Bayesian inference in *ab initio* nuclear theory

Andreas Ekström
Chalmers University of Technology
andreas.ekstrom@chalmers.se

XIX International Meeting on "Selected Topics in Nuclear and Atomic Physics"
Fiera di Primiero, October 4-8, 2024

CHALMERS
UNIVERSITY OF TECHNOLOGY

Swedish
Research
Council

Evolving nuclear physics landscape

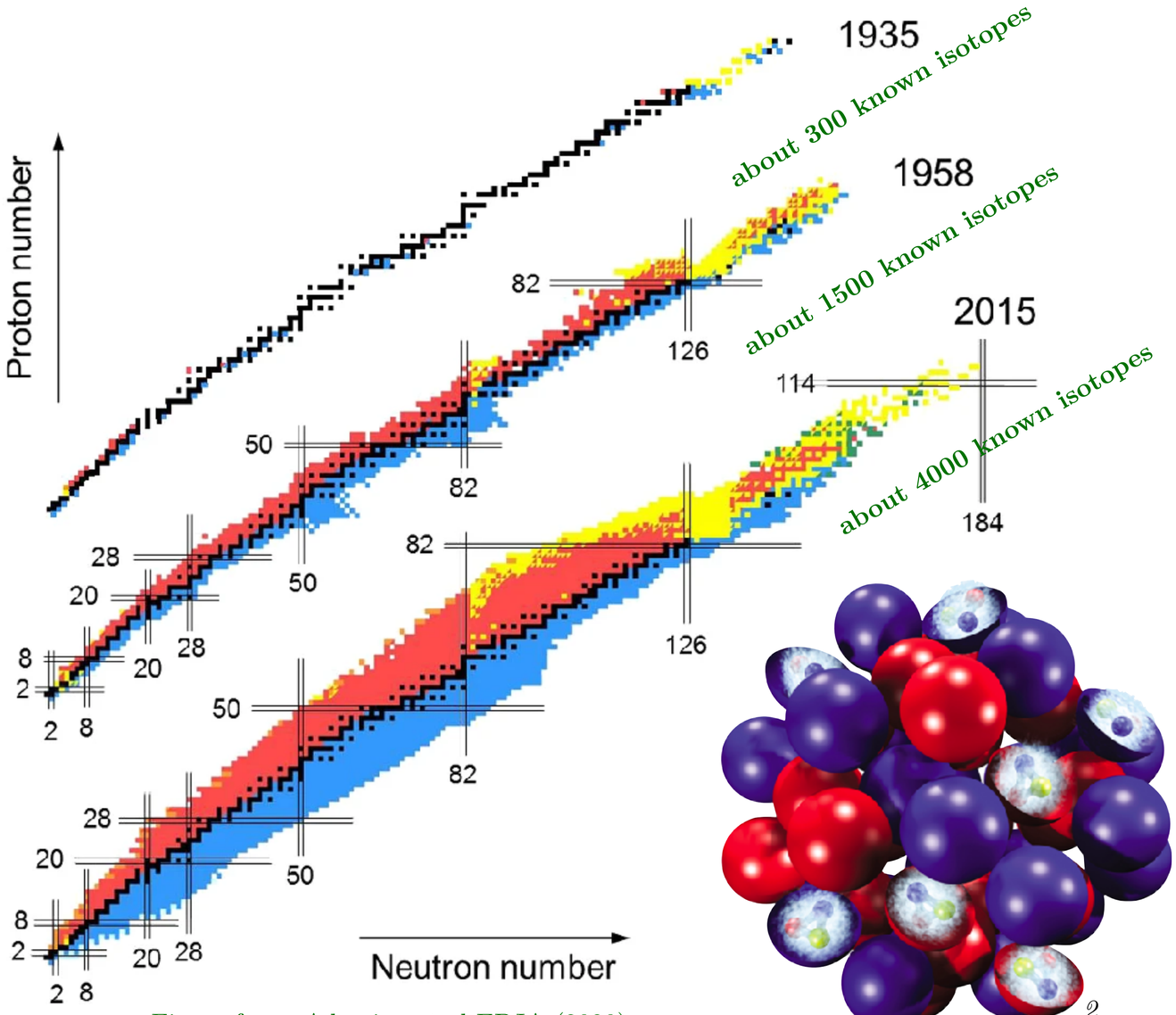
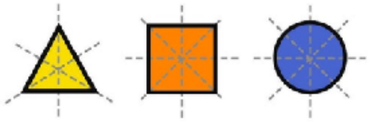
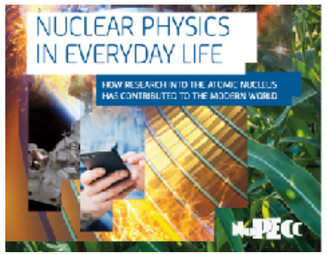


Figure from: Adamian et al EPJA (2020)

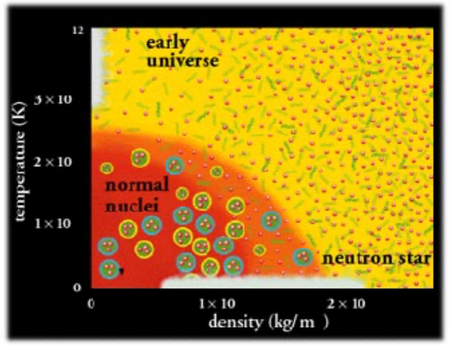
fundamental symmetries



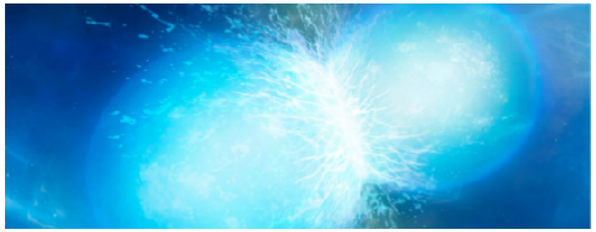


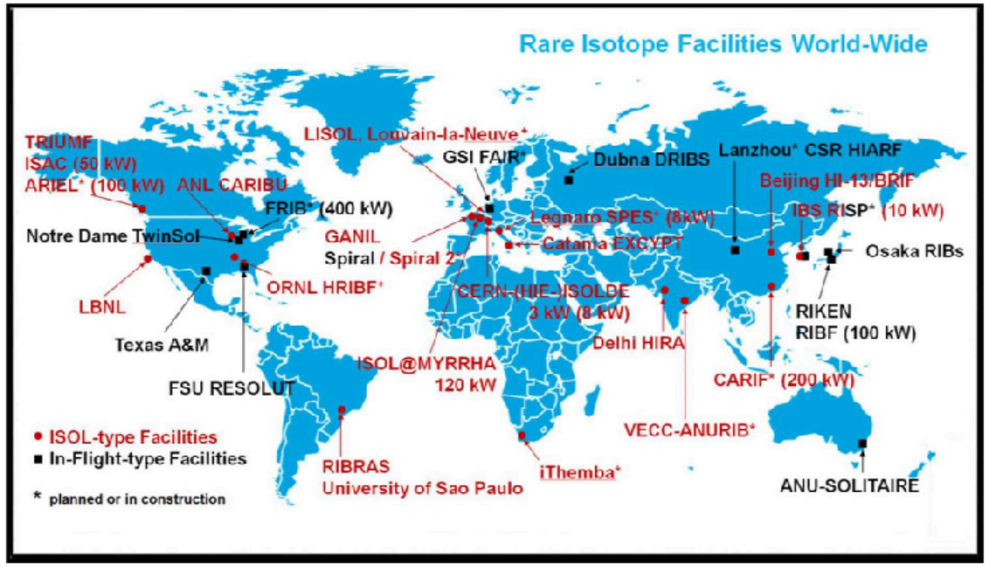
nuclear astrophysics

Nuclear equation of state

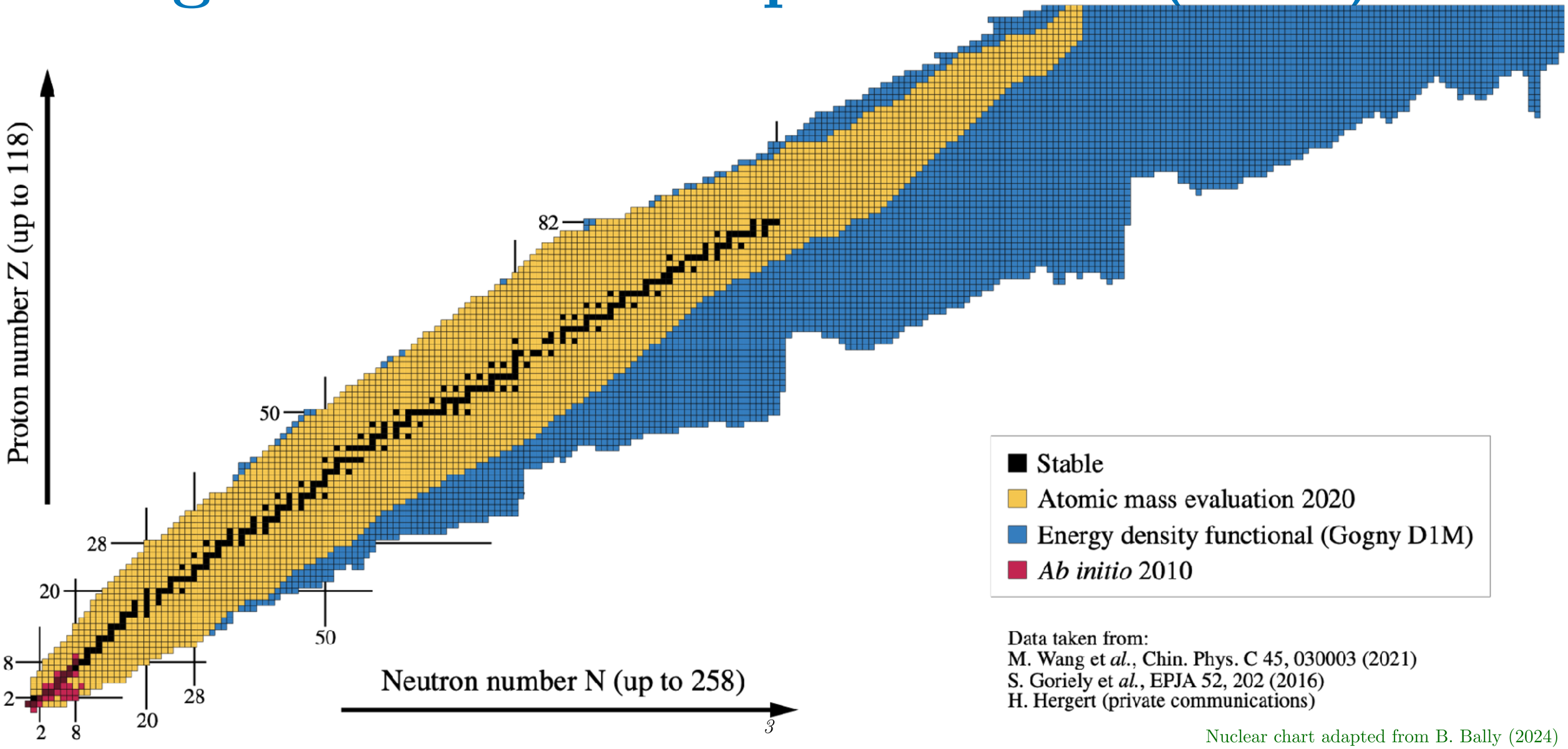


neutron star matter





Progress in *ab initio* predictions (2010)



Data taken from:
M. Wang et *al.*, Chin. Phys. C 45, 030003 (2021)
S. Goriely et *al.*, EPJA 52, 202 (2016)
H. Hergert (private communications)

Nuclear chart adapted from B. Bally (2024)

Progress in *ab initio* predictions (2024)

Breakthroughs in
- effective field theory
- many-body method
- statistical infere

Constraining neutron star matter
Huth, Pang, Tews, Le Fèvre, Schwenk, et al Nature (2022)

Topics to discuss in four lectures:

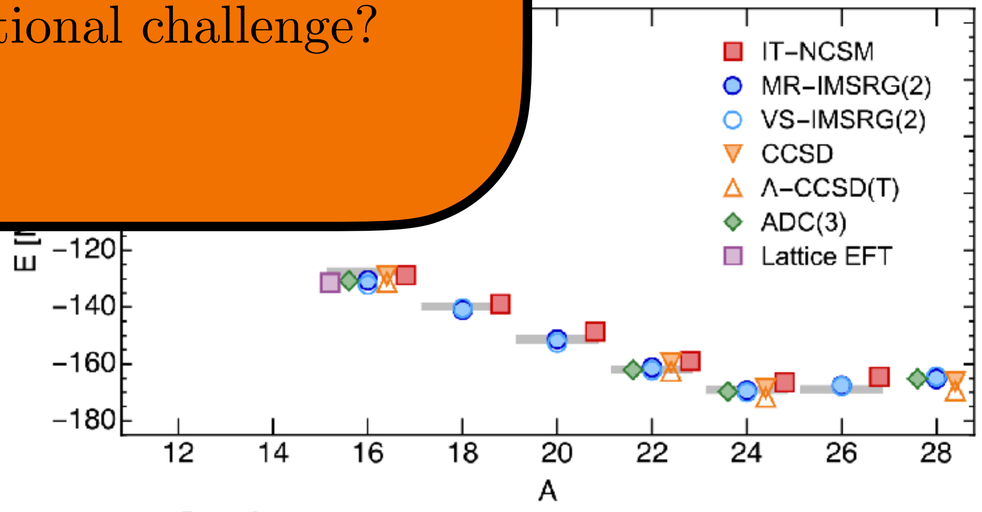
1) What is *ab initio* in nuclear physics?

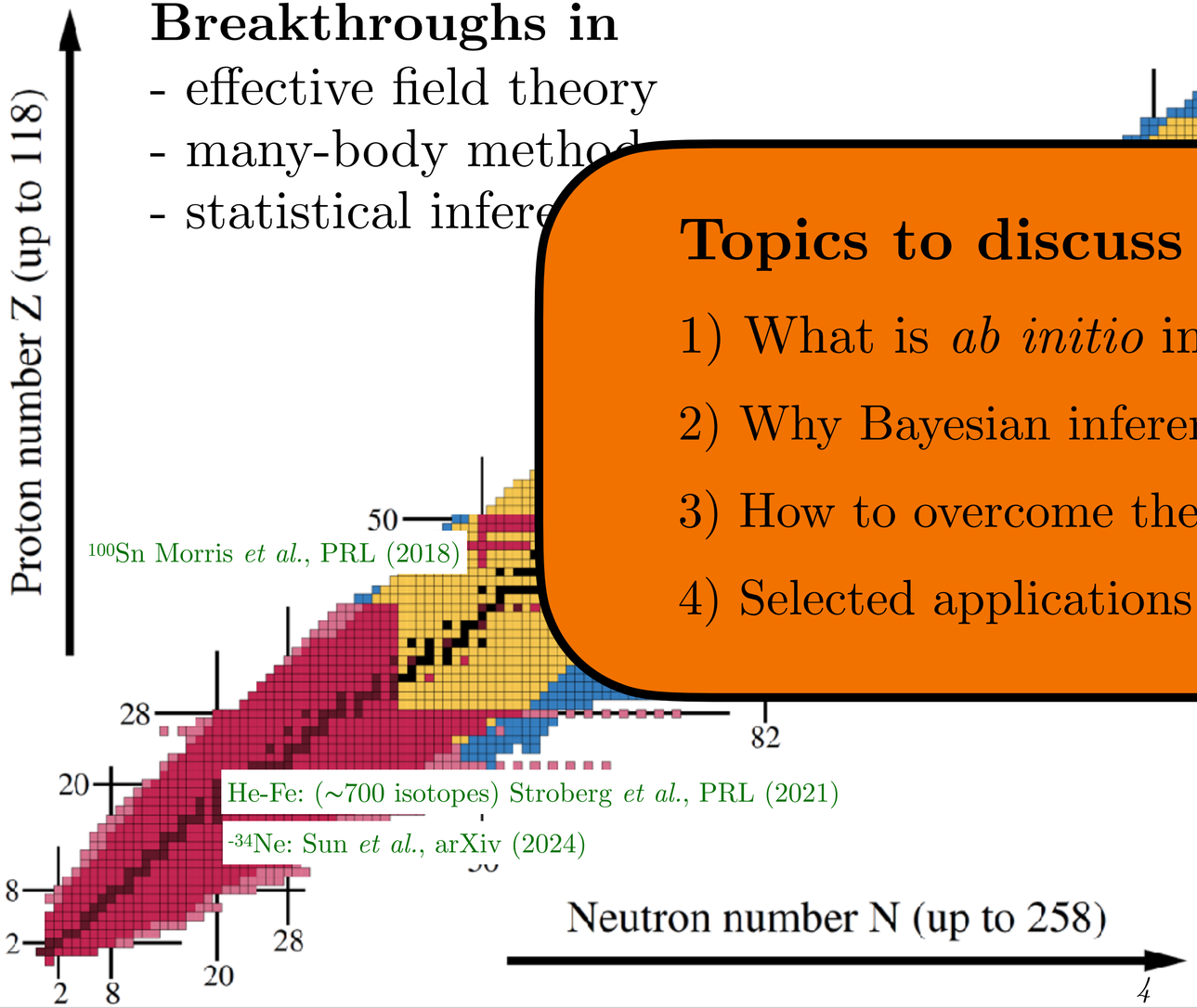
2) Why Bayesian inference?

3) How to overcome the computational challenge?

4) Selected applications

rtainty dominates





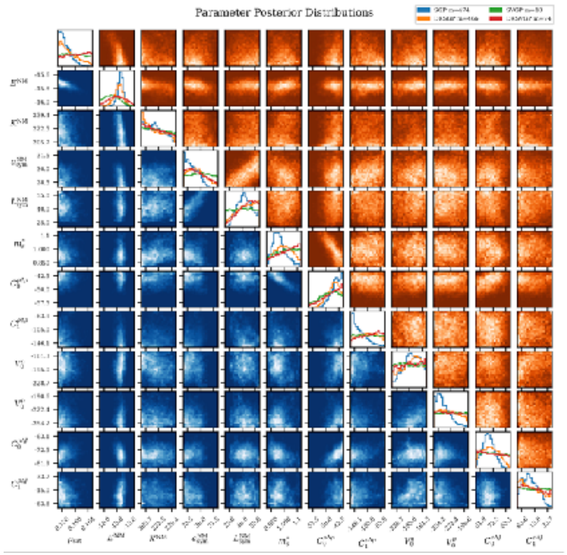
^{100}Sn Morris *et al.*, PRL (2018)

He-Fe: (∼700 isotopes) Stroberg *et al.*, PRL (2021)

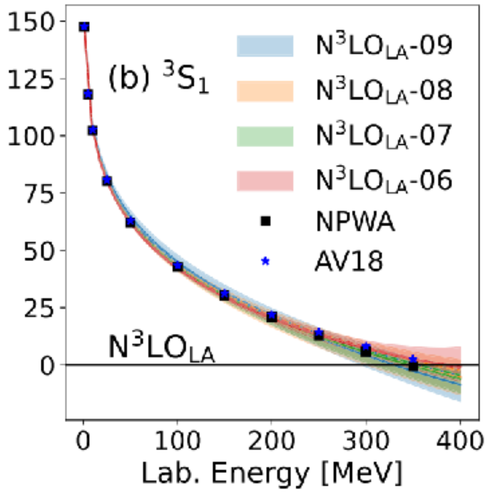
$^{-34}$Ne: Sun *et al.*, arXiv (2024)

Nuclear chart adapted from B. Bally (2024)

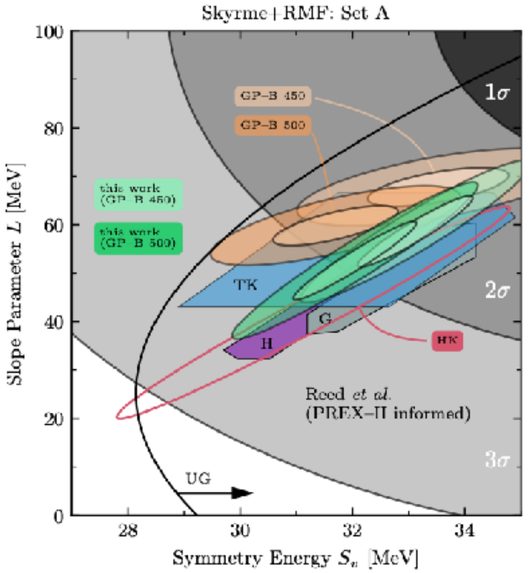
There's a lot of UQ in NP (2024 snapshots)



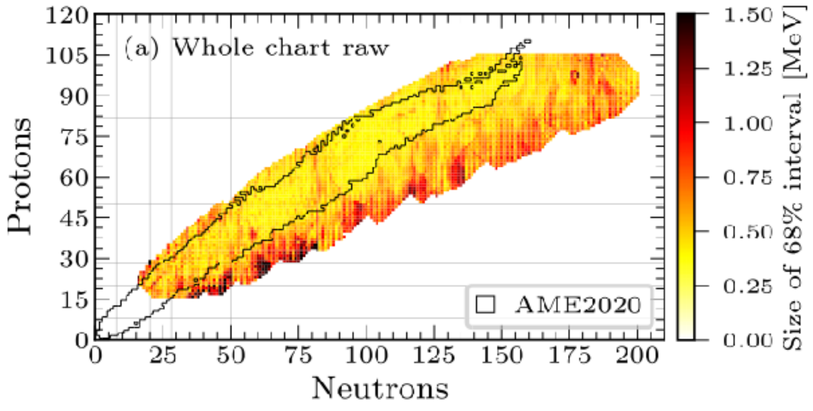
nuclear DFT: Stetzler arXiv (2024)
Posterior using GP emulators



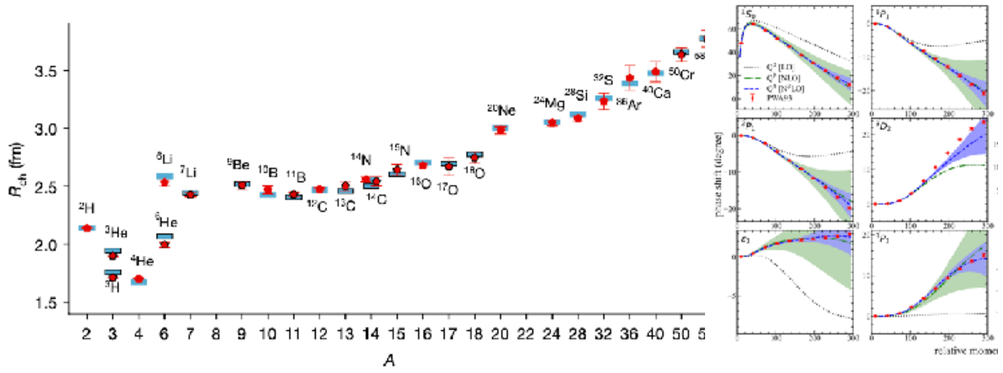
χEFT: Somasundaram PRC (2024)
Bayesian inference and MC



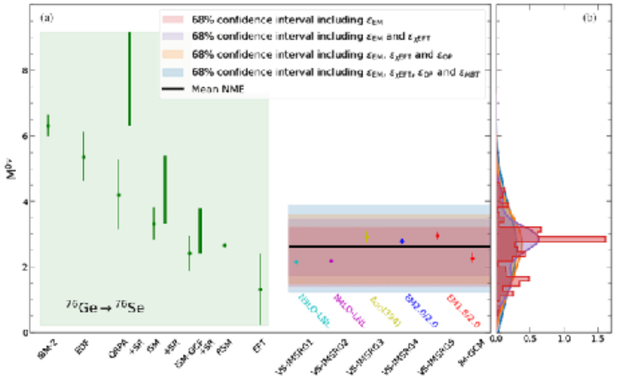
Nuclear Matter: Drischler arXiv (2024)
Bayesian mixture model



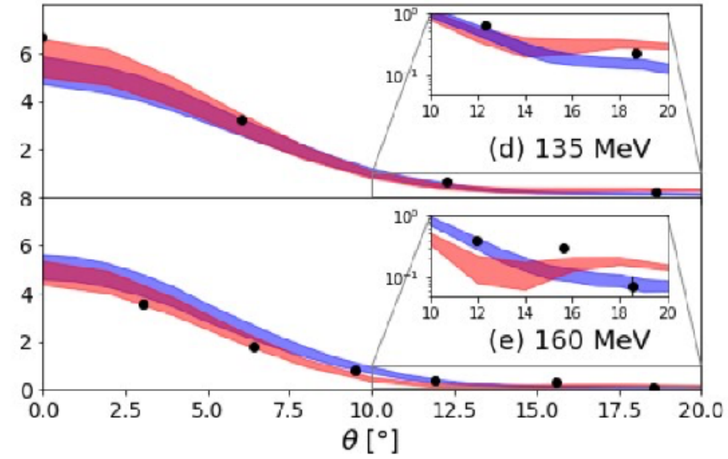
Nuclear mass models: Saito PRC (2024)
Bayesian model averaging



Lattice EFT predictions: Elhatisari Nature (2024)
History matching, MCMC, emulators



UQ fo $0\nu\beta\beta$ decay: Belley PRL (2024)
Posteriors, history matching, resampling, GP emulators



Nuclear reactions: Smith PRC (2024)
posteriors predictive distributions

ISNET

Information and Statistics in Nuclear Experiment and Theory

A flurry of Bayesian statistics and Machine Learning activity in the nuclear physics community.
Join ISNET and engage with fellow nuclear physicists and statisticians!

Topics covered:

- Uncertainty quantification and statistical analysis
- Emulators and optimization
- Model mixing and data mining
- Machine learning
- Bayesian inference
- Statistics techniques in nuclear experiments
- New frontiers of nuclear physics

https://napp.fudan.edu.cn/indico/event/757/

https://isnet-series.github.io/ for upcoming meetings, email list, resources

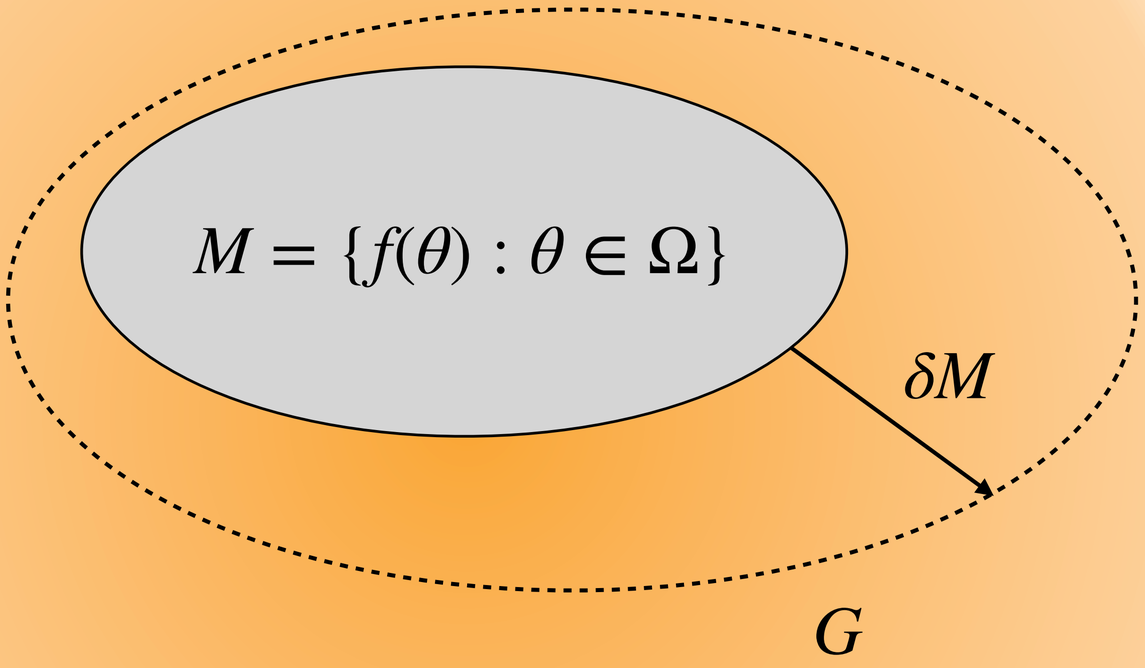
More about ISNET in next issue of NuPECC Nuclear Physics News

Lecture 1
What is *ab initio* in nuclear physics?

Remember: models are not reality

- we simplify and/or do not know the full story

We consider a **model** of some nuclear observable as a function $y = f(\theta),\ \theta \in \Omega \subset \mathbb{C}^d$, where Ω is some specified parameter space



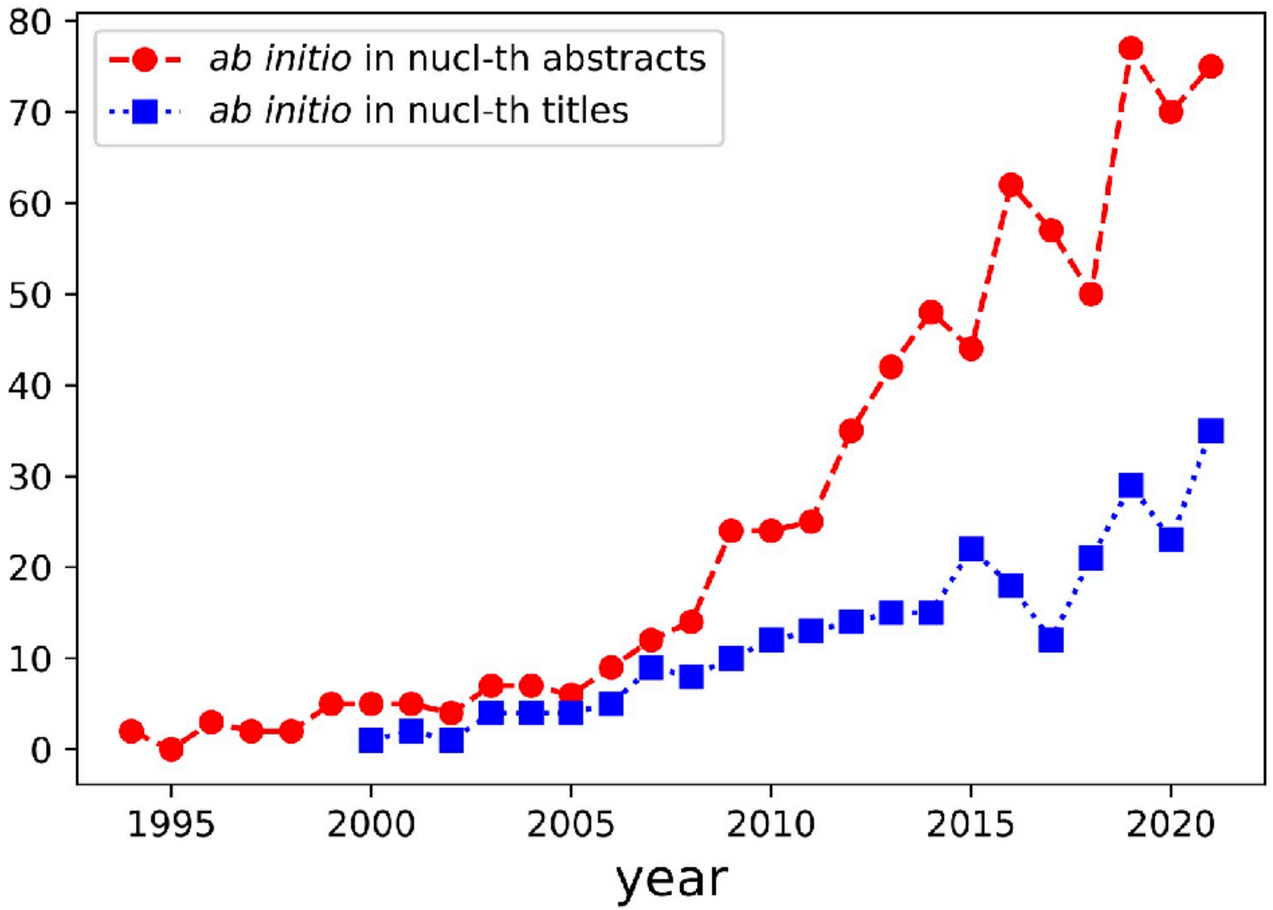
"All models are wrong"
i.e., data $D \sim G$ *where* $G \notin M$

G. E. P. Box (1976)

We should estimate the model discrepancy δM *and* improve our model(s) to (hopefully) make meaningful inferences and predictions.

Kennedy and O'Hagan (2001)
Brynjarsdóttir and O'Hagan (2014)

Ab initio: a label in nuclear theory since the mid 1990's



frontiers | Frontiers in Physics

TYPE Review
PUBLISHED 20 February 2023
DOI 10.3389/fphy.2023.1129094

OPEN ACCESS

EDITED BY
Paul Stevenson,
University of Surrey, United Kingdom

REVIEWED BY
Andreas Nogga,
Helmholtz Association of German Research Centres (HZ), Germany
Tobias Frederico,
Instituto de Tecnologia da Aeronáutica (ITA), Brazil
Richard Furnstahl,
The Ohio State University, United States

*CORRESPONDENCE
A. Ekström,
andreas.ekstrom@chalmers.se

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

RECEIVED 21 December 2022
ACCEPTED 07 February 2023
PUBLISHED 20 February 2023

CITATION
Ekström A, Forssén C, Hagen G, Jansen GR, Jiang W and Papenbrock T (2023), What is *ab initio* in nuclear theory? *Front. Phys.* 11:1129094.
doi: 10.3389/fphy.2023.1129094

What is *ab initio* in nuclear theory?

A. Ekström[1]*, C. Forssén[1], G. Hagen[2,3], G. R. Jansen[2,4], W. Jiang[1] and T. Papenbrock[2,3]

[1]Department of Physics, Chalmers University of Technology, Göteborg, Sweden, [2]Physics Division, Oak Ridge National Laboratory, Oak Ridge, TN, United States, [3]Department of Physics and Astronomy, University of Tennessee, Knoxville, TN, United States, [4]National Center for Computational Sciences, Oak Ridge National Laboratory, Oak Ridge, TN, United States

Ab initio has been used as a label in nuclear theory for over two decades. Its meaning has evolved and broadened over the years. We present our interpretation, briefly review its historical use, and discuss its present-day relation to theoretical uncertainty quantification.

KEYWORDS

***ab initio* nuclear theory, effective field theory, Bayesian inference, many-body methods, uncertainty quantication**

1 Introduction

The literal meaning of the latin term *ab initio* implies that one starts from the beginning. In computations of atomic nuclei, this means that the relevant degrees of freedom should be quarks and gluons. However, the history of physics tells us that we do not need to know everything to describe something and that we have some freedom in choosing the starting point. As such, we do not necessarily have to employ Standard Model degrees of freedom. In fact, many nuclear properties were successfully analyzed in terms of hadronic degrees of freedom before we even knew about the existence of quarks [1–3]. Today, we know how to explain this using renormalization group (RG) ideas [4, 5]. One may wonder about the exact meaning of the *ab initio* method and what should constitute the beginning. However, it is safe to say that a hallmark of this approach is its promise of precise and accurate predictions, with quantified uncertainties, across the multiple energy scales relevant to nuclei. Examples range from low-energy collective phenomena such as deformation and rotation [6–10], to loosely bound and unbound nuclei [11–16], and to lepton nucleus scattering in the quasi-elastic energy regime [17–19]. We expect the *ab initio* method to reliably extrapolate, in a controlled and systematic way, to regions outside the ones used for inferring the model parameters. Following the ideas from effective field theory (EFT) [20], *we interpret the ab initio method to be a systematically improvable approach for quantitatively describing nuclei using the finest resolution scale possible while maximizing its predictive capabilities.* A key part of this interpretation is the possible tension between the two latter aspects. In a nuclear physics context, we therefore let nucleons, and possibly other relevant hadronic degrees of freedom, define the beginning. Lattice quantum chromodynamics (QCD) might one day be the optimal starting point for predicting nuclear phenomena. Presently, Lattice QCD continues to provide useful input for EFTs based on hadronic degrees of freedom. However, it currently lacks predictive power for describing atomic nuclei [21–25].

We acknowledge that the *ab initio* method is interpreted differently by different people; see, e.g., Refs. [26–33]. In nuclear physics, the evolution of *ab initio* and its wide application

Ab initio®?

not a registered trademark. Different meaning to different people ...

The New York Times

Published Jan. 12, 2022 Updated Jan. 14, 2022

Is Gruyère Still Gruyère if It Doesn't Come From Gruyères?

A federal judge says yes, siding with U.S. cheese producers who say gruyère can be produced anywhere, not just in Switzerland and France.

|...|

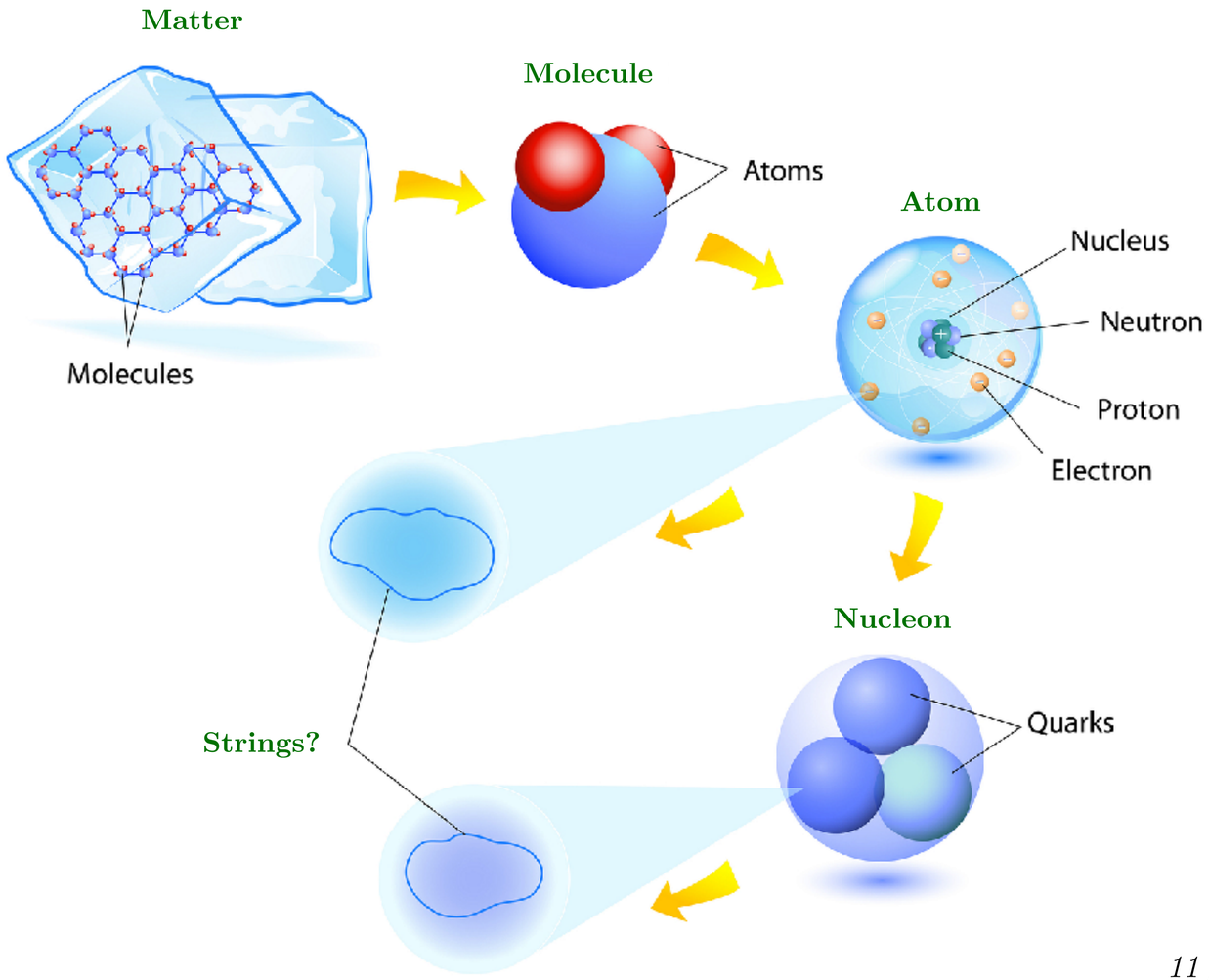
Perhaps most importantly, according to Swiss guidelines, gruyère must be made in the region around Gruyères, Switzerland, which has produced the cheese since the 12th century.

In the United States, however, gruyère can be made anywhere, according to a federal court ruling that was made public last week. It was the latest development in a long-running legal tangle between American cheese producers and producers in Switzerland and France over what makes gruyère gruyère. ?

ab initio, adv. & adj. From the beginning (=strings)?

/ˌab ɨˈnɪʃiəʊ/

ab-uh-NISH-ee-oh



The history of physics tells us that we do not need to know everything to describe something and that we have some freedom in choosing the starting point.

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

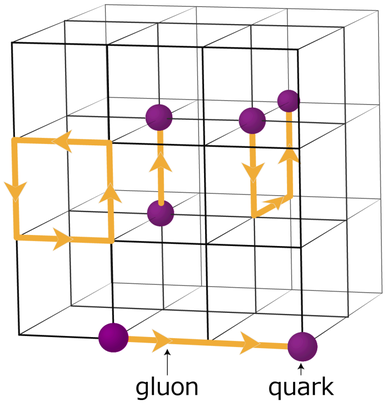
A key part of this interpretation is the possible *tension* between the two latter aspects.

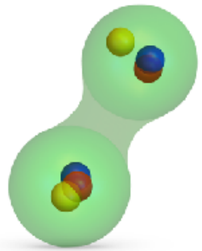
In a nuclear physics context, we therefore let hadrons (nucleons / pions) define the beginning.

[11] A. Ekström, C.Forssen, G.Hagen, G. R. Jansen, W. Jiang, and T. Papenbrock, Frontiers (2022)

Lattice quantum chromodynamics (QCD)

-non-perturbative QCD predictions for few-nucleon systems





- energy-spectra and complex nuclear dynamics
- electroweak and beyond-Standard-Model currents
- LQCD data for constraining models of nuclei

😟 Sign-problem yields exponential degradation of signal-to-noise ratio at large Euclidean time. This becomes worse as the pion mass is reduced towards the physical point and the number of nucleons increase.

Necessary to obtain control over the lattice volume and physical pion mass extrapolations to make meaningful comparison between different LQCD results as well as with experiments.

Currently: ~ 20 MeV systematic uncertainty in the NN binding energy.

Z. Davoudi et al Phys. Rep. (2021)

Ab initio offers an inferential advantage

$$y_{\text{exp}}(\vec{x}) = y_{\text{th}}(\vec{\alpha};\vec{x}) + \delta y_{\text{th}}(\vec{\alpha};\vec{x}) + \delta y_{\text{exp}}(\vec{x})$$

model parameters | **model discrepancy** | **control parameters**

Nuclear *ab initio:* a *systematically improvable* approach for quantitatively describing nuclei using the finest resolution scale possible while maximizing its predictive capabilities.

$$H|\Psi_i\rangle = E_i|\Psi_i\rangle$$

***χEFT* to approximate low-energy QCD**

$$H(\vec{\alpha}) = T + V^{(0)}(\vec{\alpha}_{(0)}) + V^{(1)}(\vec{\alpha}_{(1)}) + V^{(2)}(\vec{\alpha}_{(2)}) + \ldots$$

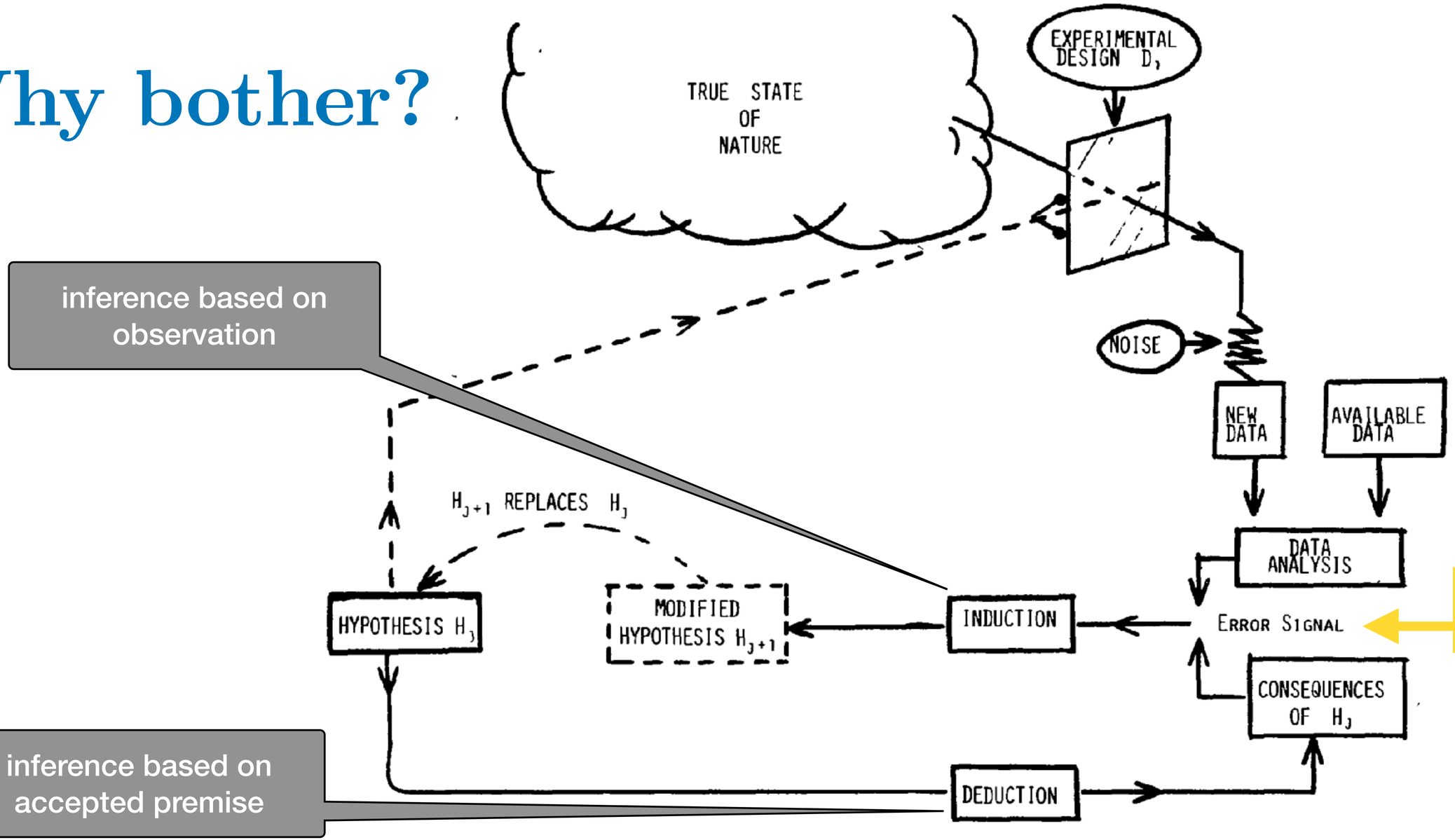
***A*-body methods with controllable approximation**

$$|\Psi\rangle = |\Phi^{(0)}\rangle + |\Phi^{(1)}\rangle + |\Phi^{(2)}\rangle + \ldots$$

This systematicity creates an *inferential advantage*. We can test our assumptions about the model (M) and its discrepancy as we increase the fidelity of M.

A. Ekström, C.Forssen, G.Hagen, G. R. Jansen, W. Jiang, and T. Papenbrock, Frontiers (2022)

Why bother?



[a] The experimental design is here shown as a movable window looking onto the true state of nature. Its positioning at each stage is motivated by current beliefs, hopes, and fears.

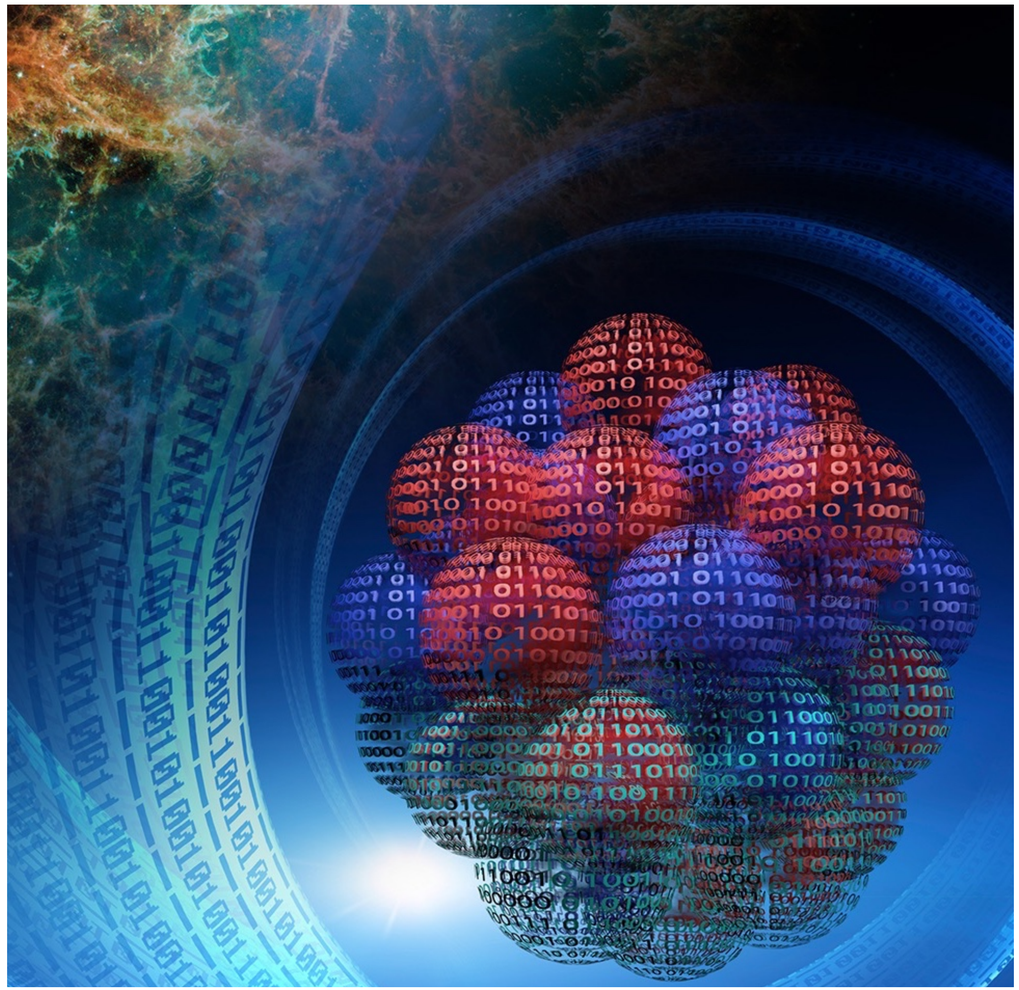
Science and Statistics

George E. P. Box

Journal of the American Statistical Association, Vol. 71, No. 356. (Dec., 1976), pp. 791-799.

Several interesting questions

- **Some** interaction parametrizations work better than others, but only for selected types of observables and in limited regions of the nuclear chart, e.g. NNLOsat, 1.8/2.0 (EM), ΔNNLOgo. **How to improve?**
- **What** is the predictive power of *ab initio* nuclear theory?
- **Can** we reach the same level of accuracy as experiment?
- **To what extent** can nuclei be described in effective field theories (EFTs) of quantum chromodynamics?
- **How to best use computational statistics**: (experimental design, sensitivity analysis, MCMC, Bayesian inference, model mixing, history matching, ABC, …)

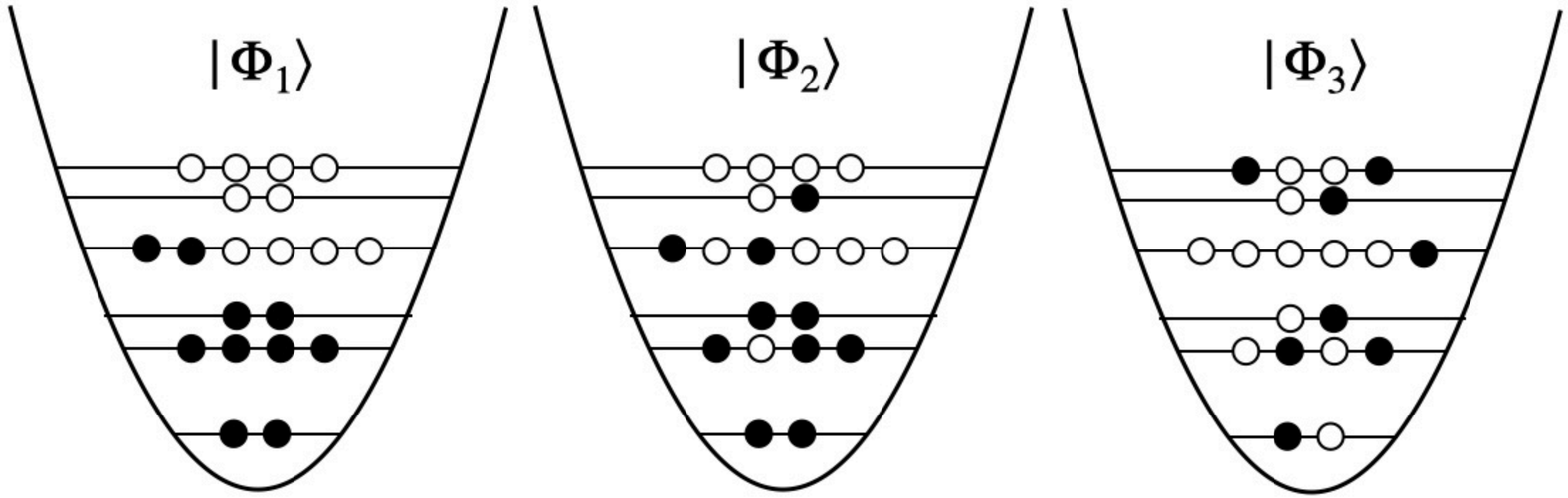


Ab initio many-body methods

- some examples of systematicity: the no-core shell model

$$H|\Psi\rangle = E|\Psi\rangle$$

We operate in a basis of many-nucleon states, e.g., slater determinants $|\Phi_1\rangle, |\Phi_2\rangle, |\Phi_3\rangle, \ldots$
Controlled approximation of maximum energy of excitations.



Barrett, Navratil, Vary (2013)

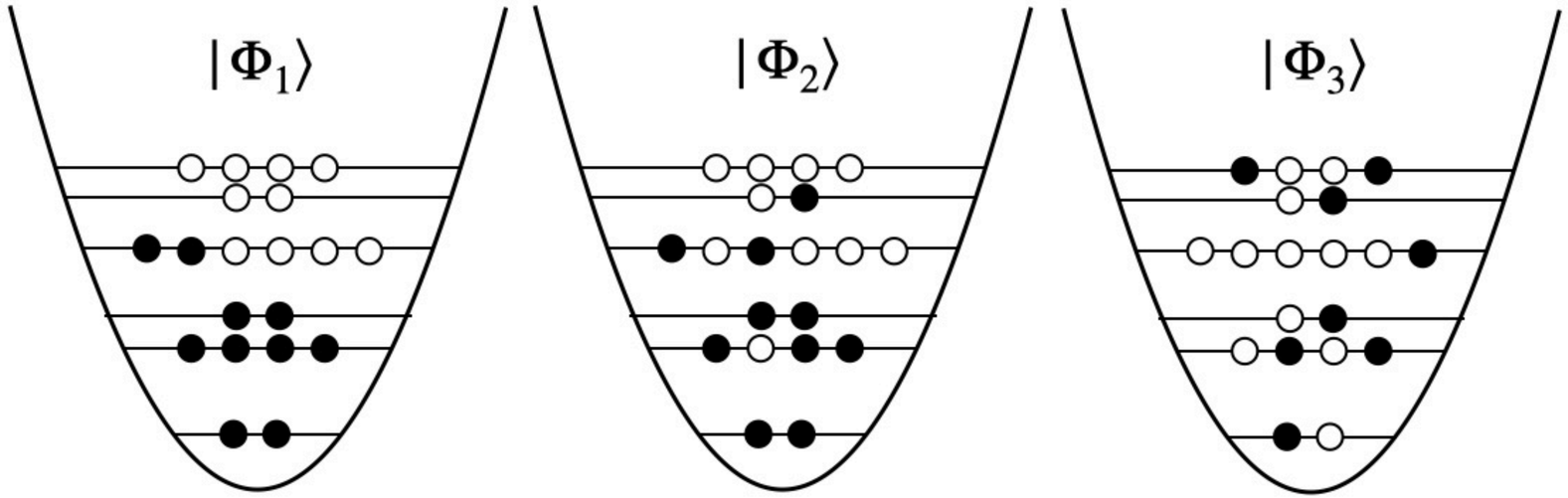
Ab initio many-body methods

- some examples of systematicity: the no-core shell model

Insert a complete set of states $1 = \sum_j |\Phi_j\rangle\langle|\Phi_j|$ and left-multiply with $\langle\Phi_i|$

$$\sum_j \langle\Phi_i|H|\Phi_j\rangle\langle\Phi_j|\Psi\rangle = E\langle\Phi_i|\Psi\rangle$$

We represent this as a finite matrix diagonalization problem and obtain a variational bound on E. A larger basis will approximate the A-nucleon state better. **Exponential cost...**

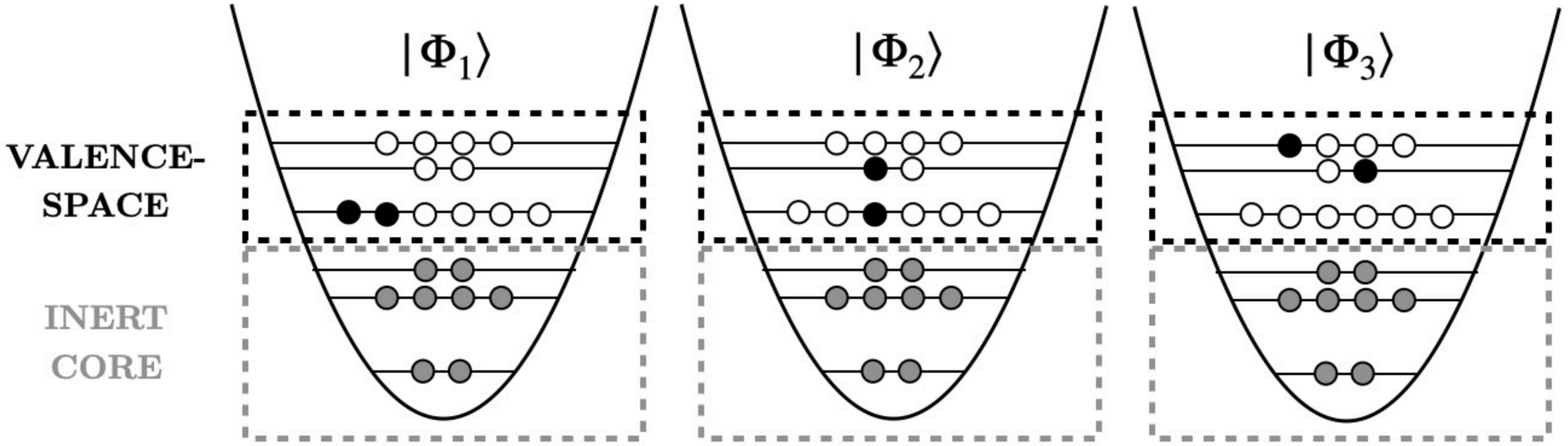


Barrett, Navratil, Vary (2013)

Ab initio many-body methods

- some examples of systematicity: the canonical shell model

caveat: there are methods to systematically project χEFT interactions to a limited valence-space. Some effective interactions are however of more phenomenological character, e.g., USD-B/GXPF1A/..., with matrix elements calibrated reproduce selected nuclear structure in a small portion of the nuclear chart. *Unclear* how to systematically improve such approaches.



Goeppert Mayer, Haxel, Jensen, Suess, Wigner, Talmi, de Shalit, Brown, Brown, Wildenthal, Kuo, and many others

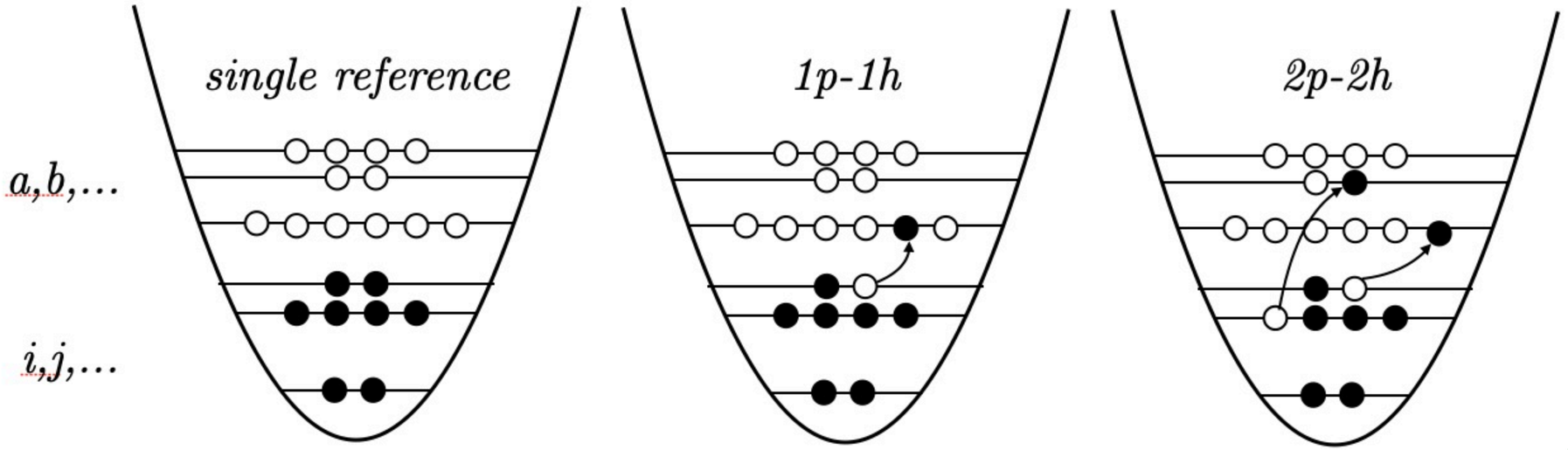
Ab initio many-body methods

- some examples of systematicity: the coupled-cluster method

Few-particle / few-hole excitations systematically generated on top of a reference state as $|\Psi\rangle = e^T |\Phi\rangle$, where the cluster operator $T = T_1 + T_2 + \ldots + T_A$ with

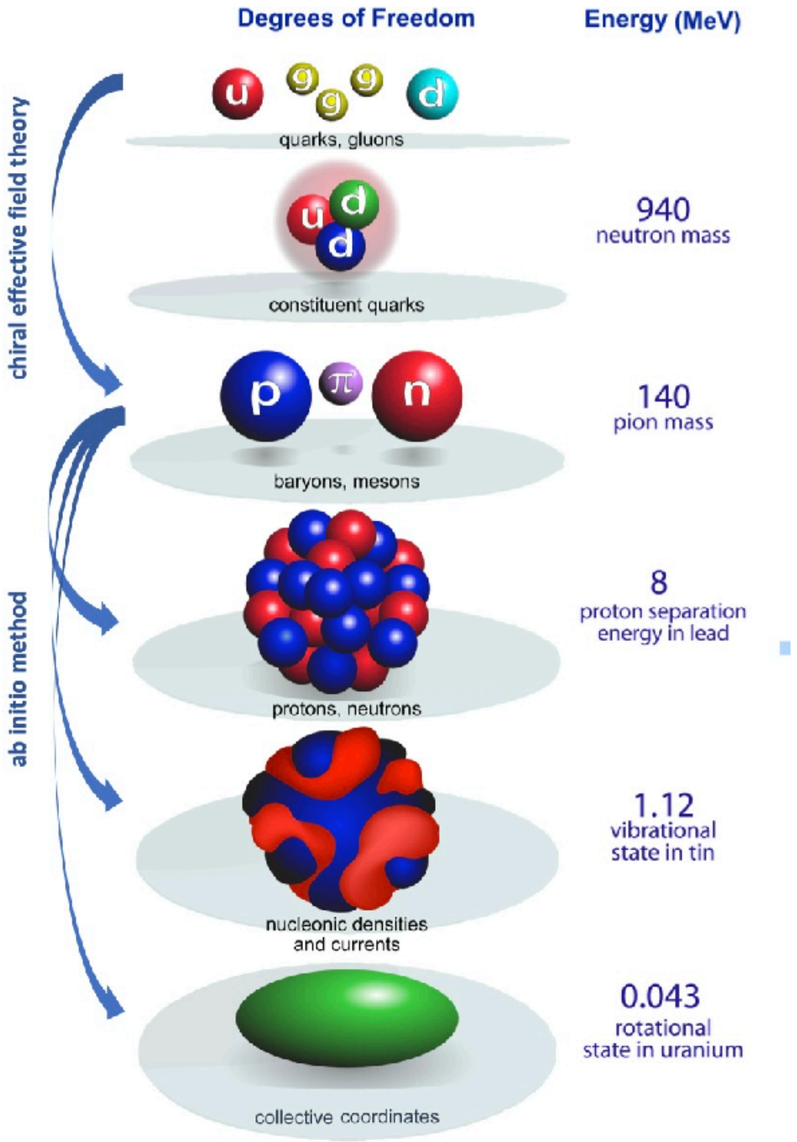
$$T_1 = \sum_{ia} t_i^a a_a^\dagger a_i \text{ and } T_2 = \sum_{ijab} t_{ij}^{ab} a_a^\dagger a_b^\dagger a_j a_i$$

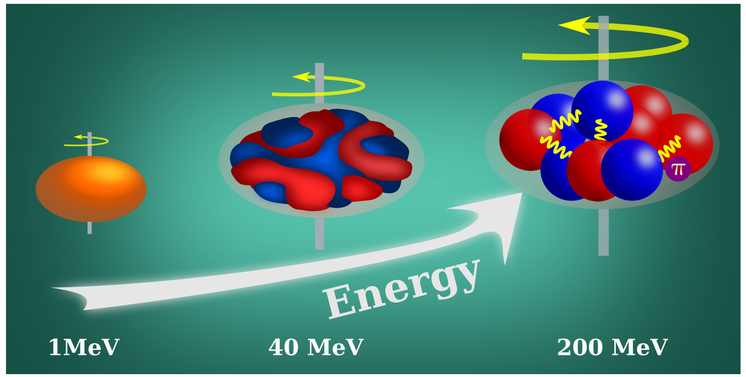
considers all nucleons, yet scales polynomially ($A^3 u^5$). Performs very well in the vicinity of closed (sub) shells. Recent developments: axially deformed states and collective transitions

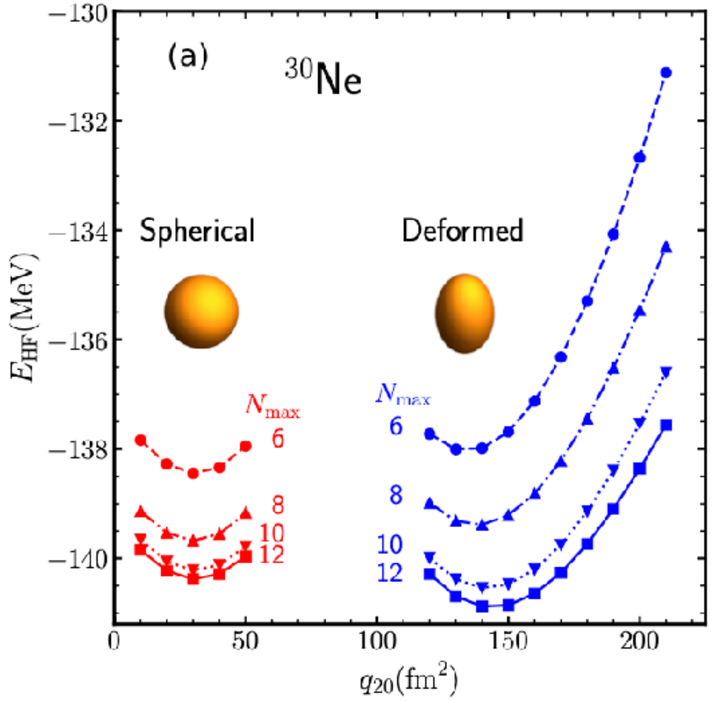


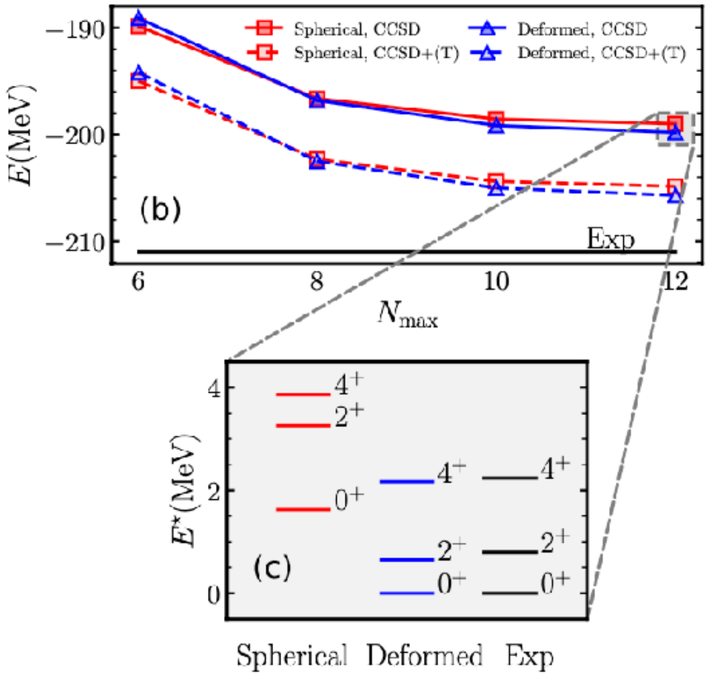
Coester, Kümmel (1950s)
Dean and Hjorth-Jensen (2003)
Hagen et al (2013)

Progress in *ab initio* methods

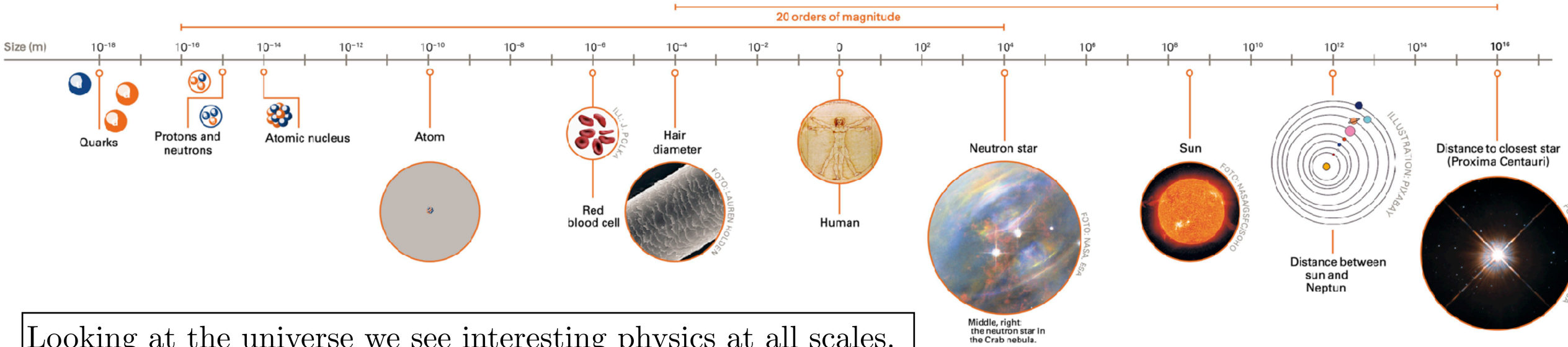




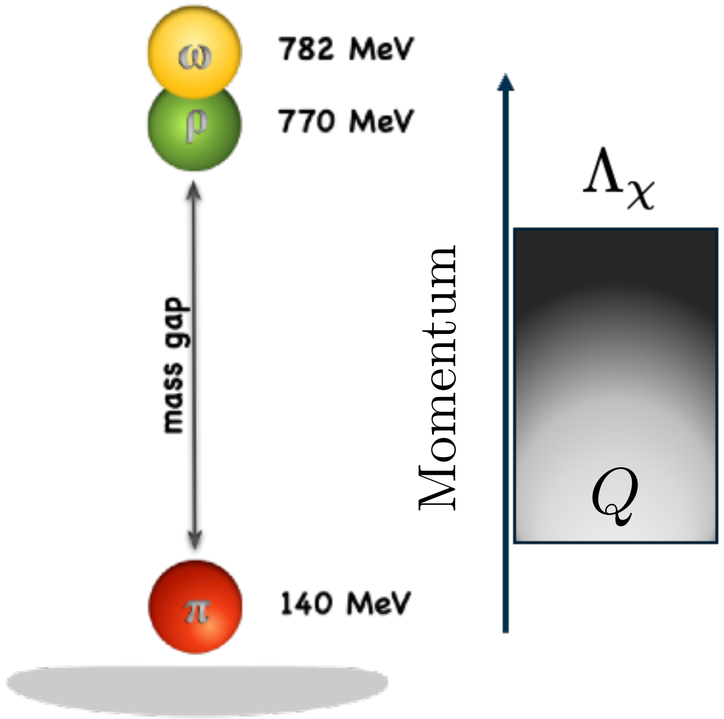




Z. H. Sun et al (2024)



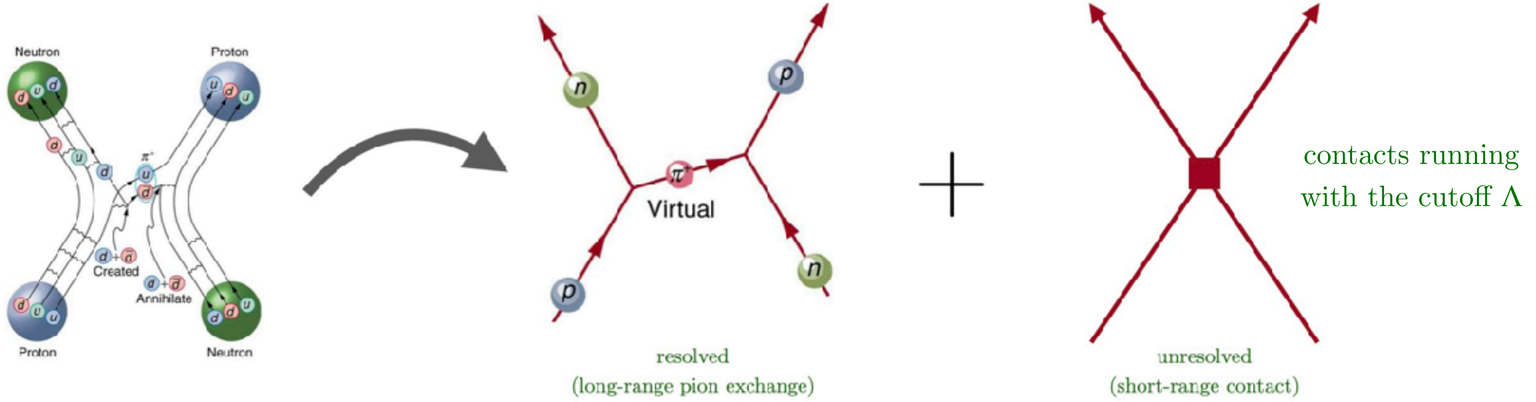
Looking at the universe we see interesting physics at all scales.
When exploring new energy scales we find new phenomena.



EFT enables us to isolate some phenomena from all the rest such that we can describe something without having to resolve everything.

- identify a separation of energy scales; low-energy ('physics we are interested in') and high-energy ('underlying physics')
- write effective Lagrangian of pions & nucleons as allowed by the symmetries of the full theory (QCD).
- chiral symmetry requires pion interactions to be proportional to Q, i.e., weak at low energies
- power counting facilitates an expansion in a ratio Q/Λ_χ
- short-distance physics ("quarks and gluons") beyond hard scale represented as contact interactions.

The nuclear interaction from χEFT



$$[H_0 + V(\vec{\alpha})]|\Psi\rangle = E|\Psi\rangle$$

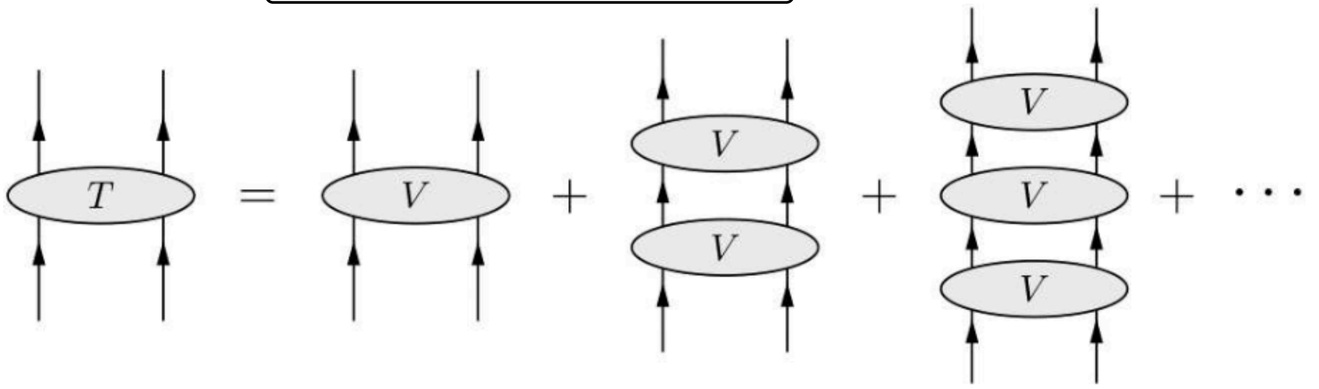
We typically solve the Schrödinger equation. But let's view the procedure from the (formally) equivalent Lippmann-Schwinger equation

$$T = V + V[E - H_0]^{-1}T$$

$$T|\phi\rangle \equiv V|\Psi\rangle$$

Introducing the (unknown) T-matrix shifts the challenge from finding the wavefunction to finding the scattering matrix *S=1-T*

potential from χEFT

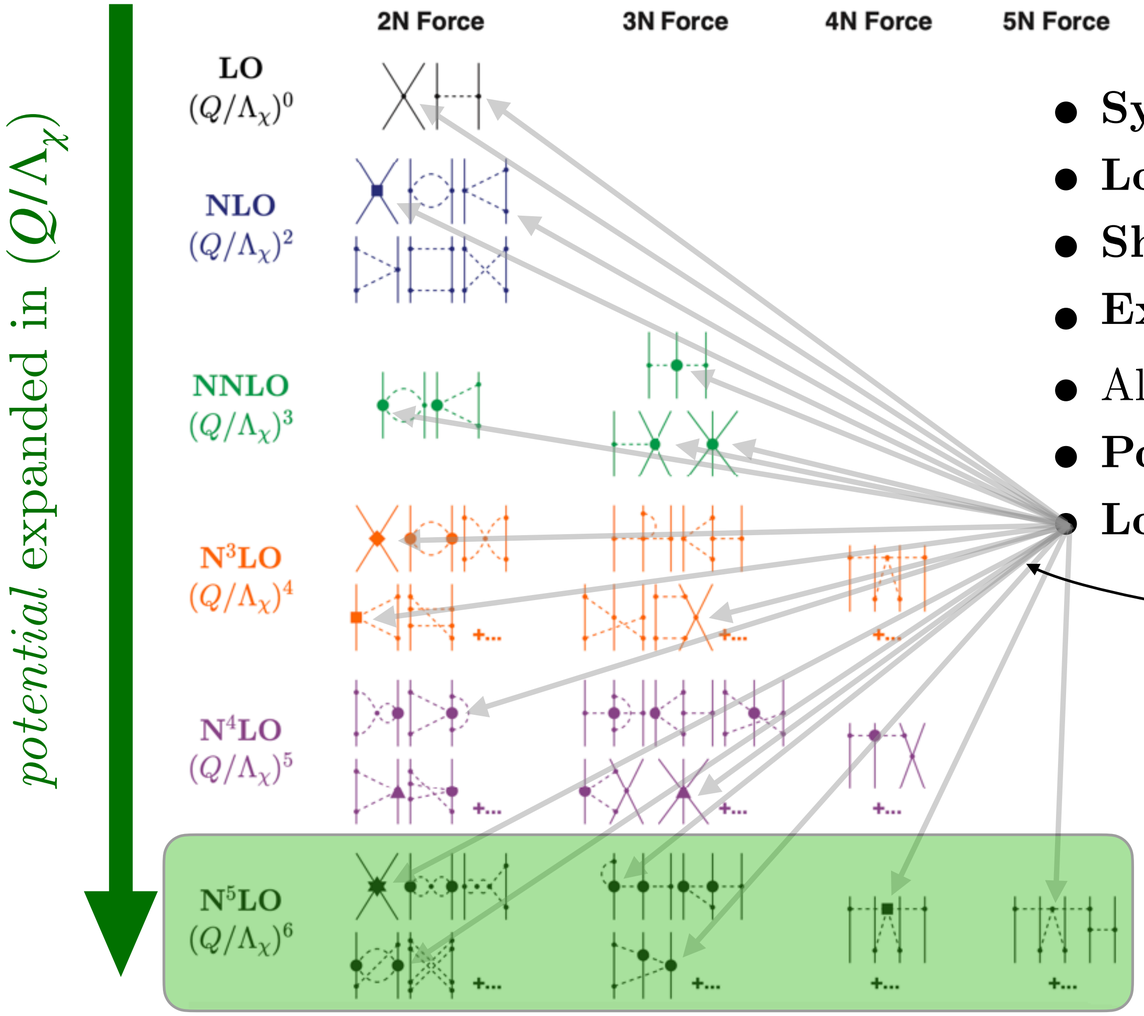


"V iterated to infinite order"

UV-divergences treated using cutoff regularization $V(p',p) \to f(p';\Lambda)V(p',p)f(p;\Lambda)$.
Typically, many-body methods converge numerically for $\Lambda \lesssim 500$ MeV

Chiral effective field theory (χEFT)

- Weinberg power counting (WPC)



- **Symmetries** of QCD dictate contents of effective Lagrangian
- **Long-ranged** physics governed by pion exchanges
- **Short-ranged** physics determined by a set of contact interactions
- **Expansion** in (Q/Λ_χ) [soft scale $(\sim m_\pi)$ over hard scale $(\sim m_N)$]
- All operators must be regulated ⇒ **cutoff dependence**
- **Power counting** organizes contributions (diagrams) at each order
- **Low-energy** constants (LECs) must be fit from data once

That's many parameters..
where did my predictive power go?

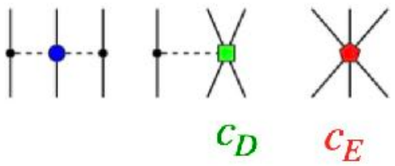
$$y_{\rm th}^{(k)} = y_{\rm ref} \sum_{\nu=0}^{k} c_\nu \left(\frac{Q}{\Lambda_\chi}\right)^\nu \quad \text{and} \quad \delta y_{\rm th}^{(k)} = y_{\rm ref} \sum_{\nu=k+1}^{\infty} c_\nu \left(\frac{Q}{\Lambda_\chi}\right)^\nu$$

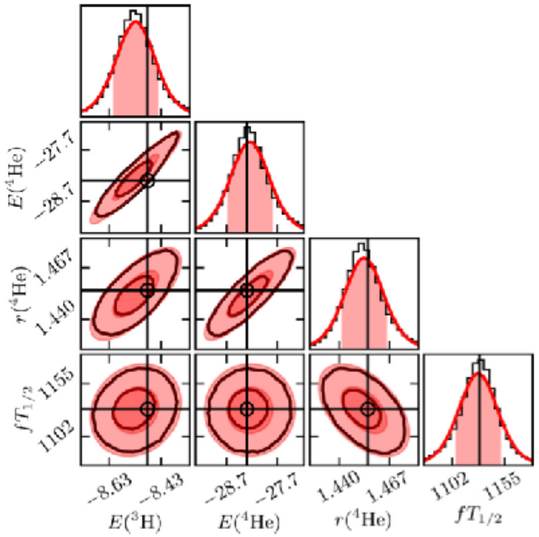
EFT prediction | **EFT truncation error**

Weinberg, van Kolck, Kaplan, Savage, Wise, Weise, Kaiser, Bernard, Meißner, Epelbaum, Machleidt, Entem, ...

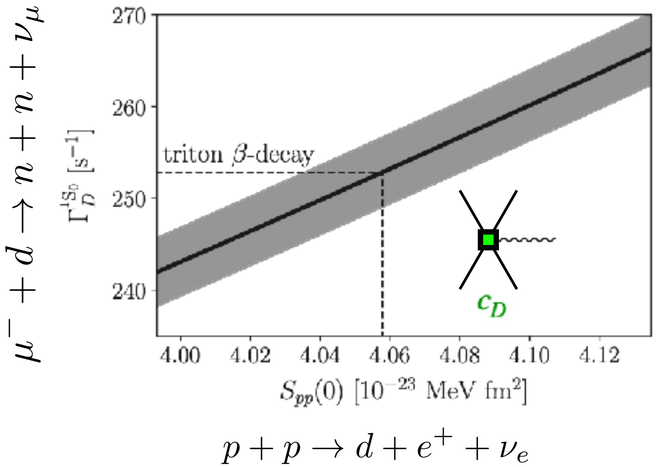
Where did my predictive power go?

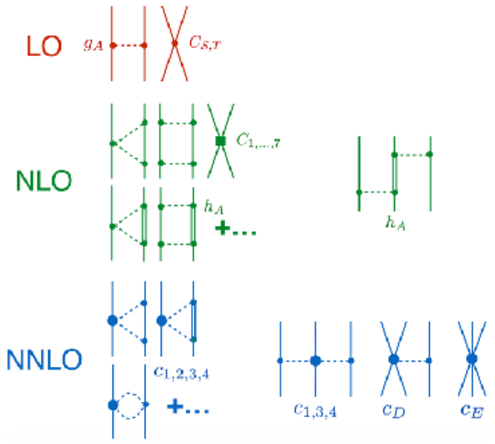
symmetries for the win!

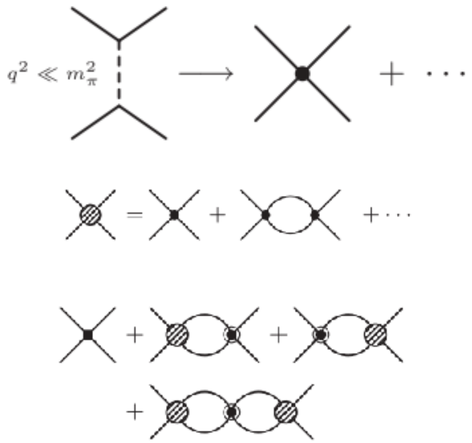


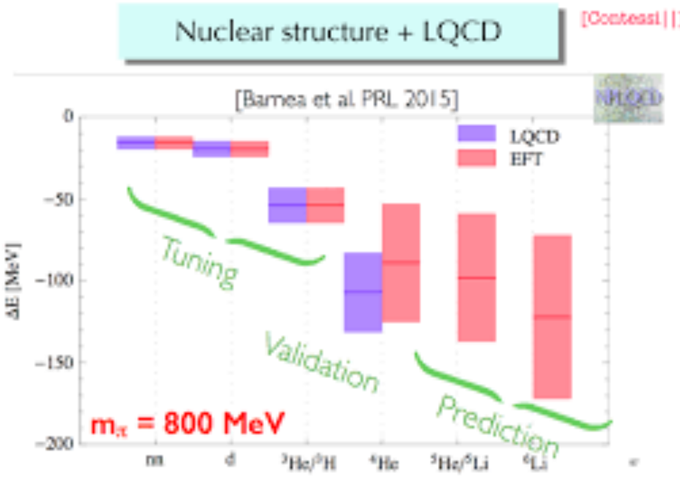


- Two- and three-nucleon forces, as well as electroweak currents, emerge in the same theoretical framework and share fundamental pion-nucleon couplings and symmetries with QCD.
- The pion-nucleon sub-process in the nuclear interaction is accessible via studies of pion-nucleon scattering
- Effects of nucleon excitations, e.g., the Δ(1232) resonance, can be accounted for via an additional small-scale expansion
- Integrating out the pion yields a lower-energy (contact) EFT that can be matched to χEFT
- EFT can extrapolate LQCD results







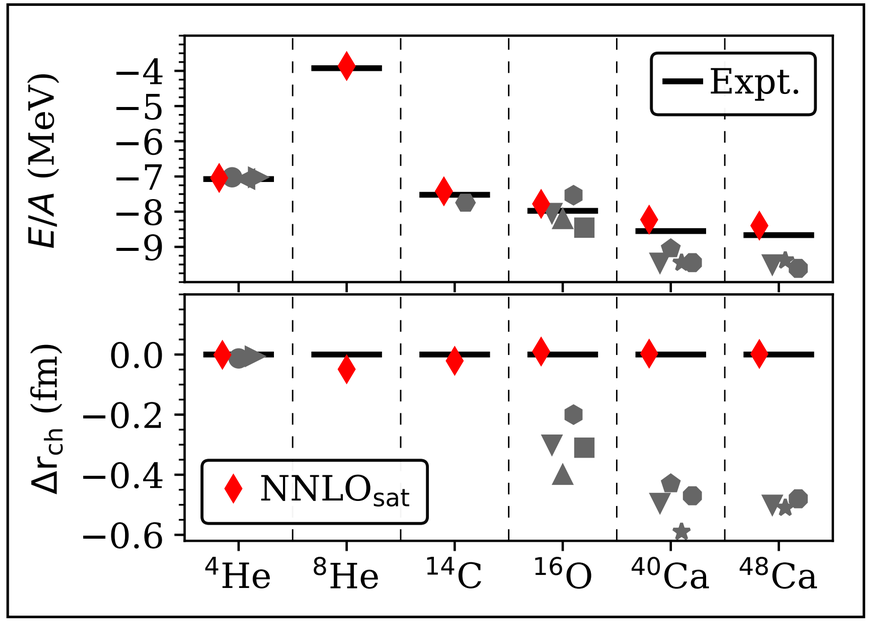


Some interactions work better than others

- we don't know why...

  Different regularization schemes, power counting, and calibration choices:

  - Idaho-N3LO(500)
  - NNLOsat(450)
  - NNLOopt(500)
  - Δ-NNLO$_{\mathrm{GO}}$(394/450)
  - 1.8/2.0 (EM)
  - EMN LO-N4LO
  - EKM LO-N4LO
  - Local Norfolk interactions
  - Long-Yang (450,500,550)
  - Non-implausible N2LO sets
  - ... and many more



Ekström et al Phys. Rev. C (2015)

Renormalization group invariance and χEFT

- *Modified* Weinberg power counting (MWPC)

amplitude expanded in (Q/Λ_χ)

non-perturbative one-pion-exchange

$^3S_0, {}^3S_1 - {}^3D_1, {}^3P_{0,1}, {}^3P_2 - {}^3F_2, {}^1P_1$

perturbative contriutions

LO g_A $C_{^1S_0}, C_{^3S_1}, D_{^3P_0}, D_{^3P_2}$

LO 3NF?
Yang et al (2023)

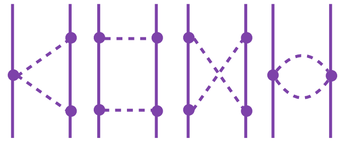
NLO $D_{^1S_0}$

N²LO $E_{^1S_0}, D_{^3S_1}, D_{^3S_1-{}^3D_1}, D_{^1P_1}, D_{^3P_1}, E_{^3P_0}, E_{^3P_2}, E_{^3P_2-{}^3F_2}$

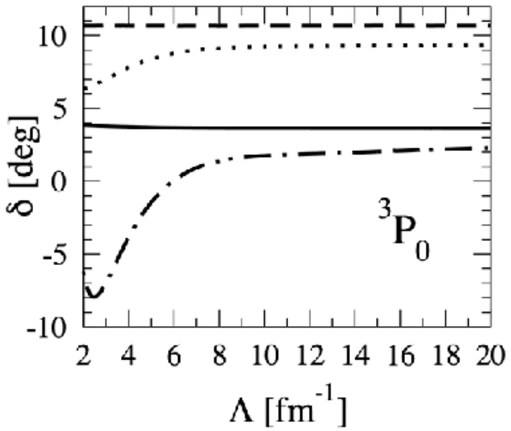
N³LO $c_{1,3,4}$ $F_{^1S_0}$

All other partial waves

g_A



Promoting counterterms to absorb cutoff dependence due to singular one-pion exchange

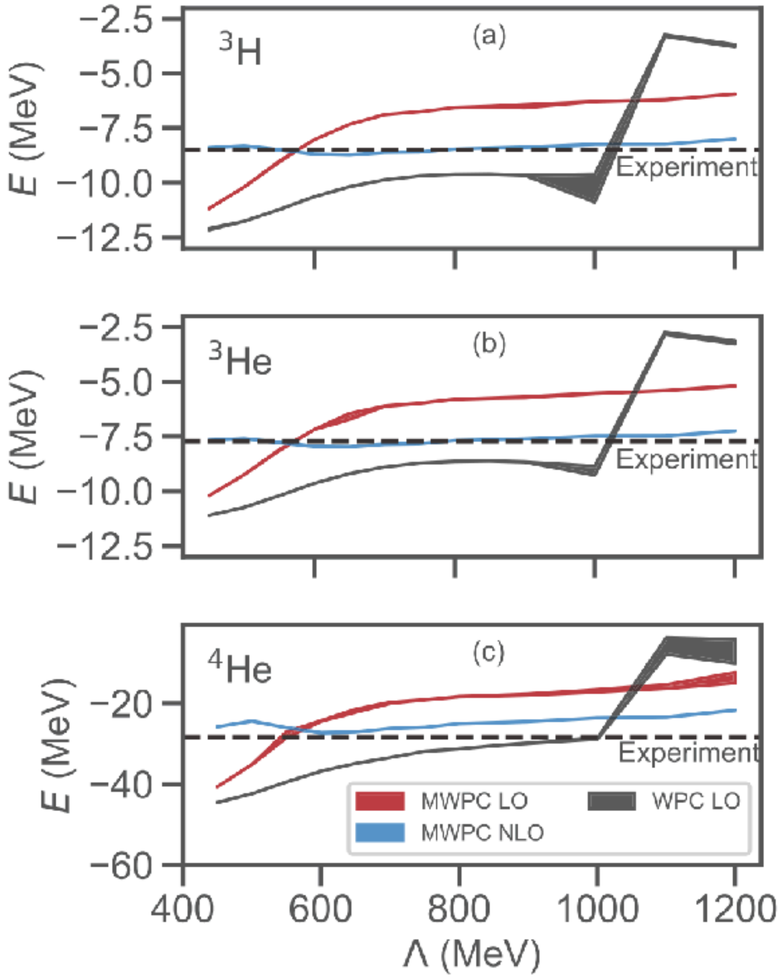


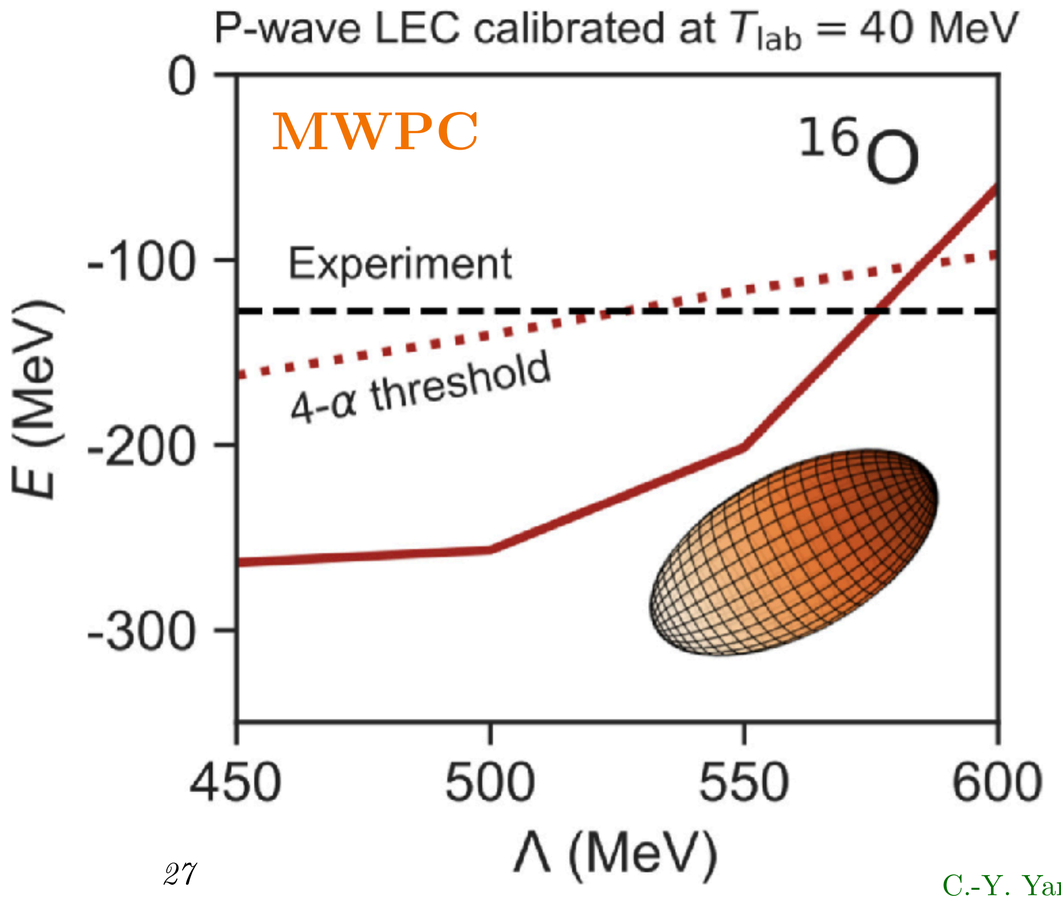
Nogga, Timmermans, van Kolck (2005)

Several other PCs exist: Kolck, Kaplan, Savage, Wise, Long, Valderrama, Griesshammer, Yang, Birse, Arriola, Phillips, ...

Deficiencies at leading order in χEFT

Atomic nuclei with A>4 unphysical





C.-Y. Yang, et al. Phys. Rev. C (2021)

Next lecture

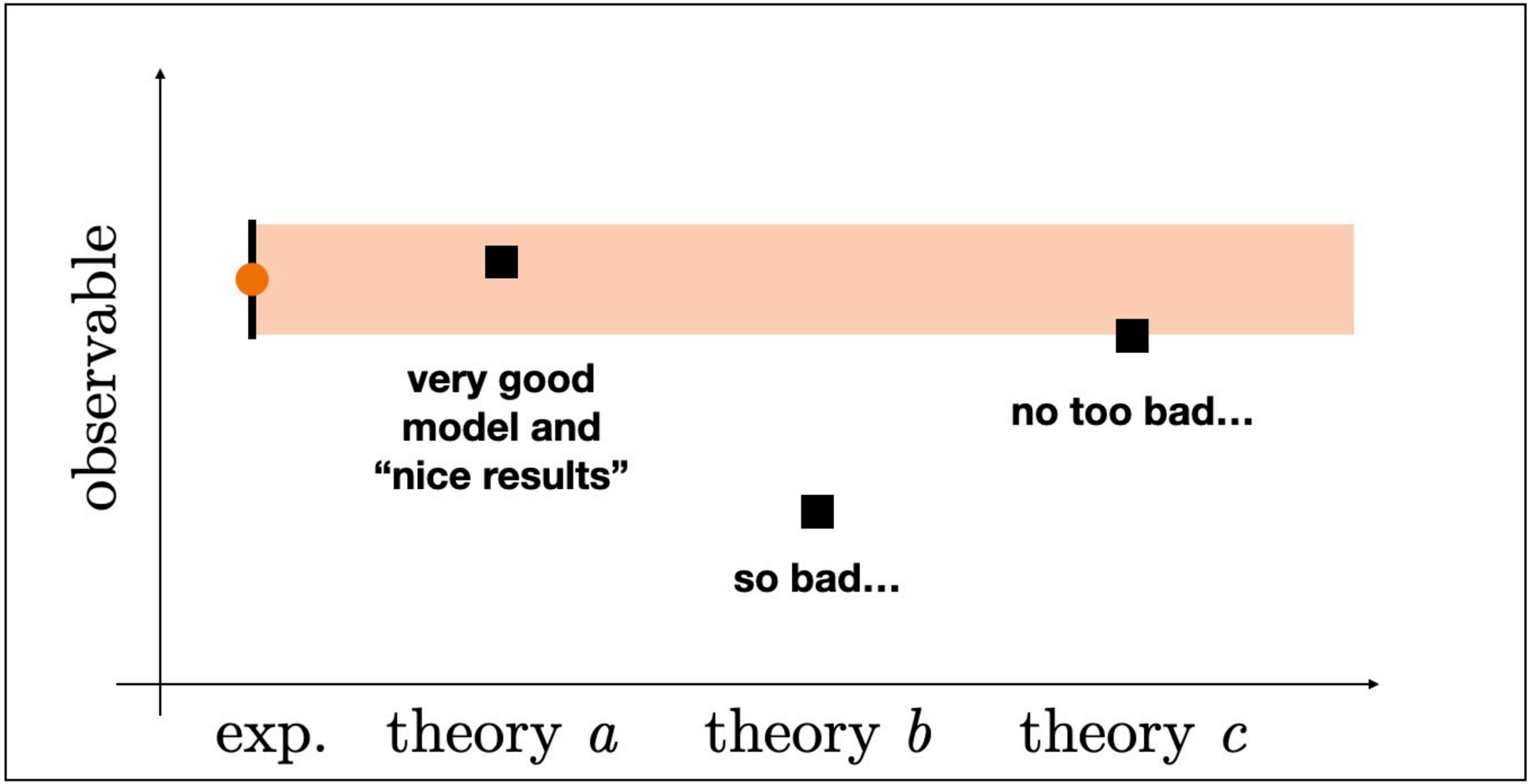
How do we quantify uncertainties in *ab initio* nuclear nuclear?

What is the role of effective field theory (EFT) in improving nuclear physics predictions?

Lecture 2
Why do *Bayesian* inference?

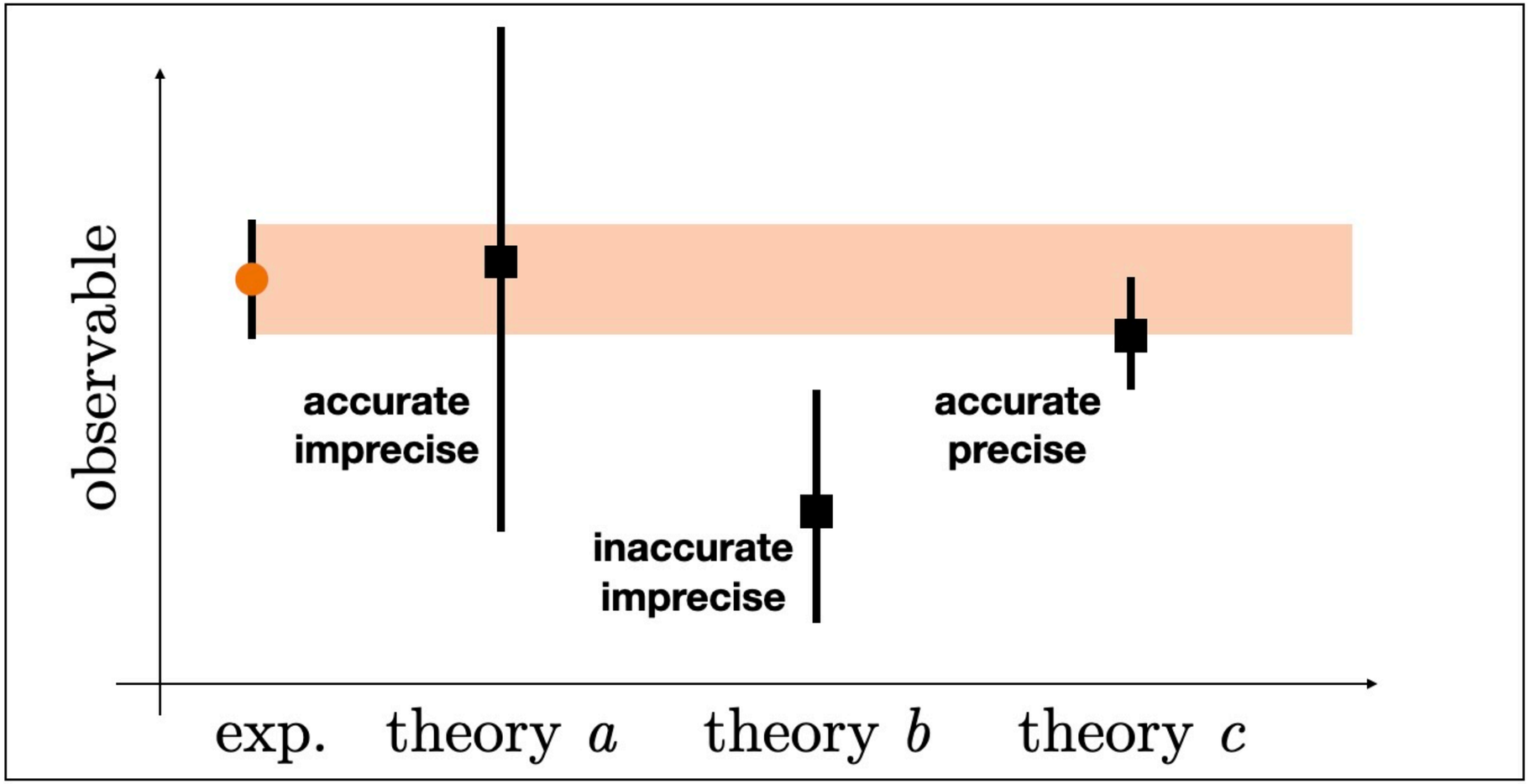
Quantified uncertainties are useful

One could be arrogant and say: *"This is theory, there are no uncertainties"*



Quantified uncertainties are useful

uncertainties allows for interesting questions: do we measure again, do we need other data, do predictions correlate, are we seeing new physics, which theory to use where, should we combine (mix) theories, ...?



A rationale for statistics

- In Physics, we are always[1] uncertain about propositions.
- Statistics is the study of uncertainties.
- We can use probability to measure and combine uncertainties.
- Probabilities are personal (subjective) [other people might disagree about this statement].
- We are interested in solving a fundamental problem of inference; quantifying the posterior predictive probability (ppd) $p(\mathscr{F} \,|\, \mathscr{D}, I)$

  in english: we wish to predict future (unseen) data $\mathscr{F}$ based on (conditioned on) past (seen) data $\mathscr{D}$ and available background knowledge I. We are uncertain about $\mathscr{F}$, and must describe it in terms of probability p. The first step to enable quantitative calculations of this ppd is to construct a theory T.

- Theory T plays a key role since it enables the past data to be "forgotten", i.e., $p(\mathscr{F} \,|\, \mathscr{D}, I) = p(\mathscr{F} \,|\, T, I)$.
- In the realm of T, we derive a model M (including a discrepancy term) to make quantitative statements.
- Models come with parameters $\overrightarrow{\alpha}$, whose values we are uncertain of.
- We compute the ppd for some unseen nuclear observable y_{th} based on seen data $\mathscr{D}$ as: $p(y_{\text{th}} \,|\, \mathscr{D}, M, I) = \int p(y_{\text{th}} \,|\, \overrightarrow{\alpha}, M, I) p(\overrightarrow{\alpha} \,|\, \mathscr{D}, M, I)\, d\overrightarrow{\alpha}$
- We quantify the probabilities that measure our uncertainty in the model parameters: $p(\overrightarrow{\alpha} \,|\, \mathscr{D}, M, I)$.

[1] Cromwells rule: leave a little probability for the moon being made of green cheese; it can be as small as 1 in a million, but have it there since otherwise an army of astronauts returning with samples of the said cheese will leave you unmoved

D. V. Lindley, The Statistician (2000)

Bayes' rule: from likelihood & prior to posterior

- Collect N data points that we gather in a data vector $\mathscr{D}$
- To explain the data, propose some model M, depending on parameters $\vec{\alpha}$
- Apply Bayes' rule

$$p(\vec{\alpha} \,|\, \mathscr{D}, M, I) = \frac{p(\mathscr{D} \,|\, \vec{\alpha}, M, I) \cdot p(\vec{\alpha} \,|\, M, I)}{p(\mathscr{D} \,|\, M, I)}$$

Posterior **Likelihood** **Prior**

Marginal likelihood

most likely not Rev. T. Bayes

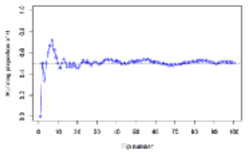
- The **prior** encodes our knowledge about the parameter values before analyzing the data
- The **likelihood** is the probability of the data given a set of parameters
- The **marginal likelihood** (or model evidence) provides normalization of the posterior
- The **posterior** is the complete inference and resulting probability density for the parameters $\vec{\alpha}$

Probability: a brief taxonomy

$$\int_X f \, d\mu$$

Mathematical (Kolmogorov) irrespective of interpretation. A mathematical framework that places probability in measure theory [only mathematical]

Classical (Laplace) based on symmetry and a finite number of equally likely outcomes [very limited and idealised scope]



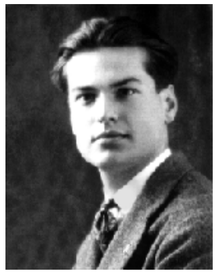
Frequency (Venn, von Mises, Fisher, Neyman) probability is uniquely determined by the relative ratio of occurrence. [restricted to repeatable phenomena and thus of limited scope]

Bayesian (Bayes, Price, Laplace) applies to an event about which the agent is uncertain [prior dependence, not an issue but a virtue if you ask me...]

objective (Keynes, Carnap, Jeffries, Cox, Jaynes) the probability (rational belief) of a proposition X given proposition Y is the "degree to which Y *logically* entails X". Given the same evidence, all rational human beings will entertain the same degree of belief in a hypothesis [how to always find this simultaneous logical connection to evidence]

subjective (Ramsey, de Finetti, Savage) probability *is* a subjective expression of confidence held by an individual (or team) subject only to coherence requirements. Probability is a normative and time indexed measure of anything *you* don't know. [subjectivity, how to elicit personal probabilities]

$$P_{\text{🚶}}^{\tau}(X_1 = x_1; \mathsf{H}(\tau))$$



Bruno de Finetti

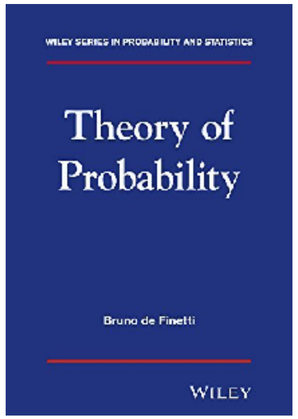
b. Innsbruck, 13 June 1906 – d. Rome, 20 July 1985

An Italian mathematician, probabilist, and actuary (Generali) who made profound contributions to the theory of probability.

During the 1940s he taught mathematical analysis in **Padova.**

He independently formulated the subjective analysis of probability, also devised by Frank P. Ramsey (Cambridge), and later championed by Savage, Lindley, and others.

"You should start with the *concept* of probability and then derive the theory"



My thesis, paradoxically, and a little provocatively, but nonetheless genuinely, is simply this:

PROBABILITY DOES NOT EXIST

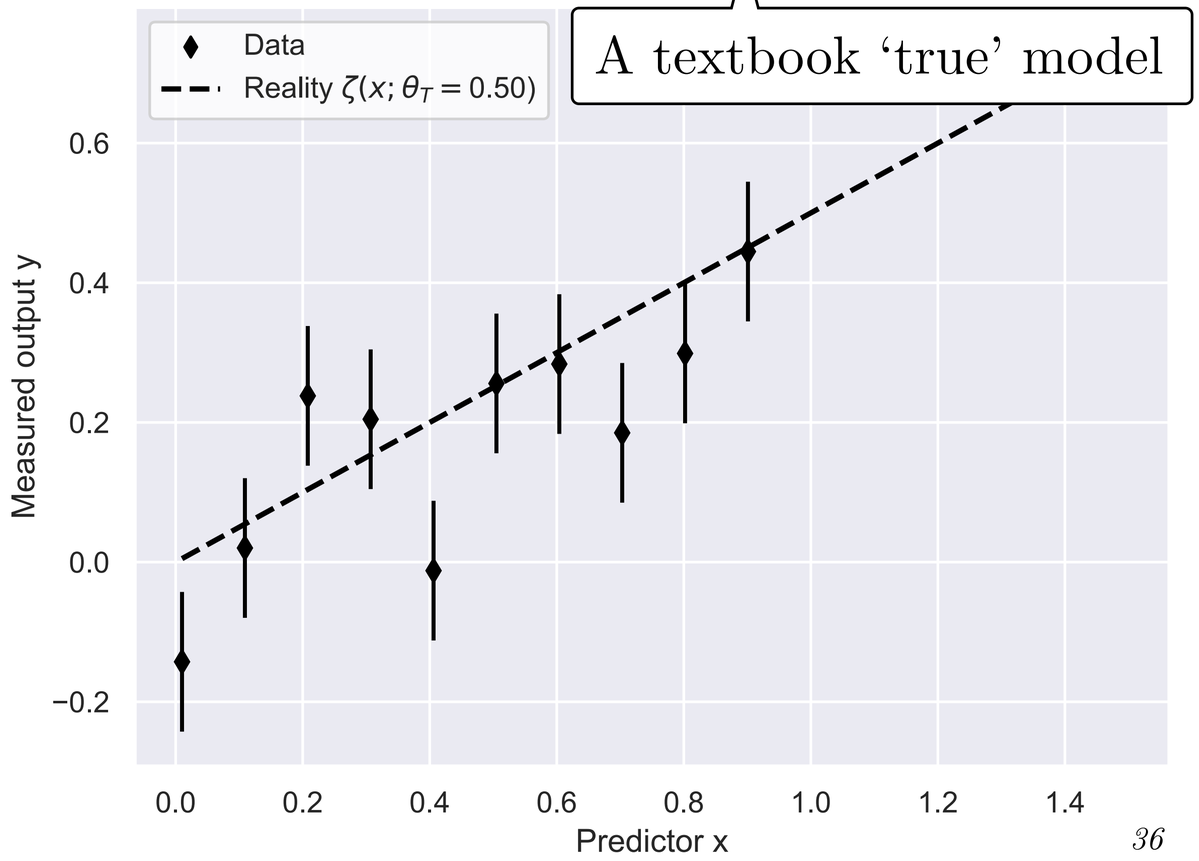
The abandonment of superstitious beliefs about the existence of Phlogiston, the Cosmic Ether, Absolute Space and Time, ..., or Fairies and Witches, was an essential step along the road to scientific thinking. Probability, too, if regarded as something endowed with some kind of objective existence, is no less a misleading misconception, an illusory attempt to exteriorize or materialize our true probabilistic beliefs.

[actual]

Linear regression: frequentist approach

- a complicated ordeal...

Let's assume data generated by $y_{\text{exp}}(x) = \theta_T \cdot x + \mathcal{N}(0,\sigma_T^2)$ with $\theta_T = 0.5$ and $\sigma_T = 0.1$, and posit a linear model $y_{\text{th}}(x) = \theta \cdot x$ for explaining 10 data points D. We now ask; what is θ?



In a frequentist analysis, we approach this by maximizing the data likelihood. The resulting ordinary least squares estimator is given by

$$\hat{\theta} = (\Phi^T\Phi)^{-1}(\Phi\theta - D) \approx 0.41$$

Now, we would like to know the uncertainty in our estimate of θ.

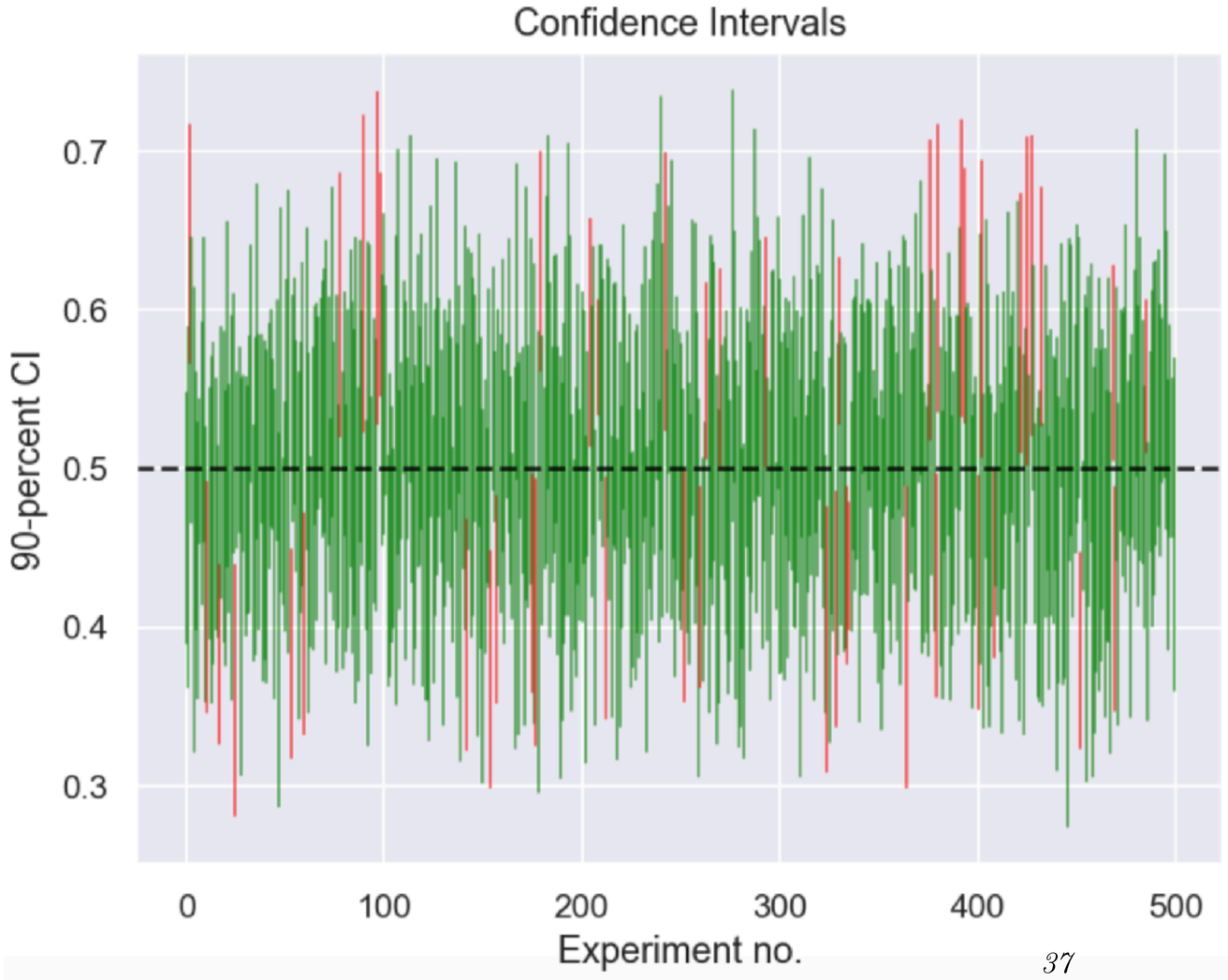
But in a frequentist approach we are only concerned wih data. So, we extract a 90% *confidence interval.*

$$I(\theta) = [0.29, 0.53]$$

What is the interpretation of this interval?

The p% confidence interval

overlaps with the true θ value p% of the time, *if* the model is right

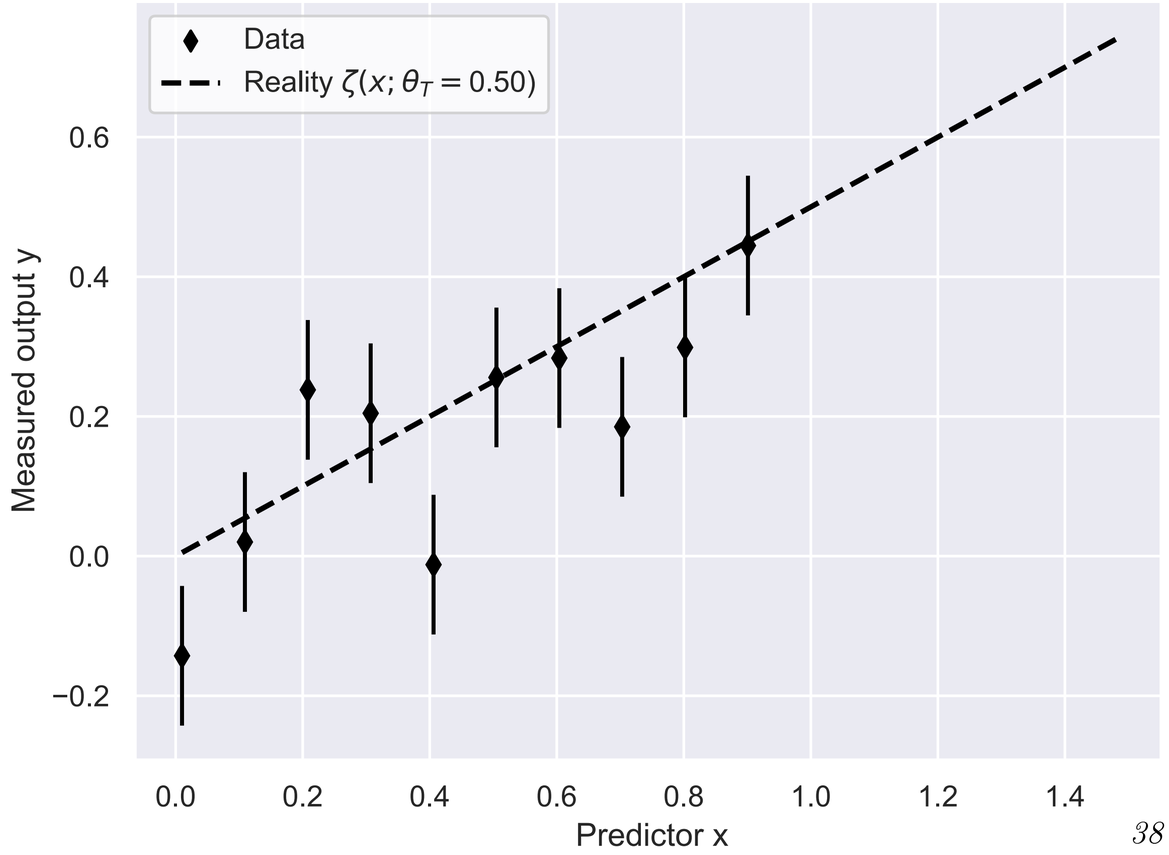


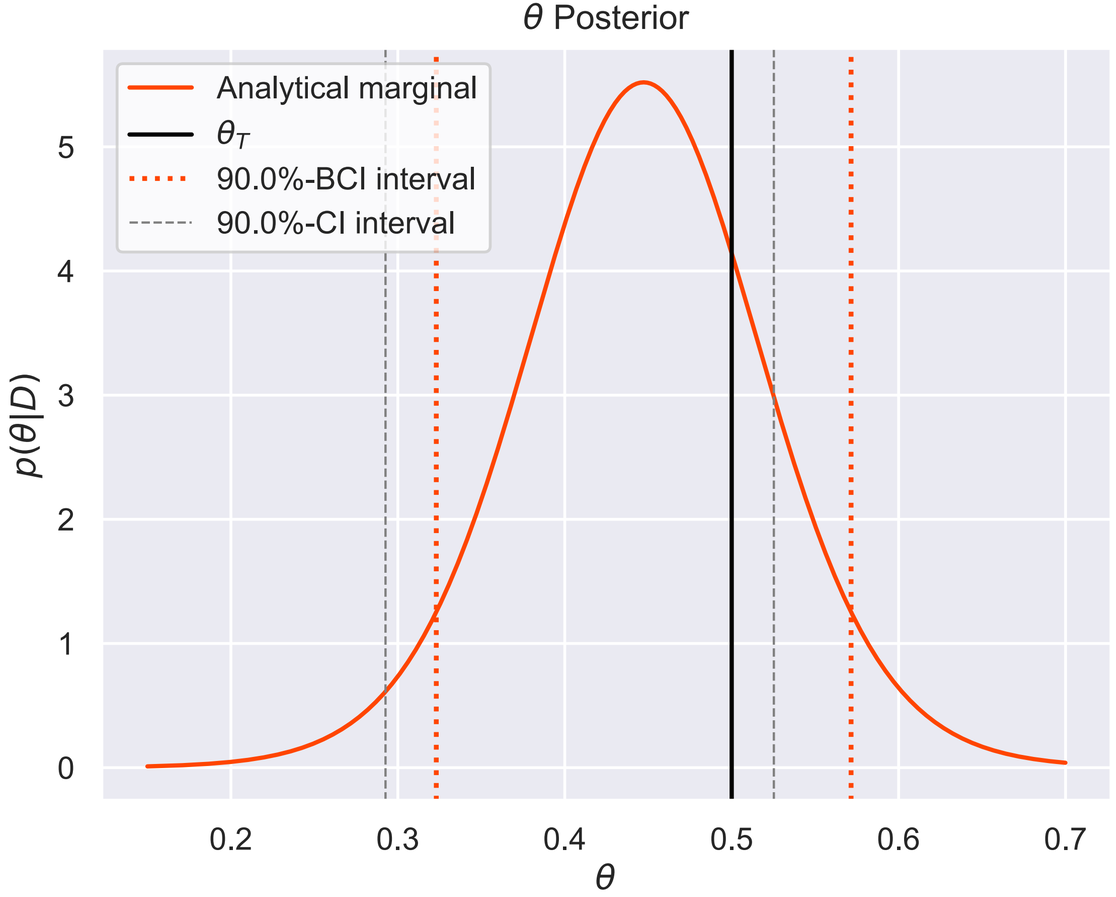
- It's a frequentist statement about the interval I.
- It doesn't apply to the data that you have, only the long-term frequency of future data sets.
- It is often misinterpreted to be a statement about the parameter.
- It only applies if the model is correct (all models are wrong...).

Linear regression: Bayesian approach

- conceptually easy (statistics made fun!)

Let's assume data generated by $y_{\mathrm{exp}}(x) = \theta_T \cdot x + \mathcal{N}(0,\sigma_T^2)$ with $\theta_T = 0.5$ and $\sigma_T = 0.1$,
and posit a linear model $y_{\mathrm{th}}(x) = \theta \cdot x$ for explaining 10 data points D. We now ask; what is θ?

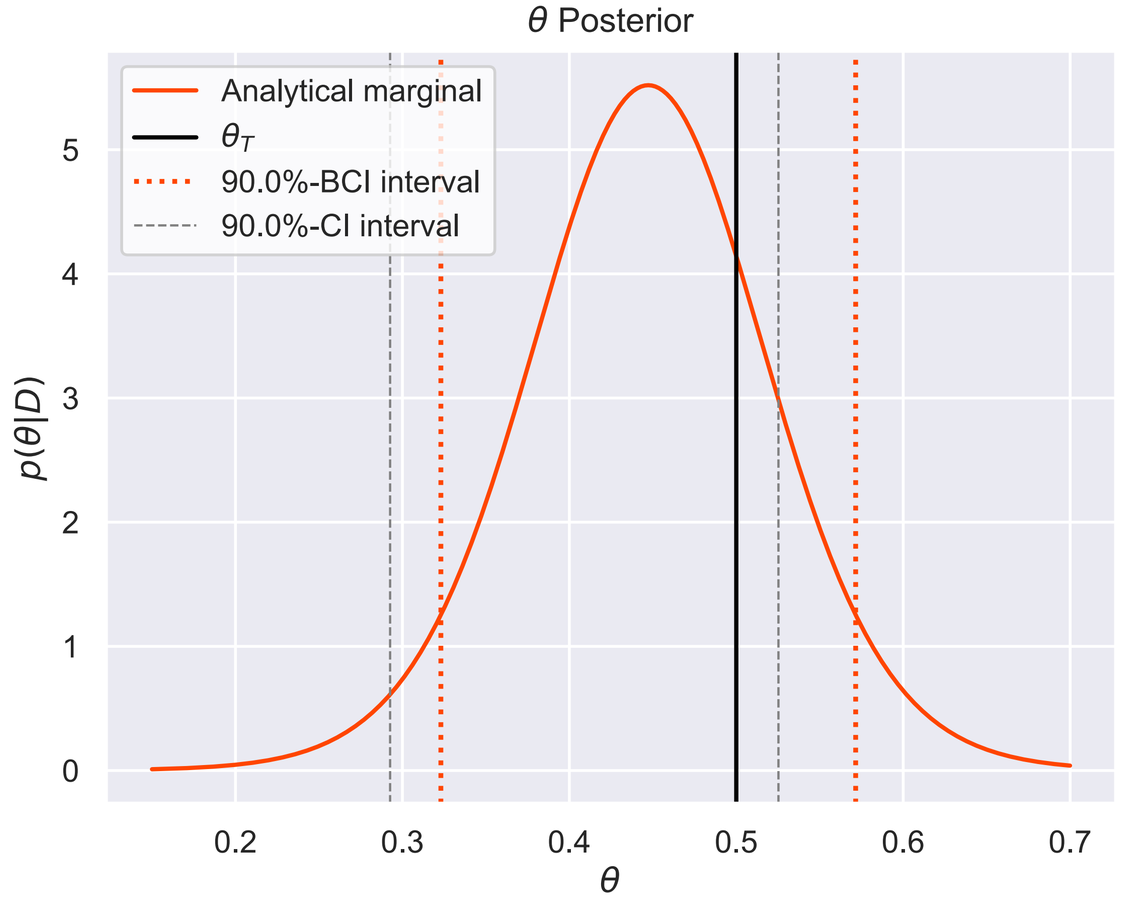




Linear regression: Bayesian approach

- conceptually easy (statistics made fun!)

Let's assume data generated by $\zeta(x) = x \cdot \theta_T + \mathcal{N}(0, \sigma^2)$ with $\theta_T = 0.5$ and $\sigma = 0.1$,
and posit a linear model $y = x \cdot \theta$ for explaining 10 data points D. We now ask; what is θ?



- The posterior $p(\theta \,|\, D, M, I)$ is a probabilistic statement about θ given the data D.
- It applies to the data we have.
- We can define a Bayesian credible interval covering some $p\%$ probability mass. Easier to interpret.
- It is conditioned on the model M, i.e., the statement hinges on the model being correct (still, all models are wrong...)

What to do when 'all models are wrong'

- probabilistic belief specification on the model discrepancy

We should add a model discrepancy term (the difference between the model and reality)

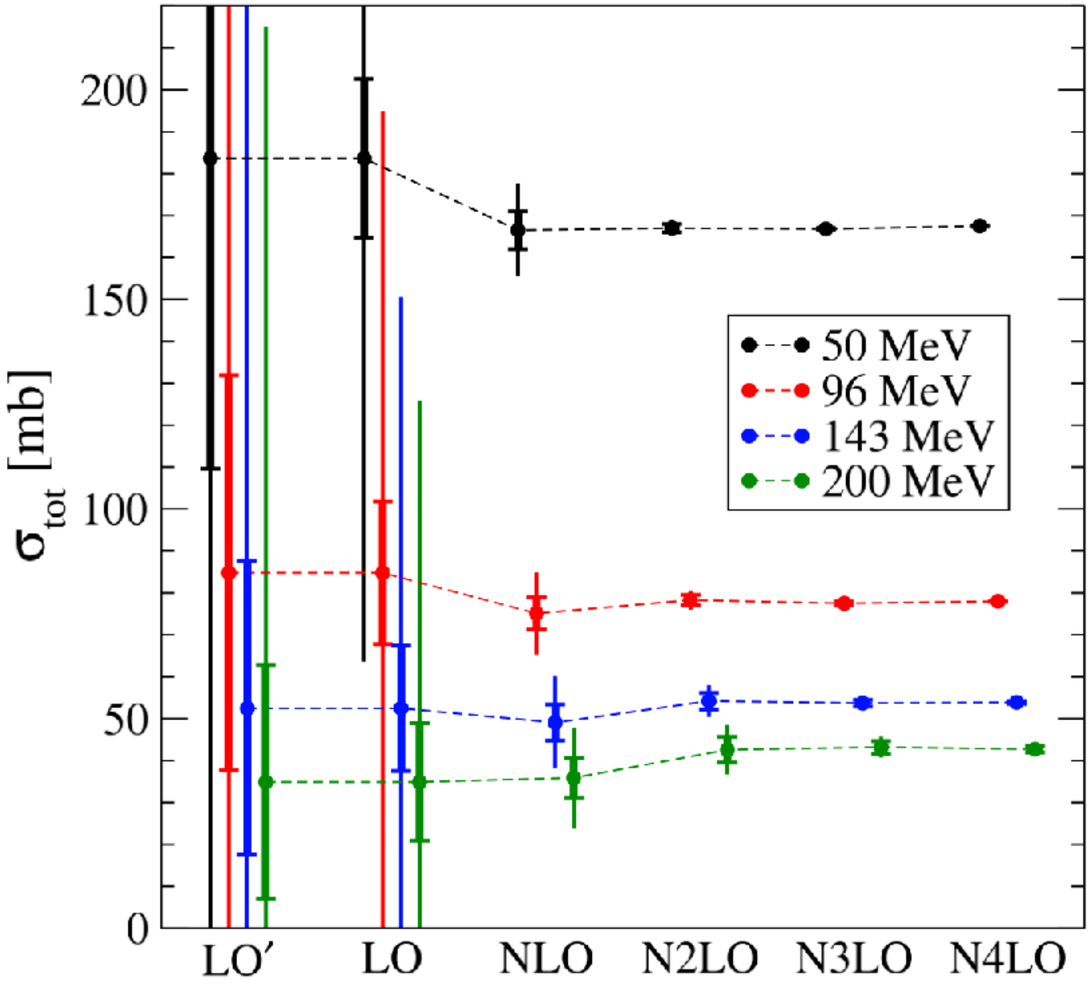
$$y_{\mathrm{exp}}(\vec{x}) = y_{\mathrm{th}}(\overrightarrow{\alpha};\vec{x}) + \delta y_{\mathrm{th}}(\overrightarrow{\alpha};\vec{x}) + \delta y_{\mathrm{exp}}(\vec{x})$$

We are of course uncertain about the model discrepancy, otherwise we'd (typically) 'include' it in the model. So, we need a probabilistic description of the discrepancy term.

The systematicity of the *ab initio* method is beneficial in this regard.

The model discrepancy in χEFT

- first example: the truncation error in neutron proton scattering



Given a potential description V_{np}, we compute the total elastic scattering cross section trivially.

How do we quantify the error due to missing (unresolved) physics in our potential V_{np}? With EFTs we have an obvious way forward:

$$y_{\rm th}^{(k)} = y_{\rm ref} \sum_{\nu=0}^{k} c_\nu \left(\frac{Q}{\Lambda_\chi}\right)^\nu \qquad \delta y_{\rm th}^{(k)} = y_{\rm ref} \sum_{\nu=k+1}^{\infty} c_\nu \left(\frac{Q}{\Lambda_\chi}\right)^\nu$$

$$c_i = \frac{y_{\rm th}^{(i)} - y_{\rm th}^{(i-1)}}{y_{\rm ref}(Q/\Lambda_\chi)^i}$$

Clearly, the question is: how do we estimate the unseen EFT expansion coefficients $c_{i>k}$?

neglected for now: *i) the EFT expansion is asymptotic, ii) the soft and hard scales are typically uncertain too, iii) the correlation structure of the expansion coefficients is likely non-trivial, and iv) power counting...*

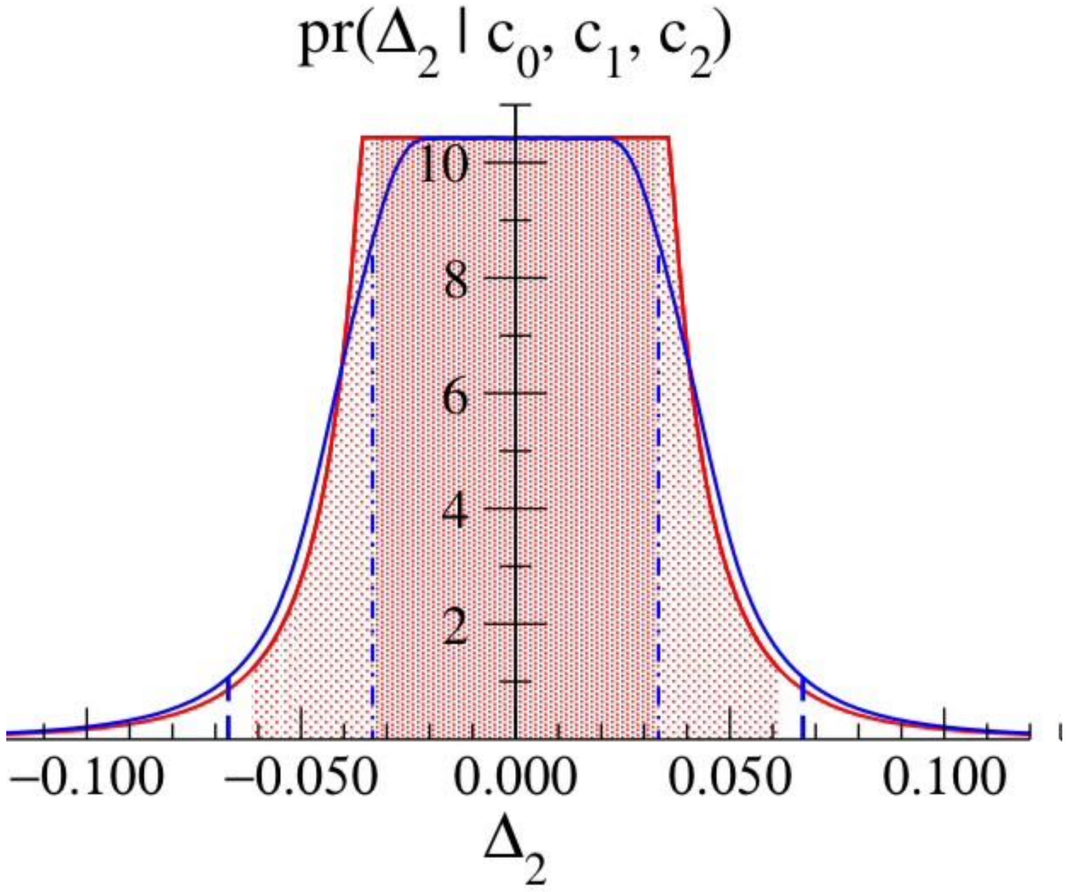
Furnstahl et al Phys Rev C (2015)

The model discrepancy in χEFT

- the truncation error

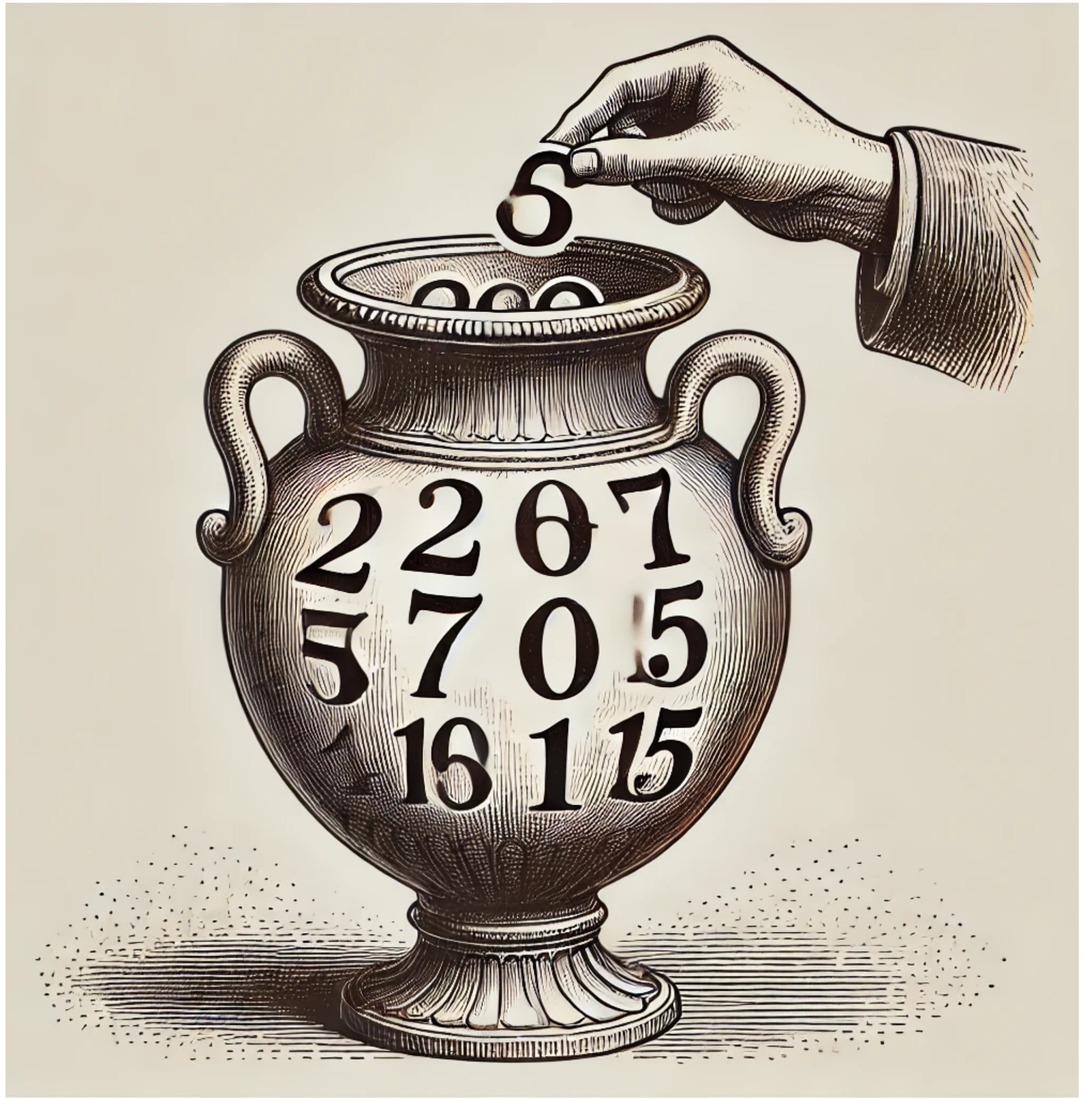
We are uncertain about the value of c_4, i.e., we seek $p(c_4 \,|\, c_0, c_2, c_3, I)$

A Bayesian argument could go like this: let's treat all c_i as random variables whose values are drawn independently from a normal distribution with variance $\bar{c}^2$. We further assume the possible values of $\bar{c}^2$ to follow some *a priori* specified distribution in accordance with our beliefs about the EFT expansion for y. We then learn about $\bar{c}^2$ from c_0, c_2, c_3, which in turn yields a probabilistic description of c_4.



Furnstahl et al Phys Rev C (2015)

Truncation errors: an analogy



We are interested in drawing numbers from an urn containing positive numbers. We draw numbers with replacement (putting the numbers back into the urn after each draw).

We have observed n draws $x_1, x_2, x_3, \ldots, x_n$. (corresponds to n orders in χEFT). For example 3,4,2,5,1,5.

What is the number we draw next?

In statistical terms, we seek the posterior predictive distribution (ppd) for the next draw x_{n+1} given $x_1, \ldots, x_n$

Truncation errors: an analogy

- prior distribution for the largest number A in the urn

We don't know A, i.e., we are uncertain about its value. Let's therefore assume a scale-invariant probability density (for which $\mathbb{P}(A \in [ka, kb])$ is independent of k.) called the log-uniform distribution

$$p(A) \propto \frac{1}{A}$$

This prior reflects our belief that A could span several orders of magnitude with equal probability. We can normalize this prior by defining upper and lower bounds which we remove in a limiting procedure at the end of the calculation. Let's forget about such rigor and assume finite limits that always cover the range of x_i values that we draw

Truncation errors: an analogy

- likelihood of observed draws $x_1, x_2, \ldots, x_n$

Given A, we assume each draw x_i is uniformly distributed between 0 and A

$$p(x_i \,|\, A) = \frac{1}{A} \text{ for } 0 \leq x_i \leq A.$$

Assuming independent draws, the probability of observing the entire dataset $x_1, x_2, \ldots, x_n$ given A is

$$p(x_1, x_2, \ldots, x_n \,|\, A) = \prod_{i=1}^{n} p(x_i \,|\, A) = \frac{1}{A^n} \text{ for } A \geq x_{\max}.$$

Here, $x_{\max} = \max(x_1, x_2, \ldots, x_n)$.

Truncation errors: an analogy

- using Bayes' theorem to evaluate the posterior

According to Bayes' theorem we have

$$p(A \,|\, x_1, x_2, \ldots, x_n) \propto p(x_1, x_2, \ldots, x_n \,|\, A) \cdot p(A).$$

Substituting for our prior and likelihood (of parameter A) we obtain

$$p(A \,|\, x_1, x_2, \ldots, x_n) \propto \frac{1}{A^n} \cdot \frac{1}{A} \text{ for } A \geq x_{\text{max}}.$$

This expression represents the posterior distribution for A given the observed data. We are *learning* more about A as we draw more numbers.

Truncation errors: an analogy

- using Bayes' theorem to evaluate the posterior

$$p(A \,|\, x_1, x_2, \ldots, x_n) \propto \frac{1}{A^n} \cdot \frac{1}{A} \text{ for } A \geq x_{\max}.$$

This expression represents the posterior distribution for A given the observed data. We are *learning* more about A as we draw more numbers.

We normalize the posterior by integrating over all values of A

$$\int_{x_{\max}}^{\infty} \frac{1}{A^{n+1}}\, dA = \frac{1}{n x_{\max}^n} \Rightarrow p(A \,|\, x_1, x_2, \ldots, x_n) = \frac{n x_{\max}^n}{A^{n+1}} \text{ for } A \geq x_{\max}$$

Truncation errors: an analogy

- computing the posterior predictive distribution

We seek

$$p(x_{n+1}\,|\,x_1, x_2, \ldots, x_n) = \int dA\, p(x_{n+1}, A\,|\,x_1, x_2, \ldots, x_n) = \int dA\, p(x_{n+1}\,|\,A) p(A\,|\,x_1, x_2, \ldots, x_n).$$

where we 'marginalized in' the A dependence and in the second step utilized that the draws are independent of each other. As for the previous draws we have

$$p(x_{n+1}\,|\,A) = \frac{1}{A} \text{ for } A \geq x_{n+1}.$$

Substituting the expressions for $p(x_{n+1}\,|\,A)$ and $p(A\,|\,x_1, x_2, \ldots, x_n)$, we obtain

$$p(x_{n+1}\,|\,x_1, x_2, \ldots, x_n) = n x_{max}^n \int_t^\infty dA\, \frac{1}{A^{n+2}}.$$

Truncation errors: an analogy

- computing the posterior predictive distribution

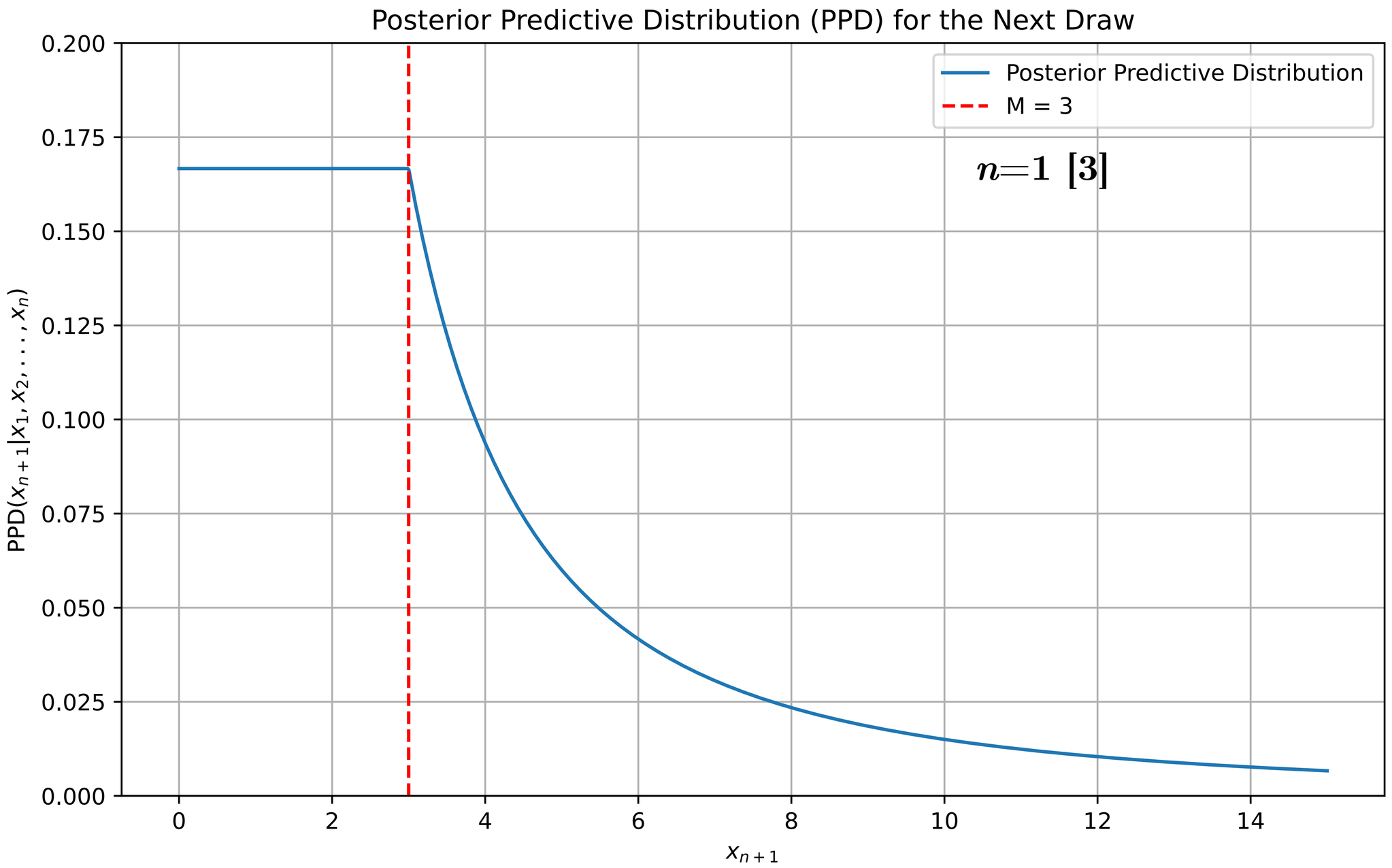
- The likelihood $p(x_{n+1}\,|\,A)$ is only non-zero for $A \geq x_{n+1}$
- The posterior $p(A\,|\,x_1, x_2, \ldots, x_n)$ is only non-zero for $A \geq x_{\max}$
- So, we have two cases depending on whether $x_{n+1} \leq x_{\max}$ or $x_{n+1} > x_{\max}$, i.e., two possible lower bounds $t = x_{n+1}$ and $t = x_{\max}$, corresponding to drawing a number greater than any previous number or not.

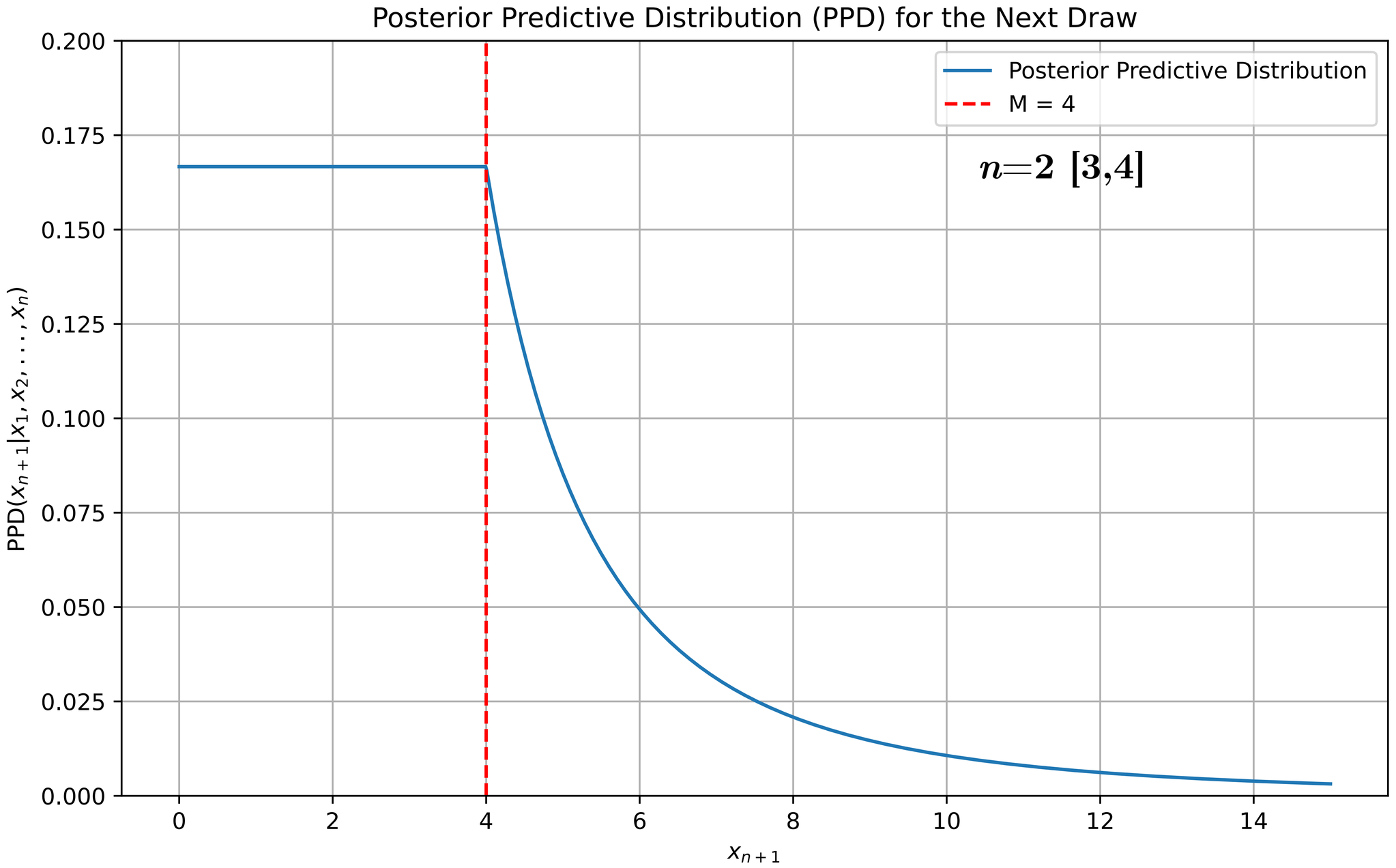
The ppd for x_{n+1} is:

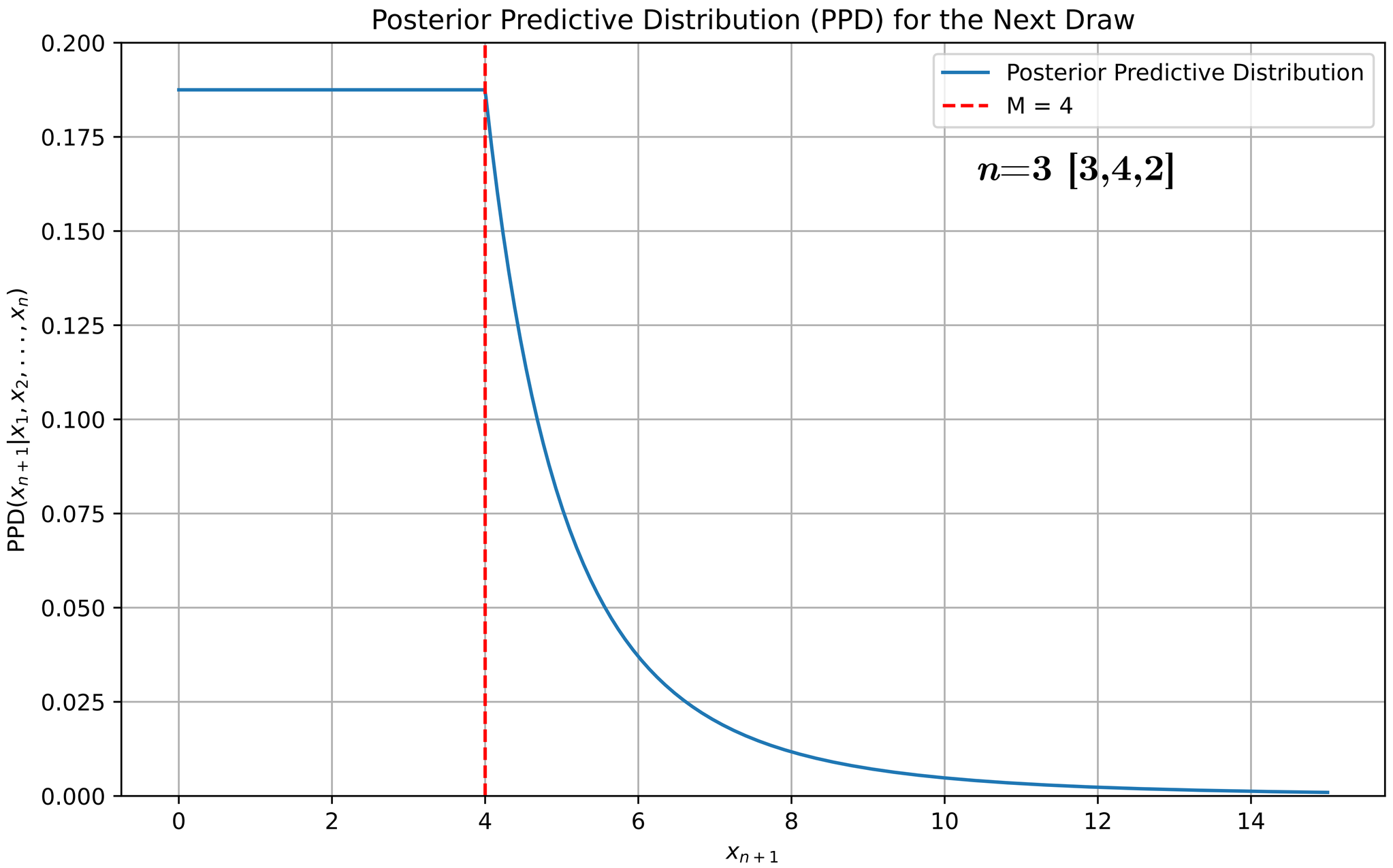
$$p(x_{n+1}|x_1, x_2, \ldots, x_n) = \frac{n}{n+1}\frac{1}{x_{\max}} \begin{cases} 1 & \text{if } x_{n+1} \leq x_{\max} \\ \left(\frac{x_{\max}}{x_{n+1}}\right)^{n+1} & \text{if } x_{n+1} > x_{\max} \end{cases}$$

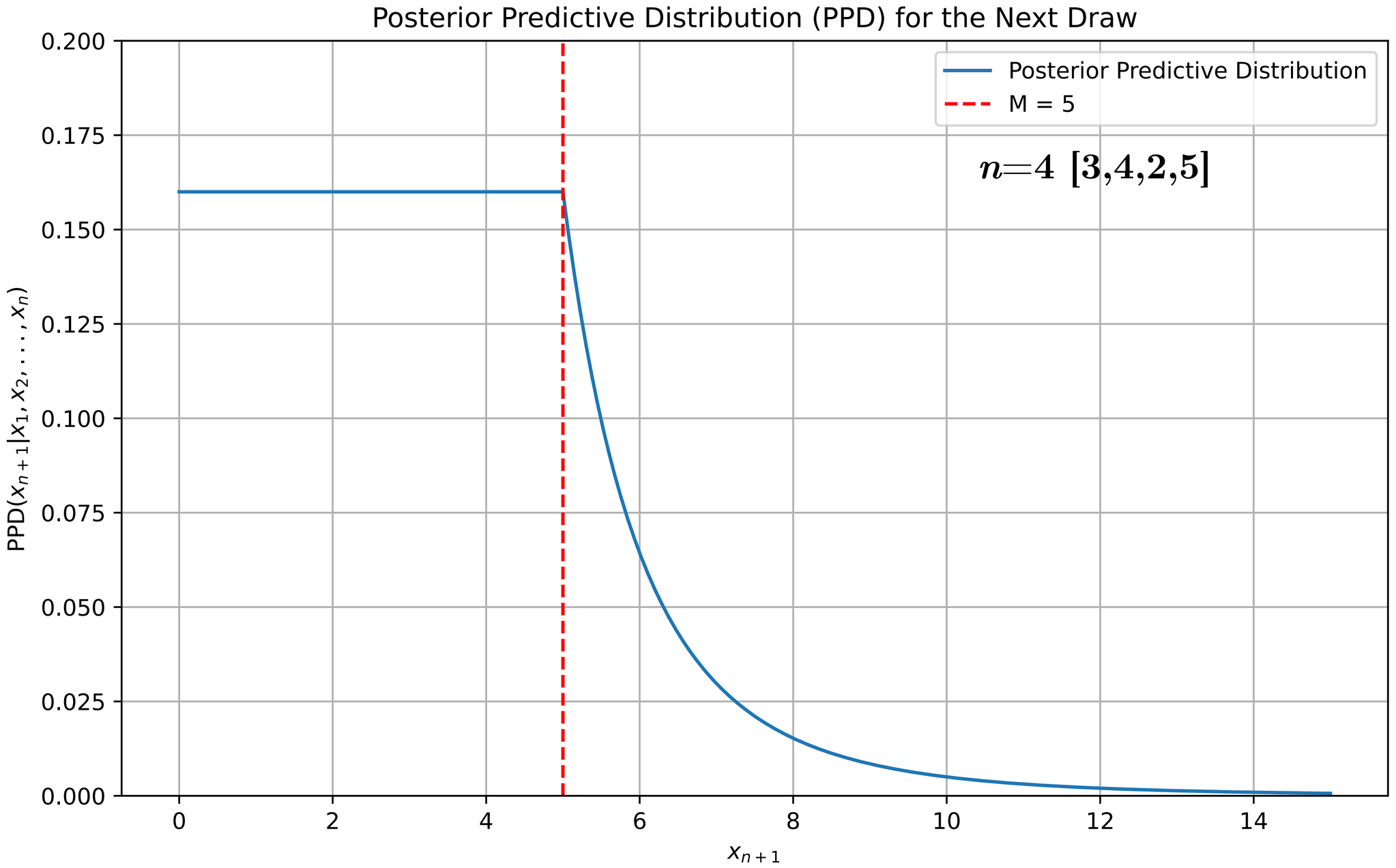
Uniform for $x_{n+1} \leq x_{\max}$ and power-law decay for $x_{n+1} > x_{\max}$

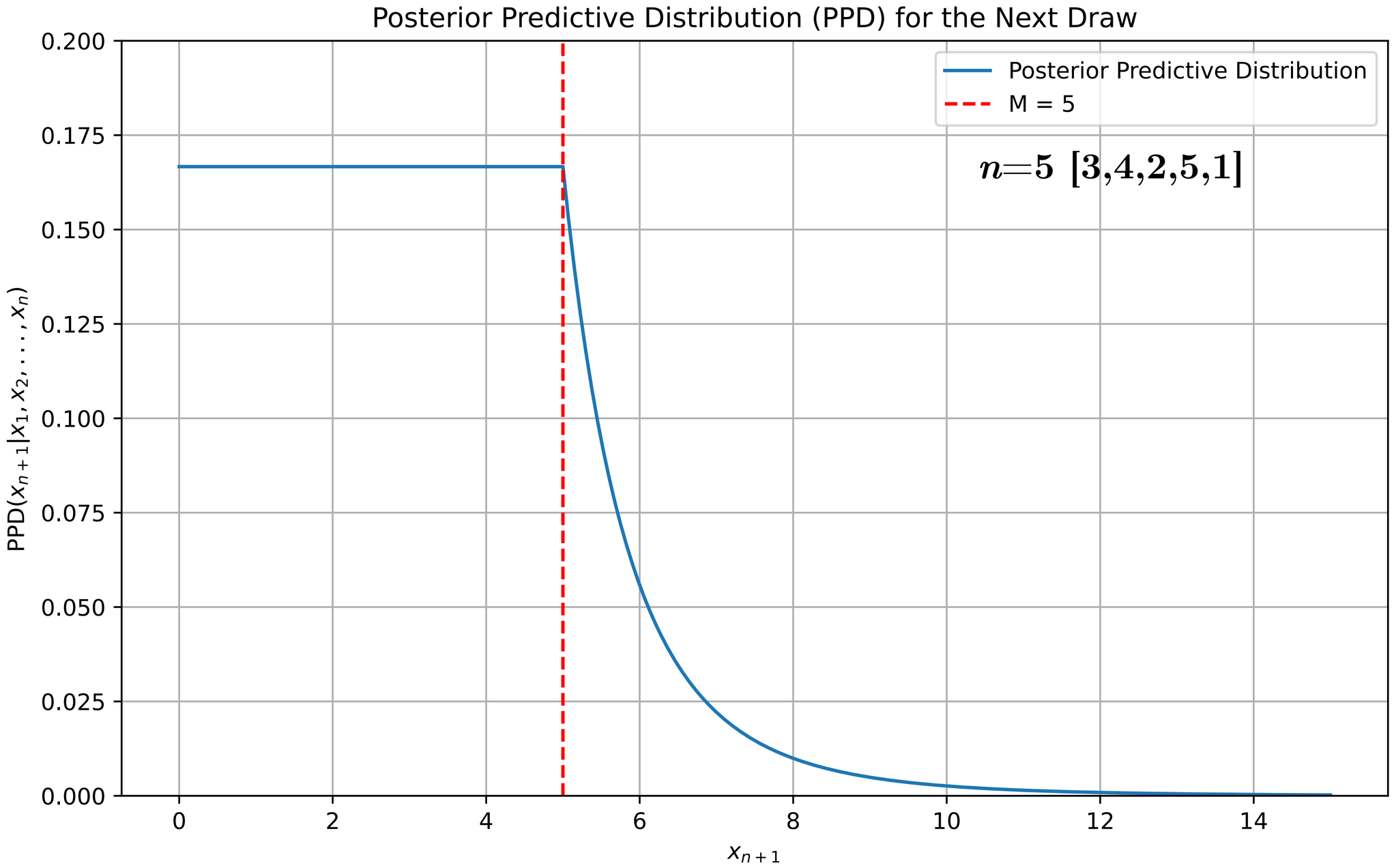
Furnstahl et al Phys Rev C (2015)

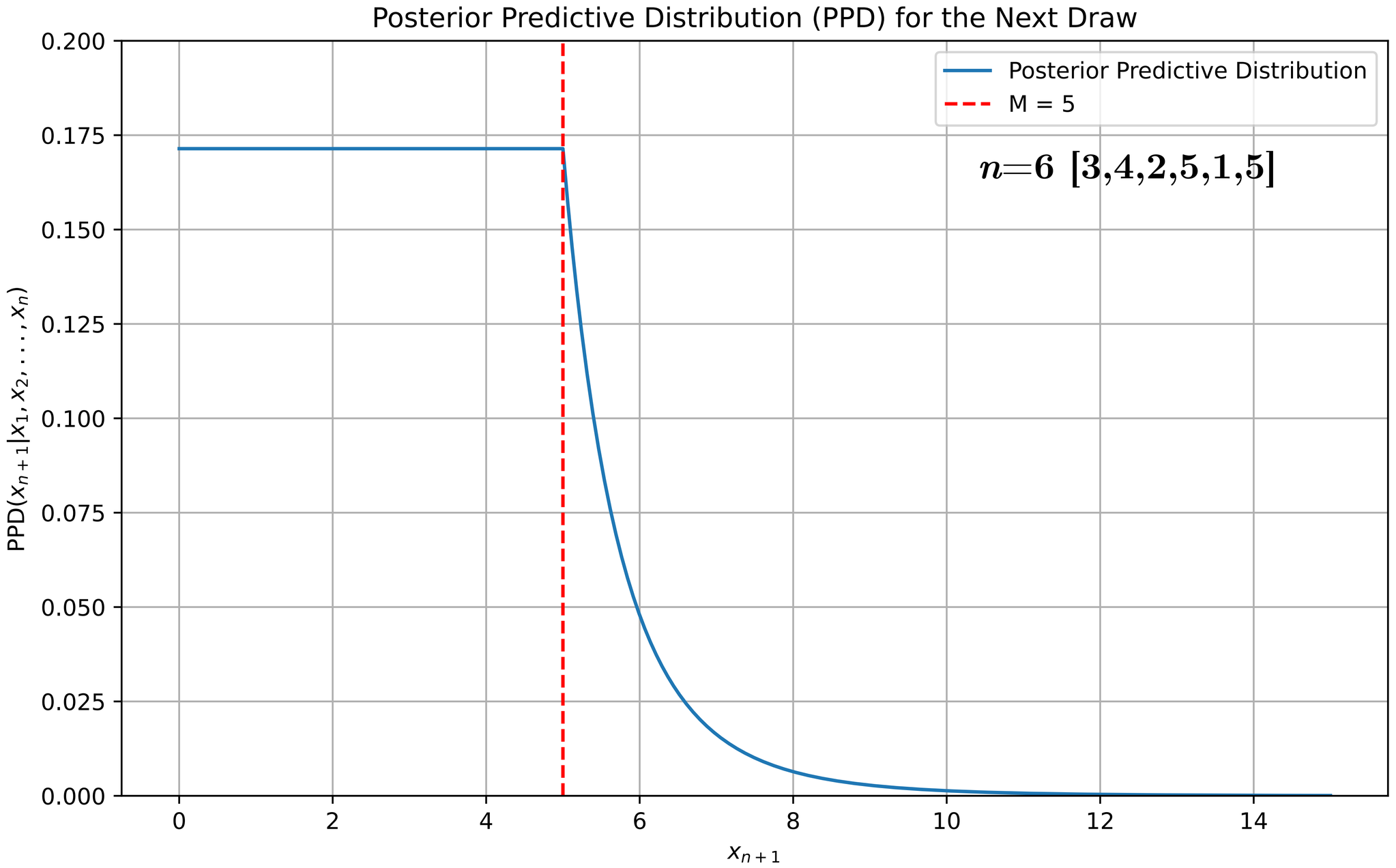

Posterior Predictive Distribution (PPD) for the Next Draw
Posterior Predictive Distribution
M = 3
$n=1$ [3]
$\mathrm{PPD}(x_{n+1}|x_1, x_2, \ldots, x_n)$
x_{n+1}
0.200
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
0
2
4
6
8
10
12
14

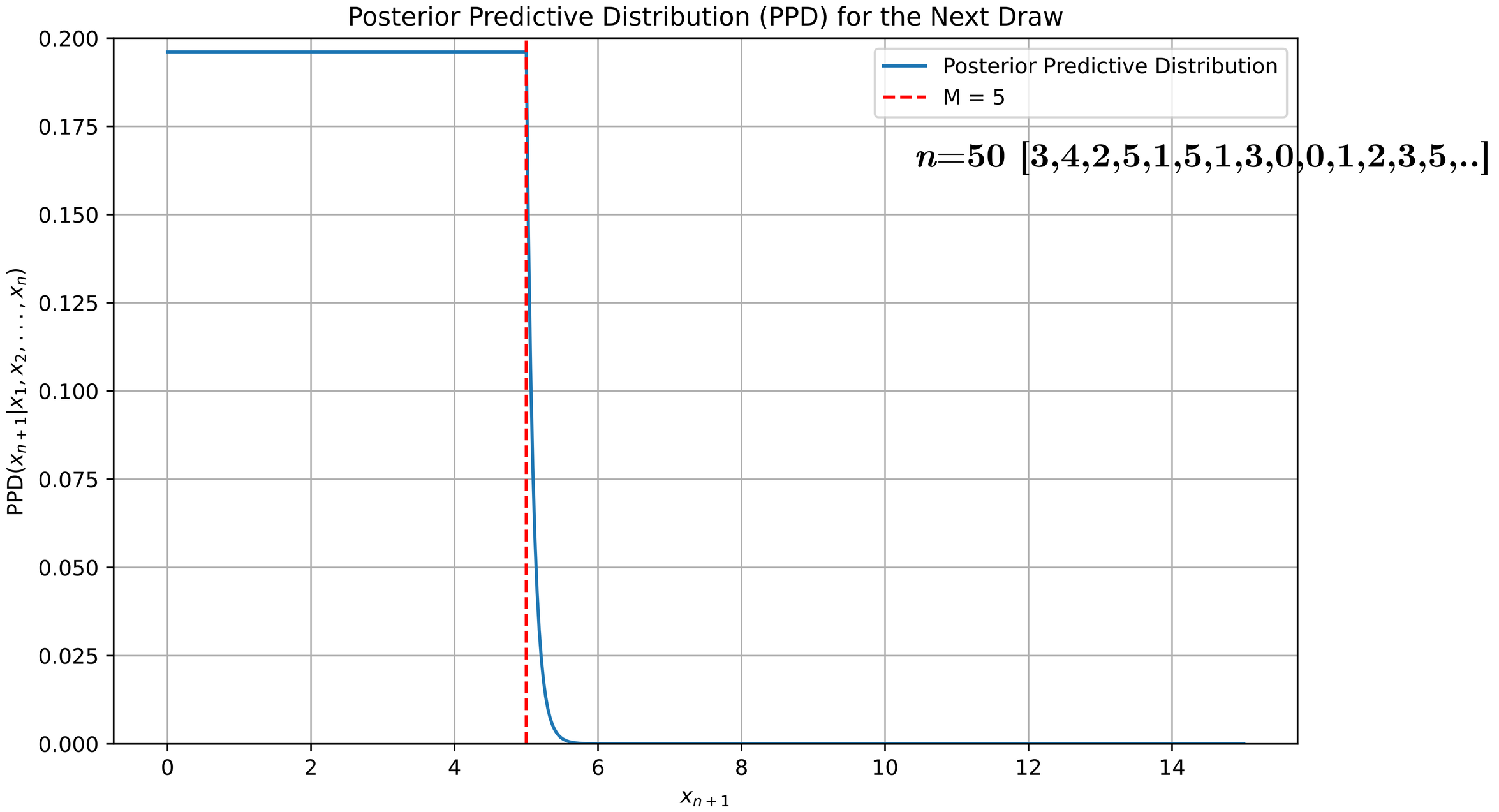

Posterior Predictive Distribution (PPD) for the Next Draw
Posterior Predictive Distribution
M = 4
$n=2$ [3,4]
$\mathrm{PPD}(x_{n+1}|x_1, x_2, \ldots, x_n)$
x_{n+1}
0.200
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
0
2
4
6
8
10
12
14


Posterior Predictive Distribution (PPD) for the Next Draw
Posterior Predictive Distribution
M = 4
n=3 [3,4,2]
$\mathrm{PPD}(x_{n+1}|x_1, x_2, \ldots, x_n)$
x_{n+1}
0.200
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
0
2
4
6
8
10
12
14


Posterior Predictive Distribution (PPD) for the Next Draw
Posterior Predictive Distribution
M = 5
n=4 [3,4,2,5]
$\mathrm{PPD}(x_{n+1}|x_1, x_2, \ldots, x_n)$
x_{n+1}
0.200
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
0
2
4
6
8
10
12
14

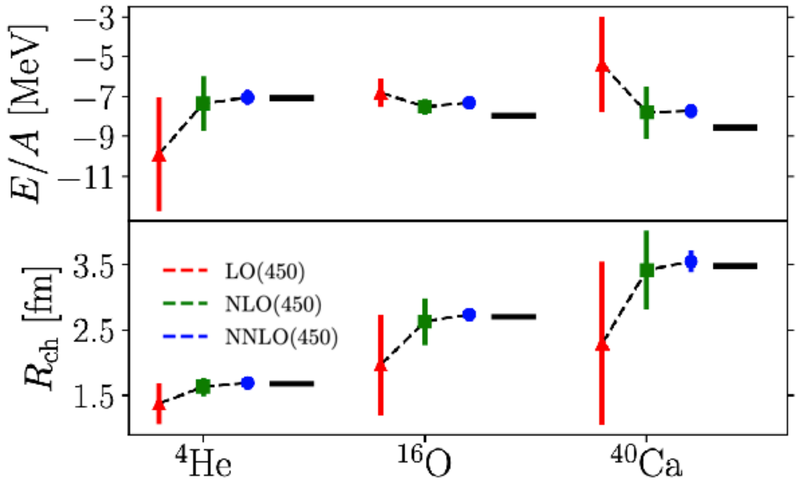

Posterior Predictive Distribution (PPD) for the Next Draw
Posterior Predictive Distribution
M = 5
n=5 [3,4,2,5,1]
0.200
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
PPD($x_{n+1}|x_1, x_2, \ldots, x_n$)
0
2
4
6
8
10
12
14
x_{n+1}


Posterior Predictive Distribution (PPD) for the Next Draw
Posterior Predictive Distribution
M = 5
n=6 [3,4,2,5,1,5]
PPD($x_{n+1}|x_1, x_2, \ldots, x_n$)
x_{n+1}
0.200
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
0
2
4
6
8
10
12
14


Posterior Predictive Distribution (PPD) for the Next Draw
Posterior Predictive Distribution
M = 5
n=50 [3,4,2,5,1,5,1,3,0,0,1,2,3,5,..]
$\mathrm{PPD}(x_{n+1}|x_1, x_2, \ldots, x_n)$
x_{n+1}
0.200
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
0
2
4
6
8
10
12
14

Estimating the EFT truncation-error

ab initio coupled cluster calculations of A=4,16,40 nuclei



What one typically see in papers

$$y_{\rm th}^{(k)} = y_{\rm ref}\sum_{\nu=0}^{k} c_\nu \left(\frac{Q}{\Lambda_\chi}\right)^\nu \qquad \delta y_{\rm th}^{(k)} = y_{\rm ref}\sum_{\nu=k+1}^{\infty} c_\nu \left(\frac{Q}{\Lambda_\chi}\right)^\nu \qquad c_i = \frac{y_{\rm th}^{(i)} - y_{\rm th}^{(i-1)}}{y_{\rm ref}(Q/\Lambda_\chi)^i}$$

$$\delta y_{\rm th}^{(\rm NjLO)} = y_{\rm ref} Q^{j+2} \max(c_0, c_1, c_2, \ldots, c_{j+1})$$

is an $(\frac{n_c}{n_c+1}) \times 100\,\%$ degree of belief interval about the NjLO prediction

One can show that placing a Gaussian prior $p(c_i \,|\, I) = \mathcal{N}(0,\bar{c}^2)$ on the expansion coefficients leads to a Gaussian posterior for the truncation error $p(\delta y_{\rm th}^{(k)} \,|\, \bar{c}^2, Q, \Lambda_\chi, I) = \mathcal{N}(0,\sigma_{\rm th}^2)$ where the variance of the EFT truncation error is given by

$$\sigma_{\rm th}^2 = \bar{c}^2 y_{\rm ref}^2 \frac{[Q/\Lambda_\chi]^{2(k+1)}}{1-[Q/\Lambda_\chi]}$$

one can further show that placing a (conjugate) inverse gamma prior on the variance $\bar{c}^2$ yields a student t distribution for the EFT truncation error

Melendez et al Phys Rev C (2019)
Ekström et al Phys Rev C (2018)
Furnstahl et al Phys Rev C (2015)
Epelbaum et al Phys Rev Lett (2014)

The BUQEYE Cheatsheet for Pointwise Truncation Errors (arXiv:1904.10581)

From observable y, extract coefficients

$$\begin{aligned} \vec{y}_k &\equiv \{y_0, y_1, \cdots, y_k\} \\ \Rightarrow \vec{c}_k &\equiv \{c_0, c_1, \cdots, c_k\} \end{aligned} \tag{A1}$$

Choose ν_0 and τ_0. Update hyperparameters

$$\nu = \nu_0 + n_c \tag{A7}$$

$$\nu\tau^2 = \nu_0\tau_0^2 + \vec{c}_k^{\,2} \tag{A8}$$

Compute posterior

$$\text{pr}(y \,|\, \vec{y}_k, Q) \sim t_\nu\left[y_k, y_{\text{ref}}^2 \frac{Q^{2(k+1)}}{1-Q^2}\tau^2\right] \tag{A13}$$

```python
import numpy as np
y_ref = 20.0; Q = 0.3; k = 3
y_k = [21.7, 27.3, 25.4, 26.2]
c_k = np.array([y_k[0] / y_ref] + [
  (y_k[n] - y_k[n-1]) / (y_ref * Q**n)
  for n in range(1, k+1)])

nu_0 = 1; tau_0 = 1 # ~Uninformative
nu = nu_0 + len(c_k)
tau_sq = \
  (nu_0 * tau_0**2 + c_k @ c_k) / nu

from scipy.stats import t
scale = y_ref * Q**(k+1) * \
  ( tau_sq / (1 - Q**2) )**0.5
y = t(nu, y_k[-1], scale)
dob = y.interval(0.95) # (25.7, 26.7)
```

Next lecture

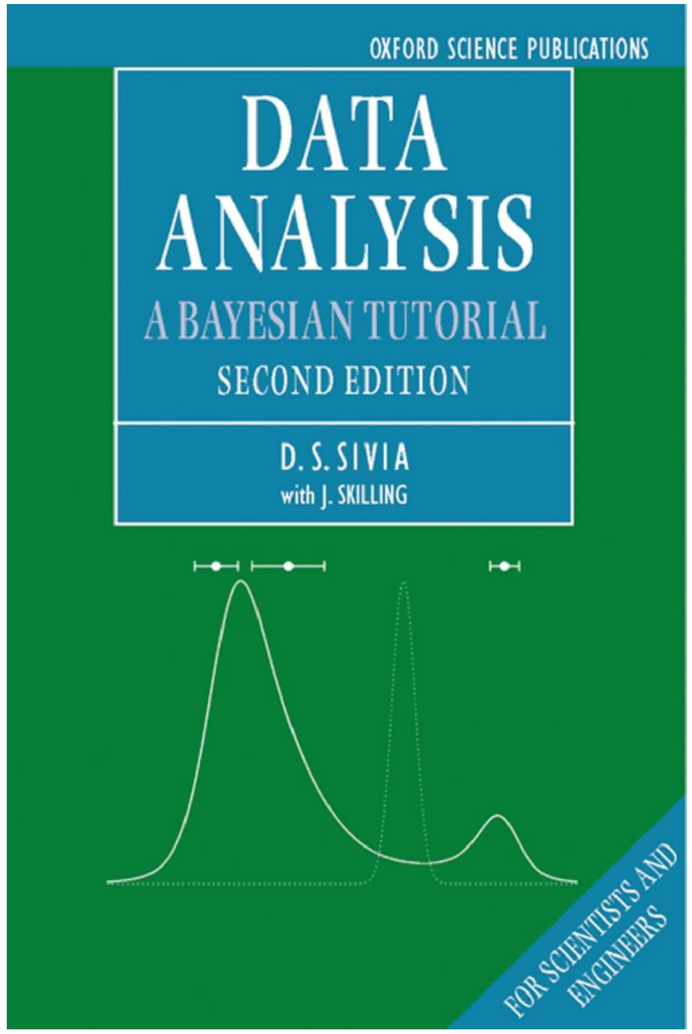
What are the computational statistics challenges in *ab initio* methods?

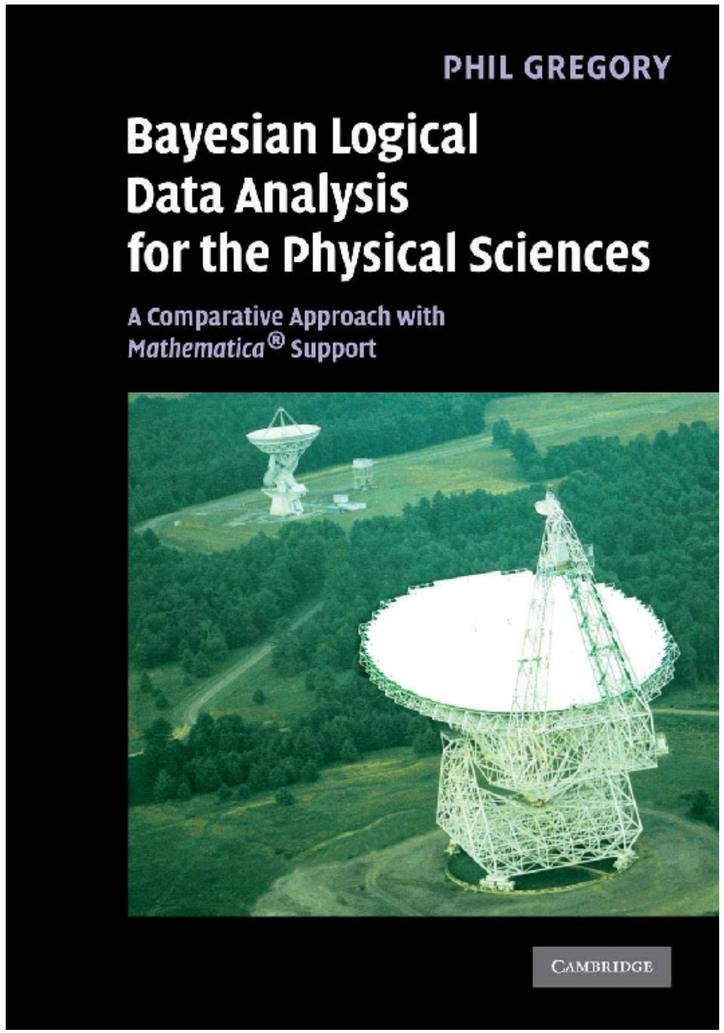
How can we emulate many-body calculations based on no-core shell model and coupled-cluster theory?

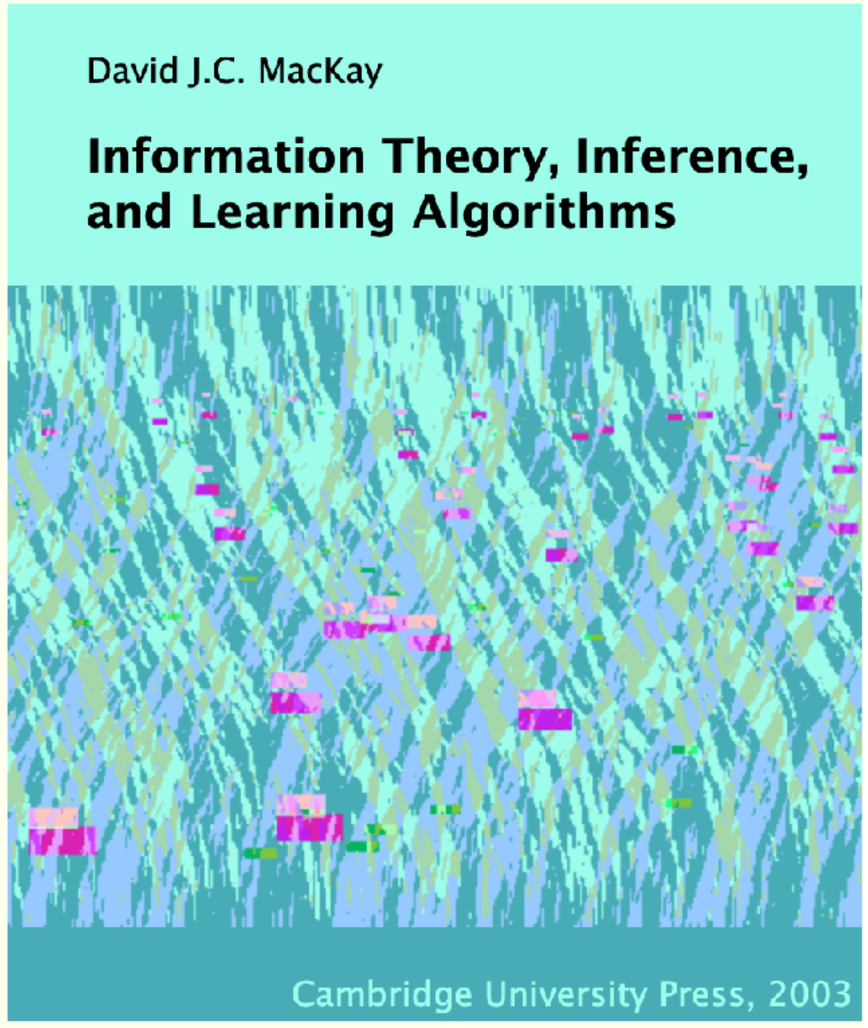
Lecture 3

Overcoming the computational challenge

Several good books out there





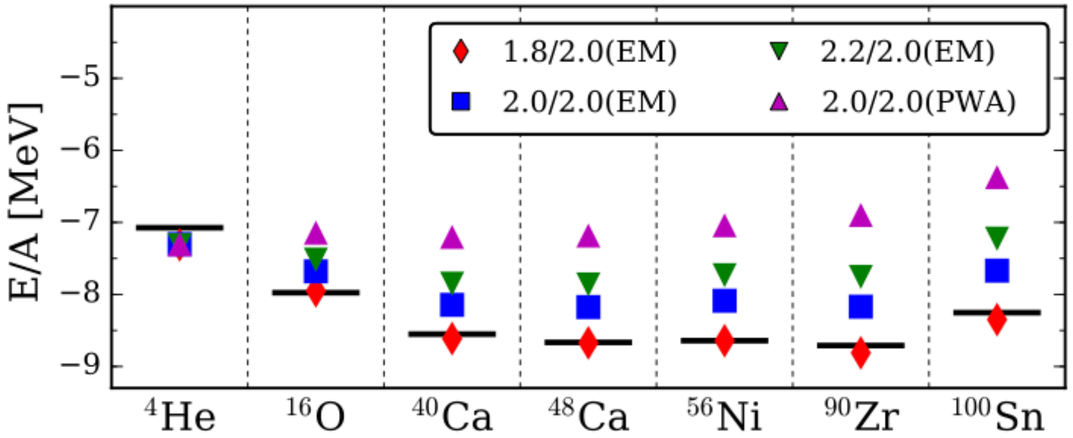


https://www.inference.org.uk/mackay/itila/

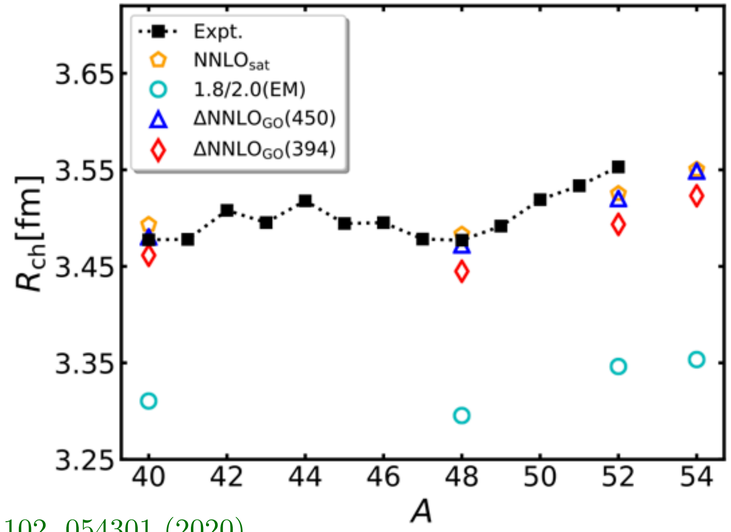
reminder: some interactions work better than others

- we don't know why...

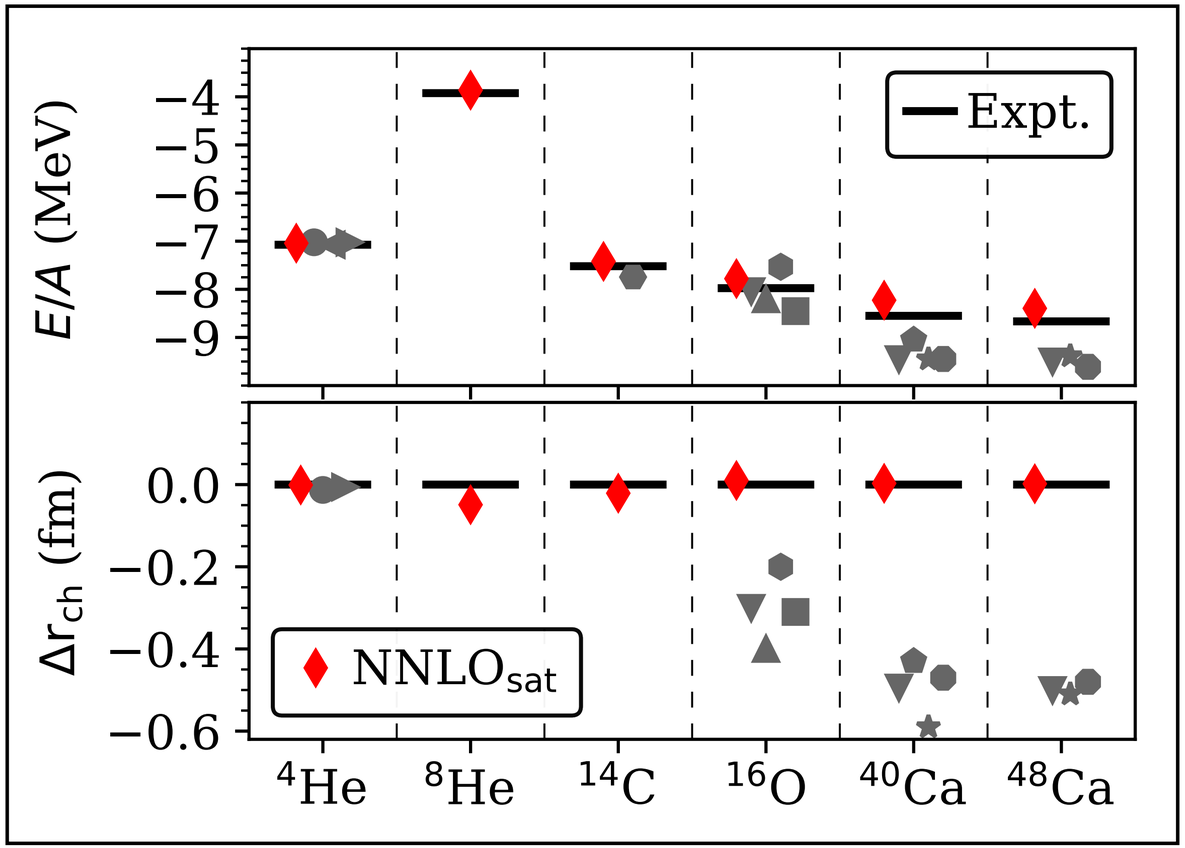
$$[T + V(\vec{\alpha})]\,|\Psi\rangle = E\,|\Psi\rangle$$



T. D. Morris et al. PRL (2018)



Jiang et al Phys. Rev. C 102, 054301 (2020)



Ekström et al Phys. Rev. C (2015)

Bayesian parameter estimation

- inferring the low-energy constants in χEFT

the whole thing should be included in predictions

$$y_{\text{exp}}(\vec{x}) = y_{\text{th}}(\vec{\alpha};\vec{x}) + \delta y_{\text{th}}(\vec{\alpha};\vec{x}) + \delta y_{\text{exp}}(\vec{x})$$

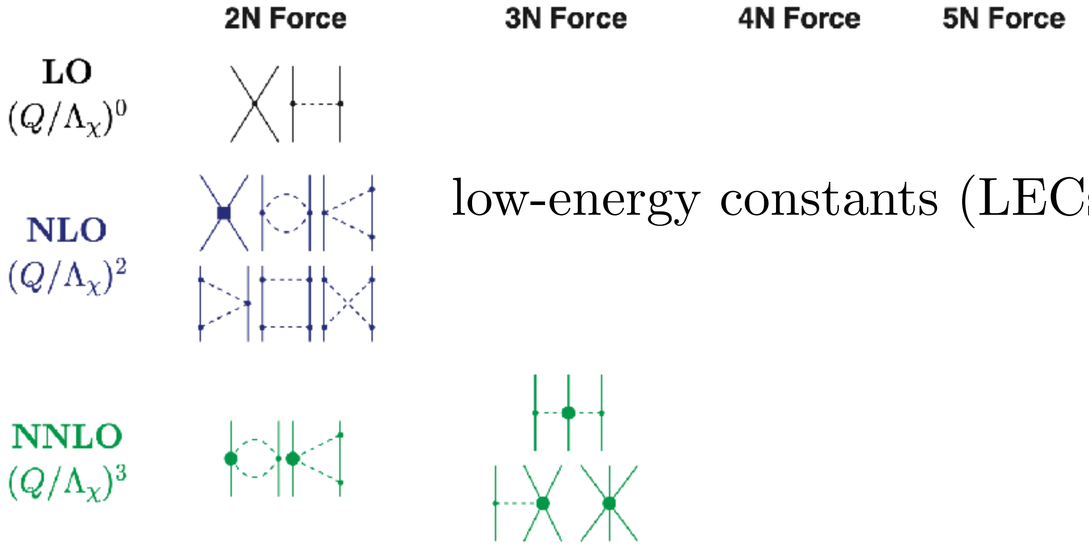
parameters

discrepancy

Posterior **Likelihood** **Prior**

$$p(\vec{\alpha}\,|\,\mathscr{D},M,I) = \frac{p(\mathscr{D}\,|\,\vec{\alpha},M,I)\cdot p(\vec{\alpha}\,|\,M,I)}{p(\mathscr{D}\,|\,M,I)}$$

Marginal likelihood

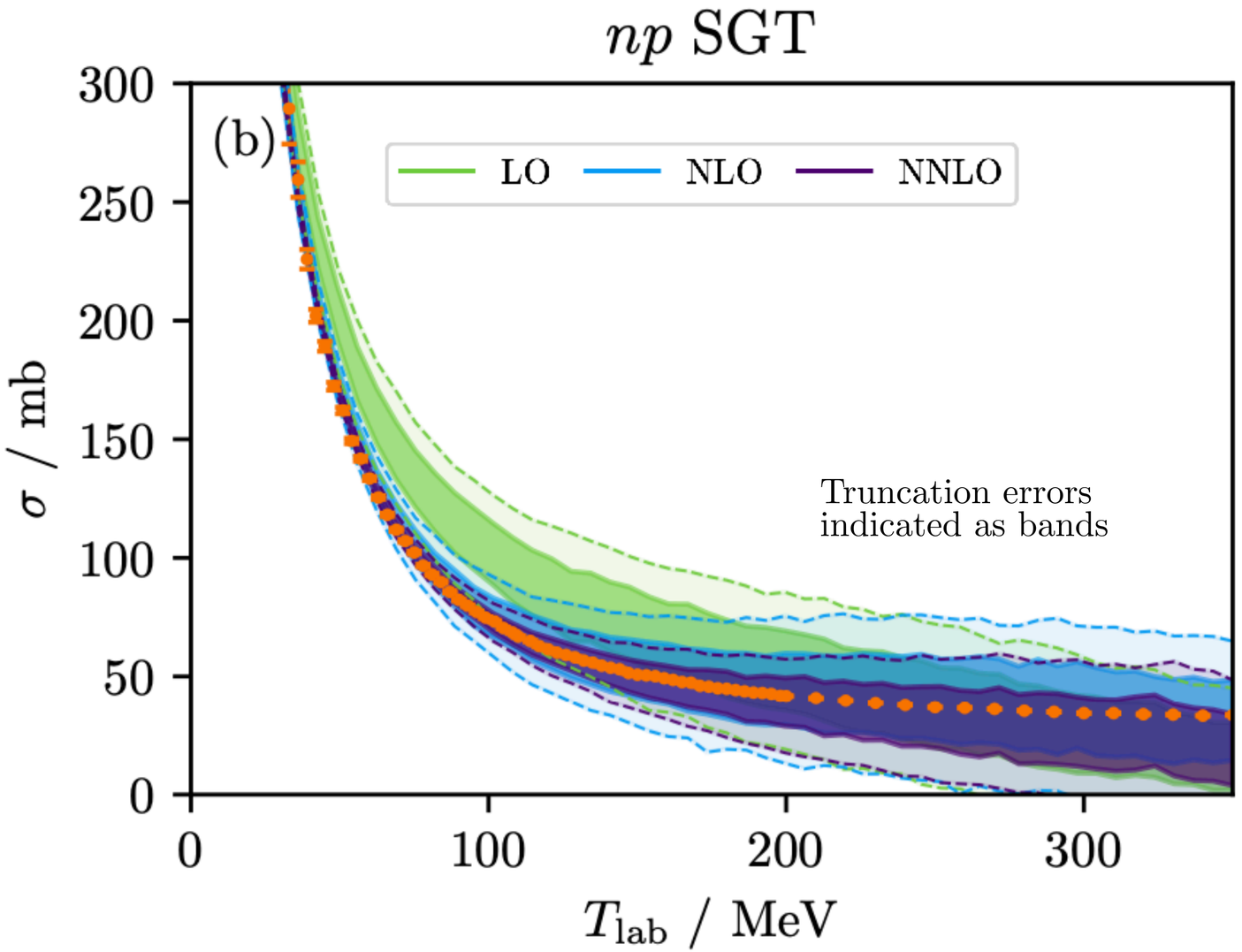


low-energy constants (LECs)

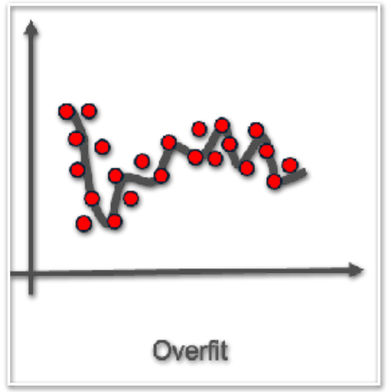
We need to:

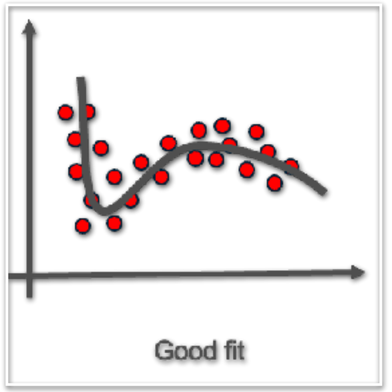
- Define our prior
- Evaluate the likelihood
- Sample the posterior using MCMC

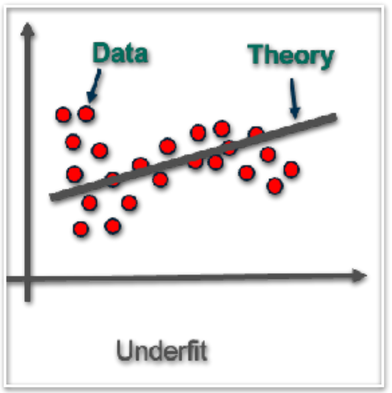
The EFT error estimate is critical in calibration



Svensson, Ekström, Forssen Phys Rev C (2022)







Maximizing the likelihood, i.e., minimizing a χ^2 could lead to overfitting. This approach would also inherit all the complications of a frequentist intepretation of probability...

With Bayes: regularisation via the truncation error and the priors in general

Formulating the likelihood

- the traditional calibration data is NN scattering cross sections

	Training data (D)			Validation data ($\tilde{D}$)		
Obs	np	pp	Total (%)	np	pp	Total (%)
SGT	315	0	315 (7.2)	84	0	84 (4.2)
SGTL	11	0	11 (0.3)	4	0	4 (0.2)
SGTT	16	0	16 (0.4)	3	0	3 (0.1)
DSG	1221	756	1977 (45.2)	457	159	616 (30.5)
A	5	47	52 (1.2)	0	24	24 (1.2)
AP	0	5	5 (0.1)	0	0	0 (0.0)
AT	30	0	30 (0.7)	35	0	35 (1.7)
AXX	0	143	143 (3.2)	0	120	120 (5.9)
AYY	64	151	215 (4.8)	46	151	197 (9.8)
AZX	0	137	137 (3.1)	0	120	120 (5.9)
AZZ	45	39	84 (1.9)	27	10	37 (1.8)
CKP	0	1	1 (0.0)	0	1	1 (0.0)
D	13	56	69 (1.6)	14	42	56 (2.8)
D0SK	8	0	8 (0.2)	14	0	14 (0.7)
DT	39	0	39 (0.9)	39	0	39 (1.9)
MSKN	0	8	8 (0.2)	0	8	8 (0.4)
MSSN	0	8	8 (0.2)	0	8	8 (0.4)
NNKK	8	0	8 (0.2)	0	0	0 (0.0)
NSKN	12	0	12 (0.3)	13	0	13 (0.6)
NSSN	4	0	4 (0.1)	14	0	14 (0.7)
P	0	489	489 (11.2)	0	260	260 (12.9)
PB	590	0	590 (13.5)	253	0	253 (12.5)
PT	38	0	38 (0.9)	19	0	19 (0.9)
R	5	50	55 (1.3)	0	50	50 (2.5)
RP	0	22	22 (0.5)	0	5	5 (0.2)
RPT	1	0	1 (0.0)	1	0	1 (0.0)
RT	29	0	29 (0.7)	37	0	37 (1.8)
All	2454	1912	4366 (100)	1060	958	2018 (100)

'Granada database'

$$p(\mathscr{D} \mid \vec{\alpha}, M, I) \propto \exp\left(-\frac{1}{2}\vec{r}^{T}(\Sigma_{\text{exp}} + \Sigma_{\text{th}})^{-1}\vec{r}\right) \quad \text{where } \vec{r} = y^{k}_{\text{th}}(\vec{\alpha}) - \mathscr{D}$$

Assuming truncation error at different scattering energies and angles are uncorrelated leads to a diagonal covariance matrix Σ_{th}

$$\begin{bmatrix} \Sigma & 0 & \dots & 0 & 0 \\ 0 & \Sigma & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & \dots & \Sigma & 0 \\ 0 & 0 & \dots & 0 & \Sigma \end{bmatrix}$$

requires solving S.E. for all values of $\vec{\alpha}$

$$[T + V(\vec{\alpha})]\,|\Psi\rangle = E\,|\Psi\rangle$$

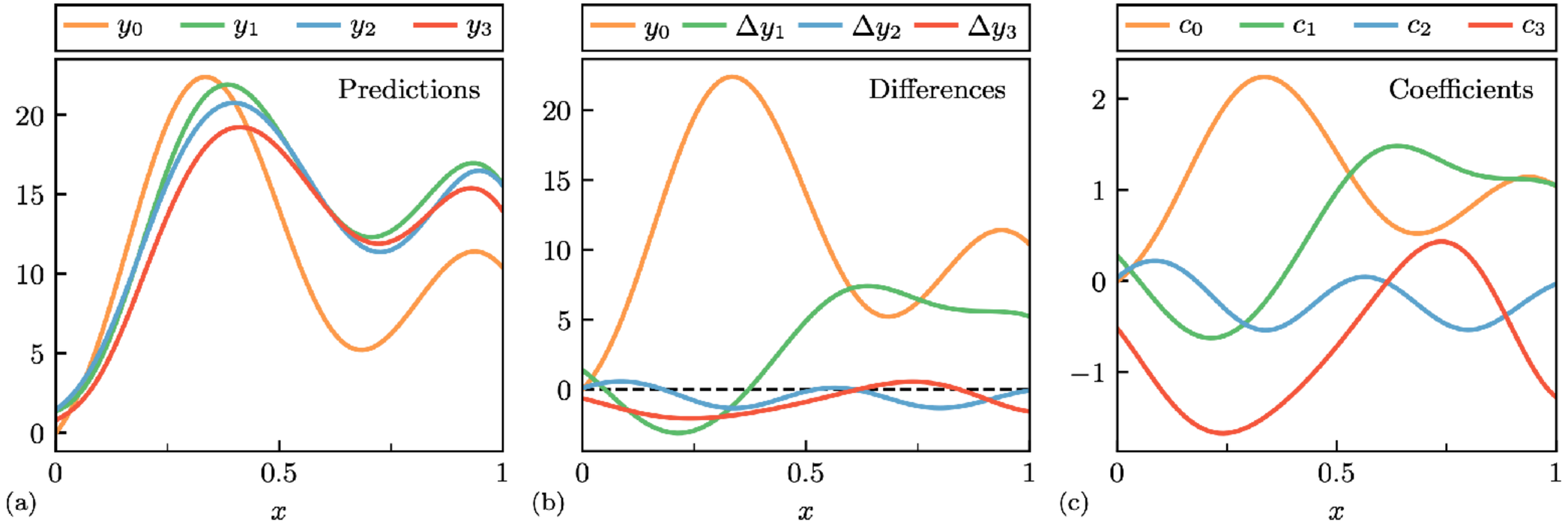
Not a computational problem for A=2

Svensson, Ekström, Forssen Phys Rev C (2022)

The model discrepancy in χEFT

$$\delta y_{\rm th}^{(k)} = y_{\rm ref} \sum_{\nu=k+1}^{\infty} c_\nu \left(\frac{Q}{\Lambda_\chi}\right)^\nu$$

- Correlated truncation errors



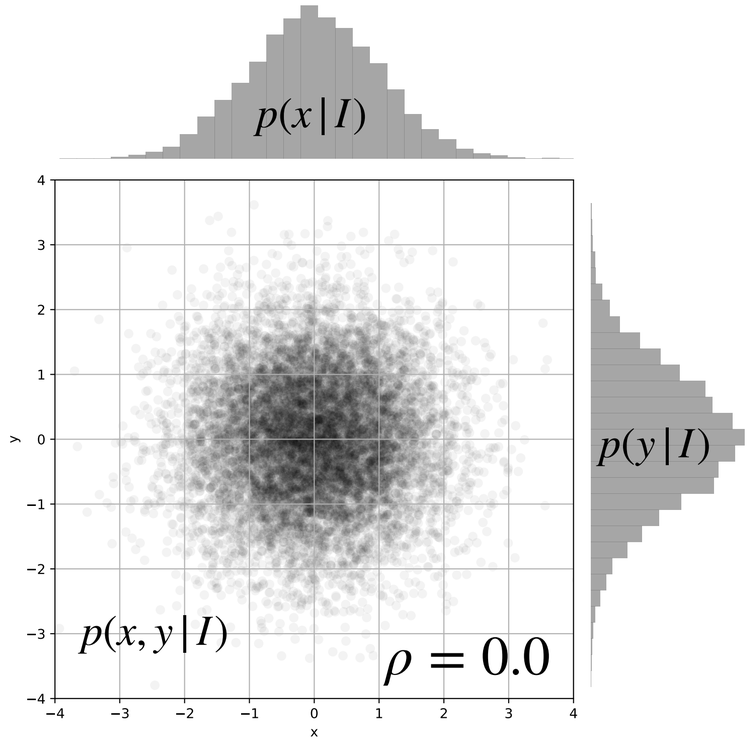
$$y_{\rm th}^{(k)} = y_{\rm ref} \sum_{\nu=0}^{k} c_\nu \left(\frac{Q}{\Lambda_\chi}\right)^\nu$$

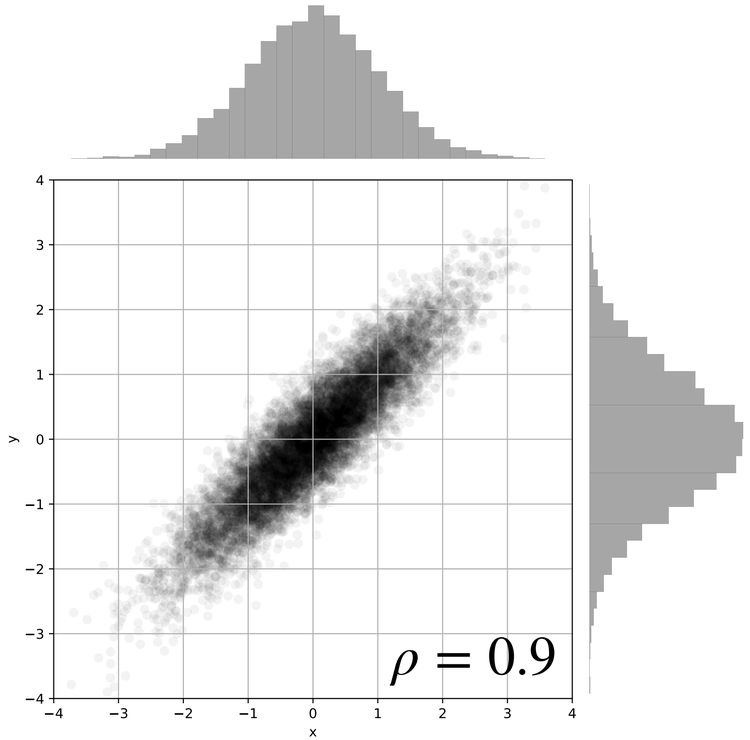
$$c_i = \frac{y_{\rm th}^{(i)} - y_{\rm th}^{(i-1)}}{y_{\rm ref}(Q/\Lambda_\chi)^i}$$

Melendez et al Phys Rev C (2019)

The model discrepancy in χEFT

- Correlated truncation errors



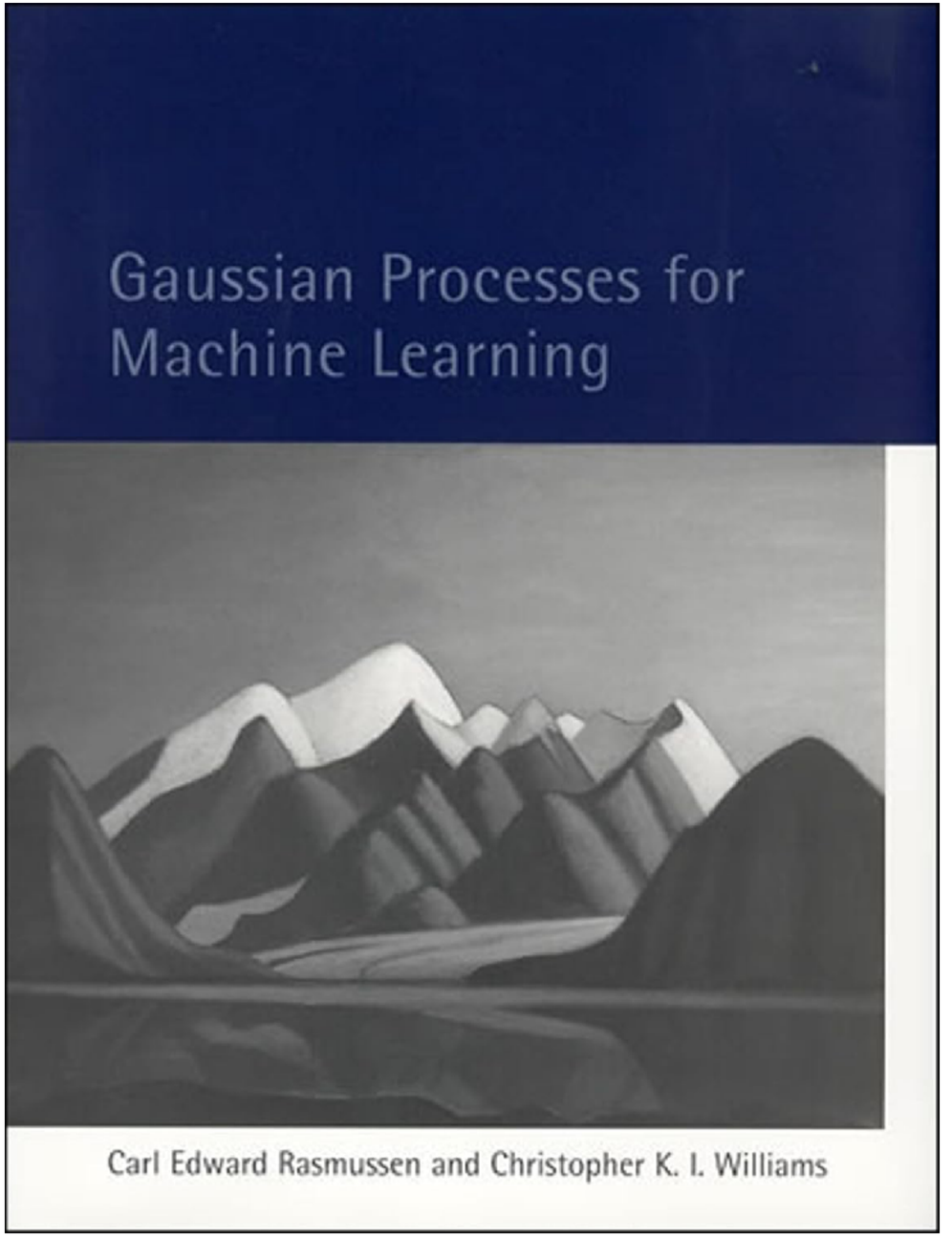


Bivariate normal distribution $x, y \sim \mathcal{N}(\vec{\mu}, \Sigma)$ with mean vector $\vec{\mu} = [0,0]^T$ and covariance matrix

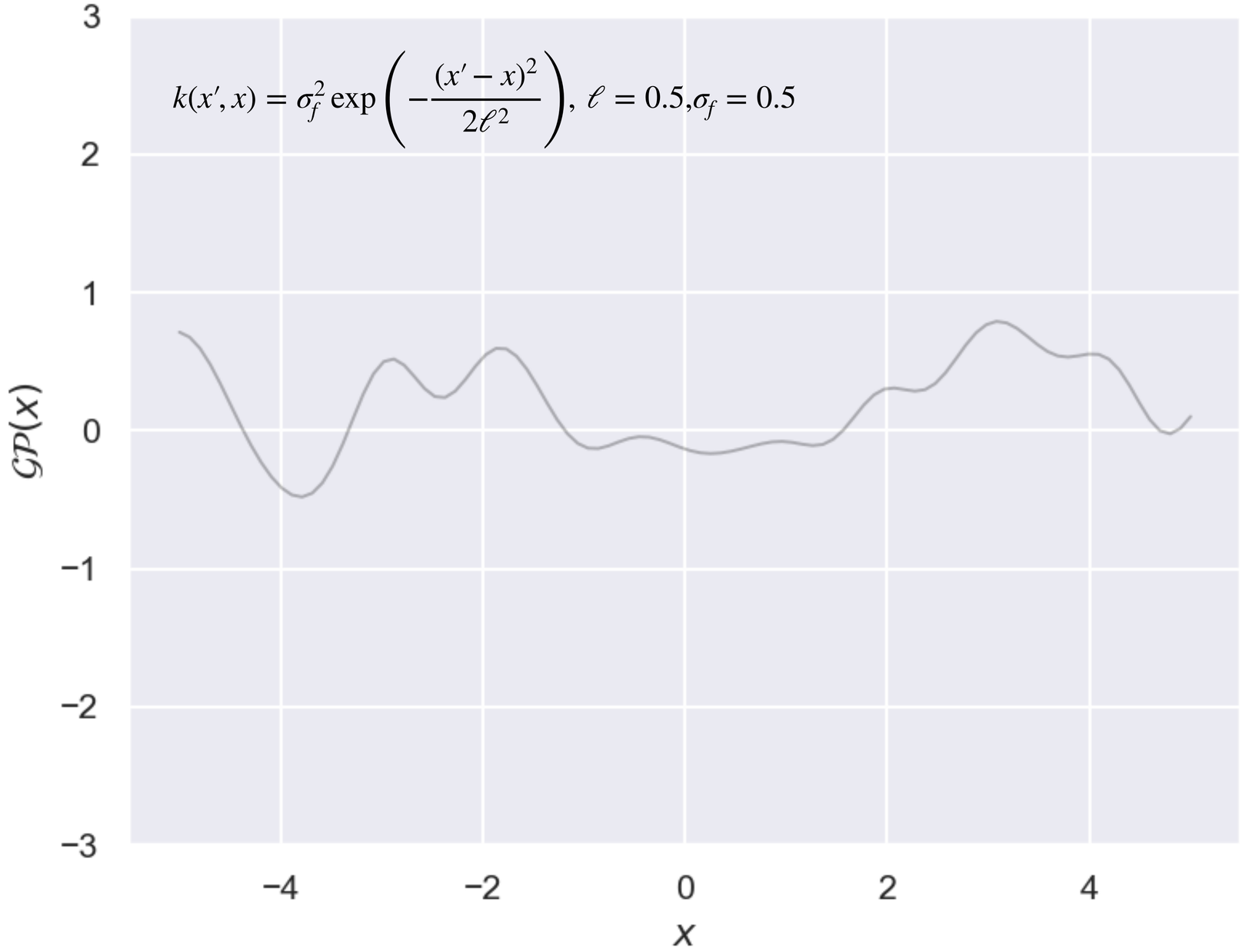
$$\Sigma = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}$$

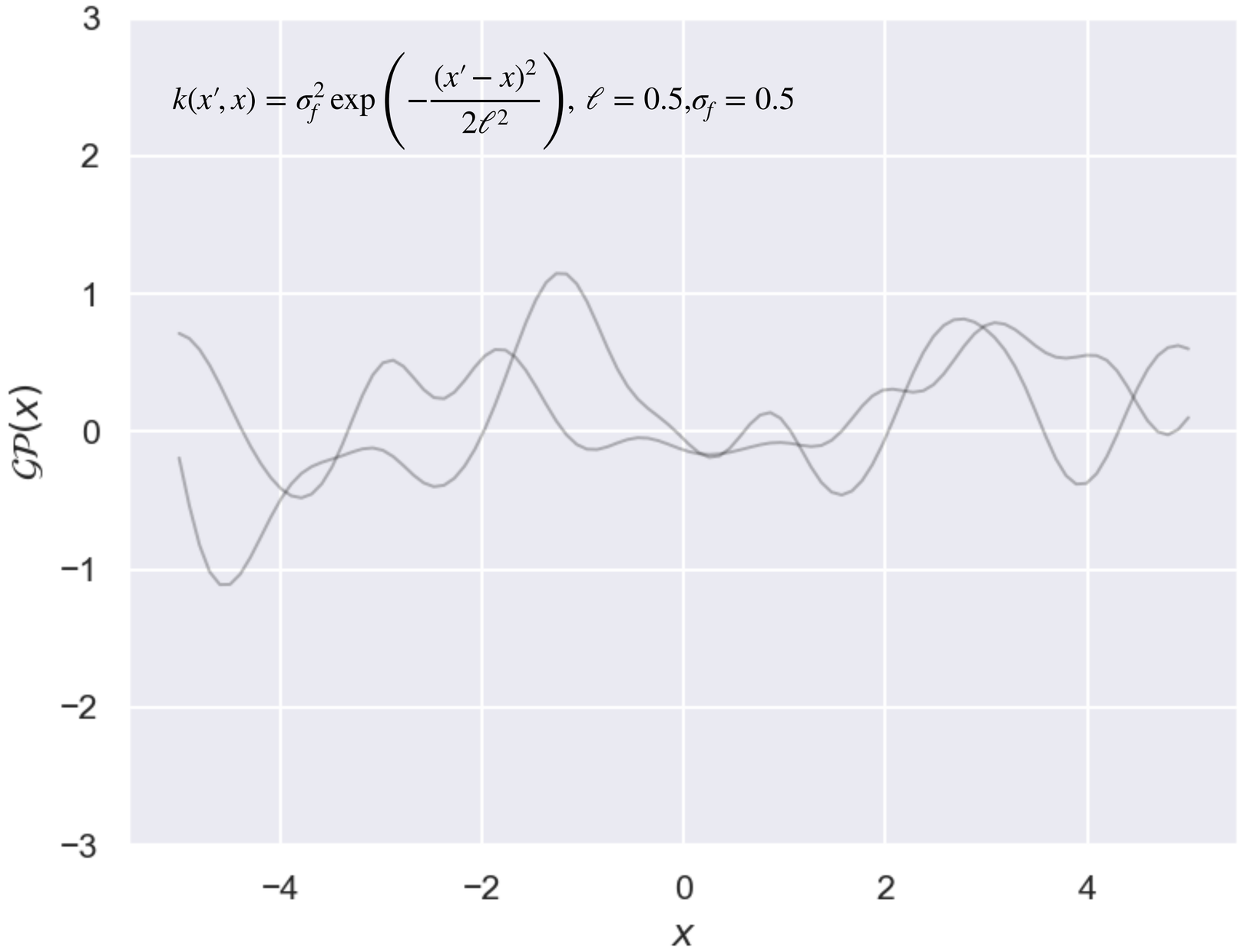
For two random variables, x and y, we have $-1 \leq \rho \leq +1$. For more than two variables, all pairs have a correlation structure $\rho_{ij} = \rho_{ji}$.

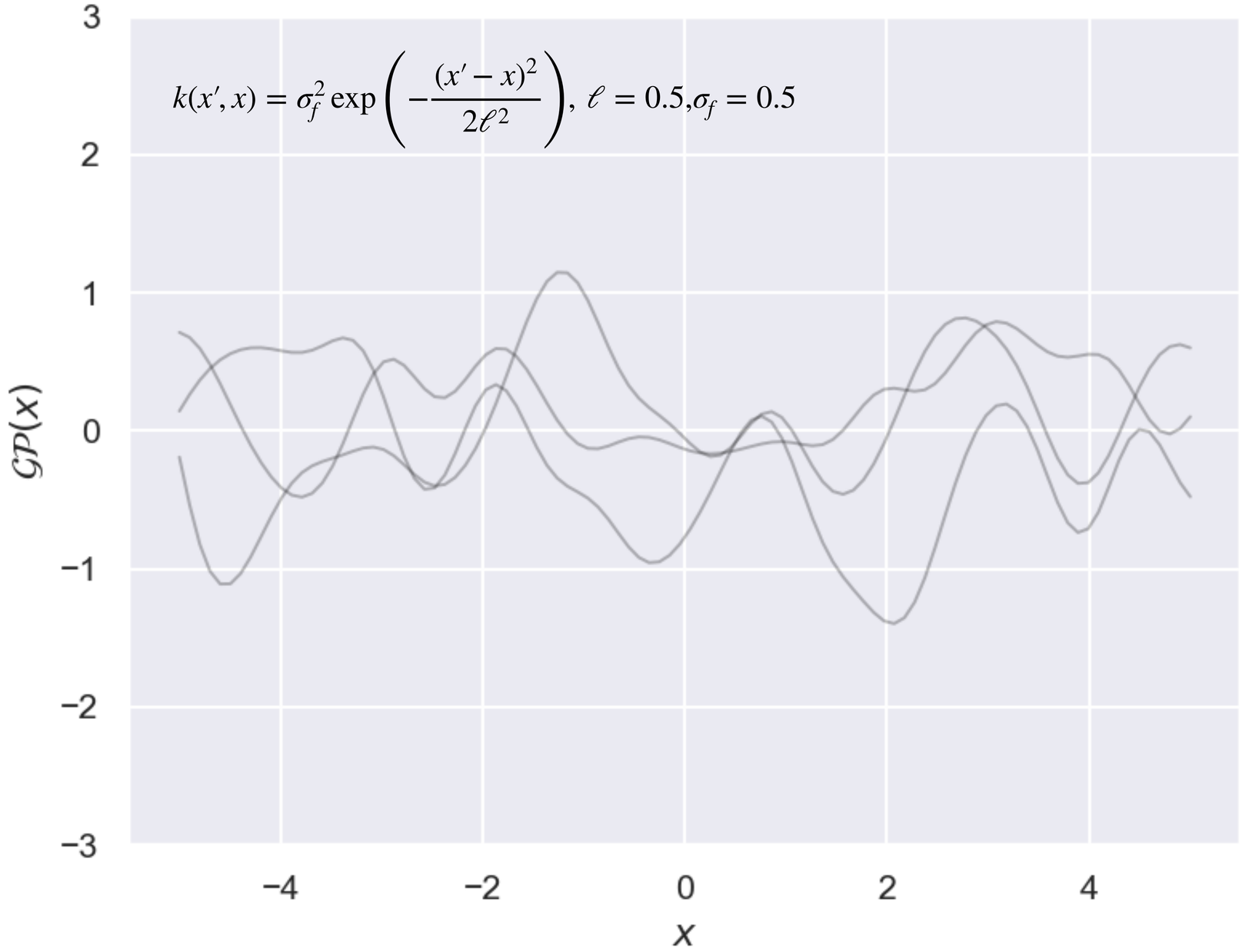
A Gaussian process: a collection of random variables, any finite number of which have a joint Gaussian distribution. A kernel function ***k(x',x)*** defines the covariance ('smoothness'/'correlation') between points x' and x: $(f(x_1), f(x_2), f(x_3), \ldots, f(x_N)) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{k})$. Gaussians are closed under addition which enables closed form conditionals of Gaussian processes.

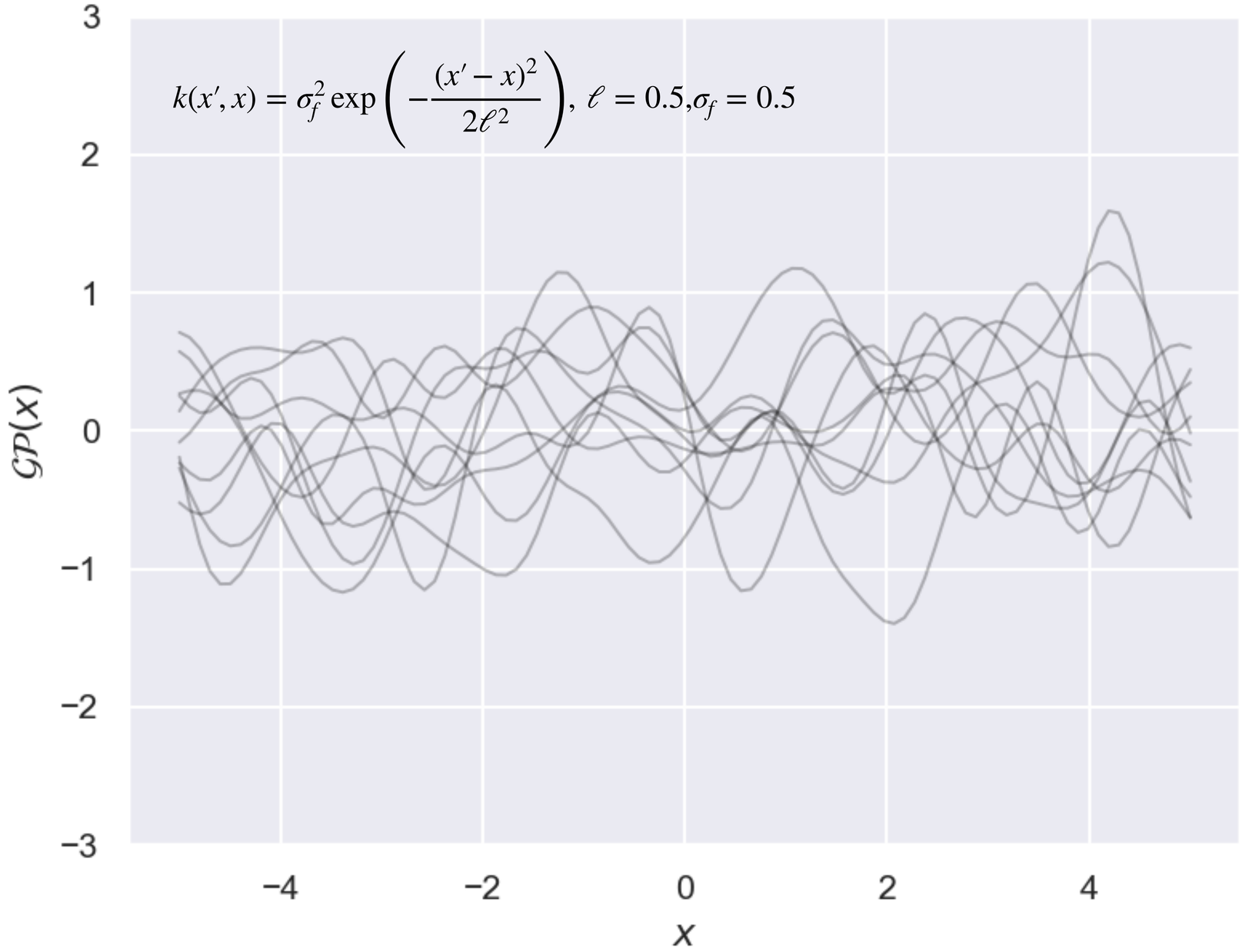


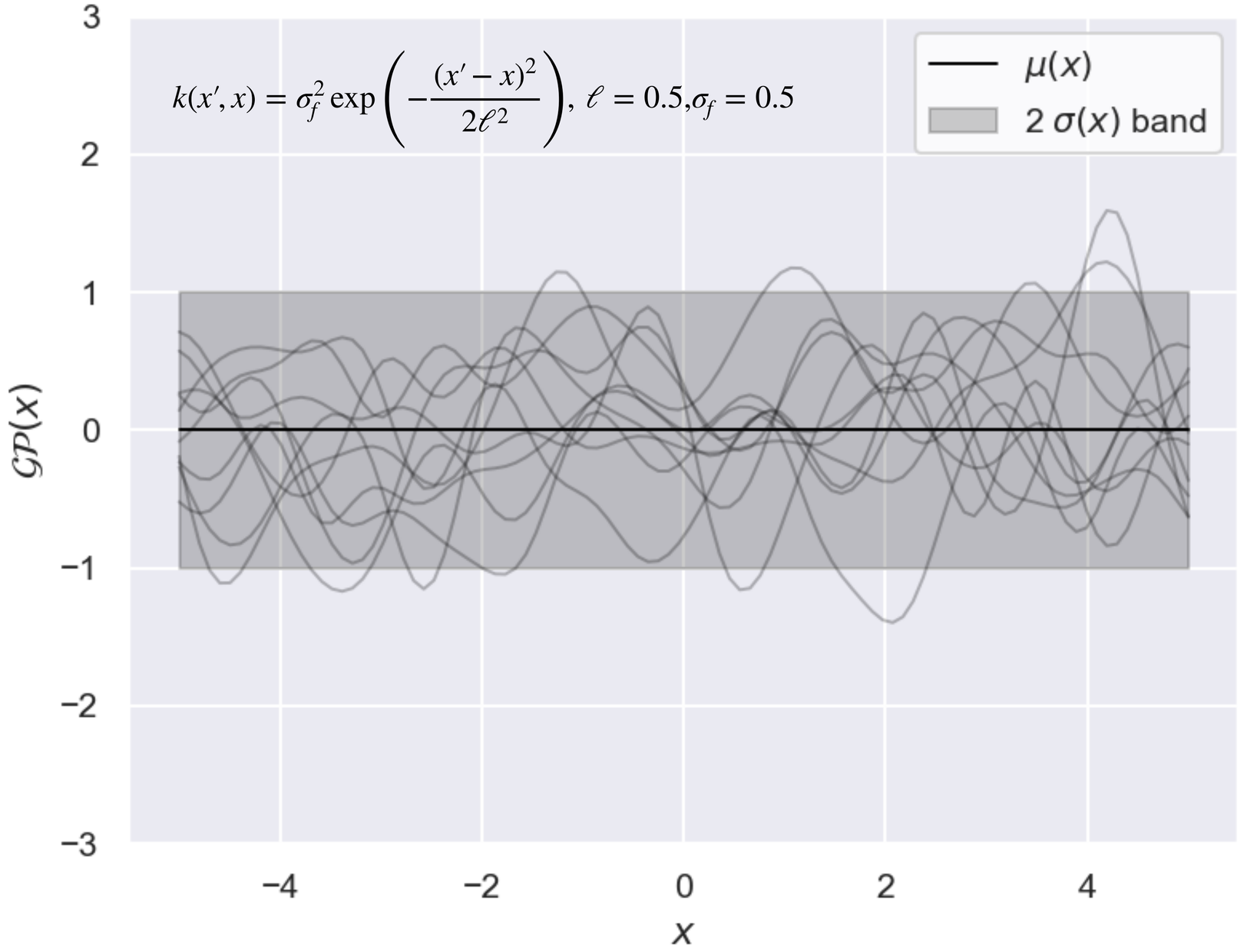
https://gaussianprocess.org/gpml/

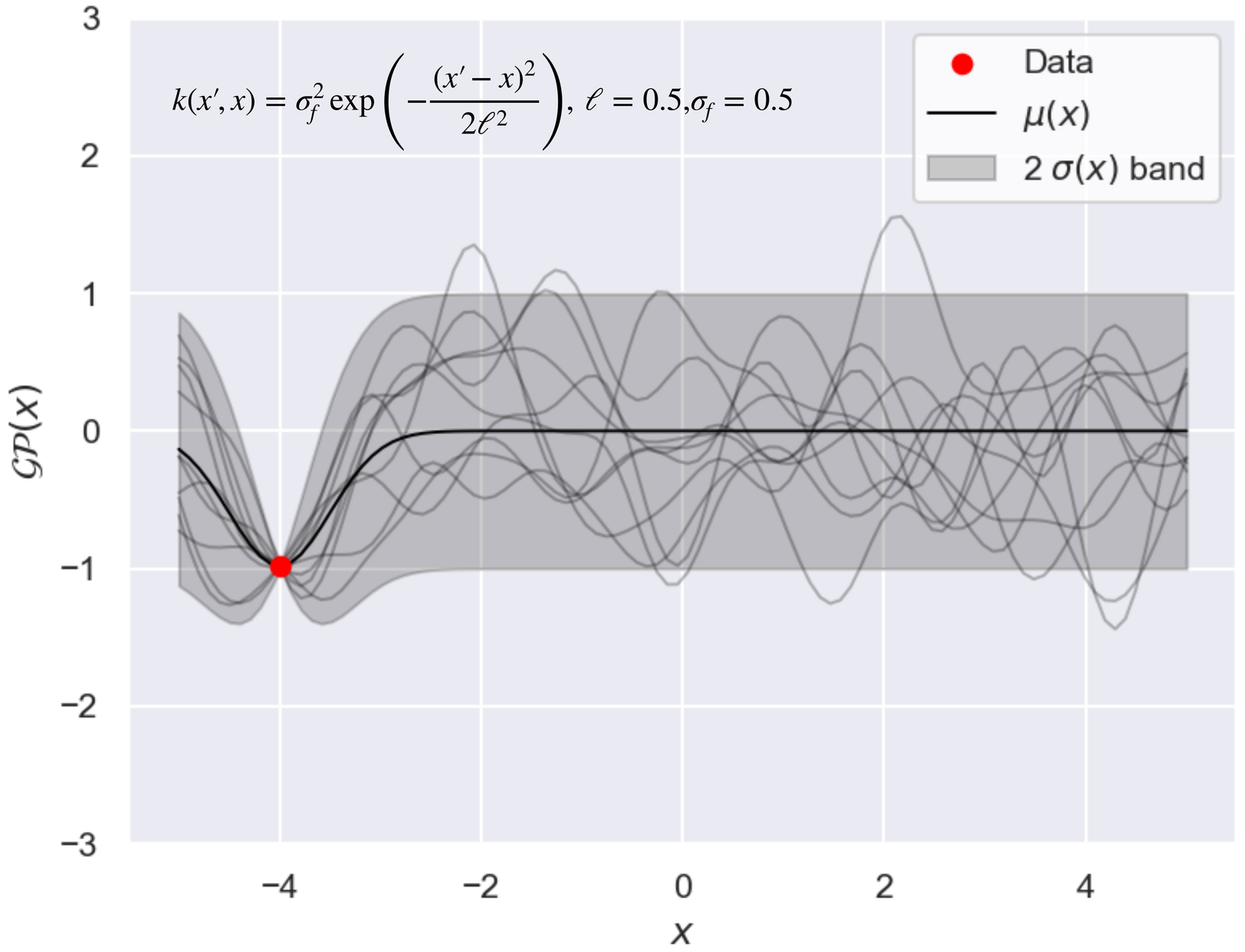

$k(x', x) = \sigma_f^2 \exp\left(-\frac{(x' - x)^2}{2\ell^2}\right),\ \ell = 0.5, \sigma_f = 0.5$
3
2
1
0
−1
−2
−3
$\mathcal{GP}(x)$
−4
−2
0
2
4
x

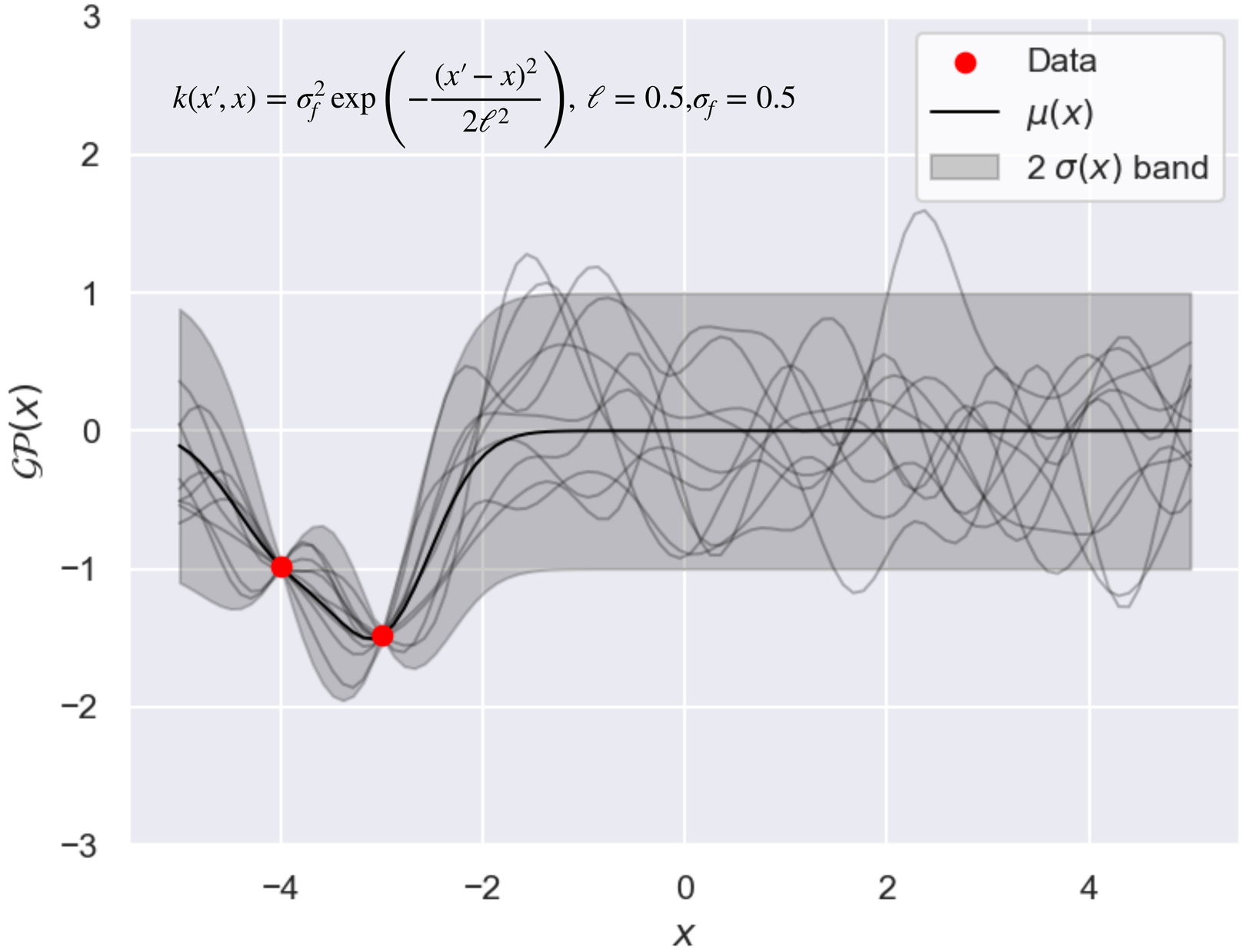

$k(x', x) = \sigma_f^2 \exp\left(-\frac{(x' - x)^2}{2\ell^2}\right),\ \ell = 0.5, \sigma_f = 0.5$
3
2
1
0
−1
−2
−3
$\mathcal{GP}(x)$
−4
−2
0
2
4
x

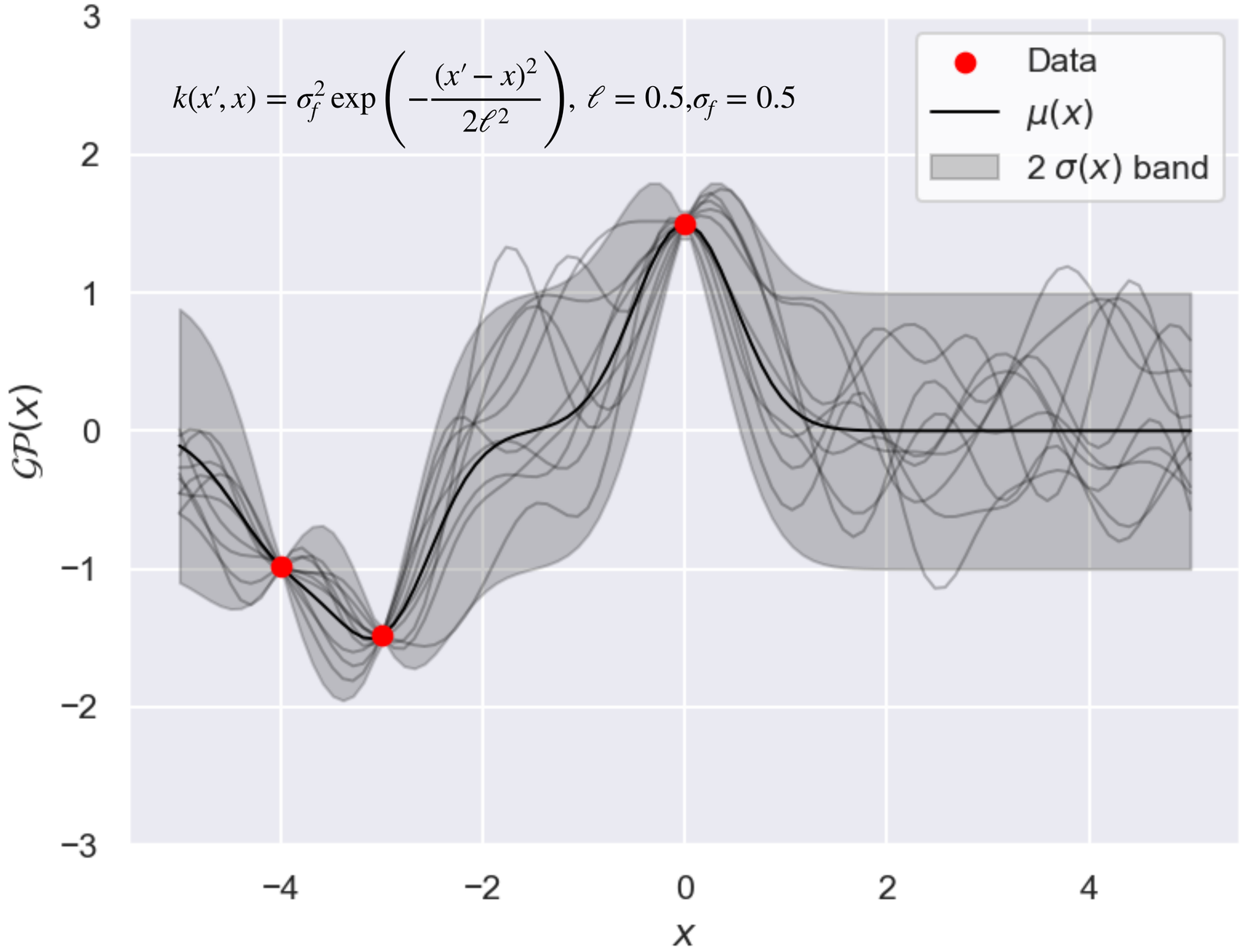

$k(x', x) = \sigma_f^2 \exp\left(-\frac{(x' - x)^2}{2\ell^2}\right),\ \ell = 0.5, \sigma_f = 0.5$
3
2
1
0
−1
−2
−3
$\mathcal{GP}(x)$
−4
−2
0
2
4
x

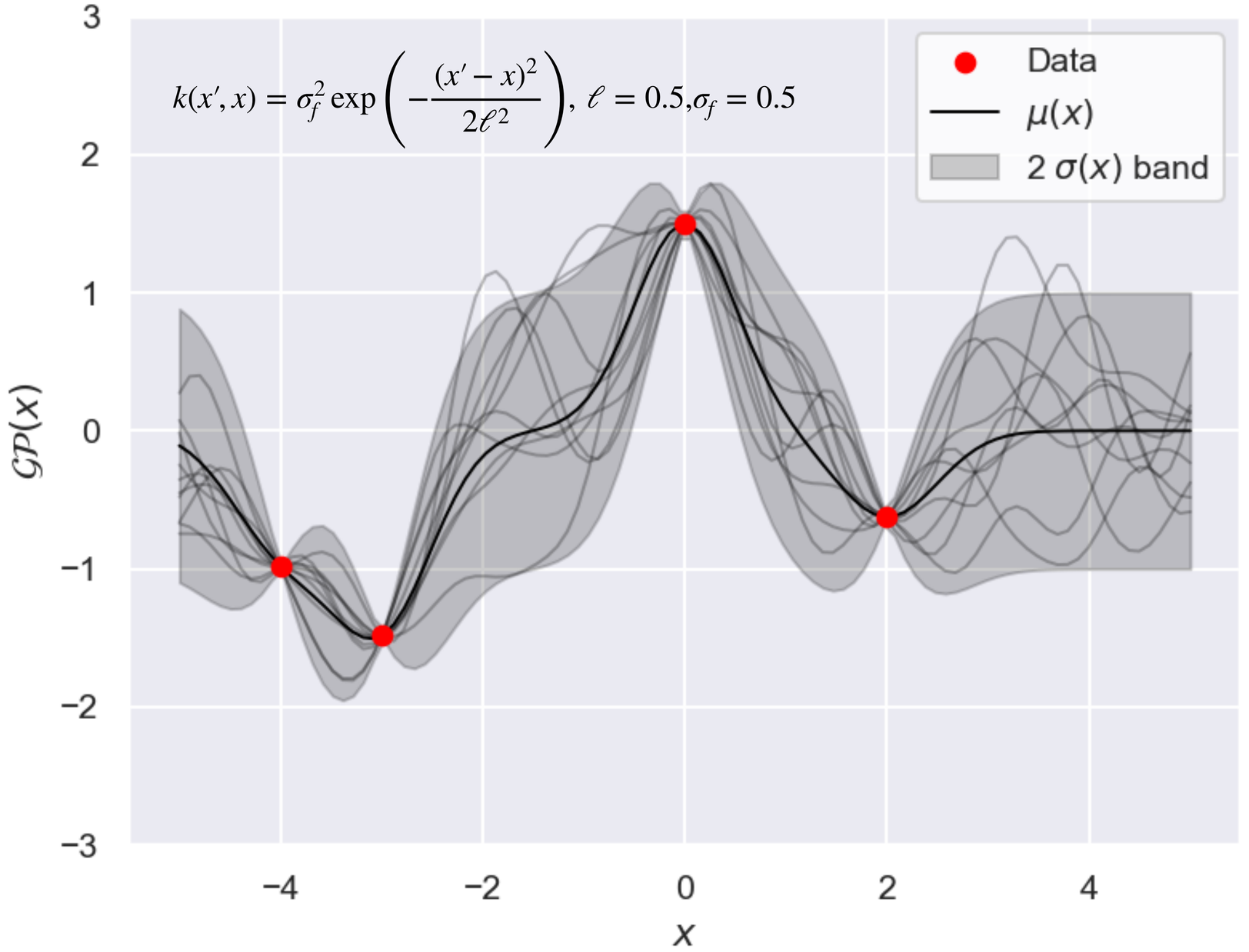

$k(x', x) = \sigma_f^2 \exp\left(-\frac{(x' - x)^2}{2\ell^2}\right),\ \ell = 0.5, \sigma_f = 0.5$
3
2
1
0
−1
−2
−3
$\mathcal{GP}(x)$
−4
−2
0
2
4
x

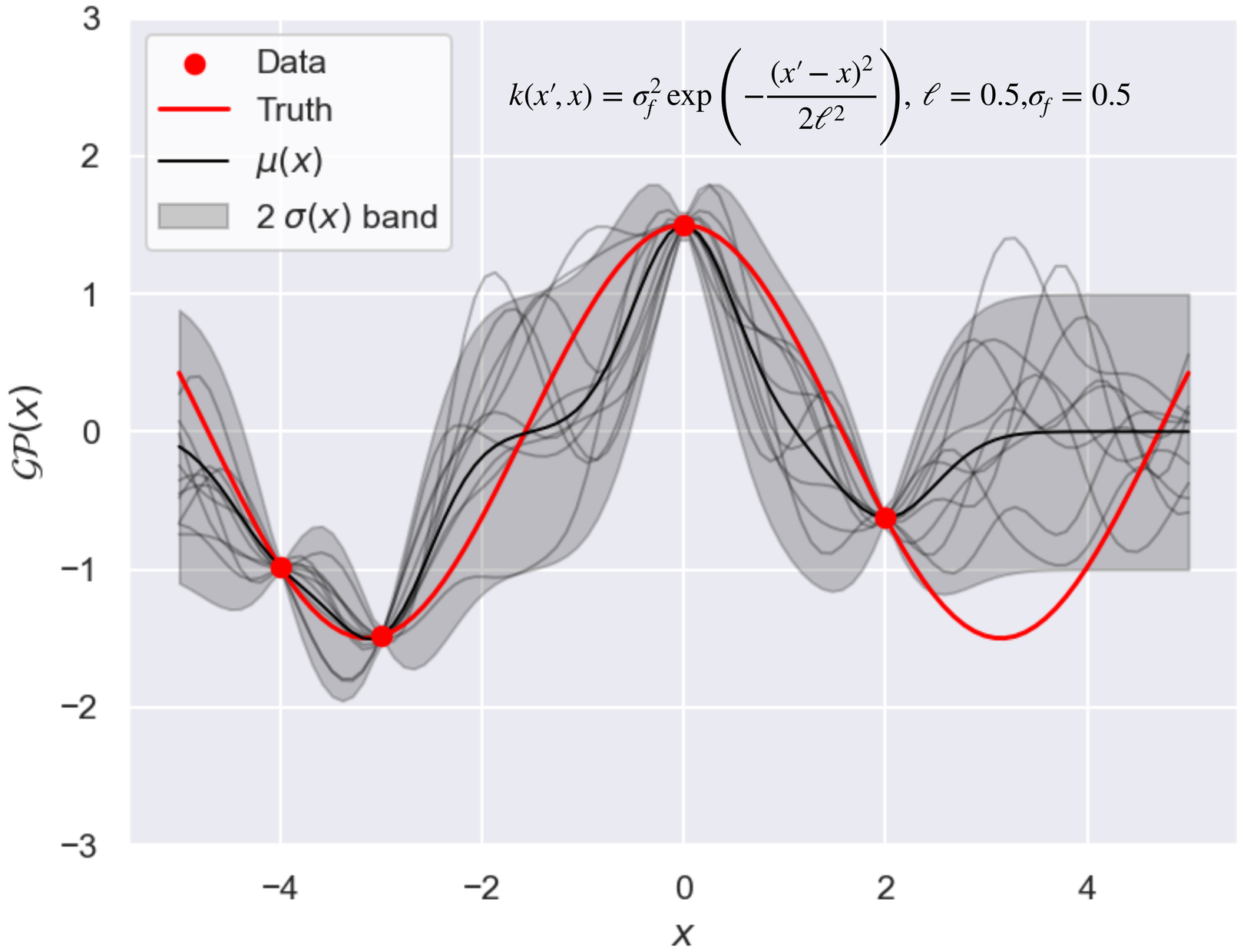

$k(x', x) = \sigma_f^2 \exp\left(-\frac{(x'-x)^2}{2\ell^2}\right),\ \ell = 0.5, \sigma_f = 0.5$
$\mu(x)$
2 $\sigma(x)$ band
$\mathcal{GP}(x)$
x
3
2
1
0
−1
−2
−3
−4
−2
0
2
4


$k(x', x) = \sigma_f^2 \exp\left(-\frac{(x'-x)^2}{2\ell^2}\right),\ \ell = 0.5, \sigma_f = 0.5$
Data
$\mu(x)$
2 $\sigma(x)$ band
$\mathcal{GP}(x)$
x
3
2
1
0
−1
−2
−3
−4
−2
0
2
4

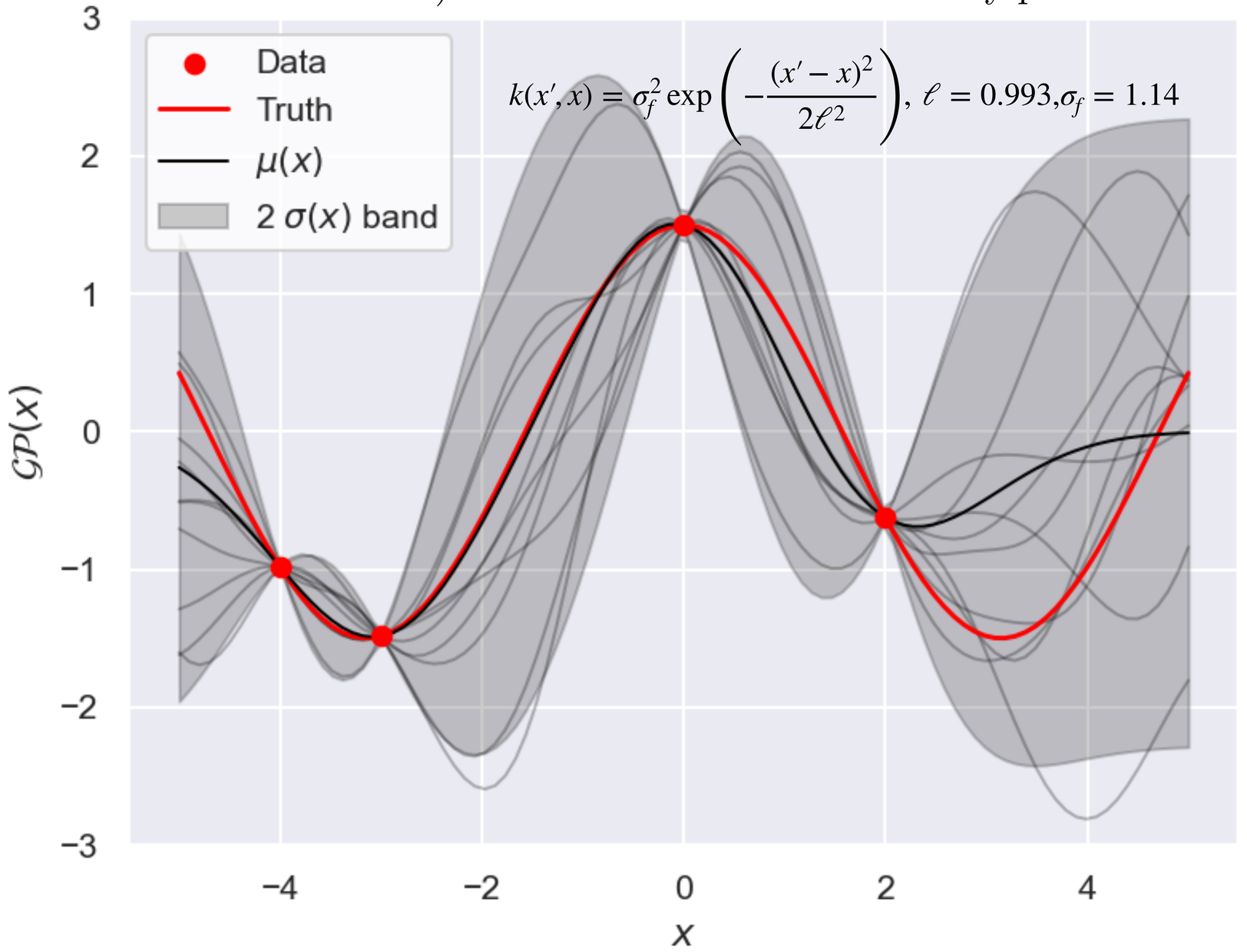

$k(x', x) = \sigma_f^2 \exp\left(-\frac{(x'-x)^2}{2\ell^2}\right),\ \ell = 0.5, \sigma_f = 0.5$
Data
$\mu(x)$
2 $\sigma(x)$ band
$\mathcal{GP}(x)$
x
3
2
1
0
−1
−2
−3
−4
−2
0
2
4

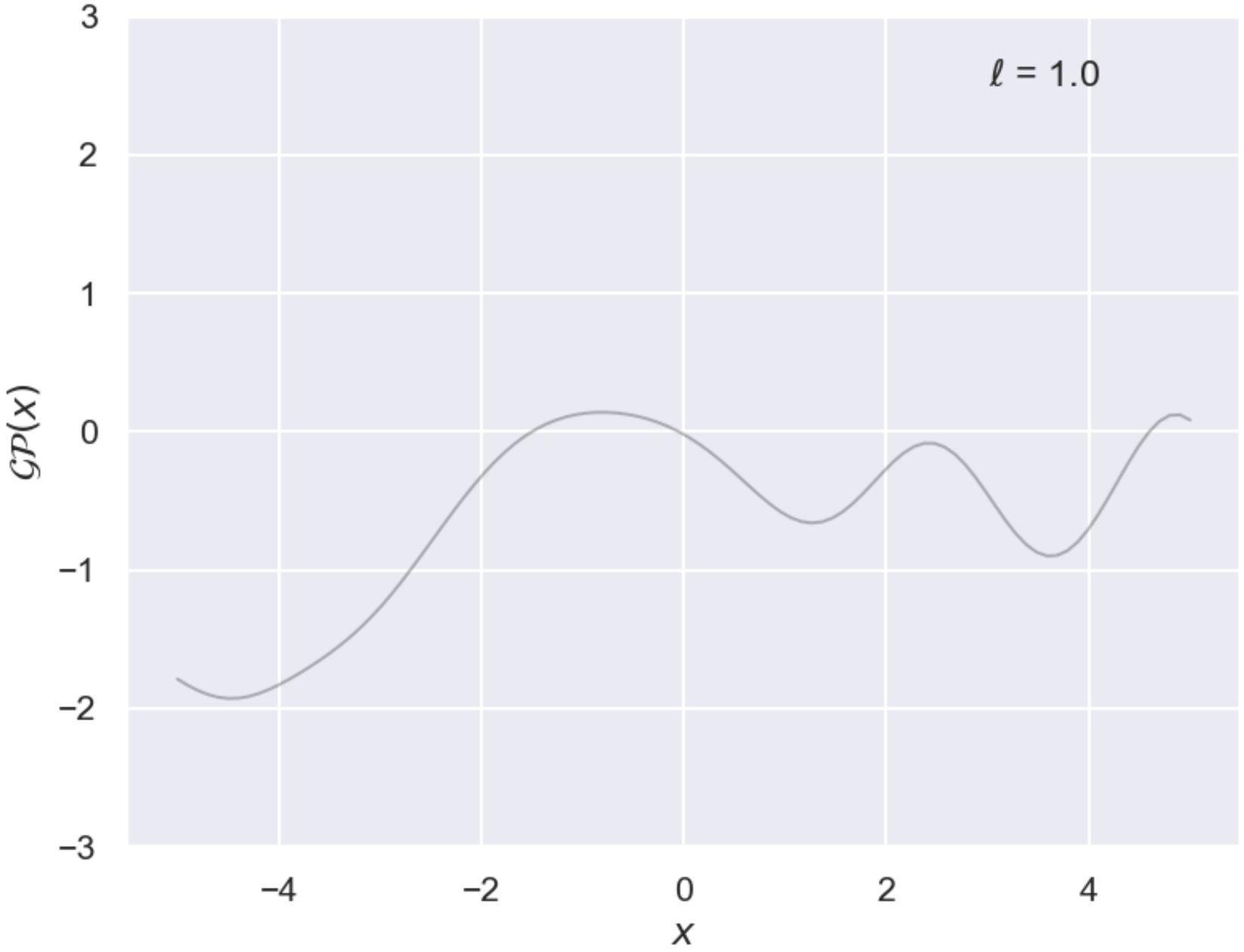

$k(x', x) = \sigma_f^2 \exp\left(-\frac{(x'-x)^2}{2\ell^2}\right),\ \ell = 0.5, \sigma_f = 0.5$
Data
$\mu(x)$
2 $\sigma(x)$ band
$\mathcal{GP}(x)$
x
3
2
1
0
−1
−2
−3
−4
−2
0
2
4

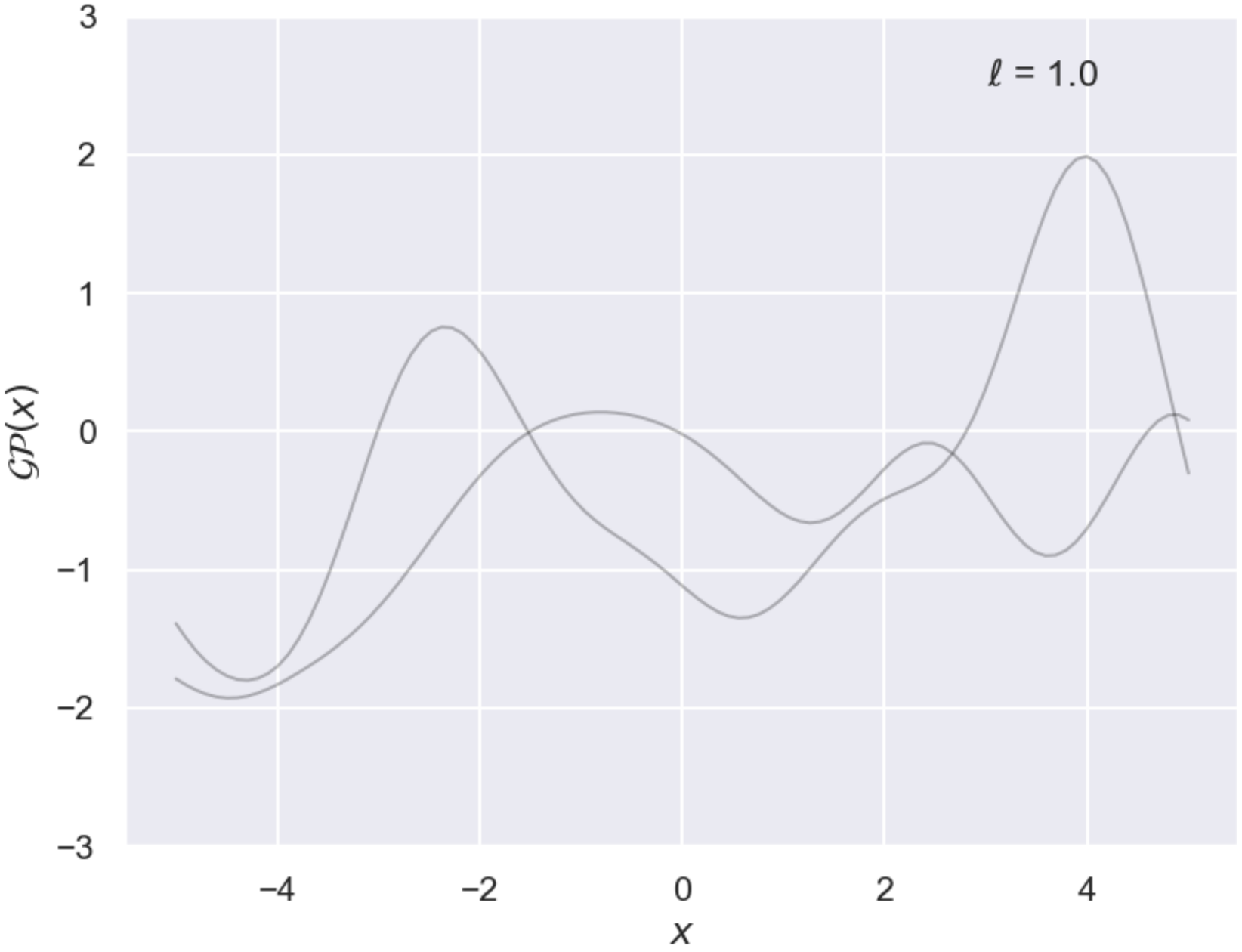

$k(x', x) = \sigma_f^2 \exp\left(-\frac{(x'-x)^2}{2\ell^2}\right),\ \ell = 0.5, \sigma_f = 0.5$
Data
$\mu(x)$
2 $\sigma(x)$ band
$\mathcal{GP}(x)$
x
3
2
1
0
−1
−2
−3
−4
−2
0
2
4

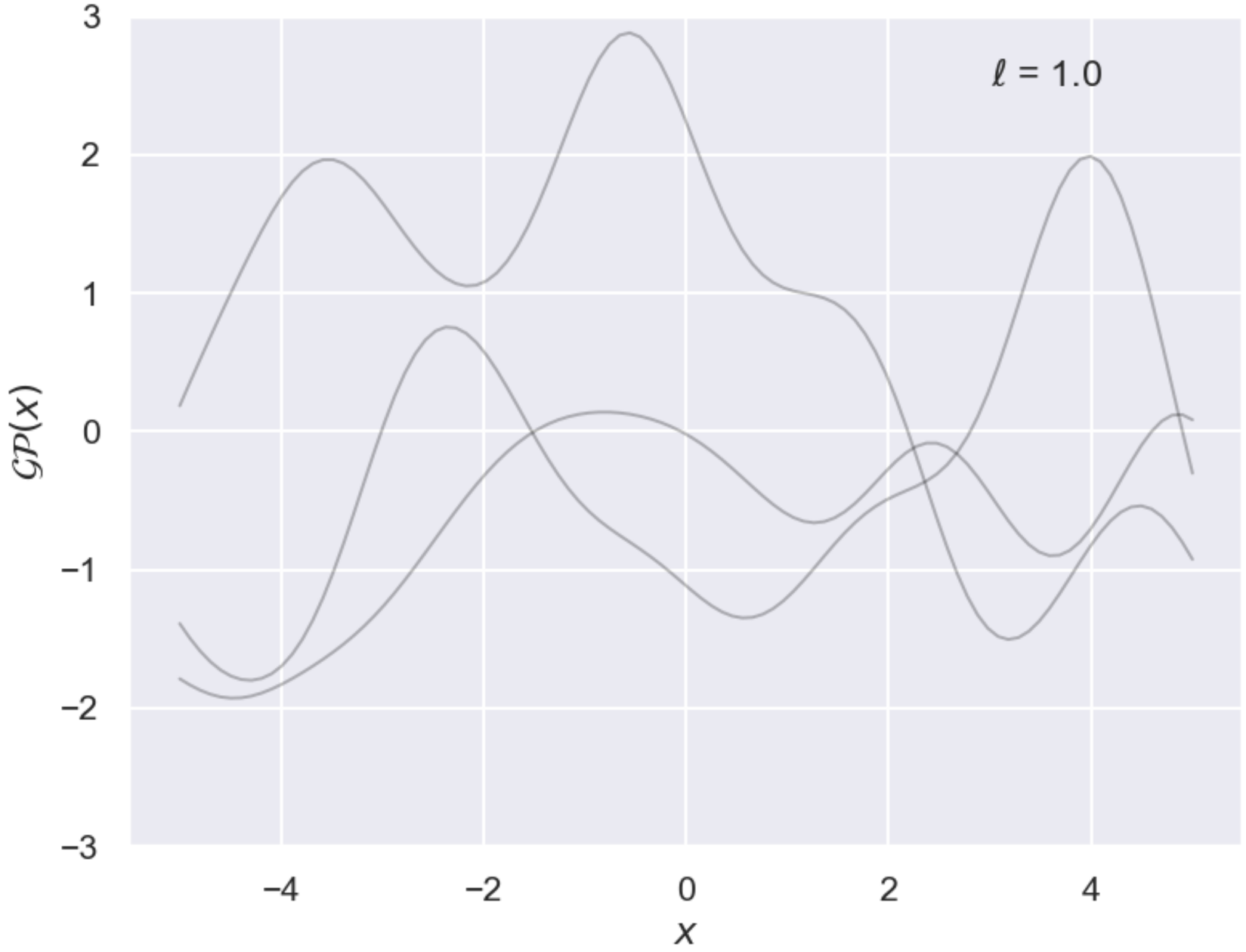

Data
Truth
$\mu(x)$
2 $\sigma(x)$ band
$k(x', x) = \sigma_f^2 \exp\left(-\frac{(x'-x)^2}{2\ell^2}\right)$, $\ell = 0.5, \sigma_f = 0.5$
$\mathcal{GP}(x)$
x
3
2
1
0
−1
−2
−3
−4
−2
0
2
4

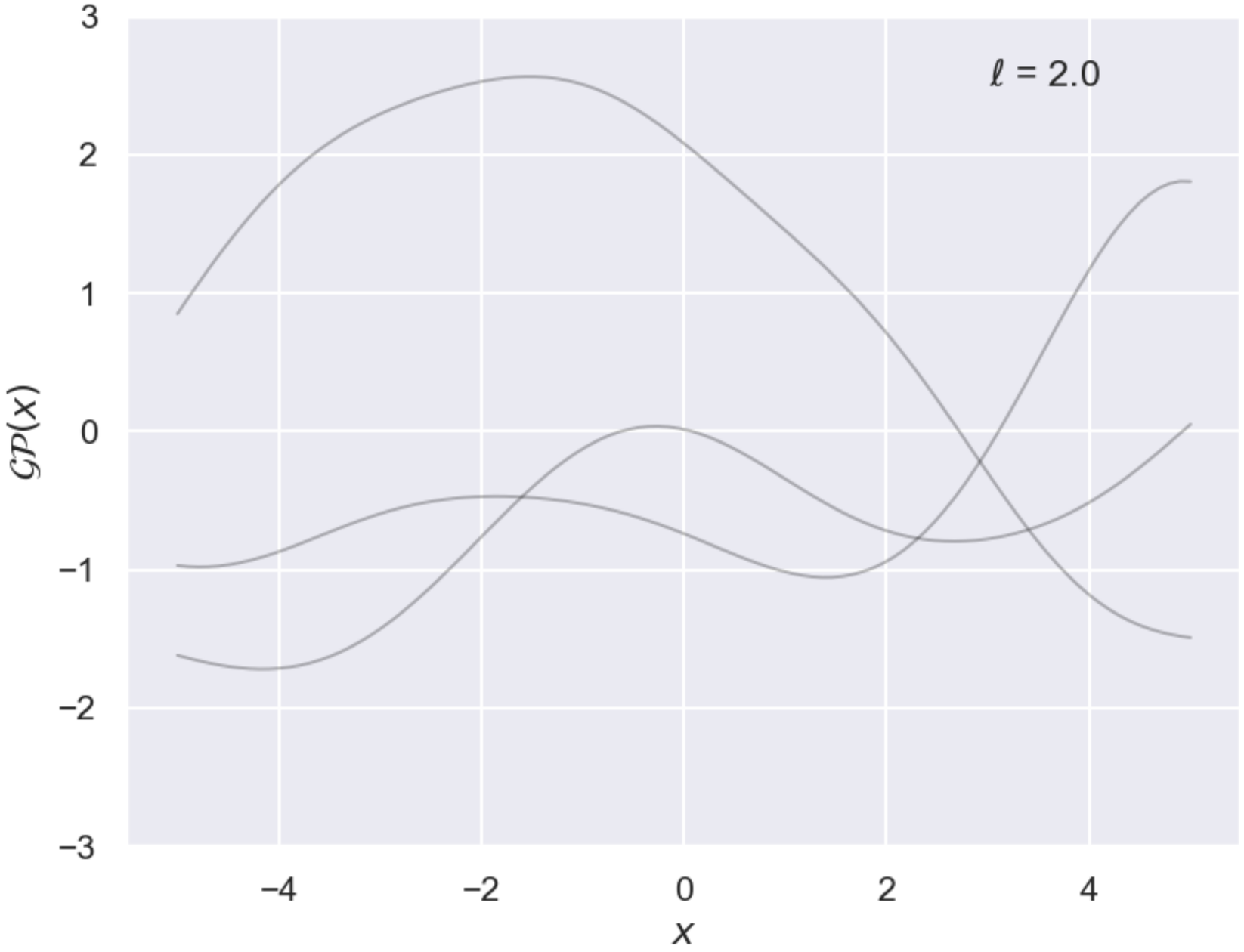
We can learn (MLE or MCMC inference) the hyperparameters (length scale and variance) of the GP conditioned on theory predictions

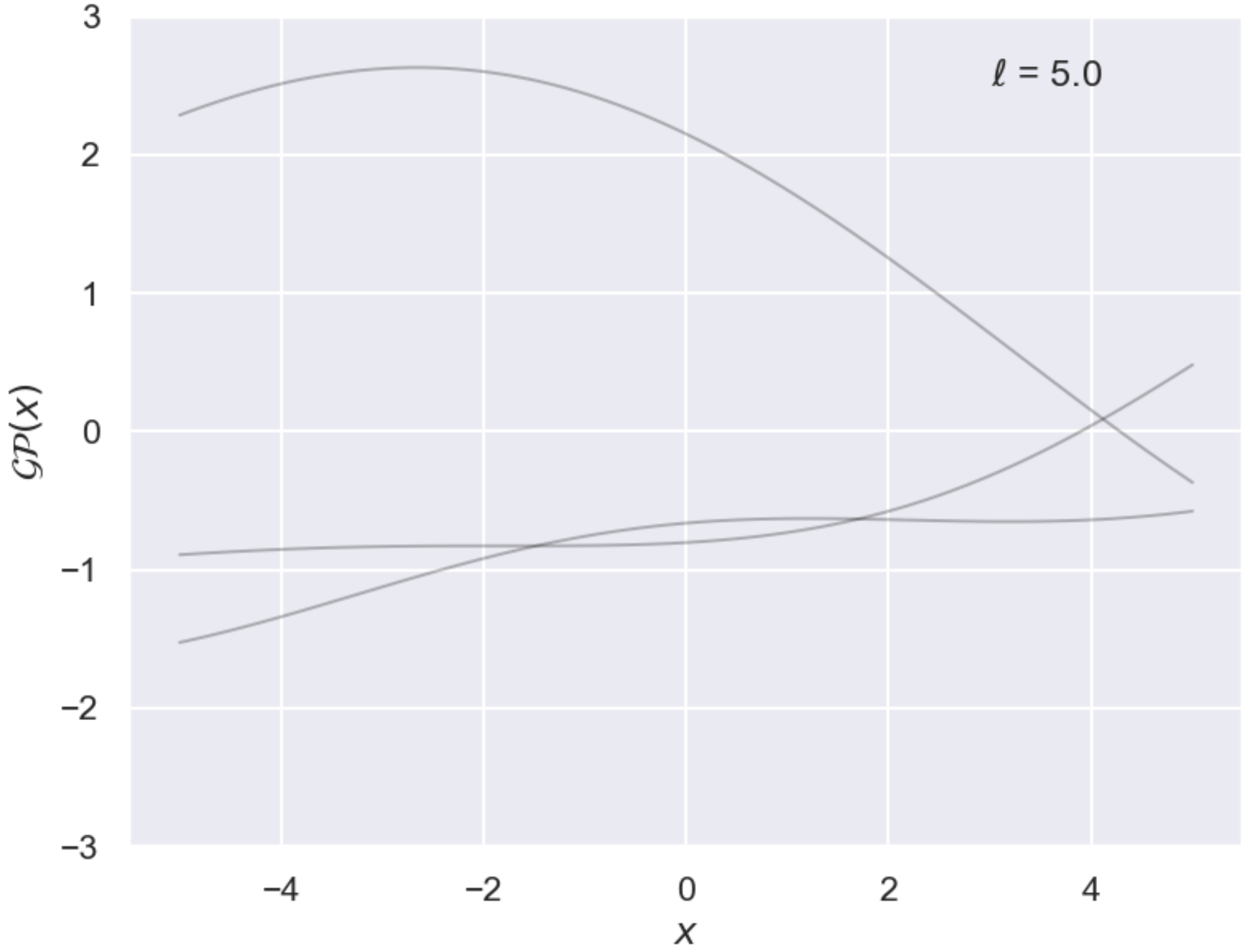


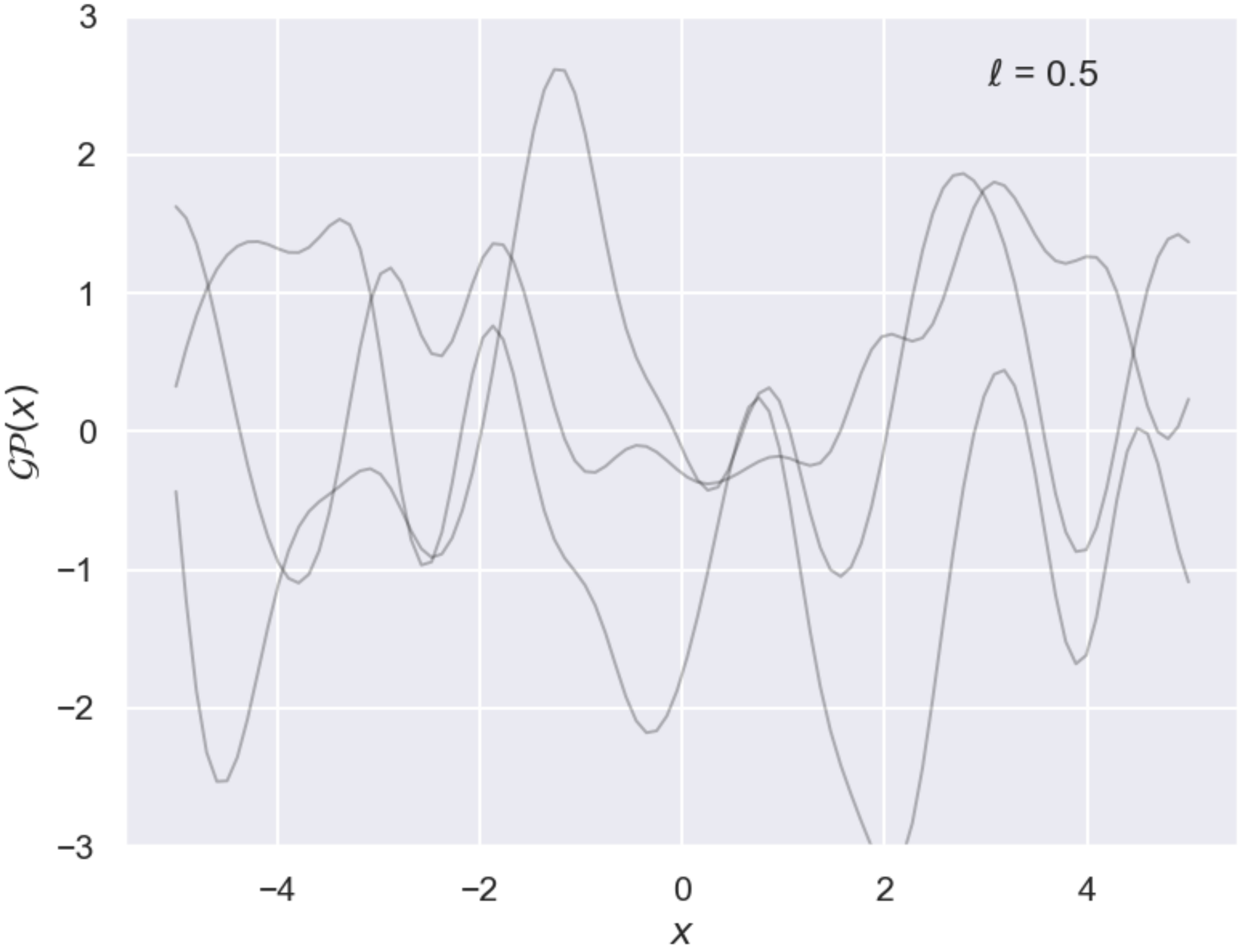

ℓ = 1.0
$\mathcal{GP}(x)$
3
2
1
0
−1
−2
−3
−4
−2
0
2
4
x

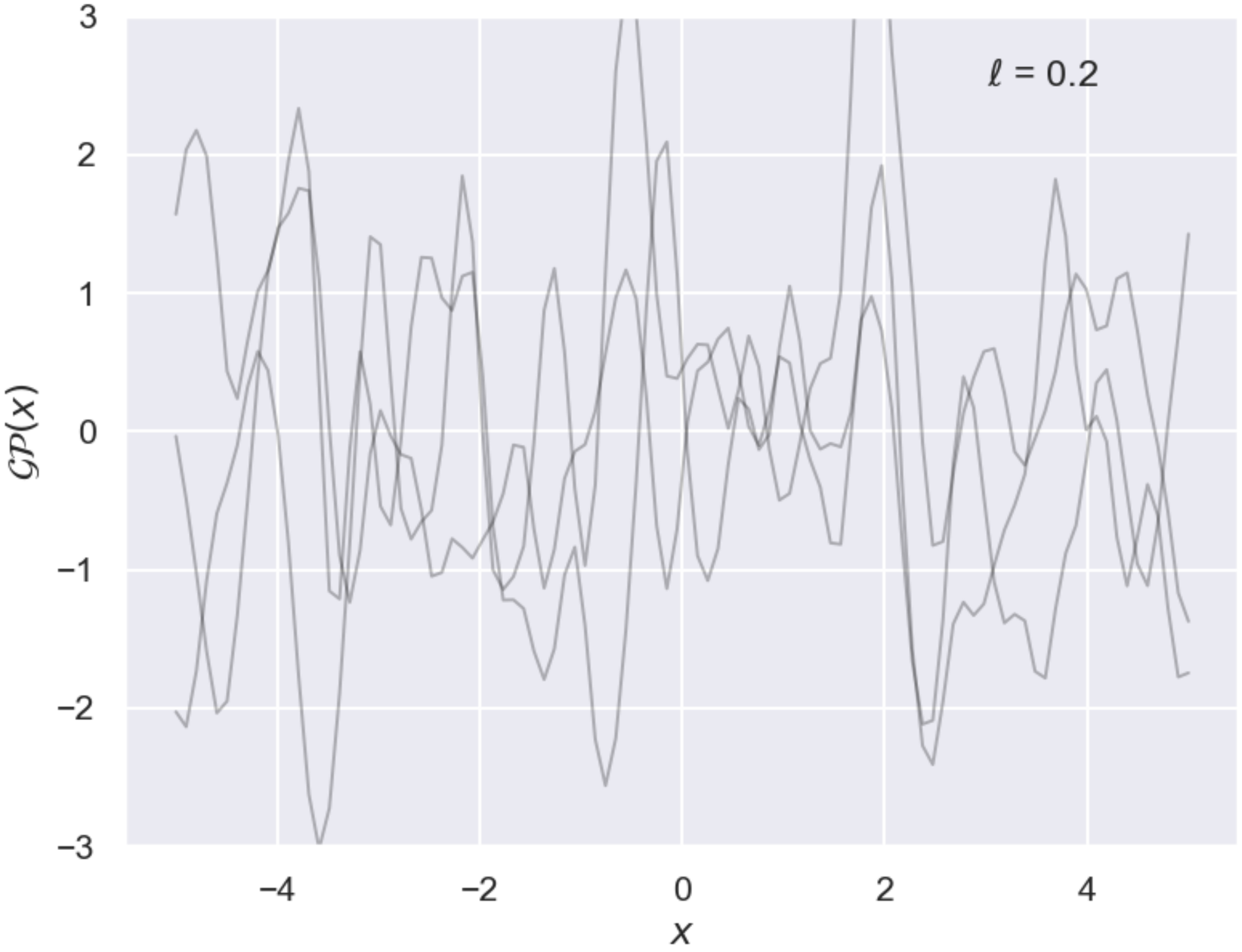

ℓ = 1.0
3
2
1
0
−1
−2
−3
$\mathcal{GP}(x)$
−4
−2
0
2
4
x

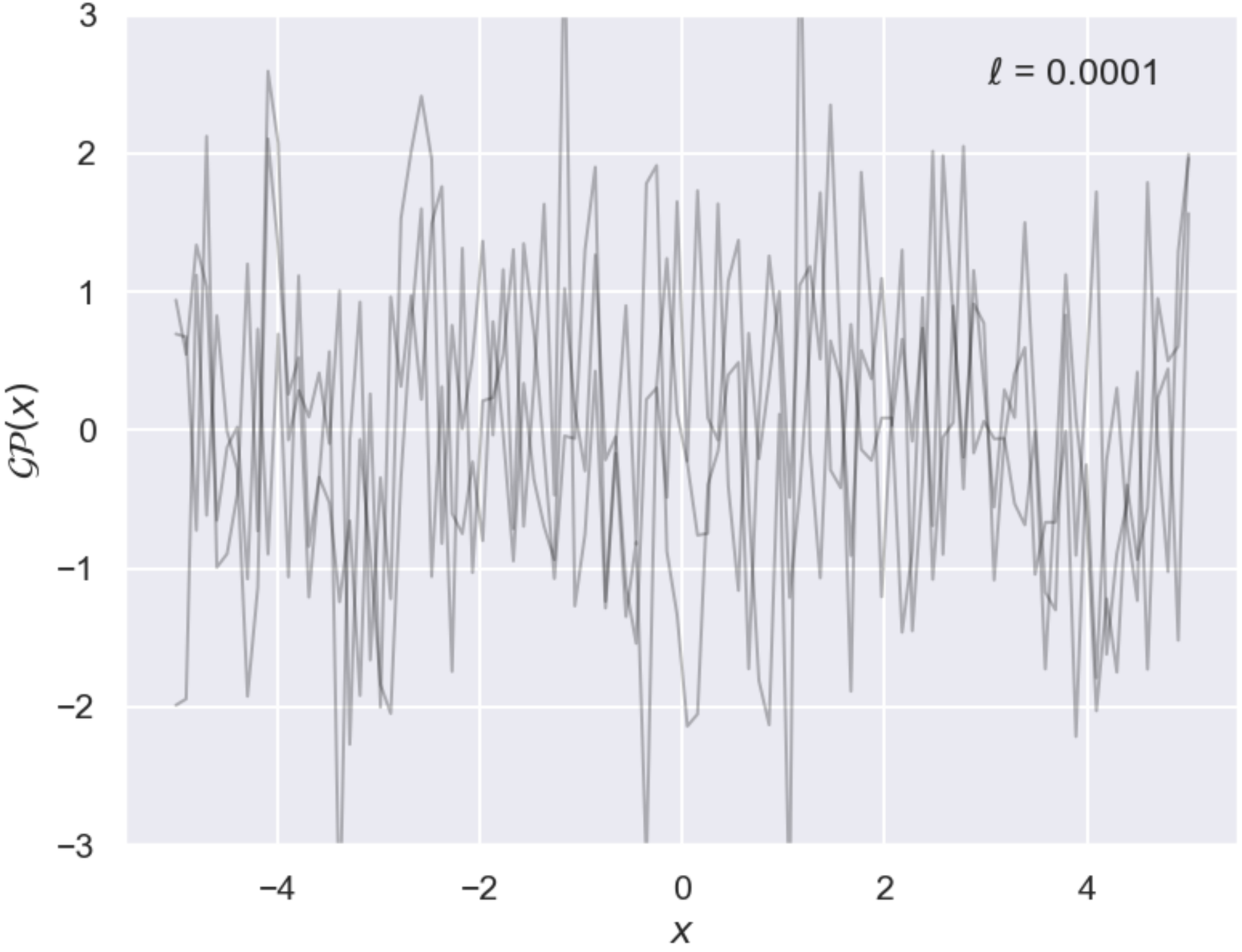

ℓ = 1.0
$\mathcal{GP}(x)$
x
3
2
1
0
−1
−2
−3
−4
−2
0
2
4


ℓ = 2.0
3
2
1
0
−1
−2
−3
$\mathcal{GP}(x)$
−4
−2
0
2
4
x

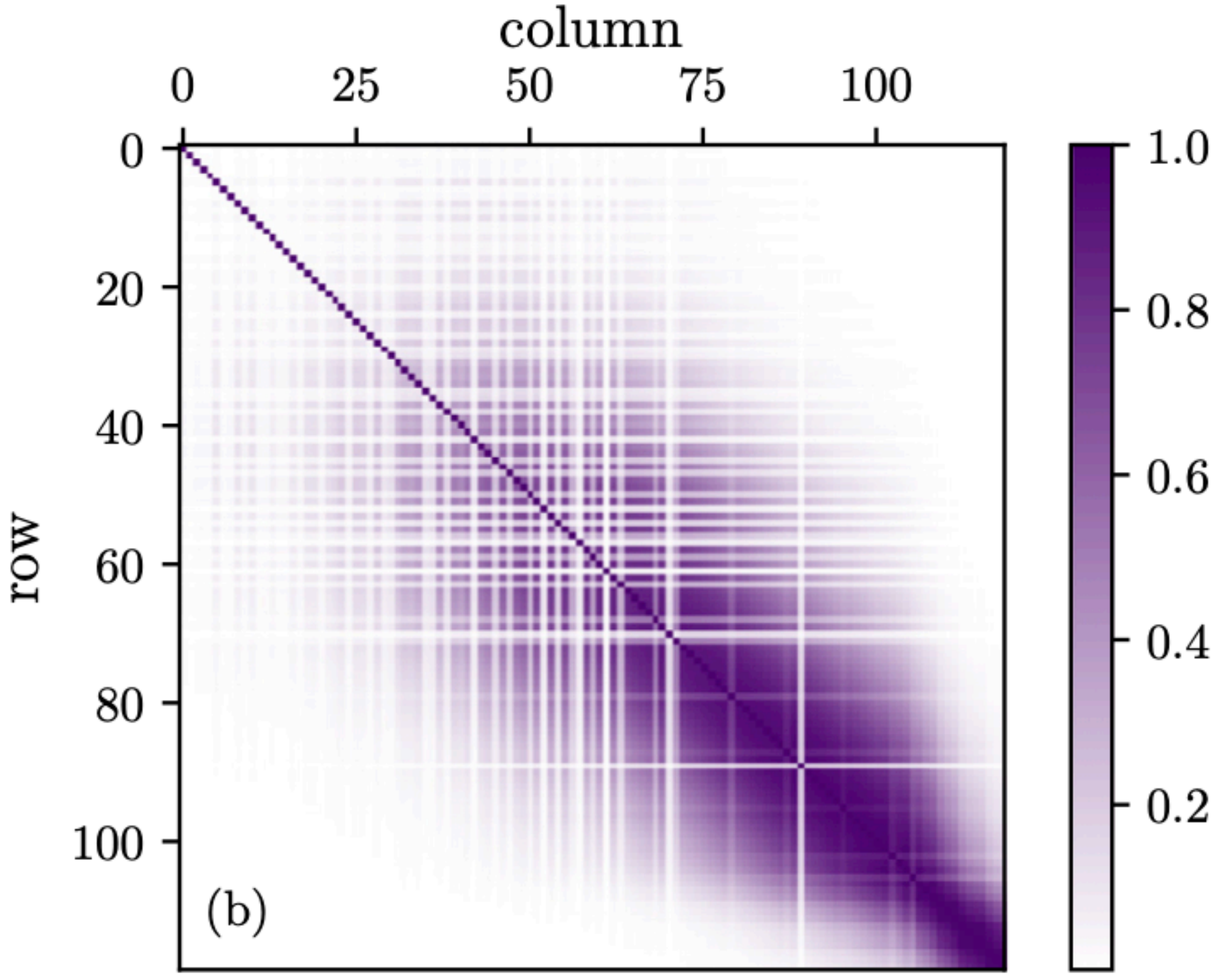

ℓ = 5.0
3
2
1
0
−1
−2
−3
$\mathcal{GP}(x)$
−4
−2
0
2
4
x


ℓ = 0.5
3
2
1
0
−1
−2
−3
$\mathcal{GP}(x)$
−4
−2
0
2
4
x


ℓ = 0.2
3
2
1
0
−1
−2
−3
$\mathcal{GP}(x)$
−4
−2
0
2
4
x


ℓ = 0.0001
3
2
1
0
−1
−2
−3
$\mathcal{GP}(x)$
−4
−2
0
2
4
x

Learning correlations



Notation	Definition	Acronym	$N_{d,y}$	$N_{T_{\rm lab},y}$	$n_{\rm eff}$	$\widehat{\ell_{T_{\rm lab}}}$ (MeV)		$\widehat{c}^2$
$\sigma_{\rm tot}$	total cross section	SGT	119	113	30.8	49		0.56^2
σ_T	$\sigma_{\rm tot}(\uparrow\downarrow) - \sigma_{\rm tot}(\uparrow\uparrow)$	SGTT	3	3	—	—		—
σ_L	$\sigma_{\rm tot}(\leftrightarrows) - \sigma_{\rm tot}(\rightrightarrows)$	SGTL	4	4	3.8	47		1.95^2
Notation	Tensor	Illustration	Acronym	$N_{d,y}$	$N_{T_{\rm lab},y}$	$n_{\rm eff}$	$\widehat{\ell_{T_{\rm lab}}}$ (MeV)	$\widehat{\ell}_\theta$ (deg)	$\widehat{c}^2$
$\sigma(\theta)$	I_{0000}		DSG	1207	68	352.9	73	39	0.61^2
$A(\theta)$	D_{s0k0}		A	5	1	5.0	68	37	0.65^2
$A_t(\theta)$	K_{0ks0}		AT	30	2	25.5	74	33	0.49^2
$A_{yy}(\theta)$	A_{00nn}		AYY	58	4	13.7	66	36	1.27^2
$A_{zz}(\theta)$	A_{00kk}		AZZ	45	2	16.7	70	28	0.98^2
$D(\theta)$	D_{n0n0}		D	13	1	4.5	70	27	0.65^2
$D_t(\theta)$	K_{0nn0}		DT	36	5	31.8	57	29	0.67^2
$D_x^z(\theta)$	D_{0s0k}		D0SK	8	1	3.1	60	29	0.53^2
$A_y(\theta)$	P_{n000}		P	503	28	298.7	59	28	0.31^2
$N_{zz}^y(\theta)$	N_{0nkk}		NNKK	8	1	8.0	45	24	0.22^2
$N_{xy}^x(\theta)$	N_{0skn}		NSKN	12	1	10.9	81	35	0.87^2
$N_{xy}^x(\theta)$	N_{0ssn}		NSSN	4	1	4.0	76	31	0.62^2
$R(\theta)$	D_{s0s0}		R	5	1	5.0	74	32	0.67^2
$R_t(\theta)$	K_{0ss0}		RT	29	3	25.2	83	26	0.55^2
$R_t'(\theta)$	K_{0sk0}		RPT	1	1	1.0	61	35	0.53^2

Svensson, Ekström, Forssen Phys Rev C (2024)

Learning correlations

Low-energy NN scattering: correlated UQ

$$y_{\rm exp}(\vec{x}) = y_{\rm th}^{(k)}(\vec{\alpha};\vec{x}) + \delta y_{\rm exp}(\vec{x}) + \delta y_{\rm th}^{(k)}(\vec{x}).$$

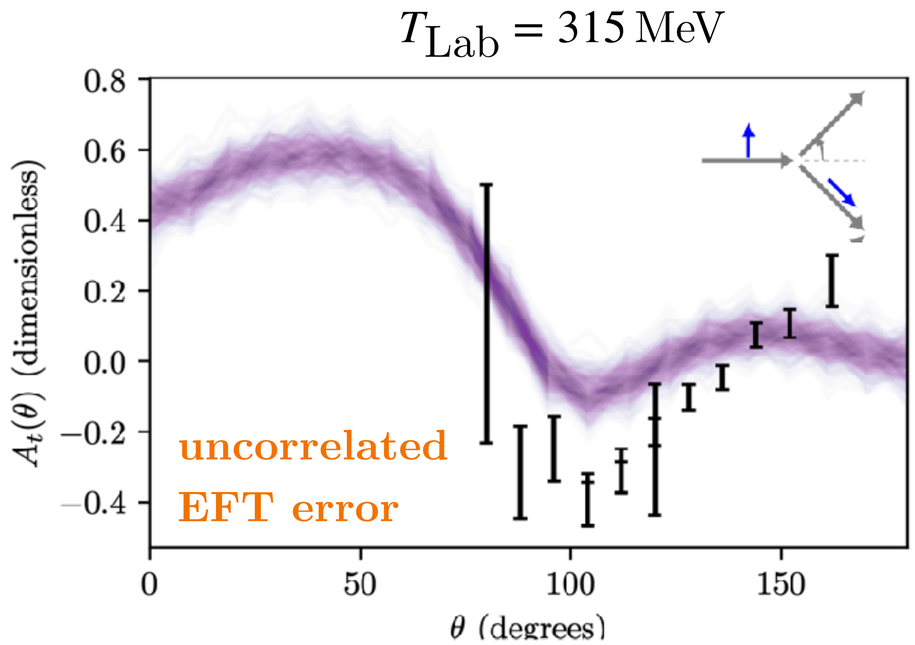
Model correlations across kinematics $\vec{x} = (T_{\rm lab}, \theta)$,

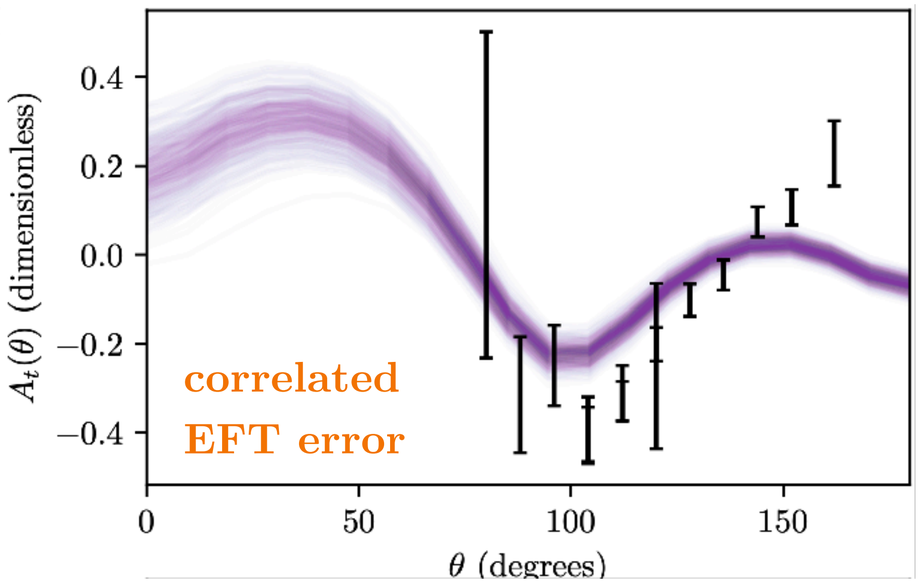
$$(\boldsymbol{\Sigma}_{{\rm th},y})_{mn} = {\rm cov}\left[\delta y_{\rm th}^{(k)}(\vec{x}_m), \delta y_{\rm th}^{(k)}(\vec{x}_n)\right]$$ **GP-model**

Correlation lengths for np scattering: 45–83 MeV & 24–39 deg

Accounting for correlations:

- reduces the effective number of independent NN data with factors 8 and 4 at NLO and NNLO.
- leaves LEC posterior mode location more or less unchanged, but doubles the posterior width.
- yields a smoother and more realistic uncertainty estimate.



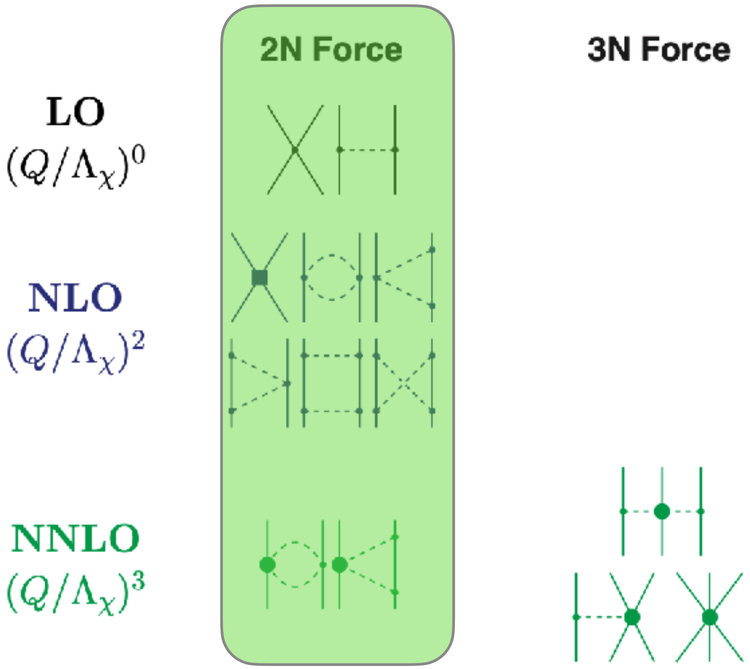


I. Svensson, et al. Phys. Rev. C (2024)

J. A. Melendez, et al. Phys. Rev. C (2019)

Defining priors

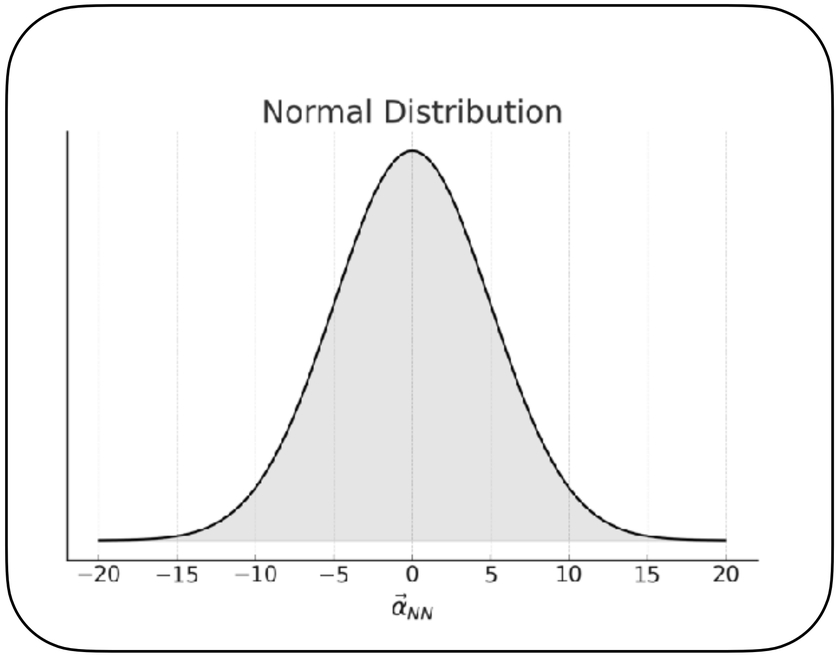
- encoding your domain knowledge

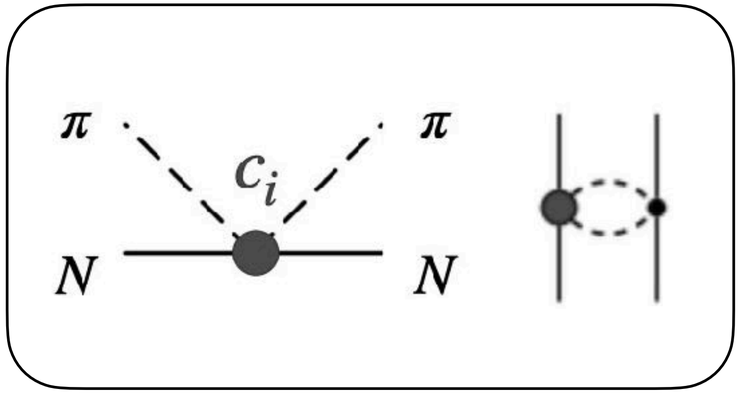


We have two types of LECs: ‘short-range’ and ‘long-range’.

Assuming no prior dependence between these sectors, we write

$$p(\vec{\alpha}\,|\,I) = p(\vec{\alpha}_{NN}\,|\,I) \cdot p(\vec{\alpha}_{\pi N}\,|\,I)$$





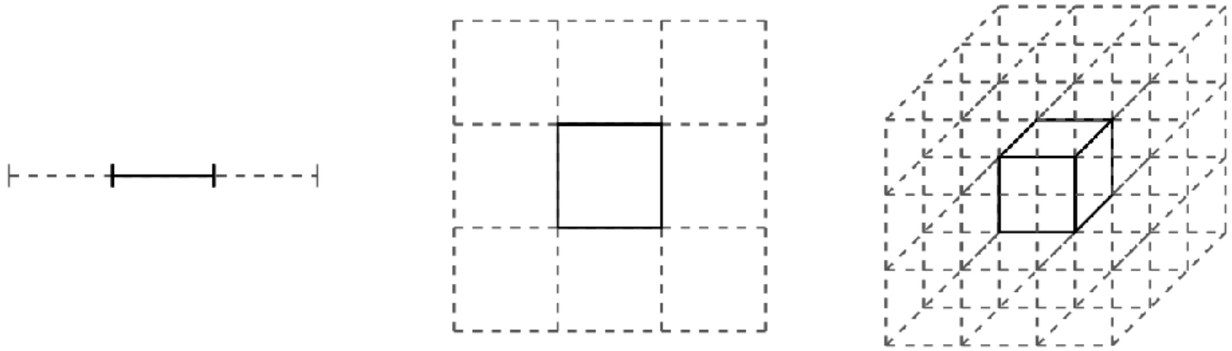
Roy-Steiner analysis of πN scattering amplitudes

Svensson, Ekström, Forssen Phys Rev C (2022)

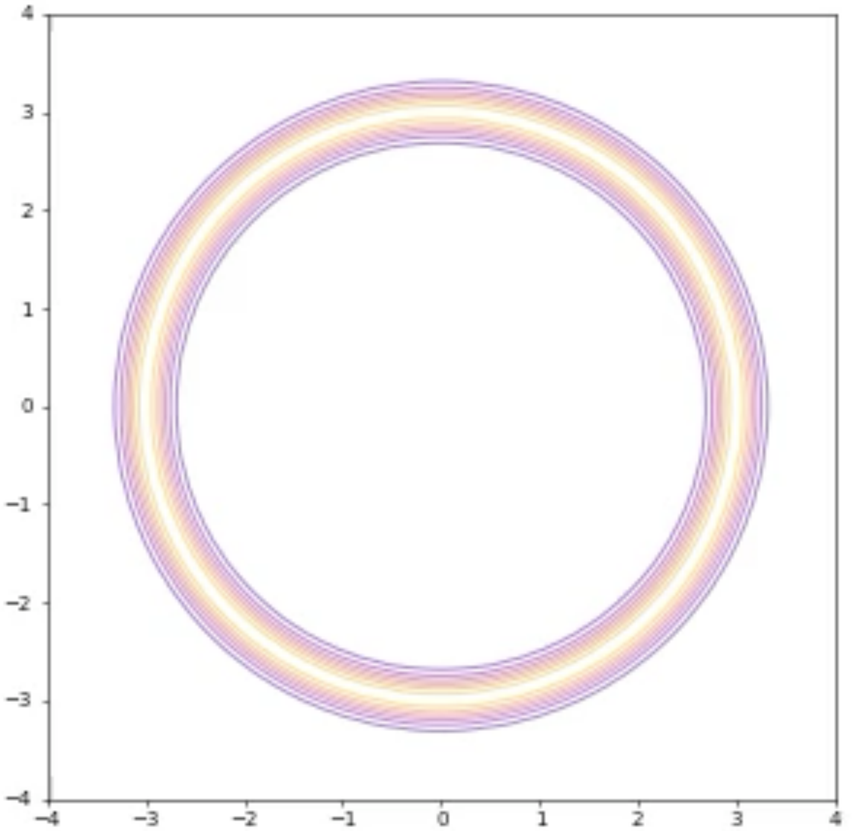
Sampling the posterior

- Markov chain Monte Carlo

$$p(\vec{\alpha} \,|\, \mathscr{D}, M, I) = \frac{p(\mathscr{D} \,|\, \vec{\alpha}, M, I) \cdot p(\vec{\alpha} \,|\, M, I)}{p(\mathscr{D} \,|\, M, I)}$$



credit: M. Betancourt arXiv:1701.02434

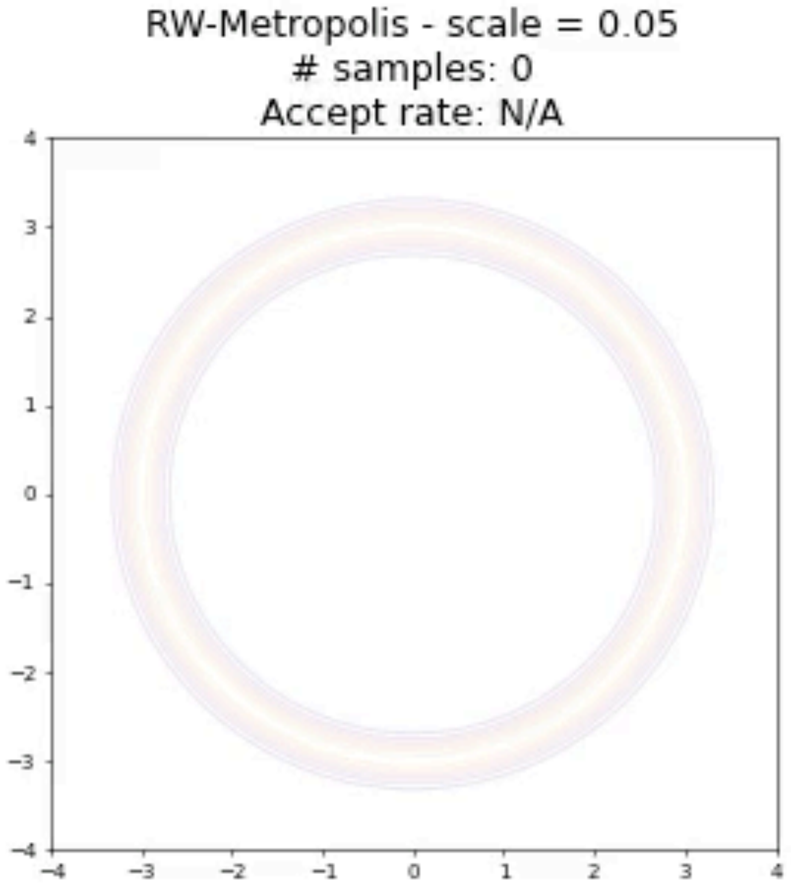


With 10^4 posterior samples per chain, and 10 chains, at each order the HMC sampling passes all convergence tests.

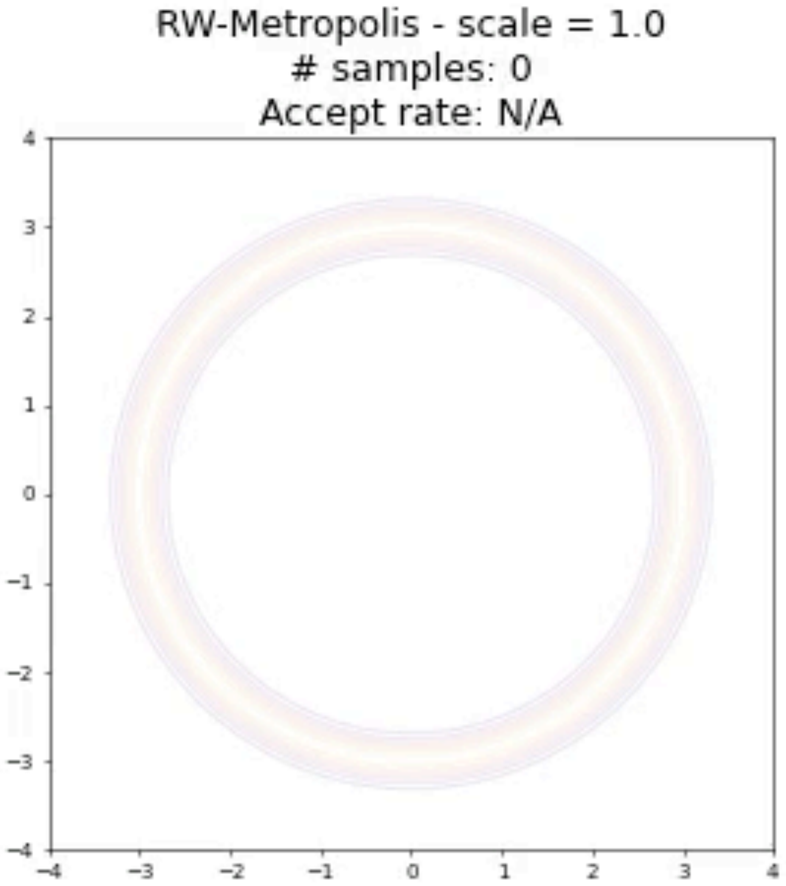
Although we have to compute derivatives and integrate Hamilton's equations at ~ 20 (time) steps for each HMC step, each posterior sample is very informative. This leads to an overall advantage of using HMC.

S. Duane, et al. Phys. Lett B **195,** 216 (1986)
I. Svensson, et al Phys. Rev. C **105**, 014004(2022)

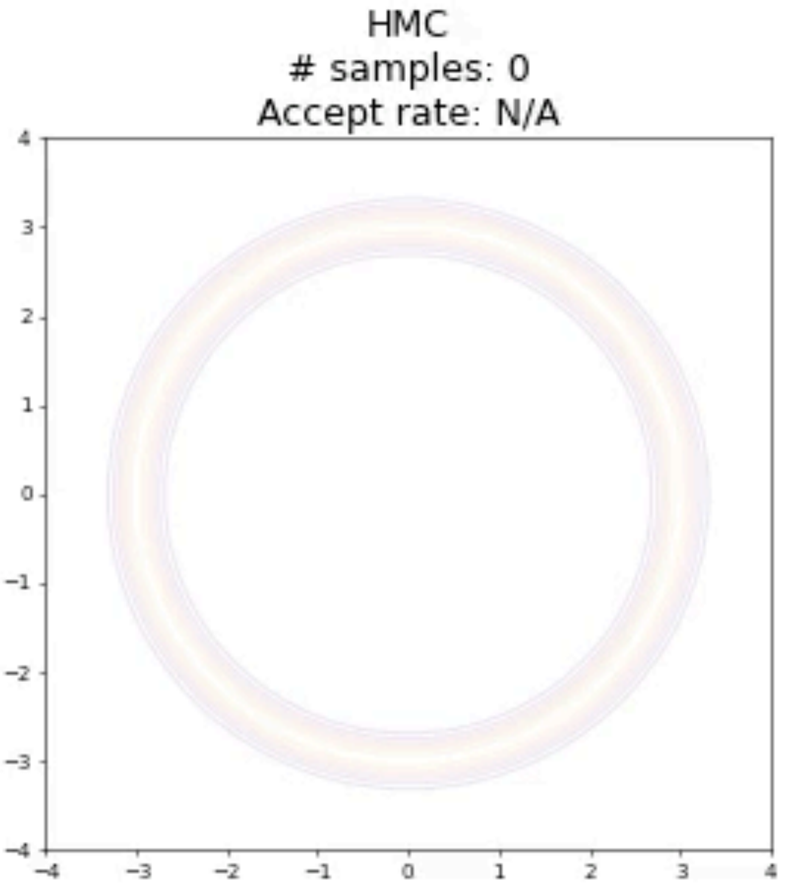
high acceptance
little exploration



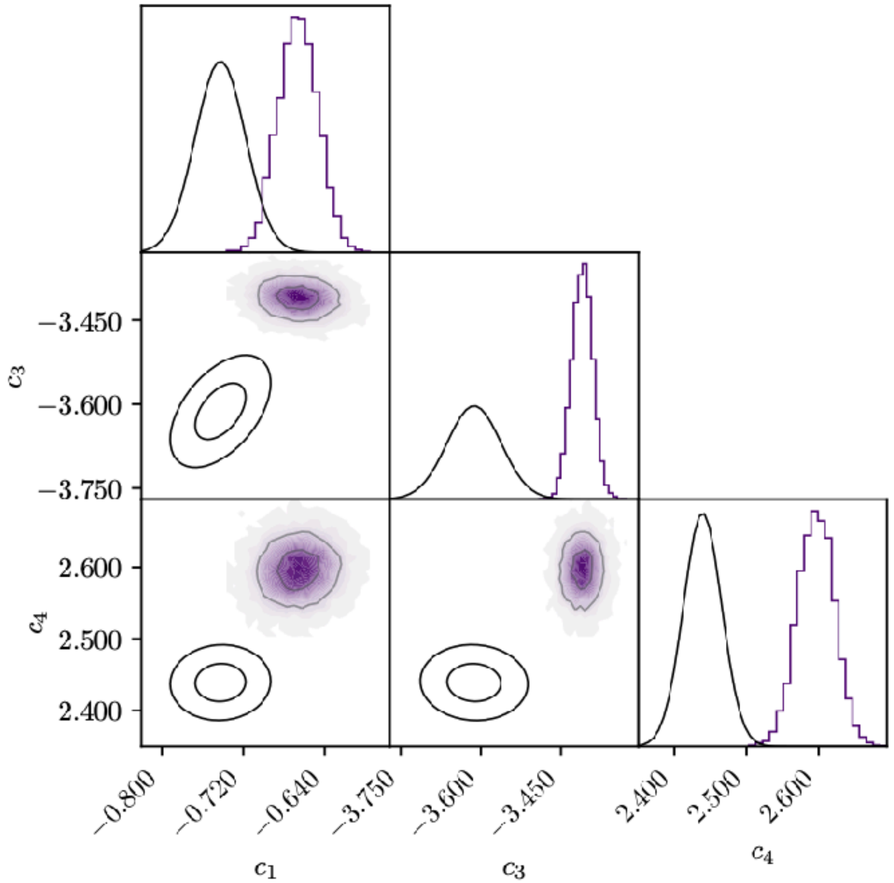
low acceptance
more exploration



high acceptance
more exploration



The *NN* LEC posterior up to N3LO



Subleading πN posterior shift

S. B. S. Miller et al, Phys. Rev. C **106**, 024001 (2022)

Posterior predictive distribution

Low-energy nd scattering cross sections (NN-only)

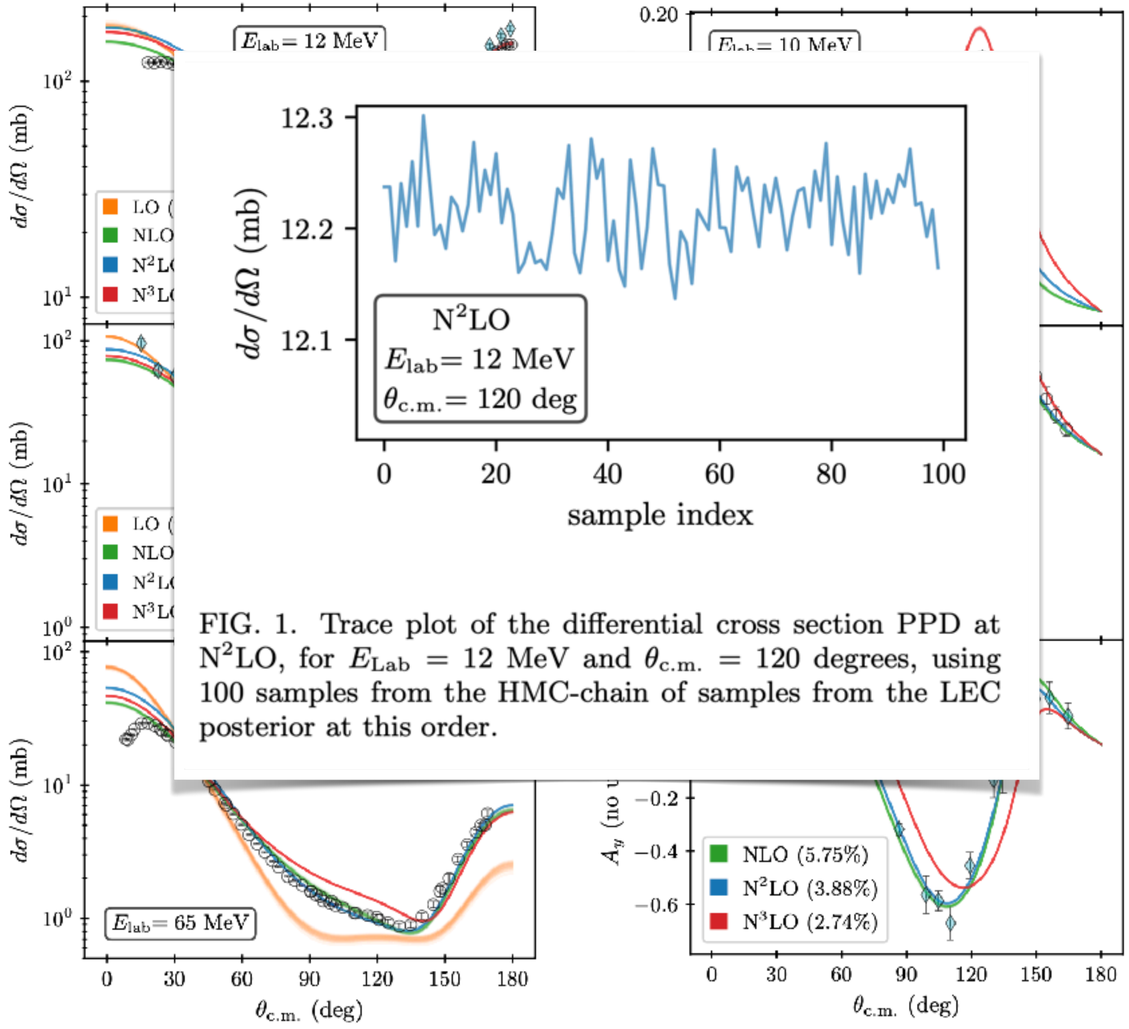
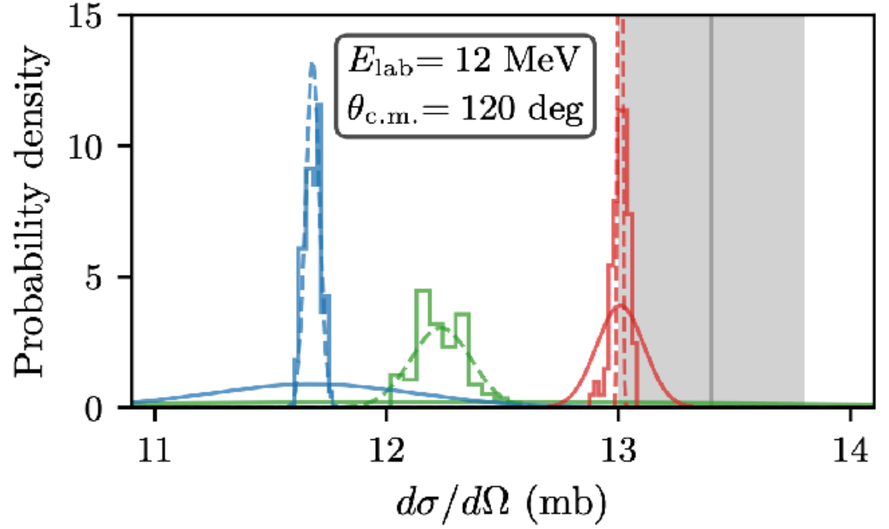


FIG. 1. Trace plot of the differential cross section PPD at N^{2}LO, for $E_{\rm Lab}$ = 12 MeV and $\theta_{\rm c.m.}$ = 120 degrees, using 100 samples from the HMC-chain of samples from the LEC posterior at this order.

Draw 100 samples from the LEC posteriors at LO, ...,N3LO and compute *nd* cross sections.

Does not help resolve the A_y puzzle.



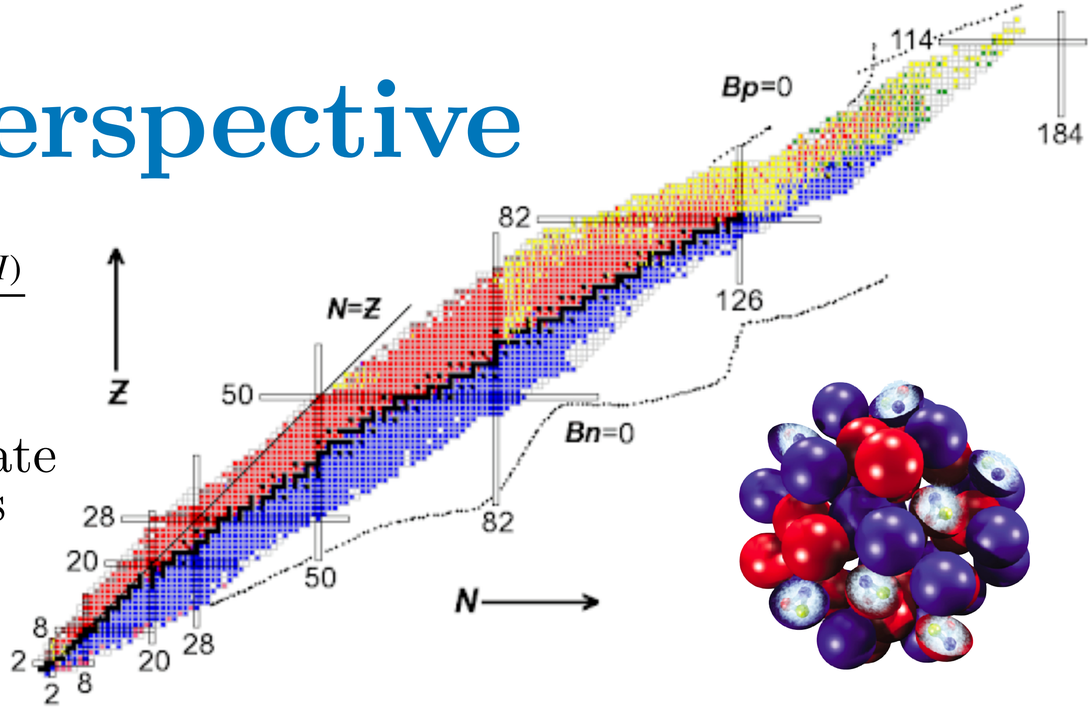
At N3LO, our estimates of the EFT truncation error is comparable to the LEC PPD. Both smaller than experiment (gray).

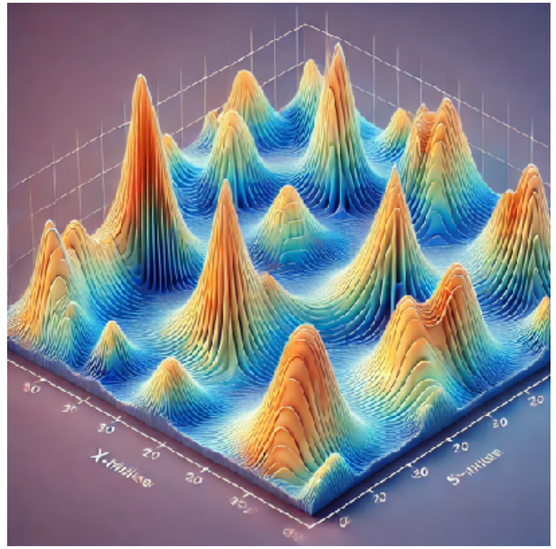
S. B. S. Miller et al, PRC (2022)

A bigger perspective

$$p(\vec{\alpha}\,|\,\mathscr{D},M,I) = \frac{p(\mathscr{D}\,|\,\vec{\alpha},M,I)\cdot p(\vec{\alpha}\,|\,M,I)}{p(\mathscr{D}\,|\,M,I)}$$

We can use A>3 to calibrate chiral NN+3N interactions





We would like to test our *ab initio* model across the nuclear chart.

Can we find regions in the parameter space that reproduce seen data?

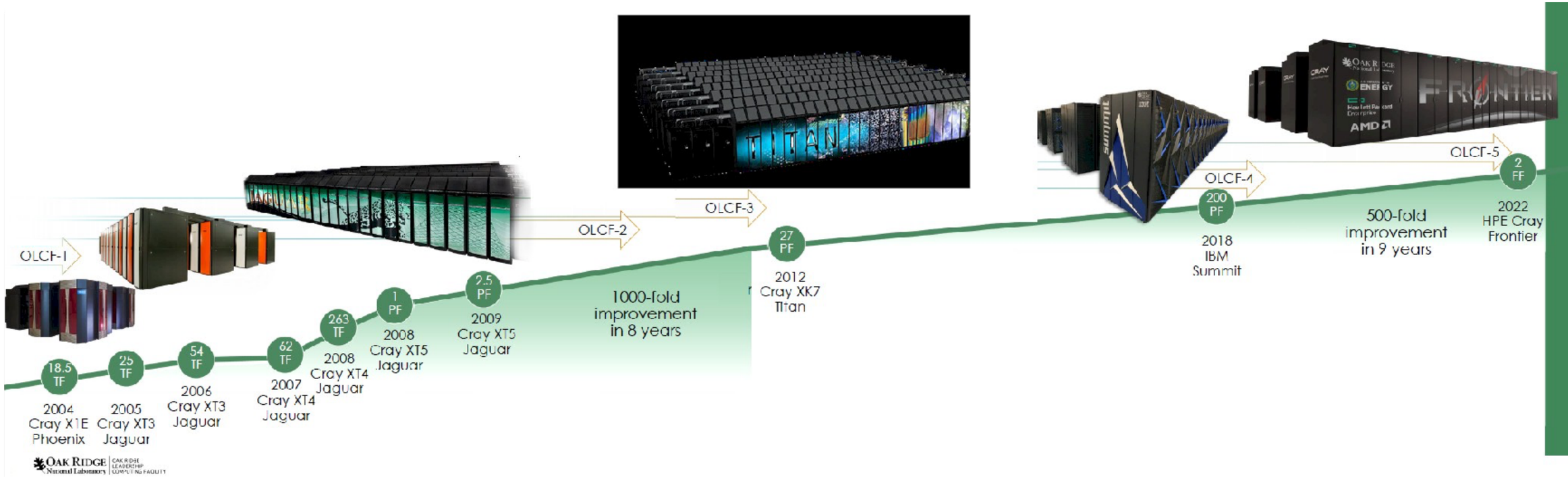
😰

- computational cost of likelihood ⟶ - Emulators
- multimodality and MCMC ⟶ - History matching
- formulating model discrepancy ⟶ - Needs more work

😀

Computing nuclei: an HPC problem

Solving the Schrödinger equation for a large collection of strongly interacting nucleons typically requires substantial high-performance computing resources. Naively, the computational cost to solve the Schrödinger equation grows exponentially with nucleon number and basis size. Polynomially scaling methods exist but are still computationally expensive.



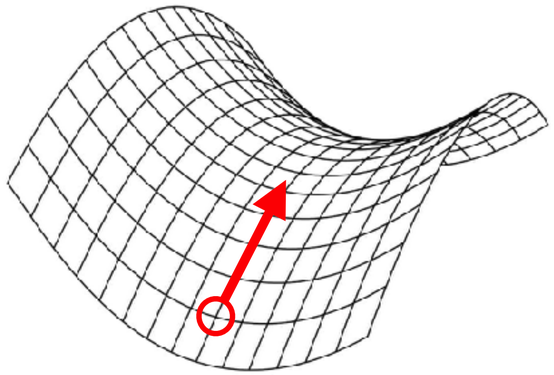

The likelihood is computationally expensive and the parameter space is huge

Eigenvector continuation (EC)

$$H(\alpha) = H_0 + \alpha H_1$$

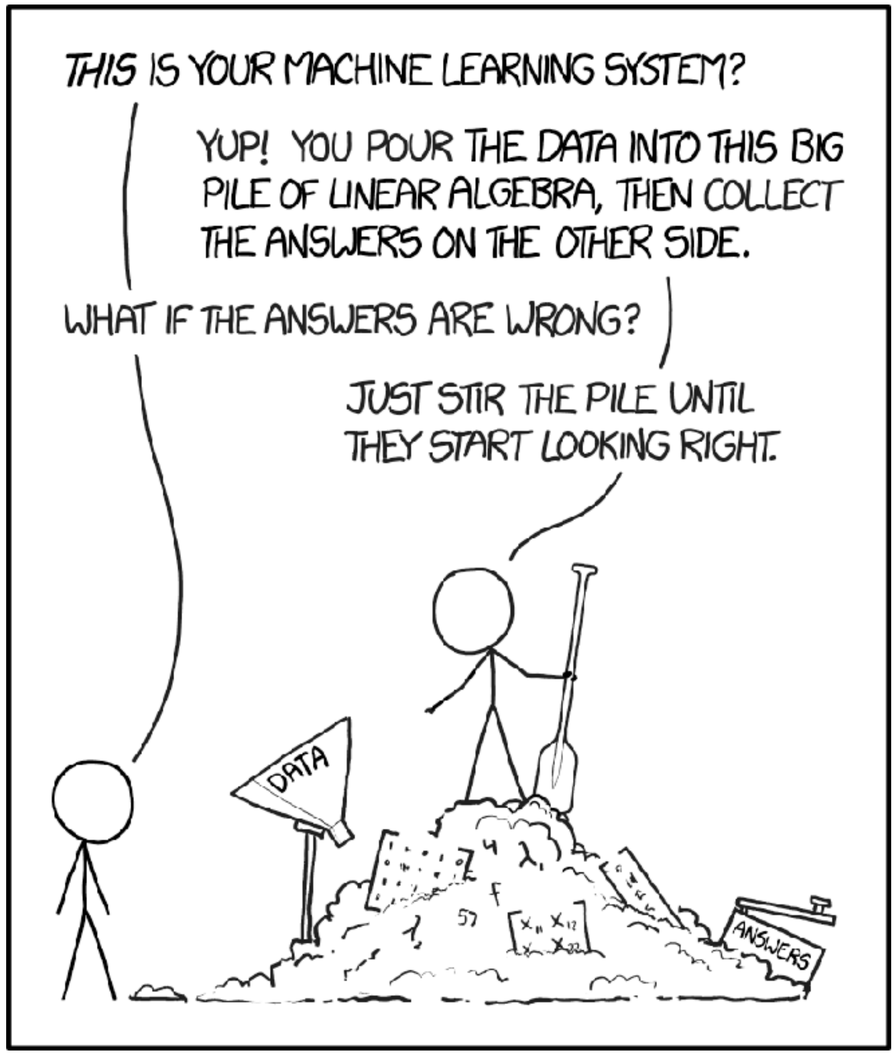
continuous parameter

The key insight is that while an eigenvector resides in a linear space with enormous dimensions, the eigenvector trajectory generated by smooth changes of the Hamiltonian matrix is well approximated by a very low-dimensional manifold in many applications.



D. Frame, et al. Phys. Rev. Lett. **121**, 032501 (2018)

EC for emulation ("supervised learning")



$$y_{\rm th}(\vec{\alpha};\vec{x}) = \tilde{y}_{\rm th}(\vec{\alpha};\vec{x}) + \delta\tilde{y}_{\rm th}(\vec{\alpha};\vec{x})$$

SLOW FAST

We have a Hamiltonian that depends smoothly on parameters. For chiral potentials we can often write this parameter dependence as

$$\hat{H}(\vec{\alpha}) = \hat{H}_0 + \sum_{i=1}^{N} \alpha_i \hat{H}_i$$

Pick N_{EC} (training) points in the parameter space where we solve for "exact" solutions (using e.g. NCSM or CC)

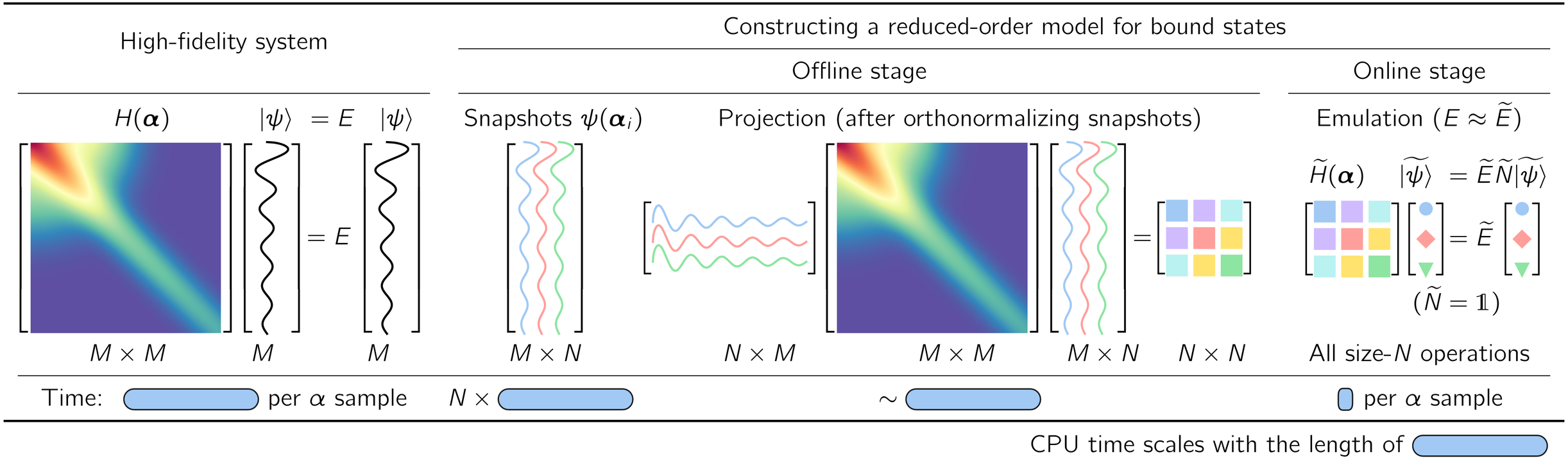
Project onto a basis of N_{EC} "exact "wave functions and diagonalize the generalized eigenvalue problem

$$\widetilde{H}(\vec{\alpha}_\odot)|\widetilde{\Psi}_0\rangle = \widetilde{E}_0(\vec{\alpha}_\odot)\mathbf{N}|\widetilde{\Psi}_0\rangle$$

S. König, et al. Phys. Lett. B 810, 135814 (2020)

A. Ekström and G. Hagen Phys. Rev. Lett. 123, 252501 (2019)

EC for fast and accurate many-body emulation



$$\hat{H}(\vec{\alpha}) = \hat{H}_0 + \sum_{i=1}^{N} \alpha_i \hat{H}_i$$

S. König, et al. Phys. Lett. **B 810**, 135814 (2020)

A. Ekström and G. Hagen Phys. Rev. Lett. **123**, 252501 (2019)

T. Duguet, A. Ekström, R. J. Furnstahl, S. König, D. Lee, Rev. Mod Phys **96**, 031002 (2024)

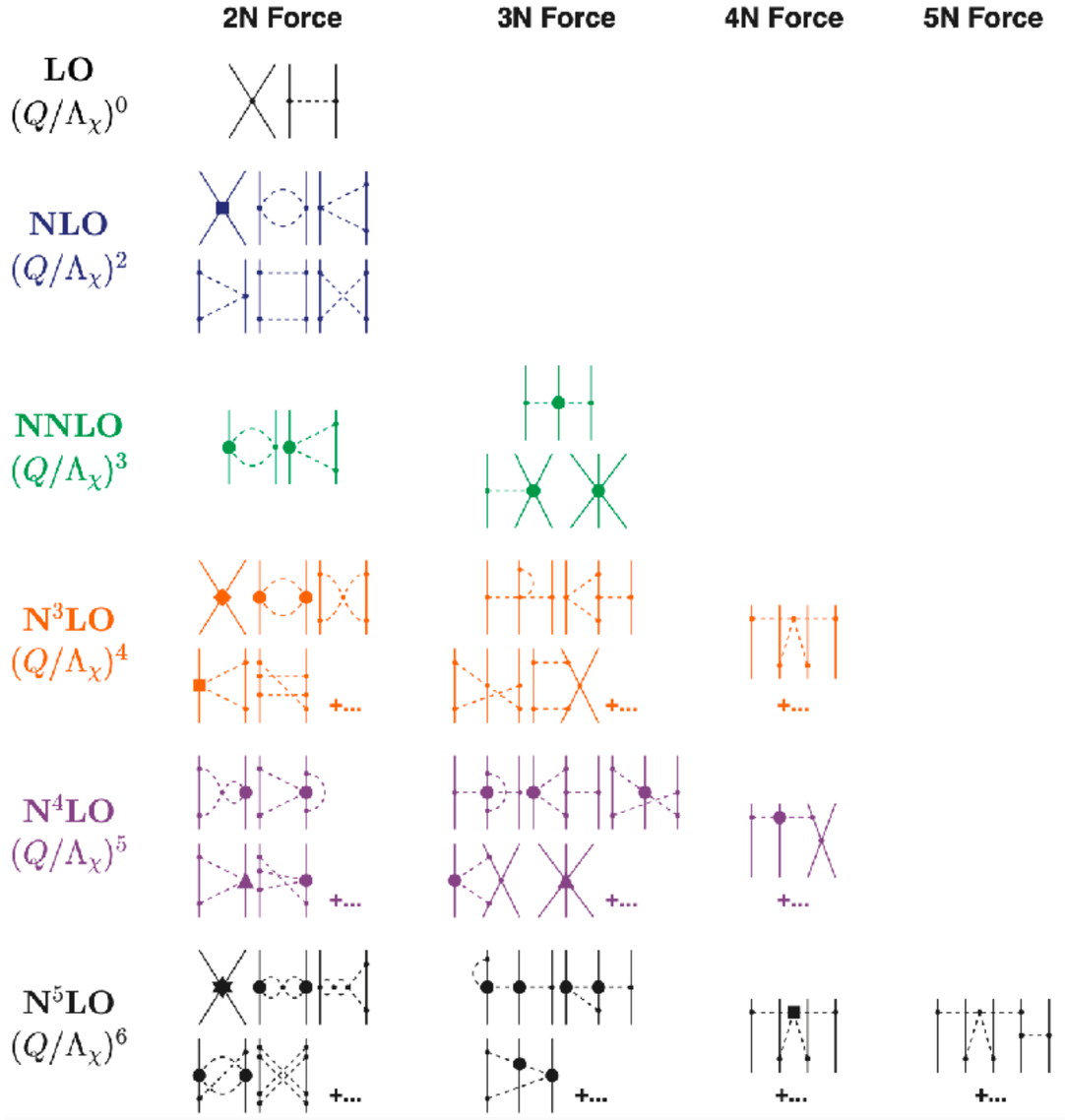
Linear structures in Weinberg power counting

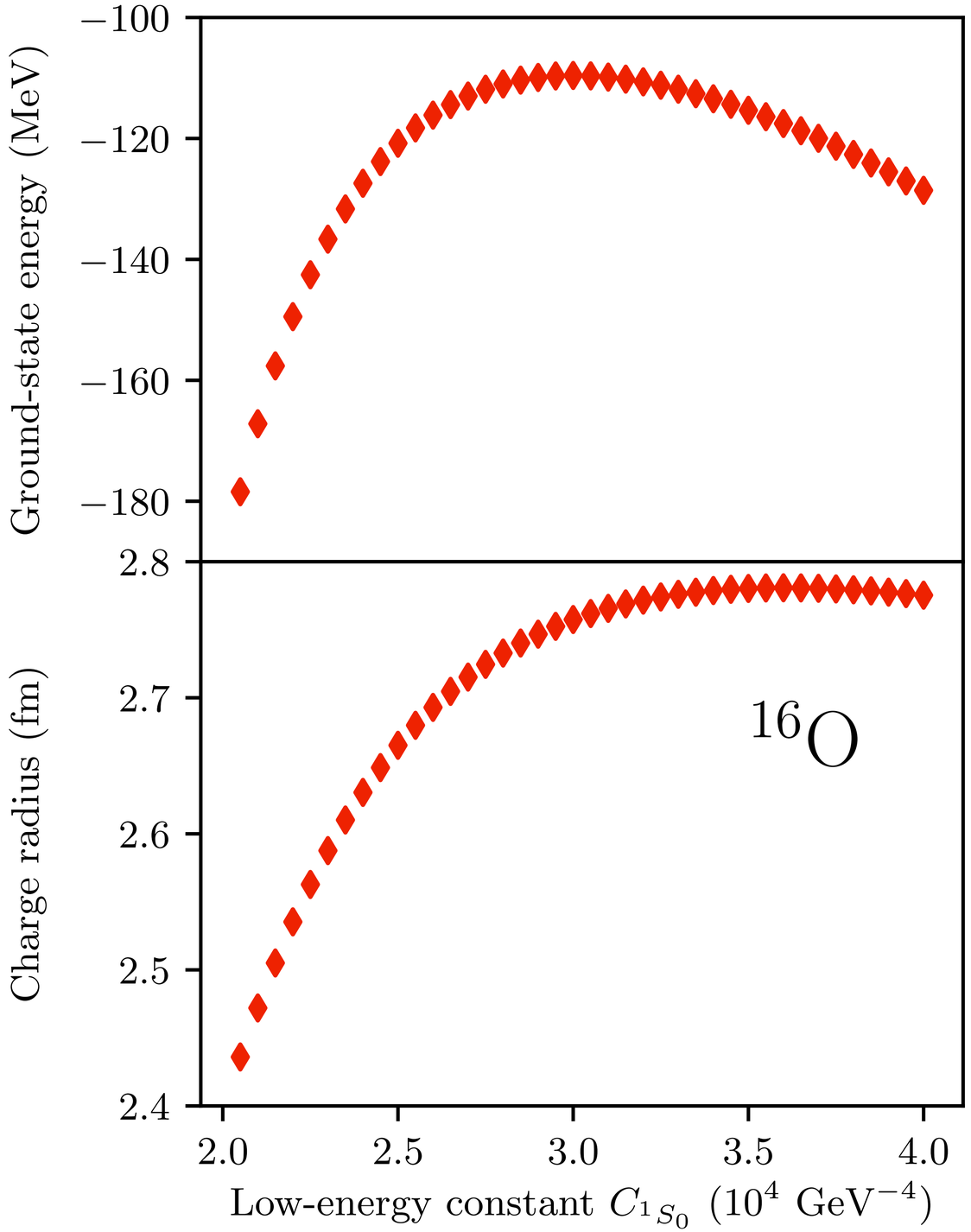
Speedup: for chiral potentials we can often write the parameter dependence as

$$V(\alpha) = 1\pi + \sum_i h_i(\alpha)\, 2\pi + \sum_i g_i(\alpha)$$

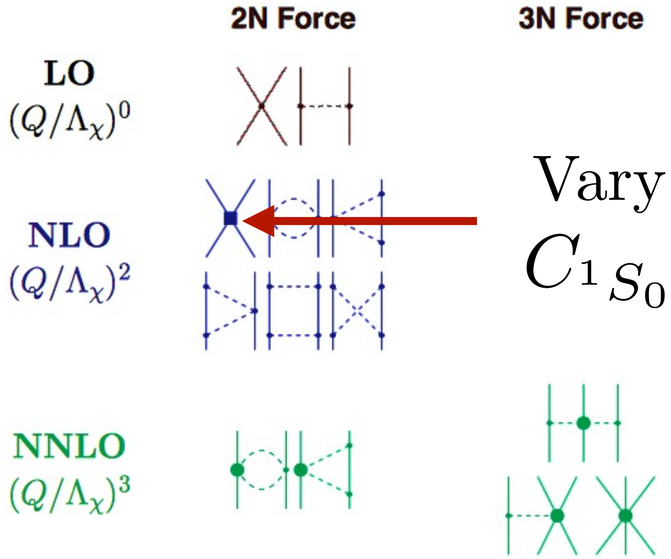
The Hamiltonian up to NNLO is trivially linear in the relevant EFT coupling constants.

Starting at N3LO, there are quadratic dependencies of the $\alpha_i \alpha_j$ due to higher order loop diagrams. Still, it can be factored out of the potential terms



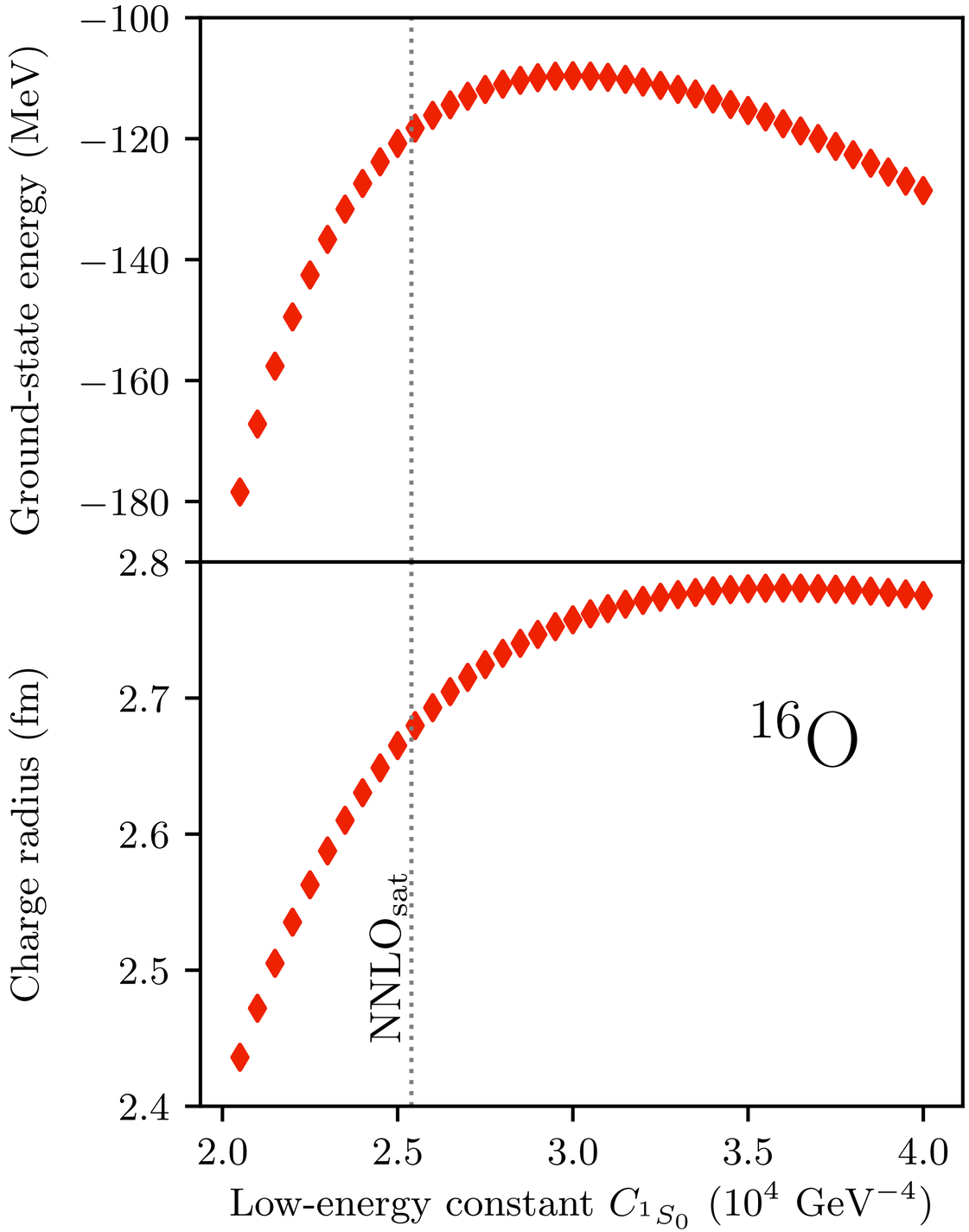


Exact coupled cluster calculations at the singles and doubles level

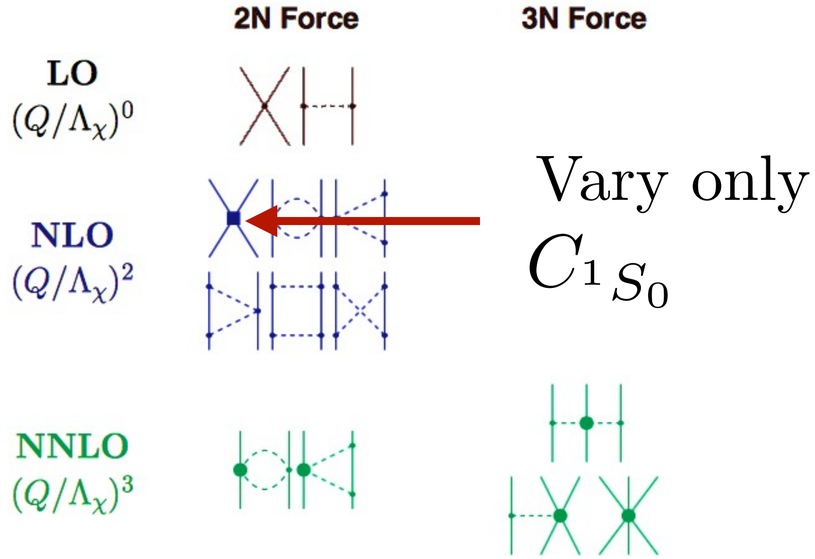


Vary only C_{1S_0}

A. Ekström and G. Hagen Phys. Rev. Lett. **123**, 252501 (2019)

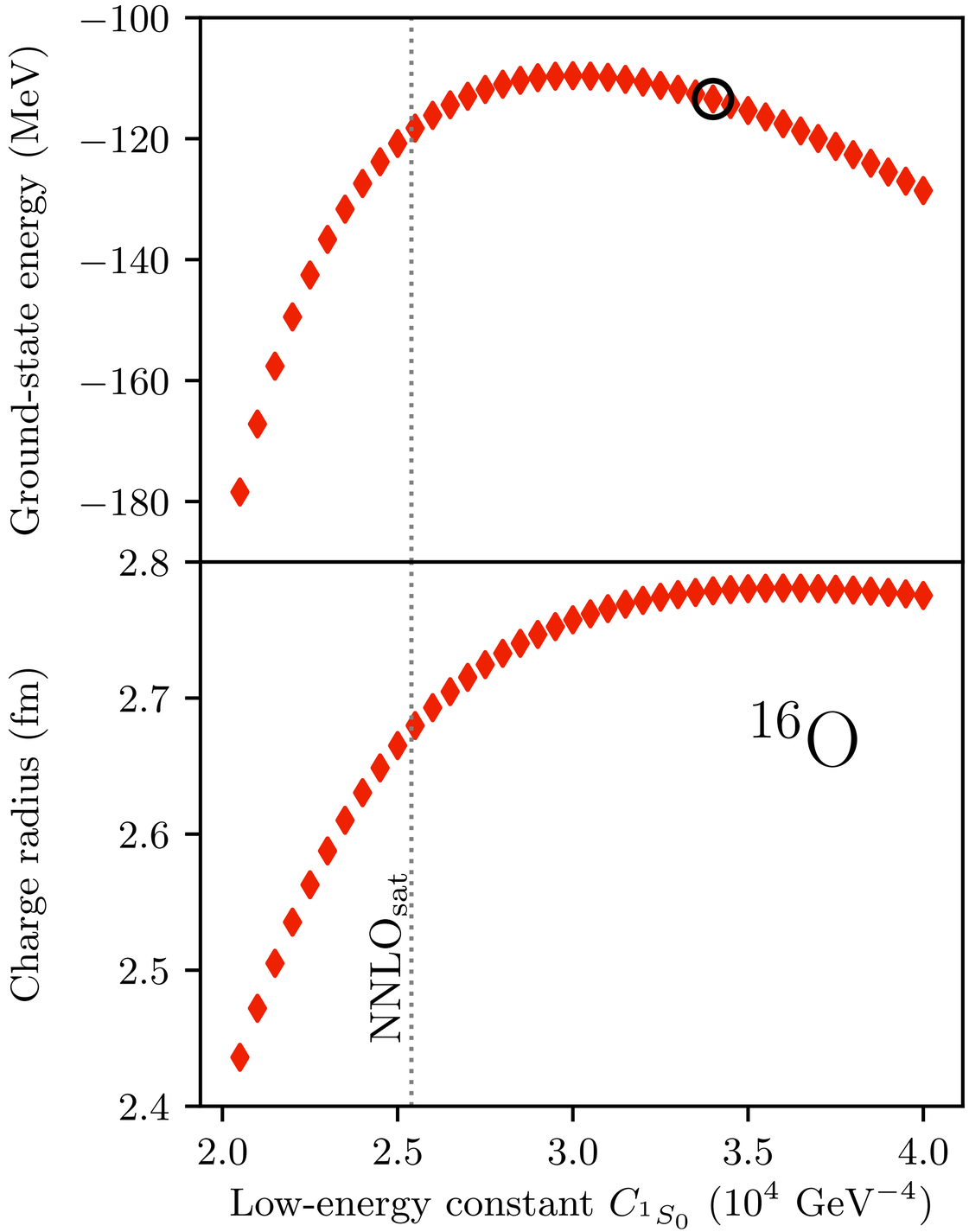


Exact coupled cluster calculations at the singles and doubles level

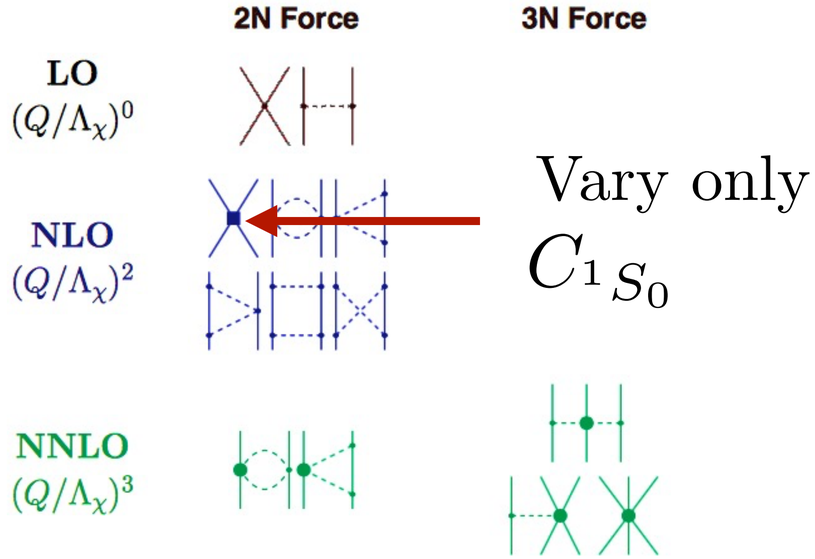


Vary only C_{1S_0}

A. Ekström and G. Hagen Phys. Rev. Lett. **123**, 252501 (2019)

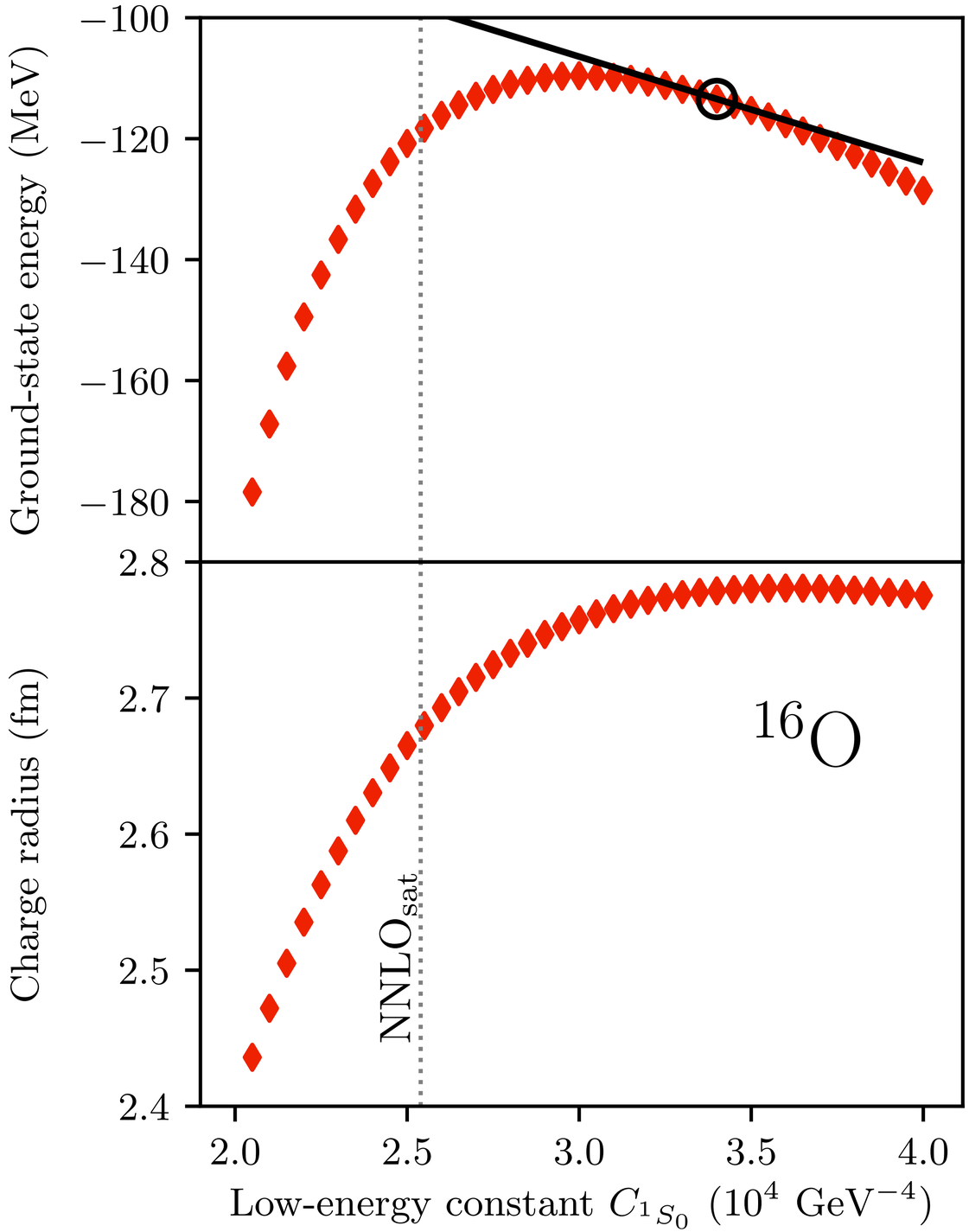


Exact coupled cluster calculations at the singles and doubles level



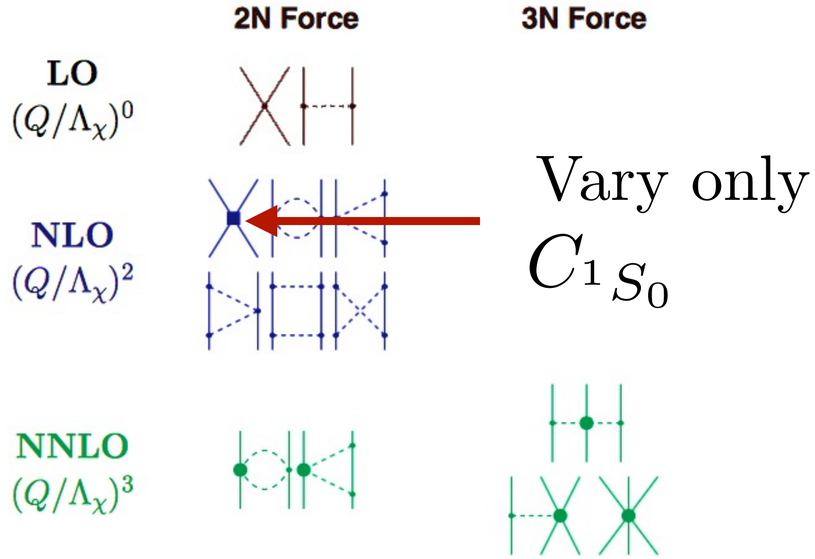
Vary only C_{1S_0}

A. Ekström and G. Hagen Phys. Rev. Lett. **123**, 252501 (2019)



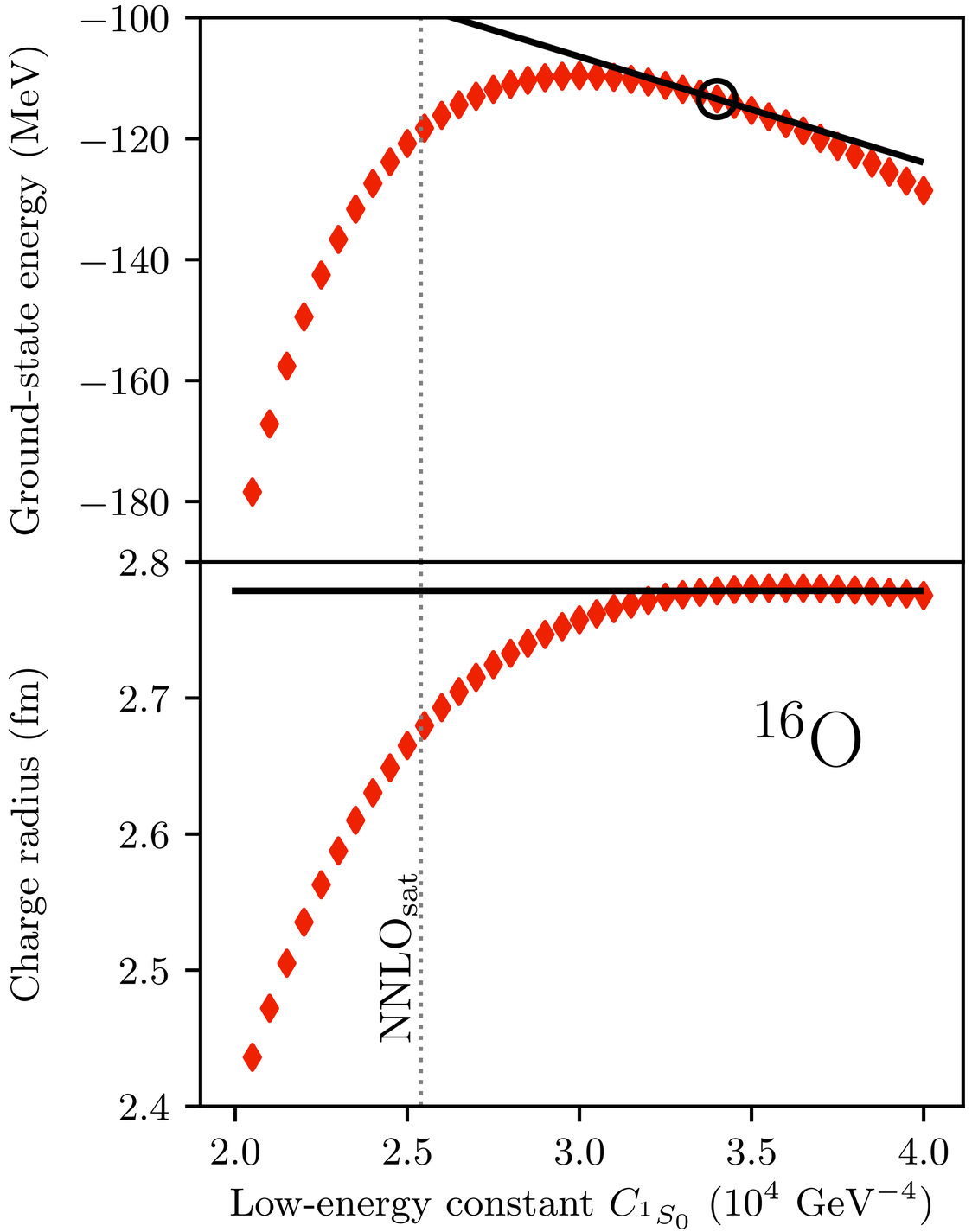
Exact coupled cluster calculations at the singles and doubles level

SPCC



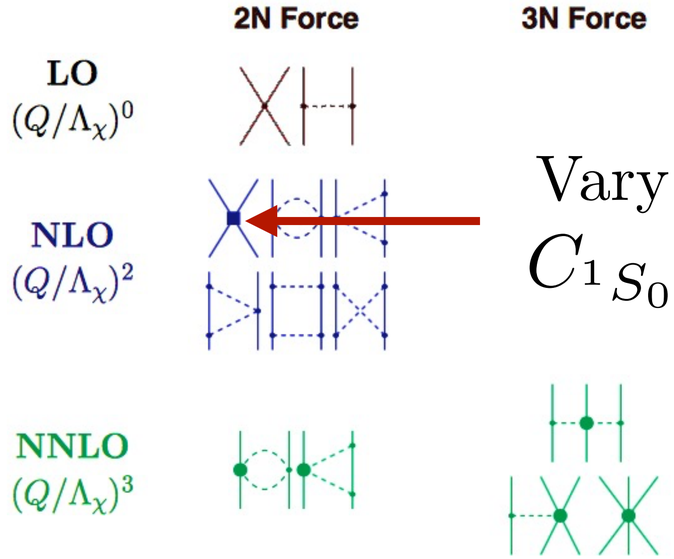
Vary only C_{1S_0}

A. Ekström and G. Hagen Phys. Rev. Lett. **123**, 252501 (2019)



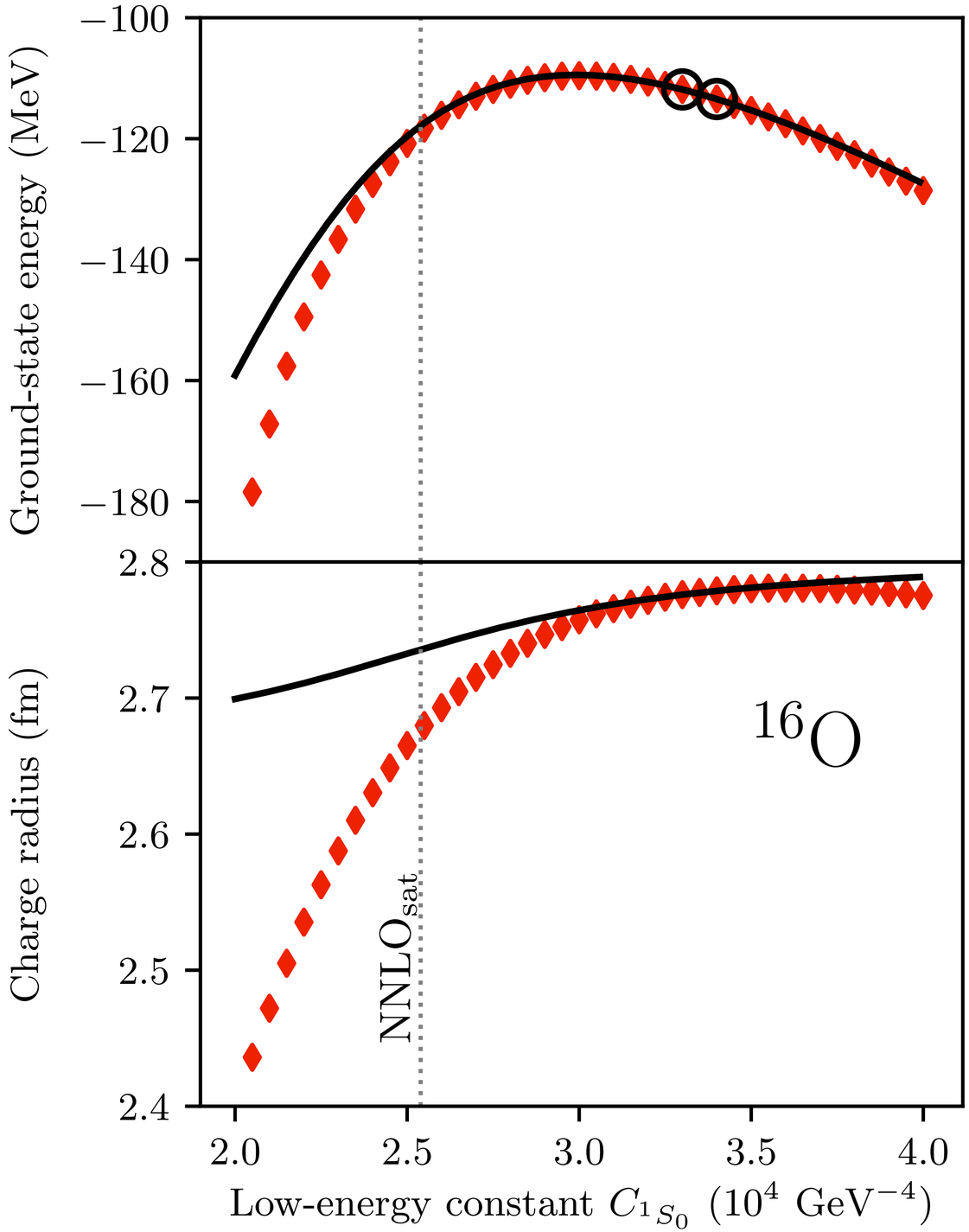
Exact coupled cluster calculations at the singles and doubles level

SPCC



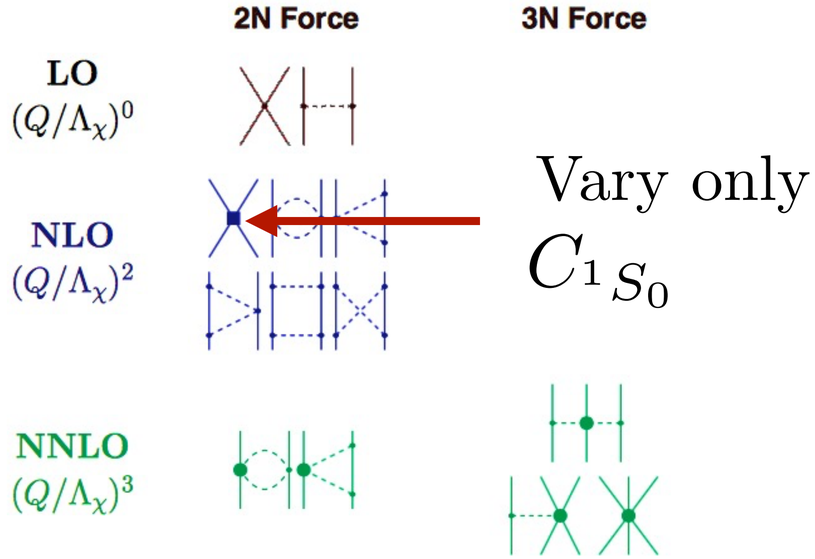
Vary only C_{1S_0}

A. Ekström and G. Hagen Phys. Rev. Lett. **123**, 252501 (2019)



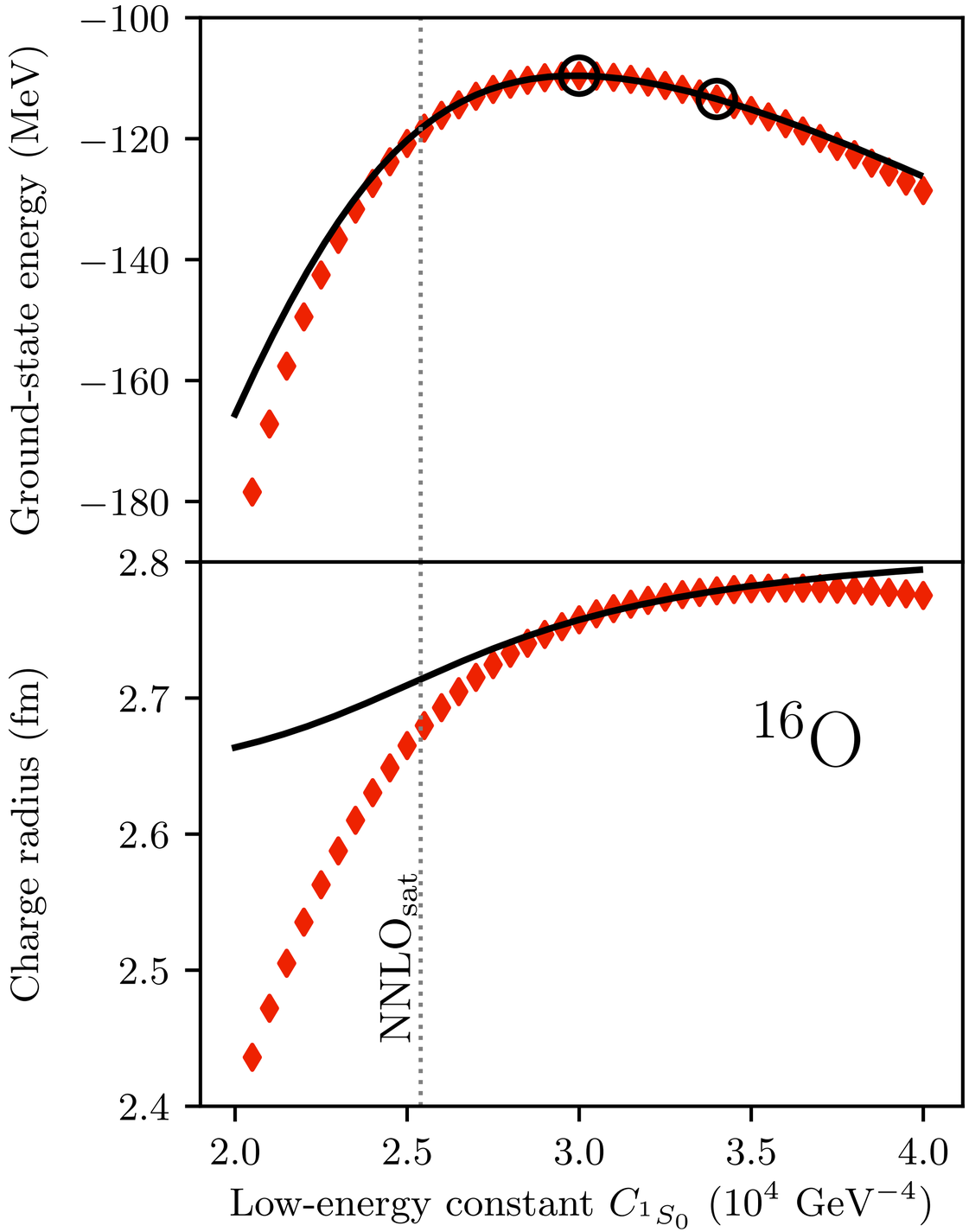
Exact coupled cluster calculations at the singles and doubles level

SPCC



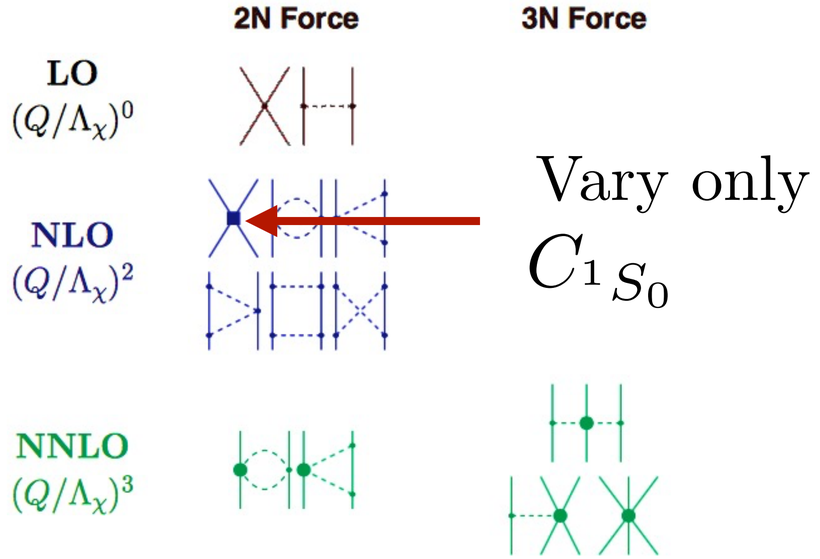
Vary only C_{1S_0}

A. Ekström and G. Hagen Phys. Rev. Lett. **123**, 252501 (2019)



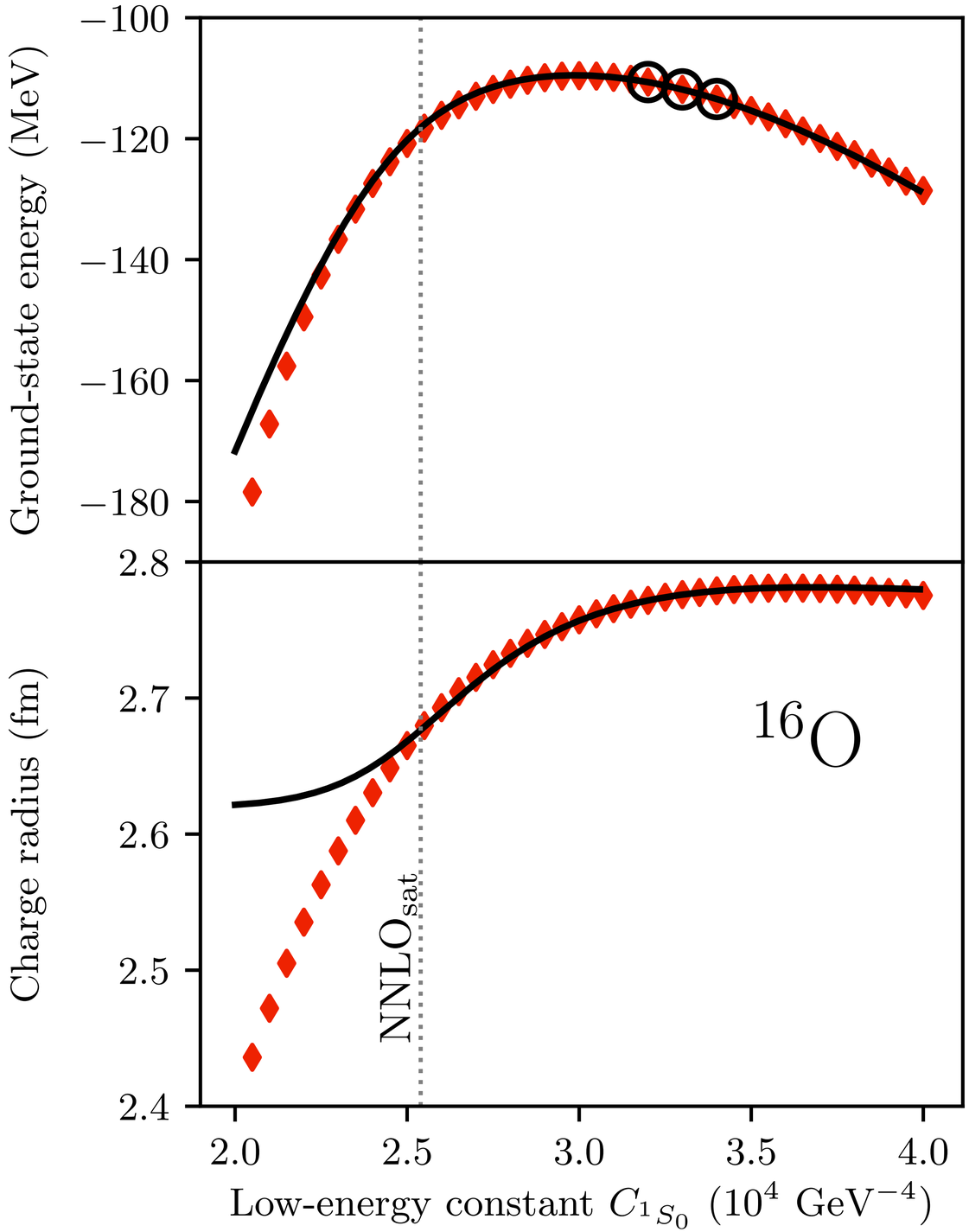
Exact coupled cluster calculations at the singles and doubles level

SPCC



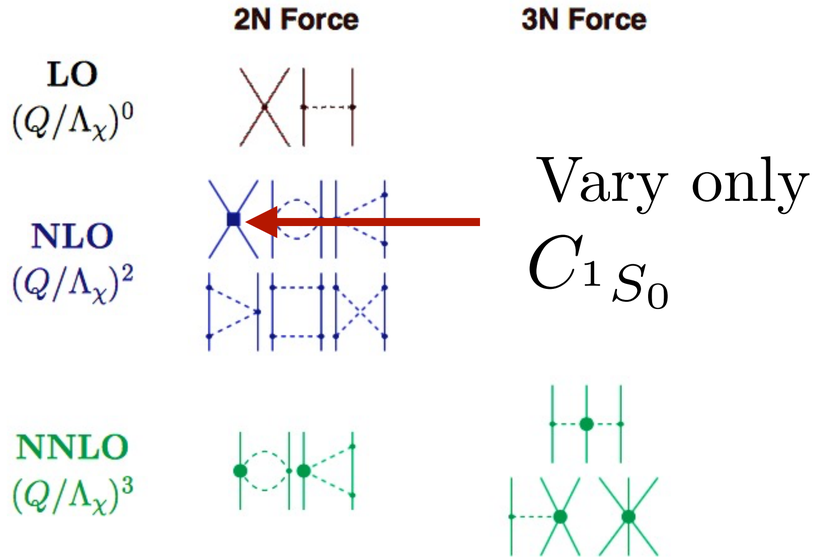
Vary only C_{1S_0}

A. Ekström and G. Hagen Phys. Rev. Lett. **123**, 252501 (2019)



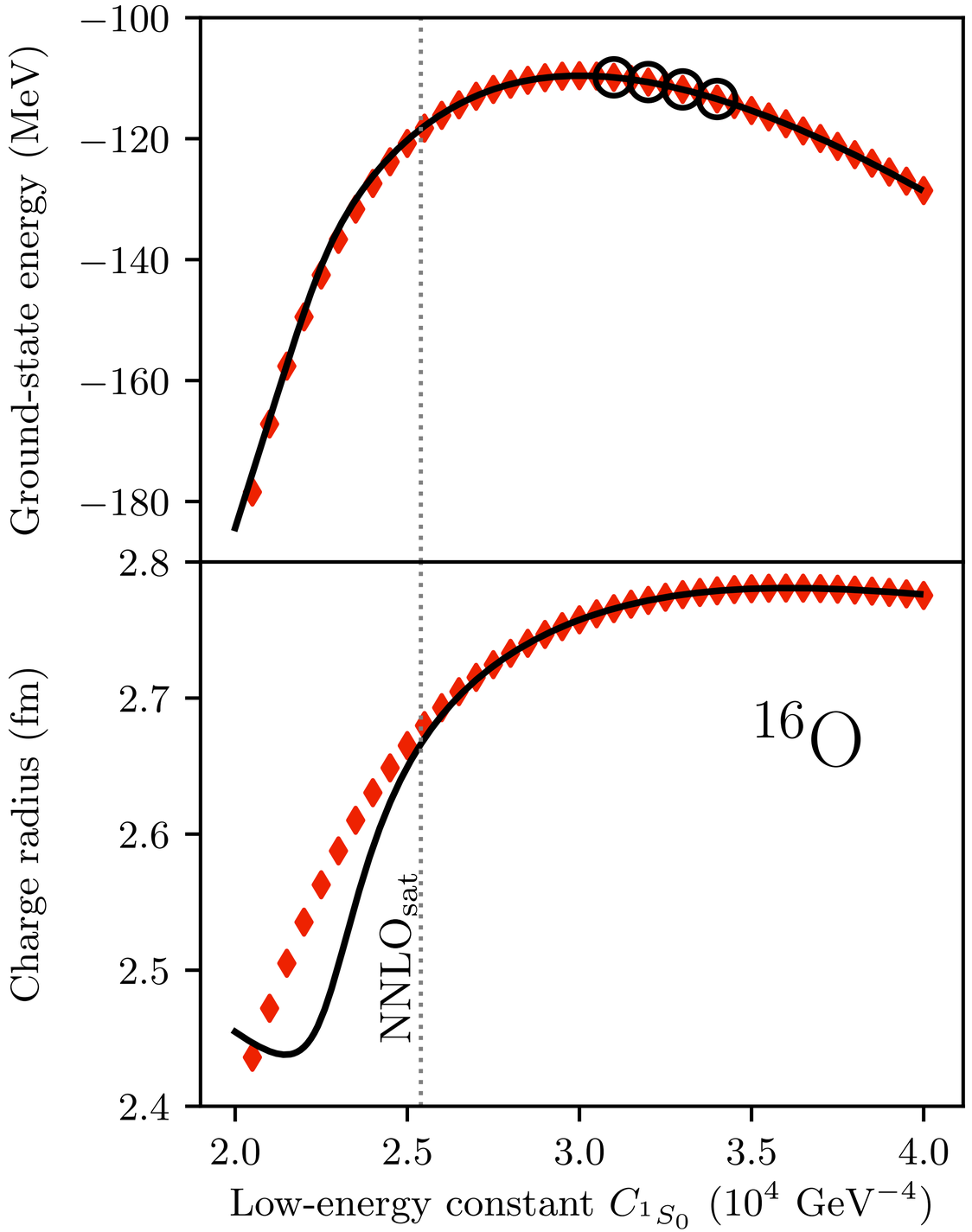
Exact coupled cluster calculations at the singles and doubles level

SPCC



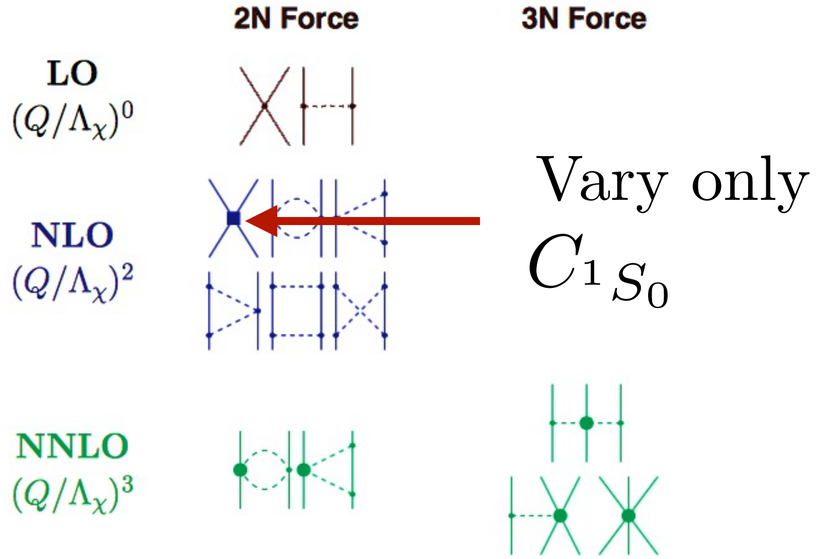
Vary only C_{1S_0}

A. Ekström and G. Hagen Phys. Rev. Lett. **123**, 252501 (2019)



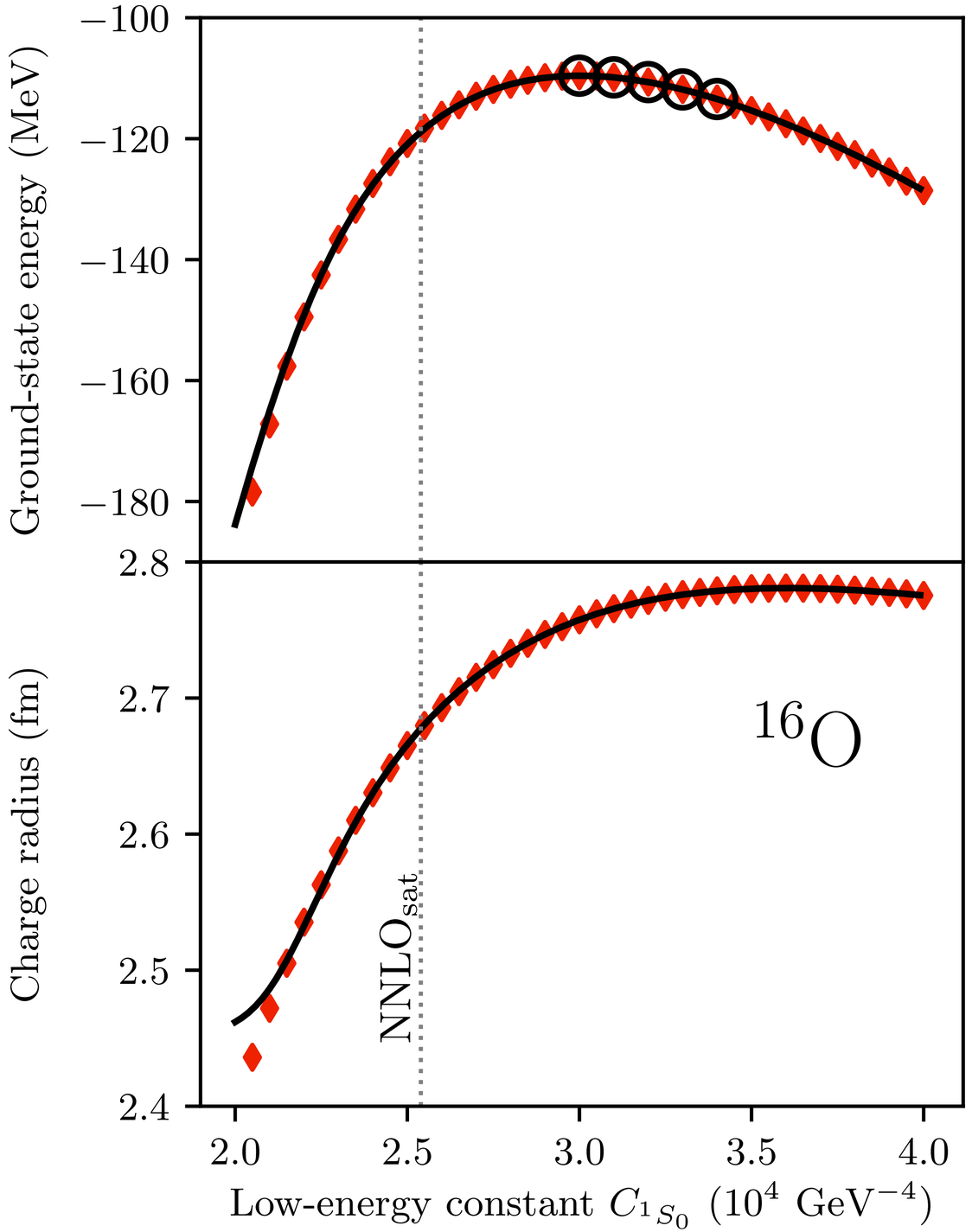
Exact coupled cluster calculations at the singles and doubles level

SPCC



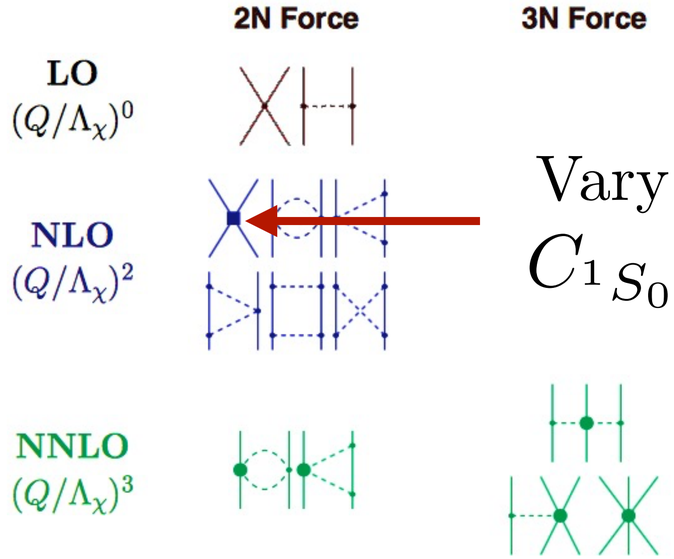
Vary only C_{1S_0}

A. Ekström and G. Hagen Phys. Rev. Lett. **123**, 252501 (2019)



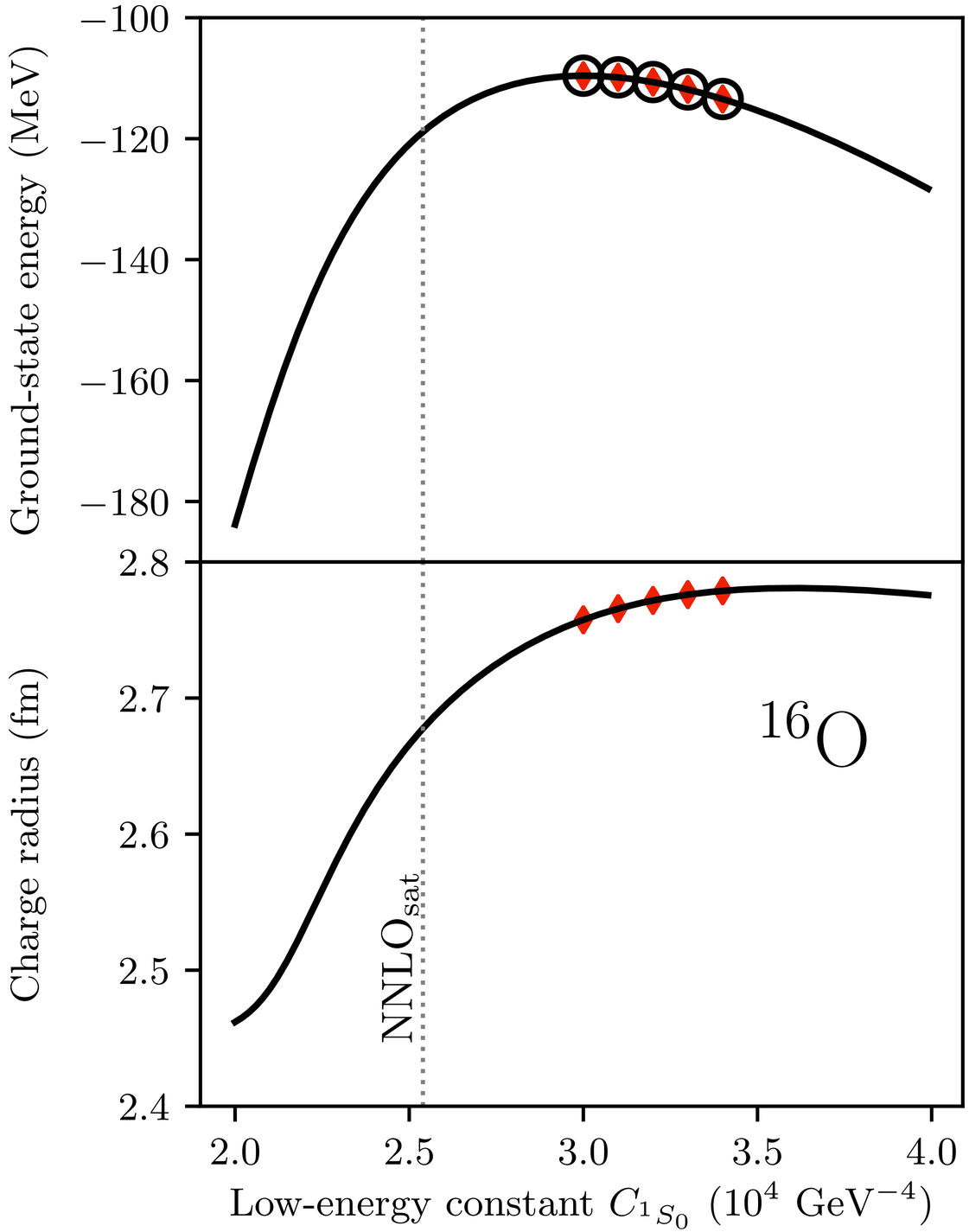
Exact coupled cluster calculations at the singles and doubles level

SPCC



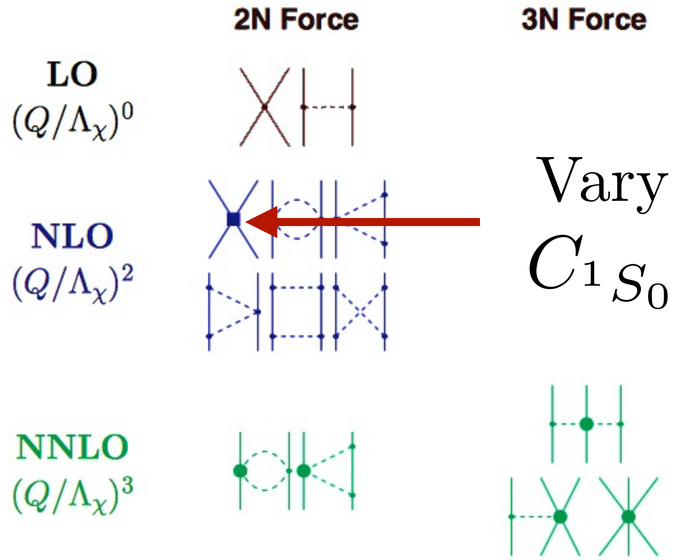
Vary only C_{1S_0}

A. Ekström and G. Hagen Phys. Rev. Lett. **123**, 252501 (2019)



Exact coupled cluster calculations at the singles and doubles level

SPCC



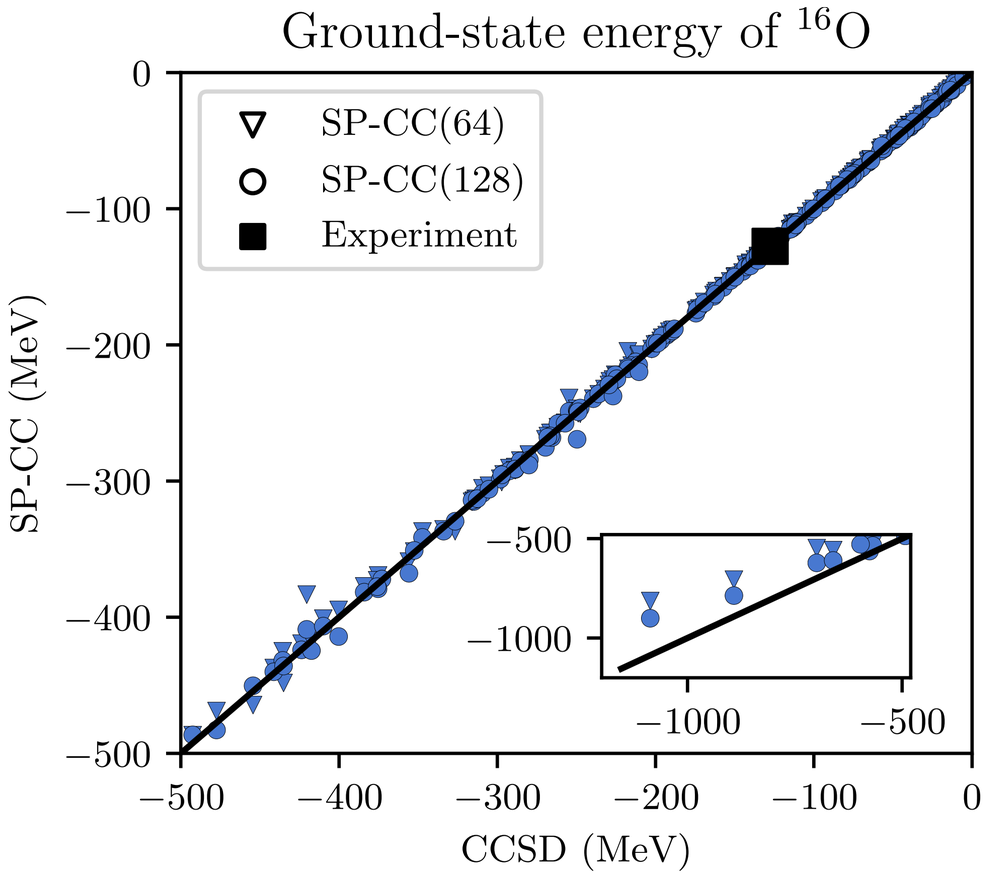
Vary only C_{1S_0}

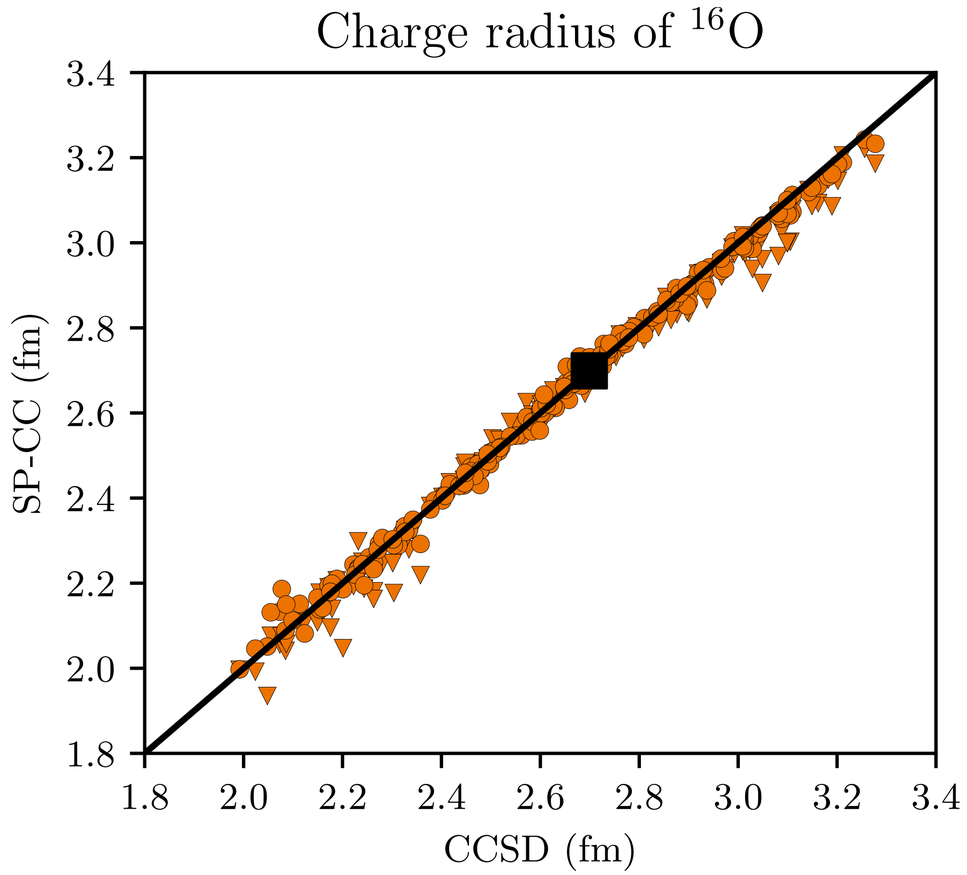
A. Ekström and G. Hagen Phys. Rev. Lett. **123**, 252501 (2019)

The EC emulators are accurate

$$y_{\mathrm{exp}}(\vec{x}) = \tilde{y}_{\mathrm{th}}(\vec{\alpha};\vec{x}) + \delta\tilde{y}_{\mathrm{th}}(\vec{\alpha};\vec{x}) + \delta y_{\mathrm{th}}(\vec{\alpha};\vec{x}) + \delta y_{\mathrm{exp}}(\vec{x})$$

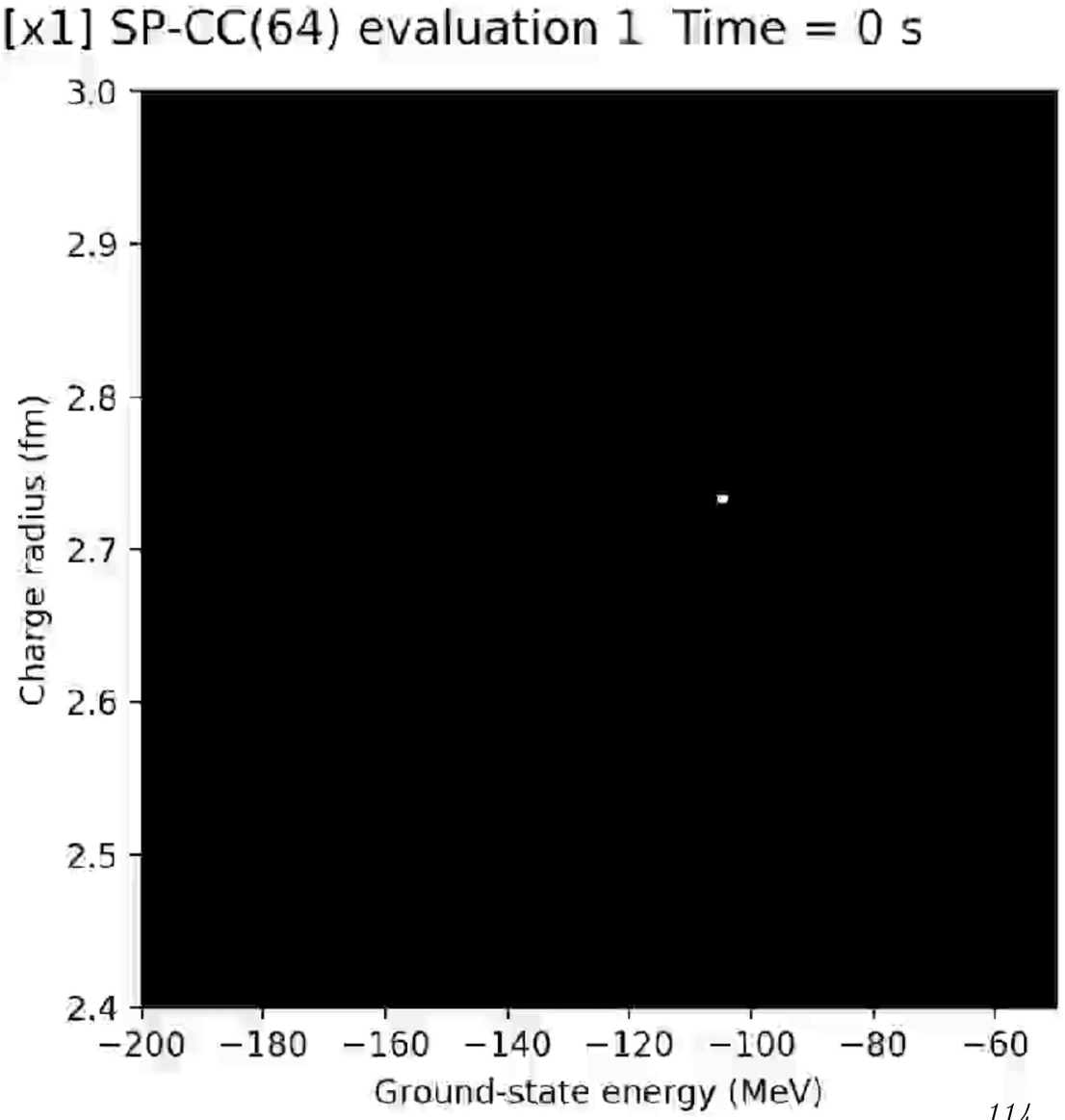
emulator emulator-error





A. Ekström and G. Hagen Phys. Rev. Lett. **123**, 252501 (2019)

The EC emulators are fast



This animation demonstrates the real-time speed and accuracy of emulated predictions for the radius & energy of the ground-state in oxygen-16 for different values of the interaction parameters (the LECs).

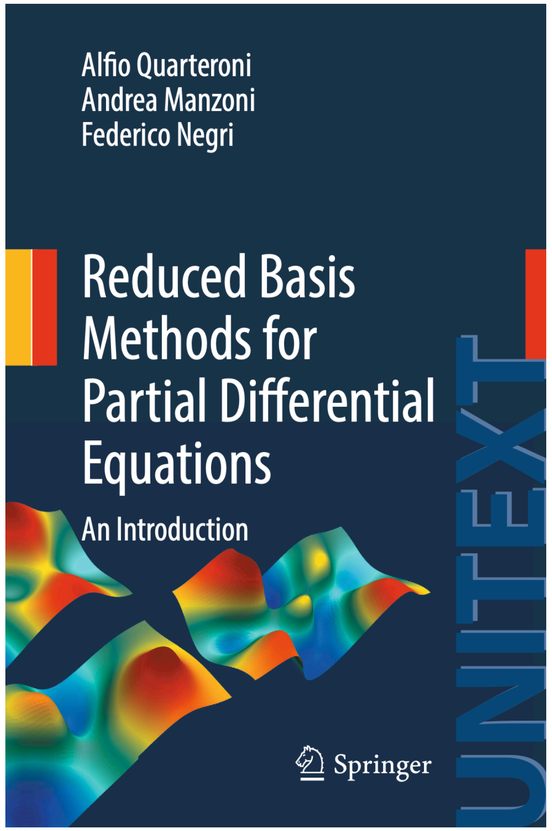
Accuracy: roughly the pixel size.

Speedup: 20-years of single node computation can be replaced by a 1 hour run on a laptop.

A. Ekström and G. Hagen Phys. Rev. Lett. **123**, 252501 (2019)

Model order reduction

- EC; a reduced basis method (RBM)



REVIEWS OF MODERN PHYSICS, VOLUME 96, JULY–SEPTEMBER 2024

Colloquium: **Eigenvector continuation and projection-based emulators**

Thomas Duguet
IRFU, CEA, Université Paris-Saclay, 91191 Gif-sur-Yvette, France and KU Leuven, Department of Physics and Astronomy, Instituut voor Kern- en Stralingsfysica, 3001 Leuven, Belgium

Andreas Ekström
Department of Physics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden

Richard J. Furnstahl
Department of Physics, The Ohio State University, Columbus, Ohio 43210, USA

Sebastian König
Department of Physics, North Carolina State University, Raleigh, North Carolina 27695, USA

Dean Lee
Facility for Rare Isotope Beams and Department of Physics and Astronomy, Michigan State University, Michigan 48824, USA

(published 14 August 2024)

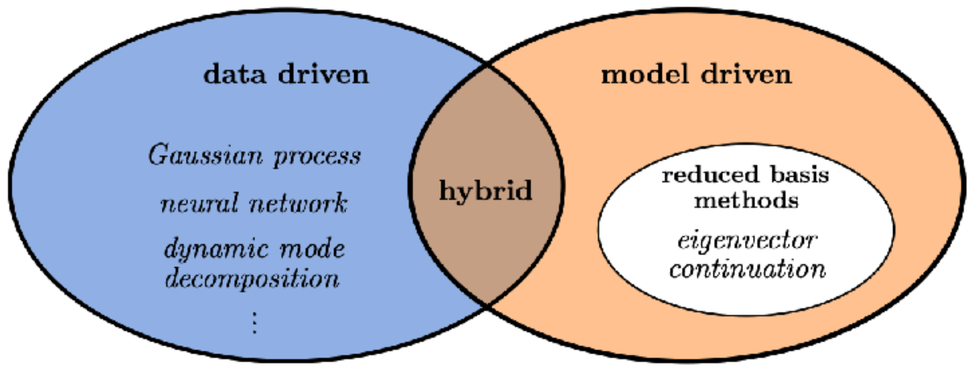
Eigenvector continuation is a computational method for parametric eigenvalue problems that uses subspace projection with a basis derived from eigenvector snapshots from different parameter sets. It is part of a broader class of subspace-projection techniques called reduced-basis methods. In this Colloquium, the development, theory, and applications of eigenvector continuation and projection-based emulators are presented. The basic concepts are introduced, the underlying theory and convergence properties are discussed, and recent applications for quantum systems and future prospects are presented.

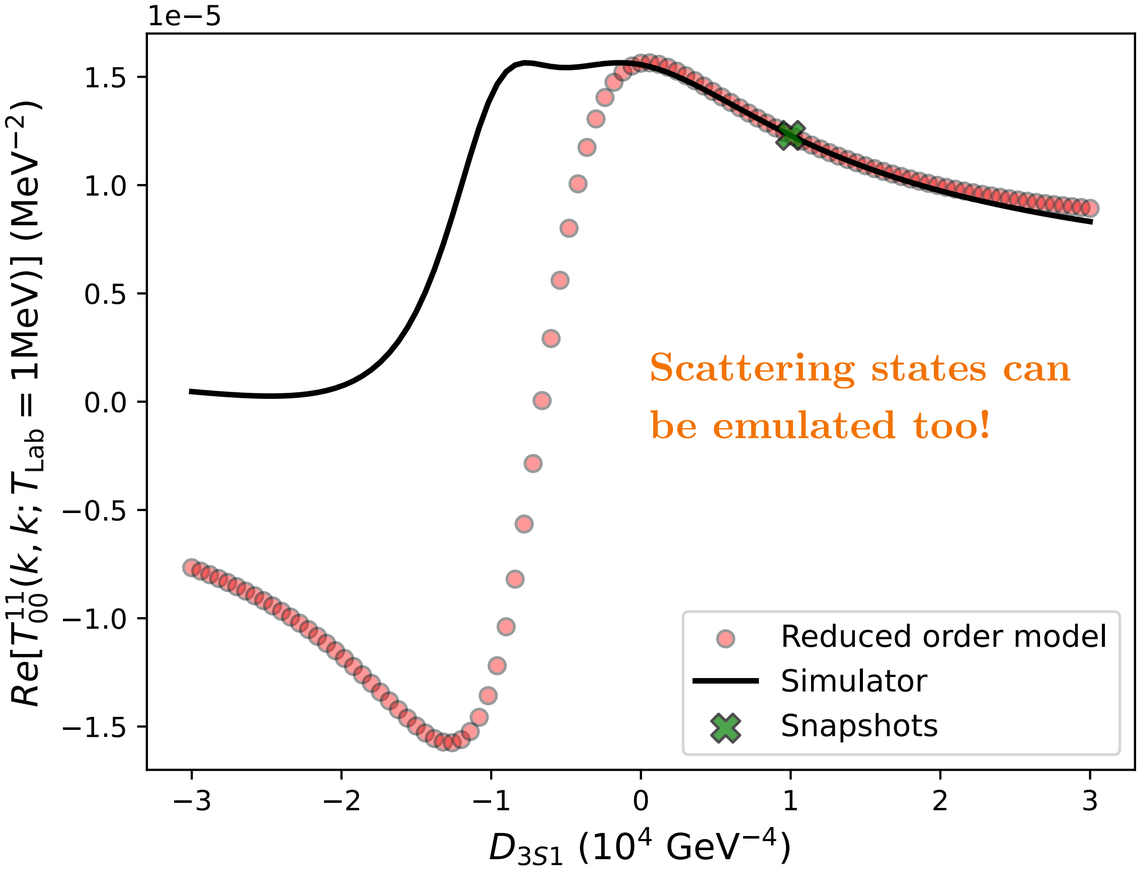
DOI: 10.1103/RevModPhys.96.031002

CONTENTS

I. MOTIVATION

The challenges of nuclear few- and many-body physics have been addressed theoretically with a wide range of accurate but often computationally expensive *high fidelity* methods. However, when we need to change the parameters characterizing the problem, such as Hamiltonian coupling constants, it can become computationally prohibitive to repeat high-fidelity calculations many times and challenging to reliably extrapolate. An alternative is to replace the high-fidelity model with an *emulator*, which is an approximate computer model (set in the literature is sometimes referred to as a surrogate. We focus in this Colloquium on the recent development and application of emulators that exploit a technique called eigenvector continuation (EC) and its extensions. Our illustrative examples are primarily drawn from nuclear structure and reactions, for which there has been an explosion of EC applications in the last few years. We emphasize, however, the general scope of the methods, which are broadly applicable to physics problems.

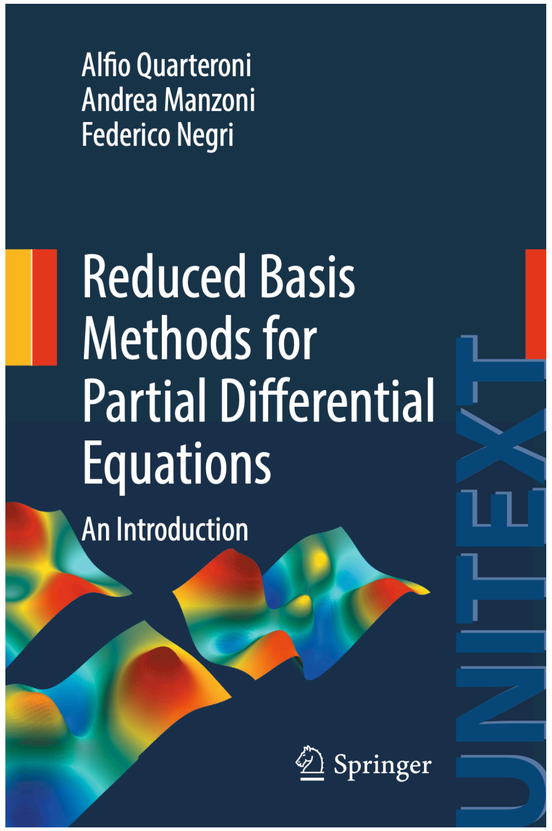




T. Duguet, A. Ekström, R. J. Furnstahl, S. König, D. Lee, Rev. Mod Phys **96**, 031002 (2024)

Model order reduction

- EC; a reduced basis method (RBM)




REVIEWS OF MODERN PHYSICS, VOLUME 96, JULY–SEPTEMBER 2024

Colloquium: Eigenvector continuation and projection-based emulators

Thomas Duguet
IRFU, CEA, Université Paris-Saclay, 91191 Gif-sur-Yvette, France and KU Leuven, Department of Physics and Astronomy, Instituut voor Kern- en Stralingsfysica, 3001 Leuven, Belgium

Andreas Ekström
Department of Physics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden

Richard J. Furnstahl
Department of Physics, The Ohio State University, Columbus, Ohio 43210, USA

Sebastian König
Department of Physics, North Carolina State University, Raleigh, North Carolina 27695, USA

Dean Lee
Facility for Rare Isotope Beams and Department of Physics and Astronomy, Michigan State University, Michigan 48824, USA

(published 14 August 2024)

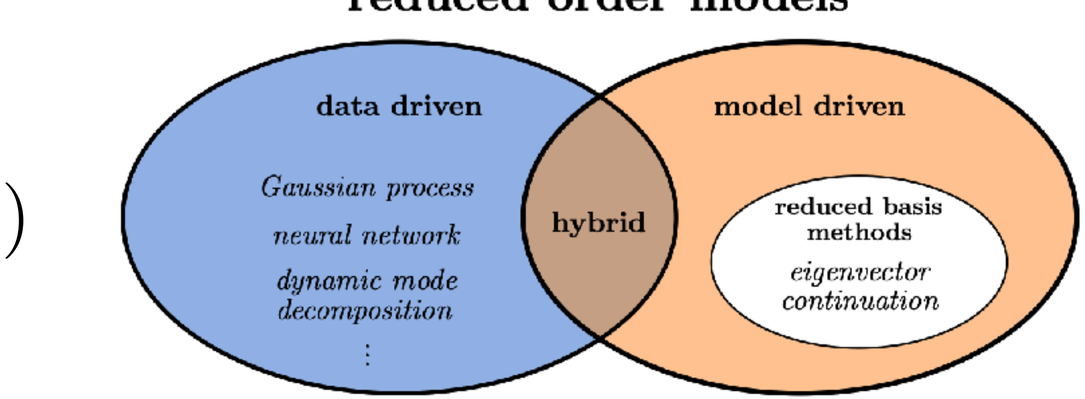
Eigenvector continuation is a computational method for parametric eigenvalue problems that uses subspace projection with a basis derived from eigenvector snapshots from different parameter sets. It is part of a broader class of subspace-projection techniques called reduced-basis methods. In this Colloquium, the development, theory, and applications of eigenvector continuation and projection-based emulators are presented. The basic concepts are introduced, the underlying theory and convergence properties are discussed, and recent applications for quantum systems and future prospects are presented.

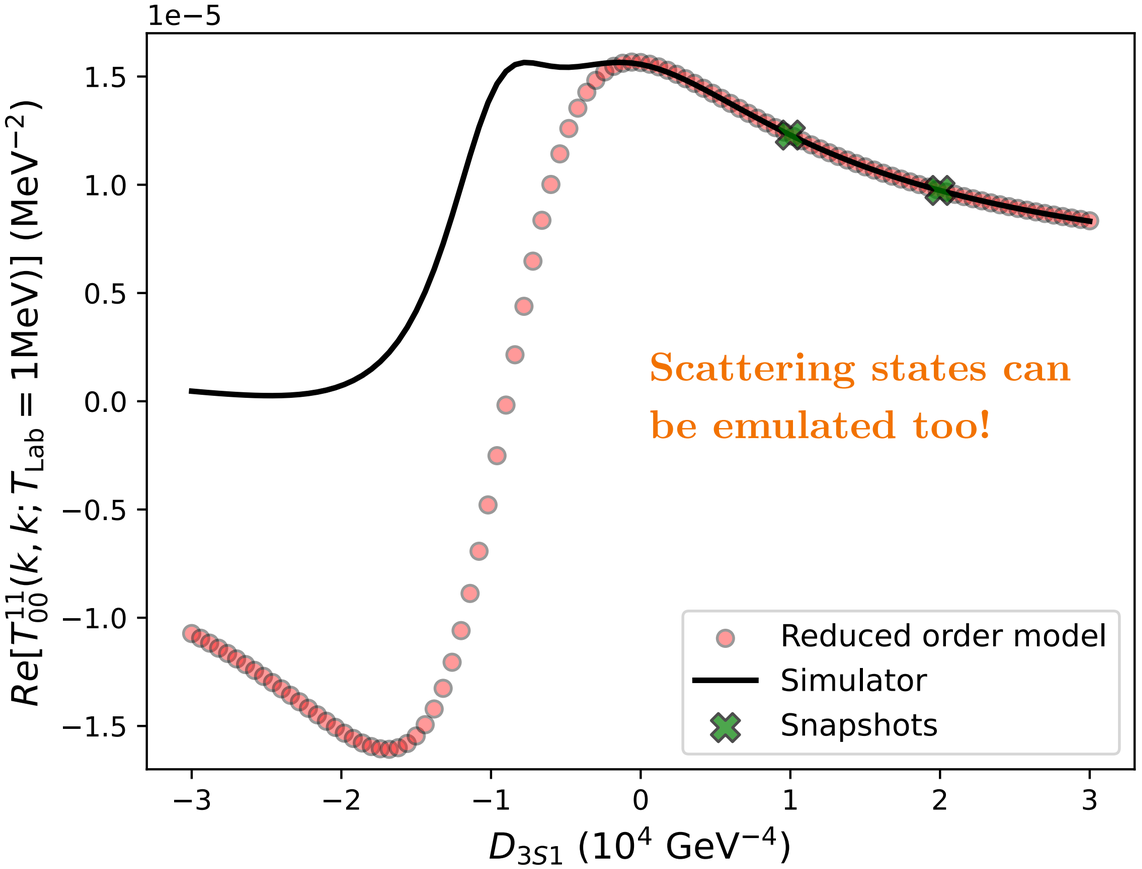
DOI: 10.1103/RevModPhys.96.031002

CONTENTS

I. MOTIVATION

The challenges of nuclear few- and many-body physics have been addressed theoretically with a wide range of accurate but often computationally expensive *high fidelity* methods. However, when we need to change the parameters characterizing the problem, such as Hamiltonian coupling constants, it can become computationally prohibitive to repeat high-fidelity calculations many times and challenging to reliably extrapolate. An alternative is to replace the high-fidelity model with an *emulator*, which is an approximate computer model (set in the literature is sometimes referred to as a surrogate. We focus in this Colloquium on the recent development and application of emulators that exploit a technique called eigenvector continuation (EC) and its extensions. Our illustrative examples are primarily drawn from nuclear structure and reactions, for which there has been an explosion of EC applications in the last few years. We emphasize, however, the general scope of the methods, which are broadly applicable to physics problems.

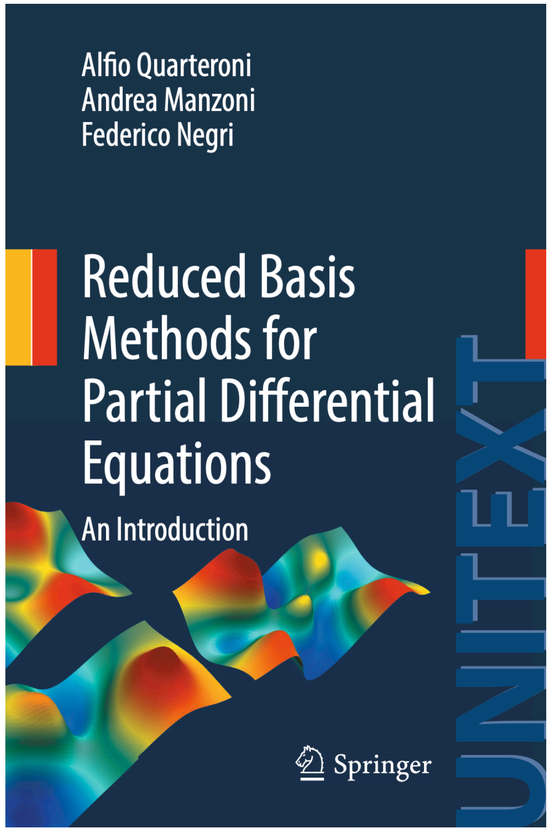




T. Duguet, A. Ekström, R. J. Furnstahl, S. König, D. Lee, Rev. Mod Phys **96**, 031002 (2024)

Model order reduction

- EC; a reduced basis method (RBM)




REVIEWS OF MODERN PHYSICS, VOLUME 96, JULY–SEPTEMBER 2024

Colloquium: **Eigenvector continuation and projection-based emulators**

Thomas Duguet
IRFU, CEA, Université Paris-Saclay, 91191 Gif-sur-Yvette, France and KU Leuven, Department of Physics and Astronomy, Instituut voor Kern- en Stralingsfysica, 3001 Leuven, Belgium

Andreas Ekström
Department of Physics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden

Richard J. Furnstahl
Department of Physics, The Ohio State University, Columbus, Ohio 43210, USA

Sebastian König
Department of Physics, North Carolina State University, Raleigh, North Carolina 27695, USA

Dean Lee
Facility for Rare Isotope Beams and Department of Physics and Astronomy, Michigan State University, Michigan 48824, USA

(published 14 August 2024)

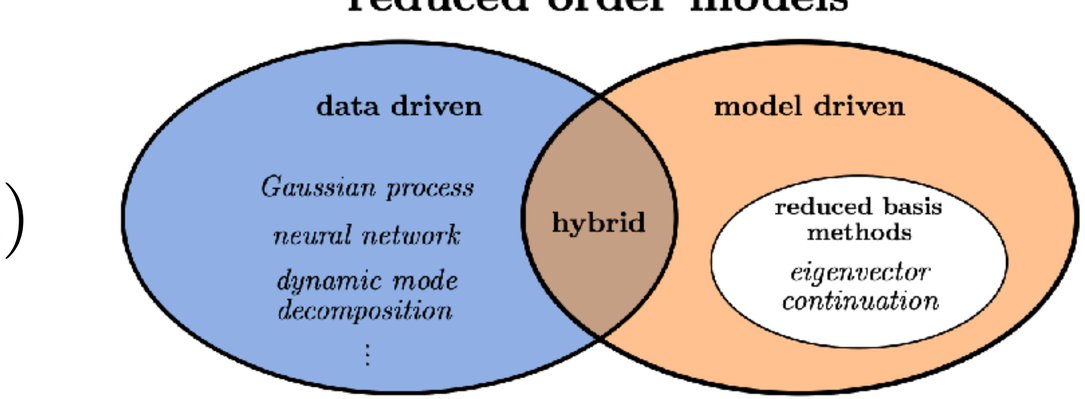
Eigenvector continuation is a computational method for parametric eigenvalue problems that uses subspace projection with a basis derived from eigenvector snapshots from different parameter sets. It is part of a broader class of subspace-projection techniques called reduced-basis methods. In this Colloquium, the development, theory, and applications of eigenvector continuation and projection-based emulators are presented. The basic concepts are introduced, the underlying theory and convergence properties are discussed, and recent applications for quantum systems and future prospects are presented.

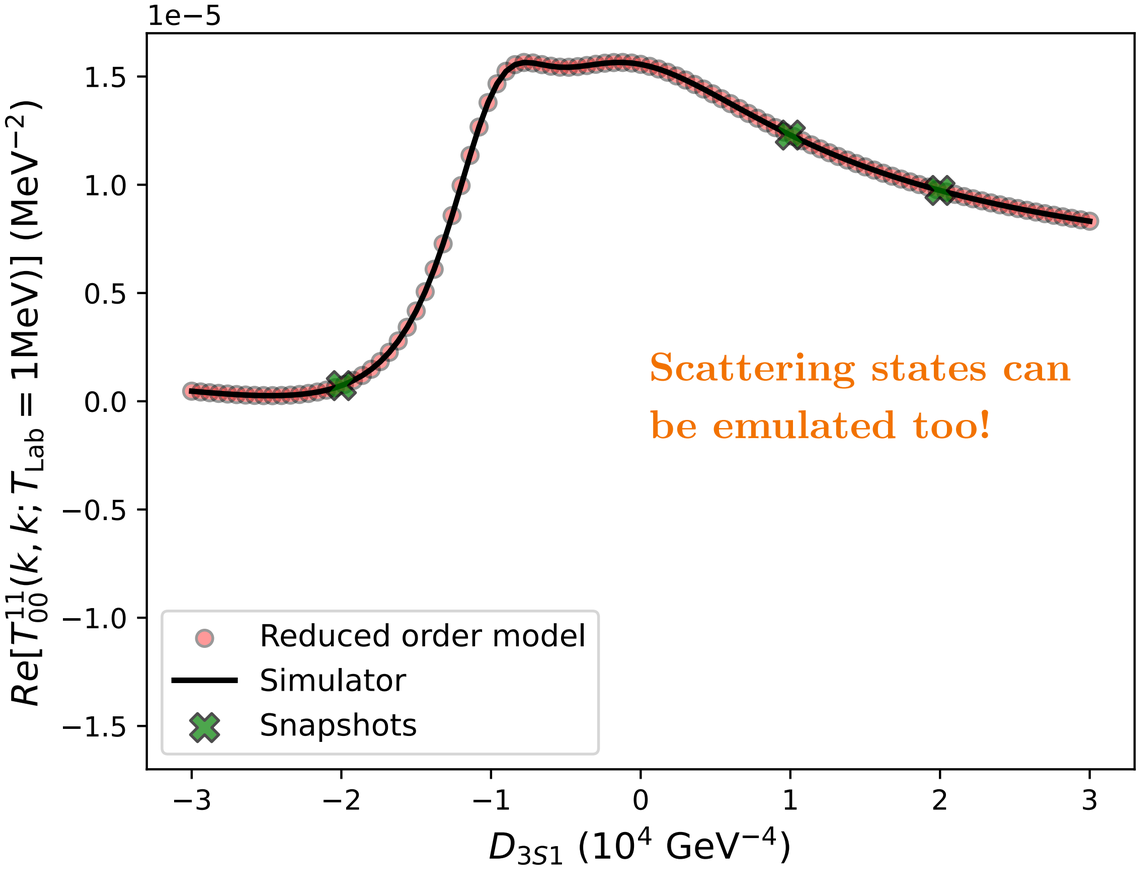
DOI: 10.1103/RevModPhys.96.031002

CONTENTS

I. MOTIVATION

The challenges of nuclear few- and many-body physics have been addressed theoretically with a wide range of accurate but often computationally expensive *high fidelity* methods. However, when we need to change the parameters characterizing the problem, such as Hamiltonian coupling constants, it can become computationally prohibitive to repeat high-fidelity calculations many times and challenging to reliably extrapolate. An alternative is to replace the high-fidelity model with an *emulator*, which is an approximate computer model (set in the literature is sometimes referred to as a surrogate. We focus in this Colloquium on the recent development and application of emulators that exploit a technique called eigenvector continuation (EC) and its extensions. Our illustrative examples are primarily drawn from nuclear structure and reactions, for which there has been an explosion of EC applications in the last few years. We emphasize, however, the general scope of the methods, which are broadly applicable to physics problems.

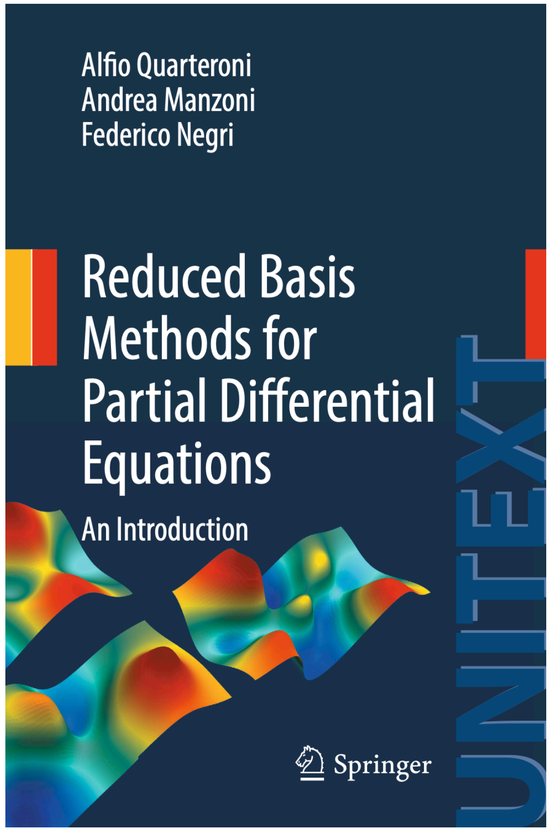




T. Duguet, A. Ekström, R. J. Furnstahl, S. König, D. Lee, Rev. Mod Phys **96**, 031002 (2024)

Model order reduction

- EC; a reduced basis method (RBM)




REVIEWS OF MODERN PHYSICS, VOLUME 96, JULY–SEPTEMBER 2024

Colloquium: Eigenvector continuation and projection-based emulators

Thomas Duguet
IRFU, CEA, Université Paris-Saclay, 91191 Gif-sur-Yvette, France and KU Leuven, Department of Physics and Astronomy, Instituut voor Kern- en Stralingsfysica, 3001 Leuven, Belgium

Andreas Ekström
Department of Physics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden

Richard J. Furnstahl
Department of Physics, The Ohio State University, Columbus, Ohio 43210, USA

Sebastian König
Department of Physics, North Carolina State University, Raleigh, North Carolina 27695, USA

Dean Lee
Facility for Rare Isotope Beams and Department of Physics and Astronomy, Michigan State University, Michigan 48824, USA

(published 14 August 2024)

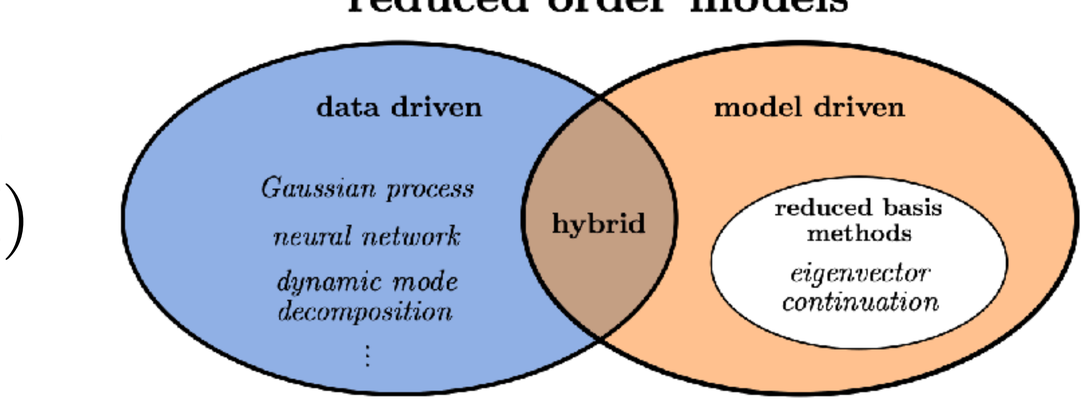
Eigenvector continuation is a computational method for parametric eigenvalue problems that uses subspace projection with a basis derived from eigenvector snapshots from different parameter sets. It is part of a broader class of subspace-projection techniques called reduced-basis methods. In this Colloquium, the development, theory, and applications of eigenvector continuation and projection-based emulators are presented. The basic concepts are introduced, the underlying theory and convergence properties are discussed, and recent applications for quantum systems and future prospects are presented.

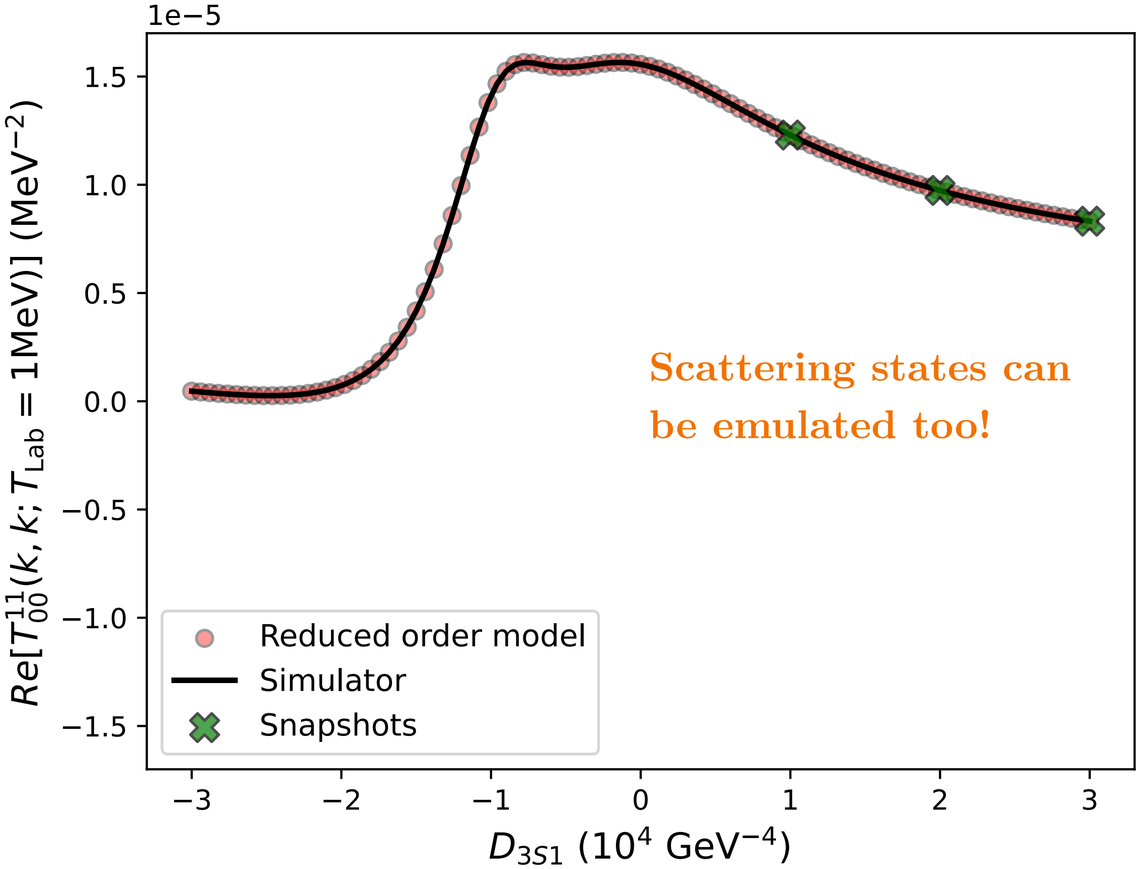
DOI: 10.1103/RevModPhys.96.031002

CONTENTS

I. MOTIVATION

The challenges of nuclear few- and many-body physics have been addressed theoretically with a wide range of accurate but often computationally expensive *high fidelity* methods. However, when we need to change the parameters characterizing the problem, such as Hamiltonian coupling constants, it can become computationally prohibitive to repeat high-fidelity calculations many times and challenging to reliably extrapolate. An alternative is to replace the high-fidelity model with an *emulator*, which is an approximate computer model (set in the literature is sometimes referred to as a surrogate. We focus in this Colloquium on the recent development and application of emulators that exploit a technique called eigenvector continuation (EC) and its extensions. Our illustrative examples are primarily drawn from nuclear structure and reactions, for which there has been an explosion of EC applications in the last few years. We emphasize, however, the general scope of the methods, which are broadly applicable to physics problems.

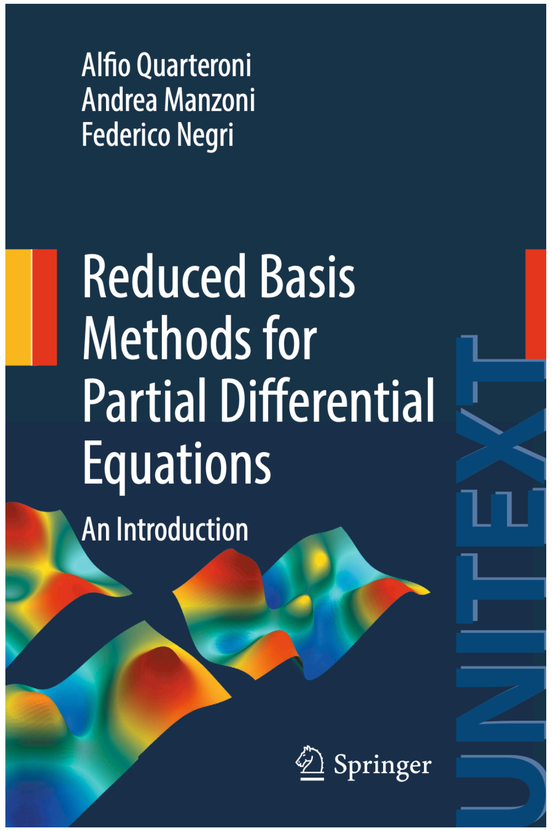




T. Duguet, A. Ekström, R. J. Furnstahl, S. König, D. Lee, Rev. Mod Phys **96**, 031002 (2024)

Model order reduction

- EC; a reduced basis method (RBM)




REVIEWS OF MODERN PHYSICS, VOLUME 96, JULY–SEPTEMBER 2024

Colloquium: **Eigenvector continuation and projection-based emulators**

Thomas Duguet
IRFU, CEA, Université Paris-Saclay, 91191 Gif-sur-Yvette, France and KU Leuven, Department of Physics and Astronomy, Instituut voor Kern- en Stralingsfysica, 3001 Leuven, Belgium

Andreas Ekström
Department of Physics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden

Richard J. Furnstahl
Department of Physics, The Ohio State University, Columbus, Ohio 43210, USA

Sebastian König
Department of Physics, North Carolina State University, Raleigh, North Carolina 27695, USA

Dean Lee
Facility for Rare Isotope Beams and Department of Physics and Astronomy, Michigan State University, Michigan 48824, USA

(published 14 August 2024)

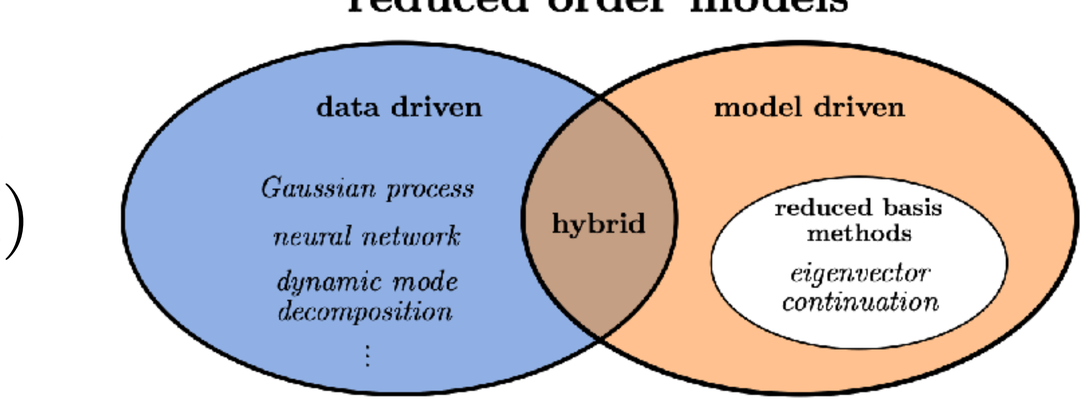
Eigenvector continuation is a computational method for parametric eigenvalue problems that uses subspace projection with a basis derived from eigenvector snapshots from different parameter sets. It is part of a broader class of subspace-projection techniques called reduced-basis methods. In this Colloquium, the development, theory, and applications of eigenvector continuation and projection-based emulators are presented. The basic concepts are introduced, the underlying theory and convergence properties are discussed, and recent applications for quantum systems and future prospects are presented.

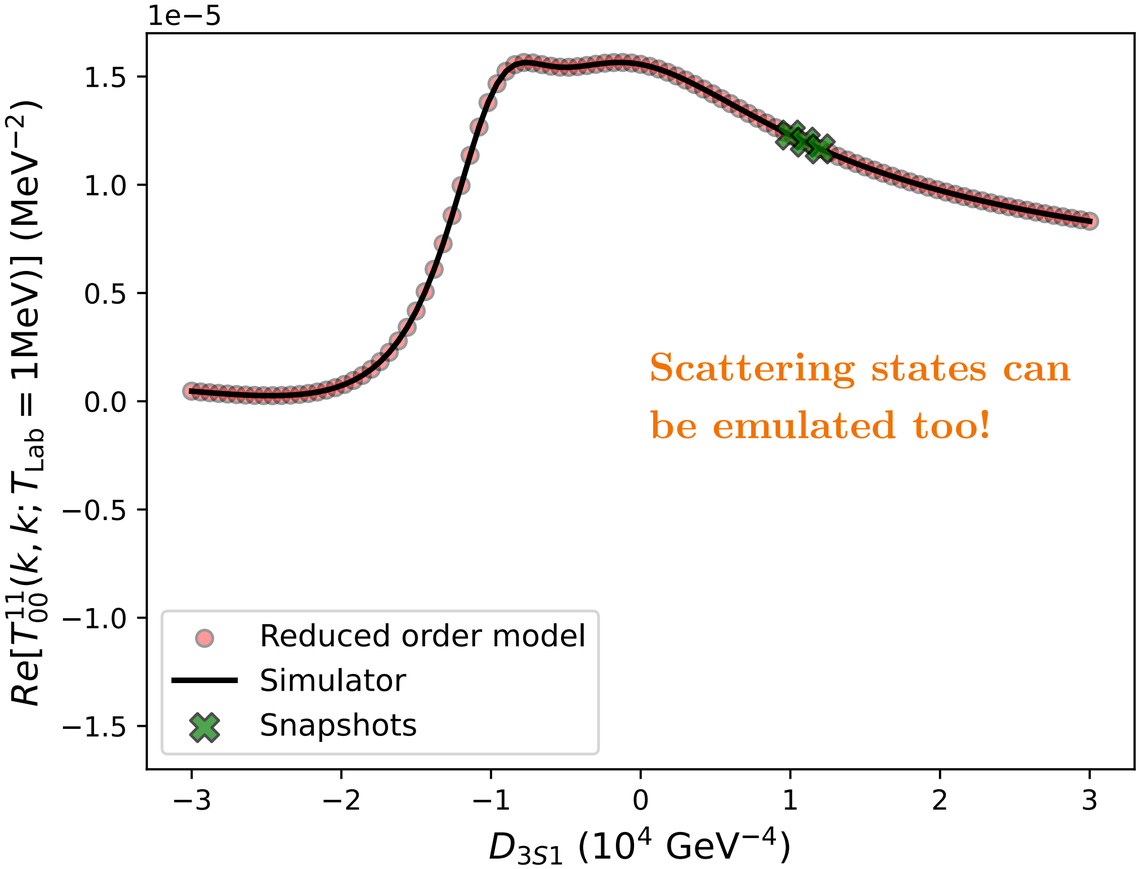
DOI: 10.1103/RevModPhys.96.031002

CONTENTS

I. MOTIVATION

The challenges of nuclear few- and many-body physics have been addressed theoretically with a wide range of accurate but often computationally expensive *high fidelity* methods. However, when we need to change the parameters characterizing the problem, such as Hamiltonian coupling constants, it can become computationally prohibitive to repeat high-fidelity calculations many times and challenging to reliably extrapolate. An alternative is to replace the high-fidelity model with an *emulator*, which is an approximate computer model (set in the literature is sometimes referred to as a surrogate. We focus in this Colloquium on the recent development and application of emulators that exploit a technique called eigenvector continuation (EC) and its extensions. Our illustrative examples are primarily drawn from nuclear structure and reactions, for which there has been an explosion of EC applications in the last few years. We emphasize, however, the general scope of the methods, which are broadly applicable to physics problems.

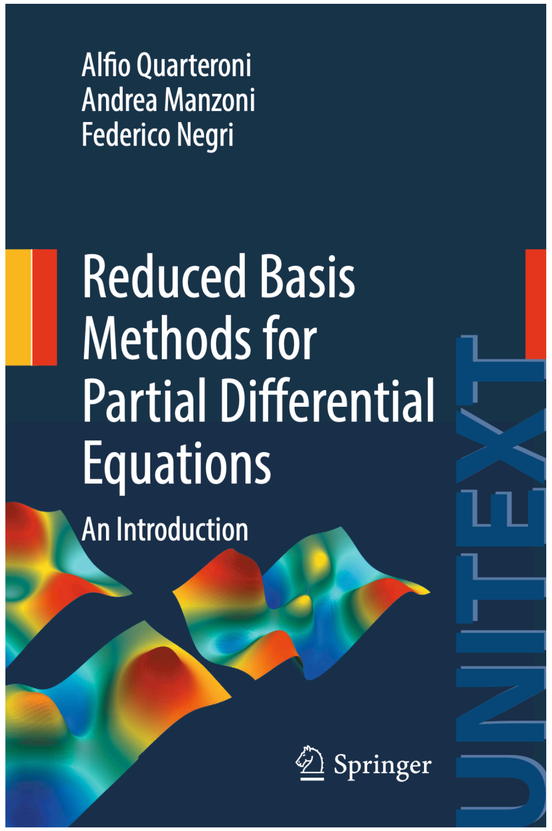




T. Duguet, A. Ekström, R. J. Furnstahl, S. König, D. Lee, Rev. Mod Phys **96**, 031002 (2024)

Model order reduction

- EC; a reduced basis method (RBM)



REVIEWS OF MODERN PHYSICS, VOLUME 96, JULY–SEPTEMBER 2024

Colloquium: **Eigenvector continuation and projection-based emulators**

Thomas Duguet
IRFU, CEA, Université Paris-Saclay, 91191 Gif-sur-Yvette, France and KU Leuven, Department of Physics and Astronomy, Instituut voor Kern- en Stralingsfysica, 3001 Leuven, Belgium

Andreas Ekström
Department of Physics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden

Richard J. Furnstahl
Department of Physics, The Ohio State University, Columbus, Ohio 43210, USA

Sebastian König
Department of Physics, North Carolina State University, Raleigh, North Carolina 27695, USA

Dean Lee
Facility for Rare Isotope Beams and Department of Physics and Astronomy, Michigan State University, Michigan 48824, USA

(published 14 August 2024)

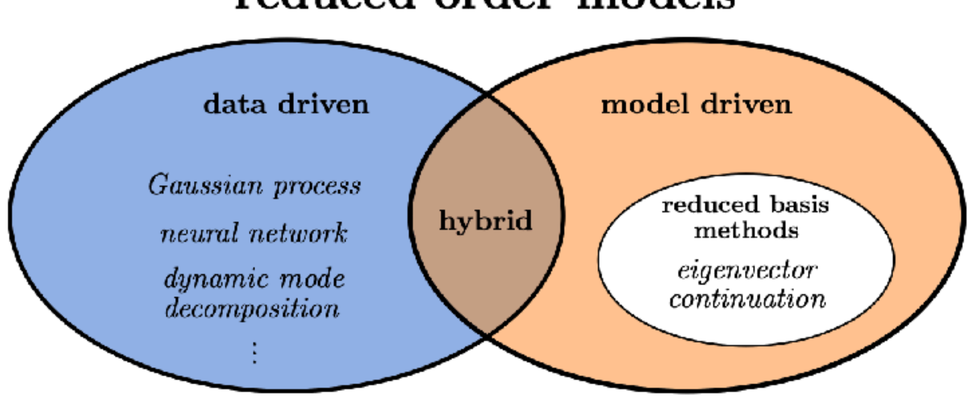
Eigenvector continuation is a computational method for parametric eigenvalue problems that uses subspace projection with a basis derived from eigenvector snapshots from different parameter sets. It is part of a broader class of subspace-projection techniques called reduced-basis methods. In this Colloquium, the development, theory, and applications of eigenvector continuation and projection-based emulators are presented. The basic concepts are introduced, the underlying theory and convergence properties are discussed, and recent applications for quantum systems and future prospects are presented.

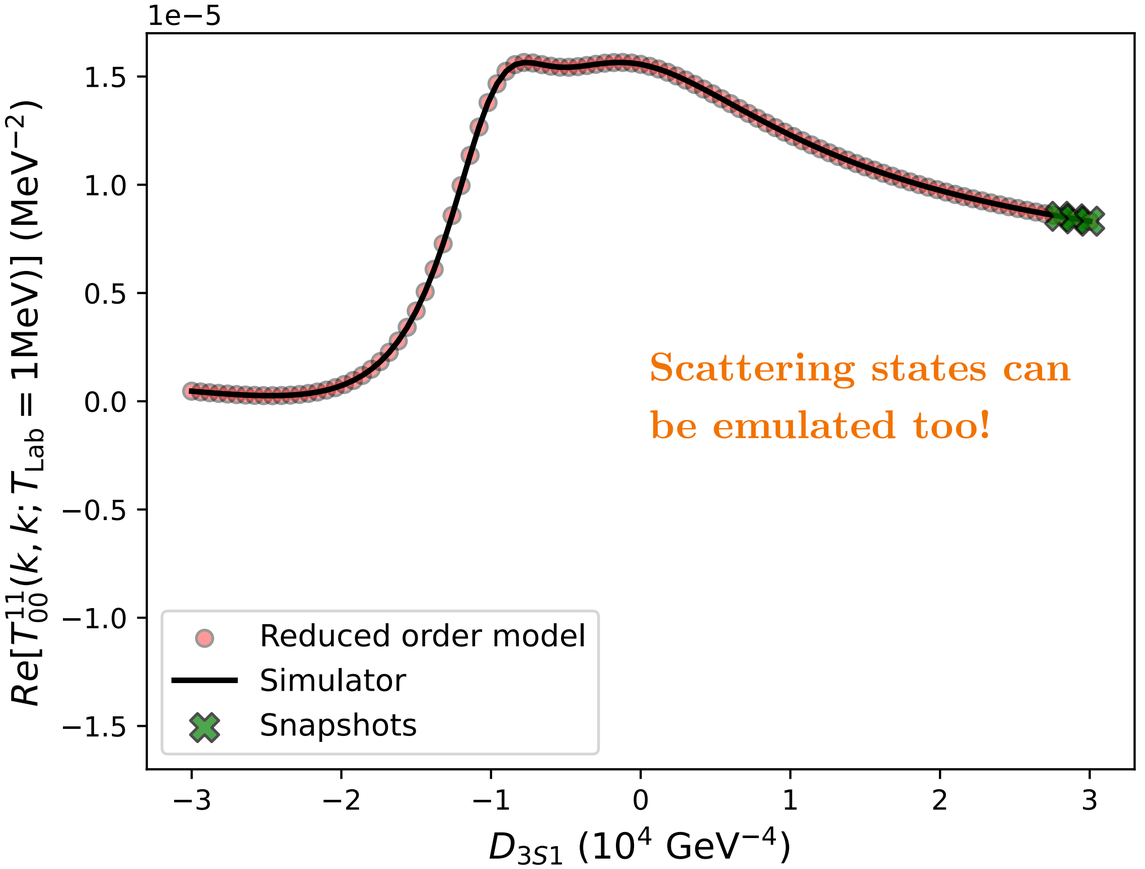
DOI: 10.1103/RevModPhys.96.031002

CONTENTS

I. MOTIVATION

The challenges of nuclear few- and many-body physics have been addressed theoretically with a wide range of accurate but often computationally expensive *high fidelity* methods. However, when we need to change the parameters characterizing the problem, such as Hamiltonian coupling constants, it can become computationally prohibitive to repeat high-fidelity calculations many times and challenging to reliably extrapolate. An alternative is to replace the high-fidelity model with an *emulator*, which is an approximate computer model (set in the literature is sometimes referred to as a surrogate. We focus in this Colloquium on the recent development and application of emulators that exploit a technique called eigenvector continuation (EC) and its extensions. Our illustrative examples are primarily drawn from nuclear structure and reactions, for which there has been an explosion of EC applications in the last few years. We emphasize, however, the general scope of the methods, which are broadly applicable to physics problems.

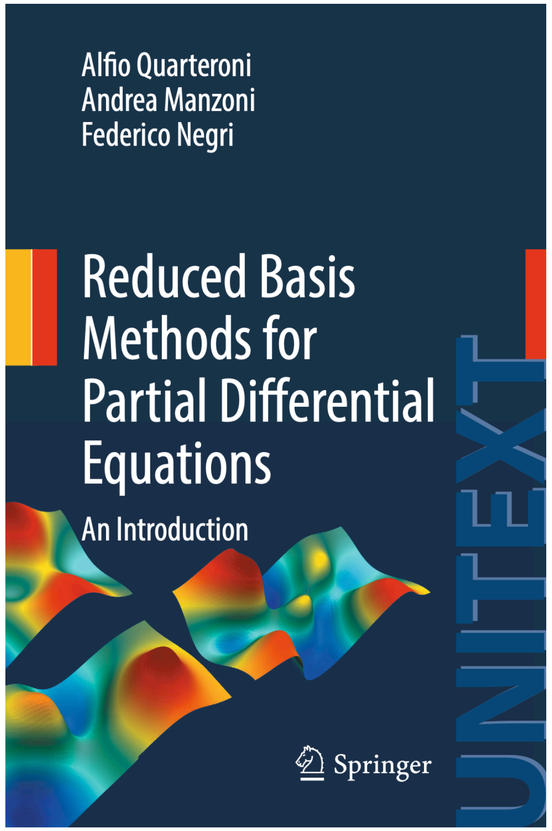




T. Duguet, A. Ekström, R. J. Furnstahl, S. König, D. Lee, Rev. Mod Phys **96**, 031002 (2024)

Model order reduction

- EC; a reduced basis method (RBM)



REVIEWS OF MODERN PHYSICS, VOLUME 96, JULY–SEPTEMBER 2024

Colloquium: Eigenvector continuation and projection-based emulators

Thomas Duguet
IRFU, CEA, Université Paris-Saclay, 91191 Gif-sur-Yvette, France and KU Leuven, Department of Physics and Astronomy, Instituut voor Kern- en Stralingsfysica, 3001 Leuven, Belgium

Andreas Ekström
Department of Physics, Chalmers University of Technology, SE-412 96 Göteborg, Sweden

Richard J. Furnstahl
Department of Physics, The Ohio State University, Columbus, Ohio 43210, USA

Sebastian König
Department of Physics, North Carolina State University, Raleigh, North Carolina 27695, USA

Dean Lee
Facility for Rare Isotope Beams and Department of Physics and Astronomy, Michigan State University, Michigan 48824, USA

(published 14 August 2024)

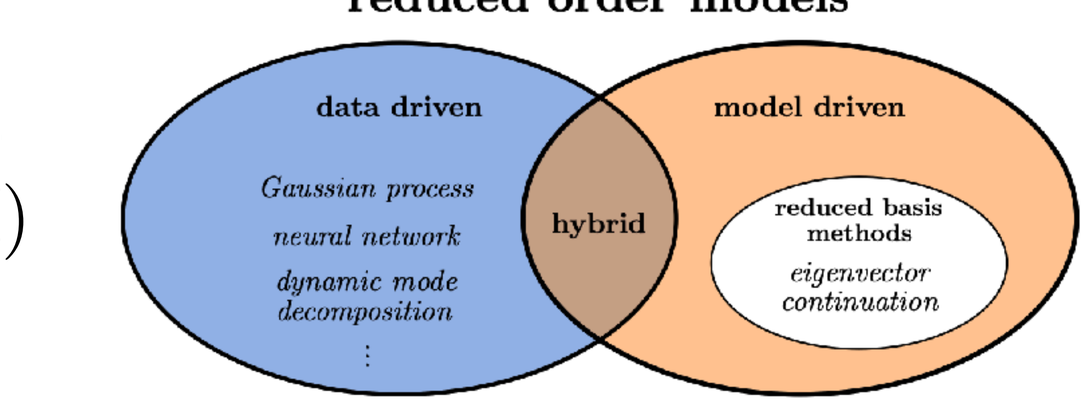
Eigenvector continuation is a computational method for parametric eigenvalue problems that uses subspace projection with a basis derived from eigenvector snapshots from different parameter sets. It is part of a broader class of subspace-projection techniques called reduced-basis methods. In this Colloquium, the development, theory, and applications of eigenvector continuation and projection-based emulators are presented. The basic concepts are introduced, the underlying theory and convergence properties are discussed, and recent applications for quantum systems and future prospects are presented.

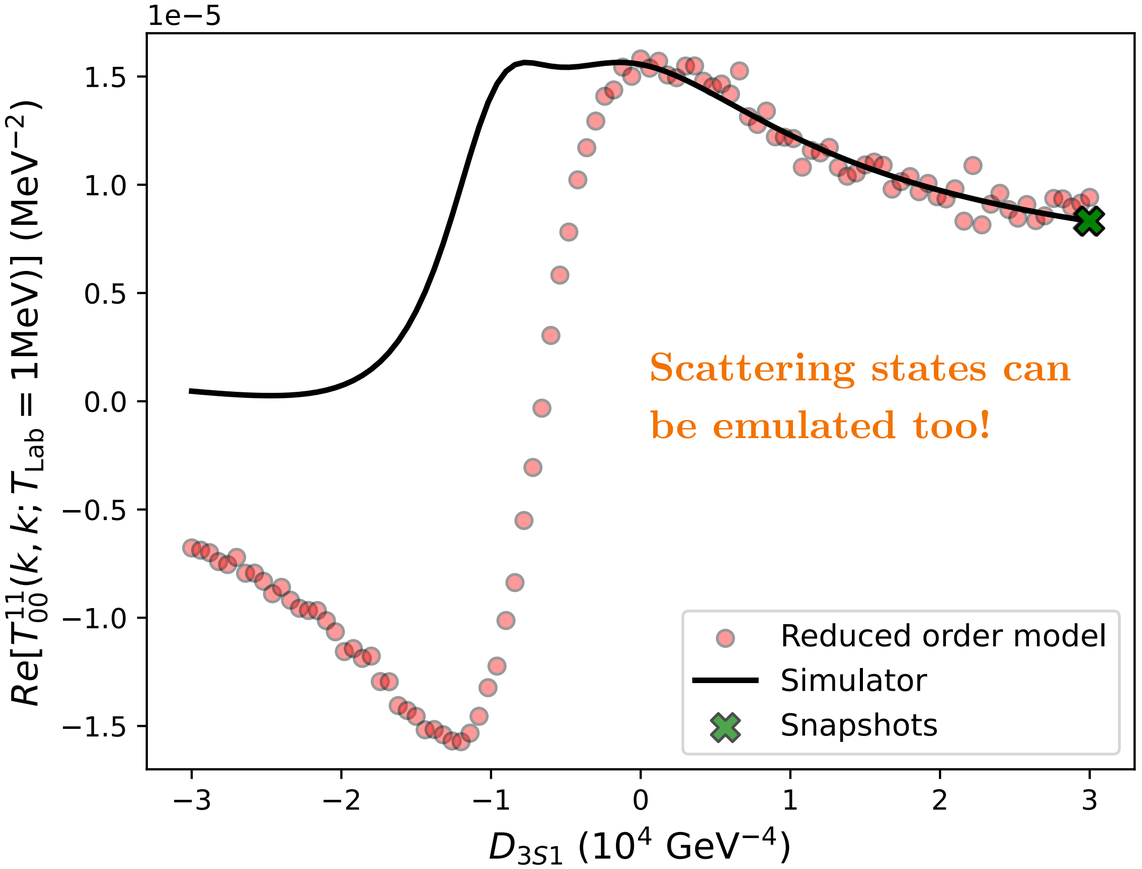
DOI: 10.1103/RevModPhys.96.031002

CONTENTS

I. MOTIVATION

The challenges of nuclear few- and many-body physics have been addressed theoretically with a wide range of accurate but often computationally expensive *high fidelity* methods. However, when we need to change the parameters characterizing the problem, such as Hamiltonian coupling constants, it can become computationally prohibitive to repeat high-fidelity calculations many times and challenging to reliably extrapolate. An alternative is to replace the high-fidelity model with an *emulator*, which is an approximate computer model (set in the literature is sometimes referred to as a surrogate. We focus in this Colloquium on the recent development and application of emulators that exploit a technique called eigenvector continuation (EC) and its extensions. Our illustrative examples are primarily drawn from nuclear structure and reactions, for which there has been an explosion of EC applications in the last few years. We emphasize, however, the general scope of the methods, which are broadly applicable to physics problems.

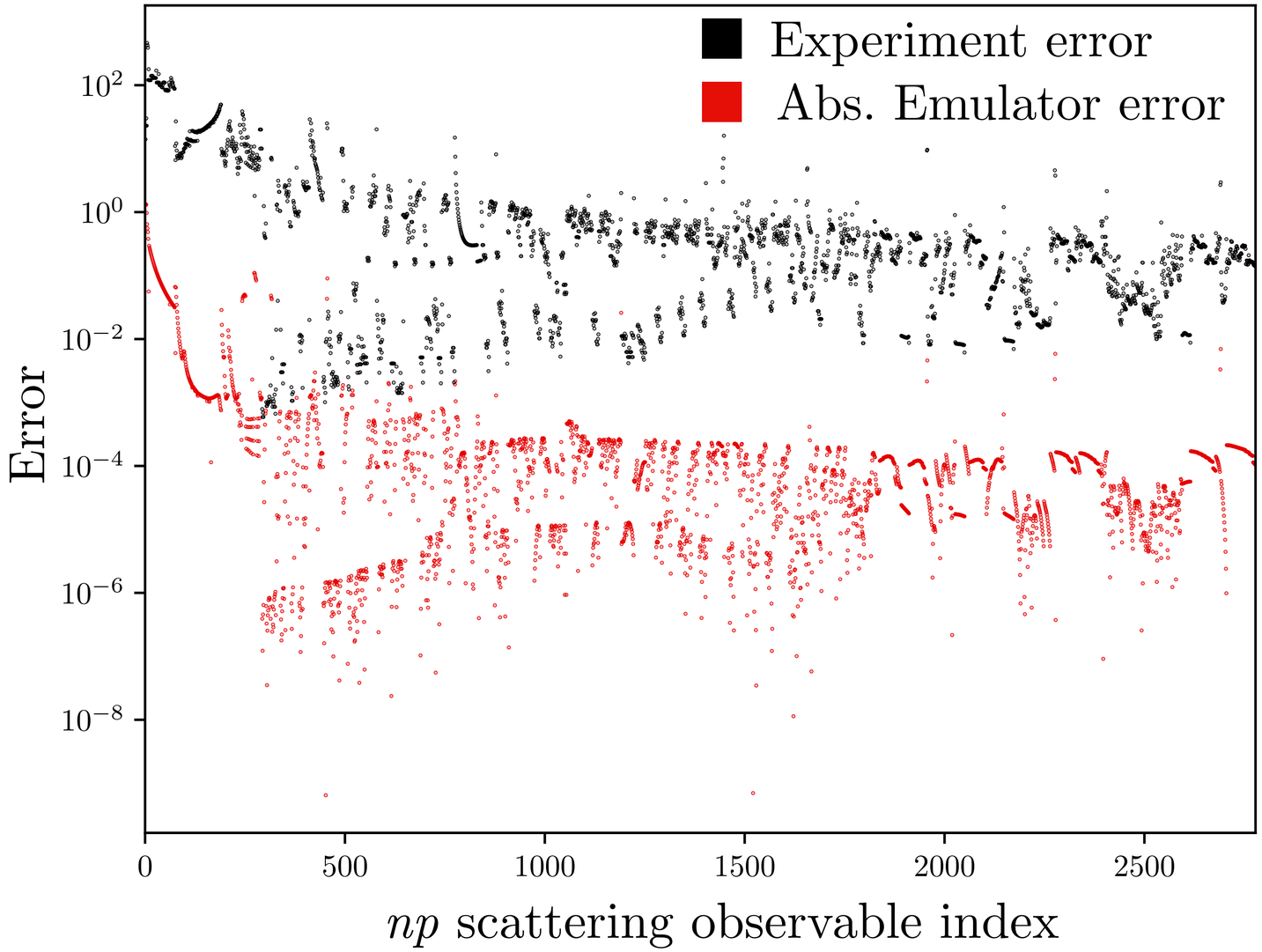




T. Duguet, A. Ekström, R. J. Furnstahl, S. König, D. Lee, Rev. Mod Phys **96**, 031002 (2024)

Benchmarking the *NN* scattering emulator

Deltafull NNLO np sector, 8 training points



It takes **0.4 seconds** to evaluate the entire likelihood.

Python implementation accelerated by Google JAX jit-compilation. Provides derivatives via AD as well. *(overhead x2)*

Svensson, Ekström, Forssen Phys Rev C (2024)

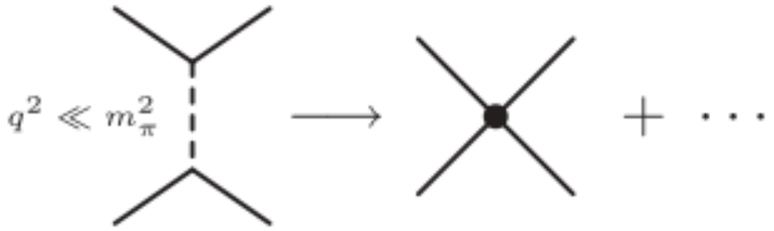
Next lecture

What insights can we gain from uncertainty quantification in *ab initio* nuclear physics?

How can we analyze the gap between theory predictions and experimental results?

Lecture 4
Selected applications

Inferring the breakdown scale of pionless EFT

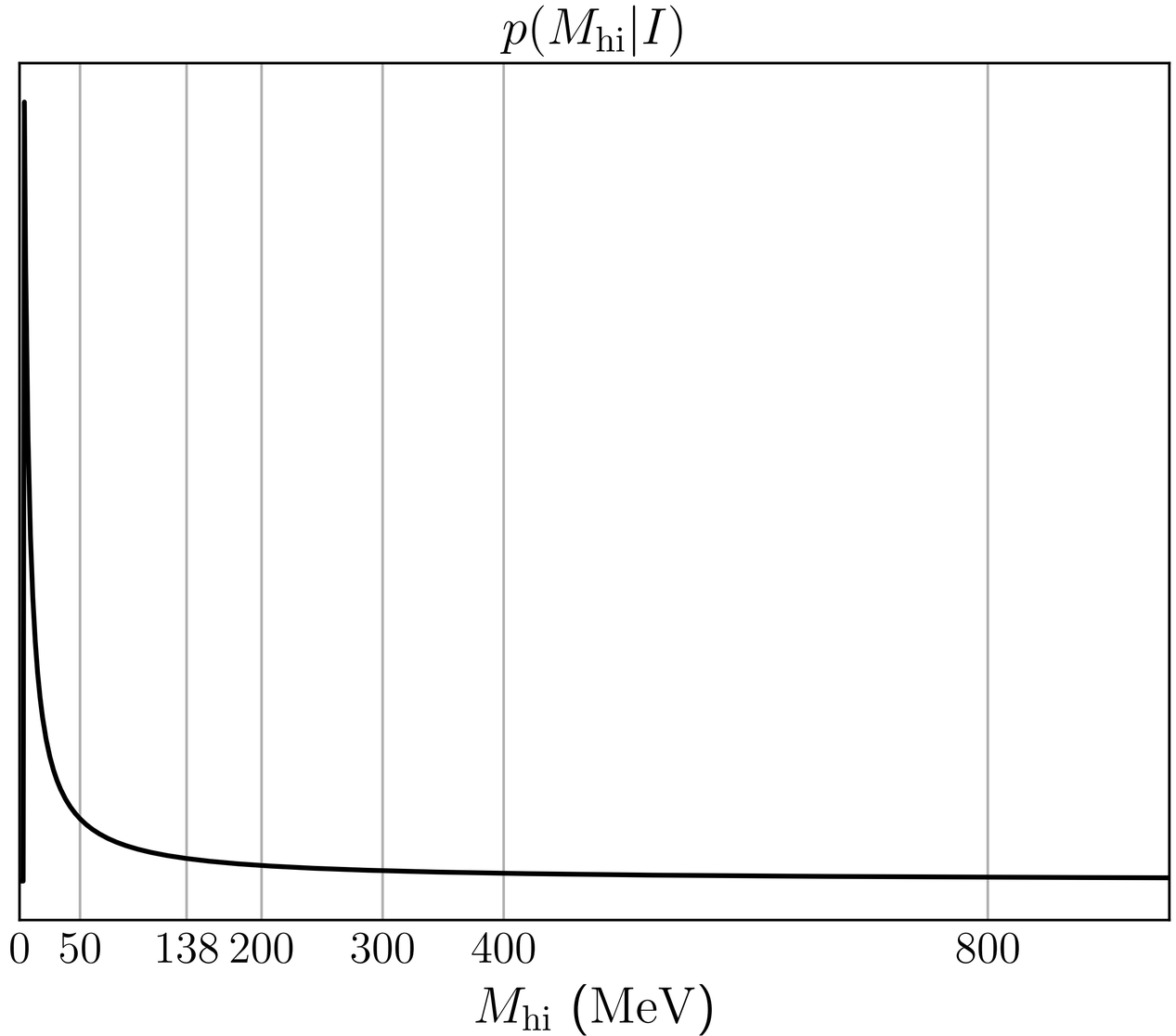


The breakdown scale is due to excluded massive degrees of freedom

$$y_{\rm th}^{(k)} = y_{\rm ref} \sum_{\nu=k+1}^{\infty} c_\nu \left(\frac{Q}{M_{\rm hi}} \right)^\nu$$

$$p(M_{\rm hi}|\boldsymbol{\sigma}, I) \propto p(\boldsymbol{\sigma}|M_{\rm hi}, I) p(M_{\rm hi}|I)$$ log-uniform prior

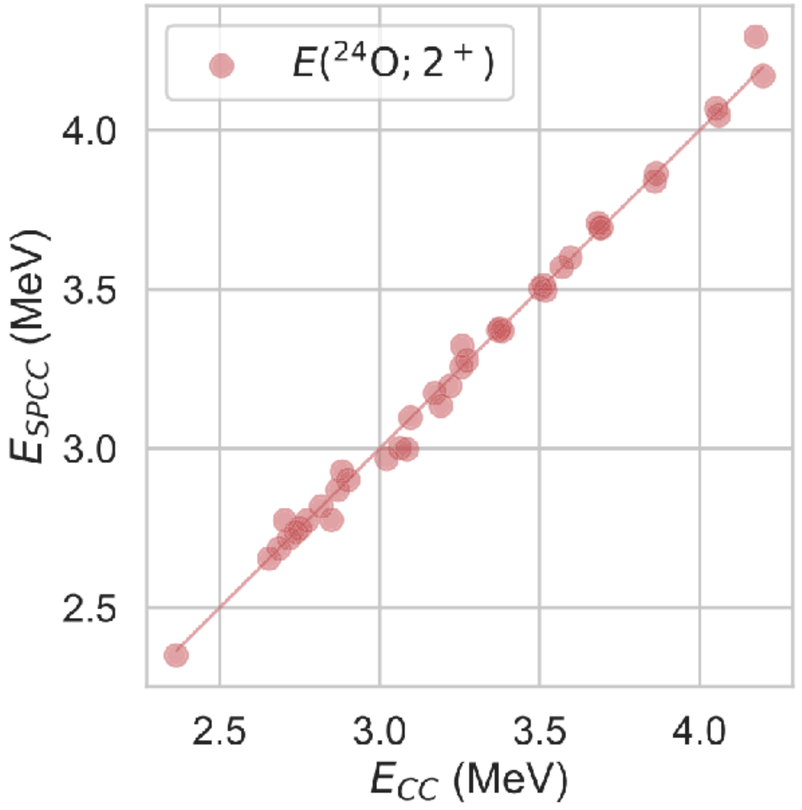
$$p(M_{\rm hi}|\boldsymbol{\sigma}, I) \propto p(M_{\rm hi}|I) \prod_{i=1}^{K} \left(\tau_i^{\nu_i} \prod_{n\in[1,2]} Q(k_i)^n \right)^{-1}$$



A. Ekström and L. Platter arXiv (2024) Melendez et al Phys Rev C (2019)

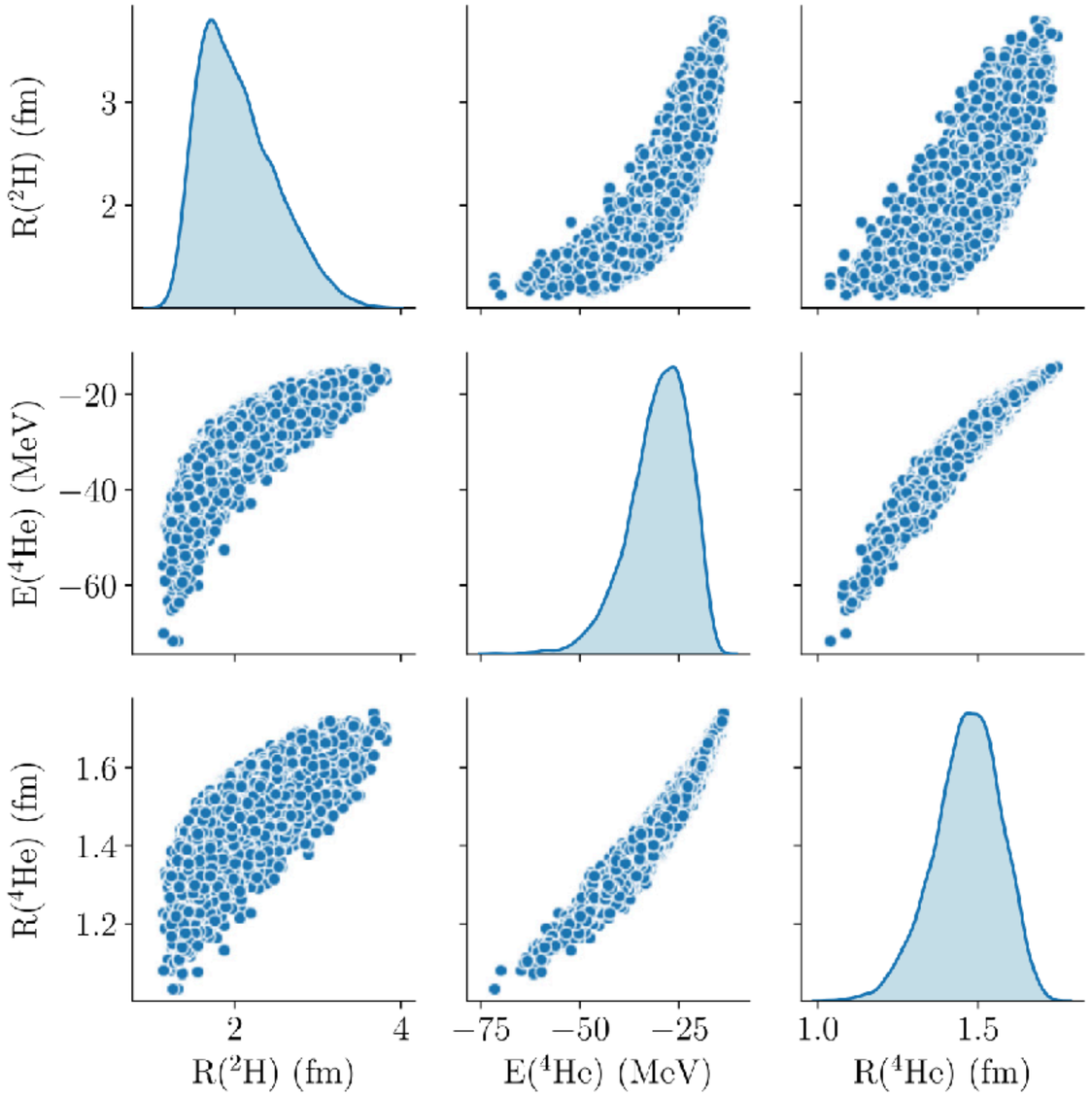
A computational statistics laboratory

- using fast and accurate emulators



CCSDT-3

EC for emulating no-core shell model predictions

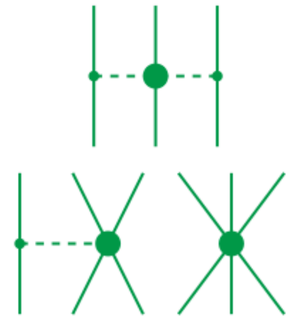


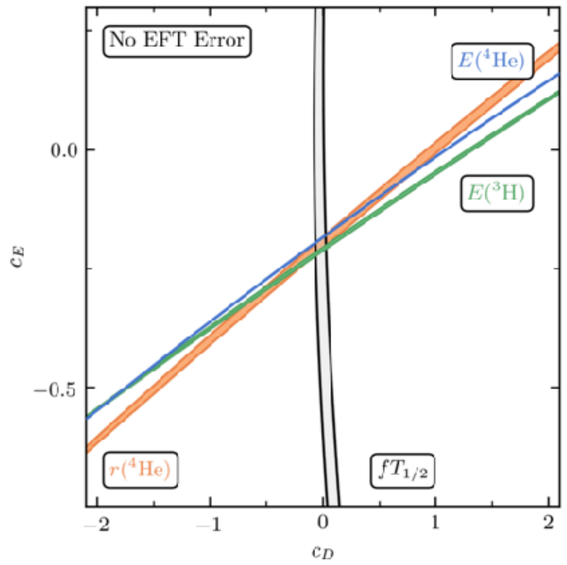
10^4 different $\vec{\alpha}$-values drawn randomly from a 10%-hypercube of the $\mathrm{NNLO_{sat}}$ interaction

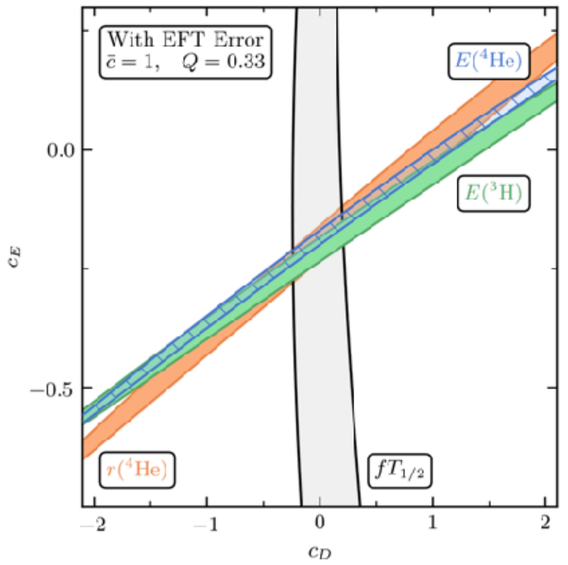
Less than 1 minute on a standard laptop

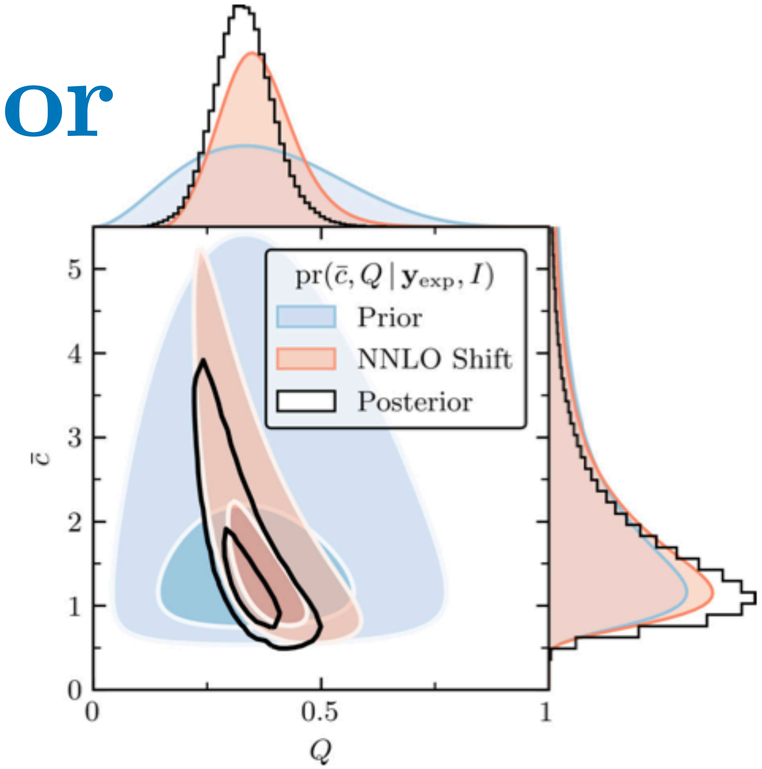
S. König, et al. Phys. Lett. **B 810**, 135814 (2020)

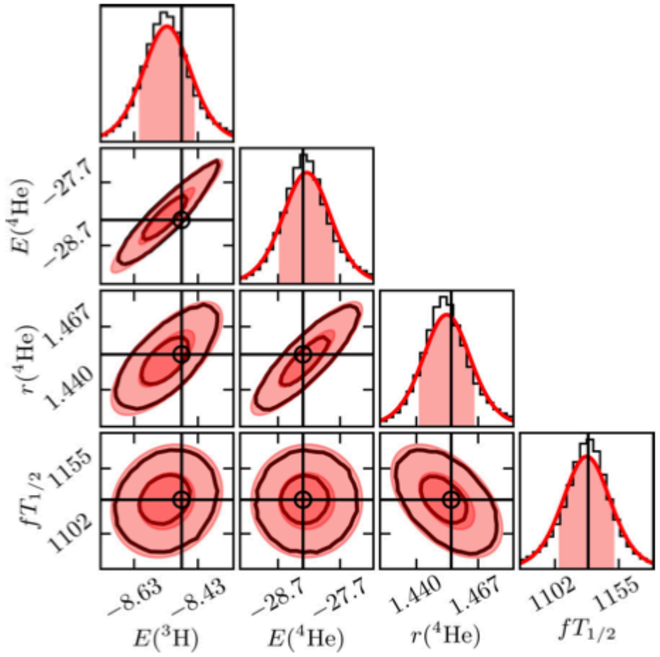
Analyzing the few-nucleon sector











A consistent solution for the 3H binding energy, the 4He binding energy and radius, and the 3H β-decay rate can only be obtained if χEFT truncation errors are included in the analysis.

$$p(\vec{\alpha}, \bar{c}^2, Q \,|\, \mathscr{D}, I) \propto p(\mathscr{D} \,|\, \vec{\alpha}, \Sigma_{\mathrm{th}}, \Sigma_{\mathrm{exp}}, \Sigma_{\mathrm{meth}}, I) p(\bar{c}^2 \,|\, Q, \vec{\alpha}, I) p(Q \,|\, \vec{\alpha}, I) p(\vec{\alpha} \,|\, I)$$

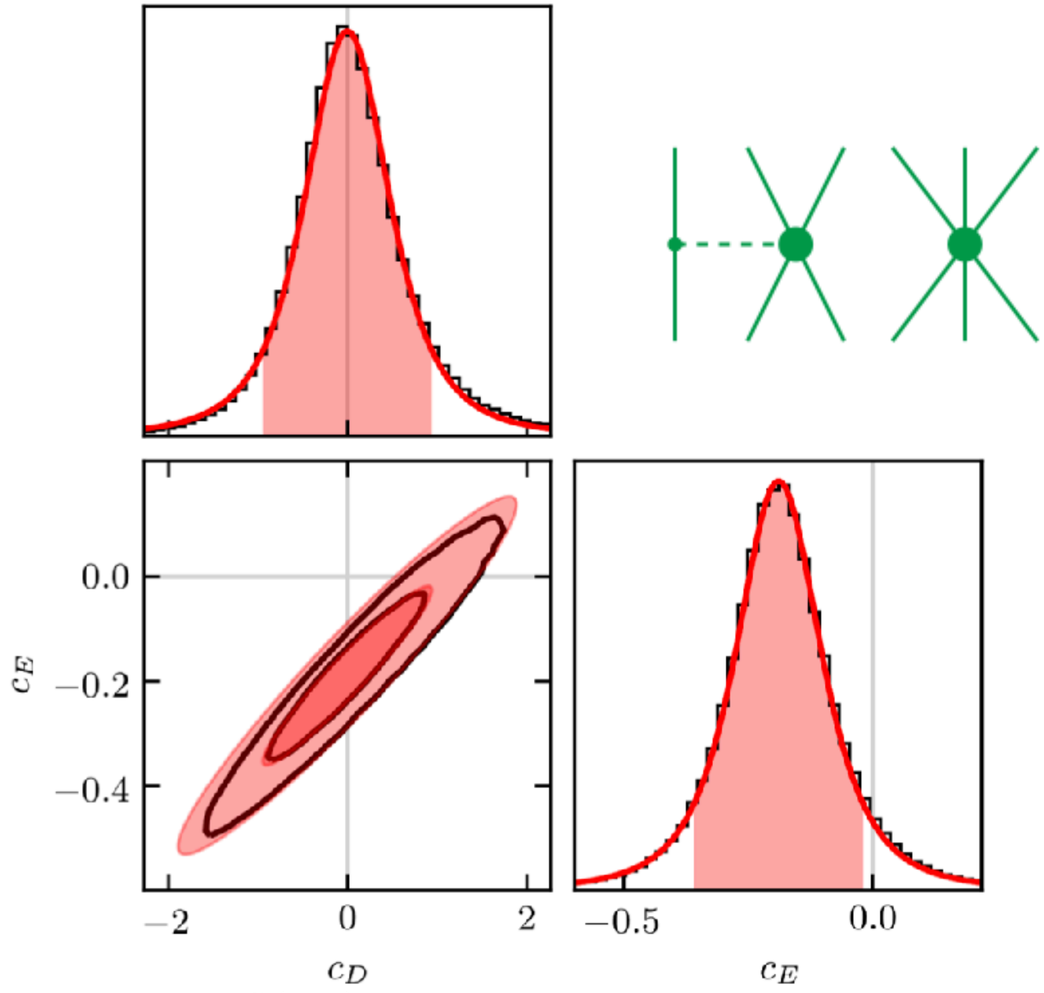
S. Wesolowski et al Phys Rev C (2021)

JupiterNCSM *no-core shell model with NN+3N* https://github.com/thundermoose/JupiterNCSM

fastwigxj *fast Wigner symbol computation* http://fy.chalmers.se/subatom/fastwigxj/

Posterior predictive distribution

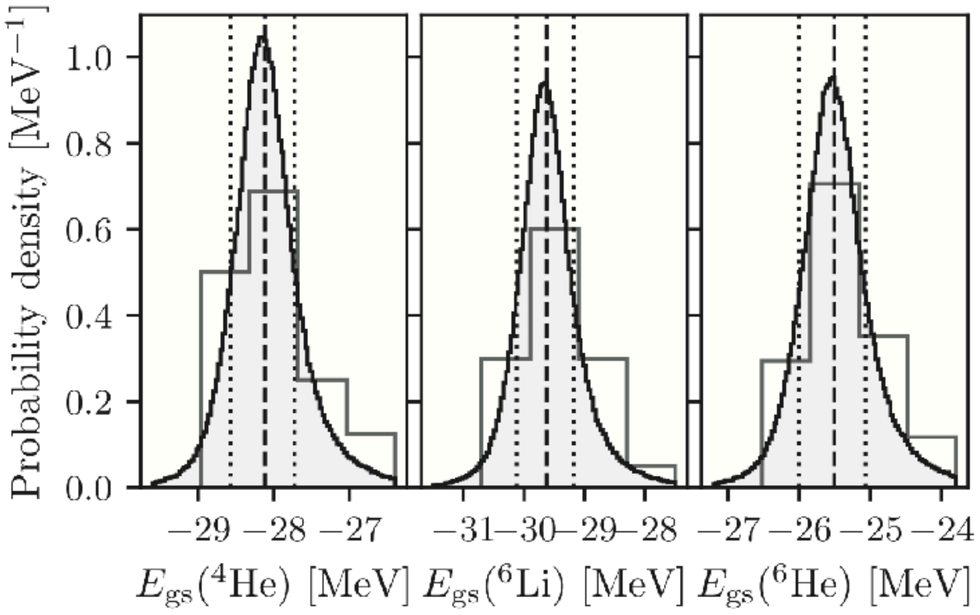
Low-sample representations can be useful



$$\mathrm{pr}(\boldsymbol{y}_{\mathrm{NCSM}} \mid \mathcal{D}, I) = \int d\vec{a}\, \mathrm{pr}(\boldsymbol{y}_{\mathrm{NCSM}}, \vec{a} \mid \mathcal{D}, I)$$

$$= \int d\vec{a}\, \mathrm{pr}(\boldsymbol{y}_{\mathrm{NCSM}} \mid \vec{a}, \mathcal{D}, I)\mathrm{pr}(\vec{a} \mid \mathcal{D}, I),$$

$$\mathrm{PPD}_{\mathrm{NCSM}} = \{\boldsymbol{y}_{\mathrm{NCSM}}(\vec{a}) : \vec{a} \sim \mathrm{pr}(\vec{a} \mid \mathcal{D}, I)\}$$

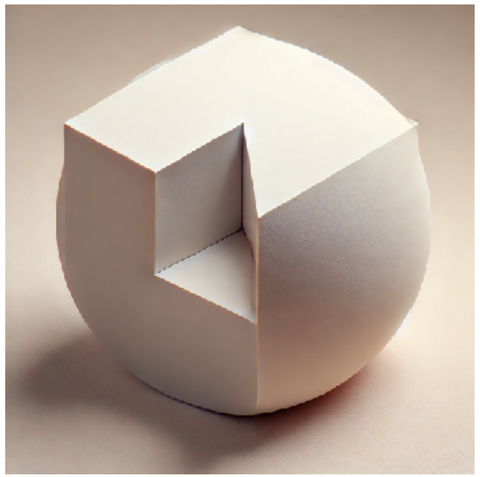


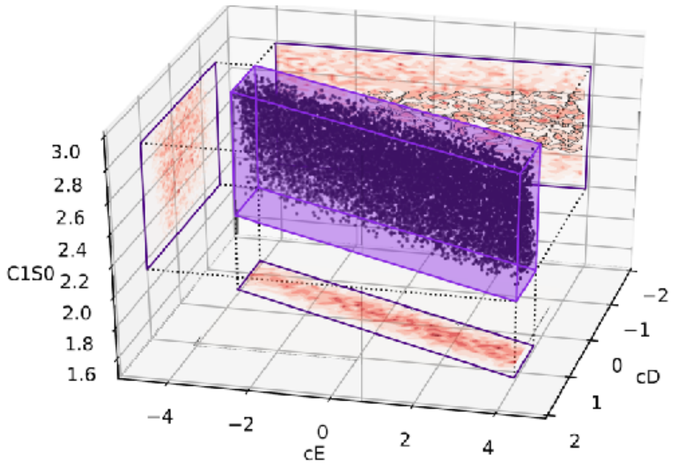
S. Wesolowski, et al. PRC (2021)

T. Djärv et al Phys. Rev. C **105**, 014005 (2022)

History matching

- removing parts of parameter space that do not reproduce data





Easier to claim *implausibility* than to quantify posterior probability

$$\Theta(\vec{\alpha}) \text{ vs. } p(\vec{\alpha} \,|\, \mathscr{D}, M, I) \text{ where } \Theta(\vec{\alpha}) = \begin{cases} 1 \text{ implausible} \\ 0 \text{ non} - \text{implausible} \end{cases}$$

History matching is a Bayes linear method operating only at the level of means and variances. We define an implausibility measure

$$I^2(\vec{\alpha}) = \max_{y^{(i)}_{\text{exp}} \in \mathscr{D}} \frac{|\,\mathbb{E}[\tilde{y}^{(i)}_{\text{th}}(\vec{\alpha})] - y^{(i)}_{\text{exp}}\,|^2}{\text{Var}[\tilde{y}^{(i)}_{\text{th}}(\vec{\alpha}) - y^{(i)}_{\text{exp}}]}$$

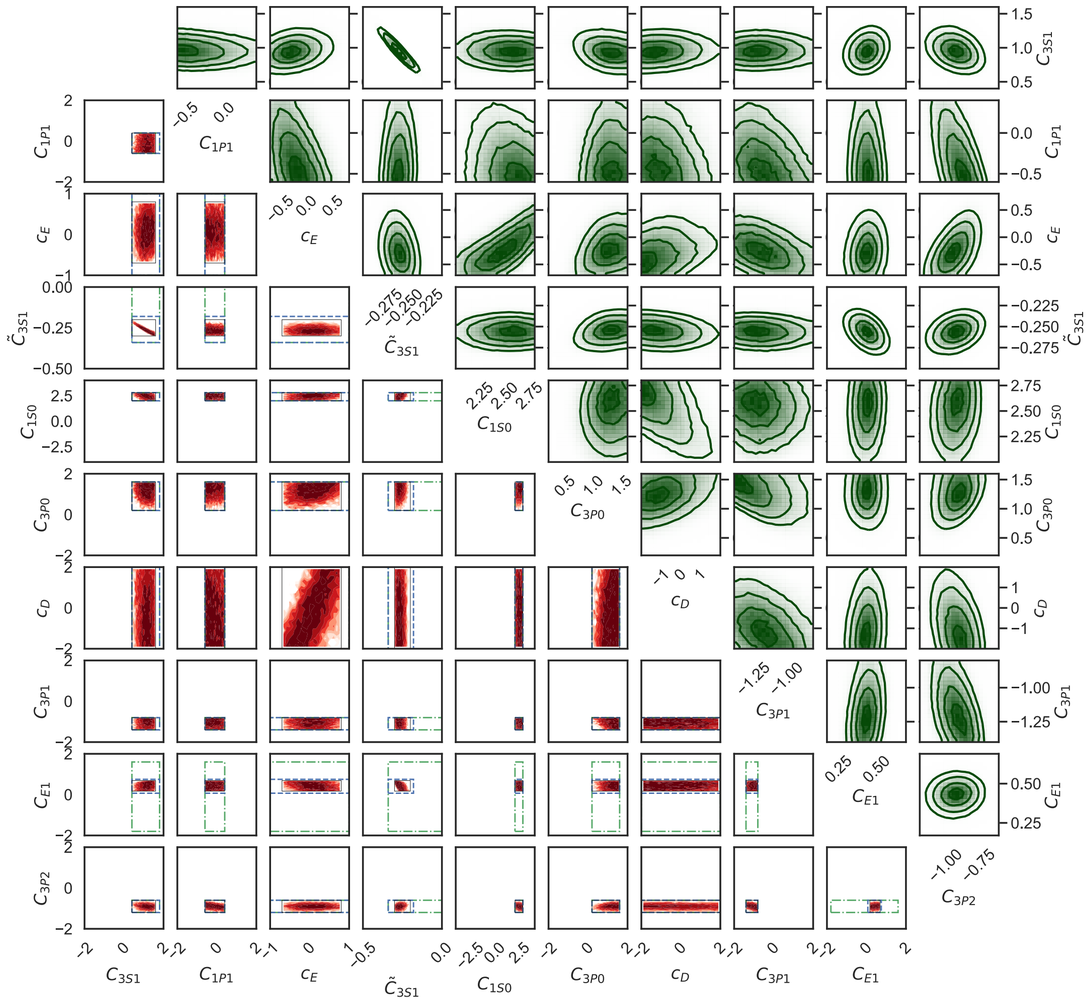
and iteratively rule out $\vec{\alpha}$ for which $I(\vec{\alpha}) > c$, where we often use $c = 3$ appealing to Pukelheim's 3-sigma rule

$$P(\,|X - \mu\,| > k\sigma) \leq \frac{4}{9k^2} \underset{k=3}{=} \frac{4}{81} = 0.049 \text{ (Vysochanskii - Petunin inequality)}$$

Ian Vernon et al, Bayesian Analysis (2010)

Ab initio analysis of 28-O

Bayesian posterior pdf



history matching

- **History matching** identifies the parameter region where we expect the LEC posterior distribution to reside.
- **MCMC + emulators** to draw 10^8 samples of the LEC posterior at ΔNNLO with NN+3N interaction.

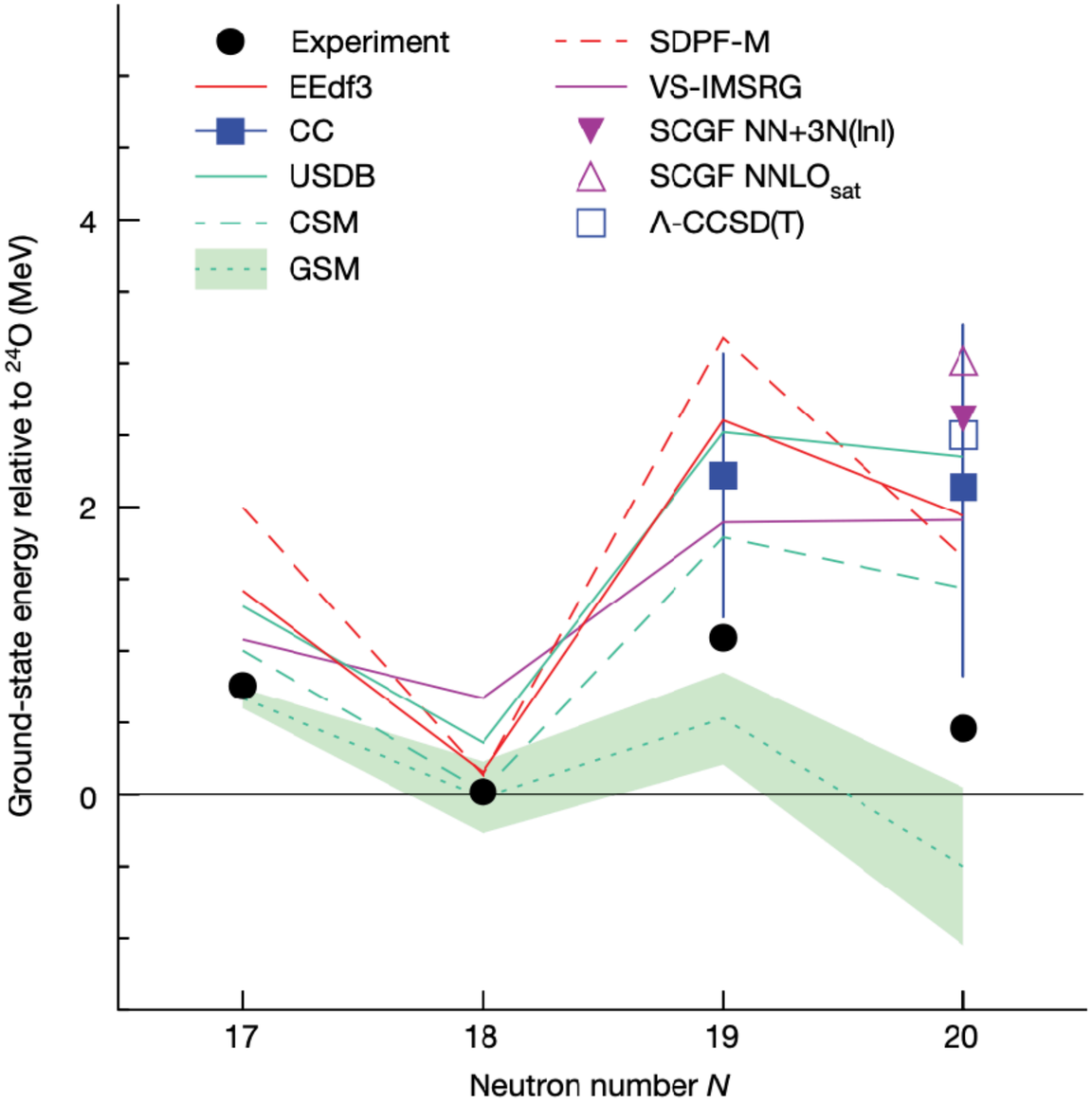
$$\mathrm{p}(\vec{\alpha}|A = 2 - 24)$$

- We assume uniform prior + uncorrelated normal likelihood. Effects of heavier-tail (Student's *t)* correlations ($\rho \approx 0.6$) not great.
- Very informative to update the parameter posterior with ΔE(25O,24O). Subsequently draw 121 parameter samples that we employ in our prediction of oxygen-27/28.

$$\vec{\alpha} \sim \mathrm{p}(\vec{\alpha}|A = 2 - 25)$$

Y. Kondo et al. Nature (2023)

Ab initio analysis of 28-O



The experimental data point (red star) is away from the posterior maximum. Only a few finely-tuned chiral interactions are able to reproduce low-energy and exotic oxygen structure

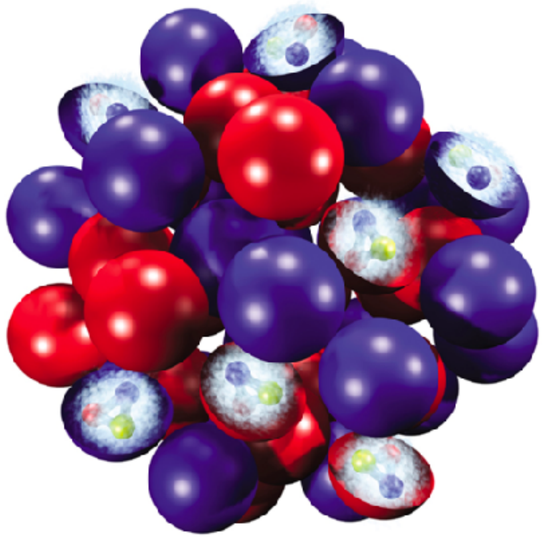
We can state with 98% certainty that 28-O is unbound $\Delta E(^{28,24}\mathrm{O}) \equiv E(^{28}\mathrm{O}) - E(^{24}\mathrm{O}) = 2.1^{+1.2}_{-1.3}$ MeV

$$z = M(\theta) + \epsilon_{\text{method}} + \epsilon_{\text{model}} + \epsilon_{\text{exp}}.$$

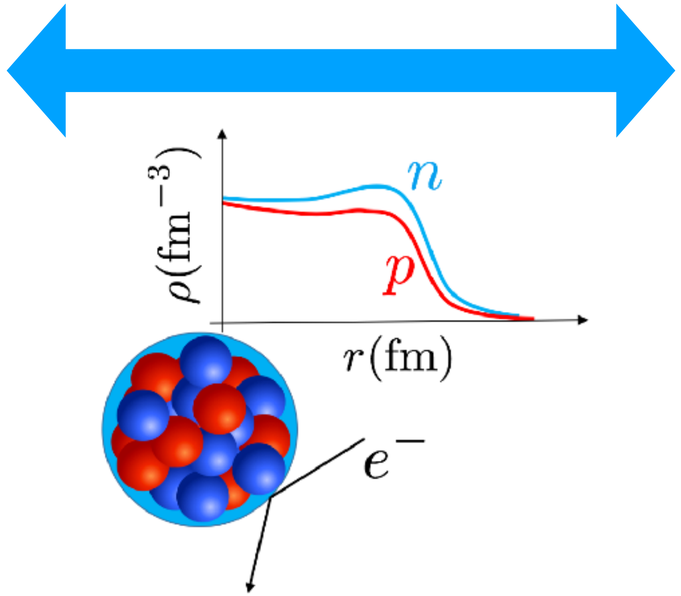
Target	z	ϵ_{exp}	ϵ_{model}	ϵ_{method}	$\epsilon_{\text{emulator}}$
$E(^2\mathrm{H})$	−2.2298	0.0	0.05	0.0005	0.001%
$r_p^2(^2\mathrm{H})$	3.9030	0.0	0.02	0.0005	0.001%
$Q(^2\mathrm{H})$	0.27	0.01	0.003	0.0005	0.001%
$E(^3\mathrm{H})$	−8.4818	0.0	0.17	0.0005	0.005%
$E(^4\mathrm{He})$	−28.2956	0.0	0.55	0.0005	0.005%
$r_p^2(^4\mathrm{He})$	2.1176	0.0	0.045	0.0005	0.05%
$E(^{16}\mathrm{O})$	127.62	0.0	0.75	1.5	0.5%
$r_p^2(^{16}\mathrm{O})$	6.660	0.0	0.16	0.05	1%
$\Delta E(^{22,16}\mathrm{O})$	−34.41	0.0	0.4	0.5	1%
$\Delta E(^{24,22}\mathrm{O})$	−6.35	0.0	0.4	0.5	4%
$E_{2^+}(^{24}\mathrm{O})$	4.79	0.0	0.5	0.25	2%
$\Delta E(^{25,24}\mathrm{O})$	0.77	0.02	0.4	0.25	—

Y. Kondo et al. Nature (2023)

Linking the neutron skin of 208Pb to nuclear forces



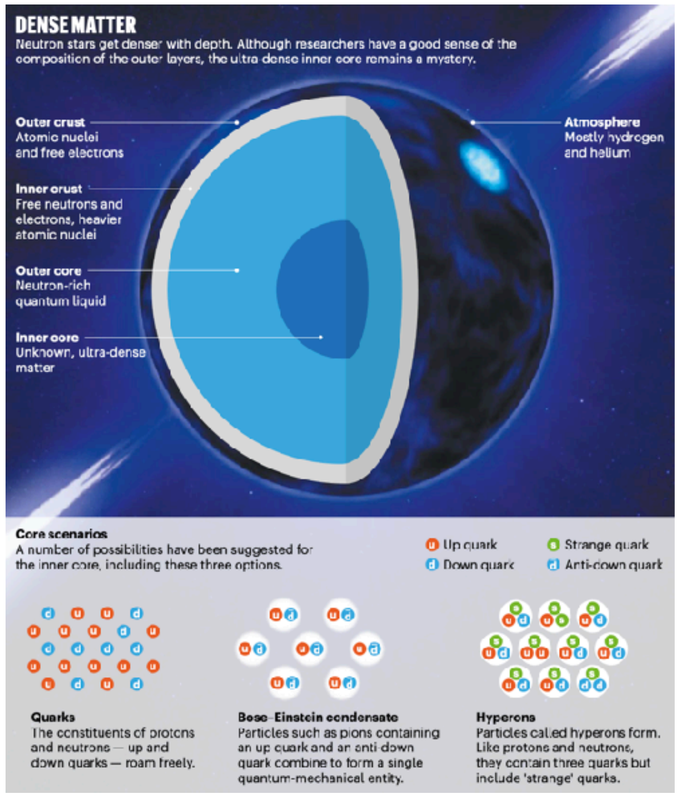
The neutron skin-thickness depends on the pressure (P) of neutron-rich matter: the greater the pressure, the thicker the skin as neutrons are pushed out against surface tension. For finite nuclei, Symmetry energy (S) is lower at surface region. Energy to be gained by moving some excess neutrons outwards. But this increases surface tension (the surface grows)



(1) Challenging to find a "terrestrial slab" of pure neutron matter to probe its EOS.

$$P \approx \frac{L\rho_0}{3}$$

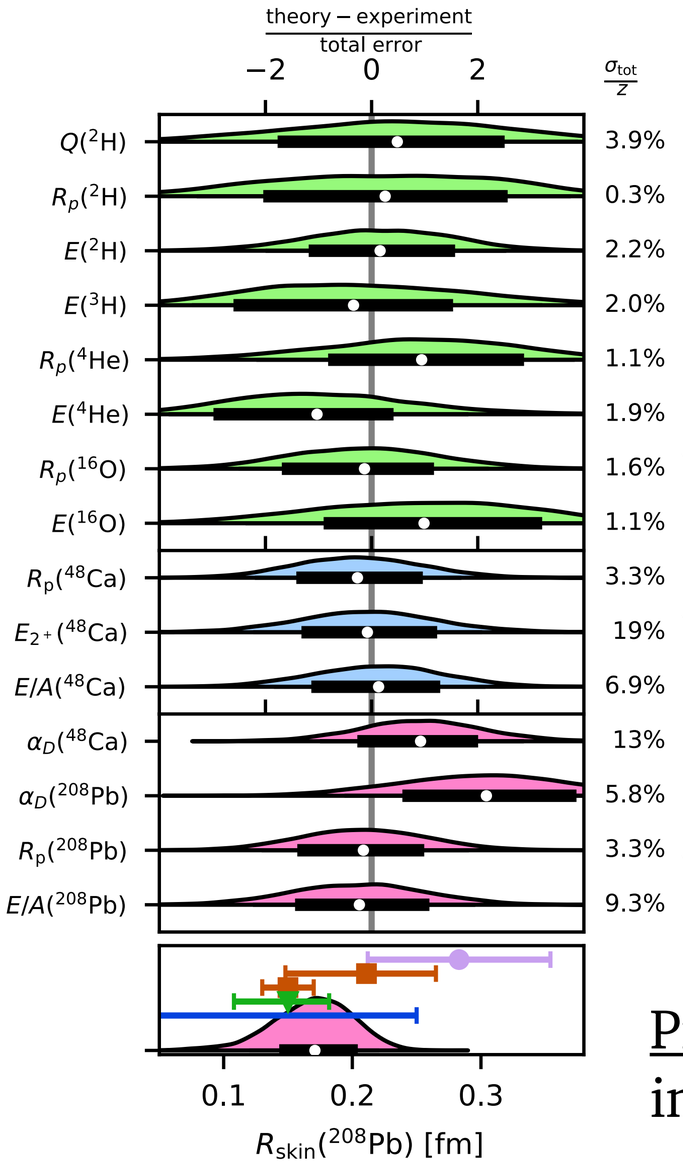
(2) Instead: measure the skin thickness.
(3) Use a theoretical model to analyze relation between PNM pressure and the skin thickness.



The same pressure supports a neutron star against gravity. Thus models with thicker neutron skins often produce neutron stars with larger radii.

B. A. Brown Phys. Rev. Lett. 85, 5296 (2000)
C. J. Horowitz and J. Piekarewicz Phys. Rev. Lett. 86, 5647 (2001)

Linking the neutron skin of 208Pb to nuclear forces



History Matching

We explore 10^9 different interaction parameterizations

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

Calibration

Likelihood weighting

Validation

Inspect ab initio model and error estimates

History-matching observables

Observable	z	ε_{exp}	ε_{model}	ε_{method}	ε_{em}	PPD
$E(^2H)$	-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	—	$3.36^{+0.14}_{-0.13}$

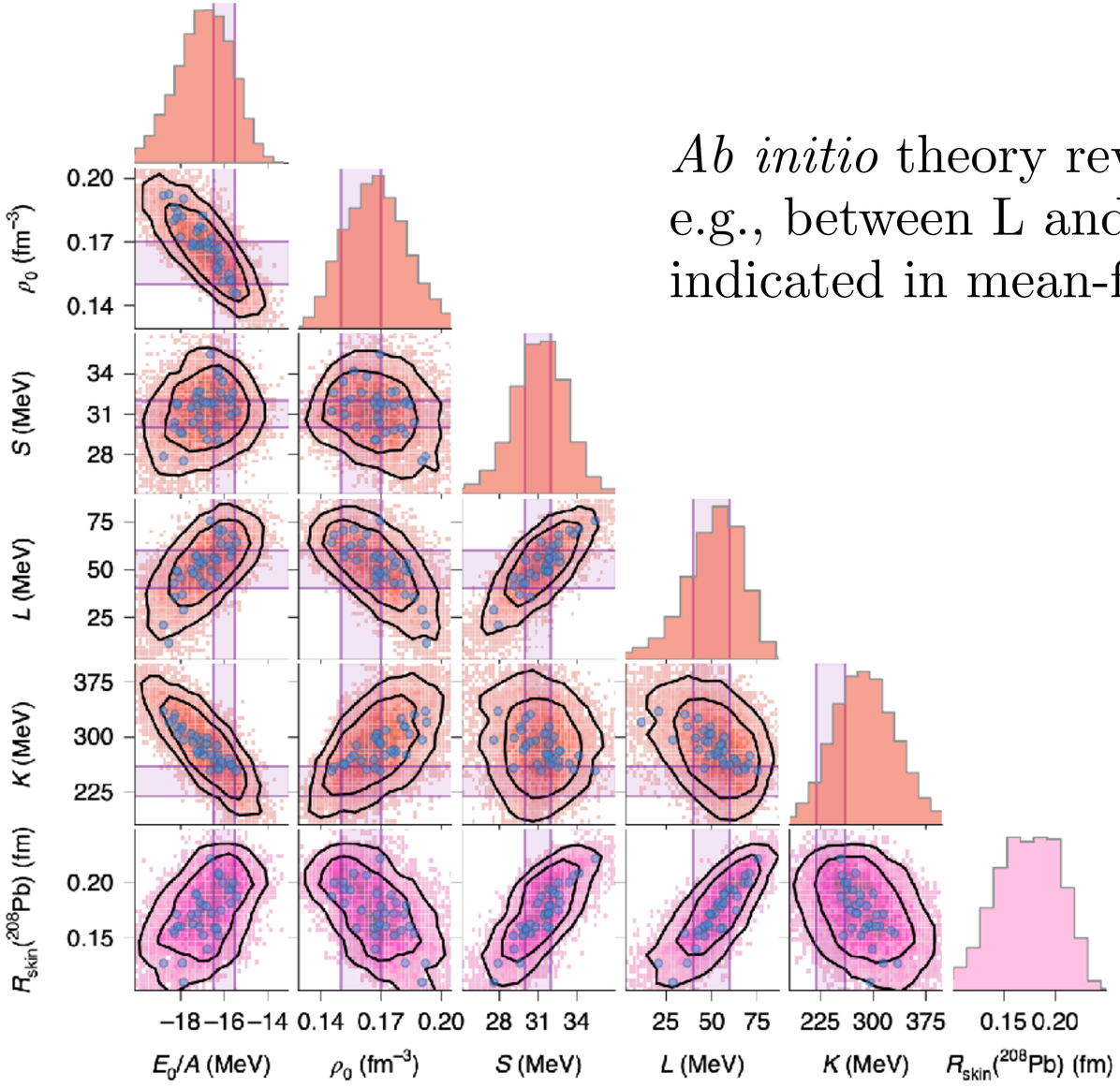
Validation observables

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	—	$22.6^{+2.1}_{-1.8}$

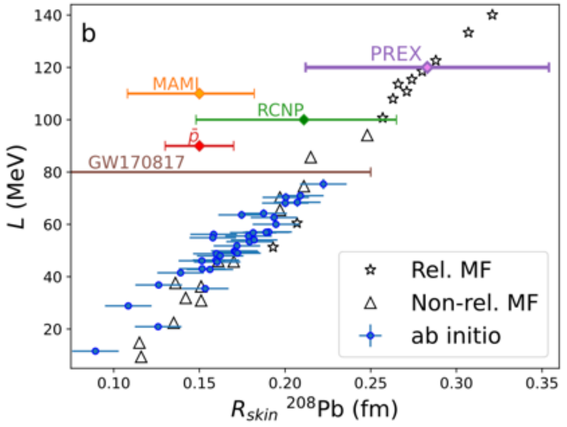
Prediction: small skin thickness 0.14-0.20 fm in mild (1.5sigma) tension with PREX.

B. Hu et al Nature Physics(2022)

Linking the neutron skin of 208Pb to nuclear forces



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

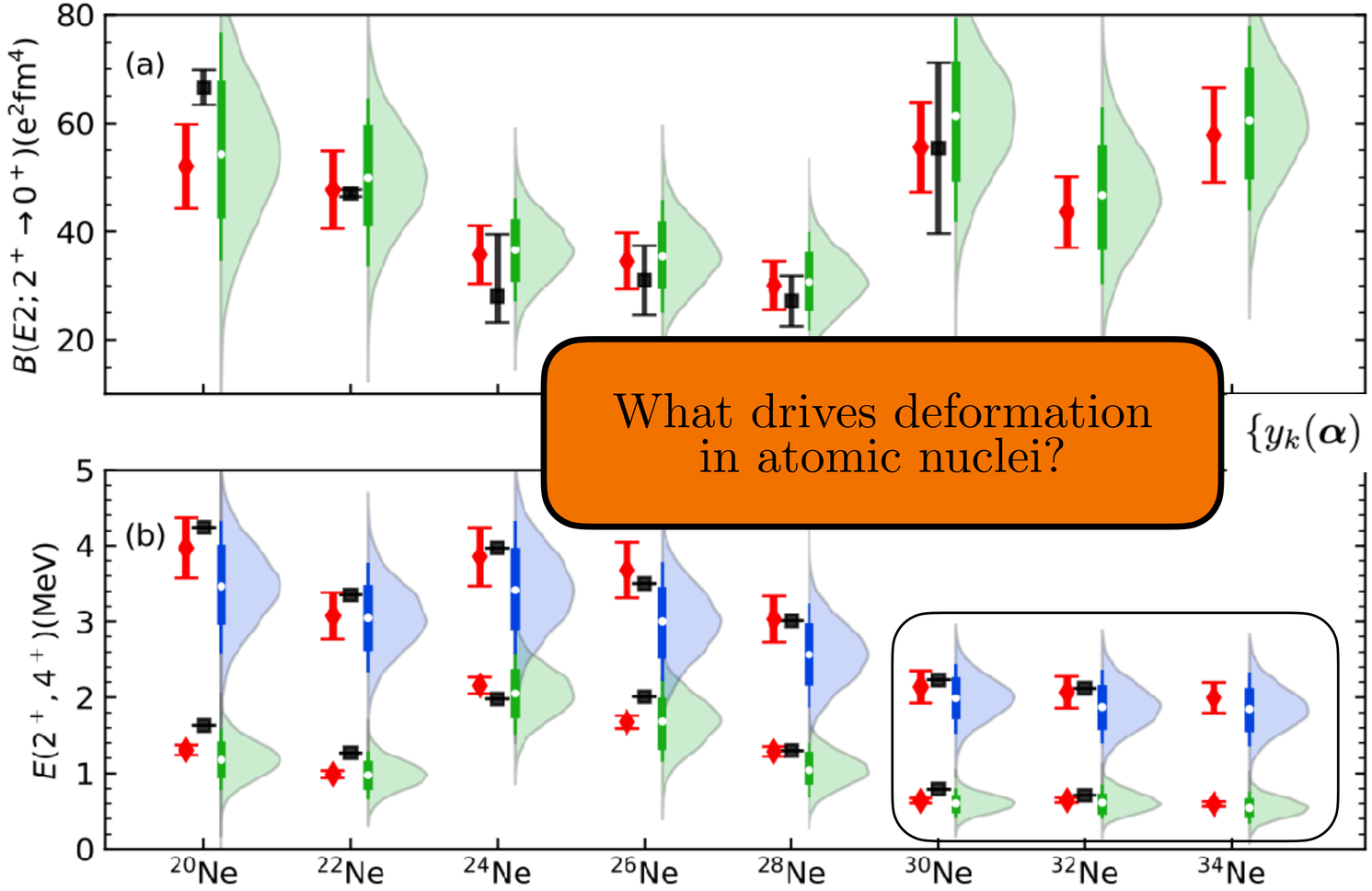


Nuclear matter properties			
Observable	median	68% CR	90% CR
E_0/A	-16.9	$[-17.9, -15.4]$	$[-19.1, -14.9]$
ρ_0	0.167	$[0.150, 0.181]$	$[0.142, 0.194]$
S	31.1	$[29.1, 33.2]$	$[27.6, 34.6]$
L	52.7	$[38.3, 68.5]$	$[23.9, 76.2]$
K	287	$[242, 331]$	$[216, 362]$
Neutron skins			
Observable	median	68% CR	90% CR
$R_{\rm skin}(^{48}{\rm Ca})$	0.164	$[0.141, 0.187]$	$[0.123, 0.199]$
$R_{\rm skin}(^{208}{\rm Pb})$	0.171	$[0.139, 0.200]$	$[0.120, 0.221]$

B. Hu et al Nature Physics(2022)

Linking deformation and chiral nuclear forces

- good reproduction of data without effective charges



What drives deformation in atomic nuclei?

PPDs from importance resampled non-implausible points from 28-O study

$$\{y_k(\boldsymbol{\alpha}) + \varepsilon_{\mathrm{MB}} + \varepsilon_{\mathrm{EFT}} : \boldsymbol{\alpha} \sim p(\boldsymbol{\alpha}|\mathcal{D}_{\mathrm{cal}})\}$$

Rigid rotor $R_{42} = 10/3$

$$E(J) = \frac{J(J+1)}{2I}$$

Z. H. Sun et al arXiv (2024)

Global sensitivity analysis

- what is responsible for the variance in the output?

A sensitivity analysis addresses the question *'How much does each model parameter contribute to the uncertainty in the prediction?'*

Variance-based methods for GSA decompose the variance of a certain model output in terms of each input and their combinations.

Global methods deal with the uncertainties of the outputs due to input variations over the whole domain.

Bottleneck: Converging the MC sampling of the variance integrals require approximately 10^6 samples

A. Saltelli et al. Global sensitivity analysis: the primer, John Wiley & Sons

Global sensitivity analysis

- what is responsible for the variance in the output?

Consider a model $Y = f(\boldsymbol{X})$ where $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ follow some (iid) distribution.

- If we set $X_i = x_i^\star$ (true), how would $\mathrm{Var}(Y)$ change?
- This is measured by the conditional variance $\mathrm{Var}_{\boldsymbol{X}_{-i}}(Y \mid X_i = x_i^\star)$
- However, since we don't know $x_i^\star$, we average $\mathbb{E}_i[\mathrm{Var}_{\boldsymbol{X}_{-i}}(Y \mid X_i)]$

By the law of total variance $\mathrm{Var}(Y) = \mathrm{Var}_i(\mathbb{E}_{\boldsymbol{X}_{-i}}[Y \mid X_i]) + \mathbb{E}_i[\mathrm{Var}_{\boldsymbol{X}_{-i}}(Y \mid X_i)]$

$$\mathbb{E}_{\boldsymbol{X}_{-i}}[Y \mid X_i] = \int d\boldsymbol{X}_{-i}\, f(X_1, \ldots, X_n) p(\boldsymbol{X}_{-i} \mid X_i)$$

large/small

small/large

From this we *define* the *main effect* $S_i = \dfrac{\mathrm{Var}_i(\mathbb{E}_{\boldsymbol{X}_{-i}}[Y \mid X_i])}{\mathrm{Var}(Y)} \in [0,1]$

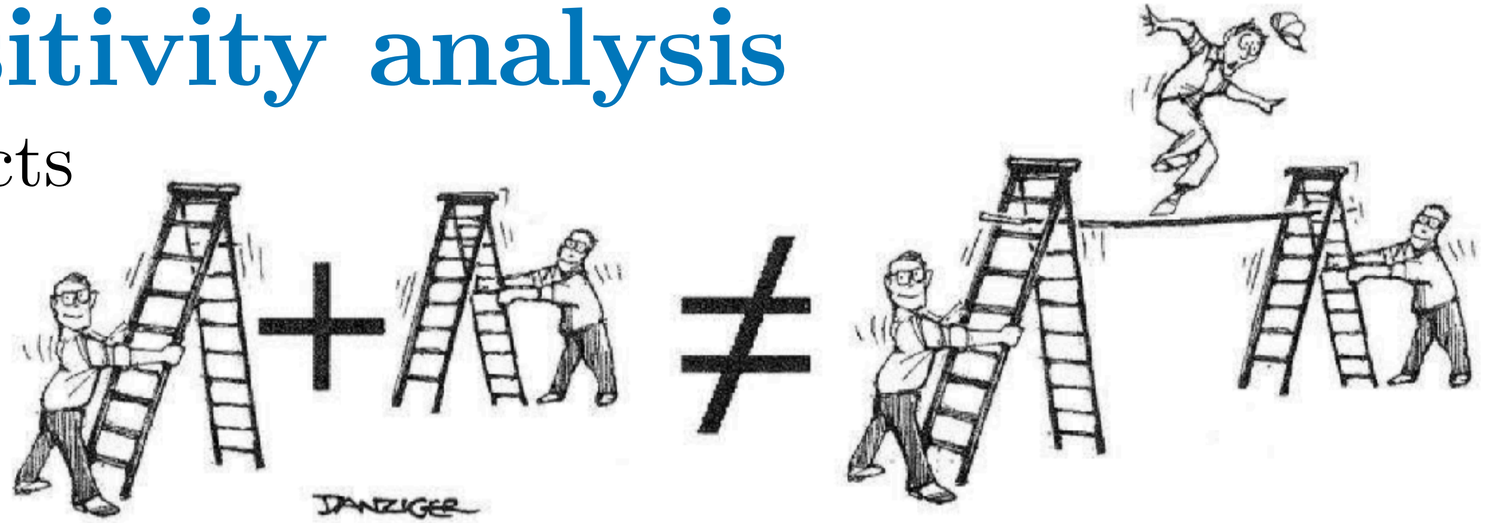
This is a very convenient scalar quantity to convey essential information about a model

A large value of S_i indicates large first-order sensitivity to X_i
A small value of S_i indicates small first-order sensitivity to X_i

A. Saltelli et al. Global sensitivity analysis: the primer, John Wiley & Sons

Global sensitivity analysis

- higher-order effects



Variances are computed with MC sampling: a show-stopper for expensive models.

We can define high-order sensitivities $S_{ij} = \dfrac{\mathrm{Var}_{ij}(\mathbb{E}_{\boldsymbol{X}_{-ij}}[Y|X_i, X_j])}{\mathrm{Var}(Y)} - S_i - S_j$ etc.

Computing S_{ijk} and even higher-order sensitivites quickly becomes a formidable task. Fortunately, it is possible to compute the total effect S_{Ti} of model parameter i, including its first order effect and sensitivities to all orders as

$$S_{Ti} = 1 - \frac{\mathrm{Var}_{\boldsymbol{X}_{-i}}(\mathbb{E}_i[Y|\boldsymbol{X}_{-i}])}{\mathrm{Var}(Y)} = S_i + S_{ij} + S_{ijk} + \ldots \quad (S_{Ti} \geq S_i, \text{ equality for additive models})$$

A. Saltelli et al. Global sensitivity analysis: the primer, John Wiley & Sons

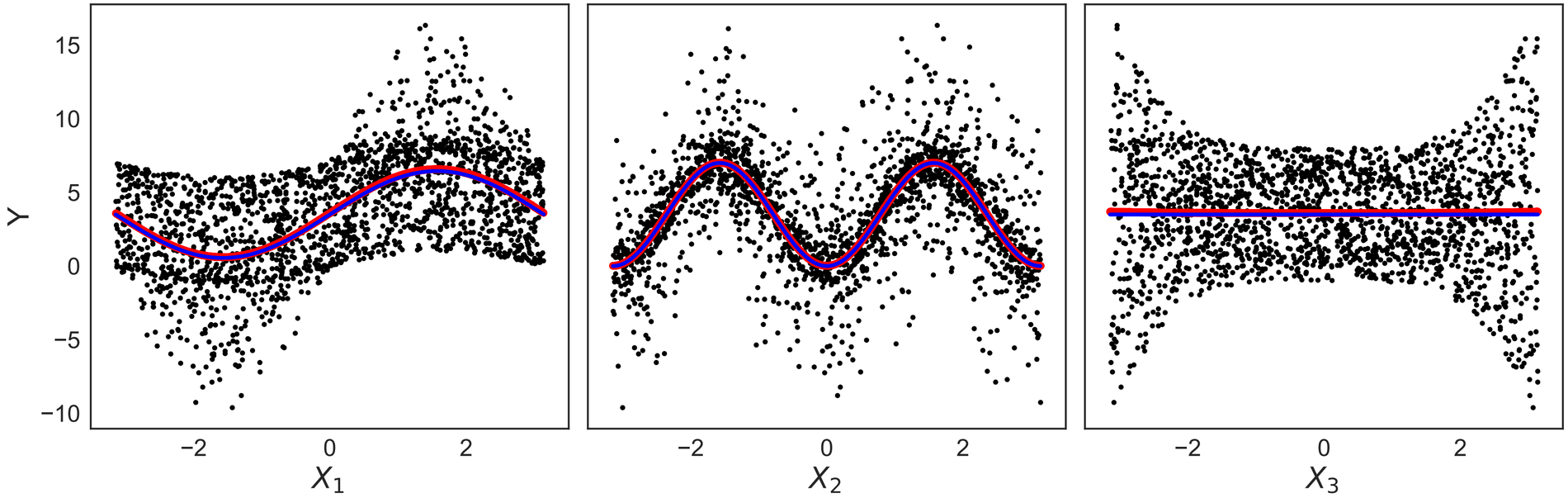
Global sensitivity analysis

- an example

The *Ishigami* function $f(X_1, X_2, X_3) = \sin X_1 + A \sin^2 X_2 + B X_3^4 \sin X_1, \quad X_i \sim \mathcal{U}(-\pi, +\pi), \quad A = 7.0, \quad B = 0.1$

- Non-linear, non-additive, X_3 has no additive effect on Y but interacts only with X_1 -

1st order *and* 2nd order via X_3 | no interaction. purely 1st order | 2nd order via X_1



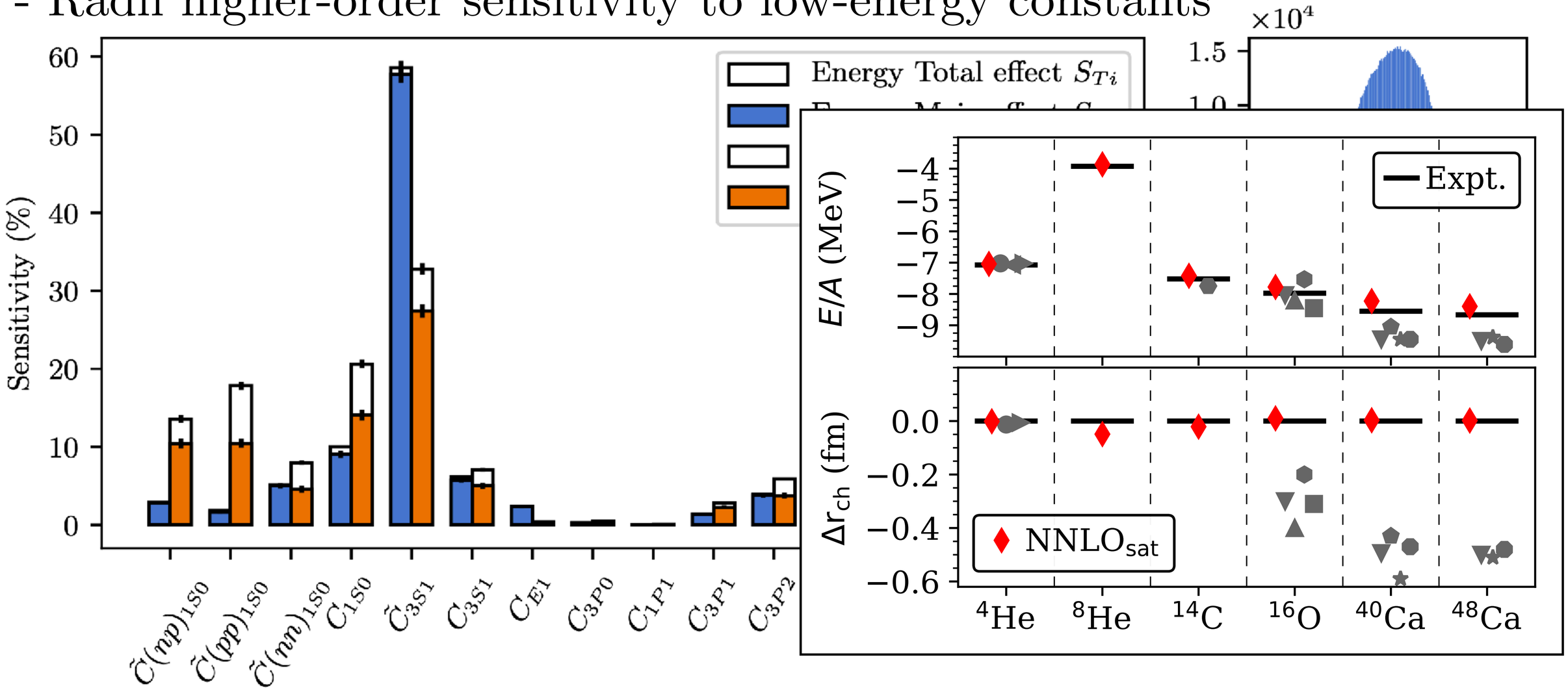
$S_1 = 0.31$
$ST_1 = 0.56$

$S_2 = 0.44$
$ST_2 = 0.44$

$S_3 = 0.00$
$ST_3 = 0.24$
$S_{13} = 0.24$

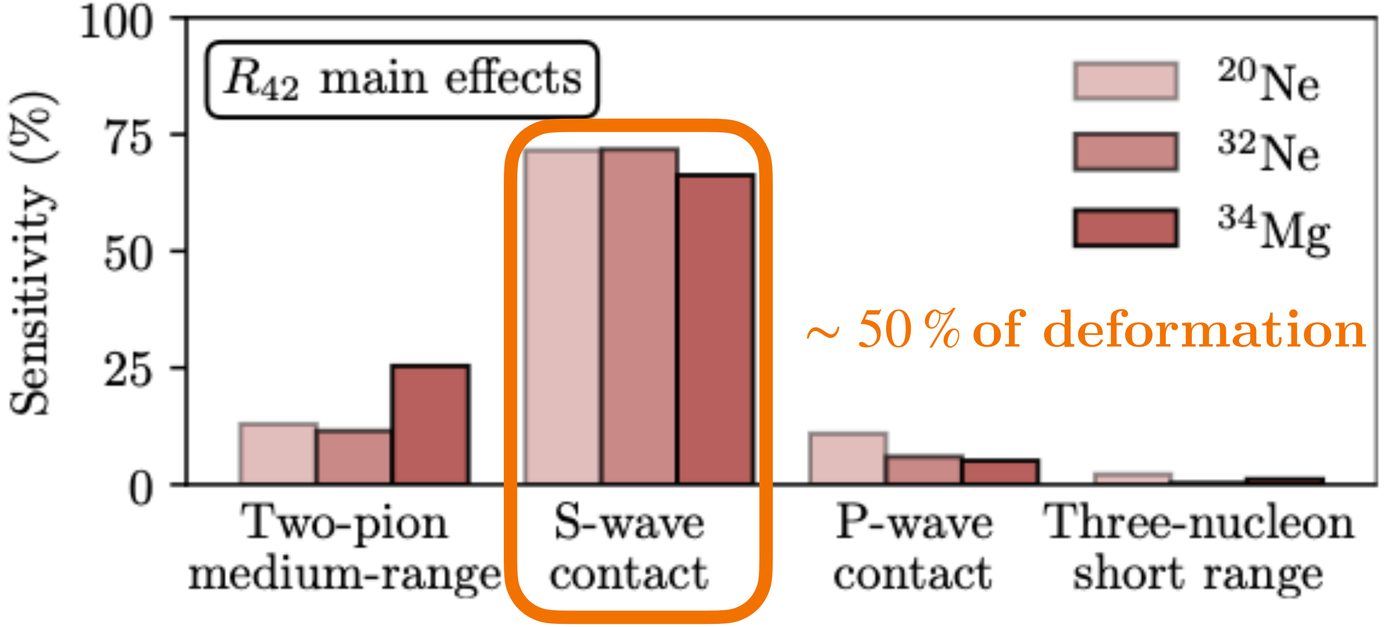
Global sensitivity analysis of an atomic nucleus

- Radii higher-order sensitivity to low-energy constants



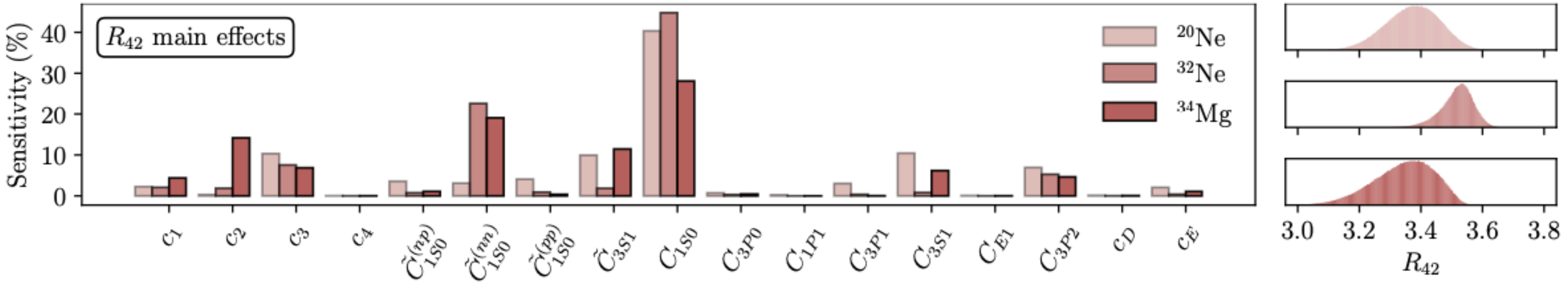
Ekström and Hagen Phys Rev Lett (2019)

Linking deformation and chiral nuclear forces



$$S_i = \frac{\mathrm{Var}_i(\mathbb{E}_{X-i}[Y \mid X_i])}{\mathrm{Var}(Y)} \in [0,1]$$ 10^6 Monte Carlo samples

- Adding short-range repulsion appears to increase deformation, probably via reduced pairing.
- Increasing medium-range 2π-exchange increases deformation, presumbably by adding attraction in higher partial waves



Z. H. Sun et al (2024)

If there was a fifth lecture

How can recent advances in *ab initio* theory and Bayesian inference generate more knowledge?

Bayesian model mixing

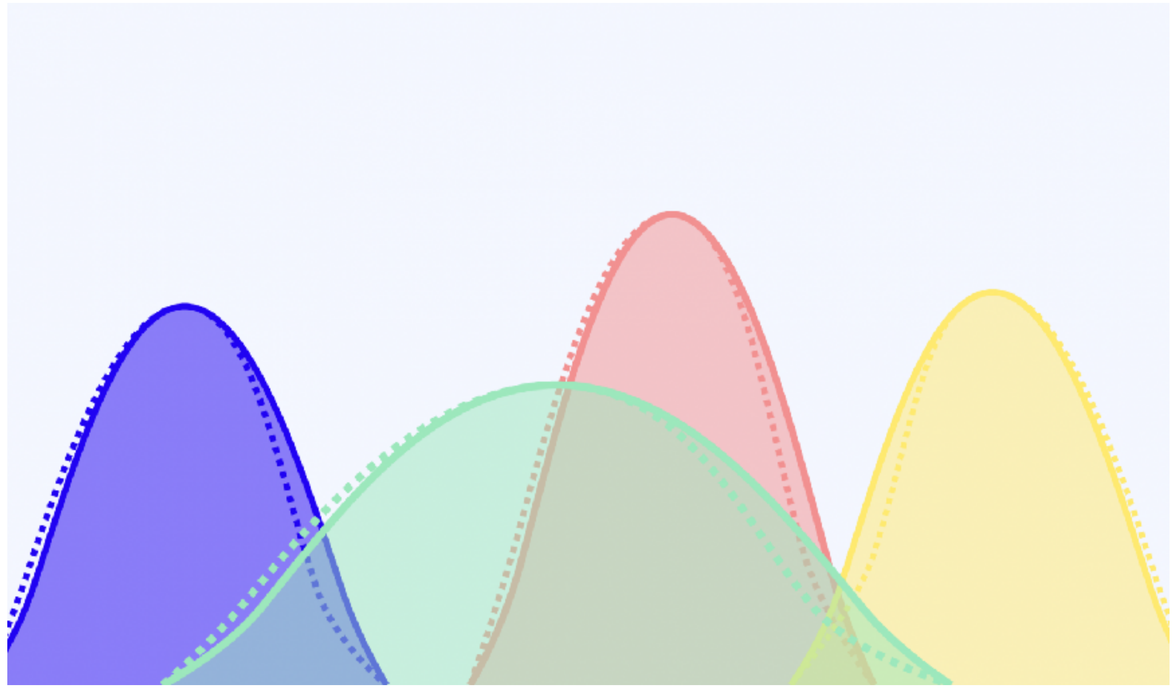
There are several competing models. Which one to choose?

Let's not choose in the traditional sense.

The theory of Bayesian model mixing (and related methods) enables construction of composite models in a statistically meaningful way. "We average over all considered models".

Consider a combination of, e.g., pionless and pionfull EFT, or mixing *ab initio* and DFT predictions,...

$$p(y\,|\,x,\mathscr{D}) = \sum_{i=1}^{K} w_i(x) p(y\,|\,x,\mathscr{D},M_k)$$



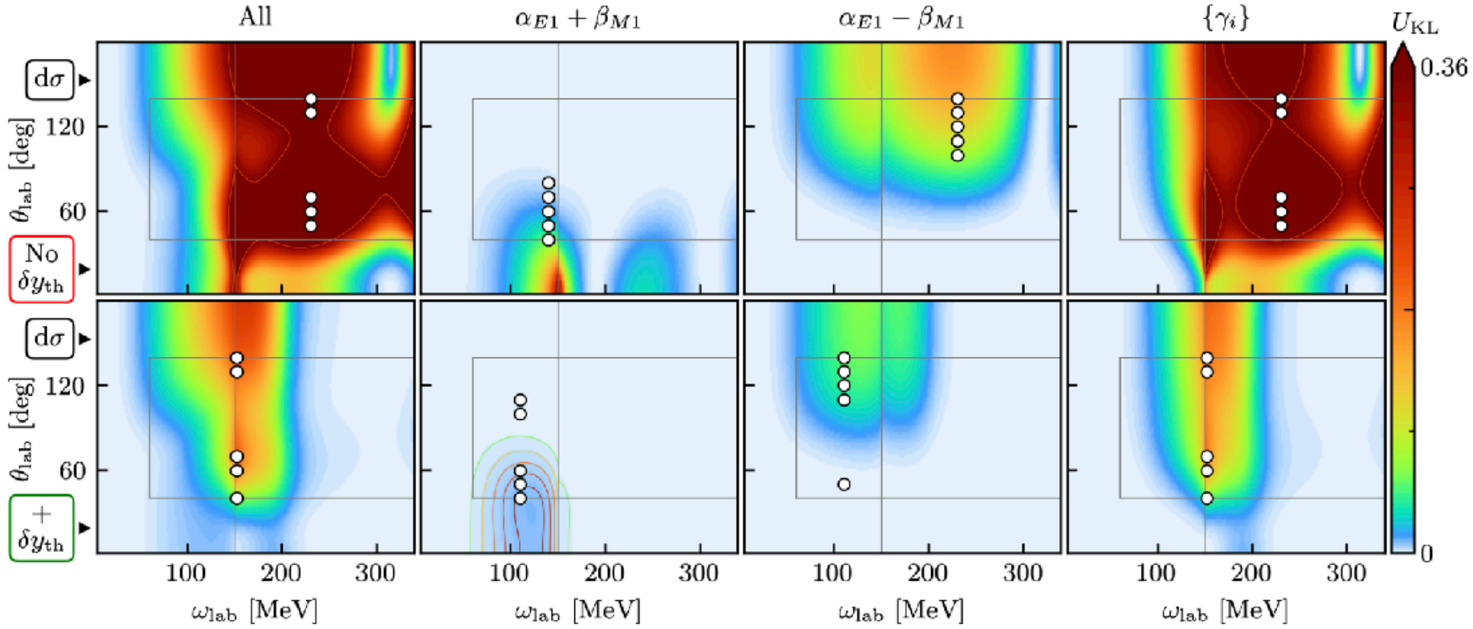
Yuling Yao, Aki Vehtari, Daniel Simpson, Andrew Gelman, Bayesian Anal. **13,** 917 (2018)

Bayesian experimental design

Which experiment(s) to do, given: 0) your priors, 1) the present experimental data, 2) estimated model uncertainty, 3) current experimental constraints, and 4) your desired benefit

$$d^* = \arg\max_d \int\int U(d, \mathscr{D}, \vec{\alpha})\, p(\vec{\alpha}, \mathscr{D} \,|\, d, I)\, d\vec{\alpha}\, d\mathscr{D}$$

Specify a utility U reflecting the purpose of the experiment, regard the design choice as a decision problem, and select the design that maximizes the expected utility. Numerically very demanding...



Melendez et al EPJA (2021)