

# The S-matrix bootstrap. Part III: higher dimensional amplitudes

---

Miguel F. Paulos,<sup>a</sup> Joao Penedones,<sup>b</sup> Jonathan Toledo,<sup>b</sup> Balt C. van Rees<sup>c</sup> and Pedro Vieira<sup>d,e</sup>

<sup>a</sup>*Theoretical Physics Department, CERN, Geneva, Switzerland*

<sup>b</sup>*Institute of Physics, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland*

<sup>c</sup>*Centre for Particle Theory, Department of Mathematical Sciences, Durham University, Lower Mountjoy, Stockton Road, Durham DH1 3LE, U.K.*

<sup>d</sup>*Perimeter Institute for Theoretical Physics, Waterloo, ON N2L 2Y5, Canada*

<sup>e</sup>*ICTP South American Institute for Fundamental Research, IFT-UNESP, São Paulo, SP 01440-070, Brazil*

*E-mail:* [miguel.paulos@gmail.com](mailto:miguel.paulos@gmail.com), [jpenedones@gmail.com](mailto:jpenedones@gmail.com),  
[jonathan.campbell.toledo@gmail.com](mailto:jonathan.campbell.toledo@gmail.com), [baltvanrees@gmail.com](mailto:baltvanrees@gmail.com),  
[pedrogvieira@gmail.com](mailto:pedrogvieira@gmail.com)

**ABSTRACT:** We consider constraints on the S-matrix of any gapped, Lorentz invariant quantum field theory in 3+1 dimensions due to crossing symmetry, analyticity and unitarity. We extremize cubic couplings, quartic couplings and scattering lengths relevant for the elastic scattering amplitude of two identical scalar particles. In the cases where our results can be compared with the older S-matrix literature they are in excellent agreement. We also extremize a cubic coupling in 2+1 dimensions which we can directly compare to a universal bound for a QFT in AdS. This paper generalizes our previous 1+1 dimensional results of [1] and [2].

**KEYWORDS:** Scattering Amplitudes, Nonperturbative Effects, Space-Time Symmetries

**ARXIV EPRINT:** [1708.06765](https://arxiv.org/abs/1708.06765)

---

**Contents**

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Two dimensions redux and unit circles</b>	<b>3</b>
<b>3</b>	<b>Higher dimensions</b>	<b>7</b>
3.1	Ansatz	9
<b>4</b>	<b>Results</b>	<b>11</b>
4.1	Cubic coupling	11
4.2	Quartic coupling	13
4.3	Exploring scattering lengths	16
4.4	Bonus feature: three spacetime dimensions and QFT in AdS	19
4.4.1	S-matrix bootstrap approach	19
4.4.2	QFT in AdS approach	19
4.4.3	Results	20
<b>5</b>	<b>Discussion</b>	<b>21</b>
<b>A</b>	<b><math>x(s)</math> vs <math>\rho_s, \rho_t</math> in 1 + 1 dimensions</b>	<b>23</b>
<b>B</b>	<b>Constraint surface in <math>\rho</math>-coordinates</b>	<b>24</b>
<b>C</b>	<b>Mandelstam representation</b>	<b>25</b>
<b>D</b>	<b>Partial wave integrals</b>	<b>26</b>
D.1	Pole contributions	26
D.2	$\rho_s \rho_t \rho_u$ contributions	28
D.3	Large energy	29
D.4	Large spin	31
D.5	Threshold expansion and elastic unitarity	33
<b>E</b>	<b>Non-relativistic limit</b>	<b>34</b>
<b>F</b>	<b>Semidefinite programming implementation</b>	<b>35</b>
<b>G</b>	<b>Slow convergence on a simple 2D example</b>	<b>37</b>

---

## 1 Introduction

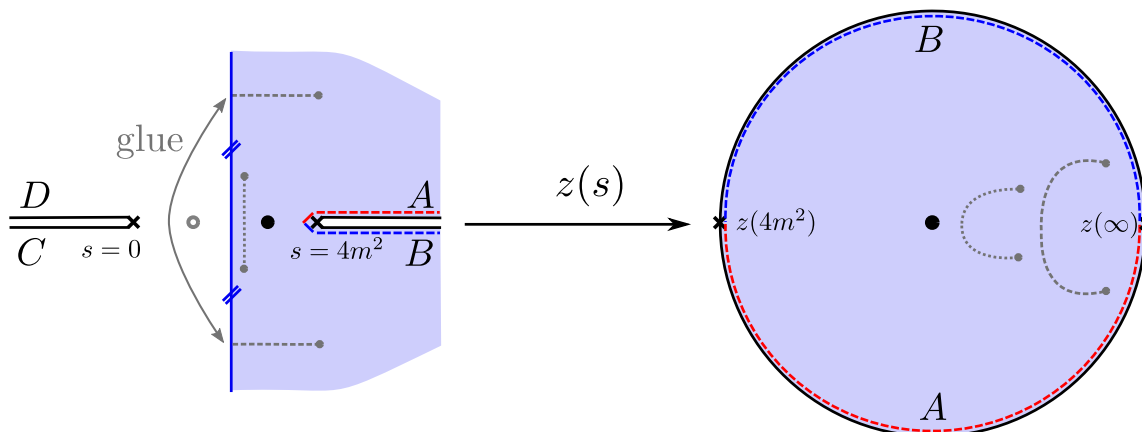
In [1] and [2] we initiated a bootstrap analysis of massive quantum field theories. In particular, we obtained bounds on couplings of a quantum field theory compatible with a given spectrum of stable particles.

Physically, one expects such bounds to exist since increasing the interaction strength will typically increase the attraction between particles. As such, we expect to have maximum values for couplings beyond which the masses of bound states must decrease, or new bound-states should emerge from the continuum, or both.

Mathematically, this problem is also very natural once we make the non-trivial assumption that scattering amplitudes are described by functions that are analytic away from the usual physical poles and cuts. The point is that analytic functions always attain their maximum at a boundary of their domain of definition. In the context of scattering amplitudes, these boundaries are the cuts generated by multiparticle intermediate states. For *physical kinematics* the amplitude along the cut is constrained by the conditions that probabilities add up to one — i.e. by unitarity. For this reason we focus on the two body scattering of the lightest particle in the theory since then all the usual cuts of the amplitude correspond to physical kinematics. In 1+1 dimensions where unitarity can be directly applied at the level of the S-matrix (simply,  $|S(s)| \leq 1$  for  $s$  along the cuts) we are faced with a clean problem in the theory of complex functions of a single variable. As we have an analytic function on a domain with a boundary along which it is bounded, so we are able to constrain its values inside this region and in particular the various physical couplings which we define as residues of factorization poles. Section 2 contains a derivation of the two dimensional bound which is a significant refinement of that in [2].

In this paper we move the focus to higher dimensions which contains a plethora of very interesting and difficult elements absent in the simpler 1+1 dimensional case. An essential difference is that the most convenient way to formulate unitarity requires introducing *partial waves* and these are not bounded by unitarity along their entire boundary (only along the so-called “right cut”). Therefore the simple complex analysis argument of 1+1 dimensions cannot directly apply. Furthermore, the analyticity and crossing symmetry requirements involve the *amplitudes* rather than the partial waves, which forces one to use both descriptions of the scattering event. Still, it is possible to overcome these technical obstacles. We shall introduce a kind of uniformization coordinates where the full space of physical kinematics is mapped to (a few) unit circles. This will allow us to Taylor expand the amplitudes in a convergent and manifestly crossing symmetric way in the full physical plane and then to numerically impose unitarity along the physical boundaries.

We start by revisiting the two dimensional results with this new approach in section 2, setup the higher dimensional problem in section 3 and present and analyze the corresponding numerical results in section 4. In section 4.4 we compare our numerical results with the completely orthogonal approach of [1] which is based on QFT in AdS and in particular does not require any analyticity assumptions. We conclude in section 5. A number of appendices are included to complement the main text presentation.



**Figure 1.** Mapping from the cut  $s$ -plane to the unit disk given in equation (2.1). The mapping associates the points  $z(2 + iy) = z(2 - iy)$  and maps the half plane  $\text{Re}(s) > 2$  to the full unit disk. The grey, dashed curves on the left map to those on the right and are included to help the reader visualize the mapping.

## 2 Two dimensions redux and unit circles

In this section we revisit the much simpler two dimensional problem. In two dimensions we can solve things analytically, and so it is a great training ground for developing intuition and testing any new numerical approaches. Nonetheless, for the braver readers eager to learn about the higher dimensional story, this section can be skipped without compromising the logic of the paper.

Most of the mathematical analysis of [2] boils down to minor variations of the following simple problem:

**Q:** Consider all real analytic functions  $f(z) = [f(z^*)]^*$  with no singularities inside the unit disk apart from a simple pole at  $z = 0$  and which are bounded on the unit circle as  $|f(e^{i\phi})| \leq 1$ .<sup>1</sup> What is the maximum possible residue at  $z = 0$  and which function has that residue?

**A:** The maximum residue is 1 and the corresponding function is  $f = 1/z$ .

Indeed  $g(z) = f(z)/(1/z)$  has no singularities inside the disk and obeys  $|g(z)| \leq 1$  at the boundary of the unit disk. By the so-called maximum modulus principle, it satisfies  $|g(z)| \leq 1$  *everywhere* inside the disk. Its value at the origin — which is nothing but the residue of  $f$  — is therefore at most 1. This maximum value is attained when  $g$  is constant everywhere, that is when  $g(z) = 1$  corresponding to  $f(z) = 1/z$ .

To see how this simple problem relates to the analysis in [2, 3] consider the  $2 \rightarrow 2$  S-matrix  $S(s)$  for scattering of identical neutral particles of mass  $m$  considered in [2]. Assume also that there is a single bound-state showing up in this S-matrix element and for simplicity assume its mass  $m_b > \sqrt{2}m$ . Because of crossing symmetry  $S(s) = S(4m^2 - s)$

<sup>1</sup>In addition,  $f(z)$  should not have an essential singularity at the boundary of the disk such that  $|f(z)|$  diverges as we approach the boundary from any direction inside the disk.

and we can focus on the region  $\text{Re}(s) > 2m^2$  without any loss of generality. In this half plane we have a threshold cut starting at  $s = 4m^2$ , the bound-state pole at  $s = m_b^2$  and no other singularities. Consider then the change of variable

$$z = \frac{\sqrt{s(4m^2 - s)} - m_b\sqrt{4m^2 - m_b^2}}{\sqrt{s(4m^2 - s)} + m_b\sqrt{4m^2 - m_b^2}} \quad \text{or} \quad \left(\frac{s - 2m^2}{2m^2}\right)^2 = 1 + \frac{m_b^2}{m^2}\left(\frac{m_b^2}{4m^2} - 1\right)\left(\frac{z + 1}{z - 1}\right)^2 \tag{2.1}$$

which maps this half plane into the unit disk, the bound-state pole into the origin of that disk and finally the threshold cut — where unitarity is to be imposed — to the boundary of the disk, see figure 1. In terms of  $z$  the S-matrix is therefore exactly constrained by the conditions of the previous point; it has a pole at  $z = 0$  and obeys  $|S(z)| \leq 1$  at the boundary of the disk.<sup>2</sup> Its maximum residue — which is where we measure the (square of the) coupling to the bound-state — is therefore 1 and the corresponding optimal S-matrix is therefore  $S(z) = 1/z$ .

To recover the results of [2] — see e.g. formula (36) therein — we simply need to take into account the Jacobian to go from  $z$  to  $s$ , the simple kinematical multiplicative factors relating the S-matrix and the T-matrix and a factor of  $m^4$  to render the coupling dimensionless. All other results of [2] for more complicated bound-state spectra can be treated through simple generalizations of this simple example!<sup>3</sup>

Although redundant at this point, it is instructive for what will come next in higher dimensions to set up this exactly solvable problem numerically. We define a function  $S(z)$  in the unit circle as a pole plus a convergent Taylor expansion which we truncate at some large power  $z^M$ . Then we simply maximize the residue with the constraint that in a tightly spaced grid of  $K$  points on the unit circle unitarity is satisfied. In `Mathematica`, the simple code below does the job:

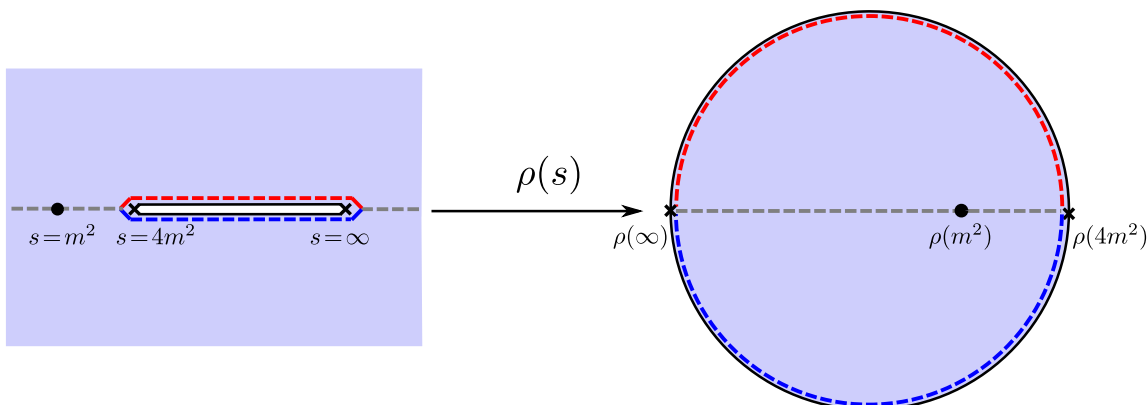
```
M=20; K=50;
S[z_] = residue/z + Sum[c[n] z^n, {n, 0, M}];
variables = {residue}~Join~Table[c[n], {n, 0, M}];
constraints = Table[S[Exp[I x]] S[Exp[-I x]] <= 1, {x, 0, \[Pi], \[Pi]/K}];
FindMaximum[{residue, constraints}, variables]
```

This nicely yields  $residue \simeq 1$  and  $c_n \simeq 0$  with great numerical accuracy which can be always improved. The reader is encouraged to copy/paste this and try by him/herself. It should take about 2 or 3 seconds to run.

As a last warm-up it is very useful to solve this very same problem in a third way since this last approach is the closest to what we will do in higher dimensions. In this last

<sup>2</sup>Note that this condition also holds on the lower half of the disk due to real analyticity.

<sup>3</sup>Strictly speaking the map to the unit circle is not even needed here. It suffices to assume there is no essential singularity at infinity so that the unitarity cut is the boundary of the region where  $S(s)$  takes values. Then  $S(s)/z(s)^{-1}$  is free of singularities in the physical region and obeys  $|S(s)| \leq 1$  on the cuts which are the boundaries of this region. Hence it can at most be one inside by the maximum modulus principle and the bound on the residue of  $S$  follows. This is the argument in [3]. We still found the unit circle discussion to be useful as a warm-up to the higher dimensional case.



**Figure 2.** Mapping from the cut  $s$ -plane to the unit disk given in equation (2.2).

approach to the problem we start by thinking of the S-matrix as being a function of both  $s$  and  $t$  as if they were independent variables; they are not since  $s + t + u = 4m^2$  and  $u = 0$  in two dimensions.<sup>4</sup> Then  $S(s, t)$  is a function with a cut for  $s > 4m^2$ , another cut for  $t > 4m^2$  as well as poles for single-particle processes in the  $s$ - and  $t$ -channels. Next we use a very convenient change of variable which maps the full complex plane with those cuts removed into the unit disk. This is the map

$$s \mapsto \rho_s = \frac{\sqrt{4m^2 - s_0} - \sqrt{4m^2 - s}}{\sqrt{4m^2 - s_0} + \sqrt{4m^2 - s}}, \quad s = \frac{s_0(1 - \rho_s)^2 + 16m^2\rho_s}{(1 + \rho_s)^2}. \quad (2.2)$$

where  $s_0 < 4m^2$  is a free parameter that we can choose according to convenience. In the present case, it is convenient to choose  $s_0 = 2m^2$  so that  $\rho_s = 0$  corresponds to the crossing symmetric point  $s = t = 2m^2$ . A similar map is also very useful in conformal bootstrap studies [4]. It is illustrated in figure 2. The top of the cut maps to the upper boundary of the unit disk and the bottom of the cut maps to the lower boundary of the disk. The interval  $[0, 4m^2]$  maps to the interval  $\rho \in [2\sqrt{2} - 3, 1]$  so this is where we find the poles associated to stable particles.

Apart from the poles corresponding to single particle exchanges,  $S(\rho_s, \rho_t)$  is analytic for both  $\rho_s$  and  $\rho_t$  inside the unit disk and thus we can write

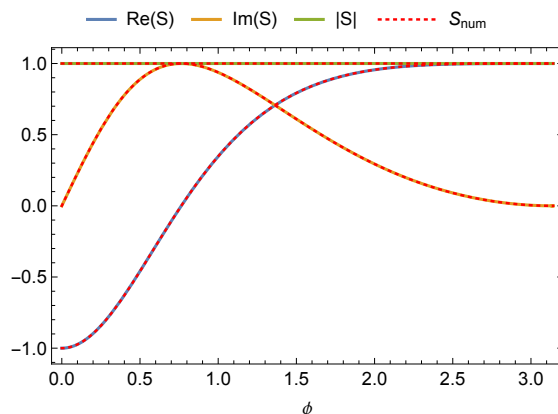
$$S(s, t) = -\frac{\hat{g}^2}{s - m_b^2} - \frac{\hat{g}^2}{t - m_b^2} + \sum_{a,b=0}^{\infty} c_{ab} \rho_s^a \rho_t^b \quad (2.3)$$

Crossing symmetry is guaranteed provided the coefficients of the convergent Taylor expansion are symmetric,  $c_{ab} = c_{ba}$ . Since we are going to evaluate the S-matrix on the constraint surface  $s + t = 4m^2$  we can simplify this ansatz further. In terms of  $\rho_s$  and  $\rho_t$  this constraint yields

$$\rho_s^2 \rho_t + \rho_t^2 \rho_s + 4\rho_s \rho_t + \rho_s + \rho_t = 0 \quad (2.4)$$

This means the representation (2.3) has a big redundancy. We can always add to it polynomials in the left hand side of the constraint (2.4). To remove this ambiguity, we can set

<sup>4</sup>More precisely, either  $u = 0$  or  $t = 0$  corresponding to backward and forward scattering.



**Figure 3.** Comparison of the exact optimal S-matrix (given by  $1/z(s)$  with  $z$  given by (2.1)) to numerical results using the ansatz (2.3) with the  $a, b$  series truncated at maximum degree  $N = 5$  and  $m_{BS}^2 = 3$ . We plot the physical region  $\rho = e^{i\phi}$  with  $\phi \in [0, \pi]$ . The numerical results (red dashed) are indistinguishable from the exact results.

to zero many of constants  $c_{ab}$  (in appendix B we explained in detail which  $c_{ab}$  can be set to zero).

Numerically, we set a cut-off in the sum (2.3) and impose unitarity for  $s > 4$  which corresponds to the upper half circle where  $\rho_s = e^{i\phi}$  with  $\phi \in [0, \pi]$ . We evaluate  $|S(s, t)|^2$  in a uniform grid in the  $\phi$  interval which gives a set of quadratic constraint equations on the  $c_{ab}$  and the residues of the poles. We optimize  $\hat{g}^2$  in the usual way using FindMaximum for example. The outcome of this third approach is in perfect agreement with our previous analytical and numerical results as illustrated in figure 3.

To summarize: in two dimensions we can find the optimal S-matrix with largest possible residue analytically.<sup>5</sup> We do so by dividing the S-matrix by a clever guess and using the maximum modulus principle to show that this ratio should be one. We recovered the same analytic results numerically in two ways. In the first one we start from a parametrization of the kinematics where we can Taylor expand the S-matrix and then truncate that expansion to obtain a finite algebraic problem which we can put on a computer. The second numerical approach is a small variation where we think of the S-matrix as a function of  $s$  and  $t$  as if they were independent and then consider a double Taylor expansion in each of them.

What we implicitly used in the last method can be called an analytic *extension* — note that it is *not* an analytic *continuation* as we are increasing the number of variables and not just moving into the complex plane keeping the number of variables fixed. In this

<sup>5</sup>Notice that if we allow essential singularities at  $s = \infty$  then there is no upper bound on  $\hat{g}^2$ . To see that consider the ansatz

$$S(s, t) = -\frac{\hat{g}^2}{s - m_b^2} \exp \left[ -\left( \frac{s - m_b^2}{\delta m^2} \right)^{2n} \right] + (s \leftrightarrow t), \tag{2.5}$$

where  $\delta m^2 = \frac{1}{2} \min(4m^2 - m_b^2, m_b^2)$ . For any value of  $\hat{g}^2$ , we can find a (large) positive integer  $n$  such that this ansatz satisfies the unitarity constraint  $|S(s, 4m^2 - s)| \leq 1$  for  $s > 4m^2$ . We thank Etienne Granet for raising this point. We exclude such essential singularities at  $s = \infty$  because they are incompatible with causality (see for instance appendix D of [5]).

extension we promoted the S-matrix to a more general function of two variables which has no singularities in the cut  $s$  and  $t$  planes.<sup>6</sup> Equivalently, in terms of the  $\rho$  variables, we assumed the existence of an extension into a function  $S(\rho_s, \rho_t)$  which has no singularities in the polydisk  $\{\rho_s, \rho_t \text{ such that } |\rho_s| \leq 1 \text{ and } |\rho_t| \leq 1\}$  while all we know a priori is that such a regular function exists only in the intersection of the polydisk with the constraint (2.4). Why do we have the right to assume that such an extension exists at all? For instance, it could happen that such an extension would inevitably introduce new singularities in the full polydisk domain which would then invalidate the convergence of the double expansion (2.3). Numerically, using this extension method we seem to find perfect agreement with the analytic results so somehow we should be safe. Indeed, the polydisk is a so-called Stein manifold<sup>7</sup> and the constraint (2.4) is an holomorphic embedding and as such defines a submanifold inside the polydisk which is also Stein. As discussed in greater detail below, there is a rather remarkable mathematical result which states that regular analytic extensions from Stein sub-manifolds inside Stein manifolds to the full Stein manifold do exist! The perfect numerical agreement is thus to be expected.

Of course, in two dimensions this discussion is a clear use of excessive force. On the other hand, in higher dimensions we will also make use of such analytic extensions and there we will not have the luxury of the analytic results to cross-check our numerics. The theorem alluded to above generalizes to that case as well and is key in providing confidence for the higher dimensional numerics.

There is also another more pedestrian explanation of why the double Taylor expansion numerics had to work which we present in appendix A; however, contrary to the discussion above, it makes use of particular features of the two dimensional problem and is not that useful as a warm up to the higher dimensional case.

### 3 Higher dimensions

We now move on to scattering amplitudes in  $d+1$  spacetime dimensions. Consider again the elastic scattering process of two identical real scalar particles of mass  $m$ . In our conventions the S-matrix element is

$$\langle \mathbf{p}_3, \mathbf{p}_4 | S | \mathbf{p}_1, \mathbf{p}_2 \rangle = \mathbb{1} + i(2\pi)^{d+1} \delta^{(d+1)}(p_1 + p_2 - p_3 - p_4) M(s, t, u) \quad (3.1)$$

with normalization such that

$$\mathbb{1} = (2\pi)^{2d} 4E_{\mathbf{p}_1} E_{\mathbf{p}_2} \left( \delta^{(d)}(\mathbf{p}_1 - \mathbf{p}_3) \delta^{(d)}(\mathbf{p}_2 - \mathbf{p}_4) + (3 \leftrightarrow 4) \right) \quad (3.2)$$

where  $E_{\mathbf{p}} = \sqrt{m^2 + \mathbf{p}^2}$ . The Mandelstam invariants are given by

$$s = (p_1 + p_2)^2 \quad t = (p_1 - p_3)^2 \quad u = (p_1 - p_4)^2 \quad (3.3)$$

---

<sup>6</sup>Of course we still have the poles associated with stable particles but these can be easily treated separately as in (2.3). Here, we focus on the parametrization of the analytic part of the S-matrix.

<sup>7</sup>The unit disk is an open Riemann surface and those are Stein manifolds. Products of Stein manifolds are also Stein so the polydisk is also Stein.



which of course obey  $s + t + u = 4m^2$ , and we henceforth work in units such that  $m^2 = 1$ . We often write  $M(s, t) \equiv M(s, t, 4 - s - t)$ . In the channel under consideration  $s$  is the squared center-of-mass energy and the scattering angle is given by

$$x = \cos(\theta) = 1 + \frac{2t}{s-4} = -1 - \frac{2u}{s-4} \tag{3.4}$$

Physical values of the Mandelstam invariants are therefore  $4 \leq s$  and  $4 - s \leq t \leq 0$ . We can project onto channels with definite angular momentum by introducing the *partial amplitudes*:

$$S_\ell(s) = 1 + i \frac{(s-4)^{\frac{d-2}{2}}}{\sqrt{s}} \int_{-1}^1 dx (1-x^2)^{\frac{d-3}{2}} P_\ell^{(d)}(x) M(s, t)|_{t \rightarrow \frac{1}{2}(s-4)(x-1)} \tag{3.5}$$

where  $P_\ell^{(d)}(x)$  is proportional<sup>8</sup> to the Gegenbauer polynomials. In our conventions,

$$P_\ell^{(3)}(x) = \frac{1}{32\pi} P_\ell(x), \quad P_\ell^{(2)}(x) = \frac{1}{8\pi} \cos(\ell\theta), \tag{3.6}$$

with  $P_\ell(x)$  the usual Legendre polynomials, normalized such that  $P_\ell(1) = 1$ . We note that  $S_\ell(s) = 1$  for odd  $\ell$  because Bose symmetry implies invariance under the reflection  $\theta \rightarrow \pi - \theta$ .

Although the S-matrix element (3.1) has all kind of distributional properties, the amplitude  $M(s, t, u)$  is a regular function (see e.g. [6, section 4.3]). We will assume that  $M(s, t, u)$  obeys three further constraints:

- **Crossing symmetry:**  $M(s, t, u)$  is completely symmetric in its arguments. The symmetry  $u \leftrightarrow t$  follows from the aforementioned Bose symmetry, but the other generator of the crossing symmetry group can only be found from a more sophisticated analysis and requires the LSZ prescription.
- **Analyticity:**  $M(s, t, 4 - s - t)$  is analytic for arbitrary complex  $s$  and  $t$ , except for potential bound-state poles at  $s = m_b^2$  with  $0 < m_b^2 < 4$ , a cut along the real axis starting at  $s = 4$ , and the images of these singularities under the crossing symmetry transformations. It further obeys the usual reality condition  $M(s^*, t^* 4 - s^* - t^*) = M^*(s, t, 4 - s - t)$ . We note that the analyticity assumption is actually rather optimistic, since this ‘maximal’ analyticity has not been proven from axiomatic field theory.<sup>9</sup> On the other hand some a posteriori justification is provided

---

<sup>8</sup>In general spacetime dimension, we have

$$P_\ell^{(d)}(x) = \frac{l! \Gamma(\frac{d-2}{2})}{4(4\pi)^{\frac{d}{2}} \Gamma(d+l-2)} C_\ell^{(d-2)/2}(x).$$

<sup>9</sup>Certain analyticity properties are known to be valid very generally, derived either to all orders in perturbation theory or from axiomatic field theory; the latter case sometimes requires the Wightman axioms and other times merely requires the validity of the LSZ prescription and causality. Typically one can prove two-variable analyticity for all  $s$  (modulo the known poles and cuts) but only for some finite range of values of  $t$  or of  $x$  which in particular includes the physical values. A standard result is that the proven analyticity is sufficient to analytically continue the amplitude from the  $s$ -channel to the  $t$  or  $u$  channels, establishing crossing symmetry [7]. We refer to [8, 9] and references therein for more extensive discussions.

by the remarkable agreement between some of our results and those obtained without maximal analyticity in the older literature. We therefore believe that this assumption is sufficiently mild to generate physically meaningful results. We offer some further comments on this point in section 4.4 and the conclusions section below.

- **Unitarity:** from  $S^\dagger S = 1$  we find that the unitarity constraint for elastic scattering takes the form

$$|S_\ell(s)| \leq 1 \tag{3.7}$$

for all  $s \geq 4$  and  $\ell \in \{0, 2, 4, \dots\}$ . Generically no other channels are available for a finite window of values of  $s$ , starting at 4 and ending at a higher threshold (like  $s = 9$  for three-particle scattering). In such a window the above inequality should in fact be saturated. In this work we will not impose such saturation, but our numerics in principle allows for it.

The aim of the S-matrix bootstrap program (as we envisage it) is to use these general conditions to obtain concrete constraints on the behavior of the function  $M(s, t, u)$  or the partial amplitudes  $S_\ell(s)$  at interesting points. Many results from the previous century can be found in the textbook [10] and the reviews [8, 11].

The recent works [12, 13] pursue a bootstrap analysis of scattering amplitudes of weakly interacting higher spin theories, where the amplitudes are meromorphic functions of the Mandelstam invariants. Analytically, they beautifully explore the large  $s$  and  $t$  regime of weakly interacting higher spin scattering amplitudes and observe remarkable universality there. In contrast, our analysis is fully non-perturbative and the only poles of the scattering amplitudes are associated with stable particles (below the 2-particle continuum). Nevertheless it would be very interesting to investigate the same large  $s$  and  $t$  regime within our numerical approach.

### 3.1 Ansatz

In this subsection we explore the consequences of our analyticity assumption in some detail. As a toy model we can start with a single-variable function  $f(z)$  which is analytic in a simple domain  $D \subset \mathbb{C}$ . If we define  $\rho : D \rightarrow \Delta$  as a biholomorphic map between  $D$  and the unit disk  $\Delta = \{\rho \in \mathbb{C} : |\rho| < 1\}$ , then any such  $f(z)$  has a Taylor series expansion of the form

$$f(z) = \sum_{n=0}^{\infty} c_n \rho(z)^n \tag{3.8}$$

which converges as long as  $|\rho(z)| < 1$ . Our multi-variable problem is unfortunately not so easy, since for  $M(s, t)$  the moving cuts imply that the domain of analyticity in one variable, say  $s$ , depends on the other variable  $t$ . We will remedy this as follows. First we relax the constraint  $s+t+u = 4$  and consider three-variable functions  $M(s, t, u)$ . Then we transform the variables  $(s, t, u) \rightarrow (\rho_s, \rho_t, \rho_u)$  using the map (2.2) which is, with  $m^2 = 1$ ,

$$s \mapsto \rho_s = \frac{\sqrt{4-s_0} - \sqrt{4-s}}{\sqrt{4-s_0} + \sqrt{4-s}}, \quad s = \frac{s_0(1-\rho_s)^2 + 16\rho_s}{(1+\rho_s)^2}. \tag{3.9}$$

In this case, it is convenient to choose  $s_0 = \frac{4}{3}$  so that  $\rho_s = \rho_t = \rho_u = 0$  corresponds to the crossing symmetric point  $s = t = u = \frac{4}{3}$ . Now, since the transformation  $\rho_s$  maps the  $s$ -plane minus the right cut starting at  $s = 4$  to the unit disk, we see that in the  $\rho$  variables all the cuts lie *outside* the polydisk  $\Delta^3$  defined by  $|\rho_s| < 1$ ,  $|\rho_t| < 1$  and  $|\rho_u| < 1$ . The only remaining singularities are then the poles and it is natural to write

$$M(s, t, u) = -\frac{g^2}{s - m_b^2} - \frac{g^2}{t - m_b^2} - \frac{g^2}{u - m_b^2} + \sum_{a,b,c=0} \alpha_{abc} \rho_s^a \rho_t^b \rho_u^c \quad (3.10)$$

where the triple  $\rho$  series converges inside  $\Delta^3$ , and for definiteness we have put in the poles for a single scalar bound state of mass  $m_b$ . The demands of crossing symmetry are implemented by demanding that the coefficients  $\alpha_{abc}$  are totally symmetric in their indices. When restricted to the surface defined by  $s+t+u = 4$  the ansatz (3.10) obeys the analyticity and crossing symmetry constraints. It is perhaps more surprising that the converse is also true: any function obeying the analyticity constraints on the surface  $s + t + u = 4$  can be extended to a function on  $\Delta^3$ , analytic modulo the poles, and therefore can be written in the form (3.10). This follows from a mathematical theorem known as Cartan’s theorem B, which is a statement about the vanishing of higher cohomologies of coherent analytic sheaves on Stein manifolds (see e.g. [14]) — in the case at hand this implies that there is no obstruction to an extension away from the surface  $s + t + u = 4$ .<sup>10</sup>

The triple  $\rho$  expansion in equation (3.10) is the starting point for our numerical work. Our approach is to restrict the expansion to a finite sum by imposing

$$a + b + c \leq N_{\max} \quad (3.11)$$

and then further restricting to the constraint surface  $s + t + u = 4$  which is given by a polynomial equation

$$\rho_s^2 \rho_t^2 \rho_u + \rho_s^2 \rho_u^2 \rho_t + \rho_t^2 \rho_u^2 \rho_s + (\text{lower degree terms}) = 0 \quad (3.12)$$

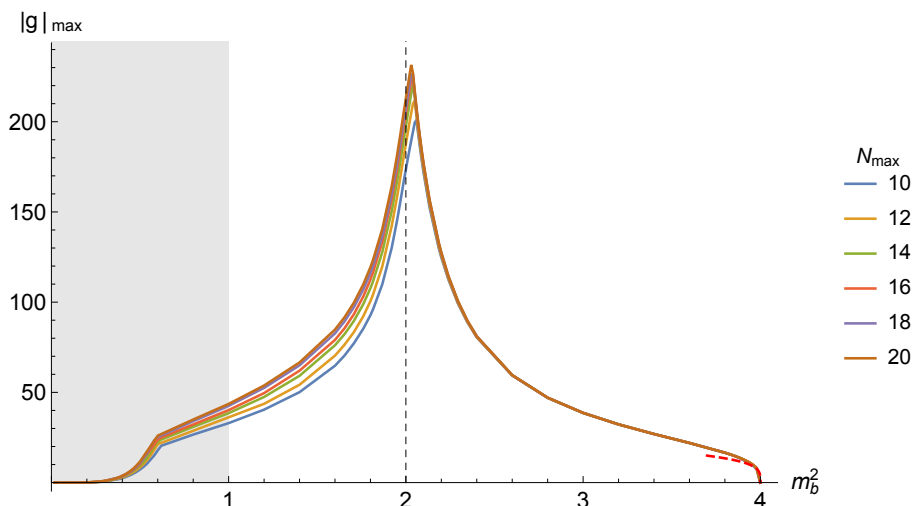
and which in practice allows us to eliminate many terms in (3.10) (in appendix B we explain in detail which terms can be set to zero). The remaining freedom in our ansatz then consists of the finitely many remaining  $\alpha_{abc}$  together with the bound state parameters; since this is a finite-dimensional space we can use a computer to numerically explore the space of scattering amplitudes. Of course we want to keep  $N_{\max}$  as large as possible. As we will see, in fortunate cases the numerical results stabilize already for feasible values of  $N_{\max}$ , while in other cases we can extrapolate.<sup>11</sup>

It will be the job of the computer to impose the unitarity constraints, which are quadratic constraints in the parameters  $g^2$  and  $\alpha_{abc}$ . Rather than checking the infinity

---

<sup>10</sup>In contrast to the Mandelstam representation, notice that our ansatz (3.10) ‘solves’ the constraints of analyticity and crossing symmetry without demanding specific asymptotic behavior for large values of the Mandelstam invariants. We offer more comments on the relation between our ansatz and the Mandelstam representation in appendix C.

<sup>11</sup>As discussed further in appendix D.3, the unitarity constraints imply that the large energy behavior is somewhat restricted if we keep  $N_{\max}$  finite, but we do not expect this to affect the physics in our results.



**Figure 4.** Largest possible value  $|g|_{\max}$  as a function of  $m_b^2$ , using a triple rho expansion of the amplitude for the given values of  $N_{\max}$  and after imposing the unitarity constraints for spins up to  $\ell_{\max} = 20$ . As explained in the text, the shaded area is physically incompatible with our analyticity assumption. We added the analytic result of appendix E as the dashed line near  $m_b^2 = 4$ .

of constraints for all  $s$  and  $\ell$ , we impose a cutoff and check that unitarity constraints are obeyed only for  $\ell \leq \ell_{\max}$  and along a grid of values for  $s$ . Experimentally we observe that our results remain meaningful if  $\ell_{\max}$  is not much smaller than  $N_{\max}$  and if the grid is sufficiently refined. In appendix F we discuss the dependence on these parameters in more detail, and outline the numerical implementation.

## 4 Results

In this section we present our numerical results for several maximization problems using the S-matrix bootstrap method explained above. For most of this section we restrict our attention to 3+1 dimensional QFTs, i.e.  $d = 3$  in our notation. In the final subsection 4.4, we consider 2 + 1 dimensional QFTs.

### 4.1 Cubic coupling

For our first result we consider a scattering amplitude with a single pole corresponding to the exchange of a scalar particle of mass  $m_b$ , exactly as in our ansatz (3.10), and maximize the value of the residue  $g^2$  as a function of  $m_b$ .<sup>12</sup>

In figure 4 we plot the maximum absolute value of the coupling  $|g|$  defined as the residue of the pole, with the different curves corresponding to different values of  $N_{\max}$ . We have obtained this plot by maximizing  $|g|$  for a sequence of values of  $m_b$  and the indicated curve is an interpolation through our data points. The plot is rather rich; we discuss its key features one by one.

<sup>12</sup>For  $m_b \neq m$  this in particular implies that there is by assumption no three-point coupling where all particles have mass  $m$ . This could be due to a symmetry but we do not have to commit to an underlying mechanism here.

- Convergence with  $N_{\max}$ . For  $m_b \gtrsim \sqrt{2}$  we see that  $|g|_{\max}$  is nearly stationary as we vary  $N_{\max}$ , whereas for  $m_b \lesssim \sqrt{2}$  we observe more significant improvements with  $N_{\max}$ . We have no explanation for this disparate behaviour (although we suspect it to be related to some subtler higher energy behaviour to which our ansatz is struggling to converge — see also discussion section 5 and appendix G). Numerically we find that we can extrapolate to infinite  $N_{\max}$  and appear to get a finite answer in either domain. We expect this value to correspond to an upper bound on  $|g|$  for *any* scattering amplitude that obeys the constraints of the previous section.<sup>13</sup>
- Peak near  $m_b \sim \sqrt{2}$ . The clear peak is reminiscent of two-dimensional scattering amplitudes, where it was easily explained because in that case the  $s$ - and  $u$ -channel poles cancel precisely at  $m_b = \sqrt{2}$  and the number  $|g|$  becomes meaningless — so no upper bound can be obtained.<sup>14</sup> In greater than two dimensions the cross-channel poles are smeared into a cut by the projection onto the partial waves. One can easily see from (3.5) that this cut starts at  $s = 4 - m_b^2$  thus we find in the partial amplitudes the  $s$ -channel pole starts to overlap with the  $t$ - and  $u$ -channel cut when  $m_b^2 \leq 2$ . While there is a singularity at the branch point of this cut with the correct sign to “screen” the  $s$ -channel pole, this singularity is not strong enough to fully cancel the pole as in 1 + 1 dimensions. The singularity is a  $\log(s - 4 + m_b^2)$  in 3 + 1 and  $(s - 4 + m_b^2)^{-1/2}$  in 2 + 1 (see appendix D for the explicit expressions). We thus expect the peak in figure 4 to remain finite as  $N_{\max} \rightarrow \infty$ . This is borne out by some crude extrapolations (not shown).
- Behavior near threshold,  $m_b \sim 2$ . As explained in appendix E, when  $m_b - 2$  is parametrically small we can analytically constrain the behavior of  $|g|_{\max}$  as a function of  $m_b$ . This result is plotted in the figure as the dashed red line segment. Figure 5 shows a closer analysis of this limit. We see that it accurately traces our numerical results, with the agreement improving as  $m_b$  approaches 2.
- Behavior for  $m_b < 1$ . In this region the scattered particle is no longer the lightest particle in the theory and on physical grounds we expect the two-particle cut in  $A(s, t, u)$  to begin at  $2m_b$  rather than at  $2m$ . For small enough  $m_b$  this is corroborated by our numerics since  $|g|_{\max} \sim 0$  so no pole can be present without modifying our ansatz. It would be interesting to understand in more detail the kink near  $m_b \approx 0.5$ .

For  $m_b = 1$  we can identify the pole with an exchange of the external particle. Reference [15] (see also [10]) discusses an analytic upper bound on  $|g|$  for that case which in our conventions takes the value:

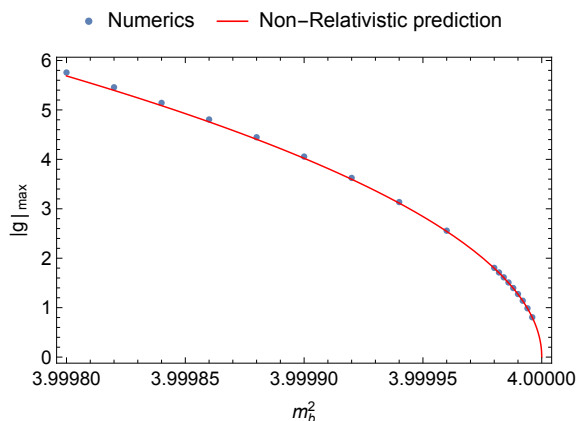
$$|g| \lesssim 16\pi\sqrt{1.5 \cdot 10^6} \approx 61562.4 \tag{4.1}$$

which is far weaker than our current bounds.<sup>15</sup>

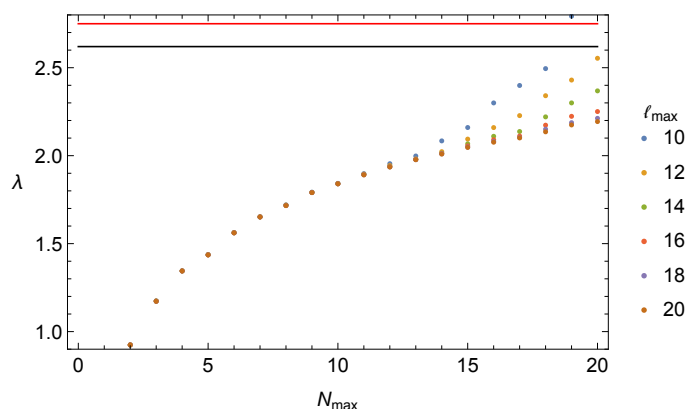
<sup>13</sup>As for any of the results in this paper, it might very well be possible to derive even stronger bounds by including the constraints from other processes involving more particles.

<sup>14</sup>In our ansatz (3.10) this is easily observed by recalling that  $t = 0$  in two dimensions, so also  $u = 4 - s$ . There is also only one partial wave with  $\ell = 0$ .

<sup>15</sup>In [10] the author conceded that “[it] is a large number, but of course [the] calculation was only carried through to show that there exists an upper bound.” We are however not aware of any better previous bounds in the literature.



**Figure 5.** Comparison of numerics with the non-relativistic prediction  $|g|_{\max} \sim 256\pi\sqrt{2 - m_b/m}$  derived in appendix E. These numerics were performed with  $s_0 = m_b^2$  so that the bound state pole always maps to the centre of the  $\rho$  disk. This greatly expedites the convergence in this limit. For example one can already see convergence with  $N_{\max} = 2$  and  $\ell_{\max} = 4$  which are the parameter values used for this plot.

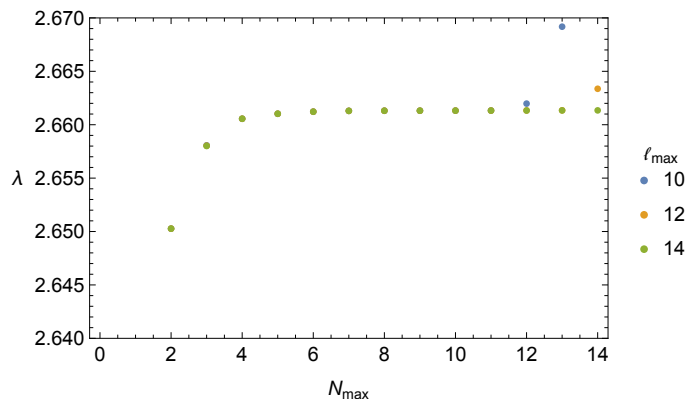


**Figure 6.** A first attempt at obtaining a maximal value of the quartic coupling  $\lambda \equiv \frac{1}{32\pi} M\left(\frac{4}{3}, \frac{4}{3}, \frac{4}{3}\right)$ , using the ansatz (3.10) with  $g = 0$ . We impose the unitarity constraint (3.7) for all  $\ell \leq \ell_{\max}$ . Convergence requires larger  $\ell_{\max}$  for higher values of  $N_{\max}$ . With this ansatz, the maximal quartic coupling continues to increase significantly with  $N_{\max}$  even for  $N_{\max} = 20$ . The black line indicates the value 2.262 achieved in the solution of [16], while the red line indicates the rigorous upper bound 2.75 of [17]. For large enough  $\ell_{\max}$  and  $N_{\max}$  our curves must eventually form a plateau between these two lines, however the convergence is so poor that this cannot be inferred from the plot.

## 4.2 Quartic coupling

Our second set of results concerns the scattering amplitudes  $M(s, t, u)$  without any bound state poles, as for example would be the case in  $\pi^0$  scattering. We will constrain the value of the amplitude at the symmetric but unphysical point  $s = t = u = 4/3$  and therefore define:

$$\lambda \equiv \frac{1}{32\pi} M\left(\frac{4}{3}, \frac{4}{3}, \frac{4}{3}\right) \tag{4.2}$$



**Figure 7.** Maximal value of the quartic coupling  $\lambda \equiv \frac{1}{32\pi} M(\frac{4}{3}, \frac{4}{3}, \frac{4}{3})$ , now using the ansatz (3.10) with  $g = 0$ , supplemented with the term (4.4). With this improved ansatz, the maximal quartic coupling effectively saturates for  $N_{\max} \gtrsim 6$ . A few values of  $l_{\max}$  are shown to demonstrate that the value of the plateau is independent of this cutoff — the data points for various  $l_{\max}$  are indistinguishable until around  $N_{\max} \gtrsim 12$  where the plateau is lost for  $l_{\max} = 10$  (this is just the usual loss of the plateau when  $N_{\max}$  becomes too large compared to  $l_{\max}$ ).

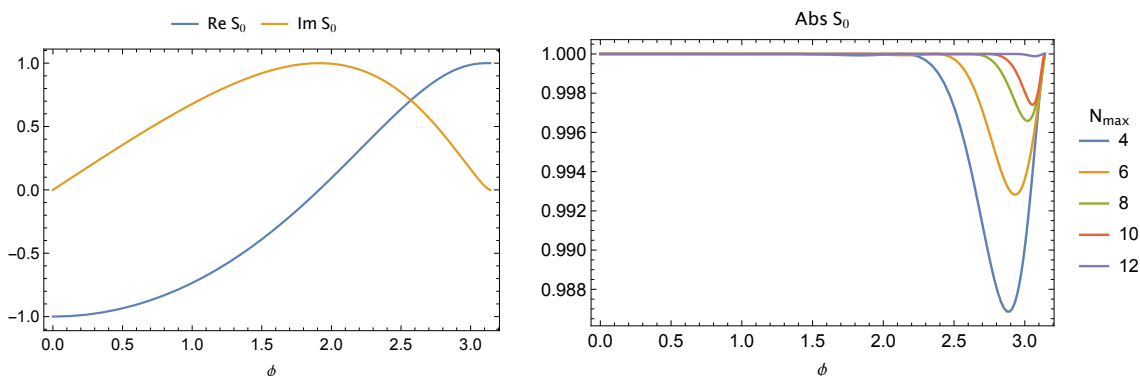
Historically  $\lambda$  was taken to be a measure of the quartic pion interaction strength. In previous works [17] it was constrained both from above and below, in our conventions:

$$- 8.2 \leq \lambda \leq 2.75 \tag{4.3}$$

These constraints stem only from the use of axiomatically proven analyticity, crossing and unitarity. Another data point is provided by the explicit “amplitudes” constructed by Auberson and Mennessier, one with  $\lambda = 2.62$  [16] and one with  $\lambda = -1.69$  [18], both of which obey analyticity, crossing and unitarity. This provides a lower bound for any upper bound and vice versa. It is particularly remarkable that there exists a fairly narrow interval [2, 62, 2.75] in which the best upper bound must reside.

Let us first discuss the case of the upper bound. Figure 6 shows the largest possible value on  $\lambda$  using the ansatz (3.10) (with  $g = 0$ ). One can see that the convergence with  $N_{\max}$  is quite slow which suggests the presence of a singularity near or on the boundary of the  $\rho$  discs. Indeed, as pointed out in [16, 17] the amplitude which achieves the upper bound naturally has a singularity of the form  $(s - 4)^{-1/2}$  corresponding to a bound state sitting precisely at threshold. Physically this is intuitive: the positive sign of the amplitude corresponds to an *attractive* interaction.<sup>16</sup> The situation in which the interaction is as attractive as possible without introducing new bound states occurs just at the point where a resonance is pulled all the way to the threshold. Mathematically it is natural that to make the amplitude as big as possible at the symmetric point it should be made as big as possible at threshold. Figure 7 shows the bound on  $\lambda$  with the threshold bound state

<sup>16</sup>For example in a non-relativistic approximation this would correspond to an attractive delta function potential [19].



**Figure 8.** Real and imaginary parts of  $S_0$  for  $N_{\max} = 12$  and  $\ell_{\max} = 20$  (left plot). Absolute value of  $S_0$  for  $\ell_{\max} = 20$  and several values of  $N_{\max}$  (right plot).

included in the ansatz. This amounts to adding

$$\alpha \left( \frac{1}{\rho_s - 1} + \frac{1}{\rho_t - 1} + \frac{1}{\rho_u - 1} \right) \quad (4.4)$$

to the ansatz (3.10) where now  $\alpha$  is another parameter to be varied. This singularity does not cause a violation of unitarity because it is canceled by the phase-space volume factor in (3.7). More precisely, we find that the  $\ell = 0$  partial amplitude near threshold behaves like

$$S_0(s) = 1 + \frac{1}{16\sqrt{6}\pi} \alpha + O(\sqrt{s-4}). \quad (4.5)$$

and therefore

$$-32\sqrt{6}\pi \leq \alpha \leq 0. \quad (4.6)$$

The unitarity constraints for the higher spin partial amplitudes do not lead to further restrictions on  $\alpha$ .

Once the threshold bound state (4.4) is included we find that convergence is now quite rapid as indicated by the plateau in figure 7 already seen at modest values of  $\ell_{\max}$  and  $N_{\max}$ . The height of the plateau is  $2.6613\dots$  and since

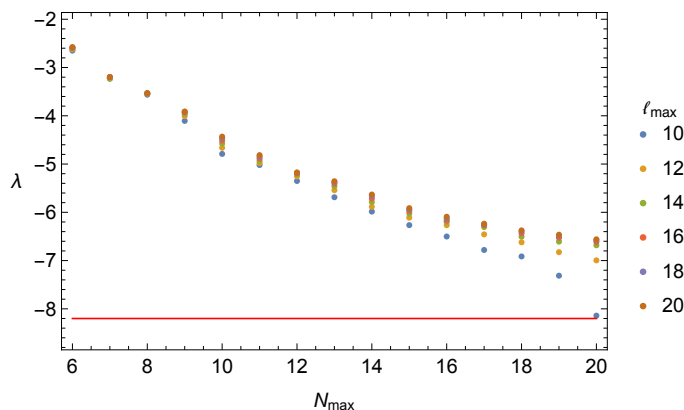
$$2.62 < 2.6613\dots < 2.75. \quad (4.7)$$

it falls beautifully below the rigorous bound of [17] but above the solution constructed in [16]. Given the flexibility of our ansatz we expect this value to represent the strictest possible bound that derives from unitarity, crossing and analyticity of a single amplitude.

An interesting feature of the optimal solution is what appears to be a tendency toward saturation of unitarity. In right plot in figure 8 one can see that  $|S_0|$  increasingly saturates unitarity for increasing values of  $N_{\max}$ . A related fact is that we observe numerically  $\alpha = -32\sqrt{6}\pi$  to great accuracy indicating that unitarity is saturated at threshold. Unitarity saturation is also observed in the higher partial waves.

Let us now consider the lower extremum for which our results are shown in figure 9. As in the previous case (without the threshold singularity) the convergence is quite slow in





**Figure 9.** Minimal value of the quartic coupling  $\lambda \equiv \frac{1}{32\pi} M(\frac{4}{3}, \frac{4}{3}, \frac{4}{3})$  achieved with the ansatz (3.10) (with  $g = 0$ ). With this ansatz, the minimal quartic coupling continues to decrease significantly with  $N_{\max}$  even for  $N_{\max} = 20$ .

$N_{\max}$ . Unfortunately the addition of a threshold bound-state of the form (4.4) cannot save us here, since we would need  $\alpha > 0$  to lower the value of  $\lambda$  but according to (4.6) this is not allowed by unitarity of the spin 0 partial amplitude at threshold. Physically this makes sense — if  $\lambda < 0$  then this indicates a repulsive force which does not favour the creation of bound states nor moving resonances down to the threshold value. Unfortunately we were not able to identify the relevant singularity in this case and thus were not able to improve the slow convergence.

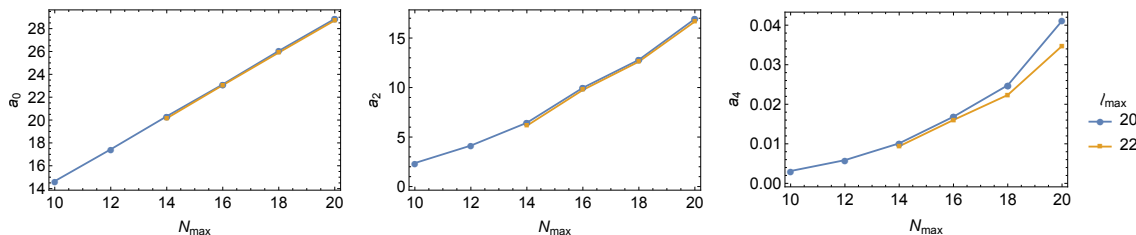
Notwithstanding these convergence issues, we did already significantly improve the lowest possible value of  $-1.69$  that was explicitly constructed in [18]. As the authors of that paper already noted, the discrepancy between their  $-1.69$  and the lower bound  $-8.2$  of [17] means that either the lower bound is quite far from optimal, or that the behaviour of the amplitude which provides this bound is quite “wild” so as to not be contained within the space of functions they explored. Our results indicate that the latter scenario is the correct one since we do seem to be approaching a value in the ball park of the lower bound in (4.3).

### 4.3 Exploring scattering lengths

Another set of observables that received interest in days long gone were the scattering lengths  $a_\ell$ . These are defined as the behavior of the partial waves when  $s$  approaches its threshold value 4. We will restrict ourselves to four spacetime dimensions, i.e.  $d = 3$ , where it is typically defined as

$$a_\ell := \lim_{s \rightarrow 4} \frac{S_\ell(s) - 1}{i(s - 4)^{\ell+1/2}}. \tag{4.8}$$

with the limit taken from above in order to make direct contact with experiment. The power of  $s - 4$  in the denominator arises as follows. One assumes that  $\lim_{s \rightarrow 4} M(s, t)$  is finite for all  $t$  in some neighborhood of zero. Analyticity in  $t$  then allows one to write down a Taylor series expansion in  $t$  whose radius of convergence remains strictly positive as  $s \rightarrow 4$ . Substituting  $t = \frac{1}{2}(s - 4)(x - 1)$  and doing the  $x$  integral in (3.5) to project onto the partial waves of spin  $\ell$  then gives a finite scattering length for all  $\ell$  precisely with



**Figure 10.** Exploring large values of the scattering length. We plot the largest obtained spin 0, 2 and 4 scattering lengths as a function of the number of terms in our ansatz as parametrized by  $N_{\max}$ . For the larger values of  $N_{\max}$  we include two values of  $\ell_{\max}$ . The results are in line with the expectation of unbounded growth as  $N_{\max} \rightarrow \infty$ .

the given prefactor (recall that we are considering  $d = 3$ ). The factor of  $i$  is included to make the scattering length real if  $M(s, t)$  is real-analytic. In this section we will investigate constraints on the scattering length for amplitudes without bound state poles, so we will be using the ansatz (3.10) without the pole terms.

Let us begin with the largest possible values of the scattering length. We first recall that, in ordinary quantum mechanics, scattering lengths are known to diverge when a resonance crosses the threshold value  $s = 4$ . In the  $\rho$ -variables in  $d = 3$  this can be seen by considering scattering amplitudes that locally take the form

$$-\mu \frac{P_\ell^{(3)}(x)}{\rho_s - 1 - \epsilon} + \dots \tag{4.9}$$

with the dots denoting subleading terms, which include permutations to make the amplitude crossing symmetric and other terms to make the amplitude unitary for  $s$  away from 4. From unitarity near  $s = 4$  we obtain the constraint

$$0 \leq \mu \leq \frac{2\ell + 1}{(8\pi)^2 \sqrt{4 - s_0}} \tag{4.10}$$

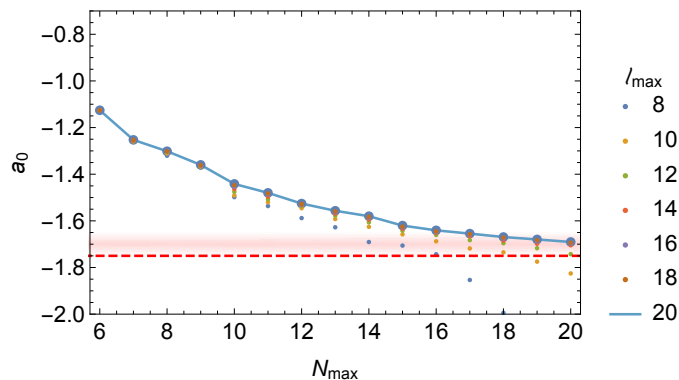
where we recall that  $s_0$  in our ansatz is equal to  $4/3$  and we used that  $\int_{-1}^1 dx P_\ell^{(3)}(x)^2 = [512\pi^2(2\ell + 1)]^{-1}$  in our conventions. The important observation here is that unitarity bounds  $\mu$  independently of the value of  $\epsilon$ , whereas the contribution to the spin  $\ell$  scattering length is given by

$$\frac{(16\pi)^2 \mu}{(2\ell + 1)\epsilon} \tag{4.11}$$

so by sending  $\epsilon$  to zero from above we can get an infinitely large positive scattering length. Notice that  $\epsilon < 0$  creates a pole on the physical sheet and this is disallowed by our ansatz.<sup>17</sup>

The unboundedness from above is borne out by our numerical results. In figure 10 we plot the largest possible values we can obtain for the spin 0, 2 and 4 scattering lengths with our usual ansatz (3.10), again with  $g = 0$ . We observe no convergence to a finite value as we increase  $N_{\max}$ .

<sup>17</sup>In fact, for negative but small  $\epsilon$  and  $\ell = 0$  this amplitude reproduces precisely the extremal behavior for a bound state near threshold discussed in section 4.1 and in appendix E.



**Figure 11.** Lowest possible value of the scattering length as a function of  $N_{\max}$ . Red dashed line: precise lower bound obtained in [21]. Red shaded band: approximate lower bound obtained in [20].

We can also consider the lowest possible values of the scattering lengths. For spin 0 the best known lower bound dates from 1980 and is given by [20]

$$a_0 \gtrsim -1.7, \tag{4.12}$$

which slightly improves on a more precise bound obtained five years earlier in [21]:

$$a_0 \geq -1.75. \tag{4.13}$$

These result were the culmination of a series of works, starting with the observations in [22] which were followed by a series of intermediate improvements in e.g. [10, 20, 23, 24].<sup>18</sup> Our numerical results are shown in figure 11 and are clearly converging in the neighborhood of the above lower bounds. This shows that the lower bound can more or less be saturated (with an amplitude that falls within our ansatz), which is actually a new result: the best known constructible value was -0.88 [18].

In fact, it may appear that we get dangerously close to the value  $-1.7$  and that further increasing  $N_{\max}$  may push us over the edge. However for this particular bound the convergence with  $l_{\max}$  is quite slow and the value corresponding to infinite  $l_{\max}$  may in fact increase a little bit. It would be interesting to perform a precision study with larger values of  $l_{\max}$  and  $N_{\max}$  and to simultaneously re-compute with higher precision the lower bound of  $-1.7$  obtained in [20]. We leave this to the future.

For the higher spin scattering lengths one can use the Froissart-Gribov representation, see e.g. [11], to arrive at the simple lower bound:

$$a_\ell \geq 0 \quad \forall \ell \geq 2 \tag{4.14}$$

This is borne out by our numerics but we do not show the results since a plot consisting of nothing but zeroes is not very interesting.<sup>19</sup>

<sup>18</sup>Papers like [18] contain a reference to an unpublished lower bound of -1.65 that had supposedly been obtained in 1978 by Caprini and Dita, the authors of [20]. It was confirmed to us by Irinel Caprini in personal communication that this value is incorrect.

<sup>19</sup>We would like to remark that for sufficiently high  $N_{\max}$  (say, 20) we need to impose unitarity for relatively large values of  $l_{\max}$  (say, 24) before the lowest possible allowed value of  $a_\ell$  gets pinned at zero.

#### 4.4 Bonus feature: three spacetime dimensions and QFT in AdS

In our previous work [1] we outlined another method for constraining QFT data, based on putting a QFT in AdS. The main idea is to investigate the *boundary correlation functions*, which behave exactly like CFT correlation functions (except there is no stress tensor) and are therefore amenable to an ordinary conformal bootstrap analysis. As we explained in [1], the translation between boundary and bulk quantities parallels the standard AdS/CFT dictionary, for example  $m^2 R^2 = \Delta(\Delta - d)$ , and furthermore we found precise formulae that dictate how the boundary correlation functions morph into flat-space scattering amplitudes upon sending the AdS curvature to zero. In [1] we numerically tested these equations in 1+1 dimensions and found a quantitative match between the two approaches to the S-matrix bootstrap.

For this paper we set out to repeat this exercise for QFTs in 2+1 dimensions. We focused on the 2+1 dimensional version of the maximal possible coupling that we discussed in section 4.1. This setup was called scenario I in [1]. We discuss the salient points of the methodology before presenting the results.

##### 4.4.1 S-matrix bootstrap approach

For the S-matrix bootstrap, the only difference in the implementation between the 3+1 dimensional analysis of section 4.1 and the present one is that we were no longer able to compute the partial amplitudes (3.5) analytically. The method explained in appendix D fails because the factor  $(1 - x^2)^{\frac{d-3}{2}}$  in (3.5) introduces an additional square-root cut in 2+1 dimensions ( $d = 2$  in the conventions of this paper). Thus we are forced to evaluate the partial amplitudes by brute force use of Mathematica's `NIntegrate`. Although slow, this approach is manageable with the use of multiple computing cores. This leads us to the:

- **First approach:** maximal three-point coupling  $g^2$  for any flat-space QFT, obtained by assuming a flat-space scattering amplitude captured by our ansatz (3.10) and obeying the unitarity condition (3.7), as a function of  $m_b/m$ .

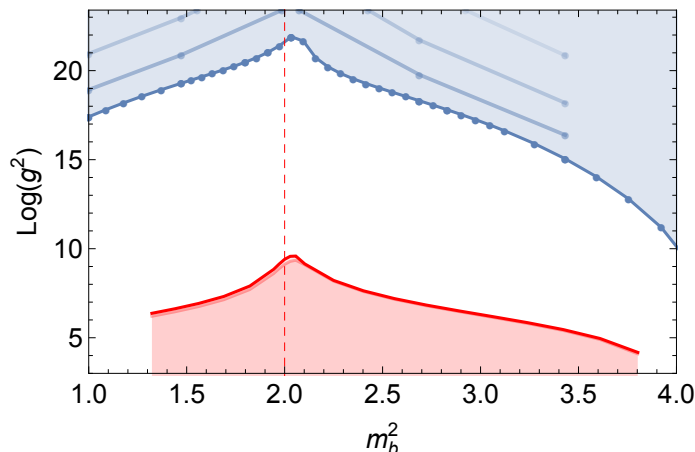
##### 4.4.2 QFT in AdS approach

For the QFT in AdS approach we refer to [1] for a detailed exposition of the method, except that presently we consider two-dimensional rather than one-dimensional conformal four-point functions. This implies that there is an extra cross ratio, since  $z$  is no longer kinematically equal to  $\bar{z}$ , and conformal blocks are labelled by a pair  $(\Delta, \ell)$  rather than just the scaling dimension  $\Delta$ . The combined effect of these modifications is simply that the numerics is computationally much more demanding.<sup>20</sup>

Now, in [1] we obtained a precise match in 1+1 dimensions by taking the raw numerical QFT in AdS results and performing a double extrapolation: first to “infinite computational

---

<sup>20</sup>The introduction of spin does lead to one new subtlety, namely the magnitude of the two-particle gap for spinning particles. If there is a single scalar particle corresponding to a boundary operator with dimension  $\Delta$  then we chose to set the two-particle gap at  $2\Delta + \ell$  as in free field theory. Notice that the flat-space limit merely dictates that the gap tends to  $2\Delta$  for very large  $\Delta$ , but there is freedom in choosing the subleading terms.



**Figure 12.** Investigating the coupling between two particles of mass  $m = 1$  and a third particle of mass  $m_b$  in 2+1 dimensional QFT. **First approach:** maximum flat-space coupling for a QFT obtained with our ansatz (3.10). We plot two curves with  $N_{\max}$  equal to 10 (bottom) and 18 (top) which lie almost on top of each other. **Second approach:** upper bound on the bulk coupling for a QFT in AdS with a radius of curvature  $R \approx 16$ . The four curves connect the sets of data points which were obtained with functionals with 10 (top), 36, 78 and 136 (bottom) components.

power” and then to infinite  $\Delta$  corresponding to the flat-space limit. For our 2+1 dimensional results we unfortunately run into trouble at the first step: our numerical results, obtained for  $1 \leq \Delta \leq 20$  with functionals with up to 136 components, were not amenable to reliable extrapolations. We therefore chose to present the result *directly for a QFT in AdS*. We chose  $\Delta = 17$  as a representative value.<sup>21</sup> Altogether this gives the:

- **Second approach:** maximal three-point (bulk) coupling  $g^2$  for a QFT in AdS, obtained by assuming boundary correlation functions consistent with unitarity and a spectrum with the natural two-particle gaps, again as a function of  $m_b/m$ .

### 4.4.3 Results

The resulting bounds are shown in figure 12. Notice the logarithmic scale.<sup>22</sup> It is clear that the upper bound obtained from QFT in AdS is way larger than the largest value obtained from the S-matrix bootstrap, but the AdS results have not converged yet and one may hope that the numerical upper bound can decrease much further. The good news, however, is the remarkably similar *shape* of the two curves, both having a somewhat asymmetric peak slightly above  $m_b^2 = 2$ . In this sense we see a repetition of the results in 1+1 dimensions, namely that we can obtain similar bounds on the residue of a pole in a scattering amplitudes using two drastically different methods.

Physically, it is important to realize that our QFT in AdS approach is completely devoid of any assumptions about the analyticity of the flat-space scattering amplitude. If one agrees that the result in figure 12 provides evidence of the equivalence between the two

<sup>21</sup>For  $\Delta = 17$  we find that  $m^2 R^2 = \Delta(\Delta - 2) = 255$  so the reduced compton wavelength of the particle is about 16 times the AdS radius of curvature in our setup - in this sense space is already quite flat.

<sup>22</sup>On a regular scale the shape of the peak is very similar to the one shown in figure 4.

approaches, then either our S-matrix bounds on the coupling do not require the amount of analyticity that we have imposed or the analyticity (at least of the extremal scattering amplitudes) is a property that we may hope to derive from the QFT in AdS construction. Either option would be very interesting and should be investigated further.<sup>23</sup>

Although  $\Delta = 17$  was the largest value for which we had a full set of results, let us briefly discuss the result for  $0 < \Delta \leq 20$ . In line with the results in [1], the absolute value of the numerical bounds decreases quickly upon decreasing  $\Delta$ . For  $\Delta \gtrsim 4$  the curve always has a peak hovering around  $m_b^2 = 2$ , which broadens a bit upon decreasing  $\Delta$ . For  $0 < \Delta \lesssim 4$  the peak moves more or less linearly towards  $m_b^2 = 4$  as  $\Delta \rightarrow 0$ . In the future it would be interesting to invest more computational resources and explore in more detail both this behavior and the general convergence of the bounds.

## 5 Discussion

Here we continued our exploration of the space of S-matrices of gapped quantum field theories initiated in [1, 2]. We present a fresh approach to an old question of constraining S-matrix elements based on unitarity, crossing and analyticity. The former two properties are firmly established properties of the S-matrix whose meaning requires no clarification. By analyticity we mean the rather simplistic (but perhaps most natural) assumption that  $M(s, t, u)$  is an analytic function of each of its variables with no singularities in their respective cut planes. We make no assumption about the properties of the S-matrix outside of this union of cut planes — i.e. off the physical sheet.

Of course there are many open questions in S-matrix theory pertaining to analyticity. Are all singularities in the complex Mandelstam variables  $s, t, u$  associated to Landau diagrams (as expected based on perturbation theory) or should we be open to more exotic possibilities especially in strongly coupled theories? What is the most general possible large energy behaviour of scattering amplitudes? Finally, if we bravely cross the gates and delve into the various Riemann sheets of non-perturbative scattering amplitudes by crossing its various cuts in the physical sheet, what kind of scary Chimeras await us down there?

We tried to be optimistic — by assuming the minimal expected singularities in the physical sheet — and cautious at the same time — by assuming as little as possible about the uncontrollable world of the other unphysical sheets or the large energy behaviour of scattering amplitudes. In short we mapped the physical sheet into a few unit disks and assumed little about the behaviour of amplitudes on the boundary of those disks which is where both the large energy behaviour as well as the various physical thresholds lie. Inside these disks we assumed that the only singularities were poles associated to stable bound states.

In the future, it would be interesting to develop new numerical investigations relying on more rigorous analyticity assumptions. Perhaps our results are not too sensitive to this distinction, or perhaps we will encounter exotic S-matrices which make use of the allowed

---

<sup>23</sup>In 1+1 dimensions the status of analyticity is a little different. Although we are not aware of any full-fledged two-dimensional proofs, since  $t = 0$  kinematically one may say that analyticity in two dimensions is similar to forward analyticity in higher dimensions. The analyticity properties of  $M(s, t = 0)$  can often be proven from axiomatic field theory [8].

non-analyticity to allow for a wider range of values. Both would be very interesting! To this end, it is worth noting that in the case of the quartic pion coupling and the lower bound on the spin zero scattering length we can say with confidence that we are in the former scenario — our results approach the bounds obtained in [17, 18] and in [20] which are based on rigorously proven analyticity properties. More evidence for the first scenario is the at least qualitative match between our maximal coupling and the upper bound on the same observable for a QFT in AdS, since the latter computation relied on no analyticity properties whatsoever. Finally we can point to the consistency of our approach with a Mandelstam representation expansion discussed in appendix C.

As for the behaviour at the boundary of the disks the idea here is that we can be agnostic about it and let regular Taylor expansions in the bulk converge towards whatever they *want* to. Of course, without inputting the correct singularities at the boundary of the disk, the numerics should still work but their convergence will suffer considerably. We encountered two examples of this already in the main text. The first is the quartic coupling numerics whose convergence increased substantially once we allowed for a bound state singularity at threshold. Another example is probably the four dimensional bound state coupling numerics when the bound-state mass is less than  $\sqrt{2}$  times the mass of the lowest particle. The numerics are converging much slower for that range as clearly seen in the left curves in figure 4. We suspect in this case it is rather related to a non-trivial large energy behaviour of the S-matrix which the ansatz has a hard time reproducing.<sup>24</sup> It would be interesting to investigate this further.

It is also at the boundary of these disks where we read physical amplitudes with any  $s > 4m^2$  and negative  $t$ . Multi-particle production will show up as further cuts at larger  $s$  such as  $9m^2$ ,  $16m^2$ , etc and infinitely many others like  $(m + m')^2$ , etc if there are other stable particles. We saw no signs of these singularities in our numerics. As we for example show in figure 8, our optimal S-matrices do not seem to open multi-particle production cuts in any significant way. A priori this sounds very strange. How could we have no particle production of four particles from two particles if - by crossing one particle to the past - that amplitude is related to a  $3 \rightarrow 3$  process which obviously must exist?<sup>25</sup> Indeed, it is known [26] that particle production is mandatory. It can not be strictly zero or it would lead to important contradictions. Unfortunately, the same work [26] — or any other work as far as we know — does not put a lower bound on *how much* particle production one *must have* and as such we could not reach a sharp contradiction with the numerics which by definition can never rule out an arbitrarily low particle production.<sup>26</sup>

Nonetheless, absence of particle production is unphysical in spacetime dimension greater than 2. We would like to describe more realistic theories where particle production

---

<sup>24</sup>See appendix G for a two dimensional example which we believe might be the counterpart of what we are observing here.

<sup>25</sup>Of course in 1+1 there is a well know loophole in this argument which allows for integrable theories [25]. This loophole is not possible in higher dimensions.

<sup>26</sup>Actually one can show that a certain amount of production must persist in the limit of infinite spin [27]. However, to our knowledge, there is no theorem saying, for example, that the first  $L$  partial waves exhibit no production.



naturally arises. One way of forcing such particle production in a natural way is to study multiple S-matrix elements where we consider a system of scattering elements involving not only the lightest particle but also the next-to-lightest etc. We are currently working on this and finding some very encouraging preliminary results in two dimensions where the bounds are often significantly improved and the corresponding S-matrices do exhibit particle production and thus must correspond to genuinely non-integrable theories in contrast to our previous work [2].

The analyticity properties of scattering amplitudes of several particles of different mass are more intricate than what we considered here. The optimistic scenario is that all singularities on the physical sheet follow from Landau diagrams describing propagation of on-shell particles. This *Landau analyticity* is far from being rigorously established but it is a reasonable physical conjecture to start from. Even with this assumption, we will have to deal with anomalous thresholds (singularities that arise from Landau diagrams that are not on a line). A simple example is the scattering amplitude of particles of mass greater than  $\sqrt{2}$  times the mass lightest particle. We plan to analyse this issue in the future, starting in 1+1 dimensions.

### Acknowledgments

We thank Zoltan Bajnok, Benjamin Basso, Irinel Caprini, Patrick Dorey, Davide Gaiotto, Zohar Komargodski, Martin Kruczenski, Andre Martin, Rafael Porto, Francesco Riva, Slava Rychkov, Amit Sever, Alexander Zamolodchikov and Alexander Zhiboedov for numerous enlightening discussions and suggestions. Research at the Perimeter Institute is supported in part by the Government of Canada through NSERC and by the Province of Ontario through MRI. This research received funding from the grant CERN/FIS-NUC/0045/2015. This work was additionally supported by a grant from the Simons Foundation (JP: #488649, BvR: #488659, PV: #488661) JP is supported by the National Centre of Competence in Research SwissMAP funded by the Swiss National Science Foundation.

### A $x(s)$ vs $\rho_s, \rho_t$ in 1 + 1 dimensions

Consider the map

$$x(s) = \frac{2 - \sqrt{4 - s}\sqrt{s}}{s - 2}$$

which maps the full  $s$ -plane minus the cuts  $s > 4$  and  $s < 0$  into the unit disc  $|x(s)| \leq 1$  and the map

$$\rho_s = \frac{2 - \sqrt{4 - s}}{2 + \sqrt{4 - s}}$$

which maps the full  $s$ -plane minus a single cut  $s > 4$  into the unit disc  $|\rho_s| \leq 1$ . An analytic function in the  $s$ -plane minus the cuts  $s > 4$  and  $s < 0$  — such as the S-matrix once we subtract out its known poles — can be written as

$$f(s) = \sum_{n=0}^{\infty} c_n x(s)^n \tag{A.1}$$



Now, we have

$$x(s) = \frac{\rho_s - \rho_t}{1 - \rho_s \rho_t} \quad \text{where} \quad t = 4 - s \quad (\text{A.2})$$

which admits a convergent expansion in powers of  $\rho_s$  and  $\rho_t$  provided they are both inside the unit list (and hence so is their product in the denominator). Hence the function  $f(s)$  can also be cast as

$$f(s) = \sum_{n=0}^{\infty} c_{nm} \rho_s^n \rho_t^m \quad \text{where} \quad t = 4 - s \quad (\text{A.3})$$

As such, our 1 + 1 numerics had to work.

## B Constraint surface in $\rho$ -coordinates

The on-shell condition imposes

$$0 = s + t + u - 4m^2. \quad (\text{B.1})$$

If we write this constraint in terms of the  $\rho_s, \rho_t$  and  $\rho_u$  variables with arbitrary  $s_0$  (cf. eq. (2.2)) we get a somewhat lengthy expression of the form

$$0 = \left( s_0 - \frac{4}{3}m^2 \right) \rho_s^2 \rho_t^2 \rho_u^2 + \dots + \left( s_0 - \frac{4}{3}m^2 \right). \quad (\text{B.2})$$

Specializing to the case  $s_0 = \frac{4}{3}m^2$ , the point  $\rho_s = \rho_t = \rho_u = 0$  satisfies the on-shell condition. Defining then the symmetrized monomials:

$$\rho^{(a,b,c)} = \rho_s^a \rho_t^b \rho_u^c + \text{perms}, \quad (\text{B.3})$$

the constraint equation becomes ( $m = 1$ ):

$$0 = \rho^{(1,2,2)} - 4\rho^{(1,1,2)} + \rho^{(1,2,0)} + 12\rho^{(1,1,1)} - 4\rho^{(1,1,0)} - \rho^{(1,0,0)}. \quad (\text{B.4})$$

We can now obtain all such constraints by multiplying this equation by other symmetrized monomials. As an example, multiplying by  $\rho^{(1,0,0)}$  we get a new identity,

$$0 = \rho^{(0,0,2)} + 2\rho^{(0,1,1)} - 4\rho^{(0,1,2)} + \rho^{(0,1,3)} + 2\rho^{(0,2,2)} - 12\rho^{(1,1,1)} + 14\rho^{(1,1,2)} - 4\rho^{(1,1,3)} - 8\rho^{(1,2,2)} + \rho^{(1,2,3)} + 3\rho^{(2,2,2)}. \quad (\text{B.5})$$

We can use these identities to systematically reduce the number of monomials in our ansatz as explained in figure 13. Note that in two spacetime dimensions we can set  $u = 0$  which simplifies the constraint equation to (2.4).

Terms we can eliminate in Two Dimensions:

$C_{1,0}$	
$C_{2,0}$	$C_{1,1}$
$C_{3,0}$	$C_{2,1} \rightarrow 0$
$C_{4,0}$	$C_{3,1}$
$C_{5,0}$	$C_{4,1} \rightarrow 0$
$C_{6,0}$	$C_{5,1}$
$C_{7,0}$	$C_{6,1} \rightarrow 0$
$C_{8,0}$	$C_{7,1}$
$C_{9,0}$	$C_{8,1} \rightarrow 0$
$C_{10,0}$	$C_{9,1}$
$C_{11,0}$	$C_{10,1} \rightarrow 0$

	$C_{2,2} \rightarrow 0$
	$C_{3,2} \rightarrow 0$
	$C_{4,2} \rightarrow 0$
	$C_{5,2} \rightarrow 0$
	$C_{6,2} \rightarrow 0$
	$C_{7,2} \rightarrow 0$
	$C_{8,2} \rightarrow 0$
	$C_{9,2} \rightarrow 0$
	$C_{10,2} \rightarrow 0$
	$C_{11,2} \rightarrow 0$

Terms we Can Eliminate in Higher Dimensions:

$\alpha_{1,0,0}$																				
$\alpha_{2,0,0}$	$\alpha_{1,1,0}$																			
$\alpha_{3,0,0}$	$\alpha_{2,1,0}$	$\alpha_{1,1,1}$																		
$\alpha_{4,0,0}$	$\alpha_{3,1,0}$	$\alpha_{2,2,0}$	$\alpha_{2,1,1}$																	
$\alpha_{5,0,0}$	$\alpha_{4,1,0}$	$\alpha_{3,2,0}$	$\alpha_{3,1,1}$	$\alpha_{2,2,1} \rightarrow 0$																
$\alpha_{6,0,0}$	$\alpha_{5,1,0}$	$\alpha_{4,2,0}$	$\alpha_{4,1,1}$	$\alpha_{3,2,1}$	$\alpha_{2,2,2} \rightarrow 0$															
$\alpha_{7,0,0}$	$\alpha_{6,1,0}$	$\alpha_{5,2,0}$	$\alpha_{5,1,1}$	$\alpha_{4,2,1}$	$\alpha_{3,3,1} \rightarrow 0$	$\alpha_{3,2,2} \rightarrow 0$														
$\alpha_{8,0,0}$	$\alpha_{7,1,0}$	$\alpha_{6,2,0}$	$\alpha_{6,1,1}$	$\alpha_{5,2,1}$	$\alpha_{4,3,1} \rightarrow 0$	$\alpha_{4,2,2} \rightarrow 0$	$\alpha_{3,3,2} \rightarrow 0$													
$\alpha_{9,0,0}$	$\alpha_{8,1,0}$	$\alpha_{7,2,0}$	$\alpha_{7,1,1}$	$\alpha_{6,2,1}$	$\alpha_{5,3,1}$	$\alpha_{4,4,1} \rightarrow 0$	$\alpha_{5,2,2} \rightarrow 0$	$\alpha_{4,3,2} \rightarrow 0$	$\alpha_{3,3,3} \rightarrow 0$											
$\alpha_{10,0,0}$	$\alpha_{9,1,0}$	$\alpha_{8,2,0}$	$\alpha_{8,1,1}$	$\alpha_{7,2,1}$	$\alpha_{6,3,1}$	$\alpha_{5,4,1} \rightarrow 0$	$\alpha_{6,2,2} \rightarrow 0$	$\alpha_{5,3,2} \rightarrow 0$	$\alpha_{4,4,2} \rightarrow 0$	$\alpha_{4,3,3} \rightarrow 0$										
$\alpha_{11,0,0}$	$\alpha_{10,1,0}$	$\alpha_{9,2,0}$	$\alpha_{9,1,1}$	$\alpha_{8,2,1}$	$\alpha_{7,3,1}$	$\alpha_{6,4,1} \rightarrow 0$	$\alpha_{7,2,2} \rightarrow 0$	$\alpha_{6,3,2} \rightarrow 0$	$\alpha_{5,4,2} \rightarrow 0$	$\alpha_{5,3,3} \rightarrow 0$	$\alpha_{4,4,3} \rightarrow 0$									

**Figure 13.** When centering the  $\rho$  variables around general points, we can eliminate all constants  $c_{a,b}$  with  $a,b > 1$  in two dimensions and all constants  $\alpha_{a,b,c}$  with  $a,b,c > 1$  in higher dimensions. By centering the  $\rho$  variables around  $s_0 = 2$  in two dimensions and around  $s_0 = 4/3$  in higher dimensions, the kinematical constraints simplify further allowing us to eliminate a few more terms in the Taylor expansions as explained in the text. An option for which terms we can eliminate is illustrated in the tables above up to  $N = 11$  where the level  $N = a + b$  or  $N = a + b + c$  is the total powers of  $\rho$  in the multiple Taylor expansion. The number of terms we should keep at each level is  $\frac{N}{2} + \frac{(-1)^N}{4} + \frac{3}{4}$  in two dimensions and  $\frac{N^2}{12} + \frac{N}{2} + \frac{(-1)^N}{8} + \frac{2}{9} \cos\left(\frac{2\pi N}{3}\right) + \frac{47}{72}$  in higher dimensions.

### C Mandelstam representation

The double dispersion representation proposed by Mandelstam [28] implies that the amplitude can be written as follows

$$M(s, t, u) = B(s, t) + B(s, u) + B(t, u), \tag{C.1}$$

where

$$B(s, t) = \int ds' dt' \frac{C(s', t')}{(s' - s)(t' - t)}. \tag{C.2}$$

If there are no stable particles below threshold, the double discontinuity  $C(s, t)$  has support inside the region  $s > 4m^2$  and  $t > 4m^2$ . In practice, this form of the double dispersion relation is not valid and one needs to include subtractions. A simple trick to derive the form of the dispersion relation with  $n$  subtractions is to use the identity

$$\frac{1}{s' - s} = \frac{(s - s_0)^n}{(s' - s)(s' - s_0)^n} + \sum_{k=0}^{n-1} \frac{(s - s_0)^k}{(s' - s_0)^{k+1}} \tag{C.3}$$

in equation (C.2) for both factors in the denominator. This leads to

$$\begin{aligned}
 B(s, t) &= (s - s_0)^n (t - t_0)^n \int ds' dt' \frac{C(s', t')}{(s' - s)(t' - t)(s' - s_0)^n (t' - t_0)^n} \\
 &+ \sum_{k=0}^{n-1} (s - s_0)^k (t - t_0)^n \int dt' \frac{c_k(t')}{(t' - t)(t' - t_0)^n} \\
 &+ \sum_{k=0}^{n-1} (t - t_0)^k (s - s_0)^n \int ds' \frac{c_k(s')}{(s' - s)(s' - s_0)^n} \\
 &+ \sum_{k,l=0}^{n-1} (s - s_0)^k (t - t_0)^l c_{k,l}
 \end{aligned} \tag{C.4}$$

where

$$c_k(t) = \int ds' \frac{C(s', t)}{(s' - s_0)^{k+1}}, \quad c_{k,l} = \int ds' dt' \frac{C(s', t')}{(s' - s_0)^{k+1} (t' - t_0)^{l+1}}. \tag{C.5}$$

In general the integrals (C.5) do not converge. The subtracted dispersion relation is (C.4) considering  $c_k(t)$  and  $c_{k,l}$  as independent functions from the double discontinuity  $C(s, t)$ .

Stable particles correspond to delta-function pieces in the single discontinuities  $c_k(s)$ .<sup>27</sup> Besides these delta-functions, the support of  $c_k(s)$  is  $s \geq 4m^2$ . Therefore, the analytic properties of equation (C.4) imply that

$$B(s, t) = \text{Poles} + \sum_{a,b=0}^{\infty} \alpha_{(ab)} \rho_s^a \rho_t^b, \tag{C.6}$$

with a convergent double  $\rho$  series in the product of two unit disks. This is a more restricted form of formula (3.10) where we set to zero all coefficients  $\alpha_{abc}$  with  $a > 0, b > 0$  and  $c > 0$ .

In order to test the validity of Mandelstam representation, we reconsidered the problem discussed in section 4.2 using the more restricted ansatz

$$B(s, t) = \frac{\alpha}{2} \left( \frac{1}{\rho_s - 1} + \frac{1}{\rho_t - 1} \right) + \sum_{a,b=0}^{N_{\max}} \alpha_{(ab)} \rho_s^a \rho_t^b. \tag{C.7}$$

In figure 14, we show the maximal value of the quartic coupling  $\lambda$  obtained with this ansatz. The maximal value  $\lambda \approx 2.6613\dots$  is obtained for  $N_{\max} \gtrsim 6$ . This result suggests that in the limit of large  $N_{\max}$  both ansätze cover the same space of functions.

## D Partial wave integrals

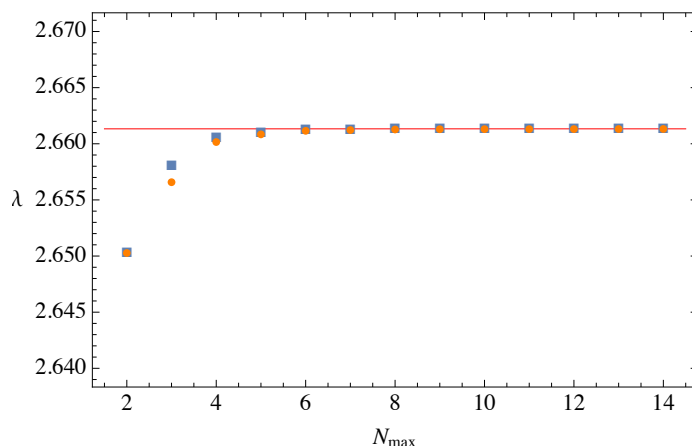
### D.1 Pole contributions

Here we will consider the contribution to partial waves coming from poles of the scattering amplitude. Consider

$$M(s, t, u)_{\text{poles}} = -\frac{g^2}{s - m_b^2} - \frac{g^2}{t - m_b^2} - \frac{g^2}{u - m_b^2}. \tag{D.1}$$

---

<sup>27</sup>Therefore, we should use  $n \geq j + 1$  where  $j$  is the maximal spin of the stable particles. In this way, the second and third line of (C.4) can reproduce the degree  $j$  polynomial residue of the pole produced by the stable particle.



**Figure 14.** Comparison of upper bound on pion coupling using ansatz (3.10) with  $g = 0$  and the threshold singularity (4.4) included (blue squares) versus ansatz (C.7) (orange dots). In both cases we use  $\ell_{\max} = 14$ . The plateau converges to  $2.6613\dots$  in both cases.

It is easy to compute the partial wave decomposition of this expression. For  $d = 3$  we get

$$\frac{S_\ell(s) - 1}{2i} = -\frac{g^2}{32\pi} \frac{\sqrt{s-4}}{\sqrt{s}} \left[ \frac{\delta_{\ell,0}}{s - m_b^2} - \frac{4}{s - 4m^2} Q_\ell(x_b) \right] \quad (\text{D.2})$$

with  $x_b = x(s, t = m_b^2)$  and  $Q_\ell(z)$  the Legendre function of the second kind with branch cut along  $z \in (-1, 1)$ . For  $d = 2$  we instead get

$$\frac{S_\ell(s) - 1}{2i} = -\frac{g^2}{16\sqrt{s}} \left[ \frac{\delta_{\ell,0}}{s - m_b^2} - \frac{2}{m_b \sqrt{s - 4m^2 + m_b^2}} \left( \frac{m_b - \sqrt{s - 4m^2 + m_b^2}}{m_b + \sqrt{s - 4m^2 + m_b^2}} \right)^\ell \right]. \quad (\text{D.3})$$

Now consider the contribution to the amplitude from a threshold bound state. The pole part is

$$M(s, t, u)_{\text{th.pole}} = -\frac{2\alpha}{\sqrt{6}\sqrt{4m^2 - s}} - \frac{2\alpha}{\sqrt{6}\sqrt{4m^2 - t}} - \frac{2\alpha}{\sqrt{6}\sqrt{4m^2 - u}}. \quad (\text{D.4})$$

If we focus on the case  $d = 3$ , we must compute integrals of the form:

$$\int_{-1}^1 dx \frac{P_\ell(x)}{\sqrt{4m^2 - t(x)}} \quad (\text{D.5})$$

with  $t(x) = -\frac{1}{2}(s - 4)(1 - x)$ . Introducing the generating function for the Legendre polynomials

$$\sum_{n=0}^{+\infty} z^n P_n(x) = \frac{1}{\sqrt{1 - 2xz + z^2}}, \quad (\text{D.6})$$

it is not difficult to obtain

$$\int_{-1}^1 dx \frac{P_\ell(x)}{\sqrt{4m^2 - t(x)}} = \frac{4}{2\ell + 1} \frac{(\sqrt{s} - 2m)^\ell}{(\sqrt{s} + 2m)^{\ell+1}} \quad (\text{D.7})$$

Adding up contributions from  $s, t, u$  the partial amplitudes are

$$\frac{S_\ell(s) - 1}{2i} = -\frac{\alpha}{16\sqrt{6}\pi} \frac{\sqrt{s - 4m^2}}{\sqrt{s}} \left( \frac{i \delta_{\ell,0}}{\sqrt{s - 4m^2}} + \frac{4}{2\ell + 1} \frac{(\sqrt{s} - 2m)^\ell}{(\sqrt{s} + 2m)^{\ell+1}} \right) \quad (\text{D.8})$$

## D.2 $\rho_s \rho_t \rho_u$ contributions

Here we will show how to obtain the contribution to the partial amplitudes from terms of the form  $\rho_s^a \rho_t^b \rho_u^c$  analytically in  $d = 3$ . While the calculation is somewhat tedious, the underlying concept is simple: the integral that we want to do has only one cut (of square-root type) in the integrand and thus with a simple trigonometric change of variables the integrand can be converted to a rational function and computed by partial fractions (or some more clever method).

The non-trivial integrals to perform take the form

$$I_{b,c}^\ell = \int_{-1}^1 dx P_\ell(x) \rho(t)^b \rho(u)^c \quad (\text{D.9})$$

with, as in (2.2) with  $m = 1$ ,

$$\rho(s) = \frac{1 - \sqrt{1 - \frac{s-s_0}{4-s_0}}}{1 + \sqrt{1 - \frac{s-s_0}{4-s_0}}} \quad (\text{D.10})$$

In applications we typically set  $s_0 = 4/3$ . We next introduce our first inspired change of variables from  $x$  to  $\phi$  which is given by

$$x = -\frac{s+4}{s-4} \cos(2\phi). \quad (\text{D.11})$$

In these variables we get:

$$I_{b,c}^\ell = 4 \left( \frac{s+4}{s-4} \right) \int d\phi P_\ell(x(\phi)) \sin(\phi) \cos(\phi) \left( \frac{1 - r \cos \phi}{1 + r \cos \phi} \right)^b \left( \frac{1 - r \sin \phi}{1 + r \sin \phi} \right)^c \quad (\text{D.12})$$

where we also introduced

$$r^2 \equiv \frac{4+s}{4-s_0}. \quad (\text{D.13})$$

We should now do the usual change of variables,

$$\phi = 2 \arctan(y) \quad (\text{D.14})$$

This gives

$$I_{b,c}^\ell = \frac{4^2 r^2}{r^2 - \frac{8}{4-s_0}} \int dy P_\ell(x(y)) \frac{y(1-y^2)}{(1+y^2)^3} \left( \frac{(1-r) + y^2(1+r)}{(1+r) + y^2(1-r)} \right)^b \left( \frac{1 - 2ry + y^2}{1 + 2ry + y^2} \right)^c. \quad (\text{D.15})$$

We have

$$x(y) = \frac{r^2}{r^2 - \frac{8}{4-s_0}} \frac{1 - 6y^2 + y^4}{(1+y^2)^2}, \quad (\text{D.16})$$

and the integral runs from  $y_i$  to  $y_f$  with

$$y_i = \frac{\sqrt{4-s_0}}{2} \left( r - \sqrt{r^2 - \frac{4}{4-s_0}} \right), \quad y_f = \left( \frac{r - \frac{2}{\sqrt{4-s_0}}}{r + \frac{2}{\sqrt{4-s_0}}} \right)^{\frac{1}{2}}. \quad (\text{D.17})$$

The trick now is to rewrite the integration region using the discontinuity of a logarithm,

$$\int_{y_1}^{y_2} dy f(y) = \frac{1}{2\pi i} \int_{y_1}^{y_2} dy f(y) \text{Disc} \log \left( \frac{y-y_2}{y-y_1} \right) = \frac{1}{2\pi i} \int_{(y_1, y_2)} dy f(y) \log \left( \frac{y-y_2}{y-y_1} \right) \quad (\text{D.18})$$

where  $(y_1, y_2)$  is a clockwise contour wrapping the line segment from  $y_1$  to  $y_2$ . In our case  $f(y)$  is a rational function, therefore we can pull the contour to infinity so that it picks up the poles of  $f(y)$  to obtain exact expressions.

### D.3 Large energy

Let us consider the large energy limit  $s \rightarrow \infty$  of our ansatz. Since unitarity is imposed for each spin  $\ell$ , we are interested in the limit  $s \rightarrow \infty$  with fixed scattering angle  $\theta$ . In this limit, we find

$$\rho_s^a \rho_t^b \rho_u^c = (-1)^{a+b+c} \left[ 1 + \frac{2\sqrt{4-s_0}}{\sqrt{s}} \left( ia - \frac{\sqrt{2}b}{\sqrt{1-x}} - \frac{\sqrt{2}c}{\sqrt{1+x}} \right) - \frac{4\sqrt{2}(4-s_0)a}{s} \left( \frac{ib}{\sqrt{1-x}} + \frac{ic}{\sqrt{1+x}} + \text{real} \right) + O\left(s^{-\frac{3}{2}}\right) \right]. \quad (\text{D.19})$$

The contribution from the pole terms in our ansatz are real and of order  $1/s$  in this limit and therefore can be neglected. The leading term in (D.19) only contributes to the spin 0 partial wave. The large  $s$  expansion of  $S_0(s)$  is given by

$$S_0(s) = 1 + \frac{is^{\frac{d-3}{2}}}{2^{2d-1}\pi^{\frac{d}{2}-1}\Gamma\left(\frac{d}{2}\right)} \sum_{a,b,c} \alpha_{abc} (-1)^{a+b+c} \left[ 1 + \frac{2ia\sqrt{4-s_0} + \text{real}}{\sqrt{s}} + O\left(\frac{1}{s}\right) \right]. \quad (\text{D.20})$$

Unitarity implies that (for  $d > 4$  the inequality must be saturated)

$$\sum_{a,b,c} \alpha_{abc} (-1)^{a+b+c} a \geq 0. \quad (\text{D.21})$$

If  $d > 2$  then unitarity also implies that

$$\sum_{a,b,c} \alpha_{abc} (-1)^{a+b+c} = 0. \quad (\text{D.22})$$

For  $d = 2$ , the correct condition is

$$32\sqrt{4-s_0} \sum_{a,b,c} \alpha_{abc} (-1)^{a+b+c} a \geq \left[ \sum_{a,b,c} \alpha_{abc} (-1)^{a+b+c} \right]^2. \quad (\text{D.23})$$

For  $\ell > 0$  (even) we find

$$S_\ell(s) = 1 - i s^{\frac{d-4}{2}} I_\ell \sum_{a,b,c} \alpha_{abc} (-1)^{a+b+c} a \left[ 1 + \frac{i 2\sqrt{4-s_0} b + \text{real}}{\sqrt{s}} + O\left(\frac{1}{s}\right) \right], \quad (\text{D.24})$$

where<sup>28</sup>

$$I_\ell = 4\sqrt{2}\sqrt{4-s_0} \int_{-1}^1 dx (1-x^2)^{\frac{d-3}{2}} P_\ell^{(d)}(x) \frac{1}{\sqrt{1+x}} > 0. \quad (\text{D.25})$$

Therefore, unitarity implies (for  $d > 5$  the inequality must be saturated)

$$\sum_{a,b,c} \alpha_{abc} (-1)^{a+b+c} ab \leq 0. \quad (\text{D.26})$$

For  $d > 3$  unitarity also implies

$$\sum_{a,b,c} \alpha_{abc} (-1)^{a+b+c} a = 0. \quad (\text{D.27})$$

For  $d = 3$  we find

$$40\pi \sum_{a,b,c} \alpha_{abc} (-1)^{a+b+c} ab \leq - \left[ \sum_{a,b,c} \alpha_{abc} (-1)^{a+b+c} a \right]^2. \quad (\text{D.28})$$

where we used that  $I_\ell < I_2 = \frac{\sqrt{4-s_0}}{10\pi}$  for  $\ell > 2$ .

Where applicable, we have verified the above constraints a posteriori for our numerical solutions and found them satisfied to very good numerical accuracy.

As a final comment, we remark that the unitarity constraints dictate that  $\lim_{s \rightarrow \infty} S_\ell(s) = 1$  for any amplitude within our ansatz with finite  $N_{\text{max}}$ .<sup>29</sup> This property is likely to be too restrictive, and it is therefore worthwhile to try to improve our ansatz with more singular terms compatible with unitarity and analyticity. As a first attempt we added an extra term of the form  $(\rho_s + 1)(\rho_t + 1)^{-1}$  plus  $s, t, u$  permutations, which allows  $\lim_{s \rightarrow \infty} (S_\ell(s) - 1)$  to be non-zero — this modification however did not significantly change any of the results displayed above. In the future we plan to add other more singular terms and investigate their effect in more detail.<sup>30</sup> Finally, the restricted behavior at large  $s$  might also be a source of slow convergence when  $N_{\text{max}} \rightarrow \infty$  we have observed in some cases. This idea is also corroborated by the two dimensional analysis in appendix G.

<sup>28</sup>For  $d \leq 2$  the integral  $I_\ell$  is divergent. The origin of this divergence is that we can only use the large  $s$  form of the integrand for  $(1-x)s \gg 1$  and  $(1+x)s \gg 1$ . The effect of this can be taken into account by including the  $s$ -dependence  $I_\ell \sim s^{\frac{2-d}{2}}$  for  $d < 2$  and  $I_\ell \sim \log s$  for  $d = 2$ .

<sup>29</sup>With the exception of  $d = 5$  where it is possible to obtain  $\lim_{s \rightarrow \infty} S_\ell(s) \neq 1$  for  $\ell > 0$ .

<sup>30</sup>We also deem it likely that there exists a higher-dimensional version of the two-dimensional construction discussed in footnote 5 that would lead to unbounded couplings, but we again expect the associated essential singularity to be in conflict with causality.

## D.4 Large spin

The partial waves can also be written in terms of an hypergeometric function,

$$P_\ell^{(d)}(x) = \frac{2^{1-2d}\pi^{\frac{1}{2}-\frac{d}{2}}}{\Gamma\left(\frac{d-1}{2}\right)} {}_2F_1\left(-\ell, d+\ell-2; \frac{d-1}{2}; \frac{1-x}{2}\right). \quad (\text{D.29})$$

It is convenient to define

$$Q_\ell^{(d)}(x) = -\frac{\Gamma(l+1)(x-1)^{2-d-\ell}}{\pi^{\frac{d}{2}-1}2^{2d+\ell-1}\Gamma\left(\frac{d}{2}+\ell\right)} {}_2F_1\left(d+\ell-2, \frac{d+2\ell-1}{2}; d+2\ell-1; \frac{2}{1-x}\right) \quad (\text{D.30})$$

such that

$$\text{Disc}\left[(x^2-1)^{\frac{d-3}{2}}Q_\ell^{(d)}(x)\right] = 2\pi i(1-x^2)^{\frac{d-3}{2}}P_\ell^{(d)}(x), \quad -1 < x < 1. \quad (\text{D.31})$$

Notice that for integer  $d$  the function  $Q_\ell^{(d)}(x)$  has no monodromy around  $x = \infty$ . We will work in the sheet where  $Q_\ell^{(d)}(x)$  only has a branch cut from  $x = -1$  to  $x = 1$ . The factor  $(x^2-1)^{\frac{d-3}{2}} = x^{d-3}(1-x^{-2})^{\frac{d-3}{2}}$  has the same analytic properties. Then we can write

$$\int_{-1}^1 dx (1-x^2)^{\frac{d-3}{2}} P_\ell^{(d)}(x) M(s, x) = \frac{1}{2\pi i} \oint_C dx (x^2-1)^{\frac{d-3}{2}} Q_\ell^{(d)}(x) M(s, x), \quad (\text{D.32})$$

where the contour  $C$  encircles the real segment  $[-1, 1]$  clockwise and  $M(s, x)$  denotes the amplitude  $M(s, t)|_{t \rightarrow \frac{1}{2}(s-4)(x-1)}$ . Since

$$Q_\ell^{(d)}(x) \approx -\frac{1}{2^{\frac{3d}{2}}\pi^{\frac{d-2}{2}}\ell^{\frac{d-2}{2}}(x^2-1)^{\frac{d-2}{4}}(x+\sqrt{x^2-1})^{\ell+\frac{d-2}{2}}}, \quad (\text{D.33})$$

for large  $\ell$  and  $x^2 > 1$ , we can expand the contour and drop the contribution from infinity. At large spin, the integral will be dominated by the singularity of  $M(s, x)$  closer to the origin  $x = 0$ . Generically, this will come from the poles associated with stable particles. More precisely,

$$\int_{-1}^1 dx (1-x^2)^{\frac{d-3}{2}} P_\ell^{(d)}(x) M(s, x) = \frac{1}{2\pi i} \oint_C dx (x^2-1)^{\frac{d-3}{2}} Q_\ell^{(d)}(x) M(s, x) \quad (\text{D.34})$$

$$= -\frac{1}{2\pi i} \left[ \int_{-\infty}^{-x_1(s)} dx + \int_{x_1(s)}^{\infty} dx \right] (x^2-1)^{\frac{d-3}{2}} Q_\ell^{(d)}(x) [M(s, x+i\epsilon) - M(s, x-i\epsilon)] \quad (\text{D.35})$$

$$= -\frac{1}{i\pi} \int_{x_1(s)}^{\infty} dx (x^2-1)^{\frac{d-3}{2}} Q_\ell^{(d)}(x) [M(s, x+i\epsilon) - M(s, x-i\epsilon)] \quad (\text{D.36})$$

where  $x_1(s)$  is determined from

$$t(s, x) = m_1^2 \quad \Rightarrow \quad x_1(s) = 1 + \frac{2m_1^2}{s-4m^2}. \quad (\text{D.37})$$



In fact, the pole  $\frac{-g^2}{t-m_1^2}$  contributes

$$\int_{-1}^1 dx (1-x^2)^{\frac{d-3}{2}} P_\ell^{(d)}(x) M(s, x) \approx -\frac{2g^2}{s-4m^2} (x_1(s)^2 - 1)^{\frac{d-3}{2}} Q_\ell^{(d)}(x_1(s)), \quad (\text{D.38})$$

which decays exponentially with  $l$ . Notice that this gives a purely imaginary contribution to  $S_\ell(s)$  (see equation (3.5)), which by itself would violate unitarity. However, unitarity can be restored with a small real contribution of the order of the square of (D.38). At large  $l$ , this requires that we match the exponential behaviour

$$\left(x_1(s) + \sqrt{x_1(s)^2 - 1}\right)^2 = x_2(s) + \sqrt{x_2(s)^2 - 1} \quad \Leftrightarrow \quad \frac{m_2^2}{4m_1^2} = 1 + \frac{m_1^2}{s-4m^2}. \quad (\text{D.39})$$

In other words, unitarity can be restored with another particle or particles of total invariant mass squared  $m_2^2 \geq 4m_1^2$ . This is what happens in perturbation theory where the box diagram restores unitarity of the tree-level exchanges.

Let us now study the contribution from the polynomial terms  $\rho_s^a \rho_t^b \rho_u^c$  in our ansatz. The discontinuity of  $M$  for  $x > 1$  comes from

$$\rho(t(s, x+i\epsilon))^b - \rho(t(s, x-i\epsilon))^b \approx 2ib \sqrt{\frac{2s-8m^2}{4m^2-s_0}} \sqrt{x-x_*(s)} \quad (\text{D.40})$$

where

$$x_*(s) = \frac{s+4m^2}{s-4m^2} \quad (\text{D.41})$$

and we only kept the leading behaviour of the discontinuity near its lower end  $x_*(s)$ . Similarly, we can approximate

$$Q_\ell^{(d)}(x) \approx Q_\ell^{(d)}(x_*(s)) \exp\left[-\ell \frac{s-4m^2}{4m\sqrt{s}} (x-x_*(s))\right] \quad (\text{D.42})$$

and find

$$\int_{-1}^1 dx (1-x^2)^{\frac{d-3}{2}} P_\ell^{(d)}(x) \rho_s^a \rho_t^b \rho_u^c \approx -\frac{1}{i\pi} \rho_s^a \rho_{-s}^c (x_*(s)^2 - 1)^{\frac{d-3}{2}} Q_\ell^{(d)}(x_*(s)) \times \quad (\text{D.43})$$

$$\times \int_{x_*(s)}^\infty dx \exp\left[-\ell \frac{s-4m^2}{4m\sqrt{s}} (x-x_*(s))\right] 2ib \sqrt{\frac{2s-8m^2}{4m^2-s_0}} \sqrt{x-x_*(s)} \quad (\text{D.44})$$

$$= -\frac{b}{\sqrt{\pi} \ell^{\frac{3}{2}}} \frac{8m^{\frac{3}{2}} s^{\frac{3}{4}}}{s-4m^2} \sqrt{\frac{2}{4m^2-s_0}} \rho_s^a \rho_{-s}^c (x_*(s)^2 - 1)^{\frac{d-3}{2}} Q_\ell^{(d)}(x_*(s)) \quad (\text{D.45})$$

$$\approx \frac{F(s)}{\ell^{\frac{d+1}{2}} \left(x_*(s) + \sqrt{x_*(s)^2 - 1}\right)^\ell} b \rho_s^a \rho_{-s}^c \quad (\text{D.46})$$

where  $F(s) > 0$  for  $s > 4m^2$ . Notice that at large  $\ell$  the leading contribution comes from  $t \approx 4m^2$  which implies that  $\rho_u \rightarrow \rho_{-s}$ . Unitarity implies  $\text{Re } S_\ell(s) \leq 1$  which at large  $\ell$  becomes

$$\sum_{a,b,c} \alpha_{abc} b (\text{Im } \rho_s^a) \rho_{-s}^c \geq 0. \quad (\text{D.47})$$

Notice that this condition is independent of the spin  $\ell$  and of the spacetime dimension  $d$ . This justifies our numerical procedure of truncating the unitarity conditions at some value of the spin  $\ell_{\max} \gg 1$ . Writing  $\rho_s = e^{i\phi}$  with  $\phi \in [0, \pi]$ , equation (D.47) can be written as

$$\sum_{a,b,c} \alpha_{abc} (-1)^c b \sin(a\phi) \left[ \frac{\sqrt{1 + y_0 \cos \frac{\phi}{2}} - \cos \frac{\phi}{2}}{\sqrt{1 + y_0 \cos \frac{\phi}{2}} + \cos \frac{\phi}{2}} \right]^c \geq 0, \quad \forall \phi \in [0, \pi], \quad (\text{D.48})$$

where  $y_0 = \frac{4m^2 + s_0}{4m^2 - s_0} > -1$ .

The constraints (D.47) are linear constraints on the numerical coefficients and can easily be taken into account in our numerical code (again by sampling for a discrete set of values of  $s$ ). We have run several of our analyses both with and without this additional constraint. As expected, the effect of the additional term decreases with the maximum spin  $\ell_{\max}$  for which we manifestly check the unitarity constraints. For the values  $\ell_{\max}$  used in our plots the effect of including (D.47) is always small and amounts to maybe to a one percent change in the final result.

## D.5 Threshold expansion and elastic unitarity

Here we shall discuss the threshold behaviour of amplitudes satisfying our ansatz. We start with the expression for the amplitude,

$$M(s, t, u) = \sum_{a,b,c=0}^{+\infty} \alpha_{abc} \rho_s^a \rho_t^b \rho_u^c + \text{poles} \quad (\text{D.49})$$

At threshold the poles become constants and are irrelevant. This is not so for threshold poles which are discussed separately below. Define  $w := \sqrt{s-4}$ . Then for  $s \rightarrow 4^+$  above the cut we have

$$\rho_s = 1 + 2 \sum_{n=1}^{+\infty} \left( \frac{i}{\sqrt{4-s_0}} \right)^n w^n = 1 + \frac{2i}{\sqrt{4-s_0}} w + \dots \quad (\text{D.50a})$$

$$\rho_t^b \rho_u^c = \sum_{k=0}^{\infty} w^{2k} \left( \sum_{n+m=k} c_{n,m} (1-x)^n (1+x)^m \right) \quad (\text{D.50b})$$

Recall that in our conventions the partial waves take the form:

$$S_\ell(s) = 1 + i \frac{(s-4)^{\frac{d-2}{2}}}{\sqrt{s}} \int_{-1}^1 dx (1-x^2)^{\frac{d-3}{2}} P_\ell^{(d)}(x) M(s, t)|_{t \rightarrow \frac{1}{2}(s-4)(x-1)}. \quad (\text{D.51})$$

The leading contribution for the spin  $\ell$  partial wave corresponds to the  $k = \ell$  term in the above, leading to

$$S_\ell(s) = 1 - b_\ell w^{d-1+2\ell} + i a_\ell w^{d-2+2\ell} + \dots, \quad (\text{D.52})$$

with real  $a_\ell, b_\ell$ . These are linear combinations of the coefficients  $\alpha_{abc}$  in our ansatz. Unitarity near threshold imposes:

$$b_\ell \geq 0, \quad d \geq 2, \quad \ell \geq 0; \quad (\text{D.53a})$$

$$a_0 = 0, \quad d = 2, \quad \ell = 0; \quad (\text{D.53b})$$

$$b_0 \geq a_0^2/2, \quad d = 3, \quad \ell = 0. \quad (\text{D.53c})$$

Near threshold we have the expansion of  $S_\ell(s)$  in terms of the phase shift,

$$S_\ell(s) = e^{2i\delta_\ell(s)} \sim 1 + 2i\delta_\ell(s) - 2\delta_\ell(s)^2 + \dots \quad (\text{D.54})$$

Absence of particle production would imply reality of  $\delta_\ell(s)$ , and hence a measure of the inelasticity of the amplitude at the threshold is

$$\frac{\text{Re}[1 - S_\ell(s)]}{[\text{Im} S_\ell(s)]^2} = O[(s - 4)^{-\ell}]. \quad (\text{D.55})$$

We see that for positive spin we generically get a divergent result in the threshold limit. This means that our ansatz does not automatically give an amplitude which becomes purely elastic as we approach threshold, unlike what we would expect on physical grounds. In order for purely elastic scattering to hold, we would have had to impose order  $\ell$  linear constraints on the coefficients of the threshold expansion of the spin  $\ell$  partial wave. We did *not* impose these in our numerical computations. However, experimentally we do find that as the number of parameters in our ansatz is increased, the coefficients in the threshold expansion seem to decrease.

## E Non-relativistic limit

Consider a scalar  $\phi$  of mass  $m$  interacting with itself via the exchange of a second heavy scalar  $\Phi$  with mass  $m_b = 2m - \epsilon$  with small  $\epsilon$ . We can think of  $\Phi$  as a loosely bound state of two  $\phi$  particles with binding energy  $\epsilon$ . The two body amplitude for  $\phi + \phi$  scattering contains a pole at  $s = m_b^2$  due to virtual production of a  $\Phi$  which is just below the two-particle threshold at  $s = (2m)^2$ . The residue of this pole  $g^2$  is the square of the  $\phi\phi\Phi$  coupling. Now consider low energy  $\phi + \phi$  scattering and write  $s = (2m + E)^2$  where  $E$  is the centre of mass energy after subtraction of the rest mass. The  $s$ -channel pole of the amplitude is given by<sup>31</sup>

$$M^{\text{pole}} = \frac{m^{5-d}g^2}{s - m_b^2} \sim \frac{m^{5-d}g^2/\epsilon}{4m(E/\epsilon + 1)} \quad (\text{E.1})$$

where we have assumed small  $E$  and  $\epsilon$ . The  $l = 0$  phase shift inherits this pole through the relation

$$\frac{\sqrt{s}}{i(s - 4m^2)^{\frac{d-2}{2}}}(S_0(s) - 1) = \int_{-1}^1 dx (1 - x^2)^{\frac{d-3}{2}} P_0^{(d)}(x) M(s, t)|_{t \rightarrow \frac{1}{2}(s-4)(x-1)} \quad (\text{E.2})$$

Plugging (E.1) into (E.2) and zooming in on the pole of the phase shift at  $s = (2m - \epsilon)^2$  we have

$$2^{3-d}m(m\epsilon)^{1-d/2}S_0^{\text{pole}}(E/\epsilon) \sim \frac{2^{1-2d}\pi^{1-\frac{d}{2}}}{\Gamma(\frac{d}{2})} \frac{m^{5-d}g^2/\epsilon}{4m(E/\epsilon + 1)} \quad (\text{E.3})$$

We write the pole of the phase shift as  $g_{NR}^2/(E/\epsilon + 1)$  where  $g_{NR}^2$  is the residue in units of the binding energy  $\epsilon$ . We then have

$$g^2 \rightarrow 2^{4+d}\pi^{\frac{d}{2}-1}\Gamma(d/2)g_{NR}^2(\epsilon/m)^{2-\frac{d}{2}} \quad (\text{E.4})$$

---

<sup>31</sup>The factor  $m^{5-d}$  is to make the coupling  $g^2$  dimensionless.

We will show below that there is a bound on the non-relativistic coupling  $g_{NR}^2 \leq 2^2$ . Note that this correctly predicts the behaviour

$$g_{1+1}^2 \leq 2^7 (\epsilon/m)^{3/2} \quad (\text{E.5})$$

in 1 + 1 dimensions [2]. Moreover, this limit has been studied extensively in 3 + 1 dimensions ( $d = 3$ ) [29, 30]. These authors find (adding a factor of 2 to their results to account for identical particles)

$$g_{3+1}^2 \leq 2^8 \pi \sqrt{\epsilon/m} \quad (\text{E.6})$$

and thus we find perfect agreement with (E.4).

Let us now derive the bound on  $g_{NR}^2$  quoted above. Recall that we are considering a very weakly bound state with binding energy  $\epsilon$ . We wish to obtain the behaviour of  $g_{\max}^2(\epsilon/m)$  for small  $\epsilon/m$ . Thus we concentrate on “slow” physics at energies  $E \sim \epsilon$  (recall  $E$  is the centre of mass energy after removal of the rest mass). Formally, in the phase shift we consider  $s \rightarrow \bar{s}\epsilon^2$  and consider finite  $\bar{s}$  as  $\epsilon \rightarrow 0$ . Any singularities of the phase shift that are a finite distance (in  $s$ ) from the two-particle threshold — e.g. the left cut and inelastic thresholds — will be infinitely far away in  $\bar{s}$  and thus only contribute through positive powers of  $\epsilon$ . We can thus neglect these singularities to obtain the leading behaviour of  $g_{\max}^2(\epsilon/m)$  and consider a non-relativistic phase shift  $S_{NR}(\bar{E})$  with only a right-hand cut starting at  $\bar{E} = 0$  and a single bound-state pole at  $\bar{E} = -1$ , where  $\bar{E} = E/\epsilon$ . Since this phase shift is bounded by unitarity along the cut and cannot grow faster than a constant at infinity then the residue of the pole can easily be bounded by maximum modulus theorem. Perhaps the cleanest way to derive the precise value of the bound is to consider the change of coordinates

$$x(E) = \frac{1 - (-\bar{E})^{1/2}}{1 + (-\bar{E})^{1/2}}, \quad \bar{E}(x) = -\frac{(x-1)^2}{(x+1)^2} \quad (\text{E.7})$$

which maps the E-plane minus the positive real axis to the unit disk and maps the bound state pole to the origin

$$\frac{g_{NR}^2}{\bar{E}(x) - 1} \sim \frac{g_{NR}^2}{4x} \quad (\text{E.8})$$

Now note that the function  $f(x) = x S_{NR}(x)$  is analytic throughout the unit disk and obeys  $|f| \leq 1$  on the boundary due to unitarity. Thus maximum modulus theorem implies  $1 \geq f(0) = g_{NR}^2/4$  which is the desired bound.

## F Semidefinite programming implementation

Consider an ansatz as in (3.10), truncated such that  $a+b+c \leq N_{\max}$ . After eliminating the redundant monomials as described in appendix B, we are left with a finite subset of the  $\alpha_{abc}$ , which together with the coupling  $g^2$  completely determine the amplitude. Let us group these real coefficients into a vector that we call  $\vec{\eta}$ , so we can schematically write (3.10) as

$$M(s, t, 4 - s - t) = \vec{\eta} \cdot \overrightarrow{M(s, t)} \quad (\text{F.1})$$

with  $\overrightarrow{M(s,t)}$  the vector of functions of  $s$  and  $t$  that each coefficient multiplies. We then substitute into the partial amplitude projection (3.5) and get, schematically,

$$S_\ell(s) = 1 + i\vec{\eta} \cdot \overrightarrow{f_\ell(s)} \quad (\text{F.2})$$

with  $\overrightarrow{f_\ell(s)}$  defined in the obvious way as the integral of  $\overrightarrow{M(s,t)}$  against the Gegenbauer polynomials with the right prefactor. The unitarity constraints  $|S_\ell(s)|^2 \leq 1$  now dictate that for all physical  $\ell$  and  $s$  we must have

$$\left(1 - \vec{\eta} \cdot \vec{I}\right)^2 + (\vec{\eta} \cdot \vec{R})^2 \leq 1 \quad \Leftrightarrow \quad U \equiv 2\vec{\eta} \cdot \vec{I} - (\vec{\eta} \cdot \vec{I})^2 - (\vec{\eta} \cdot \vec{R})^2 \geq 0 \quad (\text{F.3})$$

with  $\vec{R} = \text{Re}[\overrightarrow{f_\ell(s)}]$  and  $\vec{I} = \text{Im}[\overrightarrow{f_\ell(s)}]$ . This constraint can be re-phrased as a semidefiniteness condition. Indeed, consider the matrix

$$M := \begin{pmatrix} 1 + \vec{\eta} \cdot \vec{R} & 1 - \vec{\eta} \cdot \vec{I} \\ 1 - \vec{\eta} \cdot \vec{I} & 1 - \vec{\eta} \cdot \vec{R} \end{pmatrix} \quad (\text{F.4})$$

The eigenvalues of this matrix are precisely

$$\lambda_\pm = 1 \pm \sqrt{1 - U} \quad (\text{F.5})$$

As befits a Hermitian matrix, they are always real since  $U \leq 1$  by construction. It is now clear that

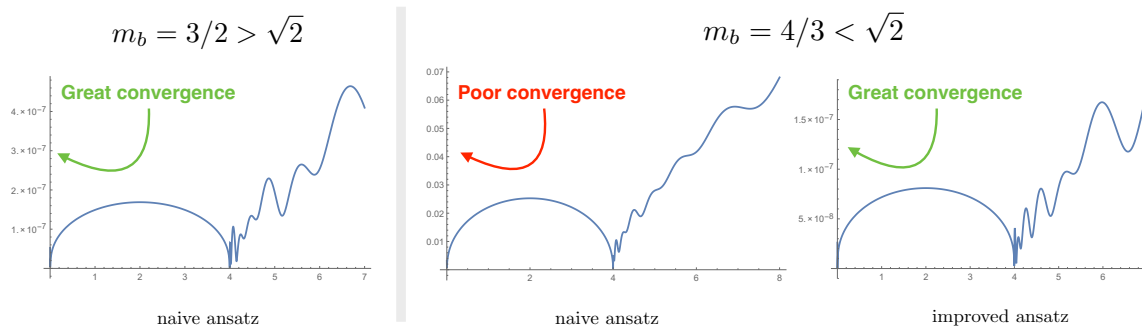
$$M \succeq 0 \Leftrightarrow U \geq 0. \quad (\text{F.6})$$

and the unitarity constraints are therefore precisely those of a semidefinite program.

We need to choose a grid of values of  $s$  and a finite set of spins  $\ell$  for which to test the unitarity constraints. We found it sufficient to take approximately 200 values of  $s$ , interspersed uniformly along the upper half of the unit circle in the  $\rho_s$  variable defined in the main text. We observed no significant change in the results by taking a more refined  $s$  grid, or by distributing the points differently along the unit circle. The maximal value of the spin  $\ell_{\max}$  is indicated in the various plots. Notice that  $\ell_{\max}$  needs to be sufficiently big since otherwise the extremal value completely destabilizes — see for example the data points in figure 6 with  $\ell_{\max} = 10$  for large  $N_{\max}$ . In practice we observed convergence by taking  $\ell_{\max}$  at least as large as  $N_{\max}$ , and for the scattering length computations we needed at least  $N_{\max} + 4$ . Increasing  $\ell_{\max}$  beyond these values did not affect our results.

In our numerical computations we did find it necessary to retain very high precision, generally at least 1000 binary digits. This appears to stem from the approximate redundancy that remains even after imposing the polynomial constraint B. To illustrate this we can for example compute a derivative like

$$\begin{aligned} \frac{\partial^2}{\partial s^2} \left( \sum_{a,b,c} \alpha_{abc} \rho^a(s) \rho^b(t) \rho^c(4-s-t) \right) \Big|_{s=t=4/3} &= \\ &= \frac{9}{256} \left( \alpha_{100} + \alpha_{001} + \frac{1}{2} \alpha_{200} + \frac{1}{2} \alpha_{002} - \frac{1}{2} \alpha_{101} \right) \end{aligned} \quad (\text{F.7})$$



**Figure 15.** Plot of  $|(S_{\text{num}} - S_{\text{analytic}})/S_{\text{analytic}}|$ , that is of the relative mismatch in the numerical solution in this two dimensional example where the analytic solution is available. In all these plots we use  $\Lambda = 20$  and check unitarity in a small grid of 40 points. With these parameters, `mathematica`'s built-in `FindMaximum` suffices and produces an outcome in about two or three seconds. We see on the left that for  $m_b > \sqrt{2}m$  the agreement is spectacular with the most naive ansatz (G.4) while in the middle we see that the agreement is much worse (a few percent off) with the same ansatz when  $m_b < \sqrt{2}m$ . On the right we see that this is neatly fixed - leading again to a perfect convergence - by simply adopting an improved ansatz as in (G.5).

In a typical solution we find that this derivative is rather modest in magnitude, of order  $10^2$  or so, whereas the individual coefficients can be very large, of order  $10^{24}$  in some solutions. These kind of cancellations require high precision.

We have performed all the numerical computations in section 4 with `sdpb` [31]. Details of the computations like parameter settings are available from the authors upon request.

## G Slow convergence on a simple 2D example

In this appendix we revisit once more the two dimensional problem considered in section 2 but this time done in the language of the  $M$  amplitude rather than  $S$ . In two dimensions the two are simply related by

$$S(s, t) - 1 = \frac{1}{2\sqrt{st}} \times M(s, t), \quad s + t = 4m^2. \quad (\text{G.1})$$

and unitarity then reads

$$\text{Im}(M(s, t)) - \frac{1}{4\sqrt{-st}} |M(s, t)|^2 \leq 0 \quad \text{for} \quad s = 4m^2 - t > 4m^2. \quad (\text{G.2})$$

This discussion will provide us with a simple example of numerics which work yet converge very slowly until we slightly improve our ansatz and thus completely solve this convergence issue.

To be concrete we consider here the case where there is a single bound-state with mass  $m_b$  whose coupling we maximize. The S-matrix with the largest coupling and such bound-state is given by [2]

$$S_{\text{max } g} = \text{sign}(m_b - \sqrt{2}m) \times \frac{\sqrt{s(s - 4m^2)} + \sqrt{m_b^2(4m^2 - m_b^2)}}{\sqrt{s(s - 4m^2)} - \sqrt{m_b^2(4m^2 - m_b^2)}} \quad (\text{G.3})$$

At high energies the S-matrix approaches  $+1$  for  $m_b > \sqrt{2}m$  and  $-1$  for  $m_b < \sqrt{2}m$  and this leads to a very different behavior when translated to the amplitude  $M$ . In particular, for a light bound state  $m_b < \sqrt{2}m$  we see that the amplitude  $M$  in (G.1) must diverge at high energies so that the right hand side approaches  $-2$ . This is hard for an ansatz ala (2.3) to achieve, that is it would require that the sum in

$$M_{\text{naive}}(s, t) = -\frac{\hat{g}^2}{s - m_b^2} - \frac{\hat{g}^2}{t - m_b^2} + \sum_{a,b=0}^{\Lambda} c_{ab} \rho_s^a \rho_t^b \tag{G.4}$$

to develop a divergence as  $s = 4m^2 - t \rightarrow \infty$  which corresponds to  $\rho_s, \rho_t \rightarrow -1$ . Such non-analytic behavior at the boundary of the unit disc *can* be achieved *but* a numerically sufficiently accurate approximation requires very large  $\Lambda$ .

In this case there is however a very obvious improvement which is to simply allow for a divergence at large energies which is after all allowed by unitarity and write down instead an ansatz of the form

$$M_{\text{improved}}(s, t) = -\frac{\hat{g}^2}{s - m_b^2} - \frac{\hat{g}^2}{t - m_b^2} + \sum_{a,b=0}^{\Lambda-2} c_{ab} \rho_s^a \rho_t^b + \frac{\beta}{\sqrt{\rho_s + 1}\sqrt{\rho_t + 1}} + \frac{\tilde{\beta}}{(\rho_s + 1)(\rho_t + 1)} \tag{G.5}$$

This immediately allows for a more general high energy behavior and thus an extreme improvement in convergence as illustrated in figure 15.

The moral of this story seems to be that we better allow for flexible ansätze which can easily capture various analytic properties of scattering amplitudes if we want to achieve optimal convergence. In this simple two dimensional example, allowing for an ansatz with a more flexible high energy behavior led to a drastic improvement in the numerics.

**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] M.F. Paulos, J. Penedones, J. Toledo, B.C. van Rees and P. Vieira, *The S-matrix bootstrap. Part I: QFT in AdS*, *JHEP* **11** (2017) 133 [[arXiv:1607.06109](https://arxiv.org/abs/1607.06109)] [[INSPIRE](#)].
- [2] M.F. Paulos, J. Penedones, J. Toledo, B.C. van Rees and P. Vieira, *The S-matrix bootstrap II: two dimensional amplitudes*, *JHEP* **11** (2017) 143 [[arXiv:1607.06110](https://arxiv.org/abs/1607.06110)] [[INSPIRE](#)].
- [3] M. Creutz, *Rigorous bounds on coupling constants in two-dimensional field theories*, *Phys. Rev.* **D 6** (1972) 2763 [[INSPIRE](#)].
- [4] M. Hogervorst and S. Rychkov, *Radial coordinates for conformal blocks*, *Phys. Rev.* **D 87** (2013) 106004 [[arXiv:1303.1111](https://arxiv.org/abs/1303.1111)] [[INSPIRE](#)].
- [5] 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](https://arxiv.org/abs/1407.5597)] [[INSPIRE](#)].

- [6] S. Weinberg, *The quantum theory of fields. Volume 1: foundations*, Cambridge University Press, Cambridge, U.K. (2005) [INSPIRE].
- [7] J. Bros, H. Epstein and V. Glaser, *A proof of the crossing property for two-particle amplitudes in general quantum field theory*, *Commun. Math. Phys.* **1** (1965) 240 [INSPIRE].
- [8] A. Martin, *Rigorous results from theory and unitarity*, CERN-TH-1181, CERN, Geneva, Switzerland (1970).
- [9] J.D. Bjorken and S.D. Drell, *Relativistic quantum fields*, (1965) [INSPIRE].
- [10] A. Martin, *Scattering theory: unitarity, analyticity and crossing*, *Lect. Notes Phys.* **3** (1969) 1 [INSPIRE].
- [11] F.J. Yndurain, *Rigorous constraints, bounds and relations for scattering amplitudes*, *Rev. Mod. Phys.* **44** (1972) 645 [INSPIRE].
- [12] S. Caron-Huot, Z. Komargodski, A. Sever and A. Zhiboedov, *Strings from massive higher spins: the asymptotic uniqueness of the Veneziano amplitude*, *JHEP* **10** (2017) 026 [arXiv:1607.04253] [INSPIRE].
- [13] A. Sever and A. Zhiboedov, *On fine structure of strings: the universal correction to the Veneziano amplitude*, *JHEP* **06** (2018) 054 [arXiv:1707.05270] [INSPIRE].
- [14] F. Forstneric, *Stein manifolds and holomorphic mappings: the homotopy principle in complex analysis*, Springer, (2011).
- [15] V.F. Müller, *An upper bound on the coupling constant of scalar mesons*, *Nuovo Cim.* **A 42** (1966) 185.
- [16] G. Auberson, L. Epele and F.R.A. Simao, *Almost optimality of an axiomatic bound for  $\pi^0\pi^0$  scattering*, *Nucl. Phys.* **B 133** (1978) 266 [INSPIRE].
- [17] C. Lopez and G. Mennessier, *Bounds on the  $\pi^0\pi^0$  amplitude*, *Nucl. Phys.* **B 118** (1977) 426 [INSPIRE].
- [18] G. Auberson and G. Mennessier, *On the optimality of axiomatic constraints for  $\pi^0\pi^0$  scattering: the case of the lower bounds*, *Nucl. Phys.* **B 162** (1980) 440 [INSPIRE].
- [19] S.R. Coleman, *More about the massive Schwinger model*, *Annals Phys.* **101** (1976) 239 [INSPIRE].
- [20] I. Caprini and P. Dita, *A new method for deriving rigorous results on  $\pi\pi$  scattering*, *J. Phys.* **A 13** (1980) 1265 [INSPIRE].
- [21] C. Lopez and G. Mennessier, *A new absolute bound on the  $\pi^0\pi^0$  S-wave scattering length*, *Phys. Lett.* **B 58** (1975) 437 [INSPIRE].
- [22] L. Lukaszuk and A. Martin, *Absolute upper bounds for  $\pi\pi$  scattering*, *Nuovo Cim.* **A 52** (1967) 122 [INSPIRE].
- [23] B. Bonnier and R.V. Mau, *Connection between the Wigner inequalities and analyticity and unitarity*, *Phys. Rev.* **165** (1968) 1923 [INSPIRE].
- [24] B. Bonnier, *Derivation and implications of rigorous absolute bounds for  $\pi\pi$  partial waves up to 1 GeV*, *Nucl. Phys.* **B 95** (1975) 98 [INSPIRE].
- [25] P. Dorey, *Exact S matrices*, in *Conformal field theories and integrable models. Proceedings, Eotvos Graduate Course*, Budapest, Hungary, 13–18 August 1996, pg. 85 [hep-th/9810026] [INSPIRE].



- [26] S.O. Aks, *Proof that scattering implies production in quantum field theory*, *J. Math. Phys.* **6** (1965) 516.
- [27] A.J. Dragt, *Amount of four-particle production required in S-matrix theory*, *Phys. Rev.* **156** (1967) 1588.
- [28] S. Mandelstam, *Analytic properties of transition amplitudes in perturbation theory*, *Phys. Rev.* **115** (1959) 1741 [INSPIRE].
- [29] M.A. Ruderman and S. Gasiorowicz, *Limits on coupling constants in field theories with finite sources*, *Nuovo Cim.* **8** (1958) 861.
- [30] V. Gribov, Y.B. Zel'dovich and A. Perelomov, *On the maximal charge for a given mass of the bound state*, *Sov. Phys. JETP* **13** (1961) 836 [*Zh. Eksp. Teor. Fiz.* **40** (1961) 1190] [INSPIRE].
- [31] D. Simmons-Duffin, *A semidefinite program solver for the conformal bootstrap*, *JHEP* **06** (2015) 174 [arXiv:1502.02033] [INSPIRE].