Application of machine learning to nuclear reactions

Student Darius Likandrovas
Supervisor Dr. Arnoldas Deltuva

Supported by Lietuvos mokslo taryba, No. S-CERN-24-2

Vilnius University
Institute of Theoretical Physics and Astronomy

darius.likandrovas@ff.vu.lt

July 2, 2026

Overview

Introduction

- Momentum-space integral equation methods allow scattering problems to be solved reliably by using the screened Coulomb interaction

$$w_R(r) = \frac{\alpha Z_1 Z_2}{r} \exp\left[-(r/R)^n\right],$$

  together with a renormalization procedure, which makes it possible to apply standard scattering theory.
- At low energies, the convergence of momentum-space integral equation methods with respect to R is slow, so the calculations become less accurate.

Aim and objectives

Aim

- To apply kernel ridge regression (KRR) with a Gaussian kernel to extrapolate phase shifts to the unscreened Coulomb-interaction limit $R \to \infty$, equivalently $1/R \to 0$.

Objectives

- To prepare proton–deuteron phase-shift data for the quartet channel $J^{\Pi} = 3/2^{+}$ for extrapolation with respect to the inverse Coulomb screening radius $1/R$.
- To apply the KRR model to predict phase shifts at larger screening radii and to investigate how the results depend on the selected sets of input screening radii.
- To calculate the Coulomb-modified effective-range expansion parameters a_c and r_c from the extrapolated phase shifts and compare them with the values reported by the Pisa group, $\tilde{a}_c = 13.90$ fm and $\tilde{r}_c = 2.02$ fm.

Screening and renormalization of the Coulomb interaction

The Coulomb potential is replaced by a screened potential

$$w_R(r) = \frac{\alpha Z_1 Z_2}{r} \exp[-(r/R)^n], \qquad n \geq 1.$$

The physical Coulomb interaction is recovered in the limit $R \to \infty$.

Although the screened interaction has a finite range, the scattering amplitude contains an R-dependent phase. Therefore, after solving the problem at finite R, the amplitude must be renormalized before taking the limit $R \to \infty$.

Two-body scattering

$$\langle \mathbf{p}_f | t | \mathbf{p}_i \rangle = \langle \mathbf{p}_f | t_C | \mathbf{p}_i \rangle + \lim_{R\to\infty} z_R^{-1/2}(p_f) \langle \mathbf{p}_f | t^{(R)} - t_R | \mathbf{p}_i \rangle z_R^{-1/2}(p_i).$$

Elastic proton–deuteron scattering

$$\langle \mathbf{q}_f | U | \mathbf{q}_i \rangle = \langle \mathbf{q}_f | T_C^{\mathrm{c.m.}} | \mathbf{q}_i \rangle + \lim_{R\to\infty} z_R^{-1/2}(q_f) \langle \mathbf{q}_f | U^{(R)} - T_R^{\mathrm{c.m.}} | \mathbf{q}_i \rangle z_R^{-1/2}(q_i).$$

Data and regression problem

Let the training data be

$$\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N}, \qquad x_i \in \mathcal{X}, \qquad y_i \in \mathbb{R}.$$

Here x_i is an input vector, y_i is the corresponding target value, and $\mathcal{X}$ denotes the non-empty input space.
The aim is to construct a function

$$f : \mathcal{X} \to \mathbb{R}$$

which maps each input x_i to a predicted output $f(x_i)$.
The function is chosen such that its predictions are close to the known target values

$$f(x_i) \approx y_i, \qquad i = 1, \ldots, N.$$

Mercer kernel

Definition 1

A kernel function

$$k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$$

is called symmetric if

$$k(x, z) = k(z, x) \qquad \forall x, z \in \mathcal{X}.$$

A symmetric kernel function is called a Mercer kernel if, for any $N \in \mathbb{N}$, any choice $x_1, \ldots, x_N \in \mathcal{X}$ and any vector $c = (c_1, \ldots, c_N)^T \in \mathbb{R}^N$, one has

$$\sum_{i=1}^{N} \sum_{j=1}^{N} c_i c_j k(x_i, x_j) \geq 0.$$

The associated Gram matrix is

$$K_{ij} = k(x_i, x_j), \qquad K = (K_{ij})_{i,j=1}^{N}.$$

Feature map

Definition 2

Let $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ be a Mercer kernel. A feature map is a mapping

$$\phi : \mathcal{X} \to H$$

into a Hilbert space H such that

$$k(x, z) = \langle \phi(x), \phi(z) \rangle_H \qquad \forall x, z \in \mathcal{X}.$$

The feature map gives a Hilbert-space representation of the input set. Even if $\mathcal{X}$ is only an abstract set, the mapped objects $\phi(x) \in H$ can be added, scaled and compared by the inner product of H.
For the training points $x_1, \ldots, x_N$, the Gram matrix is

$$K_{ij} = k(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle_H.$$

The feature-space coordinates $\phi(x)$ are never computed explicitly.

Reproducing kernel Hilbert space

A Mercer kernel k defines a Hilbert space of functions

$$H_k \subset \{f : \mathcal{X} \to \mathbb{R}\}.$$

Definition 3

The space H_k is called a reproducing kernel Hilbert space with kernel k if

$$k(\cdot, x) \in H_k \qquad \forall x \in \mathcal{X}$$

and

$$f(x) = \langle f, k(\cdot, x) \rangle_{H_k} \qquad \forall f \in H_k,\ \forall x \in \mathcal{X}.$$

The last identity is called the reproducing property. It means that function values can be computed by inner products in H_k.

KRR variational problem

Kernel ridge regression is formulated as a minimization problem over the RKHS H_k. The function to be determined is

$$f \in H_k.$$

For a given regularization parameter $\lambda > 0$, the function $\hat{f} \in H_k$ is chosen by solving

$$\hat{f} = \arg\min_{f \in H_k} \left[\sum_{i=1}^{N} \big(y_i - f(x_i)\big)^2 + \lambda \|f\|_{H_k}^2 \right].$$

Representer theorem

Theorem 1

Let H_k be an RKHS and consider the KRR minimization problem. Then its solution has the form

$$\hat{f}(\cdot) = \sum_{i=1}^{N} \alpha_i k(\cdot, x_i).$$

Thus the solution lies in the finite-dimensional subspace

$$\mathcal{M} = \text{span}\{k(\cdot, x_1), \ldots, k(\cdot, x_N)\}.$$

Therefore the minimization over functions

$$f \in H_k$$

is reduced to finding the coefficients

$$\alpha_1, \ldots, \alpha_N.$$

Idea of the proof

Let

$$\mathcal{M} = \text{span}\{k(\cdot, x_1), \ldots, k(\cdot, x_N)\}.$$

Since $\mathcal{M}$ is a finite-dimensional subspace of H_k, any function $f \in H_k$ can be decomposed as

$$f = f_{\|} + f_{\perp}, \qquad f_{\|} \in \mathcal{M}, \qquad f_{\perp} \perp \mathcal{M}.$$

The reproducing property gives the value of f at each training point

$$f(x_i) = \langle f, k(\cdot, x_i) \rangle_{H_k}.$$

Since $k(\cdot, x_i) \in \mathcal{M}$ and $f_{\perp} \perp \mathcal{M}$, it follows that

$$f(x_i) = \langle f_{\|}, k(\cdot, x_i) \rangle_{H_k} + \langle f_{\perp}, k(\cdot, x_i) \rangle_{H_k} = f_{\|}(x_i).$$

Thus the data fitting term depends only on $f_{\|}$. On the other hand

$$\|f\|_{H_k}^2 = \|f_{\|}\|_{H_k}^2 + \|f_{\perp}\|_{H_k}^2.$$

Therefore, the component $f_{\perp}$ does not improve the fit, but only increases the regularization term. Hence, the solution must satisfy

$$f_{\perp} = 0.$$

Finite-dimensional KRR problem

Using the representer theorem, the KRR solution has the form

$$\hat{f}(x_i) = \sum_{j=1}^{N} K_{ij}\alpha_j, \qquad K_{ij} = k(x_i, x_j).$$

The RKHS norm is

$$\|\hat{f}\|_{H_k}^2 = \alpha^T K \alpha.$$

Therefore, the KRR problem reduces to the finite-dimensional minimization

$$F(\alpha) = \|y - K\alpha\|^2 + \lambda \alpha^T K \alpha.$$

Differentiating with respect to α gives

$$\nabla_\alpha F(\alpha) = 2K\left[(K + \lambda I)\alpha - y\right].$$

Setting the gradient to zero gives the KRR linear system

$$(K + \lambda I)\alpha = y.$$

Prediction formula

For a new point $x_* \in \mathcal{X}$, the KRR prediction is

$$\hat{f}(x_*) = \sum_{i=1}^{N} \alpha_i k(x_*, x_i).$$

Thus, the prediction requires only the kernel values between the new point x_* and the training points $x_1, \ldots, x_N$.

Model application scheme

For each energy E, the dependence of the phase shift on the inverse screening radius is considered

$$\delta(E, 1/R).$$

The model input consists of the energy and several phase-shift values at selected finite screening radii

$$x_i = (E_i, \delta_i(1/R_1), \delta_i(1/R_2), \ldots, \delta_i(1/R_M)) .$$

The model predicts the phase-shift correction between the largest input screening radius $R_0 = \max\{R_1, \ldots, R_M\}$ and a larger radius R_∞,

$$y_i = \delta(E_i, 1/R_\infty) - \delta(E_i, 1/R_0).$$

The extrapolated phase shift is then obtained as

$$\hat{\delta}(E_i, 1/R_\infty) = \delta(E_i, 1/R_0) + \hat{y}_i.$$

Data

In this work, the proton–deuteron scattering quartet channel $J^{\Pi} = 3/2^{+}$ is considered.
The following low-energy values are used

$$E \in \{0.05, 0.10, 0.20, 0.40, 0.70, 1.00\}\ \mathrm{MeV}.$$

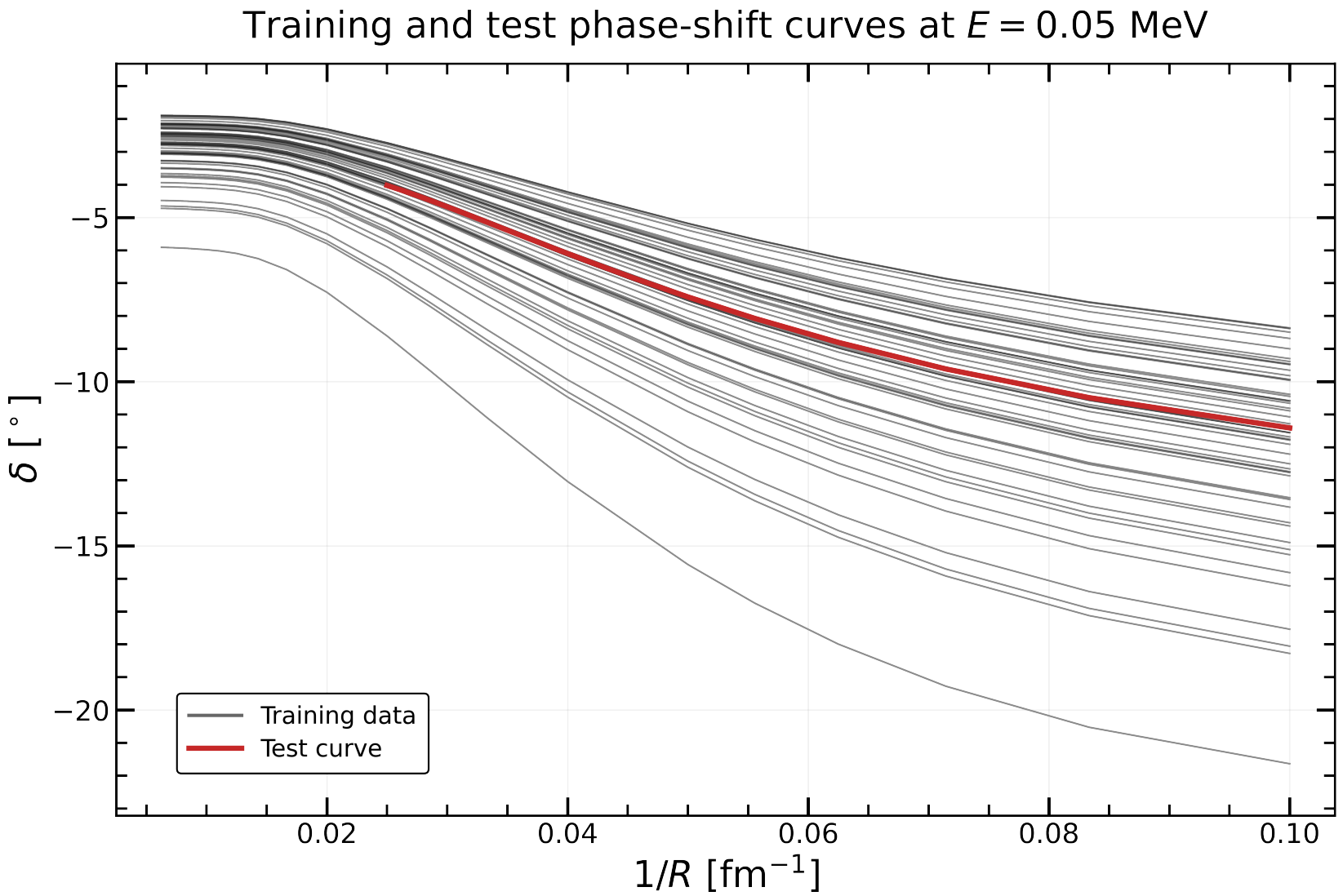
The phase-shift curves are given on a grid of finite Coulomb screening radii

$$R_n = 10 + 2n\ \mathrm{fm}, \qquad n = 0, 1, \ldots, 15.$$

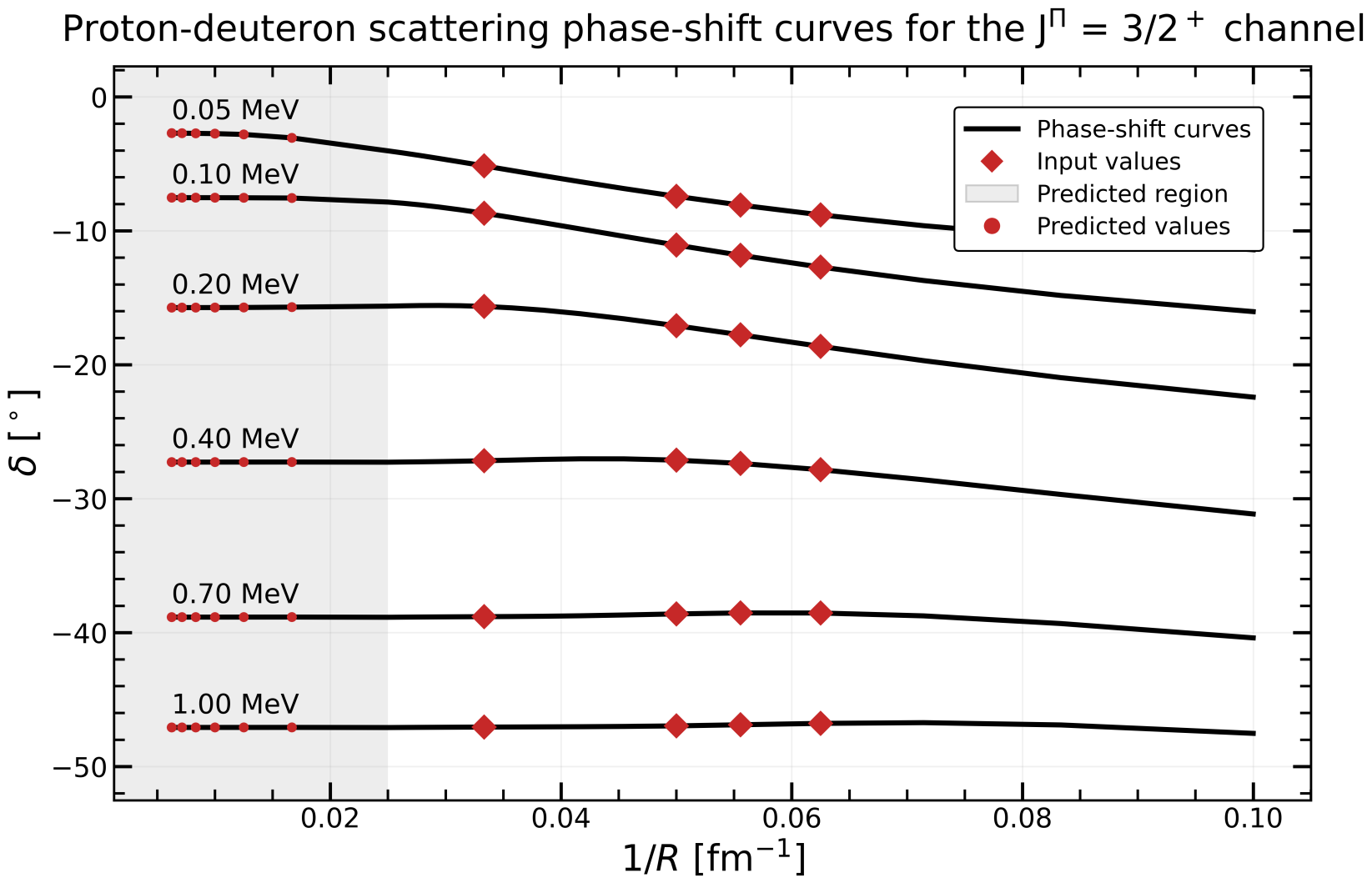
The extrapolation is performed to larger screening radii

$$R_\infty \in \{140, 160\}\ \mathrm{fm}.$$

Training and test curves



Phase-shift extrapolation



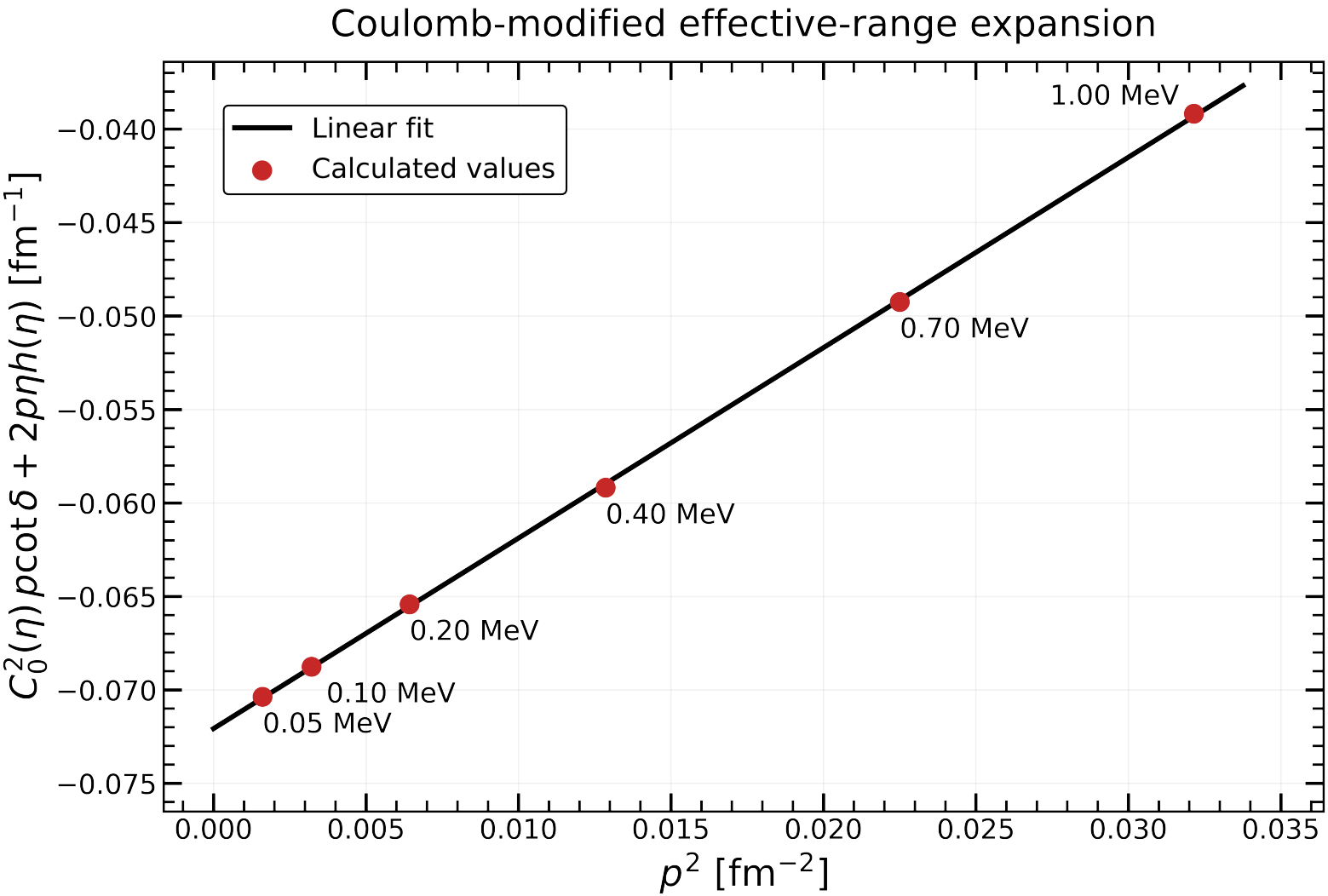
Coulomb-modified effective-range expansion

At low energies, the Coulomb-modified effective-range expansion is applied to the $J^\Pi = 3/2^+$ channel

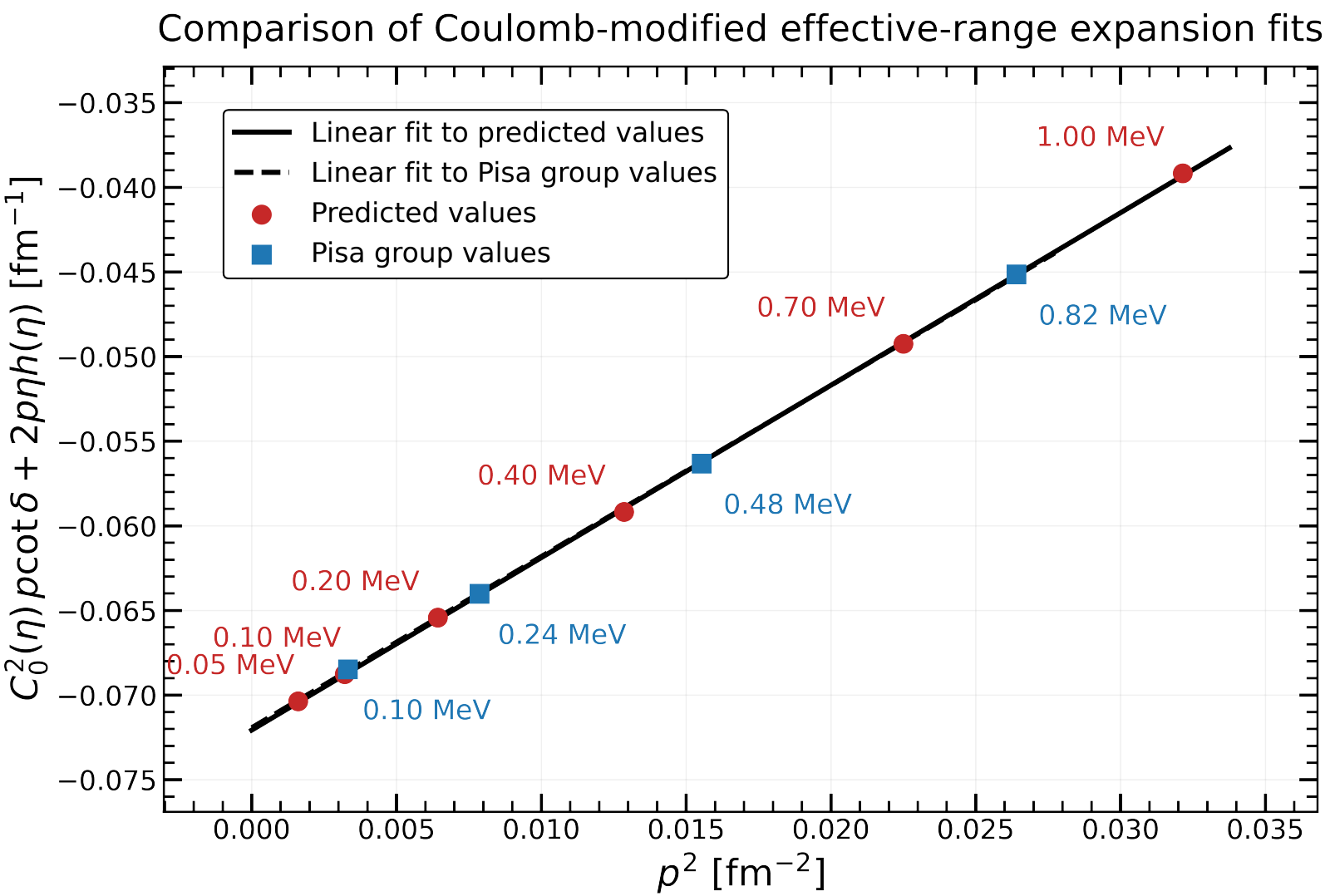
$$C_0^2(\eta) p \cot\delta + 2p\eta h(\eta) = -\frac{1}{a_c} + \frac{1}{2} r_c p^2.$$

Here a_c is the Coulomb-modified scattering length, r_c is the Coulomb-modified effective range, η is the Sommerfeld parameter, and $C_0^2(\eta)$ and $h(\eta)$ are standard Coulomb functions.

Application of the expansion



Comparison with the Pisa group



Sets of input screening radii

In the analysis, 18 different sets of input screening radii were tested. For each set, the parameters a_c and r_c were determined in six different ways.
The parameter variation was described by a central value and a variation interval. For $x = a_c$ or $x = r_c$, these quantities are defined as

$$x = \bar{x} \pm \Delta x, \qquad \bar{x} = \frac{x_{\max} + x_{\min}}{2}, \qquad \Delta x = \frac{x_{\max} - x_{\min}}{2}.$$

The deviation from the Pisa group values and the stability measure are defined as

$$D_{\mathrm{ref}} = \sqrt{(\bar{a}_c - \tilde{a}_c)^2 + (\bar{r}_c - \tilde{r}_c)^2}, \qquad D_{\mathrm{stab}} = \sqrt{(\Delta a_c)^2 + (\Delta r_c)^2}.$$

R [fm]	$\bar{a}_c$ [fm]	$\bar{r}_c$ [fm]	D_{ref} [fm]	D_{stab} [fm]
$(32, 26, 20, 14)$	13.89 ± 0.03	2.01 ± 0.04	0.012	0.050
$(18, 16, 14, 12)$	13.88 ± 0.04	2.01 ± 0.05	0.022	0.071
$(22, 20, 18, 16, 14)$	13.89 ± 0.04	2.01 ± 0.05	0.028	0.082
$(26, 24, 22, 20)$	13.89 ± 0.07	1.99 ± 0.07	0.032	0.103
$(26, 24, 22, 20, 18)$	13.89 ± 0.06	1.99 ± 0.06	0.036	0.092

Conclusions

- The KRR model was applied to phase-shift extrapolation with respect to $1/R$, predicting the phase-shift correction from the largest input screening radius R_0 to R_∞.
- A stability analysis was performed using 18 different input sets.
- For the extrapolation to $R_\infty = 160$ fm, the obtained values were $\bar{a}_c = 13.89 \pm 0.03$ fm and $\bar{r}_c = 2.01 \pm 0.04$ fm.
- These values agree, within the estimated uncertainty intervals, with the values calculated from the Pisa group phase shifts $\tilde{a}_c = 13.90$ fm and $\tilde{r}_c = 2.02$ fm.
- The calculated phase-shift curves are consistent with the physical meaning of the limit $R \to \infty$. As $1/R$ decreases, the dependence on the screening radius becomes weaker and the curves approach the asymptotic regime, although nontrivial oscillations remain in the pre-asymptotic region.