



**CHALMERS**  
UNIVERSITY OF TECHNOLOGY

## **Bayesian probability updates using sampling/importance resampling: Applications in nuclear theory**

Downloaded from: <https://research.chalmers.se>, 2022-12-10 10:53 UTC

Citation for the original published paper (version of record):

Jiang, W., Forssén, C. (2022). Bayesian probability updates using sampling/importance resampling: Applications in nuclear theory. *Frontiers in Physics*, 10. <http://dx.doi.org/10.3389/fphy.2022.1058809>

N.B. When citing this work, cite the original published paper.



## OPEN ACCESS

EDITED BY  
Michele Viviani,  
National Institute of Nuclear Physics of  
Pisa, Italy

REVIEWED BY  
Saori Pastore,  
Washington University in St. Louis,  
United States  
Bijaya Acharya,  
Oak Ridge National Laboratory (DOE),  
United States

\*CORRESPONDENCE  
Weiguang Jiang,  
wjiang@uni-mainz.de

SPECIALTY SECTION  
This article was submitted to Nuclear  
Physics,  
a section of the journal  
Frontiers in Physics

RECEIVED 30 September 2022  
ACCEPTED 20 October 2022  
PUBLISHED 03 November 2022

CITATION  
Jiang W and Forssén C (2022), Bayesian  
probability updates using sampling/  
importance resampling: Applications in  
nuclear theory.  
*Front. Phys.* 10:1058809.  
doi: 10.3389/fphy.2022.1058809

COPYRIGHT  
© 2022 Jiang and Forssén. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution  
or reproduction is permitted which does  
not comply with these terms.

# Bayesian probability updates using sampling/importance resampling: Applications in nuclear theory

Weiguang Jiang\* and Christian Forssén

Department of Physics, Chalmers University of Technology, Göteborg, Sweden

We review an established Bayesian sampling method called sampling/importance resampling and highlight situations in nuclear theory when it can be particularly useful. To this end we both analyse a toy problem and demonstrate realistic applications of importance resampling to infer the posterior distribution for parameters of  $\Delta$ NNLO interaction model based on chiral effective field theory and to estimate the posterior probability distribution of target observables. The limitation of the method is also showcased in extreme situations where importance resampling breaks.

## KEYWORDS

bayesian inference, probability updates, importance resampling, uncertainty quantification, *ab initio* nuclear theory, low-energy constants

## 1 Introduction

Bayesian inference is an appealing approach for dealing with theoretical uncertainties and has been applied in different nuclear physics studies [1–16]. In the practice of Bayesian analyses, a sampling procedure is usually inevitable for approximating the posterior probability distribution of model parameters and for performing predictive computations. Various Markov chain Monte Carlo (MCMC) methods [17–21] are often used for this purpose, even for complicated models with high-dimensional parameter spaces. However, MCMC sampling typically requires many likelihood evaluations, which is often a costly operation in nuclear theory, and there is a need to explore other sampling techniques. In this paper, we review an established method called sampling/importance resampling (S/IR) [22–24] and demonstrate its use in realistic nuclear physics applications where we also perform comparisons with MCMC sampling.

In recent years, there has been an increasing demand for precision nuclear theory. This implies a challenge to not just achieve accurate theoretical predictions but also to quantify accompanying uncertainties. The use of *ab initio* many-body methods and nuclear interaction models based on chiral effective field theory ( $\chi$ EFT) has shown a potential to describe finite nuclei and nuclear matter based on extant experimental data (e.g. nucleon-nucleon scattering, few-body sector) with controlled approximations [25–29]. The interaction model is parametrized in terms of low-energy constants (LECs), the number of which is growing order-by-order according to the rules of a

corresponding power counting [30–32]. Very importantly, the systematic expansion allows to quantify the truncation error and to incorporate this knowledge in the analysis [4–6, 10–14]. Indeed, Bayesian inference is an excellent framework to incorporate different sources of uncertainty and to propagate error bars to the model predictions. Starting from Bayes’ theorem

$$\text{pr}(\boldsymbol{\theta}|\mathcal{D}) \propto \mathcal{L}(\boldsymbol{\theta})\text{pr}(\boldsymbol{\theta}), \tag{1}$$

where  $\text{pr}(\boldsymbol{\theta}|\mathcal{D})$  is the posterior probability density function (PDF) for the vector  $\boldsymbol{\theta}$  of LECs (conditional on the data  $\mathcal{D}$ ),  $\mathcal{L}(\boldsymbol{\theta}) \equiv \text{pr}(\mathcal{D}|\boldsymbol{\theta})$  is the likelihood and  $\text{pr}(\boldsymbol{\theta})$  is the prior. Then for any model prediction one needs to evaluate the expectation value of a function of interest  $y(\boldsymbol{\theta})$  (target observables) according to the posterior. This involves integrals such as

$$\int d\boldsymbol{\theta} y(\boldsymbol{\theta})\text{pr}(\boldsymbol{\theta}|\mathcal{D}), \tag{2}$$

which can not be analytically solved for realistic cases. Fortunately, integrals such as Eq. 2 can be approximately evaluated using a finite set of samples  $\{\boldsymbol{\theta}_i\}_{i=1}^N$  from  $\text{pr}(\boldsymbol{\theta}|\mathcal{D})$ . MCMC sampling methods are the main computational tool for providing such samples, even for high-dimensional parameter volumes [16]. However the use of MCMC in nuclear theory typically requires massive computations to record sufficiently many samples from the Markov chain. There are certainly situations where MCMC sampling is not ideal, or even becomes infeasible:

- 1) When the posterior is conditioned on some calibration data for which our model evaluations are very costly. Then we might only afford a limited number of full likelihood evaluations and our MCMC sampling becomes less likely to converge.
- 2) Bayesian posterior updates in which calibration data is added in several different stages. This typically requires that the MCMC sampling must be carried out repeatedly from scratch.
- 3) Model checking where we want to explore the sensitivity to prior assignments. This is a second example of posterior updating.
- 4) The prediction of target observables for which our model evaluations become very costly and the handling of a large number of MCMC samples becomes infeasible.

These are situations where one might want to use the S/IR method [23, 24], which can exploit the previous results of model evaluations to allow posterior probability updates at a much smaller computational cost compared to the full MCMC method. In the following sections we first review the S/IR method and then present both toy and realistic applications in which its performance is compared with full MCMC sampling. Finally, we illustrate limitations of the method by considering cases where S/IR fails and we highlight the importance of the so-called effective number of samples. More difficult scenarios, in

which the method fails without a clear warning, are left for the concluding remarks.

## 2 Sampling/importance resampling

The basic idea of S/IR is to utilize the inherent duality between samples and the density (probability distribution) from which they were generated [23]. This duality offers an opportunity to indirectly recreate a density (that might be hard to compute) from samples that are easy to obtain. Here we give a brief review of the method and illustrate with a toy problem.

Let us consider a target density  $h(\boldsymbol{\theta})$ . In our applications this target will be the posterior PDF  $\text{pr}(\boldsymbol{\theta}|\mathcal{D})$  from Eq. 1. Instead of attempting to directly collect samples from  $h(\boldsymbol{\theta})$ , as would be the goal in MCMC approaches, the S/IR method uses a detour. We first obtain samples from a simple (even analytic) density  $g(\boldsymbol{\theta})$ . We then resample from this finite set using a resampling algorithm to approximately recreate samples from the target density  $h(\boldsymbol{\theta})$ . There are (at least) two different resampling methods. In this paper we only focus on one of them called weighted bootstrap (more details of resampling methods can be found in Refs. [22, 23]).

Assuming we are interested in the target density  $h(\boldsymbol{\theta}) = f(\boldsymbol{\theta}) / \int f(\boldsymbol{\theta}) d\boldsymbol{\theta}$ , the procedure of resampling *via* weighted bootstrap can be summarized as follows:

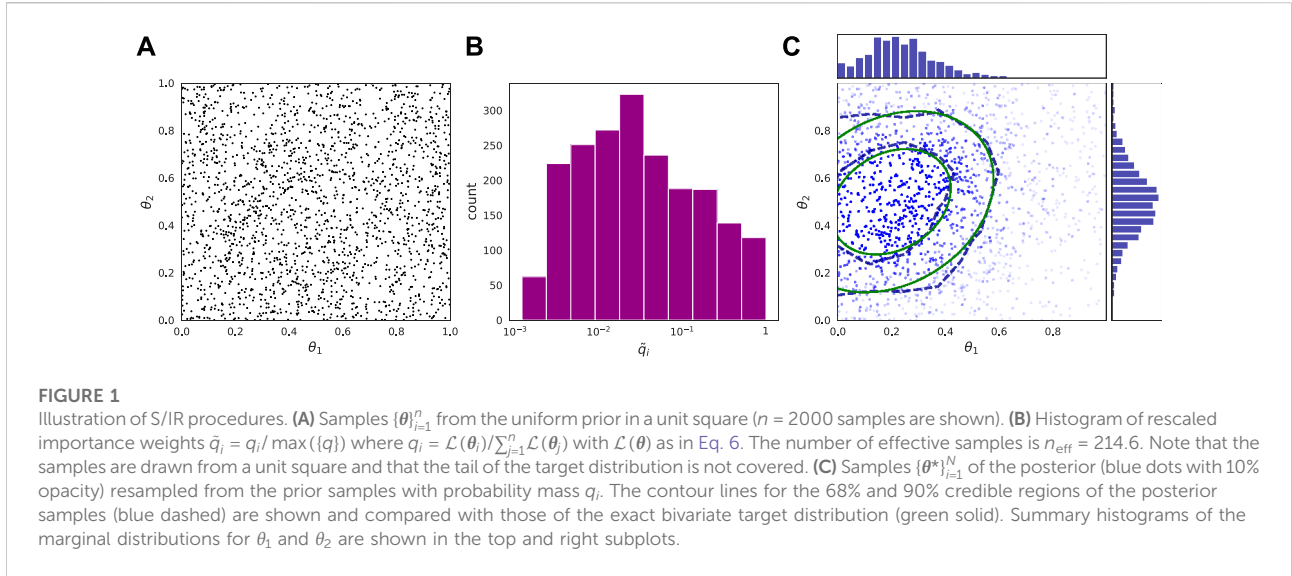
- 1) Generate the set  $\{\boldsymbol{\theta}_i\}_{i=1}^n$  of samples from a sampling density  $g(\boldsymbol{\theta})$ .
- 2) Calculate  $\omega_i = f(\boldsymbol{\theta}_i) / g(\boldsymbol{\theta}_i)$  for the  $n$  samples and define importance weights as:  $q_i = \omega_i / \sum_{j=1}^n \omega_j$ .
- 3) Draw  $N$  new samples  $\{\boldsymbol{\theta}_i^*\}_{i=1}^N$  from the discrete distribution  $\{\boldsymbol{\theta}_i\}_{i=1}^n$  with probability mass  $q_i$  on  $\boldsymbol{\theta}_i$ .
- 4) The set of samples  $\{\boldsymbol{\theta}_i^*\}_{i=1}^N$  will then be approximately distributed according to the target density  $h(\boldsymbol{\theta})$ .

Intuitively, the distribution of  $\boldsymbol{\theta}^*$  should be good approximation of  $h(\boldsymbol{\theta})$  when  $n$  is large enough. Here we justify this claim *via* the cumulative distribution function of  $\theta^*$  (for the one-dimensional case)

$$\begin{aligned} \text{pr}(\theta^* \leq a) &= \sum_{i=1}^n q_i \cdot H(a - \theta_i) = \frac{1}{n} \frac{\sum_{i=1}^n \omega_i \cdot H(a - \theta_i)}{\sum_{i=1}^n \omega_i} \\ n \rightarrow \infty \frac{\mathbb{E}_g \left[ \frac{f(\theta)}{g(\theta)} \cdot H(a - \theta) \right]}{\mathbb{E}_g \left[ \frac{f(\theta)}{g(\theta)} \right]} &= \frac{\int_{-\infty}^a f(\theta) d\theta}{\int_{-\infty}^{\infty} f(\theta) d\theta} = \int_{-\infty}^a h(\theta) d\theta, \end{aligned} \tag{3}$$

with  $\mathbb{E}_g[X(\theta)] = \int_{-\infty}^{\infty} X(\theta)g(\theta) d\theta$  the expectation value of  $X(\theta)$  with respect to  $g(\theta)$ , and  $H$  Heaviside step function such that

$$H(a - \theta) = \begin{cases} 1 & \text{if } \theta \leq a, \\ 0 & \text{if } \theta > a. \end{cases} \tag{4}$$



The above resampling method can be applied to generate samples from the posterior PDF  $h(\theta) = \text{pr}(\theta|\mathcal{D})$  in a Bayesian analysis. It remains to choose a sampling distribution,  $g(\theta)$ , which in principle could be any continuous density distribution. However, recall that  $h(\theta)$  can be expressed in terms of an unnormalized distribution  $f(\theta)$ , and using Bayes' theorem 1) we can set  $f(\theta) = \mathcal{L}(\theta)\text{pr}(\theta)$ . Thus, choosing the prior  $\text{pr}(\theta)$  as the sampling distribution  $g(\theta)$  we find that the importance weights are expressed in terms of the likelihood,  $q_i = \mathcal{L}(\theta_i) / \sum_{j=1}^n \mathcal{L}(\theta_j)$ . Assuming that it is simple to collect samples from the prior, the costly operation will be the evaluation of  $\mathcal{L}(\theta_i)$ . Here we make the side remark that an effective and computationally cost-saving approximation can be made if we manage to perform a pre-screening to identify (and ignore) samples that will give a very small importance weight. We also note that the above choice of  $g(\theta) = \text{pr}(\theta)$  is purely for

simplicity and one can perform importance resampling with any  $g(\theta)$ .

In Figure 1 we follow the above procedure and give a simple example of S/IR to illustrate how to get samples from a posterior distribution. We consider a two-dimensional parametric model with  $\theta = (\theta_1, \theta_2)$ . Given data  $\mathcal{D}$  obtained under the model we have:

$$\text{pr}(\theta_1, \theta_2|\mathcal{D}) = \frac{\mathcal{L}(\theta_1, \theta_2)\text{pr}(\theta_1, \theta_2)}{\iint \mathcal{L}(\theta_1, \theta_2)\text{pr}(\theta_1, \theta_2) d\theta_1 d\theta_2} \quad (5)$$

For simplicity and illustration, the joint prior distribution for  $\theta_1, \theta_2$  is set to be uniform over the unit square as shown in Figure 1A. In this example we also assume that the data  $\mathcal{D}$  follows a multivariate Student-t distribution such that the likelihood function is

$$\mathcal{L}(\theta_1, \theta_2) = \frac{\Gamma[(\nu + p)/2]}{\Gamma(\nu/2)^{\nu/p/2} \pi^{p/2} |\Sigma|^{1/2}} \left[ 1 + \frac{1}{\nu} (\theta - \mu)^T \Sigma^{-1} (\theta - \mu) \right]^{-(\nu+p)/2}, \quad (6)$$

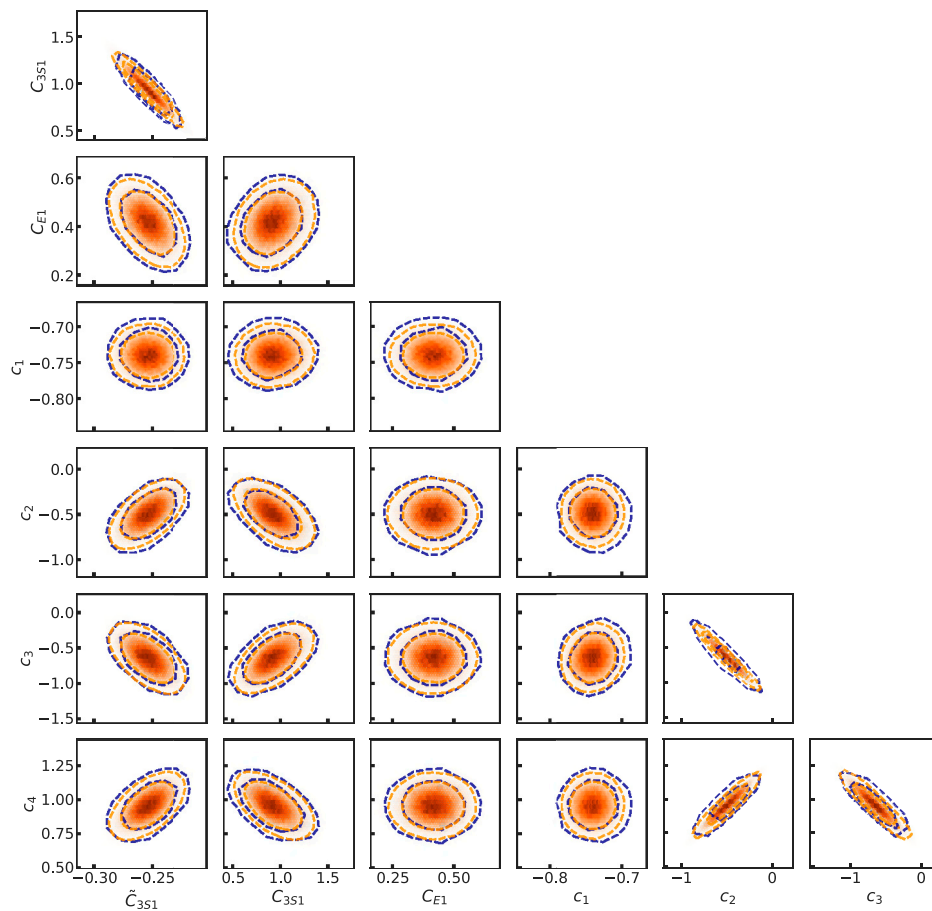
where the dimension  $p = 2$ , the degrees of freedom  $\nu = 2$ , the mean vector  $\mu = (0.2, 0.5)$  and the scale matrix  $\Sigma = [[0.02, 0.005], [0.005, 0.02]]$ .

The importance weights  $q_i$  are then computed for  $n = 2000$  samples drawn from the prior (these prior samples are shown in Figure 1A). The resulting histogram of importance weights is shown in Figure 1B. Here the weights have been rescaled as  $\tilde{q}_i = q_i / \max(\{q\})$  such that the sample with the largest probability mass corresponds to 1 in the histogram. We also define the effective number of samples,  $n_{\text{eff}}$ , as the sum of rescaled importance weights,  $n_{\text{eff}} = \sum_{i=1}^n \tilde{q}_i$ . Finally, in Figure 1C we show  $N = 20,000$  new samples  $\{\theta_i^*\}_{i=1}^N$  that are drawn from the prior samples  $\{\theta_i\}_{i=1}^n$  according to the probability mass  $q_i$  for each  $\theta_i$ . The blue and green contour lines represent (68% and 90%) credible regions for the resampled distribution and for the

**TABLE 1** Target values,  $z$ , and error assignments,  $\varepsilon$ , for observables used in the model calibration and for predictions. Energies in [MeV], point-proton radii in [fm], and the deuteron quadrupole moment in [ $e^2\text{fm}^2$ ].

**Calibration observables**

Observable	$z$	$\varepsilon_{\text{exp}}$	$\varepsilon_{\text{model}}$	$\varepsilon_{\text{method}}$	$\varepsilon_{\text{em}}$ (%)
$E(^2\text{H})$	-2.2298	0	0.05	0.0005	0.001
$R_p(^2\text{H})$	1.976	0	0.005	0.0002	0.0005
$Q(^2\text{H})$	0.27	0.01	0.003	0.0005	0.001
Predicted observables					
$E(^3\text{H})$	-8.4818	0	0.17	0.0005	0.01
$E(^4\text{He})$	-28.2956	0	0.55	0.0005	0.01
$R_p(^4\text{He})$	1.455	0	0.016	0.0002	0.003



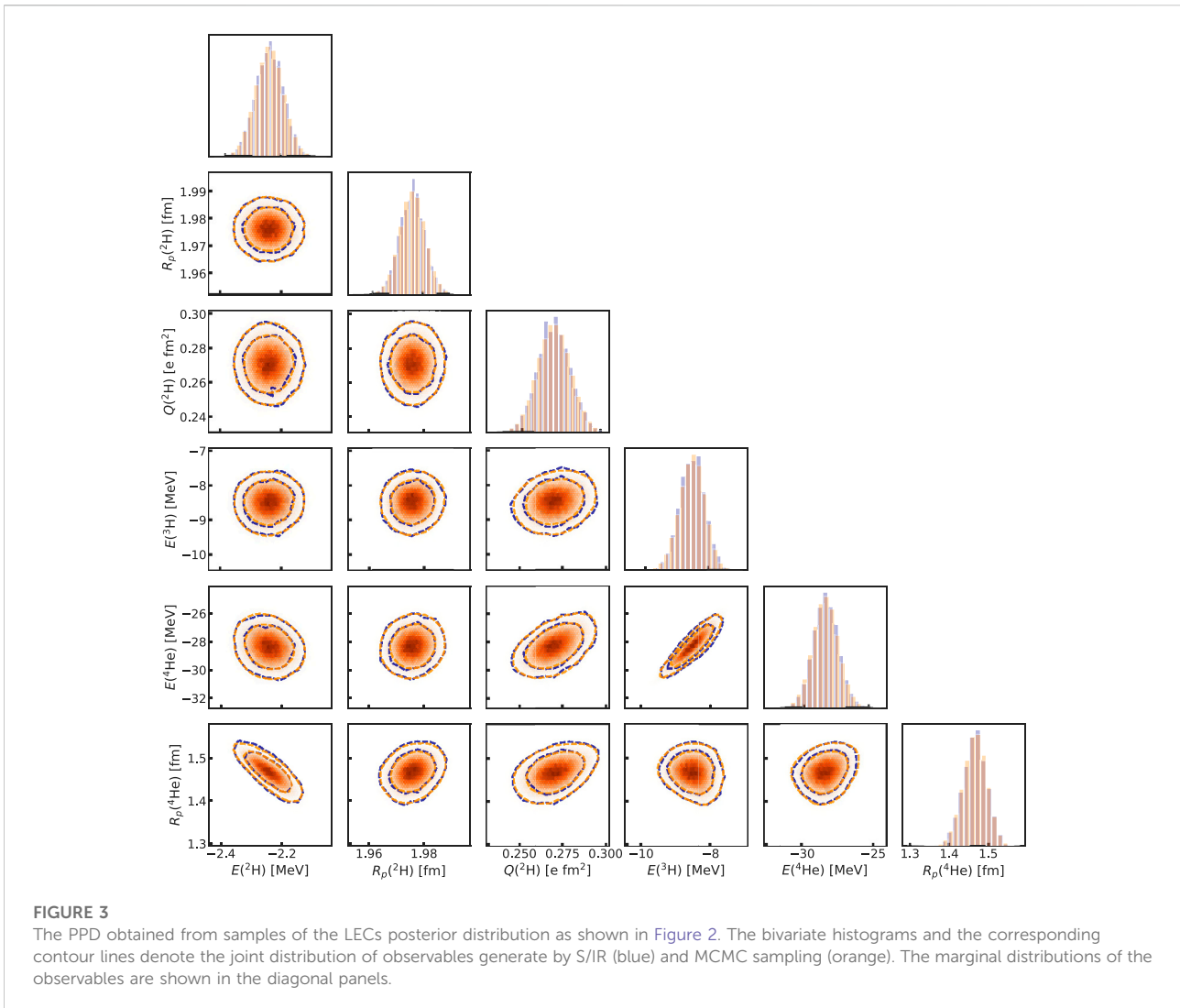
**FIGURE 2** The joint posterior of LECs sampled with S/IR (blue) compared with MCMC sampling (orange). The LECs are shown in units of  $10^4 \text{ GeV}^{-1}$ ,  $10^4 \text{ GeV}^{-2}$  and  $10^4 \text{ GeV}^{-4}$  for  $c_i$ ,  $\tilde{C}_i$  and  $C_i$ , respectively. The likelihood observables and assigned errors are given in Table 1. The contour lines indicate 68% and 90% credible regions.

Student-t distribution, respectively. This result demonstrates that the samples generated by the S/IR method give a very good approximation of the target posterior distribution.

### 3 Nuclear physics applications

Now that we have reviewed the basic idea of the S/IR method, we move on to present realistic applications of the resampling technique in nuclear structure calculations. Here we study Bayesian inference involving the  $\Delta$ NNLO chiral interaction [33] with explicit inclusion of delta isobar degree of freedom at next-to-next-to-leading order. In Weinberg’s power counting the  $\Delta$ NNLO interaction model is parametrized by 17 LECs, with four pion-nucleon LECs ( $c_{1,2,3,4}$ ) that are inferred from pion-nucleon scattering data and 13 additional LECs that should be inferred from extant experimental data of low-energy nucleon-nucleon scattering and bound-state nuclear observables.

For this application we treat only a subset of the parameters as active and keep the other LECs fixed at values taken from the  $\Delta$ NNLO<sub>GO</sub>(450) interaction [34]. Specifically, we consider deuteron observables and use seven active model parameters:  $c_{1,2,3,4}$ ,  $\tilde{C}_{3S1}$ ,  $C_{3S1}$ ,  $C_{E1}$ . Our Gaussian likelihood contains three data with independent errors: the deuteron ground state energy  $E$ , its point-proton radius  $R_p$  and one-body quadrupole moment  $Q$  with experimental and theoretical targets from Refs. [35–37]. Note that the target point-proton radii were transformed from experimental charge radii using the same relation as in Ref. [38]. For the target  $Q$  we use the theoretical result obtained by the CD-Bonn [37] model. With these simplified conditions, we perform S/IR as well as MCMC sampling to study 1) the posterior PDF for the LECs and 2) posterior predictive distributions (PPDs) for selected few-body observables. This application therefore allows a straightforward comparison of the two different sampling methods in a realistic setting. We note that the inclusion of all 17 LECs as active parameters would have required more

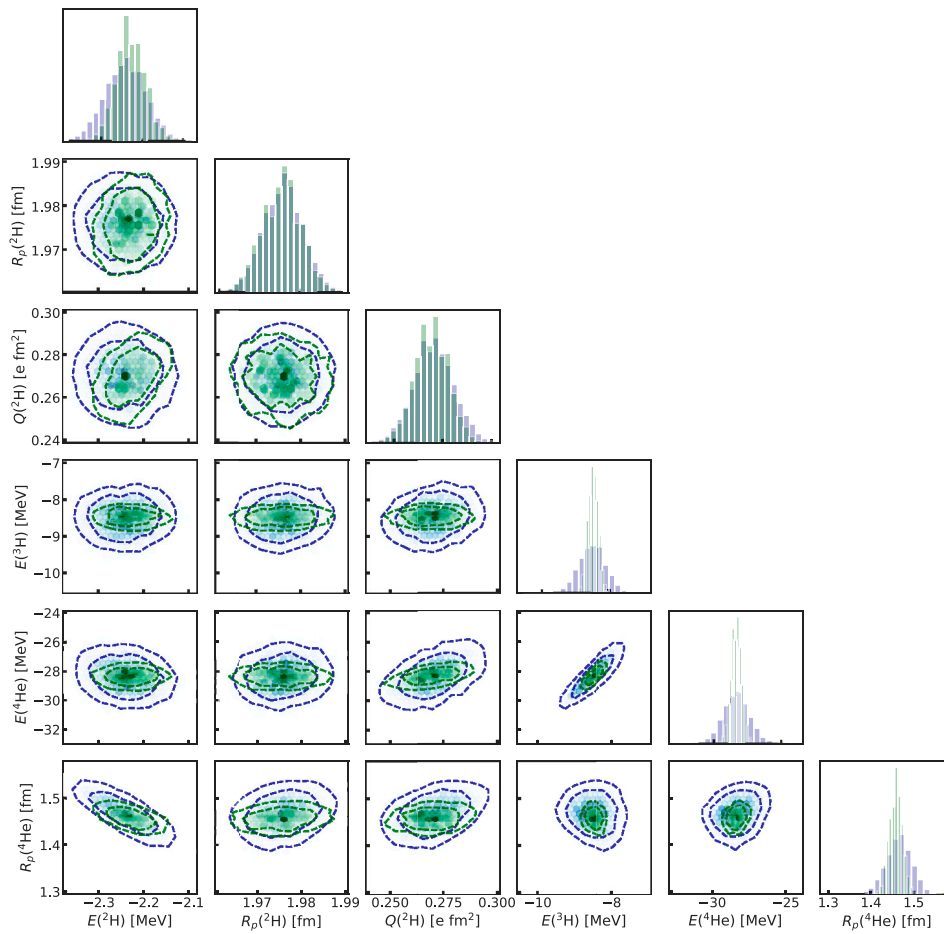


careful tuning of the MCMC sampling algorithm and corresponding convergence studies.

It is the computation of observables, e.g., for likelihood evaluations, which is usually the major, time-consuming bottleneck in Bayesian analyses using MCMC methods. In this application, the statistical analysis is enabled by the use of emulators which mimic the outputs of many-body solvers but are faster by orders of magnitude. The emulators employed here for the ground-state observables of the deuteron, and later for few-body observables, are based on eigenvector continuation [39–41]. These emulators allow to reduce the computation time from seconds to milliseconds while keeping the relative error (compared with full no-core shell model calculation) within 0.001%. Unfortunately, emulators are not yet available for all nuclear observables. The MCMC sampling of posterior PDFs, or the evaluation of expectation integrals such as Eq. 2, will typically not work for models with observables that require heavy calculations.

The experimental target values and error assignments for the calibration observables used to condition the posterior PDF are listed in the upper half of Table 1. In this study we assume a normally-distributed likelihood, and consider different sources of error when calibrating the model predictions with experimental data. The errors are assumed to be independent. They include experimental,  $\epsilon_{\text{exp}}$ , model ( $\chi$ EFT truncation) discrepancy,  $\epsilon_{\text{model}}$ , many-body method,  $\epsilon_{\text{method}}$ , and emulator,  $\epsilon_{\text{em}}$ , errors. The  $\chi$ EFT truncation errors are estimated based on order-by-order calculations as in Ref. [33]. More details on the determination of the error scales can be found in Ref. [42].

Furthermore, we take advantage of previous studies and incorporate information about  $c_{1,2,3,4}$  from a Roy-Steiner analysis of pion-nucleon scattering data [43] and identify a non-implausible domain for  $\tilde{C}_{3S1}$ ,  $C_{3S1}$ ,  $C_{E1}$  from a history matching approach in Ref. [42]. With this prior knowledge we



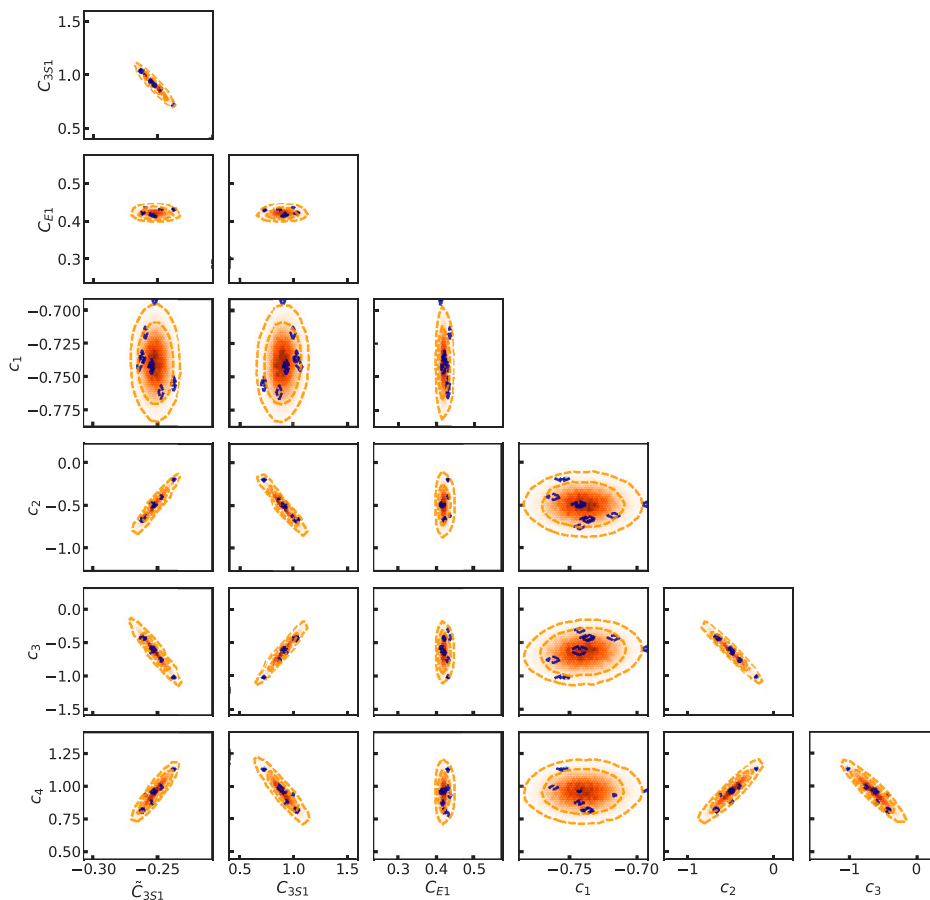
**FIGURE 4**  
 The posterior predictive distribution from sampling over two different posterior distributions.  $PPD_{A=2}$  (blue) is calibrated by the deuteron observables while  $PPD_{A=2,3,4}$  (green) is calibrated by the deuteron,  $^3\text{H}$  and  $^4\text{He}$  observables. The marginal distributions of the observables are shown in the diagonal panels.

set up the prior distribution of the seven LECs as the product of a multivariate Gaussian for  $c_{1,2,3,4}$  and a uniform distribution for  $\tilde{C}_{3SI}, C_{3SI}, C_{E1}^1$ . We note that the use of history matching is very beneficial for both S/IR and MCMC sampling. For S/IR it allows to select a sampling distribution that promises a large overlap with the target distribution and it identifies prior samples that are likely to have large weights in the resampling step. For MCMC, the non-improbable samples from history matching serve as good starting points for the walkers and thereby give faster convergence.

1 Specifically we use the non-improbable domain that was identified in wave 2 of the history matching performed in Ref. [42]. This wave only included deuteron observables.

### 3.1 Posterior sampling

Once we have the prior and the likelihood function we are able to draw samples from the posterior PDF and to analyze the *ab initio* description of few-nucleon systems with the present interaction model. The joint posterior of the LECs is shown in Figure 2, where we compare bivariate, marginal distributions from S/IR and MCMC sampling. For the MCMC sampling we employed an open-source Python toolkit called emcee [44] that performs affine-invariant ensemble sampling. We use 150 walkers that are warmed up with 5,000 initial steps and then move for  $5 \times 10^5$  steps. This amounts to  $7.6 \times 10^7$  likelihood evaluations. The positions of the walkers are recorded every 500 steps which gives  $1.5 \times 10^5$  samples from the posterior distribution of the LECs. On the other hand, for S/IR we first acquire  $2 \times 10^4$  samples from the prior distribution and perform the same number of likelihood evaluations to get the importance weights. From this limited set we then draw  $1.5 \times$



**FIGURE 5**  
 The posterior of LECs sampled with S/IR (blue) compared with MCMC sampling (orange) for a situation when the deuteron calibration observables are associated with errors that have been reduced by an order of magnitude (see text for details). The LECs are shown in units of  $10^4 \text{ GeV}^{-1}$ ,  $10^4 \text{ GeV}^{-2}$  and  $10^4 \text{ GeV}^{-4}$  for  $c_i$ ,  $\tilde{c}_i$  and  $C_i$ , respectively.

$10^5$  samples using resampling (the same final number as in MCMC). Note that several prior samples occur more than once in the final sample set. Here the number of effective samples for S/IR is  $n_{\text{eff}} = 1589.9$ . As we can see from Figure 2, the contour lines of both sampling methods are in good agreement and, e.g., the correlation structure of the LEC pairs are equally well described. The histograms of S/IR and MCMC samples are both plotted in the figure but are almost impossible to distinguish.

As a second stage we use the inferred model to perform model checking of the calibration observables and to predict the  $^3\text{H}$  ground-state energy and the  $^4\text{He}$  ground-state energy and point-proton radius (see Table 1). For this purpose the PPD is defined as the set

$$\{\mathbf{y}_{\text{th}}(\boldsymbol{\theta}); \boldsymbol{\theta} \sim \text{pr}(\boldsymbol{\theta} | \mathcal{D})\}, \quad (7)$$

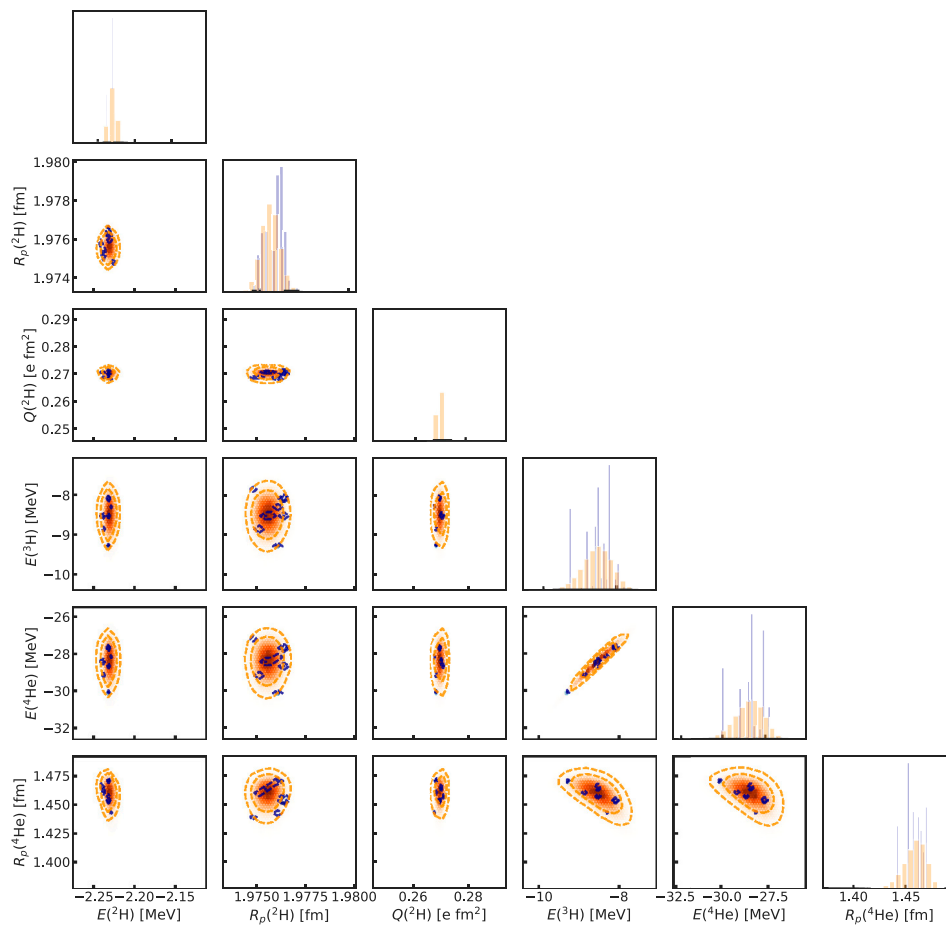
where  $\mathbf{y}_{\text{th}}(\boldsymbol{\theta})$  is the theoretical predictions of selected observables using the model parameter vector  $\boldsymbol{\theta}$ . Figure 3 illustrate the PPD of the three deuteron observables using S/IR

(blue) and MCMC sampling (orange). The marginal histograms of the observable predictions are shown in the diagonal panels of the corner plot. In this study both sampling methods give very similar distributions for all observables. Note that the predictive distributions for the three deuteron observables can be considered as model checking since they appeared in the likelihood function and therefore conditioned the LEC posterior. The  $^3\text{H}$  and  $^4\text{He}$  observables, on the other hand, are predictions in this study. Their distributions are characterized by larger variances compared to the deuteron predictions.

### 3.2 Posterior probability updates

As mentioned in the introduction, the S/IR method requires a minimum amount of computation to produce new samples when the posterior PDF is updated for various reasons. Here we present





**FIGURE 6**  
The PPD generated using S/IR (blue) and MCMC sampling (orange) for the posterior distributions shown in Figure 5. Marginal histograms of the observables are shown in the diagonal panels.

one likely scenario where the posterior is changed due to different choices of calibration data (for instance the inclusion of newly-accessible observables). Let us start from the previously described calibration of our interaction model with three selected deuteron observables. If we add  $^3\text{H}$  and  $^4\text{He}$  observables into the calibration (experimental target values and error assignments as in Table 1) to further condition the model, the likelihood function needs to be updated accordingly. The sampling of the posterior PDF should be repeated from the beginning and the new samples should be used to construct PPDs. However, using S/IR we resample from the same set of prior samples—only with different importance weights. The same set of samples also appear in the sampling of PPDs. To distinguish the original and the updated posteriors we use the notation  $PPD_{A=2}$  to denote predictions with only deuteron observable as calibration data and  $PPD_{A=2,3,4}$  with  $^3\text{H}$  and  $^4\text{He}$  added to the likelihood. These two different PPDs, generated by S/IR, are shown in Figure 4. Note that the  $PPD_{A=2}$  (blue) is the same as in Figure 3, and is shown here as a benchmark. As expected we observe that the description of  $^3\text{H}$

and  $^4\text{He}$  observables is more accurate and more precise (smaller variations) with  $PPD_{A=2,3,4}$  (green) as compared with  $PPD_{A=2}$  (blue). We also find that the deuteron ground state energy is slightly improved with the updated posterior. This can be explained by the anti-correlation between  $R_p(^4\text{He})$  and  $E(^2\text{H})$ . The additional constraints imposed by  $R_p(^4\text{He})$  through the likelihood function propagates to  $E(^2\text{H})$  via the correlation structure.

### 3.3 S/IR limitations

So far we have focused on the feasibility and advantage of the S/IR approach. However, there are some important limitations and we recommend users to be mindful of the number of effective samples. In Figure 4, we found that our S/IR sampling of  $PPD_{A=2}$  has  $n_{\text{eff}} = 1589.9$ , while for  $PPD_{A=2,3,4}$  it drops to  $n_{\text{eff}} = 314.9$ . This can be understood by the resampling from a fixed set of prior samples. The more

complex the likelihood function, the less effective the samples. As seen in Figure 4, the contour lines of  $PPD_{A=2,3,4}$  is less smooth than those obtained from  $PPD_{A=2}$  due to the smaller number of effective samples. The S/IR method will eventually break when  $n_{\text{eff}}$  becomes too small. An intermediate remedy could be the use of kernel density estimators, although that approach typically introduces an undesired sensitivity to the choice of kernel widths.

A similar situation occurs when the target observables are characterized by very small error assignments. This leads to a sharply peaked likelihood function and a decreased overlap with the prior samples. The resulting large variance of importance weights implies that the final set representing the posterior distribution will be dominated by a very small number of samples. Here we show such an example where resampling no longer works. We attempted to reconstruct a PPD with only deuteron observables in the calibration, but where all error assignments in Table 1 had been reduced by an order of magnitude. The results of this analysis are shown in Figures 5, 6 which display the PDFs and PPDs, respectively, generated by S/IR (blue) and compared with MCMC (orange). The S/IR method does not perform well in this case. With  $n_{\text{eff}} = 4.4$  the PDF and PPD generated by S/IR are represented by a few samples. The MCMC sampling, on the other hand, does manage to identify the updated distribution.

Unfortunately one can also envision more difficult scenarios in which S/IR could fail without any clear signatures. For example, if the prior has a very small overlap with the posterior there is a risk that many prior samples get a similar importance weight (such that the number of effective samples is large) but that one has missed the most interesting region. Again, history matching is a very useful tool in the analysis as it can be used to ensure that we are focusing on the LECs domain that covers the mode(s) of the posterior.

## 4 Summary

In this paper we reviewed an established sampling method known as S/IR. Specifically, we applied importance resampling using the weighted bootstrap algorithm and sampled the posterior PDF for selected LECs of the  $\Delta$ NNLO interaction model conditioned on deuteron observables. The resulting PDF and PPD were compared with those obtained from MCMC sampling and a very good agreement was found. We also demonstrated Bayesian updating using S/IR by the addition of  $^3\text{H}$  and  $^4\text{He}$  observables to the calibration data set. As expected, the predictions of  $^3\text{H}$  and  $^4\text{He}$  observables were improved, but also the description of the deuteron ground-state energy which could be explained by the correlation structure between  $E(^2\text{H})$  and  $R_p(^4\text{He})$ . Finally, we illustrated some limitations of the S/IR

method that were signaled by small numbers of effective samples. We found that such situations occurred when the likelihood became too complex for the limited model, or when prior samples failed to resolve a very peaked posterior that resulted from small tolerances. We also argued that prior knowledge of the posterior landscape is very useful to avoid possible failure scenarios that might not be signaled by the number of effective samples.

## Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

## Author contributions

WJ and CF contributed equally in this paper.

## Funding

This work was supported by the European Research Council under the European Unions Horizon 2020 research and innovation program (Grant No. 758027) and the Swedish Research Council (Grant Nos 2017-04234 and 2021-04507). The computations and data handling were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at Chalmers Centre for Computational Science and Engineering (C3SE), and the National Supercomputer Centre (NSC) partially funded by the Swedish Research Council through Grant No. 2018-05973.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Schindler MR, Phillips DR. Bayesian methods for parameter estimation in effective field theories. *Ann Phys (N Y)* (2009) 324:682–708. doi:10.1016/j.aop.2008.09.003
- Caesar C, Simonis J, Adachi T, Aksyutina Y, Alcantara J, Altstadt S, et al. Beyond the neutron drip line: The unbound oxygen isotopes  $^{25}\text{o}$  and  $^{26}\text{o}$ . *Phys Rev C* (2013) 88:034313. doi:10.1103/PhysRevC.88.034313
- Furnstahl RJ, Klco N, Phillips DR, Wesolowski S. Quantifying truncation errors in effective field theory. *arXiv e-prints* (2015).
- Wesolowski S, Furnstahl RJ, Melendez JA, Phillips DR. Exploring bayesian parameter estimation for chiral effective field theory using nucleon–nucleon phase shifts. *J Phys G: Nucl Part Phys* (2019) 46:045102. doi:10.1088/1361-6471/aaf5fc
- Melendez JA, Furnstahl RJ, Phillips DR, Pratola MT, Wesolowski S. Quantifying correlated truncation errors in effective field theory. *Phys Rev C* (2019) 100:044001. doi:10.1103/PhysRevC.100.044001
- Epelbaum E, Golak J, Hebeler K, Kamada H, Krebs H, Meißner UG, et al. Towards high-order calculations of three-nucleon scattering in chiral effective field theory. *Eur Phys J A* (2020) 56:92. doi:10.1140/epja/s10050-020-00102-2
- Yang L, Lin C, Zhang Y, Wen P, Jia H, Wang D, et al. Bayesian analysis on interactions of exotic nuclear systems. *Phys Lett B* (2020) 807:135540. doi:10.1016/j.physletb.2020.135540
- Phillips DR, Furnstahl RJ, Heinz U, Maiti T, Nazarewicz W, Nunes FM, et al. Get on the band wagon: A bayesian framework for quantifying model uncertainties in nuclear dynamics. *J Phys G: Nucl Part Phys* (2021) 48:072001. doi:10.1088/1361-6471/abf1df
- Drischler C, Furnstahl RJ, Melendez JA, Phillips DR. How well do we know the neutron-matter equation of state at the densities inside neutron stars? A bayesian approach with correlated uncertainties. *Phys Rev Lett* (2020) 125:202702. doi:10.1103/PhysRevLett.125.202702
- Drischler C, Melendez JA, Furnstahl RJ, Phillips DR. Quantifying uncertainties and correlations in the nuclear-matter equation of state. *Phys Rev C* (2020) 102:054315. doi:10.1103/PhysRevC.102.054315
- Maris P, Epelbaum E, Furnstahl RJ, Golak J, Hebeler K, Huther T, et al. Light nuclei with semilocal momentum-space regularized chiral interactions up to third order. *Phys Rev C* (2021) 103:054001. doi:10.1103/PhysRevC.103.054001
- Wesolowski S, Svensson I, Ekström A, Forssén C, Furnstahl RJ, Melendez JA, et al. Rigorous constraints on three-nucleon forces in chiral effective field theory from fast and accurate calculations of few-body observables. *Phys Rev C* (2021) 104:064001. doi:10.1103/PhysRevC.104.064001
- Djårv T, Ekström A, Forssén C, Johansson HT. Bayesian predictions for  $A=6$  nuclei using eigenvector continuation emulators. *Phys Rev C* (2022) 105:014005. doi:10.1103/PhysRevC.105.014005
- Svensson I, Ekström A, Forssén C. Bayesian parameter estimation in chiral effective field theory using the Hamiltonian Monte Carlo method. *Phys Rev C* (2022) 105:014004. doi:10.1103/PhysRevC.105.014004
- Acharya B, Bacca S. Gaussian process error modeling for chiral effective-field-theory calculations of  $np \leftrightarrow dy$  at low energies. *Phys Lett B* (2022) 827:137011. doi:10.1016/j.physletb.2022.137011
- Svensson I, Ekström A, Forssén C. Bayesian estimation of the low-energy constants up to fourth order in the nucleon-nucleon sector of chiral effective field theory. *arXiv 2206.08250* (2022). doi:10.48550/arXiv.2206.08250
- Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of state calculations by fast computing machines. *J Chem Phys* (1953) 21:1087–92. doi:10.1063/1.1699114
- Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* (1970) 57:97–109. doi:10.1093/biomet/57.1.97
- Hitchcock DB. A history of the metropolis–hastings algorithm. *The Am Statistician* (2003) 57:254–7. doi:10.1198/0003130032413
- von Toussaint U. Bayesian inference in physics. *Rev Mod Phys* (2011) 83:943–99. doi:10.1103/RevModPhys.83.943
- Brooks S, Gelman A, Jones G, Meng X. *Handbook of Markov chain Monte Carlo*. Florida, United States: Chapman & Hall/CRC Handbooks of Modern Statistical Methods CRC Press (2011).
- Rubin DB. Using the sir algorithm to simulate posterior distributions. *Bayesian Stat* (1988) 3:395–402.
- Smith AFM, Gelfand AE. Bayesian statistics without tears: A sampling-resampling perspective. *Am Stat* (1992) 46:84–8. doi:10.2307/2684170
- Bernardo J, Smith A. Bayesian theory. In: *Wiley series in probability and statistics*. New Jersey, United States: John Wiley & Sons Canada, Limited (2006).
- Kolck UV. Effective field theory of nuclear forces. *Prog Part Nucl Phys* (1999) 43:337–418. doi:10.1016/S0146-6410(99)00097-6
- Bogner SK, Kuo TTS, Schwenk A. Model-independent low momentum nucleon interaction from phase shift equivalence. *Phys Rep* (2003) 386:1–27. doi:10.1016/j.physrep.2003.07.001
- Epelbaum E, Hammer HW, Meißner UG. Modern theory of nuclear forces. *Rev Mod Phys* (2009) 81:1773–825. doi:10.1103/RevModPhys.81.1773
- Bogner S, Furnstahl R, Schwenk A. From low-momentum interactions to nuclear structure. *Prog Part Nucl Phys* (2010) 65:94–147. doi:10.1016/j.pnpnp.2010.03.001
- Machleidt R, Entem D. Chiral effective field theory and nuclear forces. *Phys Rep* (2011) 503:1–75. doi:10.1016/j.physrep.2011.02.001
- Weinberg S. Nuclear forces from chiral Lagrangians. *Phys Lett B* (1990) 251:288–92. doi:10.1016/0370-2693(90)90938-3
- Weinberg S. Effective chiral Lagrangians for nucleon-pion interactions and nuclear forces. *Nucl Phys B* (1991) 363:3–18. doi:10.1016/0550-3213(91)90231-L
- Kaplan DB, Savage MJ, Wise MB. A new expansion for nucleon-nucleon interactions. *Phys Lett B* (1998) 424:390–6. doi:10.1016/S0370-2693(98)00210-X
- Ekström A, Hagen G, Morris TD, Papenbrock T, Schwartz PD.  $\Delta$  isobars and nuclear saturation. *Phys Rev C* (2018) 97:024332. doi:10.1103/PhysRevC.97.024332
- Jiang WG, Ekström A, Forssén C, Hagen G, Jansen GR, Papenbrock T. Accurate bulk properties of nuclei from  $A = 2$  to  $\infty$  from potentials with  $\Delta$  isobars. *Phys Rev C* (2020) 102:054301. doi:10.1103/PhysRevC.102.054301
- Wang M, Huang WJ, Kondev FG, Audi G, Naimi S. The AME 2020 atomic mass evaluation (II). Tables, graphs and references. *Chin Phys C* (2021) 45:030003. doi:10.1088/1674-1137/abddaf
- Angeli I, Marinova K. Table of experimental nuclear ground state charge radii: An update. *Data Nucl Data Tables* (2013) 99:69–95. doi:10.1016/j.adt.2011.12.006
- Machleidt R. High-precision, charge-dependent Bonn nucleon-nucleon potential. *Phys Rev C* (2001) 63:024001. doi:10.1103/PhysRevC.63.024001
- Ekström A, Jansen GR, Wendt KA, Hagen G, Papenbrock T, Carlsson BD, et al. Accurate nuclear radii and binding energies from a chiral interaction. *Phys Rev C* (2015) 91:051301. doi:10.1103/PhysRevC.91.051301
- Frame D, He R, Ipsen I, Lee D, Lee D, Rrapaj E. Eigenvector continuation with subspace learning. *Phys Rev Lett* (2018) 121:032501. doi:10.1103/PhysRevLett.121.032501
- König S, Ekström A, Hebeler K, Lee D, Schwenk A. Eigenvector continuation as an efficient and accurate emulator for uncertainty quantification. *Phys Lett B* (2020) 810:135814. doi:10.1016/j.physletb.2020.135814
- Ekström A, Hagen G. Global sensitivity analysis of bulk properties of an atomic nucleus. *Phys Rev Lett* (2019) 123:252501. doi:10.1103/PhysRevLett.123.252501
- Hu B, Jiang W, Miyagi T, Sun Z, Ekstrom A, Forssen C, et al. *Ab initio* predictions link the neutron skin of  $^{208}\text{Pb}$  to nuclear forces. *Nat Phys* (2022) 18:1196–200. doi:10.1038/s41567-022-01715-8
- Siemens D, Ruiz de Elvira J, Epelbaum E, Hoferichter M, Krebs H, Kubis B, et al. Reconciling threshold and subthreshold expansions for pion–nucleon scattering. *Phys Lett B* (2017) 770:27–34. doi:10.1016/j.physletb.2017.04.039
- Foreman-Mackey D, Hogg DW, Lang D, Goodman J. emcee: The MCMC hammer. *Publications Astronomical Soc Pac* (2013) 125:306–12. doi:10.1086/670067