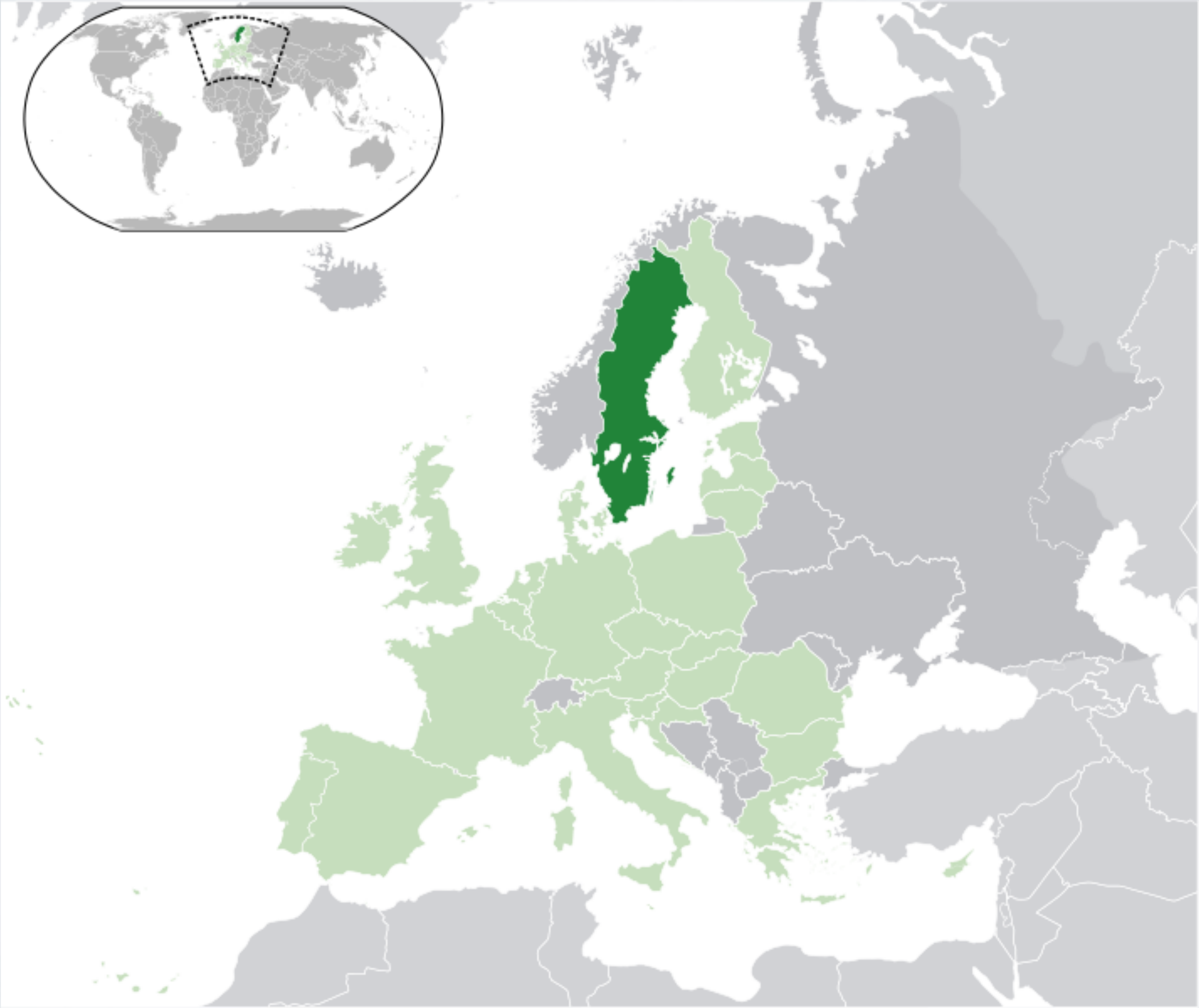
BAYESIAN METHODS IN NUCLEAR PHYSICS

Christian Forssén,
Department of Physics and Astronomy,
Chalmers, Sweden

4th PhyNuBe school: Innovative Technologies, Theories, and Methodologies in Nuclear Physics,

About me: Sweden



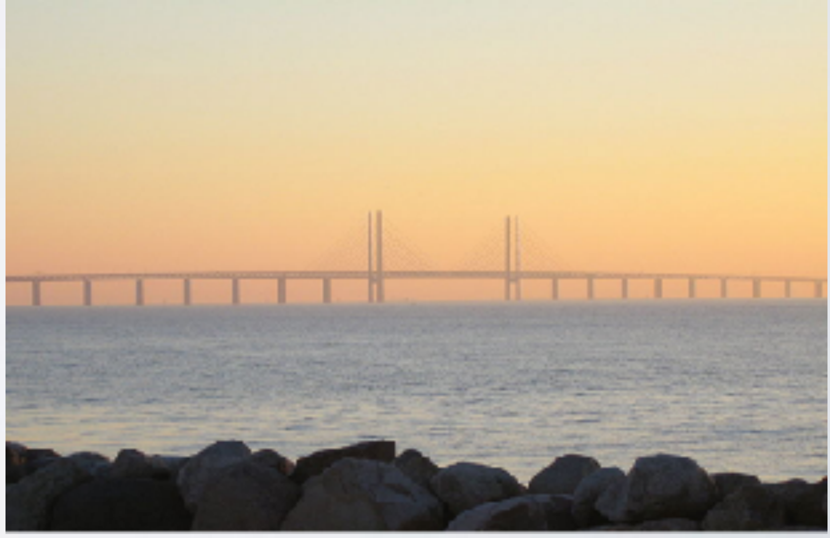
Sweden

North (Stora Sjöfallet)

West coast
(the front)

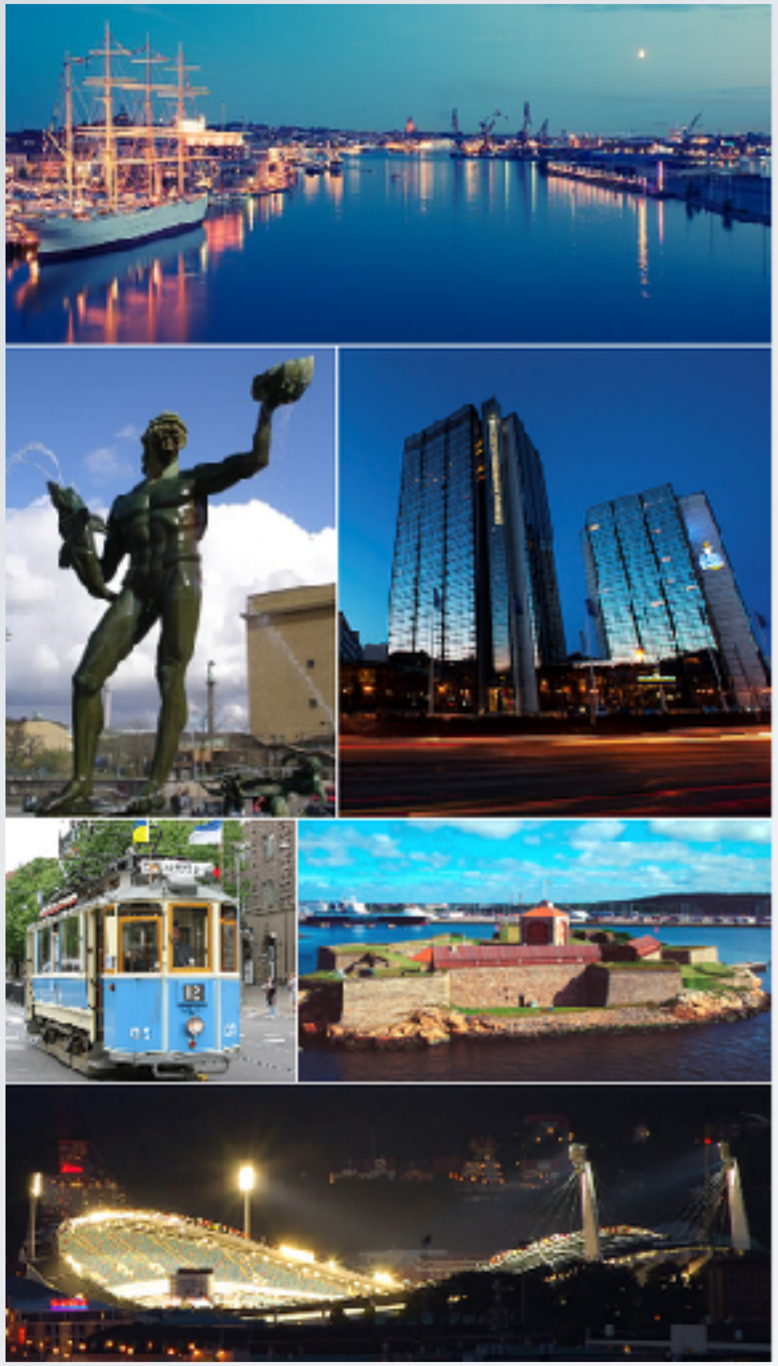
East coast
(the back)

South
(Öresund bridge)



Midsommar

Chalmers University of Technology



Chalmers

# students	~11,000
# phd students	~1,000
# faculty	~800

Theoretical subatomic physics @
Department of Physics and Astronomy

- Nuclear theory
- Neutrino physics
- Particle phenomenology
- Astroparticle theory
  - 4 faculty members
  - 1 tenure track
  - 3 postdocs, 6 phd students
  - + bachelor + master students

Gothenburg, Sweden
pop. ~600,000

About me: Research

The general goal of my theoretical research is to develop *ab initio* approaches to study strongly-interacting few- and many-body systems.

~2003

No-core shell model

~2013

Chiral EFT interactions

~2018

Bayesian methods for UQ

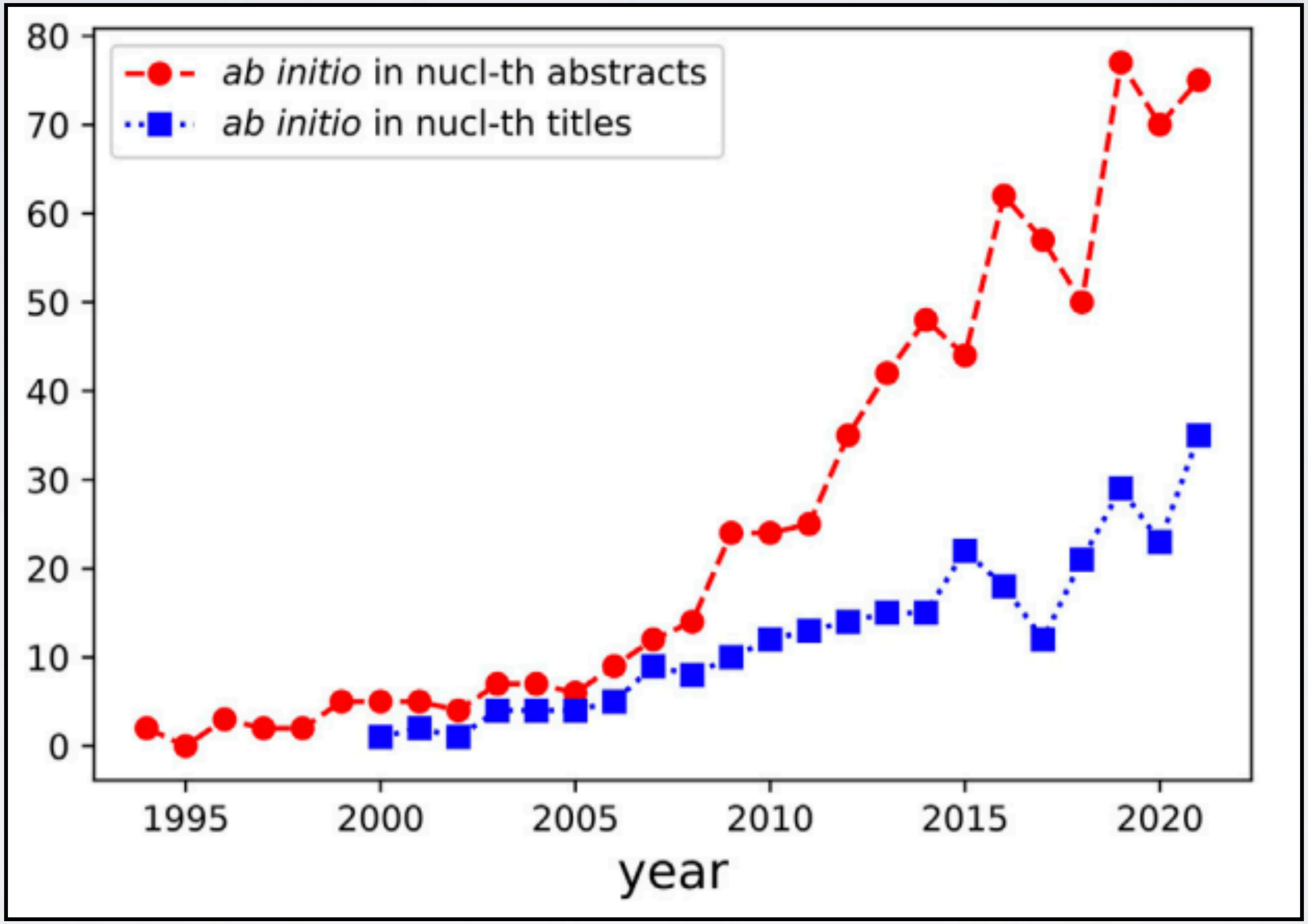
Many-body methods
+ emulators

Effective field theories

Bayesian methods

- **Nuclear interactions**
- **Nuclear structure**
- **Fundamental symmetries**
- **Astrophysics**

What is ab initio in nuclear theory



What is ab initio in nuclear theory

We(*) interpret the *ab initio* method as a "***systematically improvable* approach for quantitatively describing nuclei using the *finest resolution scale* possible while maximizing its *predictive capabilities*.**"

Fine resolution scale

Predictive capability

Systematically improvable

In a nuclear physics context, we let nucleons define the beginning.

Lattice QCD might one day be the optimal starting point.

The systematic aspects of the *ab initio* method creates an inferential advantage.

- What is the precision that can be achieved?

- Does this approach predict emergent phenomena?

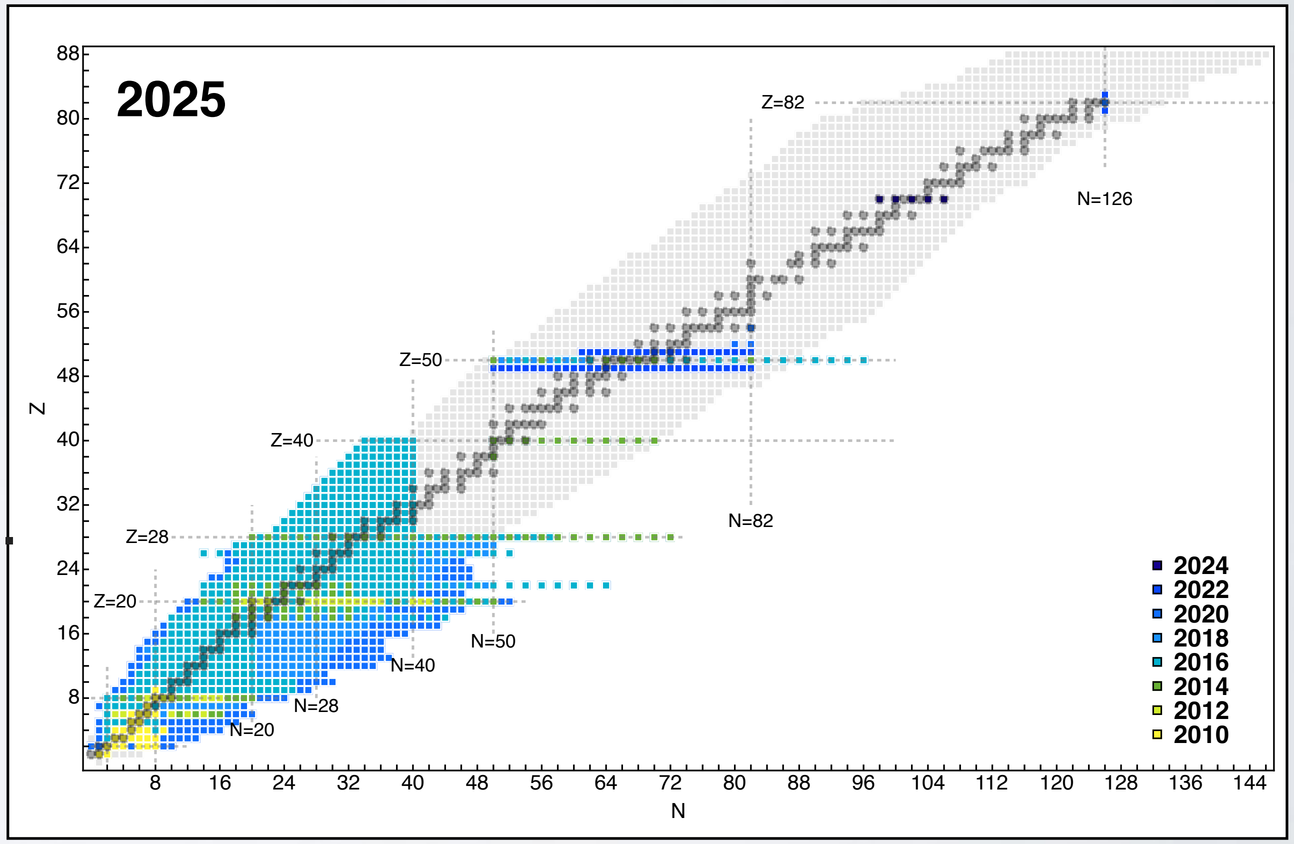


Image credit: H. Hergert

(*)A. Ekström, cf, G.Hagen, G. R. Jansen, W.G. Jiang, and T. Papenbrock, Frontiers in Phys. (2023)

These lectures

Learning outcomes

- Understand different uses of the **probability measure**.
- Derive and understand **Bayes theorem**.
- Set up simple **statistical models** to perform Bayesian inductive inference on physics models conditional on experimental data.
- Become familiar with best practices and the four steps of a **Bayesian workflow**.
- Gain insights into **Markov chain Monte Carlo** (MCMC) sampling which is an enabling technology of Bayesian methods.

Jupyter book

`https://nucleartalent.github.io/LFD_for_Physicists`

Lecture plan

Monday, May 18
09:30-10:30
Lecture 1: *Basics of Bayesian statistics and parameter estimation*
16:00-19:00
Tutorial: *Getting familiar with Bayes*

Tuesday, May 19
09:30 -10:30
Lecture 3: *The Bayesian workflow and MCMC sampling*
16:00-19:00
Tutorial: *Getting familiar with Bayes*

Let’s begin!

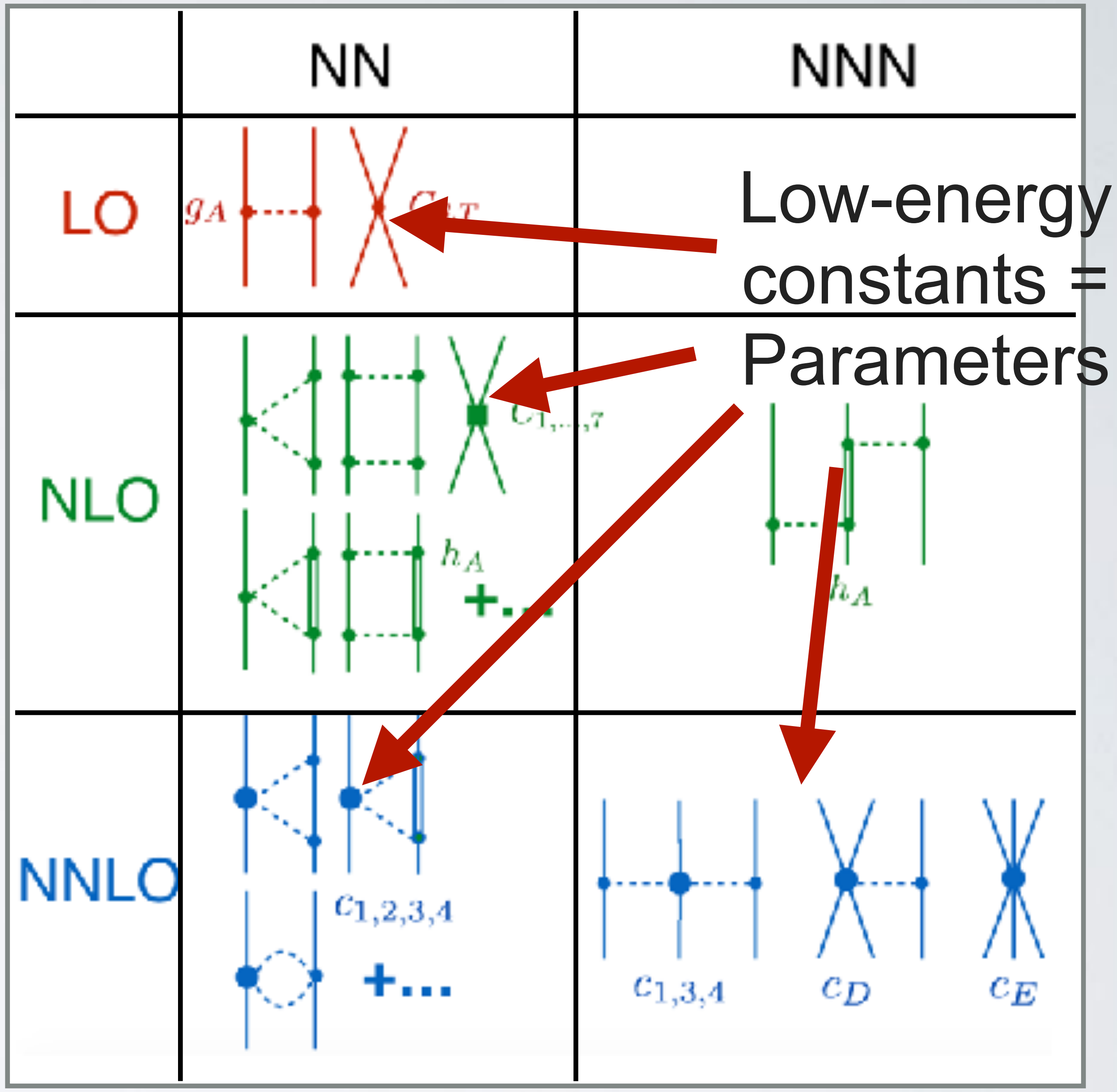
Use your favorite device:

socrative.com

Student login: CHFORSSEN

Chiral effective field theory

- Systematic expansion of nuclear forces in relevant momentum Q over breakdown scale Λ_b:
  - Based on **symmetries** of QCD
  - Pions and nucleons as explicit **degrees of freedom**
  - Λ_b : Chiral symmetry breaking scale, rho-meson mass, delta-nucleon mass difference
  - Power counting scheme enables **hierarchy** of nuclear interaction diagrams
  - Systematic expansion ➡ truncation allows **uncertainty estimates**!
  - Low-energy constants **inferred** from low-energy nuclear observables

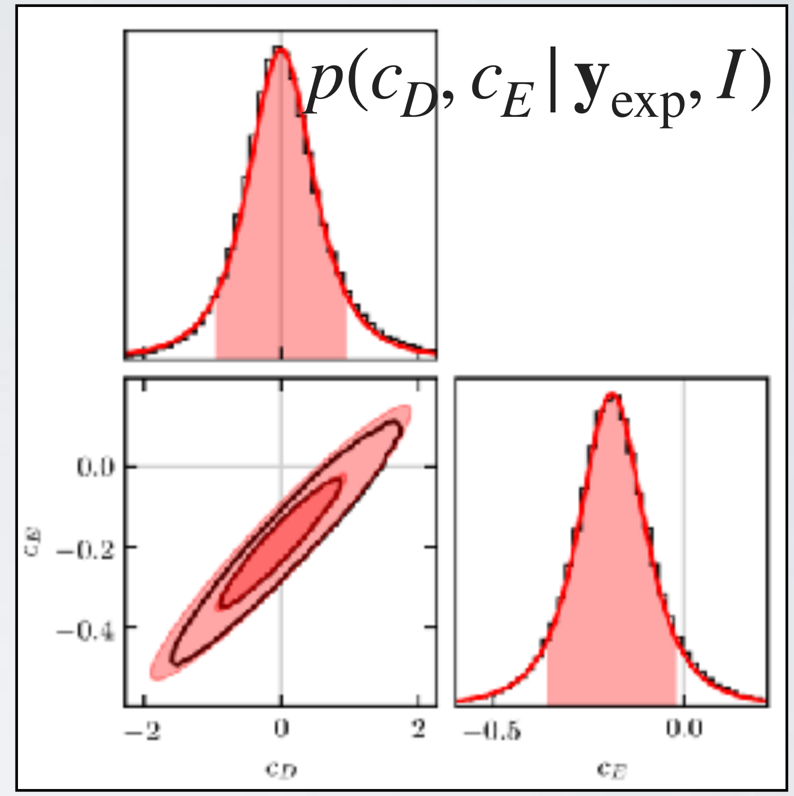


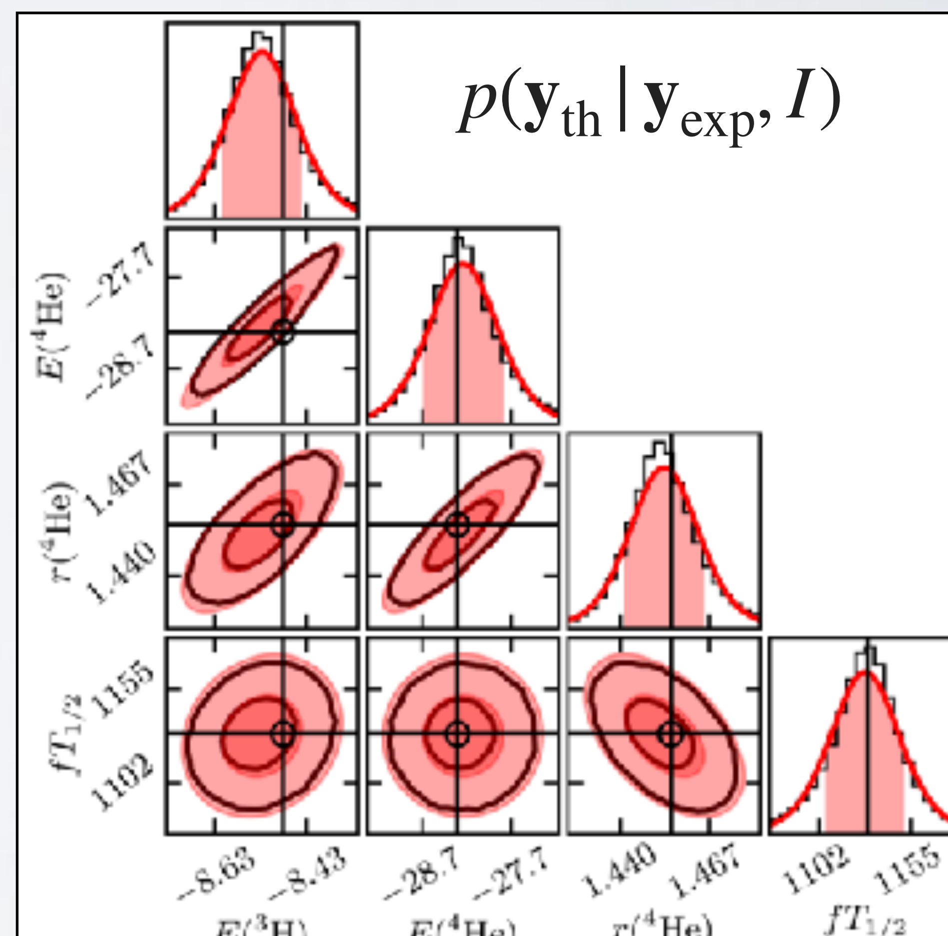
Weinberg, van Kolck, Kaiser, Bernard, Meißner, Epelbaum, Machleidt, Entem, …

Physics example #1

Rigorous constraints on three-nucleon forces in chiral effective field theory from fast and accurate calculations of few-body observables

S. Wesolowski[1,*], I. Svensson[2,†], A. Ekström[2,‡], C. Forssén[2,§], R. J. Furnstahl[3,‖], J. A. Melendez[3,¶], and D. R. Phillips[4,5,6,#]





- We seek the **posterior probability distribution** $p(c_D, c_E \,|\, \mathbf{y}_{\rm exp}, I)$
- The **data,** $\mathbf{y}_{\rm exp}$, include: ^{3}H ground state energy, ^{3}H β-decay half-life, ^{4}He ground state energy, ^{4}He charge radius
- The **information,** I, includes: data, method, and EFT truncation error models, input NN LECs and covariance, fixed πN LECs
- We set **naturalness priors** on the c_D, c_E LECs and EFT expectation priors on the truncation error model.

Bayesian workflow

Four-step Bayesian workflow in brief

1. Formulate informative priors before new data is used.
2. Define a statistical model relating the physics model and data, including all errors.
3. Compute the posterior probabilities.
4. Do model checking.

Steps 1 and 2: Statistical experimentation

Examples from: “Rigorous constraints on three-nucleon forces in chiral EFT…”

Statistical model

$p(\boldsymbol{y}_{\rm exp} \,|\, \vec{a}, \Sigma, I) = \mathcal{N}(\boldsymbol{y}_{\rm th}, \Sigma)$, with $\Sigma = \Sigma_{\rm exp} + \Sigma_{\rm method} + \Sigma_{\rm EFT}$

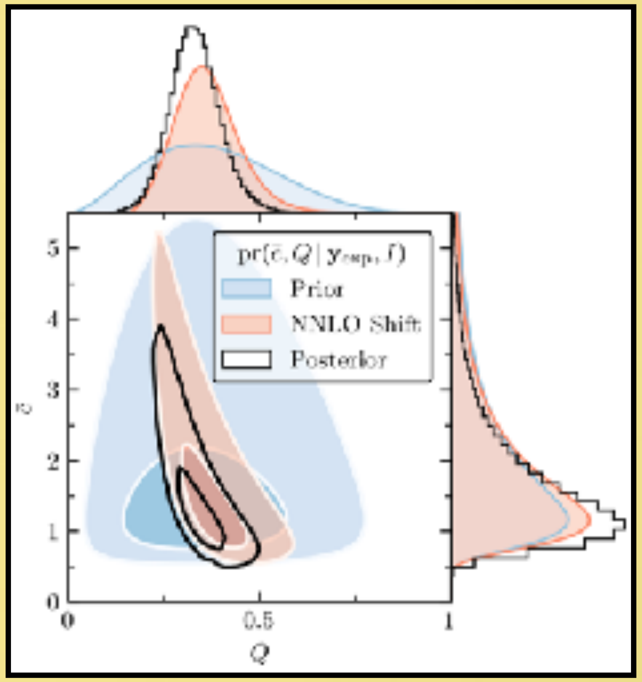
where the EFT truncation error is the largest source of uncertainty and modeled by

$$\left(\Sigma_{\rm th}\right)_{ij} = \left[\frac{(y_{\rm ref}\bar{c}Q^{k+1})^2}{1 - Q^2}\right]\delta_{ij}$$

Priors for all parameters

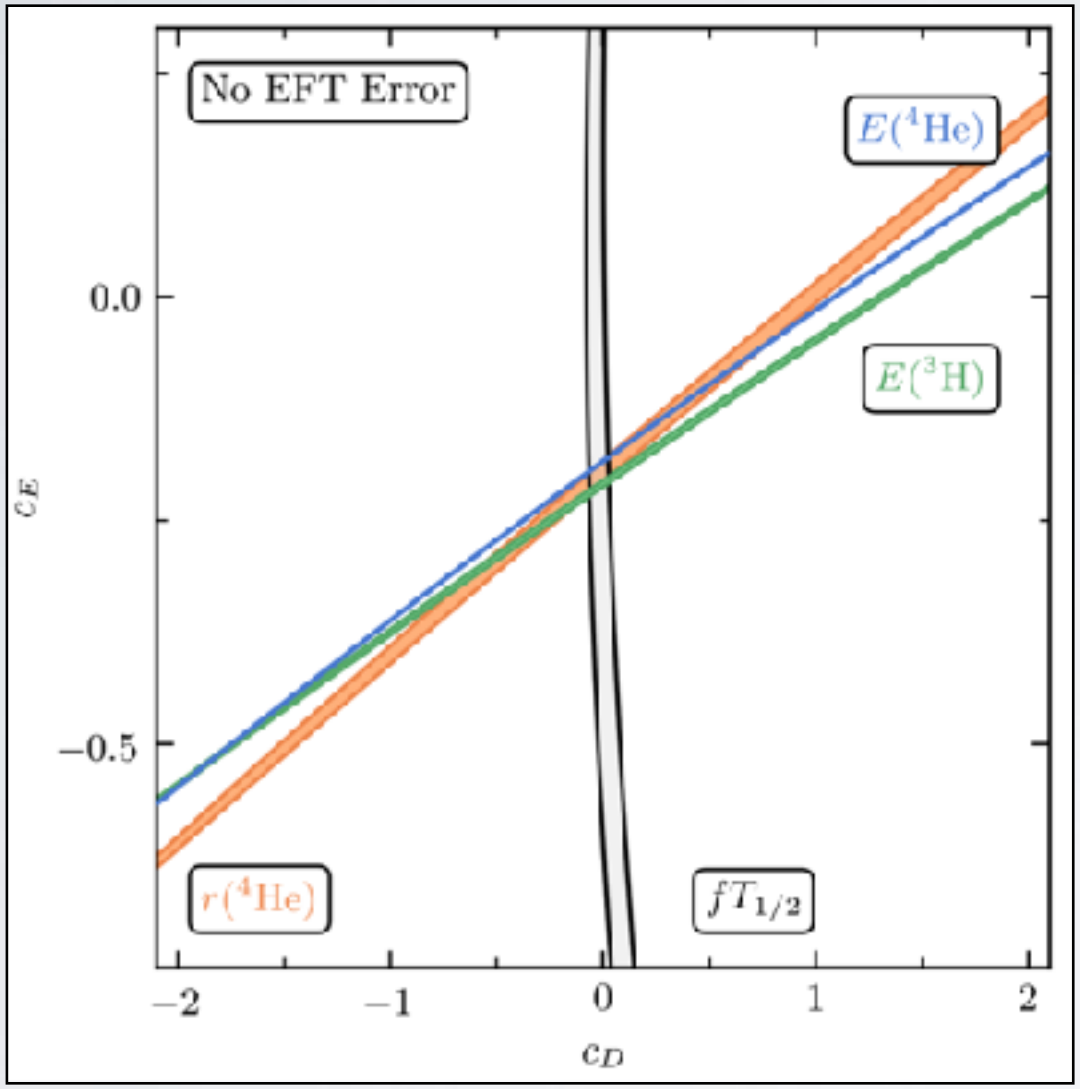
$$p(\vec{a}, \Sigma \,|\, I) = p(c_D, c_E \,|\, I)\; p(\vec{a}_{NN} \,|\, I)\; p(\bar{c} \,|\, I)\; p(Q \,|\, I),$$

where we, e.g., incorporate a naturalness expectation via
$p(c_D, c_E \,|\, I) = \mathcal{N}(0, \bar{a}^2 I_{2\times 2})$

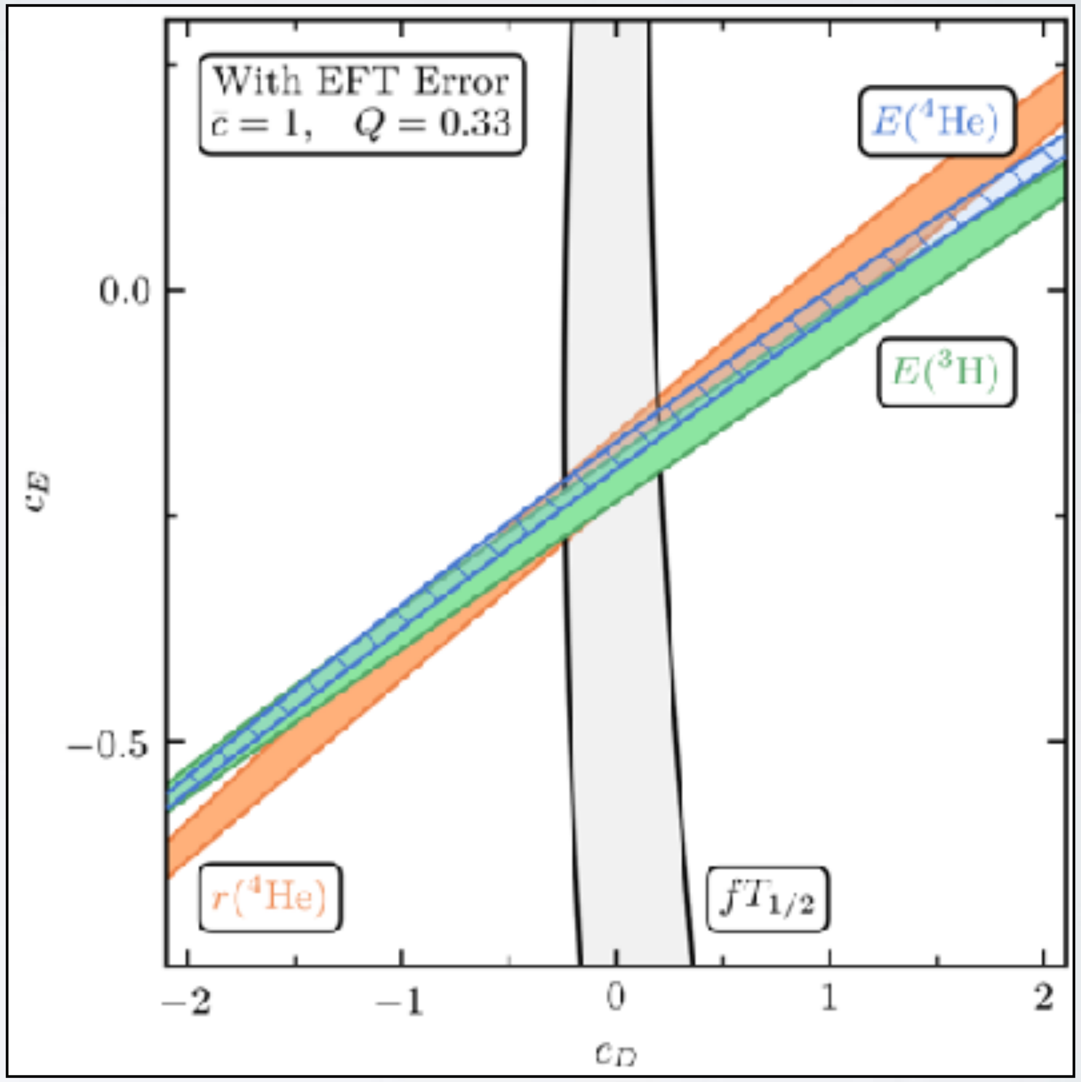


One-by-one observable analysis

What constraints are provided by each of our available observables?



Note: no mutual overlap is possible without EFT truncation error!



Now, with a model for the EFT truncation error

Steps 3 and 4: Posterior computation and model checking

Examples from: “Rigorous constraints on three-nucleon forces in chiral EFT…”

Posterior computation

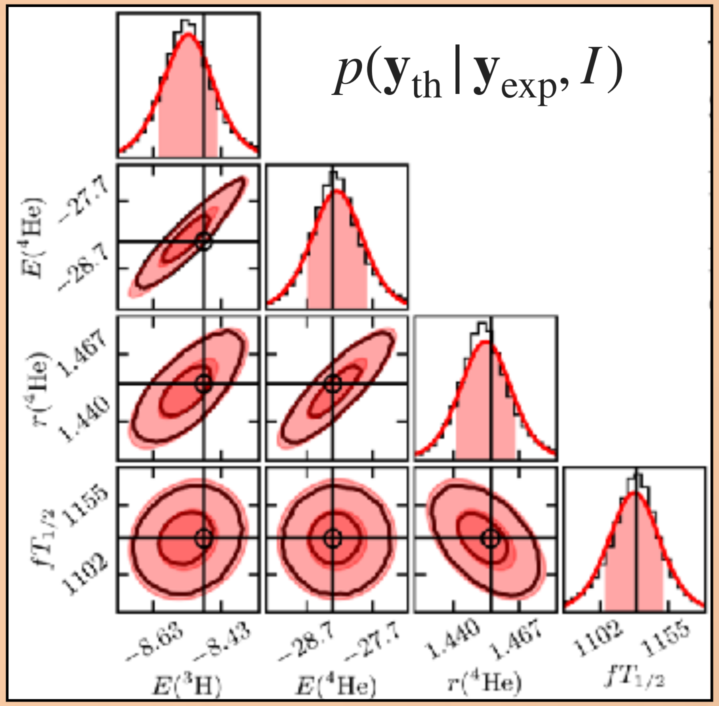
$$p(\vec{a}, \bar{c}, Q \,|\, \boldsymbol{y}_{\text{exp}}, I) = p(\boldsymbol{y}_{\text{exp}} \,|\, \vec{a}, \Sigma, I)\; p(c_D, c_E \,|\, I)\; p(\vec{a}_{NN} \,|\, I)\; p(\bar{c} \,|\, I)\; p(Q \,|\, I)$$

which is impossible to evaluate on a fine grid (14 dimensions);
every evaluation requires a model prediction of all observables in the likelihood.

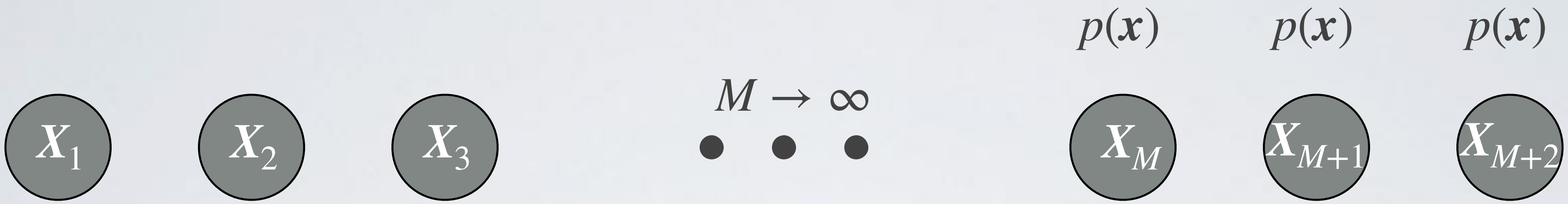
The enabling technologies are MCMC sampling and emulators.

Model checking with posterior predictions

PPD = $\left\{ \boldsymbol{y}_{\text{th}}(\vec{a}) \;:\; \vec{a} \sim p(\vec{a} \,|\, \boldsymbol{y}_{\text{exp}}, I) \right\}$



Markov chain Monte Carlo sampling



$$p_{X_{n+1}|X_n}(\boldsymbol{x}'|\boldsymbol{x}) = T(\boldsymbol{x}, \boldsymbol{x}')$$

- Under certain conditions, and with a clever construction one can design the chain such that $p(\boldsymbol{x}) = p(\boldsymbol{x} \,|\, \mathscr{D}, I)$
- Step proposal + acceptance: $T(\boldsymbol{x}, \boldsymbol{x}') = S(\boldsymbol{x}, \boldsymbol{x}')\ \mathbb{A}(\boldsymbol{x}, \boldsymbol{x}')$

with $\mathbb{A}(\boldsymbol{x}, \boldsymbol{x}') = \min\left(\dfrac{p(\boldsymbol{x}' \,|\, \mathscr{D}, I)}{p(\boldsymbol{x} \,|\, \mathscr{D}, I)}, 1\right)$ (for symmetric S)

See **interactive demo** in chapter 18.3.

Convergence of MCMC

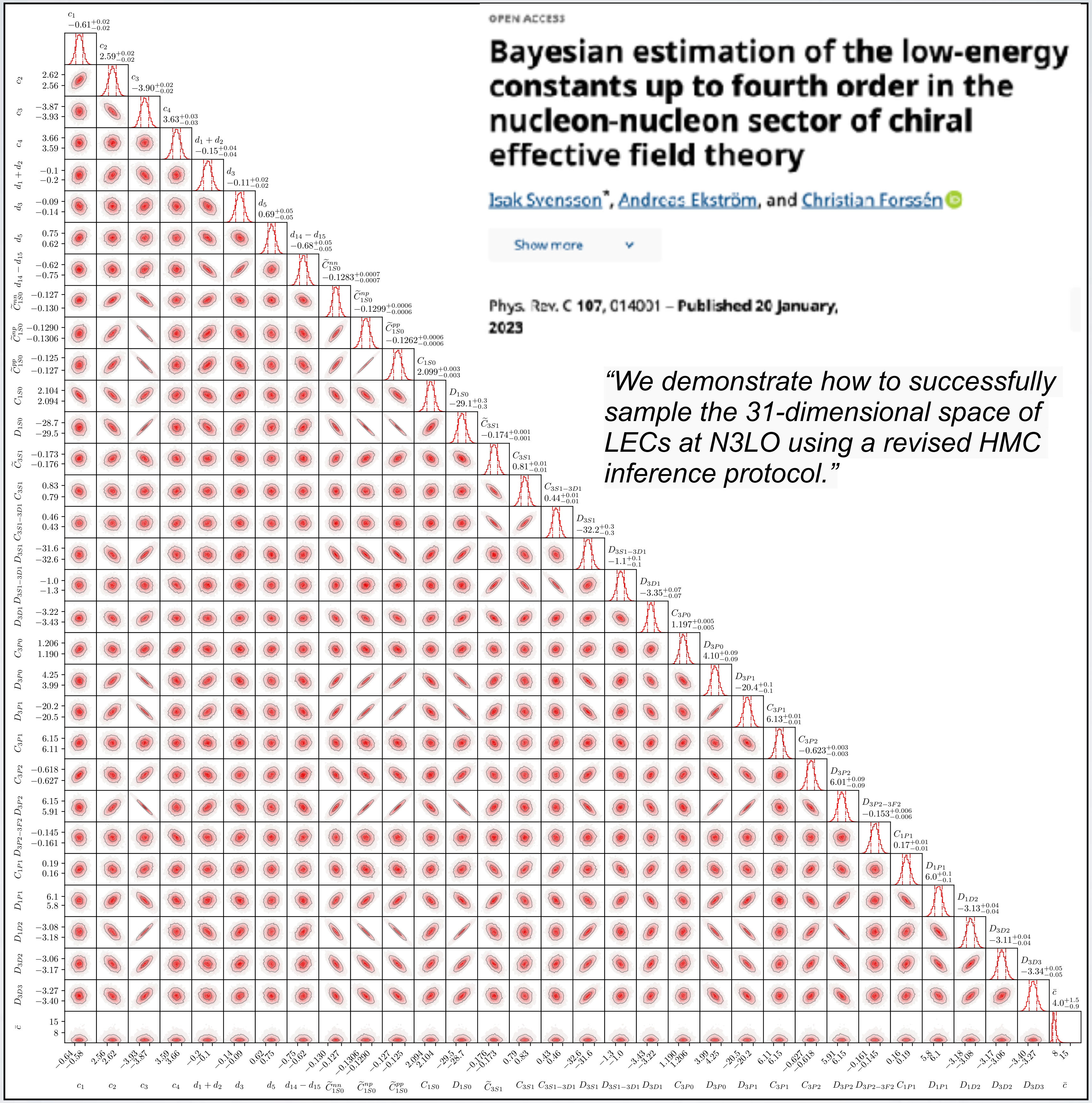
MCMC sampling can go wrong…

- Not converged (**not equilibrated**), which implies that the limiting distribution is not reached.
- **Pseudoconvergence**. It looks converged, but the chain is stuck temporarily in a local mode.
- Highly **correlated** samples. Not really wrong, but highly inefficient. The effective sample size is: ESS = N / T, where T is the autocorrelation time and N is the number of samples.

Recommendations:

- Inspect trace plots (chain position as a function of time).
- Estimate the autocorrelation time. Certain algorithms are better at producing uncorrelated samples.
- Tuning of sampling algorithm; try different ones.
- Run multiple sequences of the Markov chain with different starting points.
- Compare within-chain and between-chain variances (Gelman-Rubin test).

Physics example #2



OPEN ACCESS

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

Show more

Phys. Rev. C 107, 014001 – Published 20 January, 2023

"We demonstrate how to successfully sample the 31-dimensional space of LECs at N3LO using a revised HMC inference protocol."

Physics example #3

nature physics
ARTICLES
https://doi.org/10.1038/s41567-022-01715-8
Check for updates

OPEN

Ab initio predictions link the neutron skin of ^{208}Pb to nuclear forces

Baishan Hu[1,11], Weiguang Jiang[2,11], Takayuki Miyagi[1,3,4,11], Zhonghao Sun[5,6,11], Andreas Ekström[2], Christian Forssén[2] ✉, Gaute Hagen[1,5,6], Jason D. Holt[1,7], Thomas Papenbrock[5,6], S. Ragnar Stroberg[8,9] and Ian Vernon[10]

- The **thickness of the neutron skin** in finite nuclei shows a correlation with the **stiffness of the neutron matter equation-of-state**.

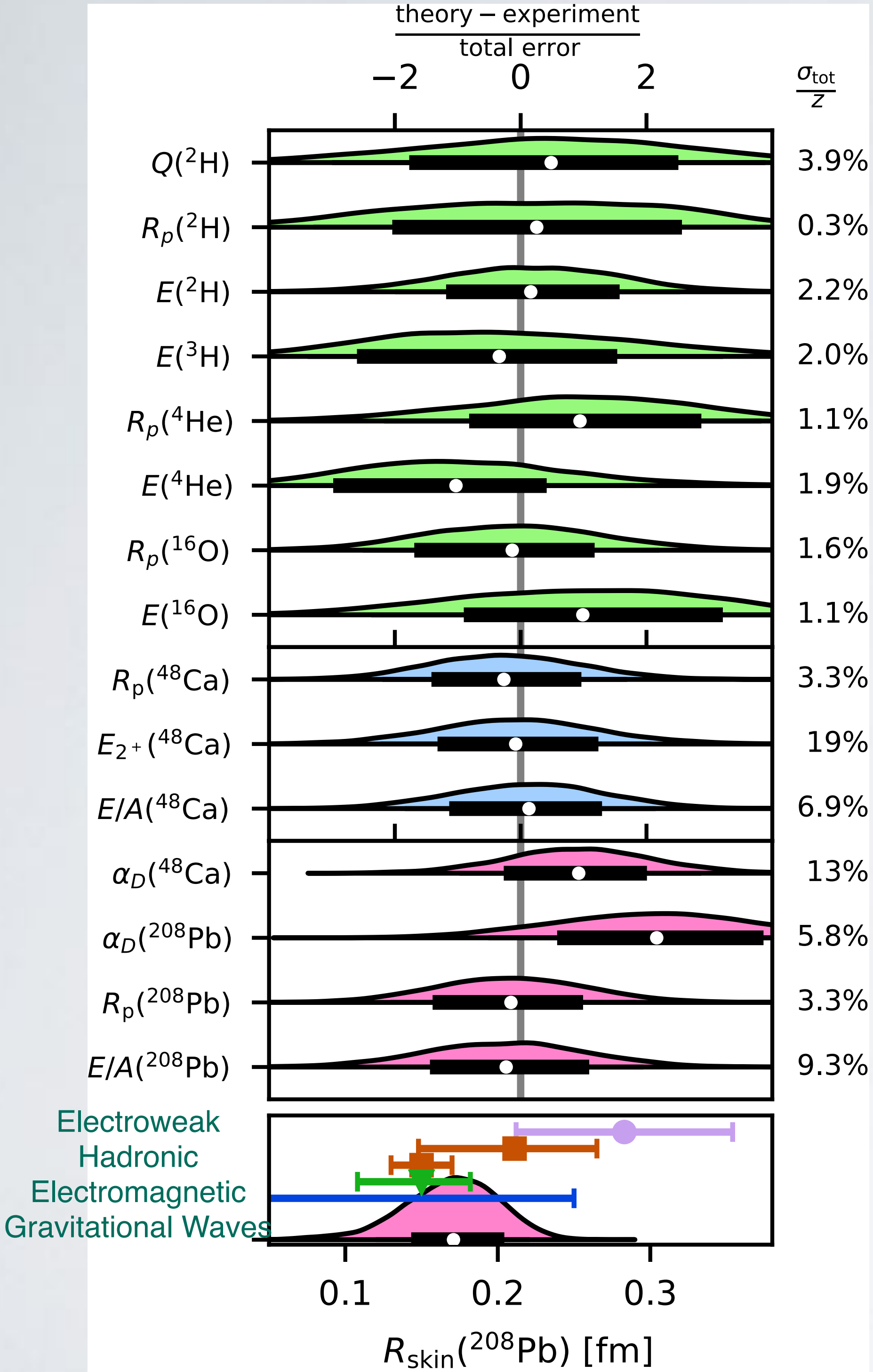
- Starting from $\chi\mathrm{EFT}$ **with explicit Δ isobar** (➡ higher breakdown scale)

- Construct an extensive **error model** (EFT truncation, method convergence, finite-size errors).

- Employ **iterative history-matching** for global parameter search. Study *ab initio* model performance, and provide a finite number of non-implausible samples.

- We produce Bayesian **posterior predictive** distributions for nuclear observables up to ^{208}Pb and for infinite nuclear matter properties.

Ab initio predictions link the skin of ^{208}Pb to nuclear forces



History Matching

We explore 10^9 different interaction parameterizations

Confronted with A=2-16 data + NN scattering information

Find 34 non-implausible interactions

Calibration

Importance resampling

Validation

Inspect ab initio model and error estimates

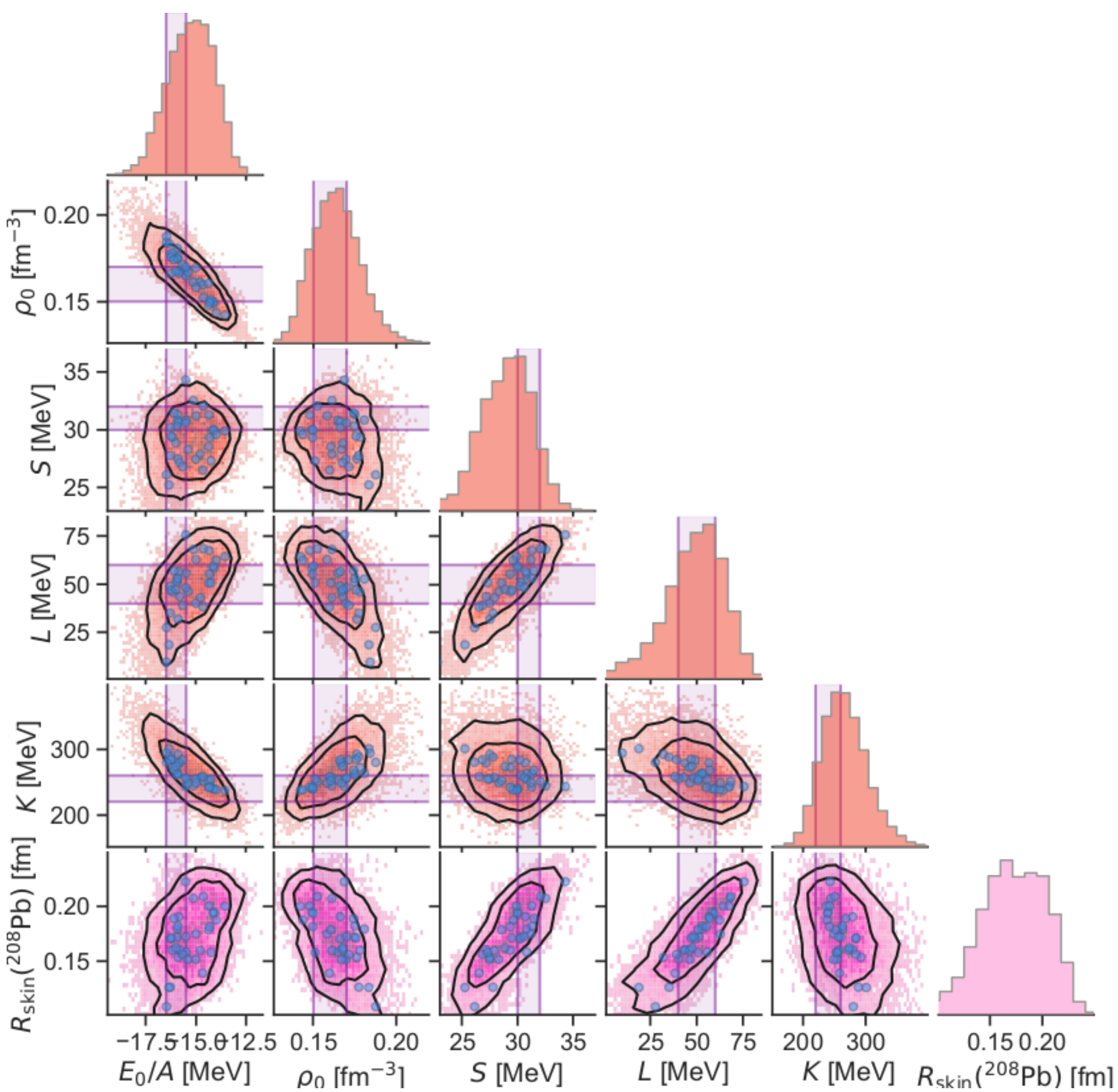
Observable	z	ε_{exp}	ε_{model}	ε_{method}	ε_{em}	PPD
History-matching observables						
$E(^2H)$	-2.2246	0.0	0.05	0.0005	0.001%	$-2.22^{+0.07}_{-0.07}$
$R_p(^2H)$	1.976	0.0	0.005	0.0002	0.0005%	$1.98^{+0.01}_{-0.01}$
$Q(^2H)$	0.27	0.01	0.003	0.0005	0.001%	$0.28^{+0.02}_{-0.02}$
$E(^3H)$	-8.4821	0.0	0.17	0.0005	0.01%	$-8.54^{+0.34}_{-0.37}$
$E(^4He)$	-28.2957	0.0	0.55	0.0005	0.01%	$-28.86^{+0.86}_{-1.01}$
$R_p(^4He)$	1.455	0.0	0.016	0.0002	0.003%	$1.47^{+0.03}_{-0.03}$
$E(^{16}O)$	127.62	0.0	1.0	0.75	0.5%	$-126.2^{+3.0}_{-2.8}$
$R_p(^{16}O)$	2.58	0.0	0.03	0.01	0.5%	$2.57^{+0.06}_{-0.06}$
Calibration observables						
Observable	z	ε_{exp}	ε_{model}	ε_{method}	ε_{em}	PPD
$E/A(^{48}Ca)$	-8.667	0.0	0.54	0.25	—	$-8.58^{+0.72}_{-0.72}$
$E_{2^+}(^{48}Ca)$	3.83	0.0	0.5	0.5	—	$3.79^{+0.86}_{-0.96}$
$R_p(^{48}Ca)$	3.39	0.0	0.11	0.03	—	$3.36^{+0.14}_{-0.13}$
Validation observables						
Observable	z	ε_{exp}	ε_{model}	ε_{method}	ε_{em}	PPD
$E/A(^{208}Pb)$	-7.867	0.0	0.54	0.5	—	$-8.06^{+0.99}_{-0.88}$
$R_p(^{208}Pb)$	5.45	0.0	0.17	0.05	—	$5.43^{+0.21}_{-0.23}$
$\alpha_D(^{48}Ca)$	2.07	0.22	0.06	0.1	—	$2.30^{+0.31}_{-0.26}$
$\alpha_D(^{208}Pb)$	20.1	0.6	0.59	0.8	—	$22.6^{+2.1}_{-1.8}$

B. Hu et al (Nature Phys. 2022)

Prediction: small skin thickness 0.14-0.20 fm in mild (1.5 sigma) tension with PREX.

Predictions for the skin thickness and nuclear matter

Ab initio theory reveals correlations, e.g., between L and R_{skin} previously demonstrated in mean-field models.



Nuclear matter properties			
Observable	median	68% CR	90% CR
E_0/A	-15.2	$[-16.3, -14.0]$	$[-17.2, -13.5]$
ρ_0	0.163	$[0.147, 0.175]$	$[0.140, 0.186]$
S	29.1	$[26.6, 31.3]$	$[25.132.8]$
L	50.3	$[37.2, 68.1]$	$[22.6, 75.8]$
K	264	$[227, 297]$	$[210, 328]$
Neutron skins			
Observable	median	68% CR	90% CR
$R_{skin}(^{48}\mathrm{Ca})$	0.164	$[0.141, 0.187]$	$[0.123, 0.199]$
$R_{skin}(^{208}\mathrm{Pb})$	0.171	$[0.139, 0.200]$	$[0.120, 0.221]$

B. Hu et al (Nature Phys. 2022)