}^{\infty} \frac{1+(-1)^J}{2} \rho_J^{++}(t) d_{0,0}^J(x). \tag{4.7}
$$

Applying this to all helicity configurations in the matrix of amplitudes (4.5) we get

$$
\begin{aligned}
\text{Disc}_t \, \mathcal{M}(s,t) = &\sum_{J=0}^{\infty} \frac{1+(-1)^J}{2}
\begin{pmatrix}
\rho_J^{++}(t) d_{0,0}^J(x) & 0 & 0 & \left[\tilde{\rho}_J^{++}(t)\right]^* d_{0,0}^J(x) \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
\tilde{\rho}_J^{++}(t) d_{0,0}^J(x) & 0 & 0 & \rho_J^{++}(t) d_{0,0}^J(x)
\end{pmatrix} \\
&+ \sum_{J=4}^{\infty}
\begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & \rho_J^{+-}(t) d_{4,4}^J(x) & \rho_J^{+-}(t)(-1)^J d_{4,-4}^J(x) & 0 \\
0 & \rho_J^{+-}(t)(-1)^J d_{4,-4}^J(x) & \rho_J^{+-}(t) d_{4,4}^J(x) & 0 \\
0 & 0 & 0 & 0
\end{pmatrix} \\
&+ \sum_{J=4}^{\infty} \frac{1+(-1)^J}{2}
\begin{pmatrix}
0 & [\tilde{\rho}_J^{+-}(t)]^* d_{4,0}^J(x) & [\tilde{\rho}_J^{+-}(t)]^* d_{4,0}^J(x) & 0 \\
\tilde{\rho}_J^{+-}(t) d_{4,0}^J(x) & 0 & 0 & [\tilde{\rho}_J^{+-}(t)]^* d_{4,0}^J(x) \\
\tilde{\rho}_J^{+-}(t) d_{4,0}^J(x) & 0 & 0 & [\tilde{\rho}_J^{+-}(t)]^* d_{4,0}^J(x) \\
0 & \tilde{\rho}_J^{+-}(t) d_{4,0}^J(x) & \tilde{\rho}_J^{+-}(t) d_{4,0}^J(x) & 0
\end{pmatrix},
\end{aligned}
\tag{4.8}
$$

where we introduced $x \equiv \cos\theta = 1 + \frac{2s}{t}$ and we used Wigner d-matrices $d_{\lambda,\lambda'}^J(x)$ which we list explicitly in appendix E. The formula above can be derived by explicitly analyzing the effect of an exchange by a particle, or equivalently an irreducible representation, of given mass $\sqrt{t}$ and spin J. The relevant three-point amplitudes are fixed up to a number and their products are encoded in the various spectral densities $\rho_J(m^2)$. The indices $++$ and $+-$ denote the helicities of the corresponding incoming gravitons.

To discuss dispersion relations we also need to understand the properties of the u-channel discontinuity which is defined as

$$
\text{Disc}_u \, \mathcal{M}(s,t) \equiv \frac{\mathcal{M}(s,-s-t-i\epsilon) - \mathcal{M}(s,-s-t+i\epsilon)}{2i}. \tag{4.9}
$$

By the same argument it takes the following form

$$
\mathrm{Disc}_u\,\mathcal{M}(s,t) = \sum_{J=0}^{\infty} \frac{1+(-1)^J}{2}
\begin{pmatrix}
0 & 0 & 0 & \left[\tilde{\rho}_J^{++}(t)\right]^* d_{0,0}^J(x) \\
0 & \rho_J^{++}(t)d_{0,0}^J(x) & 0 & 0 \\
0 & 0 & \rho_J^{++}(t)d_{0,0}^J(x) & 0 \\
\tilde{\rho}_J^{++}(t)d_{0,0}^J(x) & 0 & 0 & 0
\end{pmatrix}
$$

$$
+ \sum_{J=4}^{\infty}
\begin{pmatrix}
\rho_J^{+-}(t)d_{4,4}^J(x) & 0 & 0 & 0 \\
0 & 0 & \rho_J^{+-}(t)(-1)^J d_{4,-4}^J(x) & 0 \\
0 & \rho_J^{+-}(t)(-1)^J d_{4,-4}^J(x) & 0 & 0 \\
0 & 0 & 0 & \rho_J^{+-}(t)d_{4,4}^J(x)
\end{pmatrix}
$$

$$
+ \sum_{J=4}^{\infty} \frac{1+(-1)^J}{2}
\begin{pmatrix}
0 & [\tilde{\rho}_J^{+-}(t)]^* d_{4,0}^J(x) & [\tilde{\rho}_J^{+-}(t)]^* d_{4,0}^J(x) & 0 \\
\tilde{\rho}_J^{+-}(t)d_{4,0}^J(x) & 0 & 0 & [\tilde{\rho}_J^{+-}(t)]^* d_{4,0}^J(x) \\
\tilde{\rho}_J^{+-}(t)d_{4,0}^J(x) & 0 & 0 & [\tilde{\rho}_J^{+-}(t)]^* d_{4,0}^J(x) \\
0 & \tilde{\rho}_J^{+-}(t)d_{4,0}^J(x) & \tilde{\rho}_J^{+-}(t)d_{4,0}^J(x) & 0
\end{pmatrix},
$$

$$(4.10)$$

where the only effect of crossing is the different form of the diagonal matrix elements.

The spectral densities that enter the unitarity relation are given in terms of the product of the couplings to the corresponding intermediate states. The diagonal terms being given by the absolute value square of the couplings are nonnegative

$$
\rho_J^{++}(t) \geqslant 0, \quad \rho_J^{+-}(t) \geqslant 0. \tag{4.11}
$$

The off-diagonal terms satisfy simple Cauchy–Schwartz inequalities[15]

$$
|\tilde{\rho}_J^{+-}(t)|^2 \leqslant \rho_J^{++}(t)\rho_J^{+-}(t), \quad |\tilde{\rho}_J^{++}(t)| \leqslant \rho_J^{++}(t). \tag{4.12}
$$

Below we discuss bounds on the possible form of the low-energy coefficients stemming from unitarity and growth of the amplitude at infinity. In particular, it is convenient for us to consider combinations of amplitudes that have nonnegative semi-definite discontinuities both in the *t*- and *u*-channel. We consider the following convenient choices that are sufficient for our purposes

$$
\mathcal{M}_h(s,t) \equiv \begin{pmatrix} t^4 f(s,u) + u^4 f(s,t) - s^4 f(t,u) & h^*(s,u) \\ h(s,u) & t^4 f(s,u) + u^4 f(s,t) - s^4 f(t,u) \end{pmatrix}, \tag{4.13}
$$

$$
\mathcal{M}_g(s,t) \equiv \begin{pmatrix} t^4 f(s,u) + u^4 f(s,t) & 2s^2 t^2 u^2 g^*(s,u) \\ 2s^2 t^2 u^2 g(s,u) & t^4 f(s,u) + u^4 f(s,t) \end{pmatrix}. \tag{4.14}
$$

These matrices are *t*–*u* crossing symmetric

$$
\mathcal{M}_{h,g}(s,t) = \mathcal{M}_{h,g}(s,u). \tag{4.15}
$$

Using the form of the discontinuity dictated by unitarity above together with the Cauchy–Schwartz inequalities (4.12) one can check that

$$
\partial_s^n \mathrm{Disc}_t\,\mathcal{M}_{h,g}|_{s=0} = \partial_s^n \mathrm{Disc}_u\,\mathcal{M}_{h,g}|_{s=0} \succeq 0, \quad n \geqslant 0, \tag{4.16}
$$

---

[15] In terms of couplings to the intermediate state of energy $\sqrt{t}$ and spin J these inequalities simply state that $\left|\sum_i \lambda_{++i}\lambda_{+-i}^*\right|^2 \leqslant (\sum_i |\lambda_{++i}|^2)(\sum_i |\lambda_{+-i}|^2)$ and $|\sum_i \lambda_{++i}^2| \leqslant \sum_i |\lambda_{++i}|^2$ respectively.

where the notation '$\succeq 0$' means that the matrix is positive semi-definite.[16] We obtain these inequalities from the following properties of the Wigner d-matrices together with (4.12)

$$\partial_s^n d_{0,0}^J(x)|_{s=0} \geqslant 0, \tag{4.17}$$

$$\partial_s^n \left(d_{4,4}^J(x) - (-1)^J d_{4,-4}^J(x) \right)|_{s=0} \geqslant 0, \tag{4.18}$$

$$\partial_s^n \left(\rho_J^{++}(t) d_{0,0}^J(x) + \rho_J^{+-}(t) d_{4,4}^J(x) - 2|\tilde{\rho}_J^{+-}(t)| d_{4,0}^J(x) \right)|_{s=0} \geqslant 0, \tag{4.19}$$

where $s = 0$ corresponds to $x = 1$ (see definition of x below equation (4.8)). The first property (4.17) is a well-known property of Legendre polynomials. We have not attempted to prove (4.18) and (4.19) that rely on the properties of the relevant $d_{\lambda,\lambda'}^J(x)$, however we checked them explicitly up to $J = 30$. We also checked using formulas from appendix E the conditions above for any J and $n = 0, 1, 2$; these are the cases that we consider below in more detail. For forward scattering a closely related discussion can be found in reference [17].

The matrices above admit simple eigenvectors in the parity-preserving case, where we have $g(s, u) = g^*(s, u)$ and $h(s, u) = h^*(s, u)$. In this case the matrices above have eigenvalues

$$\mathcal{M}_{h,\pm}(s, t) = u^4 f(s, t) + t^4 f(s, u) - s^4 f(t, u) \pm h(s, u),$$

$$\mathcal{M}_{g,\pm}(s, t) = u^4 f(s, t) + t^4 f(s, u) \pm 2s^2 t^2 u^2 g(s, u). \tag{4.20}$$

Unitarity (4.16) then becomes the statement about nonnegativity of the discontinuity of equation (4.20). In this case we can apply dispersion relations directly to the functions (4.20).

4.3. Causality

Causality and unitarity put constraints on the high-energy behavior of the amplitude. The corresponding bounds are well known in the case of gapped QFTs [69], but are on a less-rigorous footing in gravitational theories.

At tree level the gravitational amplitudes are expected to satisfy

$$\lim_{|t| \to \infty} |\mathcal{M}_{\text{tree}}(s, t)| \leqslant t^2, \quad s < 0. \tag{4.21}$$

This result is very intuitive but hard to establish rigorously. It naturally emerges from various considerations [11, 70, 71].

The situation is less clear nonperturbatively, however the simple qualitative picture of high-energy scattering in gravity together with unitarity and causality again imply that a similar bound exists. The bound is usually assumed to be

$$\lim_{|t| \to \infty} |\mathcal{M}_{\text{full}}(s, t)| < t^2, \quad s < 0. \tag{4.22}$$

This can be used to write dispersion relations in $D \geqslant 5$ where the amplitudes are IR finite and make sense nonperturbatively. In $D = 4$ the situation is less clear due to the IR divergences, but presumably a similar bound exists for the IR safe observables. It is also possible to satisfy (4.22) at tree level but this requires an infinite number of particles of arbitrary high spin to be exchanged. A famous example of this type is given by the tree-level string amplitudes.

In this paper, we are interested in particular in the properties of the one-loop scattering amplitudes in $D = 4$. These do not have to satisfy (4.21) or (4.22), therefore we do not use

---

[16] A Hermitian $n \times n$ matrix M is called positive semi-definite if $z^* M z \geqslant 0$ for all $z \in \mathbb{C}^n$.

them. Indeed, let us imagine that the bound (4.21) is saturated at tree-level. This corresponds to exchange of a particle of spin 2. When we go to one-loop we can exchange a pair of spin-2 particles and the resulting amplitude will grow like t^3. Therefore it is natural that the following bound holds at one loop

$$\lim_{|t|\to\infty} |\mathcal{M}_{1-\text{loop}}(s,t)| \leqslant t^3, \quad s < 0. \tag{4.23}$$

The explicit amplitudes computed in the paper indeed satisfy equation (4.23). Below, in writing dispersion relations we assume the Regge bound to be as given in equation (4.23) and choose the number of subtractions accordingly. Note that since equation (4.23) is weaker than equation (4.21), bounds derived in this way apply both to the tree and one-loop amplitudes. They also apply to full nonperturbative amplitudes in $D \geqslant 5$.

It is a special feature of gravitational amplitudes that an infinite number of particles has to be exchanged at tree-level in order for the amplitudes to satisfy (4.21). This follows from the fact that such exchanges at tree-level correspond to non-minimal couplings, which are known to lead to violations of equation (4.21) [11, 13, 14]. The only exception to this rule is general relativity, which corresponds to the minimal self-coupling of the graviton.

The tree-level Regge growth bound (4.21) is the reason why it is so hard to construct tree-level gravitational amplitudes different from the ones of general relativity. This famous achievement of string theory leads to causal and unitarity amplitudes with infinitely many poles that corresponds to exchanges of infinitely many particles of arbitrarily large spin. In fact, every such modification must contain strings [72].[17]

4.4. Dispersive sum rules

Given the gravitational amplitudes that satisfy unitarity and the Regge bounds we can consider various dispersion relations. One class of dispersion relations that we find useful recasts the vanishing of the (subtracted) amplitude in the Regge limit in terms of its discontinuity.

Such relations are known as superconvergence relations [14], or dispersive sum rules [16]. We follow the latter terminology and consider the following integrals

$$B_k^+(s) = \oint_\infty \frac{\mathrm{d}t}{2\pi i} \mathcal{M}(s,t) \frac{1}{t} \frac{1}{(t(s+t))^k} = 0, \quad k \geqslant 2, \tag{4.24}$$

where the condition $k \geqslant 2$ originates from the Regge bound (4.23) and guarantees that the arc at infinity produces a vanishing contribution.

By deforming the contour we get the following formula

$$B_k^+(s) : \oint_{t_0} \frac{\mathrm{d}t}{2\pi i} \mathcal{M}(s,t) \frac{1}{t} \frac{1}{(t(s+t))^k}$$

$$= \int_{m_{\text{gap}}^2}^\infty \frac{\mathrm{d}t}{\pi} \frac{1}{(t(s+t))^k} \left(\frac{1}{t} \text{Disc}_t \, \mathcal{M}(s,t) + \frac{1}{s+t} \text{Disc}_u \, \mathcal{M}(s,t) \right), \tag{4.25}$$

where the integrals are depicted in figure 7.

We evaluate the lhs of equation (4.25) using the low-energy expansion, and we use the rhs to establish some nontrivial properties that the coefficients of the low-energy expansion should satisfy. More precisely, we consider the expansion of (4.25) around $s = 0$.

---

[17] Curiously, as explained in some detail in appendix D, if we allow for accumulation points in the spectrum we can consider much simpler amplitude functions. These functions, however, contain an infinite number of particles of a given mass.
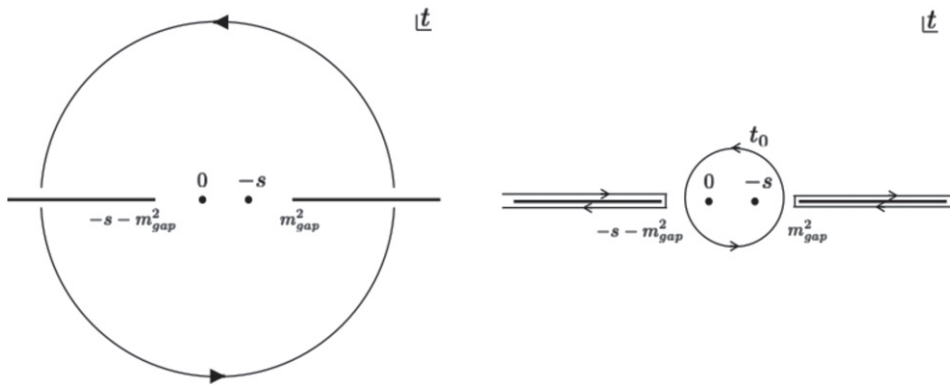
Figure 7. The vanishing of the arc integral at infinity on the left panel, cf (4.24), can be restated as an equality of the integrals on the right panel (see (4.25).)

A few comments are in order. Firstly, note that the minimal tree-level graviton-exchange term in $f(s,u) \sim \frac{1}{stu}$ does not contribute to the sum rules of interest here, $B^+_{k \geqslant 2}(s)$. In particular, we are able to expand the sum rules in powers of s. Assuming a more stringent Regge bound, $B^+_1(s)$ can be studied [7] and used to bound various coefficients in terms of G. We do not consider this sum rule in the present paper. Secondly, the non-minimal coupling β_{R^3} contributes to B^+_2 and we utilize this fact in section 6.2 in order to bound β_{R^3} in terms of other EFT data.

4.5. The theory islands

What is the space of gravitational EFTs that admit a consistent UV completion? Answering this question nonperturbatively is a formidable task, but we can consider the simpler question of understanding perturbatively in G possible consistent UV completions of gravity. It is in this spirit that in this paper we focus on the leading-order-in-G correction to general relativity.

Imagine that we label all possible perturbatively consistent theories of gravity by an index i so that the four-graviton scattering amplitudes in a given theory are given by the matrix $\mathcal{M}_i$ (see equation (4.3)).[18] It is clear then that by taking a superposition of such amplitudes with non-negative coefficients we again get an amplitude that satisfies all the constraints discussed in the sections above

$$\mathcal{M}_{\text{theory}} = \sum_i c_i \mathcal{M}_i, \quad c_i \geqslant 0. \tag{4.26}$$

By expanding such amplitudes at low energies, as in section 4.1, we get a set of higher-derivative corrections. The low-energy Wilson coefficients then form a convex cone which we can think of as the *theory space*,[19] which is generated by vectors $\mathcal{M}_i$.[20]

It has been recently established that causality, crossing and unitarity constrain Wilson coefficients both from above and from below [3–6]. In the space of couplings formed by

---

[18] For simplicity we keep the index discrete, but of course it can be continuous, e.g. denoting the mass of the particle exchanged in the loop. It does not affect the discussion.

[19] A convex cone is a subset of vector space closed under linear combinations with positive coefficients.

[20] Nonperturbatively in G, equation (4.26) is not necessarily consistent since is can violate nonperturbative unitarity $|S_J| \leqslant 1$. Since we are interested in the leading-order correction in the regime when gravity is weakly coupled, this issue is irrelevant.

weakly-coupled theories it is thus natural to talk about *the theory island*, namely the region spanned by known perturbative UV completions of gravity. By definition we have

$$\text{Theory island} \subseteq \text{EFThedron}, \tag{4.27}$$

where by the EFThedron we call the set of amplitudes satisfying the constraints reviewed in this section (in particular we only impose perturbative unitarity). It is interesting to ask to what extent we can populate the space of allowed couplings by known perturbative UV completions. Recall that by a perturbative UV completion we call an *S*-matrix that satisfies unitarity, causality and crossing perturbatively in Newton's constant *G* and for any process $m \to n$.

There are two classes of perturbative UV completions we consider here: tree-level corrected theories and one-loop corrected theories. In tree-level corrected theories the leading-order correction to general relativity enters at tree-level, in other words the higher-derivative corrections in the Lagrangian are suppressed by a new scale, e.g. string scale, which can be much lower than the Planck scale. Perturbative string theories are famous examples of this type. It is a well-known fact that gravitational amplitudes in string theory have a great degree of universality, as discussed recently in reference [63]. In appendix B we review the cases of superstring, heterotic and bosonic strings at tree level.[21] Alternatively, we can consider theories where the higher-derivative operators come with an extra power of *G*. The one-loop amplitudes computed in the present paper are precisely of this type. By choosing the mass spectrum and spins of the particles propagating in the loop we can get various amplitudes.

Combining these two examples we consider the following set of amplitudes in the present paper

$$\mathcal{M}_{\text{theory}}^{\text{here}} = c_{(\text{ss})}\mathcal{M}_{\text{tree}}^{(\text{ss})} + c_{(\text{hs})}\mathcal{M}_{\text{tree}}^{(\text{hs})} + c_{(\text{bs})}\mathcal{M}_{\text{tree}}^{(\text{bs})} + \sum_{S=0,\frac{1}{2},1,\frac{3}{2},2} c_S \mathcal{M}_S^{1-\text{loop}}. \tag{4.28}$$

Obviously, we do not claim that this space is complete. Let us emphasize that in the analysis below we allow taking arbitrary superpositions of string amplitudes with various tensions as well as one-loop amplitudes with arbitrary choices of the masses and spins, $S \leqslant 2$, of particles circulating in the loop.

On the string side, for example we can consider the ϵ deformation of the superstring amplitude discussed in the conclusions of reference [5]. It is not clear however that these deformations can be promoted to fully consistent perturbative *S*-matrices. Moreover, we checked that adding this correction does not affect the theory island considered in the present paper in a noticeable way.

On the field-theory side we can consider gapped strongly-coupled theories of matter coupled to gravity. One interesting example is large-*N* QCD coupled to gravity [15]. In this case we expect the tree-level corrections due to non-minimally coupled glueballs and one-loop corrections due to minimally-coupled glueballs to both enter at the same order G^2. Finally, we do not discuss here amplitudes in models with extra dimensions [73–75]. Of course, it would be very interesting to extend our analysis to include these examples as well.

In writing the contribution of matter we can imagine integrating out a multiple number of particles of given spin with various masses. For the constraints discussed below one can check that the effect of this freedom can be absorbed into rescaling of the coefficients c_S. In particular, the bounds that we describe below hold for any choice of masses and number of species for particles that we integrate out.

---

[21] Strictly speaking the bosonic string is not part of the theory space due to the presence of the tachyon in the spectrum, but it is useful to keep it to check various formulas.

In the following sections we use these amplitudes to populate the theory island. Surprisingly, we find that

$$\text{Theory island} \ll \text{EFThedron}, \tag{4.29}$$

in a sense that should become clear below. In other words, the set of known weakly-coupled theories occupy a much smaller space in the space of couplings than is allowed by the general constraints coming from the analysis of $2 \rightarrow 2$ scattering. This is, of course, in accord with the ongoing landscape vs swampland debate but is also different from that. Indeed, the theory island as defined here does not guarantee the existence of the nonperturbative completion of gravity. We only study the consistency of the leading-order corrections to the Einstein–Hilbert theory perturbatively in G and already in this setting we seem to find a much smaller space of possibilities than follows from general constraints.

5. Deriving bounds: elastic amplitude

In this section we analyze bounds on the low-energy expansion of the elastic amplitude $f(s, u)$ using the techniques of references [3–5].[22] We focus on couplings of the same dimensionality and derive two-sided bounds on them by explicitly identifying the facets of the relevant polytopes and then taking their crossing-symmetric section.

As observed in reference [5], identifying the relevant boundaries are particularly simple in case of the one-channel dispersion relations when the polytopes in question are cyclic.[23] In case of the two-channel dispersion relations the cyclicity property is lost. However, as reference [5] observed, a set of new boundaries that appear in this case involve low-spin partial waves and can be explicitly identified upon inspection. This is indeed what we observed in the examples below. To the best of our knowledge there is no proof that the list of boundaries obtained in this way is complete (and therefore that the bounds are optimal), and we do not attempt such a proof in the present work.

We will not attempt to relate couplings of different dimensionality to each other either. This problem was recently analyzed in reference [3, 4, 6] and the expected dimensional-analysis scaling of various couplings with order $\mathcal{O}(1)$ coefficients was rigorously established. It would be very interesting to apply these techniques to the gravitational amplitudes discussed here but we defer this to the future.

5.1. Strategy

We first briefly describe the basic strategy to derive bounds [5]. We consider the vector of low-energy couplings of the same dimensionality $\vec{F}$ which via dispersion relations is given by the sum of s- and u-channel partial waves

$$\vec{F} = \sum_J \left(c_{J,s} \vec{V}_{J,s} + c_{J,u} \vec{V}_{J,u} \right), \quad c_{J,s}, c_{J,u} \geqslant 0. \tag{5.1}$$

In this formula $c_{J,s}$ and $c_{J,u}$ encode spectral densities discussed in section 4.2 and $\vec{V}_{J,s}$, $\vec{V}_{J,u}$ are known functions related to partial waves—see e.g. (5.8) for the precise formula. However for the present discussion we can simply think of $\vec{V}_{J,s}$ and $\vec{V}_{J,u}$ as some abstract, given vectors

---

[22] See also reference [76] for a closely related discussion.

[23] We refer the reader to reference [5] for the definition of this term.

and $c_{J,s}$, $c_{J,u}$ being non-negative numbers. We would like to characterize the space (5.1), which is a polytope. A convenient way to do it is by identifying its facets

$$\vec{F} \cdot \vec{W}_i \geqslant 0, \tag{5.2}$$

where $\vec{W}_i$ is a normal to a given facet. From equation (5.1) we see that all the facets are of the form[24]

$$(W_i)_{I_0} = \epsilon_{I_0 I_1 \ldots I_d} V_{a_1}^{I_1} \ldots V_{a_d}^{I_d}, \tag{5.3}$$

where we assumed that the vectors $\vec{V}_a$ are $(d+1)$-dimensional and we contracted them using the $(d+1)$-dimensional ϵ-tensor.[25] The index a labels both spin and channel. Therefore, we can characterize the space (5.1) by a set of determinants $\langle F, a_1, \ldots, a_d \rangle \geqslant 0$. The task is then to find all $(a_1, \ldots, a_d)$, such that equation (5.2) holds. In general, this is a formidable task since the space of vectors in equation (5.1) is infinite dimensional.

As explained in reference [5], remarkably, the set of vectors $\vec{V}_{J,s}$ defines a cyclic polytope, which for us simply means that the set of its boundaries can be written down explicitly very easily. They essentially take the form of determinants built out of pairs of consecutive vectors in the sum over spins. For example, for $d = 4$ the relevant determinants take the form $(i, i+1, j, j+1)$ with $j > i$, where i and j label various spins in the sum (5.1).

The cyclicity property is lost when considering the sum of the s- and u-channel as in equation (5.1). By inspecting the resulting polytope reference [5] observed that it is 'almost cyclic' in the sense that the only boundaries which are not of the type $(i, i+1, j, j+1)$ in the example considered above involve low-spin partial waves in equation (5.1) and can be found by direct inspection. This is what we also find and this is the strategy of finding bounds that we follow in this paper. In particular, we do not prove that the set of bounds found in this way is complete (even though we suspect it to be the case). It would be very interesting to rigorously demonstrate this.

After the allowed region (5.2) is identified we impose crossing symmetry, which is the subspace defined by linear relations between various components of the coupling vector $\vec{F}$. By taking the crossing-symmetric slice of the allowed space of couplings we get two-sided bounds on the low-energy Wilson coefficients.

5.2. Non crossing-symmetric dispersive representation of low-energy couplings

Consider again the double-minus amplitude

$$\mathcal{M}_4(1^-, 2^-, 3^+, 4^+) = (\langle 12 \rangle [34])^4 \, f(t, u). \tag{5.4}$$

---

[24] Strictly speaking, this statement is only true for finite-dimensional sums in equation (5.1). For infinite-dimensional sums we can also have limiting points. For us the limiting point is $J = \infty$. We find, however, that its existence does not play a role in the analysis in the following sections. The reason for it is that in practice we search for the boundaries (5.2) by first truncating the sum over spins up to some J_{max} and then extrapolating to $J_{max} = \infty$.

[25] Here d has nothing to do with the dimensionality of spacetime. Instead, it is the dimensionality of the relevant subset of the low-energy couplings that we would like to bound.

Following reference [5], we introduce the expansion[26]

$$f(t, u) = f(t, -s - t) = \left(\frac{\kappa}{2}\right)^2 \frac{1}{stu} + |\beta_{R^3}|^2 \frac{tu}{s} - |\beta_\phi|^2 \frac{1}{s} + \sum_{k \geqslant j \geqslant 0} a_{k,j} s^{k-j} t^j. \qquad (5.5)$$

From equation (5.5) we see that for fixed k all $a_{k,j}$ have the same dimensionality. Crossing symmetry $f(t, u) = f(u, t)$ leads to linear relations among $a_{k,j}$ which will play an important role below.

Of course, in equation (4.1) we are expanding the same function, so the $f_{i,j}$ introduced there and the $a_{k,j}$ in equation (5.5) are all related to each other in a trivial fashion. We get for the first few coefficients

$$a_{0,0} = f_{0,0}, \qquad a_{1,0} = -f_{1,0}, \qquad a_{1,1} = 0. \qquad (5.6)$$

The vanishing of $a_{1,1}$ is a consequence of t–u crossing symmetry of $f(t, u)$ which is not manifest in equation (5.5).

The function of interest $f(t, u)$ admits a simple dispersive representation

$$\begin{aligned}
f(t, -s - t) = \oint \frac{\mathrm{d}s'}{2\pi i} \frac{f(t, -s' - t)}{s - s'} = {} & \left(\frac{\kappa}{2}\right)^2 \frac{1}{stu} + |\beta_{R^3}|^2 \frac{tu}{s} - |\beta_\phi|^2 \frac{1}{s} \\
& - \int_{m_{\mathrm{gap}}^2}^\infty \frac{\mathrm{d}m^2}{\pi} \left(\sum_{J=0}^\infty \frac{1 + (-1)^J}{2} \frac{\rho_J^{++}(m^2) d_{0,0}^J (1 + \frac{2t}{m^2})}{m^8} \frac{1}{s - m^2} \right. \\
& \left. + \sum_{J=4}^\infty \frac{\rho_J^{+-}(m^2) d_{4,4}^J \left(1 + \frac{2t}{m^2}\right)}{(t + m^2)^4} \frac{1}{-s - t - m^2} \right).
\end{aligned} \qquad (5.7)$$

In deriving (5.7) we used the Regge bound (4.23) to drop the arcs at infinity. When applied to equation (5.4) it implies that $|f(t, -s - t)| \leqslant 1/|s|$ at large s and fixed t. Indeed, this is the case for all the one-loop amplitudes that we consider.

An important feature of equation (5.7) is that crossing symmetry $f(t, u) = f(u, t)$ is not manifest (hence the title of this section). Therefore imposing it leads to interesting constraints. In fact, as we will see below it leads to two-sided bounds on the couplings.

By expanding equation (5.7) at small s and t we get the dispersive representations for the couplings $a_{k,j}$. They take the following form

$$a_{k,j} = \int_{m_{\mathrm{gap}}^2}^\infty \frac{\mathrm{d}m^2}{\pi} \frac{1}{(m^2)^{k+1}} \left(\sum_{J=0}^\infty \frac{1 + (-1)^J}{2} \frac{\rho_J^{++}(m^2)}{m^8} P_{++}^j(\mathcal{J}^2) + \sum_{J=4}^\infty \frac{\rho_J^{+-}(m^2)}{m^8} P_{+-}^{k,j}(\mathcal{J}^2) \right), \qquad (5.8)$$

where we introduced the spin Casimir $\mathcal{J}^2 = J(J + 1)$. We call equation (5.8) the dispersive representation of low-energy couplings. We again emphasize that crossing symmetry leads to linear relations among $a_{k,j}$ which are not manifest in equation (5.8).

---

[26] Note that we performed an $s \leftrightarrow t$ transformation compared to equation (4.1).

In the formula above we introduced polynomials $P_{++}^{j}(\mathcal{J}^2)$ and $P_{+-}^{k,j}(\mathcal{J}^2)$ which have the following properties[27]

$$
\begin{aligned}
P_{++}^0(\mathcal{J}^2) &= 1, \qquad P_{+-}^{k,0}(\mathcal{J}^2) = (-1)^k, \\
P_{++}^1(\mathcal{J}^2) &= \mathcal{J}^2, \qquad P_{+-}^{k,1}(\mathcal{J}^2) = (-1)^{k+1}(\mathcal{J}^2 - 20 - k), \\
\lim_{J \to \infty} P_{++}^j(\mathcal{J}^2) &= \frac{1}{\Gamma(j+1)^2} J^{2j} + \ldots, \\
\lim_{J \to \infty} P_{+-}^{k,j}(\mathcal{J}^2) &= \frac{(-1)^{k+j}}{\Gamma(j+1)^2} J^{2j} + \ldots
\end{aligned}
\tag{5.9}
$$

Based on the formulas above we see that couplings $a_{k,j}$ written in equation (5.8) roughly probe the k'th moment of the spectral density with respect to m^2 and $(2j)$'th moment with respect to spin. In particular, we expect that higher-j coefficients to be more sensitive to higher-spin spectral densities and higher-k couplings to have the large $\frac{m^2}{m_{\text{gap}}^2}$ region more suppressed. This is indeed what we will find in the explicit examples below.

Crossing symmetry leads to the set of sum rules which were dubbed null constraints in reference [4]. Effectively, they express the low-spin partial-wave data in terms of the higher-spin data. Consider for example the relation $a_{1,1} = 0$, listed in equation (5.6). Using the formulas above we can write it as follows

$$
\begin{aligned}
\int_{m_{\text{gap}}^2}^\infty \frac{\mathrm{d}m^2}{(m^2)^2} \frac{\rho_4^{+-}(m^2)}{m^8} = \int_{m_{\text{gap}}^2}^\infty \frac{\mathrm{d}m^2}{(m^2)^2} &\left(\sum_{J=2}^\infty \frac{1+(-1)^J}{2} \frac{\rho_J^{++}(m^2)}{m^8} \mathcal{J}^2 \right. \\
&\left. + \sum_{J=5}^\infty \frac{\rho_J^{+-}(m^2)}{m^8}(\mathcal{J}^2 - 21) \right),
\end{aligned}
\tag{5.10}
$$

where all the terms in the rhs of equation (5.10) are non-negative. In other words, crossing symmetry allows us to express the moment of $\rho_4^{+-}(m^2)$ in terms of all other spectral densities.

Let us next describe the first few bounds for $k \leqslant 6$. We do not attempt to derive the bounds across the couplings of different dimensionalities as it was done in references [3, 4, 6, 7]. Instead we focus on the geometry of couplings of the same dimensionality as was studied in reference [5]. For $k = 1, 3, 5$ we do not get any nontrivial bounds. For even k the results are presented below.

k = 0:
We first start with $k = 0$. From (5.7) it immediately follows that

$$
a_{0,0} \geqslant 0.
\tag{5.11}
$$

Moreover, we have

$$
a_{k,0} = \frac{1}{\pi} \int_{m_{\text{gap}}^2}^\infty \frac{\mathrm{d}m^2}{m^{2k+10}} \left(\sum_{J=0}^\infty \frac{1+(-1)^J}{2} \rho_J^{++}(m^2) + (-1)^k \sum_{J=4}^\infty \rho_J^{+-}(m^2) \right).
\tag{5.12}
$$

---

[27] In the $++$ channel the closed-form expression takes the form $P_{++}^j(\mathcal{J}^2) = \frac{1}{\Gamma(j+1)^2} \prod_{n=1}^j (\mathcal{J}^2 - n(n-1))$.

This immediately implies that $a_{k,0}$ for even k can be interpreted as even moments μ_k

$$\mu_k \equiv \frac{1}{\pi} \int_{m_{\mathrm{gap}}^2}^{\infty} \frac{\mathrm{d}m^2}{m^{2k+10}} \left(\sum_{J=0}^{\infty} \frac{1+(-1)^J}{2} \rho_J^{++}(m^2) + \sum_{J=4}^{\infty} \rho_J^{+-}(m^2) \right). \tag{5.13}$$

As a result the bounds of reference [77] apply. For odd k, we obviously have $|a_k| \leqslant \mu_k$.
k = 2:
Next consider $k = 2$. Here the couplings of the same dimensionality are $\mathbf{a}_2 = (a_{2,0}, a_{2,1}, a_{2,2})$. One obvious constraint is

$$a_{2,0} \geqslant 0. \tag{5.14}$$

On top of that we get the following list of constraints which specify the boundary of the region of allowed couplings,[28]

$$\{ \langle \mathbf{a}_2, i_s, i_s+2 \rangle_{i_s \geqslant 2}, \langle \mathbf{a}_2, (i_u+1)^{(2)}, i_u^{(2)} \rangle_{i_u \geqslant 4}, \langle \mathbf{a}_2, 4_u^{(2)}, 2_s \rangle \} \geqslant 0. \tag{5.15}$$

Here we follow the notation of reference [5], which is defined as

$$\langle \mathbf{a}_2, i, j \rangle \equiv \det \begin{pmatrix} a_{2,0} & v_{i,0} & v_{j,0} \\ a_{2,1} & v_{i,1} & v_{j,1} \\ a_{2,2} & v_{i,2} & v_{j,2} \end{pmatrix}, \tag{5.16}$$

where we also define the following vectors

$$v_{J,m}^{(s)} = (J_s)_m \equiv \frac{1}{m!} \partial_t^m d_J^{0,0}(1+2t)|_{t=0}, \tag{5.17}$$

$$v_{J,k,m}^{(u)} = (J_u^{(k)})_m \equiv \frac{1}{m!} \partial_t^m \frac{1}{(k-m)!} \partial_s^{k-m} \frac{1}{(1+s+t)} \frac{d_J^{4,4}(1+2t)}{(1+t)^4} |_{s,t=0}, \tag{5.18}$$

where in the formula above we set the arbitrary mass scale $m^2 = 1$ since it leads to trivial overall rescaling of the vectors that enter into the determinant (5.16) and as such does not affect the positivity bounds (5.15). Note that the vector $v_{i,k}^{(u)}$ depends on the order k of the corrections that we are studying, which we denote by $i_u^{(k)}$. For example, to evaluate (5.15) we should set $k = 2$ in equation (5.18). The constraints (5.15) have an intuitive explanation. By plugging the dispersive representation of the couplings into the determinant $\langle \mathbf{a}_2, i, j \rangle$ we get a sum over partial waves where the coefficients of partial wave i and j are zero. The plane $\langle \mathbf{a}_2, i, j \rangle$ then separates the region in the coupling space where these coefficients are positive and negative.

We see that equation (5.15) involves cyclic constraints from the s-channel $\langle \mathbf{a}_2, i_s, i_s+2 \rangle_{i_s \geqslant 2}$, similarly from the u-channel $\langle \mathbf{a}_2, (i_u+1)^{(2)}, i_u^{(2)} \rangle_{i_u \geqslant 4}$, as well as a mixed constraint $\langle \mathbf{a}_2, 4_u^{(2)}, 2_s \rangle$ that involves low-spin partial waves. This phenomenon observed in reference [5] continues to hold for higher k as well. It allows us to easily identify the set of conditions that carve out the region of allowed couplings analytically rather straightforwardly by inspecting the low-spin boundaries. Of course, in practice we only look for possible boundaries up to some finite spin J_{max} and assume that the observed pattern continues all the way to $J_{\mathrm{max}} = \infty$.

In deriving the bounds above we did not impose crossing symmetry. At the level $k = 2$ we are working it implies that

$$\mathrm{Crossing}: a_{2,1} = a_{2,2}. \tag{5.19}$$

---

[28] The constraint (5.14) can be understood as arising from the limiting point when i_s or $i_u \to \infty$.
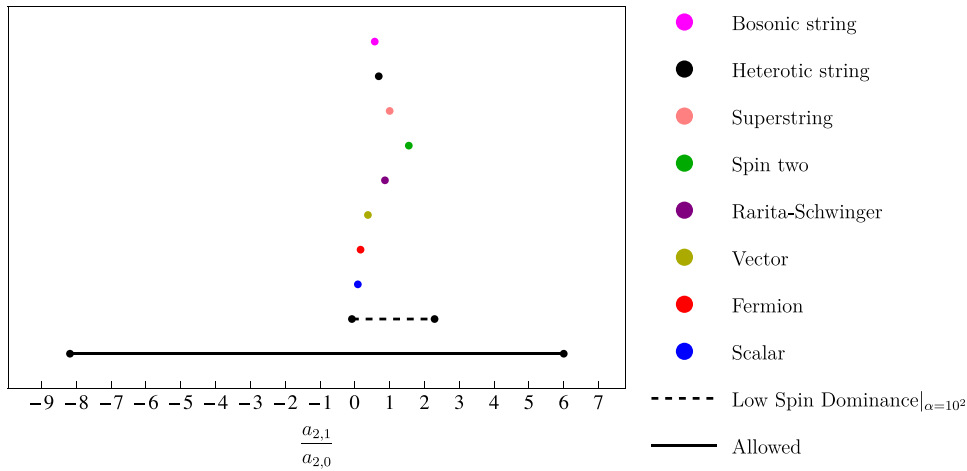
Figure 8. The allowed region for $\frac{a_{2,1}}{a_{2,0}}$ given by equation (5.20) is depicted in black. The explicit amplitudes that emerge from integrating out the one-loop matter or tree-level string theories are depicted in various colors. Assuming LSD in the form (5.26) with $\alpha = 10^2$ one can derive stronger bounds which we depict by the dashed line.

Taking the crossing-symmetric slice of the constraints (5.15) gives the two-sided bound on the Wilson coefficients

$$-\frac{90}{11} \leqslant \frac{a_{2,1}}{a_{2,0}} = \frac{a_{2,2}}{a_{2,0}} \leqslant 6 \,. \tag{5.20}$$

It is very striking that for the explicit UV completions studied here we find a much narrow window of possibilities that we depict in figure 8. We now try to understand the origin of this fact. To do this it will be useful to use the null constraint $a_{2,1} - a_{2,2} = 0$ to derive a bound similar to equation (5.20). By adding the null constraint to the dispersive representations for $a_{2,1}$ and $a_{2,2}$ it is straightforward to show that

$$-\frac{108}{13} \frac{\langle \rho_5^{+-} \rangle_2}{\langle \rho_0^{++} \rangle_2 + \langle \rho_4^{+-} \rangle_2 + \langle \rho_5^{+-} \rangle_2} \leqslant \frac{a_{2,1}}{a_{2,0}}$$

$$= \frac{a_{2,2}}{a_{2,0}} \leqslant \frac{6 \langle \rho_2^{++} \rangle_2 + \frac{16}{7} \langle \rho_4^{+-} \rangle_2}{\langle \rho_0^{++} \rangle_2 + \langle \rho_4^{+-} \rangle_2 + \langle \rho_2^{++} \rangle_2}, \tag{5.21}$$

where we introduced $\langle \ldots \rangle$ for the relevant integrals over intermediate energies m^2

$$\langle \rho_J \rangle_k \equiv \frac{1}{\pi} \int_{m_{\text{gap}}^2}^{\infty} \frac{dm^2}{m^{2k+10}} \rho_J(m^2) \,. \tag{5.22}$$

To derive the upper bound in equation (5.21) we used the dispersive representation for $a_{2,1} - \frac{2}{7}(a_{2,2} - a_{2,1})$. To derive the lower bound we considered $a_{2,1} + \frac{4}{13}(a_{2,2} - a_{2,1})$.

The bound (5.21) is rigorous. Its advantage however is that it only contains low-spin spectral densities. By adding the null constraint $a_{2,1} - a_{2,2} = 0$ with an appropriate coefficient we made sure that all the higher-spin contributions are sign-definite and the bound (5.21) then follows. It is clear then how can we come close to the saturation of the bounds (5.20). The upper bound saturation requires that $\langle \rho_2^{++} \rangle_2$ is dominant, whereas the lower-bound saturation

requires that $\langle \rho_5^{+-} \rangle_2$ dominates. Note that each of these is not the minimal spin that appears in the corresponding channel, which are $\langle \rho_0^{++} \rangle_2$ and $\langle \rho_4^{+-} \rangle_2$ correspondingly.

This is not what happens in the known physical theories as we will show in more detail. In the examples that we analyze in this paper the higher-spin contributions to the spectral densities are suppressed compared to the minimal spin ones. The fact that at large spin the spectral densities decay exponentially in spin is something familiar from the fact that the scattering is local in the impact parameter space. Analyzing the concrete examples we see that this hierarchy continues all the way to minimal spin. We call this phenomenon LSD of partial wave and mathematically we can express it as follows:

$$\text{Low-spin dominance (weak)} : \langle \rho_4^{+-} \rangle_k \geqslant \langle \rho_{J>4}^{+-} \rangle_k, \quad \langle \rho_0^{++} \rangle_k \geqslant \langle \rho_{J>0}^{++} \rangle_k. \quad (5.23)$$

In fact, equation (5.23) is very conservative and in the explicit examples the suppression of higher-spin partial waves is much stronger than equation (5.23). Nevertheless, already using equations (5.21) and (5.23) we can strengthen the bound (5.20) to obtain

$$(5.21) + (5.23) : -\frac{54}{13} \leqslant \frac{a_{2,1}}{a_{2,0}} = \frac{a_{2,2}}{a_{2,0}} \leqslant 3. \quad (5.24)$$

We clarify that here we derived equation (5.24) to illustrate the point that LSD implies stronger bounds. We have not tried to find the optimal bound that follows from equation (5.23) and dispersive representations of the couplings. It would be interesting to study this more systematically.

It is also instructive to see what happens if we assume LSD in the strong form

$$\text{Low-spin dominance (strong)} : \langle \rho_4^{+-} \rangle_k \gg \langle \rho_{J>4}^{+-} \rangle_k, \quad \langle \rho_0^{++} \rangle_k \gg \langle \rho_{J>0}^{++} \rangle_k.$$
$$(5.25)$$

We will discuss in more detail the expected form of the hierarchy below in the section dedicated to spectral densities, but for now it is sufficient to say that we will find that equation (5.25) correctly captures the region occupied by the known theories. To make (5.25) more precise let us introduce its quantitative version

$$\text{Low-spin dominance}(\alpha - \text{factor}) : \langle \rho_4^{+-} \rangle_k \geqslant \alpha \langle \rho_{J>4}^{+-} \rangle_k,$$
$$\langle \rho_0^{++} \rangle_k \geqslant \alpha \langle \rho_{J>0}^{++} \rangle_k. \quad (5.26)$$

For illustrating bounds, below we choose for concreteness $\alpha = 10^2$. This choice is not accidental as the perturbative examples considered in this paper, tree-level string amplitudes and one-loop matter amplitudes, happen to be of this type.[29]

Let us now consider the bounds coming from equation (5.21) with the additional assumption of $\alpha = 10^2$ LSD (5.26). In this way we obtain

$$(5.21) + (5.26)|_{\alpha=10^2} : -0.083 \leqslant \frac{a_{2,1}}{a_{2,0}} = \frac{a_{2,2}}{a_{2,0}} \leqslant 2.286. \quad (5.27)$$

It is also interesting to consider what happens if we consider $\alpha \to \infty$. In this case one can again use the trick of adding a null constraint $(a_{2,2} - a_{2,1})$ with some non-zero coefficient to

---

[29] One exception is the heterotic string amplitude where we have $\langle \rho_2^{++} \rangle_k \sim \frac{1}{10} \langle \rho_0^{++} \rangle_k$.

dispersive representation of the couplings to show that

$$\text{LSD}_{\alpha \to \infty} : 0 \leqslant \frac{a_{2,1}}{a_{2,0}} = \frac{a_{2,2}}{a_{2,0}} \leqslant 2 \,. \tag{5.28}$$

Remarkably, we find that the region occupied by known theories in figure 8 lies inside equation (5.27) and even equation (5.28). We will see below that the phenomenon of the data being well explained by the strong version of LSD continues for higher k's as well.

k = 4:

For $k = 4$ we first consider a subset of couplings $\hat{\mathbf{a}}_4 = (a_{4,0}, a_{4,1}, a_{4,2})$. From the dispersive representation of the couplings it immediately follows that

$$a_{4,0} \geqslant 0 \,. \tag{5.29}$$

Analyzing the sums over the partial waves we get the following list of constraints[30]

$$\left\{ \langle \hat{\mathbf{a}}_2, i_s, i_s + 2 \rangle_{i_s \geqslant 2}, \langle \hat{\mathbf{a}}_4, i_u^{(4)} + 1, i_u^{(4)} \rangle_{i_u \geqslant 5}, \langle \hat{\mathbf{a}}_4, 5_u^{(4)}, 2_s \rangle \right\} \geqslant 0 \,. \tag{5.30}$$

These include the usual cyclic constraints, as well as an extra fixed constraint. We plot the allowed region in the as shaded (green) in figure 9. A characteristic feature of the allowed region is that it is unbounded.

In deriving the bounds above we have not used crossing symmetry which leads to improved two-sided bounds. To do that we can consider a complete set of couplings at $k = 4$, namely $\mathbf{a}_4 = (a_{4,0}, a_{4,1}, a_{4,2}, a_{4,3}, a_{4,4})$. It is then straightforward to check by inspection that the following set of constraints define the boundary region in the space of couplings:

$$\begin{aligned}
&\left\{ \langle \mathbf{a}_4, i_s, i_s + 2, j_s, j_s + 2 \rangle_{j > i \geqslant 2}, \langle \mathbf{a}_4, i_u^{(4)}, i_u^{(4)} + 1, j_u^{(4)}, j_u^{(4)} + 1 \rangle_{j > i \geqslant 5}, \right. \\
&\langle \mathbf{a}_4, i_s, i_s + 2, j_u^{(4)} + 1, j_u^{(4)} \rangle_{i \geqslant 4, j \geqslant 5}, \langle \mathbf{a}_4, 2_s, 4_s, j_u^{(4)} + 1, j_u^{(4)} \rangle_{j \geqslant 7}, \\
&\langle 2_s, \mathbf{a}_4, i_s, i_s + 2, 5_u^{(4)} \rangle_{i \geqslant 4}, \langle 5_u^{(4)}, \mathbf{a}_4, j_u^{(4)}, j_u^{(4)} + 1, 2_s \rangle_{j \geqslant 7}, \\
&\langle \mathbf{a}_4, 4_u^{(4)}, 5_u^{(4)}, 6_u^{(4)}, 7_u^{(4)} \rangle, \langle \mathbf{a}_4, 2_s, 4_s, 4_u^{(4)}, 5_u^{(4)} \rangle, \langle \mathbf{a}_4, 2_s, 4_s, 7_u^{(4)}, 4_u^{(4)} \rangle, \\
&\left. \langle \mathbf{a}_4, 4_u^{(4)}, 5_u^{(4)}, 7_u^{(4)}, 2_s \rangle, \langle 4_u^{(4)}, \mathbf{a}_4, 5_u^{(4)}, 6_u^{(4)}, 4_s \rangle, \langle 4_u^{(4)}, \mathbf{a}_4, 6_u^{(4)}, 7_u^{(4)}, 4_s \rangle \right\} \geqslant 0 \,.
\end{aligned} \tag{5.31}$$

It is easy to recover the previous constraints (5.30) from equation (5.31) by taking some of the spins to infinity. For example $\lim_{i_s \to \infty} \langle 2_s, \mathbf{a}_4, i_s, i_s + 2, 5_u^{(4)} \rangle_{i \geqslant 4} \geqslant 0$ reduces to $\langle \hat{\mathbf{a}}_4, 5_u^{(4)}, 2_s \rangle \geqslant 0$.

A somewhat new feature of this case compared to the ones considered above that mixed $s–u$ channel boundaries come in infinite families. However, these families are again either cyclic, as in $\langle \mathbf{a}_4, i_s, i_s + 2, j_u^{(4)} + 1, j_u^{(4)} \rangle_{i \geqslant 4, j \geqslant 5}$ or involve low-spin partial wave only and therefore easily identifiable. We find it quite remarkable that the allowed region can be found analytically!

We then consider the section of this region by the crossing symmetry relations that take the form

$$\text{Crossing} : a_{4,3} = 2(a_{4,2} - a_{4,1}), \qquad a_{4,4} = a_{4,2} - a_{4,1} \,. \tag{5.32}$$

As a result we get the region of allowed couplings depicted in red in figure 9. To generate the plot we considered bounds (5.31) where we truncated the maximal spin to $i_{\max}, j_{\max}, k_{\max} = 20$.

---

[30] By taking $i_s \text{ or } i_u \to \infty$ limit in equation (5.30) we recover the simple bound (5.29).
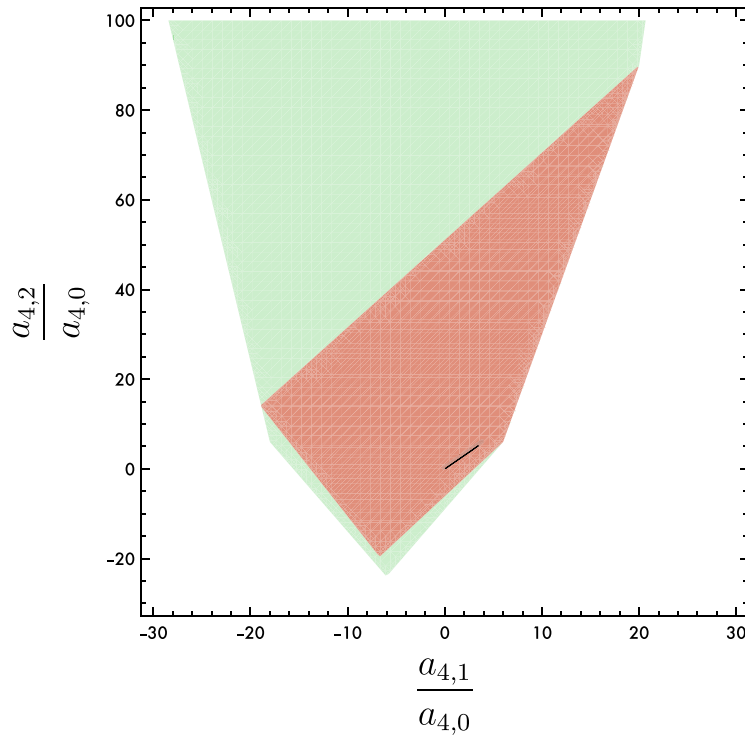
Figure 9. The allowed region for $(\frac{a_{4,1}}{a_{4,0}}, \frac{a_{4,2}}{a_{4,0}})$. The lightly shaded (green) region corresponds to the bounds (5.30). The darkly shaded (red) region corresponds to the bounds (5.31). The black theory island (which is so narrow that on this scale it looks like a line segment) is the region covered by known amplitudes. The small gray-shaded region that surrounds the black theory island corresponds to the LSD $\alpha = 10^2$ bounds.

We also checked explicitly that the all determinants obtained in this way are non-negative given dispersive representation for the couplings truncated to $J_{\max} = 200$.

The bounds coming from imposing equation (5.31) together with crossing symmetry (5.32) are two-sided, both from above and from below. The theory island occupied by the known theories forms a tiny black slit inside the allowed region.

As we did for $k = 2$ we can again understand the structure of the island using the idea of LSD. To this extent we can use the null constraints coming from crossing (5.32) to derive the following rigorous bounds

$$
-\frac{7.01\langle\rho_5^{+-}\rangle_4 + 20.18\langle\rho_6^{+-}\rangle_4}{\langle\rho_0^{++}\rangle_4 + \langle\rho_4^{+-}\rangle_4 + \langle\rho_5^{+-}\rangle_4 + \langle\rho_6^{+-}\rangle_4}
$$
$$
\leqslant \frac{a_{4,1}}{a_{4,0}} \leqslant \frac{6\langle\rho_2^{++}\rangle_4 + 20\langle\rho_4^{++}\rangle_4 + 4.01\langle\rho_4^{+-}\rangle_4}{\langle\rho_0^{++}\rangle_4 + \langle\rho_4^{+-}\rangle_4 + \langle\rho_2^{++}\rangle_4 + \langle\rho_4^{++}\rangle_4}, \tag{5.33}
$$
$$
-\frac{20.16\langle\rho_5^{+-}\rangle_4}{\langle\rho_0^{++}\rangle_4 + \langle\rho_4^{+-}\rangle_4 + \langle\rho_5^{+-}\rangle_4 + \langle\rho_6^{+-}\rangle_4}
$$
$$
\leqslant \frac{a_{4,2}}{a_{4,0}} \leqslant \frac{6\langle\rho_2^{++}\rangle_4 + 90\langle\rho_4^{++}\rangle_4 + 6.43\langle\rho_4^{+-}\rangle_4 + 15.32\langle\rho_6^{+-}\rangle_4}{\langle\rho_0^{++}\rangle_4 + \langle\rho_4^{+-}\rangle_4 + \langle\rho_2^{++}\rangle_4 + \langle\rho_4^{++}\rangle_4 + \langle\rho_6^{+-}\rangle_4}. \tag{5.34}
$$

Again we see that the corners of the red regions in figure 9 can be obtained from equation (5.33) by assuming that the sub-leading spins in the corresponding channels are dominant. This is not what happens in the physical theories and again assuming weak LSD (5.23) we can get a tighter bounds in which physical theories reside. In practice, we see that physical theories occupy even smaller region which can be understood using the strong version of LSD (5.25).

A striking feature of figure 10 is that the known theories align closely with the straight line with the slope $3/2$. To understand this we can derive the bound analogous to equation (5.33) for the difference $a_{4,2} - \frac{3}{2}a_{4,1}$. It takes the following form

$$
-\frac{3\langle\rho_2^{++}\rangle_4 + 15\langle\rho_5^{+-}\rangle_4}{\langle\rho_0^{++}\rangle_4 + \langle\rho_4^{+-}\rangle_4 + \langle\rho_2^{++}\rangle_4 + \langle\rho_5^{+-}\rangle_4}
$$
$$
\leqslant \frac{a_{4,2} - \frac{3}{2}a_{4,1}}{a_{4,0}} \leqslant \frac{60\langle\rho_4^{++}\rangle_4 + 0.47\langle\rho_4^{+-}\rangle_4 + 45.58\langle\rho_6^{+-}\rangle_4}{\langle\rho_0^{++}\rangle_4 + \langle\rho_4^{+-}\rangle_4 + \langle\rho_4^{++}\rangle_4 + \langle\rho_6^{+-}\rangle_4}. \quad (5.35)
$$

We emphasize that at this point the bound (5.35) is rigorous and no additional assumptions have been made. To derive an upper bound we have considered the dispersive representation for $a_{4,2} - \frac{3}{2}a_{4,1} + \frac{1687}{7205}(a_{4,3} - 2a_{4,4}) + \frac{411}{524}(a_{4,3} - 2(a_{4,2} - a_{4,1}))$, where we made use of crossing (5.32). The lower bound follows directly from the dispersive representation of $a_{4,2} - \frac{3}{2}a_{4,1}$.

To make use of it, we can now assume $\alpha = 10^2$ LSD and apply it to the bounds (5.33) and (5.35). In this way we get

$$
(5.33) + (5.34) + (5.35) + \mathrm{LSD}_{\alpha=10^2} : -0.27 \leqslant \frac{a_{4,1}}{a_{4,0}} \leqslant 4.01, \quad (5.36)
$$

$$
-0.20 \leqslant \frac{a_{4,2}}{a_{4,0}} \leqslant 6.52,
$$

$$
-0.15 \leqslant \frac{a_{4,2} - \frac{3}{2}a_{4,1}}{a_{4,0}} \leqslant 0.92. \quad (5.37)
$$

We plot the result in figure 10 and again observe that the known examples neatly land in the predicted region. Moreover, it is straightforward to see that if we increase α the bound (5.35) becomes not optimal and instead we get

$$
\mathrm{LSD}_{\alpha\to\infty} : \frac{a_{4,2} - \frac{3}{2}a_{4,1}}{a_{4,0}} = 0, \quad 0 \leqslant \frac{a_{4,1}}{a_{4,0}} \leqslant 4, \ 0 \leqslant \frac{a_{4,2}}{a_{4,0}} \leqslant 6 \quad (5.38)
$$

which is the line along which our explicit examples cluster. Intuitively, this result can be understood as follows: we write down explicitly the dispersive representation for the couplings $a_{4,i}$,

$$
a_{4,0} = \int_{m_{\mathrm{gap}}^2}^{\infty} \frac{\mathrm{d}m^2}{\pi} \frac{\rho_{++0}(m^2)}{(m^2)^9} + \int_{m_{\mathrm{gap}}^2}^{\infty} \frac{\mathrm{d}m^2}{\pi} \frac{\rho_{+-4}(m^2)}{(m^2)^9} + \text{higher spin},
$$

$$
a_{4,1} = 4 \int_{m_{\mathrm{gap}}^2}^{\infty} \frac{\mathrm{d}m^2}{\pi} \frac{\rho_{+-4}(m^2)}{(m^2)^9} + \text{higher spin}, \quad (5.39)
$$

$$
a_{4,2} = 6 \int_{m_{\mathrm{gap}}^2}^{\infty} \frac{\mathrm{d}m^2}{\pi} \frac{\rho_{+-4}(m^2)}{(m^2)^9} + \text{higher spin},
$$

where 'higher spin' denotes infinitely many partial-wave contributions. Suppose that the lowest-spin partial waves in each channel dominates. Note that the leading spin $J = 0$ contribution in the ρ_{++J} channel drops out from $a_{4,1}$ and $a_{4,2}$. Notice that if we set the 'higher-spin'

49
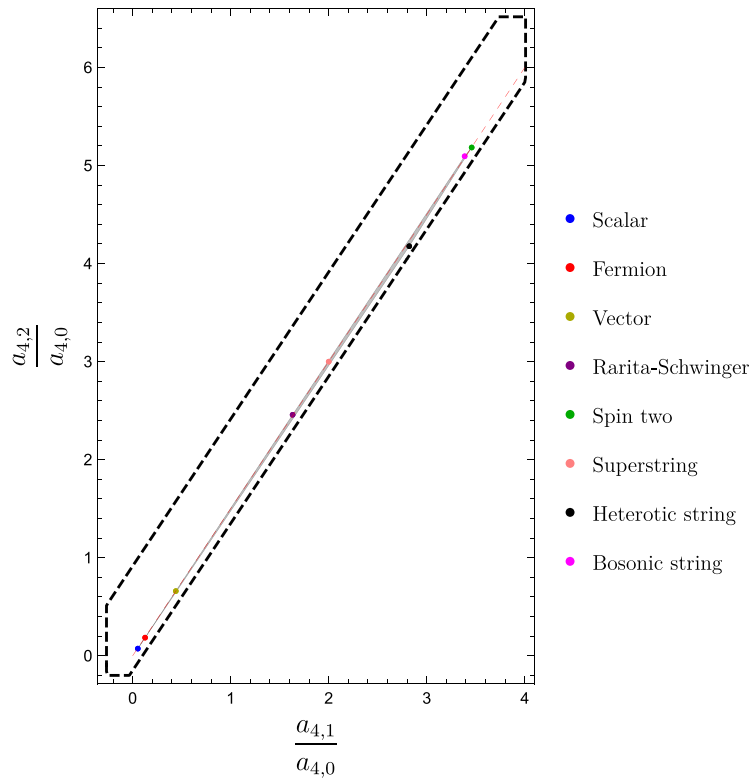
Figure 10. A scaled version of the theory island from figure 9. The dashed black line bounds the region which is found from equations (5.33) and (5.35) using the LSD assumption (5.26) with $\alpha = 10^2$. The dashed red line $a_{4,2} = \frac{3}{2}a_{4,1}$ corresponds to $\alpha = \infty$ bound (5.38). Remarkably, all known theories lie in this small region where they populate a small region around the $\alpha = \infty$ curve.

terms to 0 in the formulas above we get the region (5.38). *A priori* it is not clear that this follows from sending $\alpha \to \infty$ in equation (5.26) since we still have infinitely many spins that can potentially compensate for the smallness of $1/\alpha$. Crossing symmetry, however, guarantees that it is indeed the case and higher-spin contributions cannot compensate for smallness of $1/\alpha$ as can be seen from the explicit analysis using the finite number of partial-wave bounds similar to equation (5.35).

Note that there are two distinct ways in which points can be close to the straight line in figure 10. First, as explained above by assuming the strong version of LSD in $a_{4,1}$ and $a_{4,2}$. Another mechanism is to have $a_{4,0} \gg a_{4,1}, a_{4,2}$ which happens when the contribution of ρ_{++0} to $a_{4,0}$ dominates over the partial waves that enter to $a_{4,1}$ and $a_{4,2}$. In this case the points appear very close to the origin on the $(\frac{a_{4,1}}{a_{4,0}}, \frac{a_{4,2}}{a_{4,0}})$ plot and the slope defined by the ratio $\frac{a_{4,2}}{a_{4,1}}$ is not directly visible, nor does it need to be $\frac{3}{2}$. In fact, $a_{4,1}$ and $a_{4,2}$ have nontrivial higher-spin corrections in equation (5.39), i.e. ρ_2^{++} is sizable compared to ρ_4^{+-} and contributes ruining the $3/2$ ratio for $\frac{a_{4,2}}{a_{4,1}}$. This is precisely what happens for the scalar and fermion contributions in figure 10. In this case, by looking at figure 10 it is not apparent that equation (5.39) is not an accurate description of these points since they are close to the origin, which is on the line. When this happens the simple model of dropping higher-spin contributions in equation (5.39)

is too crude. For example, consider the one-loop massive scalar amplitude which lies very close to the origin in figure 10. In this case to accurately capture $a_{4,1}$ we also need to include $\rho_2^{++}(m^2)$ (which is comparable to $\rho_4^{+-}(m^2)$ in this case) in the formulas (5.39). For $a_{4,2}$ to get 1% precision $\rho_4^{++}(m^2)$ required as well. We re-iterate that in deriving the bounds using LSD we did not make any assumptions about the contributions of infinitely many higher-spin terms in equation (5.39) and instead used crossing symmetry to derive rigorous and tight bounds based on equation (5.26). In contrast, when we use equation (5.39) and try to estimate the contribution of higher-spin partial waves we observe that depending on the model and the coupling at hand, the number of terms required to be kept in the expansion to reach good precision varies.

We observe that the relation (5.38) generates a hierarchy $\sim 10^{-2}$ between certain coefficients in the low-energy EFT. A naive low-energy observer could have been puzzled by the fact that $|a_{4,2} - \frac{3}{2}a_{4,1}| \ll a_{4,0}$. We see that this hierarchy is generated by unitarity, namely it appears due to the dominance of the dispersive integrals by the low-spin partial waves. We discuss this point further in section 5.5.

k = 6:

The analysis gets more and more complicated as we go to higher k. For $k = 6$ we get the s-dimensional coupling vector $\mathbf{a_6} = (a_{6,0}, a_{6,1}, a_{6,2}, a_{6,3}, a_{6,4}, a_{6,5}, a_{6,6})$. Out of seven couplings only four are independent due to the crossing-symmetry relations

$$\text{Crossing}: a_{6,4} = 5(a_{6,1} - a_{6,2}) + 3a_{6,3},$$
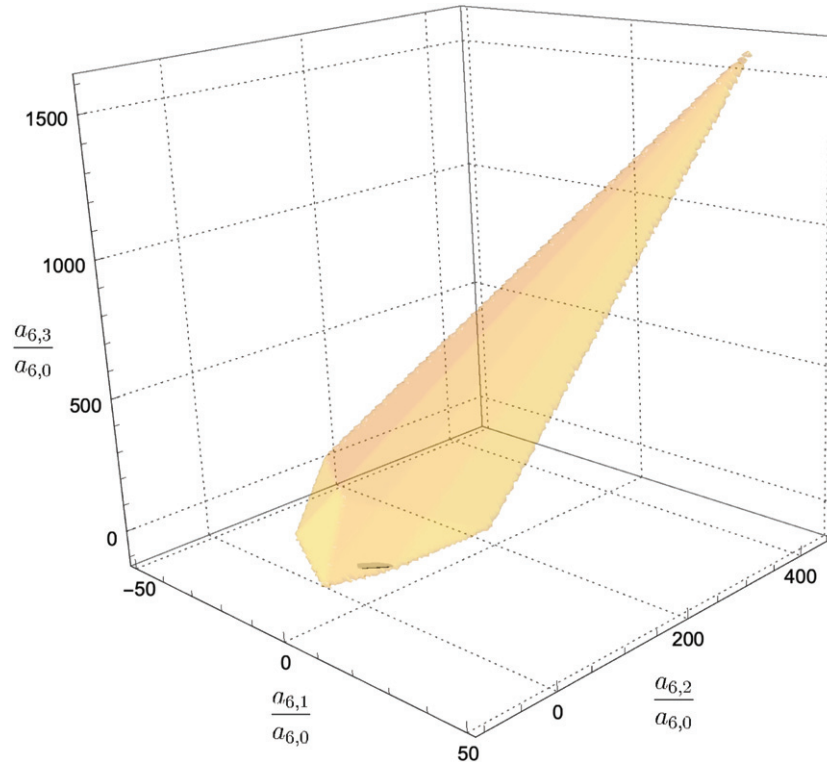$$a_{6,5} = 6(a_{6,1} - a_{6,2}) + 3a_{6,3}, \tag{5.40}$$
$$a_{6,6} = 2(a_{6,1} - a_{6,2}) + a_{6,3}.$$

Working out the boundary of the coupling region is more laborious, but follows the same pattern that we observed before

$$\{\langle \mathbf{a_6}, i_s, i_s + 2, j_s, j_s + 2, k_s, k_s + 2\rangle_{k>j>i\geqslant 2},$$
$$\langle \mathbf{a_6}, i_u^{(6)} + 1, i_u^{(6)}, j_u^{(6)}, j_u^{(6)} + 1, k_u^{(6)}, k_u^{(6)} + 1\rangle_{k>j>i\geqslant 5}, \tag{5.41}$$
$$\langle \mathbf{a_6}, 5_u^{(6)}, 4_u^{(6)}, 6_u^{(6)}, 7_u^{(6)}, 8_u^{(6)}, 9_u^{(6)}\rangle, \text{ mixed } s-u \text{ constraints}\} \geqslant 0,$$

where we list the explicit mixed constraints that we found in appendix C. After the boundaries are identified we take the crossing-symmetric slice (5.40). The resulting region of allowed couplings is plotted in figure 11. The region covered by the known theories is again a small island in the space of couplings.

As in the analysis above the island occupied by the explicit examples can be understood using the idea of strong LSD. As before the first step is to derive a set of rigorous bounds in terms of the low-spin partial waves. We do not present the complete analysis here but only present some of the relevant bounds

$$-\frac{9\langle\rho_2^{++}\rangle_6 + 25\langle\rho_5^{+-}\rangle_6}{\langle\rho_0^{++}\rangle_6 + \langle\rho_4^{+-}\rangle_6 + \langle\rho_2^{++}\rangle_6 + \langle\rho_5^{+-}\rangle_6}$$
$$\leqslant \frac{a_{6,2} - \frac{5}{2}a_{6,1}}{a_{6,0}}$$
$$\leqslant \frac{40\langle\rho_4^{++}\rangle_6 + 315\langle\rho_6^{++}\rangle_6 + 32.56\langle\rho_6^{+-}\rangle_6 + 220.41\langle\rho_7^{+-}\rangle_6}{\langle\rho_0^{++}\rangle_6 + \langle\rho_4^{+-}\rangle_6 + \langle\rho_4^{++}\rangle_6 + \langle\rho_6^{++}\rangle_6 + \langle\rho_6^{+-}\rangle_6 + \langle\rho_7^{+-}\rangle_6}, \tag{5.42}$$

Figure 11. The allowed region for couplings $(\frac{a_{6,1}}{a_{6,0}}, \frac{a_{6,2}}{a_{6,0}}, \frac{a_{6,3}}{a_{6,0}})$ derived from constraints (5.41). The black little island (barely visible) is the space occupied by known perturbative amplitudes. The small gray shaded region that surrounds it corresponds to the LSD $\alpha = 10^2$ bound. The details on how the plot was generated can be found in appendix C.

$$
- \frac{20\langle\rho_2^{++}\rangle_6 + 66.67\langle\rho_5^{+-}\rangle_6}{\langle\rho_0^{++}\rangle_6 + \langle\rho_4^{+-}\rangle_6 + \langle\rho_2^{++}\rangle_6 + \langle\rho_5^{+-}\rangle_6}
$$

$$
\leqslant \frac{a_{6,3} - \frac{10}{3}a_{6,1}}{a_{6,0}}
$$

$$
\leqslant \frac{73.34\langle\rho_4^{++}\rangle_6 + 1540\langle\rho_6^{++}\rangle_6 + 163.49\langle\rho_6^{+-}\rangle_6 + 495.64\langle\rho_7^{+-}\rangle_6}{\langle\rho_0^{++}\rangle_6 + \langle\rho_4^{+-}\rangle_6 + \langle\rho_4^{++}\rangle_6 + \langle\rho_6^{++}\rangle_6 + \langle\rho_6^{+-}\rangle_6 + \langle\rho_7^{+-}\rangle_6}. \tag{5.43}
$$

From the formulas above and a similar analysis for $a_{6,i}$ we also see that

$$
\text{LSD}_{\alpha\to\infty}: \quad \frac{a_{6,2} - \frac{5}{2}a_{6,1}}{a_{6,0}} = 0, \quad \frac{a_{6,3} - \frac{10}{3}a_{6,1}}{a_{6,0}} = 0,
$$

$$
0 \leqslant \frac{a_{6,1}}{a_{6,0}} \leqslant 6, \quad 0 \leqslant \frac{a_{6,2}}{a_{6,0}} \leqslant 15, \tag{5.44}
$$

$$
0 \leqslant \frac{a_{6,3}}{a_{6,0}} \leqslant 20.
$$

As for $k = 4$ we can understood the result above by simply dropping higher-spin contributions in the dispersive representations for the couplings and keeping only the lowest-spin partial waves in each channel

$$a_{6,0} = \int_{m_{\text{gap}}^2}^{\infty} \frac{\mathrm{d}m^2}{\pi} \frac{\rho_{++0}(m^2)}{(m^2)^{11}} + \int_{m_{\text{gap}}^2}^{\infty} \frac{\mathrm{d}m^2}{\pi} \frac{\rho_{+-4}(m^2)}{(m^2)^{11}} + \text{higher spin}, \tag{5.45}$$

$$a_{6,1} = 6 \int_{m_{\text{gap}}^2}^{\infty} \frac{\mathrm{d}m^2}{\pi} \frac{\rho_{+-4}(m^2)}{(m^2)^{11}} + \text{higher spin}, \tag{5.46}$$

$$a_{6,2} = 15 \int_{m_{\text{gap}}^2}^{\infty} \frac{\mathrm{d}m^2}{\pi} \frac{\rho_{+-4}(m^2)}{(m^2)^{11}} + \text{higher spin}, \tag{5.47}$$

$$a_{6,3} = 20 \int_{m_{\text{gap}}^2}^{\infty} \frac{\mathrm{d}m^2}{\pi} \frac{\rho_{+-4}(m^2)}{(m^2)^{11}} + \text{higher spin}. \tag{5.48}$$

We emphasize again that while dropping infinitely many terms is not justified, the bound (5.44) is rigorous.

Assuming the strong version of the LSD and using equation (5.42) we again find the small region around the island occupied by the explicit examples, as shown in figures 11 and 12.

k = 8:

For $k = 8$ we do not perform the analysis of finding the allowed region but simply report on the data in the specific theories. There are five independent couplings at this level which we choose to be $a_{8,0 \leqslant j \leqslant 4}$. Other couplings $a_{8,5 \leqslant j \leqslant 8}$ can be obtained by crossing. The dispersive representation of the couplings take the following form

$$a_{8,0} = \int_{m_{\text{gap}}^2}^{\infty} \frac{\mathrm{d}m^2}{\pi} \frac{\rho_{++0}(m^2)}{(m^2)^{13}} + \int_{m_{\text{gap}}^2}^{\infty} \frac{\mathrm{d}m^2}{\pi} \frac{\rho_{+-4}(m^2)}{(m^2)^{13}} + \text{higher spin},$$

$$a_{8,1} = 8 \int_{m_{\text{gap}}^2}^{\infty} \frac{\mathrm{d}m^2}{\pi} \frac{\rho_{+-4}(m^2)}{(m^2)^{13}} + \text{higher spin},$$

$$a_{8,2} = 28 \int_{m_{\text{gap}}^2}^{\infty} \frac{\mathrm{d}m^2}{\pi} \frac{\rho_{+-4}(m^2)}{(m^2)^{13}} + \text{higher spin}, \tag{5.49}$$

$$a_{8,3} = 56 \int_{m_{\text{gap}}^2}^{\infty} \frac{\mathrm{d}m^2}{\pi} \frac{\rho_{+-4}(m^2)}{(m^2)^{13}} + \text{higher spin},$$

$$a_{8,4} = 70 \int_{m_{\text{gap}}^2}^{\infty} \frac{\mathrm{d}m^2}{\pi} \frac{\rho_{+-4}(m^2)}{(m^2)^{13}} + \text{higher spin},$$

where we keep the leading spin contributions in both channels. To visualize the data we consider two three-dimensional slices of the space of couplings with coordinates $(\frac{a_{8,1}}{a_{8,0}}, \frac{a_{8,2}}{a_{8,0}}, \frac{a_{8,3}}{a_{8,0}})$ and $(\frac{a_{8,2}}{a_{8,0}}, \frac{a_{8,3}}{a_{8,0}}, \frac{a_{8,4}}{a_{8,0}})$. Formulas (5.49) define a line in this space upon neglecting the higher-spin partial wave contributions. We depict the result in figure 13 and it is again completely analogous to our observations for lower k's. It would be interesting to extend the analysis done for lower k to this case as well.
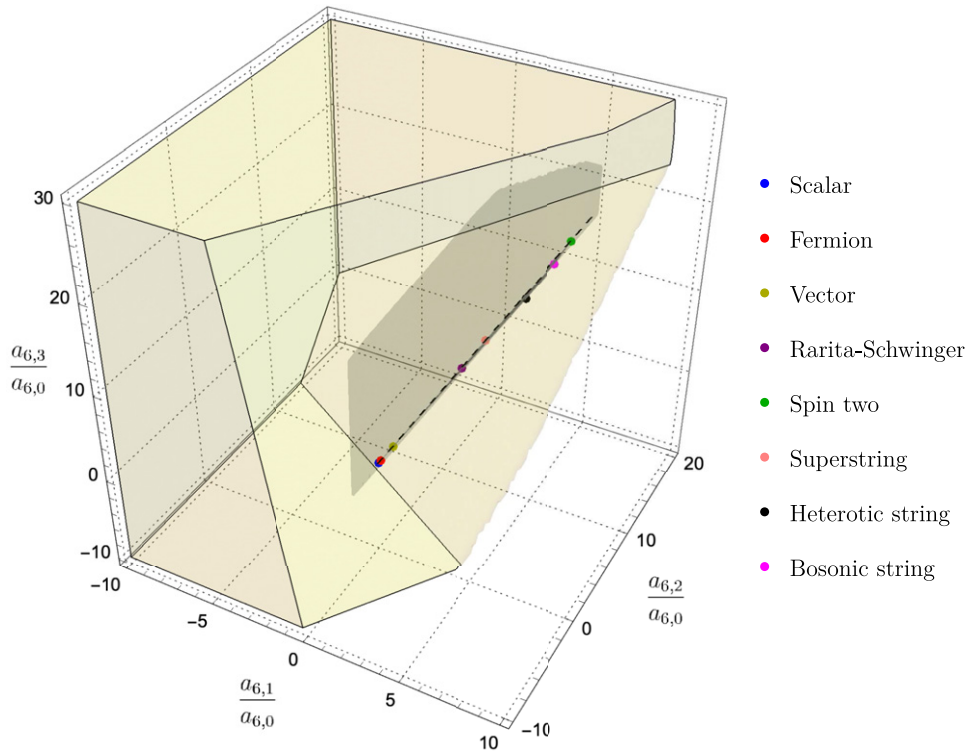
Figure 12. The scaled version of the theory island from figure 11. The gray region is derived using the bounds (5.42) (and similar bounds for $a_{6,i}$ which we do not write down explicitly) and the LSD assumption (5.26) with $\alpha = 10^2$. The dashed black line $\frac{a_{6,2}}{a_{6,1}} = \frac{5}{2}$ and $\frac{a_{6,3}}{a_{6,1}} = \frac{10}{3}$ corresponds to $\text{LSD}_{\alpha \to \infty}$ bounds given in equation (5.44).

5.3. Crossing-symmetric dispersive representation of low-energy couplings

We can also analyze the same amplitude using dispersive representation with a different pair of channels. We start by recalling that

$$\mathcal{M}_4(1^+, 2^-, 3^-, 4^+) = (\langle 23 \rangle [14])^4 f(s, u). \tag{5.50}$$

We would like to use dispersion relations at fixed t to derive the dispersive representation for the low-energy couplings. For this purpose it is natural to introduce the following parameterization of the low-energy expansion of $f(s, u)$

$$f\left(z - \frac{t}{2}, -z - \frac{t}{2}\right) = \left(\frac{\kappa}{2}\right)^2 \frac{1}{t(\frac{t^2}{4} - z^2)} + |\beta_{R^3}|^2 \frac{(\frac{t^2}{4} - z^2)}{t} - |\beta_\phi|^2 \frac{1}{t}$$

$$+ \sum_{k \geqslant q \geqslant 0}^{\infty} \tilde{a}_{k,q} z^{k-q} t^q, \qquad z = s + \frac{t}{2}, \quad k - q \in 2\mathbb{Z}_{\geqslant 0}. \tag{5.51}$$

Crossing symmetry $f(s, u) = f(u, s)$ acts as $z \to -z$ and therefore constrains $k - q$ to be even.
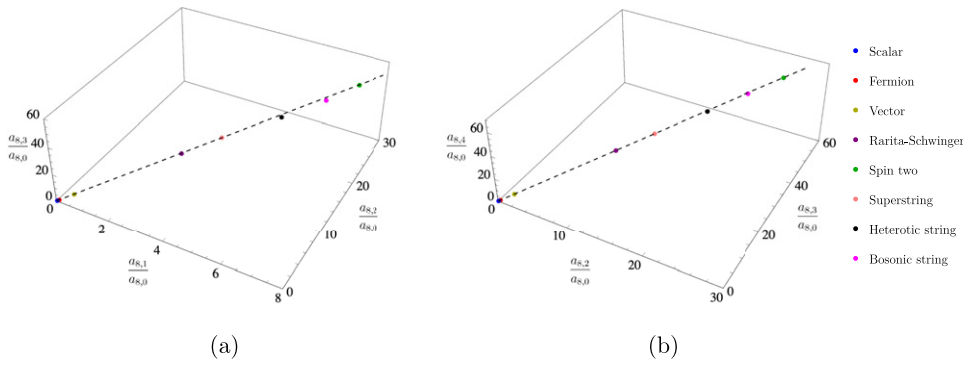
Figure 13. A plot for the data for $a_{8,j}$ in various theories. In panel (a) we plot $(\frac{a_{8,1}}{a_{8,0}}, \frac{a_{8,2}}{a_{8,0}}, \frac{a_{8,3}}{a_{8,0}})$ and in panel (b) $(\frac{a_{8,2}}{a_{8,0}}, \frac{a_{8,3}}{a_{8,0}}, \frac{a_{8,4}}{a_{8,0}})$. The dashed line corresponds to the low-spin dominant line defined by equation (5.49) upon neglecting the higher-spin contributions to the partial-wave expansion.

To derive the bounds we can try to use dispersive representation for $\tilde{a}_{k,q}$, where we keep t fixed and deform the z integral

$$
\tilde{a}_{k,q} = \frac{1}{q!} \partial_t^q \oint \frac{dz}{2\pi i} \frac{1}{z^{k-q+1}} \left[f(z - \frac{t}{2}, -z - \frac{t}{2}) - \left(\frac{\kappa}{2}\right)^2 \frac{1}{t(\frac{t^2}{4} - z^2)} \right.
$$
$$
\left. - |\beta_{R^3}|^2 \frac{(\frac{t^2}{4} - z^2)}{t} + |\beta_\phi|^2 \frac{1}{t} \right]\Bigg|_{t=0} ,
\tag{5.52}
$$

where the contour integral encircles the origin and we explicitly subtracted the contributions that are singular at $t = 0$. As discussed in the previous section, unitarity constrains the form of the discontinuity of $f(z - \frac{t}{2}, -z - \frac{t}{2})$. Therefore, in evaluating equation (5.52) we can open the contour and assuming we can drop the arcs at infinity to arrive at the following representation

$$
\tilde{a}_{k,q} = \frac{2}{q!} \partial_t^q \int_{m_{\text{gap}}^2}^\infty \frac{dm^2}{\pi} \sum_{J=4}^\infty \frac{\rho_J^{+-}(m^2)}{(m^2 + \frac{t}{2})^{k-q+1}} \frac{(-1)^J d_{4,-4}^J \left(1 + \frac{2t}{m^2}\right)}{t^4} \Bigg|_{t=0} ,
\tag{5.53}
$$
$$
k - q \in 2\mathbb{Z}_{\geqslant 0} .
$$

The factor of 2 originates from the sum over the s-channel and the u-channel discontinuities. For odd $k - q$ they cancel each other and we get zero. An important factor $1/t^4$ originates from the fact that unitarity constraints are formulated in terms of $\mathcal{M}_{+--+}$ which includes the prefactor $(\langle 23\rangle[14])^4$—see section 4.2 for details. The factor $(-1)^J$ can be understood from the fact that with a given choice of helicity the discontinuity is positive for the forward limit $u = 0$ (as opposed to $t = 0$).[31] In using the representation (5.53) we should not forget that it was derived assuming that the Regge behavior of the amplitude is such that the arcs at infinity can be dropped. In particular, given that $f(z - \frac{t}{2}, -z - \frac{t}{2}) \sim z^{J_0}$ for large $|z|$, and taking into account the subtractions of terms that are singular at $t = 0$ the representation (5.53) is valid only for $k - q > \max\{J_0, 2\}$. As opposed to the previous section all $\tilde{a}_{k,q}$, whose dispersive

---

[31] Consistency between equations (5.7) and (5.53), namely matching of the ρ_J^{+-} discontinuities, requires that $d_{4,4}^J(x) = (-1)^J d_{4,-4}^J(-x)$, which is indeed the case.

representation is given in equation (5.53), are independent and there are no extra constraints coming from crossing.

The presence of $(-1)^J$ in the sum (5.53) prevents us from deriving useful bounds from the representation (5.53). We, however, present the data for $\tilde{a}_{k,q}$ obtained from the explicit amplitudes since it reveals an interesting aspect of the discussion in the previous section.

For example, consider $k = 6$. The mapping between $a_{6,j}$ and $\tilde{a}_{6,j}$ takes the following form

$$
\begin{aligned}
\tilde{a}_{6,0} &= 2(a_{6,1} - a_{6,2}) + a_{6,3}, \\
\tilde{a}_{6,2} &= \frac{1}{4}(-10a_{6,1} + 10a_{6,2} - 3a_{6,3}), \\
\tilde{a}_{6,4} &= \frac{1}{16}(30a_{6,1} - 14a_{6,2} + 3a_{6,3}),
\end{aligned}
\tag{5.54}
$$

and we will not need $\tilde{a}_{6,6}$ for our purposes. The remarkable fact about $\tilde{a}_{6,j}$ is that the dashed LSD line from figure 12 maps to the point $(\frac{15}{4}, \frac{15}{16})$ in the $(\frac{\tilde{a}_{6,2}}{\tilde{a}_{6,0}}, \frac{\tilde{a}_{6,4}}{\tilde{a}_{6,0}})$ plane, which can also be found by keeping the lowest-spin contribution in equation (5.53)

$$
\text{LSD}_{\alpha \to \infty} : \left(\frac{\tilde{a}_{6,2}}{\tilde{a}_{6,0}}, \frac{\tilde{a}_{6,4}}{\tilde{a}_{6,0}} \right) = \left(\frac{15}{4}, \frac{15}{16} \right) \simeq (3.75, 0.94).
\tag{5.55}
$$

Therefore, the plane $(\frac{\tilde{a}_{6,2}}{\tilde{a}_{6,0}}, \frac{\tilde{a}_{6,4}}{\tilde{a}_{6,0}})$ happens to be precisely orthogonal to the line in figure 12 and it is well-suited to study the fine structure of the distribution of points around the line.

Looking at the plot 12, we see that both the scalar and the fermion lie pretty far from the naive LSD point. The reason we did not detect this on figure 12 is that for the scalar and fermion: $a_{6,0} \gg a_{6,1}, a_{6,2}, a_{6,3}$. In this way both points are very close to the LSD line in figure 12, but for the trivial reason of being close to the origin. In figure 14 we resolve the origin by switching to variables (5.54) that do not depend on $a_{6,0}$.

Working with $\tilde{a}_{k,q}$ emphasizes an important aspect of LSD. While assuming strong LSD leads to much tighter bounds on ratios $\frac{a_{k,j}}{a_{k,0}}$ it does not lead to improved bounds for $\frac{\tilde{a}_{k,q}}{\tilde{a}_{k,0}}$. The reason can be understood as follows: for any large, but finite α, the admissible range for $\frac{a_{k,j}}{a_{k,0}}$ includes 0 and a small part of the negative axis. It is then easy to see from the definition (5.54) that arbitrarily small vicinity of the origin in the $\frac{a_{k,j}}{a_{k,0}}$ includes all possible values for $\frac{\tilde{a}_{k,q}}{\tilde{a}_{k,0}}$. This precludes using LSD to derive stronger bounds on $\frac{\tilde{a}_{k,q}}{\tilde{a}_{k,0}}$. This simple fact highlights the point that not all EFT coupling bases are equally illuminating for understanding the underlying structure.

A similar analysis can be performed for $k = 8$. By switching to the $(\frac{\tilde{a}_{8,2}}{\tilde{a}_{8,0}}, \frac{\tilde{a}_{8,4}}{\tilde{a}_{8,0}}, \frac{\tilde{a}_{8,6}}{\tilde{a}_{8,0}})$ coordinates we note that the line from figure 13 maps to a single point. The same point can be obtained by keeping only $J = 4$ contribution in (5.53)

$$
\text{LSD}_{\alpha \to \infty} : \left(\frac{\tilde{a}_{8,2}}{\tilde{a}_{8,0}}, \frac{\tilde{a}_{8,4}}{\tilde{a}_{8,0}}, \frac{\tilde{a}_{8,6}}{\tilde{a}_{8,0}} \right) = (7, \frac{35}{8}, \frac{7}{16}) = (7, 4.375, 0.4375).
\tag{5.56}
$$

Therefore by plotting the data in these coordinates we can resolve the points close to the origin in figure 13 which were located there due to the fact that $a_{8,0} \gg a_{8,i>0}$. One can check again that $(\frac{\tilde{a}_{8,2}}{\tilde{a}_{8,0}}, \frac{\tilde{a}_{8,4}}{\tilde{a}_{8,0}}, \frac{\tilde{a}_{8,6}}{\tilde{a}_{8,0}})$ is independent of $a_{8,0}$ and therefore this suppression does not take place anymore. The result is depicted in figure 15.
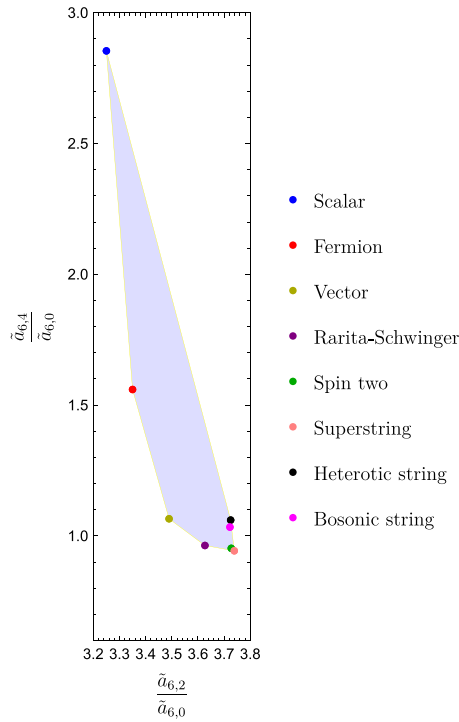
Figure 14. The transverse view of the theory island from figure 12. We indicate various theories that form the vertices of the island. To reach a particular point inside the UV island we need to take a superposition of various amplitudes.

5.4. Spectral densities and low-spin dominance

In the discussion above we considered various bounds that follow from the dispersive representation of the low-energy couplings and observed the phenomenon of LSD in the known physical theories. It manifests itself in the fact that the low-energy couplings occupy a very small region in the space of couplings allowed on general grounds. To see this phenomenon more clearly it is instructive to look at the spectral densities of the amplitudes directly.

Given a known expression for the amplitude it is not hard to compute various spectral densities. Indeed, it amounts to taking a discontinuity of the amplitude and integrating it against the proper Wigner d-function. These satisfy a familiar orthogonality relation

$$\int_{-1}^{1} dx \, d_{\lambda,\lambda'}^{J}(x) d_{\lambda,\lambda'}^{\tilde{J}}(x) = \frac{2}{2J+1} \delta_{J,\tilde{J}} . \tag{5.57}$$

For tree-level amplitudes the spectral density is a sum of delta-functions which correspond to the masses of exchanged particles. For the one-loop amplitudes the discontinuity has a continuous support above the two-particle threshold $4m^2$. In this case we perform the integration (5.57) numerically. We focus on the first few spin spectral densities in both ρ_J^{++} and ρ_J^{+-} channels. To go from spectral densities to the coefficients in the low-energy expansion we need to compute the moments as in equations (5.39), (5.45), and (5.49).

The results for a few selected cased are listed in figures 16–19. In all cases we see that the minimal-spin partial waves dominate in the corresponding channel. It is also instructive to plot
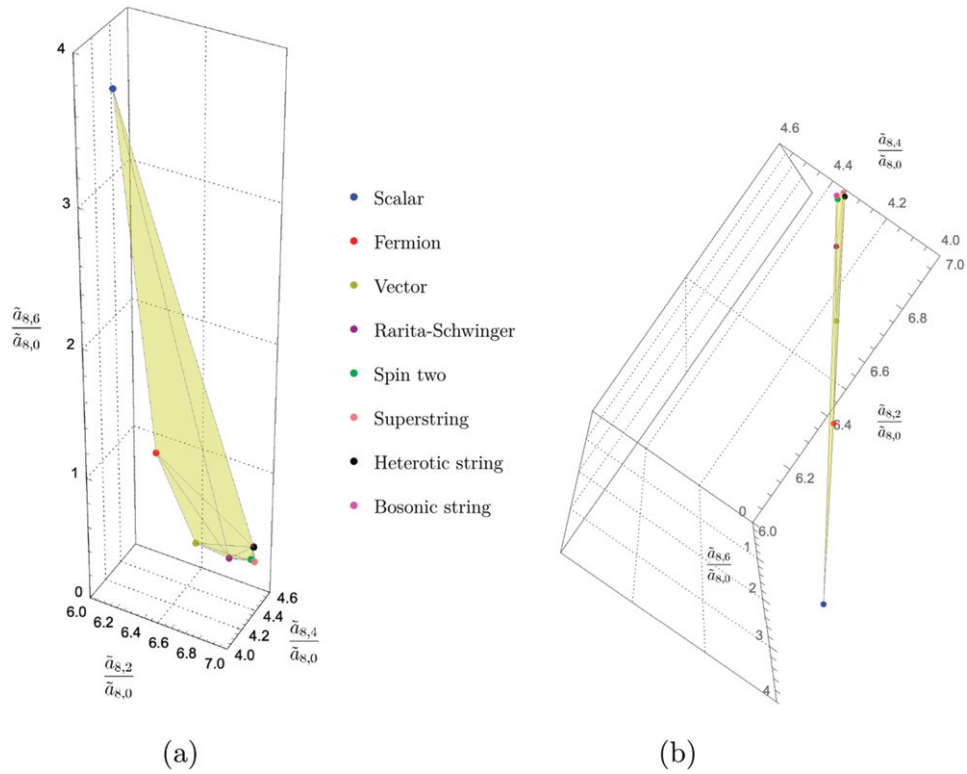
Figure 15. Various data points for $(\frac{\tilde{a}_{8,2}}{\tilde{a}_{8,0}}, \frac{\tilde{a}_{8,4}}{\tilde{a}_{8,0}}, \frac{\tilde{a}_{8,6}}{\tilde{a}_{8,0}})$. Panels (a) and (b) present the same plot as seen from different vantage points. In particular, panel (b) makes it clear that the points essentially lie on a plane.

the moments $\langle \rho_J \rangle_k$ which we present in figure 20. The moments clearly satisfy LSD used in the previous section to derive stronger bounds.

In fact the moments of spectral density $\langle \rho_J \rangle_k$ in the examples we consider not only exhibit the dominance of the lowest-spin partial waves but also rapid decay at higher J. This latter feature is expected to be completely general. Indeed, the convergence of the sum rules (5.8) requires that $\langle \rho_J \rangle_k$ decay faster than any polynomial at large J. This latter decay can be traced to locality of scattering in the impact-parameter space.

Indeed, in the perturbative regime at large J we have

$$\rho_J(s) \sim \mathrm{Im}\, \delta(s, b), \qquad b = \frac{2J}{\sqrt{s}}, \tag{5.58}$$

where $\delta(s, b)$ is the phase shift and b is the impact parameter. For fixed s and large impact parameters we expect to have

$$\mathrm{Im}\, \delta(s, b) \sim \mathrm{e}^{-m_{\mathrm{gap}} b}, \quad m_{\mathrm{gap}} b \gg 1, \tag{5.59}$$

which controls decay of $\rho_J(s)$ at large J. In string theory the leading-order behavior is different and is controlled by the transverse spreading of strings $\mathrm{Im}\, \delta(s, b) \sim \mathrm{e}^{-\frac{b^2}{2\alpha' \log \frac{s\alpha'}{4}}}$ —see reference [78].
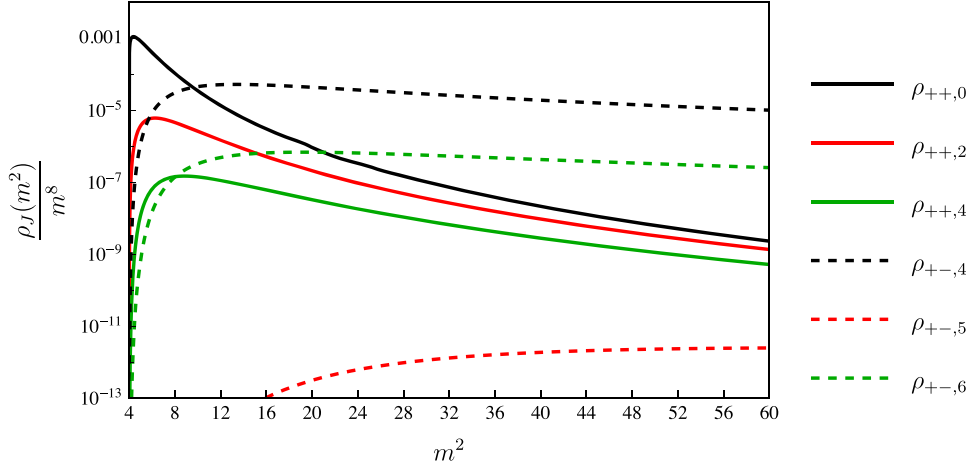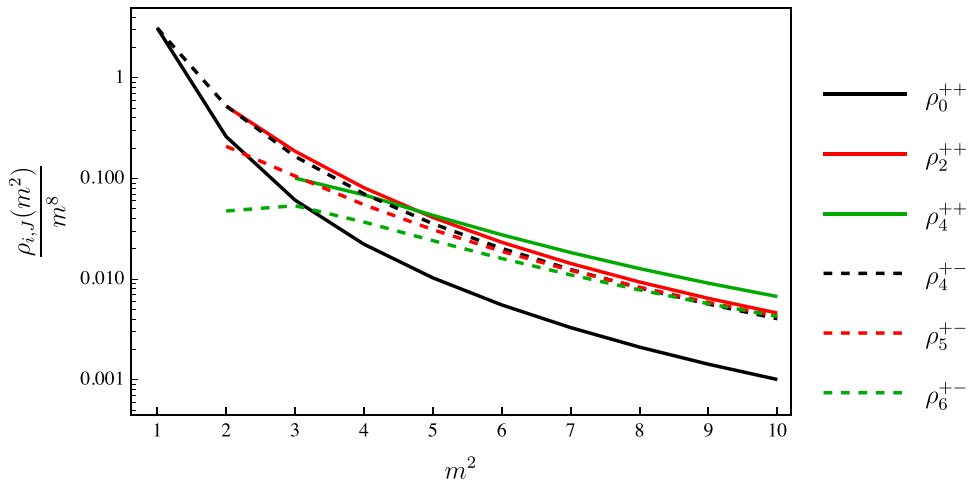
Figure 16. Spectral densities for the one-loop minimally-coupled scalar. We observe that $\rho_0^{++} \gg \rho_{J\geqslant 4}^{+-}, \rho_{J\geqslant 2}^{++}$ and that $\rho_4^{+-} \sim \rho_2^{++}$ close to the two-particle threshold. This is fully consistent with the features of the plots for various couplings in the previous section.
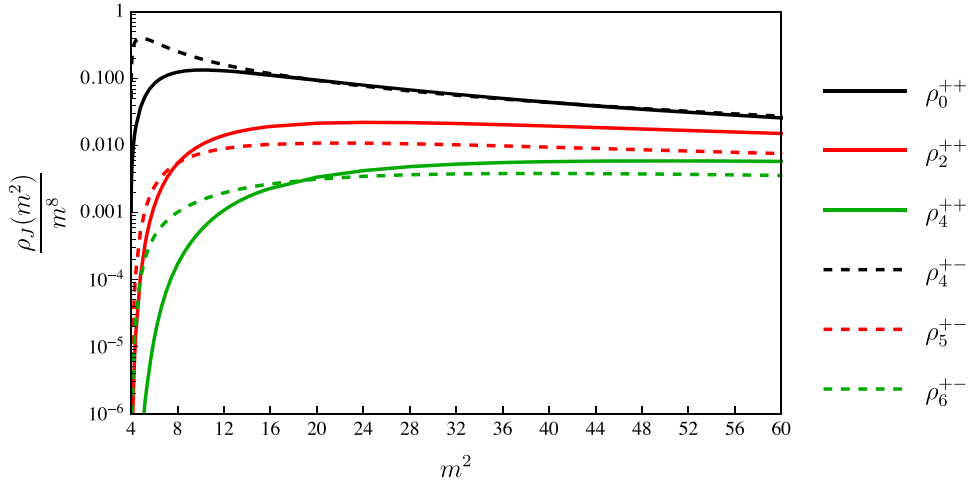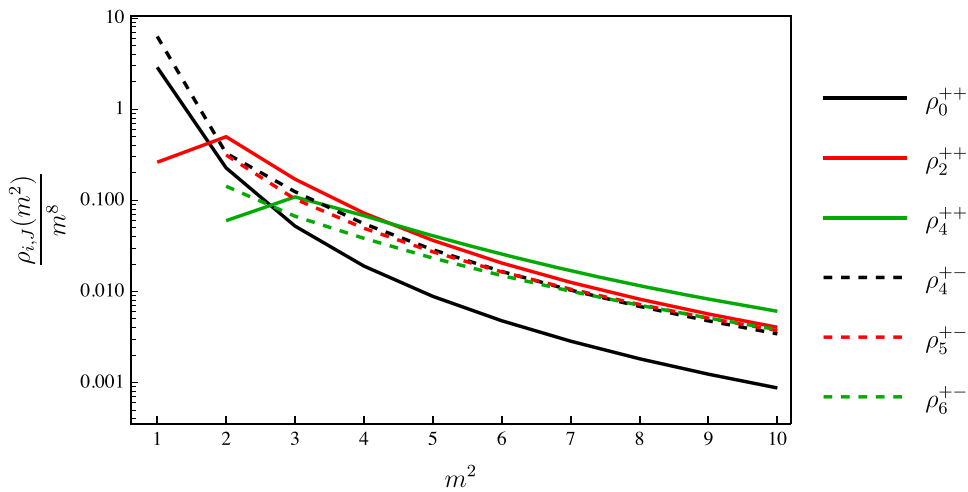

Figure 17. Spectral densities for the one-loop minimally coupled spin-2 particle. We observe that $\rho_4^{+-} \gg \rho_{J>4}^{+-}, \rho_{J\geqslant 2}^{++}$ and that $\rho_4^{+-} \sim \rho_0^{++}$. Therefore in the space of couplings this amplitude is expected to lie on the LSD line.

A priori the large impact parameter discussion is not necessarily relevant for understanding the large-J behavior of $\langle\rho_J\rangle_k$ which involves computing the moment (5.22) over all energies (as opposed to keeping s fixed as we take the large-J limit). However, in analyzing $\langle\rho_J\rangle_k$ for the amplitudes considered in the present paper we experimentally observed that the integral over energies is peaked at energies

$$\langle\rho_J\rangle_k : \frac{s_*}{m_{\text{gap}}^2} \sim J. \tag{5.60}$$

Figure 18. Spectral densities for the tree-level scattering of gravitons in the superstring theory. We observe that $\rho_4^{+-} \gg \rho_{J>4}^{+-}, \rho_{J\geqslant2}^{++}$ and that $\rho_4^{+-} \sim \rho_0^{++}$. Therefore in the space of couplings this amplitude is expected to lie on the LSD line.

Figure 19. Spectral densities for the tree-level scattering of gravitons in the heterotic string theory. We observe that $\rho_4^{+-} \gg \rho_{J>4}^{+-}, \rho_{J\geqslant2}^{++}$ and that $\rho_4^{+-} \sim \rho_0^{++}$. Therefore in the space of couplings this amplitude is expected to lie on the LSD line.

Plugging this into the formula for the impact parameter (5.58) we find that the dominant impact parameters are $m_{\text{gap}}b_* \sim J^{1/2} \gg 1$ and therefore the large-J behavior of the moments $\langle\rho_J\rangle$ is still controlled by large impact-parameter physics. We tested this picture against the data presented in figure 20 and found a qualitative agreement. It would be interesting to study the large-J limit of $\langle\rho_J\rangle_k$ more systematically. Of course, this discussion does not explain the fact that the hierarchical structure among partial waves continues all the way to the lowest spins in the examples we analyzed. It is this latter fact was crucial for the analysis in the previous section.
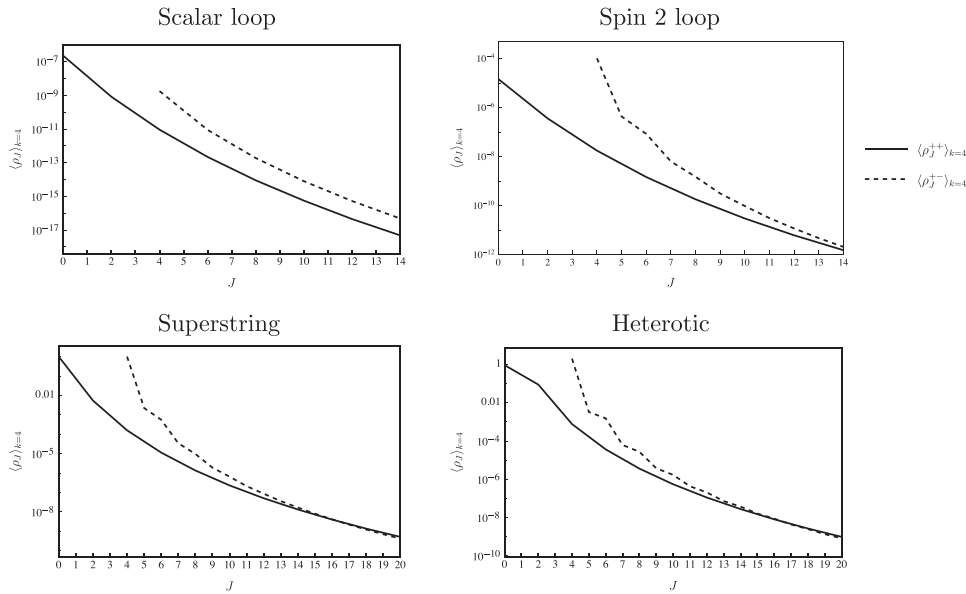
Figure 20. Moments of the spectral density $\langle \rho_J \rangle_4$ (5.22) as a function of spin J for various examples. For the scalar one loop result ρ_J^{+-} with J being odd are negligible and are not presented in the figure. The line in the corresponding panel is obtained by connecting the even spin values of the spectral density moments.

An important question is to understand how general is the picture that we observed in the tree-level string amplitudes and one-loop matter amplitudes. *A priori* these amplitudes look very different from each other, but at the level of the partial-wave analysis discussed in this section they exhibit remarkably similar behavior and strong version of LSD. This suggests that the hierarchical structures we observed in this paper could be a general property of consistent weakly-coupled gravitational S-matrices, but we do not have a proof yet.
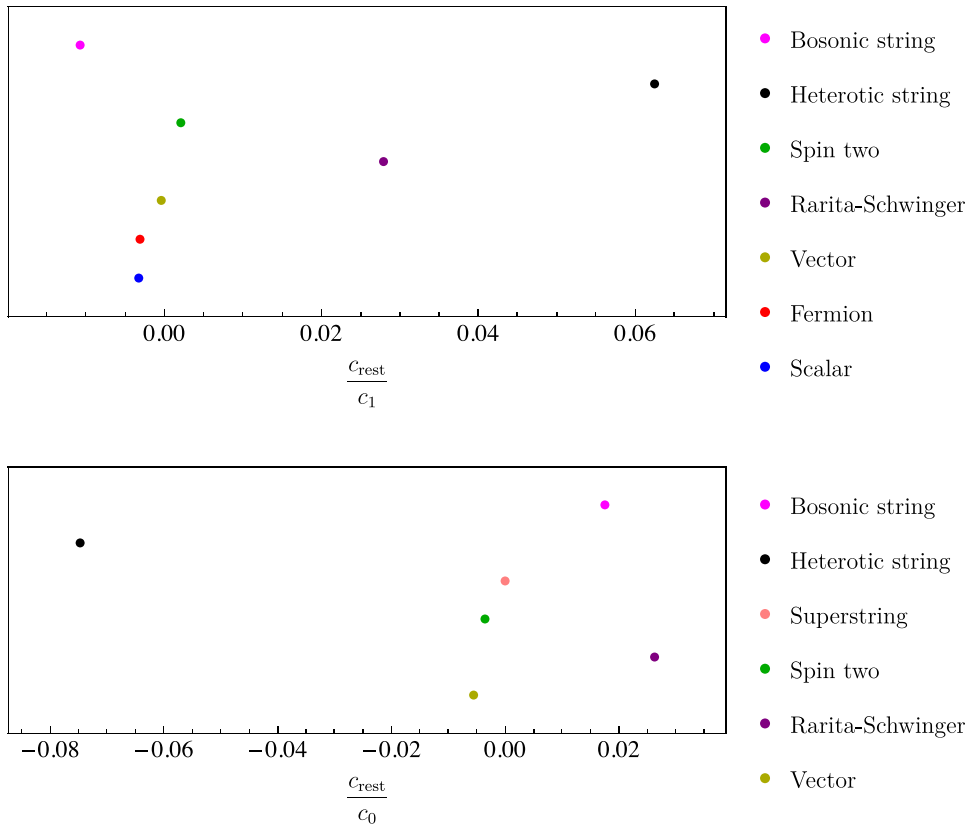
5.5. A hierarchy from unitarity

The LSD discussion above illustrates an interesting phenomenon of emergent hierarchy between EFT coefficients in the absence of an underlying symmetry. Consider for example $D^8 R^4$ type corrections discussed in the $k = 4$ section above. In writing down the relevant correction to the amplitude an EFT practitioner guided by the considerations of the supersymmetric decomposition (2.32) can write the following ansatz

$$f_{D^8 R^4}(t, u) = c_0 \, (s^2 + t^2 + u^2)^2 + c_1 \, s^4 + c_{\text{rest}} \, (t^2 + u^2)^2. \tag{5.61}$$

The first term, which is proportional to c_0, is the completely crossing-symmetric term of the type that appears in the new spin-2 part of the amplitude given in equation (H5), corresponding to a massive $\mathcal{N} = 8$ supersymmetric multiplet, obtained, for example, by dimensional reduction from five dimensions. In the $\mathcal{N} = 8$ theory the full crossing symmetry of such terms is a consequence of $\mathcal{N} = 8$ supersymmetric Ward identities [49]. The c_1 term corresponds to the additional terms needed in the new spin-3/2 part of the amplitude (H4) beyond the fully crossing symmetric terms already appearing in the new spin-2 part of the amplitude. Finally, the

Figure 21. Examples of a hierarchy from unitarity. The horizontal axis shows illustrates that the dominance of low-spin spectral densities naturally introduces a 10^{-2} hierarchy between various EFT coefficients in the absence of any symmetry. The vertical axis carries no meaning other than separating the points.

third term containing c_{rest} is the remaining independent term with $t{-}u$ crossing symmetry. This term is distinguished from the c_1 term by its differing behavior for $t \to \infty$ with s fixed.

We then express the coefficient above in terms of more familiar $a_{k,j}$ defined in equation (5.5) and used in the discussion of various bounds. We get the following result

$$
\begin{aligned}
c_0 &= \frac{1}{4}(2a_{4,1} - a_{4,2}), \\
c_1 &= \frac{1}{4}(4a_{4,0} - 5a_{4,1} + 2a_{4,2}), \\
c_{\text{rest}} &= \frac{1}{2}\left(a_{4,2} - \frac{3}{2}a_{4,1}\right).
\end{aligned}
\tag{5.62}
$$

Remarkably, we see that c_{rest} vanishes along the LSD line that follows from equation (5.39). Plugging in the values of $a_{k,j}$ for the theories discussed here we find a 10^{-2} hierarchy with $c_{\text{rest}} \ll c_1$ in the absence of any symmetry, as shown in figure 21. For spin $S = 1, \frac{3}{2}, 2$ particle in the loop we also have $c_{\text{rest}} \ll c_0$.

6. Deriving bounds: multiple polarizations

In this section we consider dispersion relations that combine information from the different helicity configurations. More precisely, we consider dispersive sum rules (4.24) which we apply to the matrices built out of various scattering amplitudes (4.13) and (4.14). We derive bounds on the inelastic amplitudes (single-minus and all-plus amplitudes) in terms of the elastic double-minus amplitude.

6.1. $s = 0$

The simplest bounds comes from setting $s = 0$ in equation (4.25). The relevant equation takes the form

$$\oint_{t_0} \frac{dt}{2\pi i} \mathcal{M}_h(s,t) \frac{1}{t} \frac{1}{(t(s+t))^k} \bigg|_{s=0} = \int_{m_{\text{gap}}^2}^{\infty} \frac{dt}{\pi} \frac{2}{t^{2k+1}} \text{Disc}_t \, \mathcal{M}_h(0,t) \succeq 0. \tag{6.1}$$

By plugging the low-energy expansion of the amplitude in the lhs of the equation above and requiring that its eigenvalues are non-negative we get

$$2f_{2k-4,0} \geqslant |h_{2k,0}|, \quad k = 2, 4, 6, \ldots \tag{6.2}$$

The combination above for different k form a set of moments as can be seen from the rhs of (6.1). The difference compared to the recent analysis in reference [77] is that in our case we have a positive semi-definite spectral density matrix instead of a function. By contracting such matrix moments with an arbitrary polarization vector reduces the problem to the one considered in reference [77].

The situation simplifies in the parity preserving case when $h^*(s,u) = h(s,u)$. In this case the eigenvectors $\mathcal{M}_{h,\pm}(s,t)$ given in equation (4.20) generate the low-energy expansion that satisfies the moment problem conditions considered in great detail in reference [77]. More precisely, we define a set of moments as $\mu_{k-1} \equiv 2f_{2k-4,0} - h_{2k,0}$ and consider the Hankel matrix $H_{ij} = \mu_{i+j-1}$.

Using the results in appendix H for the one-loop amplitude due to minimally-coupled scalar in the loop the first five moments take the following form

$$H^{S=0} = \frac{1}{140} \left(\frac{\kappa}{2}\right)^4 \frac{1}{(4\pi)^2} \begin{pmatrix} \dfrac{1}{135m^4} & \dfrac{1}{24\,024m^8} & \dfrac{1}{1969\,110m^{12}} \\ \dfrac{1}{24\,024m^8} & \dfrac{1}{1969\,110m^{12}} & \dfrac{1}{109\,745\,064m^{16}} \\ \dfrac{1}{1969\,110m^{12}} & \dfrac{1}{109\,745\,064m^{16}} & \dfrac{1}{4833\,678\,850m^{20}} \end{pmatrix}. \tag{6.3}$$

In agreement with the general prediction the one-loop moment matrix has nonnegative minors.

6.2. Away from $s = 0$: first derivative

We can use (4.16) to derive the bounds by taking the derivative of $B_k^+(s)$ with respect to s before setting $s = 0$. For example let us consider the first derivative with respect to s. In this way we get

$$\partial_s \oint_{t_0} \frac{dt}{2\pi i} \mathcal{M}_{h,g}(s,t) \frac{1}{t} \frac{1}{(t(s+t))^k} \bigg|_{s=0}$$

$$= \int_{m_{\text{gap}}^2}^{\infty} \frac{dt}{\pi} \frac{1}{t^{2k+2}} \left(-(2k+1)\text{Disc}_t \, \mathcal{M}_{h,g}(0,t) + 2t \, \partial_s \text{Disc}_t \, \mathcal{M}_{h,g}(0,t) \right). \tag{6.4}$$

The rhs is not positive semi-definite and therefore we cannot derive the bound on the lhs in a similar fashion. We can, however, consider the following linear combination

$$
\frac{2k+1}{2m_{\text{gap}}^2} \oint_{t_0} \frac{\mathrm{d}t}{2\pi i} \mathcal{M}_{h,g}(s,t) \frac{1}{t} \frac{1}{(t(s+t))^k} + \partial_s \oint_{t_0} \frac{\mathrm{d}t}{2\pi i} \mathcal{M}_{h,g}(s,t) \frac{1}{t} \frac{1}{(t(s+t))^k} \bigg|_{s=0}
$$

$$
= \int_{m_{\text{gap}}^2}^{\infty} \frac{\mathrm{d}t}{\pi} \frac{1}{t^{2k+2}} \left(\left[\frac{t}{m_{\text{gap}}^2} - 1 \right] (2k+1) \text{Disc}_t \, \mathcal{M}_{h,g}(0,t) \right.
$$

$$
\left. + 2t \, \partial_s \text{Disc}_t \, \mathcal{M}_{h,g}(0,t) \right) \succeq 0 \, . \tag{6.5}
$$

We now apply equation (6.5) to $\mathcal{M}_h$. By plugging the low-energy expansion (see equations (3.3) and (4.1)) in the lhs of the formula above and imposing that the eigenvalues of the resulting matrix are nonnegative we can derive various bounds. Let us consider for example $k = 2$. We get the following bound on the inelastic amplitude in terms of the elastic one

$$
\left| h_{5,1} + \frac{5}{2} \frac{h_{4,0}}{m_{\text{gap}}^2} \right| \leqslant 5 \left(f_{1,0} + \frac{f_{0,0}}{m_{\text{gap}}^2} \right) - 2|\beta_{R^3}|^2 \, . \tag{6.6}
$$

To apply the dispersive sum rule above to the bosonic string amplitude we first subtracted the contribution of the tachyon exchange to get $\delta f^{(\text{bs})}(s,u) = f^{(\text{bs})}(s,u) + \left(\frac{\kappa}{2}\right)^2 \frac{1}{1+t}$ and $\delta h^{(\text{bs})}(s,u) = h^{(\text{bs})}(s,u) + \left(\frac{\kappa}{2}\right)^2 \left(\frac{s^4}{1+s} + \frac{t^4}{1+t} + \frac{u^4}{1+u} \right)$. The resulting functions satisfy all the properties needed to apply (6.5) for $k = 2$. We then get, using the formulas from appendix B,

$$
\delta f_{0,0}^{(\text{bs})} = \left(\frac{\kappa}{2}\right)^2 (3 + 2\zeta(3)), \qquad \delta f_{1,0}^{(\text{bs})} = \left(\frac{\kappa}{2}\right)^2 (-3),
$$

$$
\delta h_{4,0}^{(\text{bs})} = \left(\frac{\kappa}{2}\right)^2 2, \qquad \delta h_{5,1}^{(\text{bs})} = \left(\frac{\kappa}{2}\right)^2 (-3), \qquad |\beta_{R^3}^{(\text{bs})}|^2 = \left(\frac{\kappa}{2}\right)^2 \tag{6.7}
$$

which indeed satisfy equation (6.6) where $m_{\text{gap}}^2 = 1$ in the string case.

For the one-loop minimally coupled scalar $|\beta_{R^3}|^2 \sim \kappa^6$ and appears only at two loops. For other coefficients that enter into equation (6.6) we get, using equations (3.21) and (3.26),

$$
f_{0,0}^{S=0} = \left(\frac{\kappa}{2}\right)^4 \frac{1}{(4\pi)^2} \frac{1}{6300m^4}, \qquad f_{1,0}^{S=0} = -\left(\frac{\kappa}{2}\right)^4 \frac{1}{(4\pi)^2} \frac{1}{41\,580m^6},
$$

$$
h_{4,0}^{S=0} = \left(\frac{\kappa}{2}\right)^4 \frac{1}{(4\pi)^2} \frac{1}{3780m^4}, \qquad h_{5,1}^{S=0} = -\left(\frac{\kappa}{2}\right)^4 \frac{1}{(4\pi)^2} \frac{1}{7920m^6}, \tag{6.8}
$$

so that, together with $m_{\text{gap}}^2 = 4m^2$, (6.6) is again satisfied.

More generally, the formula above bounds the correction to the three-point function of the graviton β_{R^3} from above in terms of the EFT data. To make it more manifest we can rewrite the above

$$
2|\beta_{R^3}|^2 \leqslant 5 \left(f_{1,0} + \frac{f_{0,0}}{m_{\text{gap}}^2} \right) - \left| h_{5,1} + \frac{5}{2} \frac{h_{4,0}}{m_{\text{gap}}^2} \right| \leqslant \frac{10 f_{0,0}}{m_{\text{gap}}^2}, \tag{6.9}
$$

where in the last inequality we used the fact that $|f_{1,0}| \leqslant \frac{1}{m_{\text{gap}}^2} f_{0,0}$ which readily follows from equations (5.6) and (5.8). We can restate it more succinctly in terms of the Wilson coefficients

of the gravitational EFT

$$|\beta_{R^3}|^2 \leqslant 5\frac{\beta_{R^4}^+}{m_{\text{gap}}^2},\tag{6.10}$$

where m_{gap}^2 denotes the mass gap at which the massive degrees of freedom that induce the higher-derivative corrections appear. Recall that $f_{0,0} = \beta_{R^4}^+$ was defined in equation (3.19). Bounds similar to equation (6.9) can be derived by considering the superconvergence sum rules for $k > 2$. We do not list them here.

In fact it is not difficult to strengthen the bound (6.10) by considering the following unsubtracted dispersive sum rule (analogous to $k = 0$ in equation (4.24))

$$\oint_\infty \frac{\mathrm{d}t}{2\pi i}\frac{1}{t}f(s,-s-t) = 0.\tag{6.11}$$

The universal tree-level gravitational piece $\left(\frac{\kappa}{2}\right)^2\frac{1}{stu}$, which is singular at $s = 0$, does not contribute to equation (6.11) therefore we can consider

$$\left(\frac{1}{m_{\text{gap}}^2} + \partial_s\right)(6.11)\bigg|_{s=0} : \frac{\beta_{R^4}^+}{m_{\text{gap}}^2} - |\beta_{R^3}|^2 = \int_{m_{\text{gap}}^2}^\infty \frac{\mathrm{d}t}{\pi}\frac{1}{t}\left(\left[\partial_s + \frac{1}{m_{\text{gap}}^2}\right]\text{Disc}_t f\right.$$

$$\left.+\left[\partial_s + \frac{t - m_{\text{gap}}^2}{m_{\text{gap}}^2 t}\right]\text{Disc}_u f\right)\bigg|_{s=0} \geqslant 0,\tag{6.12}$$

where nonnegativity of the rhs can be readily checked for each partial wave separately by plugging the discontinuities of $f(s,u)$, given in equation (5.7), into the formula above.[32] In this way we immediately get a bound

$$|\beta_{R^3}|^2 \leqslant \frac{\beta_{R^4}^+}{m_{\text{gap}}^2}.\tag{6.13}$$

The bound (6.13) is a step towards making the analysis of reference [11] quantitatively precise. At least for the R^3 correction to the graviton three-point coupling which is considered here, it translates the problem to bounding the coefficient $\beta_{R^4}^+$ in terms of G/m_{gap}^2.[33] This problem was beautifully solved recently in reference [7] for $D = 10$ maximal supergravity in a perturbative setting similar to ours, and it was addressed nonperturbatively in reference [18]. The method used in reference [7] is not directly applicable in four dimensions due to the IR divergences, but at least in $D \geqslant 5$ where the $2 \to 2$ amplitude is nonperturbatively well-defined it is natural to expect that one will be able to get a bound on $\beta_{R^4}^+$ in terms of G/m_{gap}^2. It would be very interesting to demonstrate it explicitly.

Indeed, assuming the nonperturbative Regge bound (4.22), we can consider the (-2) subtracted dispersion relations for $f(s,u)$ which expresses $-8\pi G/s$ in terms of the contribution of heavy states, cf equation (3.37) in reference [7]. Existence of such a dispersion relation in the absence of supersymmetry is crucially due to the fact that we consider gravitons as external states.

---

[32] More precisely, it follows from $\partial_s^n\left(\frac{d_{4,4}^J(1+\frac{2s}{m^2})}{(m^2+s)^4}\right)\bigg|_{s=0} > 0$ for $n = 0, 1$ which can be readily checked using formulas from appendix E.

[33] Recall from equation (2.15) that $\left(\kappa/2\right)^2 = 8\pi G$.

6.3. Away from $s = 0$: second derivative

Using exactly the same technique as above one can check that

$$
\frac{2k^2 - 1}{4} \oint_{t_0} \frac{\mathrm{d}t}{2\pi i} \mathcal{M}_{h,g}(s,t) \frac{1}{t} \frac{1}{(t(s+t))^k} + \frac{2k+1}{2t_0} \partial_s
$$
$$
\times \oint_{t_0} \frac{\mathrm{d}t}{2\pi i} \mathcal{M}_{h,g}(s,t) \frac{1}{t} \frac{1}{(t(s+t))^k}
$$
$$
+ \partial_s^2 \oint_{t_0} \frac{\mathrm{d}t}{2\pi i} \mathcal{M}_{h,g}(s,t) \frac{1}{t} \frac{1}{(t(s+t))^k} \bigg|_{s=0} \succeq 0. \tag{6.14}
$$

Consider $k = 2$ and plug the low-energy expansion of $\mathcal{M}_g(s,t)$ in the formula above. We get a matrix whose eigenvalues should be non-negative. In this way we can bound the constant term in the $++-+$ amplitude

$$
|g_{0,0}| \leqslant \frac{7}{4} \frac{f_{0,0}}{(m_{\mathrm{gap}}^2)^2} + \frac{5}{4} \frac{5f_{1,0} - 4|\beta_{R^3}|^2}{m_{\mathrm{gap}}^2} + \frac{3}{2} \left(f_{2,0} + 3f_{2,1} \right). \tag{6.15}
$$

As in the previous section we can check (6.15) in the bosonic string theory. The extra data compared to the previous section takes the form

$$
g_{0,0}^{(\mathrm{bs})} = \left(\frac{\kappa}{2}\right)^2 (2\zeta_3), \qquad \delta f_{2,0}^{(\mathrm{bs})} = \left(\frac{\kappa}{2}\right)^2 (3 + 2\zeta(5)),
$$
$$
\delta f_{2,1}^{(\mathrm{bs})} = \left(\frac{\kappa}{2}\right)^2 (2 + 4\zeta(3) + 2\zeta(5)). \tag{6.16}
$$

Plugging (6.16) into equation (6.15) we see that it is indeed satisfied.

For the one-loop minimally coupled scalar we get correspondingly

$$
g_{0,0}^{S=0} = \left(\frac{\kappa}{2}\right)^4 \frac{1}{(4\pi)^2} \frac{1}{6306\,300m^8}, \qquad f_{2,0}^{S=0} = \left(\frac{\kappa}{2}\right)^4 \frac{1}{(4\pi)^2} \frac{3}{560\,560m^8},
$$
$$
f_{2,1}^{S=0} = \left(\frac{\kappa}{2}\right)^4 \frac{1}{(4\pi)^2} \frac{31}{9081\,072m^8}. \tag{6.17}
$$

Again, we checked that equation (6.15) is satisfied, where we set $m_{\mathrm{gap}}^2 = 4m^2$.

Using the bounds from the previous section and (6.15) we can bound $g_{0,0}$ in terms of $\beta_{R^4}^+$ defined in equation (3.1) as follows

$$
|g_{0,0}| \leqslant \frac{815}{44} \frac{\beta_{R^4}^+}{m_{\mathrm{gap}}^4}. \tag{6.18}
$$

In deriving (6.18) we first expressed $f_{i,j}$, defined in equation (4.1), in terms of $a_{k,j}$, defined in equation (5.5), and then used equations (5.12) and (5.20). Assuming LSD we get a stronger bound

$$
\mathrm{LSD}_{\alpha \to \infty} : |g_{0,0}| \leqslant \frac{25}{4} \frac{\beta_{R^4}^+}{m_{\mathrm{gap}}^4}, \tag{6.19}
$$

where in deriving this we used the stronger LSD bound (5.28).

7. Conclusions

Our paper naturally consisted of two parts with the first part providing theoretical data that we interpreted in the second part in terms of bounds on coefficients of gravitational EFTs. In the first part, using amplitudes methods, we obtained the one-loop four-graviton amplitude with a minimally-coupled massive particle of spin up to $S = 2$ circulating in the loop. By series expanding these amplitude in large mass, we obtain theoretical data for Wilson coefficients which we analyze in the second part. Combining this data with similar theoretical data obtained from tree-level string-theory amplitudes we found that the Wilson coefficients fall on small islands compared to the general bounds coming from consistency of $2 \rightarrow 2$ scattering determined along the lines of references [3–5]. It is quite striking that the EFT coefficients derived from both string-theory and one-loop-massive amplitudes land on the same small islands. Remarkably this can be explained as a consequence of LSD in the partial-wave expansions.

7.1. Obtaining one-loop amplitudes

In order to compute the one-loop amplitudes used to generate EFT data we applied standard amplitudes methods, including spinor helicity [25], generalized unitarity [20, 21], the double copy [22, 23, 38] and integration by parts [33, 51]. Using generalized unitarity we obtained all integral coefficients except for a few whose integrals have no unitarity cuts in any channel. We fixed the coefficients of the latter integrals by using the known UV properties of the amplitude. To fully utilize this information we made use of overcomplete integral basis that contains higher-dimensional integrals, but whose coefficients do not depend on the spacetime dimension or equivalently the dimensional-regularization parameter ϵ. In addition, we also demonstrated that these coefficients do not depend on the mass m of the particles circulating in the loops. The existence of such a basis imposes constraints that makes it reasonably straightforward to determine any pieces not captured by the s, t and u channel unitarity cuts. In this basis we used the fact that the coefficient of the potential $1/\epsilon$ logarithmic UV divergence is zero [48] to completely fix the remaining integral coefficients. A basis without ϵ-dependent coefficients always exists for one-loop problems [79], however the lack of m-dependence appears to be special to our case of a closed massive loop with external massless particles.

It would of course be interesting whether there is some way to generalize our approach to more complicated situations with external legs of differing masses. To generalize our approach to the generic case of a massive one-loop amplitude one could use information from the higher-than-logarithmic divergences which are accessible in dimensional regularization by shifting the dimension downwards as discussed in section 2.5. We showed that knowledge of all these divergences is sufficient to fully constrain the remaining ambiguities, something we expect to be true more generally. We expect constraints from UV divergences and from requiring proper decoupling in the large mass limit to be sufficient to remove any ambiguities in terms that are not fixed by the unitarity cuts, up to the usual ambiguities tied to scheme choices and renormalization. This may provide an alternative method for obtaining complete one-loop amplitudes using on-shell techniques [24].

We exposed a useful supersymmetric decomposition for graviton amplitudes with a massive particle in the loop. A similar decomposition exists in gauge theory [42]. This decomposition expressed the amplitude with a particle of spin S in the loop in terms of amplitudes with lower-spin particles and simpler to calculate pieces. These pieces correspond to amplitudes with a massive BPS multiplet circulating in the loop.

Having constructed the one-loop four-graviton amplitudes with massive particles in the loop, it was then straightforward to expand in large mass, generating amplitudes matching those

from a gravitational EFT. Because the original amplitudes are sensible, satisfying appropriate Regge behavior and unitarity, this EFT is also sensible and can be used to test the constraints derived from dispersion relations.

Besides using these results in our study of bounds on EFT coefficients, we also noted a simple relations between UV divergences in higher dimensions and the coefficients in the $1/m$ expansion. This relation requires a particular analytic continuation of the amplitude to higher dimensions where we keep the number of physical states at their four-dimensional values. It would be interesting to see if similar relations hold more generally, including higher loops, and whether this connection can be exploited in studies of bounds on EFT coefficients.

7.2. EFT bounds

In order to derive bounds on the Wilson coefficients, we first reviewed in section 4 the basic properties of unitarity, crossing, and bounds on the Regge limit in the context of perturbative $2 \to 2$ scattering of gravitons in four dimensions. These allowed us to study dispersion relations and derive bounds on the Wilson coefficients.

In section 5 we focused on the double-minus amplitude and derived bounds on the Wilson coefficients along the lines of references [3–5]. We expressed the low-energy expansion coefficients as dispersive integrals in equation (5.8). We then identified the two-sided bounds on the coefficients following the observation in reference [5] that the boundaries of the allowed region inherit the cyclicity property from the one channel dispersion relation. In this way, by explicitly extracting the mixed constraints from the mixed $s-u$ channel partial waves we identified the new boundaries of the allowed region. We do not rigorously prove that our identification of the bounds is optimal, and we leave filling this gap for future work. We then introduced the idea of LSD which expresses a relationship between moments of the spectral functions of various spins, and we used crossing symmetry or null constraints [4] to derive rigorous bounds in terms of a few low-spin partial waves. We found that all the amplitudes considered in this paper lie on small islands whose location and shape can be determined using the LSD principle.

In section 6 we studied dispersion relations for the graviton amplitudes of various helicities. The key observation is that we can apply dispersion relations to matrices (4.5), composed of the different helicity amplitudes, whose discontinuity is positive semi-definite. Applying dispersive sum rules along the lines of reference [4] to these matrices we derived constraints on the Wilson coefficients which appear in the single-minus and all-plus gravitational scattering amplitudes. Notably we placed a bound on the R^3 coefficient that corrects the graviton three-point amplitude in terms of the R^4 coefficient, making a step towards making the analysis of reference [11] quantitatively precise.

The idea that the set of possibilities to UV complete gravity is much sparser than one naively would have thought lies at the heart of the swampland program [80]. Usually this sparseness is associated with non-perturbative aspects of the UV completion. In the present paper we analyzed the problem in a perturbative setting. By minimally coupling low-spin matter to gravity we generated S-matrices which we expect should satisfy the axioms of causality and unitarity up to an arbitrary order in G and in an arbitrary n-point amplitude. We can ask therefore a weaker version of the swampland question: what is the set of perturbatively consistent weakly-coupled gravitational S-matrices? It is in the framework of this question that the analysis of the present work can be placed. Unexpectedly, we found that in known examples of perturbatively consistent gravitational S-matrices the low-energy couplings lie in regions much smaller than predicted on general grounds of causality and positivity of $2 \to 2$ graviton scattering. Looking at the amplitudes for tree-level string theory and minimally coupled one-loop matter it is not *a priori* obvious at all that they should be 'close' to each other. Dispersive representation of the

low-energy expansion of both makes this similarity apparent. Indeed, as discussed in section 5 both amplitudes satisfy strong versions of LSD that localizes their low-energy data to regions that are much smaller than indicated by the general analysis of unitarity and crossing.

An obvious question is: how general is this universality? We motivated it by pointing out that the examples considered in the present paper define consistent *S*-matrices for any $n \to m$ scattering, whereas the bounds were derived by considering $2 \to 2$ scattering only. To answer this question we would need to better understand the landscape of consistent gravitational amplitudes, including any unitary perturbative or nonperturbative QFT coupled to gravity, as well as amplitudes in theories with extra dimensions. In particular, it would be very interesting to understand if and how consistency of $n \to m$ scattering can be used to rigorously establish stronger bounds on $2 \to 2$ scattering, potentially bringing us closer to the small theory islands observed in the present work. It would also be very interesting to understand the example of large-*N* QCD coupled to gravity [15].

In this paper, we only considered the leading-order effect from integrating out massive degrees of freedom. We did not address the question of bounds that can be applied to IR-safe observables in four dimensions, nor have we included the loops of massless particles which generate more general logarithmic corrections in the low-energy expansion of the amplitude. We leave these important questions to future work. The logarithmic running was discussed in references [5, 17, 77], with the basic idea being that instead of expanding the amplitude around $s, t = 0$ one considers dispersive representations of the EFT couplings defined at some scale. To deal with the IR divergences we may consider dressed states (see e.g. reference [67, 68] for a recent discussion), for which the full implications of unitarity and crossing are still to be fully understood. Another interesting problem is to repeat the analysis of the present paper in the context of $\mathrm{AdS}_4/\mathrm{CFT}_3$ where the problem of IR divergences does not arise, but dispersive techniques discussed in the present paper still hold [81].

In summary, motivated by low-energy theoretical data obtained from one-loop field-theory and tree-level string-theory amplitudes, we put forward the idea that EFTs that describe sensible weak gravitational theories live on small islands that can be understood in terms of partial-wave LSD. It will be important to understand the extent to which this can be extended to constrain gravitational theories.

Acknowledgments

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Appendix A. Minimal coupling

In this appendix we present in more detail our definition of minimal coupling of gravity to a massive spinning particle. For massive spinning particles, the requirement of general-coordinate invariance and mass-dimension-four operators leaves an ambiguity for integer spin-particles in defining the minimal coupling to gravity. In this appendix we discuss this ambiguity and the choice we make in this paper. Our choice corresponds to the one that does not violate the unitarity and causality constraints of references [12, 39, 40, 82, 83]. We start by reviewing the case of electromagnetism (EM) where a similar situation exists, before we move on to consider the gravitational case.

This issue appears only for integer-spin particles. The source of the ambiguity is the existence of a gauge-invariant operator that has the same mass dimension as the kinetic term. For half-integer-spin particles, the corresponding operator is of higher-mass dimension and hence would not count as minimal coupling. Therefore, the cases that require attention are the coupling of a spin-1 or spin-2 particle to gravity. Since the analysis of the coupling of a spin-1 particle to EM is similar, we start by reviewing this case.

Here we fix the ambiguity by requiring that the interactions smoothly match onto the massless limit. To explain our choice, it suffices to evaluate the three-point amplitude between two massive spinning particles and a photon or a graviton.[34] Specifically, we demand that by taking the massless limit of these amplitudes and identifying the spin-1 particles as gauge bosons and the spin-2 particles as gravitons, we recover the corresponding amplitudes in Yang–Mills and pure gravity. The former is indeed realized for the coupling of the *W* boson to the photon [84]. The double-copy construction used in this paper is smooth in the massless limit, and hence selects the above prescription.

For this discussion, we follow the formalism of references [85], which allows us to discuss particles of arbitrary spin.[35] After obtaining the relevant three-point amplitudes, we specialize to the cases of interest. The part of the Lagrangian we give is just the one necessary to obtain the three-point amplitudes relevant for this discussion. We emphasize that these Lagrangians would need to be modified with auxiliary-field terms in order to reproduce our one-loop calculation.

For the case of a massive spinning particle ϕ coupled to EM we have

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + D_{\mu}^{\dagger}\bar{\phi}D^{\mu}\phi - m^2\bar{\phi}\phi + e(g-1)F_{\mu\nu}\bar{\phi}M^{\mu\nu}\phi + \cdots, \qquad (A1)$$

where e is the charge of the particle, g is the gyromagnetic ratio and the ellipsis denote possible auxiliary-field terms. For the Lorentz generator M in the representation of ϕ we may use

$$(M^{\mu\nu})_{c(s)}{}^{d(s)} = 2is\delta_{(c_1}^{[\mu}\eta^{\nu](d_1}\delta_{c_2}^{d_2}\ldots\delta_{c_s)}^{d_s)}, \qquad (A2)$$

where the indices $c(s)$ and $d(s)$ stand for the symmetrized sets of vector indices $\{c_1,\ldots,c_s\}$ and $\{d_1,\ldots,d_s\}$ and symmetrizations include division by the number of terms.

We may now obtain the three-point amplitude between two massive spinning particles and a photon, depicted in figure 22. Neglecting the overall normalization we find

$$\mathcal{A}_3^{\text{EM}-s} \sim \varepsilon_1 \cdot \varepsilon_2\,\varepsilon_3 \cdot p_1 + i(g-1)\varepsilon_1 \cdot M[\varepsilon_3, p_3] \cdot \varepsilon_2, \qquad (A3)$$

---

[34] In calculating this amplitude one implicitly uses complex kinematics.
[35] Descriptions of higher-spin particles date back to Pauli and Fierz [86].
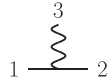
Figure 22. The three-point amplitude necessary to study minimal coupling. The straight lines represent massive spinning particles, while the wiggly line denotes either a photon or a graviton.

where labels 1 and 2 denote the massive spinning particles, while particle 3 is a photon. We use the notation

$$A \cdot M[p,q] \cdot B \equiv p_\mu q_\nu (M^{\mu\nu})_{c(s)}{}^{d(s)} A^{c(s)} B_{d(s)}. \tag{A4}$$

Plugging in the spin-1 representation for M we find

$$\mathcal{A}_3^{\text{EM}-1} \sim \varepsilon_1 \cdot \varepsilon_2 \, \varepsilon_3 \cdot p_1 + (g-1)(\varepsilon_3 \cdot \varepsilon_1 \, \varepsilon_2 \cdot p_3 - \varepsilon_2 \cdot \varepsilon_3 \, \varepsilon_1 \cdot p_3). \tag{A5}$$

Identifying all three particles as gauge bosons in the massless limit, i.e. demanding $\mathcal{A}_3^{\text{EM}-1} = \mathcal{A}_3^{\text{YM}}$, fixes $g = 2$.

The case of a massive spinning particle coupled to gravity is similar. In this case, we have

$$\mathcal{L} = -\frac{1}{16\pi G} R + \frac{1}{2} g^{\mu\nu} \nabla_\mu \phi \nabla_\nu \phi - \frac{1}{2} m^2 \phi\phi + \frac{H}{8} R_{\kappa\lambda\mu\nu} \, \phi M^{\kappa\lambda} M^{\mu\nu} \phi + \cdots, \tag{A6}$$

where H is an arbitrary dimensionless coefficient and we again denote possible auxiliary-field terms by the ellipsis. Plugging in the spin-1 representation we see that the Riemann tensor in the term proportional to H contracts into the Ricci tensor or scalar which under the equations of motion is equivalent to a ϕ^2 term. Therefore, the H term may be replaced by using an appropriate field redefinition by a ϕ^4-type term, which is not relevant for our discussion. Looking at the spin-2 representation and taking the massless limit, we recover the background-field-gauge gravitational Lagrangian of reference [87] for $H = 1$ upon identifying the spin-2 particle with the graviton.

We may alternatively reach the same conclusions by looking at the three-point amplitude between two massive spinning particles and a graviton. This amplitude, shown in figure 22, is given by

$$\mathcal{M}_3^{\text{GR}-s} \sim -\varepsilon_1 \cdot \varepsilon_2 (\varepsilon_3 \cdot p_1)^2 + i(\varepsilon_3 \cdot p_1)\varepsilon_1 \cdot M[\varepsilon_3, p_3] \cdot \varepsilon_2 + \frac{H}{2} \varepsilon_1 \cdot M[\varepsilon_3, p_3] \cdot M[\varepsilon_3, p_3] \cdot \varepsilon_2. \tag{A7}$$

For the spin-1 representation of M we observe that the last term vanishes. For the spin-2 representation we have

$$\mathcal{M}_3^{\text{GR}-2} \sim -\varepsilon_1 \cdot \varepsilon_2 (\varepsilon_3 \cdot p_1)^2 - 2\,\varepsilon_1^{\mu\kappa} \varepsilon_{2,\kappa}^\nu (\varepsilon_{3,\mu} p_{3,\nu} - \varepsilon_{3,\nu} p_{3,\mu})(\varepsilon_3 \cdot p_1)$$
$$- \text{H}\,\varepsilon_1^{\mu\nu} \varepsilon_2^{\alpha\beta} (\varepsilon_{3,\mu} p_{3,\nu} - \varepsilon_{3,\nu} p_{3,\mu})(\varepsilon_{3,\alpha} p_{3,\beta} - \varepsilon_{3,\beta} p_{3,\alpha}). \tag{A8}$$

Identifying the spin-2 particles with gravitons, i.e. setting $\mathcal{M}_3^{\text{GR}-2}$ equal to the three-point amplitude in pure gravity, we find $H = 1$.

With this choice our one-loop amplitudes do not violate unitarity or causality constraints [11, 12, 39, 83]. Hence, they serve as toy models of causal UV completions.

Appendix B. Tree-level string amplitudes

Here we collect the relevant string four-graviton tree-level amplitudes [38]

$$
\mathcal{M}_4^{\text{(ss)}}(1^+, 2^-, 3^-, 4^+) = -\left(\frac{\kappa}{2}\right)^2 (\langle 23\rangle[14])^4 \frac{\Gamma(-s)\Gamma(-t)\Gamma(-u)}{\Gamma(1+s)\Gamma(1+t)\Gamma(1+u)},
$$

$$
\mathcal{M}_4^{\text{(hs)}}(1^+, 2^-, 3^-, 4^+) = -\left(\frac{\kappa}{2}\right)^2 (\langle 23\rangle[14])^4 \frac{\Gamma(-s)\Gamma(-t)\Gamma(-u)}{\Gamma(1+s)\Gamma(1+t)\Gamma(1+u)} \left(1 - \frac{su}{t+1}\right), \quad \text{(B1)}
$$

$$
\mathcal{M}_4^{\text{(bs)}}(1^+, 2^-, 3^-, 4^+) = -\left(\frac{\kappa}{2}\right)^2 (\langle 23\rangle[14])^4 \frac{\Gamma(-s)\Gamma(-t)\Gamma(-u)}{\Gamma(1+s)\Gamma(1+t)\Gamma(1+u)} \left(1 - \frac{su}{t+1}\right)^2,
$$

where we set the closed string tension $\alpha' = 4$ and (ss), (hs), (bs) stand for superstring, heterotic string and bosonic string respectively. In the case of the bosonic string we have additional independent nonvanishing helicity configurations,

$$
\mathcal{M}_4^{\text{(bs)}}(1^+, 2^+, 3^-, 4^+) = -\left(\frac{\kappa}{2}\right)^2 ([12][14]\langle 13\rangle)^4 \frac{\Gamma(-s)\Gamma(-t)\Gamma(-u)}{\Gamma(1+s)\Gamma(1+t)\Gamma(1+u)},
$$

$$
\mathcal{M}_4^{\text{(bs)}}(1^+, 2^+, 3^+, 4^+) = -\left(\frac{\kappa}{2}\right)^2 \left(\frac{[12][34]}{\langle 12\rangle\langle 34\rangle}\right)^2 \frac{4s^2 t^2 u^2 (1 - \frac{stu}{2})^2}{(1+s)^2(1+t)^2(1+u)^2} \quad \text{(B2)}
$$

$$
\times \frac{\Gamma(-s)\Gamma(-t)\Gamma(-u)}{\Gamma(1+s)\Gamma(1+t)\Gamma(1+u)}.
$$

In section 5 we considered the dispersive representation of $f(s,u)$ and used it to derive various bounds on the Wilson coefficients. We focused on the polynomial expansion of the amplitudes at low energies around $s = t = 0$ which are generated by exchanges of massive states above a certain gap m_{gap}^2. To focus on such contributions let us write down explicitly the part of the amplitude due to the exchange by light states (in the case of strings these are tachyon, dilaton, graviton)

$$
f^{\text{(ss)}}(s,u) = -\left(\frac{\kappa}{2}\right)^2 \frac{\Gamma(-s)\Gamma(-t)\Gamma(-u)}{\Gamma(1+s)\Gamma(1+t)\Gamma(1+u)} = \left(\frac{\kappa}{2}\right)^2 \frac{1}{stu} + \delta f^{\text{(ss)}}(s,u),
$$

$$
f^{\text{(hs)}}(s,u) = -\left(\frac{\kappa}{2}\right)^2 \frac{\Gamma(-s)\Gamma(-t)\Gamma(-u)}{\Gamma(1+s)\Gamma(1+t)\Gamma(1+u)} \left(1 - \frac{su}{t+1}\right)
$$

$$
= \left(\frac{\kappa}{2}\right)^2 \left(\frac{1}{stu} - \frac{1}{t}\right) + \delta f^{\text{(hs)}}(s,u), \quad \text{(B3)}
$$

$$
f^{\text{(bs)}}(s,u) = -\left(\frac{\kappa}{2}\right)^2 \frac{\Gamma(-s)\Gamma(-t)\Gamma(-u)}{\Gamma(1+s)\Gamma(1+t)\Gamma(1+u)} \left(1 - \frac{su}{t+1}\right)^2
$$

$$
= \left(\frac{\kappa}{2}\right)^2 \left(\frac{1}{stu} - \frac{1}{t+1} + \frac{su-2}{t}\right) + \delta f^{\text{(bs)}}(s,u),
$$

where Disc $\delta f(s,u)$ is nonzero only for $s, t, u \geqslant m_{\text{gap}}^2 = 1$ and $\delta f(s,u)$ admits the low-energy expansion (4.1) with the dispersive representation of the couplings (5.8).

To write the dispersive representations for string amplitudes let us recall their Regge limit behavior. In the *t*-channel we have

$$\lim_{|t|\to\infty} |f^{(\text{ss,hs,bs})}(s,u)| \leqslant \frac{\text{const}}{|t|^2}, \quad s \leqslant 0. \tag{B4}$$

Therefore, we can write dispersion representation for $f(s,u)$ without subtractions and $\delta f(s,u)$ will take the form (5.7) and the corresponding bounds of section 5.2 apply to $\delta f(s,u)$. The situation is different in the *s*-channel relevant for section 5.3. In this case we have

$$\lim_{|s|\to\infty} |f^{(\text{ss})}(s,u)| \leqslant \frac{\text{const}}{|s|^2}, \quad t \leqslant 0, \tag{B5}$$

$$\lim_{|s|\to\infty} |f^{(\text{hs})}(s,u)| \leqslant \text{const}, \quad t \leqslant 0, \tag{B6}$$

$$\lim_{|s|\to\infty} |f^{(\text{bs})}(s,u)| \leqslant \text{const}|s|^2, \quad t \leqslant 0. \tag{B7}$$

Therefore, for the heterotic and bosonic strings we would need to consider dispersion relations with subtractions. In section 5.3 instead we directly considered the dispersive representation of the relevant couplings (5.52) in the expansion of $\delta f(s,u)$. As can be seen from (B3) the *s*-channel Regge limit of $\delta f(s,u)$ coincides with the ones of the corresponding $f(s,u)$. In particular, the set of couplings that can admit a dispersive representation in the *s*-channel in the three cases correspond to $J_0^{(\text{ss})} = -2$, $J_0^{(\text{hs})} = 0$, $J_0^{(\text{bs})} = 2$, where J_0 is the Regge intercept see the discussion after (5.53).

In section 6 we consider superconvergence relations applied to various helicity amplitudes. The relevant Regge limits for $g^{(\text{bs})}(s,u)$ and $h^{(\text{bs})}(s,u)$ can be easily read off from equation (B2) and take the following form

$$\lim_{|t|\to\infty} |g^{(\text{bs})}(s,u)| \leqslant \frac{\text{const}}{|t|^2}, \quad s \leqslant 0, \tag{B8}$$

$$\lim_{|t|\to\infty} |h^{(\text{bs})}(s,u)| \leqslant \text{const}|t|^2, \quad s \leqslant 0. \tag{B9}$$

All the Regge bounds discussed in this section are in agreement with the general tree-level Regge bound (4.21).

Appendix C. Bounding the coupling space at $k = 6$

In the main body of the paper we did not list the complete set of linear constraints that characterizes the space of admissible couplings in the case of $k = 6$. For completeness we list full set of constraints that we did not present in the main text. To reduce cluttering we avoid writing the subindices that we used in main text and specify which channel should be used in evaluating the determinant explicitly. The constraints are:

$\langle suuuuu \rangle$:

$$\begin{aligned} \{ &\langle 2,4,5,6,7,9 \rangle, \ \langle 4,4,5,7,8,9 \rangle, \ -\langle 6,4,5,6,7,8 \rangle, \\ &- \langle 6,4,5,6,8,9 \rangle, \ -\langle 6,4,6,7,8,9 \rangle, \\ &- \langle 2,5,6,7,k,k+1 \rangle_{k\geqslant 9}, \ -\langle 2,5,j,j+1,k,k+1 \rangle_{k>j\geqslant 9}, \\ &- \langle 4,5,7,8,k,k+1 \rangle_{k\geqslant 9}, \ -\langle 4,5,8,9,k,k+1 \rangle_{k\geqslant 10} \} \geqslant 0. \end{aligned} \tag{C1}$$

$\langle ssuuuu \rangle$:

$$\{-\langle 2,4,5,7,k,k+1 \rangle_{k \geqslant 9}, \; -\langle 2,4,4,5,6,7 \rangle, \; -\langle 2,4,4,5,7,9 \rangle,$$
$$-\langle 2,6,4,5,8,9 \rangle, \; -\langle 4,6,4,5,6,7 \rangle, \; -\langle 4,6,4,5,7,8 \rangle,$$
$$-\langle 4,6,5,6,7,8 \rangle, \; \langle 2,4,4,5,8,9 \rangle, \; \langle 2,6,4,5,6,9 \rangle,$$
$$\langle 2,6,4,6,7,9 \rangle, \; \langle 4,6,4,7,8,9 \rangle, \; \langle 4,8,5,6,7,8 \rangle,$$
$$\langle 2,4,5,9,k,k+1 \rangle_{k \geqslant 10}, \; \langle 2,6,6,7,k,k+1 \rangle_{k \geqslant 9},$$
$$\langle 2,6,5,6,k,k+1 \rangle_{k \geqslant 9}, \; \langle 6,8,5,6,k,k+1 \rangle_{k \geqslant 8},$$
$$\langle 2,4,j,j+1,k,k+1 \rangle_{k>j \geqslant 9},$$
$$\langle 4,6,j,j+1,k,k+1 \rangle_{k>j \geqslant 7}, \; \langle 6,8,j,j+1,k,k+1 \rangle_{k>j \geqslant 6},$$
$$\langle i,i+2,j,j+1,k,k+1 \rangle_{i \geqslant 8;k>j \geqslant 5}\} \geqslant 0\,. \tag{C2}$$

$\langle sssuuu \rangle$:

$$\{-\langle 2,4,6,4,5,6 \rangle, \; -\langle 2,4,6,4,6,7 \rangle, \; -\langle 2,4,6,4,7,9 \rangle,$$
$$\langle 2,4,6,4,5,8 \rangle, \; \langle 2,4,6,4,8,9 \rangle,$$
$$\langle 2,4,8,5,6,7 \rangle, \; \langle 4,6,8,5,6,8 \rangle, \; \langle 4,6,8,6,7,8 \rangle, \; \langle 2,4,6,7,k,k+1 \rangle_{k \geqslant 9},$$
$$\langle 2,i,i+2,5,j,j+1 \rangle_{i \geqslant 8;j \geqslant 6}, \; \langle 4,i,i+2,5,7,8 \rangle_{i \geqslant 8}\} \geqslant 0\,. \tag{C3}$$

$\langle sssssu \rangle$:

$$\{\langle 2,4,6,8,5,6 \rangle, \; -\langle 2,4,j,j+2,5,8 \rangle_{6 \leqslant j \leqslant 14}, \; \langle 2,4,j,j+2,5,7 \rangle_{j \geqslant 8},$$
$$-\langle 2,4,j,j+2,8,9 \rangle_{6 \leqslant j \leqslant 14}, \; -\langle 2,4,j,j+2,6,7 \rangle_{8 \leqslant j \leqslant 24},$$
$$-\langle 2,6,j,j+2,5,6 \rangle_{j \geqslant 8}, \; -\langle 2,4,j,j+2,k,k+1 \rangle_{j \geqslant 6,k \geqslant 9},$$
$$-\langle 4,6,j,j+2,k,k+1 \rangle_{j \geqslant 8,k \geqslant 6},$$
$$-\langle i,i+2,j,j+2,k,k+1 \rangle_{j>i \geqslant 6,k \geqslant 5}\} \geqslant 0\,. \tag{C4}$$

$\langle ssssu \rangle$:

$$\{-\langle 2,4,6,j,j+2,6 \rangle_{j \geqslant 8}, \; -\langle 2,i,i+2,j,j+2,5 \rangle_{j>i \geqslant 6}\} \geqslant 0\,. \tag{C5}$$

In the above formulas, s indicates that in evaluating the determinant one should use i_s vectors in the corresponding position, whereas u means that the corresponding vector should be $j_u^{(6)}$. Note also that i_s take only even integer values. To summarize, the precise meaning of $\langle sssuuu \rangle$: $\langle 2,4,6,7,k,k+1 \rangle_{k \geqslant 9} \geqslant 0$ is

$$\langle \mathbf{a}_6, 2_s, 4_s, 6_s, 7_u^{(6)}, k_u^{(6)}, k_u^{(6)}+1 \rangle_{k \geqslant 9} \geqslant 0\,. \tag{C6}$$

On the crossing-symmetric slice we found that the strongest constraints arise from the $\langle ssuuuu \rangle$ case. To generate figure 11 the maximal spin in the $\langle ssuuuu \rangle$ bounds above is set to $i_{\max}, j_{\max}, k_{\max} = 20$. We used dispersive representations for the couplings truncated to $J_{\max} = 100$ to check that all the determinants are non-negative. We found that the final allowed region for the couplings is not sensitive to the precise value of $k_{\max}$.

Appendix D. Amplitude with an accumulation point in the spectrum

In this appendix we analyze a toy model that violates the LSD property. Nevertheless, we find that the model still winds up on the LSD island. Consider the amplitude function

$$f(t, u) = -\frac{1}{(t - m_1^2)(u - m_1^2)(s - m_2^2)} \tag{D1}$$

$$= \frac{1}{t - m_1^2} \frac{1}{m_1^2 + m_2^2 + t} \left(\frac{1}{s - m_2^2} + \frac{1}{u - m_1^2} \right),$$

where in the second line we rewrote the amplitude in the dispersive representation using partial fractions. For scattering of external scalars, a similar model with $m_1 = m_2$ was considered in [4]. This amplitude saturates the tree-level Regge bound and it has an accumulation point in its spectrum, by which we mean that the residue of the amplitude at either m_1 or m_2 involves infinitely many particles of all spins in the partial-wave expansion. Such models should not be considered physical, but nevertheless it is useful to illustrate features when LSD is violated.

Expanding the s- and u-channel residues in the corresponding partial waves we find that the amplitude is unitary for

$$m_2 \geqslant m_1. \tag{D2}$$

Let us therefore set $m_1 = 1$ and study the model as a function of $m_2 \geqslant 1$.

It is easy to find $\rho_J^{++}(m_2^2)$ explicitly with the following result

$$-\frac{1}{t - 1} \frac{1}{1 + m_2^2 + t} \bigg|_{t = -\frac{m_2^2}{2}(1-x)} = \sum_{J=0}^{\infty} \rho_J^{++}(m_2^2) \, d_{0,0}^J(x), \tag{D3}$$

where

$$\rho_J^{++}(m_2^2) = \frac{4(2J+1)}{m_2^2(2 + m_2^2)} Q_J \left(\frac{2 + m_2^2}{m_2^2} \right), \tag{D4}$$

is determined by projecting the left-hand-side of equation (D3) onto the Legendre polynomials. The function $Q_J(z)$ is the four-dimensional Legendre Q-function[36] that can be found in equation (2.44) of reference [88]. By increasing m_2 we can make the spectral densities of non-minimal spin dominant. One way to understand this is by noting that the lhs of (D3) develops a singularity at $x = \pm 1$ when $m_2 = \infty$ which translates into enhancement of higher-spin contributions. Therefore, (D1) is an explicit example where LSD does not hold.

In the other channel we have

$$-\frac{1}{t - 1} \frac{1}{1 + m_2^2 + t} \bigg|_{t = -\frac{1}{2}(1-x)} = \sum_{J=4}^{\infty} \frac{\rho_J^{+-}(1)}{\left(\frac{1+x}{2} \right)^4} d_{4,4}^J(x). \tag{D5}$$

It is then straightforward to check that $\rho_J^{+-}(1)$ satisfies LSD with $\alpha \simeq 10$ for any value of m_2.

Curiously, if we now consider the values of the couplings $a_{k,j}$ for various values of m_2 they all end up being located at the LSD islands. For m_2 close to 1 it is manifest in the properties of the spectral densities described above. At large values of m_2 when the s-channel spectral

---

[36] Note that the Mathematica LegendreQ function is defined somewhat differently. The precise relation is $Q_J(z) = \text{LegendreQ}[J, z + i0] - i\frac{\pi}{2} \text{LegendreP}[J, z]$ for $z > 1$.

density $\rho_J^{++}(m_2^2)$ violates the LSD property the reason is that its contribution to the low-energy couplings $a_{k,j}$ is suppressed by an extra factor of $\left(\frac{1}{m_2^2}\right)^{k-j}$ coming from the expansion of $\frac{1}{s-m_2^2}$ at small s, compared to the u-channel contribution that does satisfy $\alpha \simeq 10$ LSD. In other words, for $m_2 \gg 1$ we have $\langle \rho_J^{+-} \rangle_k \gg \langle \rho_J^{++} \rangle_k$.

To summarize, while (D1) with an accumulation point in the spectrum does provide an example of an amplitude function which violates the LSD assumption at $\frac{m_2}{m_1} \gg 1$, from the low-energy couplings point of view it still ends up in the LSD island region because the $\rho_J^{++}(m_2^2)$ spectral function that violates the LSD property ends up being irrelevant. While this model should not be considered physical, it does illustrate the idea that potential violations of the LSD principle do not affect the location of the island if there is a separation of scales between the lowest-mass state and higher-mass physics that sources the violation.

Appendix E. Wigner *d*-matrices

Here we list convenient formulas for Wigner *d*-matrices used in bulk of the paper

$$d_{4,4}^J(x) = 2^{-J}(x+1)^J \, {}_2F_1\left(-J-4, 4-J; 1; \frac{x-1}{x+1}\right), \tag{E1}$$

$$d_{4,-4}^J(x) = \frac{2^{-J-7}}{315} \frac{\Gamma(J+5)}{\Gamma(J-3)}(1-x)^4(x+1)^{J-4} \, {}_2F_1\left(4-J, 4-J; 9; \frac{x-1}{x+1}\right), \tag{E2}$$

$$d_{4,0}^J(x) = \frac{2^{-J-3}}{3}\sqrt{\frac{\Gamma(J+5)}{\Gamma(J-3)}}(1-x)^2(x+1)^{J-2} \, {}_2F_1\left(4-J, -J; 5; \frac{x-1}{x+1}\right), \tag{E3}$$

$$d_{0,0}^J(x) = {}_2F_1\left(-J, J+1; 1; \frac{1-x}{2}\right). \tag{E4}$$

By expanding the formulas above around $x = 1$ it is easy to check formulas (4.17)–(4.19) for the first few n's. In Mathematica notation the $d_{\lambda_1,\lambda_2}^J(x)$ functions are given by WignerD[$\{J, \lambda_2, \lambda_1\}$, ArcCos[$x$]]. For a detailed derivation of the partial-wave expansion for spinning external particles see reference [65].

Appendix F. Explicit values of one-loop four-graviton amplitudes

In this appendix we collect the final results for the integrated one-loop amplitudes. For the double-minus configuration, we give the results for the $\mathcal{M}_4^{\{S\}}$ defined in equation (2.32). These results are also collected in a Mathematica ancillary file [89]. The supersymmetric decomposition (2.32) directly gives the amplitude for any massive particle up to spin 2 circulating in the loop. For the single-minus and all-plus helicity configurations, we present the amplitude with a scalar particle circulating in the loop. The amplitudes with a higher-spin particle circulating in the loop are all proportional to this one as shown in equation (2.34). In this appendix we give the results for the amplitudes in terms of scalar integral functions whose values we cite in appendix G.

F.1. The double-minus configuration

We first give our results formally to all orders in the dimensional-regularization parameter ϵ. We write them in terms of an overcomplete basis that contains higher-dimensional integrals as we find this form to be the most concise. This basis is chosen so that the coefficients of the integrals are free of both ϵ and m. As we explained in appendix G.2, using dimension-shifting relations the higher-dimension integrals are directly expressible in terms of standard $(4 - 2\epsilon)$-dimensional ones. We collect the explicit values of the $(4 - 2\epsilon)$-dimensional integrals to leading orders in ϵ [90] in appendix G.1.

We manifest Bose symmetry under a $2 \leftrightarrow 3$ relabeling (which implies $s \leftrightarrow u$) by writing the amplitudes as

$$\mathcal{M}_4^{\{S\}}(1^+, 2^-, 3^-, 4^+) = -\frac{1}{(4\pi)^{2-\epsilon}} \left(\frac{\kappa}{2}\right)^2 \mathcal{M}_4^{\text{tree}} \left(F_1^{\{S\}}(s, u) + F_2^{\{S\}}(s, u) + F_2^{\{S\}}(u, s)\right), \quad \text{(F1)}$$

where $F_1^{\{S\}}(s, u) = F_1^{\{S\}}(u, s)$ with kinematics in the Euclidean region, and the tree-level amplitude $\mathcal{M}_4^{\text{tree}}$ is given in equation (2.55).

For the $\mathcal{M}_4^{\{0\}}$ pieces we have

$$
\begin{aligned}
F_1^{\{0\}}(s, u) = {}& \frac{13s^4 + 52s^3u + 75s^2u^2 + 52su^3 + 13u^4}{96t^4} I_1 \\
& - \frac{8s^4 + 40s^3u + 55s^2u^2 + 40su^3 + 8u^4}{8sut^3} I_1^{6-2\epsilon} \\
& - \frac{(s^2 - su + u^2)^2 t}{64s^2u^2} I_2(t) + \frac{1}{32}\left(16 - \frac{7s}{u} - \frac{7u}{s}\right) I_2^{6-2\epsilon}(t) \\
& - \frac{45}{16t} I_2^{8-2\epsilon}(t) + \frac{s^8 + s^7u + su^7 + u^8}{128s^3u^3} I_3(t) \\
& - \frac{5(s^5 + u^5)}{64s^2u^2} I_3^{6-2\epsilon}(t) + \frac{25(s^2 - su + u^2)}{32su} I_3^{8-2\epsilon}(t) \\
& + \frac{105}{16t} I_3^{10-2\epsilon}(t) - \frac{s^5u^5}{256t^7} I_4(s, u) + \frac{7s^4u^4}{32t^6} I_4^{6-2\epsilon}(s, u) \\
& - \frac{105s^3u^3}{32t^5} I_4^{8-2\epsilon}(s, u) + \frac{105s^2u^2}{8t^4} I_4^{10-2\epsilon}(s, u) - \frac{105su}{16t^3} I_4^{12-2\epsilon}(s, u), \quad \text{(F2)}
\end{aligned}
$$

$$
\begin{aligned}
F_2^{\{0\}}(s, u) = {}& -\frac{s^3(s^2 + 2su + 2u^2)(s^4 + 4s^3u + 5s^2u^2 + 2su^3 + u^4)}{64u^2t^6} I_2(s) \\
& - \frac{s^2(7s^4 + 30s^3u + 50s^2u^2 + 40su^3 - 12u^4)}{32ut^5} I_2^{6-2\epsilon}(s) \\
& - \frac{45s^4 + 118s^3u + 294s^2u^2 + 96su^3 + 16u^4}{16st^4} I_2^{8-2\epsilon}(s) \\
& + \frac{s^6(s^6 + 7s^5u + 21s^4u^2 + 35s^3u^3 + 35s^2u^4 + 21su^5 + 7u^6)}{128u^3t^7} I_3(s) \\
& + \frac{s^4(5s^5 + 25s^4u + 50s^3u^2 + 50s^2u^3 + 25su^4 + 32u^5)}{64u^2t^6} I_3^{6-2\epsilon}(s) \\
& + \frac{5s^3(5s^3 + 15s^2u + 15su^2 - 32u^3)}{32ut^5} I_3^{8-2\epsilon}(s)
\end{aligned}
$$

$$+ \frac{3s^2(35s + 128u)}{16t^4} I_3^{10-2\epsilon}(s) - \frac{s^5 t}{256u^3} I_4(s, t)$$

$$- \frac{s^4}{32u^2} I_4^{6-2\epsilon}(s, t) - \frac{9s^3}{32tu} I_4^{8-2\epsilon}(s, t) - \frac{15s^2}{8t^2} I_4^{10-2\epsilon}(s, t) - \frac{105su}{16t^3} I_4^{12-2\epsilon}(s, t).$$

$$\tag{F3}$$

For the $\mathcal{M}_4^{\{1/2\}}$ pieces we have

$$F_1^{\{1/2\}}(s, u) = \frac{4s^2 + 7su + 4u^2}{8t^2} I_1 + \frac{(s - u)^2 t}{16su} I_2(t) + \frac{3}{4} I_2^{6-2\epsilon}(t)$$

$$+ \frac{(s^5 + u^5)t}{32s^2u^2} I_3(t) + \frac{s^3 + u^3}{4su} I_3^{6-2\epsilon}(t) - \frac{15}{8} I_3^{8-2\epsilon}(t)$$

$$- \frac{s^4 u^4}{64t^5} I_4(s, u) + \frac{15s^3 u^3}{32t^4} I_4^{6-2\epsilon}(s, u)$$

$$- \frac{45s^2 u^2}{16t^3} I_4^{8-2\epsilon}(s, u) + \frac{15su}{8t^2} I_4^{10-2\epsilon}(s, u),$$

$$\tag{F4}$$

$$F_2^{\{1/2\}}(s, u) = \frac{s^3(s + 2u)(s^2 + 2su + 2u^2)}{16ut^4} I_2(s) + \frac{3s^3 + 7s^2u + 12su^2 + 2u^3}{4t^3} I_2^{6-2\epsilon}(s)$$

$$- \frac{s^4(s + 2u)(s^4 + 3s^3u + 4s^2u^2 + 2su^3 + u^4)}{32u^2 t^5} I_3(s)$$

$$- \frac{s^3(2s^3 + 6s^2u + 6su^2 - 5u^3)}{8ut^4} I_3^{6-2\epsilon}(s) - \frac{3s^2(5s + 16u)}{8t^3}$$

$$\times I_3^{8-2\epsilon}(s) + \frac{s^4 t}{64u^2} I_4(s, t) + \frac{3s^3}{32u} I_4^{6-2\epsilon}(s, t)$$

$$+ \frac{9s^2}{16t} I_4^{8-2\epsilon}(s, t) + \frac{15su}{8t^2} I_4^{10-2\epsilon}(s, t).$$

$$\tag{F5}$$

For the $\mathcal{M}_4^{\{1\}}$ pieces we have

$$F_1^{\{1\}}(s, u) = -\frac{t}{4} I_2(t) + \frac{s^4 + s^3u + su^3 + u^4}{8su} I_3(t) + \frac{3t}{4} I_3^{6-2\epsilon}(t)$$

$$- \frac{s^3 u^3}{16t^3} I_4(s, u) + \frac{3s^2 u^2}{4t^2} I_4^{6-2\epsilon}(s, u) - \frac{3su}{4t} I_4^{8-2\epsilon}(s, u),$$

$$\tag{F6}$$

$$F_2^{\{1\}}(s, u) = -\frac{s(s^2 + 2su + 2u^2)}{4t^2} I_2(s) + \frac{s^4(s^2 + 3su + 3u^2)}{8ut^3} I_3(s)$$

$$+ \frac{s^2(3s + 8u)}{4t^2} I_3^{6-2\epsilon}(s) - \frac{s^3 t}{16u} I_4(s, t) - \frac{s^2}{4} I_4^{6-2\epsilon}(s, t) - \frac{3su}{4t} I_4^{8-2\epsilon}(s, t).$$

$$\tag{F7}$$

For the $\mathcal{M}_4^{\{3/2\}}$ pieces we have

$$F_1^{\{3/2\}}(s, u) = -\frac{t^2}{2} I_3(t) - \frac{s^2 u^2}{4t} I_4(s, u) + \frac{su}{2} I_4^{6-2\epsilon}(s, u),$$

$$\tag{F8}$$

$$F_2^{\{3/2\}}(s,u) = -\frac{s^2(s+2u)}{2t} I_3(s) + \frac{s^2 t}{4} I_4(s,t) + \frac{su}{2} I_4^{6-2\epsilon}(s,t) \,. \tag{F9}$$

Finally, for the $\mathcal{M}_4^{\{2\}}$ pieces we have

$$F_1^{\{2\}}(s,u) = -stu \, I_4(s,u) \,, \tag{F10}$$

$$F_2^{\{2\}}(s,u) = -stu \, I_4(s,t). \tag{F11}$$

We note that the progression of the new pieces from more complicated contributions to simpler ones as the spin increases is a direct consequence of the supersymmetric decomposition (2.32). As the spin increases the new pieces have lower and lower power counts corresponding to increasing supersymmetry. The final pieces (F10) and (F11) correspond to $D = (4 - 2\epsilon)$ scalar box integrals with no powers of loop momentum in the numerator. This may be compared to the $\mathcal{M}_4^{\{0\}}$ contribution which has eight powers of loop momentum in the numerator. This high power count results in, for example, the $D = (12 - 2\epsilon)$ box integrals appearing in equations (F2) and (F3).

Next we expand the above results to leading order in the dimensional-regularization parameter ϵ. Using equation (G11) we express the higher-dimensional integrals in terms of $(4 - 2\epsilon)$-dimensional ones whose explicit values through $\mathcal{O}(\epsilon^0)$ are collected in appendix G.1. In the following expressions, the integrals are understood as truncated to this order.

Both $F_1^{\{S\}}$ and $F_2^{\{S\}}$ are UV divergent. However, when put together in equation (F1), the UV divergence cancels as expected [52]. We expose this cancellation by separating the bubble integral in equation (G3) into a divergent part $I_2(0)$ and a finite part defined in equation (G5). After canceling the $1/\epsilon$-pole we write the amplitude as

$$\mathcal{M}_4^{\{S\}}(1^+, 2^-, 3^-, 4^+) = -\frac{1}{(4\pi)^2}\left(\frac{\kappa}{2}\right)^2 \mathcal{M}_4^{\text{tree}}\left(f_1^{\{S\}}(s,u) + f_2^{\{S\}}(s,u) + f_2^{\{S\}}(u,s)\right), \tag{F12}$$

where $f_1^{\{S\}}(s,u) = f_1^{\{S\}}(u,s)$ and $f_{1,2}^{\{S\}}$ are UV finite.

For the $\mathcal{M}_4^{\{0\}}$ pieces corresponding to equations (F2) and (F3) we have

$$\begin{aligned}
f_1^{\{0\}}(s,u) = -&\frac{1}{360t^5}\left(540m^4 sut^2 + su\left(2s^4 + 23s^3 u + 222s^2 u^2\right.\right.\\
&+ 23su^3 + 2u^4\right) - 2m^2 t\left(8s^4 + 5s^3 u - 366s^2 u^2\right.\\
&+ 5su^3 + 8u^4\right)\bigg) - \frac{su}{2t^7}\left(s^4 u^4 + 8m^2 s^3 u^3 t + 20m^4 s^2 t^2 u^2\right.\\
&+ 16m^6 sut^3 + 2m^8 t^4\bigg) I_4(s,u),
\end{aligned} \tag{F13}$$

$$\begin{aligned}
f_2^{\{0\}}(s,u) = \frac{u}{60st^6}&\left(2m^4 t^2(73s^3 - 147s^2 u - 48su^2 - 8u^3)\right.\\
&+ 2m^2 st(9s^4 + 78s^3 u - 105s^2 u^2 - 28su^3 - 4u^4)\\
&+ s^2(s-u)(s^4 + 9s^3 u + 46s^2 u^2 + 9su^3 + u^4)\bigg) I_2^{\text{fin}}(s)\\
&- \frac{s^2 u}{t^7}(su + 2m^2 t)(s^2 u^2 + 4m^2 stu + 2m^4 t^2) I_3(s) - \frac{m^8 su}{t^3} I_4(s,t),
\end{aligned} \tag{F14}$$

where the integrals are defined in equations (G3), (G6) and (G7). For the $\mathcal{M}_4^{\{1/2\}}$ pieces corresponding to equations (F4) and (F5) we have

$$f_1^{\{1/2\}}(s, u) = -\frac{(s^2 + 14su + u^2 + 24m^2t)su}{24t^3}$$
$$-\frac{(su + 2m^2t)(s^2u^2 + 4m^2stu + m^4t^2)su}{2t^5}I_4(s, u),\tag{F15}$$

$$f_2^{\{1/2\}}(s, u) = \frac{\left((s - u)(s^2 + 8su + u^2)s + 4m^2(-4s^3 + 2s^2u + 7su^2 + u^3)\right)u}{12t^4}I_2^{\text{fin}}(s)$$
$$-\frac{(su + 3m^2t)(su + m^2t)s^2u}{t^5}I_3(s) - \frac{m^6su}{t^2}I_4(s, t).\tag{F16}$$

For the $\mathcal{M}_4^{\{1\}}$ pieces corresponding to equations (F6) and (F7) we have

$$f_1^{\{1\}}(s, u) = -\frac{su}{2t} - \frac{(s^2u^2 + 4m^2stu + 2m^4t^2)su}{2t^3}I_4(s, u),\tag{F17}$$

$$f_2^{\{1\}}(s, u) = \frac{(s - u)su}{2t^2}I_2^{\text{fin}}(s) - \frac{(su + 2m^2t)s^2u}{t^3}I_3(s) - \frac{m^4su}{t}I_4(s, t).\tag{F18}$$

For the $\mathcal{M}_4^{\{3/2\}}$ pieces corresponding to equations (F8) and (F9) we have

$$f_1^{\{3/2\}}(s, u) = -\frac{(su + 2m^2t)su}{2t}I_4(s, u),\tag{F19}$$

$$f_2^{\{3/2\}}(s, u) = -\frac{s^2u}{t}I_3(s) - m^2su\,I_4(s, t).\tag{F20}$$

Finally, for the $\mathcal{M}_4^{\{2\}}$ pieces corresponding to equations (F10) and (F11) we have

$$f_1^{\{2\}}(s, u) = -stu\,I_4(s, u),\tag{F21}$$

$$f_2^{\{2\}}(s, u) = -stu\,I_4(s, t).\tag{F22}$$

F.2. The all-plus configuration

As noted in equation (2.34), the result for particles of any spin $0 \leqslant S \leqslant 2$ circulating in the loop is proportional to the $S = 0$ case. The all-orders-in-ϵ form of this amplitude using the higher-dimensional integral basis is

$$\mathcal{M}_4^{S=0}(1^+, 2^+, 3^+, 4^+) = \frac{1}{(4\pi)^{2-\epsilon}}\left(\frac{\kappa}{2}\right)^4\left(\frac{[12][34]}{\langle 12\rangle\langle 34\rangle}\right)^2\frac{1}{2}\left(F_3(s, t, u)\right.$$
$$\left. + F_4(s, t) + F_4(t, u) + F_4(u, s)\right),\tag{F23}$$

where

$$F_3(s, t, u) = \frac{(s^2 + t^2 + u^2)^2}{64stu}I_1 - \frac{15}{4}I_1^{6-2\epsilon},\tag{F24}$$

$$F_4(s, t) = \frac{u^2(s^3 + t^3)^2}{32s^3t^3}I_2(u) + \frac{u^3(7s^2 - 16st + 7t^2)}{16s^2t^2}I_2^{6-2\epsilon}(u)$$

$$+ \frac{45u^2}{8st} I_2^{8-2\epsilon}(u) + \frac{u^4(s^7 + t^7)}{64s^4t^4} I_3(u) + \frac{5u^3(s^5 + t^5)}{32s^3t^3} I_3^{6-2\epsilon}(u)$$

$$+ \frac{25u^2(s^3 + t^3)}{16s^2t^2} I_3^{8-2\epsilon}(u) - \frac{105u^2}{8st} I_3^{10-2\epsilon}(u) + \frac{s^4t^4}{128u^4} I_4(s, t)$$

$$+ \frac{s^3t^3}{16u^3} I_4^{6-2\epsilon}(s, t) + \frac{9s^2t^2}{16u^2} I_4^{8-2\epsilon}(s, t) + \frac{15st}{4u} I_4^{10-2\epsilon}(s, t)$$

$$+ \frac{105}{8} I_4^{12-2\epsilon}(s, t). \tag{F24}$$

There is no corresponding tree-level amplitude for the all-plus helicity. Instead, we choose the above spinor-helicity combination to be completely Bose symmetric. Given this choice, the combination in the parenthesis also has this property. Furthermore, we arrange our functions such that F_3 is completely Bose symmetric, while $F_4(s, t) = F_4(t, s)$.

The expression simplifies significantly if we expand in ϵ and drop the $\mathcal{O}(\epsilon)$ pieces. We have

$$\mathcal{M}_4^{S=0}(1^+, 2^+, 3^+, 4^+) = \frac{1}{(4\pi)^2} \left(\frac{\kappa}{2}\right)^4 \left(\frac{[12][34]}{\langle 12 \rangle \langle 34 \rangle}\right)^2 \frac{1}{2} \left(-\frac{1}{120} \left(120m^4\right.\right.$$

$$+ s^2 + t^2 + u^2\right) + 2m^8 \left(I_4(s, t) + I_4(t, u)\right.$$

$$\left.\left. + I_4(u, s)\right) + \mathcal{O}(\epsilon)\right). \tag{F26}$$

Because the corresponding tree-level amplitude vanishes this amplitude is infrared finite. As for the other helicities it is UV finite because of the lack of a viable counterterm [52]. Another interesting property is that for $m \to 0$ it has no logarithms. The all-minus amplitude follows from parity and is given by swapping angle and square brackets.

F.3. The single-minus configuration

As for the all-plus case, for the single-minus configuration we only need the $S = 0$ case. The single-minus amplitude in a form valid to all orders in ϵ is

$$\mathcal{M}_4^{S=0}(1^+, 2^+, 3^-, 4^+) = \frac{1}{(4\pi)^{2-\epsilon}} \left(\frac{\kappa}{2}\right)^4 \left([12]\langle 13 \rangle [14]\right)^4 \frac{1}{2} \left(F_5(s, t, u)\right.$$

$$\left. + F_6(s, t) + F_6(t, u) + F_6(u, s)\right). \tag{F27}$$

We choose the little group combination to have complete Bose symmetry. The function F_5 is completely Bose symmetric, while $F_6(s, u) = F_6(u, s)$. We find

$$F_5(s, t, u) = \frac{(s^2 + t^2 + u^2)^2}{64(stu)^3} I_1 + \frac{246(stu)^2 - 9(s^2 + t^2 + u^2)^3}{4(stu)^4} I_1^{6-2\epsilon}, \tag{F28}$$

$$F_6(s, u) = \frac{s^6 + u^6}{32s^5u^5} I_2(t) - \frac{s^4 + 2s^3u - 2s^2u^2 + 2su^3 + u^4}{16s^4u^4t} I_2^{6-2\epsilon}(t)$$

$$- \frac{3(9s^4 + 22s^3u + 42s^2u^2 + 22su^3 + 9u^4)}{8s^3u^3t^4} I_2^{8-2\epsilon}(t)$$

$$+ \frac{s^9 + 2s^8u + s^7u^2 + s^2u^7 + 2su^8 + u^9}{64s^6u^6} I_3(t)$$

$$+ \frac{3(s^6 + s^5 u + s u^5 + u^6)}{32 s^5 u^5} I_3^{6-2\epsilon}(t)$$

$$+ \frac{31 s^6 + 93 s^5 u + 93 s^4 u^2 + 126 s^3 u^3 + 93 s^2 u^4 + 93 s u^5 + 31 u^6}{16 s^4 u^4 t^3} I_3^{8-2\epsilon}(t)$$

$$+ \frac{3(29 s^4 + 52 s^3 u - 18 s^2 u^2 + 52 s u^3 + 29 u^4)}{8 s^3 u^3 t^4} I_3^{10-2\epsilon}(t)$$

$$+ \frac{s^2 u^2}{128 t^6} I_4(s, u) - \frac{su}{16 t^5} I_4^{6-2\epsilon}(s, u) - \frac{15}{16 t^4} I_4^{8-2\epsilon}(s, u)$$

$$- \frac{15}{4 s u t^3} I_4^{10-2\epsilon}(s, u) + \frac{105}{8 s^2 t^2 u^2} I_4^{12-2\epsilon}(s, u). \tag{F29}$$

Next, we expand the above amplitude to leading order in ϵ. We write

$$\mathcal{M}_4^{S=0}(1^+, 2^+, 3^-, 4^+) = \frac{1}{(4\pi)^2} \left(\frac{\kappa}{2}\right)^4 \left([12]\langle 13\rangle[14]\right)^4 \frac{1}{2} \left(f_5(s, t, u) + f_6(s, t)\right.$$

$$\left. + f_6(t, u) + f_6(u, s)\right). \tag{F30}$$

As in the double-minus configuration, we extract the UV poles from the bubble integrals in order to manifest the UV-divergence cancellation. In this way we may express the amplitude in terms of the UV-finite functions f_5 and f_6. We choose these functions such that f_5 is completely Bose symmetric and $f_6(s, u) = f_6(u, s)$. We have

$$f_5(s, t, u) = \frac{1}{360(stu)^4} \left((s^2 + t^2 + u^2)(stu)^2 - 15 m^2 (stu)(s^2 + t^2 + u^2)^2 \right.$$

$$\left. + m^4 \left(90(s^2 + t^2 + u^2)^3 - 2520(stu)^2\right)\right), \tag{F31}$$

and

$$f_6(s, u) = \frac{m^4}{s^3 u^3 t^4} \left(2 s^4 + 5 s^3 u + 5 s u^3 + 2 u^4\right) I_2^{\text{fin}}(t)$$

$$+ \frac{2 m^4}{s^4 t^4 u^4} \left(s^7 + 4 s^6 u + 6 s^5 u^2 + 4 s^4 u^3 + 4 s^3 u^4 + 6 s^2 u^5 \right.$$

$$+ 4 s u^6 + u^7 - m^2 (2 s^5 u + 6 s^4 u^2 + 6 s^3 u^3 + 6 s^2 u^4 + 2 s u^5))$$

$$\times I_3(t) + \frac{m^4}{s^2 u^2 t^4} \left(s^2 u^2 + 4 m^2 (stu) + 2 m^4 t^2\right) I_4(s, u). \tag{F32}$$

As for the all-plus amplitude amplitude, because there is no corresponding tree-level amplitude there are no infrared singularities and again in the $m \to 0$ limit the expression is free of logarithms. The single-plus helicity configuration follows from parity.

F.4. Pure gravity

For completeness, we also give the corresponding one-loop amplitude with a massless graviton circulating in the loop. We obtain this result by taking the massless limit of the amplitude with a massive spin-2 particle circulating in the loop, after accounting for the additional states (see equation (2.29)). Our results match the ones previously obtained in reference [19].

Referring to this amplitude as $\mathcal{M}_4^{\mathrm{GR}}$, for the all-plus configuration we have

$$\mathcal{M}_4^{\mathrm{GR}}(1^+, 2^+, 3^+, 4^+) = \frac{-1}{(4\pi)^2}\left(\frac{\kappa}{2}\right)^4 \left(\frac{[12][34]}{\langle 12\rangle\langle 34\rangle}\right)^2 \frac{s^2 + t^2 + u^2}{120}, \qquad (F33)$$

while for the single-minus configuration we find [87]

$$\mathcal{M}_4^{\mathrm{GR}}(1^+, 2^+, 3^-, 4^+) = \frac{1}{(4\pi)^2}\left(\frac{\kappa}{2}\right)^4 ([12]\langle 13\rangle[14])^4 \frac{s^2 + t^2 + u^2}{360(stu)^2}. \qquad (F34)$$

For the double-minus configuration the amplitude takes the form

$$\mathcal{M}_4^{\mathrm{GR}}(1^+, 2^-, 3^-, 4^+)$$

$$= \frac{r_\Gamma}{(4\pi)^{2-\epsilon}}\left(\frac{\kappa}{2}\right)^2 stu\mathcal{M}_4^{\mathrm{tree}}\left[\frac{2}{\epsilon}\frac{1}{stu}\left(s\log\left(\frac{-s}{\mu^2}\right) + t\log\left(\frac{-t}{\mu^2}\right)\right.\right.$$

$$\left. + u\log\left(\frac{-u}{\mu^2}\right)\right) + \frac{2}{stu}\left(u\log\left(\frac{-t}{\mu^2}\right)\log\left(\frac{-s}{\mu^2}\right)\right.$$

$$\left. + t\log\left(\frac{-s}{\mu^2}\right)\log\left(\frac{-u}{\mu^2}\right) + s\log\left(\frac{-t}{\mu^2}\right)\log\left(\frac{-u}{\mu^2}\right)\right)$$

$$+ \frac{4s^6 + 14s^5u + 28s^4u^2 + 35s^3u^3 + 28s^2u^4 + 14su^5 + 4u^6}{t^8}$$

$$\times\left(\log^2\left(\frac{s}{u}\right) + \pi^2\right)$$

$$+ \frac{(s-u)(261s^4 + 809s^3u + 1126s^2u^2 + 809su^3 + 261u^4)}{30t^7}$$

$$\times\log\left(\frac{s}{u}\right) + \frac{1682s^4 + 5303s^3u + 7422s^2u^2 + 5303su^3 + 1682u^4}{180t^6}\right],$$

$$\qquad (F35)$$

where

$$r_\Gamma = \frac{\Gamma(1+\epsilon)\Gamma^2(1-\epsilon)}{\Gamma(1-2\epsilon)}. \qquad (F36)$$

Here μ is an infrared dimensional regularization scale. Since massless gravitons circulate in the loop there is an infrared divergence. On the other hand there is no UV divergence because there is no available counterterm [52]. In this expression we use the FDH scheme [28]. For the massless case the analytic continuation from the Euclidean region to the physical one is simple and accomplished by taking $\log(-s) \to \log(s) - i\pi$. We have explicitly verified that our results for the graviton in the loop match the ones calculated in reference [19], up to the opposite sign for the $\mathcal{M}_4^{\{1/2\}}$ piece already noted in reference [54].

Appendix G. Values of one-loop integrals

In this appendix we give the values of the integrals appearing in the amplitudes collected in appendix F. We first present the $(4 - 2\epsilon)$-dimensional integrals and then discuss the higher-dimensional integrals. Furthermore, we provide an algorithmic procedure for obtaining an expression for the amplitudes with no ϵ or mass dependence in the integral coefficients.

G.1. Explicit values of one-loop integrals

We now collect the values for the integrals appearing in the above amplitudes [90]. We present the integrals in the unphysical Euclidean region where $s, t, u < 0$ and then discuss the analytic continuation to the physical region. We define a generic D-dimensional n-point integral by

$$I_n^D = i(-1)^{n+1}(4\pi)^{D/2} \int \frac{\mathrm{d}^D p}{(2\pi)^D} \frac{1}{(p^2 - m^2)((p-p_1)^2 - m^2) \cdot ((p-p_{n-1})^2 - m^2)}, \quad \text{(G1)}$$

where the p_i's are linear combinations of the external momenta. The integral I_n^D is also labeled by the specific choice of the p_i's. For example, we use $I_2^D(s)$ for a D-dimensional bubble integral that has an invariant mass square of $s = (k_1 + k_2)^2$ flowing through its external legs. Similarly, we use $I_3^D(s)$ and $I_4^D(s, t)$ for a D-dimensional triangle and box respectively, where we use all scales that may appear in the integral as arguments of the corresponding function. When dealing with a $(4 - 2\epsilon)$-dimensional integral we suppress the superscript writing $I_n^{4-2\epsilon} \equiv I_n$.

For the purposes of this paper it is sufficient to present the explicit expressions for the $(4 - 2\epsilon)$-dimensional, one-through four-point integrals up to $\mathcal{O}(\epsilon^0)$. The tadpole (one-point) integral takes the form

$$I_1 = m^{2-2\epsilon} \frac{\Gamma(1+\epsilon)}{\epsilon(\epsilon-1)}. \quad \text{(G2)}$$

The bubble (two-point) integral with a kinematic invariant s is given by

$$I_2(s) = I_2(0) + I_2^{\text{fin}}(s) + \mathcal{O}(\epsilon), \quad \text{(G3)}$$

where

$$I_2(0) = m^{-2\epsilon} \frac{\Gamma(1+\epsilon)}{\epsilon} = \frac{1}{\epsilon} + \mathcal{O}(\epsilon^0), \quad \text{(G4)}$$

and

$$I_2^{\text{fin}}(s) = 2 + x^{(s)} \log\left(\frac{x^{(s)} - 1}{x^{(s)} + 1}\right), \quad \text{(G5)}$$

with $x^{(s)} \equiv \sqrt{1 - 4m^2/s}$. The triangle (three-point) integral is

$$I_3(s) = -\frac{1}{2s} \log^2\left(\frac{x^{(s)} + 1}{x^{(s)} - 1}\right) + \mathcal{O}(\epsilon). \quad \text{(G6)}$$

Finally, the box (four-point) integral is given by

$$
\begin{aligned}
I_4(s,t) = \frac{2}{stx^{(s\,t)}} & \left[2\log^2\left(\frac{x^{(s\,t)} + x^{(s)}}{x^{(s\,t)} + x^{(t)}}\right) + \log\left(\frac{x^{(s\,t)} - x^{(s)}}{x^{(s\,t)} + x^{(s)}}\right) \right. \\
& \times \; \log\left(\frac{x^{(s\,t)} - x^{(t)}}{x^{(s\,t)} + x^{(t)}}\right) - \frac{\pi^2}{2} + \sum_{i=s,\,t}\left(2\mathrm{Li}_2\left(\frac{x^{(i)} - 1}{x^{(s\,t)} + x^{(i)}}\right)\right. \\
& \left. - 2\mathrm{Li}_2\left(-\frac{x^{(s\,t)} - x^{(i)}}{x^{(i)} + 1}\right) - \log^2\left(\frac{x^{(i)} + 1}{x^{(s\,t)} + x^{(i)}}\right)\right)\bigg],
\end{aligned}
\tag{G7}
$$

where $x^{(s\,t)} \equiv \sqrt{1 - 4m^2/s - 4m^2/t}$. To evaluate the expressions in physical regions, e.g. $s > 0$, $t, u < 0$, we need to account for the $i\epsilon$ prescription which for all our integrals is obtained by shifting the mass by $m^2 \to m^2 - i\epsilon$.

In order to match to the low-energy EFT, we expand the above integrals in the large-mass limit. It is straightforward to expand the tadpole, bubble and triangle integrals in this limit. For the box integral we use

$$
\begin{aligned}
I_4(s,t) = \; & \frac{1}{6m^4} + \frac{s+t}{60m^6} + \frac{2s^2 + st + 2t^2}{840m^8} + \frac{(s+t)(3s^2 - 2st + 3t^2)}{7560m^{10}} \\
& + \frac{12s^4 + 3s^3 t + 2s^2 t^2 + 3st^3 + 12t^4}{166\,320m^{12}} \\
& + \frac{(s+t)(10s^4 - 8s^3 t + 9s^2 t^2 - 8st^3 + 10t^4)}{720\,720m^{14}} \\
& + \frac{60s^6 + 10s^5 t + 4s^4 t^2 + 3s^3 t^3 + 4s^2 t^4 + 10st^5 + 60t^6}{21\,621\,600m^{16}} \\
& + \frac{105s^7 + 15s^6 t + 5s^5 t^2 + 3s^4 t^3 + 3s^3 t^4 + 5s^2 t^5 + 15st^6 + 105t^7}{183\,783\,600m^{18}} \\
& + \frac{280s^8 + 35s^7 t + 10s^6 t^2 + 5s^5 t^3 + 4s^4 t^4 + 5s^3 t^5 + 10s^2 t^6 + 35st^7 + 280t^8}{2\,327\,925\,600m^{20}} \\
& + \frac{252s^9 + 28s^8 t + 7s^7 t^2 + 3s^6 t^3 + 2s^5 t^4 + 2s^4 t^5 + 3s^3 t^6 + 7s^2 t^7 + 28st^8 + 252t^9}{9\,777\,287\,520m^{22}} \\
& + \frac{1}{449\,755\,225\,920m^{24}}\left(2520s^{10} + 252s^9 t + 56s^8 t^2 + 21s^7 t^3 + 12s^6 t^4 + 10s^5 t^5\right. \\
& \left. + 12s^4 t^6 + 21s^3 t^7 + 56s^2 t^8 + 252st^9 + 2520t^{10}\right) \\
& + \frac{1}{1\,873\,980\,108\,000m^{26}}\left(2310s^{11} + 210s^{10} t + 42s^9 t^2 + 14s^8 t^3 + 7s^7 t^4 + 5s^6 t^5\right. \\
& \left. + 5s^5 t^6 + 7s^4 t^7 + 14s^3 t^8 + 42s^2 t^9 + 210st^{10} + 2310t^{11}\right) \\
& + \frac{1}{101\,194\,925\,832\,000m^{28}}\left(27\,720s^{12} + 2310s^{11} t + 420s^{10} t^2\right. \\
& + 126s^9 t^3 + 56s^8 t^4 + 35s^7 t^5 + 30s^6 t^6 + 35s^5 t^7 + 56s^4 t^8 \\
& \left. + 126s^3 t^9 + 420s^2 t^{10} + 2310st^{11} + 27\,720t^{12}\right) + \mathcal{O}(m^{-30}).
\end{aligned}
\tag{G8}
$$

This expansion is included in the ancillary files [89].

G.2. Higher-dimension integrals

In constructing the amplitudes we used an overcomplete basis of integrals containing both $(4 - 2\epsilon)$- and higher-dimensional integrals, which as we noted has the advantage of removing ϵ and m dependence from the integral coefficients. We now explain the construction of this form of the amplitudes and how one returns to the usual integral basis containing only the $(4 - 2\epsilon)$-dimensional scalar integrals introduced in equation (2.7). In the $(4 - 2\epsilon)$-dimension form the coefficients of the integrals have explicit ϵ and m dependence. As we discussed in section 2.6, the basis including higher-dimension integrals is useful for exploiting known properties of the amplitude in order to fix the coefficients a_0 and b_0 in equation (2.7), which we cannot obtain from the generalized-unitarity cuts.

Higher-dimension integrals occur naturally in the course of evaluating the loop integrands. Integrals with powers of the higher-dimensional components of loop momentum μ (defined in equation (2.42)) in the numerator may be expressed directly in terms of higher-dimensional integrals. Following reference [20] we have,

$$\int \frac{\mathrm{d}^4\ell}{(2\pi)^4} \frac{\mathrm{d}^{-2\epsilon}\mu}{(2\pi)^{-2\epsilon}} \frac{(\mu^2)^r}{D_{abcd}} = \mathcal{P}(\epsilon, r)(4\pi)^r \int \frac{\mathrm{d}^{4+2r-2\epsilon}L}{(2\pi)^{4+2r-2\epsilon}} \frac{1}{D_{abcd}}, \tag{G9}$$

where the loop momentum on the right-hand side is integrated over a $(4 + 2r - 2\epsilon)$-dimensional space, D_{abcd} is defined in equation (2.23), and

$$\mathcal{P}(\epsilon, 0) = 1, \mathcal{P}(\epsilon, r) = -\epsilon(1 - \epsilon)(2 - \epsilon) \ldots (r - 1 - \epsilon), \quad r > 0. \tag{G10}$$

The resulting higher-dimensional integrals may be expressed in terms the $(4 - 2\epsilon)$-dimensional ones using the dimension-shifting formula [91]:

$$I_n^{D+2} = \frac{1}{(n - D - 1)c_0} \left[2I_n^D - \sum_{i=1}^{n} c_i I_{n-1}^{D(i)} \right], \tag{G11}$$

which holds for any spacetime dimension D and $n \leqslant 5$. I_n^D refers to an n-gon integral in D dimensions, defined in equation (G1). We use $I_{n-1}^{D(i)}$ for the integral obtained by I_n^D by removing the propagator between legs $(i - 1)$ and i. The c_i are combinations of kinematic factors given by

$$c_i = \sum_{j=1}^{n} S_{ij}^{-1}, \qquad c_0 = \sum_{i=1}^{n} c_i, \tag{G12}$$

where the matrix S for the cases of interest to us is given by

$$S_{ij} = m^2 - \frac{1}{2}p_{ij}^2, \quad \text{with} \quad p_{ij} = p_{i-1} - p_{j-1}. \tag{G13}$$

For example, a $(6 - 2\epsilon)$-dimensional box integral is expressed in terms of a $(4 - 2\epsilon)$-dimensional box and four $(4 - 2\epsilon)$-dimensional triangles, illustrated in figure 23.

Using equation (G11), we can reduce an expression that contains higher-dimensional integrals to one that does not. For the reverse process, i.e. in order to eliminate all ϵ and m dependence in a one-loop amplitude $\mathcal{M}_4^{1-\text{loop}}$, we adopt the following strategy: we start by
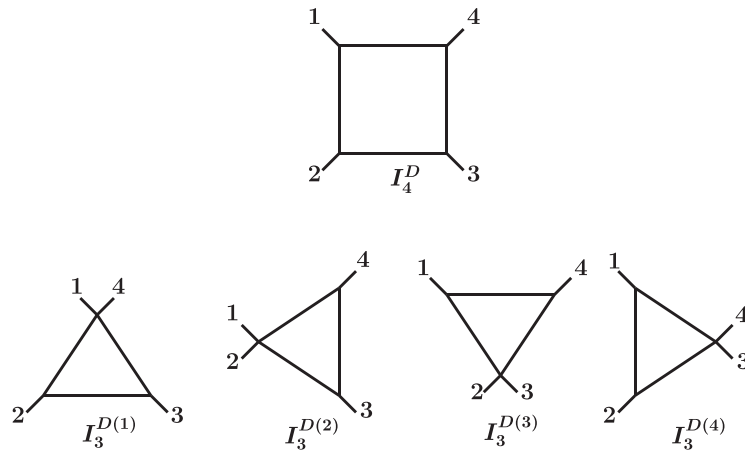
Figure 23. The D-dimensional box integral and its four triangle-integral daughters.

inspecting the coefficients of the boxes, which take the following schematic form

$$\mathcal{M}_4^{1-\text{loop}} = \frac{P(\epsilon)Q(s,t,m)}{(4-(4-2\epsilon)-1)(4-(6-2\epsilon)-1)(4-(8-2\epsilon)-1)(4-(10-2\epsilon)-1)}I_4 + \cdots$$

(G14)

where $P(\epsilon)$ is at most an order-four polynomial in ϵ that does not cancel any of the poles of the expression and $Q(s,t,m)$ is some rational function. The maximum degree in ϵ is directly tied to the maximal power of loop momentum that can appear in the numerator for minimally-coupled gravity. Here we assume the highest possible power of loop momentum, which corresponds to a spin-0 particle circulating in the loop. I_4 in equation (G14) stands for a box integral in our amplitude, whose arguments we do not specify since we are being schematic. This discussion applies to all box integrals in our amplitude. The ellipsis contains other master integrals and their coefficients. Looking at equation (G11), we may identify this term as coming from the $I_4^{12-2\epsilon}$ integral,

$$\mathcal{M}_4^{1-\text{loop}} = P(\epsilon)Q'(s,t,m)I_4^{12-2\epsilon} + \cdots,$$

(G15)

where $Q'(s,t,m)$ is some new rational function. Note that the ellipsis also changes as dictated by equation (G11). Now we may set $D = 4 - 2\epsilon$ in equation (G11) and rewrite it as follows,

$$\epsilon I_n^{6-2\epsilon} = \frac{1}{2c_0}\left[2I_n - \sum_{i=1}^{n} c_i I_{n-1}^{(i)} - c_0(n-5)I_n^{6-2\epsilon}\right],$$

(G16)

where $I_n^{4-2\epsilon} \equiv I_n$. Similarly, setting $D = 6 - 2\epsilon$ in equation (G11) we get

$$\epsilon I_n^{8-2\epsilon} = \frac{1}{2c_0}\left[2I_n^{6-2\epsilon} - \sum_{i=1}^{n} c_i I_{n-1}^{6-2\epsilon(i)} - c_0(n-7)I_n^{8-2\epsilon}\right],$$

(G17)

etc. Combining the two we may trade for example $\epsilon^2 I_n^{8-2\epsilon}$ for an expression containing $I_n^{8-2\epsilon}$, $I_n^{6-2\epsilon}$, I_n and lower-point integrals with ϵ dependence only in the coefficients of the lower-point

integrals. In this fashion we turn equation (G15) into

$$\mathcal{M}_4^{1-\text{loop}} = \sum_{r=0}^{4} Q_r(s,t) I_4^{4+2r-2\epsilon} + \cdots \tag{G18}$$

with some different rational functions $Q_r(s,t)$. In this way we eliminate all explicit ϵ and m dependence in the coefficients of the boxes (we discuss the m dependence momentarily). Only after this step is completed, the poles in ϵ in the coefficients of the triangles have a similar interpretation, i.e. as coming from higher-dimensional triangles. This is because there is a feed down from the coefficients of the boxes to those of the triangles due to equation (G11). The fact that the poles of the triangles align correctly is a nontrivial check of our calculation. We repeat this process sequentially for all lower-point integrals to completely remove the explicit ϵ and m dependence in the coefficients.

This procedure always succeeds. The reason is that there exists an alternative process of reducing the integrals to master integrals that does not introduce any explicit ϵ dependence, but instead introduces these higher-dimensional integrals (for an extensive discussion we refer the reader to appendix I of reference [79]). The existence of such a process guarantees the success of a procedure like the one outlined above. In addition, we may understand why there is no m dependence in the coefficients of the master integrals in this basis as follows: we imagine performing the calculation in a covariant gauge, in which all propagators in the integrals have the canonical form $(p^2 - m^2)$, and there are no other poles in the loop momentum or the mass. We use

$$m^2 = L^2 - (L^2 - m^2), \qquad \mu^2 = -L^2 + \ell^2 \tag{G19}$$

to trade all m and μ dependence in the coefficients for tensor integrals and lower-point integrals. Once all m and μ dependence has been eliminated in this way, we may reduce the tensor integrals using the IBP reduction procedure described in appendix I of reference [79]. The resulting expression contains higher-dimensional integrals without ϵ or m dependence.

Appendix H. High-order expansion of the one-loop four-graviton amplitudes in the large-mass limit

In this appendix, we present the large-mass expansion of our one-loop four-graviton amplitudes through $\mathcal{O}(m^{-20})$. The same results are collected in a Mathematica ancillary file [89]. We give the amplitudes in terms of loop integrals in appendix F. We obtain the results of this appendix by expanding the ones of appendix F in the large-mass limit. We give the values and expansion of the integrals in appendix G.1. The present representation corresponds to low-energy effective description of the gravitational theories under consideration.

The data contained here should be useful for systematic investigations at higher orders in the $1/m$ expansion than carried out in this paper. The examples in sections 5–6 based on low orders in the expansion suggest that Wilson coefficients in physical theories lie on small islands in the allowed parameter space. Rather strikingly, these islands are one dimensional to a good approximation, which is due to the LSD property. Rather remarkably the string-theory data we use also satisfy similar properties, so the Wilson coefficients obtained from string theory also populate these islands. We hope that the data presented here will facilitate further investigations of these features.

Using the amplitudes expanded in the large-mass limit one may obtain the low-energy effective description of the theory, along the lines of section 3. Besides the operators present in

the action of equation (3.1), one should include operators of the schematic form $D^{2k}R^4$. The operators $D^{2k}R^4$ correspond to the terms of $\mathcal{O}(m^{-(4+2k)})$.

We organize the amplitudes in a supersymmetric decomposition (2.32) in terms of the new contributions for a given spin, $\mathcal{M}_4^{\{S\}}$. We may then assemble these pieces into the contributions for a particle of a given spin circulating in the loop using equation (2.32). Regarding the double-minus configuration, starting with the spin-0 contribution and moving to the additional new pieces through spin 2, we have,

$$
\begin{aligned}
f^{\{0\}}(s,u) = \mathcal{K} \Bigg(& \frac{1}{6300m^4} + \frac{t}{41\,580m^6} \\
& + \frac{81(s^2+u^2)+155su}{15\,135\,120m^8} + \frac{t\left(161(s^2+u^2)+324su\right)}{151\,351\,200m^{10}} \\
& + \frac{3556(s^4+u^4)+14\,035(s^3u+su^3)+21\,030s^2u^2}{15\,437\,822\,400m^{12}} \\
& + \frac{t\left(2052(s^4+u^4)+8218(s^3u+su^3)+12\,287s^2u^2\right)}{41\,902\,660\,800m^{14}} \\
& + \frac{4634(s^6+u^6)+27\,650(s^5u+su^5)+69\,026(s^4u^2+s^2u^4)+91\,987s^3u^3}{430\,200\,650\,880m^{16}} \\
& + \frac{t\left(87\,780(s^6+u^6)+526\,770(s^5u+su^5)+1314\,684(s^4u^2+s^2u^4)+1752\,653s^3u^3\right)}{37\,104\,806\,138\,400m^{18}} \\
& + \frac{2551\,824(s^8+u^8)+20\,357\,964(s^7u+su^7)+71\,183\,961(s^6u^2+s^2u^6)}{4823\,624\,797\,992\,000m^{20}} \\
& + \frac{142\,285\,437(s^5u^3+s^3u^5)+177\,823\,240s^4u^4}{4823\,624\,797\,992\,000m^{20}} \Bigg),
\end{aligned}
\tag{H1}
$$

$$
\begin{aligned}
f^{\{1/2\}}(s,u) = \mathcal{K} \Bigg(& \frac{1}{1120m^4} + \frac{t}{8400m^6} + \frac{15(s^2+u^2)+28su}{554\,400m^8} \\
& + \frac{t\left(153(s^2+u^2)+313su\right)}{30\,270\,240m^{10}} \\
& + \frac{665(s^4+u^4)+2596(s^3u+su^3)+3890s^2u^2}{605\,404\,800m^{12}} \\
& + \frac{t\left(581(s^4+u^4)+2345(s^3u+su^3)+3495s^2u^2\right)}{2572\,970\,400m^{14}} \\
& + \frac{29\,106(s^6+u^6)+172\,676(s^5u+su^5)+430\,955(s^4u^2+s^2u^4)+574\,230s^3u^3}{586\,637\,251\,200m^{16}} \\
& + \frac{t\left(34\,440(s^6+u^6)+207\,564(s^5u+su^5)+517\,436(s^4u^2+s^2u^4)+690\,109s^3u^3\right)}{3226\,504\,881\,600m^{18}} \\
& + \frac{39\,270(s^8+u^8)+312\,240(s^7u+su^7)+1091\,604(s^6u^2+s^2u^6)}{16\,491\,024\,950\,400m^{20}} \\
& + \frac{2181\,716(s^5u^3+s^3u^5)+2726\,549s^4u^4}{16\,491\,024\,950\,400m^{20}} \Bigg),
\end{aligned}
\tag{H2}
$$

$$f^{\{1\}}(s,u) = \mathcal{K} \left(\frac{1}{180m^4} + \frac{t}{1680m^6} + \frac{22(s^2 + u^2) + 39su}{151\,200m^8} \right.$$

$$+ \frac{t\left(20(s^2 + u^2) + 43su\right)}{831\,600m^{10}}$$

$$+ \frac{825(s^4 + u^4) + 3125(s^3u + su^3) + 4684s^2u^2}{151\,351\,200m^{12}}$$

$$+ \frac{t\left(315(s^4 + u^4) + 1308(s^3u + su^3) + 1930s^2u^2\right)}{302\,702\,400m^{14}}$$

$$+ \frac{1036(s^6 + u^6) + 6027(s^5u + su^5) + 15\,036(s^4u^2 + s^2u^4) + 20\,030s^3u^3}{4410\,806\,400m^{16}}$$

$$+ \frac{t\left(7056(s^6 + u^6) + 43\,316(s^5u + su^5) + 107\,555(s^4u^2 + s^2u^4) + 143\,715s^3u^3\right)}{146\,659\,312\,800m^{18}}$$

$$+ \frac{11\,760(s^8 + u^8) + 92\,232(s^7u + su^7) + 322\,372(s^6u^2 + s^2u^6)}{1075\,501\,627\,200m^{20}}$$

$$+ \left. \frac{644\,205(s^5u^3 + s^3u^5) + 805\,050s^4u^4}{1075\,501\,627\,200m^{20}} \right), \tag{H3}$$

$$f^{\{3/2\}}(s,u) = \mathcal{K} \left(\frac{1}{24m^4} + \frac{t}{360m^6} + \frac{9(s^2 + u^2) + 14su}{10\,080m^8} \right.$$

$$+ \frac{t\left(8(s^2 + u^2) + 21su\right)}{75\,600m^{10}} + \frac{10(s^4 + u^4) + 34(s^3u + su^3) + 51s^2u^2}{332\,640m^{12}}$$

$$+ \frac{t\left(225(s^4 + u^4) + 1075(s^3u + su^3) + 1518s^2u^2\right)}{50\,450\,400m^{14}}$$

$$+ \frac{105(s^6 + u^6) + 558(s^5u + su^5) + 1391(s^4u^2 + s^2u^4) + 1852s^3u^3}{86\,486\,400m^{16}}$$

$$+ \frac{t\left(224(s^6 + u^6) + 1533(s^5u + su^5) + 3729(s^4u^2 + s^2u^4) + 5035s^3u^3\right)}{1102\,701\,600m^{18}}$$

$$+ \frac{5292(s^8 + u^8) + 38\,416(s^7u + su^7) + 134\,211(s^6u^2 + s^2u^6)}{97\,772\,875\,200m^{20}}$$

$$+ \left. \frac{268\,122(s^5u^3 + s^3u^5) + 335\,050s^4u^4}{97\,772\,875\,200m^{20}} \right), \tag{H4}$$

$$f^{\{2\}}(s,u) = \mathcal{K} \left(\frac{1}{2m^4} + \frac{s^2 + su + u^2}{120m^8} + \frac{stu}{504m^{10}} + \frac{(s^2 + su + u^2)^2}{3780m^{12}} + \frac{(s^2 + su + u^2)stu}{7920m^{14}} \right.$$

$$+ \frac{75(s^6 + u^6) + 225(s^5u + su^5) + 559(s^4u^2 + s^2u^4) + 743s^3u^3}{7207\,200m^{16}}$$

$$+ \frac{3(s^2 + su + u^2)^2 stu}{400\,400m^{18}}$$

$$+ \left. \frac{(s^2 + su + u^2)(56(s^6 + u^6) + 168(s^5u + su^5) + 557(s^4u^2 + s^2u^4) + 834s^3u^3)}{122\,522\,400m^{20}} \right). \tag{H5}$$

We give the relation between $\mathcal{M}_4(1^+, 2^-, 3^-, 4^+)$ and $f(s, u)$ in equation (3.3) and define $\mathcal{K}$ in equation (3.22).

For the all-plus and single-minus configurations it suffices to give the result for the spin-0 contribution since we obtain the remaining amplitudes via equation (2.34). For the all-plus configuration we have

$$
\begin{aligned}
h^{S=0}(s, u) = \mathcal{K} \Bigg(& \frac{stu}{504m^2} + \frac{(s^2 + su + u^2)^2}{3780m^4} + \frac{(s^2 + su + u^2)stu}{7920m^6} \\
& + \frac{75(s^6 + u^6) + 225(s^5u + su^5) + 559(s^4u^2 + s^2u^4) + 743s^3u^3}{7207\,200m^8} \\
& + \frac{3(s^2 + su + u^2)^2 stu}{400\,400m^{10}} \\
& + \frac{(s^2 + su + u^2)\left(56(s^6 + u^6) + 168(s^5u + su^5) + 557(s^4u^2 + s^2u^4) + 834s^3u^3\right)}{122\,522\,400m^{12}} \\
& + \frac{\left(392(s^6 + u^6) + 1176(s^5u + su^5) + 2481(s^4u^2 + s^2u^4) + 3002s^3u^3\right)stu}{888\,844\,320m^{14}} \\
& + \frac{(s^2 + su + u^2)^2\left(105(s^6 + u^6) + 315(s^5u + su^5) + 1412(s^4u^2 + s^2u^4) + 2299s^3u^3\right)}{4888\,643\,760m^{16}} \\
& + \frac{(s^2 + su + u^2)\left(150(s^6 + u^6) + 450(s^5u + su^5) + 1049(s^4u^2 + s^2u^4) + 1348s^3u^3\right)stu}{5766\,092\,640m^{18}} \\
& + \frac{1}{33\,731\,641\,944\,000m^{20}}\left(35\,640(s^{12} + u^{12})\right. \\
& + 213\,840(s^{11}u + su^{11}) + 1174\,365(s^{10}u^2 + s^2u^{10}) + 3911\,625(s^9u^3 + s^3u^9) \\
& \left.+ 8797\,526(s^8u^4 + s^4u^8) + 14\,072\,594(s^7u^5 + s^5u^7) + 16\,416\,696s^6u^6)\right),
\end{aligned}
\tag{H6}
$$

while for the single-minus configuration we find

$$
\begin{aligned}
g^{S=0}(s, u) = \mathcal{K} \Bigg(& \frac{1}{5040m^2 stu} + \frac{1}{6306\,300m^8} + \frac{(s^2 + su + u^2)}{441\,080\,640m^{12}} \\
& + \frac{stu}{2715\,913\,200m^{14}} + \frac{(s^2 + su + u^2)^2}{22\,406\,283\,900m^{16}} + \frac{(s^2 + su + u^2)stu}{64\,250\,746\,560m^{18}} \\
& + \frac{\left(27(s^6 + u^6) + 81(s^5u + su^5) + 197(s^4u^2 + s^2u^4) + 259s^3u^3\right)}{25\,057\,791\,158\,400m^{20}} \Bigg).
\end{aligned}
\tag{H7}
$$

The relation of $\mathcal{M}_4(1^+, 2^+, 3^+, 4^+)$ and $\mathcal{M}_4(1^+, 2^+, 3^-, 4^+)$ to $h(s, u)$ and $g(s, u)$ is found in equation (3.3).

ORCID iDs

Alexander Zhiboedov ⓘ https://orcid.org/0000-0002-7065-1904

References

[1] Aharonov Y, Komar A and Susskind L 1969 *Phys. Rev.* **182** 1400–3

Adams A, Arkani-Hamed N, Dubovsky S, Nicolis A and Rattazzi R 2006 *J. High Energy Phys.* JHEP10(2006)014

Distler J, Grinstein B, Porto R A and Rothstein I Z 2007 *Phys. Rev. Lett.* **98** 041601

Gruzinov A and Kleban M 2007 *Class. Quantum Grav.* **24** 3521–4

Vecchi L 2007 *J. High Energy Phys.* JHEP11(2007)054

Low I, Rattazzi R and Vichi A 2010 *J. High Energy Phys.* JHEP04(2010)126

Nicolis A, Rattazzi R and Trincherini E 2010 *J. High Energy Phys.* JHEP05(2010)095

Nicolis A, Rattazzi R and Trincherini E 2011 *J. High Energy Phys.* JHEP11(2011)128 (erratum)

Cheung C and Remmen G N 2014 *J. High Energy Phys.* JHEP12(2014)087

Baumann D, Green D, Lee H and Porto R A *Phys. Rev.* D **93** 023523

Bellazzini B, Cheung C and Remmen G N 2016 *Phys. Rev.* D **93** 064076

Bellazzini B 2017 *J. High Energy Phys.* JHEP02(2017)034

Cheung C and Remmen G N 2017 *Phys. Rev. Lett.* **118** 051601

de Rham C, Melville S, Tolley A J and Zhou S Y 2017 *Phys. Rev.* D **96** 081702

de Rham C, Melville S, Tolley A J and Zhou S Y 2018 *J. High Energy Phys.* JHEP03(2018)011

de Rham C and Tolley A J 2020 *Phys. Rev.* D **102** 084048

[2] Bai D and Xing Y-H 2018 *Nucl. Phys.* B **932** 15–28

[3] Tolley A J, Wang Z Y and Zhou S Y 2020 arXiv:2011.02400 [hep-th]

[4] Caron-Huot S and Van Duong V 2020 arXiv:2011.02957 [hep-th]

[5] Arkani-Hamed N, Huang T C and Huang Y T 2020 arXiv:2012.15849 [hep-th]

[6] Sinha A and Zahed A 2020 arXiv:2012.04877 [hep-th]

[7] Caron-Huot S, Mazac D, Rastelli L and Simmons-Duffin D 2021 arXiv:2102.08951 [hep-th]

[8] Benincasa P and Cachazo F 2007 arXiv:0705.4305 [hep-th]

[9] Weinberg S 1964 *Phys. Rev.* **135** B1049–56

Weinberg S 1965 *Phys. Rev.* **138** B988–B1002

[10] Huang Y. t., Liu J Y, Rodina L and Wang Y 2020 arXiv:2008.02293 [hep-th]

[11] Camanho X O, Edelstein J D, Maldacena J and Zhiboedov A 2016 *J. High Energy Phys.* JHEP02(2016)020

[12] Meltzer D and Perlmutter E 2018 *J. High Energy Phys.* JHEP07(2018)157

[13] Belin A, Hofman D M and Mathys G 2019 *J. High Energy Phys.* JHEP08(2019)032

[14] Kologlu M, Kravchuk P, Simmons-Duffin D and Zhiboedov A 2019 arXiv:1904.05905 [hep-th]

[15] Kaplan J and Kundu S 2020 arXiv:2009.08460 [hep-th]

[16] Caron-Huot S, Mazac D, Rastelli L and Simmons-Duffin D 2020 arXiv:2008.04931 [hep-th]

[17] Trott T 2020 arXiv:2011.10058 [hep-ph]

[18] Guerrieri A, Penedones J and Vieira P 2021 arXiv:2102.02847 [hep-th]

[19] Dunbar D C and Norridge P S 1995 *Nucl. Phys.* B **433** 181–206

[20] Bern Z and Morgan A G 1996 *Nucl. Phys.* B **467** 479

[21] Bern Z, Dixon L, Dunbar D C and Kosower D A 1994 *Nucl. Phys.* B **425** 217–60

Bern Z, Dixon L, Dunbar D C and Kosower D A 1995 *Nucl. Phys.* B **435** 59–101

[22] Bern Z, Carrasco J J M and Johansson H 2008 *Phys. Rev.* D **78** 085011

Bern Z, Carrasco J J M and Johansson H 2010 *Phys. Rev. Lett.* **105** 061602

[23] Bern Z, Carrasco J J, Chiodaroli M, Johansson H and Roiban R 2019 arXiv:1909.01358 [hep-th]

[24] Britto R and Mirabella E 2012 *J. High Energy Phys.* JHEP01(2012)045

[25] Berends F A, Kleiss R, De Causmaecker P, Gastmans R and Wu T T 1981 *Phys. Lett.* B **103** 124–8

De Causmaecker P, Gastmans R, Troost W and Wu T T 1982 *Nucl. Phys.* B **206** 53–60

Kleiss R and Stirling W J 1985 *Nucl. Phys.* B **262** 235–62

Gunion J F and Kunszt Z 1985 *Phys. Lett.* B **161** 333

Xu Z, Zhang D-H and Chang L 1987 *Nucl. Phys.* B **291** 392–428

[26] Mangano M L and Parke S J 1991 *Phys. Rep.* **200** 301–67

Dixon L J 1996 arXiv:hep-ph/9601359

[27] Collins J C 1984 *Renormalization* (Cambridge: Cambridge University Press)

[28] Bern Z and Kosower D A 1992 *Nucl. Phys.* B **379** 451–561

Bern Z, De Freitas A, Dixon L J and Wong H L 2002 *Phys. Rev.* D **66** 085002

[29] Bern Z, Dixon L and Kosower D A 1998 *Nucl. Phys.* B **513** 3–86

Bern Z, Freitas A D and Dixon L 2002 *J. High Energy Phys.* JHEP03(2002)018

Broggio A, Gnendiger C, Signer A, Stöckinger D and Visconti A 2016 *J. High Energy Phys.* JHEP01(2016)078

[30] Bern Z, Dixon L and Kosower D A 1996 *Annu. Rev. Nucl. Part. Sci.* **46** 109–48

[31] Bern Z and Huang Y-t. 2011 *J. Phys. A: Math. Theor.* **44** 454003
[32] Passarino G and Veltman M 1979 *Nucl. Phys.* B **160** 151–207
     Bern Z, Dixon L and Kosower D A 1994 *Nucl. Phys.* B **412** 751–816
[33] Smirnov A V 2008 *J. High Energy Phys.* JHEP10(2008)107
     Smirnov A V and Chuharev F S 2019 arXiv:1901.07808 [hep-ph]
[34] Eden R J, Landshoff P V, Olive D I and Polkinghorne J C 1966 *The Analytic S-Matrix* (Cambridge: Cambridge University Press)
[35] Bern Z, Dixon L and Kosower D 2004 *J. High Energy Phys.* JHEP08(2004)012
[36] Bern Z, Carrasco J J M, Dixon L J, Johansson H and Roiban R 2010 *Phys. Rev.* D **82** 125040
     Bern Z, Carrasco J J M, Johansson H and Roiban R 2012 *Phys. Rev. Lett.* **109** 241602
     Bern Z, Carrasco J J, Chen W M, Edison A, Johansson H, Parra-Martinez J, Roiban R and Zeng M 2018 *Phys. Rev.* D **98** 086021
[37] Forde D 2007 *Phys. Rev.* D **75** 125019
     Berger C F, Bern Z, Dixon L J, Febres Cordero F, Forde D, Ita H, Kosower D A and Maitre D 2008 *Phys. Rev.* D **78** 036003
[38] Kawai H, Lewellen D C and Tye S-H H 1986 *Nucl. Phys.* B **269** 1–23
[39] Porrati M 1993 *Phys. Lett.* B **304** 77–80
     Cucchieri A, Porrati M and Deser S 1995 *Phys. Rev.* D **51** 4543–9
     Giannakis I, Liu J T and Porrati M 1999 *Phys. Rev.* D **59** 104013
     Deser S and Waldron A 2002 *Nucl. Phys.* B **631** 369–87
     Cortese I, Rahman R and Sivakumar M 2014 *Nucl. Phys.* B **879** 143–61
[40] Bonifacio J, Hinterbichler K, Joyce A and Rosen R A 2018 *J. High Energy Phys.* JHEP06(2018)075
[41] Bern Z, Dixon L and Kosower D A 1993 *Phys. Rev. Lett.* **70** 2677–80
     Bern Z, Dixon L and Kosower D A 1995 *Nucl. Phys.* B **437** 259–304
[42] Bern Z and Morgan A G 1994 *Phys. Rev.* D **49** 6155–63
[43] Fayet P 1979 *Nucl. Phys.* B **149** 137
     Quevedo F, Krippendorf S and Schlotterer O 2010 arXiv:1011.1491 [hep-th]
     Boels R H 2010 *J. High Energy Phys.* JHEP05(2010)046
[44] Henn J M, Naculich S G, Schnitzer H J and Spradlin M 2010 *J. High Energy Phys.* JHEP04(2010)038
     Henn J M, Naculich S G, Schnitzer H J and Spradlin M 2010 *J. High Energy Phys.* JHEP08(2010)002
     Henn J M 2011 *J. Phys. A: Math. Theor.* **44** 454011
     Craig N, Elvang H, Kiermaier M and Slatyer T 2011 *J. High Energy Phys.* JHEP12(2011)097
     Kiermaier M 2011 arXiv:1105.5385 [hep-th]
     Elvang H, Freedman D Z and Kiermaier M 2012 *J. High Energy Phys.* JHEP06(2012)015
     Herderschee A, Koren S and Trott T 2019 *J. High Energy Phys.* JHEP10(2019)092
     Herderschee A, Koren S and Trott T 2019 *J. High Energy Phys.* JHEP08(2019)107
[45] Alday L F, Henn J M, Plefka J and Schuster T 2010 *J. High Energy Phys.* JHEP01(2010)077
[46] Morgan A G 1995 *Phys. Lett.* B **351** 249–56
[47] Zinovev Y M 1987 *Sov. J. Nucl. Phys.* **46** 540
     Berkovits N and Leite M M 1997 *Phys. Lett.* B **415** 144–8
     Berkovits N and Leite M M 1999 *Phys. Lett.* B **454** 38–42
     Gregoire T, Schwartz M D and Shadmi Y 2004 *J. High Energy Phys.* JHEP07(2004)029
[48] 't Hooft G and Veltman M J G 1972 *Nucl. Phys.* B **44** 189–213
[49] Grisaru M T, Pendleton H N and van Nieuwenhuizen P 1977 *Phys. Rev.* D **15** 996
     Grisaru M T and Pendleton H N 1977 *Nucl. Phys.* B **124** 81–92
     Parke S J and Taylor T R 1985 *Phys. Lett.* B **157** 81
     Parke S J and Taylor T R 1986 *Phys. Lett.* B **174** 465 (erratum)
     Bern Z, Dixon L, Dunbar D C, Perelstein M and Rozowsky J S 1998 *Nucl. Phys.* B **530** 401–56
[50] Arkani-Hamed N, Huang T C and Huang Y t 2017 arXiv:1709.04891 [hep-th]
[51] Chetyrkin K G and Tkachov F V 1981 *Nucl. Phys.* B **192** 159
     Laporta S 2000 *Int. J. Mod. Phys.* A **15** 5087–159
[52] 't Hooft G and Veltman M J G 1974 *Ann. Inst. H. Poincare Phys. Theor.* A **20** 69–94
[53] Weinberg S 1965 *Phys. Rev.* **140** B516
     Naculich S G and Schnitzer H J 2011 *J. High Energy Phys.* JHEP05(2011)087
     Akhoury R, Saotome R and Sterman G 2011 *Phys. Rev.* D **84** 104040

[54] Abreu S, Febres Cordero F, Ita H, Jaquier M, Page B, Ruf M S and Sotnikov V 2020 *Phys. Rev. Lett.* **124** 211601
[55] van Nieuwenhuizen P and Wu C C 1977 *J. Math. Phys.* **18** 182
[56] Broedel J and Dixon L J 2012 *J. High Energy Phys.* JHEP10(2012)091
[57] Sennett N, Brito R, Buonanno A, Gorbenko V and Senatore L 2020 *Phys. Rev.* D **102** 044056
[58] Endlich S, Gorbenko V, Huang J and Senatore L 2017 *J. High Energy Phys.* JHEP09(2017)122
[59] Fulling S A, King R C, Wybourne B G and Cummins C J 1992 *Class. Quantum Grav.* **9** 1151–98
[60] Dunbar D C, Godwin J H, Jehu G R and Perkins W B 2018 *Phys. Lett.* B **780** 41–7
[61] Brandhuber A and Travaglini G 2020 *J. High Energy Phys.* JHEP01(2020)010
[62] Accettulli Huber M, Brandhuber A, De Angelis S and Travaglini G 2020 arXiv:2006.02375 [hep-th]
[63] Chowdhury S D, Gadde A, Gopalka T, Halder I, Janagal L and Minwalla S 2020 *J. High Energy Phys.* JHEP02(2020)114
[64] Alberte L, de Rham C, Jaitly S and Tolley A J 2020 arXiv:2012.05798 [hep-th]
[65] Hebbar A, Karateev D and Penedones J 2020 arXiv:2011.11708 [hep-th]
[66] Hebbar A, Häring K, Karateev D, Meineri M and Penedones J work in progress
[67] Strominger A 2017 arXiv:1703.05448 [hep-th]
[68] Arkani-Hamed N, Pate M, Raclariu A M and Strominger A 2020 arXiv:2012.04208 [hep-th]
[69] Martin A 1965 *Nuovo Cimento* A **42** 930–53
[70] Maldacena J, Shenker S H and Stanford D 2016 *J. High Energy Phys.* JHEP08(2016)106
[71] Chandorkar D, Chowdhury S D, Kundu S and Minwalla S 2021 arXiv:2102.03122 [hep-th]
[72] Caron-Huot S, Komargodski Z, Sever A and Zhiboedov A 2017 *J. High Energy Phys.* JHEP10(2017)026
[73] Arkani-Hamed N, Dimopoulos S and Dvali G R 1998 *Phys. Lett.* B **429** 263–72
[74] Randall L and Sundrum R 1999 *Phys. Rev. Lett.* **83** 3370–3
[75] Randall L and Sundrum R 1999 *Phys. Rev. Lett.* **83** 4690–3
[76] Zhang C and Zhou S Y 2020 *Phys. Rev. Lett.* **125** 201601
     Yamashita K, Zhang C and Zhou S Y 2021 *J. High Energy Phys.* JHEP01(2021)095
[77] Bellazzini B, Elias Miró J, Rattazzi R, Riembau M and Riva F 2020 arXiv:2011.00037 [hep-th]
[78] Brower R C, Polchinski J, Strassler M J and Tan C-I 2007 *J. High Energy Phys.* JHEP12(2007)005
[79] Bern Z and Chalmers G 1995 *Nucl. Phys.* B **447** 465
[80] Arkani-Hamed N, Motl L, Nicolis A and Vafa C 2007 *J. High Energy Phys.* JHEP06(2007)060
     Ooguri H and Vafa C 2007 *Nucl. Phys.* B **766** 21–33
     Brennan T D, Carta F and Vafa C 2017 *PoS TASI2017* p 015
     Palti E 2019 *Fortschr. Phys.* **67** 1900037
     van Beest M, Calderón-Infante J, Mirfendereski D and Valenzuela I 2021 arXiv:2102.01111 [hep-th]
[81] Bissi A, Dey P and Hansen T 2020 *J. High Energy Phys.* JHEP04(2020)092
     Carmi D and Caron-Huot S 2020 *J. High Energy Phys.* JHEP09(2020)009
     Mazáč D, Rastelli L and Zhou X 2019 arXiv:1910.12855 [hep-th]
     Penedones J, Silva J A and Zhiboedov A 2020 *J. High Energy Phys.* JHEP08(2020)031
[82] Afkhami-Jeddi N, Kundu S and Tajdini A 2018 *J. High Energy Phys.* JHEP10(2018)156
[83] Afkhami-Jeddi N, Kundu S and Tajdini A 2019 *J. High Energy Phys.* JHEP04(2019)056
[84] Peskin M E and Schroeder D V 1995 *An Introduction to Quantum Field Theory* (Boca Raton, FL: CRC Press)
[85] Singh L P S and Hagen C R 1974 *Phys. Rev.* D **9** 898–909
     Bern Z, Luna A, Roiban R, Shen C H and Zeng M 2020 arXiv:2005.03071 [hep-th]
[86] Pauli W and Fierz M 1939 *Helv. Phys. Acta* **12** 297–300
[87] Bern Z, Dunbar D C and Shimada T 1993 *Phys. Lett.* B **312** 277–84
[88] Correia M, Sever A and Zhiboedov A 2020 arXiv:2006.08221 [hep-th]
[89] See the ancillary files of this manuscript
[90] Brown L M and Feynman R P 1952 *Phys. Rev.* **85** 231–44
     Brown L M 1961 *Nuovo Cimento* **22** 178–82
     Petersson B 1965 *J. Math. Phys.* **6** 1955–9
     Melrose D B 1965 *Nuovo Cimento* A **40** 181–213
     't Hooft G and Veltman M J G 1979 *Nucl. Phys.* B **153** 365–401
     van Neerven W L and Vermaseren J A M 1984 *Phys. Lett.* B **137** 241–4
     Stuart R G 1988 *Comput. Phys. Commun.* **48** 367–89
     Oldenborgh G J and Vermaseren J A M 1990 *Z. Phys.* C **46** 425–37

Denner A, Nierste U and Scharf R 1991 *Nucl. Phys.* B **367** 637–56
Davydychev A I 1993 arXiv:hep-ph/9307323
[91] Bern Z, Dixon L and Kosower D A 1993 *Phys. Lett.* B **302** 299
Bern Z, Dixon L J and Kosower D A 1993 *Phys. Lett.* B **318** 649 (erratum)
Bern Z, Dixon L and Kosower D A 1994 *Nucl. Phys.* B **412** 751–816