

## Sharp boundaries for the swampland

---

Simon Caron-Huot,<sup>a</sup> Dalimil Mazáč,<sup>b</sup> Leonardo Rastelli<sup>c</sup> and David Simmons-Duffin<sup>d</sup>

<sup>a</sup>*Department of Physics, McGill University,  
3600 Rue University, Montréal, QC Canada*

<sup>b</sup>*Institute for Advanced Study,  
Princeton, NJ 08540, U.S.A.*

<sup>c</sup>*C. N. Yang Institute for Theoretical Physics, Stony Brook University,  
100 Nicolls Rd, Stony Brook, NY 11794, U.S.A.*

<sup>d</sup>*Walter Burke Institute for Theoretical Physics, Caltech,  
1200 E California Blvd, Pasadena, CA 91125, U.S.A.*

*E-mail:* [schuot@physics.mcgill.ca](mailto:schuot@physics.mcgill.ca), [dmazac@ias.edu](mailto:dmazac@ias.edu),  
[leonardo.rastelli@stonybrook.edu](mailto:leonardo.rastelli@stonybrook.edu), [dsd@caltech.edu](mailto:dsd@caltech.edu)

**ABSTRACT:** We reconsider the problem of bounding higher derivative couplings in consistent weakly coupled gravitational theories, starting from general assumptions about analyticity and Regge growth of the S-matrix. Higher derivative couplings are expected to be of order one in the units of the UV cutoff. Our approach justifies this expectation and allows to prove precise bounds on the order one coefficients. Our main tool are dispersive sum rules for the S-matrix. We overcome the difficulties presented by the graviton pole by measuring couplings at small impact parameter, rather than in the forward limit. We illustrate the method in theories containing a massless scalar coupled to gravity, and in theories with maximal supersymmetry.

**KEYWORDS:** Effective Field Theories, Scattering Amplitudes, Classical Theories of Gravity

**ARXIV EPRINT:** [2102.08951](https://arxiv.org/abs/2102.08951)

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Dispersive sum rules</b>	<b>4</b>
2.1	Assumptions and first consequences	4
2.2	Review: bounds from the forward limit	9
<b>3</b>	<b>Bounds with gravity</b>	<b>11</b>
3.1	General idea	11
3.2	Numerical implementation	14
3.3	Bounds on $g_3$ and $g_4$ without gravity	15
3.4	Bounds on $g_2$ and $g_3$ with gravity	15
3.5	Bounds on higher contact coefficients with gravity	19
3.6	Some solutions to the constraints from string theory	20
3.7	Comments on $D = 4$ and infrared divergences	22
3.8	Maximal supergravity: bounding graviton scattering	22
<b>4</b>	<b>Conclusions</b>	<b>24</b>
<b>A</b>	<b>Details on numerics</b>	<b>26</b>
A.1	Impact parameter space inequalities	27
A.2	Outer approximation/adaptive refinement	28
A.3	Choices of parameters and numerical results	29
<b>B</b>	<b>Bounds using an extended range of <math>u</math></b>	<b>30</b>

---

## 1 Introduction

The gravitational coupling grows at short distances, giving rise to the fundamental question: how is gravity UV completed? In the absence of experiments directly probing the Planck scale, a more relevant question is: what is the space of low-energy effective field theories (EFTs) that contain gravitons and admit a UV completion? There are compelling reasons to believe that the rules of the game are more stringent in the presence of dynamical gravity. A variety of conceptual arguments, reinforced by surveys of the string theory landscape, have led to intriguing “swampland” conjectures [1–4] (such as the weak-gravity conjecture [5]), which are proposed criteria for an EFT to be embeddable in a consistent theory of quantum gravity.

Even *without* dynamical gravity, it has long been recognized that “not anything goes” in effective field theory. In a local QFT in flat space, the constraints of unitarity and

causality imply inequalities for the low-energy effective couplings [6, 7]. A series of recent papers [8–12] initiated a systematic analysis of the inequalities that follow from consideration of  $2 \rightarrow 2$  scattering processes. The analysis begins with canonical assumptions about the S-matrix, such as analyticity, crossing and, crucially, Regge boundedness. For a fixed momentum transfer  $u < 0$ , the amplitude  $\mathcal{M}(s, u)$  is assumed to grow slower than  $|s|^2$  at large  $|s|$  in the upper-half complex plane.<sup>1</sup> This is sufficient to derive twice-subtracted dispersion relations [15, 16]. These can be in turn used to express the parameters of the EFT (valid at energies  $E \ll M$ , where  $M$  is the UV cut-off) in terms of the UV data that enter for  $E \geq M$ . One is agnostic about the UV physics, except that it contributes to the partial wave expansion of  $\mathcal{M}(s, u)$  with a definite sign, thanks to unitarity. These positive sum rules constrain the allowed low-energy couplings. Notably, one can establish two-sided bounds [10, 11] that put dimensional analysis on a firm footing. A low-energy parameter of mass dimension  $-\alpha$  must scale like  $\sim 1/M^\alpha$ , possibly further suppressed by a small coupling but never larger, and with a rigorous estimate of the  $O(1)$  numerical factor.

It is very natural to ask whether the same approach may work in EFTs containing gravitons, i.e. massless spin-two particles. Can one develop a systematic and quantitative approach to the swampland, relying only on general properties of the S-matrix in asymptotically flat spacetime? In the presence of dynamical gravity, the analyticity and boundedness properties of the S-matrix are admittedly more speculative, and we will treat them as postulates. We will in particular assume that the same Regge bound  $\lim_{s \rightarrow \infty} \mathcal{M}(s, u)/s^2 = 0$  (for fixed  $u < 0$ ) holds also in this case. Note that this is a little stronger than the  $O(s^2)$  bound assumed in the “Classical Regge Growth” conjecture [17–19], which is believed to hold for any consistent *tree-level* S-matrix; we need to require that at the nonperturbative level the Regge growth is strictly smaller than  $s^2$ . The best heuristic justification for such a behavior comes from physical arguments that the scattering amplitude in impact parameter space should be analytic and bounded, as a consequence of unitarity and causality. The current best technical justification comes from viewing the S-matrix in asymptotic  $D$ -dimensional Minkowski space as the flat-space limit of a scattering process in asymptotic  $\text{AdS}_D$  space, according to the prescription of [20–22]. On the AdS side, we can appeal to the rigorous non-perturbative bound on the Regge behavior of the dual CFT correlator [23], which when translated to flat-space variables corresponds to an  $O(s)$  bound.<sup>2</sup>

As a proof of principle, we focus on the simplest example:  $2 \rightarrow 2$  scattering of identical massless scalars coupled to gravity. We assume that the low-energy EFT is weakly coupled, but make no assumptions about the UV physics beyond analyticity and unitarity. An important example to which our analysis applies is string theory with fixed but small

---

<sup>1</sup>This behavior is implied by the stronger Froissart-Martin bound [13, 14],  $\mathcal{M}(s, u) \lesssim s \log^{D-2} s$ , which (at least in a gapped theory) is a consequence of unitarity if one assumes polynomial boundedness and analyticity, which in turn can be established from first principles in a local QFT. See appendix A.2 of [8] for a nice recent discussion.

<sup>2</sup>At leading order for large  $N$ , the bound on the Regge behavior of the CFT correlator is known as the chaos bound [24]. The chaos bound translates precisely to the  $O(s^2)$  bound of the Classical Regge Growth conjecture, as has been carefully argued in the recent paper [19]. It would be very interesting to see whether the arguments of [19] can be extended to the nonperturbative regime and give a rigorous justification for our assumption that the Regge growth is *strictly better* than  $O(s^2)$ .

coupling  $g_s \ll 1$ . The cutoff  $M$  is the string scale, i.e. the mass of the first exchanged massive string mode. The theory is weakly coupled at and below the cutoff, but it eventually becomes strongly coupled as we approach the Planck scale  $M/(g_s)^{\frac{2}{D-2}}$ .

It is straightforward to derive dispersive sum rules in this setup, following the blueprint reviewed in [11]. Similar positive sum rules incorporating gravity were written down before, see e.g. [8, 25–29] (or [30] for the similar case of U(1) gauge theory). There is however a notorious obstacle in deriving bounds that involve the Newton constant  $G$ : the graviton propagator diverges in the limit of forward scattering ( $u \rightarrow 0$  in our conventions), seemingly invalidating the application of the positivity constraints. We circumvent this problem by measuring couplings at small impact parameter  $b \ll 1/M$ , while keeping the momentum transfer  $u \sim -M^2$ . The physical picture of a scattering experiment at small impact parameter is the same as in [17], but now in the more systematic framework of the S-matrix bootstrap where we look for numerical bounds (not only parametric). Apart from curing the graviton divergence, this approach has the virtue of making dimensional analysis transparent. Finally, while in this paper we treat the EFT at tree level, it is in principle straightforward to include EFT loop corrections, which will introduce low-energy cuts extending all the way to  $u = 0$ . Our method remains valid even if the amplitude is non-analytic at  $u = 0$  and it is thus ideally suited to handle loops.

As in previous work, the positivity constraints lead to a linear programming problem that can be implemented numerically to carve out the space of EFT couplings. See figure 4 and figure 7 for some sample exclusion plots for ratios of the scalar higher derivative EFT coefficients  $g_k$  over the Newton constant  $G$ , in units of the UV cutoff. We restrict our analysis to spacetime dimension  $D > 4$ , because for  $D = 4$  the integral transform to impact parameter space suffers from an IR divergence arising from the massless graviton pole. We believe that it should be possible to extend our approach to  $D = 4$ , working with suitable IR-safe observables.

We also consider bounds on graviton scattering in the presence of maximal supersymmetry. Thanks to supersymmetry,  $2 \rightarrow 2$  graviton scattering can be expressed in terms of an auxiliary fully crossing symmetric scalar amplitude with improved Regge behavior. We derive a numerical upper bound on the coefficient of the leading  $R^4$  curvature correction, and check that it is obeyed by type II string theory.

In an upcoming paper [31], we will study the analogous problem of bounding higher derivative couplings in a gravitational theory in AdS space. A prominent physical question [32] is to show that a large  $N$  CFT with a large gap admits a local bulk dual, i.e. such that higher-derivative corrections are suppressed by inverse powers of the gap. The formalism of dispersive CFT sum rules [33–37]<sup>3</sup> allows for an almost direct uplift of our flat space results to AdS. While the logic is very similar, the details of the AdS story are technically more involved. The flat space analysis presented here will serve as an indispensable warmup. One distinct advantage of the AdS setup is that all the requisite analyticity and boundedness properties are rigorous consequences of the CFT axioms.

The rest of the paper is organized as follows. In section 2, we state our assumptions and review the approach of [11] to bounding EFT coefficients, which relies on the Taylor

<sup>3</sup>See also [38–42] for more work concerning dispersive sum rules in CFTs.

expansion of the dispersive sum rules around the forward limit. In section 3, we develop our new strategy, based on localization at small impact parameter, and apply it to simple examples of gravitational EFTs. We conclude in section 4. In appendix A, we present some technical details of the numerical implementation of the linear program. Finally, in appendix B we consider bounds coming from extending the range of the momentum transfer beyond  $|u| = M^2$ .

## 2 Dispersive sum rules

In this section we state our assumptions and set the stage for the main argument in section 3. We will also review how to bound couplings in non-gravitational theories using dispersion relations expanded around the forward limit. We will closely follow the presentation of [11].

### 2.1 Assumptions and first consequences

**IR Effective Field Theory.** Consider a real scalar coupled to gravity and to unknown heavy states of mass greater than some energy scale  $M$ , in asymptotically flat  $D$ -dimensional spacetime. As our interest lies in the effects of the heavy scale  $M$ , we take the scalar to be massless, though our techniques are general and can be applied equally well to massive scalars. The  $2 \rightarrow 2$  massless scalar scattering amplitude takes the form

$$\begin{aligned} \mathcal{M}_{\text{low}}(s, u) = & 8\pi G \left[ \frac{st}{u} + \frac{su}{t} + \frac{tu}{s} \right] - \lambda_3^2 \left[ \frac{1}{s} + \frac{1}{t} + \frac{1}{u} \right] - \lambda_4 \\ & + g_2(s^2 + t^2 + u^2) + g_3(stu) + g_4(s^2 + t^2 + u^2)^2 + \dots, \end{aligned} \tag{2.1}$$

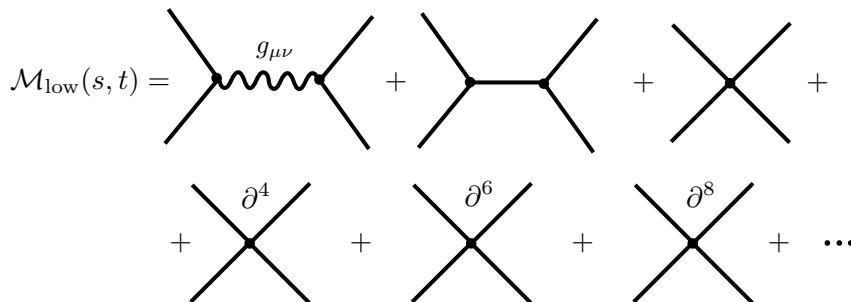
where

$$s = -(p_1 + p_2)^2, \quad t = -(p_1 - p_4)^2, \quad u = -(p_1 - p_3)^2 \tag{2.2}$$

are Mandelstam invariants, satisfying  $s + t + u = 0$ .  $G$  is Newton's constant and the first term represents the exchange of gravitons in the three channels. Apart from coupling to gravity, we have also allowed for the possibility of scalar self-interactions. A schematic representation of the scattering amplitude is shown in figure 1.

We assume that the low-energy physics is weakly coupled. To make this precise, we can assume that there is a family of S-matrices parametrized by a coupling  $\epsilon > 0$  such that all the interactions, such as  $G$ ,  $\lambda_3^2$ ,  $g_2$  etc. in (2.1), are  $O(\epsilon)$  as  $\epsilon \rightarrow 0$ . Eq. (2.1) is then the low-energy form of the scattering amplitude at the leading order in  $\epsilon$ , i.e. at tree level. For example, this is the situation in string theory, the role of  $\epsilon$  being played by the string coupling. Low-energy loop corrections are  $O(\epsilon^2)$  effects and could be included in principle but will be beside our main concern. In the rest of this paper, we will work at the leading order in  $\epsilon$  and therefore these corrections will be absent.

Under this weak-coupling assumption, the EFT coefficients are symmetrical polynomials in the Mandelstam variables, corresponding to crossing-symmetric contact diagrams. We have chosen the natural basis of monomials  $\mathcal{M}_{a,b}(s, t) = A^a B^b$ , where  $A = s^2 + t^2 + u^2$ ,



**Figure 1.** A schematic representation of the low-energy amplitude (2.1). The external scalar particles interact through graviton exchange, scalar exchange and a series of higher-derivative contact interactions.

$B = stu$ . The spin of the monomial  $\mathcal{M}_{a,b}(s, t)$  is  $J = 2a + 2b$  and the scaling dimension is  $4a + 6b$ .<sup>4</sup> Thus, there are  $(J + 2)/2$  independent contacts with spin equal to  $J$ , namely

$$A^{\frac{J}{2}}B^0, \quad A^{\frac{J}{2}-1}B^1, \quad \dots, \quad A^0B^{\frac{J}{2}}, \quad (2.3)$$

with scaling dimensions

$$2J, \quad 2J + 2, \quad \dots, \quad 3J. \quad (2.4)$$

Note that there can be multiple contacts with different spins but the same scaling dimension, the lowest example being  $A^3$  and  $B^2$ . The subscript  $k$  in the EFT coefficients  $g_k$  indicates (half) their scaling dimension, which specifies the contacts diagrams unambiguously for  $k \leq 5$ ; for higher value of  $k$  we introduce additional labels (e.g.,  $g_6$  and  $g'_6$ ) to distinguish contacts with degenerate scaling dimension.

**UV unitarity constraints.** Assuming there are no new degrees of freedom below the heavy scale  $M$ , we want to know: what are the implications of high-energy unitarity on the coefficients  $g_k$ ? It is natural to consider ratios of couplings such as

$$\frac{g_n}{g_2} \quad \text{or} \quad \frac{g_n}{G}, \quad (2.5)$$

which stay finite as  $\epsilon \rightarrow 0$ . Naive EFT scaling, or dimensional analysis, suggests that

$$\frac{g_k}{g_2} = \frac{c_k}{M^{2(k-2)}} \quad (2.6)$$

where  $c_k$  are dimensionless coefficients of order one. This is certainly the case if the EFT is obtained by integrating out a single heavy particle of mass  $M$ . However, we will be agnostic about the detailed physics above the heavy scale  $M$ , and only assume that the amplitude is unitarity and causal. In particular, it will *not* be important to assume that the heavy physics is weakly coupled at arbitrarily high energy scales. On the other hand, we will need (2.1) to be valid up to energies of order  $M$ , which does require the physics there to be weakly coupled. For example, in string theory with small but finite  $g_s$ , the appropriate

<sup>4</sup>The spin of a contact diagram is defined as the maximal spin appearing in its partial wave decomposition. We define the scaling dimension as the number of derivatives in the corresponding term in the Lagrangian. This differs from the standard definition by a constant shift.

cutoff  $M$  is the string scale, where the theory is still weakly coupled. (However, the theory does become strongly-coupled at the Planck scale.)

Unitarity is simplest to state in a decomposition of the amplitude in angular momentum partial waves. In the s-channel physical region  $\{s > 0, -s < u < 0\}$ , Mandelstam  $s$  gives the squared center-of mass energy, while the scattering angle is

$$\cos \theta = 1 + \frac{2u}{s}. \tag{2.7}$$

The partial wave decomposition reads

$$\mathcal{M}(s, u) = s^{\frac{4-D}{2}} \sum_{J \text{ even}} n_J^{(D)} c_J(s) \mathcal{P}_J \left( 1 + \frac{2u}{s} \right), \tag{2.8}$$

where  $\mathcal{P}_J(x)$  are proportional to Gegenbauer polynomials (and reduce to Legendre polynomials for  $D = 4$ ),

$$\mathcal{P}_J(x) \equiv {}_2F_1 \left( -J, J + D - 3, \frac{D-2}{2}, \frac{1-x}{2} \right). \tag{2.9}$$

The normalization

$$n_J^{(D)} \equiv \frac{2^D \pi^{\frac{D-2}{2}}}{\Gamma(\frac{D-2}{2})} (J+1)_{D-4} (2J+D-3) \tag{2.10}$$

has been chosen (see e.g. [43, 44]) such that unitarity of the S-matrix,

$$SS^\dagger = 1, \quad S = 1 + i\mathcal{M}, \tag{2.11}$$

translates into  $|1 + ic_J(s)|^2 \leq 1$ . Defining the spectral density  $\rho_J(s) = \text{Im } c_J(s)$ , we can also write

$$\text{Im } \mathcal{M}(s, u) = s^{\frac{4-D}{2}} \sum_{J \text{ even}} n_J^{(D)} \rho_J(s) \mathcal{P}_J \left( 1 + \frac{2u}{s} \right), \tag{2.12}$$

where the unitarity constraint reads

$$0 \leq \rho_J(s) \leq 2, \quad s > 0, \quad J \text{ even}. \tag{2.13}$$

The crucial fact is that  $\text{Im } \mathcal{M}(s, u)$  is a positive sum of Gegenbauer polynomials.

We can now give a precise definition of the scale  $M$ . It is the energy where  $\mathcal{M}(s, u)$  first develops a nonzero imaginary part. In other words, we will assume that  $\rho_J(s)$  vanishes for all  $0 < s < M^2$  and all even  $J$ .<sup>5</sup>

**Dispersive sum rules.** A link between the regimes of high- and low-energy is provided by a dispersion relation. We assume that for fixed real  $u < 0$ ,

- (i) the amplitude  $\mathcal{M}(s, u)$  is analytic in  $s$  in the upper-half plane  $\text{Im } s > 0$ ,<sup>6</sup> and

---

<sup>5</sup>Note that this definition only makes sense at the leading order at weak coupling since EFT loops give rise to an imaginary part for any  $s > 0$ . However, as explained above, this is a subleading effect under our assumptions.

<sup>6</sup>The amplitude is extended to the lower-half complex  $s$  plane by  $\mathcal{M}(s^*, u^*) = \mathcal{M}^*(s, u)$ , so for fixed  $u < 0$  it is analytic in  $s$  away from the real axis.

(ii)  $\mathcal{M}(s, u)$  has spin-2 convergence in the Regge limit, meaning that the following limit vanishes along any line of constant phase,

$$\lim_{|s| \rightarrow \infty} \frac{\mathcal{M}(s, u)}{s^2} = 0 \quad (u < 0). \quad (2.14)$$

For example, in string theory, (2.14) is ensured by Reggeization of the graviton:  $\mathcal{M} = O(s^{2+\alpha' u})$ . We have discussed this crucial assumption in the Introduction.

Together these conditions imply the existence of twice-subtracted dispersion relations. The starting point is

$$\oint_{\infty} \frac{ds'}{2\pi i(s' - s)} \frac{\mathcal{M}(s', u)}{s'(s' + u)} = 0, \quad (2.15)$$

where the contour integral is over a large circle. We picked  $s' = 0$  and  $s' = -u$  as subtraction points — this is a natural choice because it maintains the  $s \leftrightarrow t$  crossing symmetry without introducing an extraneous mass scale into the problem. We now deform the contour as in figure 2. We take  $|s| \ll M^2$ . There are two kinds of contributions: a low-energy circle at  $|s'| \sim M^2$  which encloses the residues at  $s' = s$ ,  $s' = 0$  and  $s' = -u$ ; and the contributions from the s-channel and t-channel high-energy cuts, starting respectively at  $s' = M^2$  and  $s' = -M^2 - u$ . Separating the low- and high-energy contributions,

$$\begin{aligned} & \frac{\mathcal{M}_{\text{low}}(s, u)}{s(s+u)} + \text{Res}_{s'=0} \left[ \left( \frac{1}{s' - s} + \frac{1}{s' + s + u} \right) \frac{\mathcal{M}_{\text{low}}(s', u)}{s'(s' + u)} \right] = \\ & = \int_{M^2}^{\infty} \frac{ds'}{\pi} \left( \frac{1}{s' - s} + \frac{1}{s' + s + u} \right) \text{Im} \left[ \frac{\mathcal{M}_{\text{high}}(s', u)}{s'(s' + u)} \right], \end{aligned} \quad (2.16)$$

where we have used the  $s \leftrightarrow t$  symmetry to combine the contributions of the two low-energy residues at  $s' = 0$  and  $s' = -u$ , and the contributions of the two high-energy cuts. For the terms on the left-hand side we use the low-energy parametrization (2.1), while for the heavy contribution, about whose details we are agnostic, we insert the partial wave decomposition (2.12),

$$\frac{8\pi G}{-u} + 2g_2 - g_3 u + 4g_4(2u^2 + s(s+u)) + \dots = \left\langle \frac{(2m^2 + u) \mathcal{P}_J \left(1 + \frac{2u}{m^2}\right)}{(m^2 + u)(m^2 - s)(m^2 + s + u)} \right\rangle, \quad (2.17)$$

where we defined the heavy averages as

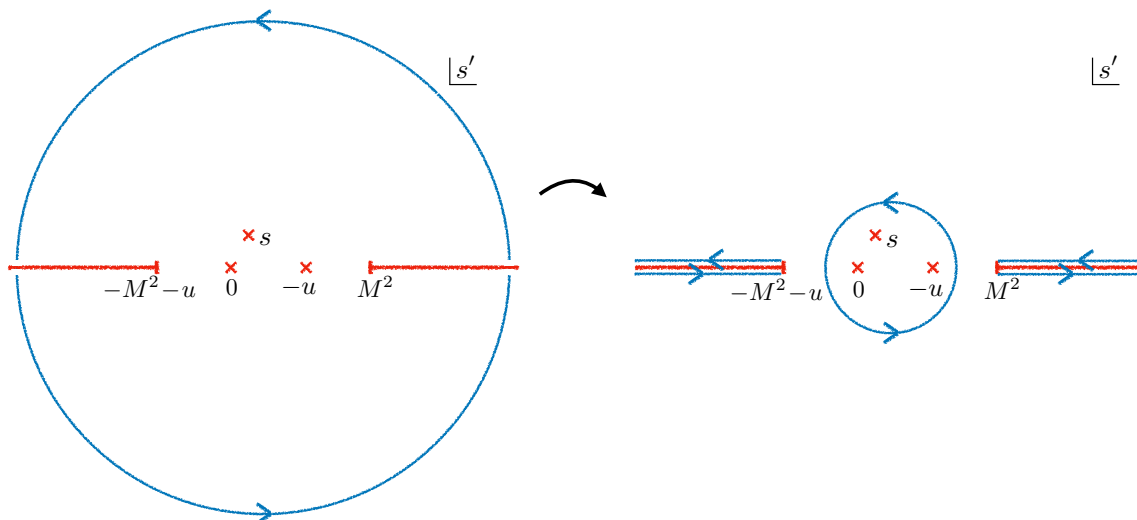
$$\langle (\dots) \rangle \equiv \frac{1}{\pi} \sum_{J \text{ even}} n_J^{(D)} \int_{M^2}^{\infty} \frac{dm^2}{m^2} m^{4-D} \rho_J(m^2) (\dots). \quad (2.18)$$

What will be important for us is that this is a positive measure since  $\rho_J(m^2) \geq 0$  by unitarity. We make no use of the upper bound  $\rho_J(m^2) \leq 2$  in this work. Note that the s- and t- channel poles have cancelled in (2.17), leaving a regular function that can be Taylor-expanded in powers of  $s/M^2$ .

We will organize these sum rules by expanding around small  $s$ . Because our subtraction preserves  $s \leftrightarrow t$  symmetry, it is easy to see that the series proceeds in integer powers of  $s(s+u)$  and that we can expand (2.17) as

$$\sum_{n=1}^{\infty} [s(s+u)]^{n-1} \mathcal{C}_{2n,u} = 0. \quad (2.19)$$





**Figure 2.** Contour deformation leading to the dispersive sum rule (2.16). We start from an integral over a large circle (2.15), which vanishes due to the spin-2 boundedness assumption (2.14). After the contour deformation, we end up with an IR contribution, represented by the small circle on the right, and a UV contribution, represented by an integral over the cuts starting at  $s' = M^2, -M^2 - u$ . In the presence of EFT loops, the IR contribution would also include cuts stretching between  $s' = -M^2 - u$  and  $s' = M^2$ , but these are subleading under our assumptions. Since the theory is assumed weakly coupled at scale  $M$ , the heavy cut is well approximated by a discrete set of poles for  $s'$  not much larger than  $M^2$ . However, this will play no role in our reasoning.

For example, the first sum rule, denoted by  $\mathcal{C}_{2,u}$ , is obtained by taking the  $s \rightarrow 0$  limit in (2.17),

$$\mathcal{C}_{2,u} = \frac{8\pi G}{-u} + 2g_2 - g_3u + 8g_4u^2 + \dots - \left\langle \frac{(2m^2 + u)\mathcal{P}_J\left(1 + \frac{2u}{m^2}\right)}{m^2(m^2 + u)^2} \right\rangle = 0, \quad (2.20)$$

while the next sum rule  $\mathcal{C}_{4,u}$  corresponds to the coefficient of  $s(s + u)$ ,

$$\mathcal{C}_{4,u} = 4g_4 + \dots - \left\langle \frac{(2m^2 + u)\mathcal{P}_J\left(1 + \frac{2u}{m^2}\right)}{m^4(m^2 + u)^3} \right\rangle = 0. \quad (2.21)$$

We can also obtain the  $\mathcal{C}_{k,u}$  sum rules more directly from the spin- $k$  subtracted version of the dispersion integral,<sup>7</sup>

$$\mathcal{C}_{k,u} \equiv \oint_{\infty} \frac{ds'}{2\pi i} \frac{1}{s'} \frac{\mathcal{M}(s', u)}{[s'(s' + u)]^{k/2}} = 0, \quad u < 0, \quad k = 2, 4, \dots \quad (2.22)$$

By the same contour deformation argument as above, we find the explicit expressions

$$\mathcal{C}_{k,u} = \text{Res}_{s'=0} \left[ \frac{2s' + u}{s'(s' + u)} \frac{\mathcal{M}_{\text{low}}(s', u)}{[s'(s' + u)]^{k/2}} \right] - \left\langle \frac{2m^2 + u}{m^2 + u} \frac{\mathcal{P}_J\left(1 + \frac{2u}{m^2}\right)}{[m^2(m^2 + u)]^{k/2}} \right\rangle. \quad (2.23)$$

<sup>7</sup>Compare with the  $s \rightarrow 0$  limit of (2.15).

The  $\mathcal{C}_{k,u}$  sum rule only receives contributions from the contacts

$$(s^2 + t^2 + u^2)^a (stu)^b \quad \text{with spin } J = 2a + 2b \geq k \text{ and } b \leq k/2. \quad (2.24)$$

The condition  $J \geq k$  is seen by closing the contour around infinity, as in (2.22), while the condition  $b \leq k/2$  is apparent by closing the contour around the origin, as in (2.23). Finally, note that only the  $\mathcal{C}_{2,u}$  sum rule is sensitive to the graviton exchange, through the term  $-8\pi G/u$ .

We now describe two distinct ways to use the  $\mathcal{C}_{k,u}$  sum rules. First, in section 2.2 we review bounds obtained by expanding around the forward limit. Then, in section 3 we introduce a new method: localizing in impact parameter space.

## 2.2 Review: bounds from the forward limit

We now briefly illustrate the strategy described in [10, 11] to derive inequalities for the EFT coefficients by expanding the sum rules (2.23) in the forward limit  $u \rightarrow 0$ . This strategy breaks down in the presence of gravity, because the  $\mathcal{C}_2(u)$  sum rule diverges in the forward limit, and so in this subsection we switch off gravity by setting  $G = 0$ .

Derivatives of  $\mathcal{C}_{k,u}$  with respect to  $u$  at  $u = 0$  compute the couplings  $g_2, g_3, \dots$ . Only couplings of spin two and higher appear in the sum rules. Spin-two couplings  $g_2$  and  $g_3$  only appear in  $\mathcal{C}_{2,u}$  and they are computed respectively by  $\mathcal{C}_{2,0}$  and  $\mathcal{C}'_{2,0}$ ,<sup>8</sup>

$$g_2 = \left\langle \frac{1}{m^4} \right\rangle, \quad g_3 = \left\langle \frac{3 - \frac{4\mathcal{J}^2}{D-2}}{m^6} \right\rangle, \quad (2.25)$$

where  $\mathcal{J}^2 = J(J + D - 3)$  is the quadratic Casimir of the massive little group  $\text{SO}(D - 1)$ . The spin-4 coupling  $g_4$  appears in  $\mathcal{C}_{2,u}$  and  $\mathcal{C}_{4,u}$  and is computed by  $\mathcal{C}''_{2,0}$  and  $\mathcal{C}_{4,0}$

$$g_4 = \left\langle \frac{1 + \frac{\mathcal{J}^2(2\mathcal{J}^2 - 5D + 4)}{2D(D-2)}}{2m^8} \right\rangle = \left\langle \frac{1}{2m^8} \right\rangle. \quad (2.26)$$

The last equality leads to the simplest example of a null constraint on the heavy data

$$\left\langle \frac{\mathcal{J}^2(2\mathcal{J}^2 - 5D + 4)}{m^8} \right\rangle = 0. \quad (2.27)$$

Null constraints arise because the low-energy amplitude is symmetric under  $s \leftrightarrow u$ .

The first sum rule in (2.25) immediately implies  $g_2 > 0$ . The remaining sum rules can be used to derive two-sided bounds on the ratios  $g_n/g_2$ . It is convenient to normalize the measure (2.18) by the  $g_2$  sum rule and define the probability measure  $\tilde{\rho}_J(m)$

$$\begin{aligned} \left\langle (\dots) \right\rangle_{g_2} &= \frac{1}{\pi g_2} \sum_{J \text{ even}} n_J^{(D)} \int_{M^2}^{\infty} \frac{dm^2}{m^2} m^{-D} \rho_J(m^2) (\dots) \\ &\equiv \sum_{J \text{ even}} \int_M^{\infty} dm \tilde{\rho}_J(m) (\dots). \end{aligned} \quad (2.28)$$

<sup>8</sup>Here and below, primes denote derivatives with respect to  $u$ .

The above sum rules become

$$1 = \langle 1 \rangle_{g_2} \quad \frac{g_3}{g_2} = \left\langle \frac{3 - \frac{4\mathcal{J}^2}{D-2}}{m^2} \right\rangle_{g_2} \quad (2.29)$$

and

$$\frac{g_4}{g_2} = \left\langle \frac{1}{2m^4} \right\rangle_{g_2} \quad 0 = \left\langle \frac{\mathcal{J}^2(2\mathcal{J}^2 - 5D + 4)}{m^4} \right\rangle_{g_2} \quad (2.30)$$

Since the measure is supported in  $m \geq M$ , it immediately follows from the second equation in (2.29) that

$$\frac{g_3}{g_2} \leq \frac{3}{M^2}. \quad (2.31)$$

One can derive a lower bound on  $g_3/g_2$  with the help of the null constraint. The  $g_2$ ,  $g_3$  and null sum rules take the form of a vector equation

$$\sum_{J \text{ even}} \int_M^\infty dm \tilde{\rho}_J(m) \begin{pmatrix} 1 \\ \frac{1}{m^2} \left( 3 - \frac{4\mathcal{J}^2}{D-2} \right) \\ \frac{\mathcal{J}^2(2\mathcal{J}^2 - 5D + 4)}{m^4} \end{pmatrix} = \begin{pmatrix} 1 \\ g_3/g_2 \\ 0 \end{pmatrix}. \quad (2.32)$$

This is now a standard linear programming problem. We are asking for what values of  $g_3/g_2$  is the vector on the r.h.s. in the positive cone spanned by the vectors on the l.h.s. This means that the allowed range for  $g_3/g_2$  is the intersection of the convex hull of points

$$\left( \frac{1}{m^2} \left( 3 - \frac{4\mathcal{J}^2}{D-2} \right), \frac{\mathcal{J}^2(2\mathcal{J}^2 - 5D + 4)}{m^4} \right) \in \mathbb{R}^2 \quad (2.33)$$

with the  $x$ -axis. It is clear that the region is finite, and easy to check that the lower bound comes from considering only  $J = 2$  and  $J = 4$  trajectories. In other words, including  $J > 4$  trajectories does not increase the size of the intersection of the convex hull with the  $x$ -axis. The bound thus comes from the intersection of the  $x$ -axis and the line connecting the point  $J = 2$ ,  $m = M$  with the point  $J = 4$ ,  $m = m_4$  and optimizing over  $m_4$ . It is clear that since the bound comes from bounded spin, it needs to have the bulk-point scaling  $M^{-2}$ . In fact, if there is any bound at all, it must have this scaling simply by dimensional analysis. Therefore one can get an analytic, if cumbersome answer:

$$-\frac{\kappa(D)}{M^2} < \frac{g_3}{g_2} \leq \frac{3}{M^2}, \quad (2.34)$$

where

$$\kappa(D) = \sqrt{\frac{(D+3)(319D^3 + 76D^2 - 292D + 32)}{24(D-2)^2(D+1)(D+4)}} + \frac{6(5D-2)}{12(D-2)}. \quad (2.35)$$

The bound can be improved by considering combinations of more functionals, see [11, 45].

We can play the same game to bound  $g_4/g_2$  but in this case including the null constraint buys us nothing since spin does not enter the sum rule for  $g_4/g_2$ . The result is simply

$$0 \leq \frac{g_4}{g_2} \leq \frac{1}{2M^4}. \quad (2.36)$$

### 3 Bounds with gravity

#### 3.1 General idea

The preceding strategy, Taylor expanding around the forward limit, suffers from three drawbacks. First, it does not manifest the expected scaling  $g_3 \sim \frac{g_2}{M^2}$  until the final stage. This is because we are evaluating the IR contribution to dispersive sum rules at small momentum transfer  $u$ :  $2g_2 - 3g_3u + \dots$ . Second, Taylor series do not generalize naturally to handle loop corrections, which have branch cuts at  $u = 0$ . Third, the strategy fails already at tree-level in the presence of gravity, since the  $\frac{8\pi G}{-u}$  pole in the  $\mathcal{C}_{2,u}$  sum rule explodes in the forward limit.

We propose that these three issues admit a common physical resolution: measure EFT couplings from small impact parameter scattering. By doing measurements at impact parameter  $b \sim 1/M$  and  $u \sim -M^2$ , the expected scaling will be automatic, branch points are avoided, and the gravity pole will be suppressed.

Let us explain the mechanism for suppressing the gravity pole in more detail, and illustrate possible applications. The physical meaning of  $u$  in dispersive sum rules  $\mathcal{C}_{k,u}$  is the magnitude of the momentum transfer. We have  $u = -(p_3 - p_1)^2$ , and  $\vec{p} = p_3 - p_1$  is the spatial momentum transfer. For high-energy scattering, fixed impact parameter scattering,  $\vec{p} = p_3 - p_1$  lies in the  $(D-2)$ -dimensional plane transverse to the incident momenta. The impact parameter  $\vec{b}$  is also a vector in  $\mathbb{R}^{D-2}$ . It is Fourier-conjugate to  $\vec{p}$ . We will write  $p = |\vec{p}|$  and  $b = |\vec{b}|$ . We will use the following conventions for going between the momentum transfer and impact parameter space. Consider a spherically symmetric momentum-space wavefunction  $f(p)$ . We define the corresponding impact parameter space wavefunction as the  $(D-2)$ -dimensional Fourier transform

$$\hat{f}(b) \equiv \int d^{D-2}\vec{p} e^{i\vec{b}\cdot\vec{p}} \frac{f(p)}{p^{D-3} \text{vol } S^{D-3}} = \Gamma\left(\frac{D-2}{2}\right) \int_0^\infty dp f(p) \frac{J_{\frac{D-4}{2}}(pb)}{(pb/2)^{\frac{D-4}{2}}}. \quad (3.1)$$

The factor  $1/p^{D-3}$  was inserted for future convenience. An integral of  $f(p)$  against the gravity pole  $\frac{8\pi G}{-u}$  can be expressed in impact parameter space as

$$\int_0^\infty dp f(p) \frac{8\pi G}{p^2} = \int_0^\infty db \hat{f}(b) \frac{8\pi G b}{D-4}. \quad (3.2)$$

Evaluation in the forward limit corresponds to  $f(p) = \delta(p)$ , which in impact parameter space is  $\hat{f}(b) = 1$ . This leads to a divergence when integrated against the graviton contribution  $\frac{8\pi G b}{D-4}$  in (3.2). By contrast, if  $\hat{f}(b)$  is localized near  $b \sim 1/M$ , the contribution of the graviton pole will automatically be suppressed by  $\sim 1/M^2$ . An example wavefunction is shown in figure 5 below. Indeed, we will soon see that in dispersive bounds  $\hat{f}(b)$  is constrained to be positive, so localization in impact parameter space is the *only* way to suppress the graviton pole relative to other contributions. To achieve such localization, the momentum-space wavefunction must have support all the way up to  $|u| \sim M^2$ .

One might worry that there is a limitation on the range of  $u$  due to the EFT series (2.20) breaking down at  $|u| \sim M^2$ . Specifically, if we evaluate dispersion relations at  $|u| \sim M^2$ ,

all contact interactions could contribute equally and it would be difficult to disentangle individual EFT coefficients. This suggests restricting to  $|u| \ll M^2$ . But that would not be good enough to get numerical bounds with the right scaling in  $M$  — it would only give parametric bounds.

The key idea for getting around this difficulty is to use low-energy crossing symmetry to eliminate all the terms starting from  $g_4 u^2$  and higher in the  $\mathcal{C}_2$  sum rule. Note that the EFT contributions to the  $\mathcal{C}_2, \mathcal{C}_4, \mathcal{C}_6$  sum rules are

$$\mathcal{C}_{2,u}|_{\text{EFT}} = \frac{8\pi G}{-u} + 2g_2 - g_3 u + 8g_4 u^2 - 2g_5 u^3 + 24g_6 u^4 - 4g_7 u^5 \dots, \quad (3.3)$$

$$\mathcal{C}_{4,u}|_{\text{EFT}} = 4g_4 - 2g_5 u + (24g_6 + g'_6)u^2 - 8g_7 u^3 + \dots, \quad (3.4)$$

$$\mathcal{C}_{6,u}|_{\text{EFT}} = 8g_6 - 4g_7 u + \dots. \quad (3.5)$$

By subtracting a linear combination of  $\mathcal{C}_{4,0}$  and  $\mathcal{C}'_{4,0} = \partial_u \mathcal{C}_{4,u}|_{u=0}$  from  $\mathcal{C}_{2,u}$ , we can cancel the  $g_4$  and  $g_5$  terms in (3.3). Next, subtracting a linear combination of  $\mathcal{C}_{6,0}$  and  $\mathcal{C}'_{6,0}$ , we can cancel the  $g_6$  and  $g_7$  terms, and so on. Repeating this procedure, we find that the following linear combination of sum rules is independent of all higher EFT coefficients:

$$\mathcal{C}_{2,u}^{\text{improved}} = \mathcal{C}_{2,u} - \sum_{n=2}^{\infty} \left( n u^{2n-2} \mathcal{C}_{2n,0} + u^{2n-1} \mathcal{C}'_{2n,0} \right). \quad (3.6)$$

Specifically, we have

$$\mathcal{C}_{2,u}^{\text{improved}}|_{\text{EFT}} = \frac{8\pi G}{-u} + 2g_2 - g_3 u, \quad (3.7)$$

with no contamination by higher contact coefficients. Note that the improved sum rule (3.6) still involves forward limits, but only of the higher-subtracted sum rules  $\mathcal{C}_4, \mathcal{C}_6, \dots$ , which do not have a graviton pole. We suspect that it should be possible to find different improvements which eliminate forward limits altogether, but (3.6) will suffice for our purposes.

The contribution of heavy states to the  $\mathcal{C}_{2,u}^{\text{improved}}$  sum rule is found by inserting the heavy contribution from (2.23) into (3.6) and performing the sum over  $n$ . This sum can be done in a closed form, yielding the following *exact* sum rule:

$$\begin{aligned} \frac{8\pi G}{-u} + 2g_2 - g_3 u &= \left\langle \frac{(2m^2 + u)\mathcal{P}_J\left(1 + \frac{2u}{m^2}\right)}{m^2(m^2 + u)^2} - \frac{u^2}{m^6} \left( \frac{(4m^2 + 3u)\mathcal{P}_J(1)}{(m^2 + u)^2} + \frac{4u\mathcal{P}'_J(1)}{m^4 - u^2} \right) \right\rangle \\ &\equiv \left\langle \mathcal{C}_{2,u}^{\text{improved}}[m^2, J] \right\rangle. \end{aligned} \quad (3.8)$$

The important feature of (3.8) is that only three EFT couplings appear on the left, yet we retain the full power of a one-parameter family of sum rules (labelled by  $u$ ), which we can use to localize at small impact parameters. In the absence of gravity,  $\mathcal{C}_{2,u}^{\text{improved}}$  is equivalent to a combination of null constraints and evaluation around the forward limit

$$\mathcal{C}_{2,u}^{\text{improved}} = \mathcal{C}_{2,0} + u\mathcal{C}'_{2,0} + u^2 \mathcal{X}_{2,u} \quad (\text{without gravity}), \quad (3.9)$$

where  $\mathcal{X}_{2,u}$  is written below in (3.15). However  $\mathcal{C}_{2,u}^{\text{improved}}$  makes sense even with gravity.

We expect that (3.8), evaluated anywhere in the range  $u \in (-M^2, 0)$ , gives a valid and convergent sum rule. When  $-u = M^2$ , the  $t$ -channel cut merges with the origin in the  $s$ -plane. Depending on the analytic structure of the  $t$ -channel cut, this may require us to modify the sum rule. Thus, we will mostly restrict to  $u \in (-M^2, 0]$ . However, for meromorphic amplitudes, we expect that (3.8) can have a larger range of validity. We discuss this idea further in appendix B.

We can now derive inequalities on the EFT couplings  $8\pi G, g_2, g_3$  by constructing functions  $f(p)$  whose integral against  $\mathcal{C}_{2,-p^2}^{\text{improved}}[m^2, J]$  is non-negative on all allowed heavy states:

$$\text{if } \int_0^M dp f(p) \mathcal{C}_{2,-p^2}^{\text{improved}}[m^2, J] \geq 0 \quad \forall m \geq M, J = 0, 2, 4 \dots \quad (3.10)$$

$$\text{then } \int_0^M dp f(p) \left[ \frac{8\pi G}{p^2} + 2g_2 + g_3 p^2 \right] \geq 0. \quad (3.11)$$

We will see that functions  $f(p)$  allowed by the condition on the first line carve out an interesting region in the  $(G, g_2, g_3)$  space. Let us interpret the first positivity condition. We claim that it *implies* the statement that  $f(p)$  is the Fourier transform of a positive function of transverse impact parameter. To see this, we take a scaling limit where  $m \rightarrow \infty$  with the ‘‘impact parameter’’  $b = \frac{2J}{m}$  held fixed. In this limit, the Gegenbauer functions become Bessel functions as follows<sup>9</sup>

$$\lim_{m \rightarrow \infty} \mathcal{P}_{\frac{mb}{2}} \left( 1 - \frac{2p^2}{m^2} \right) = \frac{\Gamma\left(\frac{D-2}{2}\right)}{(bp/2)^{\frac{D-4}{2}}} J_{\frac{D-4}{2}}(bp). \quad (3.12)$$

We then find

$$\int_0^M dp f(p) \mathcal{C}_{2,-p^2}^{\text{improved}}[m^2, J] \sim \frac{2\Gamma\left(\frac{D-2}{2}\right)}{m^4} \int_0^M dp f(p) \frac{J_{\frac{D-4}{2}}(bp)}{(bp/2)^{\frac{D-4}{2}}} \quad \text{as } m \rightarrow \infty, \quad (3.13)$$

which is indeed proportional to the transverse-plane Fourier transform of  $f(p)/p^{D-3}$ , see (3.1). This is a key finding:

*Bounds come from functions that have compact support in momentum space and are positive in impact parameter space.*

Fortunately, such functions are plentiful. Note that in addition to imposing positivity in impact parameter space, we must impose positivity of (3.10) at finite  $m \geq M$ . This further restricts the space of possible functions  $f(p)$ , but we find numerically that suitable  $f(p)$  still exist.

To obtain stronger bounds, we can supplement the  $\mathcal{C}_{2,u}^{\text{improved}}$  sum rule with additional null constraints

$$0 = \left\langle \mathcal{X}_{k,u}[m^2, J] \right\rangle, \quad k = 4, 6, \dots \quad (3.14)$$

---

<sup>9</sup>This has the following simple interpretation. Gegenbauer functions are the basis for the harmonic decomposition of functions on the sphere  $S^{D-2}$ , while Bessel functions for the Fourier decomposition of radial functions on  $\mathbb{R}^{D-2}$ . (3.12) corresponds to the flat-space limit of the sphere with momentum in  $\mathbb{R}^{D-2}$  fixed.

where [11]

$$\begin{aligned} \mathcal{X}_{k,u}[m^2, J] = & \frac{2m^2+u}{um^2(m^2+u)} \frac{m^2 \mathcal{P}_J \left(1 + \frac{2u}{m^2}\right)}{(um^2(m^2+u))^{k/2}} \\ & - \text{Res}_{u'=0} \frac{(2m^2+u')(m^2-u')(m^2+2u')}{m^2(u-u')u'(m^2-u)(m^2+u')(m^2+u+u')} \frac{m^2 \mathcal{P}_J \left(1 + \frac{2u'}{m^2}\right)}{(u'm^2(m^2+u'))^{k/2}}. \end{aligned} \quad (3.15)$$

$\mathcal{X}_{k,u}$  can be derived by starting from  $\mathcal{C}_{k,u}$  and subtracting all the EFT contributions using the forward limit of  $\mathcal{C}_{k+2,u}$ ,  $\mathcal{C}_{k+4,u}$  etc. Including the  $\mathcal{X}_{k,u}$  sum rules gives us the following linear program for bounding  $g_2$  and  $g_3$ :

$$\begin{aligned} \text{if } \forall m \geq M, J = 0, 2, 4, \dots \quad & \int_0^M dp f(p) \mathcal{C}_{2,-p^2}^{\text{improved}}[m^2, J] + \\ & + \sum_{k=4,6,\dots} \int_0^M dp h_k(p) \mathcal{X}_{k,-p^2}[m^2, J] \geq 0 \end{aligned} \quad (3.16)$$

$$\text{then } \int_0^M dp f(p) \left[ \frac{8\pi G}{p^2} + 2g_2 + g_3 p^2 \right] \geq 0,$$

where the decision variables are the functions  $f(p)$  and  $h_k(p)$ . We can choose an objective function and normalization condition to optimize different quantities. For example, to obtain the best upper bound on  $g_3$  as a function of  $g_2$  and  $8\pi G$ , we must solve the optimization problem

$$\text{minimize } \int_0^M dp f(p) \left[ \frac{8\pi G}{p^2} + 2g_2 \right] \quad \text{such that } \int_0^M dp f(p) p^2 = -1, \quad (3.17)$$

where  $f(p)$  and  $h_k(p)$  satisfy the positivity constraints in (3.16).

### 3.2 Numerical implementation

To solve the linear program (3.16) numerically, we must express  $f(p)$  and the  $h_k(p)$  as finite sums of basis functions. Importantly, we must be able to find finite linear combinations of basis functions that are positive when Fourier-transformed to impact parameter space. It is well-known that polynomials restricted to an interval can have positive Fourier transforms.<sup>10</sup> This motivates us to choose powers of  $p$  as our basis functions, for example:

$$f(p) = \sum_n a_n p^n. \quad (3.18)$$

Here,  $a_n$  are constants (decision variables) to be determined by solving the linear program, and  $n$  are powers that we can choose. The powers need not be integers, but they must obey  $n > 1$  in order for the integral of  $f(p)$  against the gravity pole  $\frac{8\pi G}{p^2}$  to converge. For technical reasons that we explain in appendix A.1, we choose the basis functions listed in

<sup>10</sup>An example is

$$\int_0^1 (1-p) \cos(pb) dp = \frac{1 - \cos b}{b^2} \geq 0.$$

$D$	basis functions for $f(p)$
5	$p^3 - p^2, p^4 - p^2, p^5 - p^2, \dots$
even $\geq 6$	$p^{3/2}, p^{5/2}, p^{7/2}, \dots$
odd $\geq 7$	$p^2, p^3, p^4, \dots$

**Table 1.** Basis functions for  $f(p)$  appearing in the linear program (3.16). The technical reasons for these choices are explained in appendix A.1.

table 1.<sup>11</sup> For the functions  $h_k(p)$ , we simply expand them in nonnegative integer powers:  $p^0, p^1, \dots$

We can now truncate the expansions of  $f(p)$  and  $h_k(p)$  in basis functions to obtain a linear program with a finite number of decision variables. We deal with the infinite number of constraints (one constraint for each  $m > M$  and  $J = 0, 2, 4, \dots$ ) using a combination of discretization and polynomial approximation, described in detail in appendix A. Importantly, we must restrict  $J$  to a finite range  $J = 0, 2, \dots, J_{\max}$ . To help the bounds converge without needing to take  $J_{\max}$  very large, we explicitly include positivity in the scaling limit (3.13) as an extra inequality. We solve the resulting optimization problems numerically using SDPB [46, 47].

### 3.3 Bounds on $g_3$ and $g_4$ without gravity

In non-gravitational theories, dispersive bounds computed using the above methods reproduce the same results obtained by expanding around the forward limit. The physical reason is that heavy averages in dispersive sum rules are dominated by small impact parameters  $b \sim 1/M$ , as observed in [11]. By using functionals that are localized on the scale  $b \sim 1/M$ , we access the same physics. As an example, in figure 3, we show a bound on  $g_3, g_4$  in  $D = 6$ , computed using our small impact parameter wavepackets. The results agree with those of [11], which used expansions around the forward limit.

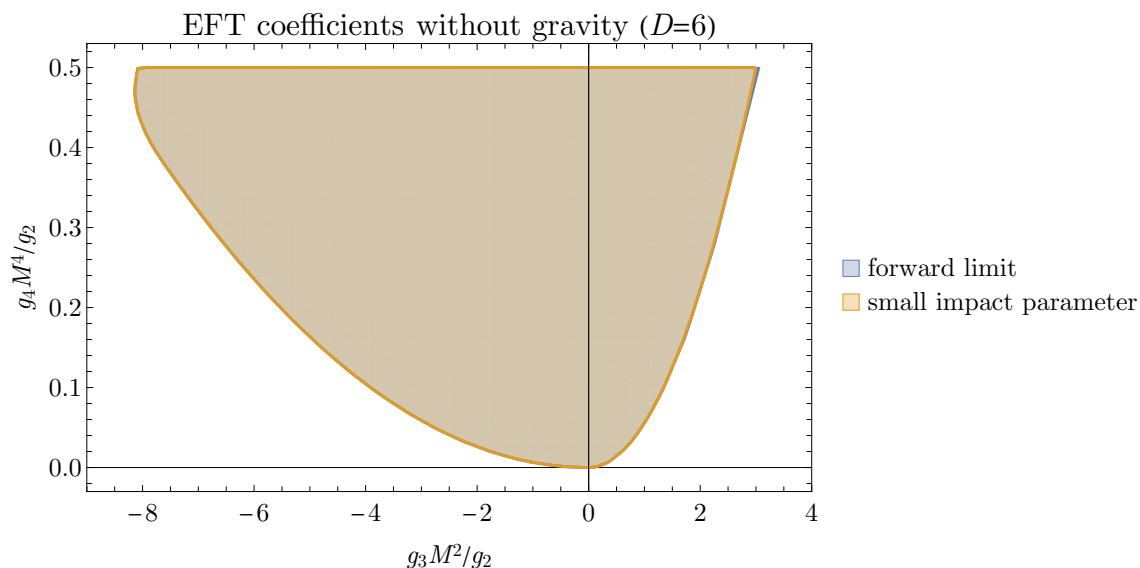
### 3.4 Bounds on $g_2$ and $g_3$ with gravity

The main advantage of our approach is that we also obtain valid bounds in gravitational theories. In figure 4, we show the allowed regions for  $g_2$  and  $g_3$  in UV-completable tree-level EFTs containing gravity in spacetime dimensions  $D = 5, \dots, 12$ . (We discuss the case  $D = 4$  in section 3.7). The bounds are computed numerically by solving (3.16) using SDPB. Note that they automatically have the expected EFT scaling in  $M$ . In short, dimensional analysis scaling is a theorem, *even in the presence of gravity!*

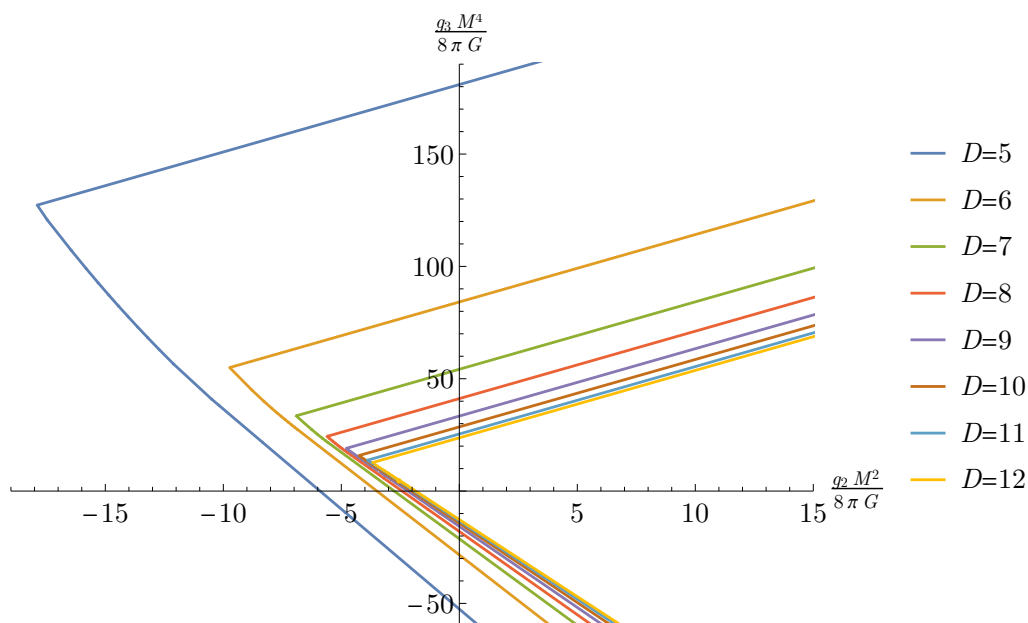
The bounds in figure 4 are computed using a 17-dimensional space of functionals built from  $\mathcal{C}_2^{\text{improved}}$  and the null constraints  $\mathcal{X}_4$  and  $\mathcal{X}_6$  (listed explicitly in table 2). Although the bounds depend on the cutoffs and approximations described in appendix A, we have chosen those cutoffs so that the results have converged within a fraction  $10^{-4}$ . The bounds

<sup>11</sup>We have also investigated other choices of pure-power basis functions and found consistent results.

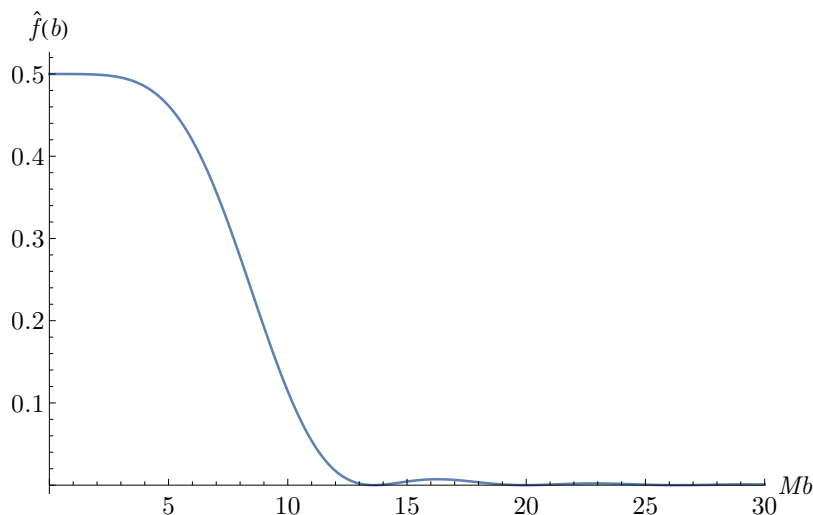




**Figure 3.** Allowed region for  $g_3$  and  $g_4$  in a non-gravitational theory in  $D = 6$  dimensions, with heavy mass scale  $M$ . We show results using two different methods: the blue region uses derivatives around the forward limit as in [11] (with a 33-dimensional space of functionals), while the yellow region uses small impact parameter wavepackets (built from the 17-dimensional space of functionals listed in table 2, together with  $\mathcal{C}_{4,u=0}$ ). The two regions are essentially identical and appear overlapping in the plot. We give more details on our numerical computations in appendix A.



**Figure 4.** Allowed regions for  $g_2$  and  $g_3$  in a theory of a scalar coupled to gravity in flat space in dimensions  $D = 5, \dots, 12$ , with heavy mass scale  $M$ . For each curve, the region to the right is allowed and the region to the left is disallowed. Each bound was computed using a 17-dimensional space of functionals, listed in table 2. We give more details on the numerical computation in appendix A. The inequalities plotted here are listed in table 3.



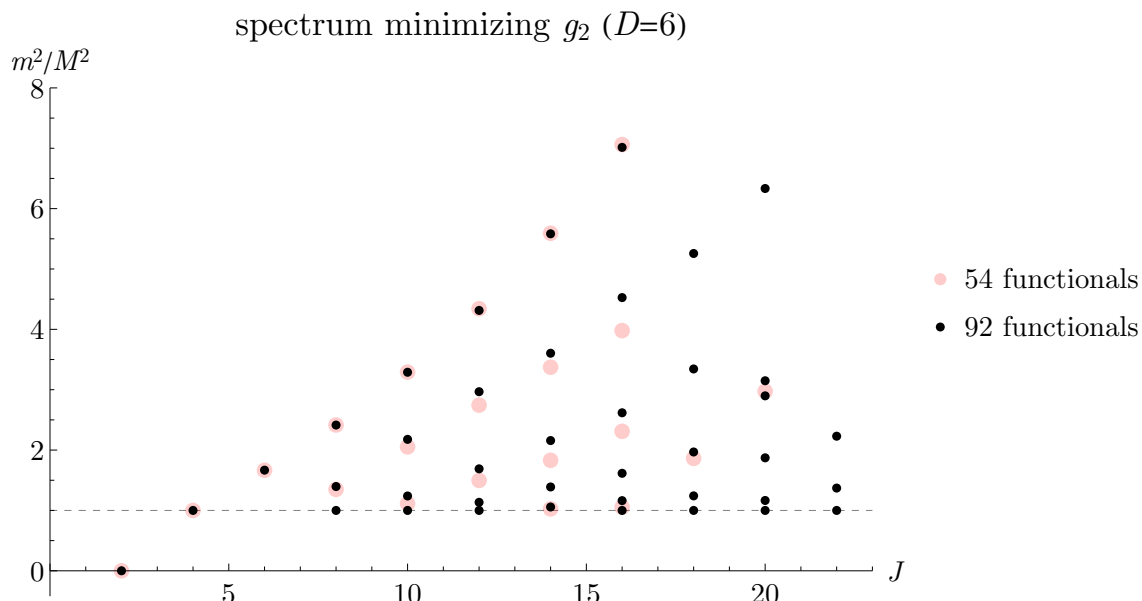
**Figure 5.** The impact parameter wavefunction  $\hat{f}(b)$  defined by (3.1) for the extremal functional that minimizes  $g_2$  in  $D = 6$ . As discussed in the text, it is localized near  $b \sim 1/M$ . The wavefunction is normalized by  $\hat{f}(0) = 1/2$  so that the contribution of  $g_2$  in (3.16) is precisely  $g_2$ . The fact that it has zero slope at  $b = 0$  guarantees that the contribution of  $g_3$  vanishes.

can be improved by choosing a larger space of functionals. We expect that the bounds shown in figure 4 are within a few percent of optimal.<sup>12</sup>

In figure 5, we show the impact parameter wavefunction  $\hat{f}(b)$  for the extremal functional that minimizes  $g_2$  in  $D = 6$ . Clearly, the numerical optimization procedure constructs sum rules dominated by  $b \sim 1/M$ .

Because our sum rules are linear and homogeneous in the EFT couplings  $8\pi G, g_2, g_3$ , we can always add an admissible amplitude without gravity to an admissible amplitude with gravity to obtain a new admissible amplitude with gravity. The allowed region in  $(g_2, g_3)$ -space without gravity is a cone  $C$  [11]. The allowed region with gravity must be a union of translations of  $C$ . Indeed, this is the case: the allowed region is similar to the non-gravitational one, but shifted so that  $g_2$  has a negative minimum value (achieved at a particular value of  $g_3$ ). Note that the bounds are stronger in larger  $D$ . This is due to the fact that the dimensional reduction of a unitary theory is unitary (more technically the fact that higher-dimensional Gegenbauer polynomials can be written as positive linear combinations of lower-dimensional Gegenbauer polynomials). Physically, it makes sense that the ratio  $g_2 M^2/G$  should not admit an upper bound:  $g_2$  and  $G$  are a priori independent couplings, measuring respectively the strength of the scalar self-interaction and the strength of gravity. We are assuming that the EFT is weakly coupled, which means that *both*  $g_2$  and  $G$  are taken to be small in units of  $M$ , but their ratio is a priori undetermined without further physical input. On the other hand, for fixed  $g_2 M^2/G$ , we expect (and will confirm) that all other dimensionless ratios  $g_k M^{2k-2}/G$  obey double-sided bounds.

<sup>12</sup>Here, we mean “optimal” for the specific infinite dimensional linear program (3.16). One could potentially obtain stronger bounds by making new assumptions about the theory in question, or by extending the range of  $u$  as discussed in appendix B.



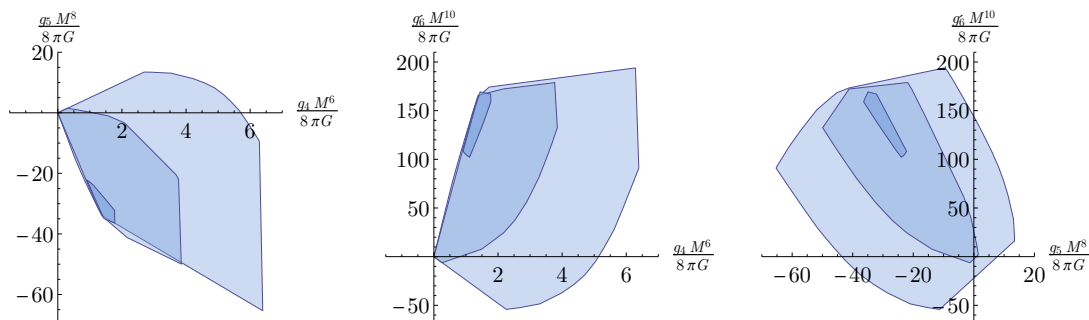
**Figure 6.** Extremal spectrum for the problem of minimizing  $g_2$  in spacetime dimension  $D = 6$ . The black points are zeros of the extremal functional, computed using a 92-dimensional space of functionals described in table 2. (We only show zeros that are stable under changing the space of functionals and other parameters.) For comparison, we show the spectrum computed using a 54-dimensional space of functionals in red. The upper trajectory has mostly converged, but the lower trajectories are still changing as we increase the space of functionals. The extremal value of  $g_2$  corresponding to this spectrum is  $g_2 M^2 / (8\pi G) = -9.57$ . In string theory, the spectrum would occupy the upper triangle, but here it is in the lower triangle.

The slope of the upper bound on  $g_3$  as a function of  $g_2$  is exactly 3. This comes from the fact that the scalar contribution to the  $\mathcal{C}_2^{\text{improved}}$  sum rule is

$$\mathcal{C}_{2,-p^2}^{\text{improved}}[m^2, J = 0] = \frac{2}{m^4} + \frac{3p^2}{m^6}, \tag{3.19}$$

which has the same form as the low-energy contribution  $2g_2 + g_3 p^2$  in (3.16). By adding scalars to the heavy spectrum, we can shift  $(g_2, g_3)$  by an arbitrary positive multiple of  $(1, 3/m^2)$ . The minimum value of  $g_2$  is achieved by a spectrum with no heavy scalars. We can compute this spectrum using the extremal functional method [48, 49]: we find functions  $f(p)$  and  $h_k(p)$  that give the optimal lower-bound on  $g_2$ . We then tabulate the values of  $m^2$  and  $J$  where the inequalities in (3.16) are saturated — i.e. the zeros of the extremal functional. We show the resulting spectrum for the case  $D = 6$  in figure 6. The extremal spectrum is remarkable (and very different from string theory) in that there is only a single state at spins  $J = 2, 4, 6$ , and a small but increasing number of states at larger  $J$ . Furthermore, the minimal value of  $m^2$  appears to be nearly flat as a function of  $J$ . It is interesting to ask whether there could be a physical theory of gravity that realizes this spectrum.

Allowed  $(g_4, g_5, g_6)$  in  $D=6$ , for any  $g_3$ , for  $\frac{g_2 M^2}{8\pi G} \in \{-9, -5, 0\}$



**Figure 7.** Bounds on the higher-dimension contact coefficients  $g_4, g_5, g_6'$  in  $D = 6$ . We compute bounds that are independent of  $g_3$  by restricting to the subspace of functionals whose EFT contribution is independent of  $g_3$ . We compute the 3-dimensional allowed region in  $(g_4, g_5, g_6')$ -space for a few example values of  $g_2$ , and then project onto pairs of axes for display. The values of  $g_2$  shown here are  $g_2 = 0$  (lightest blue),  $\frac{g_2 M^2}{8\pi G} = -5$  (medium blue),  $\frac{g_2 M^2}{8\pi G} = -9$  (darkest blue). These bounds were computed using the same parameters and functionals listed in column 1 of table 2, with the additional functionals  $\mathcal{C}_{4,-p^2}^{\text{improved}} \times \{p^0, p^1, \dots, p^5\}$ . When  $g_2$  approaches the “tip of the cone” at its minimum value  $g_2 M^2 / (8\pi G) \approx -9.6$ , the allowed region shrinks towards  $\left(\frac{g_4 M^6}{8\pi G}, \frac{g_5 M^8}{8\pi G}, \frac{g_6' M^{10}}{8\pi G}\right) \approx (1.4, -34, 165)$ .

### 3.5 Bounds on higher contact coefficients with gravity

The same method straightforwardly extends to higher EFT coefficients. Using the same strategy as in (3.6), we can define a  $\mathcal{C}_{4,u}^{\text{improved}}$  sum rule that isolates the coefficients  $g_4, g_5, g_6'$ :

$$\begin{aligned}
 4g_4 - 2ug_5 + g_6' u^2 &= \langle \mathcal{C}_{4,u}^{\text{improved}}[m^2, J] \rangle \\
 \mathcal{C}_{4,u}^{\text{improved}}[m^2, J] &= \frac{(2m^2 + u)\mathcal{P}_J(1 + \frac{2u}{m^2})}{m^4(m^2 + u)^3} - \frac{u^2(6m^6 - 4m^2u^2 + 4m^4u - 3u^3)\mathcal{P}_J(1)}{m^{12}(m^2 + u)^3} \\
 &\quad - \frac{2u^3(-3m^2u + 4m^4 - 5u^2)\mathcal{P}'_J(1)}{m^{12}(m^2 - u)(m^2 + u)^2} + \frac{4u^4\mathcal{P}''_J(1)}{m^{12}(u - m^2)(m^2 + u)}.
 \end{aligned}
 \tag{3.20}$$

The coefficient  $g_6$ , which multiplies  $(s^2 + t^2 + u^2)^3$  in the Lagrangian, was eliminated since it can be measured by the higher-subtracted sum rule  $\mathcal{C}_{6,u}$ : we will therefore not discuss it here. We then generalize (3.16) to include linear combinations of the  $\mathcal{C}_{2,-p^2}^{\text{improved}}, \mathcal{C}_{4,-p^2}^{\text{improved}}$  and  $\mathcal{X}_{k,-p^2}$  sum rules, each integrated against its own function of  $p$ .<sup>13</sup> By finding linear combinations that are positive for all  $m \geq M$  and  $J = 0, 2, 4, \dots$ , we obtain inequalities on EFT data with the correct scaling in  $8\pi G$  and  $M$ . As an example, in figure 7 we show bounds on  $g_4, g_5, g_6'$  in spacetime dimension  $D = 6$ , for some example values of  $g_2$ .

<sup>13</sup>Alternatively, since the forward limits  $u \rightarrow 0$  of spin-4 sum rules converge, we could simply use derivatives of the un-improved  $\mathcal{C}_4$ , supplemented with null constraints, as opposed to using  $\mathcal{C}_{4,u}^{\text{improved}}$ .

### 3.6 Some solutions to the constraints from string theory

It is interesting to compare our bounds to explicit solutions of crossing symmetry with gravity. The exchange of gravitons in the three channels alone violates the Regge bound (2.14) and needs to be UV completed. A simple UV completion is provided by the amplitude of four real dilatons in type II string theory [50]

$$\mathcal{M}_{\text{string}}(s, t) = -8\pi G \frac{(st)^2 + (tu)^2 + (su)^2}{m^6} \frac{\Gamma\left(-\frac{s}{m^2}\right)\Gamma\left(-\frac{t}{m^2}\right)\Gamma\left(-\frac{u}{m^2}\right)}{\Gamma\left(\frac{s}{m^2} + 1\right)\Gamma\left(\frac{t}{m^2} + 1\right)\Gamma\left(\frac{u}{m^2} + 1\right)}. \quad (3.21)$$

Here  $m \geq M$  is the mass of the lightest massive state exchanged. It is straightforward to check that along any ray of constant phase in the  $s$  plane away from the real axis, at fixed  $u$ , we have

$$|\mathcal{M}_{\text{string}}(s, u)| = O(|s|^{2+2u/m^2}), \quad (3.22)$$

in agreement with (2.14). Furthermore, the amplitude is unitary for  $2 \leq D \leq 24$ .<sup>14</sup> The low-energy expansion of (3.21) starts as follows

$$\mathcal{M}_{\text{string}}(s, t) = 8\pi G \left( \frac{tu}{s} + \frac{us}{t} + \frac{st}{u} \right) + \frac{4\pi G \zeta(3)}{m^6} (s^2 + t^2 + u^2)^2 + \dots, \quad (3.23)$$

i.e. this amplitude has

$$g_2 = 0, \quad g_3 = 0, \quad g_4 = \frac{4\pi G \zeta(3)}{m^6}, \quad (3.24)$$

meaning that the point

$$\frac{g_2 M^2}{8\pi G} = 0 \quad \frac{g_3 M^4}{8\pi G} = 0 \quad \frac{g_4 M^6}{8\pi G} = x \frac{\zeta(3)}{2} \quad (3.25)$$

is consistent with all the constraints for any  $0 < x \leq 1$  as long as  $2 \leq D \leq 24$ .

This point lies at the origin in figure 4 and does not come close to saturating the bounds on  $g_2$  and  $g_3$ . One can construct a consistent solution of the constraints lying closer to the bound by subtracting all the scalar (spin-0) exchanges from  $\mathcal{M}_{\text{string}}$  but this only has the effect of translating the solution to the lower-left of the origin in figure 4. For example, in  $D = 6$  we find:

$$\frac{g_2 m^2}{8\pi G} \approx -0.75, \quad \frac{g_3 m^4}{8\pi G} \approx -2.07, \quad (3.26)$$

bringing us somewhat closer to the numerical bound, but still far from saturation. We can also consider scattering of real dilatons in the heterotic string [52], where we find

$$\frac{g_2 m^2}{8\pi G} = \frac{3}{16}, \quad \frac{g_3 m^4}{8\pi G} = \frac{3}{4}, \quad (3.27)$$

also safely within our bounds.

---

<sup>14</sup>Unitarity for  $2 \leq D \leq 10$  follows from unitarity of type II superstring theory. The residues of (3.21) at the massive poles experimentally admit a positive expansion into Gegenbauer polynomials in the extended range  $2 \leq D \leq 24.148\dots$ . We are not aware of a proof of unitarity of (3.21) in the extended range, see [51].

String theory is of course by no means the unique way to “unitarize” tree-level two-to-two graviton exchange, especially at large impact parameters. Perhaps the simplest idea (pre-dating string theory) is to exponentiate the eikonal phase. To see this let’s transform the amplitude to impact parameter space as

$$\mathcal{M}(s, b) \equiv \frac{1}{2s} \int \frac{d^{D-2}p}{(2\pi)^{D-4}} e^{ip \cdot b} \mathcal{M}(s, u = -p^2) \quad (3.28)$$

which allows to write the eikonal amplitude as (see [53] in the context of string theory)

$$\mathcal{M}_{\text{eik}}(s, b) = -i \left( e^{is\hat{\chi}(b)} - 1 \right), \quad \hat{\chi}(b) = \frac{G\Gamma\left(\frac{D-4}{2}\right)}{(\pi b^2)^{\frac{D-4}{2}}}. \quad (3.29)$$

Here we are thinking of  $b$  large compared to the string scale and other scales. Expanding the exponential to first order reproduces tree-level graviton exchange. It is not hard to see that this model indeed saturates the  $\mathcal{C}_{2,u}^{\text{improved}}$  sum rule from eq. (3.8), with the lower endpoint removed: in Fourier space that sum rule reads

$$\hat{\chi}(b) = \frac{2}{\pi} \int_0^\infty \frac{ds}{s^2} \text{Im} \mathcal{M}(s, b) \rightarrow_{\text{eik}} \frac{2}{\pi} \int_0^\infty \frac{ds}{s^2} 2 \sin^2 \left( \frac{s\hat{\chi}(b)}{2} \right). \quad (3.30)$$

The eikonal model does not quite satisfy the assumptions in the paper and so we do not put the corresponding data in the same plot (one would have to choose a scheme for subtracting the low-energy loop contribution to the spectral density from  $s < M^2$ ), however this contribution is negligible at sufficiently large impact parameters, on which we will focus here. One can interpret the above sum rules as constraining high-energy data at fixed “impact parameter”  $b = 2J/m$ . This may be seen using eqs. (2.18) and (3.12) and integrating using orthogonality of Bessel functions:

$$\hat{\chi}(b) \approx \left\langle \frac{1}{bm^4} \delta \left( b - \frac{2J}{m} \right) \right\rangle \frac{\Gamma\left(\frac{D-2}{2}\right)}{(b/2)^{D-4}}. \quad (3.31)$$

This should be understood in the sense of distributions, with sufficient smearing in  $b$  to satisfy the finite- $p$  support condition, and it shows that low-energy gravity predicts high-energy averages at fixed  $J/m$ . These constraints are of course satisfied in tree-level string theory, but rather differently than in the eikonal approximation. Due to the famous logarithmic spreading with energy,

$$\text{Im} \mathcal{M}_{\text{string}}(s, b) \propto e^{-\frac{b^2 M^2}{8 \log s}} \quad (3.32)$$

at large impact parameters  $bM \gg 1$  string theory’s spectral density effectively vanishes below *exponentially* large  $s$ .

The spectral densities (3.30) and (3.32) could hardly differ more from each other. Yet they are both physically reasonable and are both realized in different regimes of string theory. In our view, this indicates that high-energy models should be used with great care when deriving EFT bounds, or perhaps avoided (as we do in this paper).

### 3.7 Comments on $D = 4$ and infrared divergences

Trying to play the above game for gravitational theories in  $D = 4$ , we run into a problem. In order for the integral against gravity to converge, we need  $\lim_{p \rightarrow 0} f(p)/p$  to vanish. However, when  $D = 4$ , this limit is proportional to  $\int d^2 \vec{b} \hat{f}(\vec{b})$ , which must be strictly positive.

In other words, if we restrict to functionals that are positive at all impact parameters, the action on gravity is logarithmically divergent. This reflects the infrared divergence of the gravitational potential in the Regge limit, given by the two-dimensional Fourier transform  $\int \frac{d^2 p}{p^2} e^{ip \cdot b}$ .

A simple solution is to require positivity only at distances less than an IR cutoff  $b < b_{\max}$ . In practice, this can be achieved simply by introducing a small cutoff in momentum space,  $|p| > p_{\min}$  where  $p_{\min} \ll M$ . We then consider the linear programming problem in (3.16) but replacing the action on gravity by  $f_2(p = 0)$ , which gives the coefficient of the logarithmic divergence  $\sim \log \frac{1}{p_{\min}}$ . In this “leading-log” approximation the shape of the extremal functional is independent of the cutoff. The cutoff  $p_{\min}$  can then be quantitatively related to  $b_{\max}$ , the impact parameter at which the action becomes negative, by plotting the Fourier transform of the functional. In fact, the relation can be found analytically by series-expanding at small  $p_{\min}$ , from which we find  $p_{\min}^2 = c/(M b_{\max}^3)$  where the numerical constant  $c \approx 1$  is found from the extremal functional. In this way we obtain the lower bound:<sup>15</sup>

$$g_2 \geq -\frac{8\pi G}{M^2} \times 25 \log(0.3 M b_{\max}) \quad (D = 4). \quad (3.33)$$

What physical value should we choose for the cutoff? In the context of AdS/CFT, there will be a clear choice: the AdS scale  $b_{\max} \sim R_{\text{AdS}}$ . But this could be an over-estimate since the distance  $b_{\max}$  need simply be a scale outside which we consider the amplitude to be computable, for example using the eikonal approximation. Negativity of a functional is not necessarily a problem if it occurs in a region under analytic control. In Minkowski space, it should be possible to make this analysis fully rigorous by considering coherent states of the scalar and its radiation; this would require knowing the properties of such dressed states under crossing symmetry and discontinuities. We leave this to future work.

### 3.8 Maximal supergravity: bounding graviton scattering

Can two gravitons produce heavy states with an arbitrary cross-section, or must all processes involving gravitons be suppressed by  $G$ ? Here we give partial support for the latter idea, in the special case of maximal supersymmetry.

The technical simplification is that the graviton lies in the same multiplet as a scalar. We can factor out the helicity dependence and effectively we have a massless real scalar with low-energy amplitude

$$\mathcal{M}_{\text{susy}}(s, u) = \frac{8\pi G}{stu} + g_0 + g_2(s^2 + t^2 + u^2) + \dots \quad (3.34)$$

---

<sup>15</sup>An earlier arXiv version of this paper used an incorrect relation between  $p_{\min}$  with  $b_{\max}$ , which resulted in a numerically incorrect bound.

Maximal supersymmetry effectively improves the  $s \rightarrow \infty$  behavior by four powers, so that we have the high energy bound

$$s^2 \mathcal{M}_{\text{susy}}(s, u) \rightarrow 0 \quad \text{as } s \rightarrow \infty. \quad (3.35)$$

The  $(-2)$ -subtracted dispersion relation now converges for  $u < 0$ . The corresponding  $\mathcal{C}_{-2}$  and  $\mathcal{C}_0$  sum rules read:

$$\mathcal{C}_{-2} : \frac{8\pi G}{-u} = \left\langle m^2(2m^2 + u) \mathcal{P}_J \left( 1 + \frac{2u}{m^2} \right) \right\rangle, \quad (3.36)$$

$$\mathcal{C}_0 : g_0 + 2g_2 u^2 + \dots = \left\langle \frac{(2m^2 + u) \mathcal{P}_J \left( 1 + \frac{2u}{m^2} \right)}{m^2 + u} \right\rangle. \quad (3.37)$$

Note how gravity enters its own sum rule, apparently decoupled from the rest. This is a simplifying feature of supersymmetry. Since the second line involves higher powers of  $1/m^2$  than the first, we can bound  $g_0$  in terms of gravity and the heavy mass  $M$ . Again the trick is to measure  $g_0$  using small impact parameters.

Proceeding as in (3.8), we may eliminate the higher contacts  $2g_2 u^2 + \dots$  from the left-hand side of (3.37) using higher subtracted sum rules. A shortcut is to use  $s \leftrightarrow u$  symmetry of the unsubtracted dispersion relation:

$$\left\langle \frac{m^2(2m^2 + u) \mathcal{P}_J \left( 1 + \frac{2u}{m^2} \right)}{(m^2 - s)(m^2 + s + u)} \right\rangle = \left\langle \frac{m^2(2m^2 + s) \mathcal{P}_J \left( 1 + \frac{2s}{m^2} \right)}{(m^2 - u)(m^2 + s + u)} \right\rangle \quad (s, u < 0). \quad (3.38)$$

Note that this equality involves only heavy data: the massless poles contribute  $\frac{1}{stu}$  to both sides and cancel each other. Setting  $s = 0$  we obtain a family of null constraints

$$0 = \left\langle \frac{(2m^2 + u) \mathcal{P}_J \left( 1 + \frac{2u}{m^2} \right)}{m^2 + u} - \frac{2m^4}{m^4 - u^2} \right\rangle \quad (u < 0), \quad (3.39)$$

where the first term is the average that previously appeared in (3.37). Adding the identity  $g_0 = \langle 2 \rangle$  then gives the desired analog of (3.8):

$$\begin{aligned} g_0 &= \left\langle \frac{(2m^2 + u) \mathcal{P}_J \left( 1 + \frac{2u}{m^2} \right)}{m^2 + u} - \frac{2u^2}{m^4 - u^2} \right\rangle_{\text{heavy}}. \\ &\equiv \left\langle \mathcal{C}_{0,u}^{\text{improved}}[m^2, J] \right\rangle. \end{aligned} \quad (3.40)$$

Compared with (3.37), the left-hand-side is now under complete control. We have an infinite family of ways to measure  $g_0$ , labelled by  $-M^2 < u \leq 0$ .

Using the linear programming strategy in (3.16), now integrating the gravity measurement in (3.36) against powers  $p^2, p^3, \dots, p^{10}$ , the  $g_0$  measurement in (3.40) against four powers of  $p$ , and  $\mathcal{X}_2$  against two powers of  $p$  (for a total of 14 functionals), we proved the following bound:

$$0 \leq g_0 \leq 3.000 \frac{8\pi G}{M^6} \quad (D = 10, \text{ maximal supergravity}). \quad (3.41)$$



We find it remarkable that such a bound exists at all. It shows that in a theory with 32 real supercharges all interactions must shut down as  $G \rightarrow 0$ . It is an interesting question whether a similar statement holds for the coupling of gravitons to heavy states in non-supersymmetric theories.

The bound is compatible with type II string theory, where

$$\frac{g_0 M^6}{8\pi G} = 2\zeta(3) \approx 2.40 < 3.000. \tag{3.42}$$

It would be interesting to find a model which saturates the bound. Note that the bound is sharp only for theories weakly coupled below the scale  $M$  (i.e.  $M \ll M_{\text{pl}}$ ), since we neglected EFT loops.

Recent work [54] also considered bounds on  $g_0$  with maximal supersymmetry in ten dimensions. Their set up differs from ours in that they consider bounds on  $g_0/8\pi G$  not in the units of the UV cutoff  $M$  but in Planck units  $M_{\text{pl}}$ , without assuming weak coupling  $M \ll M_{\text{pl}}$ . Therefore, in their case there is no upper bound on  $g_0 M_{\text{pl}}^6/8\pi G$  since this ratio gets arbitrarily large in weakly coupled string theory. Intriguingly, they provide evidence that besides the rigorous bound  $g_0 \geq 0$ , there should be a stronger lower bound of the form  $g_0 M_{\text{pl}}^6/8\pi G \geq c > 0$ , possibly saturated by strongly coupled string theory. In the units of the UV cutoff, this lower bound approaches zero at weak coupling and thus is compatible with (3.41).

## 4 Conclusions

In this paper, we considered higher derivative corrections in UV consistent gravitational theories in flat space. We explained how to derive bounds on these corrections using dispersion relations for the S-matrix. We focussed on the weakly coupled regime, meaning that the gravitational and any other interaction is very small. In this regime, the ratios  $g_n/8\pi G$  of higher derivative couplings to the gravitational coupling are fixed numbers. On physical grounds, it is expected that in consistent theories, these numbers should be suppressed by inverse powers of the UV cutoff  $M$ . Here we define  $M$  to be the mass of the first state which does not appear in the low-energy effective field theory. It has been known for some time [17] that theories where higher derivative corrections are large in the units of  $M$  violate causality. Nevertheless, the long-standing challenge has been to turn such parametric bounds into precise bounds on the order one coefficients. In this paper, we derived such bounds.

We solved the problem in the context of theories containing a light (massless) scalar coupled to gravity. In such theories, we can consider higher derivative contact self-interaction of the scalar particle. The leading interaction, which we denoted  $g_2$ , has four derivatives, followed by  $g_3$  with six derivatives,  $g_4$  with eight etc. In *non-gravitational* theories, causality and unitarity have long been known to imply that  $g_2$  must be positive [7]. More recently,  $g_3/g_2$ ,  $g_4/g_2$  etc. have been argued to satisfy two-sided bound in the units of  $M$  in the absence of gravity [10, 11], starting from dispersion relations expanded around the forward limit. However, the incorporation of gravity poses an obstruction to

such program since the exchange of massless gravitons gives rise to a pole at vanishing momentum transfer.

In this paper, we overcame this difficulty by localizing the dispersion relations at small impact parameters rather than at small momentum transfer. This automatically leads to the correct EFT scaling and gives rise to a robust bootstrap program for bounding the order one coefficients. We showed that  $g_2$  is bounded from below and that  $g_3$ ,  $g_4$  and higher couplings are confined to compact regions at a fixed  $g_2$ , see figures 4 and 7. The theory minimizing the  $g_2$  coupling exhibits a peculiar spectrum shown in figure 6. It would be interesting to understand if it can come from a consistent theory of gravity.

We also considered theories with maximal supersymmetry. In this case, we were able to give both upper and lower bound on the leading correction, corresponding to  $R^4$ . This shows that all interactions must shut down in a theory of gravity with maximal supersymmetry if we take  $G \rightarrow 0$ .

Our results opens up several obvious avenues for future research. The type of question addressed here in flat space has a natural analogue in AdS, and will be the subject of an upcoming work [31]. It would also be extremely interesting to derive similar bounds in the presence of a positive cosmological constant. To do that, one would first need to clarify the consequences of causality and unitarity in de Sitter space.

We have focussed on the simplest example of identical massless scalars. It will be natural to consider in our framework the scattering of more general external states — most fundamentally, graviton scattering. We have already mentioned the need for a refinement of our method in  $D = 4$  to handle the IR divergence in the impact parameter representation. The incorporation of EFT loops is another natural direction. Our results can be regarded as a step in the classification program of weakly coupled theories of gravity, in the spirit of [17]. The tools are now mature to pursue this program systematically.

## Acknowledgments

The authors would like to thank Walter Landry, Yue-Zhou Li and Julio Parra Martinez for useful discussions. The work of S.C.-H. 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, the Simons Collaboration on the Nonperturbative Bootstrap, and the Sloan Foundation. D.M. gratefully acknowledges funding provided by Edward and Kiyomi Baird as well as the grant DE-SC0009988 from the U.S. Department of Energy. The work of L.R. is supported in part by NSF grant # PHY-1915093 and by the Simons Foundation (Simons Collaboration on the Nonperturbative Bootstrap and Simons Investigator Award). D.S.-D. is supported by Simons Foundation grant 488657 (Simons Collaboration on the Nonperturbative Bootstrap) and a DOE Early Career Award under grant no. DE-SC0019085. Some of the computations in this work were performed on the Caltech High-Performance Cluster, partially supported by a grant from the Gordon and Betty Moore Foundation.

## A Details on numerics

In this appendix, we give details on our numerical implementation of the linear program (3.16), which we reproduce here for convenience.

$$\text{if: } \int_0^1 dp f(p) \mathcal{C}_{2,-p^2}^{\text{improved}}[m^2, J] + \sum_{k=4,6,\dots} \int_0^1 dp h_k(p) \mathcal{X}_{k,-p^2}[m^2, J] \geq 0$$

$$\forall m \geq 1, J = 0, 2, 4, \dots \tag{A.1}$$

$$\text{then: } \int_0^1 dp f(p) \left[ \frac{1}{p^2} + 2g_2 + g_3 p^2 \right] \geq 0, \tag{A.2}$$

We have set  $M = 1$  and  $8\pi G = 1$ ; the dependence on these quantities can be restored using dimensional analysis and homogeneity.

We are free to choose any objective function and normalization condition on the functions  $f(p)$  and  $h_k(p)$ . The solution of the resulting optimization problem then provides a valid inequality on EFT data of the form (A.2). The full allowed region in the  $(g_2, g_3)$  plane is the intersection of the allowed regions for all such inequalities. For example, to obtain the plots in figure 4, we maximized the distance from a chosen point  $(g_{2,0}, g_{3,0})$  along rays of constant angle  $\theta$  in the  $(g_2, g_3)$  plane. We chose  $(g_{2,0}, g_{3,0})$  near the tip of the expected allowed region (known from earlier experimentation) and scanned over angles  $\theta \in \{0, \frac{\pi}{20}, \dots, \frac{39\pi}{20}\}$ .

We expand  $f(p)$  and  $h_k(p)$  in pure powers of  $p$ ,

$$f(p) = \sum_n a_n p^n, \quad h_k(p) = \sum_{i=0}^{i_k} b_{k,i} p^i. \tag{A.3}$$

For each  $J$ , the integrals (A.1) against pure powers of  $p$  can be computed analytically in terms of  ${}_2F_1$  hypergeometric functions, for example

$$\int_0^1 dp p^n \mathcal{C}_{2,-p^2}^{\text{improved}}[m^2, 2]$$

$$= -\frac{4(D-1) {}_2F_1\left(1, \frac{n+1}{2}; \frac{n+3}{2}; -\frac{1}{m^2}\right)}{(D-2)m^4(n+1)} + \frac{2(3D-4)}{(D-2)m^4(n+1)} - \frac{3(3D-2)}{(D-2)m^6(n+3)}. \tag{A.4}$$

Parametrizing  $m^2 = \frac{1}{1-x}$ , we would ideally like to impose (A.1) for all  $J = 0, 2, \dots$  and  $x \in [0, 1)$ . In practice, we must restrict  $J \in \{0, 2, \dots, J_{\text{max}}\}$  and discretize  $x$ .<sup>16</sup> Our initial discretization is

$$x \in \left\{ 0, \delta_x, 2\delta_x, \dots, \left\lceil \frac{1}{\delta_x} - 1 \right\rceil \delta_x \right\} \tag{A.5}$$

for some a small parameter  $\delta_x$ , listed below in table 2. Discretizing  $x$  weakens the inequalities, potentially resulting in incorrect bounds. However, we can effectively remove this problem by adaptively refining the discretization as described in section A.2.

<sup>16</sup>Alternatively, it would be interesting to find an approximation for functions like (A.4) in terms of a positive function of  $x$  times a polynomial. This would allow us to rewrite positivity constraints in terms of positive semidefinite matrices and apply semidefinite programming, as done for CFT four-point functions [55, 56].

### A.1 Impact parameter space inequalities

Restricting  $J \leq J_{\max}$  weakens the inequalities as well. To reduce the dependence of the resulting bounds on  $J_{\max}$ , it is useful to explicitly include inequality constraints from the scaling limit  $m \rightarrow \infty$  with fixed impact parameter  $b = \frac{2J}{m}$ :

$$\Gamma\left(\frac{D-2}{2}\right) \int_0^1 dp f(p) \frac{J_{\frac{D-4}{2}}(bp)}{(bp/2)^{\frac{D-4}{2}}} \geq 0 \quad \forall b \geq 0. \quad (\text{A.6})$$

(The null constraints  $\mathcal{X}_4, \mathcal{X}_6, \dots$  are subleading in this limit, so the functions  $h_k(p)$  don't enter this condition.) Again, the integral against a pure power of  $p$  can be done analytically:

$$\Gamma\left(\frac{D-2}{2}\right) \int_0^1 dp p^n \frac{J_{\frac{D-4}{2}}(bp)}{(bp/2)^{\frac{D-4}{2}}} = \frac{{}_1F_2\left(\frac{n+1}{2}; \frac{D-2}{2}, \frac{n+3}{2}; -\frac{b^2}{4}\right)}{n+1}. \quad (\text{A.7})$$

The resulting functions have an oscillatory and non-oscillatory part at large  $b$ :<sup>17</sup>

$$\frac{{}_1F_2\left(\frac{n+1}{2}; \frac{D-2}{2}, \frac{n+3}{2}; -\frac{b^2}{4}\right)}{n+1} \sim \frac{2^n \Gamma\left(\frac{D-2}{2}\right) \Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{D-n-3}{2}\right)} \frac{1}{b^{n+1}} + \frac{2^{\frac{D-3}{2}} \Gamma\left(\frac{D-2}{2}\right) \cos\left(b - \frac{\pi(D-1)}{4}\right)}{\sqrt{\pi} b^{\frac{D-1}{2}}} + \dots \quad (\text{A.8})$$

In order for linear combinations of such functions to be positive at large  $b$ , we must either include at least one  $n$  such that  $n \leq \frac{D-3}{2}$ , or take linear combinations of functions that cancel the leading oscillatory term at large  $b$ . We return to this statement below.

To efficiently impose positivity at large  $b$ , we use the following trick. After writing  $f(p)$  as a sum of pure powers, we have

$$\Gamma\left(\frac{D-2}{2}\right) \int_0^1 dp f(p) \frac{J_{\frac{D-4}{2}}(bp)}{(bp/2)^{\frac{D-4}{2}}} = A(b) + B(b) \cos\left(b - \frac{\pi(D-1)}{4}\right) + C(b) \sin\left(b - \frac{\pi(D-1)}{4}\right), \quad (\text{A.9})$$

where the functions  $A(b), B(b), C(b)$  have well-behaved asymptotic expansions in inverse powers of  $b$ . Let us now write

$$A + B \cos \phi + C \sin \phi = \begin{pmatrix} \cos \frac{\phi}{2} & \sin \frac{\phi}{2} \\ C & A - B \end{pmatrix} \begin{pmatrix} \cos \frac{\phi}{2} \\ \sin \frac{\phi}{2} \end{pmatrix}. \quad (\text{A.10})$$

We can thus replace positivity of (A.9) with the stronger condition

$$\begin{pmatrix} A(b) + B(b) & C(b) \\ C(b) & A(b) - B(b) \end{pmatrix} \succeq 0, \quad (\text{A.11})$$

where “ $M \succeq 0$ ” means  $M$  is positive semidefinite. Because (A.11) implies (A.9), this replacement is rigorous, but may result in sub-optimal bounds. However, at large  $b$ , the  $\cos(b - \frac{\pi(D-1)}{4})$  and  $\sin(b - \frac{\pi(D-1)}{4})$  terms in (A.9) are rapidly oscillating so that (A.11) is

<sup>17</sup>The full expansion of  ${}_1F_2$  hypergeometric functions around infinity can be found in [57].

a good approximation to the original positivity condition. In (A.11), we can now expand  $A(b), B(b), C(b)$  in inverse powers of  $b$ , and truncate the expansions. If  $n \equiv \frac{D-3}{2} \pmod{1}$ , after pulling out a positive factor, we obtain a matrix polynomial inequality of  $b$ , which can be used in SDPB.<sup>18</sup>

These considerations suggest that we should expand  $f(p)$  in the functions  $p^2, p^3, \dots$  in odd  $D$  and the functions  $p^{3/2}, p^{5/2}, \dots$  in even  $D$ . This works well except in  $D = 5$ , since  $2 > \frac{5-3}{2}$ , so generic linear combinations of (A.8) cannot be positive at large  $b$ . One solution in  $D = 5$  is to cancel the leading oscillatory term at large  $b$ , which can be done by using differences  $p^3 - p^2, p^4 - p^2, \dots$ . This leads to the basis functions shown in table 1.

We encode positivity using the  $2 \times 2$  matrix (A.11) for  $b \geq B$ , where  $B$  is some cutoff. We keep  $m_{\max}$  subleading terms in the expansion of  $B(b)$  and  $C(b)$  at large  $b$ , where  $m_{\max}$  is listed below. For smaller  $b \leq B$ , we impose positivity at discretized impact parameters:

$$b \in \left\{ \epsilon_b, \epsilon_b + \delta_b, \dots, \epsilon_b + \left\lceil \frac{B - \epsilon_b}{\delta_b} - 1 \right\rceil \delta_b \right\}, \tag{A.12}$$

where  $\epsilon_b, \delta_b$  are small parameters listed below in table 2. Like in the fixed- $J$  case, we adaptively refine our discretization of  $b$ , as described in section A.2.

## A.2 Outer approximation/adaptive refinement

Our problem has several constraints that depend on a continuous parameter. For example, we would like to impose positivity of (A.1) for all  $x \in [0, 1)$  and  $J = 0, 2, 4, \dots$ , where  $m = \frac{1}{1-x}$ . By only imposing positivity at a discrete set of  $x$  as in (A.5), we run the risk that the solver could return a solution that is negative between two discretized values. In fact, this almost always happens, since the solution to any optimization problem involves some set of saturated inequalities. If the left-hand side of (A.1) is zero at some value of  $x$ , it will generically be negative on one side of that zero. Typically, the solver returns a solution that vanishes at pairs of neighboring discrete values of  $x$ , and is negative between them.

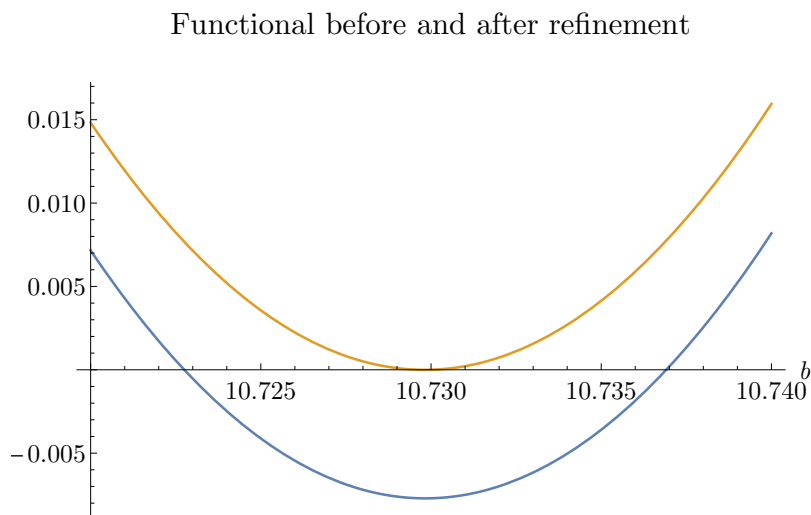
We mitigate this problem by adaptively refining the discretization. We begin with an initial discretization of  $x$  and  $b$  and run the solver. The resulting functional will be negative between pairs of points in the initial discretization. We identify these negative regions and add new positivity constraints in a finer-spaced grid covering the negative regions. Specifically, suppose that the functional dips negative between a pair  $(x_1, x_2)$  of discretized  $x$ . Let the minimum of the functional between  $x_1$  and  $x_2$  (which we estimate from a quadratic approximation) be  $x_*$ . We add new positivity constraints at the locations

$$\{x_* - Ns, x_* - (N - 1)s, \dots, x_*, \dots, x_* + (N - 1)s, x_* + Ns\}, \tag{A.13}$$

where  $s = |x_2 - x_1|/N$  with e.g.  $N = 10$ . Running the solver again with the new constraints included, the new solution will typically have a saturated inequality near  $x_*$  with a negative region reduced in size by a factor of  $N$ . We repeat this procedure until the negative regions

---

<sup>18</sup>Alternatively, if  $n \equiv \frac{D-3}{2} + \frac{p}{q} \pmod{1}$ , we can perform a change of variables  $b \rightarrow b'^q$  to again obtain a matrix polynomial of  $b'$ . However, this step multiplies the degree of the resulting polynomial by  $q$ , which can result in a performance hit in SDPB.



**Figure 8.** A functional returned by SDPB before refinement (blue) and after 3 steps of refining the discretization and re-running SDPB (orange). We plot the action of the functional in impact parameter space, zoomed in near a negative region after the initial run of SDPB. Our initial discretization used step size  $\delta_b = 1/32$ . The functional indeed has a negative region of roughly size  $\delta_b$ , with minimum value  $-0.0077$  in this region. After the refinement steps, the negative region is  $b \in (10.7298112, 10.7298134)$ , and the minimum value of the functional is  $-1.83 \times 10^{-10}$ .

are extremely small (e.g. of size  $10^{-6}$ ), and it is clear that the solution is converging to a nonnegative functional. The solution after refinement is usually quite close to the original solution. We give an example of a negative region in an initial solution and the result after refining the discretization in figure 8.

An important benefit of this method is that the bounds become essentially independent of the parameters  $\epsilon_x, \delta_x, \epsilon_b, \delta_b$  describing the initial discretization, provided they are sufficiently small.

Abstractly, the space of allowed functions  $f(p), h_k(p)$  is a convex region carved out by an infinite number of inequalities. By discretizing  $x$ , we obtain an “outer approximation” of this region in terms of a finite number of inequalities. (“Outer” because our approximate region is bigger than the true region.) By refining our discretization, we obtain a more accurate outer approximation. An efficient implementation of this method should include a way of hot-starting from the previous solution after each refinement step. We have not implemented this — instead we simply run SDPB from scratch after each refinement. Typically only a few refinement steps are needed, so this does not give a huge performance hit.

### A.3 Choices of parameters and numerical results

Our implementation of the linear program (3.16) involves several parameters. In table 2, we list the parameters used for the computations in this work. The inequalities plotted in figure 4 (in units where  $8\pi G = 1$  and  $M = 1$ ) are given in table 3.

	Figure 4	Figure 6
functionals	$\{p^n \mathcal{C}_2^{\text{improved}} \mid n \leq 7.5\}$ $\cup \{p^i \mathcal{X}_4 \mid i = 0, 1, 2, 3, 4, 5\}$ $\cup \{p^i \mathcal{X}_6 \mid i = 0, 1, 2, 3\}$	$\{p^n \mathcal{C}_2^{\text{improved}} \mid n = 1.5, 2.5, \dots, 27.5\}$ $\cup \{p^i \mathcal{X}_4 \mid i = 0, 1, \dots, 15\}$ $\cup \{p^i \mathcal{X}_6 \mid i = 0, 1, \dots, 13\}$ $\cup \dots$ $\cup \{p^i \mathcal{X}_{14} \mid i = 0, 1, \dots, 5\}$
$J_{\text{max}}$	42	150
$\delta_x$	1/400	1/800
$\epsilon_b$	1/250	1/250
$\delta_b$	1/32	1/100
$B$	40	80
$m_{\text{max}}$	2	6
non-default SDPB parameters	--precision=768	--precision=840 --dualityGapThreshold=1e-80 --primalErrorThreshold=1e-80 --dualErrorThreshold=1e-80

**Table 2.** Choices of parameters for the computations in this work. These parameters were chosen experimentally — we have not attempted to optimize them. For the computation in figure 4, we verified that further increasing  $J_{\text{max}}$  changes the bounds by less than the fraction  $10^{-4}$ . Because of the refinement procedure described in section A.2, the bounds are essentially independent of  $\delta_x, \epsilon_b, \delta_b$  in all cases. For figure 4, the precise linear combinations of  $\mathcal{C}_2^{\text{improved}}$  are  $D$ -dependent, and listed in table 1. For the computation in figure 6, we supplemented the initial discretization of  $x$  with additional constraints at  $x \in \{1 - 4/J, 1 - 4/J + 1/(50J), \dots, 1 - 1/(50J)\}$  for  $J = 36, \dots, 150$ . Finally, for the extremal functional computation in figure 6, we project onto the subspace of functionals with vanishing contribution of  $g_3$ . This reduces the space of functionals from 93 dimensions to 92 dimensions. Tables of functionals were produced using Mathematica with 300 decimal digits of precision.

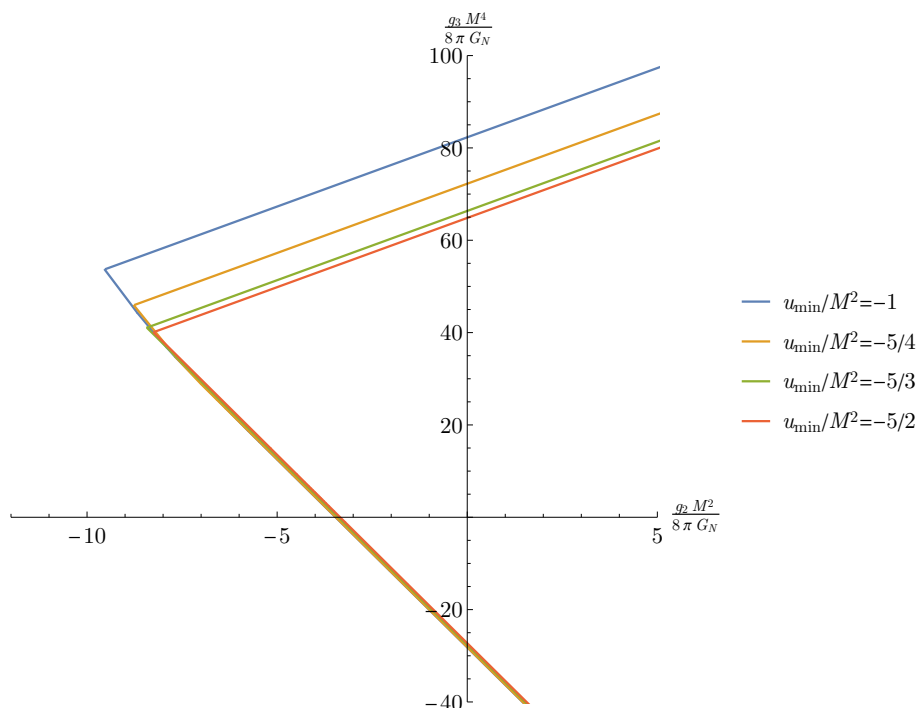
## B Bounds using an extended range of $u$

When  $u = -M^2$ , the  $t$ -channel cut merges with the origin in the  $s$ -plane. As noted in section 3, this may invalidate the  $\mathcal{C}_{2,u}^{\text{improved}}$  sum rule in general for  $u \leq -M^2$ . However, for meromorphic amplitudes, i.e. amplitudes where the  $t$ -channel “cut” is simply a collection of simple poles, there is no obvious problem with taking  $u \leq -M^2$ . Indeed, the low-energy contribution to the improved sum rules  $\mathcal{C}_{k,u}^{\text{improved}}$  and  $\mathcal{X}_{k,u}$  stays finite for general  $u < 0$ . This is unlike the situation for the unimproved sum rule  $\mathcal{C}_{k,u}$ , the left-hand side of which is an infinite sum of contact contributions with a finite radius of convergence. Similarly, the terms on the right-hand side of  $\mathcal{C}_{k,u}$  have a pole at  $u = -m^2$ , but this pole is removed in the improved sum rules. We have checked in examples coming from string theory, such as (3.21), that the improved sum rules continue to hold in the complete range  $u \in (-\infty, 0)$ .

$D = 5$	$g_2 - \frac{g_3}{3} + 60.3086 \geq 0$ $\wedge g_2 + 0.0647867g_3 + 9.64034 \geq 0$ $\wedge g_2 + 0.0750150g_3 + 8.40592 \geq 0$ $\wedge g_2 + 0.0779037g_3 + 8.09643 \geq 0$ $\wedge g_2 + 0.0802745g_3 + 7.86165 \geq 0$ $\wedge g_2 + 0.0823523g_3 + 7.66918 \geq 0$ $\wedge g_2 + 0.0842715g_3 + 7.50180 \geq 0$ $\wedge g_2 + 0.0861338g_3 + 7.34848 \geq 0$ $\wedge g_2 + 0.0879254g_3 + 7.20921 \geq 0$ $\wedge g_2 + 0.0898265g_3 + 7.07029 \geq 0$ $\wedge g_2 + 0.0919903g_3 + 6.92264 \geq 0$ $\wedge g_2 + 0.0946037g_3 + 6.75833 \geq 0$ $\wedge g_2 + 0.0980028g_3 + 6.56596 \geq 0$ $\wedge g_2 + 0.112285g_3 + 5.90160 \geq 0$ $\wedge g_2 + 0.112297g_3 + 5.90112 \geq 0$ $\wedge g_2 + 0.112318g_3 + 5.90066 \geq 0$ $\wedge g_2 + 0.112362g_3 + 5.90107 \geq 0$
$D = 6$	$g_2 - \frac{g_3}{3} + 28.0546 \geq 0$ $\wedge g_2 + 0.0940931g_3 + 4.55838 \geq 0$ $\wedge g_2 + 0.101479g_3 + 4.22726 \geq 0$ $\wedge g_2 + 0.107010g_3 + 4.00645 \geq 0$ $\wedge g_2 + 0.110538g_3 + 3.87630 \geq 0$ $\wedge g_2 + 0.113142g_3 + 3.78572 \geq 0$ $\wedge g_2 + 0.115195g_3 + 3.71765 \geq 0$ $\wedge g_2 + 0.116954g_3 + 3.66176 \geq 0$ $\wedge g_2 + 0.118572g_3 + 3.61228 \geq 0$ $\wedge g_2 + 0.120160g_3 + 3.56553 \geq 0$ $\wedge g_2 + 0.121815g_3 + 3.51866 \geq 0$
$D = 7$	$g_2 - \frac{g_3}{3} + 18.0717 \geq 0$ $\wedge g_2 + 0.107754g_3 + 3.30601 \geq 0$ $\wedge g_2 + 0.114745g_3 + 3.11426 \geq 0$ $\wedge g_2 + 0.119761g_3 + 2.98759 \geq 0$ $\wedge g_2 + 0.124098g_3 + 2.88546 \geq 0$ $\wedge g_2 + 0.127812g_3 + 2.80238 \geq 0$ $\wedge g_2 + 0.129850g_3 + 2.75853 \geq 0$
$D = 8$	$g_2 - \frac{g_3}{3} + 13.7186 \geq 0$ $\wedge g_2 + 0.110239g_3 + 2.92120 \geq 0$ $\wedge g_2 + 0.113486g_3 + 2.84526 \geq 0$ $\wedge g_2 + 0.116510g_3 + 2.77802 \geq 0$ $\wedge g_2 + 0.119462g_3 + 2.71525 \geq 0$ $\wedge g_2 + 0.122467g_3 + 2.65393 \geq 0$ $\wedge g_2 + 0.125690g_3 + 2.59077 \geq 0$ $\wedge g_2 + 0.129207g_3 + 2.52462 \geq 0$ $\wedge g_2 + 0.133224g_3 + 2.45244 \geq 0$ $\wedge g_2 + 0.134841g_3 + 2.42430 \geq 0$ $\wedge g_2 + 0.134849g_3 + 2.42422 \geq 0$ $\wedge g_2 + 0.134858g_3 + 2.42416 \geq 0$ $\wedge g_2 + 0.134874g_3 + 2.42429 \geq 0$
$D = 9$	$g_2 - \frac{g_3}{3} + 11.1250 \geq 0$ $\wedge g_2 + 0.134185g_3 + 2.24546 \geq 0$ $\wedge g_2 + 0.138540g_3 + 2.17837 \geq 0$ $\wedge g_2 + 0.138548g_3 + 2.17837 \geq 0$ $\wedge g_2 + 0.138567g_3 + 2.17856 \geq 0$
$D = 10$	$g_2 - \frac{g_3}{3} + 9.5208 \geq 0$ $\wedge g_2 + 0.138871g_3 + 2.05513 \geq 0$ $\wedge g_2 + 0.14150g_3 + 2.0196 \geq 0$
$D = 11$	$g_2 - \frac{g_3}{3} + 8.4687 \geq 0$ $\wedge g_2 + 0.143720g_3 + 1.91178 \geq 0$ $\wedge g_2 + 0.143729g_3 + 1.91169 \geq 0$ $\wedge g_2 + 0.143733g_3 + 1.91166 \geq 0$ $\wedge g_2 + 0.143738g_3 + 1.91165 \geq 0$ $\wedge g_2 + 0.143746g_3 + 1.91166 \geq 0$ $\wedge g_2 + 0.143772g_3 + 1.91187 \geq 0$
$D = 12$	$g_2 - \frac{g_3}{3} + 7.9016 \geq 0$ $\wedge g_2 + 0.144433g_3 + 1.86671 \geq 0$ $\wedge g_2 + 0.145825g_3 + 1.85029 \geq 0$

**Table 3.** The inequalities plotted in figure 4 (in units where  $8\pi G = 1$  and  $M = 1$ ).

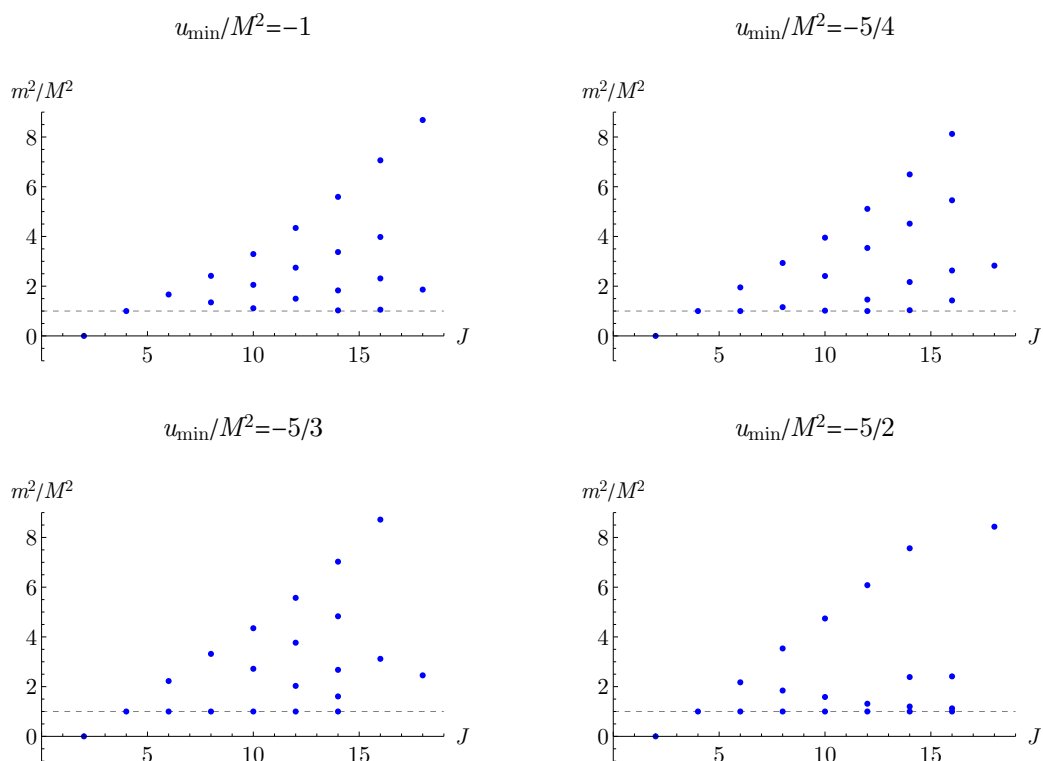




**Figure 9.** Bounds on  $g_2, g_3$  in  $D = 6$ , computed using wavepackets with support in  $u \in [u_{\min}, 0]$  for  $u_{\min}/M^2 = -1, -5/4, -5/3, -5/2$ . As  $u_{\min}$  gets more negative, the kink slides along the lower bound on  $g_3$ . These bounds were computed using a 54-dimensional space of functionals spanned by  $\{p^n \mathcal{C}_{2,-p^2}^{\text{improved}} \mid n = 1.5, \dots, 19.5\} \cup \{p^n \mathcal{X}_k \mid (n, k) \in \{(0, 4), \dots, (11, 4)\} \cup \{(0, 6), \dots, (9, 6)\} \cup \{(0, 8), \dots, (7, 8)\} \cup \{(0, 10), \dots, (5, 10)\}\}$ .

In general, we expect the improved sum rules to hold in the complete range if the exchanged spectrum contains only finitely many states below any given mass, as is the case in string theory. This is because any finite number of states can not affect the convergence of the improved sum rules and by removing all states with  $m^2 < -u$ , we effectively find ourselves in the situation  $u \in (-M^2, 0)$  where convergence is guaranteed.

Thus, it is interesting to consider how the bounds change when we use wavefunctions with larger support in  $u$ . We leave a more complete exploration of this idea to future work. In this appendix, we explore some example bounds obtained with a larger (but still compact) range of  $u$ . In figure 9, we show bounds on  $g_2, g_3$  obtained using wavefunctions with support on  $u \in [-yM^2, 0]$  for various  $y$ . As  $y$  increases, the minimal  $g_2$  “kink” slides to the right. Interestingly, the corresponding extremal spectra appear to simplify, see figure 10. For larger  $y$ , the spectrum may be converging toward a single linear Regge trajectory  $m^2 = M^2(J - 2)/2$ , with possibly another trajectory at  $m^2 = M^2$ . It is interesting to ask whether a solution to crossing symmetry with this structure can exist.



**Figure 10.** Extremal spectra for the bound that minimizes  $g_2$  in  $D = 6$ , computed using different wavefunctions with different ranges  $u \in [u_{\min}, 0]$ . As  $u_{\min}$  gets more negative, the spectrum appears to simplify, with approximately linear Regge trajectories at  $m^2 \sim M^2(J - 2)/2$  and  $m^2 = M^2$ .

**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] H. Ooguri and C. Vafa, *On the Geometry of the String Landscape and the Swampland*, *Nucl. Phys. B* **766** (2007) 21 [[hep-th/0605264](#)] [[INSPIRE](#)].
- [2] T.D. Brennan, F. Carta and C. Vafa, *The String Landscape, the Swampland, and the Missing Corner*, *PoS TASI2017* (2017) 015 [[arXiv:1711.00864](#)] [[INSPIRE](#)].
- [3] E. Palti, *The Swampland: Introduction and Review*, *Fortsch. Phys.* **67** (2019) 1900037 [[arXiv:1903.06239](#)] [[INSPIRE](#)].
- [4] M. van Beest, J. Calderón-Infante, D. Mirfendereski and I. Valenzuela, *Lectures on the Swampland Program in String Compactifications*, [arXiv:2102.01111](#) [[INSPIRE](#)].
- [5] N. Arkani-Hamed, L. Motl, A. Nicolis and C. Vafa, *The String landscape, black holes and gravity as the weakest force*, *JHEP* **06** (2007) 060 [[hep-th/0601001](#)] [[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] 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](#)].
- [8] N. Arkani-Hamed, T.-C. Huang and Y.-T. Huang, *The EFT-Hedron*, *JHEP* **05** (2021) 259 [[arXiv:2012.15849](#)] [[INSPIRE](#)].
- [9] B. Bellazzini, J. Elias Miró, R. Rattazzi, M. Riembau and F. Riva, *Positive Moments for Scattering Amplitudes*, [arXiv:2011.00037](#) [[INSPIRE](#)].
- [10] A.J. Tolley, Z.-Y. Wang and S.-Y. Zhou, *New positivity bounds from full crossing symmetry*, *JHEP* **05** (2021) 255 [[arXiv:2011.02400](#)] [[INSPIRE](#)].
- [11] S. Caron-Huot and V. Van Duong, *Extremal Effective Field Theories*, *JHEP* **05** (2021) 280 [[arXiv:2011.02957](#)] [[INSPIRE](#)].
- [12] A. Sinha and A. Zahed, *Crossing Symmetric Dispersion Relations in Quantum Field Theories*, *Phys. Rev. Lett.* **126** (2021) 181601 [[arXiv:2012.04877](#)] [[INSPIRE](#)].
- [13] M. Froissart, *Asymptotic behavior and subtractions in the Mandelstam representation*, *Phys. Rev.* **123** (1961) 1053 [[INSPIRE](#)].
- [14] A. Martin, *Unitarity and high-energy behavior of scattering amplitudes*, *Phys. Rev.* **129** (1963) 1432 [[INSPIRE](#)].
- [15] M. Gell-Mann, M.L. Goldberger and W.E. Thirring, *Use of causality conditions in quantum theory*, *Phys. Rev.* **95** (1954) 1612 [[INSPIRE](#)].
- [16] M.L. Goldberger, *Causality Conditions and Dispersion Relations. 1. Boson Fields*, *Phys. Rev.* **99** (1955) 979 [[INSPIRE](#)].
- [17] 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](#)].

- [18] 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](#)].
- [19] D. Chandorkar, S.D. Chowdhury, S. Kundu and S. Minwalla, *Bounds on Regge growth of flat space scattering from bounds on chaos*, *JHEP* **05** (2021) 143 [[arXiv:2102.03122](#)] [[INSPIRE](#)].
- [20] L. Susskind, *Holography in the flat space limit*, *AIP Conf. Proc.* **493** (1999) 98 [[hep-th/9901079](#)] [[INSPIRE](#)].
- [21] J. Polchinski, *S matrices from AdS space-time*, [hep-th/9901076](#) [[INSPIRE](#)].
- [22] J. Penedones, *Writing CFT correlation functions as AdS scattering amplitudes*, *JHEP* **03** (2011) 025 [[arXiv:1011.1485](#)] [[INSPIRE](#)].
- [23] S. Caron-Huot, *Analyticity in Spin in Conformal Theories*, *JHEP* **09** (2017) 078 [[arXiv:1703.00278](#)] [[INSPIRE](#)].
- [24] J. Maldacena, S.H. Shenker and D. Stanford, *A bound on chaos*, *JHEP* **08** (2016) 106 [[arXiv:1503.01409](#)] [[INSPIRE](#)].
- [25] 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](#)].
- [26] J. Tokuda, K. Aoki and S. Hirano, *Gravitational positivity bounds*, *JHEP* **11** (2020) 054 [[arXiv:2007.15009](#)] [[INSPIRE](#)].
- [27] 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](#)].
- [28] E. Pajer, D. Stefanyszyn and J. Supel, *The Boostless Bootstrap: Amplitudes without Lorentz boosts*, *JHEP* **12** (2020) 198 [[arXiv:2007.00027](#)] [[INSPIRE](#)].
- [29] T. Grall and S. Melville, *Positivity Bounds without Boosts*, [arXiv:2102.05683](#) [[INSPIRE](#)].
- [30] L. Alberte, C. de Rham, S. Jaitly and A.J. Tolley, *QED positivity bounds*, *Phys. Rev. D* **103** (2021) 125020 [[arXiv:2012.05798](#)] [[INSPIRE](#)].
- [31] S. Caron-Huot, D. Mazac, L. Rastelli and D. Simmons-Duffin, *AdS Bulk Locality from Sharp CFT Bounds*, [arXiv:2106.10274](#) [[INSPIRE](#)].
- [32] I. Heemskerck, J. Penedones, J. Polchinski and J. Sully, *Holography from Conformal Field Theory*, *JHEP* **10** (2009) 079 [[arXiv:0907.0151](#)] [[INSPIRE](#)].
- [33] D. Carmi and S. Caron-Huot, *A Conformal Dispersion Relation: Correlations from Absorption*, *JHEP* **09** (2020) 009 [[arXiv:1910.12123](#)] [[INSPIRE](#)].
- [34] D. Mazáč, L. Rastelli and X. Zhou, *A Basis of Analytic Functionals for CFTs in General Dimension*, [arXiv:1910.12855](#) [[INSPIRE](#)].
- [35] J. Penedones, J.A. Silva and A. Zhiboedov, *Nonperturbative Mellin Amplitudes: Existence, Properties, Applications*, *JHEP* **08** (2020) 031 [[arXiv:1912.11100](#)] [[INSPIRE](#)].
- [36] S. Caron-Huot, D. Mazac, L. Rastelli and D. Simmons-Duffin, *Dispersive CFT Sum Rules*, *JHEP* **05** (2021) 243 [[arXiv:2008.04931](#)] [[INSPIRE](#)].
- [37] R. Gopakumar, A. Sinha and A. Zahed, *Crossing Symmetric Dispersion Relations for Mellin Amplitudes*, *Phys. Rev. Lett.* **126** (2021) 211602 [[arXiv:2101.09017](#)] [[INSPIRE](#)].

- [38] D. Mazac, *Analytic bounds and emergence of AdS<sub>2</sub> physics from the conformal bootstrap*, *JHEP* **04** (2017) 146 [[arXiv:1611.10060](#)] [[INSPIRE](#)].
- [39] D. Mazac and M.F. Paulos, *The analytic functional bootstrap. Part I: 1D CFTs and 2D S-matrices*, *JHEP* **02** (2019) 162 [[arXiv:1803.10233](#)] [[INSPIRE](#)].
- [40] D. Mazac and M.F. Paulos, *The analytic functional bootstrap. Part II. Natural bases for the crossing equation*, *JHEP* **02** (2019) 163 [[arXiv:1811.10646](#)] [[INSPIRE](#)].
- [41] M.F. Paulos, *Analytic functional bootstrap for CFTs in  $d > 1$* , *JHEP* **04** (2020) 093 [[arXiv:1910.08563](#)] [[INSPIRE](#)].
- [42] D. Carmi, J. Penedones, J.A. Silva and A. Zhiboedov, *Applications of dispersive sum rules:  $\epsilon$ -expansion and holography*, *SciPost Phys.* **10** (2021) 145 [[arXiv:2009.13506](#)] [[INSPIRE](#)].
- [43] S.B. Giddings and M. Srednicki, *High-energy gravitational scattering and black hole resonances*, *Phys. Rev. D* **77** (2008) 085025 [[arXiv:0711.5012](#)] [[INSPIRE](#)].
- [44] M. Correia, A. Sever and A. Zhiboedov, *An Analytical Toolkit for the S-matrix Bootstrap*, [arXiv:2006.08221](#) [[INSPIRE](#)].
- [45] X. Li, C. Yang, H. Xu, C. Zhang and S.-Y. Zhou, *Positivity in Multi-Field EFTs*, [arXiv:2101.01191](#) [[INSPIRE](#)].
- [46] D. Simmons-Duffin, *A Semidefinite Program Solver for the Conformal Bootstrap*, *JHEP* **06** (2015) 174 [[arXiv:1502.02033](#)] [[INSPIRE](#)].
- [47] W. Landry and D. Simmons-Duffin, *Scaling the semidefinite program solver SDPB*, [arXiv:1909.09745](#) [[INSPIRE](#)].
- [48] D. Poland and D. Simmons-Duffin, *Bounds on 4D Conformal and Superconformal Field Theories*, *JHEP* **05** (2011) 017 [[arXiv:1009.2087](#)] [[INSPIRE](#)].
- [49] S. El-Showk and M.F. Paulos, *Bootstrapping Conformal Field Theories with the Extremal Functional Method*, *Phys. Rev. Lett.* **111** (2013) 241601 [[arXiv:1211.2810](#)] [[INSPIRE](#)].
- [50] J.H. Schwarz, *Superstring Theory*, *Phys. Rept.* **89** (1982) 223 [[INSPIRE](#)].
- [51] N. Arkani-Hamed, *Towards deriving string theory as the weakly coupled UV completion of gravity*, Talk at Strings 2016, Tsinghua University, Beijing, China (2016).
- [52] H. Kawai, D.C. Lewellen and S.H.H. Tye, *A Relation Between Tree Amplitudes of Closed and Open Strings*, *Nucl. Phys. B* **269** (1986) 1 [[INSPIRE](#)].
- [53] D. Amati, M. Ciafaloni and G. Veneziano, *Superstring Collisions at Planckian Energies*, *Phys. Lett. B* **197** (1987) 81 [[INSPIRE](#)].
- [54] A. Guerrieri, J. Penedones and P. Vieira, *Where is String Theory?*, [arXiv:2102.02847](#) [[INSPIRE](#)].
- [55] D. Poland, D. Simmons-Duffin and A. Vichi, *Carving Out the Space of 4D CFTs*, *JHEP* **05** (2012) 110 [[arXiv:1109.5176](#)] [[INSPIRE](#)].
- [56] F. Kos, D. Poland and D. Simmons-Duffin, *Bootstrapping Mixed Correlators in the 3D Ising Model*, *JHEP* **11** (2014) 109 [[arXiv:1406.4858](#)] [[INSPIRE](#)].
- [57] F.W.J. Olver et al. eds., *NIST Digital Library of Mathematical Functions*, release 1.1.0 of 2020-12-15, <http://dlmf.nist.gov/>.