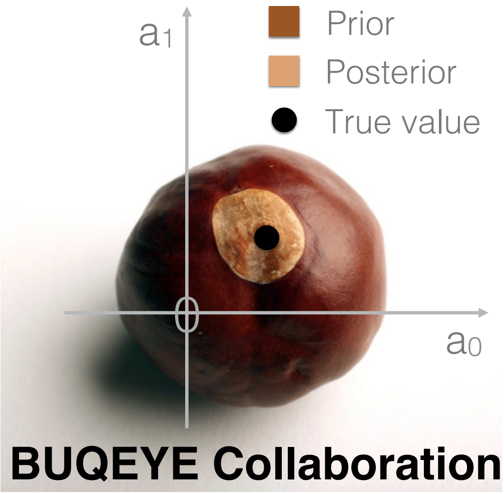
Uncertainty quantification (UQ) for chiral effective field theory

Dick Furnstahl

BSM Physics with Nucleons and Nuclei, July, 2020

THE OHIO STATE UNIVERSITY



Special thanks to my BUQEYE collaborators!

U.S. DEPARTMENT OF ENERGY

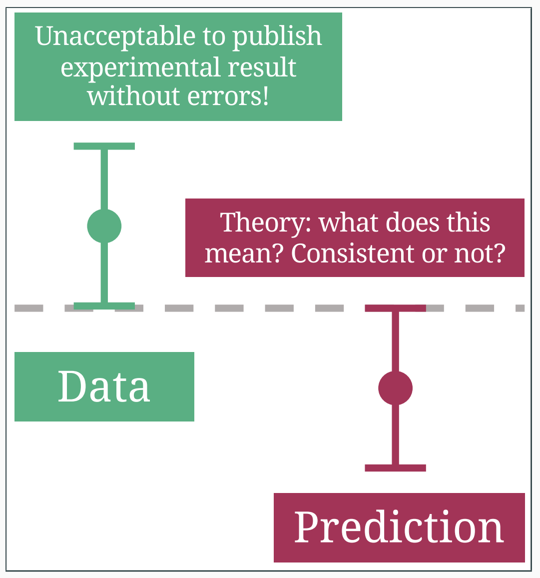
NSF

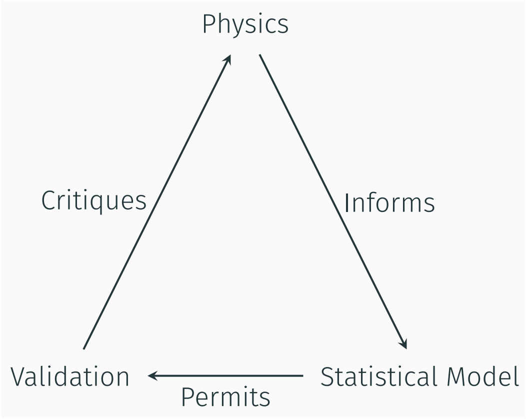
NUCLEI

Nuclear Computational Low-Energy Initiative

Preview of take-away points

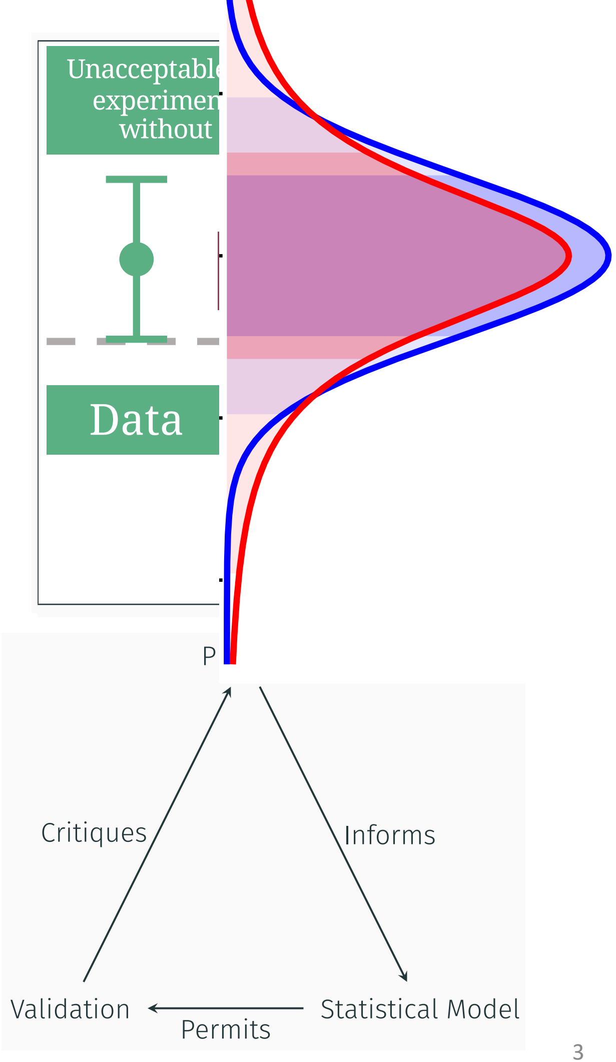
- Bayesian statistics is a powerful framework for (chiral) EFT uncertainty quantification (UQ). *Everything is a pdf.*





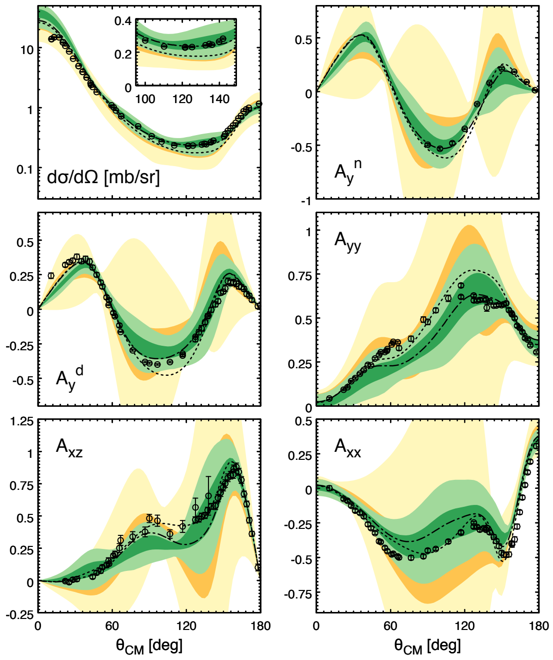
Preview of take-away points

- Bayesian statistics is a powerful framework for (chiral) EFT uncertainty quantification (UQ). *Everything is a pdf.*

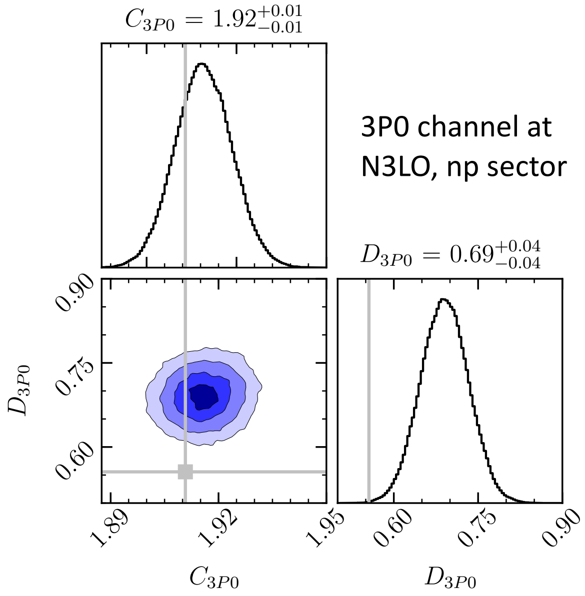


Preview of take-away points

- Bayesian statistics is a powerful framework for (chiral) EFT uncertainty quantification (UQ). *Everything is a pdf.*

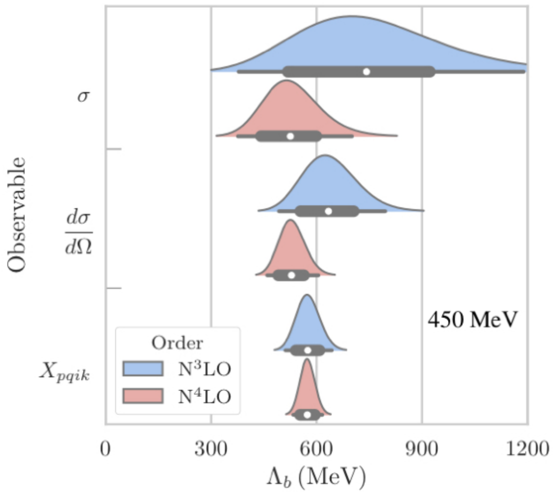


pr($\delta\mathbf{y}_{th}$ | $\mathbf{y}_{th}$, I) ⇒ pdf of *theory error* $\delta\mathbf{y}_{th}$ given theory calculations $\mathbf{y}_{th}$

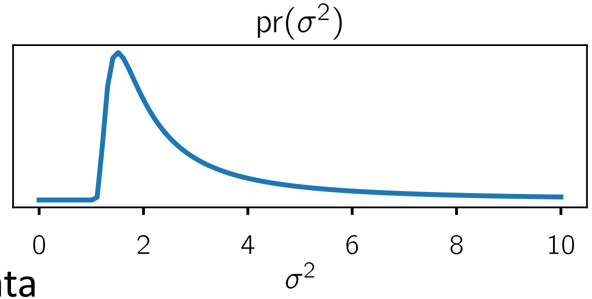


3P0 channel at N3LO, np sector

pr($\boldsymbol{\theta}$ | $\mathbf{y}_{exp}$, Σ_{exp} , Σ_{th} , I) ⇒ pdf of model parameters $\boldsymbol{\theta}$ given data $\mathbf{y}_{exp}$ and experiment/theory errors Σ



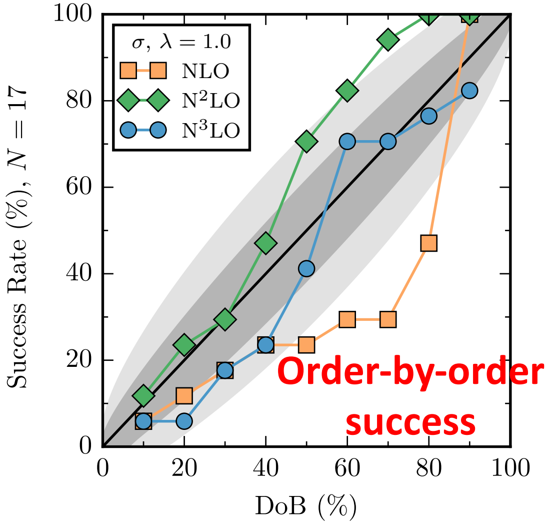
pr(Λ_b | $\mathbf{y}_{th}$, I) ⇒ pdf of breakdown scale of EFT expansion

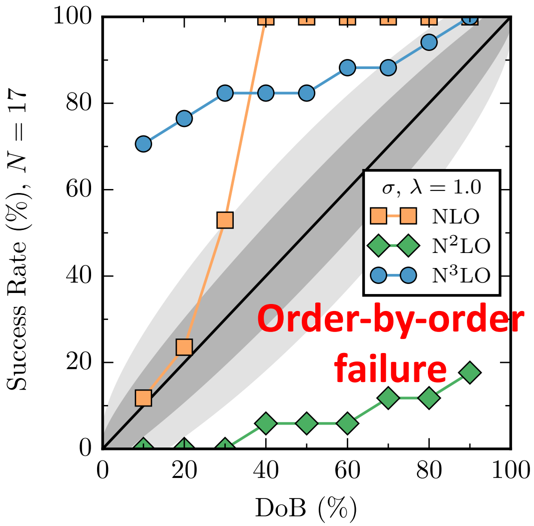


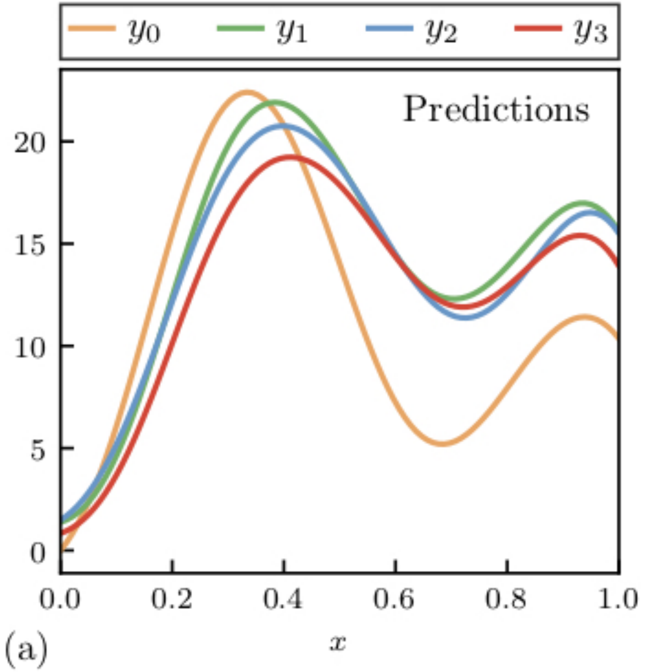
Naturalness prior!

Preview of take-away points

- Bayesian statistics is a powerful framework for (chiral) EFT uncertainty quantification (UQ). *Everything is a pdf.*
- EFT theory *discrepancy model* from the convergence pattern and naturalness *priors*; assumptions explicit.
- *Model checking* is an essential part of Bayesian UQ.



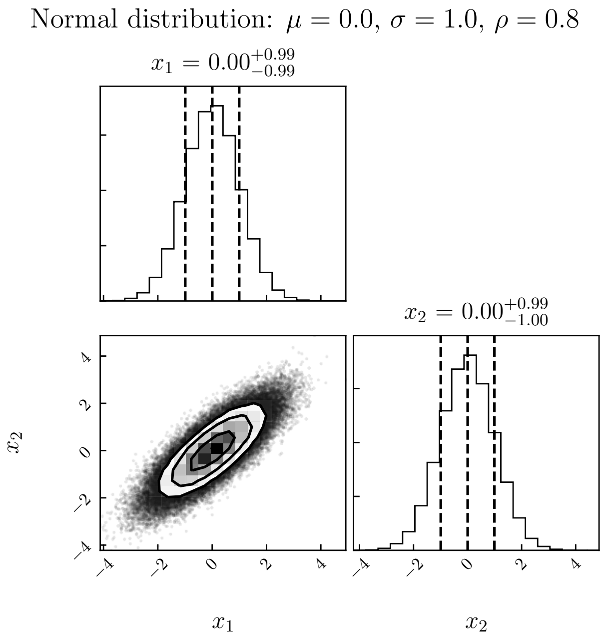


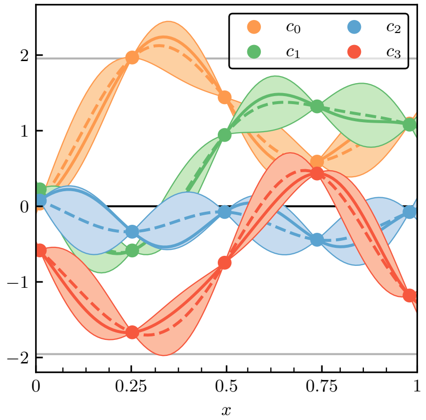


Toy model for an observable *y(x);* e.g., cross section vs. energy, order-by-order in EFT.

Preview of take-away points

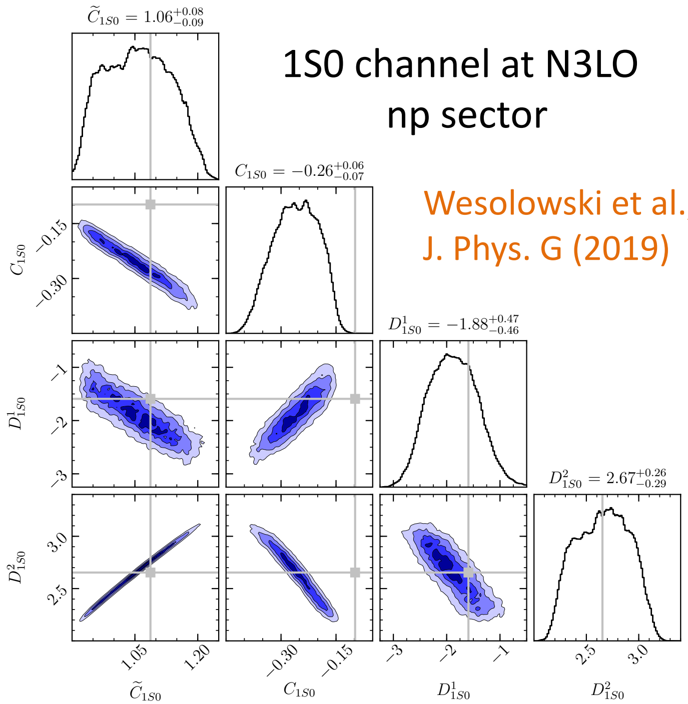
- Bayesian statistics is a powerful framework for (chiral) EFT uncertainty quantification (UQ). *Everything is a pdf.*
- EFT theory *discrepancy model* from the convergence pattern and naturalness *priors*; assumptions explicit.
- *Model checking* is an essential part of Bayesian UQ.
- *Correlations* matter (in many ways). Gaussian process truncation error model for continuous correlations.





Preview of take-away points

- Bayesian statistics is a powerful framework for (chiral) EFT uncertainty quantification (UQ). *Everything is a pdf.*
- EFT theory *discrepancy model* from the convergence pattern and naturalness *priors*; assumptions explicit.
- *Model checking* is an essential part of Bayesian UQ.
- *Correlations* matter (in many ways). Gaussian process truncation error model for continuous correlations.
- Bayesian: *sample* for parameter estimation and the propagation of uncertainties; use *emulators* (like EC)!
- Using priors and truncation errors minimizes overfitting and dependence on how much data is used; posteriors can be used for diagnostics.



1S0 channel at N3LO
np sector

Wesolowski et al., J. Phys. G (2019)

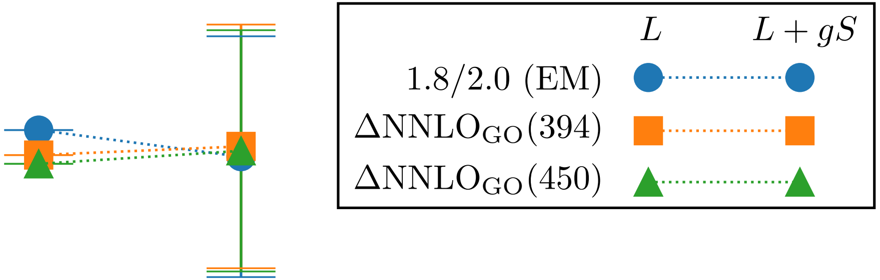
$\mathrm{pr}(\boldsymbol{\theta} \mid \mathbf{y}_{\mathrm{exp}}, \Sigma_{\mathrm{exp}}, \Sigma_{\mathrm{th}}, I) \Rightarrow$ pdf of model parameters $\boldsymbol{\theta}$ given data $\mathbf{y}_{\mathrm{exp}}$ and experiment/theory errors Σ

Xilin Zhang S@INT talk on eigenvector continuation (EC) for scattering on August 13.

Preview of take-away points

- Bayesian statistics is a powerful framework for (chiral) EFT uncertainty quantification (UQ). *Everything is a pdf.*
- EFT theory *discrepancy model* from the convergence pattern and naturalness *priors*; assumptions explicit.
- *Model checking* is an essential part of Bayesi
- *Correlations* matter (in many ways). Gaussia truncation error model for continuous corre
- Bayesian: *sample* for parameter estimation a propagation of uncertainties; use *emulators*
- Using priors and truncation errors minimizes overfitting and dependence on how much data is used; posteriors can be used for diagnostics.
- *Model mixing* is a frontier. Can we use RG invariance?

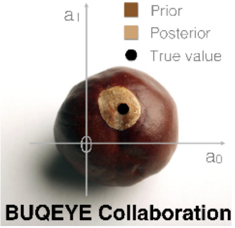
Gaute Hagen's talk on 17-Jul-2020



$$M^{0\nu} = L + gS$$

BUQEYE Collaboration ("Bayesian Uncertainty Quantification: Errors for Your EFT")

The buckeye is the state tree of Ohio! →



https://buqeye.github.io

Papers and software (including Jupyter notebooks for figures)

Prehistory:

INT program in 2003: Theories of Nuclear Forces and Nuclear Systems

M. Schindler + D. Phillips, Bayesian Methods for Parameter Estimation in EFTs (2008-9)

Ohio U. and Ohio State U. revive Bayesian EFT in 2014 → BUQEYE!

Recent highlight: "Effective Field Theory Truncation Errors and Why They Matter," Jordan Melendez PhD thesis (2020)

Types of theory error

- **Uncertainty from numerical method**
- **Truncation of Hilbert space** (e.g., lattice volume/spacing or ho model space)
- **Model discrepancy** ("All models are wrong, but some are useful" – George Box)
  - Incomplete or in parts incorrect physical model
  - Effective field theory expansion truncation error

Mostly systematic and correlated (doesn't mean they can't be treated as random!).

- Systematic: error is not reduced when more observations are averaged
- Correlated: e.g., model error in binding energy for oxygen chain all in same direction

Bayesian statistics is ideal for modeling, combining, and propagating theory errors!

Two ways to treat the theory model discrepancy

Statistical model for observable ***y***: $\boldsymbol{y}_{\rm exp} = \boldsymbol{y}_{\rm th} + \delta\boldsymbol{y}_{\rm th} + \delta\boldsymbol{y}_{\rm exp}$

Advice from statisticians: *any* model for theory discrepancy is better than no model!

1. Model the distribution of residuals: $\boldsymbol{r} \equiv \boldsymbol{y}_{\rm exp} - \boldsymbol{y}_{\rm th}$

- $(\delta\boldsymbol{y}_{\rm exp})_n$ is often a Gaussian with mean $\mu = 0$ and variance σ_n^2: error bars size σ_n
- For $\delta\boldsymbol{y}_{\rm th}$, look at pattern of residuals and *model* it (uncorrelated or correlated).

2. For effective field theories (EFT), learn from *convergence pattern*

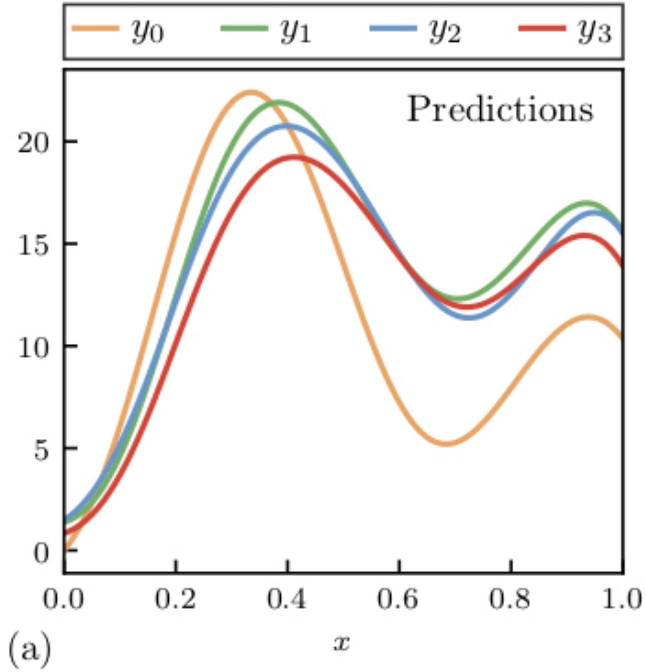
- Expect that each order will *roughly* improve by expansion parameter $Q < 1$:

Theory at order k: $\boldsymbol{y}_k = \boldsymbol{y}_{\rm ref} \sum_{n=0}^{k} c_n Q^n$ Omitted orders: $\delta\boldsymbol{y}_{\rm th} = \boldsymbol{y}_{\rm ref} \sum_{n=k+1}^{\infty} c_n Q^n$

- **Treat the c_ns as random variables and learn their distribution from calculated orders**

Precursors of Bayesian EFT truncation error modeling

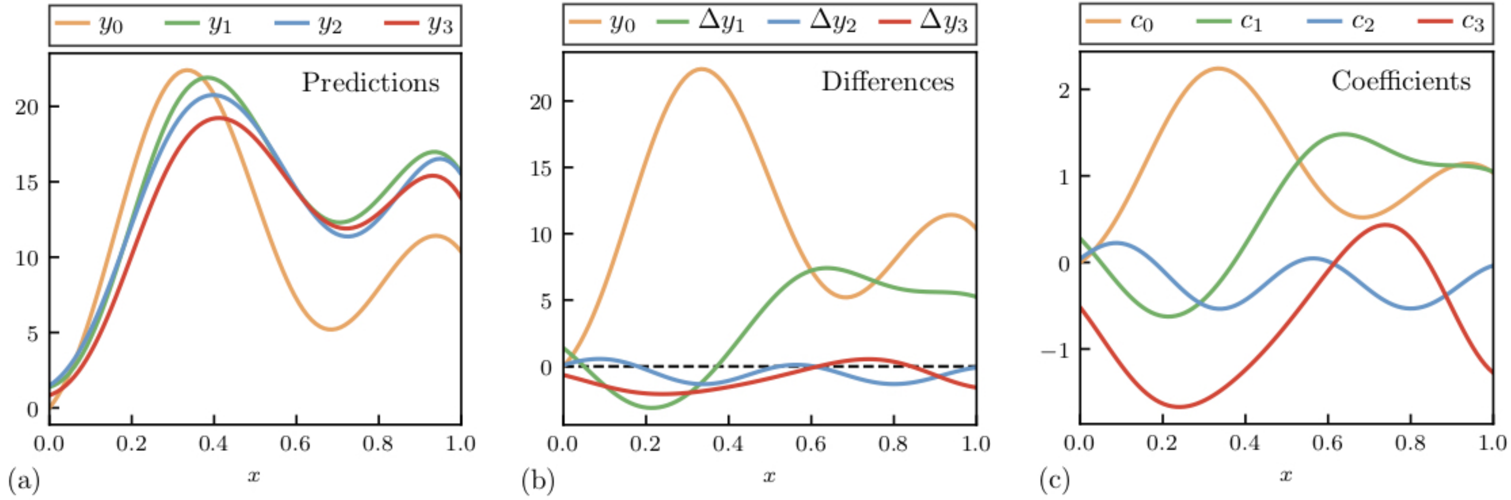
- **EFT expectations for expansions (cf. using cutoff dependence)**
  - Pionless EFT [e.g, McGovern, Grießhammer, Phillips (2013); many others]
  - Epelbaum, Krebs, Meißner, EPJA (2015)
- **Perturbative QCD expansions [Cacciari and Houdeau, JHEP 1109 (2011) 039]**
  - Bayesian error estimates using convergence pattern rather than scale dependence
  - Recent: M. Bonvini, arXiv:2006.16293



Toy model for an observable *y(x);* e.g., cross section vs. energy, order-by-order in EFT.

Bayesian approach formalizes the EFT expectations → statistical; model checking

Coefficients for Bayesian EFT truncation model (not LECs!)



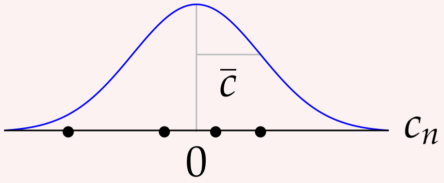
- Order-by-order predictions of *y*: $y_{\rm th}(x) = y_0 \to y_1 \to \cdots \to y_k$
- Focus on differences: $\Delta y_n = y_n - y_{n-1}$ → rescale by reference and Q^n : $c_n \equiv \dfrac{\Delta y_n}{y_{\rm ref} Q^n}$
- Treat c_ns (*not* LECs!!) as random variables and learn from calculated orders

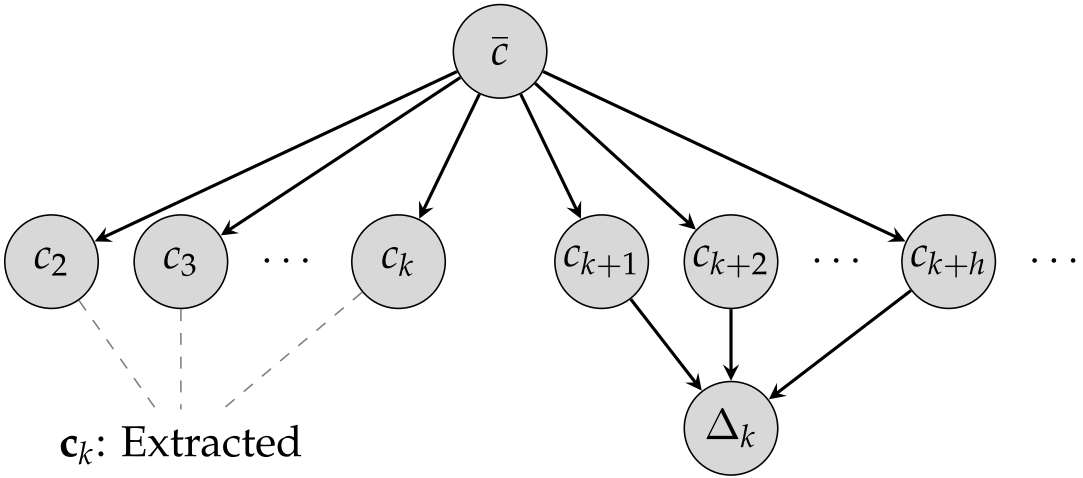
$$\boldsymbol{y}_k = \boldsymbol{y}_{\rm ref} \sum_{n=0}^{k} c_n Q^n \;\rightarrow\; \delta \boldsymbol{y}_{\rm th} = \boldsymbol{y}_{\rm ref} \sum_{n=k+1}^{\infty} c_n Q^n \qquad \chi{\rm EFT} \Rightarrow Q = \frac{\{p, m_\pi\}}{\Lambda_b}, \quad \Lambda_b \approx 600\,{\rm MeV}$$

Assumption: behavior of c_ns persists across orders with characteristic size (natural)

Key Assumption: Naturalness bound

All c_n are drawn from the *same* distribution with a natural size $\bar{c}$



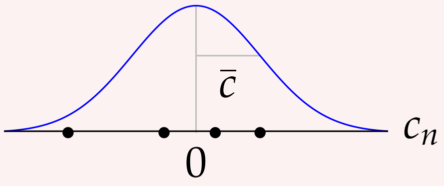


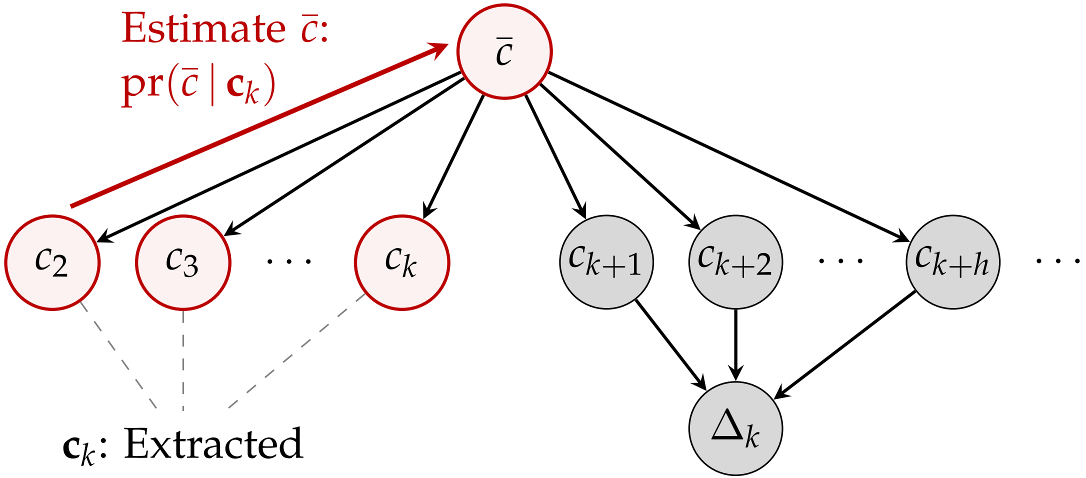
$\delta y_{\mathrm{th}} = y_{\mathrm{ref}} \Delta_k$
at order k

$\mathrm{pr}(\Delta \mid \mathbf{c}_k) \propto$

Key Assumption: Naturalness bound

All c_n are drawn from the *same* distribution with a natural size $\bar{c}$





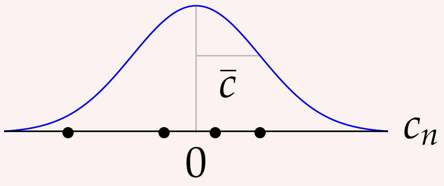
$\delta y_{\rm th} = y_{\rm ref} \Delta_k$

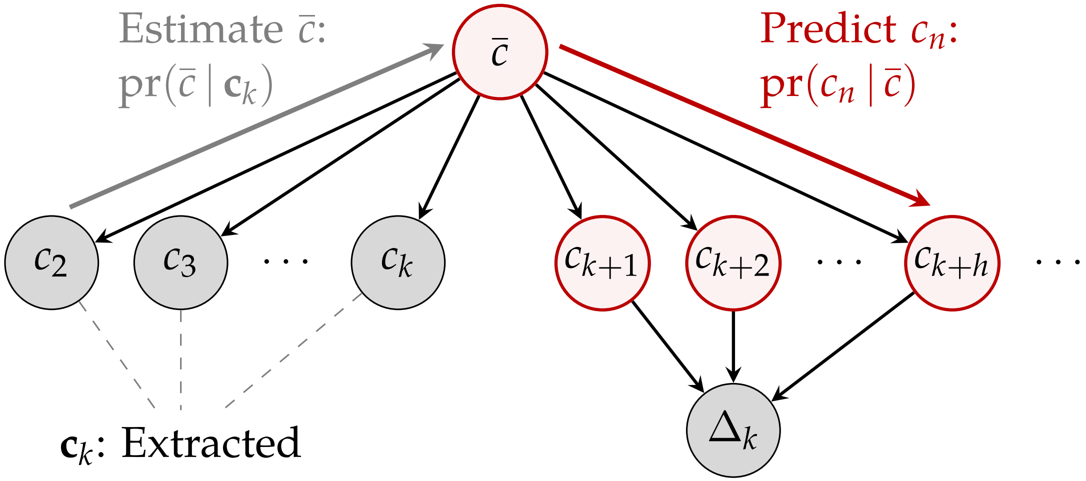
at order k

$$\mathrm{pr}(\Delta \mid \mathbf{c}_k) \propto \int \mathrm{d}\bar{c}\, \mathrm{pr}(\bar{c} \mid \mathbf{c}_k)$$

Key Assumption: Naturalness bound

All c_n are drawn from the *same* distribution with a natural size $\bar{c}$



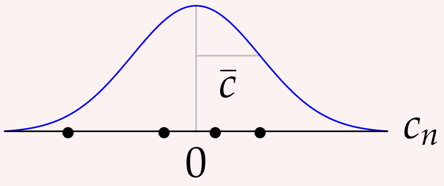


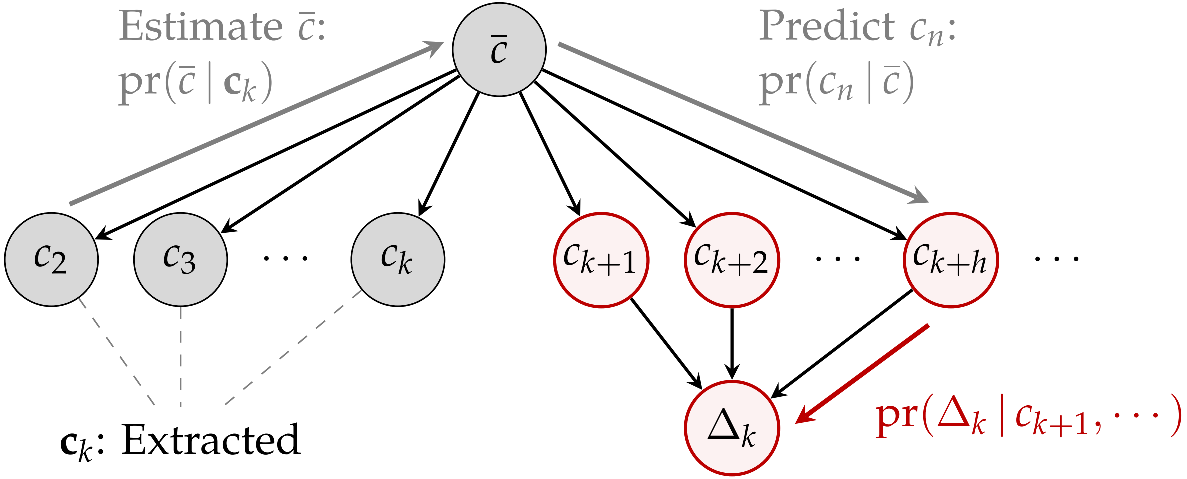
$\delta y_{\mathrm{th}} = y_{\mathrm{ref}} \Delta_k$
at order k

$$\mathrm{pr}(\Delta \mid \mathbf{c}_k) \propto \int d\bar{c}\, \mathrm{pr}(\bar{c} \mid \mathbf{c}_k) \left[\prod_{n=k+1}^{\infty} \int dc_n\, \mathrm{pr}(c_n \mid \bar{c}) \right]$$

Key Assumption: Naturalness bound

All c_n are drawn from the *same* distribution with a natural size $\bar{c}$





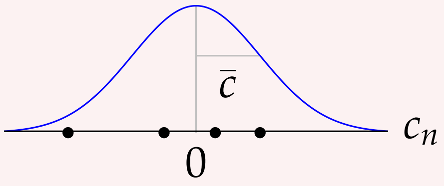
$\delta y_{\mathrm{th}} = y_{\mathrm{ref}} \Delta_k$

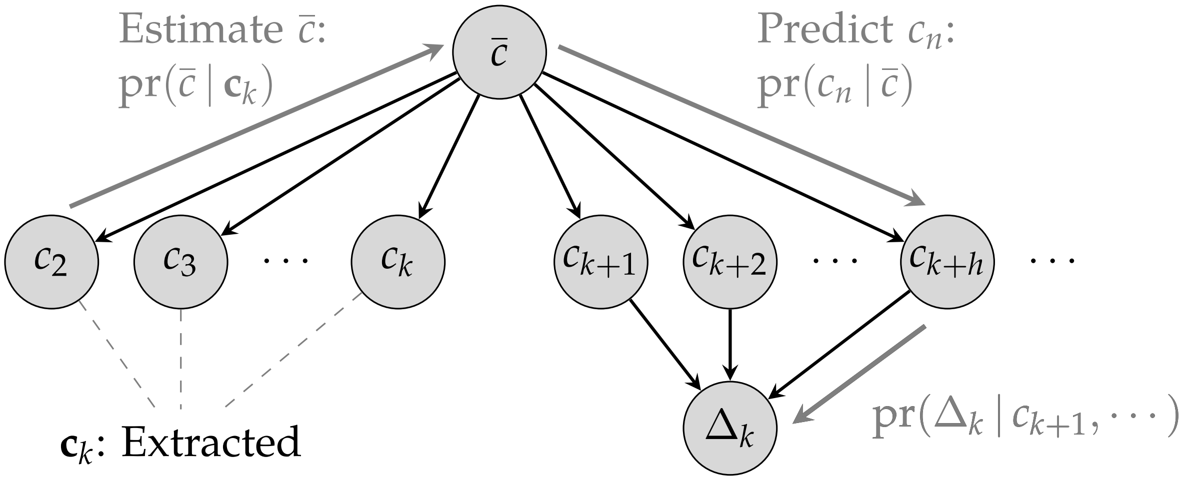
at order k

$$\mathrm{pr}(\Delta \,|\, \mathbf{c}_k) \propto \int d\bar{c}\, \mathrm{pr}(\bar{c} \,|\, \mathbf{c}_k) \left[\prod_{n=k+1}^{\infty} \int dc_n\, \mathrm{pr}(c_n \,|\, \bar{c}) \right] \mathrm{pr}(\Delta_k \,|\, c_{k+1}, \cdots)$$

Key Assumption: Naturalness bound

All c_n are drawn from the *same* distribution with a natural size $\bar{c}$



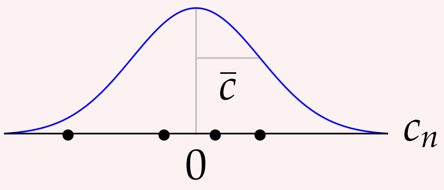


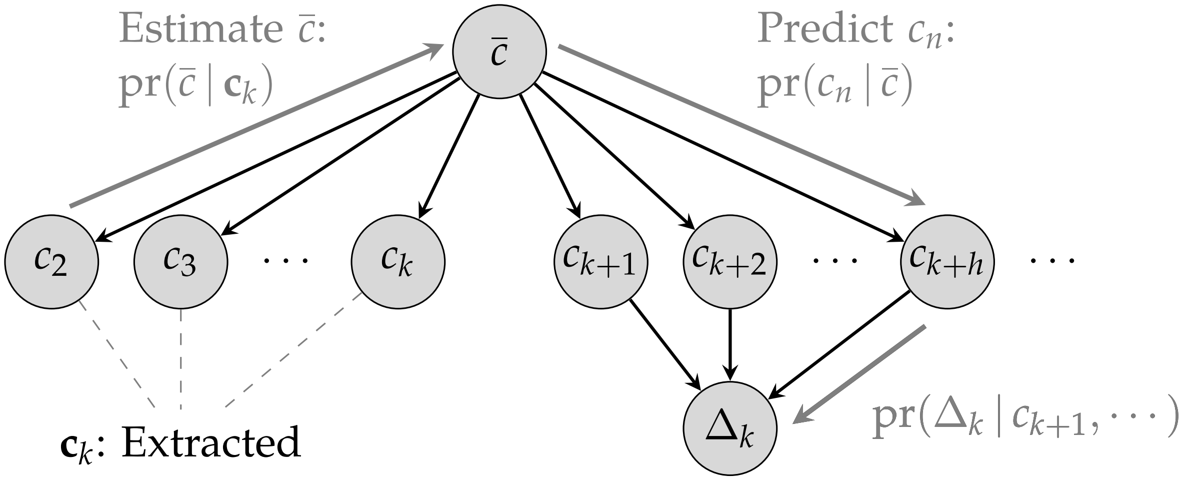
$\delta y_{\mathrm{th}} = y_{\mathrm{ref}} \Delta_k$
at order k

$$\mathrm{pr}(\Delta \mid \mathbf{c}_k) \propto \int d\bar{c}\, \mathrm{pr}(\bar{c} \mid \mathbf{c}_k) \left[\prod_{n=k+1}^{\infty} \int dc_n\, \mathrm{pr}(c_n \mid \bar{c}) \right] \mathrm{pr}(\Delta_k \mid c_{k+1}, \cdots)$$

Key Assumption: Naturalness bound

All c_n are drawn from the *same* distribution with a natural size $\bar{c}$



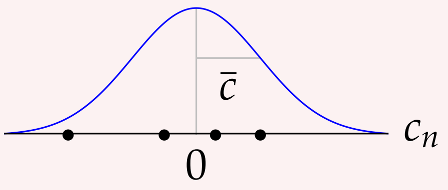


$\delta y_{\mathrm{th}} = y_{\mathrm{ref}}\Delta_k$
at order k

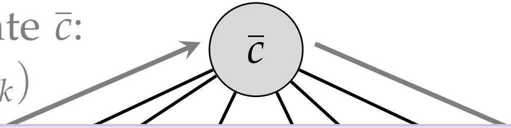
$$\mathrm{pr}(\Delta\,|\,\mathbf{c}_k) \propto \int d\bar{c} \int dc_{k+1}\,\mathrm{pr}(\Delta_k\,|\,c_{k+1})\,\mathrm{pr}(c_{k+1}\,|\,\bar{c})\;\mathrm{pr}(\bar{c}\,|\,\mathbf{c}_k)$$

Key Assumption: Naturalness bound

All c_n are drawn from the *same* distribution with a natural size $\bar{c}$



Estimate $\bar{c}$: $\mathrm{pr}(\bar{c} \mid \mathbf{c}_k)$



Predict c_n: $\mathrm{pr}(c_n \mid \bar{c})$

Tests of priors show insensitivity to details beyond leading orders → use flexible *conjugate* priors:

$$\mathrm{pr}(c_n \mid \bar{c}^2) \implies c_n \mid \bar{c}^2 \overset{iid}{\sim} \mathcal{N}(0, \bar{c}^2)$$

$$\mathrm{pr}(\bar{c}^2) \implies \bar{c}^2 \sim \chi^{-2}(\nu_0, \tau_0^2)$$

All integrals done analytically; intuitive updating.

$$\delta y_{\mathrm{th}} = y_{\mathrm{ref}} \Delta_k$$

at order k

$$\mathrm{pr}(\Delta \mid \mathbf{c}_k) \propto \int d\bar{c} \int dc_{k+1}\, \mathrm{pr}(\Delta_k \mid c_{k+1})\, \mathrm{pr}(c_{k+1} \mid \bar{c})\; \mathrm{pr}(\mathbf{c}_k \mid \bar{c})\, \mathrm{pr}(\bar{c})$$

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

From observable y, extract coefficients

$$\vec{y}_k \equiv \{y_0, y_1, \cdots, y_k\} \quad \text{(A1)}$$
$$\Rightarrow \vec{c}_k \equiv \{c_0, c_1, \cdots, c_k\}$$

Choose ν_0 and τ_0. Update hyperparameters

$$\nu = \nu_0 + n_c \quad \text{(A7)}$$
$$\nu\tau^2 = \nu_0\tau_0^2 + \vec{c}_k^{\,2} \quad \text{(A8)}$$

Compute posterior

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

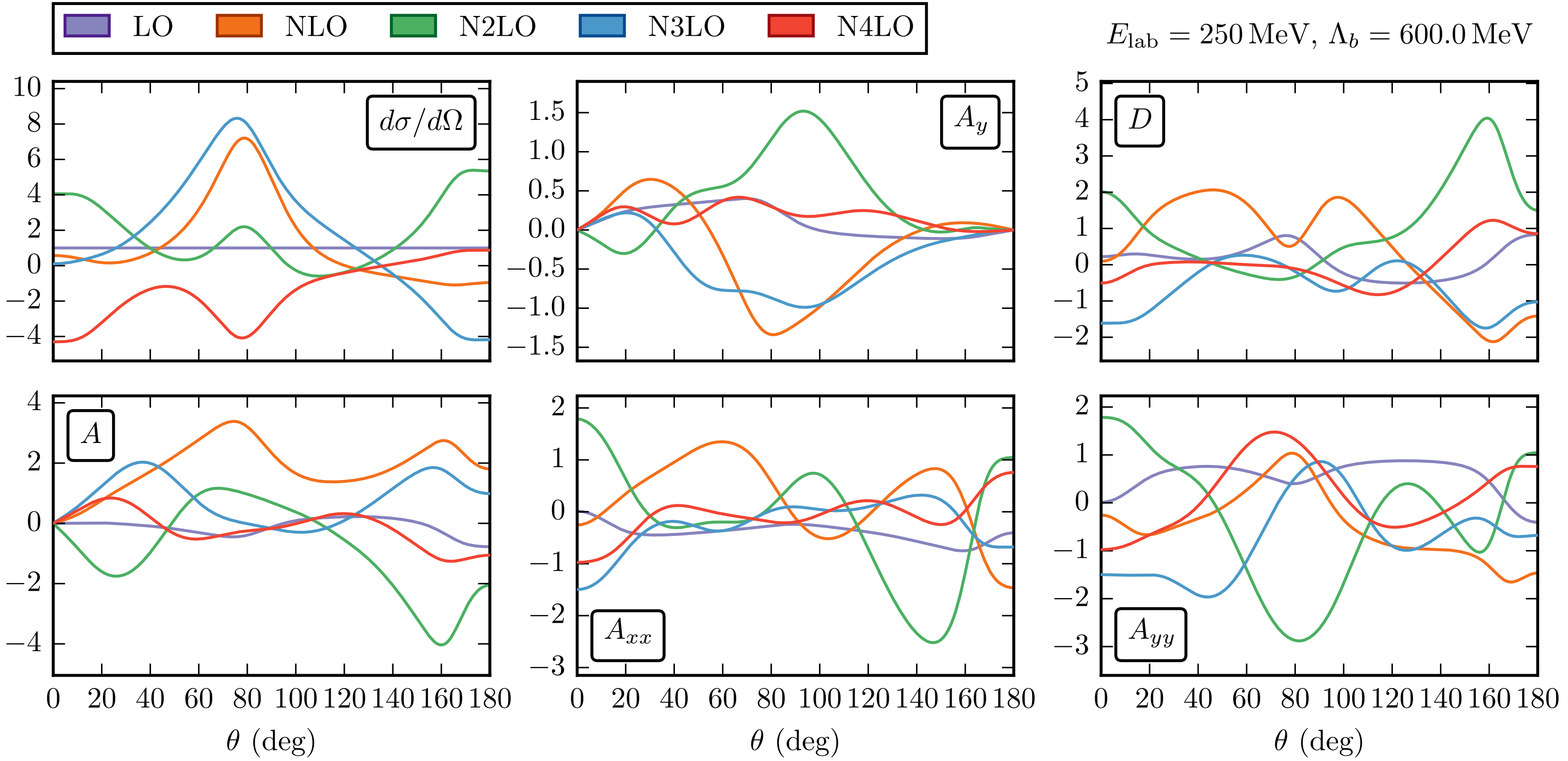
```
import numpy as np
y_ref = 20.0; Q = 0.3; k = 3
y_k = [21.7, 27.3, 25.4, 26.2]
c_k = np.array([y_k[0] / y_ref] + [
  (y_k[n] - y_k[n-1]) / (y_ref * Q**n)
  for n in range(1, k+1)])

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)
```

Note: If $n_c \gg 1$, the posterior for y becomes a normal distribution.

From https://buqeye.github.io/

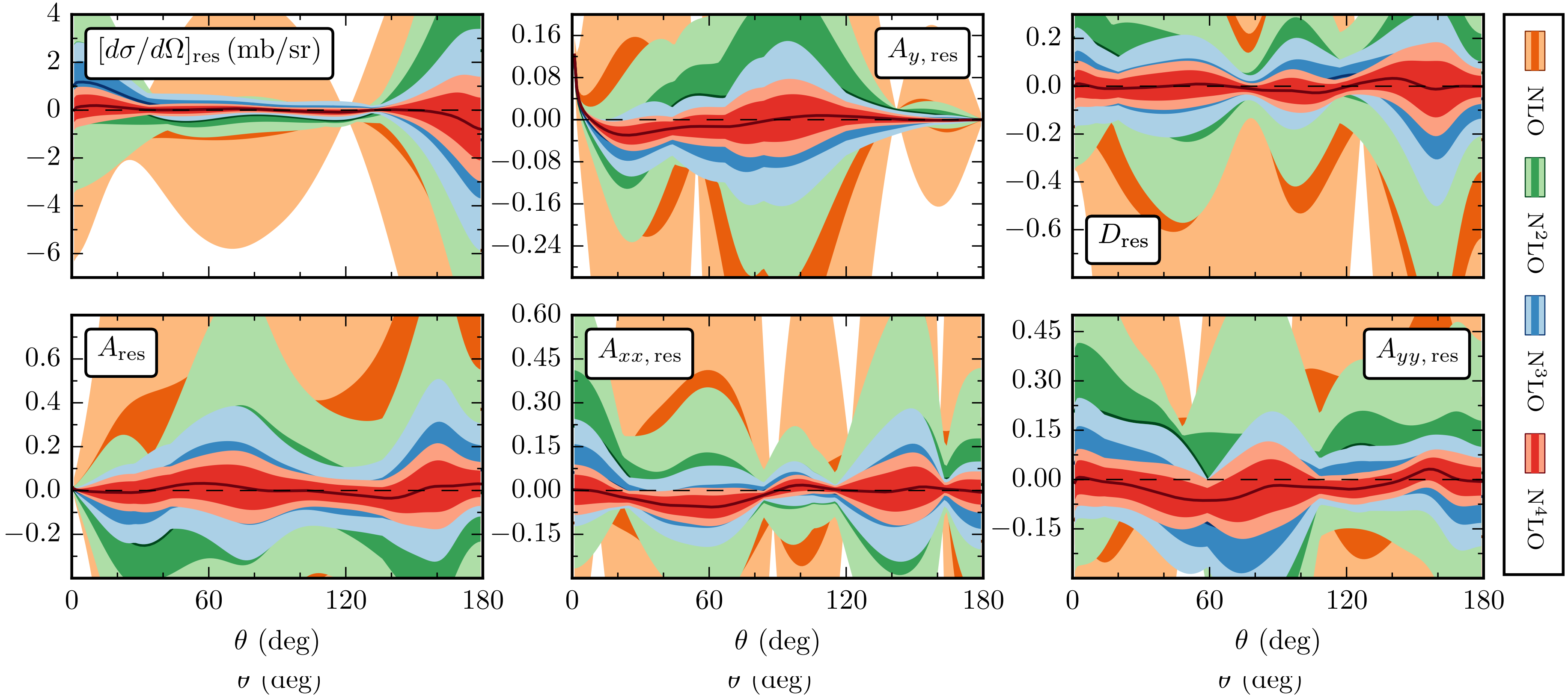
Convergence pattern example from NN observables

J. Melendez et al. PRC (2017)



Convergence pattern example from NN observables

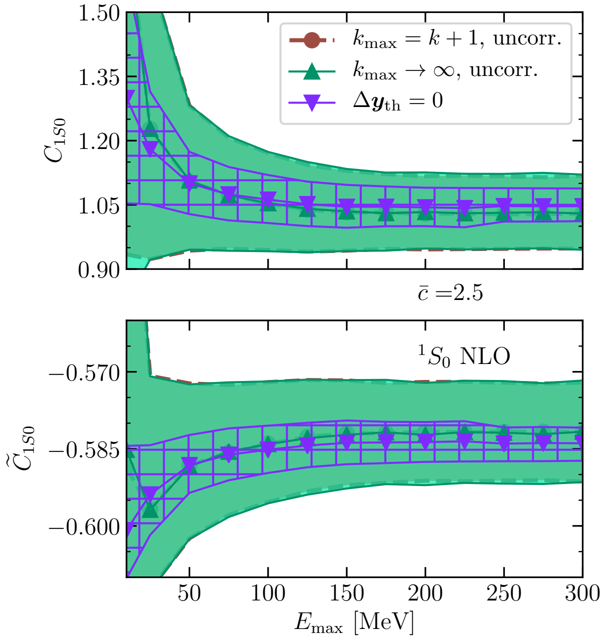
J. Melendez et al. PRC (2017)



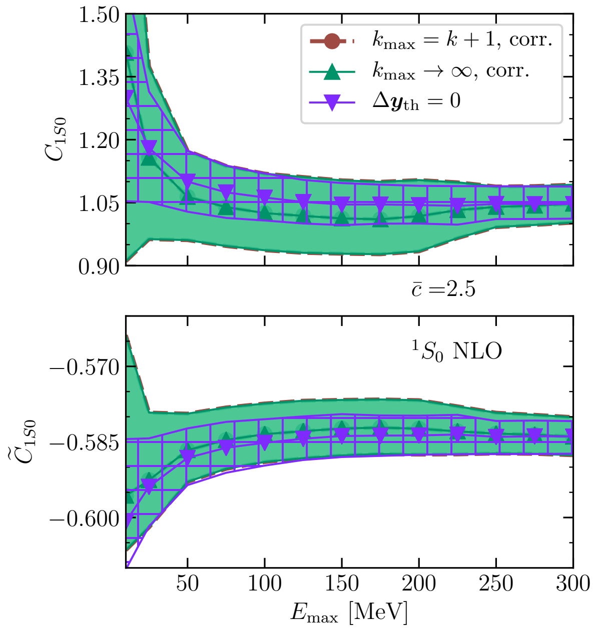
Effect of including truncation errors in parameter estimation

$$\mathrm{pr}(\vec{a}_k|\boldsymbol{y}_{\rm exp},\Sigma_{\rm exp},\Sigma_{\rm th}) \propto e^{-\frac{1}{2}\mathbf{r}^T(\Sigma_{\rm exp}+\Sigma_{\rm th})^{-1}\mathbf{r}} \times e^{-(\vec{a}_k)^2/2\bar{a}^2} \qquad \boldsymbol{r} \equiv \boldsymbol{y}_{\rm exp} - \boldsymbol{y}_{\rm th}$$

If effects of higher-order operators are absorbed by including theory errors, LEC extractions should be independent of E_{max}, the highest-energy datum used:



Uncorrelated assumption



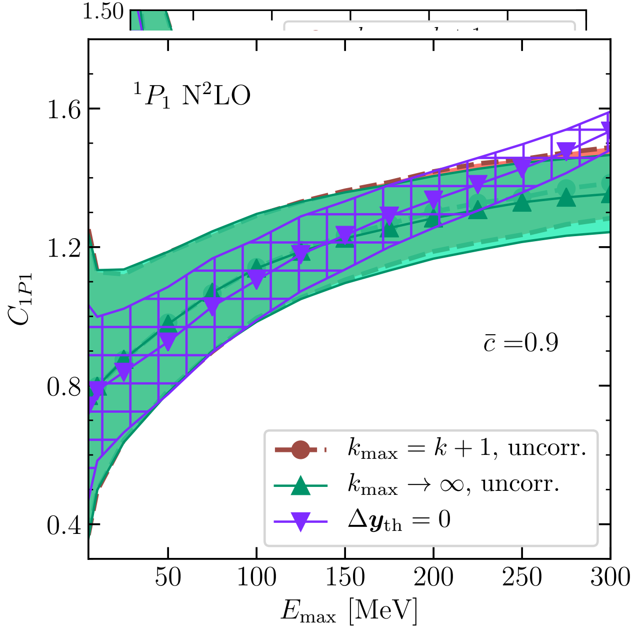
Correlated assumption

Wesolowski et al., J. Phys. G (2019)

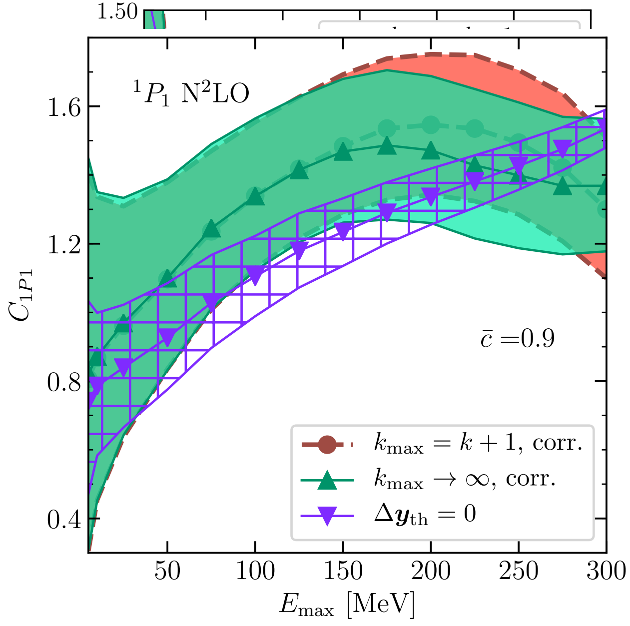
Effect of including truncation errors in parameter estimation

$$\mathrm{pr}(\vec{a}_k|\boldsymbol{y}_{\mathrm{exp}}, \Sigma_{\mathrm{exp}}, \Sigma_{\mathrm{th}}) \propto e^{-\frac{1}{2}\mathbf{r}^T(\Sigma_{\mathrm{exp}}+\Sigma_{\mathrm{th}})^{-1}\mathbf{r}} \times e^{-(\vec{a}_k)^2/2\bar{a}^2} \qquad \boldsymbol{r} \equiv \boldsymbol{y}_{\mathrm{exp}} - \boldsymbol{y}_{\mathrm{th}}$$

If effects of higher-order operators are absorbed by including theory errors, LEC extractions should be independent of E_{max}, the highest-energy datum used:



Uncorrelated assumption



Correlated assumption

Wesolowski et al., J. Phys. G (2019)

Bayesian applications by LENPIC

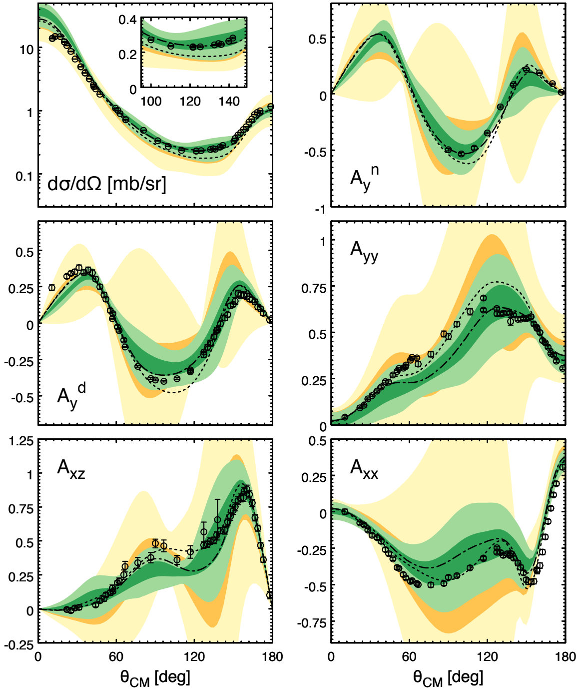
http://www.lenpic.org/

Analytic Bayesian DOB intervals based on Melendez et al., PRC **96**, (2017); cheatsheet at https://buqeye.github.io/

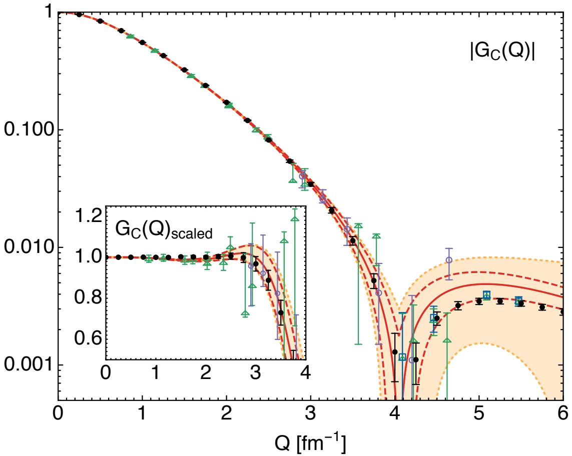
Epelbaum et al, *Towards high-order calculations of three-nucleon scattering in chiral effective field theory*, arXiv:1907.03608, EPJA **56**, 92 (2020).

- SMS potentials + consistent 3N
- 68%/95% bands for Nd obs.
- Also: prior sensitivity tests
- Error analysis → will need ≥N^{4}LO 3N forces for Nd

Observables for elastic nucleon-deuteron scattering at 135 MeV.

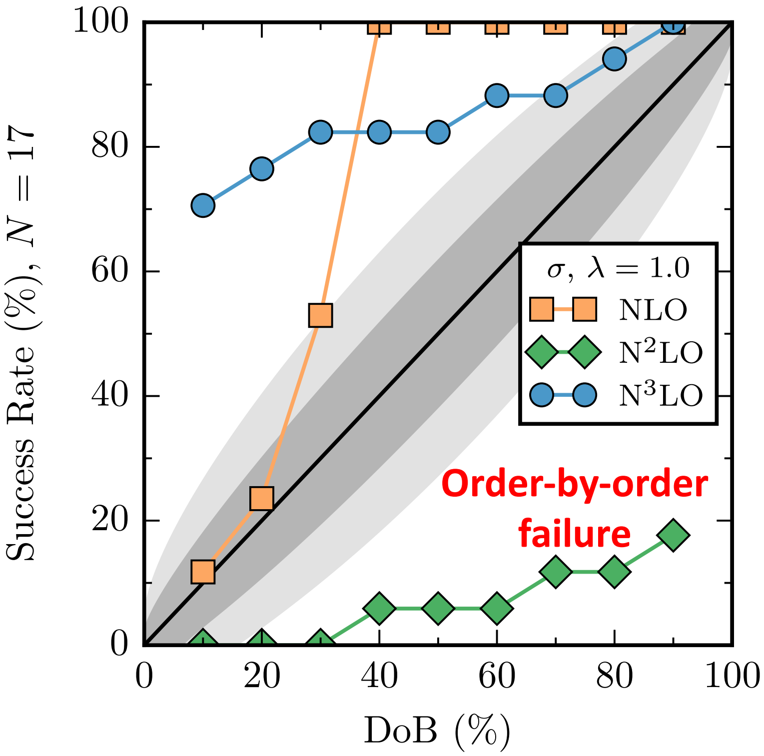


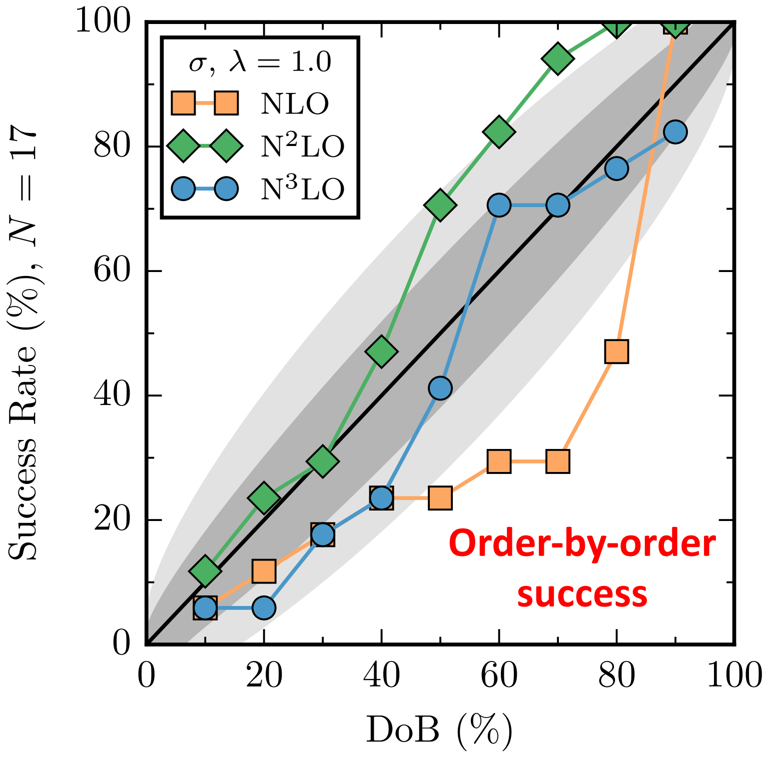
Filin et al., *Extraction of the neutron charge radius from a precision calculation of the deuteron structure radius*, arXiv:1911.04877, PRL (2020).



Colored band is 68% Bayes credible interval for truncation error from convergence pattern.

Model Checking I: Weather plots (empirical coverage)

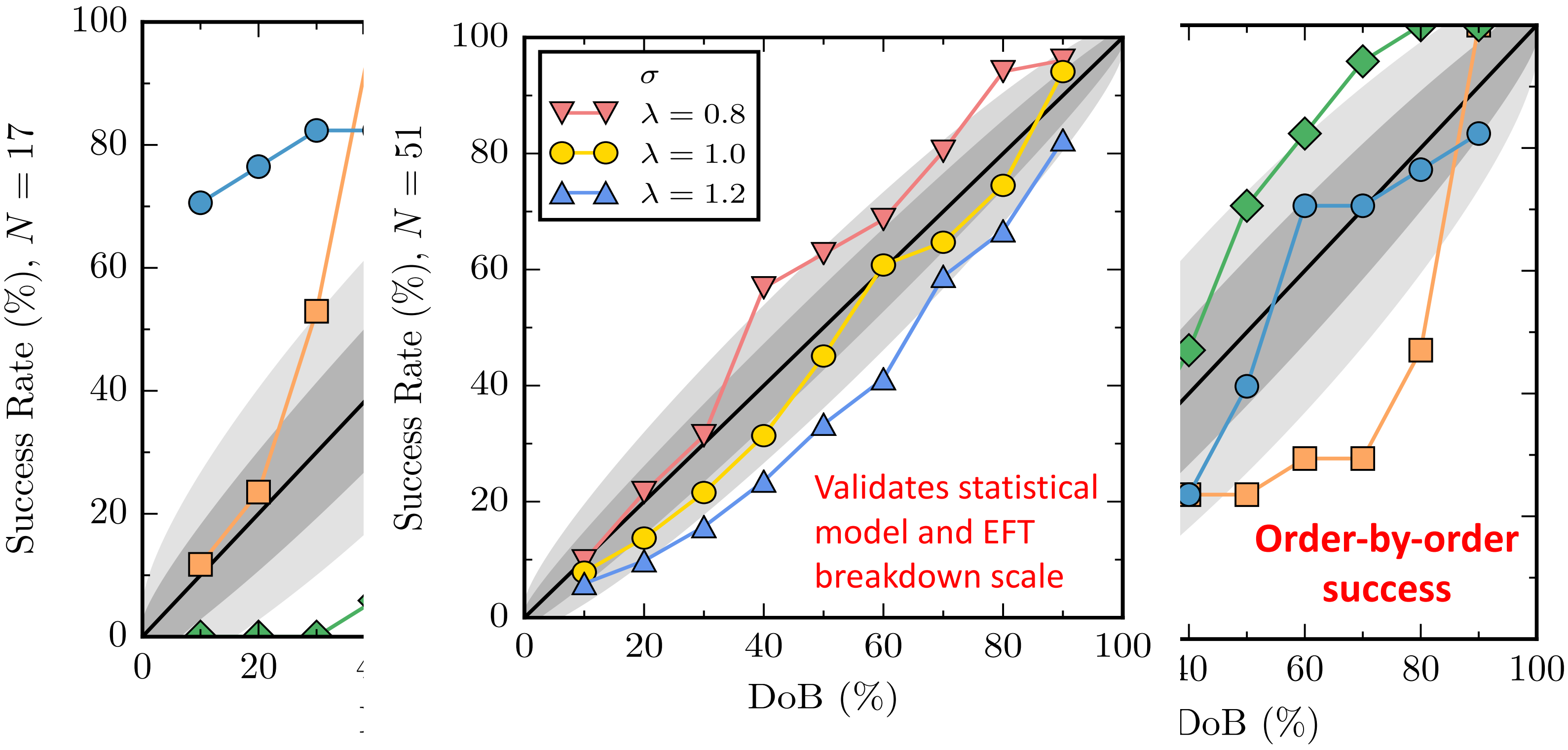




Test of EKM NN chiral EFT potentials from Melendez et al., PRC **96**, 024003 (2017)

In progress (2020): similar analysis of other NN interactions

Model Checking I: Weather plots (empirical coverage)



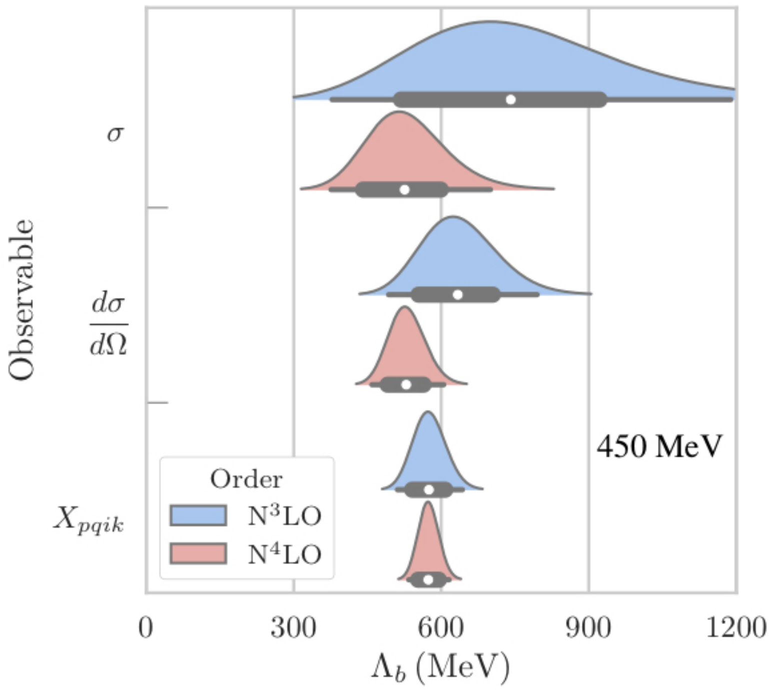
Test of EKM NN chiral EFT potentials from Melendez et al., PRC **96**, 024003 (2017)

In progress (2020): similar analysis of other NN interactions

What is the breakdown scale of our effective field theory (EFT)?

- Chiral EFT relies on an expansion parameter $Q \sim \{p, m_\pi\} / \Lambda_b$
- But what is the breakdown scale Λ_b (where the terms are all the same)?
- We promote Λ_b to a random variable with a posterior pdf (Bayesian!)
- This posterior is predicted as a byproduct of the $\delta \mathbf{y}_{th}$ estimation.

Posteriors from NN observables at many more energies/angles; correlations?

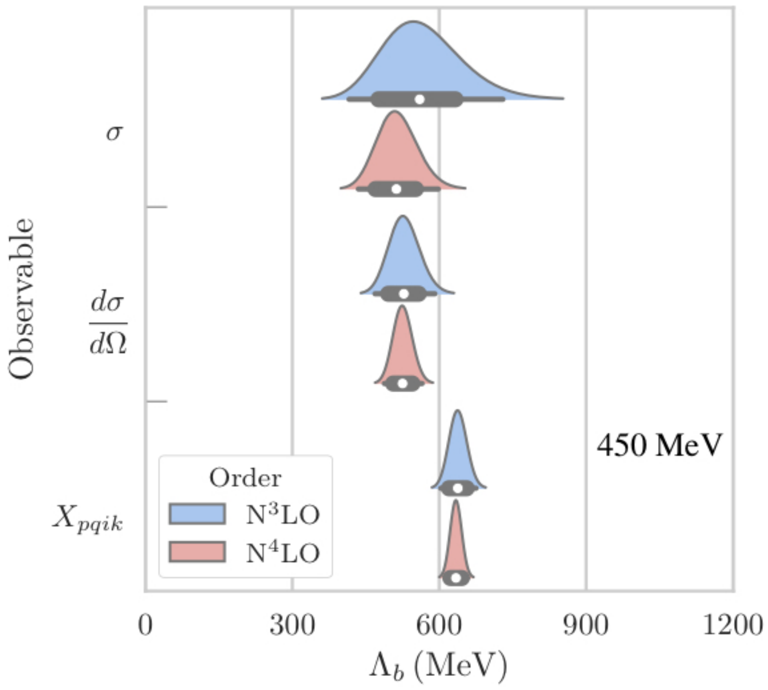


Melendez et al., PRC (2017), PRC (2019), and in preparation (2020)

What is the breakdown scale of our effective field theory (EFT)?

- Chiral EFT relies on an expansion parameter $Q \sim \{p, m_\pi\} / \Lambda_b$
- But what is the breakdown scale Λ_b (where the terms are all the same)?
- We promote Λ_b to a random variable with a posterior pdf (Bayesian!)
- This posterior is predicted as a byproduct of the $\delta \mathbf{y}_{th}$ estimation.

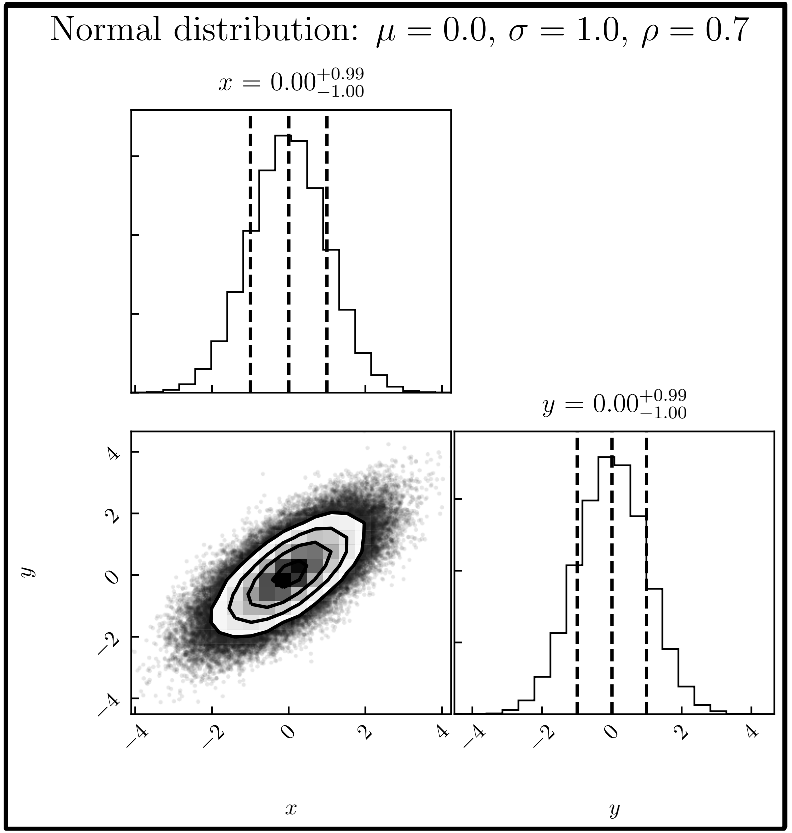
Posteriors from NN observables at many more energies/angles; correlations?



Melendez et al., PRC (2017), PRC (2019), and in preparation (2020)

Reminder about statistical correlations

- pr(*x*, *y* | *z*) “joint probability (density) of *x* and *y* given z” (*contingent* on *z*)



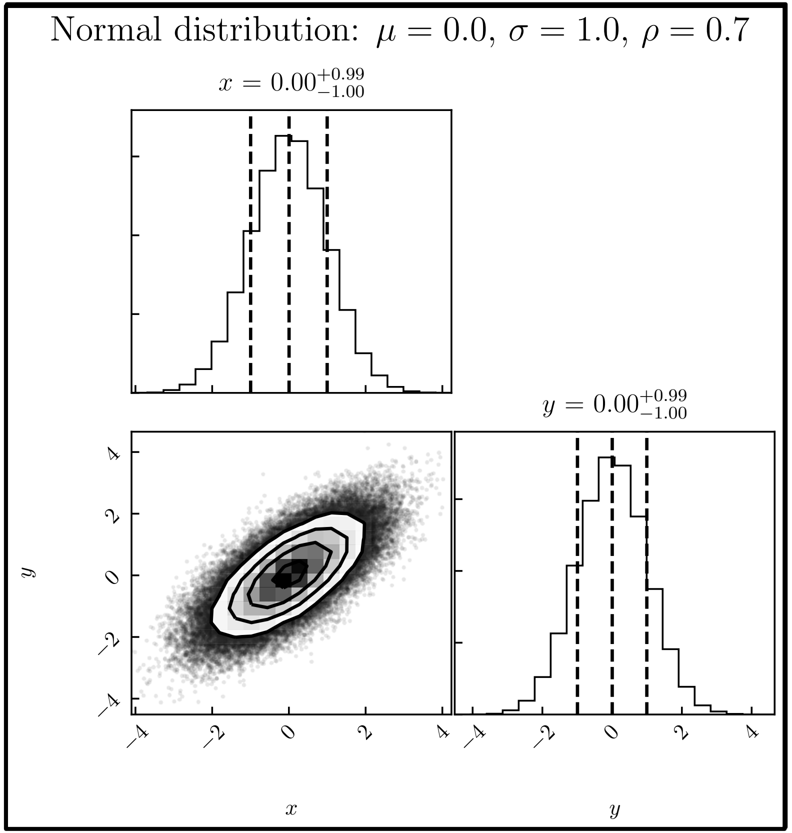
$$\mathcal{N} e^{-\frac{1}{2}\mathbf{r}^\intercal \Sigma^{-1} \mathbf{r}} = \mathcal{N} e^{-\frac{(x-\mu)^2}{2\sigma_x^2}} e^{-\frac{(y-\mu)^2}{2\sigma_y^2}}$$

$$\mathbf{r} = \begin{bmatrix} x \\ y \end{bmatrix} \qquad \Sigma = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix}$$

e.g., $X - Y \sim \mathcal{N}(\mu_x - \mu_y, \sigma_x^2 + \sigma_y^2)$

Reminder about statistical correlations

- pr(*x*, *y* | *z*) "joint probability (density) of *x* and *y* given z" (*contingent* on *z*)



$$\mathcal{N} e^{-\frac{1}{2}\mathbf{r}^\intercal \Sigma^{-1}\mathbf{r}} = \text{correlated gaussian}$$

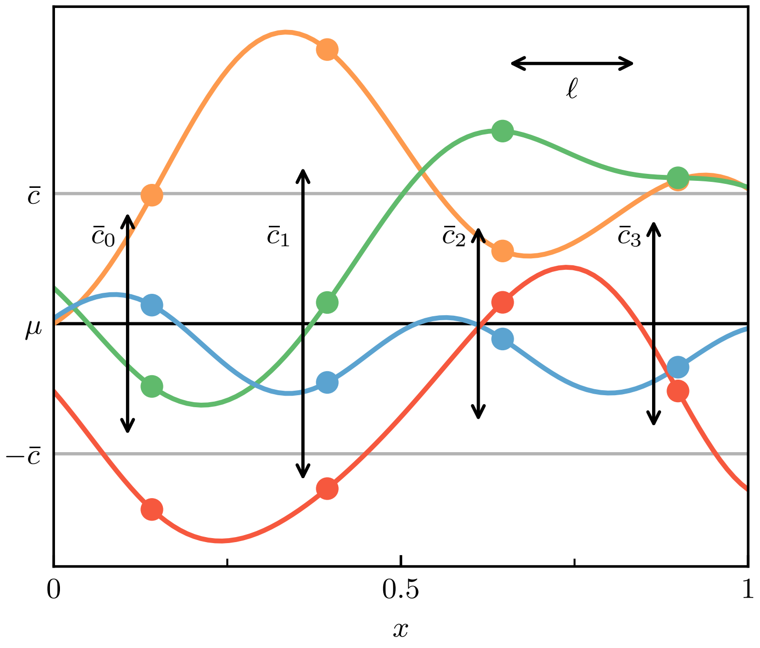
$$\mathbf{r} = \begin{bmatrix} x \\ y \end{bmatrix} \qquad \Sigma = \begin{bmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{bmatrix}$$

With two points *x* and *y*, $-1 \le \rho \le 1$ → correlation. With many points $x_1, x_2, \ldots x_N$, all pairs have a ρ_{ij} correlation to be learned.

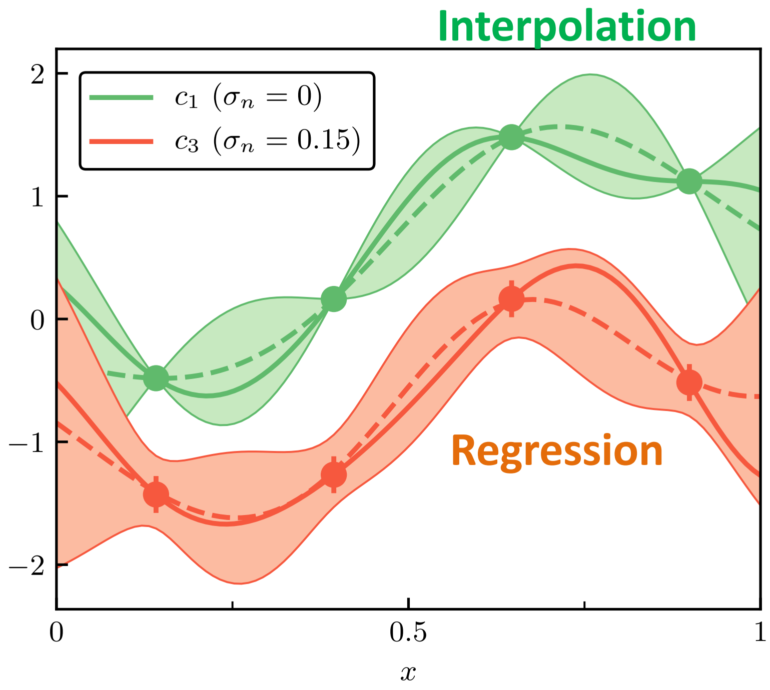
$$X - Y \sim \mathcal{N}(\mu_x - \mu_y, \sigma_x^2 + \sigma_y^2 - 2\sigma_x\sigma_y\rho)$$

Correlations with nearby points: Gaussian processes (GPs)

Start with a correlated Gaussian with two points x and y, where $-1 \leq \rho \leq 1$ gives correlation.
Now imagine many points $x_1, x_2, \ldots x_N$ all of which have a ρ_{ij} correlation given by a function.



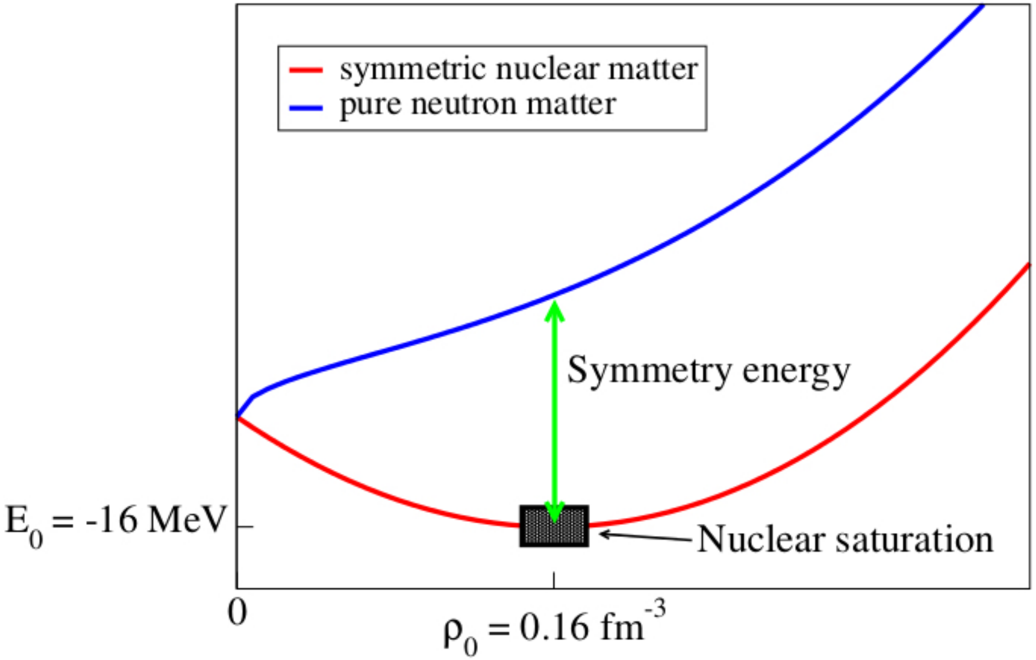
Pointwise $\bar{c}_i$s vs. GP with learned $\bar{c}$ and l

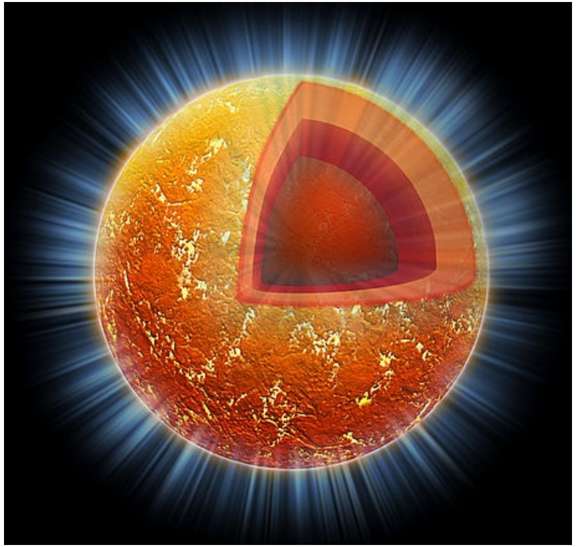


J. Melendez et al., PRC (2019)

Correlated theory errors for infinite matter (arXiv:2004.07232, 2004.07805)

See Christian Drischler S@INT talk (May 28, 2020)





Neutron star

What are the constraints on the infinite matter equation of state from microscopic calculations, accounting for correlations?

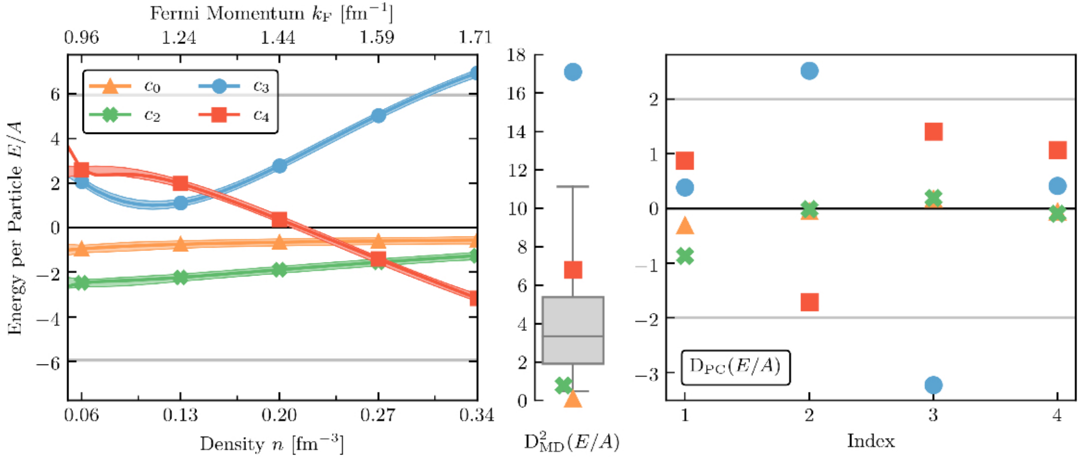
[Also see Drischler et al. (2017, 2019); Gandolfi et al. (2019), Lonardoni et al. (2019), …]

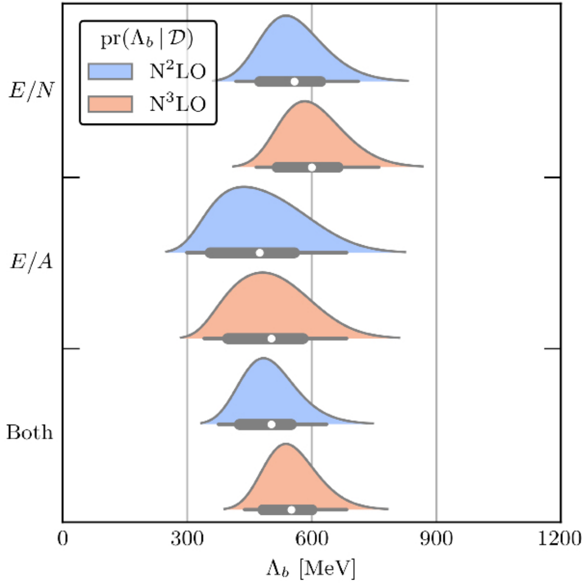
Correlated theory errors for infinite matter

Christian Drischler, Jordan Melendez, rjf, Daniel Phillips (arXiv:2004.07805)

NN potential: Entem-Machleidt-Nosyk (450/500 MeV) with 3N LECs fitted to saturation

1. Apply many-body perturbation theory at each order in χEFT.
2. Learn hyperparameters for GPs from coefficient functions.
3. Propagate uncertainties for derivative quantities (with GPs).
4. Model checking tests how well the GP model works.

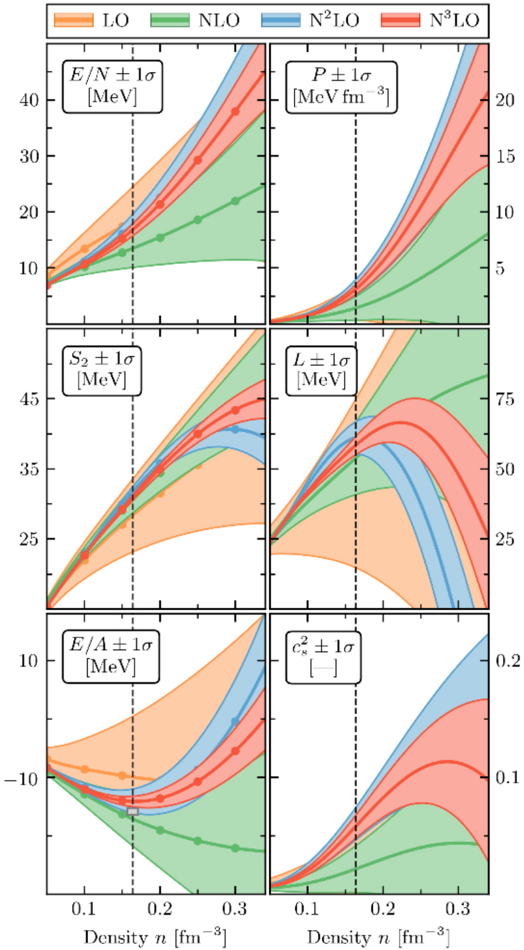




See Melendez et al., PRC (2019)

Find posteriors for breakdown scale

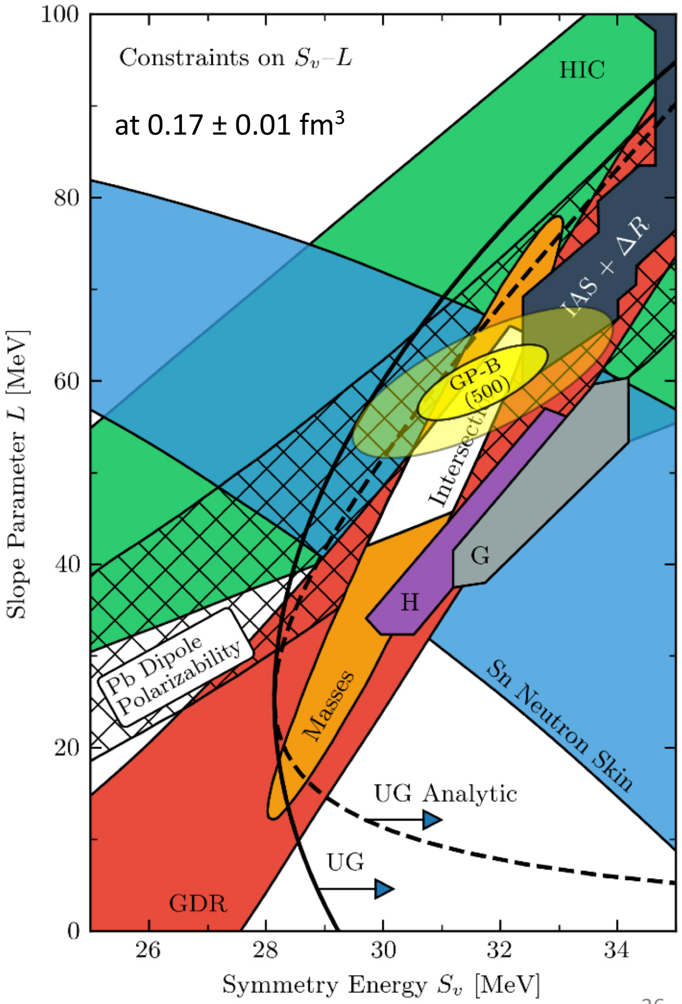
Correlated theory errors for EOS properties



C. Drischler et al. (arXiv:2004.07232)

Correlated GP treatment gives better estimates for truncation errors and clean propagation of uncertainties to derived quantities.

See also comparisons to GW and NICER posteriors!



Slide 17 from Gaute Hagen’s talk on 17-Jul-2020

Can we say something useful about the leading-order contact?

(your feedback appreciated)

RG invariance of the *NN* system requires a leading-order contact term to enter 0vbb
[V. Cirigliano, W. Dekens, J. de Vries, M. L. Graesser, E. Mereghetti, S. Pastore, U. van Kolck PRL (2019)]:

$$V_c(\mathbf{r_{12}}) = g\delta(\mathbf{r}_{12})\tau_-^{(1)}\tau_-^{(2)}, \quad g = g_0 + g_\Lambda$$

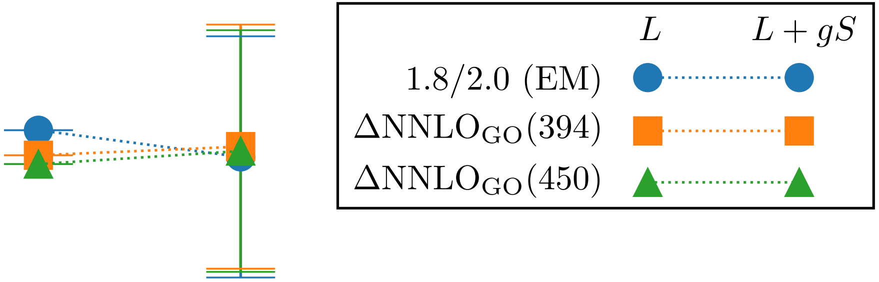
Finite term

”running” term $\approx x\log(\Lambda/\mu)$

- The contact parameter may in the future be determined from QCD
- Can one assess the ”running” part by requiring RG invariance of computed results?
- To play this game, we use a range of interactions in various nuclei and minimize:

$$\chi^2 \equiv \sum_{1\le i<j\le n} \sum_a \left(L_{i,a} + g_i S_{i,a} - L_{j,a} - g_j S_{j,a}\right)^2 \qquad M^{0\nu} = L + gS$$

- The chi-square minimization leads to a linear problem:



- The matrix *A* has an eigenvalue that is almost zero and *l*
- This could be consistent with the lack of information ab
- Assessing the ”running” term of the contact could help

Slide 17 from Gaute Hagen's talk on 17-Jul-2020

Can we say something useful about the

(your feedback apprecia

Experience from SRG running: soft potentials → factorization works very well → matrix element ≈ scale independent → ≈ 0 eigenvalue (?)

Treat as Bayesian model mixing? Rather than model *selection*, combine the predictions of all. Use RG invariance to model δy_{th}?

quires a leading-order contac

s, M. L. Graesser, E. Mereghet

$$V_c(\mathbf{r_{12}}) = g\delta(\mathbf{r}_{12})\tau_-^{(1)}\tau_-^{(2)}, \quad g = g_0 + g_\Lambda$$

Finite term

"running" term $\approx x\log(\Lambda/\mu)$

- The contact parameter may in the future be determined from QCD
- Can one assess the "running" part by requiring RG invariance of computed results?
- To play this game, we use a range of interactions in various nuclei and minimize:

$$\chi^2 \equiv \sum_{1\le i<j\le n}\sum_a (L_{i,a} + g_i S_{i,a} - L_{j,a} - g_j S_{j,a})^2 \qquad M^{0\nu} = L + gS$$

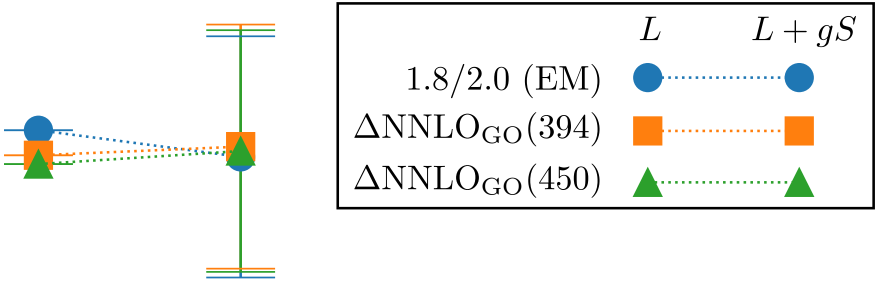
Implied model for discrepancy δy_{th}: contact term plus independent Gaussian noise with same σ, correlated through $[M^{0\nu}]_a$.

r problem:

t zero and

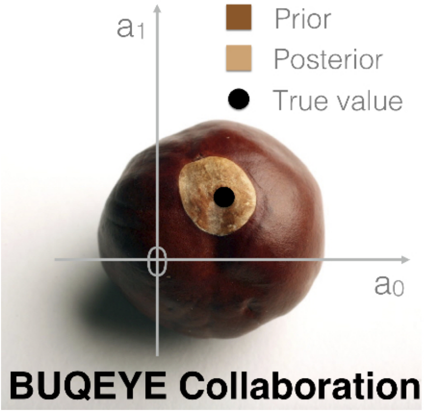
rmation ab

could help



Recap of take-away points

- Bayesian statistics is a powerful framework for (chiral) EFT uncertainty quantification (UQ). *Everything is a pdf.*
- EFT theory *discrepancy model* from the convergence pattern and naturalness *priors*; assumptions explicit.
- *Model checking* is an essential part of Bayesian UQ.
- *Correlations* matter (in many ways). Gaussian process truncation error model for continuous correlations.
- Bayesian: *sample* for parameter estimation and the propagation of uncertainties; use *emulators* (like EC)!
- Using priors and truncation errors minimizes overfitting and dependence on how much data is used; posteriors can be used for diagnostics.
- *Model mixing* is a frontier. Can we use RG invariance?



https://buqeye.github.io

Papers and software (including Jupyter notebook for figures)

New and coming soon: experimental design; EC for scattering; constraints on 3BF LECs; analysis of NN forces; exploiting correlations in nuclei.

Extra slides

Summary: ingredients of EFTs that invite Bayesian statistics

Bayesian approach: (almost) *everything* is a probability density function (pdf)!

$\mathrm{pr}(\mathrm{EFT}_1 \mid \boldsymbol{y}_{exp}, I)/\mathrm{pr}(\mathrm{EFT}_2 \mid \boldsymbol{y}_{exp}, I)$

$\Longrightarrow$ evidence ratio given data

$\mathrm{pr}(c_n(\boldsymbol{x}) \mid \vec{\boldsymbol{c}}_{\boldsymbol{k}}, I) \Longrightarrow$ pdf of expansion coefficients

$\mathrm{pr}(\boldsymbol{\theta} \mid \boldsymbol{y}_{exp}, \Sigma_{\mathrm{exp}}, I) \propto \mathrm{pr}(\boldsymbol{y}_{exp} \mid \boldsymbol{\theta}, \Sigma_{\mathrm{exp}}, I)\, \mathrm{pr}(\boldsymbol{\theta} \mid I)$

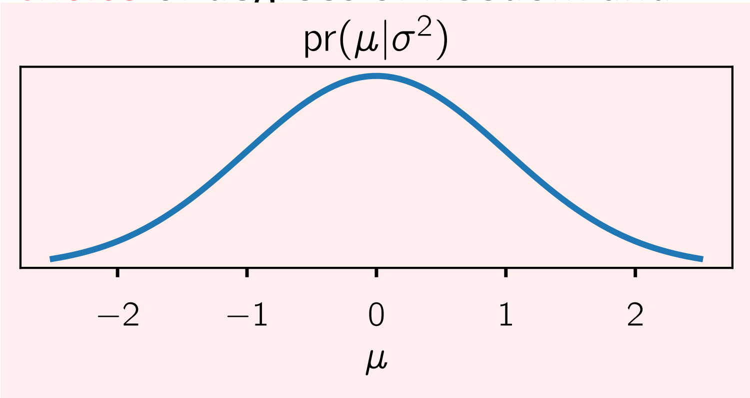
$\Longrightarrow$ pdf of LECs given data $\boldsymbol{y}_{exp}, \Sigma_{\mathrm{exp}}$ (and Σ_{th}!)

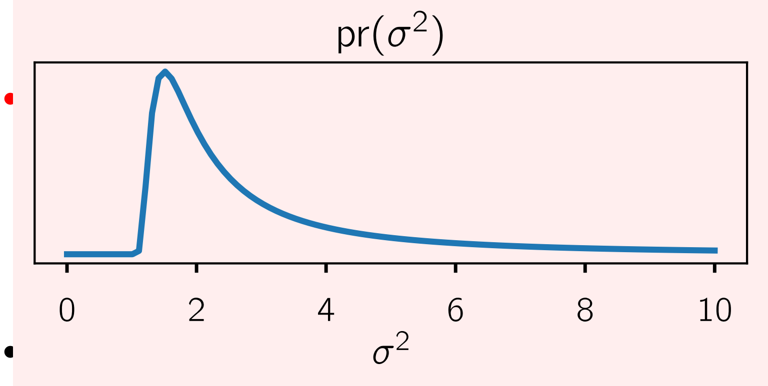
$\mathrm{pr}(\Lambda_b \mid \boldsymbol{\theta}, I) \Longrightarrow$ pdf of EFT breakdown scale

$\mathrm{pr}(\mu, \sigma \mid I) \Longrightarrow$ pdf of hyperparameters

Use rules of probability to manipulate.
Marginalize over nuisance parameters.

- *Choice* of degrees of freedom and





provide robust theory error bars

State of knowledge as probability distributions (pdfs)

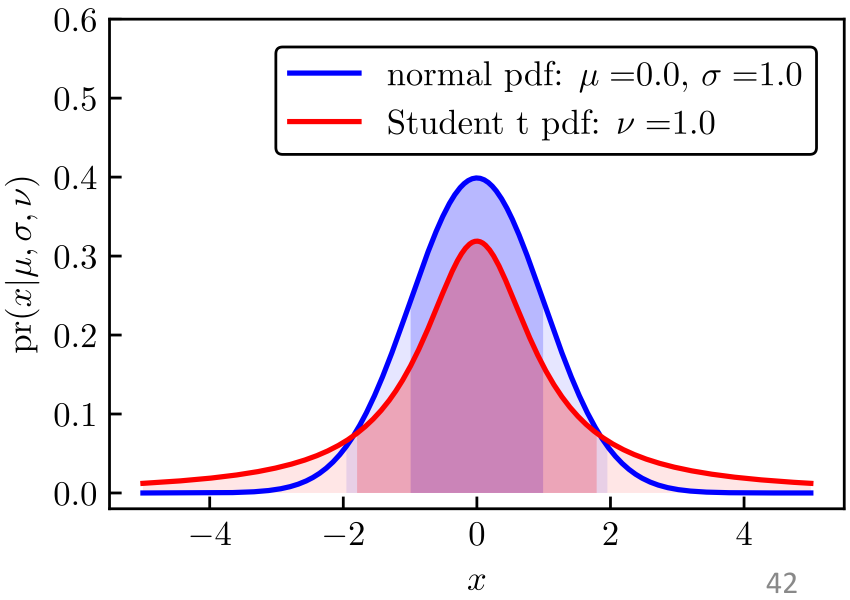
- pr(A, B | C) "joint probability (density) of A and B given C" (*contingent* on C)
- A, B, C can be observables, parameters, uncertainties, propositions, models, …
- cf. quantum mechanics $|\psi(x, y)|^2$ or $|\psi(x)|^2 = \int |\psi(x, y)|^2 \, dy$ (*marginalization*)
- Bayesian confidence (credible) interval:

$$\mathrm{pr}(a \le x \le b) = \int_a^b |\psi(x)|^2 \, dx$$

Examples of pdfs for theory UQ:

$\mathrm{Pr}(\boldsymbol{\theta} \mid \mathbf{y}_{\mathrm{exp}}, \Sigma_{\mathrm{exp}}, \Sigma_{\mathrm{th}}, I) \Rightarrow$

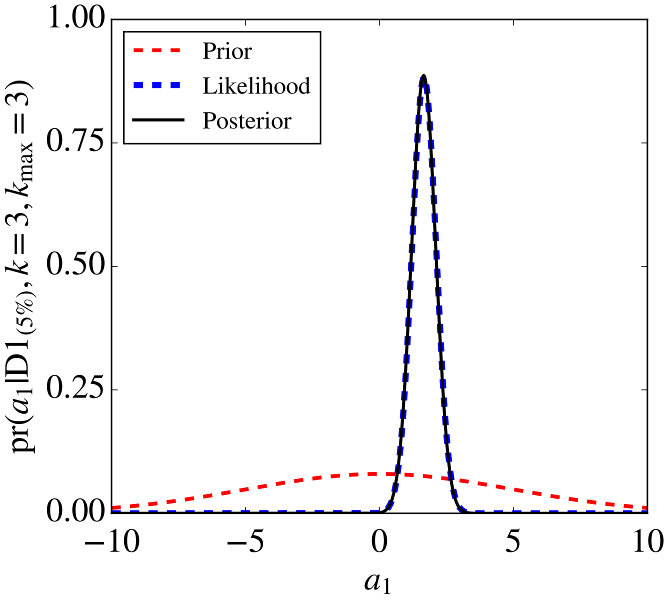
pdf of model parameters $\boldsymbol{\theta}$ given data y_{exp} and experiment/theory errors Σ, plus other information I



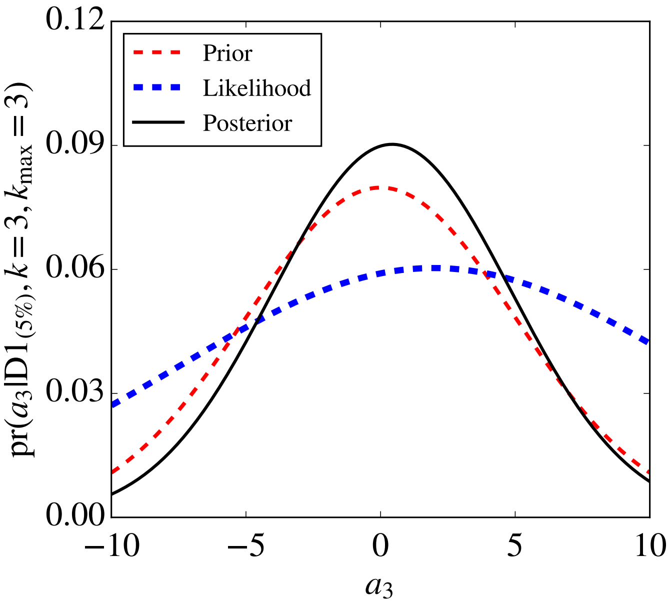
Bayes's Theorem: How to update knowledge in PDFs

$$\mathrm{pr}(A|B,I) = \frac{\mathrm{pr}(B|A,I)\,\mathrm{pr}(A|I)}{\mathrm{pr}(B|I)} \Longrightarrow \underbrace{\mathrm{pr}(\boldsymbol{\theta}|\mathbf{y}_{\mathrm{exp}},I)}_{\text{posterior}} \propto \underbrace{\mathrm{pr}(\mathbf{y}_{\mathrm{exp}}|\boldsymbol{\theta},I)}_{\text{likelihood}} \times \underbrace{\mathrm{pr}(\boldsymbol{\theta}|I)}_{\text{prior}}$$

Likelihood overwhelms prior



Prior suppresses unconstrained likelihood



Bayes's Theorem: How to update knowledge in PDFs

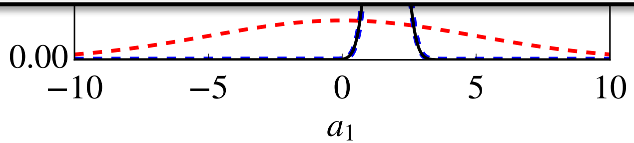
$$\mathrm{pr}(A|B,I) = \frac{\mathrm{pr}(B|A,I)\,\mathrm{pr}(A|I)}{\mathrm{pr}(B|I)} \Longrightarrow \underbrace{\mathrm{pr}(\boldsymbol{\theta}|\mathbf{y}_{\mathrm{exp}},I)}_{\text{posterior}} \propto \underbrace{\mathrm{pr}(\mathbf{y}_{\mathrm{exp}}|\boldsymbol{\theta},I)}_{\text{likelihood}} \times \underbrace{\mathrm{pr}(\boldsymbol{\theta}|I)}_{\text{prior}}$$

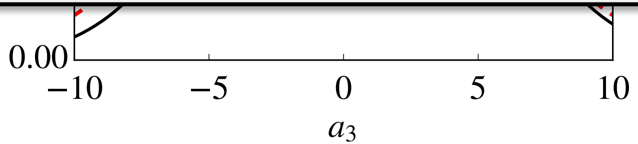
Likelihood overwhelms prior

Prior suppresses unconstrained likelihood

$$\mathrm{pr}(\mathbf{y}_{\mathrm{exp}}|\boldsymbol{\theta},I) \propto e^{-\chi^2/2} \underset{\text{uncorrelated}}{\Longrightarrow} \chi^2 = \sum_i^{\text{data}} \mathbf{r}_i \frac{1}{\sigma_i^2} \mathbf{r}_i \,, \quad \mathbf{r} \equiv \mathbf{y}_{\mathrm{exp}} - \mathbf{y}_{\mathrm{th}}(\boldsymbol{\theta})$$

$$\text{correlated: } \mathrm{pr}(\boldsymbol{\theta}|\mathrm{y}_{\mathrm{exp}},\Sigma_{\mathrm{exp}},\Sigma_{\mathrm{th}},I) \propto e^{-\frac{1}{2}\mathbf{r}^{\intercal}(\Sigma_{\mathrm{exp}}+\Sigma_{\mathrm{th}})^{-1}\mathbf{r}} \times e^{-\boldsymbol{\theta}^2/2\bar{\theta}^2}$$





Parameter estimation: Exploring projected posteriors

- First without the model discrepancy (theory error) term:

$$\boldsymbol{y}_{\rm exp} = \boldsymbol{y}_{\rm th} \qquad + \delta \boldsymbol{y}_{\rm exp}$$

- Consider high-enough order so that truncation error is small
- Regular least-squares (χ^2) likelihood times Gaussian prior for natural LECs:

$$\mathrm{pr}(\vec{a}_k | \boldsymbol{y}_{\rm exp}, \Sigma_{\rm exp}) \propto \mathrm{pr}(\boldsymbol{y}_{\rm exp} | \vec{a}_k, \Sigma_{\rm exp})\ \mathrm{pr}(\vec{a}_k)$$

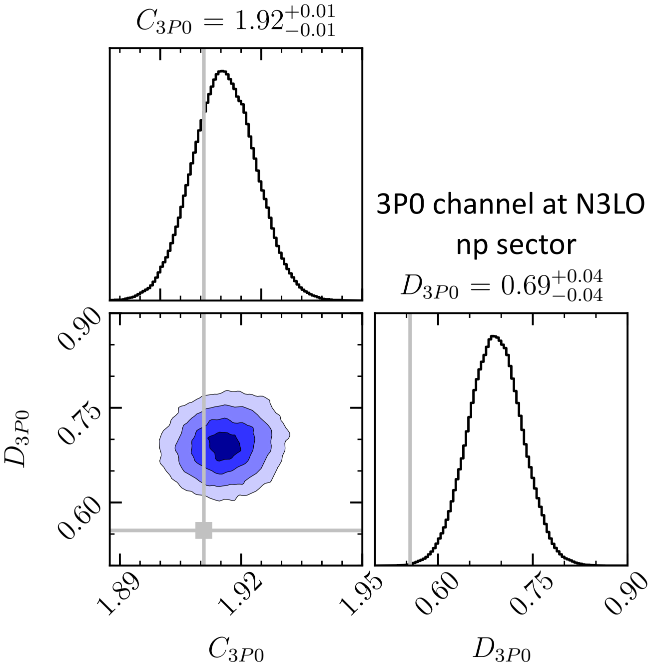
$$\propto e^{-\frac{1}{2}\mathbf{r}^T \Sigma_{\rm exp}^{-1} \mathbf{r}} \times e^{-(\vec{a}_k)^2/2\bar{a}^2}$$

Residual: $\boldsymbol{r} \equiv \boldsymbol{y}_{\rm exp} - \boldsymbol{y}_{\rm th}$

Discourages large LECs (naturalness)

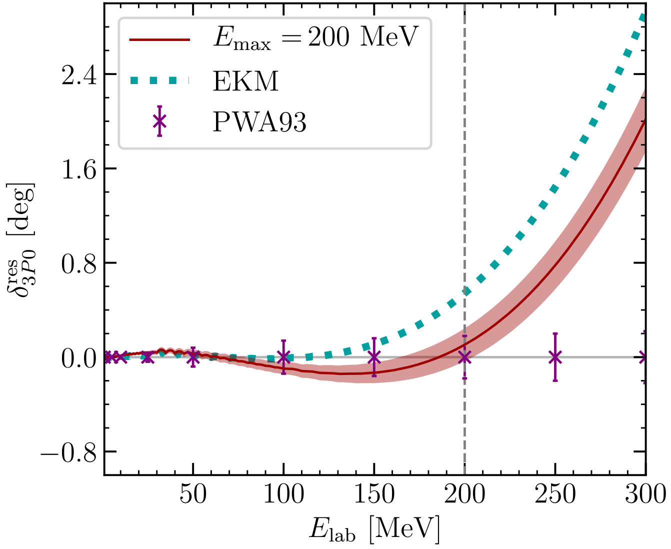
Projected posteriors as a diagnostic tool

$$\mathrm{pr}(\vec{a}|\boldsymbol{y}_{\mathrm{exp}}, \Sigma_{\mathrm{exp}}) \propto \exp\left[-\frac{1}{2}\sum_{i=1}^{N_d}\frac{r_i^2}{\sigma_i^2}\right] \times e^{-(\vec{a}_k)^2/2\bar{a}^2}$$



3P0 channel at N3LO
np sector

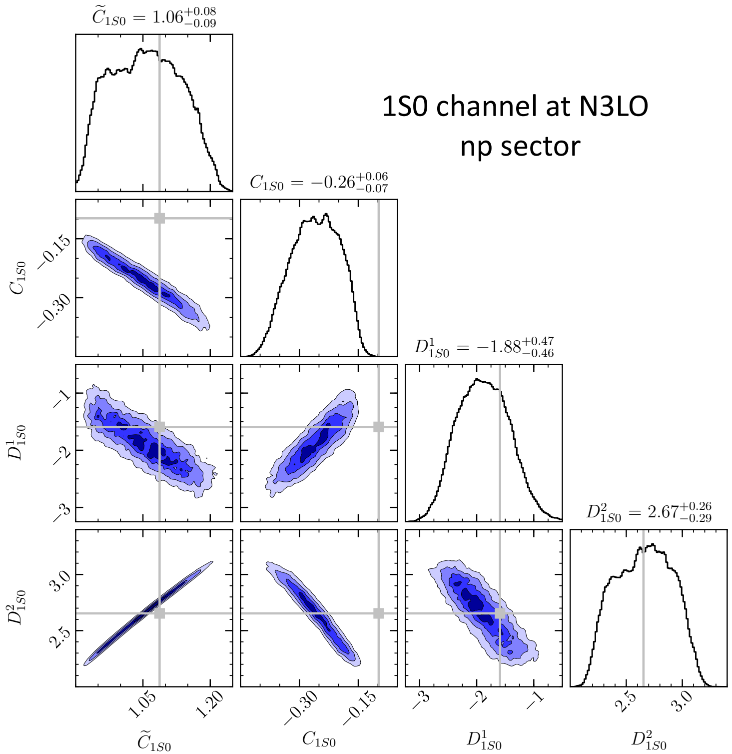
- Projected posteriors give:
  - info on correlations
  - "best" parameter values
  - indication of normality
- Uncorrelated and Gaussian
- Also: Comparison to EKM's values from optimization (need not match!)



$$\delta^{\mathrm{res}}_{\mathrm{p.w.}} = \delta^{\mathrm{pred}}_{\mathrm{p.w.}} - \delta^{\mathrm{PWA93}}_{\mathrm{p.w.}}$$

Projected posteriors as a diagnostic tool

$$\mathrm{pr}(\vec{a}|\boldsymbol{y}_{\mathrm{exp}}, \Sigma_{\mathrm{exp}}) \propto \exp\left[-\frac{1}{2}\sum_{i=1}^{N_d}\frac{r_i^2}{\sigma_i^2}\right] \times e^{-(\vec{a}_k)^2/2\bar{a}^2}$$

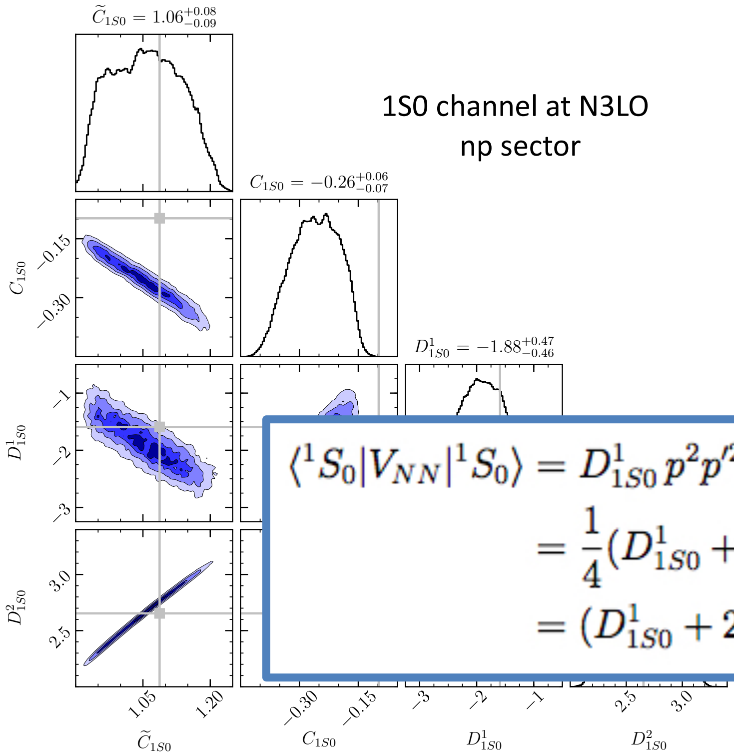


- Irregular structure of posterior
- Nothing *necessarily* wrong, but a clue
- Here: a physics issue actually explains the distorted structure
- Parameter redundancy at N3LO, an operator can be eliminated
- True in 3S1-3D1 channel as well

Wesolowski et al., J. Phys. G (2019)

Projected posteriors as a diagnostic tool

$$\mathrm{pr}(\vec{a}|\boldsymbol{y}_{\mathrm{exp}}, \Sigma_{\mathrm{exp}}) \propto \exp\left[-\frac{1}{2}\sum_{i=1}^{N_d}\frac{r_i^2}{\sigma_i^2}\right] \times e^{-(\vec{a}_k)^2/2\bar{a}^2}$$



1S0 channel at N3LO
np sector

- Irregular structure of posterior
- Nothing *necessarily* wrong, but a clue

- Here: a physics issue actually explains the distorted structure

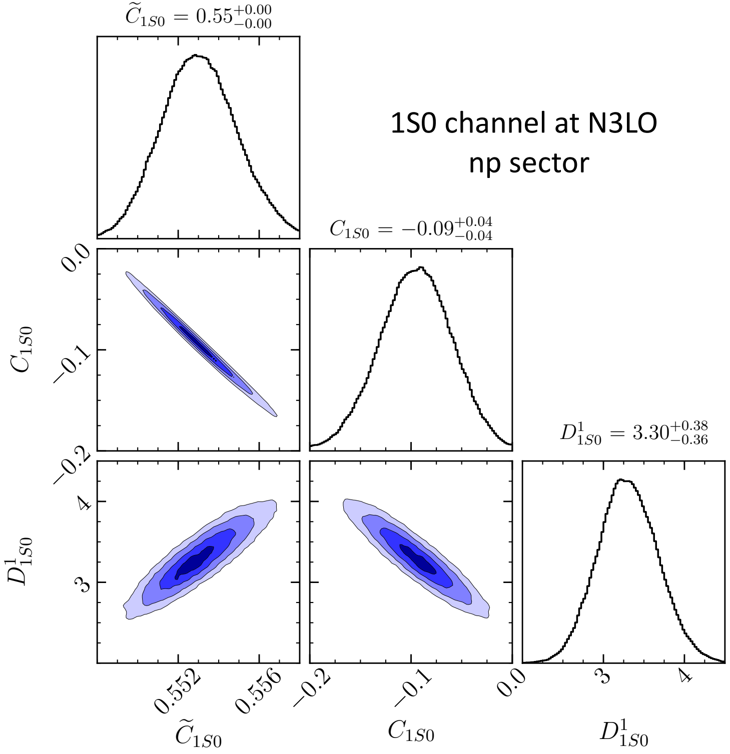
- Parameter redundancy at N3LO, an

$$\begin{aligned}\langle {}^1S_0|V_{NN}|{}^1S_0\rangle &= D^1_{1S0}\,p^2p'^2 + D^2_{1S0}\,(p^4+p'^4)\\ &= \frac{1}{4}(D^1_{1S0}+2D^2_{1S0})(p^2+p'^2)^2 - \frac{1}{4}(D^1_{1S0}-2D^2_{1S0})(p^2-p'^2)^2\\ &= (D^1_{1S0}+2D^2_{1S0})\,p^2p'^2 + D^2_{1S0}(p^2-p'^2)^2\,,\end{aligned}$$

Wesolowski et al., J. Phys. G (2019)

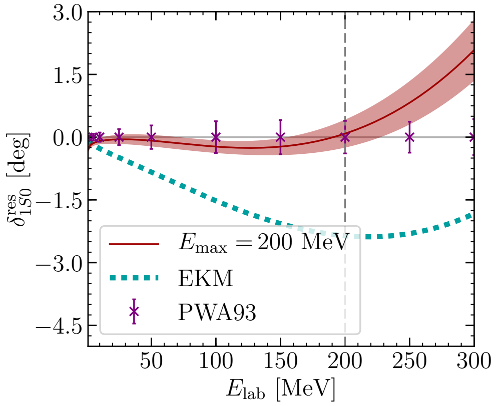
Projected posteriors as a diagnostic tool

$$\mathrm{pr}(\vec{a}|\boldsymbol{y}_{\mathrm{exp}}, \Sigma_{\mathrm{exp}}) \propto \exp\left[-\frac{1}{2}\sum_{i=1}^{N_d}\frac{r_i^2}{\sigma_i^2}\right] \times e^{-(\vec{a}_k)^2/2\bar{a}^2}$$



1S0 channel at N3LO
np sector

- Set D^2_{1S0} to 0 and use only 3 parameters
- Description of data still good (and potential softer)
- Reinert et al., EPJA (2018) improved potential



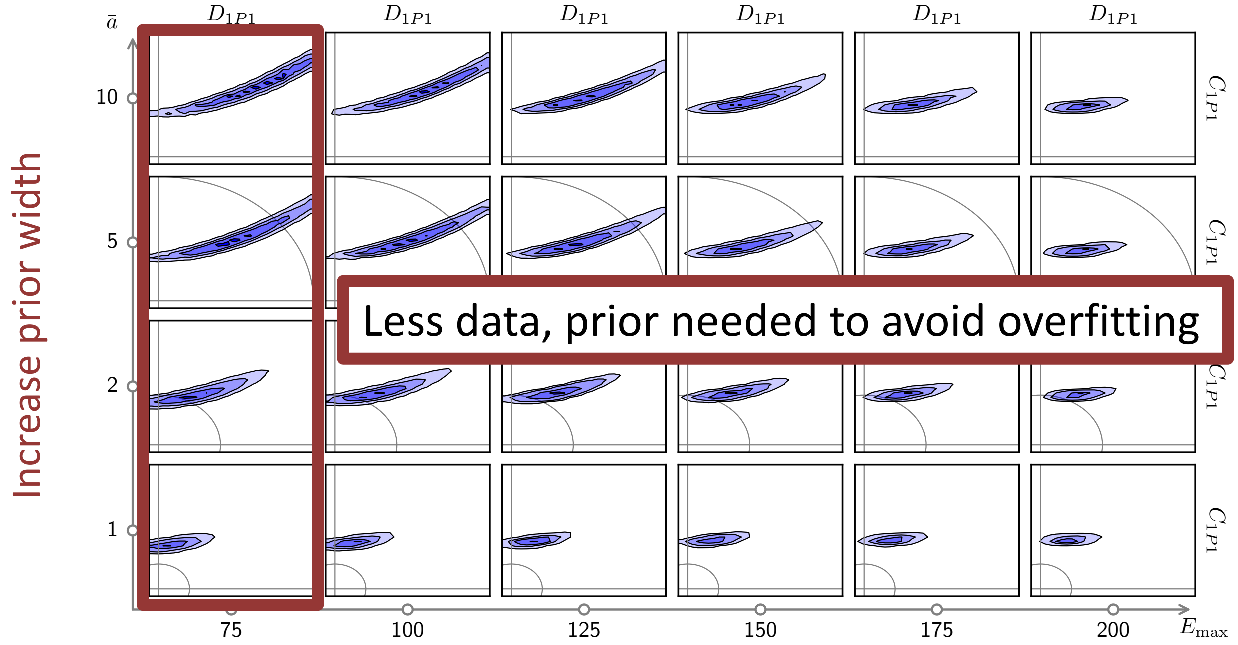
Wesolowski et al., J. Phys. G (2019)

Effect of the prior with less data

$$\mathrm{pr}(\vec{a}|\boldsymbol{y}_{\mathrm{exp}}, \Sigma_{\mathrm{exp}}) \propto \exp\left[-\frac{1}{2}\sum_{i=1}^{N_d}\frac{r_i^2}{\sigma_i^2}\right] \times e^{-(\vec{a}_k)^2/2\bar{a}^2}$$

Wesolowski et al., J. Phys. G (2019)

Repeat same problem, vary givens



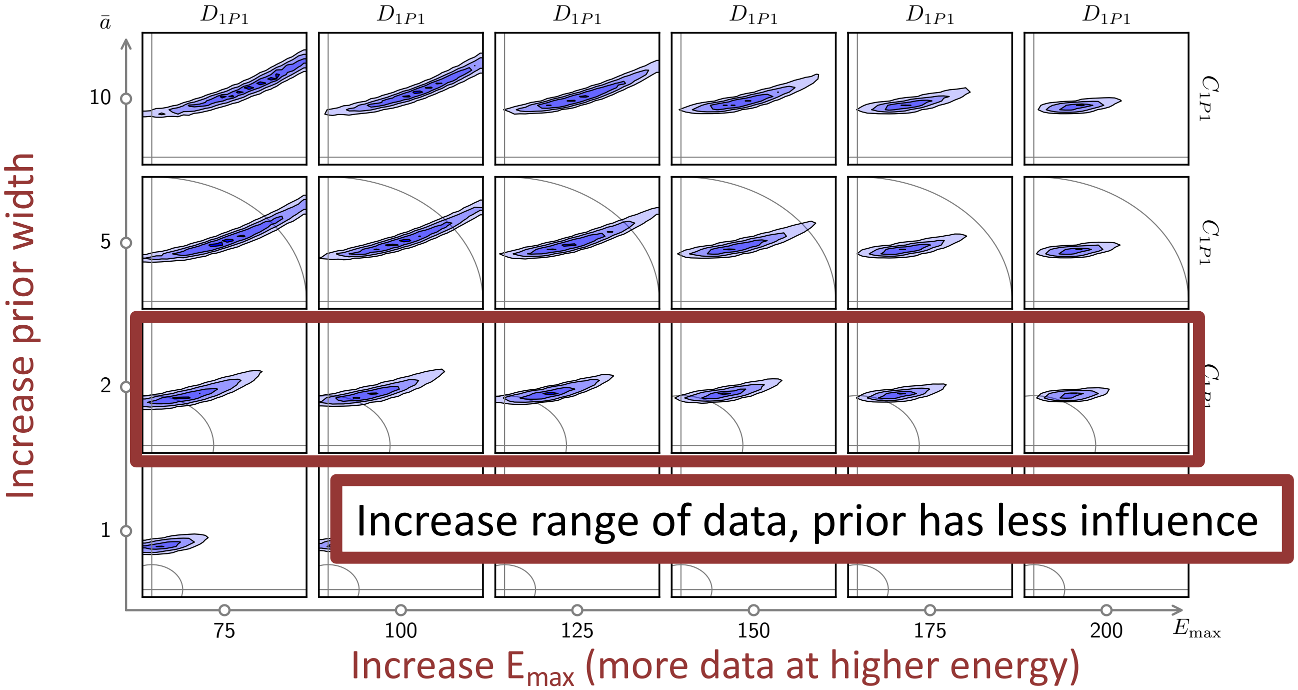
Increase E_{max} (more data at higher energy)

Effect of the prior with less data

$$\mathrm{pr}(\vec{a}|\boldsymbol{y}_{\mathrm{exp}}, \Sigma_{\mathrm{exp}}) \propto \exp\left[-\frac{1}{2}\sum_{i=1}^{N_d}\frac{r_i^2}{\sigma_i^2}\right] \times e^{-(\vec{a}_k)^2/2\bar{a}^2}$$

Wesolowski et al., J. Phys. G (2019)

Repeat same problem, vary givens



Parameter estimation including truncation errors

$$\boldsymbol{y}_{\text{exp}} = \boldsymbol{y}_{\text{th}} + \boxed{\delta\boldsymbol{y}_{\text{th}}} + \delta\boldsymbol{y}_{\text{exp}} \qquad \delta\boldsymbol{y}_{\text{th}} = \boldsymbol{y}_{\text{ref}} \sum_{n=k+1}^{\infty} c_n Q^n$$

Full posterior pdf including theory error:

$$\text{pr}(\vec{a}_k | \boldsymbol{y}_{\text{exp}}, \Sigma_{\text{exp}}, \Sigma_{\text{th}}) \propto e^{-\frac{1}{2}\mathbf{r}^T(\Sigma_{\text{exp}}+\Sigma_{\text{th}})^{-1}\mathbf{r}} \times e^{-(\vec{a}_k)^2/2\bar{a}^2}$$

Covariance matrix for assuming uncorrelated *Q* points:

$$(\Sigma_{\text{th,uncorr.}})_{ij} = (\boldsymbol{y}_{\text{ref}})_i^2\, \bar{c}^2 \sum_{n=k+1}^{k_{\text{max}}} Q_i^{2n}\delta_{ij} \xrightarrow[k_{\text{max}}\to\infty]{} \frac{(\boldsymbol{y}_{\text{ref}})_i^2\, \bar{c}^2\, Q_i^{2k+2}}{1-Q_i^2}\delta_{ij}$$

Covariance matrix assuming fully correlated *Q* points:

$$(\Sigma_{\text{th,corr.}})_{ij} = (\boldsymbol{y}_{\text{ref}})_i(\boldsymbol{y}_{\text{ref}})_j\, \bar{c}^2 \sum_{n=k+1}^{k_{\text{max}}} Q_i^n Q_j^n \xrightarrow[k_{\text{max}}\to\infty]{} \frac{(\boldsymbol{y}_{\text{ref}})_i(\boldsymbol{y}_{\text{ref}})_j\, \bar{c}^2\, Q_i^{k+1}Q_j^{k+1}}{1-Q_iQ_j}$$

Wesolowski et al., J. Phys. G (2019)

Bayes is great, but won't the sampling be too expensive?

- **Dilemma:** Bayesian sampling of posteriors is desirable, but modern nuclear theory calculations typically require supercomputer resources.
- **Solution:** Emulators! Train a computer model of the calculation using a representative set of parameters and then sample from the model instead.

- **Gaussian process emulators:** think about efficient interpolations formulas, but with error estimates.
  Eigenvector continuation: use a *really* effective variational basis (introduced for physics applications by Frame et al., PRL **121** (2018)).

Calculate $i = 1$ to N ground-state wfs: $\widehat{H}(\boldsymbol{\theta}_i)|\psi_0(\boldsymbol{\theta}_i)\rangle = E_0(\boldsymbol{\theta}_i)|\psi_0(\boldsymbol{\theta}_i)\rangle$,

Use trial wavefunction from linear combination of $|\psi_0(\boldsymbol{\theta}_i)\rangle$s:

$$|\psi_{\text{trial}}\rangle = c_1|\psi_0(\boldsymbol{\theta}_1)\rangle + c_2|\psi_0(\boldsymbol{\theta}_2)\rangle + \cdots + c_N|\psi_0(\boldsymbol{\theta}_N)\rangle$$

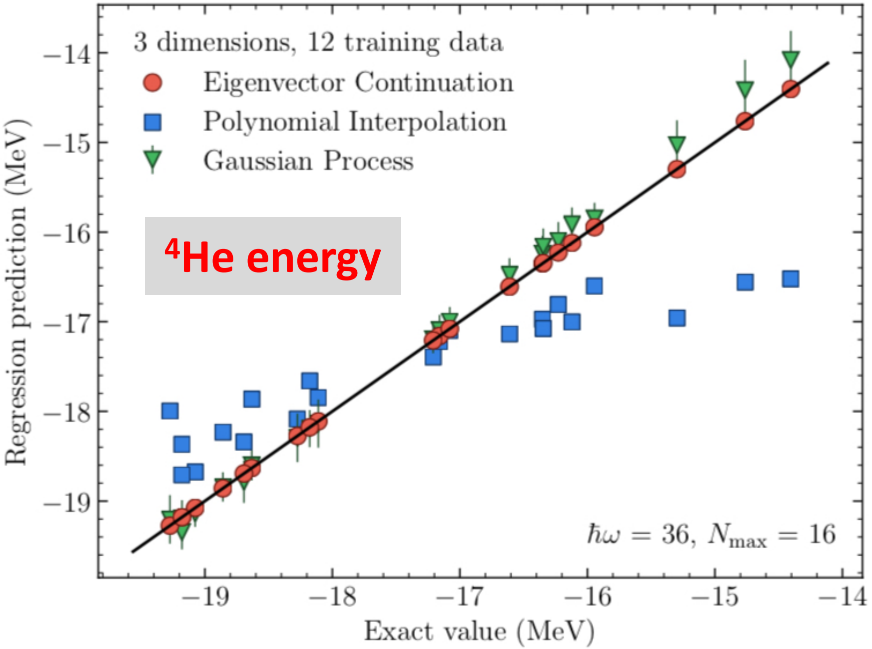
Minimize $\langle\psi_{\text{trial}}|\widehat{H}(\boldsymbol{\theta})|\psi_{\text{trial}}\rangle$ with $\langle\psi_{\text{trial}}|\psi_{\text{trial}}\rangle = 1$ to find $E_0(\boldsymbol{\theta})$ and c_is

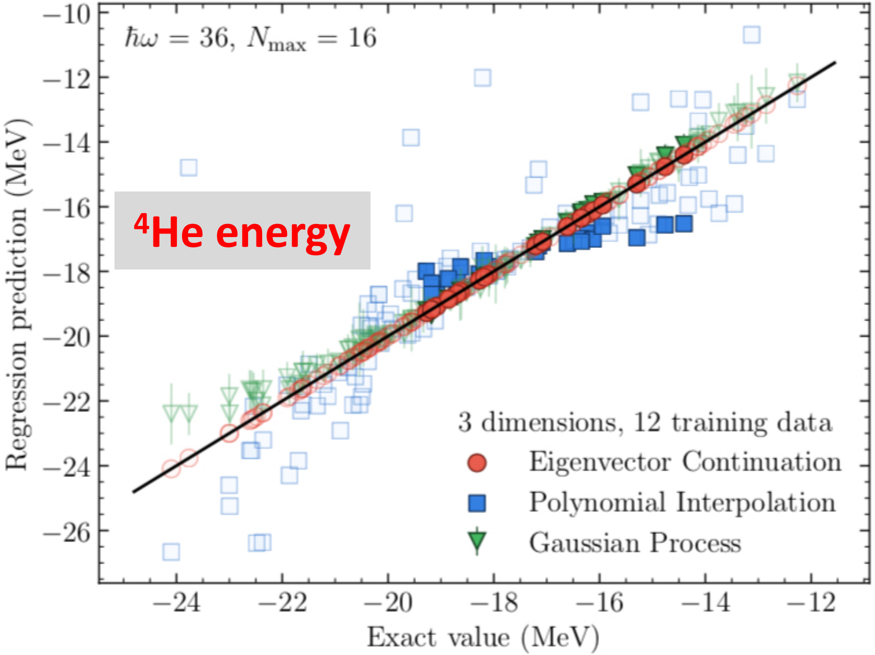
Bayes is great, but won't the sampling be too expensive?

Eigenvector Continuation as an Efficient and Accurate Emulator for Uncertainty Quantification

S. König, A. Ekström, K. Hebeler, D. Lee, A. Schwenk

arXiv:1909.08446
based on Frame et al., PRL **121** (2018))



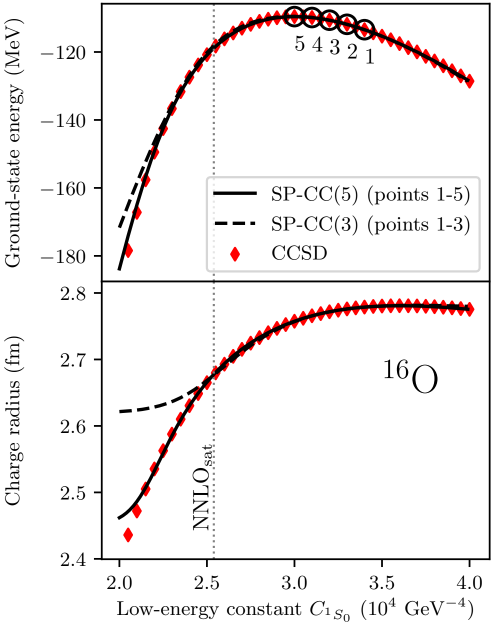


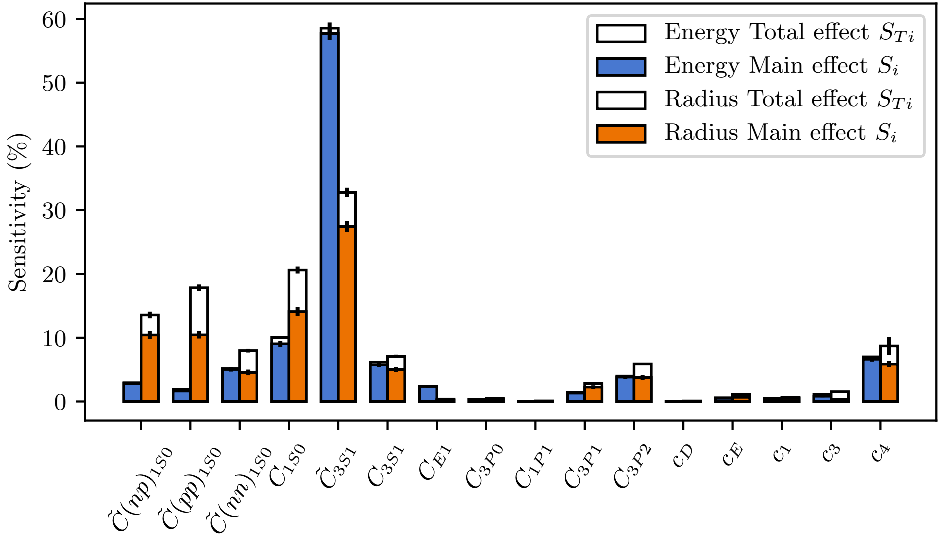
Bayes is great, but won't the sampling be too expensive?

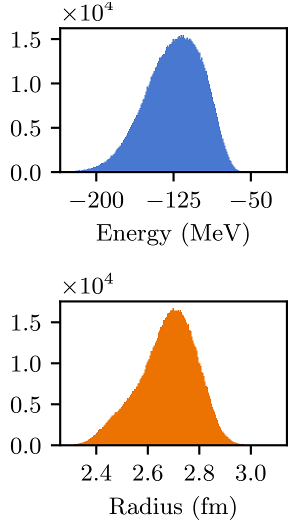
Global sensitivity analysis of bulk properties of an atomic nucleus

A. Ekström and G. Hagen

arXiv: 1910.02922







"We have to use (16 + 1) · 216 = 1,114,112 quasi MC samples to extract statistically significant main and total effects of the energy and radius for all LECs. With SP-CC(64) this took about 1 hour on a standard laptop, while an equivalent set of exact CCSD computations would require 20 years."

Gaussian process model of the coefficients

$$\mathrm{pr}(c_n \mid \boldsymbol{\theta}) \overset{\text{iid}}{\sim} \mathcal{GP}(\mu, \sigma^2 R_\ell)$$

Conjugate priors:

$$\mu \mid \sigma^2 \sim \mathcal{N}(\eta, \sigma^2 V)$$
$$\sigma^2 \sim \chi^{-2}(\nu, \tau^2)$$