Thesis for the degree of Doctor of Philosophy

Extending the reach of uncertainty quantification in nuclear theory

Isak Svensson

Department of Physics
Chalmers University of Technology
Gothenburg, Sweden 2023

Extending the reach of uncertainty quantification in nuclear theory
Isak Svensson
ISBN: 978-91-7905-823-4.

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 5289
ISSN 0346-718X

Department of Physics
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone +46 (0)31-7721000

Cover: Bayesian posterior for the low-energy constants at next-to-next-to-next-to-leading order in chiral effective field theory. The posterior is conditioned on nucleon-nucleon elastic scattering data and was sampled using Hamiltonian Monte Carlo.

Chalmers Reproservice
Gothenburg, Sweden 2023

Extending the reach of uncertainty quantification in nuclear theory
Isak Svensson
Department of Physics
Chalmers University of Technology

Abstract

The theory of the strong interaction—quantum chromodynamics (QCD)—is unsuited to practical calculations of nuclear observables and approximate models for nuclear interaction potentials are required. In contrast to phenomenological models, chiral effective field theories (χEFTs) of QCD grant a handle on the theoretical uncertainty arising from the truncation of the chiral expansion. Uncertainties in χEFT are preferably quantified using Bayesian inference, but quantifying reliable posterior predictive distributions for nuclear observables presents several challenges. First, χEFT is parametrized by unknown low-energy constants (LECs) whose values must be inferred from low-energy data of nuclear structure and reaction observables. There are 31 LECs at fourth order in Weinberg power counting, leading to a high-dimensional inference problem which I approach by developing an advanced sampling protocol using Hamiltonian Monte Carlo (HMC). This allows me to quantify LEC posteriors up to and including fourth chiral order. Second, the χEFT truncation error is correlated across independent variables such as scattering energies and angles; I model correlations using a Gaussian process. Third, the computational cost of computing few- and many-nucleon observables typically precludes their direct use in Bayesian parameter estimation as each observable must be computed in excess of 10^5 times during HMC sampling. The one exception is nucleon-nucleon scattering observables, but even these incur a substantial computational cost in the present applications. I sidestep such issues using eigenvector-continuation emulators, which accurately mimic exact calculations while dramatically reducing the computational cost. Equipped with Bayesian posteriors for the LECs, and a model for the truncation error, I explore the predictive ability of χEFT, presenting the results as the probability distributions they always were.

Publications

This thesis is based on the work contained in the following papers.

I "Rigorous constraints on three-nucleon forces in chiral effective field theory from fast and accurate calculations of few-body observables"
S. Wesolowski, I. Svensson, A. Ekström, C. Forssén, R. J. Furnstahl, J. A. Melendez, and D. R. Phillips
Preprint: arXiv:2104.04441 [nucl-th]
Phys. Rev. C **104**, 064001 (2021)

II "Bayesian parameter estimation in chiral effective field theory using the Hamiltonian Monte Carlo method"
I. Svensson, A. Ekström, and C. Forssén
Preprint: arXiv:2110.04011 [nucl-th]
Phys. Rev. C **105**, 014004 (2022)

III "Bayesian estimation of the low-energy constants up to fourth order in the nucleon-nucleon sector of chiral effective field theory"
I. Svensson, A. Ekström, and C. Forssén
Preprint: arXiv:2206.08250 [nucl-th]
Phys. Rev. C **107**, 014001 (2023)

IV "Inference of the low-energy constants in delta-full chiral effective field theory including a correlated truncation error"
I. Svensson, A. Ekström, and C. Forssén
(in preparation)

This thesis is also based on the following research software. The code can be downloaded via the provided web address.

I "montepython" - An open-source research code for Hamiltonian Monte Carlo sampling of Bayesian posteriors and other distributions
I. Svensson
Licence: GPL
https://github.com/svisak/montepython

In addition, the author has been involved in the work leading to the following publication. This paper is not included in the thesis.

I "Effective field theory analysis of the Coulomb breakup of the one-neutron halo nucleus ^{19}C"
P. Capel *et al.*
Preprint: arXiv:2301.06444 [nucl-th]

Statement of contributions

My contributions to the included papers were:

I I performed the optimization of the nucleon-nucleon-sector LECs and wrote the corresponding parts of the manuscript.

II I developed the HMC sampling code, generated the numerical data and figures, performed the analyses, and wrote the majority of the manuscript.

III I generated all numerical data except Table II, generated the figures, performed the analyses, and wrote the manuscript together with my co-authors.

IV I developed the code, generated all numerical data, generated all figures except Figs. 2, 4, and 5, and contributed to the analyses and the manuscript together with my co-authors.

Acknowledgements

I will be brief, though heartfelt. My supervisors, Andreas Ekström and Christian Forssén, are simply extraordinary. *Thank you!*

Thank you also to my examiner, Tünde Fülöp, my director of studies, Christophe Demazière, and my overseas collaborators: Daniel, Sarah, Richard, and Jordan.

Many other people have put a golden sheen on my time at Chalmers, including (but not limited to) Sean Miller, Tor Djärv, Oliver Thim, Alexandra Roussou, Weiguang Jiang, Håkan Johansson, Erik Nilsson, Chieh-Jen Yang, Andreas Heinz, Maria Vittoria Managlia, Anna Kawecka, Matthias Holl, Timon Emken, Taylor Gray, Einar Urdshals, Vanessa Zema, Avik Banerjee, Gaute Hagen, Gabriele Ferretti, Riccardo Catena, Ulf Gran, Björn Jonson, and Thomas Nilsson. I will, or already do, miss you!

A very special thanks to Douglas Di Julio, who among other things suggested that I apply for this position.

I further wish to mention a few names without context and in no particular order: Staffan Rydén, Philip Dalsbecker, Catia Alves Fernandes, Didrik Almqvist, Karin Cederberg, Charles & Owen & Isaline, Benny Peterson, Daniela Grassi, Louise Liljeblad, and all the people I meet regularly and gladly but infrequently. I am almost certain I have forgotten someone important, sorry about that!

And finally, Viktor Friberg and Johan Gabrielsson, my sisters Elin and Åsa, and of course my mother Lena and father Gunnar, who are all quite wonderful.

Moon
1. The celestial orb which revolves round the earth; the satellite of the earth; a secondary planet, whose light, borrowed from the sun, is reflected to the earth, and serves to dispel the darkness of night.

– Webster's Dictionary, 1913

List of acronyms

Symbols

NN	nucleon-nucleon
NNN	three-nucleon
Nd	nucleon-deuteron
χEFT	chiral effective field theory
πN	pion-nucleon
nn	neutron-neutron
np	neutron-proton
pp	proton-proton
1PE	one-pion exchange
2NF	two-nucleon force
3NF	three-nucleon force

A

AD	automatic differentiation

C

c.m.	center-of-mass
CC	coupled cluster
CPU	central processing unit

E

EC	eigenvector continuation
EFT	effective field theory
ERE	effective range expansion
ESS	effective sample size

G

GP	Gaussian process
GPU	graphics processing unit

H

HMC	Hamiltonian Monte Carlo
HO	harmonic oscillator

I

IMSRG	in-medium similarity renormalization group

J

JIT	just-in-time compilation

L

LEC	low-energy constant
LHS	Latin hypercube sampling
LO	leading order
LS	Lippmann-Schwinger

M

MAP	maximum a posteriori
MCMC	Markov chain Monte Carlo
MLE	maximum likelihood estimation
MOR	model order reduction

N

N3LO	next-to-next-to-next-to-leading order
NCSM	no-core shell model
NLO	next-to-leading order
NNLO	next-to-next-to-leading order

P

PDF	probability density function
PMOR	parametric model order reduction
PPD	posterior predictive distribution
PW	partial wave

Q

QCD	quantum chromodynamics

R

RB	reduced basis
RBF	radial basis function
RMS	root mean squared

S

SIR	sampling/importance resampling
SM	Standard Model

U

UQ	uncertainty quantification

Contents

Chapter 1

Introduction

Since its discovery, quantum chromodynamics (QCD) has remained our most fundamental theory for describing the dynamics of strongly interacting particles. QCD is part of the Standard Model (SM) of particle physics and describes how quarks and gluons interact via their force-carriers, the gluons themselves. The quarks are confined into composite particles—hadrons—such as the nucleon. Nuclear forces arise from residual quark and gluon interactions in a similar manner to how the Van der Waals forces arise between molecules [1]. One can pursue QCD calculations on a finite space-time lattice using the lattice QCD method [2], but such calculations are currently too computationally expensive for general use in nuclear theory [3]. Nuclear theory is more commonly formulated in terms of nucleons and the forces acting between them.

The quantum state $|\psi\rangle$ of an interacting A-nucleon system is governed by the time-independent non-relativistic A-nucleon Schrödinger equation

$$(\hat{H}_0 + \hat{V}_{NN} + \hat{V}_{NNN} + \ldots)\,|\psi\rangle = E\,|\psi\rangle\,, \tag{1.1}$$

where $\hat{H}_0$ denotes the kinetic energy operator of the free nucleons, $\hat{V}_{NN}$ ($\hat{V}_{NNN}$) is a two-nucleon (three-nucleon) potential operator, and E is the eigenenergy of the system. Theoretical predictions for nuclear observables may be found by numerically solving Eq. (1.1), a problem physicists have been working on for decades [4]. From now on, I will for simplicity denote Eq. (1.1) as "the Schrödinger equation".

Today, *ab initio* many-body methods are routinely used for solving the Schrödinger equation. *Ab initio* (literally meaning "from the beginning") is a somewhat broad term, but an *ab initio* method typically entails adopting a relevant set of degrees of freedom—often nucleons and pions—and performing calculations in a systematically improvable way [5]. Recent years have seen dramatic improvements in the capabilities of *ab initio* methods (see Ref. [6] for a recent review).

Computing nucleon-nucleon (NN) elastic scattering cross sections via the Lippmann-Schwinger (LS) equation is straightforward and has been possible for more than half a century [7], while three-nucleon (NNN) scattering presents more of a challenge and is currently attracting attention [8–11] with or without three-nucleon forces (3NFs). This is a clear example of how rapidly—naïvely exponentially—the computational complexity of Eq. (1.1) increases with A. Practitioners seeking to study few-nucleon systems frequently turn to the no-core shell model (NCSM) [12, 13], which considers the dynamics of all nucleons. The NCSM basis must in practice be truncated at some finite number of harmonic oscillator (HO) excitations above the lowest configuration, limited by a parameter $N_{\max}$. The number of basis states grows roughly exponentially with $N_{\max}$, and the dimensionality of the Hilbert space reaches the order 10^4, in Jacobi coordinates, for the $A = 3, 4$ observables in Paper I. The lowest eigenvalue and eigenstate takes about 1 minute to solve for using a Lanczos algorithm [14]. Beyond the reach of NCSM, polynomially-scaling methods like the in-medium similarity renormalization group (IMSRG) [15] and coupled cluster (CC) [16] methods push the boundaries in terms of medium- and heavy-mass nuclei, currently reaching ^{208}Pb [17]. All of these methods utilize physically motivated approximations to yield numerically exact solutions to the Schrödinger equation, for which I will use the term *simulation*. This is in contrast with so-called *emulated* solutions, which efficiently mimic simulator output at the cost of some accuracy. Emulators play a critical role for uncertainty quantification (UQ) in modern nuclear theory.

The challenge of solving the many-body Schrödinger equation notwithstanding, we need a model for the nuclear potential $\hat{V}$ and an estimate of its theoretical uncertainty. The latter aspect is the main focus of this thesis. Research on $\hat{V}$ has been underway for almost a century, initially through phenomenological models which over time have reached a very high level of refinement. Accurate phenomenological potentials such as CD-Bonn [18] are still in use, but unfortunately provide little information regarding any theoretical uncertainty in predictions. To bridge the gap between phenomenological models and costly lattice QCD calculations, we turn to effective field theories (EFTs), specifically chiral effective field theory (χEFT) [1, 19, 20]. χEFT can provide insight into the uncertainty of the theoretical predictions, while also remaining applicable in practice. Weinberg [21–24] proposed to adopt a set of relevant degrees of freedom—e.g., nucleons and pions—and construct a nuclear interaction potential from the most general effective Lagrangian $\mathcal{L}_{\text{eff}}$ that observes the symmetries of QCD, notably approximate chiral symmetry, and expanding $\mathcal{L}_{\text{eff}}$ in powers of $Q = p/\Lambda_b$. Here, p is a low-momentum scale and Λ_b is the breakdown scale of the theory. Λ_b is expected to be of the same order of magnitude as the nucleon mass, i.e., $\Lambda_b \approx 1$ GeV. The chiral expansion yields an infinite number of terms encoding different types of pion-nucleon (πN) and nucleon-contact interactions. For practical calculations, the expansion must be

truncated, and it is therefore important to employ a suitable power counting scheme wherein the diagrams are ordered according to the importance of their contributions. The power counting scheme assigns an order $k \geq 0$ to each interaction diagram: the dominant interactions comprise the leading order (LO), and higher-order corrections—suppressed by successively higher powers of Q—are sorted as next-to-leading order (NLO), next-to-next-to-leading order (NNLO), and so on. The systematic suppression of higher-order corrections yields a principal handle on the uncertainty induced by the finite truncation of the series, i.e., the EFT truncation error.

χEFT employs nucleons and pions as effective degrees of freedom in its so-called Δ-less variant [1], which we use in Papers I-III. In its Δ-full variant, used in Paper IV, the $\Delta(1232)$ resonance is also included as an explicit degree of freedom (see, e.g., Ref. [25]). χEFT is well suited for analyzing physics characterized by momenta comparable to the pion mass ($m_\pi \approx 140$ MeV) [20] and has been shown to provide a quantitative and accurate model of the strong nuclear interaction in many applications [6]. Yet open questions regarding appropriate power counting [26–29] remain to revolve issues grounded in renormalization. Here, I use standard Weinberg power counting as shown in Figs. 1.1 and 1.2. No interactions contribute at order $k = 1$ [30] and we therefore designate $k = 0, 2, 3, \ldots$ as LO, NLO, NNLO, and so on. As shown in Fig. 1.1, only 2NFs appear at the lowest orders with the well-known one-pion exchange (1PE) at LO, while Fig. 1.2 reveals that 3NFs enter at NNLO (NLO) in Δ-less (Δ-full) χEFT.

Each new order in the chiral expansion introduces a number of unknown parameters known as low-energy constants (LECs) that govern the strength of the corresponding interaction diagrams. The LECs are divided into categories based on the type of interaction, e.g., "contact LECs" for short-range nucleon-contact interactions and πN LECs for long-ranged π-exchanges. Power counting places an important expectation on the LECs in that they are supposed to be of "natural" size (i.e., of order 1) in appropriate units of the breakdown scale; a failure to fulfil this expectation is an indication that the EFT is not working as advertised. The number of LECs starts at just two at LO, but grows rapidly to 11 at NLO and above 30 at next-to-next-to-next-to-leading order (N3LO). In this thesis, I will denote a single LEC as α and all (in the context relevant) LECs as $\vec{\alpha}$.

The LECs generally have to be inferred from experimental data. They are consequently imbued with an uncertainty which, like the EFT truncation error, will propagate to observable predictions. Additional uncertainties can also appear, such as numerical errors from the solution of the Schrödinger equation. Quantifying these uncertainties and their impact on predictions is of fundamental importance for precision nuclear physics, and the motivation for this thesis.

In recent years, Bayesian inference methods have been developed as a tool for UQ in χEFT [31–36]. The Bayesian interpretation of probability enables us

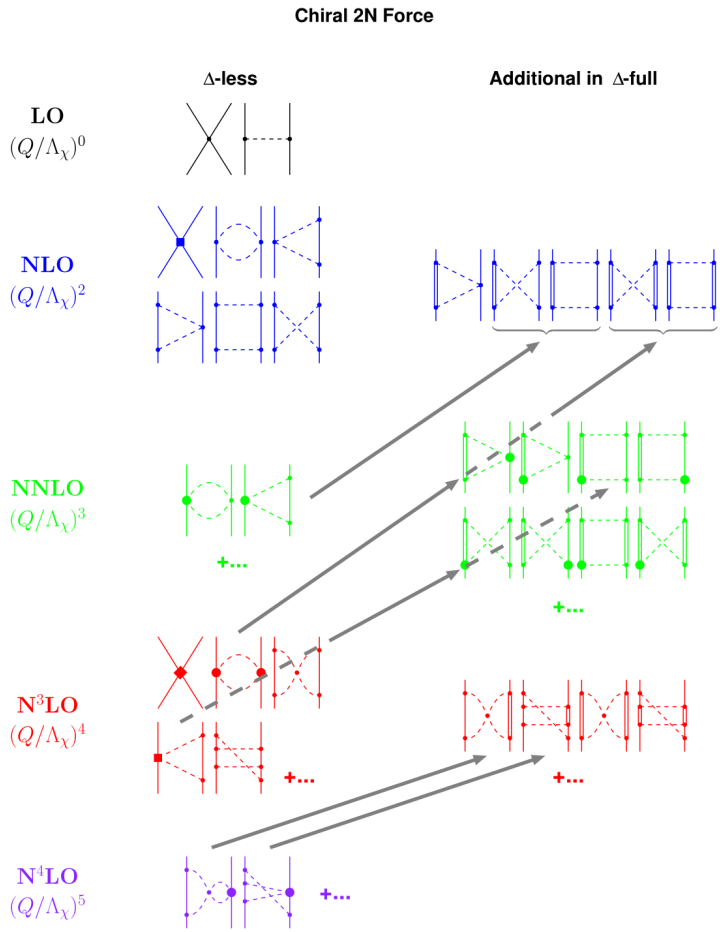


Figure 1.1: Hierarchy of two-nucleon forces (2NFs) in χEFT according to Weinberg power counting. Solid lines represent nucleons, dashed lines represent pions, and double lines represent Δ excitations. In this figure, Q denotes the low-momentum scale and Λ_χ the breakdown scale. Figure from Ref. [1]. Copyright © Elsevier 2011. Reproduced with permission.

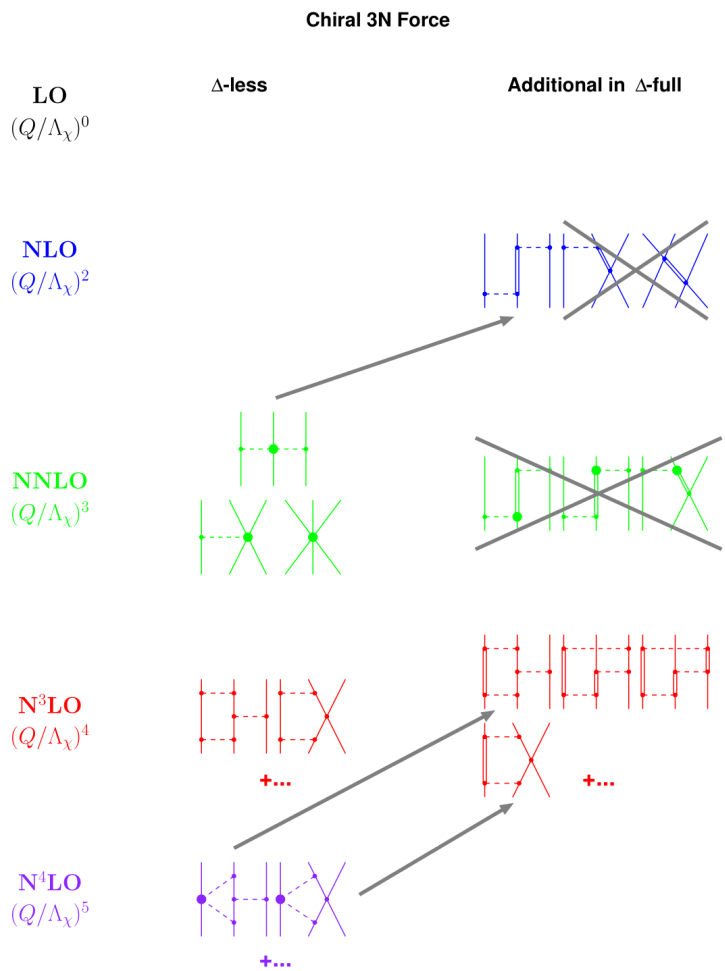


Figure 1.2: Like Fig. 1.1, but for 3NFs. Figure from Ref. [1]. Copyright © Elsevier 2011. Reproduced with permission.

to infer a probability density function (PDF), called a posterior, for the LECs at a given chiral order. This posterior can be conditional on a range of given information, such as the empirical data used for the parameter estimation and a statistical model for the EFT truncation error [32]. The Bayesian framework also allows—or rather, requires—us to encode a priori beliefs or assumptions in the so-called prior PDF. This gives us a straightforward way to encode, e.g., our expectation of natural LECs. In this thesis, the quantified uncertainties in the LEC posteriors are propagated to predictions of NN scattering observables to yield posterior predictive distributions (PPDs) with an added model discrepancy term [37] for the energy-dependent EFT truncation error.

The dimensionality of the posterior is at least equal to the number of LECs, which unavoidably means that Bayesian parameter estimation at NLO and beyond entails evaluating a high-dimensional PDF. This poses a significant computational challenge whose only realistic solution is the application of Markov chain Monte Carlo (MCMC) sampling. Unfortunately, the performance of standard MCMC algorithms degrades rapidly with increasing dimensionality and therefore require very long runtimes even at NLO. In this work, I therefore introduce an advanced MCMC sampling algorithm, Hamiltonian Monte Carlo (HMC) [38, 39], to Bayesian UQ in χEFT and show it to be adept at sampling high-dimensional parameter spaces. The use of HMC opens up the possibility of inferring joint posteriors for the LECs and other quantities of interest, such as the EFT expansion parameter Q.

Together with my collaborators, I extend the reach of UQ in nuclear theory to include Bayesian inference of 3NFs in Paper I through the use of NCSM emulators. In Paper II, we employ HMC to perform efficient inferences and predictions in the NN sector up to NNLO in Δ-less χEFT. In Paper III, we extend further to include N3LO and investigate the charge dependence of the strong nuclear interaction through PPDs of effective range expansion (ERE) parameters. In Paper IV, we combine HMC and NN scattering emulators [40, 41] with advanced modeling of correlated EFT truncation errors in Δ-full χEFT to investigate the effects on posteriors and predictions, paving the way for fully Bayesian analyses in the few-nucleon sector and establishing a stepping stone towards Bayesian *ab initio* theory for atomic nuclei.

Chapter 2

Simulating solutions to the Schrödinger equation

Solving the Schrödinger equation is a central challenge in nuclear theory. A list of commonly-used *ab initio* many-body methods to this end is presented in Chapter 1. The purpose of this chapter is to summarize the methods I use and to establish the corresponding notation for describing the NN system. Here, I only discuss simulations, i.e., numerically exact methods that yield high-fidelity results but are computationally costly in the context of Bayesian inferences. Approximate solutions—emulation—will be discussed in Chapter 4.

2.1 The non-relativistic Schrödinger equation in a partial-wave basis

The non-relativistic and time-independent Schrödinger equation for the NN system states that

$$\hat{H}\,|\psi\rangle = (\hat{H}_0 + \hat{V})\,|\psi\rangle = E\,|\psi\rangle \tag{2.1}$$

where E is the energy associated with the system in the state $|\psi\rangle$, $\hat{V}$ is the operator representing the interaction potential, and $\hat{H}_0$ is the kinetic energy operator for the relative motion of the two nucleons,

$$\hat{H}_0 = \frac{\mathbf{p}^2}{2\mu}. \tag{2.2}$$

Here, $\mathbf{p}$ is the relative momentum and μ is the reduced mass. In this work, the potentials $\hat{V}$ are represented in a plane-wave partial-wave basis. We use a basis that couples the nucleon spins $\mathbf{S} = \mathbf{s}_1 + \mathbf{s}_2$ and the relative orbital angular momentum $\mathbf{L}$ to the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$. I denote the

quantum numbers corresponding to **S**, **L**, and **J** as S, L, and J, respectively, while $p = \|\mathbf{p}\|$. Each state in the basis is further characterized by its isospin **T**, with associated quantum number T, and parity $\pi = (-1)^L$. I denote the projections of **J** and **T** on the quantization axis $\hat{\mathbf{z}}$, which I take to be the beam axis, as m and T_z, respectively. The Pauli exclusion principle dictates that only antisymmetric basis states are allowed; this is easily achieved in the NN case by only including channels where $(-1)^{L+S+T} = -1$ [42]. This factor arises as a result of applying the two-body antisymmetrization operator $\hat{A}_{12}$ to the NN states, where

$$\hat{A}_{12} = \frac{1}{2!}\left(\mathbb{I} - \hat{P}_{12}\right) \tag{2.3}$$

and $\hat{P}_{12}$ is the permutation operator that swaps particle indices 1 and 2.

Using this notation we can denote a state for two nucleons as $|LSJTT_z\pi; p\rangle$. From now on, I will leave T, T_z, and π implicit and denote an NN state as $|LSJ; p\rangle$. J and π—and hence S—are always conserved, and the interaction potential in a given partial wave (PW) can be written as

$$\langle L'SJ; p'|\, \hat{V}\, |LSJ; p\rangle = V^{SJ}_{L'L}(p', p) \tag{2.4}$$

where $\langle L'SJ; p'|$ and $|LSJ; p\rangle$ represent the final and initial states, respectively. L, however, is not always conserved as the strong nuclear force includes non-central (tensor) components, and PW channels where $L \neq L'$ are referred to as coupled channels. For NN systems, we have $|L - L'| = 0, 2$. Scattering experiments are generally defined in terms of the laboratory energy T_{lab}, which is related to the center-of-mass (c.m.) momentum p via

$$T_{\text{lab}} = \frac{\sqrt{(p^2 + m_t^2)(p^2 + m_p^2)} + p^2 - m_p m_t}{m_t} \tag{2.5}$$

where m_p (m_t) is the mass of the projectile (target) particle.

Quantum states are frequently described using spectroscopic notation shorthand ${}^{2S+1}L_J$, where $L = 0, 1, 2, 3, \ldots$ are denoted S, P, D, F, and so on in alphabetical order. For example, the singlet state $|LSJ\rangle = |000\rangle$ is denoted ${}^1S_0$. Using this notation, the lightest-mass bound state of nucleons—the deuteron—consists of a neutron and a proton bound in the coupled triplet state ${}^3S_1$-${}^3D_1$ with a ground state energy of $E({}^2\text{H}) = -2.224575(9)$ MeV [43] and no bound excited states. Computing $E({}^2\text{H})$ thus entails setting up a potential

$$V_{{}^3S_1\text{-}{}^3D_1} = \begin{pmatrix} V^{31}_{00} & V^{31}_{02} \\ V^{31}_{20} & V^{31}_{22} \end{pmatrix} \tag{2.6}$$

and solving Eq. (1.1).

In this work, the elements of $\hat{V}$ are derived from χEFT [1] where $\hat{V}$ typically depends linearly on the LECs. Thus, the potential contribution in each PW

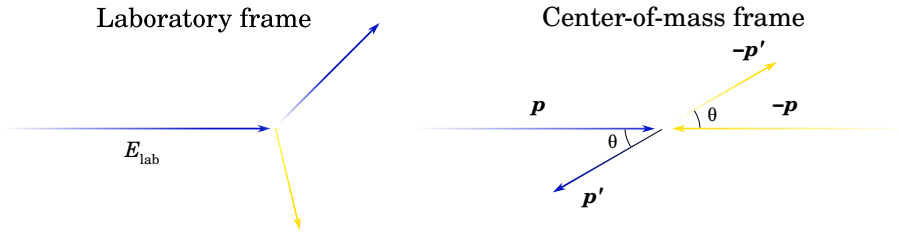


Figure 2.1: NN elastic scattering in laboratory and c.m. coordinates. The projectile and target nucleons are colored blue and yellow, respectively. T_{lab} is the laboratory energy, $\mathbf{p}$ ($\mathbf{p}'$) is the initial (final) c.m. momentum, and θ is the c.m. scattering angle. In the laboratory frame, the target nucleon is initially at rest.

channel consists of a constant part $\hat{V}_{\text{const}}$ and contributions that depend on the LECs,

$$\hat{V}(\vec{\alpha}) = \hat{V}_{\text{const}} + \sum_{i=1}^{N_{\text{LECs}}} f(\alpha_i)\hat{V}_i, \tag{2.7}$$

where N_{LECs} is the number of LECs and $f(\alpha_i) = \alpha_i$ up to NNLO in Weinberg power counting [1, 19, 20]. This property will be important in Chapter 4.

2.2 Two-nucleon elastic scattering

In order to solve the Schrödinger equation for NN scattering observables, we rewrite it in an alternate form, the LS equation. The LS equation is equivalent to the Schrödinger equation under imposition of boundary conditions stipulating that the eigenvalues vary continuously with p and the eigenstates are free waves in the absence of an interaction potential. Derivations of the LS equation can be found in any textbook discussing scattering, for example Ref. [44]. Here, I will follow the method outlined by Sakurai [45] which I find simple and instructive.

We start by letting $|\phi\rangle$ represent a plane-wave eigenstate of $\hat{H}_0$ with energy E_0:

$$\hat{H}_0 |\phi\rangle = E_0 |\phi\rangle . \tag{2.8}$$

We seek the solution to the Schrödinger equation,

$$(\hat{H}_0 + \hat{V}) |\psi\rangle = E_0 |\psi\rangle , \tag{2.9}$$

with the same eigenenergy E_0 as the free particles. Using Eq. (2.8) we can

rewrite Eq. (2.9) as

$$(E_0 - \hat{H}_0)\,|\psi\rangle = (E_0 - \hat{H}_0)\,|\phi\rangle + \hat{V}\,|\psi\rangle \tag{2.10}$$

and, applying $1/(E_0 - \hat{H}_0)$, we get

$$|\psi\rangle = |\phi\rangle + \frac{1}{E_0 - \hat{H}_0}\hat{V}\,|\psi\rangle\,, \tag{2.11}$$

where we have required that $|\psi\rangle \to |\phi\rangle$ as $\hat{V} \to 0$. This is the solution to Eq. (2.9), except that $1/(E_0 - \hat{H}_0)$ is singular. The singularity can be dealt with by introducing a small imaginary part $i\epsilon$:

$$\left|\psi^{\pm}\right\rangle = |\phi\rangle + \frac{1}{E_0 - \hat{H}_0 \pm i\epsilon}\hat{V}\left|\psi^{\pm}\right\rangle \tag{2.12}$$

where $|\psi^{+}\rangle$ ($|\psi^{-}\rangle$) represents an outgoing (incoming) wave. Eq. (2.12) is the LS equation. However, in order to arrive at a form that is useful in practice, we introduce the transition operator $\hat{T}$ according to

$$\hat{V}\left|\psi^{+}\right\rangle \equiv \hat{T}\,|\phi\rangle \tag{2.13}$$

and multiply Eq. (2.12) by $\hat{V}$ from the left to yield

$$\hat{T}\,|\phi\rangle = \hat{V}\,|\phi\rangle + \hat{V}\frac{1}{E_0 - \hat{H}_0 + i\epsilon}\hat{T}\,|\phi\rangle\,. \tag{2.14}$$

In operator form, this reads

$$\hat{T} = \hat{V} + \hat{V}\hat{G}_0\hat{T} \tag{2.15}$$

where we have introduced

$$\hat{G}_0 = \frac{1}{E - \hat{H}_0 + i\epsilon} \tag{2.16}$$

which is known as the free Green's function or, alternatively, the free resolvent. Formally solving for $\hat{T}$ yields

$$\hat{T} = \frac{\hat{V}}{1 - \hat{V}\hat{G}_0} \tag{2.17}$$

which in practice amounts to a matrix inversion. Numerically, the imaginary part of the free resolvent can be handled using the Cauchy principal value method [7]. $\hat{T}$ encodes all information about the scattering process and in knowing it we are in a position to compute scattering cross sections. In practice, nuclear potentials defined in momentum space need to be represented as

a matrix on a grid of finite momenta p, which is conveniently achieved using Gauss-Legendre quadrature. Obtaining converged solutions to the LS equation for NN scattering usually entails inverting a $\sim 100 \times 100$ matrix. In addition to the Gauss-Legendre grid points we typically want to add one point for the on-shell momentum we are interested in such that we get the corresponding $\hat{T}$-matrix element needed for computing observables. Solving the LS equation once is not challenging; however, we must repeat the calculation for every laboratory energy T_{lab} that we are interested in, and given that the world database of elastic scattering cross sections [46, 47] comprises around 6000 data points across hundreds of energies in the $0 < T_{\text{lab}} \leq 350$ MeV range, the computational cost adds up to be quite significant in the context of a Bayesian analysis; in Papers I-III, solving once for all relevant observables requires of the order of 10 s (wall time). Solving for $\hat{T}$ is the dominant cost, and the overall cost therefore scales roughly linearly with the number of unique energies.

2.2.1 Computing scattering observables from the $\hat{T}$ matrix

Once we have the on-shell $\hat{T}$-matrix element $T_{L'L}^{SJ}$, where

$$T_{L'L}^{SJ} = \langle L'SJ; p| \hat{T} |LSJ; p\rangle \tag{2.18}$$

we must perform a few steps before we arrive at cross sections [42]. The $\hat{S}$ matrix relates the incoming and outgoing states of a scattering process, and its on-shell element $S_{L'L}^{SJ}$ is related to $T_{L'L}^{SJ}$ via

$$S_{L'L}^{SJ} = 1 - 2\mu p \times 2iT_{L'L}^{SJ}. \tag{2.19}$$

Note that the factor $2\mu p$ depends on the normalization of the interaction potential $\hat{V}$. Once $S_{L'L}^{SJ}$ is known, we may compute the spin-scattering matrix $\mathbf{M}$ [48], which yields the scattered part of the final state. The partial-wave decomposition of $\mathbf{M}$ is given by [49]

$$\begin{aligned}\mathbf{M}_{m'm}^{S'S}(\theta,\phi) = \frac{\sqrt{4\pi}}{2ip} \sum_{J,L,L'}^{\infty} (-1)^{S-S'} i^{L-L'} \hat{J}^2 \hat{L} Y_{m-m'}^{L'}(\theta,\phi) \\ \times \begin{pmatrix} L' & S' & J \\ m-m' & m' & -m \end{pmatrix} \times \begin{pmatrix} L & S & J \\ 0 & m & -m \end{pmatrix} \times \langle L'SJ| \hat{S} - 1 |LSJ\rangle\end{aligned} \tag{2.20}$$

where $Y_{m-m'}^{L'}(\theta,\phi)$ is a spherical harmonic, $\hat{J} = \sqrt{2J+1}$, $\hat{L} = \sqrt{2L+1}$, and the parentheses are Wigner $3j$ symbols. θ defines the c.m. scattering angle; since we have chosen $\hat{\mathbf{z}}$ to be the beam axis and we have cylindrical symmetry, we have $(\theta,\phi) = (\theta, 0)$. For NN scattering, $\mathbf{M}_{m'm}^{S'S}$ is a 4×4 matrix to account

for singlet and triplet states. One must choose a convention which I define thus:

$$\begin{pmatrix} M^0_{00} & 0 & 0 & 0 \\ 0 & M^1_{-1-1} & M^1_{-10} & M^1_{-11} \\ 0 & M^1_{0-1} & M^1_{00} & M^1_{01} \\ 0 & M^1_{1-1} & M^1_{10} & M^1_{11} \end{pmatrix} \tag{2.21}$$

In practice, Eq. (2.20) must be truncated at some finite values of J, L, L'. In Papers I, II, and III, we truncate to $L_{\max} = 30$, while $J_{\max} = 30$ is used in Paper IV. These values are sufficient for converged results for $T_{\text{lab}} \leq 350$ MeV, where convergence is defined as the method error being negligible compared to other uncertainties.

All elastic scattering observables can be computed from $\mathbf{M}$. These observables can be broadly categorized as total cross sections σ_{tot} (denoted SGT in the SAID convention [50]), unpolarized differential cross sections $\sigma(\theta)$ (DSG), and various spin observables, such as spin polarization and correlation. Spin observables are normalized to the differential cross section and therefore take values between -1 and 1. The $\mathbf{M}$ matrix can be used to directly compute all of these observables. For example, the differential cross section is given by [51]

$$\frac{d\sigma}{d\Omega} = \frac{1}{4} \operatorname{Tr} \mathbf{M}\mathbf{M}^\dagger. \tag{2.22}$$

However, it is convenient to define the Saclay amplitudes a-e and use these to compute cross sections. They are given by

$$\begin{aligned} a &= \frac{1}{2}(M^1_{11} + M^1_{00} - M^1_{1-1}) \\ b &= \frac{1}{2}(M^1_{11} + M^0_{00} + M^1_{1-1}) \\ c &= \frac{1}{2}(M^1_{11} - M^0_{00} + M^1_{1-1}) \\ d &= \frac{1}{2\cos(\theta)}(-M^1_{11} + M^1_{00} + M^1_{1-1}) \\ e &= \frac{i}{\sqrt{2}}(M^1_{10} - M^1_{01}) \end{aligned}$$

and can be used to compute all elastic scattering cross sections; see Ref. [48] for a complete list. To give one example, the total cross section σ_{tot} is given by

$$\sigma_{\text{tot}} = \frac{2\pi}{p} \operatorname{Im}(a + b), \tag{2.23}$$

which is the optical theorem expressed in terms of Saclay amplitudes.

Computing the Saclay amplitudes is less computationally demanding than solving for $\hat{T}$. However, it is still worthwhile to avoid unnecessary recomputations by precomputing as many factors in (2.20) as possible. This will be discussed in Section 4.3.2.

Chapter 3

Bayesian inference and sampling

In recent years, the common approach [52–59] of representing the nuclear interaction and subsequent predictions of nuclear observables as a function of a single set (or a few sets) of optimized LEC values has started to shift [9, 17, 60, 61]. This shift has been precipitated by rapidly improving computational power, algorithmic improvements [6, 62], the introduction of Bayesian UQ to nuclear theory [31–36], and, even more recently, by the introduction of fast and accurate emulation of nuclear observables through eigenvector continuation (EC) emulation [63–65]. Still, comprehensive UQ including inference of posterior distributions for all relevant LECs has yet to be achieved. In this chapter I present my contributions towards this goal.

3.1 Linking experiment and theory

If we operate under the assumption that an observable y has a true value y_{true} [37], an experiment typically yields a value y_{exp} which is y_{true} plus an error δy_{exp}:

$$y_{\text{exp}} = y_{\text{true}} + \delta y_{\text{exp}}. \tag{3.1}$$

The error δy_{exp} typically follows a normal distribution [46, 47] with a standard deviation σ_{exp}, i.e, in general,

$$\text{pr}(\delta y_{\text{exp}}) = \mathcal{N}\left(0, \sigma^2_{\text{exp}}\right). \tag{3.2}$$

Theory, meanwhile, yields a $\vec{\alpha}$-dependent prediction $y_{\text{th}}(\vec{\alpha})$ with an error δy_{th}:

$$y_{\text{true}} = y_{\text{th}}(\vec{\alpha}) + \delta y_{\text{th}} \tag{3.3}$$

Thus experiment and theory are related via [35]

$$y_{\text{exp}} = y_{\text{th}}(\vec{\alpha}) + \delta y_{\text{th}} + \delta y_{\text{exp}}. \tag{3.4}$$

This relationship forms the basis of defining a likelihood function. If we assume that δy_{th} is independent of δy_{exp} and also follows a normal distribution with standard deviation σ_{th}, then

$$\text{pr}(y_{\text{exp}}|\vec{\alpha}) = \mathcal{N}\left(y_{\text{th}}(\vec{\alpha}), \sigma_{\text{th}}^2 + \sigma_{\text{exp}}^2\right) \tag{3.5}$$

which serves as a suitable likelihood function for a single datum. Alternatively, we can forgo the notion of a "true" value and directly link experiment and theory via Eq. (3.4).

3.2 Bayesian predictions

We wish to make predictions of an unknown observable y_{new} conditioned on existing data D and information I, where I encompasses all assumptions and choices such as the chiral order, computational methods in use, and our modeling of errors. As χEFT features inherent uncertainties, we write our prediction as a PDF $\text{pr}(y_{\text{new}}|D, I)$ using standard statistical notation, where $\text{pr}(A|B)$ indicates a PDF for A conditioned on B, and $\text{pr}(y_{\text{new}}|D, I)$ is a PPD for y_{new}. Our prediction depends on the LECs, i.e., $y_{\text{new}} = y_{\text{new}}(\vec{\alpha})$, which themselves are uncertain. To progress, we marginalize in $\vec{\alpha}$ [66]:

$$\text{pr}(y_{\text{new}}|D, I) = \int \text{pr}(y_{\text{new}}, \vec{\alpha}|D, I) d\vec{\alpha}. \tag{3.6}$$

By applying the product rule of probability calculus we get

$$\text{pr}(y_{\text{new}}|D, I) = \int \text{pr}(y_{\text{new}}|\vec{\alpha}, D, I)\, \text{pr}(\vec{\alpha}|D, I) d\vec{\alpha}. \tag{3.7}$$

Now, y_{new} does not depend on D since $\vec{\alpha}$ is given, a situation known as conditional independence. We can therefore simplify this to

$$\text{pr}(y_{\text{new}}|D, I) = \int \text{pr}(y_{\text{new}}|\vec{\alpha}, I)\, \text{pr}(\vec{\alpha}|D, I) d\vec{\alpha}. \tag{3.8}$$

Assuming that $\text{pr}(\vec{\alpha}|D, I)$—the posterior for the LECs—is known, we can evaluate Eq. (3.8), and we have our PPD for y_{new}. This is a univariate distribution. Note, however, that this PPD only takes into account uncertainties in the LECs $\vec{\alpha}$; it does not account for the energy-dependent truncation error in χEFT. We model the truncation error δy_{th} as a PDF $\text{pr}(\delta y_{\text{th}}|\vec{c}^2, Q, I)$ where $\vec{c}^2$ and Q characterize the convergence of an observable in χEFT (see below). Assuming we can draw samples from both $\text{pr}(y_{\text{new}}|D, I)$ and $\text{pr}(\delta y_{\text{th}}|\vec{c}^2, Q, I)$, we can predict y_{new} including a truncation error term.

The posterior for the LECs is not known, however; we must define it, and evaluate it numerically. This is a central challenge when one wishes to extract

PPDs of nuclear observables using χEFT. We can use Bayes' theorem to rewrite the posterior,

$$\mathrm{pr}(\vec{\alpha}|D, I) \propto \mathrm{pr}(D|\vec{\alpha}, I)\,\mathrm{pr}(\vec{\alpha}|I), \tag{3.9}$$

where the two factors on the right-hand side are the likelihood and the prior, respectively. The likelihood is common to both Bayesian and frequentist statistics, but the prior is unique to Bayes. As the name implies, the prior represents our knowledge of $\vec{\alpha}$ prior to seeing any data, and introduces a degree of subjectivity to a Bayesian analysis. This subjectivity can be partly mitigated by either choosing a weak prior, or by choosing a restrictive prior that is based upon solid information from previous studies or theoretical arguments [67]. The following sections are devoted to defining the likelihood and the prior. For notational simplicity, I now omit explicit conditioning on I.

3.3 What data?

Our Bayesian inferences are conditioned on data D, which we must choose. A natural candidate for LEC inference is NN elastic scattering cross sections, which are abundantly available thanks to a large number of experiments across many decades. Throughout this thesis I use the Granada database [46, 47] which comprises 6 384 neutron-proton (np) and proton-proton (pp) elastic scattering cross sections measured in the 1 eV to 350 MeV range of laboratory energies T_{lab}. During inferences, I consistently truncate the database at $T_{\mathrm{lab}} = 290$ MeV, i.e, below the pion-production threshold. The higher-energy data is used during model checking as validation data. The bulk of the database is comprised of integrated cross sections σ_{tot} (SGT in the SAID convention [50]), differential cross sections $\sigma(\theta)$ (DSG), and polarizations $A_y(\theta)$ (P). The remainder consists of various types of higher-order spin observables. For the exact details of the data used in each analysis I refer to the respective papers. It would be straightforward to include, e.g., deuteron properties as data in an inference, however the abundance of scattering observables makes this inclusion generally redundant. In principle, heavier nuclei can be included in the same way as the deuteron and could certainly provide information the deuteron cannot; however, every other nucleus presents (at the very minimum) a three-body problem that comes with additional computational challenges.

Conspicuously absent from the Granada database are neutron-neutron (nn) cross sections. The neutron's lack of electric charge makes nn experiments impractical since particle accelerators rely on electric charge for their operation. This is somewhat problematic for us, since χEFT features explicit isospin symmetry breaking at NLO and beyond [68], and we must include data that features nn interactions in order to constrain the corresponding LEC $\widetilde{C}^{nn}_{^1\mathrm{S}_0}$. Or, more accurately, we have plenty of *data* to infer $\widetilde{C}^{nn}_{^1\mathrm{S}_0}$; the issue is that nearly

all of it is $A \geq 3$ data that is non-trivial to make theoretical calculations of. One solution is to use ERE parameters, i.e, the nn ${}^1\mathrm{S}_0$ scattering length $a^{nn}_{{}^1\mathrm{S}_0}$ and/or effective range $r^{nn}_{{}^1\mathrm{S}_0}$, for which empirical values exist [1, 69]. Extracting empirical values for these ERE parameters is done by analyzing some reaction or observable that results in two neutrons with a low relative energy [69]; for example, the deuteron breakup reaction $nd \longrightarrow nnp$ can be used. In Paper II, I use the empirical scattering length $a^{nn}_{{}^1\mathrm{S}_0} = -18.9 \pm 0.4$ [69] fm to constrain $\widetilde{C}^{nn}_{{}^1\mathrm{S}_0}$.

I developed a new code for calculating scattering cross sections in Paper IV (see Chapter 4). This code does not yet take electromagnetic interactions into account, and pp cross sections are therefore excluded from D in that project, as well as np cross sections below 30 MeV since the magnetic moments of the neutron and proton interact and contribute with a small but non-negligible effect at low energies [70]. In analogy with $\widetilde{C}^{nn}_{{}^1\mathrm{S}_0}$, we can constrain $\widetilde{C}^{pp}_{{}^1\mathrm{S}_0}$ using ERE parameters where electromagnetic effects have been removed.

In the Δ-less theory, 3NFs enter at NNLO, and $A \geq 3$ and/or NNN scattering data is necessary if we wish to constrain the corresponding LECs, c_D and c_E. Data from, e.g., ^{3}H and ^{4}He—systems with more than one neutron—can in principle be used to constrain $\widetilde{C}^{nn}_{{}^1\mathrm{S}_0}$, but we also have two 3NF LECs c_D and c_E to infer, and we need at least one datum per LEC. NNN scattering data, such as nucleon-deuteron (Nd) scattering observables, is an excellent candidate for this purpose, which should be explored in the future. Recent PPDs of Nd cross sections using χEFT can be found in Refs. [8, 9], using only 2NFs, and the effects of 3NFs are explored in Refs. [10, 11]. In Paper I, however, bound-state properties of $A = 3, 4$ systems are used instead. In particular, we use the binding energies of ^{3}H and ^{4}He, the point-proton radius of ^{4}He, and the β-decay comparative half-life of ^{3}H. We also discuss the importance of including truncation errors for consistent results, see Chapter 5.

3.4 Likelihood

The traditional approach to fitting, or optimizing, the LECs in χEFT is to define a χ^2 cost function [71]

$$\chi^2 = \sum_{i=1}^{N} \left(\frac{y_{\mathrm{th},i}(\vec{\alpha}) - y_{\mathrm{exp},i}}{\sigma_{\mathrm{exp},i}} \right)^2 , \tag{3.10}$$

where $y_{\mathrm{exp},i}$ is the ith datum, $y_{\mathrm{th},i}(\vec{\alpha})$ is the corresponding computed observable value, $\sigma_{\mathrm{exp},i}$ is the standard deviation of the experimental error, and N is the number of data points. The LEC vector $\vec{\alpha}^*$ that minimizes this cost function is

$$\vec{\alpha}^* = \mathrm{argmin}(\chi^2) \tag{3.11}$$

and can be challenging to find. This approach is equivalent to maximizing [72] the χ^2 likelihood

$$\mathrm{pr}(D|\vec{\alpha}) \propto \exp\left(-\frac{1}{2}\chi^2\right), \tag{3.12}$$

and is an example of maximum likelihood estimation (MLE). MLE is a staple of frequentist statistics and has been used in several LEC inferences to extract optimized LEC values $\vec{\alpha}^*$ (see, e.g., Refs [18, 52, 57, 73–75]), on occasion along with local approximations of LEC covariance matrices in order to propagate LEC uncertainties to predictions [54, 59] (see also Paper I). However, a Bayesian approach delivers the PDF $\mathrm{pr}(\vec{\alpha}|D)$, which a frequentist approach does not.

The likelihood (3.12) with χ^2 defined by Eq. (3.10) assumes that the residuals are uncorrelated and follow a normal distribution. Furthermore, it assumes that the experimental error is the only source of uncertainty. We can introduce an independent theory error by replacing $\sigma_{\mathrm{exp},i}$ in Eq. (3.10) with σ_i, defined as [54, 71]

$$\sigma_i^2 = \sigma_{\mathrm{exp},i}^2 + \sigma_{\mathrm{th},i}^2 \tag{3.13}$$

where $\sigma_{\mathrm{th},i}$ is the standard deviation of an assigned theoretical (Gaussian-distributed) error. We can circumvent the restrictive assumption that the respective errors in σ_i^2 are internally uncorrelated by introducing a covariance matrix $\mathbf{\Sigma}$, where

$$\mathbf{\Sigma} = \mathbf{\Sigma}_{\mathrm{exp}} + \mathbf{\Sigma}_{\mathrm{th}}. \tag{3.14}$$

By introducing the residual vector

$$\vec{r} = \vec{y}_{\mathrm{th}} - \vec{y}_{\mathrm{exp}} \tag{3.15}$$

the corresponding likelihood can be written as [35]

$$\mathrm{pr}(D|\vec{\alpha}, \mathbf{\Sigma}_{\mathrm{exp}}, \mathbf{\Sigma}_{\mathrm{th}}) \propto \exp\left(-\frac{1}{2}\vec{r}^{\,T} \cdot \mathbf{\Sigma}^{-1} \cdot \vec{r}\right). \tag{3.16}$$

The likelihood is conditional on $\vec{\alpha}$ since $\vec{y}_{\mathrm{th}}$ depends on $\vec{\alpha}$.

3.4.1 Characterizing the theoretical uncertainty

How do we, then, quantify the theoretical error? A core strength of χEFT is that it gives us a principal handle on the theoretical error; the sub-leading contributions to an observable should at each order be suppressed by a factor Q (see Chapter 1), where we define Q as $Q = \max(m_\pi, p)/\Lambda_b$ or a smoothed version of this function. The asymptotic convergence of an observable in χEFT may thus be formalized as [32]

$$y_{\mathrm{th}} = y_{\mathrm{ref}} \sum_{n=0}^{\infty} c_n Q^n \tag{3.17}$$

where y_{ref} is a dimensionful reference value—for example an empirical value—and c_n are dimensionless expansion coefficients. In practice we must truncate this expansion at a finite chiral order k, resulting in

$$y_k = y_{\text{ref}} \sum_{n=0}^{k} c_n Q^n + y_{\text{ref}} \sum_{n=k+1}^{\infty} c_n Q^n \tag{3.18}$$

where the second term is the (unknown) truncation error $\delta y_{\text{th}}^{(k)}$. It must be noted that the ansatz in Eq. (3.17) could potentially be improved upon, inadequate for certain observables, or even outright erroneous; further research is necessary for validation. What is beyond doubt is that χEFT currently does come with a non-negligible truncation error. We may compute c_n up to order k as

$$c_n = \frac{y_{\text{th}}^{(n)} - y_{\text{th}}^{(n-1)}}{y_{\text{ref}} Q^n} \tag{3.19}$$

where $y_{\text{th}}^{(n)}$ is a theoretical prediction at order n, with obvious modifications for c_0. We assume that c_n is drawn from some unknown underlying distribution that is common to all orders. The known expansion coefficients $c_n, n = 0, \ldots, k$ give information about the unknown distribution for c, where we drop the subscript to indicate that we are not specifying the chiral order. We expect that the expansion coefficients are of natural size to conform with with our expectations about EFT convergence. Under the assumption that c is normally distributed with variance $\bar{c}^2$, i.e,

$$\text{pr}(c|\bar{c}^2) = \mathcal{N}(0, \bar{c}^2) \tag{3.20}$$

we find [35]

$$\text{pr}(\delta y_{\text{th}}^{(k)}|\bar{c}^2, Q) = \mathcal{N}(0, \sigma_{\text{th},k}^2) \tag{3.21}$$

where

$$\sigma_{\text{th},k}^2 = \bar{c}^2 y_{\text{ref}}^2 \frac{Q^{2(k+1)}}{1 - Q^2} \tag{3.22}$$

If we assume that the expansion coefficients are uncorrelated, we can build a (diagonal) theory covariance matrix according to

$$\boldsymbol{\Sigma}_{\text{th},ii} = \bar{c}^2 y_{\text{ref},i}^2 \frac{Q_i^{2(k+1)}}{1 - Q_i^2}. \tag{3.23}$$

We use this form for $\boldsymbol{\Sigma}_{\text{th}}$ in Papers I-III, with $\bar{c}$ determined as the root mean squared (RMS) value of a set of expansion coefficients calculated using Eq. (3.19) and a point-estimate LEC vector found using MLE with $\bar{c} = 1$. (In Paper I, we simply use $\bar{c} = 1$ throughout.) However, the expansion coefficients are undeniably correlated across $x = (T_{\text{lab}}, \theta)$ from visual inspections of calculated

coefficients, and we can generalize (3.23) by introducing a correlation matrix, or kernel function, $k(x, x')$ which encodes the correlations in the EFT truncation error between x and x'. This leads to [35, 36]

$$\boldsymbol{\Sigma}_{\mathrm{th},ij} = \bar{c}^2 k(x_i, x_j) y_{\mathrm{ref},i} y_{\mathrm{ref},j} \frac{Q_i^{k+1} Q_j^{k+1}}{1 - Q_i Q_j} \tag{3.24}$$

How, then, can we find a faithful representation of correlated EFT errors? Melendez *et al.* [36] introduced modeling of expansion coefficients as draws from a Gaussian process (GP) [76]. A GP is a stochastic process from which any finite number of random samples jointly follow a multivariate normal distribution. Introducing a mean function $m(x)$ and a covariance function $K(x, x') = \bar{c}^2 k(x, x')$, a GP $f(x)$ is written as

$$f(x) \sim \mathrm{GP}\left(m(x), K(x, x')\right). \tag{3.25}$$

In general, scattering observables are functions of both laboratory energy T_{lab} and the c.m. scattering angle θ, in which case $x \in \mathbb{R}^2$ and the GP is a function $\mathbb{R}^2 \to \mathbb{R}^1$. The exception to this is total cross sections, which are only functions of T_{lab}. Given a vector $\vec{x} \in \mathbb{R}^N$, where N is the number of inputs x, draws from the GP are by definition normally distributed, i.e.,

$$\vec{f} \mid \vec{x} \sim \mathcal{N}(m(\vec{x}), K(\vec{x}, \vec{x})), \tag{3.26}$$

where $m(\vec{x}) \in \mathbb{R}^N$ and $K(\vec{x}, \vec{x}) \in \mathbb{R}^{N \times N}$. The key quality of a GP is that, as long as we have determined m and K, we have a closed-form expression for f at any input x. In our case, the output of f is an expansion coefficient. We can compute a set of expansion coefficients c_{cal} for a given observable at a predetermined set of calibration points $\vec{x}_{\mathrm{cal}}$, calibrate the GP to c_{cal}, and learn the covariance matrix $\bar{c}^2 k(\vec{x}, \vec{x})$ at the inputs $\vec{x}$ defined by the data D. Inserting the result into Eq. (3.24) then yields the covariance matrix for the EFT truncation error since we assume a Gaussian prior.

We use the above approach in Paper IV to investigate the effect of correlated EFT truncation errors on the parameter inference. We calibrate a GP for each type of observable in the data set (i.e., disregarding possible inter-observable correlations). This leads to a block-diagonal structure for $\boldsymbol{\Sigma}$. NN elastic scattering observables typically exhibit smooth behavior as a function of x and it is therefore reasonable to model the truncation error as a smoothly varying function. A suitable and commonly used kernel in this situation is the radial basis function (RBF), or squared exponential, kernel which has the form

$$k(x, x') = \exp\left[-\frac{(x - x')^T \boldsymbol{l}^{-2} (x - x')}{2}\right], \tag{3.27}$$

where $\boldsymbol{l} = \text{diag}(l_{T\text{lab}}, l_{\theta\text{c.m.}})$ encodes the correlation lengths in T_{lab} and θ. For total cross sections, this simplifies to a single correlation length $l_{T\text{lab}}$. We assume that $m(\vec{x}) = \mu$ is a constant function.

We determine μ, $\bar{c}^2$, and $\boldsymbol{l}$ in a Bayesian way. We place a conjugate prior on the hyperparameters μ and $\bar{c}^2$, and thus have closed-form expressions for their joint posterior (see Paper IV for details). Unfortunately, no conjugate prior exists for $\boldsymbol{l}$ [36] and we therefore place a uniform prior on $l_{T\text{lab}}$ and $l_{\theta\text{c.m.}}$ according to

$$\text{pr}(l_{T\text{lab}}) = \mathcal{U}(\epsilon, 290) \tag{3.28}$$

$$\text{pr}(l_{\theta\text{c.m.}}) = \mathcal{U}(\epsilon, 180) \tag{3.29}$$

where $\epsilon = 10^{-5}$. One could perform a full Bayesian inference and MCMC sample the posterior for $(\boldsymbol{l}, \bar{c}^2, \mu)$ jointly with $\vec{\alpha}$ at the cost of a substantial increase in dimensionality. However, we are well-prepared to handle high-dimensional parameter spaces (see Section 3.6). For the present we content ourselves with employing maximum a posteriori (MAP) estimates for these quantities.

We find in Paper IV that our estimate of $\bar{c}^2$ is rather small, thus diminishing the effect of the exact modeling of the truncation error, yet that the effective number of data decreases when we account for correlated errors. See also Chapter 5 for further results and discussions.

3.5 Priors for the low-energy constants

An appropriate prior reflects our state of knowledge before incorporating any data into our analysis. Ideally, there should be no significant tension between the prior and the posterior, as this indicates conflicting prior beliefs. Such tensions could indicate inadequate error modeling, or perhaps better, hint at as-yet-undiscovered physics insights. I use a few different priors for the LECs in this thesis, and here I will summarize the central points. I refer to the respective papers for further details.

In Paper I, which focuses on the 3NF LECs c_D and c_E, we use MLE with a local approximation of the parameter covariance matrix to find a prior for the 2NF LECs. This prior is an example of an informative prior which is strongly constraining, standing in for an expensive-to-evaluate likelihood. The benefit of this approach is thus that we bypass the need for the relatively expensive calculations of NN observables during the sampling of the full posterior, which also includes the 3NF LECs. The computational load incurred by the sampling is thus greatly reduced. As will be discussed in Chapter 4, the advent of scattering emulators means that the necessity of such shortcuts has recently faded considerably [41].

In addition to c_D and c_E themselves, we also infer posteriors for $\bar{c}^2$ and Q, which characterize the truncation error in the NNN sector. We place a weakly

informative prior on Q in the form of a beta distribution $B(Q|a = 3, b = 5)$ which encodes that $0 < Q < 1$ and that we expect that $Q < 0.5$. For $\bar{c}^2$ we adopt a conjugate prior, i.e., a prior that after exposure to data yields an analytic posterior of the same form. Specifically, we use a scaled inverse chi-squared distribution that reflects our belief that $\bar{c}$ is of order 1.

We use the same priors throughout Papers II-IV. The prior for the contact LECs is an uncorrelated Gaussian of width 5 in the appropriate units (see, e.g., Paper II):

$$\mathrm{pr}(\vec{\alpha}) \propto \exp\left(-\frac{1}{2}\vec{\alpha}^2/5^2\right) \tag{3.30}$$

This weakly informative prior encodes our expectation that the LECs are of order 1 by mildly penalizing large values. We have detected little reason to change this prior over the course of this work since nearly all LECs stay well within its confines. The only exceptions are a small number of fourth-order contacts at N3LO, e.g., $D_{^1S_0}$ and $D_{^3S_1}$ which take on somewhat unnatural values; here, the prior helps prevent overfitting to high-energy data. For the πN LECs entering at NNLO and N3LO we adopt an informative prior based on a so-called Roy-Steiner analysis of πN scattering data [77, 78]. In principle, we could use a weak prior for these LECs and include πN scattering data in the likelihood, but this prior is more precise (at least in the Δ-less case) and computationally less demanding. Our Bayesian framework straightforwardly allows for such flexibility.

3.6 Sampling the posteriors with Hamiltonian Monte Carlo

With the exception of LO, the LEC posteriors defined above are high-dimensional functions of $N_{\mathrm{dim}} = 10$-31 parameters. Evaluating such functions on a grid is intractable since the number of grid points grows exponentially. Table 3.1 demonstrates this so-called curse of dimensionality for a hypothetical posterior where each evaluation takes 1 s to complete (i.e, about the actual evaluation times in this work, which span roughly 0.5 s to 15 s). The problem is exacerbated by our a priori lack of knowledge about the posterior: where are we going to place the grid points? We want to draw samples in regions of high probability mass. The vast majority of the high-dimensional parameter space contributes negligibly to any expectation value $\langle f(\vec{\alpha})\rangle$ of a function f of the LECs. We need a sampling method that simultaneously minimizes the number of evaluations needed *and* identifies the region(s) of probable LEC values. Our best option is to use MCMC sampling to address this problem.

Table 3.1: Time needed to evaluate a hypothetical multidimensional posterior on a grid. Each evaluation takes 1 s, and we take 10 samples per direction.

N_{dim}	Samples/N_{dim}	Samples	Time
1	10	10	10 s
2	10	100	1 min 40 s
3	10	1000	17 min
4	10	10000	2.7 h
10	10	10^{10}	317 years
15	10	10^{15}	32 million years
18	10	10^{18}	Twice the age of the universe

3.6.1 Genesis of MCMC: the Metropolis algorithm

MCMC saw the light of day in 1953 in the shape of the Metropolis algorithm [79], named after its inventor[1]. Following a subsequent improvement by Hastings [80] the algorithm has become known as Metropolis-Hastings.

Metropolis-Hastings and all other MCMC algorithms (see, e.g., Ref. [81] for a detailed overview) employ Markov chains to sample a PDF $\text{pr}(\vec{\alpha})$. A Markov chain is a stochastic process where the probability distribution for each sample only depends on the previous sample, so-called memorylessness. Over time, the distribution of samples will converge to a unique stationary distribution as long as the Markov chain will eventually visit every possible sample (so-called ergodicity). Metropolis-Hastings, and all other MCMC algorithms, construct Markov chains in such a way that the stationary distribution is the sought PDF $\text{pr}(\vec{\alpha})$.

Given a current sample $\vec{\alpha}$, the algorithm proceeds by proposing a new sample $\vec{\alpha}^{\,\prime}$ from a proposal distribution q. q is typically (but not necessarily) a normal distribution centered on $\vec{\alpha}$. Then, the Hastings ratio

$$r = \frac{\text{pr}(\vec{\alpha}^{\,\prime})q(\vec{\alpha}|\vec{\alpha}^{\,\prime})}{\text{pr}(\vec{\alpha})q(\vec{\alpha}^{\,\prime}|\vec{\alpha})} \tag{3.31}$$

is computed followed by the acceptance probability

$$a = \min(1, r). \tag{3.32}$$

The new sample is accepted with probability a, a criterion we may call the Metropolis criterion. If the sample is rejected, a new copy of $\vec{\alpha}$ is appended to the chain. In finite (but possibly very long) time, the Markov chain will converge to provide samples from $\text{pr}(\vec{\alpha})$.

[1] One of Metropolis's co-authors, Arianna Rosenbluth, passed away in 2020. She is credited with the first implementation, but was largely unaware of the impact the algorithm has had. `https://www.nytimes.com/2021/02/09/science/arianna-wright-dead.html`

The Metropolis-Hastings algorithm works quite well for many problems, and it certainly outperforms evaluating $\mathrm{pr}(\vec{\alpha})$ on a grid by a huge margin. However, for the problems relevant in this thesis it is not sufficient. The reason is that the samples are autocorrelated, i.e., subsequent samples are not independent of each other. As a method based on random walks in some high-dimensional space, each step must be rather small, or else nearly every proposed sample would be rejected. The small steps induce autocorrelation. If we assume that the samples are not autocorrelated, but rather completely uncorrelated, the central limit theorem states that the sampling variance of a parameter expectation value $\langle\alpha\rangle$ is given by $\mathrm{Var}\left[\langle\alpha\rangle\right] = \mathrm{Var}\left[\alpha\right]/N$, where N is the number of samples. With autocorrelated samples, however, the sampling variance is

$$\mathrm{Var}\left[\langle\alpha\rangle\right] = \tau \frac{\mathrm{Var}\left[\alpha\right]}{N} \tag{3.33}$$

where

$$\tau = 1 + 2 \lim_{N\to\infty} \sum_{h=1}^{N} \rho(h). \tag{3.34}$$

τ is known as the integrated autocorrelation time, and depends on the autocorrelation function $\rho(h)$, where h is the separation of the samples; i.e., $\rho(h)$ measures the autocorrelation between two samples that are h steps apart. With correlated samples, $\rho(h)$ is greater than 0 and thus the sampling variance $\mathrm{Var}\left[\langle\alpha\rangle\right]$ is greater than it would be without autocorrelations. This motivates us to define the effective sample size (ESS) as

$$\mathrm{ESS} = N/\tau. \tag{3.35}$$

To keep ESS high, we need to collect many samples and keep τ low. Unfortunately, τ tends to increase with N_{dim} as the sampler needs to take very short steps in high-dimensional spaces, or risk wandering into uninteresting regions. To compensate, we must increase N, which is costly. MCMC samplers are thus not immune to the curse of dimensionality.

A modern, and very popular, variation on the Metropolis-Hastings algorithm is provided by EMCEE [82], a so-called affine-invariant ensemble sampler [83]. The "ensemble" refers to a collection of simultaneous walkers at different locations in the parameter space, each of which provides their own chain of samples to the collective whole. Metropolis-Hastings does not employ this concept and might thus be called a single-walker algorithm by analogy. EMCEE is easy to use and provides overall superior performance to Metropolis-Hastings, but is still affected by strong autocorrelation in higher dimensions. There is, however, one algorithm that excels at suppressing τ in high-dimensional spaces.

3.6.2 Hamiltonian Monte Carlo

HMC [38], originally known as Hybrid Monte Carlo, was conceived by lattice QCD physicists to increase sampling efficiency in high-dimensional spaces by combating autocorrelated samples. The key idea of HMC is to consider an MCMC walker as a "ball" gliding through the parameter space by simulating Hamiltonian dynamics, rather than to do a random walk. I can conclude that this works very well in our application, with the caveat that it requires an experienced user in order to get the most out of the algorithm. It is my intention and hope that this section can help you, the reader, use HMC effectively for your sampling needs.

Throughout this thesis I use a custom implementation of HMC, which is publicly available[2]. Its implementation and use is extensively documented in Papers II and III, and I refer to those papers for further details. My implementation, and the mathematical details presented in this thesis, is mainly based on Ref. [39]. I use what might be referred to as vanilla, or standard, HMC; more advanced variants also exist, such as the No-U-Turn sampler [84], which aims to simplify the process of tuning HMC, and Riemannian Manifold HMC [85], which aims to address shortcomings related to sampling complicated functions. A complicated function might be, for instance, one that exhibits strong curvature, or bimodality where the two modes have different shapes.

HMC proceeds by reimagining the posterior—which, for simplicity, I will here denote $\mathrm{pr}(\vec{\alpha})$—as a potential energy function $U(\vec{\alpha})$. Explicitly, we have $U(\vec{\alpha}) = -\ln \mathrm{pr}(\vec{\alpha})$, for reasons that will become clear shortly. In addition, we introduce an auxiliary momentum $\vec{p}$ of the same dimensionality as $\vec{\alpha}$ and a corresponding kinetic energy function $K(\vec{p})$. The titular Hamiltonian is thus

$$H(\vec{\alpha}, \vec{p}) = K(\vec{p}) + U(\vec{\alpha}) = K(\vec{p}) - \ln \mathrm{pr}(\vec{\alpha}). \tag{3.36}$$

Please note that this Hamiltonian is unrelated to the nuclear interaction Hamiltonians discussed in other chapters! We can link $H(\vec{\alpha}, \vec{p})$ to a probability distribution using the canonical Boltzmann distribution from statistical mechanics,

$$\mathrm{pr}(\vec{\alpha}, \vec{p}) = \frac{1}{Z} \exp\left(\frac{-H(\vec{\alpha}, \vec{p})}{T}\right) = \frac{1}{Z} \exp\left(\frac{-U(\vec{\alpha})}{T}\right) \exp\left(\frac{-K(\vec{p})}{T}\right), \tag{3.37}$$

where Z is a normalization constant and T is the "temperature" of the system, here serving to render the exponent dimensionless. $\mathrm{pr}(\vec{\alpha}, \vec{p})$ will be subjected to the Metropolis criterion, where Z will cancel out, and we can thus set $Z = 1$. We see from Eq. (3.37) that the joint PDF $\mathrm{pr}(\vec{\alpha}, \vec{p})$ is the product of two distributions in $\vec{\alpha}$ and $\vec{p}$ separately, and if we set $T = 1$ and marginalize over $\vec{p}$ we reacquire $\mathrm{pr}(\vec{\alpha})$.

[2] `https://github.com/svisak/montepython.git`

The kinetic energy function, $K(\vec{p})$, is an implementation design choice. We employ the commonly used form

$$K(\vec{p}) = \frac{1}{2}\vec{p}^{\,T}\mathcal{M}^{-1}\vec{p}, \tag{3.38}$$

where $\mathcal{M}$ is a user-chosen positive-definite symmetric matrix known as the mass matrix. This form yields a Gaussian PDF for $\vec{p}$. The expression for the joint PDF $\mathrm{pr}(\vec{\alpha}, \vec{p})$ becomes

$$\mathrm{pr}(\vec{\alpha}, \vec{p}) = \exp(-H(\vec{\alpha}, \vec{p})) = \mathrm{pr}(\vec{\alpha}) \exp\left(-\frac{1}{2}\vec{p}^{\,T}\mathcal{M}^{-1}\vec{p}\right). \tag{3.39}$$

So, now we have the existing sample $\vec{\alpha}$, we have reimagined the (negative log) posterior as a potential energy surface, we have introduced a momentum $\vec{p}$, and we have a link between $H(\vec{\alpha}, \vec{p})$ and $\mathrm{pr}(\vec{\alpha}, \vec{p})$. How do we propose a new sample $(\vec{\alpha}, \vec{p})$? Naturally, we simulate Hamiltonian dynamics for a finite period of time, i.e., we solve Hamilton's equations

$$\frac{d\alpha_i}{dt} = \frac{\partial H}{\partial p_i} \tag{3.40}$$

$$\frac{dp_i}{dt} = -\frac{\partial H}{\partial \alpha_i} \tag{3.41}$$

where i indexes the dimension. A ball governed by Hamiltonian dynamics will follow a path of constant total energy, which in this case also means a path of constant probability $\mathrm{pr}(\vec{\alpha}, \vec{p})$. If we, for the moment, ignore any numerical errors in the simulation, this means that $\mathrm{pr}(\vec{\alpha}, \vec{p})$ subjected to the Metropolis criterion will always be accepted. Yet, even though the *total* energy is conserved, the potential and kinetic energies will have changed. And since we marginalize over $\vec{p}$ (by simply discarding it), we are left with a guaranteed-to-be-accepted sample $\vec{\alpha}'$ that can be located far from the previous sample. This is the brilliance of HMC. Once the proposed sample has been accepted, the momentum $\vec{p}$ is resampled and the process begins again.

Our choices for K and U mean that we can rewrite Hamilton's equations as

$$\frac{d\alpha_i}{dt} = \frac{\partial K}{\partial p_i} = (\mathcal{M}^{-1}p)_i \tag{3.42}$$

$$\frac{dp_i}{dt} = -\frac{\partial U}{\partial \alpha_i}. \tag{3.43}$$

The important thing to note here is that $\partial U/\partial \alpha_i$ is a partial derivative of the negative log posterior with respect to the ith LEC; we must therefore have access to the gradient of the negative log posterior in order to use HMC. In practice, this is achieved using automatic differentiation (AD) [86, 87] which will be discussed further in Chapter 4.

Unfortunately we cannot ignore numerical errors in the simulation, and in practice the Metropolis criterion will sometimes reject proposed samples due to unconserved total energy. To mitigate this problem, it is necessary to (i) simulate Hamiltonian dynamics using a short time step, and (ii) use a numerical integration method where the numerical error does not accumulate. The standard method in HMC is known as leapfrog integration, which is a so-called symplectic integrator. Symplectic integration means that the local discretization error is equally likely to be positive or negative in each step and that the integration preserves the time-reversibility and volume-preservation of Hamiltonian dynamics. This is crucial for MCMC as time-reversibility and volume-preservation are prerequisites for upholding detailed balance, a sufficient requirement to guarantee that the MCMC algorithm converges to the correct stationary distribution [67].

HMC features three hyperparameters that need to be tuned by the user. These are the integration timestep ϵ, the number of timesteps L, and the mass matrix $\mathcal{M}$. Detailed discussions of how to choose these parameters can be found in Papers II and III. Here, I will only summarize my methods for picking $\mathcal{M}$. Simply put, a performant $\mathcal{M}$ needs to be a decent approximation of the inverse of the LEC covariance matrix [88]. I have tested manual methods to construct $\mathcal{M}$, e.g., by using an identity matrix or by exploiting previously published LEC uncertainties. I have concluded that these methods are unusable, and that more sophisticated methods are called for.

In Paper II, I found success in extracting $\mathcal{M}$ by using a short preliminary MCMC sampling of the posterior, approximating the parameter covariance from the samples, and inverting the result. This method yielded integrated autocorrelation times τ of order 1, i.e., virtually uncorrelated samples. In fact, at LO and NLO, we find $\tau < 1$, i.e., anti-correlated samples, a condition known as antithetic sampling [89]. With $\tau < 1$, the ESS becomes *greater* than the number of actual samples, indicating extremely efficient sampling. However, this tuning method is quite cumbersome and somewhat unreliable. It also proved inadequate for successfully sampling the extremely challenging N3LO posterior. Fig. 3.1 shows autocorrelation functions for the HMC samplings performed in Paper II. For comparison, corresponding results achieved with EMCEE are also shown. It is clear that EMCEE yields much higher autocorrelations. At the same time, it must be noted that simulating Hamiltonian dynamics is expensive and accounts for the vast majority of the computational cost in HMC. Thus, HMC can only be worthwile if the decrease in autocorrelations ultimately compensates for the increased per-sample cost. We show in Paper II that this is indeed the case in our application to χEFT.

In Paper III I devised a new method inspired by MLE with local covariance approximation, as discussed in Section 3.4. By replacing the preliminary sampling with an optimization using the BFGS algorithm [91–94] I was able to extract a local approximation of the parameter covariance matrix which I used to construct $\mathcal{M}$ efficiently. Using this method, anticorrelations between samples

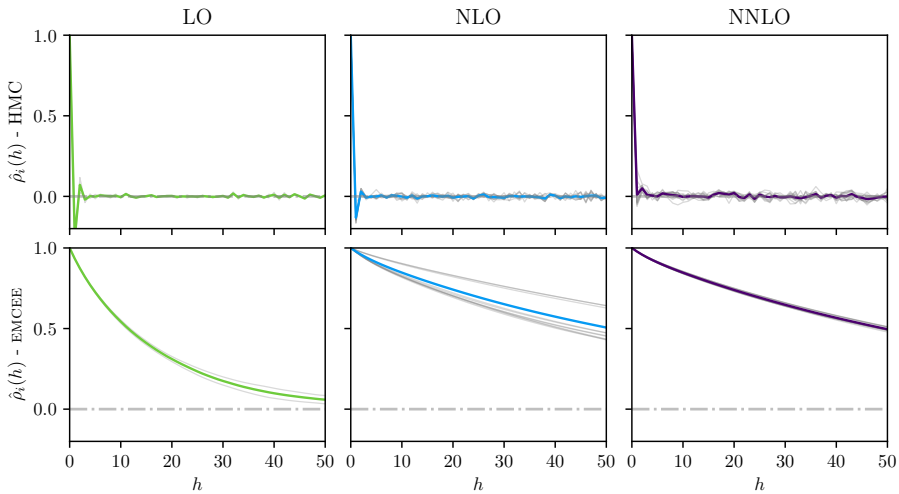


Figure 3.1: Autocorrelation functions at LO, NLO, and NNLO as achieved with HMC (top) and EMCEE (bottom). Gray lines represent individual LECs, and colored lines represent the average. The HMC algorithm was tuned using the "preliminary sampling" method (see text). Originally published in Ref. [90] (Paper II). Copyright © 2022 by the American Physical Society. Reproduced with permission.

increased somewhat, thus increasing performance. More importantly, we were able to extract a converged N3LO posterior with $\tau \approx 3$.

The integrated autocorrelation time is useful not only to judge the efficiency of the sampling, but also to assess its convergence [95]. Assessing convergence is necessary in any MCMC application, or the results are meaningless. Unless the sampled PDF is analytic or well understood (rendering MCMC redundant), it is not possible to declare an MCMC chain for a continuous PDF "converged" since the sampler cannot explore the entire parameter space in finite time and the unexplored regions may contain non-negligible probability mass. Convergence testing should thus be understood as trying to detect signs of non-convergence. Multimodal distributions in particular are prime sources of what we may call false convergence; an MCMC chain could easily have sampled one mode to "convergence" while completely missing other modes. There are sampling methods, such as nested sampling [96] and parallel tempering [97, 98], which are designed to handle multimodality, but the problem remains a difficult one in practice.

Convergence testing using τ is based on the evolution of τ estimates as the length N of the MCMC chain increases. If N is low, then τ will be underestimated; it is thus very dangerous to trust estimates of τ if $\tau \ll N$ does not hold!

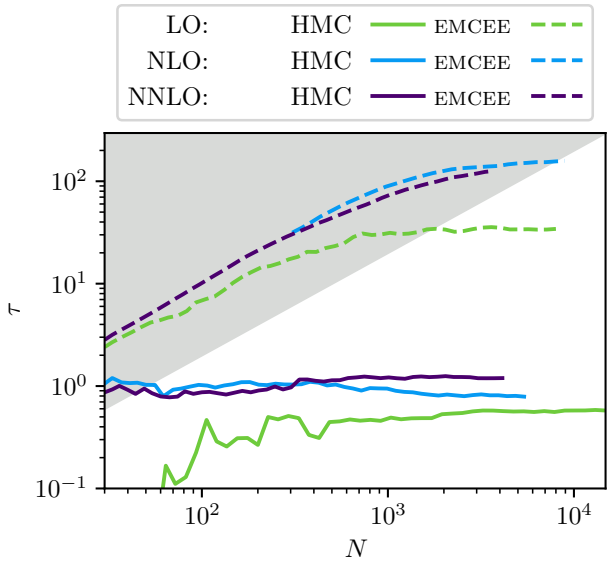


Figure 3.2: Convergence of the estimate of τ with respect to the length N of the chain. The results for EMCEE are averaged over all walkers, hence the seemingly short chains; the real number of samples can be found by multiplying the indicated number by the number of walkers (10 at LO, 50 at NLO, and 65 NNLO). The grayed area indicates the non-convergence zone, i.e., $N < 50\tau$. Originally published in Ref. [90] (Paper II). Copyright © 2022 by the American Physical Society. Reproduced with permission.

For every MCMC sampling one should thus plot the τ estimate as a function of N, and verify that τ has stabilized. Only if (i) τ has stabilized, and (ii) $N > n \cdot \tau$ for some constant $n \gg 1$ can we claim that the chain is converged. The constant n is an arbitrary cutoff, for which different suggestions exist. The EMCEE authors recommend $n = 50$, which we also apply here. Fig. 3.2 shows the evolution of τ for the samplings performed in Paper II, i.e., for samplings at LO-NNLO using both HMC and EMCEE. It is clear that the τ estimates for the HMC chains stabilize very rapidly, and cross the convergence threshold of $N > 50\tau$ within about 50 samples. We could have imposed a much stricter convergence criterion, say $n = 1\,000$, and still declared convergence. For EMCEE, the τ estimates continue to rise until about 10 000 samples at LO and roughly 500 000 samples at NLO (the estimates for EMCEE are averaged over 10 walkers at LO and 50 walkers at NLO, respectively). At NNLO, EMCEE does not reach convergence.

Many methods for assessing MCMC convergence have been devised (see Ref [99] for a recent review). Beyond the τ convergence criterion discussed above, I have also employed the standard Gelman-Rubin diagnostic $\hat{R}$ [100, 101], which relies on multiple identically-prepared chains run in parallel. The essence of $\hat{R}$ is to compare the variance within chains to the variance between chains, and from this information compute the potential variance reduction longer chains would offer. $\hat{R}$ is discussed in more detail in Paper II; here, it suffices to say that I find the τ convergence criterion to be stricter and thus more useful.

3.7 Sampling/importance resampling

The flexibility of the Bayesian framework is demonstrated in Paper III, where we infer all relevant LECs except one ($\widetilde{C}^{nn}_{^1\mathrm{S}_0}$) using HMC, and then extend the posteriors to include $\widetilde{C}^{nn}_{^1\mathrm{S}_0}$ using a sampling technique known as sampling/importance resampling (SIR) [61, 102, 103]. Here, I will outline the technique in general terms with our application to $\widetilde{C}^{nn}_{^1\mathrm{S}_0}$ inference in mind. Particular details can be found in Paper III.

SIR is a straightforward and intuitive method that is useful in low-dimensional problems, such as extending an already-resolved high-dimensional parameter inference problem by including an additional parameter. We wish to draw samples from a non-trivial PDF $h(\theta)$ defined in terms of some unnormalized distribution $f(\vec{\theta})$, i.e.,

$$h(\vec{\theta}) = \frac{f(\vec{\theta})}{\int f(\vec{\theta}) d\vec{\theta}}. \tag{3.44}$$

In our case, $f(\vec{\theta})$ is the unnormalized joint posterior for $\widetilde{C}^{nn}_{^1\mathrm{S}_0}$ and all other LECs. To proceed, we define a *sampling distribution* $g(\vec{\theta})$ which is trivial to sample from; for example, $g(\vec{\theta})$ might be a uniform or Gaussian distribution. From $g(\vec{\theta})$ we draw N samples of $\vec{\theta}$, and for each sample $\vec{\theta}_i$ we calculate weights

$$w_i = \frac{f(\vec{\theta}_i)}{g(\vec{\theta}_i)} \tag{3.45}$$

from which we define the normalized weights q_i according to

$$q_i = \frac{w_i}{\sum_{j=1}^{N} w_j}. \tag{3.46}$$

The samples $\vec{\theta}_i$, weighted by q_i, are then distributed according to $h(\vec{\theta})$ [102].

The main difficulty in using SIR is choosing an appropriate sampling distribution $g(\vec{\theta})$ which is trivial to sample from and that preferably resembles $f(\vec{\theta})$.

We cannot escape that $\dim(g) = \dim(h)$, and in our simplest case—NLO—we then have $\dim(g) = 11$. However, through HMC we already have an excellent representation of the posterior for all LECs except $\widetilde{C}^{nn}_{^1\mathrm{S}_0}$, in the form of a Markov chain with virtually uncorrelated samples. These samples are indeed trivial to sample from once they have been acquired. The only remaining question is the marginal sampling distribution for $\widetilde{C}^{nn}_{^1\mathrm{S}_0}$, and here we are aided by both χEFT and nuclear physics in a broader sense. χEFT tells us that $\widetilde{C}^{nn}_{^1\mathrm{S}_0}$ is expected to be of natural size, and we also know that the charge dependence of the strong nuclear force is a small effect ($\approx 1\%$ [1]). We can thus inspect the marginal posteriors for $\widetilde{C}^{np}_{^1\mathrm{S}_0}$ and $\widetilde{C}^{pp}_{^1\mathrm{S}_0}$, and define a marginal sampling distribution for $\widetilde{C}^{nn}_{^1\mathrm{S}_0}$ that covers roughly the same values, with some margin for safety.

SIR works very well in our application and we were able to painlessly amend our HMC-sampled posteriors with $\widetilde{C}^{nn}_{^1\mathrm{S}_0}$. The resulting joint posterior picks up parameter correlations between $\widetilde{C}^{nn}_{^1\mathrm{S}_0}$ and other parameters, most notably positive correlations with $\widetilde{C}^{np}_{^1\mathrm{S}_0}$ and $\widetilde{C}^{pp}_{^1\mathrm{S}_0}$ as expected for a small charge dependence. The results will be further discussed in Chapter 5.

Chapter 4

Emulating solutions to the Schrödinger equation

The computational cost of simulating solutions to the Schrödinger equation with numerically exact methods is a major hurdle for Bayesian inferences. It is therefore understandable that LEC inference has thus far generally been limited to MLE through optimization [52, 54, 59, 74, 104, 105], a challenging enough problem in itself. As I demonstrate in Papers II and III, it is now possible—if costly, and with the aid of modern hardware and highly optimized software—to sample joint Bayesian posteriors for the LECs up to fourth order in χEFT conditioned on NN data. Conditioning such an inference on data beyond the NN sector is, unfortunately, costly in the extreme. We must find ways to compute observables faster.

Emulators [106] provide the required leverage. Emulation, as opposed to simulation, yields approximate solutions to (in this case) the Schrödinger equation at a fraction of the computational cost. GP emulation [76] is one of the best-known methods for emulation, but its accuracy is limited. For an emulator to be useful, its error with respect to the corresponding simulation must be sufficiently small relative to other uncertainties. In 2018, Frame *et al.* introduced EC to the nuclear theory community as a method of finding extremal eigenvalues and eigenvectors of matrices too large to deal with using standard linear algebra methods [63]. König *et al.* subsequently built upon EC to construct highly efficient and accurate emulators for bound-state properties of nuclei [64]. EC emulation has since garnered considerable interest in the nuclear theory community and has been applied to NN [40, 41] and NNN scattering [107], many-body systems [65], among others [60, 108]. As it turns out, EC can be understood as an existing model order reduction (MOR) method known as the reduced basis (RB) method [109].

4.1 Model order reduction

A large matrix (here generically denoted $\mathbf{A}$) containing all relevant information about a problem is commonplace in nuclear theory. Frequently, however, much of the information contained in $\mathbf{A}$ is of little interest. The aim of MOR is to compress the full problem to a smaller form which maintains the important properties of $\mathbf{A}$ with high accuracy. Of particular interest to us is a subtype of MOR, parametric model order reduction (PMOR), which allows us to create reduced-order models which depend linearly on a set of parameters. These parameters are the LECs in the present application. EC takes the problem at hand and projects the original large-space system onto a smaller subspace, thus creating a low-dimensional representation. For the Hamiltonian, the subspace is spanned by eigenvectors of the original problem at a number of preselected values of the input parameters.

A note on terminology is in order. Emulation occurs in two distinct stages: *offline* and *online*. The offline stage, which may also be known as "training" or "calibration", encompasses the construction of the emulator and is a one-time cost which may be expensive to perform. In contrast, the online stage is when the emulator is in use; operations performed during the online stage should be computationally cheap. A major component of the offline stage is the computation of a number of *snapshots*; a snapshot is the result of a simulation, i.e., an exact solution to the problem at hand.

4.2 Variational emulators

We need two types of emulators: one for bound states, and one for scattering amplitudes. However, most of the work relevant for this thesis entails scattering emulators. Nevertheless, both emulator types are variational in nature, i.e., we are looking for approximate solutions $\widetilde{\psi} = \widetilde{\psi}(\vec{\beta})$ to a functional $S[\psi]$ where $\widetilde{\psi}$ depends on variational parameters $\vec{\beta}$ and the stationarity condition $\delta S[\widetilde{\psi}(\vec{\beta})] = 0$ is fulfilled for $\vec{\beta} = \vec{\beta_\star}$. Here, ψ represents exact solutions. Starting from the ansatz

$$\widetilde{\psi} = \sum_{i=1}^{N_s} \beta_i \psi_i = \mathbf{U}\vec{\beta} \tag{4.1a}$$

$$\mathbf{U} = [\psi_1, \ldots, \psi_{N_s}], \tag{4.1b}$$

where $\psi_i = \psi_i(\vec{\alpha}_i)$ are N_s snapshots of ψ at N_s different sets of LECs $\vec{\alpha}_i$, I will now define S and find $\vec{\beta_\star}$ for both types of emulators.

4.2.1 Neutron-proton scattering

For np scattering we wish to find $\widetilde{T}$ which emulates the on-shell $\hat{T}$-matrix element T for a given momentum p. The formalism for doing so has been developed [40, 41] by combining EC with the Kohn [110] and Newton [111] variational principles, respectively. Here, I employ the method developed in Ref. [41], which uses the Newton variant.

To find a functional S, we first write the LS equation in two equivalent forms:

$$\hat{T} = \hat{V} + \hat{V}\hat{G}_0\hat{T} \tag{4.2a}$$

$$\hat{T} = \hat{V} + \hat{T}\hat{G}_0\hat{V} \tag{4.2b}$$

Rewriting Eq. (4.2a) as

$$\hat{V} = \hat{T} - \hat{V}\hat{G}_0\hat{T} \tag{4.3}$$

and inserting into Eq. (4.2b) yields

$$\hat{T} = \hat{V} + \hat{T}\hat{G}_0(\hat{T} - \hat{V}\hat{G}_0\hat{T}) = \hat{V} + \hat{T}\hat{G}_0\hat{T} - \hat{T}\hat{G}_0\hat{V}\hat{G}_0\hat{T}. \tag{4.4}$$

If we subtract Eq. (4.4) from the sum of the two variants of the LS equation we get the Newton functional [111]

$$S[\hat{T}] = \hat{V} + \hat{V}\hat{G}_0\hat{T} + \hat{T}\hat{G}_0\hat{V} - \hat{T}\hat{G}_0\hat{T} + \hat{T}\hat{G}_0\hat{V}\hat{G}_0\hat{T}. \tag{4.5}$$

which is stationary around exact solutions to the LS equation.

As in Eq. (4.1a), we make the ansatz

$$\widetilde{T} = \sum_{i=1}^{N_s} \beta_i \hat{T}_i \tag{4.6}$$

where each $\hat{T}_i = \hat{T}_i(\vec{\alpha}_i)$ is a snapshot of the on-shell $\hat{T}$-matrix element at different sets of LECs. Inserting Eq. (4.6) into Eq. (4.5) yields

$$S[\hat{T}] = \hat{V} + \vec{\beta}^{\,T}\vec{b} - \frac{1}{2}\vec{\beta}^{\,T}\mathbf{B}\vec{\beta} \tag{4.7}$$

where the elements of the column vector $\vec{\beta}$ and the matrix $\mathbf{B}$ are

$$b_i = \hat{T}_i\hat{G}_0\hat{V} + \hat{V}\hat{G}_0\hat{T}_i \tag{4.8a}$$

$$B_{ij} = \hat{T}_i\hat{G}_0\hat{T}_j + \hat{T}_j\hat{G}_0\hat{T}_i - \hat{T}_i\hat{G}_0\hat{V}\hat{G}_0\hat{T}_j - \hat{T}_j\hat{G}_0\hat{V}\hat{G}_0\hat{T}_i. \tag{4.8b}$$

The stationarity condition $dS/d\vec{\beta} = 0$ then yields

$$\mathbf{B}\vec{\beta}_\star - \vec{b} = 0 \tag{4.9}$$

and the resulting emulator is obtained as [41]

$$\hat{T} \approx \hat{V} + \frac{1}{2}\vec{b}^{\,T}\mathbf{B}^{-1}\vec{b}. \tag{4.10}$$

The only LEC dependence here is in the potential, $\hat{V} = \hat{V}(\vec{\alpha})$. Chiral potentials are linear with respect to the LECs, see Eq. (2.7). We can therefore precompute the constituent terms of $\vec{b}$ and $\mathbf{B}$, and evaluating the emulator for a new set of LECs consists of performing simple dot products over the terms in Eqs. (4.8a) and (4.8b) where $\hat{V}$ appears, followed by an evaluation of Eq. (4.10).

4.2.2 Bound-state properties of nuclei

The premise for emulating bound-state properties of nuclei is the same as for NN scattering, but the expressions differ since this is an eigenvalue problem. We can rewrite the Schrödinger equation,

$$\hat{H}\,|\psi\rangle = E\,|\psi\rangle\,, \tag{4.11}$$

as a functional

$$S[\psi] = \langle\psi|\,\hat{H}\,|\psi\rangle - E(\langle\psi|\psi\rangle - 1). \tag{4.12}$$

As before, we seek an approximate solution that fulfils the stationary condition $\delta S[\widetilde{\psi}] = 0$. We have [109]

$$\delta S[\psi] = 2\,\langle\delta\psi|\,\hat{H} - E\,|\psi\rangle - \delta E(\langle\psi|\psi\rangle - 1) = 0. \tag{4.13}$$

With an ansatz as in Eq. (4.1a), we get

$$\delta S[\psi] = 2\delta\vec{\beta}^{\dagger}\mathbf{U}^{\dagger}(\hat{H} - E)\mathbf{U}\vec{\beta} - \delta E(\vec{\beta}^{\dagger}\mathbf{U}^{\dagger}\mathbf{U}\vec{\beta} - 1) = 0 \tag{4.14}$$

which is approximately fulfilled by the variational parameters $\vec{\beta}_\star$ when

$$\mathbf{U}^{\dagger}\hat{H}\mathbf{U}\vec{\beta}_\star = \widetilde{E}\mathbf{U}^{\dagger}\mathbf{U}\vec{\beta}_\star. \tag{4.15}$$

If we define

$$\widetilde{H} = \mathbf{U}^{\dagger}\hat{H}\mathbf{U} \tag{4.16a}$$

$$\mathbf{N} = \mathbf{U}^{\dagger}\mathbf{U}, \tag{4.16b}$$

where $\mathbf{N}$ is the norm matrix, this is recognized as a generalized eigenvalue problem in N_s (i.e., the number of snapshots) dimensions,

$$\widetilde{H}\vec{\beta}_\star = \widetilde{E}\mathbf{N}\vec{\beta}_\star, \tag{4.17}$$

with eigenvalues $\widetilde{E}$. By substituting Eq. (4.15) into Eq. (4.13) we see that $\widetilde{E}$ is an approximation for the eigenenergy E of the original Schrödinger equation. Crucially, $\dim(\widetilde{H})$ is far smaller than $\dim(\hat{H})$. The reason for this is that while the number of basis states needed to construct $\hat{H}$ can be high, its effective dimensionality—its information density—is often low [63]. Thus, the computational effort needed to find $\widetilde{E}$ is a small fraction of what is required for E.

4.3 Implementation of fast neutron-proton scattering emulators

In Paper IV, I develop Python [112] code for both simulation and emulation of $\hat{T}$-matrix elements, along with code to compute scattering observables in line with the formalism introduced in Chapter 2.

4.3.1 Construction of emulators

It is important to note that an emulator as presented in Section 4.2.1—Eq. (4.10) in particular—emulates the on-shell $\hat{T}$-matrix element for one energy and one PW only. Emulators must be constructed and evaluated for every channel at every energy for which we wish to compute a scattering cross section; a great number of emulators, and some bookkeeping, will thus be required for inference problems. My choices of data and PW truncation in Paper IV results in 177 unique energies and 182 channels, for a total of 32 214 emulators. For simplicity, I will from now on use the word "emulator" to denote the collection of emulators at a specific energy.

It is straightforward to create an emulator by directly implementing Eq. (4.10) and looping over each channel of the PW basis. However, we can gain efficiency by using the superiority of NumPy [113] operations over Python loops and exploiting the linearity (2.7) of the chiral potential $\hat{V}$ with respect to the LECs $\vec{\alpha}$. In the offline stage we can precompute

$$\mathbf{X}_{ij} = \hat{T}_i \hat{G}_0 \hat{T}_j + \hat{T}_j \hat{G}_0 \hat{T}_i \tag{4.18}$$

$$\mathbf{Y}_{ij} = \hat{T}_i \hat{G}_0 \hat{V}(\vec{1}) \hat{G}_0 \hat{T}_j + \hat{T}_j \hat{G}_0 \hat{V}(\vec{1}) \hat{G}_0 \hat{T}_i \tag{4.19}$$

$$\mathbf{Z}_i = \hat{T}_i \hat{G}_0 \hat{V}(\vec{1}) + \hat{V}(\vec{1}) \hat{G}_0 \hat{T}_i \tag{4.20}$$

for $1 \leq i, j \leq N_s$ and only store the on-shell elements (or any off-shell elements, for that matter, if they are of interest). $\mathbf{X}$ will then be an $N_s \times N_s$ matrix, $\mathbf{Y}$ a block-diagonal matrix stored as an $(N_{\text{LECs}} + 1)$-length vector of $N_s \times N_s$ matrices, and similarly $\mathbf{Z}$ an $(N_{\text{LECs}} + 1)$-length vector of N_s-length vectors. In

the online stage it is then sufficient to compute $\vec{\alpha} \cdot \mathbf{Y}$, $\vec{\alpha} \cdot \mathbf{Z}$, and

$$\mathbf{B}_{ij} = \mathbf{X}_{ij} - \mathbf{Y}_{ij} \tag{4.21a}$$

$$\vec{b}_i = \mathbf{Z}_i \tag{4.21b}$$

followed by an evaluation of Eq. (4.10). These steps are repeated for each channel at each energy. At this point, we have all on-shell $\hat{T}$-matrix elements required to compute observables.

4.3.2 Fast observables from emulated $\hat{T}$-matrix elements

The process of computing observables, as detailed in Section 2.2.1, is identical whether the $\hat{T}$-matrix elements are simulated or emulated. With emulation, however, going from $\hat{T}$ to cross sections represents a considerable fraction of the overall computational cost and in order to truly leverage the advantage of emulation we should try to compute cross sections as efficiently as possible.

The key to achieving this is precomputation of as many factors as possible in the spin-scattering matrix (2.20). Since we know at the outset which observables appear in the likelihood (see Chapter 3), the only a priori unknown factor is $\langle L'SJ|\,\hat{S} - 1\,|LSJ\rangle$ and we may compute $\mathbf{M}$ by multiplying the $\hat{T}$-matrix elements with precomputed factors and summing over all channels. From there, cross sections can be computed via the Saclay amplitudes. Doing this, I find that emulating T dominates the overall computational cost of computing scattering cross sections.

4.3.3 Snapshots of $\hat{T}$

As discussed in Section 4.2.1, a number of snapshots of the $\hat{T}$ matrix are needed to construct an emulator. It is advisable to choose LECs for the snapshots by using a space-filling method such as Latin hypercube sampling (LHS) [114] for two reasons. The first is to cover as much of the parameter space as possible, thus informing the emulator about the interesting parameter region. The second is to avoid (nearly) linearly dependent columns in (4.8b), which may impact our ability to invert $\mathbf{B}$. In order to use LHS, one must define the extent of the space. χEFT is helpful here: the expected naturalness of the LECs implies that we only need to explore a relatively small region in the vicinity of zero. Previous works also support this conclusion [52, 54, 59, 74, 104], with Ref. [105] being especially relevant here since we employ Δ-full χEFT in Paper IV.

In Paper II, I infer that the LO LECs ($\widetilde{C}_{^1S_0}$ and $\widetilde{C}_{^3S_1}$) are both confined to a minute region near $\widetilde{C}_{^1S_0} \approx \widetilde{C}_{^3S_1} \approx -0.1 \times 10^4\,\mathrm{GeV}^{-2}$, in agreement with earlier works (see, e.g., Ref. [54]). There is, to my knowledge, no precedent for large values of these parameters, and their expected natural size is $|\widetilde{C}| \sim 0.15 \times 10^4\,\mathrm{GeV}^{-2}$ [115]. Hence, my LO emulators are trained in $[-0.3, 0.3] \times 10^4\,\mathrm{GeV}^{-2}$

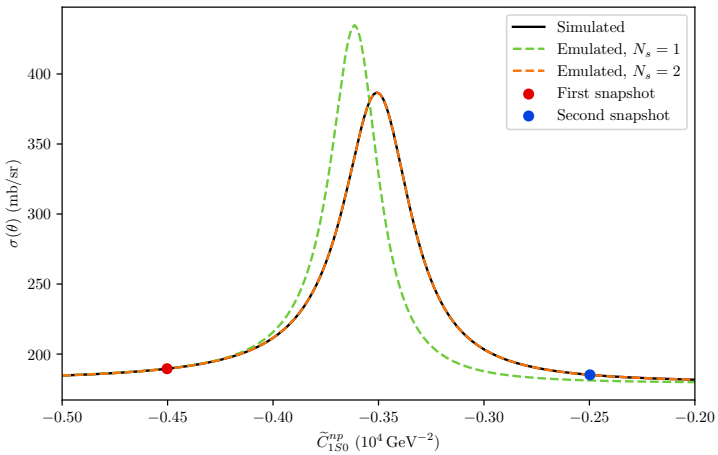


Figure 4.1: The np differential cross section $\sigma(\theta)$ at $T_{\rm lab} = 1.0$ MeV and $\theta = 25$ degrees, computed at Δ-NNLO as a function of $\widetilde{C}^{np}_{1S0}$ using simulation (black) and two different emulators trained with $N_s = 1, 2$ snapshots (green and orange, respectively). The snapshots are shown as colored dots. All other LECs are kept fixed at their Δ-NNLO$_{\rm GO}$ values [105] in both the offline and online stages.

for both parameters. The NN contact LECs at NLO and NNLO in Δ-full EFT take on somewhat larger values than at LO so we need to train in a wider region; I choose $[-4.0, 4.0]$ (in naturalness units) in all directions. The exact limits are not important as we expect the emulators to both inter- and extrapolate well [64]. The recent Δ-NLO$_{\rm GO}$ and Δ-NNLO$_{\rm GO}$ interactions [105] find all parameters to be naturally-sized with maximum parameter values around 2.5 (in the appropriate units), well within my limits. Alternatively, the Bayesian priors could be exploited to provide regions for training the emulators, but in the present case I find that my simple prescription is satisfactory.

I have found that fewer than ten snapshots are needed in order to achieve emulator errors that are orders of magnitude smaller than the experimental error for the observables we wish to compute, as discussed in the next section.

4.3.4 Emulator accuracy

A useful emulator introduces an error $\Delta y_{\rm emul}$ that is negligible in comparison to the experimental error $\delta y_{\rm exp}$ and the EFT truncation error $\delta y_{\rm th}$. Here, I will demonstrate that $\Delta y_{\rm emul} \ll \delta y_{\rm exp}$; we show in Paper IV that this holds for the truncation error as well. But let me first illustrate the emulator accuracy visually.

Figs. 4.1 and 4.2 show np differential cross sections at $T_{\rm lab} = 1.0$ MeV and

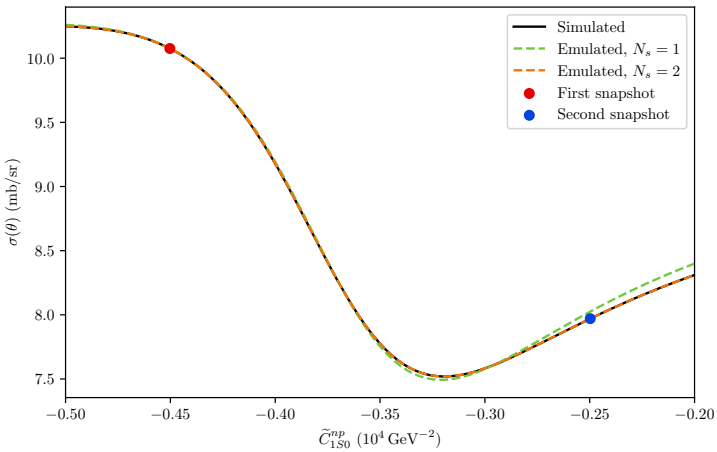


Figure 4.2: Same as Fig. 4.1, except at $T_{\rm lab} = 100.0$ MeV and $\theta = 25$ degrees.

100.0 MeV, respectively, both at a scattering angle of 25 degrees. All other details are identical. The cross sections are computed at Δ-NNLO as a function of $\widetilde{C}^{np}_{1S0}$. A simulation and two different emulations are shown. The emulators are trained in one direction using one and two snapshots, respectively. All other LECs are fixed at their Δ-NNLO$_{\rm GO}$ values. The figures demonstrate remarkable accuracy. At just two snapshots, the emulated results are indistinguishable from the simulation, even through the challenging interpolation in Fig. 4.1 where the cross section shows a strong curvature with respect to the LEC. We also see that even a single snapshot yields a reasonable agreement with simulation at the higher energy in Fig. 4.2. However, we clearly require at least two snapshots to achieve an accurate emulator, especially since all other directions are fixed in this case.

A more quantitative result can be seen in Fig. 4.3, which shows the emulator error as a fraction of the experimental error for 2 779 *np* scattering observables, where data in the $0 < T_{\rm lab} < 30$ range is included in addition to the data D. The observables are calculated with $N_s = 8$ snapshots at Δ-NNLO. The LEC values are taken from Δ-NNLO$_{\rm GO}$ [105]. The values shown in the figure are calculated as

$$\Delta y_i = \frac{|y_{\rm emulation} - y_{\rm simulation}|_i}{\Delta y_{{\rm exp},i}} \tag{4.22}$$

where i is the index of the observable and $\Delta y_{{\rm exp},i}$ is (the standard deviation of) the experimental error. The figure shows that, for most observables, the emulator error is roughly 0.1% of the experimental error and thus negligible. In the worst case, $\Delta y_{\rm emul}$ is just under 10% of $\Delta y_{{\rm exp},i}$, which is minor for a single

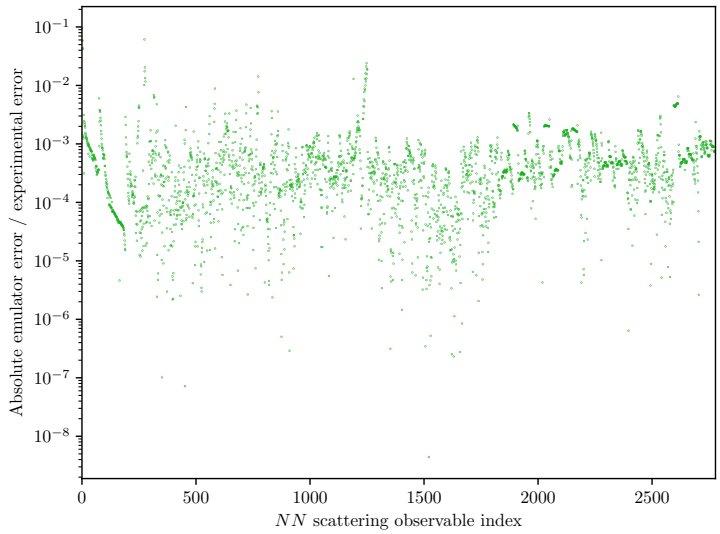


Figure 4.3: Emulator error at Δ-NNLO as a fraction of the experimental error for 2 779 np scattering observables. The emulators were trained using $N_s = 8$ snapshots with LECs in the interval $[-4.0, 4.0]$ in natural units (see Chapter 1). The snapshots were produced with LHS. The validation set of LECs is the Δ-NNLO$_{\mathrm{GO}}$ set.

datum. Fig. 4.3 further reveals that there are structures to the emulator error; this is quite expected, as many observables are measured in clusters at narrow intervals of energies and angles within the same experiment. In particular, observables that are identical except for the scattering angle are computed using the same emulator and it is natural that the error is similar.

Fig. 4.4 shows the same type of plot, but calculated at ten randomly generated sets of validation LECs instead of the Δ-NNLO$_{\mathrm{GO}}$ set. The validation sets are generated the same way as the training snapshots, but with a different random seed. The figure shows that the overall behavior of the emulators is consistent within this parameter space. I conclude that the NN scattering emulator errors are negligible and thus do not model the ensuing uncertainty in the inference of the LECs.

4.3.5 Automatic differentiation and code acceleration with JAX

As discussed in Chapter 3, we need not only the observables but their gradients with respect to the LECs in order to use HMC. Several optimization algorithms

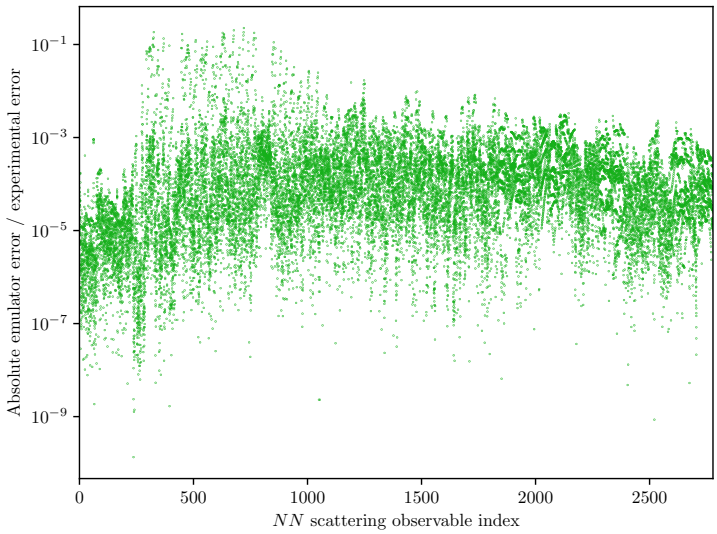


Figure 4.4: Emulator error at Δ-NNLO as a fraction of the experimental error for 2 779 *np* scattering observables. The emulators were trained using $N_s = 8$ snapshots with LECs in the interval $[-4.0, 4.0]$ in natural units (see Chapter 1). The snapshots were produced with LHS. Ten sets of validation LECs are included in the figure.

also rely on gradient information. Since our emulators are linear in $\vec{\alpha}$, it is straightforward to differentiate them with respect to the LECs. However, in HMC one needs gradients of the entire (negative log) posterior with respect to the LECs. Deriving explicit expressions for these gradients would, while possible, be laborious, in particular as the truncation error is GP-modeled and future applications might entail HMC-sampling of the hyperparameters. A superior alternative is to use automatic differentiation (AD) [86, 87]. AD exploits the fact that all mathematical calculations performed on a computer are broken down into a set of elementary operations (addition, multiplication, etc.) and elementary functions (log, cos, etc.) whose derivatives are known. By repeatedly applying the chain rule, and overloading variable types to also include derivative information, AD can compute gradients of arbitrary functions to machine precision. The computational overhead is typically about a factor of two to three [86]. In Papers I-III, where I employ a C/Fortran code to simulate NN scattering observables, an AD library called Rapsodia [87] is used in combination with some analytic expressions to extract gradients. In Paper IV, where observables are emulated in Python, I instead use JAX [116].

A brief demonstration of AD in JAX is in order. In Listing 4.1, I instruct

JAX to compute the value and the gradient of a function `lnposterior`, which of course represents the Bayesian (log) posterior for the LECs $\vec{\alpha}$ (the so-called traced variable). This function includes every aspect of the posterior, including calculation of all included observables through emulation, the LEC prior, and so on. Note that the JAX library can trace gradients through relatively advanced operations such as matrix inversions and eigenvalue problems.

Listing 4.1: Demonstration of how to use JAX to compute the gradient of a function.

```
import jax

def lnposterior(lecs, *args, **kwargs):
    '''Log posterior for the LECs.'''
    value = lnlikelihood(lecs, *args, **kwargs)
    value = value + lnprior(lecs, *args, **kwargs)
    return value

def lnprior(lecs, *args, **kwargs):
    [function definition omitted]

def lnlikelihood(lecs, *args, **kwargs):
    [function definition omitted]

# Tell JAX that we want gradients of lnposterior
lnposterior_AD = jax.value_and_grad(lnposterior)

# Evaluate for an LEC array 'lecs'
value, gradient = lnposterior_AD(lecs)
```

The variables `value` and `gradient` will now contain the value of the log posterior computed at $\vec{\alpha}$ and its gradient with respect to $\vec{\alpha}$, respectively. The syntax is such that the first argument to the differentiated function defines the variable for which the gradient should be computed. Thus, JAX is in principle extremely easy to use. However, all code executed as part of `lnposterior` must abide by some relatively minor restrictions imposed by the library, and existing code may need to be altered in order to conform. To give two examples:

- in-place modification of array elements (e.g. `arr[0] = x`) is not allowed. Instead, something like `arr = arr.at[0].set(x)` must be used.
- Most NumPy operations need to be performed using JAX's NumPy extension, jax.numpy.

I refer the reader to the documentation [117] for all details.

Unfortunately, the code as shown in Listing 4.1 is likely to carry a significant computational overhead compared to a non-differentiated version due to

the design of JAX. But JAX has a trick up its sleeve that will actually make the differentiated code *faster* than a standard Python implementation. This trick is just-in-time compilation (JIT). JIT takes the Python function and, upon first invocation, caches a compiled version of it in memory, thus increasing execution speed of subsequent calls. In JAX, AD and JIT are designed to be used together to create a very powerful tool for machine learning and other computationally demanding applications. The syntax—and caveats—for JIT is similar to the AD functionality. For example, to introduce JIT in Listing 4.1 one might add the following line: `lnposterior_AD_JIT = jax.jit(lnposterior_AD)`. The function `lnposterior_AD_JIT` may then be used the same way as `lnposterior_AD`.

JIT is very powerful and my usage of it is relatively mundane, as it is mainly intended to be used on graphics processing units (GPUs) and other hardware accelerators whereas I am using it on standard central processing units (CPUs). Nevertheless, the resulting execution speed is impressive. For example, one evaluation of a Δ-NNLO posterior conditioned on 2 779 scattering cross sections (where the energy range $0 < T_{\text{lab}} < 30$ has been included for benchmarking) across a total of 87 906 PW channels takes 1.3 s with AD on a modern CPU, or 0.4 s without AD, with emulators trained on eight snapshots. The CPU used for these benchmarks is a 16-core Intel Core i9-12900K. No explicit multithreading is done in my code.

4.4 Emulating the no-core shell model for few-body systems

In Paper I, we emulate observables for ^{3}H, ^{3}He, and ^{4}He, using the approach laid out in Section 4.2.2 applied to the NCSM method [12, 13] to analyze the χEFT truncation error. We wish to emulate

$$\hat{H}(\vec{\alpha})\,|\psi(\vec{\alpha})\rangle = E(\vec{\alpha})\,|\psi(\vec{\alpha})\rangle\,, \tag{4.23}$$

where $|\psi(\vec{\alpha})\rangle$ denotes the ground state and $E(\vec{\alpha})$ its energy and the LEC dependence is made explicit. Using $N_s = 50$ NCSM snapshots of $|\psi(\vec{\alpha})\rangle$ we construct the generalized eigenvalue problem given by Eq. (4.17), thus reducing the basis size from $\dim[\hat{H}(\vec{\alpha})] \approx 10^4$ to $\dim[\widetilde{H}(\vec{\alpha})] = 50$. $\hat{H}(\vec{\alpha})$ is linear with respect to $\vec{\alpha}$,

$$\hat{H}(\vec{\alpha}) = \hat{H}_{\text{constant}} + \sum_{i=1}^{N_{\text{LECs}}} \alpha_i \hat{V}_i, \tag{4.24}$$

where $\hat{H}_{\text{constant}}$ encompasses all LEC-independent contributions and $\hat{V}_i$ encompasses all contributions that depend on α_i. The subspace-projected Hamiltonian

is likewise linear:

$$\widetilde{H}(\vec{\alpha}) = \widetilde{H}_{\text{constant}} + \sum_{i=1}^{N_{\text{LECs}}} \alpha_i \widetilde{V}_i. \tag{4.25}$$

The subspace projections of $\widetilde{H}_{\text{constant}}$ and $\widetilde{V}_i$ need therefore only be computed in the offline stage. The online stage then entails computing the sum (4.25) and solving the eigenvalue problem Eq. (4.17).

It is straightforward to emulate observables other than ground-state energies. The expectation value of an observable $\hat{O}(\vec{\alpha})$, given by

$$\langle \hat{O}(\vec{\alpha}) \rangle = \langle \psi(\vec{\alpha}) | \, \hat{O}(\vec{\alpha}) \, | \psi(\vec{\alpha}) \rangle \, , \tag{4.26}$$

can be emulated as

$$\langle \hat{O}(\vec{\alpha}) \rangle \approx \vec{\beta_\star}(\vec{\alpha})^\dagger \left[\mathbf{U}^\dagger \hat{O}(\vec{\alpha}) \mathbf{U} \right] \vec{\beta_\star}(\vec{\alpha}), \tag{4.27}$$

where $\vec{\beta_\star}(\vec{\alpha})$ is an emulated eigenstate corresponding to $|\psi\rangle\,(\vec{\alpha})$.

The emulators developed for Paper I show errors (compared to simulation) that are orders of magnitude smaller than the corresponding experimental uncertainty and/or the error of the NCSM simulation itself. Meanwhile, the computational speedup is vast: from ~ 1 minute to 10 ms, or a factor of 6 000.

Chapter 5

Inferred posteriors and posterior predictive distributions

Broadly speaking, Bayesian analyses of parametric models yield two results: the intermediate posterior PDF for the values of the parameters, and the final PPDs for unobserved values of observables. In this chapter I will highlight and discuss selected results of both types of distributions in the context of χEFT. In many cases I will present results that are included in the papers, but with some of the details expanded upon.

5.1 A gentle introduction: leading order

The LO in standard Weinberg power counting has deficiencies that have been known for quite some time [26]. It yields predictions that struggle to meet expectations; part of which is exemplified in Paper II, where the EFT expansion coefficients for the LO-NLO shift in theoretical predictions are concerningly large, i.e., the NLO correction is greater than it is expected to be. Efforts to improve this situation are under way (see, e.g., Ref. [29] and references therein), but in the meantime LO predictions should be viewed with a measure of scepticism. These issues notwithstanding, LO can teach us important lessons about UQ in χEFT since its two-dimensional parameter space can be thoroughly explored without MCMC.

We do this in Paper II and find that the LEC posterior conditioned on NN scattering data is multimodal. We explicitly show that the posterior is bimodal due to the emergence of an np bound state—the deuteron—at $\widetilde{C}_{^3S_1} \approx -0.05 \times 10^4\,\mathrm{GeV}^{-2}$. Near the LEC value where the bound state forms, the total cross section at low energies grows extremely large and the likelihood for observed data becomes correspondingly small. Most MCMC algorithms—including HMC—will not be able to cross this barrier in practice. The same

effect occurs in the ${}^1S_0$ channel, and we therefore know that the LO posterior is (at least) quad-modal. We can, however, safely discard all parameter values where the deuteron is unbound and/or an unphysical bound state appears in the ${}^1S_0$ channel. We can generalize this reasoning to higher orders and in principle dismiss any LEC values that predict unphysical bound states between two nucleons. While we do not do so, incorporating such knowledge in the information I could certainly be worthwhile; as an example, there are 23 NN contact LECs at N3LO in Weinberg power counting, all of which can be chosen such that one or more bound states do (or do not) appear in their respective channels within the EFT low-energy range of validity. From a pure modeling perspective, dismissing values known to yield unphysical results can significantly limit the parameter space.

The disjointed regions of probability mass in the LO posterior also highlights the importance of vigilance regarding signs of multimodality. We routinely randomize starting points for the optimization algorithm and the MCMC sampler.

5.2 Posteriors for the low-energy constants

Beyond LO, the complexity of the sampling problem increases dramatically, with the next-simplest case—NLO—featuring eleven LECs. Throughout this work, we employ HMC to sample all posteriors and we have yet to encounter a posterior which we have not been able to sample efficiently with our HMC protocol. As an example, we show in Paper II that the HMC sampling at NLO is over six times more efficient than if we use the affine-invariant EMCEE [82] sampler. Given that the wall-time to reach convergence for this particular EMCEE sampling is in excess of a week, this is a major improvement. With HMC, we achieve antithetic sampling (see Chapter 3) with $\tau = 0.78$, which yields an ESS 1.3 times greater than the length of the MCMC chain. Beyond NLO, we have only achieved converged results with HMC.

The HMC sampling at NLO in Paper II is a key moment, as it is our first successful application of HMC to a challenging sampling problem. The resulting posterior is shown in Fig. 5.1. The figure, which is also included in Paper II, shows the LEC posterior at NLO in Δ-less χEFT conditioned on np and pp scattering data in the $0 < T_{\text{lab}} \leq 290$ MeV range using an uncorrelated truncation error model. The posterior has some features that we have since come to expect. First, we see strong correlation between the charge-dependent leading-order LECs $\widetilde{C}^{np}_{^1\text{S}_0}$ and $\widetilde{C}^{pp}_{^1\text{S}_0}$, which is unsurprising given that the charge-dependence of the strong nuclear force is known to be a small effect. Second, we see (in this case strong) anticorrelations between leading-order LECs $\widetilde{C}$ and higher-order LECs C acting in the same partial waves, while most other LECs are pairwise uncorrelated. For example, $\widetilde{C}_{^3S_1}$ and $C_{^3S_1}$ are strongly anticorrelated. Third, the posterior is overall very narrow, i.e., the LECs are very well

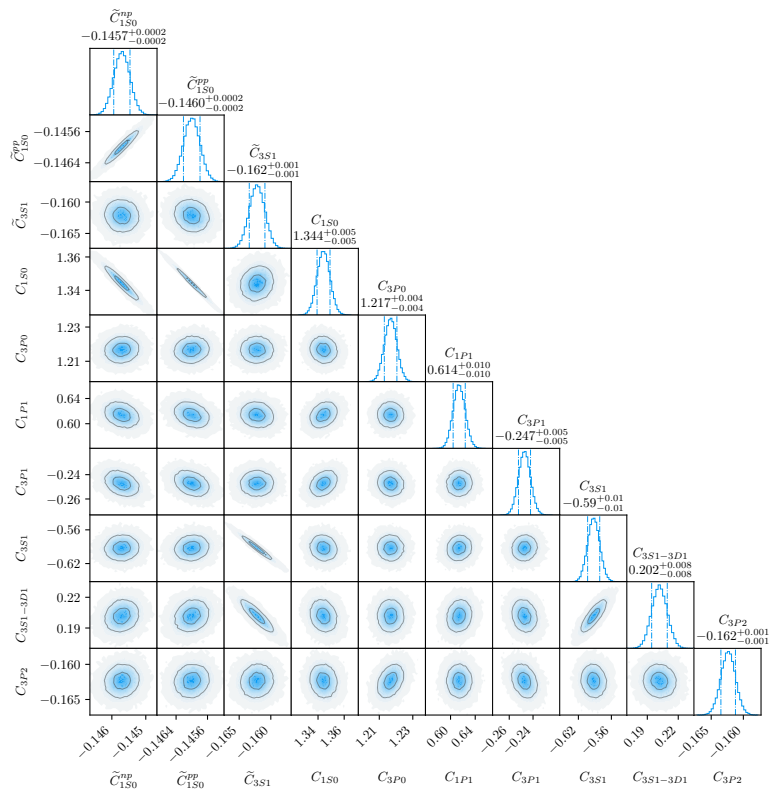


Figure 5.1: NLO posterior for the LECs in Δ-less χEFT in units of $10^4\,\mathrm{GeV}^{-2}$ ($\widetilde{C}$) and $10^4\,\mathrm{GeV}^{-4}$ (C). Originally published in Ref. [90] (Paper II). Copyright © 2022 by the American Physical Society. Reproduced with permission.

determined. This observation returns in all posteriors we have seen.

In Paper III, we present a sucessful sampling of an LEC posterior at N3LO in Δ-less χEFT. The dimensionality of this posterior puts printing it in its entirety out of the question (it is published as supplemental material in Paper III), but a subset of the LECs are shown in Fig. 5.2. The included LECs are the fourth-order contact LECs D, which do not appear at the lower orders. These LECs are more sensitive to high-momentum data, and we find that we must include data above 40 MeV in order to constrain all of them. The most notable feature in Fig. 5.2 is that three LECs—$D_{^1S_0}$, $D_{^3S_1}$, and $D_{^3P_1}$—are unnaturally sized. Precedents for unnaturally-sized $D_{^1S_0}$ exist [52], and the choice of the regulator can impact the size of the LECs [104]. However, the in my opinion most striking aspect of this N3LO result is that it was nearly as easy to accomplish as the

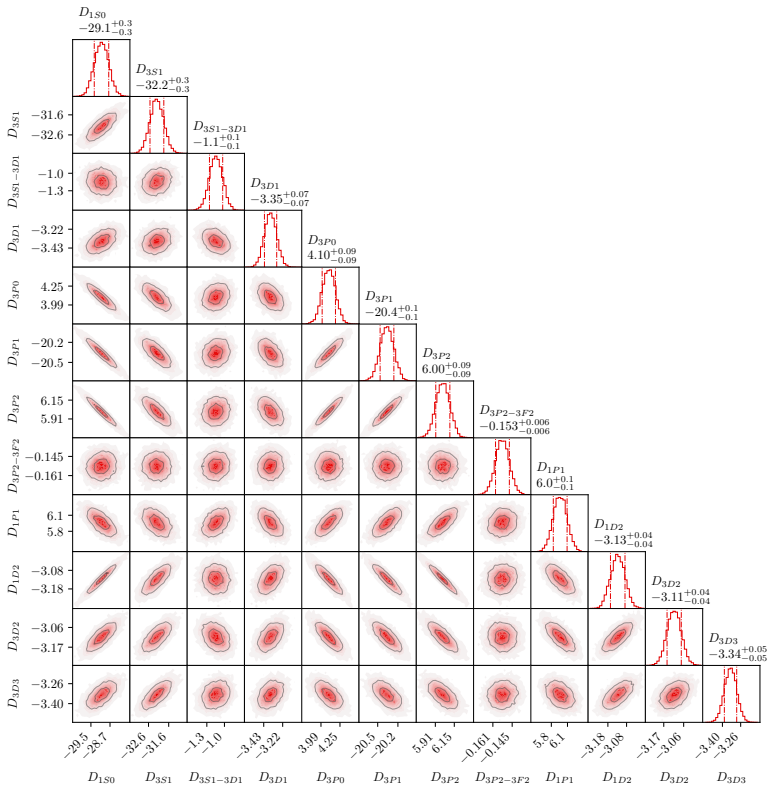


Figure 5.2: Marginal posterior for a subset of the LECs at N3LO in Δ-less χEFT. Only NN contact LECs exclusive to N3LO and higher orders are shown, i.e., eight πN LECs are excluded along with eleven lower-order contact LECs. Hence, 19 LECs are not shown. The LECs are shown in units of 10^4 GeV^{-6}.

lower-order samplings, inspiring confidence in attempting even more challenging samplings in future works.

In Paper III we also employ SIR to extend the posteriors by inferring $\widetilde{C}^{nn}_{^1\mathrm{S}_0}$ from an empirical scattering length $a^{nn}_{^1S_0} = -18.9 \pm 0.4$ fm [69], where electromagnetic effects have been removed. We place a conjugate prior on $\bar{c}^2$ for $\widetilde{C}^{nn}_{^1\mathrm{S}_0}$ and find that this results in a wider marginal posterior for this LEC than the two other charge-dependent LECs, $\widetilde{C}^{np}_{^1\mathrm{S}_0}$ and $\widetilde{C}^{pp}_{^1\mathrm{S}_0}$. We also find that MAP predictions of ^{2}H, ^{3}H, ^{3}He, and ^{4}He observables, such as ground-state energies, are in good agreement with experimental results given that we do not include 3NFs. The exceptions to this finding are predictions of ^{3,4}He and ^{3}H at N3LO; we can only speculate as to the cause, but the unnaturally-sized D LECs may

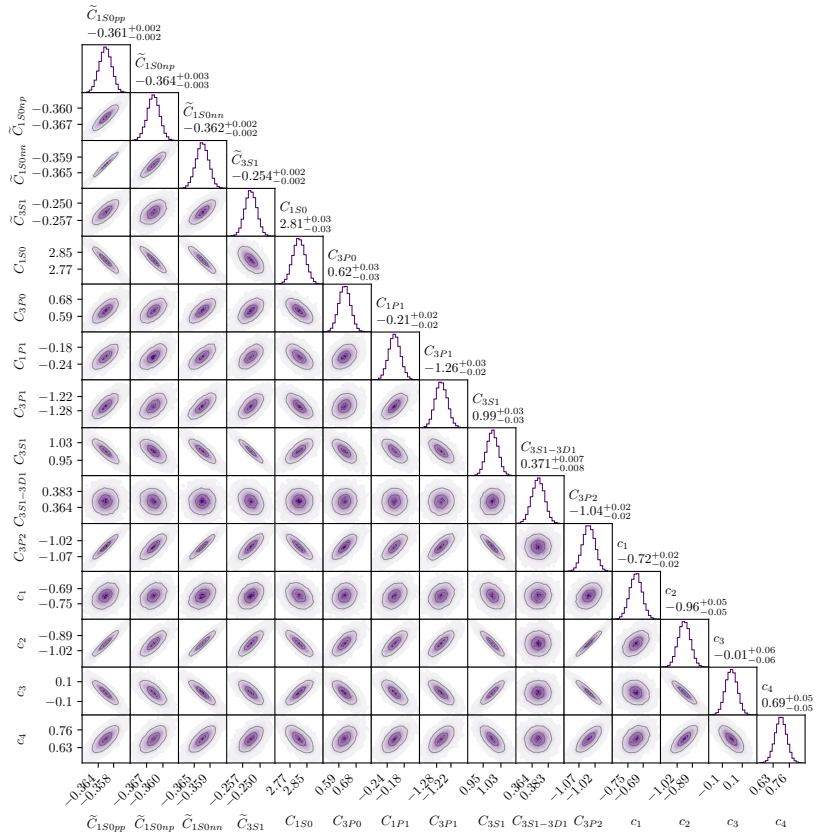


Figure 5.3: NNLO posterior for the LECs in Δ-full χEFT, in units of $10^4\,\mathrm{GeV}^{-2}$ ($\widetilde{C}$), $10^4\,\mathrm{GeV}^{-4}$ (C), and GeV^{-1} (c).

contribute. Of course, the NN and NNN LECs should ideally be simultaneously inferred from $A \geq 2$ data. We also show that this two-stage sampling procedure yields identical results to an MCMC sampling where we in the first step approximate the marginal posterior for the HMC sampled LECs by a multivariate normal distribution and then sample all LECs with HMC in a second step. The posteriors resulting from Paper III were used by Miller *et al.* [9] to investigate the impact of LEC variability on Nd elastic scattering.

In Paper IV we turn to emulation and employ the Δ-full variant of χEFT. Here, much effort goes into characterizing correlated EFT truncation errors to investigate the effect of including non-zero, finite correlation lengths. Fig. 5.3 shows a Δ-NNLO LEC posterior with a correlated model for the truncation error. A central question we wanted to answer is whether the very narrow

posteriors we find with an uncorrelated error model will become significantly wider if we take correlations in the truncation error into account. We find that the posteriors do indeed become wider—by roughly a factor of 2-3 at Δ-NLO and about 1.5 at Δ-NNLO—by comparing with a posterior that is identical except that an uncorrelated error model is used. Since the balance of experimental and theoretical errors changes with the chiral order, it is not surprising that the characterization of the truncation error has a greater impact at NLO than at NNLO. Unlike the corresponding result shown in Paper IV, I here include empirical ${}^1S_0$ scattering lengths for nn and pp scattering $a^{nn}_{{}^1S_0}$ and $a^{pp}_{{}^1S_0} = -17.3 \pm 0.40$ fm [1] in order to infer $\widetilde{C}^{nn}_{{}^1\mathrm{S}_0}$ and $\widetilde{C}^{pp}_{{}^1\mathrm{S}_0}$. As for $a^{nn}_{{}^1S_0}$, electromagnetic effects have been removed (by the originators of the empirical value) for $a^{pp}_{{}^1S_0}$. Except for the inclusion of $(\widetilde{C}^{nn}_{{}^1\mathrm{S}_0}, \widetilde{C}^{pp}_{{}^1\mathrm{S}_0})$, Fig. 5.3 is identical to the corresponding result in Paper IV, i.e., the included scattering lengths only affect the explicitly charge-dependent LECs. The truncation errors for $a^{nn}_{{}^1S_0}$ and $a^{pp}_{{}^1S_0}$ are treated as uncorrelated.

In Paper I, we employ a multivariate normal prior for the NN LECs based on an MLE with a local approximation of the LEC covariance. We use the naïve estimation $\bar{c} = 1$ for the standard deviation of the EFT expansion coefficients. Meanwhile, in Paper II, we perform a full sampling of a nearly identically specified Bayesian posterior with $\bar{c} = 2.08$. The similarities between these inferences naturally invites a comparison and present us with an opportunity to judge the fidelity of the NN prior in Paper I. Fig. 5.4 shows the two results side-by-side. We see that the position of the mode has shifted somewhat, in particular for $C_{{}^3P_0}$ and $C_{{}^1P_1}$, but the width of the one-dimensional marginal PDFs are overall similar. The latter observation is somewhat counter-intuitive since we have explicitly doubled the size of the EFT truncation error in the HMC sampled posterior and we could thus reasonably expect the posterior to be notably wider. However, this behavior appears quite consistently in all posteriors we have produced: the size $\bar{c}$ of the truncation generally influences the *location* of the PDF to a far greater extent than its *width*. However, inspecting the two-dimensional pairwise PDFs indicates that the volume of the MCMC-sampled posterior (in terms of some multidimensional credible region) is indeed substantially larger. In Paper IV we incorporate $\bar{c}^2$ as a parameter in the HMC sampling of the posterior to investigate whether we have underestimated the size truncation error. We find that this has a small but notable effect on the width of the LEC posterior.

5.3 Tension: πN prior vs. posterior

Throughout this work we employ a Roy-Steiner [77, 78] prior for the πN LECs appearing at NNLO and above, in lieu of incorporating πN scattering data in our likelihood. Assuming that all uncertainties are adequately modeled, the

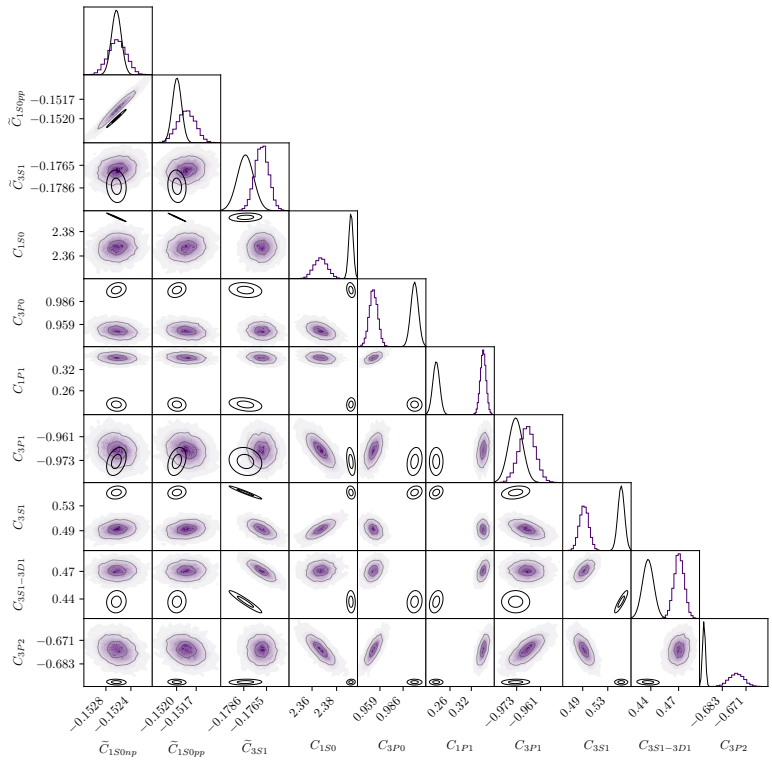
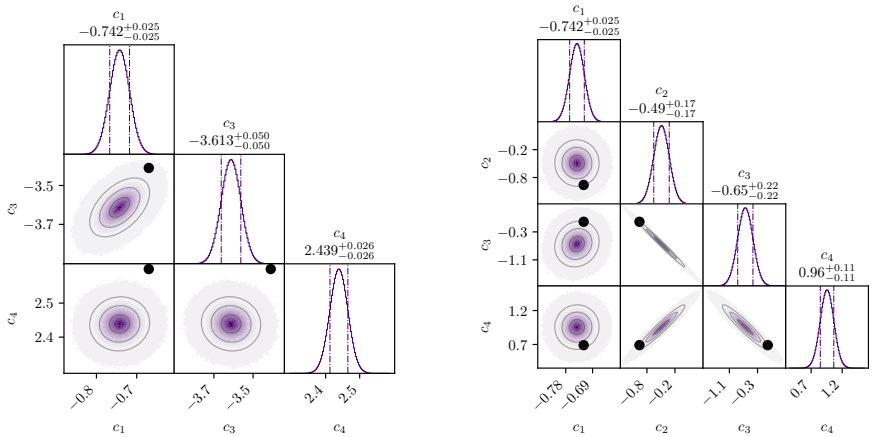


Figure 5.4: Side-by-side comparison between the prior for the NNLO NN contact LECs from Paper I (black contour lines) and the posterior in Paper II (filled purple contours). Only parameters estimated in both projects are included. The LECs are shown in units of $10^4\,\mathrm{GeV}^{-2}$ ($\widetilde{C}$) and $10^4\,\mathrm{GeV}^{-4}$ (C).

posterior for these LECs should considerably overlap with the adopted prior. However, throughout Papers II-IV, this is not the reality we face. We mostly find posteriors for the πN LECs that differ considerably from the priors we adopt, indicating tension between the respective analyses of πN and NN scattering data. We find the greatest overlap at Δ-NNLO in Paper IV, where the discrepancy is relatively minor and the MAP of our NN analysis appears within the 3σ (99.7%) credible region of the bivariate marginals of the Roy-Steiner prior. This is shown in Fig. 5.5. At NNLO and in particular N3LO in Δ-less χEFT, the discrepancy is such that the priors and posteriors are in clear conflict which must be resolved in future work. Note that the Roy-Steiner determination of the πN LECs is less precise in the Δ-full case than in the Δ-less case, due to the relatively unknown $\pi N\Delta$ axial coupling h_A [78]. We do, however, find in Pa-



(a) (b)

Figure 5.5: Roy-Steiner priors and MAP estimates in units of GeV^{-1}. The MAP estimates are shown as black dots. The Roy-Steiner priors are displayed as gray contour lines filled with purple. The contour lines encompass 68, 95, and 99.7 % of the probability mass, respectively. (a) Δ-less NNLO. (b) Δ-full NNLO.

per III that the prior and posterior overlap perfectly at NNLO when only NN scattering data in the range $0 < T_{\mathrm{lab}} < 40.5$ MeV is taken into consideration. This indicates that the conflict likely originates from high-energy data and the need for improved χEFT error models, possibly in both kinds of analyses.

5.4 Inferring three-nucleon forces

In Paper I we use the leverage of emulation of few-body observables to infer a posterior for the 3NF LECs c_D and c_E at NNLO in Δ-less χEFT. We employ conjugate priors for the truncation error parameter $\bar{c}^2$ and the expansion parameter Q, and as data we use a set of $A = 3, 4$ observables: binding energies, radii, and the β-decay half-life of the triton.

The posterior for (c_D, c_E) displays a strong mutual correlation. In addition, we find that the posterior is insensitive to the variablity of the NN LECs. Given our available data, the triton (comparative) half-life $fT_{1/2}$ is essential for a robust inference of both LECs. The other observables only constrain a linear combination of the two LECs (see, e.g., Ref. [17]), while $fT_{1/2}$ constrains c_D but not c_E, as expected [118]. We further find that the EFT truncation error needs

to be taken into account in order to find a consistent fit for the observables we consider. With our modeling of the truncation error, we find that we can produce joint PPDs for the considered observables which are consistent with the experimental values.

5.5 Posterior predictive distributions

In the end, we are interested in PPDs of observables, a selection of which I will now present. All PPDs shown here include sampling of the EFT truncation error. In Paper II, we perform a large-scale investigation into the predictive ability of our model: PPDs at LO-NNLO for *all* observables in the Granada database [46, 47] are provided as supplementary material to that paper. Overall, we find that our predictions are overly confident, i.e., the PPDs are too narrow. We trace the cause to an underestimated variance $\bar{c}^2$ of the EFT truncation error.

We quantify the quality of our predictions by use of empirical coverages, i.e., we systematically compute credible intervals for our predictions and check how frequently these intervals cover the experimental data. Ideally, a (say) 50% credible interval should cover the experimental data 50% of the time. Due to our overly confident predictions, we generally find empirical coverages that fall short of this expectation. However, by improving our estimation of $\bar{c}^2$ we find empirical coverages that indicate overall reliable PPDs. We improve the estimation of $\bar{c}^2$ by learning about the truncation error from all chiral orders available to us, i.e., up to and including N3LO, whereas at first we only learned about $\bar{c}^2$ from orders up to the order at which we made predictions.

Naturally, we carry over our learned lessons into Paper III, where similar but improved inferences are performed. Here, we do not repeat the in-depth model check, but instead focus on learning about the charge dependence of the strong nuclear force via ERE parameters. The ERE is generally given by

$$p \cot(\delta) = -\frac{1}{a} + \frac{1}{2} r p^2 + \mathcal{O}(p^4) \tag{5.1}$$

as $p \to 0$, where δ in our case is the ${}^1S_0$ NN phase shift. Fig. 5.6 shows PPDs for the ${}^1S_0$ scattering lengths a and effective ranges r for nn (left column), np (middle column), and pp (right column) scattering. The scattering length for nn, a_{nn}, serves as a model check as we condition our inference on this empirical datum and it is therefore unsurprising that the modes of all three PPDs coincide with the empirical value. All other PPDs are predictions. We note that all PPDs are consistent with the empirical results, with the possible exception of the N3LO prediction of a_{pp}, where the PPD only overlaps with the edge of the 95% confidence interval of the empirical value. This latter result is in line with the rather poor N3LO predictions for ^{3,4}He, mentioned above. We also note

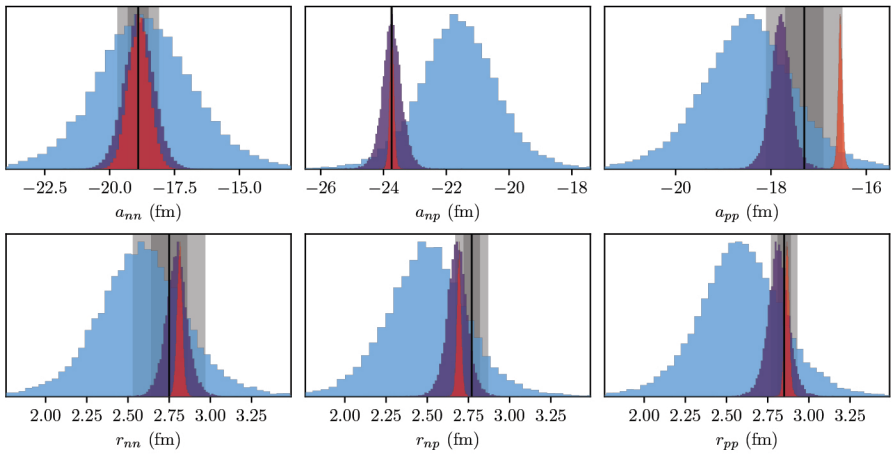


Figure 5.6: PPDs of scattering lengths and effective ranges at NLO (blue), NNLO (purple), and N3LO (red). Empirical results are shown as black lines, with corresponding 1σ (2σ) uncertainties as a dark (light) gray area. The empirical results are without electromagnetic effects and gathered from Ref. [1], except for $a_{^1S_0}^{nn}$ which we take from Ref. [69]. Note that the histograms have been scaled such that they have the same peak height. Originally published in Ref. [119] (Paper III).

that the NNLO and N3LO predictions for a_{np} are in perfect agreement with the empirical result, while the NLO prediction is less accurate, although the PPD still covers the empirical result. Of note is the existence of a total np cross section measured at 1 eV [120] in the data, a datum that corresponds closely with the np scattering length. Evidently, the NNLO and N3LO posteriors are tuned to reproduce this datum well, whereas the NLO posterior is differently balanced. Finally, all three chiral orders are in rather good agreement with the effective ranges r.

In Paper IV, we return to PPDs of elastic scattering cross sections, now equipped with a model for correlated truncation errors and a Δ-full theory. To emphasize the correlation structure we present all PPDs as 500 individual predictions and contrast with predictions from an uncorrelated error model. First, we perform a model check to ascertain that the LEC inference is sane; here, I will exemplify with the spin-correlation observable $A_{zz}(\theta)$ (denoted AZZ in the SAID convention [50]). Fig. 5.7 shows the available experimental data at 67.5 MeV and 260 MeV (note that the inference is conditioned on this data) along with our corresponding NLO PPDs, where the truncation error is (a) correlated,

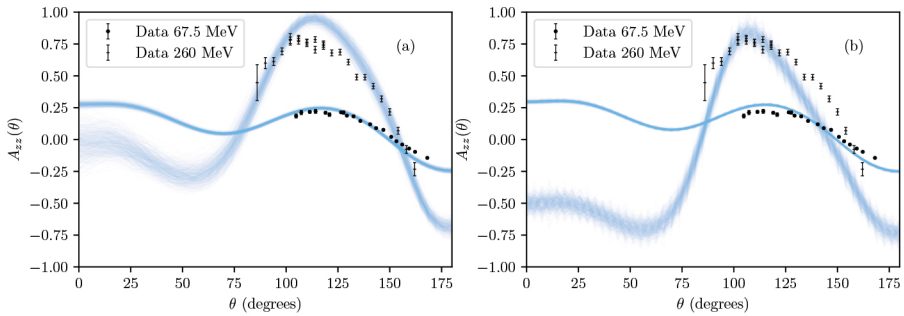


Figure 5.7: Experimental data and corresponding PPDs for the spin-correlation observable $A_{zz}(\theta)$ (AZZ) at 67.5 MeV [122] and 260 MeV [123]. The PPDs are produced at Δ-NLO with (a) a correlated error model and (b) an uncorrelated error model. 500 samples from each PPD are shown.

and (b) uncorrelated. Pay attention to the smoothness of the predictions in (a) compared to the jagged structure in (b). The figure indicates that our inference is overall sound, although the quality of the fit is somewhat better in the correlated case. Note that (a) does a better job of fitting the general features of the data, whereas (b) reproduces a few adjacent data points well at the expense of the rest of the data. The PPDs at 260 MeV show a striking difference below 90 degrees, where in the correlated case the PPD predicts that $A_{zz}(0) \approx 0$ whereas in the uncorrelated case the PPD predicts $A_{zz}(0) \approx -0.5$. Which is more accurate is impossible to say as no experimental data exists in this range. However, asking which inference is "better" or "correct" is not the pertinent question; we appear to have some unaccounted-for uncertainty in at least one of our models, and we need to determine what it is. We include further model-checking in Paper IV, where, for instance, we demonstrate that we take into account that some observables are known by symmetry arguments to be zero at 0 and/or 180 degrees. This constraint—which does not apply to $A_{zz}(\theta)$—is straightforward to impose in a GP model of the truncation error [36, 121].

Fig. 5.8 shows experimental data and PPDs at Δ-NNLO for the unpolarized differential cross section $\sigma(\theta)$ at $T_{\text{lab}} = 304.2$ MeV. The LEC inference is *not* conditioned on this data. As in Fig. 5.7, (a) shows the predictions with a correlated error model while (b) shows the equivalent uncorrelated predictions. In this case, the uncorrelated error model yields predictions that align better with the experimental results. However, it is important to note that two experimental data sets, virtually identical to the data shown in the figure, is present in the data D on which the inference is conditioned. These data sets are measured at $T_{\text{lab}} = 265.8$ MeV and $T_{\text{lab}} = 284.8$ MeV, at the same scattering angles, and by the same experimental group, as the data shown in the figure (see the supple-

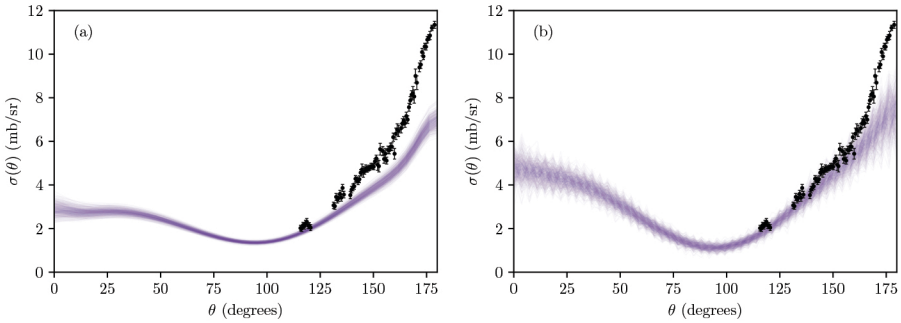


Figure 5.8: Experimental data and corresponding PPDs for the unpolarized differential cross section $\sigma(\theta)$ (DSG) at 304.2 MeV [124]. The PPDs are produced at Δ-NNLO with (a) a correlated error model and (b) an uncorrelated error model. 500 samples from each PPD are shown.

mental material for Paper II). It is reasonable to assume that this large data set strongly influences an uncorrelated error model, where every individual datum is given equal importance. Nevertheless, the only unequivocal conclusion we can draw is that we have underestimated some uncertainty in both cases. This may stem from the rather limited information we have regarding the truncation error in Δ-full χEFT: here, we do not yet have access to N3LO predictions, and we deem the shift from LO to NLO to be unrepresentative of higher-order truncation errors due to the poor quality of LO combined with the comparatively high quality of NLO in Δ-full χEFT. We are thus left with only the (generally small) NLO-NNLO correction to learn about the truncation error. As in the Δ-less case, N3LO in Δ-full χEFT introduces a rather large number of diagrams with corresponding new LECs and it is therefore reasonable to assume that the N3LO correction is greater than we can predict from the NLO-NNLO difference.

Finally, Fig. 5.9 shows PPDs produced with a correlated error model at NLO for all *np* observables at $290 < T_{\text{lab}} \leq 350$ MeV in the Granada database. The observables are organized in alphabetical order (according to the SAID convention), followed by laboratory energy in ascending order. The overall conclusion is that the size of the truncation error is somewhat underestimated.

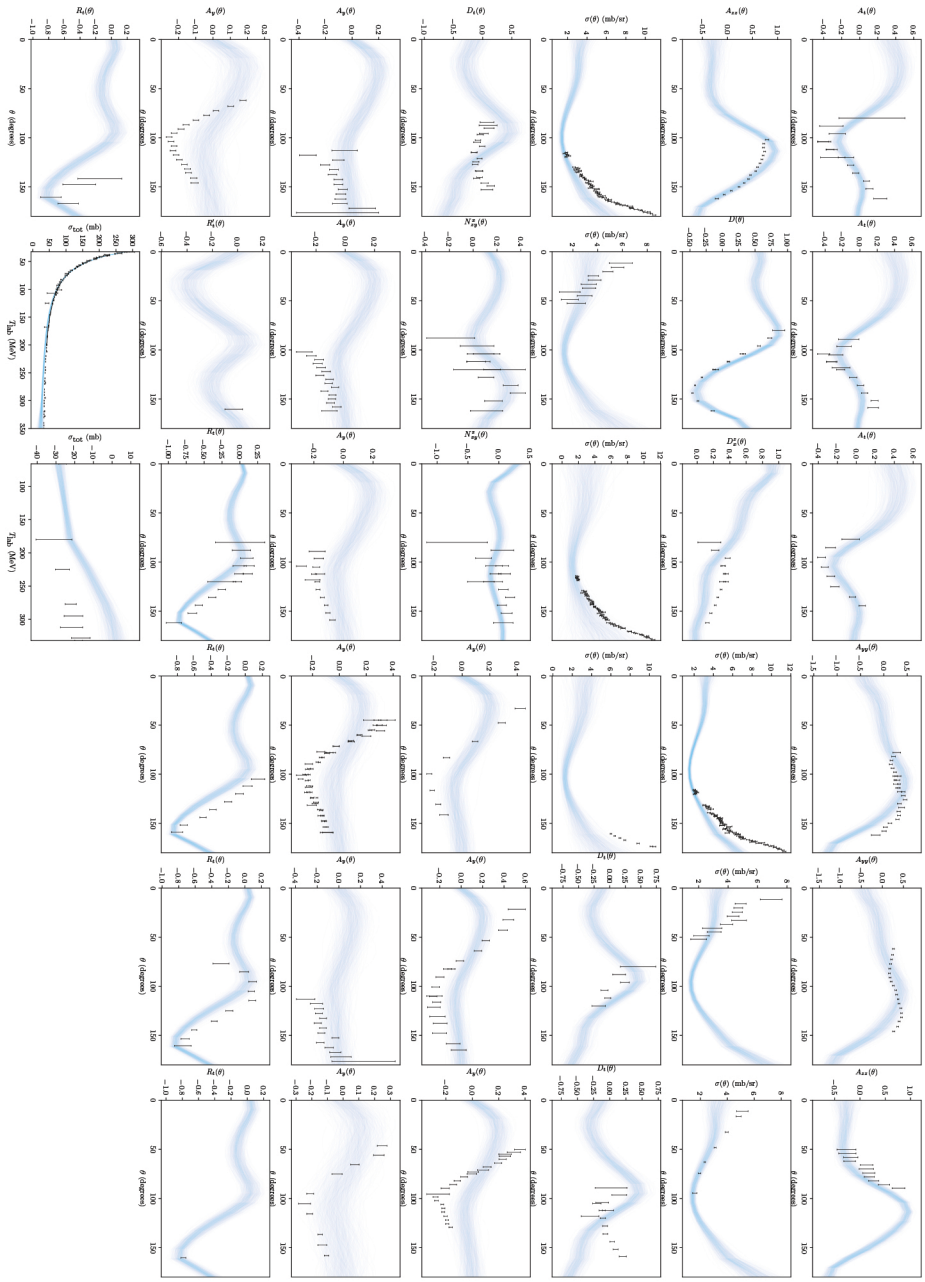


Figure 5.9: Δ-NLO predictions for all $290 < T_{\rm lab} \leq 350$ MeV np elastic-scattering observables in the Granada database [46, 47].

Chapter 6

Outlook

Theoretical predictions of nuclear structure and reaction observables are inherently PDFs, and should where possible be presented as such. In this thesis I extend the reach of UQ in nuclear theory to include LEC inference up to fourth order in χEFT via the HMC [38, 39] sampling method, 3NFs, and correlated models for the EFT truncation error via GPs. I implement fast *np* scattering emulators [41] and use them in conjunction with AD to infer LEC posteriors in the Δ-full variant of χEFT. I compute PPDs of scattering observables both as a model-checking tool and for making predictions of unseen data. The results in this thesis open up multiple enticing avenues for future research, a selection of which I will highlight here.

Simultaneous inference of two- and three-nucleon forces from $A \geq 2$ data. An obvious near-term outlook is to combine the techniques used in this thesis to perform simultaneous inferences of the NN, πN, and NNN LECs. Through emulation of few-nucleon systems and HMC sampling, all technology needed to do this is in place. Such an inference may be conditioned on NNN scattering data since the required emulators are available [107].

Improved error modeling. The behavior of the EFT truncation error is quite complex. While we account for, e.g., correlations across NN scattering energies and angles using GPs and investigate the effect of sampling the variance $\bar{c}^2$ of the truncation error jointly with the LECs, several aspects remain unaccounted for. Future work could model correlated errors across observable types, including different scattering observables and bound-state properties. Since the HMC sampling method is well-equipped to handle high-dimensional parameter spaces, we expect to be able to marginalize over multiple truncation-error hyperparameters—such as correlation lengths and $\bar{c}^2$—as opposed to using MAP estimates. Similarly, the posterior for the EFT expansion parameter Q may be inferred and marginalized over as well.

Advances in MCMC sampling. My implementation of HMC has proved to

be an efficient and reliable tool for sampling LEC posteriors, and its core operational principles have remained unchanged throughout Papers II-IV. However, further developments could certainly be beneficial. At the top of the list of potential improvements we find Riemannian manifold HMC [85], or some similar technology, which could (i) alleviate the user's burden of tuning the mass matrix, and (ii) increase the algorithm's ability to efficiently sample PDFs with complicated shapes. This is important because tuning the mass matrix is a critical part of using HMC. Implementing the No-U-Turn sampler [84] is a further outlook, albeit a less useful one in my opinion.

PPDs of NNN scattering with 3NFs. NNN scattering is not only of interest in the inference stage, but also as a target for predictions. For example, the A_y puzzle (see, e.g., Ref. [125]) could conceivably find its explanation through UQ.

PPDs of light-, medium-, and heavy-mass nuclei. With quantified PDFs for the LECs and the EFT truncation error, the theoretical uncertainty can be propagated to *ab initio* predictions of atomic nuclei. Targets of interest include neutron-rich light elements, ground-state energies of medium-mass nuclei such as isotopes of oxygen, or observables of heavy-mass nuclei such as ^{208}Pb with links to infinite nuclear matter. Open questions remain regarding the modeling of EFT truncation errors for bound states, e.g., how do we define an appropriate low-momentum scale and how do we model discrete, correlated truncation errors?

Bibliography

[1] R. Machleidt and D. R. Entem, "Chiral effective field theory and nuclear forces", Phys. Rept. **503**, 1–75 (2011).

[2] K. G. Wilson, "Confinement of quarks", Phys. Rev. D **10**, 2445–2459 (1974).

[3] Z. Davoudi, W. Detmold, K. Orginos, A. Parreño, M. J. Savage, P. Shanahan, and M. L. Wagman, "Nuclear matrix elements from lattice QCD for electroweak and beyond-Standard-Model processes", Phys. Rept. **900**, 1–74 (2021).

[4] H. A. Bethe, "Nuclear many-body problem", Phys. Rev. **103**, 1353–1390 (1956).

[5] A. Ekström, C. Forssén, G. Hagen, G. R. Jansen, W. Jiang, and T. Papenbrock, "What is ab initio in nuclear theory?", Front. Phys. **11**, 1129094 (2023).

[6] H. Hergert, "A guided tour of *ab initio* nuclear many-body theory", Front. in Phys. **8**, 379 (2020).

[7] M. I. Haftel and F. Tabakin, "Nuclear saturation and the smoothness of nucleon-nucleon potentials", Nucl. Phys. A **158**, 1–42 (1970).

[8] S. B. S. Miller, A. Ekström, and K. Hebeler, "Neutron-deuteron scattering cross sections with chiral *NN* interactions using wave-packet continuum discretization", Phys. Rev. C **106**, 024001 (2022).

[9] S. B. S. Miller, A. Ekström, and C. Forssén, "Posterior predictive distributions of neutron-deuteron cross sections", Phys. Rev. C **107**, 014002 (2023).

[10] E. Epelbaum et al., "Few- and many-nucleon systems with semilocal coordinate-space regularized chiral two- and three-body forces", Phys. Rev. C **99**, 024313 (2019).

[11] P. Maris et al., "Nuclear properties with semilocal momentum-space regularized chiral interactions beyond N2LO", Phys. Rev. C **106**, 064002 (2022).

[12] P. Navratil, G. P. Kamuntavicius, and B. R. Barrett, "Few-nucleon systems in translationally invariant harmonic oscillator basis", Phys. Rev. C **61**, 044001 (2000).

[13] B. R. Barrett, P. Navratil, and J. P. Vary, "Ab initio no core shell model", Prog. Part. Nucl. Phys. **69**, 131–181 (2013).

[14] C. Lanczos, "An iteration method for the solution of the eigenvalue problem of linear differential and integral operators", J. Res. Natl. Bur. Stand. B **45**, 255–282 (1950).

[15] H. Hergert, S. Bogner, T. Morris, A. Schwenk, and K. Tsukiyama, "The in-medium similarity renormalization group: a novel ab initio method for nuclei", Phys. Rept. **621**, 165–222 (2016).

[16] G Hagen, T Papenbrock, M Hjorth-Jensen, and D. J. Dean, "Coupled-cluster computations of atomic nuclei", Rept. Prog. Phys. **77**, 096302 (2014).

[17] B. Hu et al., "Ab initio predictions link the neutron skin of ^{208}Pb to nuclear forces", Nature Phys. **18**, 1196–1200 (2022).

[18] R. Machleidt, "The high precision, charge dependent Bonn nucleon-nucleon potential (CD-Bonn)", Phys. Rev. C **63**, 024001 (2001).

[19] E. Epelbaum, H.-W. Hammer, and U.-G. Meißner, "Modern Theory of Nuclear Forces", Rev. Mod. Phys. **81**, 1773–1825 (2009).

[20] H. W. Hammer, S. König, and U. van Kolck, "Nuclear effective field theory: status and perspectives", Rev. Mod. Phys. **92**, 025004 (2020).

[21] S. Weinberg, "Phenomenological Lagrangians", Physica A **96**, 327–340 (1979).

[22] S. Weinberg, "Nuclear forces from chiral Lagrangians", Phys. Lett. B **251**, 288–292 (1990).

[23] S. Weinberg, "Effective chiral Lagrangians for nucleon-pion interactions and nuclear forces", Nucl. Phys. B **363**, 3–18 (1991).

[24] S. Weinberg, "Three body interactions among nucleons and pions", Phys. Lett. B **295**, 114–121 (1992).

[25] H. Krebs, E. Epelbaum, and U.-G. Meißner, "Nuclear forces with Delta-excitations up to next-to-next-to-leading order. I. Peripheral nucleon-nucleon waves", Eur. Phys. J. A **32**, 127–137 (2007).

[26] A. Nogga, R. G. E. Timmermans, and U. van Kolck, "Renormalization of one-pion exchange and power counting", Phys. Rev. C **72**, 054006 (2005).

[27] H. W. Grießhammer, "Assessing theory uncertainties in eft power countings from residual cutoff dependence", PoS **CD15**, 104 (2016).

[28] C. Yang, A. Ekström, C. Forssén, and G. Hagen, "Power counting in chiral effective field theory and nuclear binding", Phys. Rev. C **103**, 054304 (2021).

[29] O. Thim, E. May, A. Ekström, and C. Forssén, "Bayesian analysis of χEFT at leading order in a modified Weinberg power counting", (2023), arXiv:2302.12624 [nucl-th].

[30] C. Ordóñez and U. van Kolck, "Chiral lagrangians and nuclear forces", Phys. Lett. B **291**, 459–464 (1992).

[31] M. R. Schindler and D. R. Phillips, "Bayesian methods for parameter estimation in effective field theories", Annals Phys. **324**, 682–708 (2009).

[32] R. J. Furnstahl, N. Klco, D. R. Phillips, and S. Wesolowski, "Quantifying truncation errors in effective field theory", Phys. Rev. C **92**, 024005 (2015).

[33] S. Wesolowski, N. Klco, R. J. Furnstahl, D. R. Phillips, and A. Thapaliya, "Bayesian parameter estimation for effective field theories", J. Phys. G **43**, 074001 (2016).

[34] J. A. Melendez, S. Wesolowski, and R. J. Furnstahl, "Bayesian truncation errors in chiral effective field theory: nucleon-nucleon observables", Phys. Rev. C **96**, 024003 (2017).

[35] S. Wesolowski, R. J. Furnstahl, J. A. Melendez, and D. R. Phillips, "Exploring Bayesian parameter estimation for chiral effective field theory using nucleon–nucleon phase shifts", J. Phys. G **46**, 045102 (2019).

[36] J. A. Melendez, R. J. Furnstahl, D. R. Phillips, M. T. Pratola, and S. Wesolowski, "Quantifying correlated truncation errors in effective field theory", Phys. Rev. C **100**, 044001 (2019).

[37] J. Brynjarsdóttir and A. O'Hagan, "Learning about physical parameters: the importance of model discrepancy", Inverse Probl. **30**, 114007 (2014).

[38] S. Duane, A. Kennedy, B. J. Pendleton, and D. Roweth, "Hybrid Monte Carlo", Phys. Lett. B **195**, 216–222 (1987).

[39] R. M. Neal, "MCMC using Hamiltonian dynamics", in *Handbook of Markov Chain Monte Carlo*, edited by S. Brooks, A. Gelman, G. L. Jones, and X.-L. Meng (CRC Press, Boca Raton, 2011), pp. 113 –162.

[40] R. J. Furnstahl, A. J. Garcia, P. J. Millican, and X. Zhang, "Efficient emulators for scattering using eigenvector continuation", Phys. Lett. B **809**, 135719 (2020).

[41] J. A. Melendez, C. Drischler, A. J. Garcia, R. J. Furnstahl, and X. Zhang, "Fast & accurate emulation of two-body scattering observables without wave functions", Phys. Lett. B **821**, 136608 (2021).

[42] W Glöckle, *The quantum mechanical few-body problem*, Texts and monographs in physics (Springer, Berlin, 1983).

[43] C. Van Der Leun and C. Alderliesten, "The deuteron binding energy", Nucl. Phys. A **380**, 261–269 (1982).

[44] J. R. Taylor, *Scattering theory: the quantum theory of nonrelativistic collisions* (Courier Corporation, 2006).

[45] J. J. Sakurai and J. Napolitano, *Modern quantum mechanics* (Cambridge University Press, 2020).

[46] R. N. Pérez, J. E. Amaro, and E. Ruiz Arriola, "Partial-wave analysis of nucleon-nucleon scattering below the pion-production threshold", Phys. Rev. C **88**, 024002 (2013).

[47] R. N. Pérez, J. E. Amaro, and E. Ruiz Arriola, "Coarse-grained potential analysis of neutron-proton and proton-proton scattering below the pion production threshold", Phys. Rev. C **88**, 064002 (2013).

[48] J. Bystricky, F. Lehar, and P. Winternitz, "Formalism of nucleon-nucleon elastic scattering experiments", Journal de Physique **39**, 1–32 (1978).

[49] V. G. J. Stoks, R. A. M. Klomp, M. C. M. Rentmeester, and J. J. de Swart, "Partial-wave analysis of all nucleon-nucleon scattering data below 350 MeV", Phys. Rev. C **48**, 792–815 (1993).

[50] *INS data analysis center: SAID*, http://gwdac.phys.gwu.edu, Accessed: 2021-03-29.

[51] N. Hoshizaki, "Formalism of nucleon-nucleon scattering", Prog. Theor. Phys. Suppl. **42**, 107–159 (1969).

[52] D. R. Entem and R. Machleidt, "Accurate charge dependent nucleon-nucleon potential at fourth order of chiral perturbation theory", Phys. Rev. C **68**, 041001(R) (2003).

[53] A. Ekström, G. R. Jansen, K. A. Wendt, G. Hagen, T. Papenbrock, B. D. Carlsson, C. Forssén, M. Hjorth-Jensen, P. Navrátil, and W. Nazarewicz, "Accurate nuclear radii and binding energies from a chiral interaction", Phys. Rev. C **91**, 051301(R) (2015).

[54] B. D. Carlsson, A. Ekström, C. Forssén, D. F. Strömberg, G. R. Jansen, O. Lilja, M. Lindby, B. A. Mattsson, and K. A. Wendt, "Uncertainty analysis and order-by-order optimization of chiral nuclear interactions", Phys. Rev. X **6**, 011019 (2016).

[55] G. Hagen et al., "Neutron and weak-charge distributions of the ^{48}Ca nucleus", Nature Phys. **12**, 186–190 (2015).

[56] V. Lapoux, V. Somà, C. Barbieri, H. Hergert, J. D. Holt, and S. R. Stroberg, "Radii and binding energies in oxygen isotopes: a challenge for nuclear forces", Phys. Rev. Lett. **117**, 052501 (2016).

[57] M. Piarulli, L. Girlanda, R. Schiavilla, A. Kievsky, A. Lovato, L. E. Marcucci, S. C. Pieper, M. Viviani, and R. B. Wiringa, "Local chiral potentials with Δ-intermediate states and the structure of light nuclei", Phys. Rev. C **94**, 054007 (2016).

[58] M. Piarulli, L. Girlanda, R. Schiavilla, R. N. Pérez, J. E. Amaro, and E. Ruiz Arriola, "Minimally nonlocal nucleon-nucleon potentials with chiral two-pion exchange including Δ resonances", Phys. Rev. C **91**, 024003 (2015).

[59] P. Reinert, H. Krebs, and E. Epelbaum, "Semilocal momentum-space regularized chiral two-nucleon potentials up to fifth order", Eur. Phys. J. A **54**, 86 (2018).

[60] T. Djärv, A. Ekström, C. Forssén, and H. T. Johansson, "Bayesian predictions for $A = 6$ nuclei using eigenvector continuation emulators", Phys. Rev. C **105**, 014005 (2022).

[61] W. Jiang and C. Forssén, "Bayesian probability updates using sampling/importance resampling: applications in nuclear theory", Front. in Phys. **10**, 1058809 (2022).

[62] T. Miyagi, S. R. Stroberg, P. Navrátil, K. Hebeler, and J. D. Holt, "Converged ab initio calculations of heavy nuclei", Phys. Rev. C **105**, 014302 (2022).

[63] D. Frame, R. He, I. Ipsen, D. Lee, D. Lee, and E. Rrapaj, "Eigenvector continuation with subspace learning", Phys. Rev. Lett. **121**, 032501 (2018).

[64] S. König, A. Ekström, K. Hebeler, D. Lee, and A. Schwenk, "Eigenvector Continuation as an Efficient and Accurate Emulator for Uncertainty Quantification", Phys. Lett. B **810**, 135814 (2020).

[65] A. Ekström and G. Hagen, "Global sensitivity analysis of bulk properties of an atomic nucleus", Phys. Rev. Lett. **123**, 252501 (2019).

[66] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Rubin, A. Vehtari, and D. B. Dunson, *Bayesian data analysis*, Texts in statistical science series (Chapman & Hall/CRC, 2014).

[67] P. Gregory, *Bayesian logical data analysis for the physical sciences: a comparative approach with Mathematica® support* (Cambridge University Press, 2005).

[68] U. van Kolck, "Isospin violation in low-energy hadronic physics", Few Body Syst. Suppl. **9**, 444–448 (1995).

[69] A. Gårdestig, "Extracting the neutron-neutron scattering length—recent developments", J. Phys. G **36**, 053001 (2009).

[70] V. G. J. Stoks and J. J. De Swart, "The magnetic moment interaction in nucleon-nucleon phase shift analyses", Phys. Rev. C **42**, 1235–1248 (1990).

[71] J. Dobaczewski, W. Nazarewicz, and P. Reinhard, "Error estimates of theoretical models: a guide", J. Phys. G **41**, 074001 (2014).

[72] D. Bertsekas, *Nonlinear programming*, Athena scientific optimization and computation series (Athena Scientific, 2016).

[73] R. B. Wiringa, V. G. J. Stoks, and R. Schiavilla, "An accurate nucleon-nucleon potential with charge independence breaking", Phys. Rev. C **51**, 38–51 (1995).

[74] A. Ekström et al., "Optimized chiral nucleon-nucleon interaction at next-to-next-to-leading order", Phys. Rev. Lett. **110**, 192502 (2013).

[75] E. Epelbaum, H. Krebs, and U.-G. Meißner, "Precision nucleon-nucleon potential at fifth order in the chiral expansion", Phys. Rev. Lett. **115**, 122301 (2015).

[76] C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning* (The MIT Press, 2005).

[77] M. Hoferichter, J. Ruiz de Elvira, B. Kubis, and U.-G. Meißner, "Roy-Steiner-equation analysis of pion-nucleon scattering", Phys. Rept. **625**, 1–88 (2016).

[78] D. Siemens, J. Ruiz de Elvira, E. Epelbaum, M. Hoferichter, H. Krebs, B. Kubis, and U.-G. Meißner, "Reconciling threshold and subthreshold expansions for pion-nucleon scattering", Phys. Lett. B **770**, 27–34 (2017).

[79] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines", J. Chem. Phys. **21**, 1087–1092 (1953).

[80] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications", Biometrika **57**, 97–109 (1970).

[81] S. Brooks, A. Gelman, G. Jones, and X.-L. Meng, eds., *Handbook of Markov Chain Monte Carlo* (Chapman and Hall/CRC Press, 2011).

[82] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, "Emcee: the MCMC hammer", PASP **125**, 306–312 (2013).

[83] J. Goodman and J. Weare, "Ensemble samplers with affine invariance", Comm. App. Math. Com. Sc. **5**, 65–80 (2010).

[84] M. D. Hoffman and A. Gelman, "The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo", J. Mach. Learn. Res. **15**, 1593–1623 (2014).

[85] M. Girolami, B. Calderhead, and S. A. Chin, *Riemannian manifold Hamiltonian Monte Carlo*, 2009, arXiv:0907.1100 [stat.CO].

[86] A. Griewank, "A mathematical view of automatic differentiation", Acta Numer. **12**, 321–398 (2003).

[87] I. Charpentier and J. Utke, "Fast higher-order derivative tensors with rapsodia", Optim. Method. Softw. **24**, 1–14 (2009).

[88] M. Betancourt, *A conceptual introduction to Hamiltonian Monte Carlo*, 2017, arXiv:1701.02434 [stat.ME].

[89] J. M. Hammersley and K. W. Morton, "A new Monte Carlo technique: antithetic variates", Math. Proc. Camb. Philos. Soc. **52**, 449–475 (1956).

[90] I. Svensson, A. Ekström, and C. Forssén, "Bayesian parameter estimation in chiral effective field theory using the Hamiltonian Monte Carlo method", Phys. Rev. C **105**, 014004 (2022).

[91] C. G. Broyden, "The convergence of a class of double-rank minimization algorithms 1. General considerations", J. Inst. Math. Appl. **6**, 76–90 (1970).

[92] R. Fletcher, "A new approach to variable metric algorithms", Comp. J. **13**, 317–322 (1970).

[93] D. Goldfarb, "A family of variable-metric methods derived by variational means", Math. Comp. **24**, 23–26 (1970).

[94] D. F. Shanno, "Conditioning of quasi-Newton methods for function minimization", Math. Comp. **24**, 647–656 (1970).

[95] A. Sokal, "Monte Carlo methods in statistical mechanics: foundations and new algorithms", in *Functional integration: basics and applications*, edited by C. DeWitt-Morette, P. Cartier, and A. Folacci (Springer US, Boston, MA, 1997), pp. 131–192.

[96] J. Skilling, "Nested sampling", AIP Conf. Proc. **735**, 395–405 (2004).

[97] R. H. Swendsen and J.-S. Wang, "Replica Monte Carlo simulation of spin-glasses", Phys. Rev. Lett. **57**, 2607–2609 (1986).

[98] C. J. Geyer, "Markov chain Monte Carlo maximum likelihood", in Comp. Sci. Stat.: Proceedings of the 23rd symposium interface (1991).

[99] V. Roy, "Convergence diagnostics for Markov chain Monte Carlo", Annu. Rev. Stat. Appl. **7**, 387–412 (2020).

[100] A. Gelman and D. B. Rubin, "Inference from iterative simulation using multiple sequences", Statist. Sci. **7**, 457 –472 (1992).

[101] S. P. Brooks and A. Gelman, "General methods for monitoring convergence of iterative simulations", J. Comput. Graphi. Stat. **7**, 434–455 (1998).

[102] A. F. M. Smith and A. E. Gelfand, "Bayesian statistics without tears: a sampling–resampling perspective", The American Statistician **46**, 84–88 (1992).

[103] D. Rubin, "Using the SIR algorithm to simulate posterior distributions", in *Bayesian statistics 3*, edited by J. M. Bernardo, M. H. DeGroot, D. V. Lindley, and A. F. M. Smith (Oxford University Press, 1988).

[104] E. Epelbaum, H. Krebs, and U.-G. Meißner, "Improved chiral nucleon-nucleon potential up to next-to-next-to-next-to-leading order", Eur. Phys. J. A **51**, 53 (2015).

[105] W. G. Jiang, A. Ekström, C. Forssén, G. Hagen, G. R. Jansen, and T. Papenbrock, "Accurate bulk properties of nuclei from $A = 2$ to ∞ from potentials with Δ isobars", Phys. Rev. C **102**, 054301 (2020).

[106] J. Oakley and A. O'Hagan, "Bayesian inference for the uncertainty distribution of computer model outputs", Biometrika **89**, 769–784 (2002).

[107] X. Zhang and R. J. Furnstahl, "Fast emulation of quantum three-body scattering", Phys. Rev. C **105**, 064004 (2022).

[108] C. Drischler, M. Quinonez, P. G. Giuliani, A. E. Lovell, and F. M. Nunes, "Toward emulating nuclear reactions using eigenvector continuation", Phys. Lett. B **823**, 136777 (2021).

[109] J. A. Melendez, C. Drischler, R. J. Furnstahl, A. J. Garcia, and X. Zhang, "Model reduction methods for nuclear emulators", J. Phys. G **49**, 102001 (2022).

[110] W. Kohn, "Density functional and density matrix method scaling linearly with the number of atoms", Phys. Rev. Lett. **76**, 3168–3171 (1996).

[111] R. Newton, *Scattering theory of waves and particles*, Dover Books on Physics (Dover Publications, 2002).

[112] G. Van Rossum and F. L. Drake, *Python 3 reference manual* (CreateSpace, Scotts Valley, CA, 2009).

[113] C. R. Harris et al., "Array programming with NumPy", Nature **585**, 357 (2020).

[114] M. D. McKay, R. J. Beckman, and W. J. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code", Technometrics **21**, 239–245 (1979).

[115] E. Epelbaum, W. Glöckle, and U.-G. Meißner, "The two-nucleon system at next-to-next-to-next-to-leading order", Nucl. Phys. A **747**, 362 –424 (2005).

[116] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, *JAX: Composable transformations of Python+NumPy programs*, version 0.3.13, 2018.

[117] The JAX authors, *JAX: High-performance array computing - JAX documentation*, https://jax.readthedocs.io/en/latest/index.html, 2023.

[118] D. Gazit, S. Quaglioni, and P. Navratil, "Three-Nucleon Low-Energy Constants from the Consistency of Interactions and Currents in Chiral Effective Field Theory", Phys. Rev. Lett. **103**, [Erratum: Phys.Rev.Lett. 122, 029901 (2019)], 102502 (2009).

[119] I. Svensson, A. Ekström, and C. Forssén, "Bayesian estimation of the low-energy constants up to fourth order in the nucleon-nucleon sector of chiral effective field theory", Phys. Rev. C **107**, 014001 (2023).

[120] E. Lomon and R. Wilson, "Neutron-proton scattering at a few MeV", Phys. Rev. C **9**, 1329–1335 (1974).

[121] I. Vernon, S. E. Jackson, and J. A. Cumming, "Known boundary emulation of complex computer models", SIAM/ASA Journal on Uncertainty Quantification **7**, 838–876 (2019).

[122] Hammans, Brogli-Gysin, Burzynski, Campbell, Haffter, Henneck, Lorenzon, Pickar, Sick, Konter, Mango, and van den Brandt, "Neutron-proton spin-correlation parameter AZZ at 68 MeV", Phys. Rev. Lett. **66**, 2293–2296 (1991).

[123] J. Arnold et al., "Measurement of spin observables in neutron-proton elastic scattering", Eur. Phys. J. C **17**, 1434 (2000).

[124] B. E. Bonner, J. E. Simmons, C. L. Hollas, C. R. Newsom, P. J. Riley, G. Glass, and M. Jain, "Measurement of *np* charge exchange for neutron energies 150-800 MeV", Phys. Rev. Lett. **41**, 1200–1203 (1978).

[125] D. Huber and J. L. Friar, "The A_y puzzle and the nuclear force", Phys. Rev. C **58**, 674–685 (1998).