

which shows that $\det(1 - B^{(\mathbf{P})}K) = \det(1 - KB^{(\mathbf{P})})$ must be real.

Before continuing, let us show that \tilde{K} and $B^{(\mathbf{P})}$ do not commute. We can write

$$\begin{aligned}\langle J'm_J L'S'a' | K | Jm_J L S a \rangle &= \delta_{J'J} \delta_{m_J' m_J} K_{L'S'a'; L S a}^{(J)}, \\ \langle J'm_J L'S'a' | B^{(\mathbf{P})} | Jm_J L S a \rangle &= \delta_{S'S} \delta_{a'a} B_{J'm_J L'; Jm_J L}^{(\mathbf{P} S a)},\end{aligned}\quad (337)$$

where the reduced matrix of K is real and symmetric, and the reduced matrix of $B^{(\mathbf{P})}$ is Hermitian. The reduced matrix of $B^{(\mathbf{P})}$ depends on a since it depends on u_a as shown in Eq. (334). Given these equations, we see that

$$\begin{aligned}\langle J'm_J L'S'a' | KB^{(\mathbf{P})} | Jm_J L S a \rangle &= K_{L'S'a'; L'' S a}^{(J')} B_{J'm_J L'; Jm_J L}^{(\mathbf{P} S a)}, \\ \langle J'm_J L'S'a' | B^{(\mathbf{P})}K | Jm_J L S a \rangle &= K_{L'' S'a'; L S a}^{(J)} B_{J'm_J L'; Jm_J L''}^{(\mathbf{P} S'a')}. \end{aligned}\quad (338)$$

From these results, one sees that these matrices do *not* commute:

$$[K, B^{(\mathbf{P})}] \neq 0. \quad (339)$$

If these matrices did commute, then we could find common eigenvectors of the two matrices, and the quantization determinant would become a product of terms, each term involving the eigenvalues of the two matrices for a particular common eigenvector. However, this is not to be. Also, from the above equations, it is also not possible to even find a form of \tilde{K} that would commute with the box matrix.

11 Summary so far

The derivation of the relationship between the finite-volume two-particle energies and the infinite-volume scattering amplitudes presented in the preceding sections was rather lengthy and complicated, so it is useful to summarize the situation so far.

We work in an L_V^3 spatial volume with periodic boundary conditions. For a given total momentum $\mathbf{P} = (2\pi/L_V)\mathbf{d}$, where \mathbf{d} is a vector of integers, we determine the total energy E in the lab frame for a particular two-particle interacting state in our lattice QCD simulations. We boost to the center-of-momentum frame and calculate

$$E_{\text{cm}} = \sqrt{E^2 - \mathbf{P}^2}, \quad \gamma = \frac{E}{E_{\text{cm}}}. \quad (340)$$

Let N_d denote the number of two-particle channels that are open, and denote the masses and spins of the two scattering particles in channel a by m_{ja} and s_{ja} , respectively, for $j = 1, 2$. If $|E_{\text{cm}}^2| \geq |m_{1a}^2 - m_{2a}^2|$, then we can calculate the following quantities in each channel:

$$\mathbf{q}_{\text{cm},a}^2 = \frac{1}{4}E_{\text{cm}}^2 - \frac{1}{2}(m_{1a}^2 + m_{2a}^2) + \frac{(m_{1a}^2 - m_{2a}^2)^2}{4E_{\text{cm}}^2}, \quad (341)$$

$$u_a^2 = \frac{L_V^2 \mathbf{q}_{\text{cm},a}^2}{(2\pi)^2}, \quad \mathbf{s}_a = \left(1 + \frac{(m_{1a}^2 - m_{2a}^2)}{E_{\text{cm}}^2}\right) \mathbf{d}. \quad (342)$$

The total energy E is then related to the dimensionless unitary scattering S -matrix through the quantization condition:

$$\det[1 + F^{(\mathbf{P})}(S - 1)] = 0. \quad (343)$$

We use an orthonormal basis of states, each labelled $|Jm_JLSa\rangle$, where J is the total angular momentum of the two particles in the cm frame, m_J is the projection of the total angular momentum onto the z -axis, L is the orbital angular momentum of the two particles in the cm frame, S in the basis vector is the total spin of the two particles (not the action nor the scattering matrix). The index a refers to particle species, the spins s_1, s_2 , intrinsic parities η_1^P, η_2^P , isospins I_1, I_2 , isospin projections I_{z1}, I_{z2} , and possibly G -parities η_1^G, η_2^G of particle 1 and 2. Other quantum numbers, such as strangeness and charm, could also be included. The $F^{(\mathbf{P})}$ matrix in this basis is given by

$$\begin{aligned} \langle J'm_{J'}, L'S'a' | F^{(\mathbf{P})} | Jm_J, LSa \rangle &= \delta_{a'a} \delta_{S'S} \frac{1}{2} \left\{ \delta_{J'J} \delta_{m_{J'}m_J} \delta_{L'L} \right. \\ &\quad \left. + \langle J'm_{J'} | L'm_{L'} S m_S \rangle \langle Lm_L S m_S | Jm_J \rangle W_{L'm_{L'}; Lm_L}^{(\mathbf{P}a)} \right\}, \end{aligned} \quad (344)$$

where $W^{(\mathbf{P}a)}$ is defined below.

Since it is easier to parametrize a real symmetric matrix than a unitary matrix, we then employ the dimensionless, real, and symmetric K -matrix. Defining the transition operator T via $S = 1 + iT$, then the K -matrix is given by

$$K = (2T^{-1} + i)^{-1}, \quad K^{-1} = 2T^{-1} + i, \quad (345)$$

and hence,

$$S = (1 + iK)(1 - iK)^{-1} = (1 - iK)^{-1}(1 + iK). \quad (346)$$

The effective range expansion suggests that we write

$$K^{-1} = \mathcal{U}^{-1} \tilde{K}^{-1} \mathcal{U}^{-1}. \quad (347)$$

where we have introduced the diagonal matrix

$$\langle J'm_{J'}, L'S'a' | \mathcal{U} | Jm_J, LSa \rangle = q_{\text{cm},a}^{L+\frac{1}{2}} \delta_{J'J} \delta_{m_{J'}m_J} \delta_{L'L} \delta_{S'S} \delta_{a'a}. \quad (348)$$

In terms of the well-behaved \tilde{K} -matrix, the quantization condition is given in Eq. (324) by

$$\det[1 - B^{(\mathbf{P})} \tilde{K}] = 0, \quad \det[1 - \tilde{K} B^{(\mathbf{P})}] = 0, \quad (349)$$

or

$$\det[\tilde{K}^{-1} - B^{(\mathbf{P})}] = 0. \quad (350)$$

The \tilde{K} matrix depends only on E_{cm} , which does not involve the decay particle masses and is the same for all decay channels. The B matrix involves $\mathbf{q}_{\text{cm},a}^2$, which differs for each decay channel. From Eq. (325), the box matrix is given in the above basis by

$$\begin{aligned} \langle J'm_{J'} L'S'a' | B^{(\mathbf{P}a)} | Jm_J LSa \rangle &= -i \delta_{a'a} \delta_{S'S} \left(\frac{2\pi u_a}{L_V} \right)^{L'+L+1} W_{L'm_{L'}; Lm_L}^{(\mathbf{P}a)} \\ &\quad \times \langle J'm_{J'} | L'm_{L'}, S m_S \rangle \langle Lm_L, S m_S | Jm_J \rangle. \end{aligned} \quad (351)$$

S is a total intrinsic spin in the above equation. Notice that $B^{(\mathbf{P})}$ is diagonal in channel space, but mixes different total angular momentum sectors, whereas the scattering S -matrix and K -matrix are diagonal in angular momentum, but have off-diagonal elements in channel space. Also, the matrix elements of $B^{(\mathbf{P})}$ depend on the total momentum \mathbf{P} , whereas the matrix elements of S do not. The box matrix depends on L, L' , but is independent of a, a' . $B^{(\mathbf{P})}$ is Hermitian if u_a^2 is real. The F and B matrices are related by

$$F^{(\mathbf{P})} = \frac{1}{2}(1 + i\mathcal{B}^{(\mathbf{P})}), \quad (352)$$

where

$$B^{(\mathbf{P})} = \mathcal{U} \mathcal{B}^{(\mathbf{P})} \mathcal{U}. \quad (353)$$

In the above expressions, the W -matrix elements are given by

$$\begin{aligned} -iW_{L'm_{L'}; Lm_L}^{(\mathbf{P}a)} &= \sum_{l=|L'-L|}^{L'+L} \sum_{m=-l}^l \frac{\mathcal{Z}_{lm}(\mathbf{s}_a, \gamma, u_a^2)}{\pi^{3/2} \gamma u_a^{l+1}} \sqrt{\frac{(2L'+1)(2l+1)}{(2L+1)}} \\ &\quad \times \langle L'0, l0 | L0 \rangle \langle L'm_{L'}, lm | Lm_L \rangle, \end{aligned} \quad (354)$$

and the Rummukainen-Gottlieb-Lüscher (RGL) shifted zeta functions \mathcal{Z}_{lm} are evaluated using

$$\begin{aligned} \mathcal{Z}_{lm}(\mathbf{s}, \gamma, u^2) &= \sum_{\mathbf{n} \in \mathbb{Z}^3} \frac{\mathcal{Y}_{lm}(\mathbf{z})}{(\mathbf{z}^2 - u^2)} e^{-\Lambda(\mathbf{z}^2 - u^2)} + \delta_{l0} \frac{\gamma\pi}{\sqrt{\Lambda}} F_0(\Lambda u^2) \\ &\quad + \frac{i^l \gamma}{\Lambda^{l+1/2}} \int_0^1 dt \left(\frac{\pi}{t}\right)^{l+3/2} e^{\Lambda t u^2} \sum_{\substack{\mathbf{n} \in \mathbb{Z}^3 \\ \mathbf{n} \neq 0}} e^{\pi i \mathbf{n} \cdot \mathbf{s}} \mathcal{Y}_{lm}(\mathbf{w}) e^{-\pi^2 \mathbf{w}^2 / (t\Lambda)}, \end{aligned} \quad (355)$$

with $\mathcal{Y}_{lm}(\mathbf{x}) = |\mathbf{x}|^l Y_{lm}(\hat{\mathbf{x}})$ and

$$\mathbf{z} = \mathbf{n} - \gamma^{-1} \left[\frac{1}{2} + (\gamma - 1) s^{-2} \mathbf{n} \cdot \mathbf{s} \right] \mathbf{s}, \quad (356)$$

$$\mathbf{w} = \mathbf{n} - (1 - \gamma) s^{-2} \mathbf{s} \cdot \mathbf{n} \mathbf{s}, \quad \mathcal{Y}_{lm}(\mathbf{x}) = |\mathbf{x}|^l Y_{lm}(\hat{\mathbf{x}}), \quad (357)$$

$$F_0(x) = -1 + \frac{1}{2} \int_0^1 dt \frac{e^{tx} - 1}{t^{3/2}}. \quad (358)$$

We choose $\Lambda \approx 1$ for fast convergence of the summation, and the integral in Eq. (355) is efficiently done using Gauss-Legendre quadrature. The function $F_0(x)$ is given in terms of the Dawson or erf function:

$$F_0(x) = \begin{cases} e^x (2\sqrt{x} D(\sqrt{x}) - 1), & (x \geq 0), \\ -e^x - \sqrt{-x\pi} \operatorname{erf}(\sqrt{-x}), & (x < 0). \end{cases} \quad (359)$$

12 Block diagonalization

So far we have determined both the matrix F and the scattering matrix S in terms of the orthonormal cm-frame basis states labelled $|Jm_J L S a\rangle$. In this basis, the quantization condition

$$\det[1 + F^{(\mathbf{P})}(S - 1)] = 0, \quad (360)$$

Defining

$$V = (1 + iM^{(d)})(1 - iM^{(d)})^{-1}, \quad (441)$$

and factoring, the above matrix becomes

$$D = \frac{1}{2}\rho^{1/2} (1 + SV) (1 - iM^{(d)})\rho^{-1/2}. \quad (442)$$

The matrix V is now well behaved. The eigenvectors of $M^{(d)}$ are also the eigenvectors of $1 \pm iM^{(d)}$, so when a divergence in an eigenvalue of $M^{(d)}$ occurs, the eigenvalue of V corresponding to the diverging eigenvalue of $M^{(d)}$ tends to unity. The divergences at non-interacting energies have been factorized off into the $(1 - iM^{(d)})$ factor. The zeros of $\det D$ are also the zeros of $\det(1 + SV)$.

We can repeat these same manipulations using our matrices. From Eq. (323), we find that

$$F^{(P)} = \frac{1}{2} (1 + i\mathcal{B}^{(P)}). \quad (443)$$

Using the quantization condition in Eq. (211) and using $\det(1 + AB) = \det(1 + BA)$ for any invertible matrices A, B , we begin with the matrix

$$1 + (S - 1)F^{(P)} = 1 + \frac{1}{2}(S - 1) (1 + i\mathcal{B}^{(P)}), \quad (444)$$

$$= 1 + \frac{1}{2}S (1 + i\mathcal{B}^{(P)}) - \frac{1}{2} (1 + i\mathcal{B}^{(P)}), \quad (445)$$

$$= \frac{1}{2}S (1 + i\mathcal{B}^{(P)}) + \frac{1}{2} (1 - i\mathcal{B}^{(P)}), \quad (446)$$

$$= \frac{1}{2}S (1 + i\mathcal{B}^{(P)}) (1 - i\mathcal{B}^{(P)})^{-1} (1 - i\mathcal{B}^{(P)}) + \frac{1}{2} (1 - i\mathcal{B}^{(P)}), \quad (447)$$

$$= \frac{1}{2} \left(1 + S (1 + i\mathcal{B}^{(P)}) (1 - i\mathcal{B}^{(P)})^{-1} \right) (1 - i\mathcal{B}^{(P)}). \quad (448)$$

The divergences at the noninteracting energies have been factorized into the $(1 - i\mathcal{B}^{(P)})$ factor on the right. So our tamed quantization condition becomes

$$\det (1 + S(1 + i\mathcal{B}^{(P)})(1 - i\mathcal{B}^{(P)})^{-1}) = 0. \quad (449)$$

As a reminder, we have (see Eq. (79))

$$S = (1 + iK)(1 - iK)^{-1} = (1 - iK)^{-1}(1 + iK). \quad (450)$$

However, the above expression is not ideal for us. First, we prefer to work directly with the box matrix $B^{(P)} = \mathcal{U}\mathcal{B}^{(P)}\mathcal{U}$ and the matrix $\tilde{K} = \mathcal{U}^{-1}K\mathcal{U}^{-1}$ or $\tilde{K}^{-1} = \mathcal{U}K^{-1}\mathcal{U}$, where the diagonal matrix \mathcal{U} of threshold factors is defined in Eq. (329). But the basic idea is to use a Cayley transformation, so this can be carried out using the \tilde{K} and $B^{(P)}$ matrices directly, instead of K and $\mathcal{B}^{(P)}$.

A Cayley transformation is a mapping between a skew-Hermitian matrix A and a unitary matrix Q which does not have -1 as an eigenvalue given by

$$Q = (1 - A)(1 + A)^{-1} = (1 + A)^{-1}(1 - A), \quad A = (1 - Q)(1 + Q)^{-1}. \quad (451)$$

Properties of this mapping include: $1 + A$ and $1 + Q$ are always invertible; the matrix A commutes with $(\mu + A)^{-1}$, where μ is a non-zero real scalar number.

(a) First, we show that $\mu \pm A$ is invertible, where μ is a non-zero real scalar number. If A is skew-Hermitian, then $A = iB$, where B is Hermitian so that $A^\dagger = -iB^\dagger = -iB = -A$. This means that A can be diagonalized by a unitary matrix U_A and all of its eigenvalues must be pure imaginary. It follows that $\mu \pm A$ is diagonalized by the same unitary matrix U_A , but now all of its eigenvalues have the form $\mu \pm i\alpha$, where α is real. Since μ is real and nonzero, none of the eigenvalues of $\mu \pm A$ can be zero, so $\det(\mu \pm A) \neq 0$, which means it must be invertible. **(b)** Next, we show that A commutes with $(\mu \pm A)^{-1}$, where μ is a nonzero real scalar. To start, we have $(\mu + A)(\mu - A) = \mu^2 - \mu A + \mu A - A^2 = (\mu - A)(\mu + A)$. Since $\mu \pm A$ is invertible, we multiply this equation on the left by $(\mu + A)^{-1}$ and on the right by $(\mu + A)^{-1}$ to obtain $(\mu - A)(\mu + A)^{-1} = (\mu + A)^{-1}(\mu - A)$. This shows that $\mu - A$ commutes with $(\mu + A)^{-1}$. But μ is a diagonal matrix, so it commutes with $(\mu + A)^{-1}$. This means A itself must commute with $(\mu + A)^{-1}$. By a similar argument, it is also easy to show that A must also commute with $(\mu - A)^{-1}$. **(c)** Using the previous two properties, we next show that Q must be unitary and that it cannot have -1 as one of its eigenvalues. To start, we see that $Q^\dagger = ((1 - A)(1 + A)^{-1})^\dagger$. Using $(B^{-1})^\dagger = (B^\dagger)^{-1}$ and $(AB)^\dagger = B^\dagger A^\dagger$, we then get $Q^\dagger = ((1 + A)^\dagger)^{-1}(1 - A)^\dagger = (1 - A)^{-1}(1 + A)$. Lastly, we previously showed that $1 + A$ and $(1 - A)^{-1}$ commute, so $Q^\dagger = (1 + A)(1 - A)^{-1}$. Using this, we then have $Q^\dagger Q = (1 + A)(1 - A)^{-1}(1 - A)(1 + A)^{-1} = 1$. Similarly, $Q Q^\dagger = (Q^\dagger Q)^\dagger = 1^\dagger = 1$. This shows that Q is unitary. Let $i\alpha_j$ denote the eigenvalues of A , where the α_j must be real. Then $\det Q = \prod_j (1 - i\alpha_j) \prod_k (1 + i\alpha_k)^{-1} = \prod_j (1 - i\alpha_j)^2 / (1 + \alpha_j^2) = \prod_j q_j$. This shows that the eigenvalues of Q are given in terms of the eigenvalues $i\alpha_j$ of A by

$$q_j = \frac{(1 - \alpha_j^2) - 2i\alpha_j}{(1 + \alpha_j^2)}, \quad |q_j| = 1. \quad (452)$$

Notice that each eigenvalue of Q has unit modulus, so $q_j = \exp(i\theta_j)$ for real θ_j . As $\alpha_j \rightarrow 0$, then $q_j \rightarrow 1$. As $\alpha_j \rightarrow \infty$, then $q_j \rightarrow -1$. Given the above expression, the value -1 is excluded from being an eigenvalue unless $\alpha_j \rightarrow \infty$. **(d)** Given the definition of Q in terms of A above, we now show how to derive the inverse relationship for A in terms of Q . Starting with the definition of Q , we have

$$\begin{aligned} 1 + Q &= 1 + (1 - A)(1 + A)^{-1} = (1 + A)(1 + A)^{-1} + (1 - A)(1 + A)^{-1}, \\ &= (1 + A + 1 - A)(1 + A)^{-1} = 2(1 + A)^{-1}, \end{aligned} \quad (453)$$

$$\begin{aligned} 1 - Q &= (1 + A)(1 + A)^{-1} - (1 - A)(1 + A)^{-1} \\ &= (1 + A - 1 + A)(1 + A)^{-1} = 2A(1 + A)^{-1} \end{aligned} \quad (454)$$

Thus,

$$(1 - Q)(1 + Q)^{-1} = 2A(1 + A)^{-1} \frac{1}{2}(1 + A) = A, \quad (455)$$

$$(1 + Q)(1 - Q)^{-1} = 2(1 + A)^{-1} \frac{1}{2}(1 + A)A^{-1} = A^{-1}. \quad (456)$$

(e) To get an expression for Q in terms of A^{-1} , assuming it exists, let $B = A^{-1}$ and use the

last relationship above:

$$\begin{aligned} 1 + B &= 1 + (1 + Q)(1 - Q)^{-1} = (1 - Q)(1 - Q)^{-1} + (1 + Q)(1 - Q)^{-1} \\ &= (1 - Q + 1 + Q)(1 - Q)^{-1} = 2(1 - Q)^{-1}, \end{aligned} \quad (457)$$

$$1 - B = 1 - (1 + Q)(1 - Q)^{-1} = -2Q(1 - Q)^{-1}. \quad (458)$$

This gives us

$$(1 - B)(1 + B)^{-1} = -2Q(1 - Q)^{-1} \frac{1}{2}(1 - Q) = -Q. \quad (459)$$

Hence,

$$Q = -(1 - A^{-1})(1 + A^{-1})^{-1}. \quad (460)$$

Starting with the quantization condition $\det(1 - \tilde{K}B^{(\mathbf{P})}) = 0$, we wish to formulate this condition in terms of the Cayley transformed matrices. Note that since \tilde{K} and $B^{(\mathbf{P})}$ are block diagonal, their Cayley transformed matrices will similarly be block diagonal. Noting that $-i\tilde{K}$ and $-iB^{(\mathbf{P})}$ are skew-Hermitian, we define the Cayley transformed box matrix and scattering matrix by

$$C_B^{(\mathbf{P})} = (1 + iB^{(\mathbf{P})})(1 - iB^{(\mathbf{P})})^{-1} = (1 - iB^{(\mathbf{P})})^{-1}(1 + iB^{(\mathbf{P})}), \quad (461)$$

$$\tilde{S} = (1 + i\tilde{K})(1 - i\tilde{K})^{-1} = (1 - i\tilde{K})^{-1}(1 + i\tilde{K}), \quad (462)$$

$$= -(1 - i\tilde{K}^{-1})(1 + i\tilde{K}^{-1})^{-1}, \quad (463)$$

then the inverse Cayley transformation is

$$-iB^{(\mathbf{P})} = (1 - C_B^{(\mathbf{P})})(1 + C_B^{(\mathbf{P})})^{-1}, \quad (464)$$

$$-i\tilde{K} = (1 - \tilde{S})(1 + \tilde{S})^{-1}, \quad (465)$$

which gives us

$$\begin{aligned} 1 - \tilde{K}B^{(\mathbf{P})} &= 1 + (-i\tilde{K})(-iB^{(\mathbf{P})}) = 1 + (1 - \tilde{S})(1 + \tilde{S})^{-1}(1 - C_B^{(\mathbf{P})})(1 + C_B^{(\mathbf{P})})^{-1} \\ &= 1 + (1 + \tilde{S})^{-1}(1 - \tilde{S})(1 - C_B^{(\mathbf{P})})(1 + C_B^{(\mathbf{P})})^{-1} \\ &= (1 + \tilde{S})^{-1}(1 + \tilde{S})(1 + C_B^{(\mathbf{P})})(1 + C_B^{(\mathbf{P})})^{-1} \\ &\quad + (1 + \tilde{S})^{-1}(1 - \tilde{S})(1 - C_B^{(\mathbf{P})})(1 + C_B^{(\mathbf{P})})^{-1} \\ &= (1 + \tilde{S})^{-1} \left[(1 + \tilde{S})(1 + C_B^{(\mathbf{P})}) + (1 - \tilde{S})(1 - C_B^{(\mathbf{P})}) \right] (1 + C_B^{(\mathbf{P})})^{-1} \\ &= 2(1 + \tilde{S})^{-1} (1 + \tilde{S}C_B^{(\mathbf{P})}) (1 + C_B^{(\mathbf{P})})^{-1}. \end{aligned} \quad (466)$$

The box matrix divergences at the noninteracting energies, causing the left-hand side of the above equation to diverge. These divergences at the noninteracting energies have now been all factorized into the $(1 + C_B^{(\mathbf{P})})^{-1}$ term on the right in the last line above. These divergences occur when an eigenvalue of $C_B^{(\mathbf{P})}$ becomes -1 . These divergences do not occur in the $(1 + \tilde{S})^{-1} (1 + \tilde{S}C_B^{(\mathbf{P})})$ terms. At energies away from the noninteracting energies, the box matrix is finite and the matrices $(1 + \tilde{S})$ and $(1 + C_B^{(\mathbf{P})})$ have nonzero determinants, making them invertible. To eliminate the divergences at the noninteracting energies of the box matrix, we simply factorize away the $(1 + C_B^{(\mathbf{P})})^{-1}$ term. The zeros of the quantization

condition we are after do not reside in the $(1 + \tilde{S})^{-1}$, so they must occur in the central matrix. Hence, our “tamed” quantization condition becomes

$$\det(1 + \tilde{S}C_B^{(\mathbf{P})}) = 0. \quad (467)$$

Again, \tilde{S} is unitary so its determinant cannot be zero and it is invertible, so an alternative quantization condition is

$$\det(\tilde{S}^{-1} + C_B^{(\mathbf{P})}) = 0. \quad (468)$$

Given the unitary natures of \tilde{S} and $C_B^{(\mathbf{P})}$, it is also possible that the determinants in Eqs. (467) and (468) might not be large, eliminating the need of using the Ω function.

The matrices inside the determinants in these equations are not Hermitian. However,

$$\begin{aligned} \det(1 + \tilde{S}C_B^{(\mathbf{P})}) &= \det(\tilde{S}\tilde{S}^{-1} + \tilde{S}C_B^{(\mathbf{P})}) = \det(\tilde{S})\det(\tilde{S}^{-1} + C_B^{(\mathbf{P})}) \\ &= \det(\tilde{S}^{-1} + C_B^{(\mathbf{P})})\det(\tilde{S}) = \det((\tilde{S}^{-1} + C_B^{(\mathbf{P})})\tilde{S}) \\ &= \det(1 + C_B^{(\mathbf{P})}\tilde{S}). \end{aligned} \quad (469)$$

Also, we know that $\tilde{S}C_B^{(\mathbf{P})}$ is unitary: $(\tilde{S}C_B^{(\mathbf{P})})^\dagger(\tilde{S}C_B^{(\mathbf{P})}) = (C_B^{(\mathbf{P})})^\dagger\tilde{S}^\dagger\tilde{S}C_B^{(\mathbf{P})} = 1$. The eigenvalues of a unitary matrix must be unimodular, so they can be written as $e^{i\theta}$ for some real θ . A unitary matrix is always unitarily diagonalizable. Let U_{SC} denote the unitary matrix that diagonalizes $\tilde{S}C_B^{(\mathbf{P})}$, then clearly U_{SC} also diagonalizes $1 + \tilde{S}C_B^{(\mathbf{P})}$. Hence,

$$\det(1 + \tilde{S}C_B^{(\mathbf{P})}) = \prod_j (1 + e^{i\theta_j}), \quad (470)$$

where θ_j are real and $e^{i\theta_j}$ are the eigenvalues of $\tilde{S}C_B^{(\mathbf{P})}$. The product of two nonzero complex numbers cannot be zero. For example, we know that for two complex number z_1 and z_2 , the modulus of the product is $|z_1 z_2| = |z_1||z_2|$ the product of the individual moduli. For the product $z_1 z_2$ to be zero, its modulus must be zero, which means that either $|z_1|$ or $|z_2|$ must be zero. This shows that every zero of $\det(1 + \tilde{S}C_B^{(\mathbf{P})})$ must come from at least one individual eigenvalue becoming zero. The product of any two nonzero eigenvalues cannot yield a zero. In other words, every zero of $\det(1 + \tilde{S}C_B^{(\mathbf{P})})$ comes from at least one $e^{i\theta_j} = -1$.

Note that this does *not* mean that whenever the real part of $\det(1 + \tilde{S}C_B^{(\mathbf{P})})$ becomes zero, the imaginary part must also be zero. The real part of $\det(1 + \tilde{S}C_B^{(\mathbf{P})})$ can be zero with the imaginary part being nonzero, and vice versa. We need *both* the real and the imaginary parts to become zero to satisfy the quantization condition.

In Sec. IIb of Ref. [30], the effect of at least one channel being closed is discussed. These authors mention that the finite-volume spectrum is sensitive to a closed channel in a very small energy region below a threshold, with the effect of the channel diminishing exponentially as $e^{-\kappa L}$, where κ is the magnitude of the imaginary scattering momentum. In the way the authors of Ref. [30] define their box matrix and scattering matrix, a loss of unitarity occurs below a threshold due to q_{cm} becoming imaginary. However, we have taken great pains to define our box matrix $B^{(\mathbf{P})}$ and scattering matrix \tilde{K} to remain Hermitian for negative \mathbf{q}_{cm}^2 . This was done by removing the threshold factors in K to produce \tilde{K} and absorbing them into $B^{(\mathbf{P})}$. See Eqs. (329), (330), and (333).

In summary, to tame the quantization condition, we should use

$$\det(1 + \tilde{S}C_B^{(\mathbf{P})}) = 0 \quad \text{or} \quad \det(\tilde{S}^{-1} + C_B^{(\mathbf{P})}) = 0, \quad (471)$$

defining the Cayley transformed box and K -matrices by

$$C_B^{(\mathbf{P})} = (1 + iB^{(\mathbf{P})})(1 - iB^{(\mathbf{P})})^{-1} = (1 - iB^{(\mathbf{P})})^{-1}(1 + iB^{(\mathbf{P})}), \quad (472)$$

$$\tilde{S} = (1 + i\tilde{K})(1 - i\tilde{K})^{-1} = (1 - i\tilde{K})^{-1}(1 + i\tilde{K}), \quad (473)$$

$$= -(1 - i\tilde{K}^{-1})(1 + i\tilde{K}^{-1})^{-1} = -(1 + i\tilde{K}^{-1})^{-1}(1 - i\tilde{K}^{-1}). \quad (474)$$

15 Fitting

Let κ_j , for $j = 1, \dots, N_K$, denote the parameters that appear in the matrix elements of either the \tilde{K} matrix or its inverse \tilde{K}^{-1} . Once a set of energies for a variety of two-particle interacting states is determined, the primary goal is then to determine the best-fit estimates of the κ_j parameters using the quantization determinant, as well as to determine the uncertainties in these estimates. In this section, we describe three methods to achieve this.

To set the stage, we first summarize the fitting procedure commonly used in lattice QCD. Observables which can be estimated directly with the Monte Carlo method as an average over an ensemble of gauge field configurations, such as temporal correlations of field operators, we refer to as *primary* observables. *Secondary* observables, such as energies and the lattice anisotropy, cannot be estimated directly with the Monte Carlo method but must be obtained indirectly from other observables. For an observable O that is either primary or secondary, let $\mathcal{E}(O)$ denote a Monte Carlo estimate of the observable obtained using the entire Markov-Chain ensemble of gauge configurations, and let $\mathcal{E}_k^{(r)}(O)$ denote an estimate of O from the k -th resampling of a resampling scheme r . The two most common resampling schemes are the jackknife $r = J$ and the bootstrap $r = B$. For a primary observable O , let $O[U_i]$ denote the value of the observable as calculated on the i -th configuration in the ensemble, and let N_e denote the number of gauge configurations in the Markov-chain ensemble, then the above quantities are given by

$$\mathcal{E}(O) = \frac{1}{N_e} \sum_{i=1}^{N_e} O[U_i], \quad (475)$$

$$\mathcal{E}_k^{(J)}(O) = \frac{1}{(N_e - 1)} \sum_{i=1, i \neq k}^{N_e} O[U_i], \quad (476)$$

$$\mathcal{E}_k^{(B)}(O) = \frac{1}{N_e} \sum_{i=1}^{N_e} O[U_{B(ki)}], \quad (477)$$

where $B(ki)$ is the i -th random draw in the k -th bootstrap resampling. The covariance of the estimates of two observables O_i and O_j can be estimated from their resampling estimates

using

$$\text{cov}(O_i, O_j) \approx \mathcal{N}^{(r)} \sum_{k=1}^{N_r} \left(\mathcal{E}_k^{(r)}(O_i) - \langle \mathcal{E}^{(r)}(O_i) \rangle \right) \left(\mathcal{E}_k^{(r)}(O_j) - \langle \mathcal{E}^{(r)}(O_j) \rangle \right), \quad (478)$$

$$\langle \mathcal{E}^{(r)}(O_i) \rangle = \frac{1}{N_r} \sum_{k=1}^{N_r} \mathcal{E}_k^{(r)}(O_i), \quad (479)$$

where N_r is the number of resamplings and the factor $\mathcal{N}^{(r)}$ depends on the resampling scheme. For the jackknife and bootstrap methods, it is given by

$$\mathcal{N}^{(J)} = \frac{(N_J - 1)}{N_J}, \quad \mathcal{N}^{(B)} = \frac{1}{N_B - 1}. \quad (480)$$

For primary observables, these covariances can also be estimated by

$$\text{cov}(O_i, O_j) \approx \frac{1}{N_e(N_e - 1)} \sum_{k=1}^{N_e} \left(O_i[U_k] - \mathcal{E}(O_i) \right) \left(O_j[U_k] - \mathcal{E}(O_j) \right), \quad (481)$$

as long as autocorrelations are small. (*sample size effects* [31])

It often occurs that a set of observables is believed to be reasonably well described by a set of model functions containing unknown parameters. In such cases, the goal is usually to find best fit estimates of these parameters. Arrange the observables into the components of a vector \mathbf{R} and the fit parameters into a vector $\boldsymbol{\alpha}$. Denote the set of model functions by the vector $\mathbf{M}(\boldsymbol{\alpha}, \mathbf{R})$ which depend on the parameters and which might depend on the observables themselves. The i -th component of $\mathbf{M}(\boldsymbol{\alpha}, \mathbf{R})$ gives the model prediction for the observable corresponding to the i -th component of \mathbf{R} . In lattice QCD, we generally determine the best fit estimates of the $\boldsymbol{\alpha}$ parameters as the values which minimize a correlated- χ^2 of residuals given by

$$\chi^2 = \mathcal{E}(r_i) \sigma_{ij}^{-1} \mathcal{E}(r_j), \quad (482)$$

where the vector of residuals is defined by $\mathbf{r}(\mathbf{R}, \boldsymbol{\alpha}) = \mathbf{R} - \mathbf{M}(\boldsymbol{\alpha}, \mathbf{R})$ and $\sigma_{ij} = \text{cov}(r_i, r_j)$ is the covariance matrix of the residuals. Since the observables are usually obtained using the same ensemble of gauge field configurations, the residuals in Eq. (482) are not statistically independent so the presence of the covariance matrix in the likelihood function is very important.

Usually, the minimization of χ^2 with respect to the parameters $\boldsymbol{\alpha}$ is accomplished using computer software, such as MINUIT2. The covariance matrix should be positive definite, so its inverse can be obtained by a Cholesky decomposition. If the model estimates of any of the observables depend on any of the other observables, then the covariance matrix must be recomputed using Eq. (478) and inverted each time the parameters $\boldsymbol{\alpha}$ are changed during the minimization process, making for a rather laborious minimization.

A significant simplification occurs if the model estimates are completely independent of the observables. In this case, $\text{cov}(r_i, r_j) = \text{cov}(R_i, R_j)$, which needs to be computed and

inverted only once at the start of the minimization. This is easy to see using

$$\begin{aligned}
\langle r_i r_j \rangle - \langle r_i \rangle \langle r_j \rangle &= \langle (R_i - M_i)(R_j - M_j) \rangle - \langle R_i - M_i \rangle \langle R_j - M_j \rangle \\
&= \langle R_i R_j - M_i R_j - M_j R_i + M_i M_j \rangle - (\langle R_i \rangle - M_i)(\langle R_j \rangle - M_j) \\
&= \langle R_i R_j \rangle - M_i \langle R_j \rangle - M_j \langle R_i \rangle + M_i M_j \\
&\quad - (\langle R_i \rangle \langle R_j \rangle - M_j \langle R_i \rangle - M_i \langle R_j \rangle + M_i M_j) \\
&= \langle R_i R_j \rangle - \langle R_i \rangle \langle R_j \rangle, \quad (\mathbf{M} \text{ indep of } \mathbf{R}).
\end{aligned} \tag{483}$$

Hence, one sees that $\text{cov}(r_i, r_j) = \text{cov}(R_i, R_j)$ when \mathbf{M} is independent of \mathbf{R} .

The minimization software usually provides not only the best fit estimates of the parameters, but also the statistical uncertainties in these estimates, including the covariances between the different fit parameter estimates. The formulas used to obtain these estimates generally assume Gaussian probability distributions. The law of large numbers makes such assumptions reasonable for a large number of gauge configurations in the Monte Carlo Markov-chain. An alternative approach is to perform the minimization of

$$\chi_k^2 = \mathcal{E}_k^{(r)}(r_i) \sigma_{ij}^{-1} \mathcal{E}_k^{(r)}(r_j), \tag{484}$$

for each resampling k and obtain the covariance of the fit parameter estimates $\text{cov}(\alpha_k, \alpha_l)$ using Eq. (478). A third approach is to carry out a bootstrap resampling of the original bootstrap estimates, recomputing estimates of σ_{ij}^{-1} , but in practice, the results of such a procedure are usually indistinguishable from those obtained using Eq. (484).

Best fit estimates of the \tilde{K} matrix parameters can be improved by utilizing results from multiple Markov-chain ensembles and lattices. One approach to performing such fits is to minimize the χ^2 of Eq. (482) taking the elements of σ_{ij} to be zero between the estimates from different ensembles, then obtain the covariances of the best fit parameter estimates from the minimization software. An alternative approach is to ensure that N_r is the same for all ensembles, then use the resamplings of all ensembles in the χ^2 of Eq. (482) with the covariance matrix estimated using Eq. (478). Given the statistical independence of the different ensembles, Eq. (478) naturally yields covariances between observable estimates from different ensembles which are very nearly zero.

Again, the primary goal is to determine the best-fit estimates of the κ_j parameters appearing in \tilde{K} or \tilde{K}^{-1} from the quantization condition, as well as to determine the uncertainties in these estimates. Having made the above introductory comments, we now describe three methods to achieve this.

15.1 Spectrum method

For each \mathbf{P} and irrep Λ , one obtains as many lab frame two-interacting-particle energies E_k as possible, staying below three-particle thresholds. For the observations R_i in the fit, an obvious choice would be to include the lab-frame energies E_k or the center-of-momentum energies $E_{\text{cm},k}$. Here, we choose the $E_{\text{cm},k}$ energies. The quantization condition with the chosen functional forms of \tilde{K} or \tilde{K}^{-1} then provides the model predictions of the observations. This involves scanning the quantization determinant in E_{cm} to find the values that result in the determinant having zero value. Evaluating the determinant requires evaluating the

box matrix elements, which requires knowing \mathbf{s}_a, u_a^2 for each channel a . To determine \mathbf{s}_a, u_a^2 , one needs to know the masses m_{1a}, m_{2a} in each decay channel, the spatial lattice volume L^3 , and the lattice aspect ratio $\xi = a_s/a_t$ if an anisotropic lattice is used. Unfortunately, these quantities must be obtained from the Monte Carlo simulations, and hence, are observations. This poses the problem that the predictions cannot be obtained solely from the parameters of the model, independent of the observations.

A simple way around this problem is to include the masses m_{1a}, m_{2a} in each decay channel, the spatial lattice volume L^3 , and the lattice aspect ratio $\xi = a_s/a_t$ as both observations and model parameters. In addition to the energy observations $E_{\text{cm},k}^{(\text{obs})}$, one also includes $m_j^{(\text{obs})}$, $L^{(\text{obs})}$, $\xi^{(\text{obs})}$, to the set of observations R_i , where $j = 1, \dots, N_p$ and N_p is the number of different particle species in all of the decay channels. At the same time, one introduces model parameters $m_j^{(\text{model})}$, $L^{(\text{model})}$, $\xi^{(\text{model})}$. Now, in scanning for the zeros of the determinant, one varies E_{cm} , evaluating \mathbf{s}_a, u_a^2 , and hence the box matrix elements, using the model parameters $m_j^{(\text{model})}$, $L^{(\text{model})}$, $\xi^{(\text{model})}$. In doing this simple trick, the model predictions are independent of the observations. This procedure is somewhat in the spirit of introducing Lagrange multipliers in a minimization.

In summary, the observations in the χ^2 minimization in this first method are

$$\text{Observations } R_i: \{ E_{\text{cm},k}^{(\text{obs})}, m_j^{(\text{obs})}, L^{(\text{obs})}, \xi^{(\text{obs})} \}, \quad (485)$$

for $k = 1, \dots, N_E$ and $j = 1, \dots, N_p$. If there are N_p particle species in all of the decay channels and N_E energies found, then there are $N_{\text{obs}} = 2 + N_p + N_E$ observations. Improved results can be obtained by increasing N_E by using several different \mathbf{P}, Λ blocks. The model parameters are

$$\text{Model fit parameters } \alpha_k: \{ \kappa_i, m_j^{(\text{model})}, L^{(\text{model})}, \xi^{(\text{model})} \}, \quad (486)$$

for $i = 1, \dots, N_K$ and $j = 1, \dots, N_p$, where N_K is the total number of parameters in the \tilde{K} -matrix elements. The total number of fit parameters is $N_{\text{param}} = 2 + N_p + N_K$.

Evaluating the predictions $M_i(\boldsymbol{\alpha})$ of the model for the N_{obs} observations is done as follows. The parameters $m_j^{(\text{model})}$, $L^{(\text{model})}$, $\xi^{(\text{model})}$ themselves give the predictions for the observations $m_j^{(\text{obs})}$, $L^{(\text{obs})}$, $\xi^{(\text{obs})}$. The model predictions corresponding to the $E_{\text{cm},k}^{(\text{obs})}$ observations are not so easily done. One needs to scan the quantization determinant in E_{cm} to find which values yield a zero value. For a given E_{cm} , one uses the parameters κ_j to evaluate the \tilde{K} matrix or its inverse, and determines the box matrix elements in terms of the RGL zeta functions using the parameters $m_j^{(\text{model})}$, $L^{(\text{model})}$, $\xi^{(\text{model})}$ to determine \mathbf{s}_a, u_a^2 . This is a rather onerous task. Computing the determinant for a given E_{cm} is quite complicated, and this must be done many times in order to bracket and then numerically find all of the needed zeros of the determinant using bisection or a Newton-Raphson type algorithm. One must then match each root found with the appropriate observed E_{cm} . Let $E_{\text{cm},k}^{(\text{model})}$ denote

each energy root found. In summary, the residuals in this method are

$$r_k = \begin{cases} E_{\text{cm},k}^{(\text{obs})} - E_{\text{cm},k}^{(\text{model})}, & (k = 1, \dots, N_E), \\ m_{k'}^{(\text{obs})} - m_{k'}^{(\text{model})}, & (k = k' + N_E, \ k' = 1, \dots, N_p), \\ L^{(\text{obs})} - L^{(\text{model})}, & (k = N_E + N_p + 1), \\ \xi^{(\text{obs})} - \xi^{(\text{model})}, & (k = N_E + N_p + 2). \end{cases} \quad (487)$$

We emphasize that, in this method, the $E_{\text{cm},k}^{(\text{model})}$ are very difficult quantities to compute using the model parameters in Eq. (486). The difficulty in carrying out this method leads us to seek simpler methods.

15.2 Determinant residual method

The difficulty in calculating the model predictions in the first method leads us to seek other simpler methods. In this second method, we introduce the quantization determinant itself as a residual. In the determinant, we use the observed box matrix elements, which requires the observed energies and the observed values for the particle masses, lattice size, and anisotropy.

Expressing the quantization condition in terms of a vanishing determinant is just a convenient way of stating that one eigenvalue becomes zero. The determinant itself is not a good quantity to use as an observable since it can become very large in magnitude for larger matrices. Instead of the determinant, we express the quantization condition using the following filter function of matrix A , having real determinant, and scalar $\mu \neq 0$:

$$\Omega(\mu, A) \equiv \frac{\det(A)}{\det[(\mu^2 + AA^\dagger)^{1/2}]}. \quad (488)$$

When one of the eigenvalues of A is zero, this function is also zero. This function can be evaluated as a product of terms, one for each eigenvalue of A . For eigenvalues of A which are much smaller than $|\mu|$, the associated term in the product tends towards the eigenvalue itself, divided by $|\mu|$. However, the key feature of this function is that for eigenvalues which are much larger than $|\mu|$, the associated term in the product goes to $e^{i\theta}$ for real θ . This function replaces the large unimportant eigenvalues with unimodular quantities so that the function does not grow with increasing matrix size. This is a much better behaved function, bounded between -1 and 1, which still reproduces the quantization condition. The constant μ can be chosen to optimize ease of numerical root finding or χ^2 minimization.

In this method, the model fit parameters are just the κ_i parameters, and the residuals are chosen to be

$$r_k = \Omega\left(\mu, 1 - B^{(\mathbf{P})}(E_{\text{cm},k}^{(\text{obs})}) \tilde{K}(E_{\text{cm},k}^{(\text{obs})})\right), \quad (k = 1, \dots, N_E), \quad (489)$$

or the matrix $\tilde{K}(E_{\text{cm},k}^{(\text{obs})})^{-1} - B^{(\mathbf{P})}(E_{\text{cm},k}^{(\text{obs})})$ could be used in the Ω function.

Clearly, the model predictions in this method are dependent on the observations themselves, so the covariance of the residual estimates must be recomputed and inverted by Cholesky decomposition throughout the minimization as the κ_j parameters are adjusted.

However, this is still much simpler than the root finding required in the spectrum method. One advantage of this method is that the complicated RGL zeta functions only need to be computed for the box matrix elements as observables; they do not need to be recomputed as model parameters are changed. Since we cannot completely remove the dependence of the model predictions on the observables in this method, there is no advantage in introducing model parameters for the energies, particle masses, and the lattice anisotropy. Hence, we do not need to recompute the box matrix elements as the model parameters are adjusted in the χ^2 minimization. The model predictions involve only the κ_j parameters and the observed energies, particle masses, and anisotropy.

If the number of basis states in the determinant computation is small enough such that the determinant can be explicitly expressed as a sum of a small number of terms, the computation of the covariance can be substantially simplified. In such cases, the center-of-momentum energies, particle masses, lattice size, and anisotropy can be introduced as both observables and model parameters, then the covariance can be expressed as a sum of terms, each being a product of a complicated quantity depending only on the observations and a very simple quantity depending only on the model parameters. Although the covariance matrix must be evaluated and inverted each time the model parameters are changed, the computation and inversion of this matrix can be done with minimal effort.

Treating the box matrix elements as observables enables a natural interface between the lattice calculation and phenomenology. If the box matrix elements and center-of-momentum energies are calculated, then together with the covariances, they contain all the information required to extract the scattering amplitudes. Non-lattice practitioners can use them without, for example, implementing the RGL zeta functions. These quantities can act as a bridge between lattice QCD computations and phenomenological applications.

16 Parametrization of the K -matrix

We now turn to the task of parametrizing the K matrix in terms of resonance parameters. We begin by examining the single channel case.

For a single channel, the S -matrix is $S = e^{2i\delta}$, so that the transition amplitude is

$$T = \frac{1}{2i}(S - 1) = \frac{1}{2i}(e^{2i\delta} - 1) = e^{i\delta} \frac{1}{2i}(e^{i\delta} - e^{-i\delta}) = e^{i\delta} \sin \delta, \quad (490)$$

and the K -matrix is

$$K = (T^{-1} + i)^{-1} = \left(\frac{e^{-i\delta}}{\sin \delta} + i \right)^{-1} = \left(\frac{\cos \delta - i \sin \delta + i \sin \delta}{\sin \delta} \right)^{-1} = \tan \delta. \quad (491)$$

The pole in K at $\delta = \pi/2$ is the signature of a resonance in this case.

For the multi-channel, multi-wave case, one common possibility is to parametrize the inverse of the \tilde{K} -matrix as a symmetric matrix of polynomials in E_{cm} :

$$\mathcal{K}_{\alpha\beta}^{(J)-1}(E_{\text{cm}}) = \sum_{k=0}^{N_{\alpha\beta}} c_{\alpha\beta}^{(Jk)} E_{\text{cm}}^k \quad (492)$$