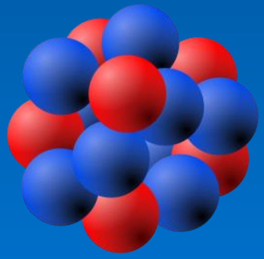
Bayesian probability updates using sampling/importance resampling: Applications in nuclear theory

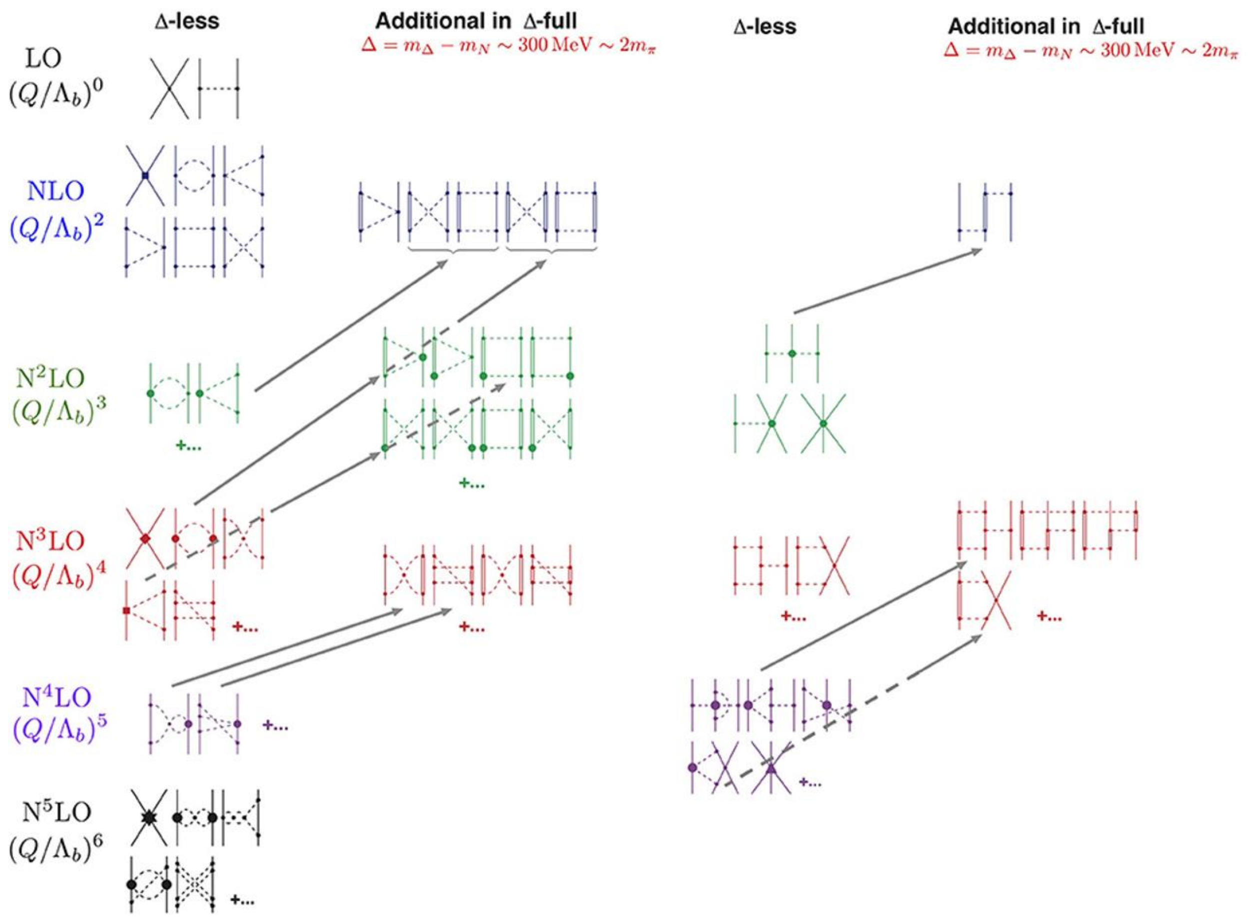


Weiguang Jiang

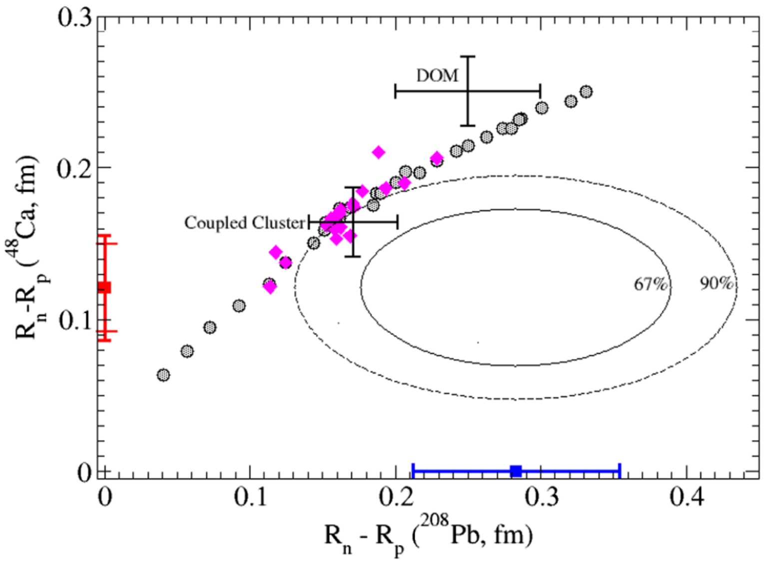
Johannes Gutenberg University of Mainz

ISNET-9, St. Louis, May 23

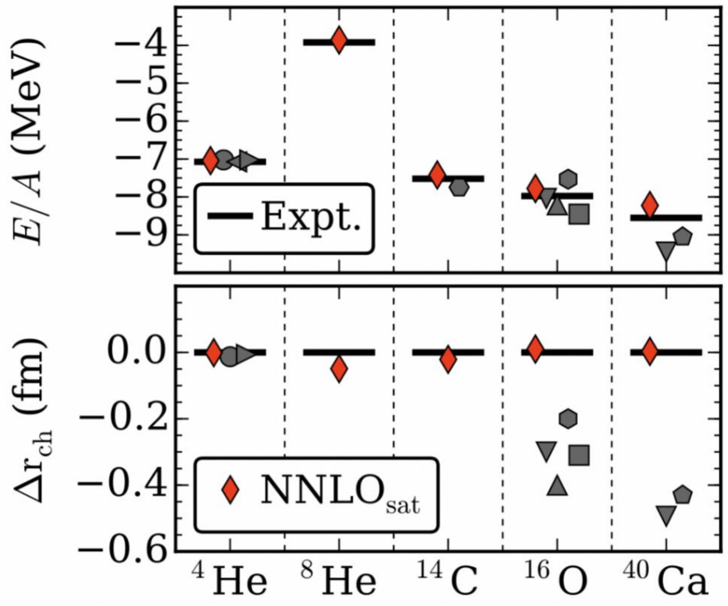
Introduction



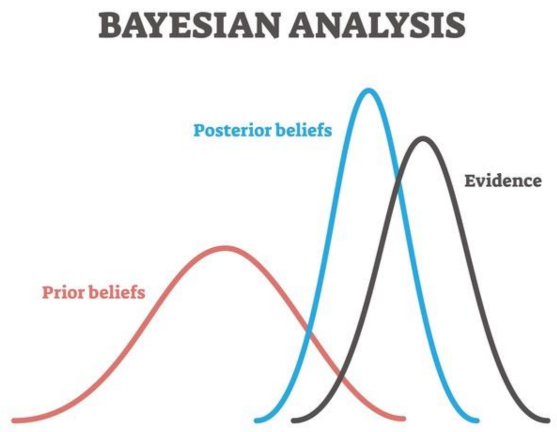
Nuclear interaction based on chiral effective field theory (EFT), parametrized in terms of low energy constants (LECs)

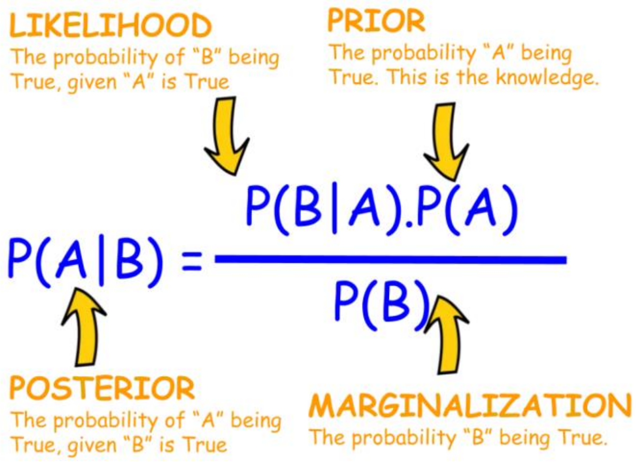


Uncertainty of the nuclear Hamiltonian (nuclear interaction)

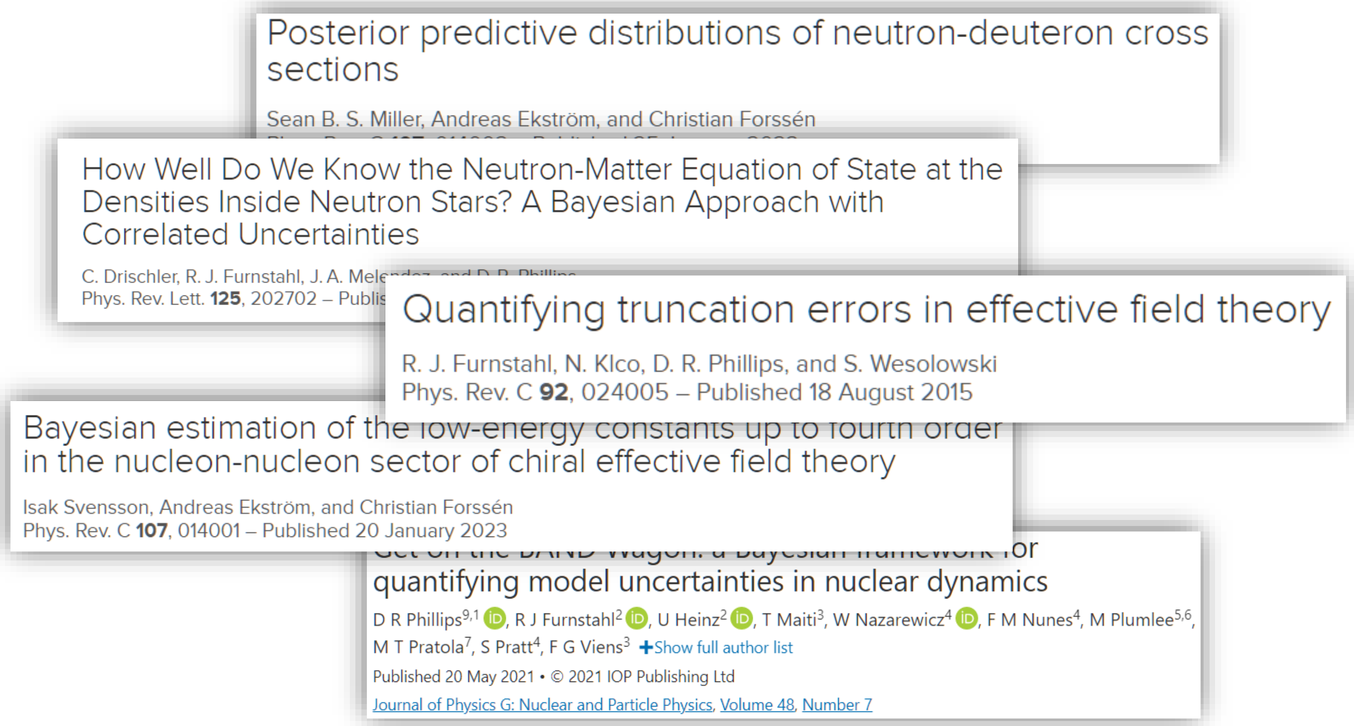


Introduction





Bayesian inference is an appealing approach for dealing with theoretical uncertainties and has been applied in different nuclear physics studies



Posterior predictive distributions of neutron-deuteron cross sections

Sean B. S. Miller, Andreas Ekström, and Christian Forssén

How Well Do We Know the Neutron-Matter Equation of State at the Densities Inside Neutron Stars? A Bayesian Approach with Correlated Uncertainties

C. Drischler, R. J. Furnstahl, J. A. Mele…
Phys. Rev. Lett. **125**, 202702 – Publis…

Quantifying truncation errors in effective field theory

R. J. Furnstahl, N. Klco, D. R. Phillips, and S. Wesolowski
Phys. Rev. C **92**, 024005 – Published 18 August 2015

Bayesian estimation of the low-energy constants up to fourth order in the nucleon-nucleon sector of chiral effective field theory

Isak Svensson, Andreas Ekström, and Christian Forssén
Phys. Rev. C **107**, 014001 – Published 20 January 2023

… quantifying model uncertainties in nuclear dynamics

D R Phillips[9,1], R J Furnstahl[2], U Heinz[2], T Maiti[3], W Nazarewicz[4], F M Nunes[4], M Plumlee[5,6], M T Pratola[7], S Pratt[4], F G Viens[3] +Show full author list

Published 20 May 2021 • © 2021 IOP Publishing Ltd

Journal of Physics G: Nuclear and Particle Physics, Volume 48, Number 7

Bayesian Probability and Sampling/Importance Resampling

Bayesian inference is an excellent framework to incorporate different sources of uncertainty and propagate errors to the model predictions.

- Posterior probability density function (PDF) in Bayes' theorem :

$$\mathrm{pr}(\theta|\mathcal{D}) \propto \underline{\mathcal{L}(\mathcal{D}|\theta)}\ \underline{\mathrm{pr}(\theta)}$$

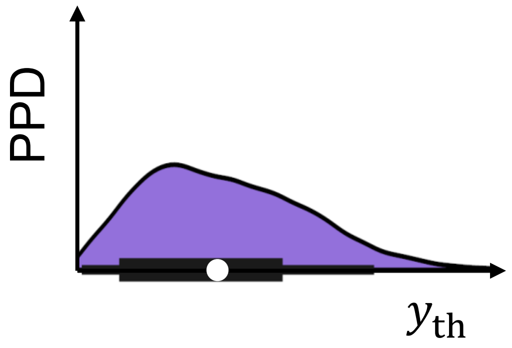
Likelihood function (usually not analytical) — Prior

Prior:
a priori hypothesis of parameterization θ (e.g. LECs under uniform distribution in a certain range)

Likelihood:
different sources of uncertainty (EFT truncation error, the many-body method error, experimental error…) go in here

- Posterior predictive distribution (PDD):

$$\mathrm{PDD} = \{y_{\mathrm{th}}(\theta) : \theta \sim \mathrm{pr}(\theta|\mathcal{D})\}$$

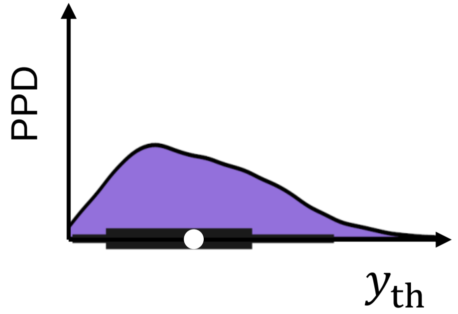


Predicting new observables

- e.g. The expectation value of certain observable $\mathrm{y}(\theta)$:

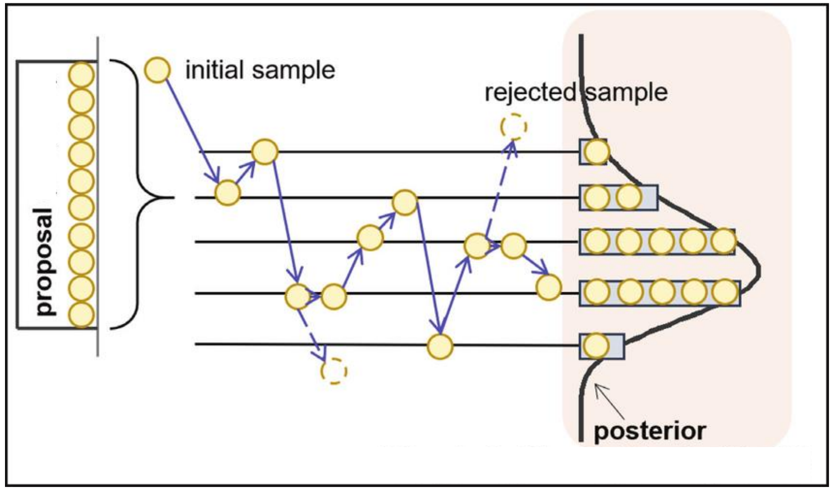
$$\int \mathrm{d}\theta\, \mathrm{y}(\theta)\, \mathrm{pr}(\theta|\mathcal{D}) \;\longrightarrow\; \mathrm{mean}(\mathrm{y}(\theta_i)) \quad \text{with } \{\theta_i \sim \mathrm{pr}(\theta|\mathcal{D})\}$$

Sampling

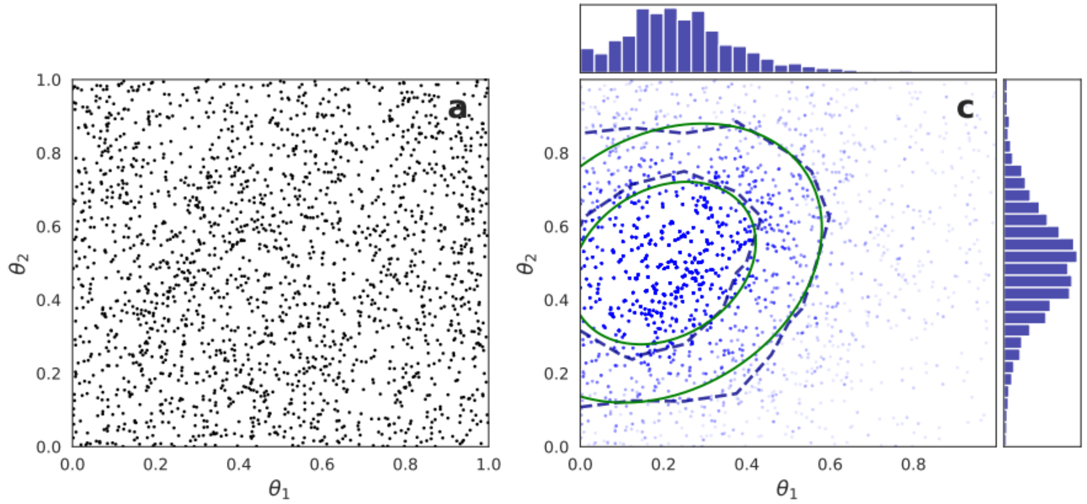


- Sampling method:

Markov chain Monte Carlo (MCMC), Sampling/Importance Resampling(SIR)…



MCMC



SIR

Applications in Nuclear Theory

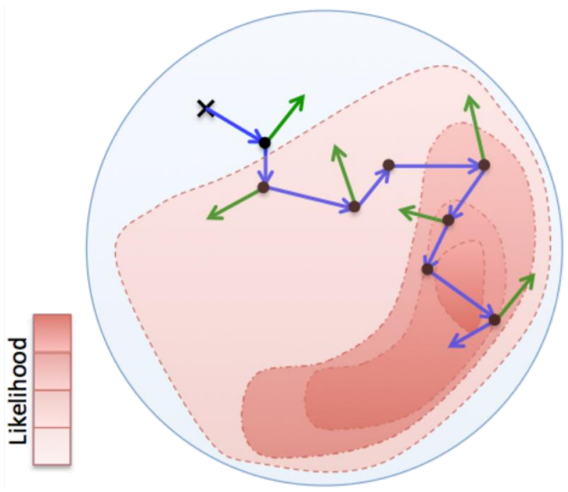
MCMC sampling typically requires many likelihood evaluations, which is often a costly operation in nuclear theory

There are certain 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.

2) Bayesian posterior updates in which calibration data is added in several different stages. Or in model checking where we want to explore the sensitivity to prior assignments. This typically requires that the MCMC sampling must be carried out repeatedly from scratch.

3) Even after we get the pdf using MCMC, the prediction of target observables for which our model evaluations could be very costly and the handling of a large number of MCMC samples becomes infeasible.

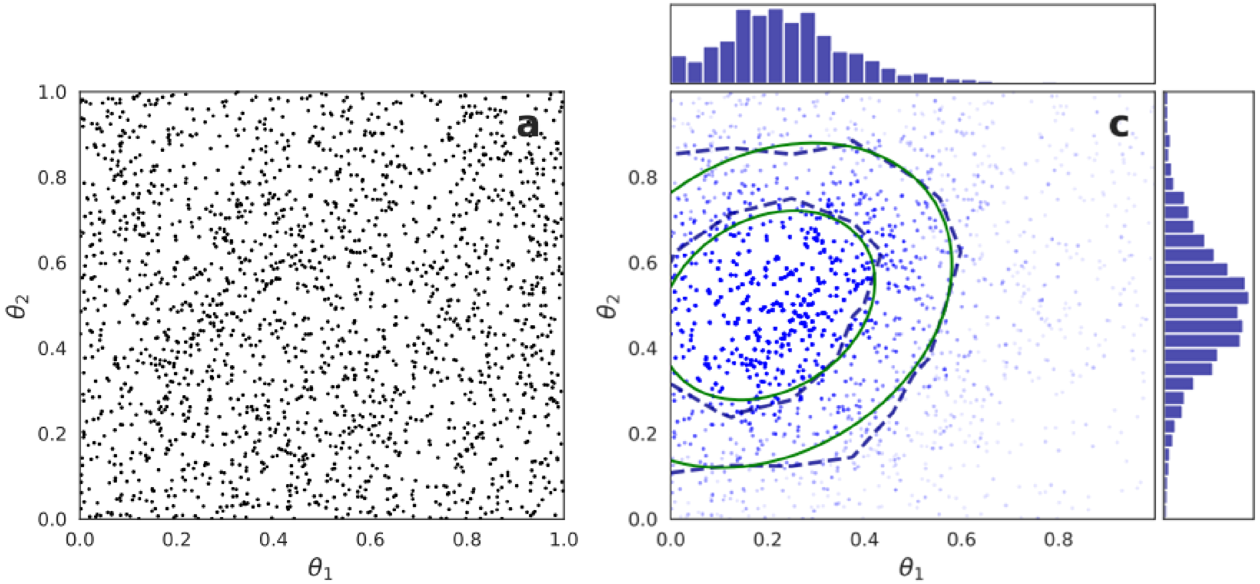


MCMC stochastic processes of "walkers"

The basic idea of SIR is to utilize the inherent duality between samples and the density (probability distribution) from which they were generated

This duality offers an opportunity to indirectly recreate a density (that might be hard to compute) from samples that are easy to obtain.

$$\{\theta_i \sim pr(\theta|\mathcal{D})\}$$



Bayesian Statistics without Tears: A Sampling-Resampling Perspective
Author(s): A. F. M. Smith and A. E. Gelfand
Source: *The American Statistician*, May, 1992, Vol. 46, No. 2 (May, 1992), pp. 84-88

Methodology: Sampling/Importance Resampling

weighted bootstrap

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: $q_i = \omega_i / \sum_{j=1}^{n} \omega_j$.

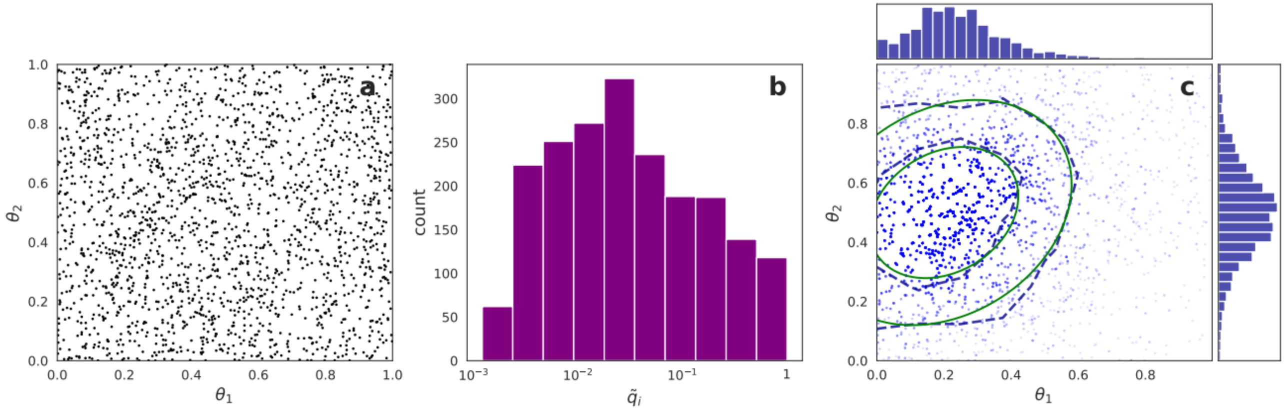
3) Draw N new samples $\{\boldsymbol{\theta}_i^*\}_{i=1}^{N}$ from the discrete distribution $\{\boldsymbol{\theta}_i\}_{i=1}^{n}$ with probability mass q_i on $\boldsymbol{\theta}_i$.

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 θ^* (for the one-dimensional case)

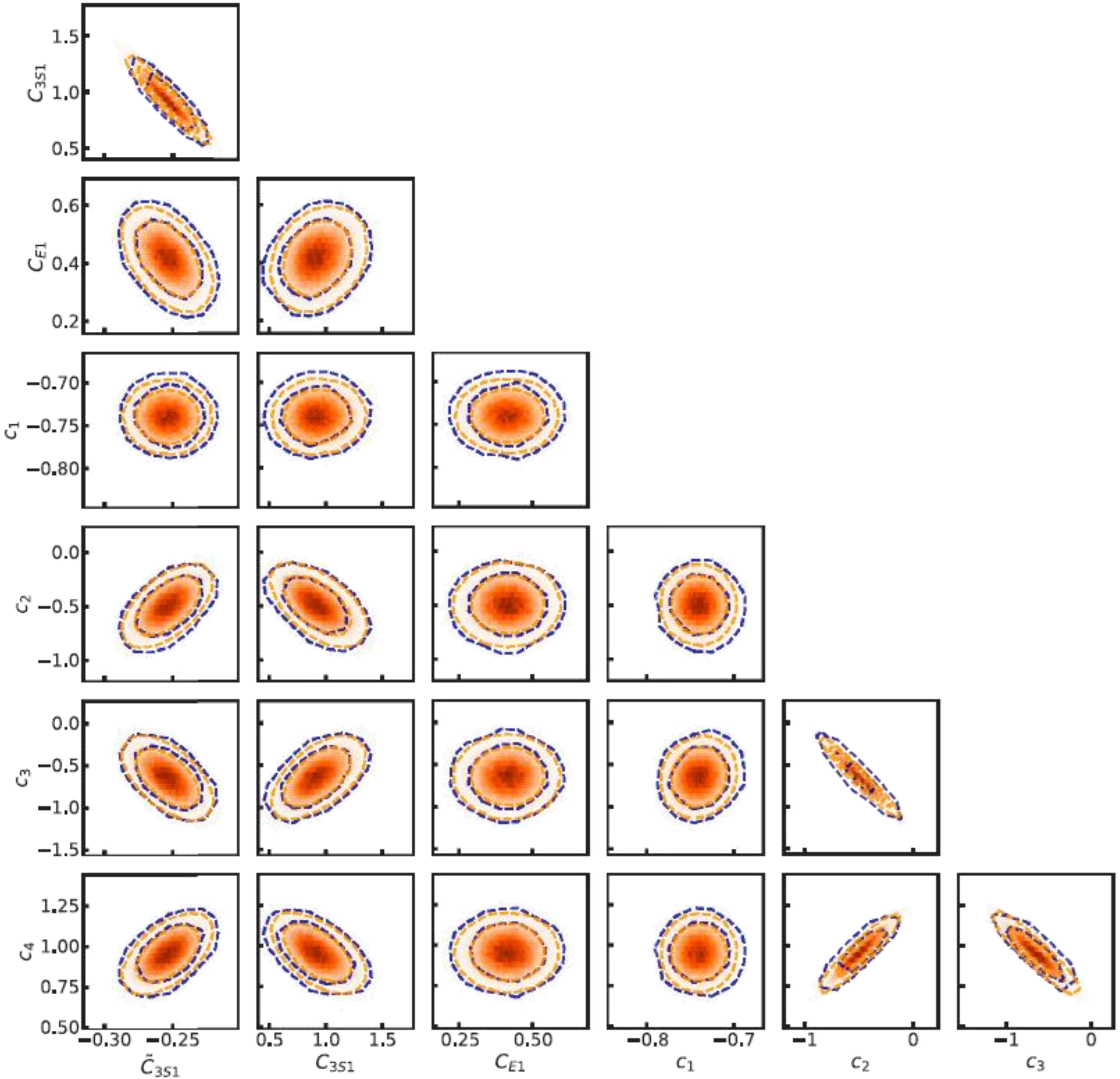
$$\mathrm{pr}\left(\theta^* \leq \mathrm{a}\right) = \sum_{i=1}^{n} q_i \cdot H\left(a - \theta_i\right) = \frac{\frac{1}{n}\sum_{i=1}^{n} \omega_i \cdot H(a-\theta_i)}{\frac{1}{n}\sum_{i=1}^{n} \omega_i}$$

$$\xrightarrow{n \to \infty} \frac{\mathbb{E}_g\left[\frac{f(\theta)}{g(\theta)} \cdot H(a-\theta)\right]}{\mathbb{E}_g\left[\frac{f(\theta)}{g(\theta)}\right]} = \frac{\int_{-\infty}^{a} f(\theta)\, d\theta}{\int_{-\infty}^{\infty} f(\theta)\, d\theta} = \int_{-\infty}^{a} h(\theta)\, d\theta,$$



Sampling (from "simple" distribution) → Importance weights → Importance resampling

Results and Discussion



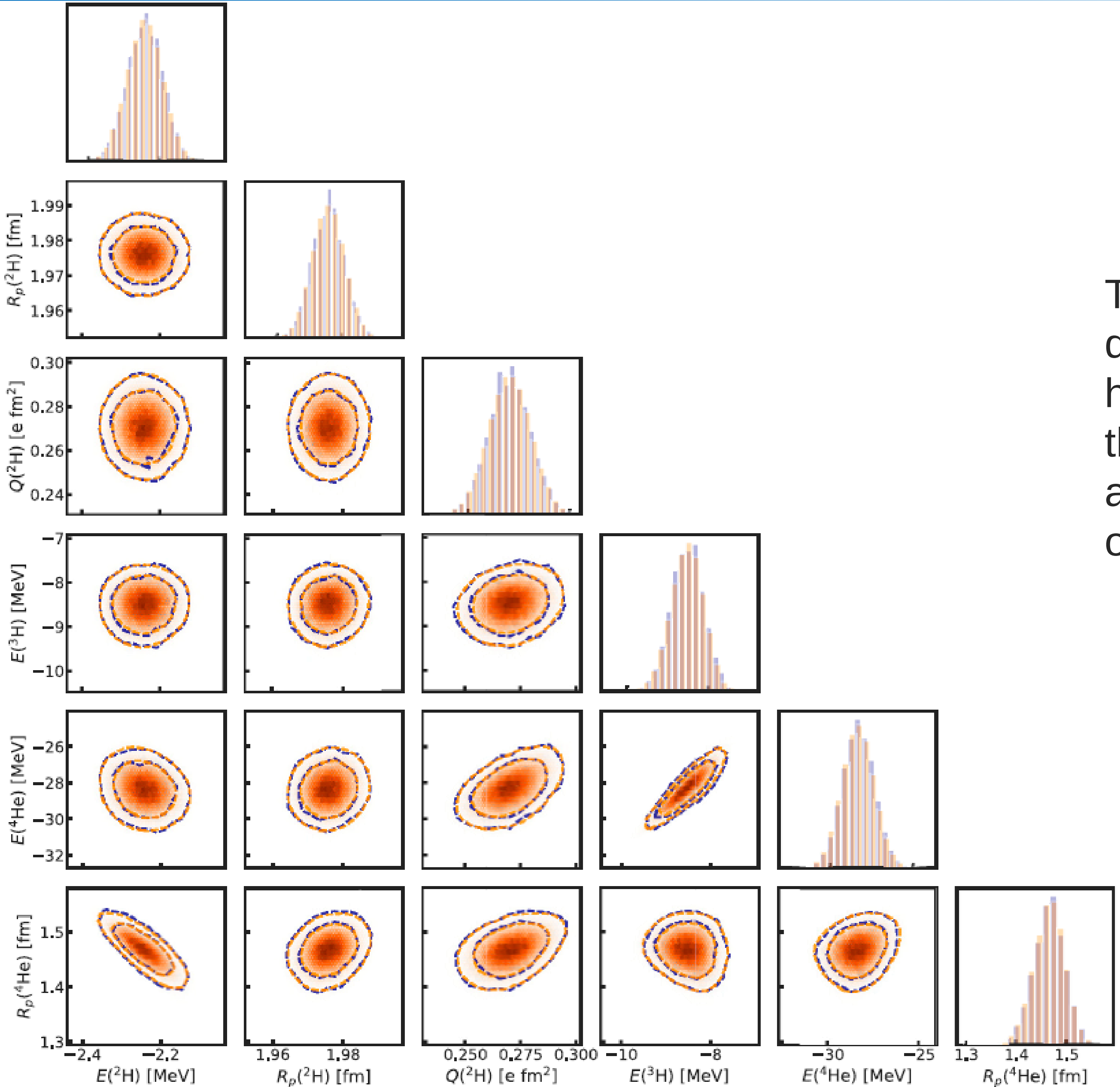
Calibration observables					
Observable	z	ε_{exp}	$\varepsilon_{\text{model}}$	$\varepsilon_{\text{method}}$	ε_{em}
$E(^2\text{H})$	-2.2298	0	0.05	0.0005	0.001%
$R_p(^2\text{H})$	1.976	0	0.005	0.0002	0.0005%
$Q(^2\text{H})$	0.27	0.01	0.003	0.0005	0.001%

seven active model parameters: $c_{1,2,3,4}$, $\tilde{C}_{3S1}$, C_{3S1}, C_{E1}

SIR: resample from 2×10^4 samples (uniform distribution)

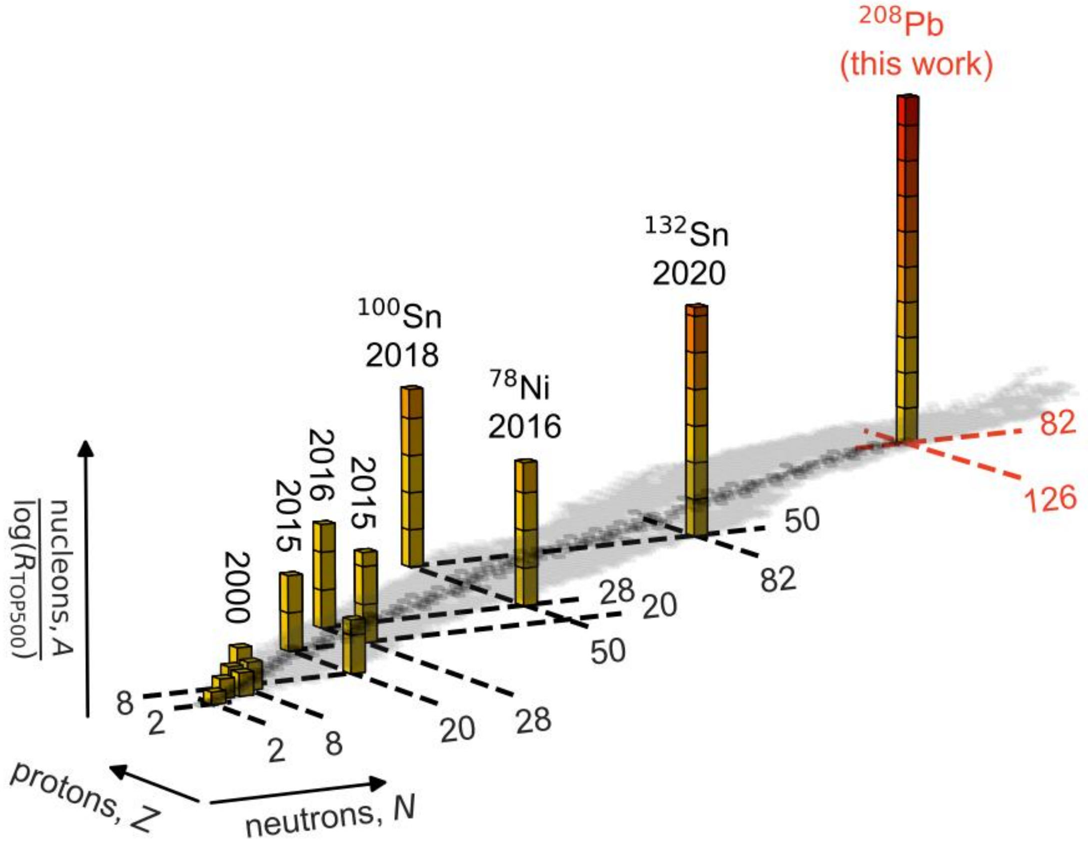
The joint **posterior** of LECs sampled with S/IR (blue) compared with MCMC sampling (orange). The likelihood observables and assigned errors are given in the above Table. The contour lines indicate 68% and 90% credible regions.

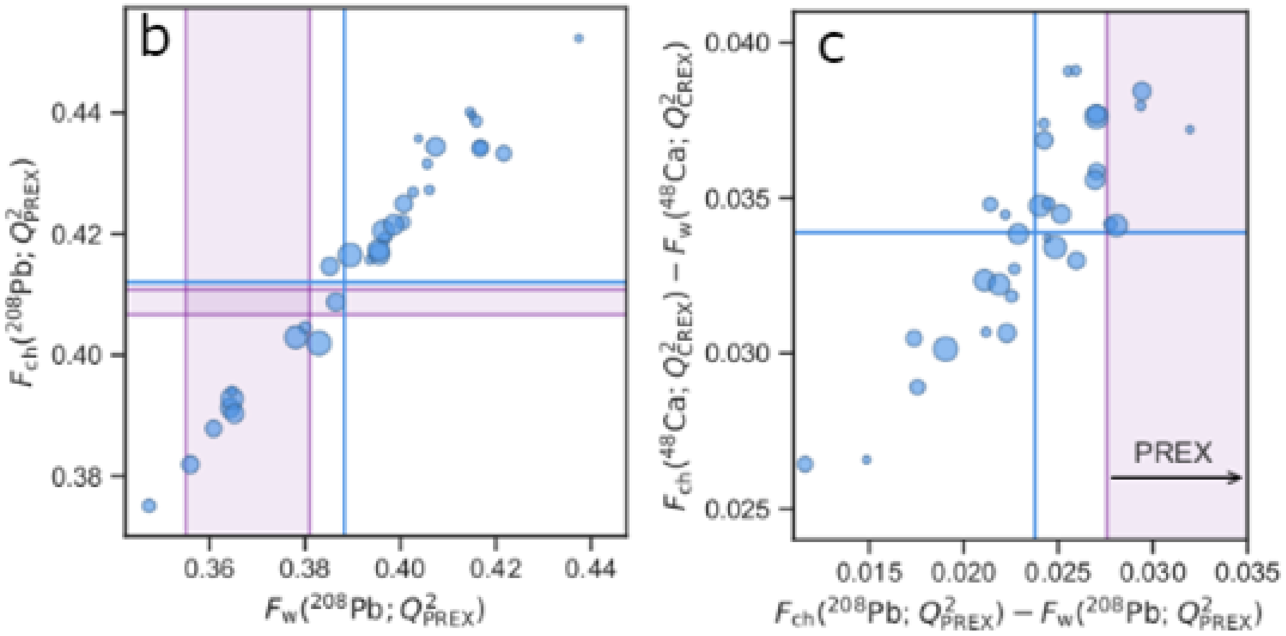
Results and Discussion



The **PPD** obtained from samples of the LECs posterior distribution as shown in previous slide. The bivariate histograms and the corresponding contour lines denote the joint distribution of observables generate by S/IR (blue) and MCMC sampling (orange). The marginal distributions of the observables are shown in the diagonal panels.

Application – neutron skin of ^{208}Pb



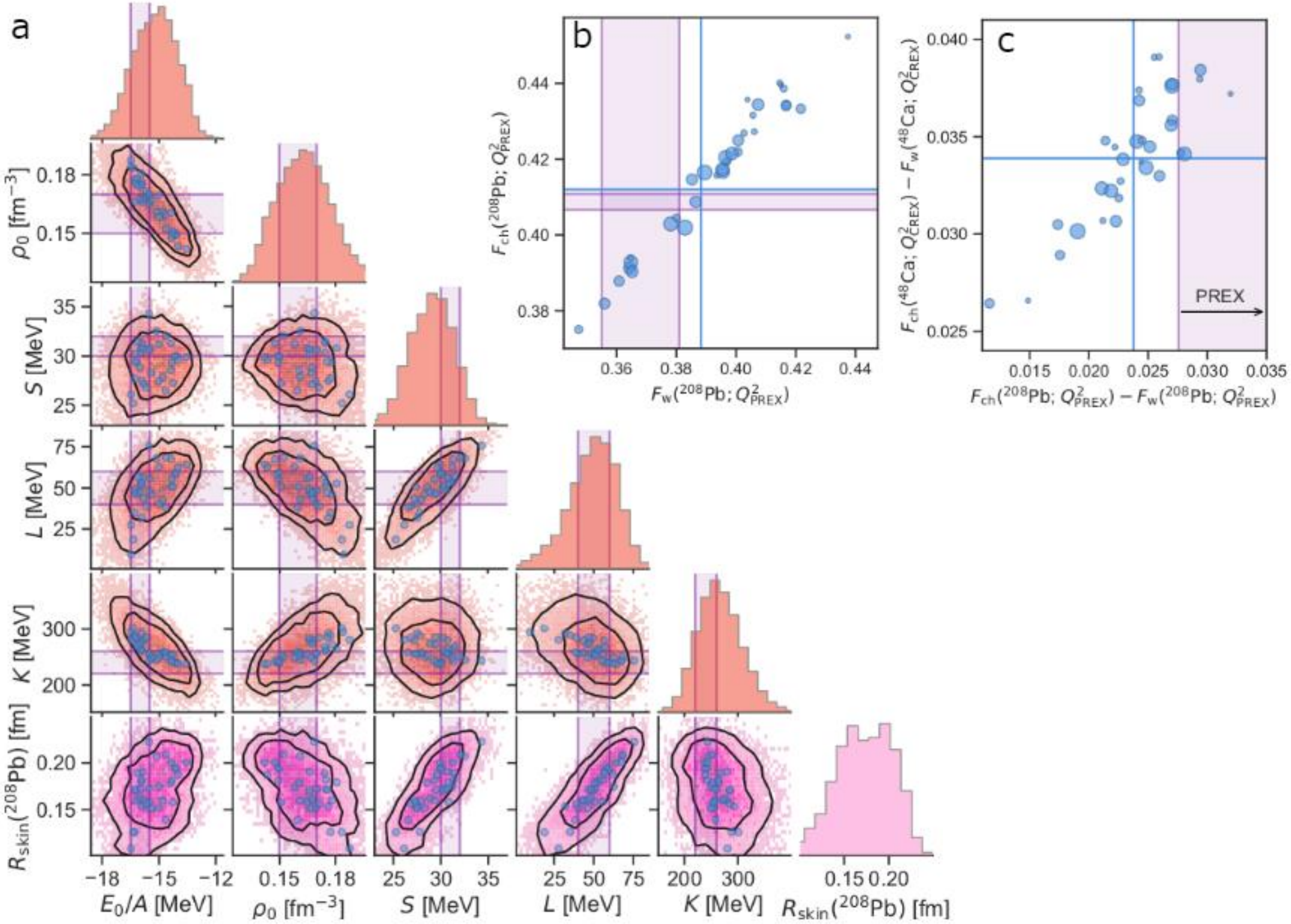


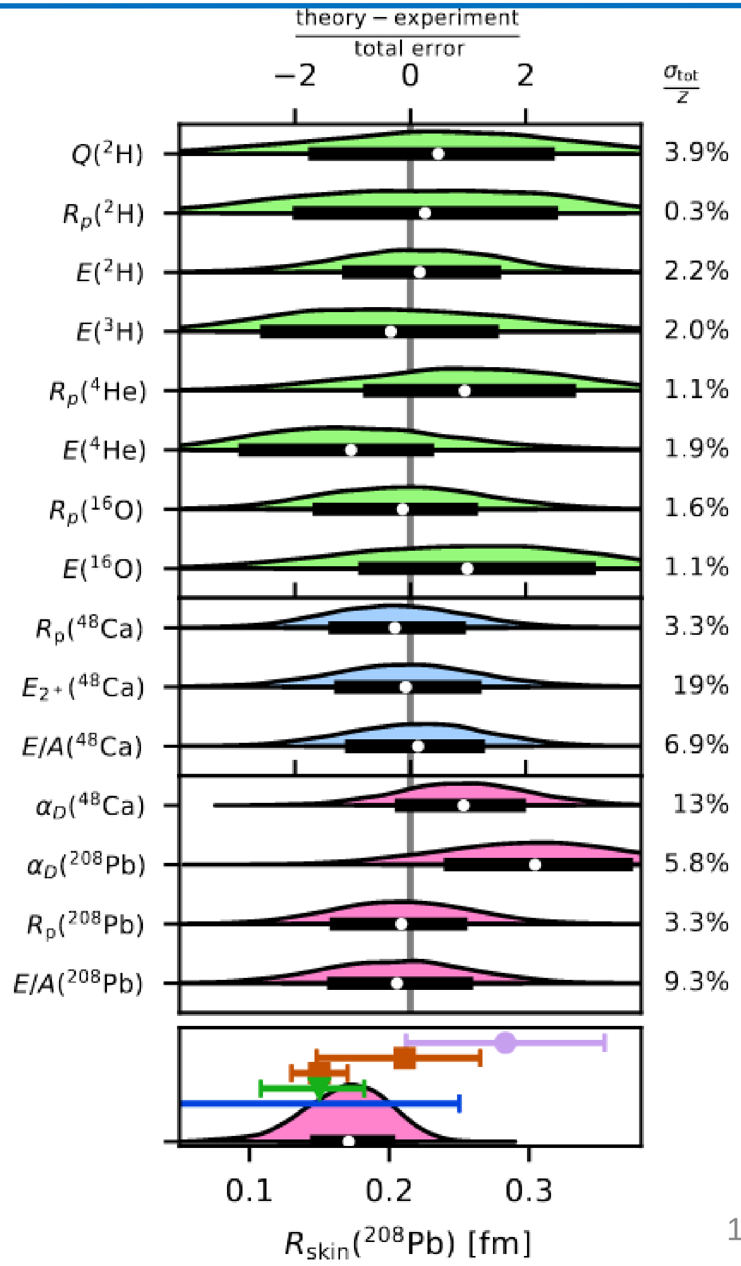
Ab initio calculation of ^{208}Pb
with 3NF up to E_{3max} = 28 by Takayuki etc.

History matching: 34 non-implausible interactions

B. S. Hu*, et al. *"Ab initio predictions link the neutron skin of 208Pb to nuclear forces," Nature Phys. 18, 1196-1200 (2022)*

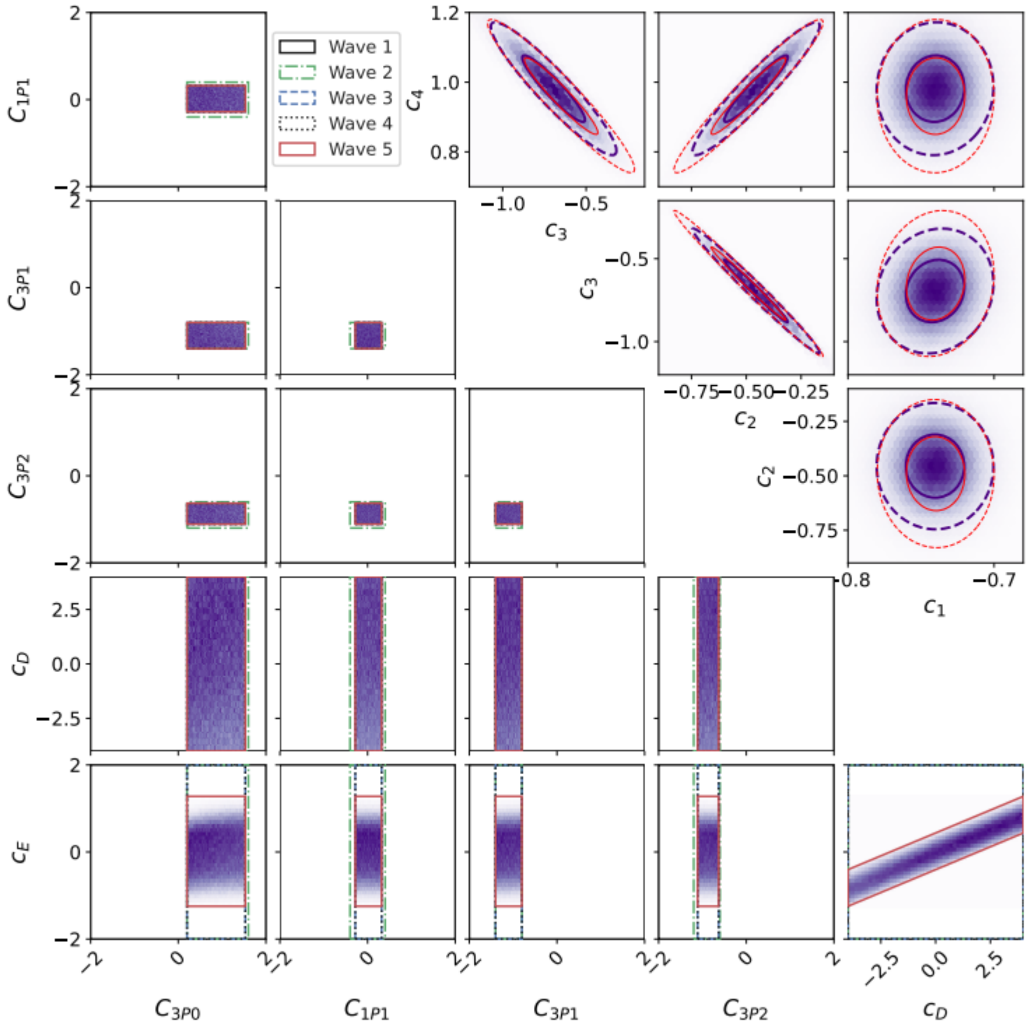
Application – neutron skin of ^{208}Pb





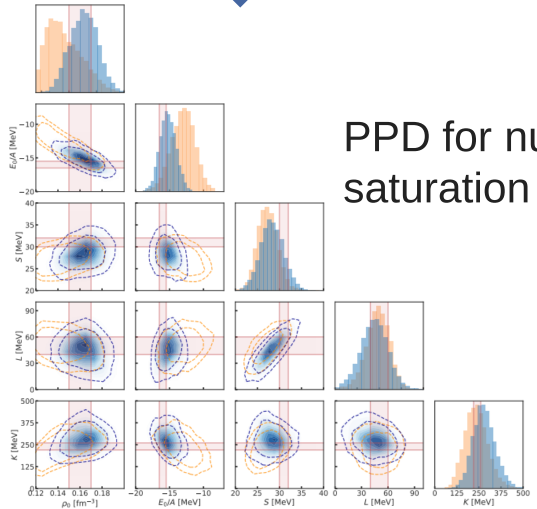
B. S. Hu*, et al. "Ab *initio predictions link the neutron skin of 208Pb to nuclear forces," Nature Phys. 18, 1196-1200 (2022)*

History matching as a good precursor to importance resampling



Observable	z	$\varepsilon_{\mathrm{exp}}$	$\varepsilon_{\mathrm{model}}$	$\varepsilon_{\mathrm{method}}$	$\varepsilon_{\mathrm{em}}$
$E(^2\mathrm{H})$	-2.2298	0.0	0.05	0.0005	0.001%
$r_p(^2\mathrm{H})$	1.976	0.0	0.005	0.0002	0.0005%
$Q(^2\mathrm{H})$	0.27	0.01	0.003	0.0005	0.001%
$E(^3\mathrm{H})$	-8.4818	0.0	0.17	0.0005	0.01%
$E(^4\mathrm{He})$	-28.2956	0.0	0.55	0.0005	0.01%
$r_p(^4\mathrm{He})$	1.455	0.0	0.016	0.0002	0.003%
Predicted observables					
$E(^6\mathrm{Li})$	-31.9940	0.0	0.55	0.2000	0.01%
$E(^{16}\mathrm{O})$	-127.62	0.0	1.00	0.75	0.5%
$r_p(^{16}\mathrm{O})$	2.58	0	0.03	0.01	0.5%

Bayesian inference with SIR



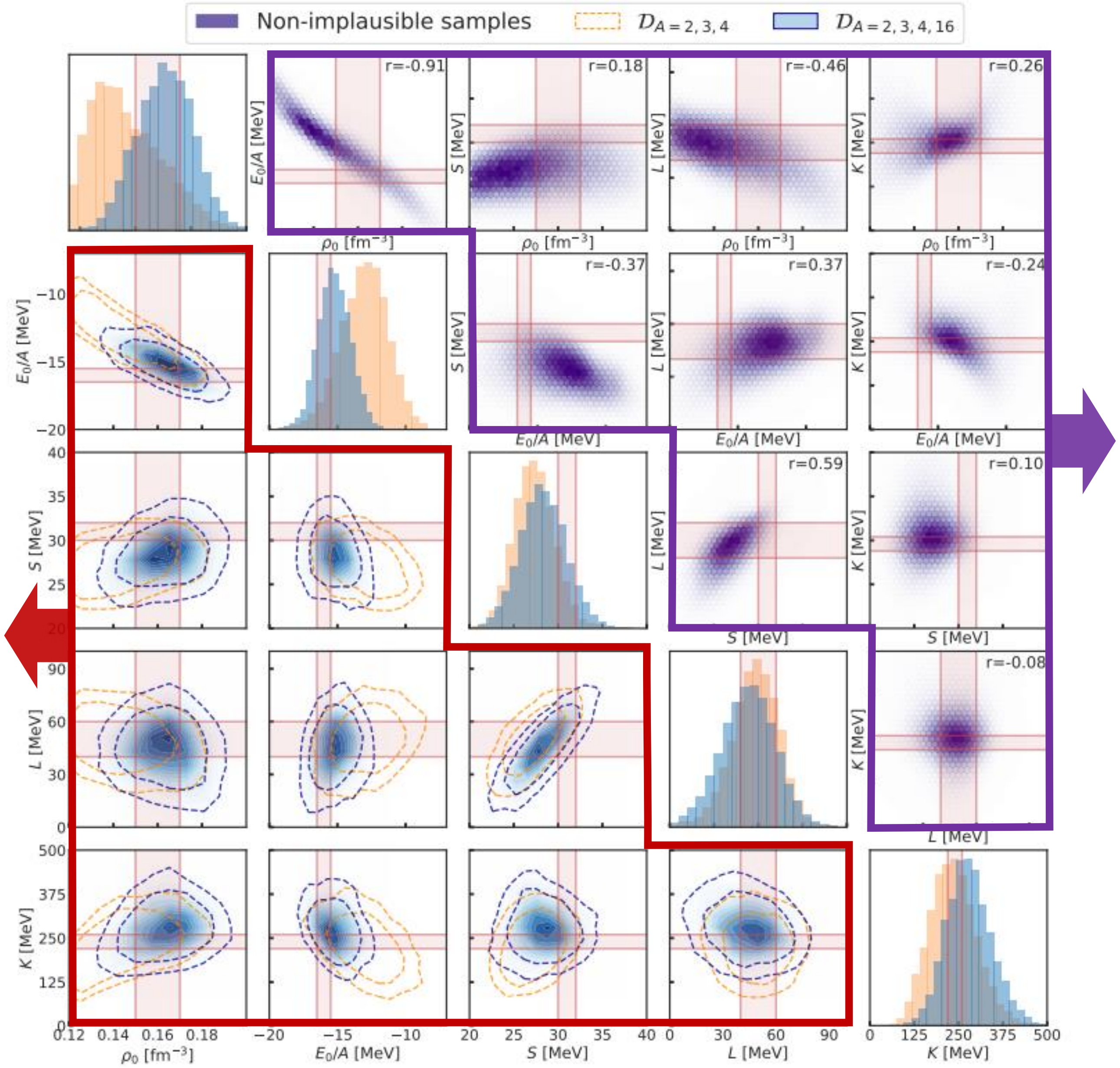
PPD for nuclear matter saturation properties

After history matching, we acquired 10^6 non-implausible interaction samples (out of 1 billion)

Application – nuclear matter

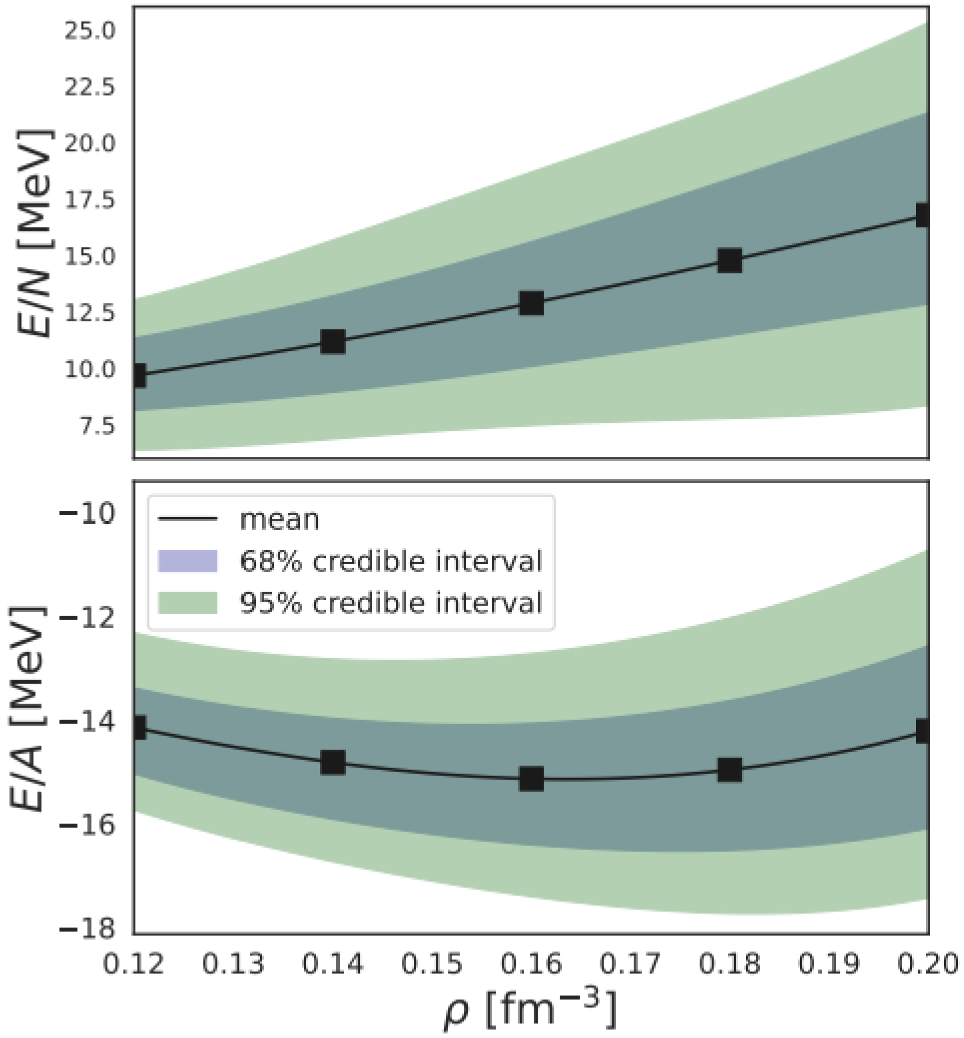
Two PPDs with different PDFs:
$D_{A=2,3,4}$, $D_{A=2,3,4,16}$

Note that the same interaction samples are used for different importance resampling stages.



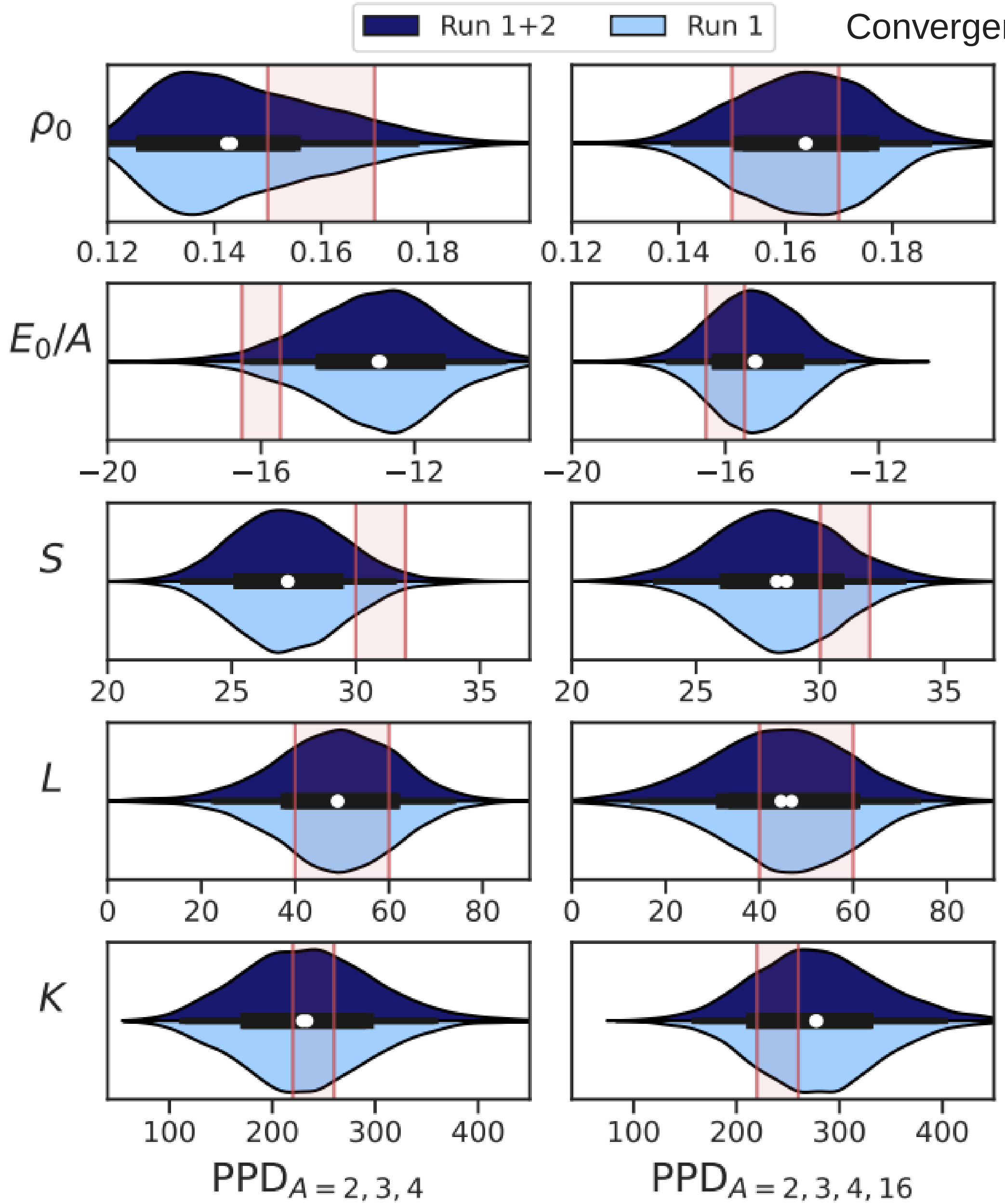
~10^6 Non-implausible interaction samples form history matching

Application – nuclear matter



The PPD for the EOS around saturation density

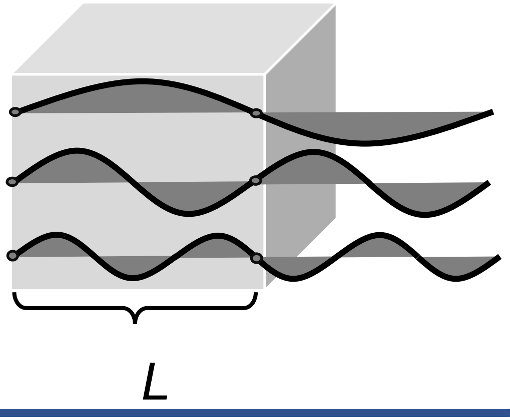
Convergence test



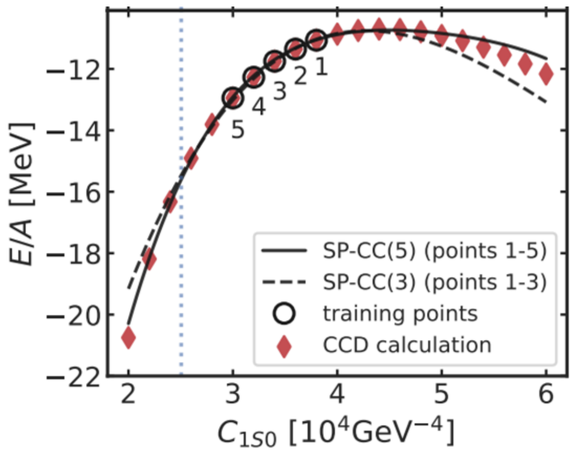
Emulating *ab initio* computations of infinite nucleonic matter

Coupled cluster nuclear matter calculation in momentum space with periodic boundary condition

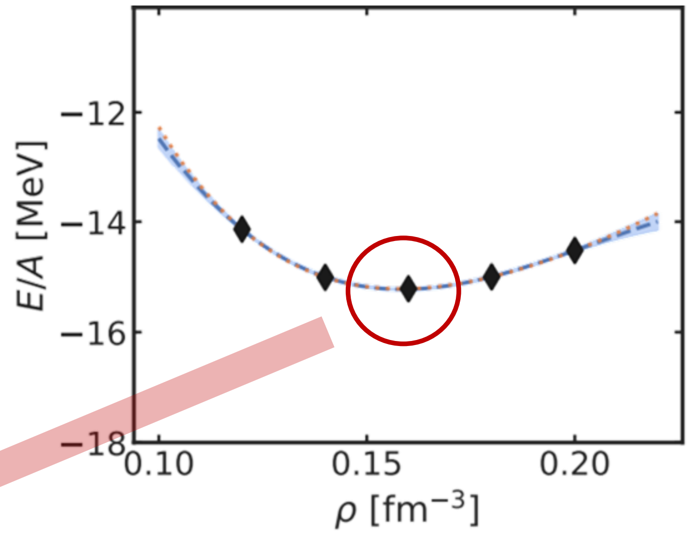
$$H_N \, e^T|\Phi_0\rangle = E \, e^T|\Phi_0\rangle$$

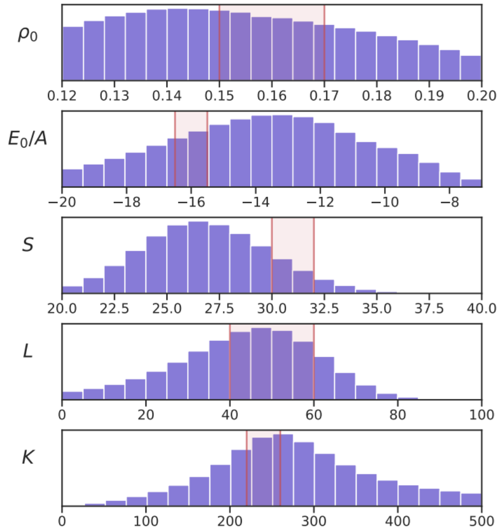


Nuclear matter emulator based on Subspace projected coupled cluster



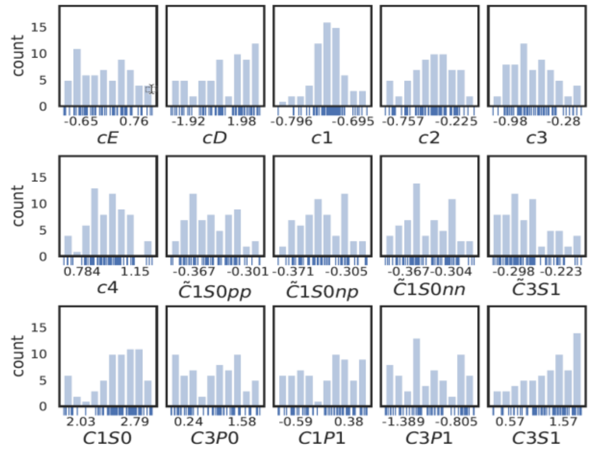
Nuclear matter equation of state for arbitrary interactions

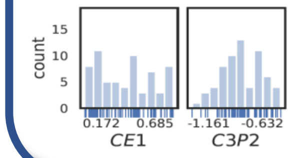




Different saturation properties

Emulator enables 10^6 times acceleration in this case
eg: for SNM (ccd ~200 CPU-hour) vs (emulator ~2ms)



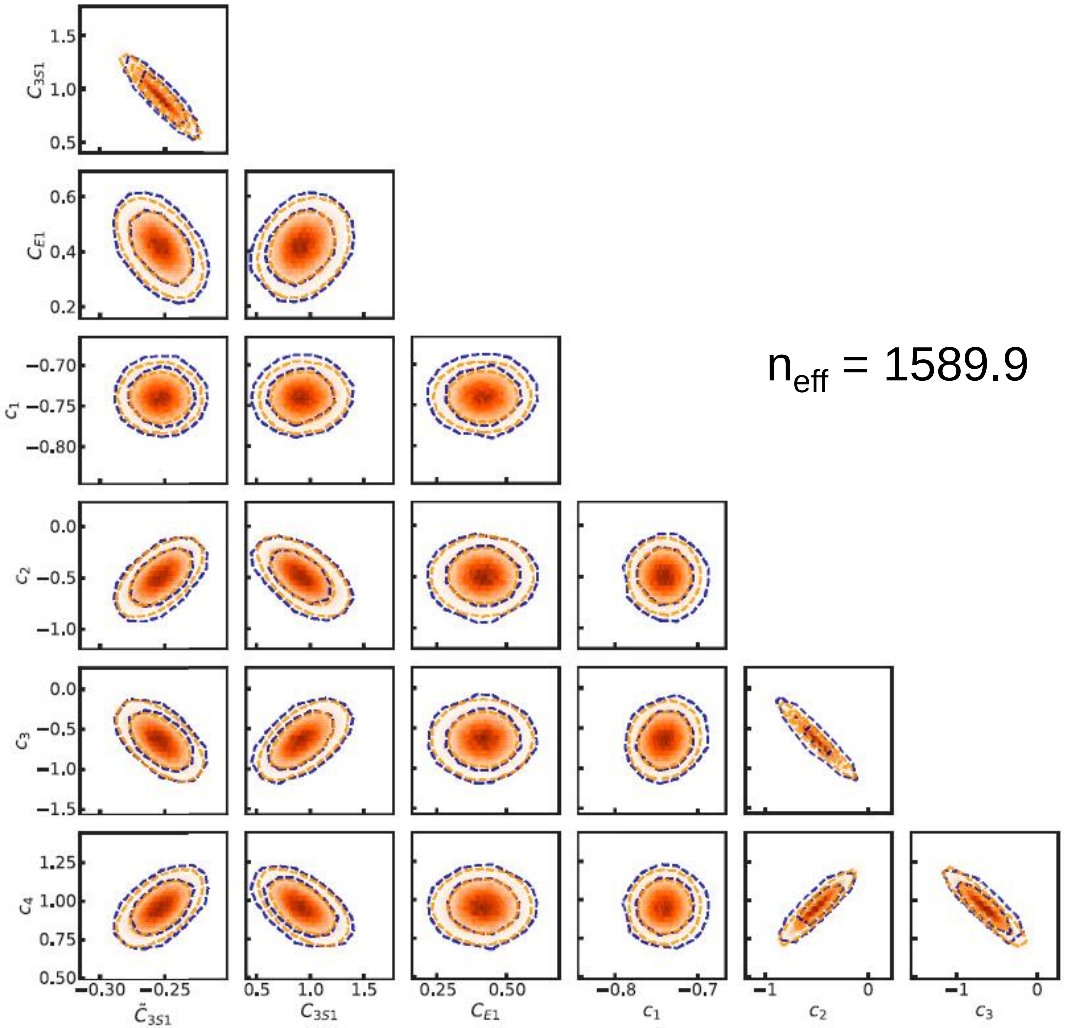


A number of complete CC calculations serves as subspace vector

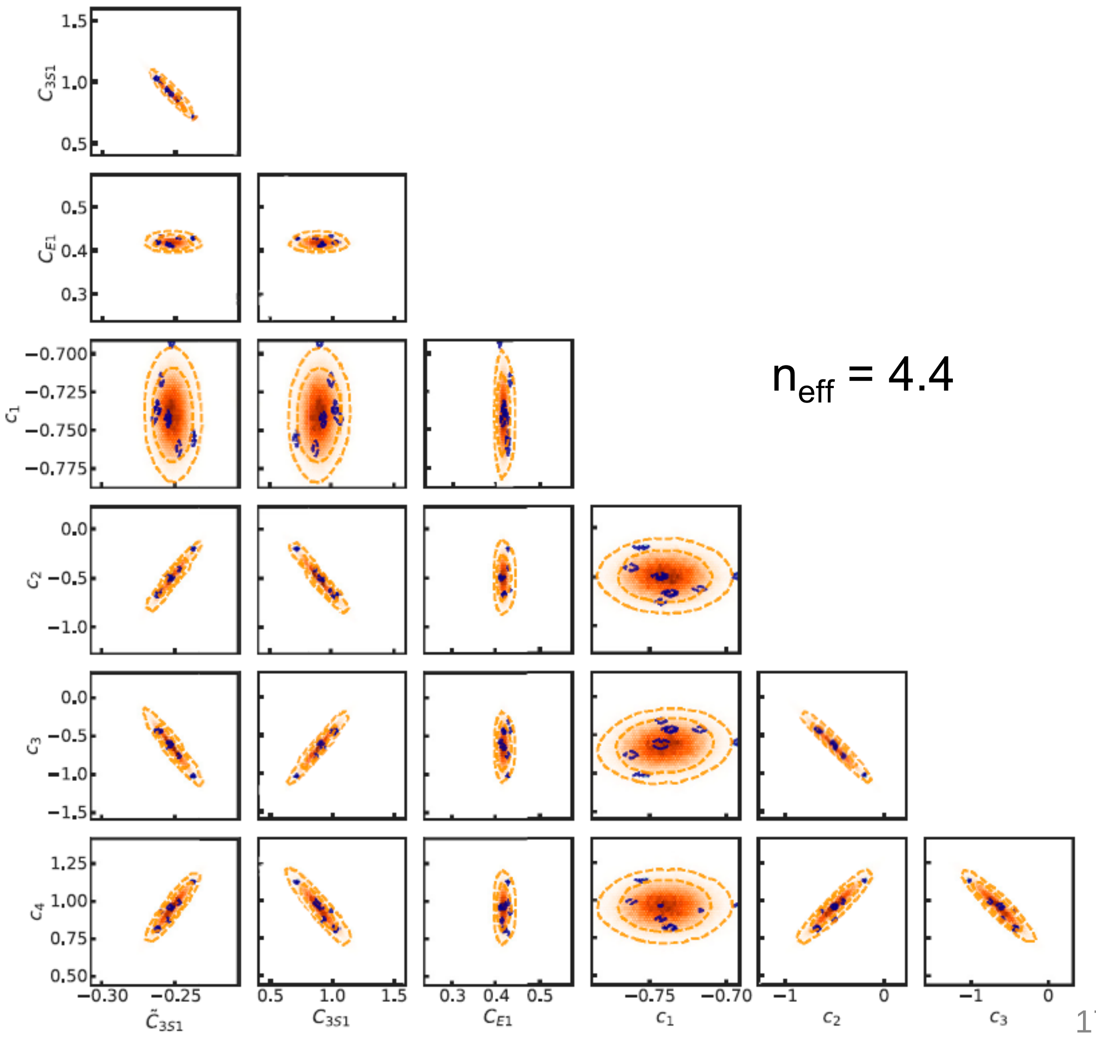
Limitations

1. More complex the likelihood function
2. Target observables are characterized by very small error assignments

Metrics: effective number of samples n_{eff} ,as the sum of rescaled importance weights, $n_{\mathrm{eff}} = \sum_{i=1}^{n}(qi/\max(q))$.



$n_{\mathrm{eff}} = 1589.9$



$n_{\mathrm{eff}} = 4.4$

Summary

- Use the Bayesian method to address the nuclear Hamiltonian (LECs) uncertainty and propagate that to predicted observables.

- In our nuclear physics applications, the Bayesian probability updates are done with sampling/importance resampling to bypass the computational difficulty.

- Limitations of sampling/importance resampling are discussed.

- Open questions: better metrics for the SIR
  other suitable sampling techniques