

Extremal effective field theories

Simon Caron-Huot and Vincent Van Duong

*Department of Physics, McGill University,
Montréal, QC, Canada*

E-mail: schuot@physics.mcgill.ca, vvd21@cam.ac.uk

ABSTRACT: Effective field theories (EFT) parameterize the long-distance effects of short-distance dynamics whose details may or may not be known. Previous work showed that EFT coefficients must obey certain positivity constraints if causality and unitarity are satisfied at all scales. We explore those constraints from the perspective of $2 \rightarrow 2$ scattering amplitudes of a light real scalar field, using semi-definite programming to carve out the space of allowed EFT coefficients for a given mass threshold M . We point out that all EFT parameters are bounded both below and above, effectively showing that dimensional analysis scaling is a consequence of causality. This includes the coefficients of $s^2 + t^2 + u^2$ and stu type interactions. We present simple $2 \rightarrow 2$ extremal amplitudes which realize, or “rule in”, kinks in coefficient space and whose convex hull span a large fraction of the allowed space.

KEYWORDS: Effective Field Theories, Scattering Amplitudes, Nonperturbative Effects

ARXIV EPRINT: [2011.02957](https://arxiv.org/abs/2011.02957)

Contents

1	Introduction	1
2	Preliminaries: scattering amplitudes and dispersion relations	3
2.1	Low energy: effective field theory	3
2.2	High energy: partial wave decomposition	4
2.3	Dispersion relations	5
2.4	The $B_k(t)$ family of sum rules	8
3	Optimization framework	11
3.1	Warm-up problem with three sum rules	11
3.2	Dual problem: general formulation	13
3.3	Example with a SDPB implementation	15
3.4	Generating null constraints	16
3.5	An ad hoc upper bound on $(\partial\phi)^4$	17
4	Numerically ruling-out: the allowed space of scalar EFTs	19
4.1	Bounds on individual coefficients	20
4.2	Two-dimensional allowed region in $(\tilde{g}_3, \tilde{g}_4)$ plane	20
4.3	Three-dimensional allowed volume in $(\tilde{g}_3, \tilde{g}_4, \tilde{g}_5)$	23
5	Analytically ruling-in: the two kinks with $\tilde{g}_4 = \frac{1}{2}$	24
6	Concluding remarks	27
A	A basis of crossing symmetry constraints	28
A.1	Derivation using dispersion relations in three channels	29
B	Bounds on operators up to order s^{10}	31

1 Introduction

The notion that degrees of freedom at different length scales decouple from each other is a cornerstone of modern physics. In this note, we consider situations where details of the short-distance physics are unknown, but one is interested in its long-distance effects as parameterized by effective field theory (EFT) coefficients. In relativistic quantum theories, it is known that “not anything goes”: if the short-distance physics part is compatible with causality and unitarity, the low-energy parameters will obey certain inequalities, discussed notably in [1]. In this paper we explore such inequalities in an effort to carve out the allowed space of local and unitary EFTs.

We will consider asymptotically flat space-time, where the S-matrix encodes long-distance or low-energy observables. We will specifically study a subset of EFT parameters, denoted g_k , captured by $2 \rightarrow 2$ scattering. As will be reviewed below, causality and unitarity imply dispersive sum rules:

$$g_k = \sum_J \int_{M^2}^{\infty} ds (\cdots)_k \rho_J(s) \tag{1.1}$$

where the spectral density $\rho_J(s)$ (proportional to the imaginary part of the amplitude) is related to the probability of a high-energy state with angular momentum J to scatter at energy $\sqrt{s} > M$, and the kernels $(\cdots)_k$ are given explicitly below and depend on the particular EFT coefficient g_k of interest. The mass M separates “light” and “heavy” states and can be interpreted as the EFT cutoff in an appropriate scheme. We will be agnostic about the high-energy sector: our *only* input will be its compatibility with unitarity and crossing symmetry. Unitarity, the statement that probabilities should lie between 0 and 1, will simply mean:

$$0 \leq \rho_J(s) \leq 2. \tag{1.2}$$

Not all spectral densities that satisfy this inequality are reasonable candidates for the imaginary part of a scattering amplitude, however. This is because Kramers-Kronig type dispersion relations can reconstruct amplitudes from ρ_J alone, but there is no guarantee that the outcome satisfies the full crossing symmetry. Crossing-symmetric $\rho_J(s)$ ’s are orthogonal to an infinite set of “null constraints”, which will be a key ingredient of this paper.

For our purposes, classifying causal and unitary EFTs amounts to finding the image, under the map (1.1), of the set of unitary and crossing-symmetric $\rho_J(s)$.

Causality constraints in quantum field theory have been discussed since the inception of the subject. Many studies were motivated by the phenomenology of the strong force [2]. To give just a few examples, dispersion relations and sum rules were used in the analysis of low-energy pion scattering [3–5], and inequalities satisfied by EFT parameters were obtained using properties of forward amplitudes in [6, 7]. This work aims to explore inequalities on EFT parameters systematically.

We focus on the simplest example: a single (non-gravitating) real scalar field. Since we view the EFT cutoff M as much larger than the mass of the light scattered particles, we take the latter to be massless. On grounds of dimensional analysis, one expects the coefficient of a $(k + d)$ -dimensional operator in the low-energy effective Lagrangian to scale like $\sim 1/M^k$, possibly further suppressed by a small coupling, but never larger. This scaling is clearly realized when one integrates out a massive field. The main question to be addressed is: *Can dimensional analysis scaling be justified by rigorous numerical bounds?* Can “accidentally large” EFT coefficients be ruled out?

We will find that the answer is positive, and we present a general framework to numerically obtain the optimal bounds. Furthermore, we will show that much of the shape of the allowed space, including two kinks, can be understood from simple analytic scattering amplitudes.

The relation between dimensional analysis scaling and causality resonates with many previous studies, for example [8–13]. Our new observation will be the seemingly universal existence of *two-sided* bounds.

This paper is organized as follows. In section 2 we review the general principles satisfied by scattering amplitudes, introducing a family of “ B_k ” sum rules expressing EFT coefficients as averages over high-energy probabilities. In section 3, we provide a general numerical optimization strategy to rule-out candidate EFTs by making use of the averaging technology. In section 4, numerical results are presented along with remarks. Section 5 bridges the numerics with the analytic results. We conclude in section 6 with a discussion about the potential use cases of the numerical framework presented and the further implications of the numerical results.

Note added: when this manuscript was being completed, the works [14] and then [15] appeared with partial overlap in the results. The second paper in particular gave a two-sided bound on the stu interaction which agrees with our eq. (3.6). Further comparisons will be interesting.

2 Preliminaries: scattering amplitudes and dispersion relations

2.1 Low energy: effective field theory

We consider $2 \rightarrow 2$ scattering of massless identical real scalars in a Poincaré invariant theory (figure 1). Treating all momenta as incoming, the amplitude is a function of Mandelstam invariants:

$$s = -(p_1 + p_2)^2, \quad t = -(p_2 + p_3)^2, \quad u = -(p_1 + p_3)^2 \quad (2.1)$$

which satisfy $s + t + u = 0$. By crossing symmetry, it is invariant under all permutations (this holds with appropriate $i0$'s in the discontinuity, as further discussed below):

$$\mathcal{M}(s, t) = \mathcal{M}(t, s) = \mathcal{M}(s, u) = \dots \quad (2.2)$$

Our first step is to parameterize the amplitude at low energies in terms of a specific effective field theory. Generally, the form of the amplitude depends on the couplings of the theory. It becomes particularly simple if the theory is weakly coupled and we restrict ourselves to the tree approximation. We thus use the tree approximation here and until subsection 2.4. In this case, the amplitude has no low-energy branch-cuts, so the EFT expansion is simply a series in small s, t, u :

$$\begin{aligned} \mathcal{M}_{\text{low}}(s, t) = & -g^2 \left[\frac{1}{s} + \frac{1}{t} + \frac{1}{u} \right] - \lambda \\ & + g_2(s^2 + t^2 + u^2) + g_3(stu) + g_4(s^2 + t^2 + u^2)^2 + g_5(s^2 + t^2 + u^2)(stu) \\ & + g_6(s^2 + t^2 + u^2)^3 + g'_6(stu)^2 + g_7(s^2 + t^2 + u^2)^2(stu) + \dots \end{aligned} \quad (2.3)$$

The first line accounts for ϕ^3 and ϕ^4 relevant interactions, while the remaining terms simply list the most general symmetric polynomials in s, t, u , to account for higher-dimension operators in the EFT. The subscript denotes the degree in Mandelstam invariants. Symmetric polynomials are easy to enumerate since their ring is freely generated by two elements: $s^2 + t^2 + u^2$ and stu (given that $s + t + u = 0$).

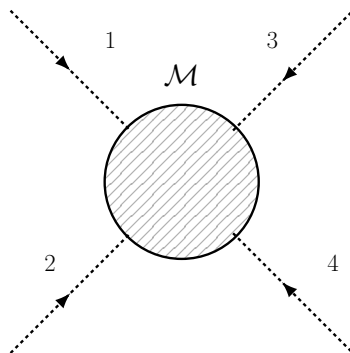


Figure 1. The $2 \rightarrow 2$ scattering process studied in this paper. For different choices of four-momenta, time can flow either horizontally or vertically (or diagonally).

A short exercise shows that the preceding amplitude is obtained from the following effective Lagrangian in the tree approximation:

$$\begin{aligned} \mathcal{L}_{\text{low}} = & -\frac{1}{2}(\partial_\mu\phi)^2 - \frac{g}{3!}\phi^3 - \frac{\lambda}{4!}\phi^4 \\ & + \frac{g_2}{2}[(\partial_\mu\phi)^2]^2 + \frac{g_3}{3}(\partial_\mu\partial_\nu\phi)^2(\partial_\sigma\phi)^2 + 4g_4[(\partial_\mu\partial_\nu\phi)^2]^2 + \dots \end{aligned} \tag{2.4}$$

As is well known, Lagrangian densities are not unique: they are defined modulo integration-by-parts and field redefinitions. One can cast any effective Lagrangian for a real scalar field into the form (2.4) by using field redefinitions to eliminate, order by order in the derivative expansion, corrections to the kinetic and cubic terms as well as appearances of $\partial^2\phi$. See for example [16] for a discussion in the Standard Model context. The amplitude (2.3) is a physical observable unaffected by such ambiguities, which is why we choose to parameterize the coefficients in terms of it.

Our goal is to constrain the EFT parameters g_k assuming existence of an high-energy completion which is causal and unitary, but not necessarily weakly coupled. Low-energy interactions involving five or more powers of ϕ will not be constrained by our methods, since they are not detected by (tree-level) $2 \rightarrow 2$ scattering. When low-energy loop corrections are included, the detailed form of eq. (2.3) will be modified, but we do not expect the number of independent EFT parameters that we can constrain to increase. A precise definition of the g_k 's that remains valid in the presence of low-energy loop corrections is given in eq. (2.18) below.

2.2 High energy: partial wave decomposition

At high energies, we will be agnostic about the form of the amplitude except for the assumption that it is causal and unitary. We follow the general framework of S-matrix theory, as reviewed for example in [2]. Let us begin with unitarity of the S-matrix, which is formally that: $S^\dagger S = 1$, where $S = 1 + i\mathcal{M}$. More precisely, one picks a physical region, say where $s > 0$ is interpreted as center-of-mass energy squared, and $-s < t < 0$ gives the momentum transfer (squared); the scattering angle is

$$\cos\theta = 1 + \frac{2t}{s}. \tag{2.5}$$

The scattering operator is a convolution with respect to angles, which is diagonalized by going to a basis of angular momentum partial waves. The unitarity condition is thus simplest to state in this basis (our conventions follow [17]):

$$\mathcal{M}(s, t) = \sum_{J \text{ even}} n_J^{(d)} f_J(s) \mathcal{P}_J \left(1 + \frac{2t}{s} \right), \quad n_J^{(d)} = \frac{(4\pi)^{\frac{d}{2}} (d + 2J - 3) \Gamma(d + J - 3)}{\pi \Gamma\left(\frac{d-2}{2}\right) \Gamma(J + 1)}, \quad (2.6)$$

where d is the space-time dimensions and $\mathcal{P}_J(x)$ are the d -dimensional version of Legendre polynomials (which appear in the $d = 4$ case):

$$\mathcal{P}_J(x) \equiv {}_2F_1 \left(-J, J + d - 3, \frac{d-2}{2}, \frac{1-x}{2} \right). \quad (2.7)$$

Details of the specific scattering process are encoded in the coefficients $f_J(s)$. In this normalization convention, unitarity of the elastic amplitude of identical real particles is [17]:

$$|S_J(s)| \leq 1, \quad S_J(s) \equiv 1 + i s^{\frac{d-4}{2}} f_J(s). \quad (2.8)$$

The elastic amplitude S_J can have absolute value less than unity due to inelastic processes. We will only need the imaginary part of the high-energy amplitude. Defining the spectral density $\rho_J(s) = s^{\frac{d-4}{2}} \text{Im} f_J(s)$, it can thus be written as

$$\text{Im} \mathcal{M}(s, t) = s^{\frac{4-d}{2}} \sum_{J \text{ even}} n_J^{(d)} \rho_J(s) \mathcal{P}_J \left(1 + \frac{2t}{s} \right) \quad (2.9)$$

where the unitarity constraint is

$$0 \leq \rho_J(s) \leq 2 \quad \forall s > 0, \forall J \text{ even}. \quad (2.10)$$

The normalization is such that $\rho_J = 1$ for complete absorption ($S_J = 0$), and $\rho_J = 2$ for an elastic phase shift π . For the most part (except for subsection 3.5) we will only use the first inequality: $0 \leq \rho_J(s)$.

2.3 Dispersion relations

The other key ingredient from S-matrix theory is the connection between low and high energies, which stems from analyticity. More precisely, we will use the following two properties of the amplitude:

1. For fixed $t < 0$ and $|s|$ sufficiently large, $\mathcal{M}(s, t)$ is analytic in s away from the real axis.
2. For fixed $t < 0$, $\lim_{|s| \rightarrow \infty} \left| \frac{\mathcal{M}(s, t)}{s^2} \right| = 0$ along any line of constant phase.

Physically, these conditions combine causality and unitarity. For an elementary explanation of their respective significance, we refer to the signal propagation model in appendix D of [8], where it is explained that propagation of a signal through a black box is causal if and only if the corresponding transfer function $S(\omega)$ is analytic in the upper-half frequency plane (with sub-exponential growth), and that $|S(\omega)|^2 \leq 1$ *throughout the upper-half-plane* if the box

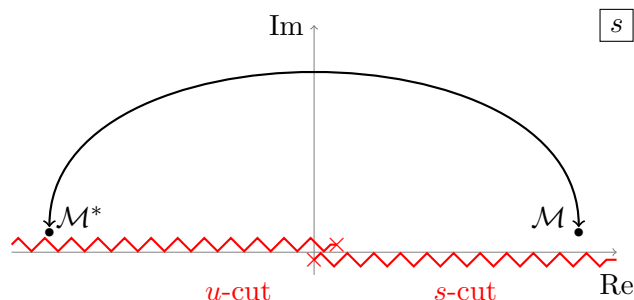


Figure 2. Analyticity in the upper-half-plane relates the s -channel amplitude and the complex conjugate (anti-time-ordered) u -channel amplitude. Note that the s - and u -channel cuts overlap in the physical region where $t < 0$, which is not a problem since the crossing path avoids small $|s|$.

furthermore preserves the squared-norm of signals. These are general facts about Fourier transforms. The original physical derivation of crossing symmetry [18] applies these facts to the expectation value of a retarded commutator in one-particle state. Schematically, one considers

$$\langle p_4 | [\phi(x_3), \phi(x_2)] \theta(x_3^0 - x_2^0) | p_1 \rangle, \quad (2.11)$$

which vanishes outside the forward light-cone rendering its Fourier transform analytic in the upper-half s -plane (intuitively, one uses that s is linear in right-moving light-cone momentum), at least for large enough $|s|$. Its boundary values on the real axis unite the s -channel amplitude and the complex conjugate of the u -channel amplitude, see figure 2.

This is the traditional understanding of crossing symmetry within the axiomatic theory [19]. The boundedness property, in particular including the extra factor of $1/s^2$ compared with the signal model, will be critical for us. We believe it can be justified physically by directly analyzing the transverse Fourier transform [20]. As far as we understand, properties 1-2 are theorems in axiomatic quantum field theory, for example in the context of pion scattering [21, 22]; we take them as axioms embodying causality and unitarity.

The two properties assumed above amount to the existence of twice-subtracted dispersion relations. Let us derive such dispersion relations explicitly. The starting point is that an integral over a large circle vanishes:

$$\oint_{\infty} \frac{ds'}{2\pi i(s' - s)} \frac{\mathcal{M}(s', t)}{(s' - s_1)(s' - s_2)} = 0 \quad (2.12)$$

where s_1 and s_2 are arbitrary subtraction points. For large enough s' , the integrand behaves like $\sim \mathcal{M}(s', t)/s'^3$, and so the integral vanishes thanks to property 2. Typically, one would formally treat all of s , s_1 and s_2 as non-real and deform the contour toward the real axis. Summing the three explicit poles and cuts then relates $\mathcal{M}(s, t)$ to its value at two subtraction points plus an integral over the discontinuity of \mathcal{M} across the real axis [23].

For our purposes it is convenient to instead treat the subtraction poles as part of the real-axis cuts. We choose $s_1 = 0$ and $s_2 = -t$ to maintain the symmetry between the s and u channels without introducing any new energy scale into the problem. The identity (2.12)

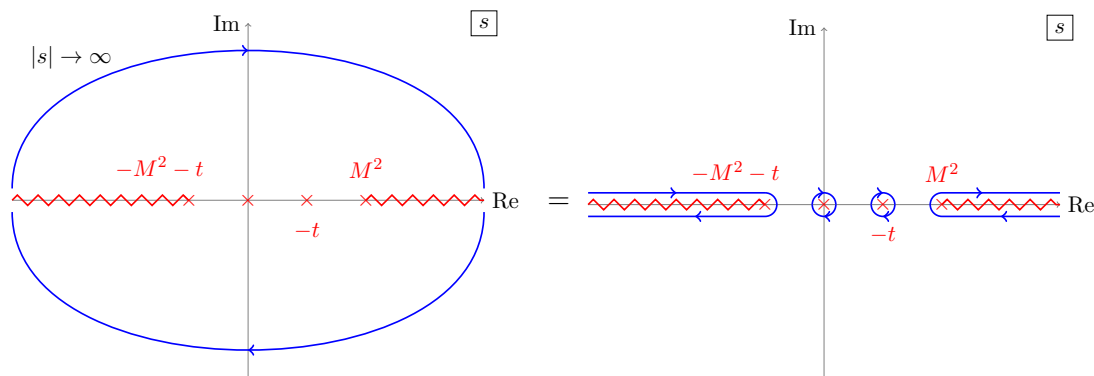


Figure 3. Contour deformation which gives the sum rule (2.14) when low-energy loops are neglected: the integral over arcs at infinity vanishes, thus relating low-energy data and heavy cuts.

then relates the residue at $s' = s$ with a discontinuity:

$$\frac{\mathcal{M}(s, t)}{s(s+t)} = \int_{-\infty}^{\infty} \frac{ds'}{\pi(s' - s)} \text{Im} \left[\frac{\mathcal{M}(s', t)}{s'(s' + t)} \right] \quad (t < 0, s \notin \mathbb{R}), \quad (2.13)$$

where we have written the discontinuity as an imaginary since the amplitude on the “wrong side of the cut” is its complex conjugate; technically $\text{Im} f(s) \equiv \frac{1}{2i} [f(s + i0) - f(s - i0)]$. We call eq. (2.13) a twice-subtracted dispersion relation because of the two powers of s' added to the denominator.

Let us see how this works in the simplest situation considered in eq. (2.3), where low-energy loops are neglected. Then branch cuts can only start at the UV cutoff M^2 . The right-hand-side of eq. (2.13) then contains low-energy poles at $s = 0$ and $s = -t$ (due both to the denominator in eq. (2.13) and poles in the amplitude), and high-energy cuts at $s > M^2$ and $s < -M^2 - t$. Separating low and high energies gives a relation:

$$\begin{aligned} \frac{\mathcal{M}_{\text{low}}(s, t)}{s(s+t)} + \text{Res}_{s'=0, -t} \left[\frac{1}{s' - s} \frac{\mathcal{M}_{\text{low}}(s', t)}{s'(s' + t)} \right] \\ = \int_{M^2}^{\infty} \frac{ds'}{\pi} \left(\frac{1}{s' - s} + \frac{1}{s' + s + t} \right) \text{Im} \left[\frac{\mathcal{M}_{\text{high}}(s', t)}{s'(s' + t)} \right]. \end{aligned} \quad (2.14)$$

We used $s \leftrightarrow u$ symmetry to combine the left and right cuts. This relation is supposed to converge for any s, t with $u < 0$. Interestingly, plugging in the EFT expansion eq. (2.3) for \mathcal{M}_{low} , one finds that both the spin-0 exchange diagram g^2 and spin-0 contact interaction λ cancel out, and what remains is pole-free (this could have been anticipated from the fact that the three residues on the left combine into a single contour over a large circle). On the right-hand-side we insert the partial wave expansion (2.9). It is useful to define heavy averages:

$$\langle F(m^2, J) \rangle \equiv \sum_{J \text{ even}} n_J^{(d)} \int_{M^2}^{\infty} \frac{dm^2}{\pi} \frac{m^{4-d}}{m^2} \rho_J(m^2) [F(m^2, J)]. \quad (2.15)$$

Eq. (2.14) becomes, for $t < 0$:

$$2g_2 - tg_3 + 4(2t^2 + s(s+t))g_4 + \dots = \left\langle \frac{2m^2 + t}{(m^2 - s)(m^2 + s + t)} \frac{\mathcal{P}_J(1 + \frac{2t}{m^2})}{m^2(m^2 + t)} \right\rangle. \quad (2.16)$$

The averaging symbol denotes a (non-normalized) positive sum over heavy states with mass $m > M$. All the results in this paper follow from Taylor-expanding both sides in s and t and using positivity of the measure $\langle \dots \rangle$. It will be useful (though non-essential) to re-organize a bit.

2.4 The $B_k(t)$ family of sum rules

It is easy to see that the Taylor expansion of both sides of eq. (2.16) maintains the symmetry under $s \rightarrow -s - t$, and therefore only even powers of s carry information. More precisely, for each even integer k the coefficient of $[s(s+t)]^{k/2-1}$ gives a one-parameter family of sum rules parameterized by t , which we call $B_k(t)$. It can be computed by taking the $s \rightarrow 0$ limit in eq. (2.12):

$$B_k(t) \equiv \oint_{\infty} \frac{ds}{2\pi i} \frac{1}{s} \frac{\mathcal{M}(s,t)}{[s(s+t)]^{k/2}} = 0 \quad (t < 0, k = 2, 4, \dots). \quad (2.17)$$

This is similar to moment sum rules $\oint_{\infty} \frac{ds}{s^{k+1}} \mathcal{M}(s,t)$ which have been used since times immemorial. Here we have simply re-organized using the $s \leftrightarrow u$ symmetry of our problem to eliminate odd moments.¹ The subscript indicates that B_k enjoys the high-energy convergence of a k -subtracted dispersion relation.

A closely related basis of sum rules was introduced recently for conformal field theory correlators [24] (see also [25]). For holographic theories, the Mellin-space form of the sum rule, called $\widehat{B}_{k,t}$ (see eq. (4.54) and section 4.8 there), precisely reduces in the flat space limit to our current $B_k(t)$. In this context, convergence for $k \geq 2$ is a consequence of the known boundedness of conformal correlators in the Regge limit.

For massless scattering, the low-energy s - and u -channel cuts of $\mathcal{M}(s,t)$ generally overlap as shown in figure 2. It is important that eq. (2.17) can be computed *without* going between the cuts. We simply deform the contour to pick heavy branch cuts at $s > M^2$ and $u > M^2$, and keep the rest as large arcs with $|s| \sim M^2$, the EFT cutoff, see figure 4. This gives a relation between physics at the scale M and that at higher-energies:

$$B_k : \oint_{|s| \approx M^2} \frac{ds}{2\pi i} \frac{1}{s} \frac{\mathcal{M}(s,t)}{[s(s+t)]^{k/2}} = \left\langle \frac{2m^2 + t}{m^2 + t} \frac{\mathcal{P}_J^{(d)} \left(1 + \frac{2t}{m^2}\right)}{[m^2(m^2 + t)]^{k/2}} \right\rangle \quad (t < 0, k = 2, 4, \dots). \quad (2.18)$$

This equation is valid even when EFT loops are included. The idea is to choose the EFT cutoff M such that loop corrections in the low-energy EFT are under control over the arcs with $|s| \approx M^2$. Eq. (2.18) thus equates an EFT-computable left-hand-side, with a high-energy average that enjoys positivity properties.

The specific relation between the left-hand-side and EFT coefficients will depend on EFT interactions. For concreteness let us thus focus again on the case where EFT loops are neglected. The left-hand-side is then just the sum of residues at $s = 0$ and $s = -t$; by

¹The identity: $\oint_{\infty} \frac{ds}{2\pi i} \frac{1}{s} [s(s+t)]^{(k-k')/2} = \delta_{k,k'}$ shows that eq. (2.17) indeed extracts the coefficient of $[s(s+t)]^{k/2}$ in \mathcal{M}_{low} .

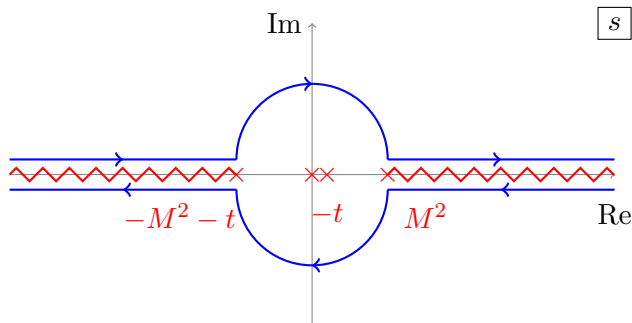


Figure 4. Integration contour to be used when low-energy loops are included; the integral vanishes as it is equivalent to arcs at infinity. This relates high-energy cuts at $s > M^2$ and $u > M^2$ with EFT-computable data near the EFT cutoff $|s| \sim M^2$.

symmetry, we can replace $\frac{1}{s}$ by $\left(\frac{1}{s} - \frac{1}{s+t}\right)$ and include a single pole, and the B_k sum rules becomes:

$$B_k : \text{Res}_{s=0} \left[\frac{2s+t}{s(s+t)} \frac{\mathcal{M}_{\text{low}}(s,t)}{[s(s+t)]^{k/2}} \right] = \left\langle \frac{2m^2+t}{m^2+t} \frac{\mathcal{P}_J^{(d)}\left(1 + \frac{2t}{m^2}\right)}{[m^2(m^2+t)]^{k/2}} \right\rangle \quad (t < 0, k = 2, 4, \dots). \tag{2.19}$$

This simplification of eq. (2.18) is only valid when neglecting EFT loops.

Let us record the first few two instances explicitly:

$$B_2 : 2g_2 - g_3t + 8g_4t^2 + \dots = \left\langle \frac{(2m^2+t) \mathcal{P}_J\left(1 + \frac{2t}{m^2}\right)}{m^2(m^2+t)^2} \right\rangle, \tag{2.20}$$

$$B_4 : 4g_4 + \dots = \left\langle \frac{(2m^2+t) \mathcal{P}_J\left(1 + \frac{2t}{m^2}\right)}{m^4(m^2+t)^3} \right\rangle. \tag{2.21}$$

The left-hand side has a regular series in t , and the right-hand side involves Gegenbauers $\mathcal{P}_J(1 + \frac{2t}{m^2})$, which can be straightforwardly expanded at small $t \ll M^2$ using eq. (2.7). Recall that averages are taken over heavy states with $m \geq M$. Matching both sides order by order in t generates a linear system in g_n 's:

$$g_2 = \left\langle \frac{1}{m^4} \right\rangle, \quad g_3 = \left\langle \frac{3 - \frac{4}{d-2} \mathcal{J}^2}{m^6} \right\rangle, \quad g_4 = \left\langle \frac{1}{2m^8} \right\rangle, \tag{2.22}$$

$$g_4 = \left\langle \frac{1 + \frac{4-5d}{2d(d-2)} \mathcal{J}^2 + \frac{1}{d(d-2)} \mathcal{J}^4}{2m^8} \right\rangle.$$

We introduced the spin Casimir $\mathcal{J}^2 = J(J+d-3)$ for convenience. Note that we truncated \mathcal{M}_{low} to order g_4 , but it is possible to work to higher orders and generate linear relations on couplings such as g_5 and so on.

The averaging notation immediately shows that $g_2, g_4 > 0$ since they are high-energy averages of positive quantities $\frac{1}{m^4}$ and $\frac{1}{2m^8}$, respectively. Furthermore, the inequalities $g_3 \leq \frac{3g_2}{M^2}$ and $g_4 \leq \frac{g_2}{2M^4}$ also follow readily since $m \geq M$ inside the average. In contrast, the sign of g_3 is not immediate due to the presence of spinning particles – the magnitude

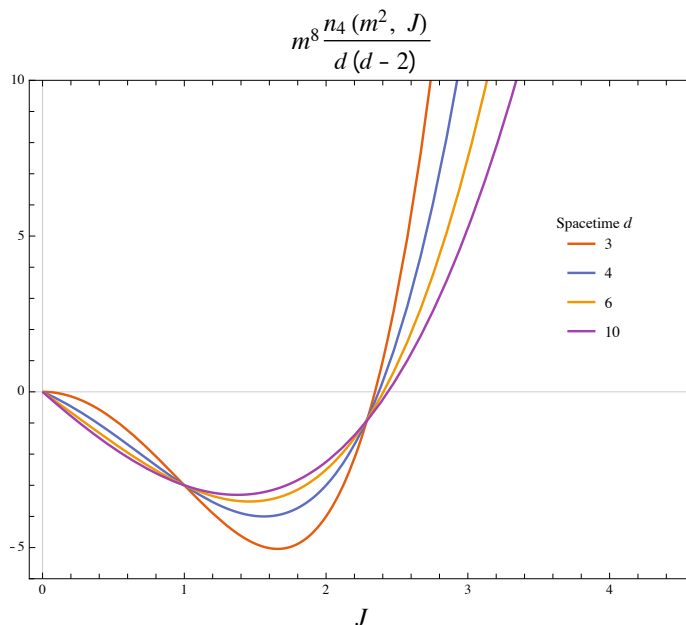


Figure 5. The null constraint $m^8 n_4(m^2, J)$, which is a function of only J . It vanishes at $J = 0$, is negative at $J = 2$, but positive at $J = 4, 6, \dots$, and thus balances spin-two against higher-spin states. The sign change in various space-time dimensions (at $J_{\text{critical}} = \frac{1}{2}(3 - d + \sqrt{d(d+4) + 1})$) is always situated between $2 < J < 3$.

of \mathcal{J}^2 requires a deeper investigation. This difficulty was noted in attempted proofs of the six-dimensional a -theorem [26].

The key to calculating a lower bound for g_3 will be the existence of two distinct averages that output g_4 . Equating them yields the first example of what turns out to be an infinite set of *null constraints*:

$$0 = \langle n_4(m^2, J) \rangle, \quad n_4(m^2, J) \equiv \frac{\mathcal{J}^2(2\mathcal{J}^2 - (5d - 4))}{m^8}. \tag{2.23}$$

This is a constraint on the probabilities $\rho_J(s)$ which define the average $\langle \cdot \rangle$. The subscript indicates the degree in $1/m^2$. Physically this stems from crossing symmetry – since there is a unique symmetric polynomial at degree 4, the coefficients of $s^2 t^2$ and s^4 must be related. There are no lower-degree examples of this phenomenon: monomials with fewer than two powers of s are killed by any double-subtracted sum rule, and odd powers of s are information-free since fixed- t dispersion relations preserve the $s \leftrightarrow -s - t$ symmetry of our problem.

Null constraints such as eq. (2.23) will be central to this work. They balance spin-two states against higher spin states: as visible from figure 5, the average vanishes for spin 0, is negative for spin 2, and positive for all other spins. This implies that, as soon as one particle of spin 2 is present, higher-spin particles must also be present, with predictable properties. (Spin two particles are singled out by the physical assumption that double-subtracted sum rule converges.)

3 Optimization framework

The B_k sum rules just introduced, coupled with positivity of high-energy averages $\langle \cdot \rangle$ provide a complete apparatus to establish potent self-consistency conditions on EFT coefficients g_k 's (defined in eq. (2.3)). We recall our physical assumptions:

- Double-subtracted dispersion relations converge
- The low-energy amplitude is crossing symmetric
- The high-energy spectral density is positive

Since we are considering averages over heavy states (with $m > M$), the coefficients (except in subsection 3.5) are naturally normalized by g_2 and the EFT cutoff M . We will therefore be bounding dimensionless ratios:

$$\tilde{g}_3 = g_3 \frac{M^2}{g_2}, \quad \tilde{g}_4 = g_4 \frac{M^4}{g_2}, \quad \tilde{g}_5 = g_5 \frac{M^6}{g_2}, \quad \dots \quad (3.1)$$

Optimal bounds on these \tilde{g}_k 's will be found by formulating a *dual problem*, in which we combine the desired averages (such as 2.18) with null constraints (such as eq. (2.23)) to obtain sign-definite sum rules. We first describe a simple example analytically, then describe a systematic implementation as a semi-definite problem amenable to publicly available software like SDPB [27].

3.1 Warm-up problem with three sum rules

As a warm-up, let us ask whether it is possible to lower-bound the \tilde{g}_3 coefficient using the B_2, B_4 sum rules previously calculated. We consider the corresponding system of three equations from (2.22) (including the null constraint obtained via g_4 data):

$$g_2 = \left\langle \frac{1}{m^4} \right\rangle, \quad g_3 = \left\langle \frac{3 - \frac{4}{d-2} \mathcal{J}^2}{m^6} \right\rangle, \quad 0 = \left\langle \frac{\mathcal{J}^2(2\mathcal{J}^2 - 5d + 4)}{m^8} \right\rangle. \quad (3.2)$$

With these definitions, let us examine a similar, but simpler set of relations:

$$h_2 = \left\langle \frac{1}{m^4} \right\rangle, \quad h_3 = \left\langle \frac{a - \mathcal{J}^2}{m^6} \right\rangle, \quad 0 = \left\langle \frac{\mathcal{J}^4 - b\mathcal{J}^2}{m^8} \right\rangle \equiv \left\langle n(m^2, J) \right\rangle. \quad (3.3)$$

These relations take on the same form as original identities when $a = \frac{3(d-2)}{4}$, $b = \frac{5d-4}{2}$ and the coupling is rescaled to $g_3 = \frac{4}{d-2} h_3$. Consequently, our warm-up problem is to lower-bound h_3 .

What makes a finite lower bound plausible is that the null constraint (the third equation) should somehow prevent large spins from contributing too much. This is an important point: the allowed range for \tilde{g}_3 is restricted by *higher derivative* crossing equations!

We now calculate a lower bound in two ways. The first – and simplest – method is to use the Cauchy-Schwarz inequality with the null constraint:

$$\left\langle \frac{\mathcal{J}^2}{m^6} \right\rangle^2 \leq \left\langle \frac{1}{m^4} \right\rangle \left\langle \frac{\mathcal{J}^4}{m^8} \right\rangle = b \left\langle \frac{1}{m^4} \right\rangle \left\langle \frac{\mathcal{J}^2}{m^8} \right\rangle. \quad (3.4)$$

Then, using the fact that $m > M$ and $\mathcal{J}^2 \geq 0$ inside the average yields $\langle \frac{\mathcal{J}^2}{m^8} \rangle \leq \frac{1}{M^2} \langle \frac{\mathcal{J}^2}{m^6} \rangle$, and by dividing both sides by that average, we obtain an upper-bound on $\langle \frac{\mathcal{J}^2}{m^6} \rangle$ as desired:

$$\left\langle \frac{\mathcal{J}^2}{m^6} \right\rangle \leq \frac{b}{M^2} \left\langle \frac{1}{m^4} \right\rangle \Rightarrow h_3 \geq -\frac{b}{M^2} h_2. \tag{3.5}$$

This has a simple physical interpretation: if we define an impact parameter $\tilde{b} = \frac{2J}{m}$, then we have effectively shown that heavy states can't contribute at impact parameters much larger than $\sim 1/M$. In terms of the original problem (3.2), we have shown that

$$-\frac{2(5d-4)}{d-2} \leq \frac{g_3 M^2}{g_2} \leq 3. \tag{3.6}$$

This shows the existence of two-sided bounds for generic couplings. This is an important qualitative result, to our knowledge originally emphasized in [11, 12]: ratios of EFT couplings, in units of the cutoff scale M , *must* be $\mathcal{O}(1)$ numbers. Numerically, however, the Cauchy-Schwarz method does not yield the optimal lower bound.

In contrast, the second – and more powerful – method is re-interpret the above task as a semi-definite problem, in order to systematically search for optimal bounds. Denote $h_i(m^2, J)$ the function whose average gives h_i . The idea is to construct positive-definite combinations of the three averages in eq. (3.3):

$$\begin{aligned} F(m^2, J) &\equiv h_3(m^2, J) + \alpha h_2(m^2, J) + \beta n(m^2, J) \\ &= \frac{a - \mathcal{J}^2}{m^6} + \frac{\alpha}{m^4} + \beta \frac{\mathcal{J}^4 - b\mathcal{J}^2}{m^8}, \end{aligned} \tag{3.7}$$

where we must find α, β such that $F(m^2, J) \geq 0$ for all $J = 0, 2, 4 \dots$ and $m \geq M$. Taking the average of any such F then proves $h_3 \geq -\alpha h_2$. The optimal bound will come from a non-negative F with minimal α .

Let us first reproduce the first Cauchy-Schwarz argument in this language, which should give $\alpha = \frac{b}{M^2}$. Assume $\beta > 0$. The argument amounts to completing squares in the \mathcal{J}^4/m^8 term:

$$F(m^2, J) = \frac{\alpha - \beta\lambda^2}{m^4} + \frac{a}{m^6} + \frac{\mathcal{J}^2}{m^6} (2\beta\lambda - 1) - \underbrace{\beta b \frac{\mathcal{J}^2}{m^8} + \beta \frac{1}{m^4} \left(\frac{\mathcal{J}^2}{m^2} - \lambda \right)^2}_{\text{positive}}. \tag{3.8}$$

For any λ this is an identical rewriting of eq. (3.7), and the Cauchy-Schwartz-like method is to choose λ, β such that the other terms are positive as well. From the limit $\mathcal{J} \rightarrow \infty$, the terms with \mathcal{J}^2 need to give a positive functions of m , which imposes that $(2\beta\lambda - 1 - \beta b/M^2) \geq 0$. To minimize α we must minimize $\beta\lambda^2$; we find that the minimum saturates the inequality, and is simply $\beta = M^2/b$ with $\beta\lambda = 1$. With this choice, our trial functional becomes

$$F(m^2, J) = \frac{\alpha - \frac{b}{M^2}}{m^4} + \frac{a}{m^6} + \underbrace{\frac{\mathcal{J}^2}{m^6} \left(\frac{1}{M^2} - \frac{1}{m^2} \right) + \frac{M^2}{bm^4} \left(\frac{\mathcal{J}^2}{m^2} - \frac{b}{M^2} \right)^2}_{\text{positive}}. \tag{3.9}$$

The minimal α for which the first terms are positive for all m^2 is then $\alpha = \frac{b}{M^2}$, precisely as anticipated! We have thus exhibited a positive functional F which proves the Cauchy-Schwarz bound in (3.5).

It is now easy to see why this bound is not optimal: F doesn't *need* to be expressible as a sum of three separately positive parts!

In $d = 4$, for example, the above argument gives $-16 \leq \tilde{g}_3$. In comparison, using the numerical search strategy detailed in the next subsections, we find that the following combination is positive for all $J = 0, 2, 4 \dots$ and $m \geq M$:

$$0 \leq F(m^2, J) = \begin{pmatrix} 10.61249 \\ 1 \\ 0.0671875 \end{pmatrix} \cdot \begin{pmatrix} g_2(m^2, J) \\ M^2 g_3(m^2, J) \\ M^4 n_4(m^2, J) \end{pmatrix}. \quad (3.10)$$

This allows to infer, by taking the average of this inequality, that

$$-10.61249 \leq \tilde{g}_3. \quad (3.11)$$

This is significantly stronger than $-16 \leq \tilde{g}_3$ that we just derived in an ad hoc manner. How can we understand the solution (3.10) analytically? The key ingredient will be that spins J are discrete, whereas our ad hoc bound treated J 's as continuous parameter.

We now calculate this bound analytically. Let us return to the functional ansatz (3.7) and try to directly constrain the unknowns α, β . Putting $J = 0$ we only deduce $\alpha \geq 0$. Putting $J = 2, 4 \dots$ gives a sequence of quadratic polynomials in $(m^2 - M^2)$, each with positive curvature. Such polynomials are non-negative if the two roots are negative, *or* if both roots are positive and equal, *or* if they are complex conjugate pairs. It seems virtually impossible to guess a priori which case is realized, however this information is readily gleaned by plotting the numerical polynomials (3.10), as shown in figure 6. We see that the $J = 2, 4$ inequalities are both saturated: the former by having a root at $m^2 = M^2$, and the latter by having a positive double root (i.e. vanishing discriminant).

These two saturated inequalities give algebraic equations that may be solved analytically; this determines the vector $(\alpha_*, 1, \beta_*)$ in eq. (3.10) to be:

$$\alpha_* = \frac{5d - 2 + \sqrt{\frac{(d+3)(319d^3 + 76d^2 - 292d + 32)}{6(d+1)(d+4)}}}{2(d-2)}, \quad \beta_* = \frac{1}{2d(d-1)} \left(\alpha_* - \frac{5d-2}{d-2} \right). \quad (3.12)$$

In $d = 4$, this gives $\alpha_* = \frac{9}{2} + \frac{7}{4}\sqrt{\frac{61}{5}} \approx 10.61249 \dots$ precisely as found numerically! This result showcases the use of numerics to guide our analytical understanding. Moreover, the preceding formula can be shown to give the correct optimal lower bound on \tilde{g}_3 in any $d \geq 3$, when including the single null constraint n_4 . (When we add more equations below, the $d = 4$ bound will converge to $\tilde{g}_3 \geq -10.346$.)

3.2 Dual problem: general formulation

We now introduce the general “dual” optimization problem which allows to carve out the space of EFT coefficients $\tilde{g}_3, \tilde{g}_4, \tilde{g}_5, \dots$ allowed by unitarity and positivity.

The data at our disposal comes from the B_k sum-rules in eq. (2.19):

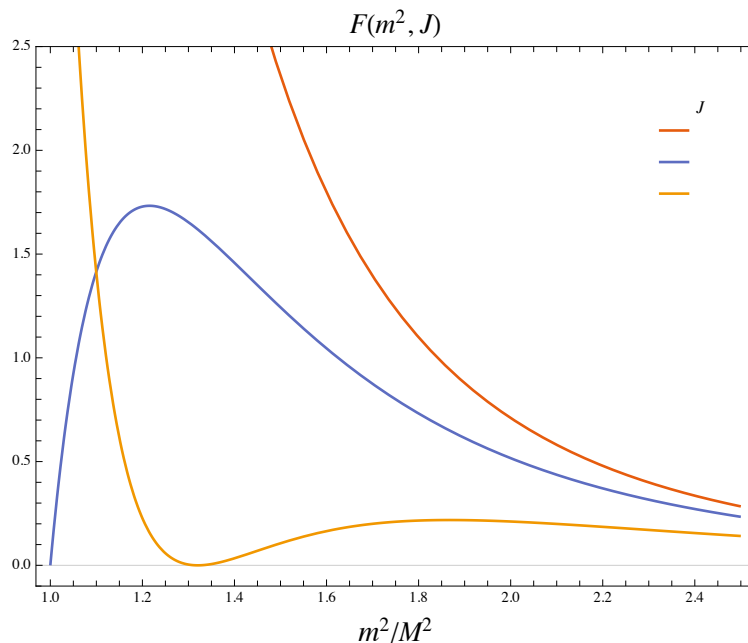


Figure 6. The non-negative function $F(m^2, J)$ returned by SDPB when maximizing the lower bound for \tilde{g}_3 . The spin 4 curve is particularly enlightening because it features a double root.

1. Representative averages $g_k(m^2, J)$ which measure each “desired” g_k :

$$g_2 = \langle g_2(m^2, J) \rangle, \quad g_3 = \langle g_3(m^2, J) \rangle, \quad g_4 = \langle g_4(m^2, J) \rangle, \quad \dots \quad (3.13)$$

2. A set of null functions $n_i(m^2, J)$ whose heavy averages vanish:

$$0 = \langle n_4(m^2, J) \rangle = \langle n_5(m^2, J) \rangle = \langle n_6(m^2, J) \rangle = \dots \quad (3.14)$$

First consider the problem of lower-bounding g_3 , given g_2 , but being agnostic about the other g_k 's with $k \geq 4$. Define a vector of functions which combines g_2, g_3 and null constraints:

$$v(m^2, J) \equiv (g_2(m^2, J), M^2 g_3(m^2, J), n_4(m^2, J), n_5(m^2, J), \dots). \quad (3.15)$$

The following (“dual”) optimization problem then determines a lower bound on g_3 :

$$\begin{cases} \text{maximize:} & A \\ \text{subject to:} & 0 \leq (-A, 1, c_4, c_5, \dots) \cdot v(m^2, J) \quad \forall m \geq M, \forall J = 0, 2, 4, \dots \end{cases} \quad (3.16)$$

Here we are maximizing over A and all possible linear combinations c_i of the null constraints $n_i(m^2, J)$. Having found such a linear combination c_i and an optimal number A , it readily follows from linearity and positivity of $\langle \cdot \rangle$ that

$$0 \leq -Ag_2 + M^2 g_3 + 0, \quad (3.17)$$

allowing the desired *optimal* bound on \tilde{g}_3 to be inferred:

$$A \leq \tilde{g}_3. \tag{3.18}$$

This equation motivates the maximization of A . Alternatively, to find an optimal *upper* bound B , it is enough to consider the analogous problem:

$$\begin{cases} \text{minimize:} & B \\ \text{subject to:} & 0 \leq (B, -1, c_4, c_5, \dots) \cdot v(m^2, J) \quad \forall m \geq M, \forall J = 0, 2, 4, \dots \end{cases} \tag{3.19}$$

It then follows that $\tilde{g}_3 \leq B$. Taking the intersection of these two sets yields a convex region defined by

$$A \leq \tilde{g}_3 \leq B. \tag{3.20}$$

These are the inequalities inferred on \tilde{g}_3 from data in the B_k sum-rules.

In theory, the null constraints n_i are part of an infinite dimensional vector space, but for numerical purposes, the dimensionality is taken to be finite and determined by the truncation order of the low-energy expansion \mathcal{M}_{low} . We will find that the optimal bounds converge rapidly as the maximal degree is increased.

This problem can be readily adapted if we are given additional assumptions. For example, to make exclusion plots in the $(\tilde{g}_3, \tilde{g}_4)$ plane, one strategy is to postulate some value of $\tilde{g}_3^{(0)}$ in the allowed range, and for each value we determine lower and upper bounds on \tilde{g}_4 . The vector v in eq. (3.15) then acquires an extra g_4 row:

$$v \mapsto (g_2, M^2 g_3, M^4 g_4, n_4, n_5, \dots). \tag{3.21}$$

Imposing positivity of $(-A, -B, 1, c_4, c_5, \dots) \cdot v(m, J)$ for all m and J yields a lower bound $\tilde{g}_4 \geq A + B\tilde{g}_3^{(0)}$, for example. In fact, since the EFT parameters enter linearly, the resulting inequality automatically carves out a half-space in the $(\tilde{g}_3, \tilde{g}_4)$ -plane:

$$\tilde{g}_4 - B\tilde{g}_3 \geq A. \tag{3.22}$$

This half-plane is tangent to the allowed region at $\tilde{g}_3 = \tilde{g}_3^{(0)}$. If this process is repeated for distinct $\tilde{g}_3^{(0)}$, a collection of planes is rapidly generated from which it is possible to carve out the convex allowed space.

This processes generalizes to higher dimensional planes (i.e. hyperplanes): the vector v always contains only those coefficients we are not being agnostic about, plus an arbitrary number of null constraints. Although we will not go beyond three-dimensional regions, we note that an efficient search algorithm in higher dimensions is described in ref. [28].

3.3 Example with a SDPB implementation

The optimization problem just formulated is in a form that is directly amenable to the SDPB solver [27]. There are just two simple substitutions to make:

1. The program accepts polynomials of $x \geq 0$; we set $m^2 \mapsto M^2(1+x)$ and remove a common denominator.

2. The program accepts finite lists of polynomial constraints. We tabulate a finite list of spins $J = 0, 2, \dots, J_{\max}$ and add a single function of x corresponding to $J \rightarrow \infty$.

The second truncation is valid as long as J_{\max} is taken sufficiently large; once convergence is achieved, further increases of J_{\max} have no effect on the bounds.

We consider now an example relevant to one of the plots in the next section, when working in $d = 4$, $J_{\max} = 40$ and Mandelstam order $n = 4$. The goal is to find a lower bounding plane on \tilde{g}_4 for fixed \tilde{g}_3 , say $\tilde{g}_3 = -10.5$ (which is slightly above the allowed lower bound (3.10) for this truncation). The polynomial vector v combines the three observed coefficients and a single null constraint n_4 , rescaled by appropriate powers of M :

$$v(x, J) \equiv M^4(1+x)^4 \times (g_2, M^2 g_3, M^4 g_4, M^4 n_4) \tag{3.23}$$

$$= \left((1+x)^2, (1+x)(3-2J(J+1)), \frac{1}{2}, 2J(J+1)(J(J+1)-8) \right). \tag{3.24}$$

In addition to tabulating v for all even $J \leq J_{\max}$, we also include the infinite- J limit, which is simply the coefficient of J^4 :

$$v(x, \infty) = (0, 0, 0, 2). \tag{3.25}$$

To lower-bound \tilde{g}_4 at the stated value of \tilde{g}_3 , we search for four-vectors y , normalized to $y \cdot (0, 0, 1, 0) = 1$, which solve the following problem:

$$\begin{cases} \text{maximize:} & h = y \cdot (-1, -(-10.5), 0, 0) \\ \text{subject to:} & 0 \leq y \cdot v(x, J) \quad \forall x > 0, J \in \{0, 2, \dots, 40, \infty\}. \end{cases} \tag{3.26}$$

The lower bound is then

$$\tilde{g}_4 \geq h \quad (\text{at } \tilde{g}_3 = -10.5). \tag{3.27}$$

The solution vector y computed by SDPB was found to be $y \approx (1.4823, 0.1810, 1, 0.01472)$, giving $h = 0.4183$. The half-plane allowed by positivity of $y \cdot v$ is thus

$$\tilde{g}_4 + 0.1810(\tilde{g}_3 + 10.5) \geq 0.4183 \tag{3.28}$$

which gives one of the boundaries used to make the $n = 4$ region in figure 10 (b) below.

3.4 Generating null constraints

So far we used a single constraint from crossing: the null average $n_4(m^2, J)$ from eq. (2.23). The bounds improve after we add more constraints. Let us describe a way to generate them.

A straightforward method is as follows. First, fix a degree in Mandelstam invariants, n . Then, list all the $\lfloor \frac{n}{2} \rfloor$ low-energy averages corresponding to this degree, namely the coefficient of t^{n-k} in the left-hand-side of the $B_k(t)$ sum rules (2.19) with $k \leq n$, using the crossing-symmetric low-energy ansatz (2.3). The right-hand-sides of those linear combinations with vanishing left-hand-side then constitute a basis of null constraints.

The first few cases, up to degree $n = 7$, are (in arbitrary normalization):

$$\begin{aligned}
 m^8 n_4 &= (4 - 5d)\mathcal{J}^2 + 2\mathcal{J}^4, \\
 m^{10} n_5 &= (23d^2 - 12d - 20)\mathcal{J}^2 + (-21d - 2)\mathcal{J}^4 + 4\mathcal{J}^6, \\
 m^{12} n_6 &= 4(13d^2 + 21d + 2)\mathcal{J}^4 - 3(d + 2)(17d^2 + 4d - 16)\mathcal{J}^2 \\
 &\quad + 2(-9d - 8)\mathcal{J}^6 + 2\mathcal{J}^8, \\
 m^{14} n_7 &= 6(45d^2 + 140d + 92)\mathcal{J}^6 + 4(-140d^3 - 619d^2 - 711d - 46)\mathcal{J}^4 \\
 &\quad + (401d^4 + 2108d^3 + 2284d^2 - 3008d - 3360)\mathcal{J}^2 \\
 &\quad + 5(-11d - 18)\mathcal{J}^8 + 4\mathcal{J}^{10}, \\
 m^{14} n'_7 &= (-27d - 14)\mathcal{J}^4 + 2(d - 1)(19d + 22)\mathcal{J}^2 + 4\mathcal{J}^6
 \end{aligned} \tag{3.29}$$

where we recall that $\mathcal{J}^2 = J(J+d-3)$. Note that *all* null averages vanish when $J = 0$: they relate spinning heavy states to one another, but spinless states are completely decoupled.

We find that there exists a single null constraint for each degree $n = 4, 5, 6$, then two constraint each for $n = 7, 8, 9$, three each for $n = 10, 11, 12$, etc.: the number of linearly independent null constraints at each degree increases by 1 for every increase of n by 3. A sequence of generating functions $X_k(t)$ which enumerates them all is discussed in appendix A.

Finally, it is important to stress that the null constraints only average to zero modulo EFT loops, since the method for finding them relies on the explicit tree-level parameterization (2.3). The interpretation of resulting inequalities as bounds on g_k is thus only strictly valid in this approximation. In an interacting EFT, the coefficients g_k depend on choices of scale and renormalization scheme, and the correct interpretation of the positive functionals F is that they give rigorous (possibly non-optimal) inequalities of the form:

$$\left\langle g_3(m^2, J) + \sum_i c_i n_i(m^2, J) \right\rangle \geq A \langle g_2(m^2, J) \rangle \tag{3.30}$$

where all averages are computed as integrals over arcs with $|s| \approx M^2$ following eq. (2.18). The method thus produces rigorous bounds on computable combinations of EFT couplings at the scale M .

3.5 An ad hoc upper bound on $(\partial\phi)^4$

The systematic method explained above bounds ratios g_k/g_2 , but how about g_2 itself? Here we present one upper bound on g_2 ; this subsection is somewhat separate from the rest since we were unable to systematically optimize the bound.

The naive intuition is that if heavy states couplings are order unity, then $g_2 \sim \frac{1}{M^d}$. An upper bound with this scaling should thus follow from the unitarity limit, $\rho_J \leq 2$. This would be the full story *if* heavy states only had a finite number of spins, however, to get an actual bound one must also control the infinite sum over spins. The idea is to combine the sum rule for g_2 with a multiple of the first null constraint n_4 :

$$g_2 = \left\langle \frac{1}{m^4} - \alpha \frac{f(J)}{m^8} \right\rangle \quad \text{with} \quad f(J) = \mathcal{J}^2(2\mathcal{J}^2 - (5d - 4)). \tag{3.31}$$

This holds for any α ; we recall that $\mathcal{J}^2 = J(J + d - 3)$ is the spin Casimir. Inserting the definition (2.15) of the average and switching to a dimensionless mass parameter $z = M^2/m^2 < 1$, this can be rewritten

$$g_2 = \frac{1}{\pi M^d} \sum_{J \text{ even}} n_J^{(d)} \int_0^1 \frac{dz}{z} z^{\frac{d}{2}} \rho_J(M^2/z) \left(1 - \alpha z^2 f(J)\right). \quad (3.32)$$

This is a good step to get an upper bound on g_2 since for $\alpha > 0$ the integrand is mostly negative at large spin J , except for a small region with z small. We use unitarity in two steps: first, we use positivity $\rho_J \geq 0$ to restrict the integral to the small- z region where the parenthesis is positive, where we can then use $\rho_J \leq 2$. Thus:

$$g_2 \leq \frac{2}{\pi M^d} \sum_{J \text{ even}} n_J^{(d)} \int_0^{z_{\max}(J)} \frac{dz}{z} z^{\frac{d}{2}} \left(1 - \alpha z^2 f(J)\right) \quad (3.33)$$

where $z_{\max}(J) = \min([\alpha f(J)]^{-\frac{1}{2}}, 1)$ determines the region where the parenthesis is positive. The important point is that at large spin this region shrinks, which will ensure convergence of the J sum. Effectively the region is bounded by impact parameter $M\tilde{b} \lesssim \alpha^{-1/4}$. At small spins the full range is generally accessible. Letting $z_{\max}(J) = 1$ for $J \leq J_*(\alpha)$, the sum splits as

$$g_2 \leq \frac{2}{\pi M^d} \left[\sum_{0 \leq J \leq J_*(\alpha)} n_J^{(d)} \left(\frac{2}{d} - \frac{2\alpha f(J)}{d+4}\right) + \frac{8\alpha^{-\frac{d}{4}}}{d(d+4)} \sum_{J > J_*(\alpha)} \frac{n_J^{(d)}}{f(J)^{\frac{d}{4}}} \right]. \quad (3.34)$$

Both sums run only over even spins. Since $n_J^{(d)} \propto J^{d-3}$ at large spin, the sum to infinity converges. This inequality is valid for any $\alpha > 0$; with increasing α , the $J = 0, 2$ terms tend to increase whereas the rest decreases: the optimal bound with this method is obtained by minimizing over α . The dependence is rather non-linear (which is why we were not able to generalize the method to include more null constraints), but evaluating the sum numerically we find (in all dimensions we tried) that the optimum occurs with $J_* = 2$. In $d = 4$, for example, the optimal value $\alpha_* = 0.025$, giving the analytic bound:

$$0 \leq \frac{g_2}{(4\pi)^2} \leq \frac{0.794}{M^4} \quad (d = 4). \quad (3.35)$$

As expected, up to a standard loop factor, the coefficient of $\frac{1}{2}(\partial\phi)^4$ can't exceed order unity in units of the heavy scale. We stress that, contrary to other bounds in this paper, upper bounds on couplings at the cutoff scale cannot be straightforwardly interpreted in terms of Lagrangian parameters, since any EFT which saturates them is by definition strongly interacting already below the cutoff, making quantum corrections non-negligible. Rather, the bound may be interpreted as follows: among all observables which are linear in the S -matrix at the scale M and which reduce to g_2 at weak coupling, there exists one which satisfies eq. (3.35): namely, eq. (3.31) with the quoted α_* . This observable resides at the scale M in the sense of eq. (3.30).

Similar bounds in different spacetime dimensions are recorded in table 1; after normalizing with a suitable d -dimension factor (closely related to $n_0^{(d)}$) we find that the obtained bound is pretty stable in dimensions.

d	3	4	5	6	7	8	9	10	26	50
bound	0.155	0.132	0.138	0.140	0.141	0.141	0.140	0.140	0.132	0.127

Table 1. Nonperturbative (possibly non-optimal) upper bounds on $\frac{g_2 M^d \Gamma(\frac{d}{2})}{(4\pi)^{\frac{d}{2}} \Gamma(d)}$ in various spacetime dimensions.

n	dim N	min \tilde{g}_3	min \tilde{g}_5	min \tilde{g}'_6
4	1	-10.6125	$-\frac{25}{6} \approx -4.17$	-15
5	2	-10.6125	$-\frac{25}{6} \approx -4.17$	
6	3	-10.3662	-4.0969	-13.436
7	5	-10.3580		
8	7	-10.3492	-4.0960	-12.926
9	9	-10.3492		
10	12	-10.3477		
11	15	-10.3474		
12	18	-10.3473	-4.0960	-12.820
13	22	-10.3470		
14	26	-10.3468		
15	30	-10.3466		
16	35	-10.3465	-4.0960	-12.811

Table 2. Convergence of the “box bounds” (4.2) with increasing Mandelstam order n , which results in larger sets of orthogonal constraints $N = \{n_4(m^2, J), \dots\}$. At order $n = 3$, all lower bounds would be $-\infty$ since no null constraints exist at that order.

4 Numerically ruling-out: the allowed space of scalar EFTs

In this section, we summarize the obtained numerical results, focusing on $d = 3 + 1$ space-time dimensions. Treating the low-energy EFT to tree-level, we determine the space of EFT coefficients g_n , where the low-energy amplitude is parameterized as

$$\begin{aligned}
 \mathcal{M}_{\text{low}} = & -g^2 \left[\frac{1}{s} + \frac{1}{t} + \frac{1}{u} \right] - \lambda \\
 & + g_2(s^2 + t^2 + u^2) + g_3(stu) + g_4(s^2 + t^2 + u^2)^2 \\
 & + g_5(s^2 + t^2 + u^2)(stu) + \dots
 \end{aligned}
 \tag{4.1}$$

Recall that it is convenient to introduce dimensionless EFT coefficients \tilde{g}_n normalized by g_2 and appropriate powers of the mass gap M introduced in equation (3.1), since the numerical analysis is performed directly on these variables. We find optimal upper and lower bounds for \tilde{g}_3 , \tilde{g}_4 , and \tilde{g}_5 , given positivity of high-energy spectral densities ρ_J , using the optimization framework introduced above.

4.1 Bounds on individual coefficients

Let us begin by confining the value of individual EFT coefficients, being completely agnostic about all the others. Table 2 shows how these bounds depend on the number of crossing symmetry constraints kept. We conclude that, in $d = 3 + 1$ space-time, normalized EFT coefficients satisfy:

$$-10.346 \leq \tilde{g}_3 \leq 3, \quad 0 \leq \tilde{g}_4 \leq \frac{1}{2}, \quad -4.0960 \leq \tilde{g}_5 \leq \frac{5}{2}. \quad (4.2)$$

The simple, rational upper bounds are saturated by the spin-0 contribution to sum rules like eq. (2.22).

While keeping constraints with higher Mandelstam degree n is feasible, runtime-wise, we found that the large-spin convergence was harder to control as we needed $J_{\max} = \mathcal{O}(1000)$. However, for $n \leq 16$, convergence is easily obtained for smaller spins. This is the reason why we stopped the table at $n = 16$; it would be interesting to understand how to stabilize the numerics at large n .

The \tilde{g}_3 lower bound is plotted as a function of $1/\dim N$ in figure 7, where $\dim N$ is the dimension of the number of null constraints accessed. The dimension of this vector space is naturally monotonically increasing with the Mandelstam order n . Ideally, one would like to extrapolate the bounds to $n \rightarrow \infty$, however the approach is somewhat irregular and we didn't find a compelling fit function. Therefore, the recorded bounds are simply taken from our largest reliable value of n .

In $d = 6$ and $d = 10$, the analogous bounds take the values:

$$d = 6 : \quad -8.13879 \leq \tilde{g}_3 \leq 3, \quad 0 \leq \tilde{g}_4 \leq \frac{1}{2}, \quad -3.03373 \leq \tilde{g}_5 \leq \frac{5}{2} \quad (4.3)$$

$$d = 10 : \quad -7.04934 \leq \tilde{g}_3 \leq 3, \quad 0 \leq \tilde{g}_4 \leq \frac{1}{2}, \quad -2.51054 \leq \tilde{g}_5 \leq \frac{5}{2} \quad (4.4)$$

These upper and lower bounds were obtained by solving the optimization problems given in equations (3.16) and (3.19), respectively. For the $n = 16$ data points in the table, we kept spins up to 200 (in addition to $J = \infty$). A full list of the bounds on the first 10 coefficients \tilde{g}_n obtained with $n = 10$ (corresponding to 12 linearly independent null constraints) are presented in appendix B.

4.2 Two-dimensional allowed region in $(\tilde{g}_3, \tilde{g}_4)$ plane

The “box bounds” in eq. (4.2) tell an incomplete story since they miss potential correlations between the coefficients. Exclusion plots on the $(\tilde{g}_3, \tilde{g}_4)$ plane can be obtained following the constrained optimization framework presented at the end of section 3.2. Specifically, upper and lower bounding planes on \tilde{g}_4 were obtained by sampling \tilde{g}_3 at various points in the range allowed by eq. (4.2). Repeating this process for a large number of sampling points generates a collection of linear inequalities, whose intersection defines a refined allowed region. Figure 8 depicts this region with over 100 sampling points along \tilde{g}_3 .

Further insight into the shape of the region can be obtained by noticing that the crossing symmetry constraints do not mix particles of spin 0 with the others, as noted in

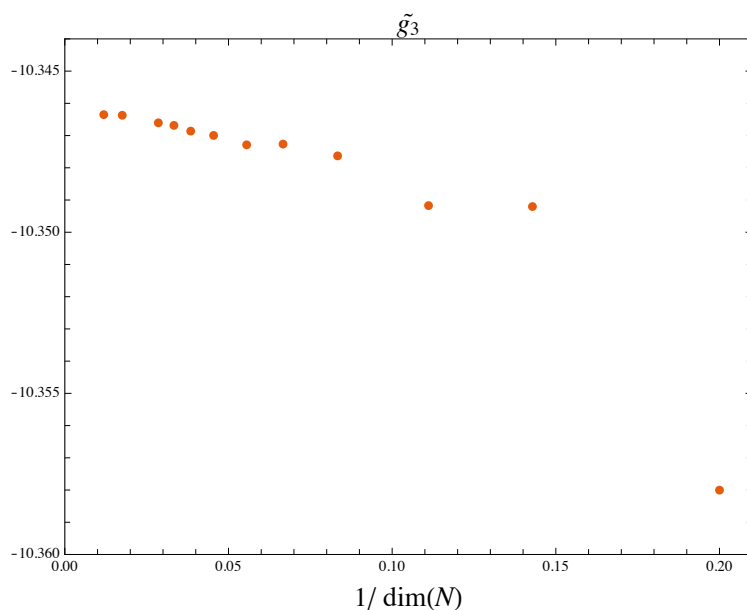


Figure 7. The \tilde{g}_3 lower bound vs. $1/\dim N$, where $\dim N$ is the dimension of the number of null constraints accessed. Plotting with respect to $1/(\dim N)^2$ or different powers of $1/n$ did not reveal any particularly compelling fitting function. The kink described in the next plot is situated at $-10.19 \approx \tilde{g}_3$.

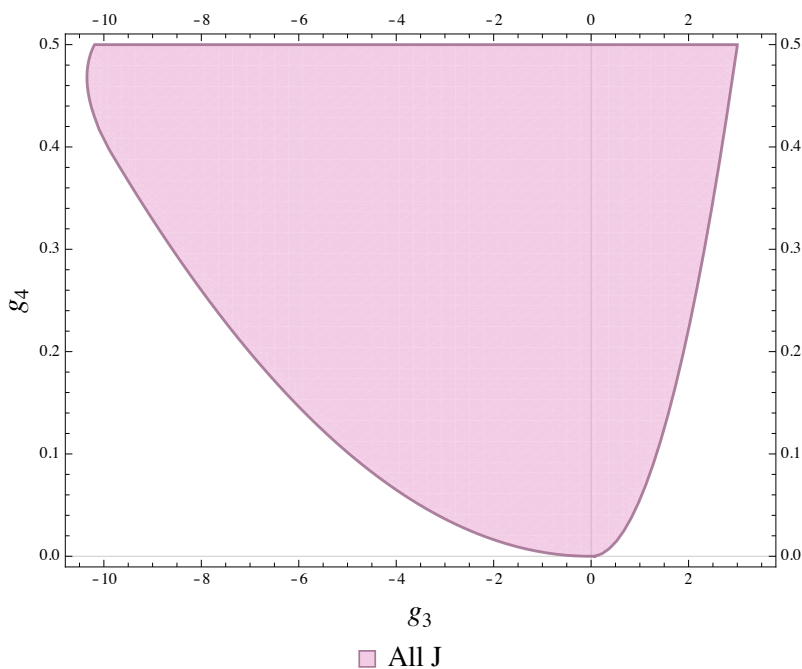


Figure 8. The $(\tilde{g}_3, \tilde{g}_4)$ allowed region. Numerics were performed at $n = 10$ Mandelstam order and $J = 0, 2, \dots, 40$. One can see that g_3 may take on negative values, while g_4 is positive. Boundaries appear smooth except for two kinks at $(-10.19, 0.5)$ and $(3, 0.5)$.

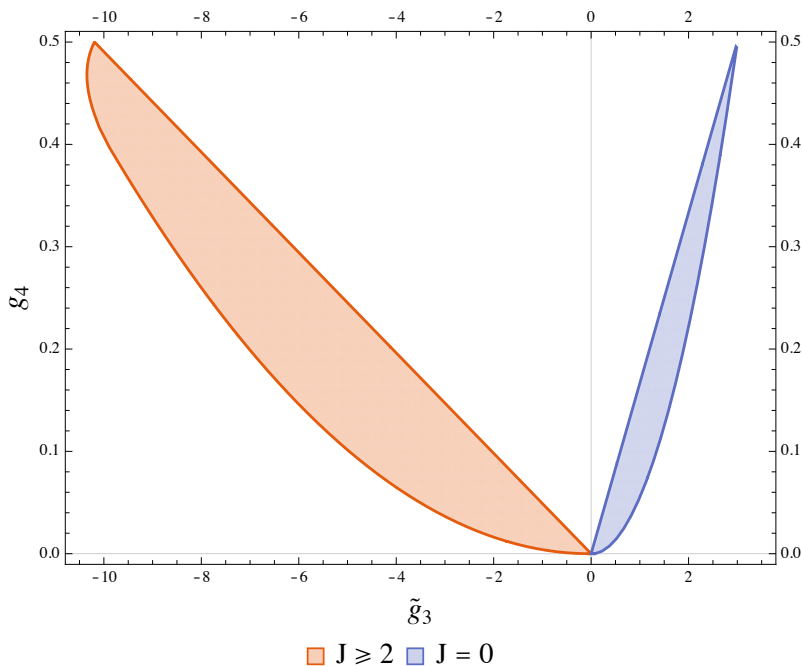


Figure 9. The $(\tilde{g}_3, \tilde{g}_4)$ allowed region, segmented into theories without spinless particles ($J \geq 2$) and theories with only spinless particles ($J = 0$). The convex hull of these two regions reproduces figure 8.

eq. (3.29). This indicates that heavy spin-0 particles satisfy crossing on their own, as will be further discussed in the next section. For numerical purposes, this decoupling allows to refine the problem: any high-energy spectrum can be written as a positive sum of its spin-0 content, plus a unitary solution to crossing that only contains particles of spin $J \geq 2$. The full allowed region is then simply the convex hull of the allowed regions for these two problems:

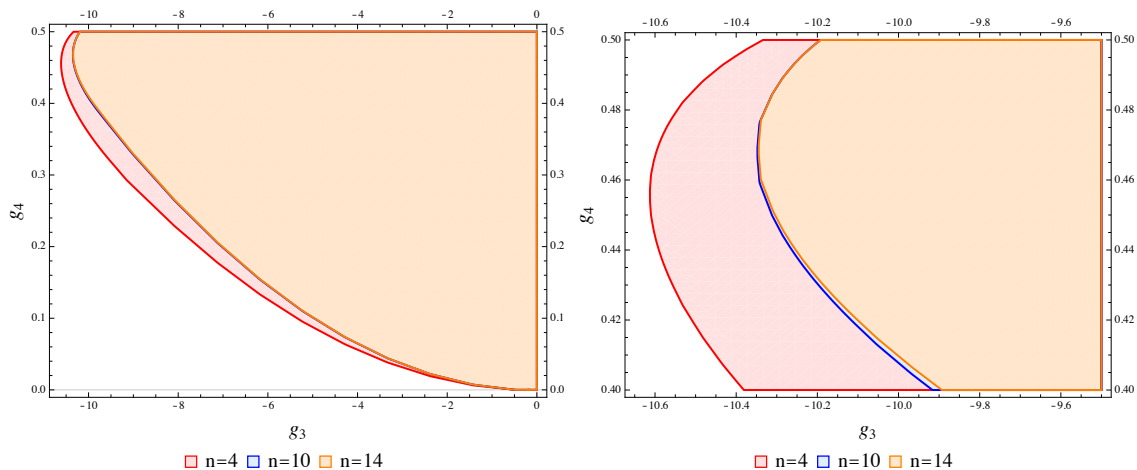
$$\text{Entire region} = \text{Convex Hull} [\text{Spin-0} + \text{Spin-}J \geq 2]. \tag{4.5}$$

As may be seen from the form of the g_3 sum rule (2.22), the two solutions are differentiated by the sign of g_3 : positive for Spin-0 and negative for Spin- $J \geq 2$.

In our implementation of the dual problem, theories with only $J \geq 2$ particles can be studied by simply dropping the positivity constraint for the functional action on $J = 0$. The allowed regions for the Spin-0 and Spin- $J \geq 2$ sub-problems are the narrow almond-shaped regions shown in figure 9.

The shape of these regions is largely explained by a simple scaling argument: given any solution to crossing, scaling-up its overall mass scale will give a new solution. Starting from any allowed point $(\tilde{g}_3, \tilde{g}_4)$, this generates an allowed path $(\alpha\tilde{g}_3, \alpha^2\tilde{g}_4)$ where $0 \leq \alpha \leq 1$. This explains the parabolic shape of the “underbellies” in figure 9. In fact the Spin-0 almond is simply the convex hull of the parabola connecting $(0, 0)$ to $(3, \frac{1}{2})$. (This is qualitatively similar to what is found in the forward limit [11, 12, 14].)

The Spin- $J \geq 2$ region is more complicated – while it also displays a parabolic underbelly near the origin, it fails to extend all the way to $\tilde{g}_4 = \frac{1}{2}$. The boundary must



(a) Convergence is rapidly achieved by adding (b) Close-up near the left kink at $(-10.19, 0.5)$. null constraints.

Figure 10. Convergence of the $(\tilde{g}_3, \tilde{g}_4)$ allowed region (shaded) with the Mandelstam degree n of crossing equations. The positive section of the horizontal axis is omitted since that region converges immediately.

thus exhibit non-analytic behaviour at the end of the parabola, however we were unable to localize a kink that remains stable with varying the Mandelstam expansion order n , suggesting milder non-analyticity (such as a discontinuous second derivative). This qualitative feature is demonstrated by close-ups on the high- \tilde{g}_4 end of the allowed region are shown in figure 10, which also indicates convergence with increasing n .

The line $\tilde{g}_4 = \frac{1}{2}$ features two kinks: at $(-10.19, 0.5)$ and $(3, 0.5)$, as shown in figure 10. Below we will find analytic expressions for the scattering amplitude at these kinks!

We chose the simplest and perhaps weakest method of calculating the exclusion region. Method of radials or normals have the potential more efficiently calculate these boundaries with higher fidelity near kinks [29]. For example, the normals maximization procedure explained extracts sharp features at regions of large curvature. Unfortunately, we were not able to use the normals method because it requires one to solve the “primal” problem instead.

4.3 Three-dimensional allowed volume in $(\tilde{g}_3, \tilde{g}_4, \tilde{g}_5)$

Finally, we consider the space of EFT coefficients $(\tilde{g}_3, \tilde{g}_4, \tilde{g}_5)$ using a similar procedure to the $2d$ process. Once the $2d$ exclusion region is obtained, points in $(\tilde{g}_3, \tilde{g}_4)$ are sampled from the exclusion region and then the optimal upper and lower bounds on \tilde{g}_5 with the associated hyperplanes are computed via the constrained optimization approach. Figure 11 shows that the $3d$ region is narrow, suggesting that S-matrix positivity is a potent constraint on the EFT.

Finally, the exclusion plot converges as the truncation order n is increased as shown in figures 10. In particular, the $n = 10$ and $n = 14$ regions overlap better than the $n = 4$ and $n = 10$ regions. Despite doubling the number of functions $n_i(m^2, J)$ in the null

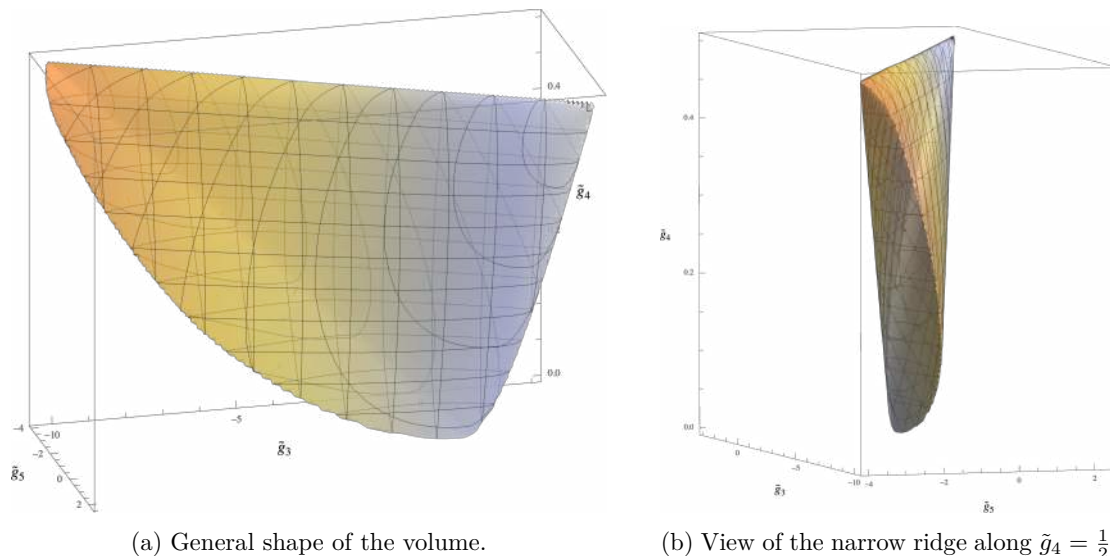


Figure 11. \tilde{g}_3 vs. \tilde{g}_4 vs. \tilde{g}_5 “tortilla chip”. Numerics were performed at $n = 10$ Mandelstam order and $J = 0, 2, \dots, 40$. The volume is surprisingly narrow, showing a strong \tilde{g}_3 - \tilde{g}_5 correlation.

constraints N to sharpen bounds between $n = 10$ and $n = 14$, the numerics show that they generate very similar regions. Therefore, it is sufficient to consider $n = 10$ in order to infer general features.

5 Analytically ruling-in: the two kinks with $\tilde{g}_4 = \frac{1}{2}$

It is interesting to ask whether the bounds on $(\tilde{g}_3, \tilde{g}_4, \dots)$ obtained in the preceding section are indeed optimal. We obtained these bounds using the so-called dual problem, which “rules-out” more and more space as one adds null constraints. This contrasts with the primal problem, whereby the goal is to “rule-in” coefficients of $2 \rightarrow 2$ S-matrices satisfying all the axioms. When results from both problems agree, the optimal solution is apprehended with complete confidence.

Systematic implementations of the primal problem have been proposed for generic field theories (without a large gap M) [30, 31]. Convergence of the dual and primal problems have also been studied for two-dimensional S-matrices [29, 32]. We will not attempt to adapt these methods to our problem, but we will study simple special theories which can be ruled in analytically.

The $(\tilde{g}_3, \tilde{g}_4)$ allowed region in figure 8 prominently displays two kinks with $\tilde{g}_4 = \frac{1}{2}$ (connected with a horizontal segment). This value is interesting, because the EFT parameter g_2 and g_4 satisfy the following sum rules:

$$g_2 = \left\langle \frac{1}{m^4} \right\rangle, \quad g_4 = \frac{1}{2} \left\langle \frac{1}{m^8} \right\rangle, \quad (5.1)$$

where $\langle \cdot \rangle$ signifies a positive sum over states with $m \geq M$ and different spin, as previously defined. The coefficient $g_4 = \frac{g_2}{2M^4}$ can only be realized if the high-energy theory contains *only* states at the single mass $m = M$!

The corresponding amplitudes are therefore rational functions with poles at s , t or $u = m^2$. Furthermore, they must satisfy the boundedness property $\lim_{s \rightarrow 0} \frac{\mathcal{M}(s,t)}{s^2} = 0$. It is easy to see that there are only two crossing-symmetric rational functions with these properties:

$$\begin{aligned} \mathcal{M}_{\text{spin-0}} &= \frac{1}{m^2 - s} + \frac{1}{m^2 - t} + \frac{1}{m^2 - u}, \\ \mathcal{M}_{\text{stu-pole}} &= \frac{m^4}{(m^2 - s)(m^2 - t)(m^2 - u)} - \gamma(d)\mathcal{M}_{\text{spin-0}}. \end{aligned} \tag{5.2}$$

Here, $\gamma(d)$ is a coefficient that will ensure unitarity and positivity of the amplitude in d -dimensional space-time. In particular, we have that

$$\gamma(d) \equiv \frac{4}{9} {}_2F_1\left(\frac{1}{2}, 1, \frac{d-1}{2}, \frac{1}{9}\right). \tag{5.3}$$

To show this, consider the spectral density of the first amplitude (its imaginary part) is supported entirely on spin-0. Following the decomposition used in figure 9, we have then subtracted a constant multiple of $\mathcal{M}_{\text{spin-0}}$ to remove the spin-0 content from $\mathcal{M}_{\text{stu-pole}}$. To find the left kink, we should thus tune $\gamma(d)$ to make the spectral density supported only on spins $J \geq 2$.

In general, the Gegenbauer polynomials satisfy an orthogonality relation [17] that allows one to extract the coefficient:

$$f_J(s) = \frac{\mathcal{N}_d}{2} \int_{-1}^1 dz (1 - z^2)^{\frac{d-4}{2}} \mathcal{P}_J(z) \mathcal{M}(s, t(z)), \tag{5.4}$$

where $z = 1 + \frac{2t}{s}$ and $\mathcal{N}_d = \frac{(16\pi)^{\frac{2-d}{2}}}{\Gamma(\frac{d-2}{2})}$ we find

$$\rho_J(s)|_{\text{stu-pole}} = 2\pi \mathcal{N}_d m^{d-4} \delta(s - m^2) \int_{-1}^1 dz (1 - z^2)^{\frac{d-4}{2}} \mathcal{P}_J(z) \left(\frac{1}{9 - z^2} - \frac{\gamma(d)}{4} \right). \tag{5.5}$$

Plugging in $\mathcal{P}_0(z) = 1$ and computing the integral, we find that the spin-0 component vanishes provided we choose:

$$\gamma(d) = \frac{4}{9} {}_2F_1\left(\frac{1}{2}, 1, \frac{d-1}{2}, \frac{1}{9}\right). \tag{5.6}$$

To verify that the partial waves are positive for $J \geq 2$, one can use the Froissart-Gribov formula (see eq. (2.53) of [17]) to analytically integrate (5.5) in terms of a single residue at $z = \pm 3$; the integral is proportional to the function defined as $Q_J^{(d)}(3)$ there, which is a positive hypergeometric function for any $d \geq 3$. We conclude that the amplitude $\mathcal{M}_{\text{stu-pole}}$, with the value (5.3), satisfies all the axioms of crossing symmetry, spin-2 Regge convergence, and positivity $\rho_J(s) \geq 0$! Of course, this does not imply that this amplitude can indeed be realized in some fully-fledged UV complete theory, only that it cannot be ruled out with our current methods.

Let us now situate the amplitudes $\mathcal{M}_{\text{spin-0}}$ and $\mathcal{M}_{\text{stu-pole}}$ in the context of our numerical exclusion plots. By series expanding at small s, t, u and comparing with the low-energy

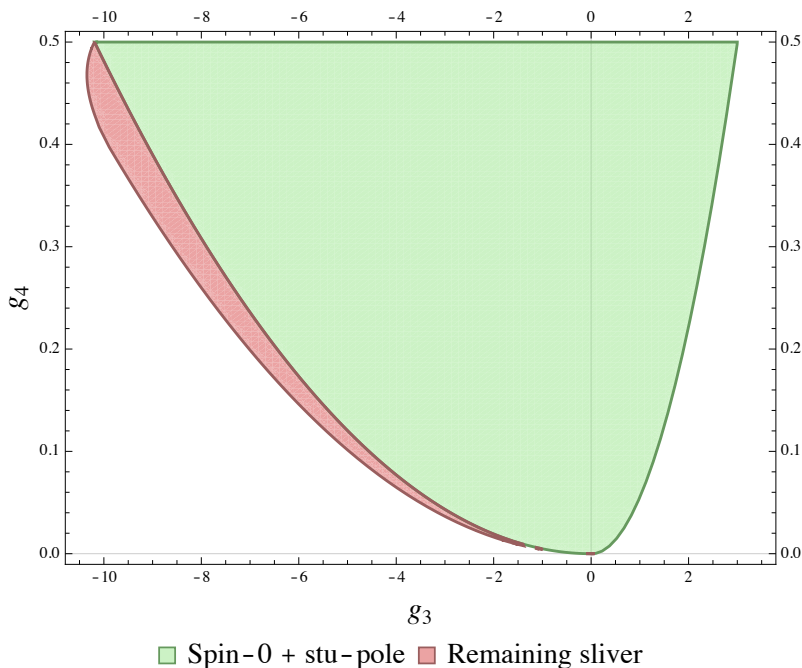


Figure 12. Remaining allowed sliver compared with the positive span of the two analytical models: $\mathcal{M}_{stu\text{-pole}}$ and $\mathcal{M}_{\text{spin-0}}$, which lie at the left and right upper kinks, respectively.

parameterization 2.6, it is straightforward to find that

$$\mathcal{M}_{\text{spin-0}} : \quad (\tilde{g}_3, \tilde{g}_4, \tilde{g}_5) = \left(\frac{3}{\tilde{m}^2}, \frac{1}{2\tilde{m}^4}, \frac{5}{2\tilde{m}^6} \right) \tag{5.7}$$

$$\mathcal{M}_{stu\text{-pole}} : \quad (\tilde{g}_3, \tilde{g}_4, \tilde{g}_5) = \left(\frac{2}{\tilde{m}^2} \frac{3\gamma(d) - 1}{2\gamma(d) - 1}, \frac{1}{2\tilde{m}^4}, \frac{1}{\tilde{m}^6} \frac{5\gamma(d) - 2}{2\gamma(d) - 1} \right), \tag{5.8}$$

where $\tilde{m} \equiv \frac{m}{M}$. At $\tilde{m} = 1$ and $d = 4$, the spin-0 model produces the upper-right kink at $(3, \frac{1}{2})$, whereas the *stu*-pole model gives $(\tilde{g}_3, \tilde{g}_4) = (6^{\frac{2 \log 2 - 1}{4 \log 2 - 3}}, \frac{1}{2}) \approx (-10.19196, 0.5)$ which precisely matches with the numerical value in figure 10! We also found agreement with the numerical position of the kink in $d = 6$ and $d = 10$. As $d \rightarrow \infty$, the kink converges to $(-6, -\frac{1}{2})$.

The fact that the amplitude $\mathcal{M}_{stu\text{-pole}}$ realizes a negative coefficient for *stu* (i.e. $g_3 < 0$) suggests that it would be impossible to prove irreversibility of six-dimensional renormalization group flow using only crossing and positivity of $2 \rightarrow 2$ scattering [26]. This contrasts the successful four-dimensional case, in which the *a*-theorem was related to positivity of g_2 [33].

The convex hull of the space of the two amplitudes $\mathcal{M}_{\text{spin-0}}$ and $\mathcal{M}_{stu\text{-pole}}$, for different values of the mass m , generates the region with parabolic boundaries shown in figure 12. The parabolic “underbellies” are simply the dimensionally-rescaled theories at the kinks. This region is analytically “ruled in”.

One notices that a sliver obtained via numerics is not in the span of these simple analytic models. Since the upper-left arc of this sliver cannot be generated by the convex hull of (possibly rescaled) discrete theories, we attribute it to a continuous one-parameter

family of “extremal theories” which terminate at the kink. It would be interesting to find an analytic expressions for this family.

One may wonder if this family of extremal theories terminates at a second kink; we did not locate any stable candidate in the numerics. This suggests that the family instead disappears inside its own convex hull. Naturally this would occur at the point where its slope becomes tangent to the parabola of dimensionally-rescaled theories originating from that point, $\frac{dg_3}{g_4} = \frac{g_3}{2g_4}$. In this scenario, only the second derivative of the boundary shape would be discontinuous, explaining the difficulty in locating it numerically.

We find it remarkable that two simple analytic models almost span the entire region obtained via numerics: to a good approximation (up to the missing sliver), *a scalar EFT is compatible with causality and unitarity if it is a positive linear combination of the models in eq. (5.2).*

6 Concluding remarks

In this work we showed, in the case of a scalar field theory, that the space of low effective field theories is sharply constrained by positivity of the S-matrix. Dimensional analysis teaches us to expect a low-energy coupling of mass dimension n to be suppressed by a factor of $1/M^n$, where M is the mass of new heavy states. Our main finding is simple: *in any causal and unitary theory, dimensional analysis scaling is a theorem.*

More precisely, for a theory of a single identical real scalar, we showed that dimensionless ratios of the form $g_k/g_2M^\#$, where g_2 is the coefficient of $\frac{1}{2}(\partial\phi)^4$, are bounded above and below by finite constants of order unity. Our method does not assume that physics above the scale M is weakly coupled, only that it is consistent with unitarity and causality. The technical assumption is the convergence of double-subtracted dispersion relations in $2 \rightarrow 2$ scattering. We then separately bounded g_2 itself in eq. (3.35).

To our (possibly incomplete) knowledge, this is the first time that two-sided bounds are obtained for interactions that vanish in the forward limit, such as the six-derivative “ stu ” contact interaction g_3 (see eqs. (2.3) and (4.2)). The key ingredient was to use “null constraints” (for example eq. (2.23)): integrals over the high-energy spectral density which must vanish by crossing symmetry, and which limit the contribution of higher-spin particles. A systematic procedure to extract optimal bounds was presented. The precise form of the null constraints are affected by low-energy self-interactions (i.e. loops within the low-energy EFT), whose effects would be interesting to investigate. It would be interesting to assess if bounds of this type are closely saturated in specific low-energy processes, for example pion scattering. In this case, a generalization to non-identical scalars might be necessary.

In this paper we ignored gravity. A graviton pole would cause the $B_2(t)$ sum rule to diverge in the forward limit, invalidating conclusions from a Taylor series around $t = 0$. However the $X_2(t)$ null constraints in eq. (A.3) should remain valid for $t < 0$ and their implications are worth investigating. Intuitively, one may anticipate that the graviton pole will somewhat weaken the bounds on scalar scattering [1, 34, 35] (see also [36]). However, following the general principle that self-consistent spinning S-matrices are harder to come by, interactions involving gravitons are likely to be sharply constrained by similar methods, possibly extending [8, 37, 38] or addressing conjectures of [39].

For non-gravitational scalar scattering, we found that crossing symmetry does not mix heavy spinning and spinless states. The allowed $2 \rightarrow 2$ low-energy S-matrices are positive sums of those two sectors. These are respectively (almost completely) spanned by the two simple analytic models in eq. (5.2), which realize theories at kinks. Perhaps this is a tantalizing hint that the collection of valid EFTs is not so vast after all.

Acknowledgments

We thank Dalimil Mazáč, Leonardo Rastelli, David Simmons-Duffin and Anh-Khoi Trinh for useful conversations and Clément Virally for technical assistance. Work of VVD is supported by the National Science and Engineering Council of Canada. Work of SCH is supported by the National Science and Engineering Council of Canada, the Canada Research Chair program, the Fonds de Recherche du Québec — Nature et Technologies, and the Simons Collaboration on the Nonperturbative Bootstrap.

A A basis of crossing symmetry constraints

In this section we present a complete basis of vanishing heavy averages, which vanish in any theory with no branch cut below M^2 (i.e. when neglecting low-energy loops). They are organized as functions $X_k(t)$ where k is an even integer representing the number of subtractions.

We first give a pedestrian argument considering the unsubtracted case, X_0 , and then make a general argument exploiting crossing symmetry.

As a warm-up, let us first consider a “superbounded” theory where $\lim_{s \rightarrow \infty} \mathcal{M}(s, t) \rightarrow 0$ for $t < 0$, so that unsubtracted dispersion relations converge, and where the low-energy amplitude is polynomial, corresponding to $g = 0$ in (2.3). In the superbounded case the first sum rule is B_0 from eq. (2.18):

$$B_0 : g_0 + 2g_2t^2 + 2g_4t^4 + \dots = \left\langle \frac{2m^2 + t}{m^2 + t} \mathcal{P} \left(1 + \frac{2t}{m^2} \right) \right\rangle \quad (\text{A.1})$$

where we relabelled the constant term in the amplitude as $g_0 = -\lambda$. Viewed as a function of t this sum rule is marginally useful because it involves infinitely many unknowns on the left-hand-side and one loses control when $-t$ becomes of order M^2 . However, each g_k with $k \geq 2$ can be computed by some other sum rule: for example, the $t \rightarrow 0$ limit of B_2 measures g_2 , $B_4(0)$ measures g_4 , etc. Dividing by an overall t for future convenience, we can package these into an infinite set of null constraints:

$$0 = \left\langle \frac{2m^2 + t}{t(m^2 + t)} \mathcal{P}_J \left(1 + \frac{2t}{m^2} \right) \right\rangle - \sum_{k=0}^{\infty} B_k^{\text{high}}(0) t^{2k-1} \quad (\text{A.2})$$

$$= \left\langle X_0(t; m^2, j) \right\rangle \quad \text{where} \quad X_0(t; m^2, J) \equiv \frac{(2m^2 + t) \mathcal{P}_J \left(1 + \frac{2t}{m^2} \right)}{t(m^2 + t)} - \frac{2m^4 \mathcal{P}_J(1)}{t(m^4 - t^2)}.$$

$X_0(t)$ defines a sensible sum rule for any $0 < -t < M^2$, and Taylor-expanding around $t = 0$ gives an infinite number of averages which vanish by crossing symmetry (if unsubtracted sum rules converge).

Repeating the same manipulations for subtracted sum rules, we find that to cancel the cancel the generic power of t in $B_2(t)$ one must also both the $B_k(0)$ and its first derivative around $t = 0$; there is then a unique solution: $\langle X_2(t; m^2, J) \rangle = 0$ where

$$X_2(t; m^2, J) = \frac{(2m^2 + t) \left(\mathcal{P}_J \left(1 + \frac{2t}{m^2} \right) - \mathcal{P}_J(1) \right)}{m^2 t^2 (m^2 + t)^2} - \frac{4\mathcal{P}'_J(1)}{tm^2(m^4 - t^2)}. \quad (\text{A.3})$$

For any $0 < -t < M^2$, this enjoys the high-energy convergence as a double-subtracted dispersion relation, whence the subscript. Setting $t = 0$ this reproduces the n_4 sum rule in eq. (3.29). Finally, we record explicitly the four-subtracted version obtained with the same method:

$$X_4(t; m^2, J) = \frac{(2m^2 + t) \left(\mathcal{P}_J \left(1 + \frac{2t}{m^2} \right) - \mathcal{P}_J(1) \right)}{m^4 t^3 (m^2 + t)^3} - \frac{2(2m^4 - 3t(m^2 + t))\mathcal{P}'_J(1)}{m^8 t^2 (m^2 - t)(m^2 + t)^2} - \frac{4\mathcal{P}''_J(1)}{m^8 t(m^4 - t^2)}. \quad (\text{A.4})$$

A.1 Derivation using dispersion relations in three channels

Let us now present the general case along with a direct derivation. The idea is to combine dispersion relations in all three channels. Consider the following identity:

$$X_k(t'): 0 = \left[\oint_{C_s} + \oint_{C_t} + \oint_{C_u} \right] \frac{ds \wedge dt}{(2\pi i)^2 st(t-t')} \frac{\mathcal{M}(s, t)}{[st(s+t)]^{k/2}}, \quad (\text{A.5})$$

where each contour is a product of circles: $C_s = \text{Circle}_{s \sim 0} \wedge \text{Circle}_{t \rightarrow \infty}$, $C_t = \text{Circle}_{t \sim 0, t'} \wedge \text{Circle}_{u \rightarrow \infty}$ and $C_u = \text{Circle}_{u \sim 0} \wedge \text{Circle}_{s \rightarrow \infty}$. The first contour implements a fixed- s sum rule, which we then evaluate at $s \rightarrow 0$, and similar for the others, but note that the fixed- t relation is evaluated at both $t = 0$ and $t = t'$. All the integrals vanish (for $k \geq 2$) due to the vanishing at large- t of $\mathcal{M}(s, t)/t^2$. (The contours around 0 are a bit dangerous, since the integrals around say large- t converges only when $s < 0$; this should not be a problem since (ignoring EFT loops) one can interpret the residue as extracting the coefficient of $1/s$ in the Laurent expansion as $s \rightarrow 0^-$.) The trick is then to deform the arcs at infinity to pick high-energy cuts and low-energy poles, as in figure 3 of the main text.

What is nice with the above contour is that all double-residues cancel because of antisymmetry of the contours. For the double residue at $(s, t) = (0, t')$ the cancellation simply follows from mismatching orientations, since the contours C_s and C_t compute the same double-residue in opposite orders:

$$0 = \text{Circle}_{s \sim 0} \wedge \text{Circle}_{t \sim t'} + \text{Circle}_{t \sim t'} \wedge \text{Circle}_{s \sim 0}. \quad (\text{A.6})$$

Alternatively, in practice, one may think of the integral $\frac{1}{2\pi i} \int_{\text{Circle}_{s \sim 0}}$ as just a residue, defined by plucking $\frac{ds}{s} \wedge$ from the left of the differential form it multiplies, and the integral vanishes because the thus-defined residue operation is antisymmetric. The nested poles at $(s, t) = (0, 0)$ require more care (one has to perform a blow-up) since all three contours contribute, but one still finds a perfect cancellation, as one may verify explicitly:

$$\left(\text{Res}_{t=0} \text{Res}_{s=0} + \text{Res}_{s=0} \text{Res}_{t=0} + \text{Res}_{t=0} \text{Res}_{s=-t} \right) \frac{\text{Poly}(s, t)}{(st(s+t))^n} ds \wedge dt = 0 \quad (\text{A.7})$$

for any polynomial numerator. This is curiously reminiscent of the Jacobi identity.

The upshot is that the integral (A.5) is orthogonal to *any* tree-level low-energy EFT amplitude (whether or not it is s - t - u symmetric), it gives purely a constraint on high-energy cuts. Similar constraints would follow for any choice of (rational) denominator. Each term in eq. (A.5) contributes cuts in two channels, for example the C_s term contributes

$$\left\langle \text{Res}_{s=0} \left(\frac{1}{m^2(t' - m^2)} + \frac{1}{(m^2 + s)(m^2 + s + t')} \right) \frac{m^2 \mathcal{P}_J \left(1 + \frac{2s}{m^2} \right)}{s(sm^2(m^2 + s))^{k/2}} \right\rangle \quad (\text{A.8})$$

where the two terms in the parenthesis come from the t and u -channel heavy cuts, respectively. If we did not assume s - t - u symmetry, we would get a relation between the three spectral densities. Here we record only the simplified result assuming that all the spectral densities are the same: $0 = \langle X_k(t; m^2, J) \rangle$ where

$$X_k(t; m^2, J) = \frac{2m^2 + t}{t(m^2 + t)} \frac{\mathcal{P}_J \left(1 + \frac{2t}{m^2} \right)}{(tm^2(m^2 + t))^{k/2}} - \text{Res}_{s=0} \left[\frac{(2m^2 + s)(m^2 - s)(m^2 + 2s)}{s(t - s)(m^2 + s)(m^2 - t)(m^2 + s + t)} \frac{\mathcal{P}_J \left(1 + \frac{2s}{m^2} \right)}{(sm^2(m^2 + s))^{k/2}} \right]. \quad (\text{A.9})$$

This is the main result of this appendix. For any even $k \geq 2$ and $-M^2 < t < 0$ we expect these to be physical sum rules (i.e. convergent in theories where double-subtracted dispersion relations converge) orthogonal to any tree-level low-energy amplitude. For $k = 0, 2, 4$ this reproduces the formulas quoted in eqs. (A.2)–(A.4).

Expanding at small t , the $X_k(t)$ sum rules admit regular Taylor series, which reproduce precisely the sum rules recorded in eq. (3.29). Namely, the coefficient of t^n in X_k has degree $\frac{3k}{2} + n + 1$ in $1/m^2$, so the first case is $X_2(0) \sim g_4$. The first time one gets two X sum rules is at weight 7, where $X_2'''(0)$ and $X_4(0)$ span g_7 and g_7' . The number of X sum rules per degree increases every 3 degree because of the $1/(stu)^{k/2}$ factor in eq. (A.5). The number of X sum rules thus agrees precisely with the counting below eq. (3.29). We conclude that the $X_k(t)$ sum rules are a complete basis of sum rules orthogonal to tree-level EFTs! As mentioned below eq. (3.30), when EFT loops are included these sum rules may average to nonzero but computable quantities.

B Bounds on operators up to order s^{10}

In table 3 we record numerical bounds on various EFT coefficients in four spacetime dimensions.

EFT coefficient	Lower bound	Upper bound
\tilde{g}_3	-10.346	3
\tilde{g}_4	0	0.5
\tilde{g}_5	-4.096	2.5
\tilde{g}_6	0	0.25
\tilde{g}'_6	-12.83	3
\tilde{g}_7	-1.548	1.75
\tilde{g}_8	0	0.125
\tilde{g}'_8	-10.03	4
\tilde{g}_9	-0.524	1.125
\tilde{g}'_9	-13.60	3
\tilde{g}_{10}	0	0.0625
\tilde{g}'_{10}	-6.32	3.75

Table 3. Bounds on coefficients $\tilde{g}_k^{(p)} = g_k^{(p)} M^{2k-4} / g_2$ for $d = 4$ spacetime dimension, where $g_k^{(p)}$ refers to the coefficient of $(s^2 + t^2 + u^2)^{\frac{k-3(2p+\delta_{k,\text{odd}})}{2}} (stu)^{2p+\delta_{k,\text{odd}}}$, which has degree k in Mandelstam invariants and contains $2p$ powers of stu more than the minimum at that degree. The upper bounds are all simple rational numbers realized by the $\mathcal{M}_{\text{spin-0}}$ model. The values (except for \tilde{g}_3) were calculated at order $n = 10$, which corresponds to the number of null constraints of $\dim N = 12$.

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

References

- [1] A. Adams, N. Arkani-Hamed, S. Dubovsky, A. Nicolis and R. Rattazzi, *Causality, analyticity and an IR obstruction to UV completion*, *JHEP* **10** (2006) 014 [[hep-th/0602178](#)] [[INSPIRE](#)].
- [2] A. Martin, *Scattering theory: unitarity, analyticity and crossing*, *Lecture Notes in Physics volume 3*, Springer (1969) [[INSPIRE](#)].
- [3] S.M. Roy, *Exact integral equation for pion pion scattering involving only physical region partial waves*, *Phys. Lett. B* **36** (1971) 353 [[INSPIRE](#)].
- [4] G. Colangelo, J. Gasser and H. Leutwyler, $\pi\pi$ scattering, *Nucl. Phys. B* **603** (2001) 125 [[hep-ph/0103088](#)] [[INSPIRE](#)].
- [5] I. Caprini, G. Colangelo, J. Gasser and H. Leutwyler, *On the precision of the theoretical predictions for $\pi\pi$ scattering*, *Phys. Rev. D* **68** (2003) 074006 [[hep-ph/0306122](#)] [[INSPIRE](#)].
- [6] T.N. Pham and T.N. Truong, *Evaluation of the derivative quartic terms of the meson chiral Lagrangian from forward dispersion relation*, *Phys. Rev. D* **31** (1985) 3027 [[INSPIRE](#)].
- [7] B. Ananthanarayan, D. Toublan and G. Wanders, *Consistency of the chiral pion pion scattering amplitudes with axiomatic constraints*, *Phys. Rev. D* **51** (1995) 1093 [[hep-ph/9410302](#)] [[INSPIRE](#)].
- [8] X.O. Camanho, J.D. Edelstein, J. Maldacena and A. Zhiboedov, *Causality constraints on corrections to the graviton three-point coupling*, *JHEP* **02** (2016) 020 [[arXiv:1407.5597](#)] [[INSPIRE](#)].
- [9] N. Afkhami-Jeddi, T. Hartman, S. Kundu and A. Tajdini, *Einstein gravity 3-point functions from conformal field theory*, *JHEP* **12** (2017) 049 [[arXiv:1610.09378](#)] [[INSPIRE](#)].
- [10] C. Cheung and G.N. Remmen, *Positive signs in massive gravity*, *JHEP* **04** (2016) 002 [[arXiv:1601.04068](#)] [[INSPIRE](#)].
- [11] N. Arkani-Hamed and Y.-T. Huang, *Positive geometry of effective field theory*, lectures given at the *Cern Winter School on supergravity, strings and gauge theory*, February 3–7, CERN, Switzerland (2020).
- [12] N. Arkani-Hamed and Y.-T. Huang, *New positivity bounds from the EFT hedron*, talk at the *24th Rencontres Itzykson of the IPHT of CEA-Saclay*, June 5–7, IPhT CEA-Saclay, France (2019).
- [13] C. de Rham, S. Melville, A.J. Tolley and S.-Y. Zhou, *Positivity bounds for scalar field theories*, *Phys. Rev. D* **96** (2017) 081702 [[arXiv:1702.06134](#)] [[INSPIRE](#)].
- [14] B. Bellazzini, J. Elias Miró, R. Rattazzi, M. Riembau and F. Riva, *Positive moments for scattering amplitudes*, [arXiv:2011.00037](#) [[INSPIRE](#)].
- [15] A.J. Tolley, Z.-Y. Wang and S.-Y. Zhou, *New positivity bounds from full crossing symmetry*, [arXiv:2011.02400](#) [[INSPIRE](#)].
- [16] I. Brivio and M. Trott, *The standard model as an effective field theory*, *Phys. Rept.* **793** (2019) 1 [[arXiv:1706.08945](#)] [[INSPIRE](#)].

- [17] M. Correia, A. Sever and A. Zhiboedov, *An analytical toolkit for the S-matrix bootstrap*, [arXiv:2006.08221](#) [[INSPIRE](#)].
- [18] M. Gell-Mann, M.L. Goldberger and W.E. Thirring, *Use of causality conditions in quantum theory*, *Phys. Rev.* **95** (1954) 1612 [[INSPIRE](#)].
- [19] J. Bros, H. Epstein and V. Glaser, *A proof of the crossing property for two-particle amplitudes in general quantum field theory*, *Commun. Math. Phys.* **1** (1965) 240 [[INSPIRE](#)].
- [20] S. Caron-Huot, *Lorentzian and analytic bootstrap lecture 3*, talk given at *2020 Bootstrap School*, June 8–12, Harvard University, U.K. (2020).
- [21] Y.S. Jin and A. Martin, *Number of subtractions in fixed-transfer dispersion relations*, *Phys. Rev.* **135** (1964) B1375 [[INSPIRE](#)].
- [22] A. Martin, *Extension of the axiomatic analyticity domain of scattering amplitudes by unitarity. 1.*, *Nuovo Cim. A* **42** (1965) 930 [[INSPIRE](#)].
- [23] R.J. Eden, P.V. Landshoff, D.I. Olive and J.C. Polkinghorne, *The analytic S-matrix*, Cambridge University Press, Cambridge U.K. (1966).
- [24] S. Caron-Huot, D. Mazac, L. Rastelli and D. Simmons-Duffin, *Dispersive CFT sum rules*, [arXiv:2008.04931](#) [[INSPIRE](#)].
- [25] J. Penedones, J.A. Silva and A. Zhiboedov, *Nonperturbative Mellin amplitudes: existence, properties, applications*, *JHEP* **08** (2020) 031 [[arXiv:1912.11100](#)] [[INSPIRE](#)].
- [26] H. Elvang, D.Z. Freedman, L.-Y. Hung, M. Kiermaier, R.C. Myers and S. Theisen, *On renormalization group flows and the a-theorem in 6d*, *JHEP* **10** (2012) 011 [[arXiv:1205.3994](#)] [[INSPIRE](#)].
- [27] D. Simmons-Duffin, *A semidefinite program solver for the conformal bootstrap*, *JHEP* **06** (2015) 174 [[arXiv:1502.02033](#)] [[INSPIRE](#)].
- [28] S.M. Chester et al., *Carving out OPE space and precise O(2) model critical exponents*, *JHEP* **06** (2020) 142 [[arXiv:1912.03324](#)] [[INSPIRE](#)].
- [29] L. Córdova, Y. He, M. Kruczenski and P. Vieira, *The O(N) S-matrix monolith*, *JHEP* **04** (2020) 142 [[arXiv:1909.06495](#)] [[INSPIRE](#)].
- [30] M.F. Paulos, J. Penedones, J. Toledo, B.C. van Rees and P. Vieira, *The S-matrix bootstrap. Part III. Higher dimensional amplitudes*, *JHEP* **12** (2019) 040 [[arXiv:1708.06765](#)] [[INSPIRE](#)].
- [31] A.L. Guerrieri, J. Penedones and P. Vieira, *Bootstrapping QCD using pion scattering amplitudes*, *Phys. Rev. Lett.* **122** (2019) 241604 [[arXiv:1810.12849](#)] [[INSPIRE](#)].
- [32] A.L. Guerrieri, A. Homrich and P. Vieira, *Dual S-matrix bootstrap. Part I. 2D theory*, *JHEP* **11** (2020) 084 [[arXiv:2008.02770](#)] [[INSPIRE](#)].
- [33] Z. Komargodski and A. Schwimmer, *On renormalization group flows in four dimensions*, *JHEP* **12** (2011) 099 [[arXiv:1107.3987](#)] [[INSPIRE](#)].
- [34] J. Tokuda, K. Aoki and S. Hirano, *Gravitational positivity bounds*, *JHEP* **11** (2020) 054 [[arXiv:2007.15009](#)] [[INSPIRE](#)].
- [35] L. Alberte, C. de Rham, S. Jaitly and A.J. Tolley, *Positivity bounds and the massless spin-2 pole*, *Phys. Rev. D* **102** (2020) 125023 [[arXiv:2007.12667](#)] [[INSPIRE](#)].

- [36] B. Bellazzini, M. Lewandowski and J. Serra, *Positivity of amplitudes, weak gravity conjecture, and modified gravity*, *Phys. Rev. Lett.* **123** (2019) 251103 [[arXiv:1902.03250](#)] [[INSPIRE](#)].
- [37] Y. Hamada, T. Noumi and G. Shiu, *Weak gravity conjecture from unitarity and causality*, *Phys. Rev. Lett.* **123** (2019) 051601 [[arXiv:1810.03637](#)] [[INSPIRE](#)].
- [38] C. de Rham, S. Melville, A.J. Tolley and S.-Y. Zhou, *Positivity bounds for massive spin-1 and spin-2 Fields*, *JHEP* **03** (2019) 182 [[arXiv:1804.10624](#)] [[INSPIRE](#)].
- [39] S.D. Chowdhury, A. Gadde, T. Gopalka, I. Halder, L. Janagal and S. Minwalla, *Classifying and constraining local four photon and four graviton S-matrices*, *JHEP* **02** (2020) 114 [[arXiv:1910.14392](#)] [[INSPIRE](#)].