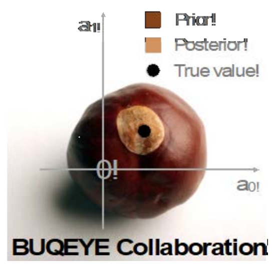
Bayesian parameter estimation for chiral interactions: NN sector

Sarah Wesolowski

New Ideas in Constraining Nuclear Forces, ECT* 2018



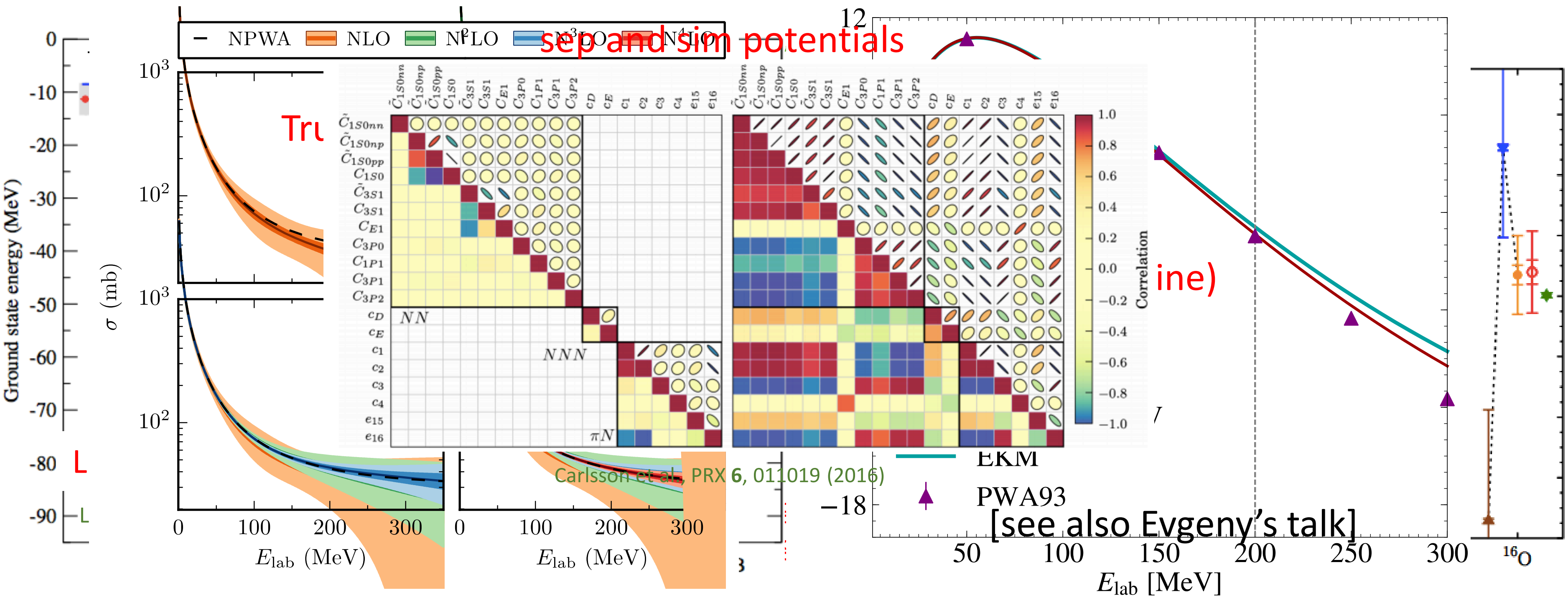
Dick Furnstahl (OSU)
Jordan Melendez (OSU)
Matt Pratola (OSU statistics)
Daniel Phillips (OU)

Supported in part by NSF, DOE, and the SciDAC NUCLEI project

Motivation

Current landscape of uncertainty quantification for interactions

- **Covariance methods**: uncertainty from fitting to data
- **Bootstrapping**: uncertainty from fitting or numerics
- **Truncation error estimates for EFTs**: codifying standard EFT protocols
- **Bayesian**: can fit the above methods into this framework



Outline

GOAL: estimate the low-energy constants (LECs) of nuclear interactions, include all information consistently, and provide statistically meaningful uncertainty estimates

E.g. overestimating error bands:
"We're going to win so much, you're going to be so sick and tired of winning."

- Case study 1: usefulness of projected posterior plots
- Case study 2: LEC stability with maximum energy for fit

- **Part 2:** modeling truncation errors using Gaussian processes
  - Brief motivation of the GP model
  - Point-wise versus curve-wise implementation
  - Posterior for the EFT breakdown scale

Recent explosion of chiral EFT interactions

Interactions from 2003-4: EM(500, 600 MeV) or family of EGM potentials; used similar non-local regulators and similar fits to NN, 3N

New generation NN+3N interactions **[references at end]**

- **Non-local:** updated EMN and new soft; Ekstrom et al. sim, sat, and with Δs
- **Local** (for QMC): Gezerlis et al. nucleons only; Piarulli et al. with Δs
- **Semi-local:** Bochum-Julich group, SCS (x-space) and SMS (p-space)

Issues: power counting, regulator artifacts, EFT convergence, fitting protocols, fine-tuning, over/under-fitting, parameter redundancies, how to do UQ?

Parameter estimation issues repeat with LECs for currents and for other EFTs

Propaganda: Follow the Bayes Way

$$\mathrm{pr}(A|B,I) = \frac{\mathrm{pr}(B|A,I)\mathrm{pr}(A|I)}{\mathrm{pr}(B|I)} \Longrightarrow \underbrace{\mathrm{pr}(x|\mathrm{data},I)}_{\text{posterior}} \propto \underbrace{\mathrm{pr}(\mathrm{data}|x,I)}_{\text{likelihood}} \times \underbrace{\mathrm{pr}(x|I)}_{\text{prior}}$$

Why use Bayesian statistics?

- Parameter estimation: conventional optimization recovered as special case
- Update expectations using Bayes' theorem when have more information
- *Assumptions are made explicit* (e.g. naturalness of LECs)
- Clear prescriptions for combining errors
- Statistics as diagnostics for *physics*
- Model checking: we can *test* if our UQ model works and study sensitivities
- Model selection: Is the Δ needed? Pionless vs. pionful formulations, …
- Particularly well suited for (any) EFT, but generally suited for theory errors

Bayesian framework for parameter estimation in EFTs

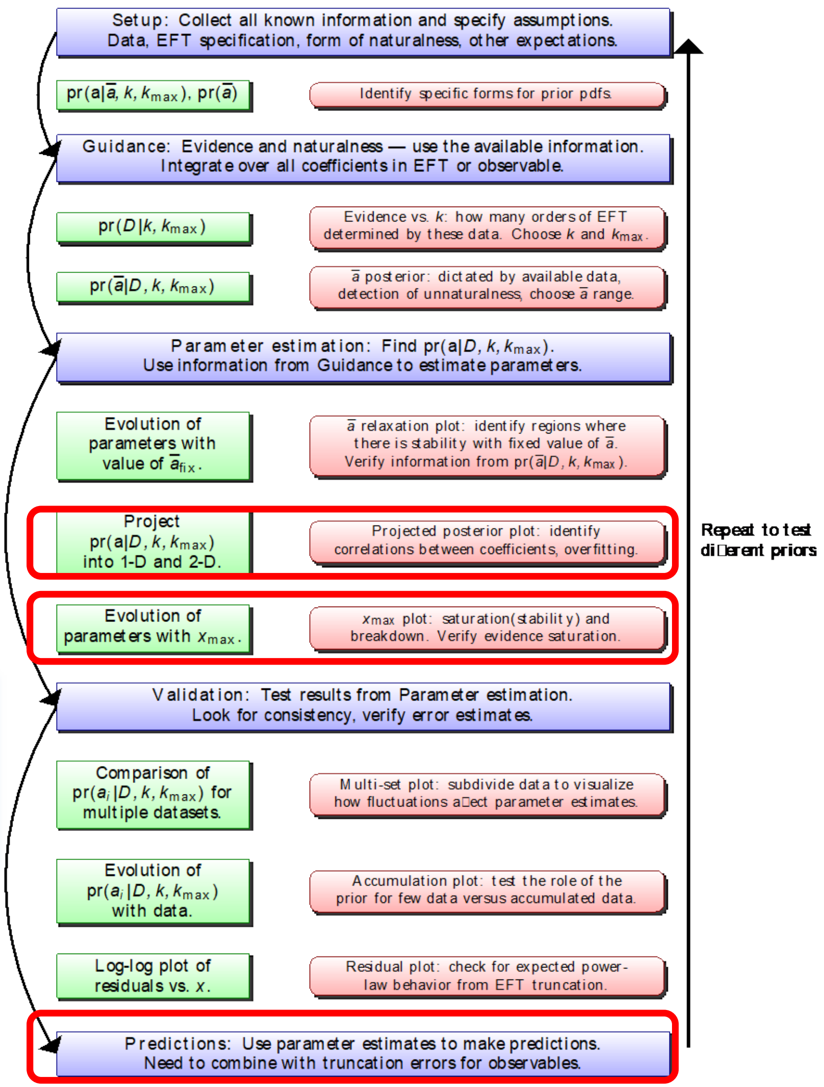
[See sw et al., J. Phys. G 43, 074001 (2016)]

Here: Use parameter estimation for NN phase shifts as test case (SCS potential of EKM)

- Good news: well studied, clear example for comparison, fairly easy computation
- Bad news: numerous and precise data ➔ differences from conventional approach are subtle. Plan: focus on illuminating cases.

1. The usefulness of projected posterior plots
2. LEC stability with maximum energy for fit
3. Combining theory uncertainties (in progress)

Based on: *Exploring Bayesian parameter estimation for chiral effective field theory using nucleon-nucleon phase shifts,* sw, R.J. Furnstahl, D. Phillips (posted soon)



Case study 1: The usefulness of projected posteriors

$$\underbrace{\mathrm{pr}(\mathbf{a}|D,I)}_{\text{posterior}} \propto \underbrace{\mathrm{pr}(D|\mathbf{a},I)}_{\text{likelihood}} \times \underbrace{\mathrm{pr}(\mathbf{a}|I)}_{\text{prior}}$$

D = data, I = background info.

Posterior pdf for vector of k^{th} order LECs $\mathbf{a}^{(k)}$ with naturalness assumption

$$\mathrm{pr}(\mathbf{a}^{(k)}|k, k_{\max}, D(E_{\max})) \propto e^{-\chi^2_{\mathrm{aug}}/2}\, e^{\mathbf{B}^T\mathbf{A}^{-1}\mathbf{B}/2}$$

account for omitted terms up to k_{max}

use data up to energy E_{max}

Case study 1: The usefulness of projected posteriors

$$\underbrace{\mathrm{pr}(\mathbf{a}|D,I)}_{\text{posterior}} \propto \underbrace{\mathrm{pr}(D|\mathbf{a},I)}_{\text{likelihood}} \times \underbrace{\mathrm{pr}(\mathbf{a}|I)}_{\text{prior}} \qquad D = \text{data}, \quad I = \text{background info.}$$

Posterior pdf for vector of k^{th} order LECs $\mathbf{a}^{(k)}$ with naturalness assumption

$$\mathrm{pr}(\mathbf{a}^{(k)}|k, k_{\max}, D(E_{\max})) \propto \boxed{e^{-\chi^2_{\text{aug}}/2}} e^{\mathbf{B}^T\mathbf{A}^{-1}\mathbf{B}/2}$$

data EFT prediction at k^{th} order

$$\boxed{\chi^2_{\text{aug}} = \sum_{i=1}^{N} \left(\frac{d_i - t^{(k)}(p_i; \mathbf{a}^{(k)})}{\sigma_i} \right)^2 + \frac{\left(\mathbf{a}^{(k)}\right)^2}{2\bar{a}^2_{\text{fix}}}}$$

Augmented χ^2 accounts for data errors σ_i *and* bounded (natural) LECs, here with simple δ-function prior at $\bar{a}_{\text{fix}}$

Case study 1: The usefulness of projected posteriors

$$\underbrace{\mathrm{pr}(\mathbf{a}|D,I)}_{\text{posterior}} \propto \underbrace{\mathrm{pr}(D|\mathbf{a},I)}_{\text{likelihood}} \times \underbrace{\mathrm{pr}(\mathbf{a}|I)}_{\text{prior}} \qquad D = \text{data}, \quad I = \text{background info.}$$

Posterior pdf for vector of k^{th} order LECs $\mathbf{a}^{(k)}$ with naturalness assumption

$$\mathrm{pr}(\mathbf{a}^{(k)}|k, k_{\max}, D(E_{\max})) \propto e^{-\chi^2_{\text{aug}}/2} \boxed{e^{\mathbf{B}^T\mathbf{A}^{-1}\mathbf{B}/2}}$$

$$\chi^2_{\text{aug}} = \sum_{i=1}^{N}\left(\frac{d_i - t^{(k)}(p_i;\mathbf{a}^{(k)})}{\sigma_i}\right)^2 + \frac{\left(\mathbf{a}^{(k)}\right)^2}{2\bar{a}^2_{\text{fix}}}$$

$$A_{jj'} = \frac{\delta_{j,j'}}{\bar{c}_{\text{fix}}{}^2} + \sum_{i=1}^{N}\frac{X_0(p_i)^2\; Q_i^{j+j'}}{\sigma_i^2}$$

$$B_j = \sum_{i=1}^{N} X_0(p_i)\, Q_i^j\; \frac{d_i - t^{(k)}(p_i;\mathbf{a}^{(k)})}{\sigma_i^2}$$

Accounts for *truncation error* from omitted higher-order terms

Ref: Stump et al., PRD 65 **014012 (appendix B)**

- Consistently include higher-order correlated errors
- Assumed to be same set of coefficients for each datum
- Minimally informative assumptions
- Easy to implement! (more consistent alternatives in progress)

Case study 1: The usefulness of projected posteriors

$$\underbrace{\mathrm{pr}(\mathbf{a}|D,I)}_{\text{posterior}} \propto \underbrace{\mathrm{pr}(D|\mathbf{a},I)}_{\text{likelihood}} \times \underbrace{\mathrm{pr}(\mathbf{a}|I)}_{\text{prior}} \qquad D = \text{data}, \quad I = \text{background info.}$$

Posterior pdf for vector of k^{th} order LECs $\mathbf{a}^{(k)}$ with naturalness assumption

$$\mathrm{pr}(\mathbf{a}^{(k)}|k, k_{\max}, D(E_{\max})) \propto \; e^{-\chi^2_{\text{aug}}/2} \, e^{\mathbf{B}^T\mathbf{A}^{-1}\mathbf{B}/2}$$

$$\chi^2_{\text{aug}} = \sum_{i=1}^{N}\left(\frac{d_i - t^{(k)}(p_i;\mathbf{a}^{(k)})}{\sigma_i}\right)^2 + \frac{\left(\mathbf{a}^{(k)}\right)^2}{2\bar{a}^2_{\text{fix}}}$$

$$A_{jj'} = \frac{\delta_{j,j'}}{\bar{c}_{\text{fix}}{}^2} + \sum_{i=1}^{N}\frac{X_0(p_i)^2\, Q_i^{j+j'}}{\sigma_i^2}$$

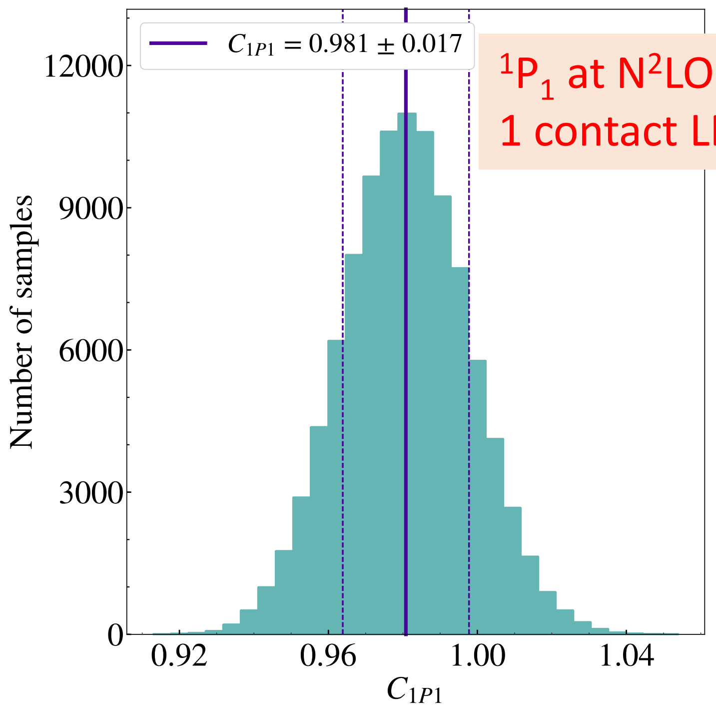
$$B_j = \sum_{i=1}^{N} X_0(p_i)\, Q_i^j\, \frac{d_i - t^{(k)}(p_i;\mathbf{a}^{(k)})}{\sigma_i^2}$$

For case study 1, we won't include truncation errors, so $k = k_{\max}$

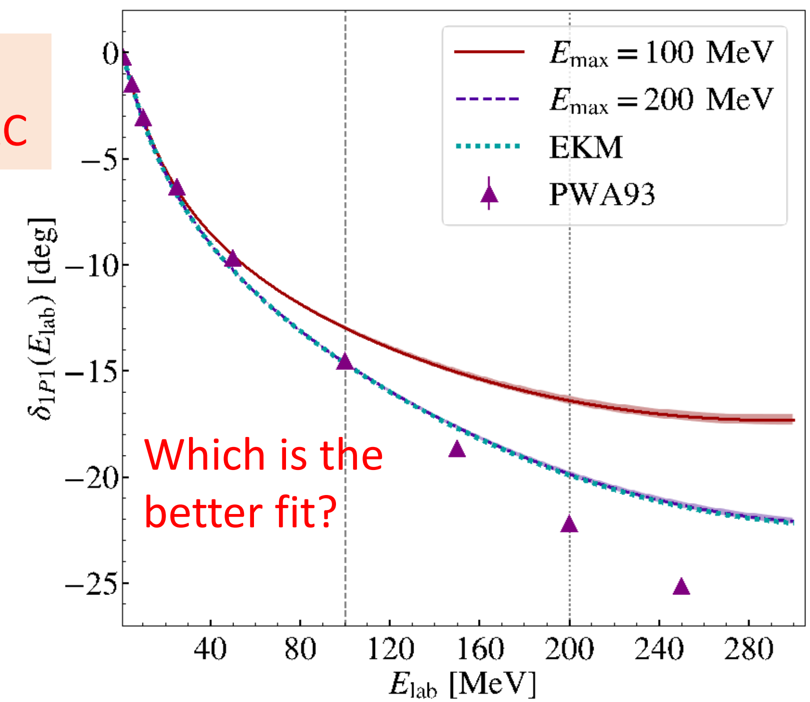
Case study 1: The usefulness of projected posteriors

$$\mathrm{pr}(\mathbf{a}^{(k)}|k, k_{\max} = k, D(E_{\max})) \propto e^{-\chi^2_{\mathrm{aug}}/2}$$

$$\chi^2_{\mathrm{aug}} = \sum_{i=1}^{N} \left(\frac{d_i - t^{(k)}(p_i; \mathbf{a}^{(k)})}{\sigma_i} \right)^2 + \frac{\left(\mathbf{a}^{(k)}\right)^2}{2\bar{a}^2_{\mathrm{fix}}}$$



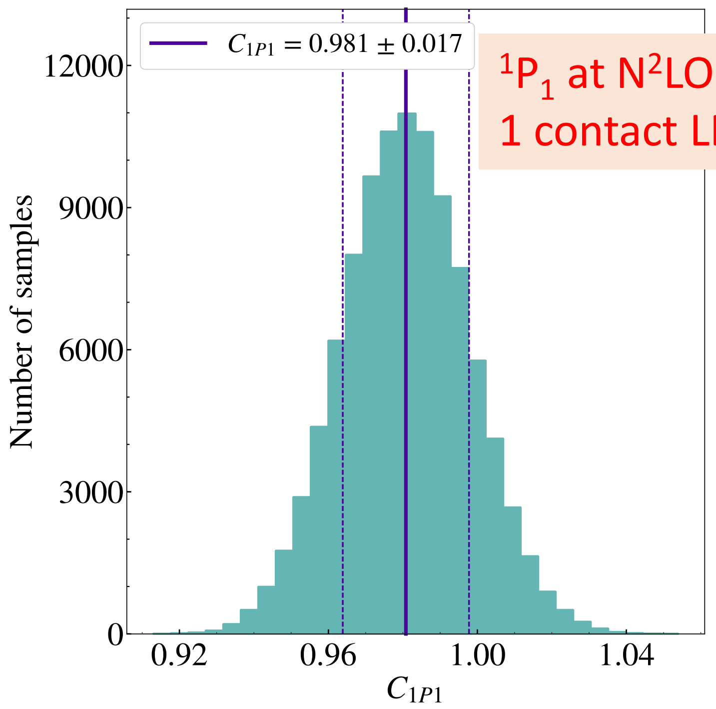
$^1P_1$ at N^2LO
1 contact LEC



Which is the better fit?

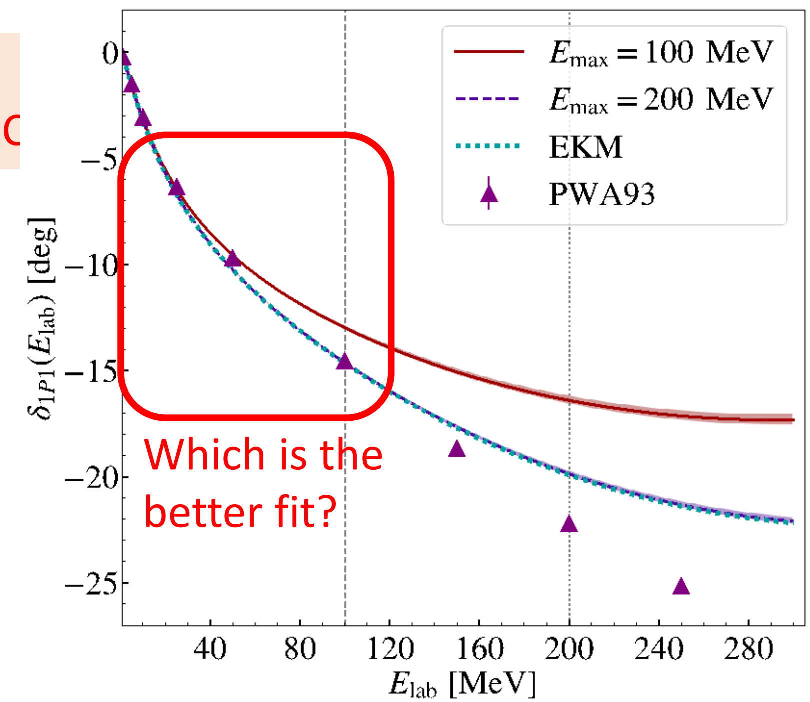
Case study 1: The usefulness of projected posteriors

$$\mathrm{pr}(\mathbf{a}^{(k)}|k, k_{\max}=k, D(E_{\max})) \propto e^{-\chi^2_{\mathrm{aug}}/2} \qquad \chi^2_{\mathrm{aug}} = \sum_{i=1}^{N}\left(\frac{d_i - t^{(k)}(p_i;\mathbf{a}^{(k)})}{\sigma_i}\right)^2 + \frac{\left(\mathbf{a}^{(k)}\right)^2}{2\bar{a}^2_{\mathrm{fix}}}$$



$^1P_1$ at N^2LO

1 contact LEC

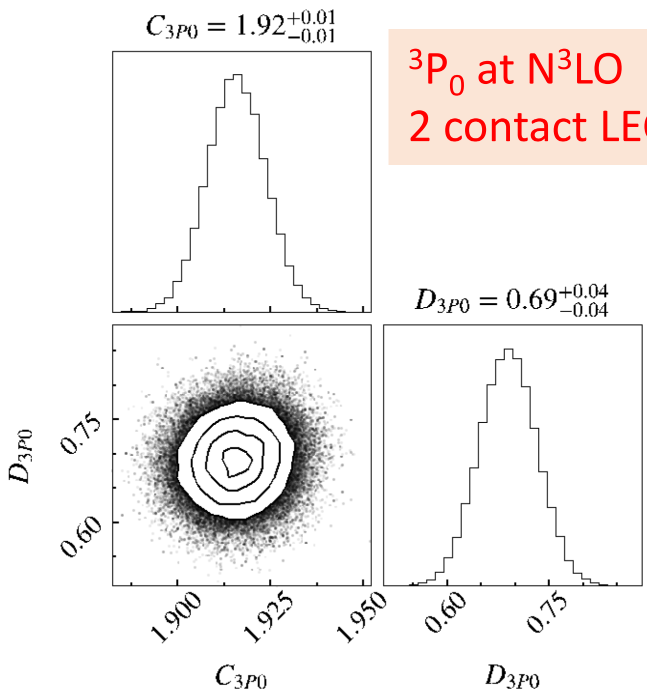


Which is the better fit?

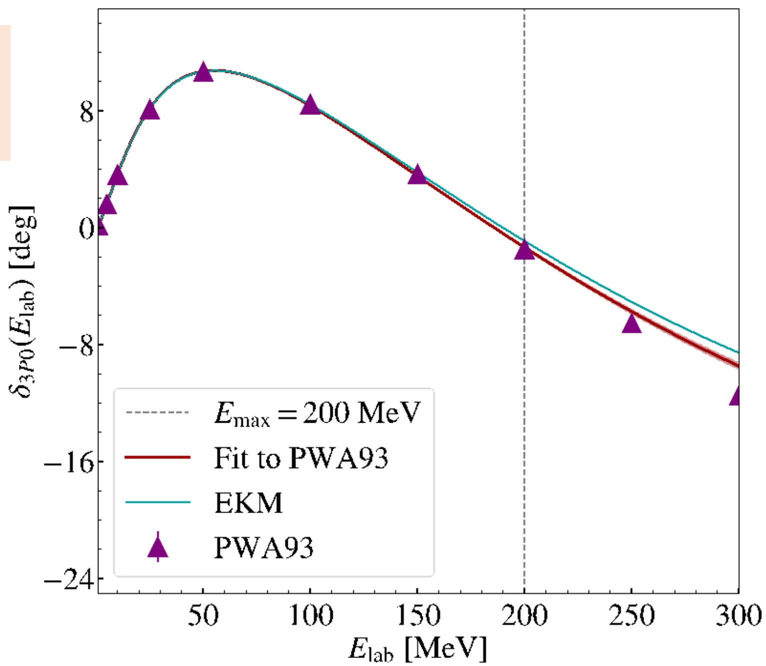
Case study 1: The usefulness of projected posteriors

$$\mathrm{pr}(\mathbf{a}^{(k)}|k, k_{\max}=k, D(E_{\max})) \propto e^{-\chi^2_{\mathrm{aug}}/2}$$

$$\chi^2_{\mathrm{aug}} = \sum_{i=1}^{N}\left(\frac{d_i - t^{(k)}(p_i;\mathbf{a}^{(k)})}{\sigma_i}\right)^2 + \frac{\left(\mathbf{a}^{(k)}\right)^2}{2\bar{a}^2_{\mathrm{fix}}}$$

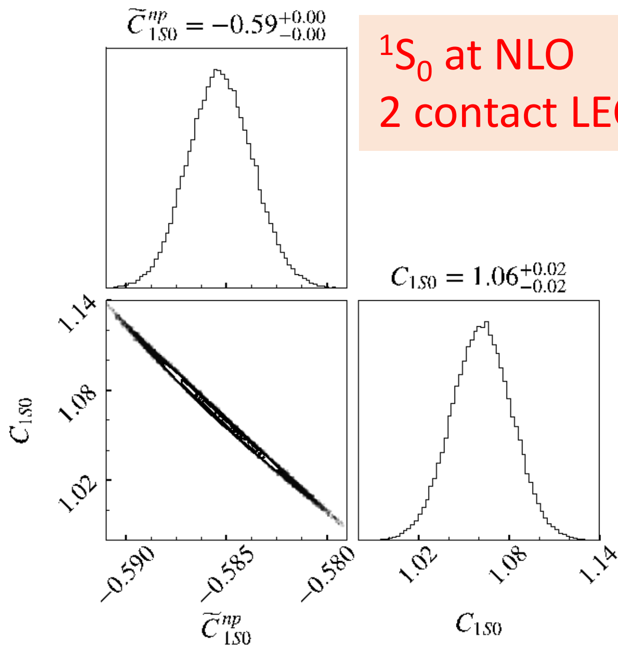


$^3P_0$ at N^3LO
2 contact LECs

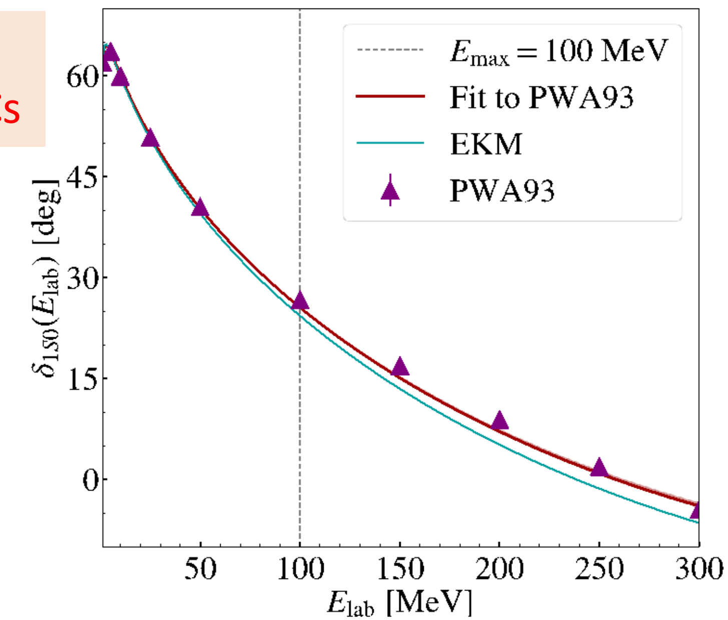


Case study 1: The usefulness of projected posteriors

$$\mathrm{pr}(\mathbf{a}^{(k)}|k, k_{\max} = k, D(E_{\max})) \propto e^{-\chi^2_{\mathrm{aug}}/2} \qquad \chi^2_{\mathrm{aug}} = \sum_{i=1}^{N} \left(\frac{d_i - t^{(k)}(p_i; \mathbf{a}^{(k)})}{\sigma_i} \right)^2 + \frac{\left(\mathbf{a}^{(k)}\right)^2}{2\bar{a}^2_{\mathrm{fix}}}$$

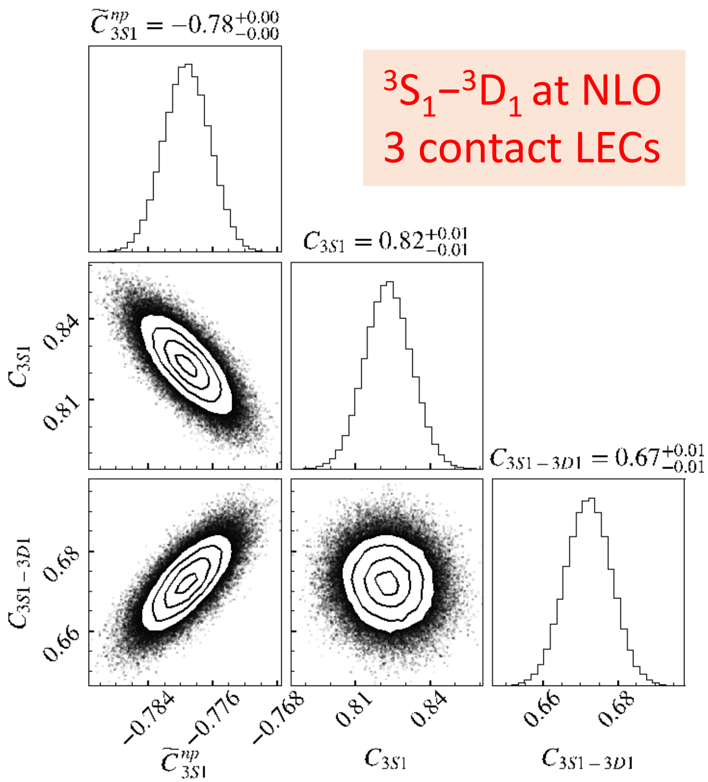


$^1S_0$ at NLO
2 contact LECs

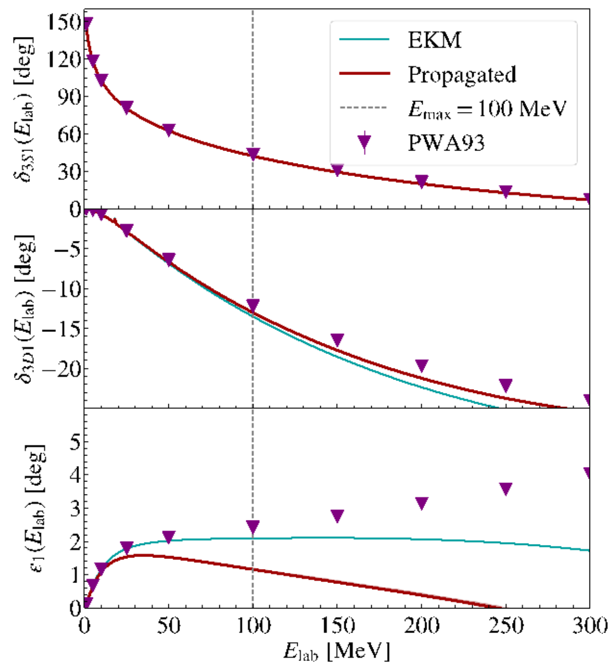


Case study 1: The usefulness of projected posteriors

$$\mathrm{pr}(\mathbf{a}^{(k)}|k, k_{\max} = k, D(E_{\max})) \propto e^{-\chi^2_{\mathrm{aug}}/2} \qquad \chi^2_{\mathrm{aug}} = \sum_{i=1}^{N}\left(\frac{d_i - t^{(k)}(p_i;\mathbf{a}^{(k)})}{\sigma_i}\right)^2 + \frac{\left(\mathbf{a}^{(k)}\right)^2}{2\bar{a}^2_{\mathrm{fix}}}$$



$^3S_1$–$^3D_1$ at NLO
3 contact LECs

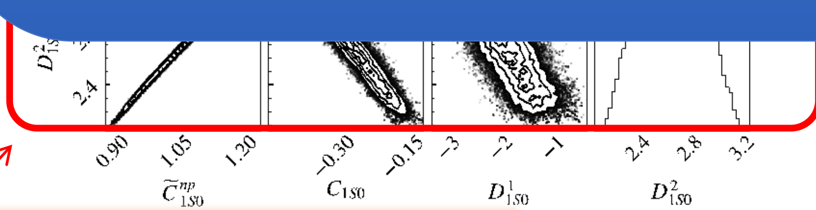


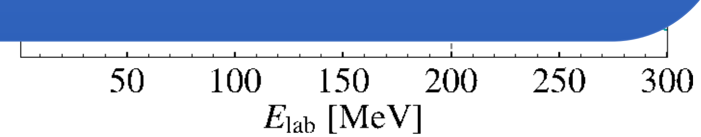
Case study 1: The usefulness of projected posteriors

$\mathrm{pr}(\mathbf{a}^{(k)}| \ldots \qquad \chi^2 = \sum^{N} \Big(d_i - t^{(k)}(p_i;\mathbf{a}^{(k)}) \Big)^2 + (\mathbf{a}^{(k)})^2$

S-wave potentials at N^{3}LO have off-shell pieces → redundancy!

$$
\begin{aligned}
D^1_{(1S0)} p^2 p'^2 + D^2_{(1S0)}(p^4 + p'^4)
&= \frac{1}{4}(D^1_{(1S0)} + 2D^2_{(1S0)})(p^2 + p'^2)^2 \\
&\quad - \frac{1}{4}(D^1_{(1S0)} - 2D^2_{(1S0)})(p^2 - p'^2)^2 \\
&= (D^1_{(1S0)} + 2D^2_{(1S0)}) p^2 p'^2 + D^2_{(1S0)}(p^2 - p'^2)^2
\end{aligned}
$$



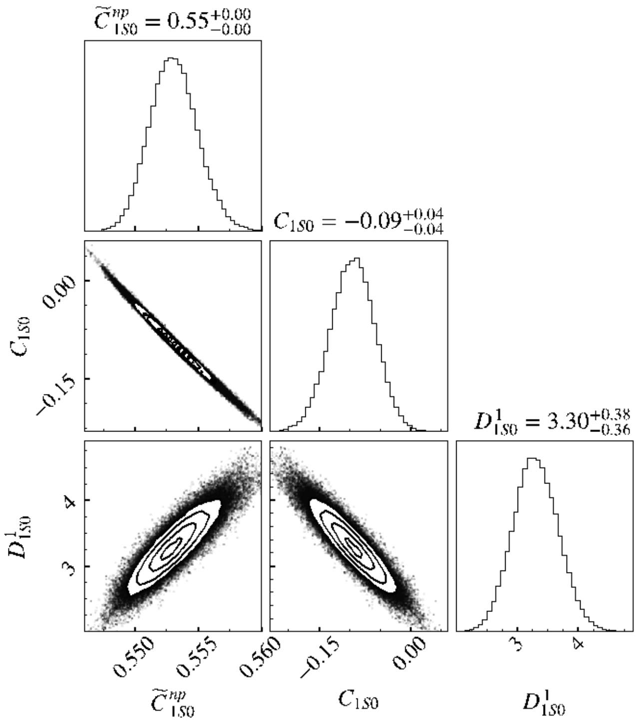


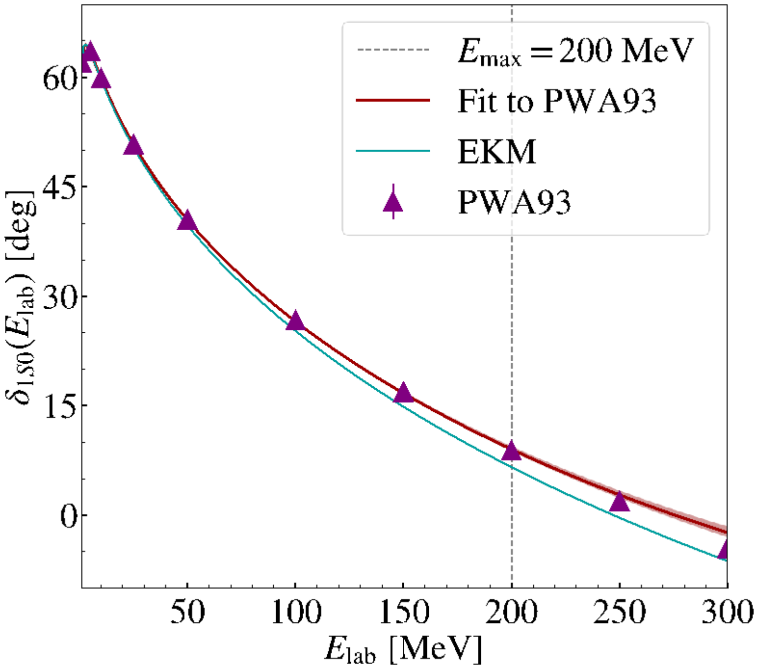
Freeze redundant operator LEC

Case study 1: The usefulness of projected posteriors

$$\mathrm{pr}(\mathbf{a}^{(k)}|k, k_{\max} = k, D(E_{\max})) \propto e^{-\chi^2_{\mathrm{aug}}/2}$$

$$\chi^2_{\mathrm{aug}} = \sum_{i=1}^{N} \left(\frac{d_i - t^{(k)}(p_i; \mathbf{a}^{(k)})}{\sigma_i} \right)^2 + \frac{\left(\mathbf{a}^{(k)}\right)^2}{2\bar{a}^2_{\mathrm{fix}}}$$





Reinert et al. find much better behaved SMS fit with three fewer parameters!

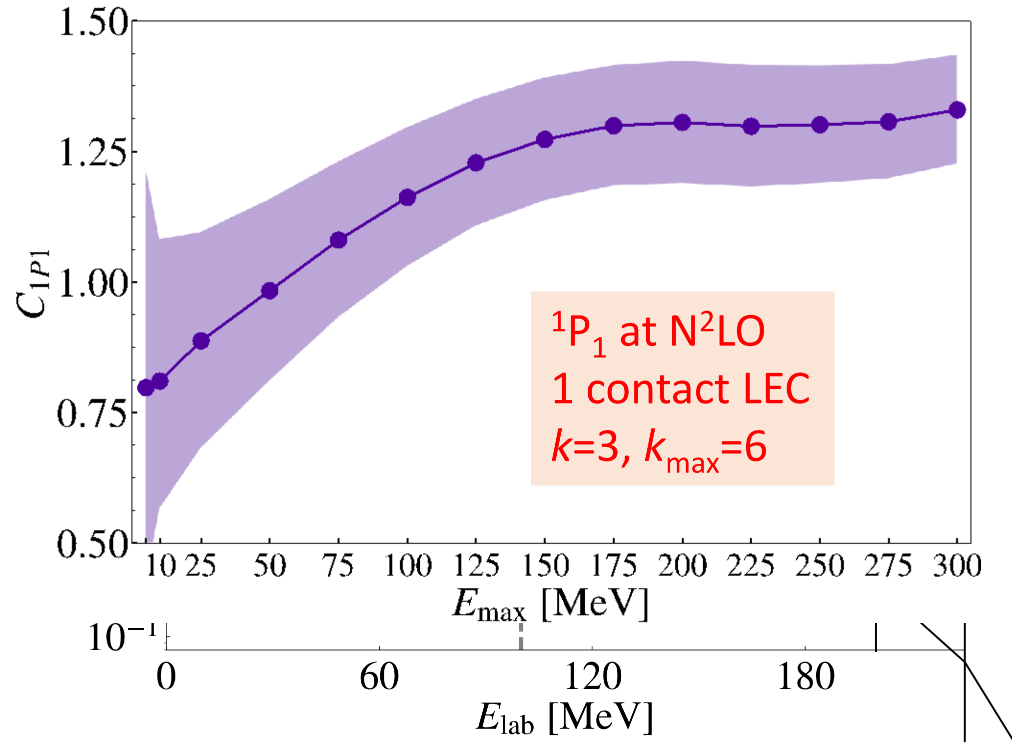
Case study 2: LEC stability with maximum energy for fit

EFT convergence

$$X(p) = X_0 \sum_{n=0}^{k} c_n Q^n, \quad \Delta_k = \sum_{n=k+1}^{k_{\max}} c_n Q^n$$

$$Q = \max(p, m_\pi)/\Lambda_b$$

- How should we choose $E_{\max}$ to fit?
- Operator expansion, so LECs should be independent of data used
- *Could distort observables* (e.g., energies work but not radii)
- Solution: account for EFT truncation
- $E_{\max}$ plots are simple proxy for Bayesian model selection
- Generally relevant for fitting!



Note: fit is to partial-wave cross section with larger uncertainty

Outline

GOAL: estimate the low-energy constants (LECs) of nuclear interactions, include all information consistently, and provide statistically meaningful uncertainty estimates

- **Part I:** parameter estimation for chiral EFT using NN phaseshifts
  - Setup: Bayesian propaganda and framework
  - How to incorporate truncation errors in parameter estimation
  - Case study 1: usefulness of projected posterior plots
  - Case study 2: LEC stability with maximum energy for fit
- **Part 2:** modeling truncation errors using Gaussian processes
  - Brief motivation of the GP model
  - Point-wise versus curve-wise implementation
  - Posterior for the EFT breakdown scale

Statistical model for truncation errors

Previous work: Furnstahl, Phillips, Klco, sw, PRC **92**, 024005 (2015)

- Previous work: used a Bayesian interpretation to develop a statistically meaningful model for truncation errors in EFTs
- Order-by-order convergence predicts size of truncation
- Can validate our predictions, diagnose convergence issues, etc.

Wondering how you can do it?

- Eq.(25) in above reference
- Generalizes common prescription
- Error bands are NOT Gaussian
- Model-checking diagnostics:
  - Furnstahl, Phillips, Klco, sw, PRC **92**, 024005 (2015)
  - Melendez, Furnstahl, sw PRC **96**, 024003 (2017)

In the limiting case of prior Set $A_\epsilon^{(1)}$, this integral can be evaluated explicitly [6],

$$\boxed{d_k^{(p)} = \bar{c}_{(k)}\, Q^{k+1} \times \begin{cases} \frac{n_c+1}{n_c}\, p\% & \text{if } p \leqslant \frac{n_c}{n_c+1}, \\ \left[\frac{1}{(n_c+1)(1-p\%)}\right]^{1/n_c} & \text{if } p > \frac{n_c}{n_c+1}, \end{cases} \quad (25)}$$

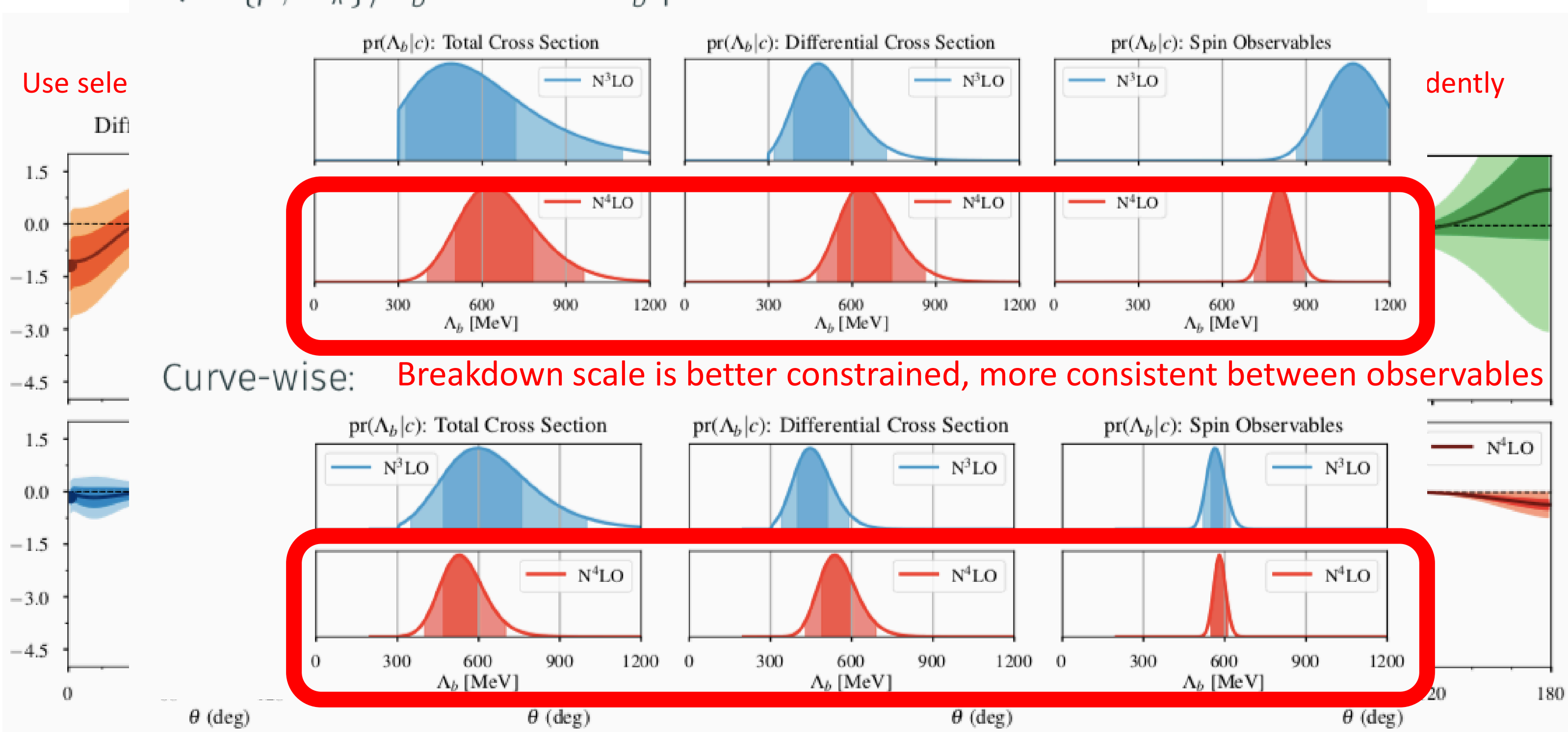
where n_c is again the number of nonzero known coefficients. Thus, with these priors, the interval of width $\bar{c}_{(k)} Q^{k+1}$ about the EFT prediction at order k is a $n_c/(n_c+1) * 100\%$ DOB interval, cf. Ref. [6]. Such a theory error bar has often been assigned in previous EFT calculations, and—as we discuss further in Sec. III—corresponds to the prescription formalized in Refs. [10,11]. It is important—e.g., in the context of error propagation—to keep in mind that this prior leads to a distribution of probability for the truncation error that is not Gaussian.

Improving the model: Gaussian processes

Gaussian process model for truncation errors:

- Predictions at nearby kinematics are not independent
- Use correlation information to improve truncation predictions
- Learn physics! diagnose breakdown scale + correlation length

$Q = \{p, m_\pi\}/\Lambda_b$. Estimate Λ_b point-wise:



Using EKM SCS R=0.9fm interaction

PRELIMINARY

Lessons from Bayesian methods in the NN sector

Tested by “fitting” (sampling posteriors) of NN LECs in partial waves; lessons are general

Case study 1: The usefulness of projected posterior plots

- Important for understanding the full information content of the data
- Most channels look Gaussian, but do statistical test before approximations!
- Use of projected posterior plots as a *physics diagnostic* illustrated by the fourth-order s-wave LECs ➔parameter degeneracy

Case study 2: LEC stability with maximum energy for fit

- What is optimal trade-off between more data to determine LECs more precisely and fit contamination at higher energies of omitted higher-order EFT terms?
- Sensitivity to E_{max} removed with Bayesian UQ ➔LECs should be independent
- Accounting for truncation errors; verify with E_{max} plots

Gaussian processes: Improving truncation error model using correlation information

- Nearby kinematics do not contain independent information
- Correlation information can be used to learn about the interaction
- The EFT breakdown can be statistically determined by examining convergence

Future Bayesian Plans (Collaborators wanted!)

- Use scattering observables directly for parameter estimation
- Identify appropriate priors for other theoretical expectations such as Wigner symmetry
- Propagate all sources of error, including from LECs, to few- and many-body observables
- Estimate the EFT expansion parameter from the expected convergence pattern of additional observable predictions (uniform matter in progress)
- Account for correlations between LECs from the πN, NN, and few-body sectors
- Assess the impact of available experimental data on the number of orders of the EFT that can be constrained via Bayesian model selection
- Employ Bayesian model checking techniques to verify for observable calculations that the EFT expansion is working ``as advertised"
- Model selection problems (which degrees of freedom to use)?

References

Some BUQEYE publications on UQ for EFT

- “A recipe for EFT uncertainty quantification in nuclear physics”, J. Phys. G 42, 034028 (2015)
- “Quantifying truncation errors in effective field theory”, Phys. Rev. C 92, 024005 (2015) [with Natalie Klco]
- “Bayesian parameter estimation for effective field theories”, J. Phys. G 43, 074001 (2016)
- “Bayesian truncation errors in chiral EFT: nucleon-nucleon observables”, Phys. Rev. C 96, 024003 (2017)

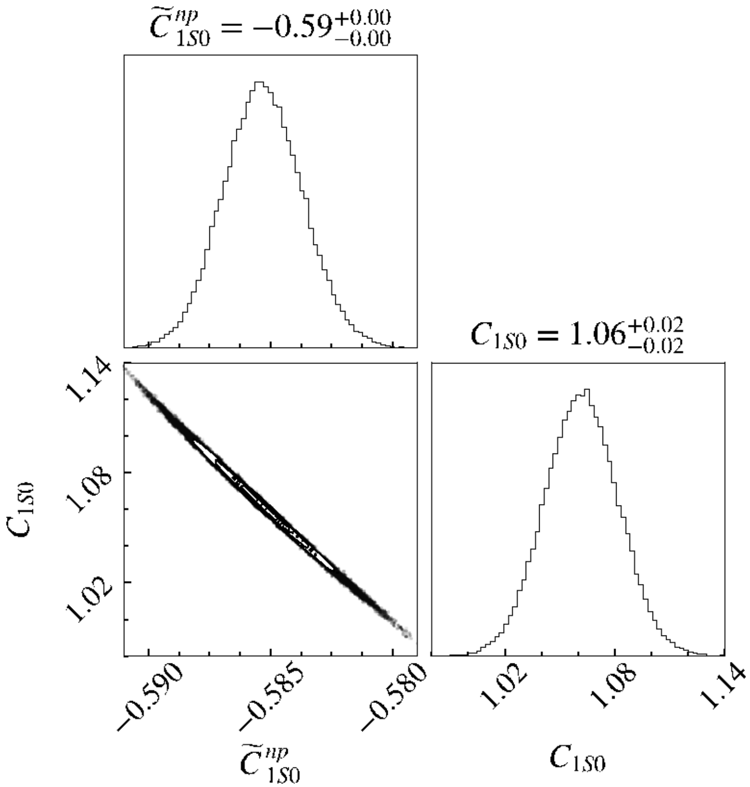
Recent publications on fitting chiral EFT interactions

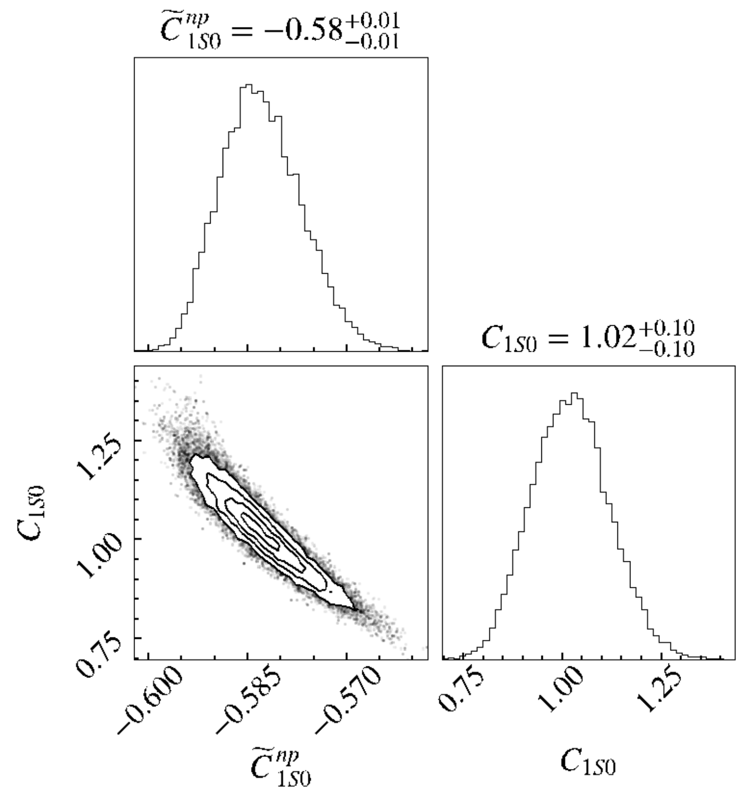
- D. R. Entem, R. Machleidt, and Y. Nosyk. “High-quality two-nucleon potentials up to fifth order of the chiral expansion,” Phys. Rev. C 96, 024004, 2017.
- B. Carlsson, A. Ekström, C. Forssén, D. F. Strömberg, O. Lilja, et al. “Uncertainty analysis and order-by-order optimization of chiral nuclear interactions,” Phys. Rev. X, 011019, 2015.
- A. Gezerlis, I. Tews, E. Epelbaum, M. Freunek , S. Gandolfi , K. Hebeler, A. Nogga, A. Schwenk. “Local chiral effective field theory interactions and quantum Monte Carlo applications,” Phys.Rev. C 90, 054323 (2014).
- M. Piarulli, L. Girlanda, R. Schiavilla, R. Navarro Ṕerez, J. E. Amaro, and E. Ruiz Arriola. “Minimally nonlocal nucleon-nucleon potentials with chiral two-pion exchange including resonances,” Phys. Rev. C, 024003, 2015.
- E. Epelbaum, H. Krebs, and U. G. Meißner. “Improved chiral nucleon-nucleon potential up to next-to-next-to-next-to-leading order,” Eur. Phys. J. A 51, 53, 2015.
- P. Reinert, H. Krebs, and E. Epelbaum. “Semilocal momentum-space regularized chiral two-nucleon potentials up to fifth order,” arXiv:1711.08821 2017.

Back-up Slides

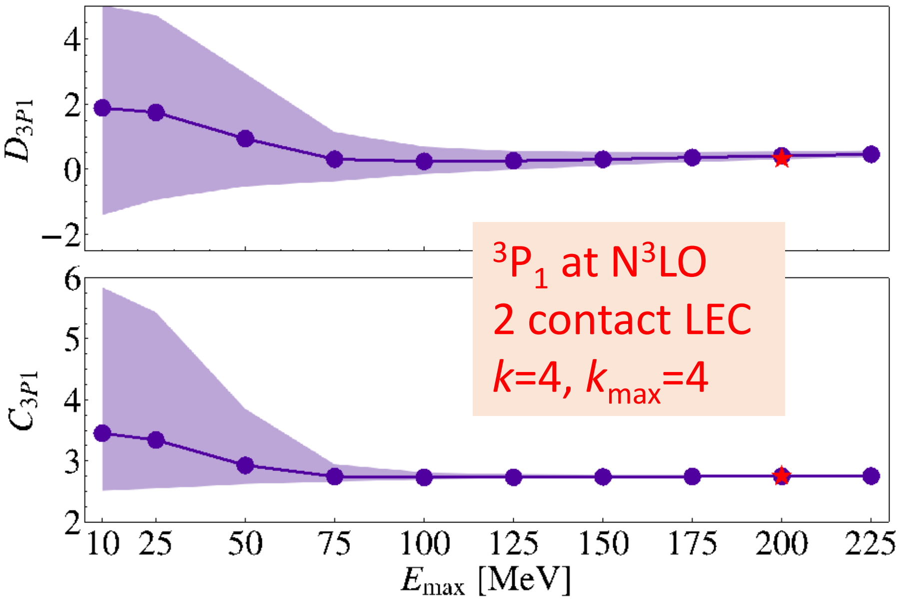
Case study 2: LEC stability with maximum energy for fit

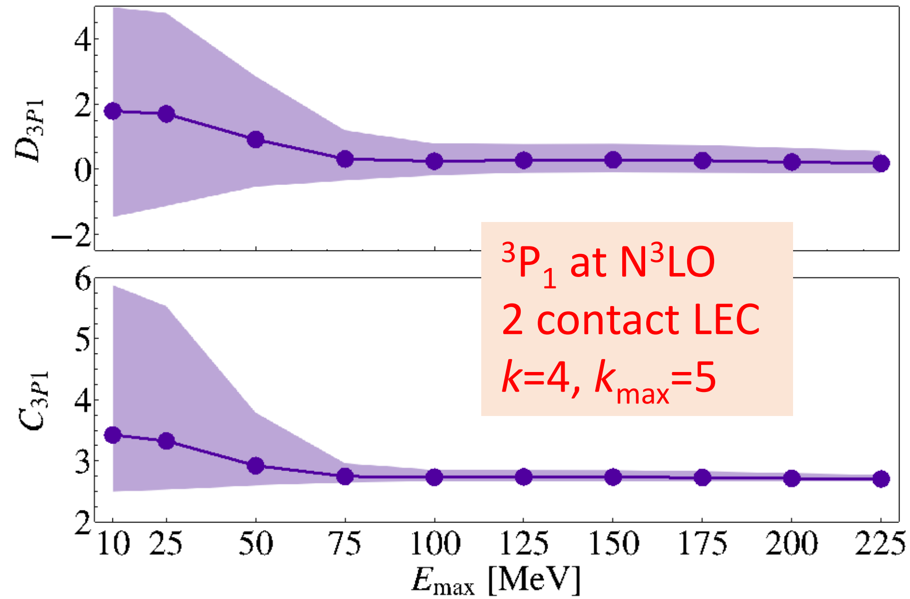
Now fitting to partial-wave *cross section:* Not completely Gaussian, but close





Case study 2: LEC stability with maximum energy for fit

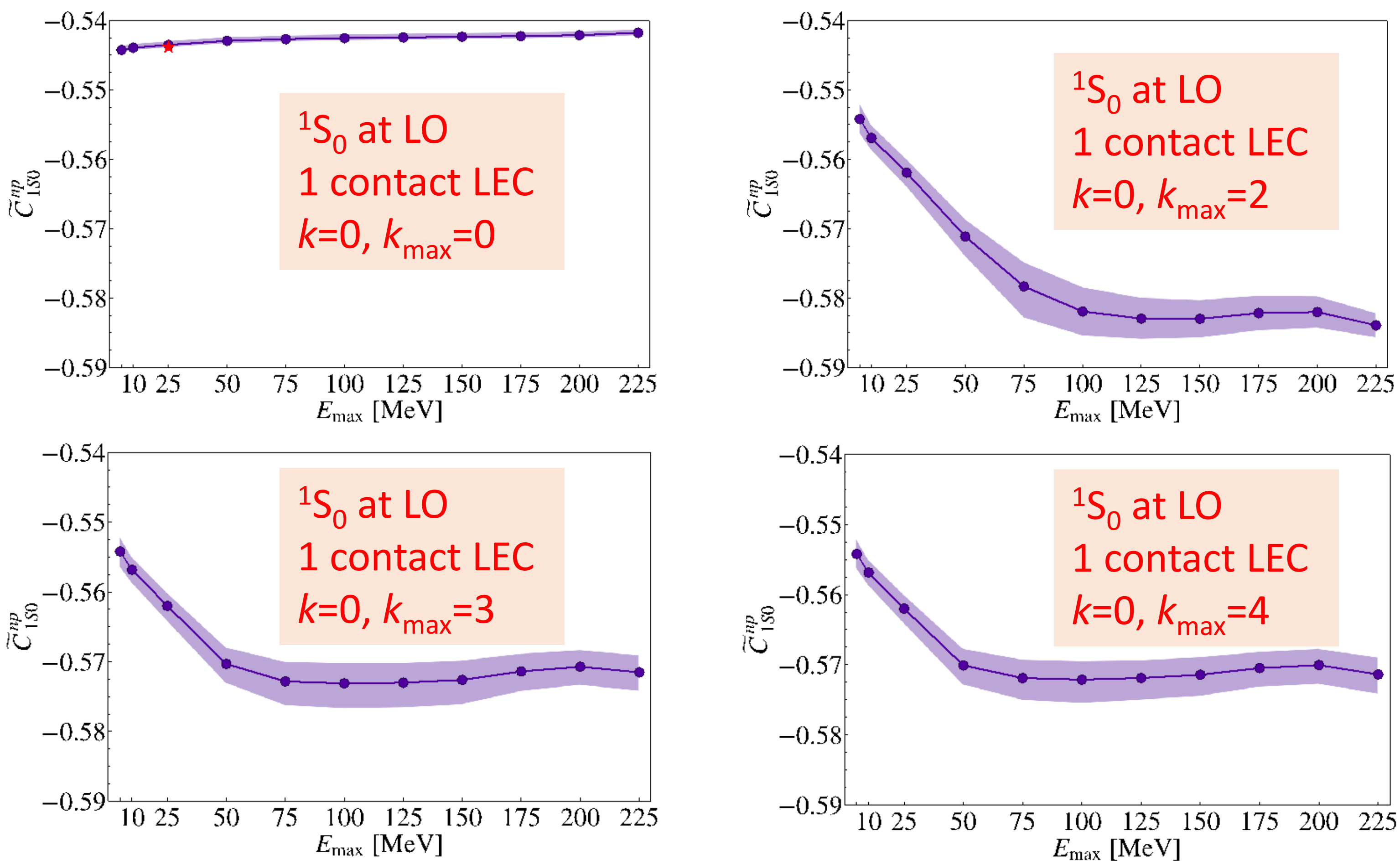




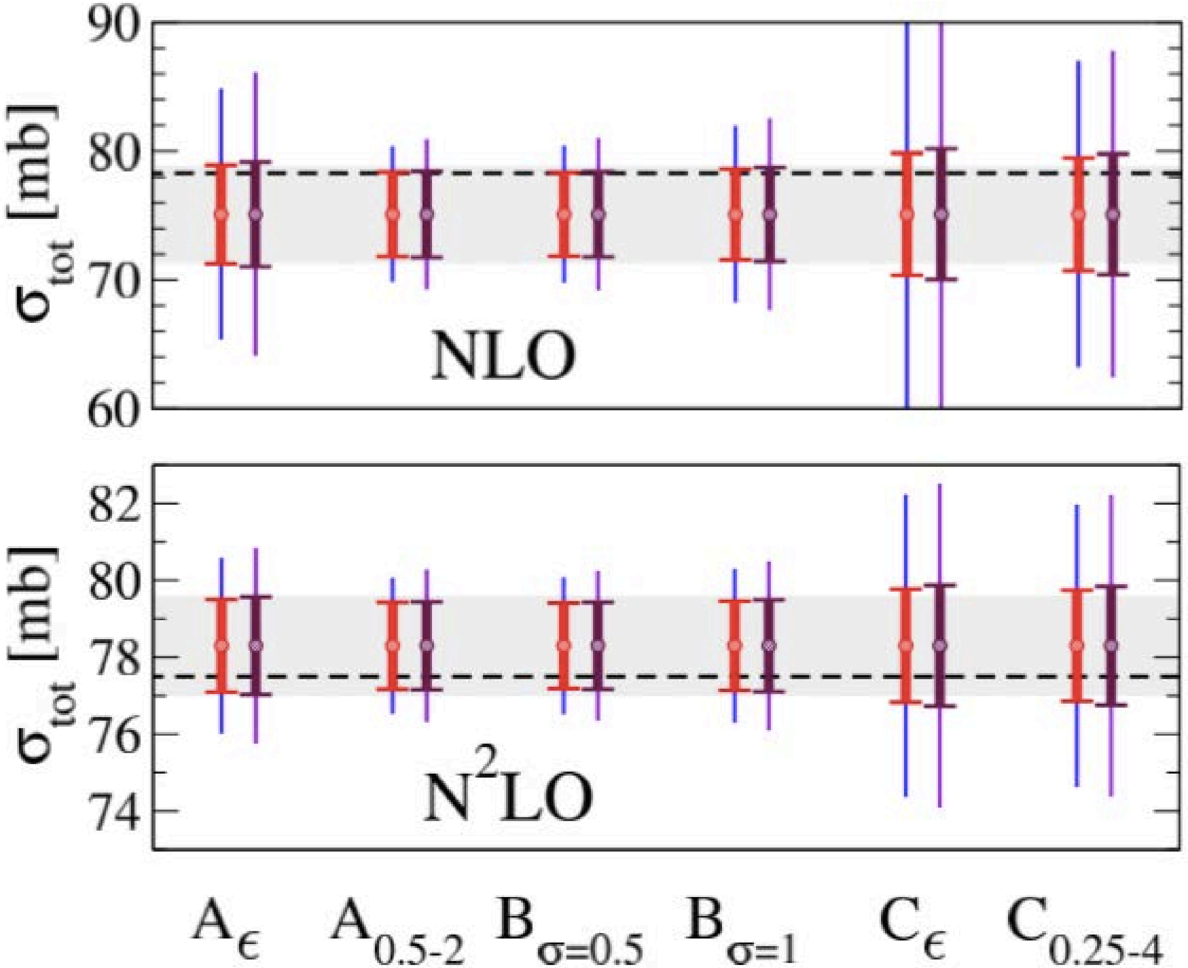
Less impact from neglected higher-order terms at higher EFT orders (larger *k*)

Here including truncation error for $^3P_1$ at N^{3}LO makes parameter estimates less E_{max} dependent and gives larger uncertainty (blow up to see details)

Case study 2: LEC stability with maximum energy for fit



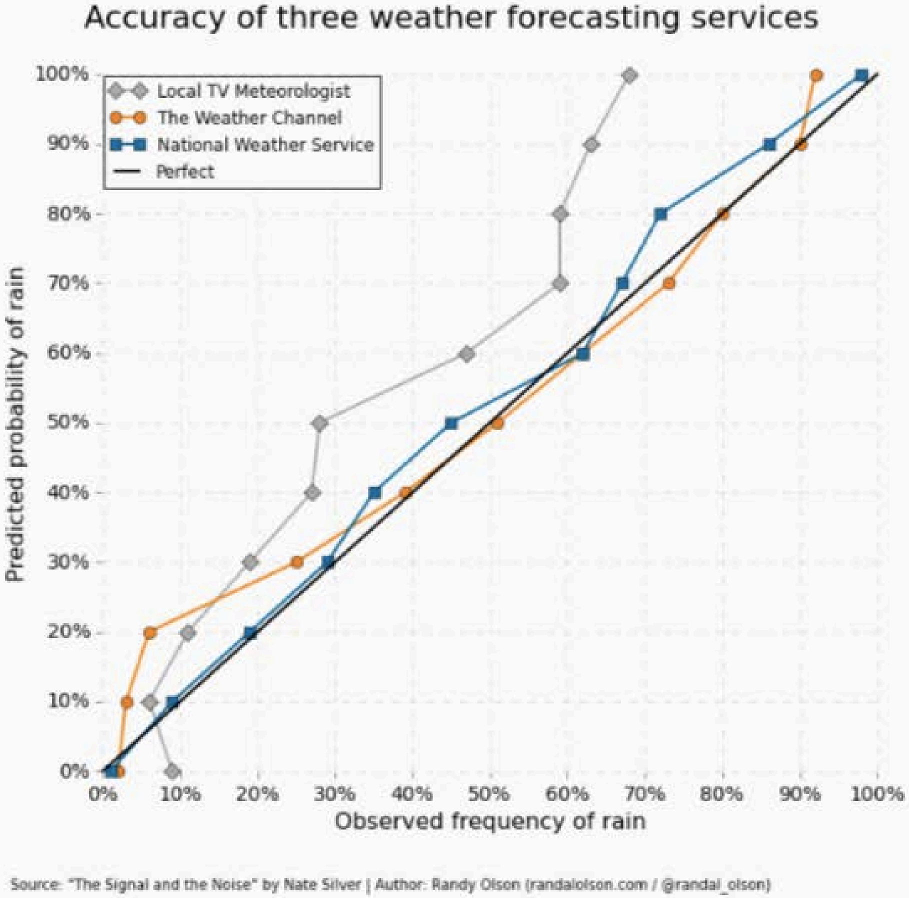
Different prior assumptions for truncation errors



EKM SCS
R=0.9fm
E_{lab} = 96MeV

Furnstahl, Phillips, Klco, sw, PRC **92**, 024005 (2015)

Credible Interval Diagnostic



- Tests model against reality
- Generalized to GPs in Bastos and O'Hagan (2009)
- Called "consistency plots" in Melendez et al. (2017)

Nuclear Matter [PRELIMINARY]

- First application to 3-body forces
- Uncertainty helps compare to empirical saturation