Published for SISSA by Springer

Received: December 10, 2021
Revised: March 1, 2022
Accepted: March 1, 2022
Published: March 24, 2022

NNLO positivity bounds on chiral perturbation theory for a general number of flavours

Benjamin Alvarez,[a] **Johan Bijnens**[b] **and Mattias Sjö**[b]

[a] *Aix Marseille Université, Université de Toulon, CNRS, CPT, Marseille, France*
[b] *Department of Astronomy and Theoretical Physics, Lund University, Box 43, SE 22100 Lund, Sweden*

E-mail: benjamin-alvarez@univ-tln.fr, bijnens@thep.lu.se, mattias.sjo@thep.lu.se

Abstract: We present positivity bounds, derived from the principles of analyticity, unitarity and crossing symmetry, that constrain the low-energy constants of chiral perturbation theory. Bounds are produced for 2, 3 or more flavours in meson-meson scattering with equal meson masses, up to and including next-to-next-to-leading order (NNLO), using the second and higher derivatives of the amplitude. We enhance the bounds by using the most general isospin combinations posible (or higher-flavour counterparts thereof) and by analytically integrating the low-energy range of the discontinuities. In addition, we present a powerful and general mathematical framework for efficiently managing large numbers of positivity bounds.

Keywords: Chiral Lagrangians, Effective Field Theories

ArXiv ePrint: 2112.04253

https://doi.org/10.1007/JHEP03(2022)159

JHEP03(2022)159

Contents

JHEP03(2022)159

JHEP03(2022)159

1 Introduction

Chiral perturbation theory (χPT) is the most widespread theory for low-energy quantum chromodynamics (QCD). It is an effective field theory (EFT) which reformulates the non-perturbative behaviour of low-energy QCD as a perturbative theory of new degrees of freedom, physically interpreted as bound states of quarks. When constructed using n light quark flavours, the degrees of freedom are the $n^2 - 1$ light pseudoscalar mesons: the pions for $n = 2$, with the kaons and eta added for $n = 3$. χPT was developed by Gasser & Leutwyler [1, 2] based on earlier work by Weinberg [3]; see [4, 5] for modern introductions with further references.

At leading order in the low-energy expansion, the only parameters of χPT are the meson mass and decay constant, but higher orders introduce a rapidly increasing number of Wilson coefficients or low-energy constants (LECs) which, while in principle derivable from the underlying QCD dynamics, must in practice be seen as unknowns. At next-to-leading order (NLO), the LECs can be measured reasonably well with experimental or lattice methods, although the precision is typically only one or two significant digits. At next-to-next-to-leading order (NNLO), only tentative results are presently available. For a review of LEC measurements, see [6].

All quantum field theories must obey the axioms of unitarity, analyticity and crossing symmetry, and normally do so by construction. However, it turns out that these axioms are not automatically satisfied by EFTs such as χPT when perturbativity is assumed at a fixed order in the expansion. Therefore, imposing the axioms actually adds new information, typically by placing bounds on the scattering amplitudes. Pioneering work was done by Martin [7] before the development of χPT as such. Bounds on NLO two-flavour χPT amplitudes, which in turn translate to bounds on the LECs, were first obtained in [8–10] and extended in [11, 12]. Further improvements were made in [13] and extended to three-flavour χPT in [14]. There is ongoing research in extending these methods, both specific to χPT and with broader scope; recent examples include [15–18].

The method of [13, 14], which serves as the basis of our method, is to apply dispersion relations (a consequence of analyticity) to a meson-meson scattering amplitude decomposed into isospin components (for higher flavours, the Clebsch-Gordan decomposition is used). Then, crossing symmetry and the optical theorem (a consequence of unitarity) are applied to give a positivity condition on the decomposed amplitude. With the amplitude calculated in terms of the LECs to some order, this results in bounds on linear combinations of LECs. More recently, stronger bounds have been obtained in [19, 20] by improving this method; put extremely simply, this was done with more sophisticated use of dispersion relations and crossing symmetry, respectively. Put similarly simply, our work instead improves the handling of the isospin decompositions and the LEC bounds themselves, although some improvements similar to [19] are also made. Perhaps more importantly, we perform the first extension to NNLO χPT with any number of flavours (two flavours was treated in [19]), albeit with the simplification that all mesons have the same mass. The LECs are independent of the chosen masses, although the bounds do depend on the mass. At NLO they depend only on the ratio of the meson mass and the subtraction scale μ, at NNLO also on the ratio of the meson mass and decay constant.

JHEP03(2022)159

Preliminary results of this work are presented in the Lund University master thesis [21]. Our work is structured as follows: section 2 introduces χPT and its LECs; section 3 (backed by appendix A) presents the $2 \to 2$ meson scattering amplitude used to obtain the bounds; section 4 (backed by appendix B) introduces the mathematical framework used to manage them; section 5 (backed by appendix C) presents the method of [13, 14] and the improvements made to it; and section 6 displays the most interesting bounds we obtain, with final remarks given in section 7.

2 Chiral perturbation theory

n-flavour χPT is based around a non-linear sigma model (NLSM), whose degrees of freedom are the $n^2 - 1$ Nambu-Goldstone bosons that arise when the chiral symmetry $G = \mathrm{SU}(n)_L \times \mathrm{SU}(n)_R$ of n-flavour massless QCD is spontaneously broken into its diagonal subgroup $H = \mathrm{SU}(n)_V$. The Goldstone bosons live in the coset space G/H, which is isomorphic to $\mathrm{SU}(n)$.

The presence of quark masses, electroweak interactions, etc. can be accounted for by including four external $n \times n$ flavour-space matrix fields — s (scalar), p (pseudoscalar), v_μ (vector) and a_μ (axial vector)[1] — into the massless QCD Lagrangian. These additions were introduced in [1, 2], and endow the Nambu-Goldstone bosons with masses and interactions that allow them to accurately model the light pseudoscalar mesons, turning the $\mathrm{SU}(n)$ NLSM into χPT proper.

The Nambu-Goldstone boson fields can be organised into a $n \times n$ flavour-space matrix field $u(\phi)$ [24, 25]. Under the chiral transformation $(g_L, g_R) \in G$, $u(\phi)$ transforms as

$$u(\phi) \longrightarrow g_R\, u(\phi)\, h\big[g_L, g_R, u(\phi)\big] = h\big[g_L, g_R, u(\phi)\big]\, u(\phi)\, g_L^\dagger, \tag{2.1}$$

where $h \in H$ is defined by this transformation. By requiring that G can be made local while leaving the extended QCD Lagrangian invariant, it can be shown that

$$\begin{aligned} \chi &\equiv 2B(s + ip) &&\longrightarrow g_R \chi g_L^\dagger, \\ \ell_\mu &\equiv v_\mu - a_\mu &&\longrightarrow g_L \ell_\mu g_L^\dagger - i\partial_\mu g_L g_L^\dagger, \\ r_\mu &\equiv v_\mu + a_\mu &&\longrightarrow g_R r_\mu g_R^\dagger - i\partial_\mu g_R g_R^\dagger, \end{aligned} \tag{2.2}$$

where B is a constant related to the leading-order (LO) meson decay constant and the $\langle \bar{q}q \rangle$ condensate.

It is possible to rewrite $u(\phi), \chi, \ell_\mu, r_\mu$ in a basis of fields that transform entirely in terms of g_L and g_R, as is done in [2] to derive the NLO χPT Lagrangian. We instead choose to follow [26–28] and rewrite them in a basis of fields that all transform as $X \to hXh^\dagger$:

$$\begin{aligned} u_\mu &\equiv i\left[u^\dagger(\partial_\mu - ir_\mu)u - u(\partial_\mu - i\ell_\mu)u^\dagger\right], \\ \chi_\pm &\equiv u^\dagger \chi u^\dagger \pm u \chi^\dagger u, \\ f_\pm^{\mu\nu} &\equiv u F_L^{\mu\nu} u^\dagger \pm u^\dagger F_R^{\mu\nu} u, \end{aligned} \tag{2.3}$$

[1] One can add more types of externals fields to χPT. Examples are symmetric or antisymmetric tensors [22, 23]. These extensions are not relevant for this work.

JHEP03(2022)159

where $F_L^{\mu\nu} \equiv \partial^\mu \ell^\nu - \partial^\nu \ell^\mu - i[\ell^\mu, \ell^\nu]$ and similarly for $F_R^{\mu\nu}$ and r^μ. These transformation properties are conserved under the covariant derivative ∇_μ defined as

$$\nabla_\mu X = \partial_\mu X + [\Gamma_\mu, X], \qquad \Gamma_\mu \equiv \frac{1}{2}\left[u^\dagger(\partial_\mu - ir_\mu)u + u(\partial_\mu - i\ell_\mu)u^\dagger\right]. \tag{2.4}$$

2.1 The χPT Lagrangian

There exists an infinite number of possible Lagrangian terms consistent with the symmetries of χPT. They can be organised into a power-counting hierarchy in the small energy-momentum scale p, where $u_\mu, \nabla_\mu = \mathcal{O}(p)$ and $\chi_\pm, f_\pm^{\mu\nu} = \mathcal{O}(p^2)$. Thus,

$$\mathcal{L}_{\chi\mathrm{PT}} = \mathcal{L}_2 + \mathcal{L}_4 + \mathcal{L}_6 + \ldots, \tag{2.5}$$

where $\mathcal{L}_{2n}$ is $\mathcal{O}(p^{2n})$; odd powers are forbidden by parity. The coefficient of each term in $\mathcal{L}_{2n}$ is a separate LEC.[2]

The LO Lagrangian is

$$\mathcal{L}_2 = \frac{F^2}{4}\langle u_\mu u^\mu + \chi_+\rangle, \tag{2.6}$$

where F is a LEC related to the LO meson decay constant, and $\langle\ldots\rangle$ indicates a trace over flavour-space indices. The LEC of the χ_+ term is $\frac{BF^2}{4}$ as defined in eq. (2.2). By requiring that the kinetic term is canonically normalised, one can fix $u(\phi) = 1 + \frac{it^a\phi^a}{F\sqrt{2}} + \ldots$, where t^a are the generators of SU(n) and Einstein's summation convention is used. The higher-order terms depend on the choice of parametrisation, which influences the computation of amplitudes but not the amplitudes themselves.

The next-to-leading-order (NLO) Lagrangian, which was first determined in [2], is in terms of our basis[3]

$$\begin{aligned}\mathcal{L}_4 = \quad & \hat{L}_0\langle u_\mu u_\nu u^\mu u^\nu\rangle + \hat{L}_1\langle u_\mu u^\mu\rangle^2 + \hat{L}_2\langle u_\mu u_\nu\rangle\langle u^\mu u^\nu\rangle + \hat{L}_3\left\langle (u_\mu u^\mu)^2\right\rangle \\ & + \hat{L}_4\langle u_\mu u^\mu\rangle\langle\chi_+\rangle + \hat{L}_5\langle u_\mu u^\mu \chi_+\rangle + \hat{L}_6\langle\chi_+\rangle^2 + \hat{L}_7\langle\chi_-\rangle^2 + \hat{L}_8\left\langle\chi_+^2 + \chi_-^2\right\rangle \\ & - i\hat{L}_9\langle f_+^{\mu\nu} u_\mu u_\nu\rangle + \hat{L}_{10}\left\langle f_+^{\mu\nu} f_{\mu\nu}^+ - f_-^{\mu\nu} f_{\mu\nu}^-\right\rangle,\end{aligned} \tag{2.7}$$

where the LECs are $\hat{L}_i$. The analogous NNLO Lagrangian with 112 LECs K_i was determined in [27]. The 1862-LEC NNNLO Lagrangian, which we do not use here, was determined in [28].

For small n, the Cayley-Hamilton identity reduces the number of independent terms, and consequently the number of LECs. At $n = 3$, it is standard to eliminate $\hat{L}_0$; the remaining LECs are conventionally labelled L_i with i preserved. At $n = 2$, it is customary to also redefine the LECs slightly, resulting in the l_i of the original Gasser-Leutwyler convention [1]. At NNLO, the 112+3 K_i (ordinary+contact terms) are reduced to 90+4 C_i at $n = 3$ and 52+4 c_i at $n = 2$ as detailed in [27]. For more details on the Lagrangians for different n, see [6, 29].

[2]Some "terms", like the one associated with $\hat{L}_{10}$ in eq. (2.7) below, actually consist of several terms. These transform into each other under the discrete symmetries of the Lagrangian, and must therefore appear with the same LEC.

[3]There are two additional *contact terms* proportional to $\left\langle\chi_+^2 - \chi_-^2\right\rangle$ and $\left\langle f_+^{\mu\nu} f_{\mu\nu}^+ + f_-^{\mu\nu} f_{\mu\nu}^-\right\rangle$. They are needed for renormalisation but make no physical contributions to the amplitudes considered here.

JHEP03(2022)159

The NLO renormalisation was first carried out in [1, 2], and the extension to NNLO in [29]; for more information on χPT renormalisation, see [30]. A slightly altered $\overline{\text{MS}}$ scheme is conventionally used, with renormalisation scale $\mu = 0.77\,\text{GeV}$. The renormalised LECs are denoted X_i^r where $X = \ell, L, \hat{L}$, etc. At $n = 2$ flavours it is conventional to use $\bar{\ell}_i$ instead, related to ℓ_i^r through

$$\ell_i^r = \frac{\gamma_i}{32\pi^2}\left[\bar{\ell}_i + \ln\left(\frac{M_{\text{phys}}^2}{\mu^2}\right)\right] \tag{2.8}$$

where M_{phys} is the chosen meson mass and γ_i are coefficients found in [1]. Effectively, eq. (2.8) sets the renormalisation scale to M_{phys} for $\bar{\ell}_i$.

3 Scattering amplitudes

In this section, and in the remainder of the paper, we will restrict ourselves to a simplified version of χPT. Firstly, we will not include the external (axial) vector fields a_μ, v_μ in the Lagrangian, which essentially amounts to ignoring electroweak corrections to the amplitude. Secondly, we will assume that all mesons have the same mass M_{phys}, as mentioned in the introduction. While this limits the phenomenological applicability of three-flavour χPT, it is a reasonable approximation that simplifies the procedure for obtaining bounds (see section 5). More importantly, the full NNLO amplitude is currently not available in the general-mass case; available results only cover $\pi\pi$ scattering in two- [30, 31] and three-flavour [32] χPT, as well as πK scattering [33], and are not expressed in terms of elementary functions. With equal masses, we normalise all Mandelstam variables so that $s + t + u = 4$.

For the general equal-mass n-flavour scattering process $a + b \to c + d$, there are nine independent flavour structures possible: the six distinct index permutations on $\left\langle t^a t^b t^c t^d \right\rangle$ and the three on $\left\langle t^a t^b \right\rangle \left\langle t^c t^d \right\rangle$. Due to charge conjugation symmetry, a permutation is not independent of its reverse. Thus, the scattering amplitude M may be decomposed as

$$\begin{aligned} M(s,t,u) = \quad & [\left\langle t^a t^b t^c t^d \right\rangle + \left\langle t^d t^c t^b t^a \right\rangle] B(s,t,u) \\ & + [\left\langle t^a t^c t^d t^b \right\rangle + \left\langle t^b t^d t^c t^a \right\rangle] B(t,u,s) \\ & + [\left\langle t^a t^d t^b t^c \right\rangle + \left\langle t^c t^b t^d t^a \right\rangle] B(u,s,t) \\ & + \delta^{ab}\delta^{cd} C(s,t,u) + \delta^{ac}\delta^{bd} C(t,u,s) + \delta^{ad}\delta^{bc} C(u,s,t), \end{aligned} \tag{3.1}$$

where s, t, u are the normalised Mandelstam variables, and crossing symmetry imposes that only two distinct functions B, C are used.[4] This is the form used in [34], where the functions B, C are given to NNLO for SU(n) equal-mass χPT. The NLO results were first obtained in [1, 35].

3.1 Other forms of the amplitude

With two flavours, the traces can be evaluated in terms of Kronecker δ's, giving[5]

$$M(s,t,u) = \delta^{ab}\delta^{cd} A(s,t,u) + \delta^{ac}\delta^{bd} A(t,u,s) + \delta^{ad}\delta^{bc} A(u,s,t), \tag{3.2}$$

[4] These functions have the symmetries $B(s,t,u) = B(u,t,s)$ and $C(s,t,u) = C(s,u,t)$, which is consistent with the symmetries of the respective flavour structures. Likewise, $A(s,t,u) = A(s,u,t)$ holds in eq. (3.2).

[5] This form can be traced back to the original current-algebra calculation [36] of the $\pi\pi$ amplitude.

JHEP03(2022)159

which is the form used in [13] (up to reordering the arguments as permitted by the symmetries of A). In terms of the functions above,

$$A(s,t,u) = C(s,t,u) + B(s,t,u) + B(t,u,s) - B(u,s,t), \tag{3.3}$$

the function A was first determined to NLO in [1].

With n flavours, the traces can be evaluated using the anticommutation relation $\left\{t^a, t^b\right\} = \frac{2}{n}\delta^{ab} + d^{abc}t^c$ to give[6]

$$\begin{aligned} M(s,t,u) = \quad & d^{abe}d^{cde}B'(s,t,u) + d^{ace}d^{bde}B'(t,u,s) + d^{ade}d^{bce}B'(u,s,t) \\ & + \delta^{ab}\delta^{cd}C'(s,t,u) + \delta^{ac}\delta^{bd}C'(t,u,s) + \delta^{ad}\delta^{bc}C'(u,s,t), \end{aligned} \tag{3.4}$$

where

$$\begin{aligned} B'(s,t,u) &= \tfrac{1}{2}\big[B(s,t,u) + B(t,u,s) - B(u,s,t)\big], \\ C'(s,t,u) &= C(s,t,u) + \tfrac{4}{n}B'(s,t,u). \end{aligned} \tag{3.5}$$

With three flavours, the Cayley-Hamilton theorem[7] allows for the removal of one term at the expense of symmetry, leaving

$$\begin{aligned} M(s,t,u) = \quad & \delta^{ab}\delta^{cd}A_1(s,t,u) + \delta^{ac}\delta^{bd}A_2(s,t,u) + \delta^{ad}\delta^{bc}A_3(s,t,u) \\ & + d^{abe}d^{cde}B_1(s,t,u) + d^{ace}d^{bde}B_2(s,t,u) \end{aligned} \tag{3.6}$$

where

$$\begin{aligned} B_1(s,t,u) &= B(t,u,s) - B(u,s,t), \qquad B_2(s,t,u) = B(t,u,s) - B(s,t,u), \\ A_1(s,t,u) &= C(s,t,u) + B(s,t,u) + \tfrac{1}{3}B_1(s,t,u), \\ A_2(s,t,u) &= C(t,u,s) + B(u,s,t) + \tfrac{1}{3}B_2(s,t,u), \\ A_3(s,t,u) &= C(u,s,t) + B(s,t,u) + B(u,s,t) - B(t,u,s). \end{aligned} \tag{3.7}$$

This is the form used in [14].

3.2 Structure of the amplitude

The functions $B(s,t,u)$ and $C(s,t,u)$ consist of one part that is polynomial in the Mandelstam variables and contains the LECs, plus the so-called *unitarity correction* that is non-polynomial in the Mandelstam variables.[8] The polynomial parts are quadratic at NLO and cubic at NNLO. At NLO, the unitarity correction does not contain any LECs; at NNLO, the unitarity correction depends on the NLO LECs.

[6] The relevant identity is

$$\left\langle t^a t^b t^c t^d \right\rangle + \left\langle t^a t^d t^c t^b \right\rangle = \tfrac{1}{2}\left(d^{abe}d^{cde} + d^{ade}d^{cbe} - d^{ace}d^{bde}\right) + \tfrac{2}{n}\left(\delta^{ab}\delta^{cd} + \delta^{ad}\delta^{cb} - \delta^{ac}\delta^{bd}\right).$$

It is most easily derived by first using $t^a t^b = \frac{1}{n}\delta^{ab} + \frac{1}{2}(d^{abc} + if^{abc})t^c$ repeatedly, and then removing all occurrences of f with the Jacobi-like identity

$$f^{abe}f^{cde} = d^{ace}d^{bde} - d^{bce}d^{ade} + \tfrac{4}{n}\left(\delta^{ac}\delta^{bd} - \delta^{ad}\delta^{bc}\right),$$

which is derived from the observation that $\left[\left[t^a, t^b\right], t^c\right] = \left\{\left\{t^b, t^c\right\}, t^a\right\} - \left\{\left\{t^c, t^a\right\}, t^b\right\}$.

[7] More specifically the $n = 3$ Cayley-Hamilton theorem, recast as the SU(3)-specific identity

$$3\left(d^{abe}d^{cde} + d^{bce}d^{ade} + d^{cae}d^{bde}\right) = 2\left(\delta^{ab}\delta^{cd} + \delta^{bc}\delta^{ad} + \delta^{ca}\delta^{bd}\right).$$

[8] This split is not uniquely defined, but we adhere to the conventions of [34].

JHEP03(2022)159

The unitarity correction at NLO depends on the function $\bar{J}$, which originates in the loop integral as shown in [1]. The NNLO unitarity correction introduces four analogous functions $k_i, i = 1, \ldots 4$ [31, 32, 34, 37]. More details about these functions can be found in appendix C.

The LEC content of the amplitude considered here is more limited than that of the full χPT Lagrangian. About half of the Lagrangian terms are dropped by not including the external (axial) vector fields, and a significant part of the NNLO Lagrangian cannot appear in a 4-particle process below NNNLO. Also, the number of LECs is reduced by the Cayley-Hamilton theorem in the 2- and 3-flavour case as described in section 2.1. Lastly, $\hat{L}_7, K_{12}, K_{24}, K_{30}, K_{34}, K_{36}, K_{41}$ and K_{42}, i.e. those whose Lagrangian terms contain $\langle \chi_- \rangle$, disappear in the equal-mass limit.[9] Even with these reductions, there are still 35 (27 at $n = 3$, 18 at $n = 2$) NNLO LECs that are involved in the amplitude at hand, in addition to 8 (7, 4) NLO LECs.

3.3 Irreducible amplitudes

The scattered particles are in the adjoint representation of SU(n). The Clebsch-Gordan decomposition of the initial and final states is therefore[10]

$$\text{Adj} \otimes \text{Adj} = R_I + R_S + R_A + R_S^A + R_A^S + R_S^S + R_A^A, \tag{3.8}$$

where R_I is the singlet representation, and the sub(super)scripts on the other representations indicate lower (upper) index pairs that are symmetric (S) or antisymmetric (A). Details on the representations and their dimensions can be found in [34, 38]. From this, it follows that the scattering amplitude can be decomposed in terms of seven corresponding irreducible amplitudes T_J. In terms of eq. (3.1), these are

$$\begin{aligned}
R_I: \quad & T_I = 2\frac{n^2-1}{n}\big[B(s,t,u) + B(t,u,s)\big] - \frac{2}{n}B(u,s,t) \\
& \quad + (n^2-1)C(s,t,u) + C(t,u,s) + C(u,s,t), \\
R_S: \quad & T_S = \frac{n^2-4}{n}\big[B(s,t,u) + B(t,u,s)\big] - \frac{4}{n}B(u,s,t) \\
& \quad + C(t,u,s) + C(u,s,t), \\
R_A: \quad & T_A = n\big[B(t,u,s) - B(s,t,u)\big] + C(t,u,s) - C(u,s,t), \\
R_S^A, R_A^S: \quad & T_{AS} = T_{SA} = C(t,u,s) - C(u,s,t), \\
R_S^S: \quad & T_{SS} = 2B(u,s,t) + C(t,u,s) + C(u,s,t), \\
R_A^A: \quad & T_{AA} = -2B(u,s,t) + C(t,u,s) + C(u,s,t).
\end{aligned} \tag{3.9}$$

[9] This can be understood by noting that χ_- has odd parity, so all terms in its expansion contain an odd number of pseudoscalar fields. If the even-parity Lagrangian term contains two traces of odd-parity objects, it can therefore only result in six-point vertices or larger, since the trace of a single field vanishes. Therefore, K_{12}, K_{24} etc. do not appear in the NNLO four-point amplitude, whereas $\hat{L}_7$ only appears in s,t,u-independent tadpole diagrams. As will be shown in section 5, we only consider s-derivatives of the amplitude, so also $\hat{L}_7$ disappears for our purposes.

[10] [35] contains an intuitive description of how the decomposition is performed.

JHEP03(2022)159

Only six amplitudes are needed, since T_{SA} and T_{AS} are identical due to crossing symmetry etc., as mentioned in [14].

In SU(3), the R_A^A representation vanishes, so only five amplitudes are needed. In [14], the representations are labelled by their dimensions, which are $1, 8, 8, 10$ and 27 in the order they appear in eqs. (3.8) and (3.9).

In SU(2), only R_I, R_A and R_S^S remain and have dimension $1, 3$ and 5, respectively. The corresponding amplitudes can be identified with the isospin components T^0, T^1 and T^2, respectively. In terms of eq. (3.2), they are

$$\begin{aligned} T^0 &= 3A(s,t,u) + A(t,u,s) + A(u,s,t), \\ T^1 &= A(t,u,s) - A(u,s,t), \\ T^2 &= A(t,u,s) + A(u,s,t). \end{aligned} \tag{3.10}$$

This well-known relation can be derived from eqs. (3.3) and (3.9).

3.4 Eigenstate amplitudes

A general amplitude can be expressed as $a_J T^J$, where the index J runs over the representations in the order they appear in eq. (3.8). For a physically applicable scattering process, however, the initial and final states should typically be taken as a product of mass eigenstates such as π, K and η. This corresponds to fixing a_J to a small selection of values so that $T(ab \to cd) = a_J(ab \to cd)T^J$. Here, as in [13, 14], we consider only elastic scattering of eigenstates, with $a_J(ab \to ab) \equiv a_J(ab)$.

With two flavours, where J runs over I, A, SS (alternatively, isospin $0, 1, 2$), the eigenstates are[11]

$$a_J(\pi^0\pi^0) = \left(\tfrac{1}{3}\ 0\ \tfrac{2}{3}\right), \qquad a_J(\pi^0\pi^\pm) = \left(0\ \tfrac{1}{2}\ \tfrac{1}{2}\right), \qquad a_J(\pi^\pm\pi^\pm) = \left(0\ 0\ 1\right), \tag{3.11}$$

and with three flavours, where J runs over I, S, A, AS, SS, they are[12]

$$\begin{aligned} a_J(\pi^0\pi^0) &= \left(\tfrac{1}{8}\ \tfrac{1}{5}\ 0\ 0\ \tfrac{27}{40}\right), & a_J(\pi^\pm\pi^\pm, K^\pm K^\pm, K^0K^0) &= \left(0\ 0\ 0\ 0\ 1\right), \\ a_J(\pi^0\pi^\pm) &= \left(0\ 0\ \tfrac{1}{3}\ \tfrac{1}{6}\ \tfrac{1}{2}\right), & a_J(K^\pm\pi^\pm, K^\pm\pi^\mp, K^0K^\pm) &= \left(0\ 0\ 0\ \tfrac{1}{2}\ \tfrac{1}{2}\right), \\ a_J(K^0\pi^\pm) &= \left(0\ \tfrac{3}{10}\ \tfrac{1}{6}\ \tfrac{1}{3}\ \tfrac{1}{5}\right), & a_J(K\pi^0) &= \left(0\ \tfrac{3}{20}\ \tfrac{1}{12}\ \tfrac{5}{12}\ \tfrac{7}{20}\right), \\ a_J(\pi\eta) &= \left(0\ \tfrac{1}{5}\ 0\ \tfrac{1}{2}\ \tfrac{3}{10}\right), & a_J(K\eta) &= \left(0\ \tfrac{1}{20}\ \tfrac{1}{4}\ \tfrac{1}{4}\ \tfrac{9}{20}\right); \end{aligned} \tag{3.12}$$

see e.g. [13, 14], respectively.[13] With four or more flavours, χPT loses its applicability as low-energy QCD since there are only three light quarks in the Standard Model. Therefore, there is little sense in considering eigenstates for n flavours, although we can note that $T(\pi^\pm\pi^\pm \to \pi^\pm\pi^\pm) = T_{SS}$ regardless of n.

[11] a_J is invariant under particle/antiparticle exchange, so $a_J(\pi^+\pi^+) = a_J(\pi^-\pi^-) \equiv a_J(\pi^\pm\pi^\pm)$. Note, however, that $a_J(\pi^\pm\pi^\mp) \neq a_J(\pi^\pm\pi^\pm)$ — they are instead related by crossing; see eq. (3.14).

[12] Here, π without superscript stands for any of $\pi^\pm$ or π^0 (and similarly for K) whenever a_J is agnostic about the particular choice. We use $a_J(ab, cd, \ldots)$ for $a_J(ab) = a_J(cd) = \ldots$.

[13] Eq. (3.12) differs from the values given in [14]: there was an error or misprint in $a_J(\pi^0\pi^\pm)$, and all eigenstates were not included, with $a_J(K^\pm\pi^\pm)$ given as $a_J(K\pi)$.

JHEP03(2022)159

One of our extensions over previous work is that we use all possible values for the a_J, rather than restricting them to eigenstates (see section 5 for what constitutes "possible"). This can be done without complications, since the mass eigenstates are completely degenerate in the equal-mass limit. However, it is still useful to view those states that remain mass eigenstates in the unequal-mass case as special. Below, by "eigenstate" we will specifically mean scattering between these states. Note that by treating general a_J, we effectively include inelastic scattering such as $a_J(\pi^0\pi^0 \to \pi^+\pi^-) = \left(\frac{1}{3}\ 0\ -\frac{1}{3}\right)$. However, it turns out that inelastic scattering is useless for our purposes by invariably failing to satisfy eq. (5.6b). This (in addition to [13, 14]) is why this section has focused mainly on elastic scattering.

3.5 Crossing symmetry

Since all amplitudes can be expressed as $a_J T^J$, crossing symmetry implies that channel crossing must take the form of a linear transformation of a_J. For $s \leftrightarrow u$ crossing, the transformation $T^I(u,t,s) = C_u^{IJ} T^J(s,t,u)$ is given by [21, 38]

$$C_u^{IJ} = \begin{pmatrix} \frac{1}{n^2-1} & 1 & -1 & \frac{4-n^2}{2} & \frac{n^2(n+3)}{4(n+1)} & \frac{n^2(n-3)}{4(n-1)} \\ \frac{1}{n^2-1} & \frac{n^2-12}{2(n^2-4)} & -\frac{1}{2} & 1 & \frac{n^2(n+3)}{4(n+1)(n+2)} & \frac{n^2(3-n)}{4(n-1)(n-2)} \\ \frac{1}{1-n^2} & -\frac{1}{2} & \frac{1}{2} & 0 & \frac{n(n+3)}{4(n+1)} & \frac{n(3-n)}{4(n-1)} \\ \frac{1}{1-n^2} & \frac{2}{n^2-1} & 0 & \frac{1}{2} & \frac{n(n+3)}{4(n+1)(n+2)} & \frac{n(n-3)}{4(n-1)(n-2)} \\ \frac{1}{n^2-1} & \frac{1}{2+n} & \frac{1}{n} & \frac{n-2}{2n} & \frac{n^2+n+2}{4(n+1)(n+2)} & \frac{n-3}{4(n-1)} \\ \frac{1}{n^2-1} & \frac{1}{2-n} & -\frac{1}{n} & \frac{n+2}{2n} & \frac{n+3}{4(n+1)} & \frac{n^2-n+2}{4(n-1)(n-2)} \end{pmatrix} \tag{3.13}$$

which also works at $n = 2, 3$ by removing appropriate rows and columns:

$$C_u^{IJ}\Big|_{\mathrm{SU}(2)} = \frac{1}{6}\begin{pmatrix} 2 & -6 & 10 \\ -2 & 3 & 5 \\ 2 & 3 & 1 \end{pmatrix} \qquad C_u^{IJ}\Big|_{\mathrm{SU}(3)} = \begin{pmatrix} \frac{1}{8} & 1 & -1 & -\frac{5}{2} & \frac{27}{8} \\ \frac{1}{8} & -\frac{3}{10} & -\frac{1}{2} & 1 & \frac{9}{8} \\ -\frac{1}{8} & -\frac{1}{2} & \frac{1}{2} & 0 & \frac{27}{40} \\ -\frac{1}{8} & \frac{2}{5} & 0 & \frac{1}{2} & \frac{9}{40} \\ \frac{1}{8} & \frac{1}{5} & \frac{1}{3} & \frac{1}{6} & \frac{7}{40} \end{pmatrix}. \tag{3.14}$$

These versions can be found in [13, 14] respectively.

4 Linear constraints

In this section, we will introduce a mathematical language of *linear constraints*. This formalism is introduced before positivity bounds (see section 5) so that they can be established in full generality. In order to make the handling of the bounds as general and powerful as possible, we dedicate this section to developing some useful mathematical definitions and results.[14]

[14] In this section employ mathematical notation that, depending on the background of the reader, may not be entirely familiar. We also define new notation for our own purposes. A glossary covering all potentially unfamiliar notation is provided in appendix B.7.

JHEP03(2022)159

4.1 Definition and combination of constraints

For a set of parameters b_i (e.g. the LECs), a linear constraint takes the general form

$$\alpha_1 b_1 + \alpha_2 b_2 + \ldots + \alpha_n b_n - c \geq 0, \tag{4.1}$$

where c, α_i are known coefficients. By treating α_i, b_i as components of vectors, this is equivalent to

$$\boldsymbol{\alpha} \cdot \boldsymbol{b} \geq c. \tag{4.2}$$

We say that $\boldsymbol{b}$ lives in the *parameter space*, whereas $\boldsymbol{\alpha}$ lives in the *constraint space*.[15] Since $\boldsymbol{\alpha}$ and c can be rescaled by any positive scalar without changing the inequality, any linear constraint can be described by the pair $\langle \boldsymbol{\alpha}, c \rangle$ with $c \in \{1, 0, -1\}$.

We say that a point $\boldsymbol{b}$ *satisfies* a constraint $\langle \boldsymbol{\alpha}, c \rangle$ if $\boldsymbol{\alpha} \cdot \boldsymbol{b} \geq c$. We denote by $\mathcal{B}(\langle \boldsymbol{\alpha}, c \rangle)$ the subset of parameter space that satisfies $\langle \boldsymbol{\alpha}, c \rangle$. For any $\boldsymbol{\alpha}$, it is clear that the origin $\boldsymbol{b} = \boldsymbol{0}$ is contained in $\mathcal{B}(\langle \boldsymbol{\alpha}, -1 \rangle)$ but not in $\mathcal{B}(\langle \boldsymbol{\alpha}, 1 \rangle)$, and lies on the boundary of $\mathcal{B}(\langle \boldsymbol{\alpha}, 0 \rangle)$ (except when $\boldsymbol{\alpha} = \boldsymbol{0}$).

The LECs will typically be subject to many linear constraints simultaneously. We will normally use the letter Ω to denote a constraint, either a single one like $\langle \boldsymbol{\alpha}, c \rangle$ or a combination of several such constraints. Given two constraints Ω, Ω', we write the constraint that imposes both of them simultaneously as $\Omega + \Omega'$. A point $\boldsymbol{b}$ satisfies $\Omega + \Omega'$ if and only if it satisfies both Ω and Ω'; thus, the $\mathcal{B}$ notation naturally generalises through $\mathcal{B}(\Omega + \Omega') \equiv \mathcal{B}(\Omega) \cap \mathcal{B}(\Omega')$. For combinations of many constraints, we will generalise $+$ into e.g. $\Omega = \sum_i \langle \boldsymbol{\alpha}_i, c_i \rangle$.

4.2 Stronger and weaker constraints

A hierarchy can be established among the constraints based on how strong (restrictive) they are. For instance, $b_1 \geq 1$ is stronger than $b_1 \geq 0$. We will write the stronger-than relation as $\Omega \geq \Omega'$, which holds if all points that satisfy Ω also satisfy Ω'. Thus, $\Omega \geq \Omega'$ is equivalent to $\mathcal{B}(\Omega) \subseteq \mathcal{B}(\Omega')$. Naturally, we say $\Omega = \Omega'$ if $\mathcal{B}(\Omega) = \mathcal{B}(\Omega')$, and say $\Omega > \Omega'$ if $\Omega \geq \Omega'$ but $\Omega \neq \Omega'$. Just like subset relations, our stronger-than relation is not a total ordering, as there exist many pairs of constraints Ω, Ω' where neither is stronger than the other. From our definitions, it trivially follows that

$$\begin{aligned} (\Omega + \Omega') &\geq \Omega, & (\Omega + \Omega') &\geq \Omega', \\ (\Omega + \Omega') &= \Omega \quad \Leftrightarrow & \Omega &\geq \Omega', \end{aligned} \tag{4.3}$$

so that if $\Omega \not\geq \Omega'$ and $\Omega' \not\geq \Omega$, their combination $\Omega + \Omega'$ is indeed a new, strictly stronger constraint. Furthermore, we see that, for all $\lambda > 0$, $\kappa > 1$, and $\boldsymbol{\alpha} \neq \boldsymbol{0}$,

$$\langle \lambda\boldsymbol{\alpha}, 1 \rangle > \langle \lambda\boldsymbol{\alpha}, 0 \rangle > \langle \lambda\boldsymbol{\alpha}, -1 \rangle\,, \tag{4.4}$$

$$\langle \kappa\boldsymbol{\alpha}, 1 \rangle < \langle \boldsymbol{\alpha}, 1 \rangle\,, \qquad \langle \lambda\boldsymbol{\alpha}, 0 \rangle = \langle \boldsymbol{\alpha}, 0 \rangle\,, \qquad \langle \kappa\boldsymbol{\alpha}, -1 \rangle > \langle \boldsymbol{\alpha}, -1 \rangle\,. \tag{4.5}$$

[15] We consistently use Roman letters for vectors in parameter space and Greek letters for vectors in constraint space. In general, parameter space may be any finite-dimensional real vector space, with constraint space considered as its dual.

JHEP03(2022)159

From the $c = 0$ version of eq. (4.5), we see that $\langle \boldsymbol{\alpha}, 0\rangle$ is not a unique representation of the constraint, since we can freely rescale $\boldsymbol{\alpha}$ without changing it. We may remove this ambiguity by constraining $\boldsymbol{\alpha}$ to be a unit vector.

There exists a constraint Ω_∞, equivalent to $\langle \mathbf{0}, 1\rangle$ or e.g. $\langle \boldsymbol{\alpha}, 1\rangle + \langle -\boldsymbol{\alpha}, 1\rangle$, that is not satisfied by any point. It follows that $\Omega_\infty \geq \Omega$ and $\Omega_\infty + \Omega = \Omega_\infty$ for any Ω. A constraint that is satisfied by all points, i.e. $\langle \mathbf{0}, -1\rangle$ or $\langle \mathbf{0}, 0\rangle$, will be called a *trivial constraint*.

4.3 Determining the relationship between constraints

We will now present a general result, which determines if a given linear constraint $\langle \boldsymbol{\beta}, c\rangle$ is weaker than an arbitrarily complicated constraint Ω. This will serve as the basis for all our uses of linear constraints.[16] For complete proofs, more details, and practical applications, see appendix B.

Consider a set of linear constraints $\langle \boldsymbol{\alpha}_i, c\rangle$ for i in some finite set $I \subset \mathbb{N}$.[17] Note that c is the same for all constraints. Then let $\omega_c \equiv \sum_{i\in I} \langle \boldsymbol{\alpha}_i, c\rangle$;[18] an example of such a constraint is given in figure 1. Then define $\mathcal{A}(\omega_c)$ as the set of all points that can be expressed as

$$\sum_{i\in I} \lambda_i \boldsymbol{\alpha}_i, \qquad \lambda_i \geq 0, \quad \sum_{i\in I} \lambda_i \in \Lambda_c, \tag{4.6}$$

where

$$\Lambda_1 = [1, \infty), \qquad \Lambda_0 = [0, \infty), \qquad \Lambda_{-1} = [0, 1]. \tag{4.7}$$

The shape of $\mathcal{A}(\omega_c)$ is illustrated in figure 2. With these definitions, the following holds:

Proposition 4.1 (determining if constraint is weaker, special case)**.** *Let $\langle \boldsymbol{\beta}, c\rangle$ be a single linear constraint, and let $\omega_c \neq \Omega_\infty$ be defined as above. Then $\langle \boldsymbol{\beta}, c\rangle \leq \omega_c$ if and only if $\boldsymbol{\beta} \in \mathcal{A}(\omega_c)$.*

This is proven in appendix B.1. If ω_c is a single linear constraint, this result reduces down to eq. (4.5). The condition $\omega_c \neq \Omega_\infty$ is necessary, since there exist corner cases where $\omega_c = \Omega_\infty$ but $\mathcal{A}(\omega_c)$ fails to cover the entire constraint space.[19] However, if $\mathcal{A}(\omega_c)$ does cover the entire space, then it is certain that $\omega_c = \Omega_\infty$.

In figure 2, we may note that $\mathcal{A}(\omega_c)$ is closely related to the *convex hull* of the $\boldsymbol{\alpha}_i$. In essence, $\mathcal{A}(\omega_c)$ is obtained by forming the hull, and then also including all points that give weaker constraints under eq. (4.5). We may also note that the convex hull can be defined as

$$\text{Hull}\left(\{\boldsymbol{\alpha}_i\}_{i\in I}\right) = \left\{ \sum_{i\in I} \lambda_i \boldsymbol{\alpha}_i \;\middle|\; \lambda_i \geq 0,\ \sum_{i\in I} \lambda_i = 1 \right\}, \tag{4.8}$$

which is very similar to eq. (4.6).

[16] Proposition 4.1, along with a version of the notation we use here, was defined in [21], although the proof was completely different. An incorrect version of proposition 4.2 was also presented without proof. To the best of our knowledge, these results are novel, although the relevant literature is vast and lies outside our area of expertise. The closest we have found is [39], although their algorithm requires knowing a point that satisfies Ω, relies on more complicated mathematical machinery, and does not include all the extensions presented further below in section 4.4 and appendix B.

[17] It is crucial that only finite combinations of constraints are considered, and it will normally be tacitly assumed that all sets like I are finite. A limited extension to infinite sets is covered in appendix B.6.

[18] We use lowercase ω here to emphasise that it is not a general constraint. A similar treatment of general Ω is given below.

[19] A trivial example of this is $\omega_1 = \langle \mathbf{0}, 1\rangle$, where $\mathcal{A}(\omega_1) = \{\mathbf{0}\}$.

JHEP03(2022)159

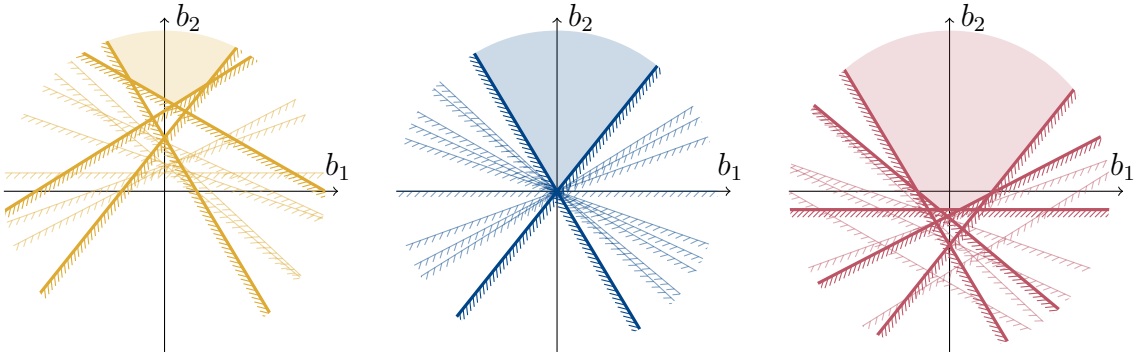

Figure 1. A cropped depiction of twelve random two-dimensional constraints $\langle \boldsymbol{\alpha}_i, c \rangle$ for $c = +1$ (left, **yellow**), $c = 0$ (middle, **blue**) and $c = -1$ (right, **red**) illustrated as the parameter-space lines $\boldsymbol{\alpha}_i \cdot \boldsymbol{b} = c$. The side of the line that is *excluded* by the constraint is hatched. The region $\mathcal{B}(\omega_c)$ for $\omega_c = \sum_i \langle \boldsymbol{\alpha}_i, c \rangle$, i.e. the set of points that satisfy all the constraints, is shaded. The lines corresponding to *relevant* constraints (i.e. those that actually delimit $\mathcal{B}(\omega_c)$; this is more closely defined in section 4.4) are drawn more strongly than the rest.

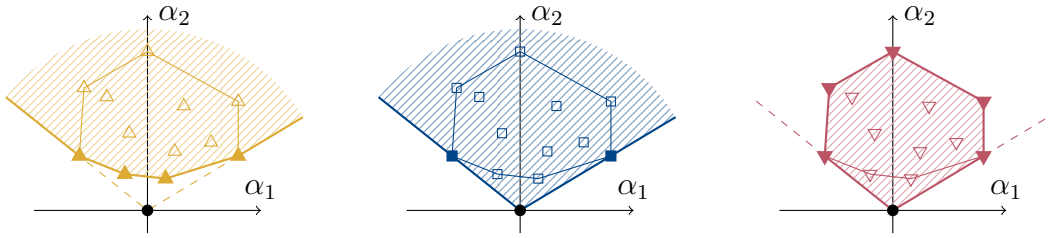

Figure 2. Examples of the regions $\mathcal{A}(\omega_1)$ (left, **yellow**), $\mathcal{A}(\omega_0)$ (middle, **blue**) and $\mathcal{A}(\omega_{-1})$ (right, **red**) in constraint space, using the same $\boldsymbol{\alpha}_i$ as in figure 1. The $\boldsymbol{\alpha}_i$ are represented as points (△, ▫, ▽, respectively) in the space, and the relevant ones are filled (again, relevancy is defined in section 4.4). The convex hulls (as defined in eq. (4.8)) of the $\boldsymbol{\alpha}_i$ are outlined. For comparison to figure 1, it is helpful to remember that $\boldsymbol{\alpha}_i$ are normal vectors to the lines shown there, and that larger $|\boldsymbol{\alpha}_i|$ correspond to lines passing *closer* to the origin. Note how, given identical $\boldsymbol{\alpha}_i$, $\mathcal{A}(\omega_0)$ is the union of $\mathcal{A}(\omega_1)$ and $\mathcal{A}(\omega_{-1})$ (this is easy to see from eq. (4.7)) whereas the set of relevant constraints is the intersection of the respective sets.

Now, let us handle the general case. The most general combination of a finite number of linear constraints can be expressed as

$$\Omega = \sum_{i \in I_1} \langle \boldsymbol{\alpha}_i, 1 \rangle + \sum_{i \in I_0} \langle \boldsymbol{\alpha}_i, 0 \rangle + \sum_{i \in I_{-1}} \langle \boldsymbol{\alpha}_i, -1 \rangle \,, \tag{4.9}$$

where I_c are some disjoint, finite, and possibly empty sets. We may compactly write this as $\Omega = \sum_{i \in I} \langle \boldsymbol{\alpha}_i, c_i \rangle$ where $I \equiv I_{+1} \cup I_0 \cup I_{-1}$ and $c_i = c$ if $i \in I_c$.

Similarly to $\mathcal{A}(\omega_c)$, let $\mathcal{A}_c(\Omega)$ be the set of all points that can be expressed as (recall that $c \in \{1, 0, -1\}$)

$$\sum_{i \in I_{-1}} \lambda_i \boldsymbol{\alpha}_i + \sum_{i \in I_0} \lambda_i \boldsymbol{\alpha}_i + \sum_{i \in I_1} \lambda_i \boldsymbol{\alpha}_i, \qquad \lambda_i \geq 0, \tag{4.10}$$

JHEP03(2022)159

with λ_i constrained by the condition

$$\sum_{i\in I_1} \lambda_i - \sum_{i\in I_{-1}} \lambda_i \geq c. \tag{4.11}$$

An illustration of $\mathcal{A}_c(\Omega)$ can be found in appendix B.4.5. With these definitions, the following holds:

Proposition 4.2 (determining if constraint is weaker, general case)**.** *Let $\langle \boldsymbol{\beta}, c\rangle$ be a linear constraint, and let $\Omega \neq \Omega_\infty$ be defined as above. Then $\langle \boldsymbol{\beta}, c\rangle \leq \Omega$ if and only if $\boldsymbol{\beta} \in \mathcal{A}_c(\Omega)$.*

This is proven in appendix B.1.4. If only one of the I_c is nonempty, this reduces down to proposition 4.1.

While it is not as useful for the purposes of proposition 4.2, one may note that eqs. (4.10) and (4.11) can be more succinctly stated as

$$\mathcal{A}_c(\Omega) = \left\{ \sum_{i\in I} \lambda_i \boldsymbol{\alpha}_i \;\middle|\; \lambda_i \geq 0, \sum_{i\in I} \lambda_i c_i \geq c \right\}. \tag{4.12}$$

This definition of $\mathcal{A}_c(\Omega)$ works also if c, c_i are not constrained to $\{-1, 0, 1\}$.

4.4 Representations and degeneracy

Checking if $\boldsymbol{b}$ satisfies Ω becomes computationally expensive if Ω is the combination of many different linear constraints. However, Ω is usually not uniquely determined by how it is expressed as a sum of linear constraints, and it is possible to vastly reduce that redundancy. To that end, we define a *representation* of a constraint Ω as any finite set $\mathcal{S}$ of linear constraints with the property[20]

$$\Omega = \sum_{\langle \boldsymbol{\alpha}, c\rangle \in \mathcal{S}} \langle \boldsymbol{\alpha}, c\rangle \,. \tag{4.13}$$

If it is implicit which representation is used for Ω, we may call the $\langle \boldsymbol{\alpha}, c\rangle \in \mathcal{S}$ the *elements of* Ω.

It is clear that there exist *minimal representations*, i.e. representations with the smallest number of elements. As we will see below, there is often a unique minimal representation, which we will label $\mathcal{R}(\Omega)$. However, there is an important exception to this: when $\mathcal{B}(\Omega)$ is contained in a hyperplane. This happens when there are some $\boldsymbol{\delta}, d$ such that $\boldsymbol{\delta} \cdot \boldsymbol{b} = d$ for all $\boldsymbol{b} \in \mathcal{B}(\Omega)$, or equivalently $\langle \boldsymbol{\delta}, d\rangle + \langle -\boldsymbol{\delta}, -d\rangle \leq \Omega$. We will call Ω *degenerate* if so is the case, and *non-degenerate* otherwise.[21] With this in mind, we can state the following result:

[20] Clearly, all Ω also admit representation as a sum of an *infinite* number of constraints. However, we will not consider such representations, and proposition 4.3 below generally only holds if Ω can be expressed as a finite sum. See appendix B.6 for a discussion about infinite sums of constraints.

[21] As defined here, Ω_∞ would be considered a special case of a degenerate constraint. In the closer study of degenerate constraints given in appendix B.2.1, it turns out to be more useful to consider Ω_∞ seperately, viewing it as neither degenerate nor non-degenerate.

JHEP03(2022)159

Proposition 4.3 (finding relevant constraints, non-degenerate case)**.** *If Ω is a non-degenerate constraint, there exists a minimal representation $\mathcal{R}(\Omega)$ that is unique up to the normalisation of its elements. Furthermore, for any representation $\mathcal{S}$ of Ω, the relation $\mathcal{R}(\Omega) \subseteq \mathcal{S}$ is true up to normalisation.*

The elements of $\mathcal{R}(\Omega)$ are exactly those $\langle \boldsymbol{\alpha}, c\rangle \leq \Omega$ for which there is some $\boldsymbol{b} \in \mathcal{B}(\Omega)$ such that $\boldsymbol{\alpha}\cdot\boldsymbol{b} = c$ and $\boldsymbol{\beta}\cdot\boldsymbol{b} > d$ for all $\langle \boldsymbol{\beta}, d\rangle \leq \Omega$ with $\langle \boldsymbol{\beta}, d\rangle \neq \langle \boldsymbol{\alpha}, c\rangle$.

This is proven in appendix B.3. Due to this uniqueness, and the fact that $\mathcal{R}(\Omega)$ is a subset of any representation, we will call the elements of $\mathcal{R}(\Omega)$ the *relevant elements of* Ω, and call all other elements of any representation *irrelevant*, since they can be discarded without altering Ω. A more practical way of finding $\mathcal{R}(\Omega)$, based on proposition 4.2, is given in appendix B.4.4.

When Ω is degenerate, there is typically no unique minimal representation, although there is still a straightforward way to find *some* minimal representation, which we will also label $\mathcal{R}(\Omega)$. This generalisation of proposition 4.3 is discussed in appendix B.2.1, along with a more general method of replacing any degenerate constraint with a non-degenerate analogue in a lower-dimensional space. Note, however, that degenerate constraints are only a corner case with little practical relevance: a small perturbation, e.g. by numerical error, to the elements of a degenerate constraint will either render it non-degenerate, or render it equal to Ω_∞.

5 Positivity bounds

Equipped with the notion of linear constraints, we are ready to move on to the main topic of this paper: positivity bounds. (For a more detailed version of this derivation, see [13]; various generalisations can be found in e.g. [19, 20].) We start by writing down the fixed-t dispersion relation for the amplitude $a_J T^J$:

$$a_J T^J(s,t) = \frac{1}{2\pi i}\oint_\gamma \mathrm{d}z \frac{a_J T^J(z,t)}{z-s}. \tag{5.1}$$

The amplitude has two branch cuts along the real axis: a right-hand cut starting at $z=4$ corresponding to the s-channel, and a left-hand cut starting at $z=-t$ corresponding to the u-channel. The discontinuity across these cuts is $T(z+i\varepsilon) - T(z-i\varepsilon) = 2i\,\mathrm{Im}\,T(z+i\varepsilon)$. For real s in the span $-t < s < 4$, we deform the contour γ as shown in figure 3. We can then reexpress the integral in terms of the discontinuities, which may require derivatives (subtractions) to make the contour at infinity vanish. Using the crossing relation derived in section 3.5 to rewrite the u-channel cut in terms of s, the result is

$$a_J \frac{\mathrm{d}^k}{\mathrm{d}s^k} T^J(s,t) = \frac{k!}{\pi}\int_4^\infty \mathrm{d}z \left[\frac{a_J}{(z-s)^{k+1}} + \frac{(-1)^k a_I C_u^{IJ}}{(z-u)^{k+1}}\right] \mathrm{Im}\,T^J(z+i\varepsilon,t). \tag{5.2}$$

The Froissart bound [40] shows that the integral converges whenever $k \geq 2$, since $T^J(z+i\varepsilon,t) = \mathcal{O}(s\ln^2 s)$.[22] We will discuss specific values for k in section 5.2; here, we keep it general.

[22] Note that $T^J(z+i\varepsilon,t)$ on the right-hand side of eq. (5.2) is the exact, non-perturbative amplitude — see e.g. [13]. Indeed, the perturbative χPT amplitude at any fixed order grows polynomially with s, so it violates the Froissart bound. We can insert the fixed-order perturbative amplitude at the left-hand side thanks to the smallness of s (and t), which guarantees good agreement with the exact one.

JHEP03(2022)159

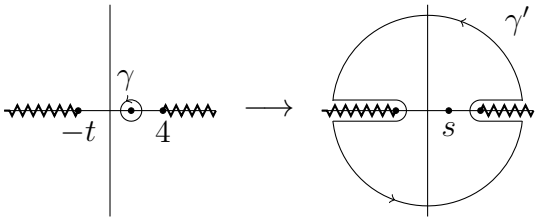

Figure 3. The contour integral in the z-plane around $z = s$ used in the dispersion relation.

Above threshold, the partial-wave expansion of the amplitude takes the form[23]

$$T^J(s,t) = \sum_{\ell=0}^{\infty} (2\ell+1) f_\ell^J(s) P_\ell\left(1 + \frac{2t}{s-4}\right), \tag{5.3}$$

where f_ℓ^J are partial wave amplitudes, P_ℓ are Legendre polynomials, and the expression in parentheses is the cosine of the scattering angle. The optical theorem then gives

$$\operatorname{Im} f_\ell^J(s) = s\beta(s)\sigma_\ell^J(s), \qquad \beta(s) \equiv \sqrt{1 - \frac{4}{s}}, \tag{5.4}$$

which is positive above threshold since the partial-wave cross-sections σ_ℓ^J are always positive. Therefore,

$$\operatorname{Im} T^J(s,t) = \sum_{\ell=0}^{\infty} (2\ell+1) s\beta(s)\sigma_\ell^I P_\ell\left(1 + \frac{2t}{s-4}\right) \tag{5.5}$$

is positive above threshold as long as P_ℓ is. Since $P_\ell(z) \geq 0$ when $z \geq 1$, eq. (5.2) therefore imposes the constraint that, for any $t \in [0,4], s \in [-t,4]$ and any representation index J,

$$a_J \frac{\mathrm{d}^k}{\mathrm{d}s^k} T^J(s,t) \geq 0 \tag{5.6a}$$

$$\text{if} \quad a_I \left\{ \delta^{IJ} \left[\frac{z-u}{z-s}\right]^{k+1} + (-1)^k C_u^{IJ} \right\} \geq 0 \quad \text{for all } z \geq 4. \tag{5.6b}$$

The region in the s,t plane where this holds is shown in figure 4. Note that $u \in [-4,4]$, so the expression in square brackets above is always positive.

Up to and including NNLO, the second derivative of T^J is linear in all LECs, so we obtain from eq. (5.6a) an expression of the form

$$\sum_i \alpha_i \hat{L}_i^r + \sum_j \beta_j K_j^r + \gamma \geq 0, \tag{5.7}$$

where the coefficients $\alpha_i, \beta_i, \gamma$ are functions of s,t and a_J, but not of the LECs. This constitutes a linear constraint, and each valid choice of s,t and a_J potentially yields a different constraint. The result of combining these constraints will be that only a limited region in parameter space ($\mathcal{B}(\Omega)$ in the notation of section 4) satisfies the positivity bounds. With some luck, the boundary of this region is close enough to the experimentally measured value to improve on its uncertainty (carefully considering also the uncertainty of the bounds).

[23] There is a limited domain of validity for this expansion, but it does not affect the range of s,t used by us. Again, see [13] for details.

JHEP03(2022)159

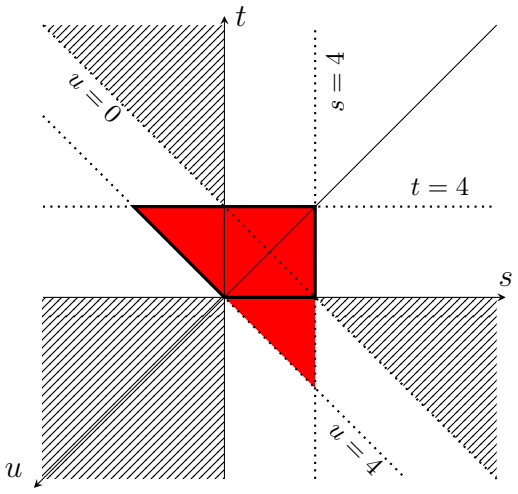

Figure 4. The plane of normalised Mandelstam variables. The **red** triangle is the region where the amplitude is real and free from singularities or branch cuts. The positivity conditions eq. (5.6) are valid inside the outlined part. The hatched regions with s, t or u positive are the physical regions for the respective channels.

5.1 Conditions on a_J

If we demand that eq. (5.6b) holds in the entire allowed s, t region, we see that the factor in square brackets can be made arbitrarily large or small by varying s, u, z. Therefore, we obtain the independent conditions $a_J \geq 0$ and $(-1)^k a_I C_u^{IJ} \geq 0$. However, we may apply the dispersion relation independently to each fixed s, t. Then, eq. (5.6b) can be made less restrictive, and a wider range of constraints on the LECs can be generated. This also includes permitting odd k for some s, t.

While we may fix s and t (which in turn fixes u), we must still allow z to cover its entire range. Therefore, finding all valid a_J for given s, t presents some practical issues. We solve this by using the technology of section 4, since eq. (5.6b) is a set of linear constraints on the vector a_I; we may write it compactly as $\boldsymbol{a} \cdot \boldsymbol{\beta}^J(z) \geq 0$. Noting that $\left(\frac{z-u}{z-s}\right)^{k+1}$ is monotonic as a function of $z \in [4, \infty)$, we see that it is always possible to write $\boldsymbol{\beta}^J(z) = \mu \boldsymbol{\beta}^J(4) + (1-\mu)\boldsymbol{\beta}^J(\infty)$ for $\mu \in [0, 1]$. By propositions 4.1 and 4.3 (see also proposition B.3), it follows that only $\boldsymbol{\beta}^J(4)$ and $\boldsymbol{\beta}^J(\infty)$ are relevant constraints on $\boldsymbol{a}$. Thus, it is sufficient to evaluate eq. (5.6b) at $z = 4$ and $z = \infty$, rather than letting z cover its entire range.

Another practical problem is that the set of allowed a_J is typically unbounded. However, eq. (5.6a) is independent of the magnitude of a_J. The obvious solution is to fix the normalisation of the vector $\boldsymbol{a}$, but this is problematic since a linear constraint on $\boldsymbol{a}$ is not necessarily a linear constraint on $\frac{1}{|\boldsymbol{a}|}\boldsymbol{a}$. Instead, we may simply rescale $\boldsymbol{a}$ so that $\sum_J a_J = 1$. This does not cover all possible $\boldsymbol{a}$ (for that, we must also look at $\sum_J a_J = 0$ and $\sum_J a_J = -1$), but it turns out that eq. (5.6b) is only satisfied by $\boldsymbol{a}$ for which this works. Using this, constraints on a_I are shown in figure 5.

JHEP03(2022)159

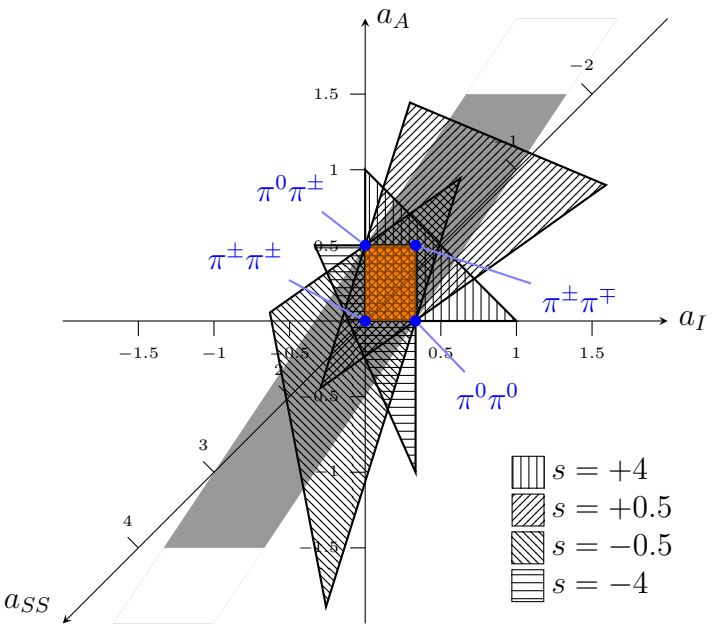

Figure 5. Illustration of which a_I, a_A, a_{SS}, normalised so that $a_I + a_A + a_{SS} = 1$, are permitted by eq. (5.6b) for $n = 2$, $k = 2$ at $t = 4$ and various fixed s. The shaded band is the region permitted by the $z = \infty$ bounds (see the discussion in section 5.1). It is independent of s and extends to infinity. The hatched triangles are the regions permitted by the $z = 4$ bounds for various s as indicated. Thus, the a_J permitted at fixed s, t is the overlap between the triangle and the band. The **orange** rectangle is the region permitted at all s, t. The **blue** points represent the eigenstates, including $a_J(\pi^\pm\pi^\mp) = C_u^{IJ} a_I(\pi^\pm\pi^\pm)$ in addition to those given in eq. (3.11). Not shown is the inelastic scattering $a_J(\pi^0\pi^0 \to \pi^+\pi^-) = \left(\frac{1}{3}\ 0\ -\frac{1}{3}\right)$, which never satisfies eq. (5.6b). For $n \geq 2$, the permitted region has an analogous shape, albeit in 4 ($n = 3$) and 5 ($n \geq 4$) dimensions, respectively. Like for $n = 2$, the eigenstate scattering amplitudes are mainly located in the corners of the always-permitted region.

5.2 The number of derivatives

As mentioned before, eq. (5.6) requires $k \geq 2$ to be valid, and $k = 2$ is sufficient; indeed, [41] claims that this value produces the best bounds. However, nothing prevents us from taking more derivatives, and with our generalised methods, we do find new relevant bounds from larger k; see e.g. figures 6 and 12 below. Also [19] makes use of higher derivatives.

At NLO, the LECs only enter through the second-order polynomial part of the amplitude, so the third and higher derivatives are parameter-independent and do not generate any bounds. This is not the case at NNLO, where the polynomial part is third-order, and where the non-polynomial unitarity correction contains NLO LECs. Therefore, $k = 3$ should yield another set of bounds on the NNLO LECs, and $k \geq 4$ should add bounds on the NLO LECs not obtainable from the NLO-only amplitude.

It also turns out that odd k cannot be used at any order in the 2-flavour case. To see this, look explicitly at eq. (5.6b) at $z = \infty$:

$$\delta^{IJ} - C_u^{IJ} = \tfrac{1}{6} c_1^I c_2^J \qquad \text{where} \quad c_1 = \begin{pmatrix} 2 & 1 & -1 \end{pmatrix}, c_2 = \begin{pmatrix} 2 & 3 & -5 \end{pmatrix} \tag{5.8}$$

JHEP03(2022)159