Case studies in Bayesian parameter estimation for chiral effective field theory

Sarah Wesolowski
Department of Mathematics and Computer Science
Salisbury University

Dick Furnstahl

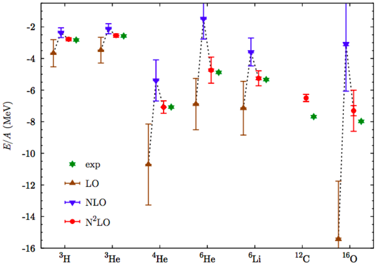
Daniel Phillips

BUQEYE Collaboration

Supported in part by DOE, NSF, SciDAC NUCLEI project

Uncertainty quantification in *ab initio* calculations

Are we there yet?

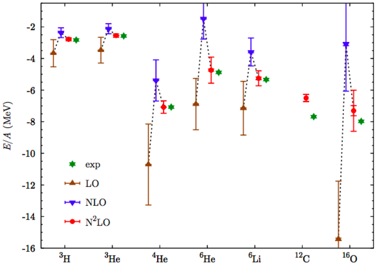


Recent QMC result for binding energy per nucleon
Lonardoni et. al, [arXiv:1709.09143], cf. Ingo's talk
Two types of error bars: method and truncation errors

- *ab initio* methods continue to improve
- Need for uncertainty quantification (UQ) for input nuclear interactions

Uncertainty quantification in *ab initio* calculations

Are we there yet?

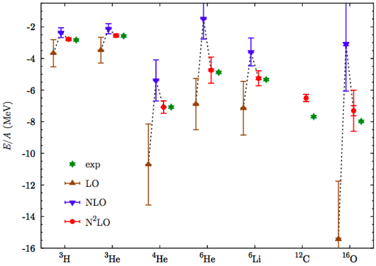


Recent QMC result for binding energy per nucleon
Lonardoni *et.* al, [arXiv:1709.09143], cf. Ingo's talk
Two types of error bars: method and truncation errors

- *ab initio* methods continue to improve
- Need for uncertainty quantification (UQ) for input nuclear interactions
- Issues remain in χEFT:
  - Regulator artifacts
  - Convergence
  - LEC fitting in NN and 3N
  - What degrees of freedom?

Uncertainty quantification in *ab initio* calculations

Are we there yet?



Recent QMC result for binding energy per nucleon
Lonardoni *et*. al, [arXiv:1709.09143], cf. Ingo's talk
Two types of error bars: method and truncation errors

- *ab initio* methods continue to improve
- Need for uncertainty quantification (UQ) for input nuclear interactions
- Issues remain in χEFT:
  - Regulator artifacts
  - Convergence
  - LEC fitting in NN and 3N
  - What degrees of freedom?
- Need full, statistically meaningful UQ!
  I.e., only as successful as the p% interval predicts

Focus on low-energy constant (LEC) estimation.

(which is entangled with other uncertainties in the calculation)

The Bayes Way

EFTs are special because they have a convergence *pattern*:

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

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

- Bayes lets us consistently incorporate EFT details: priors:

  naturalness of LECs/observable expansion, truncation errors, breakdown scale.

The Bayes Way

EFTs are special because they have a convergence *pattern*:

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

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

- Bayes lets us consistently incorporate EFT details: priors:

  naturalness of LECs/observable expansion, truncation errors, breakdown scale.

- We can tell if EFTs are working $\rightarrow$ model checking!

  Furnstahl et al, PRC (2015) and Melendez et al. PRC (2017)

The Bayes Way

EFTs are special because they have a convergence *pattern*:

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

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

- Bayes lets us consistently incorporate EFT details: priors:

  naturalness of LECs/observable expansion, truncation errors, breakdown scale.

- We can tell if EFTs are working $\rightarrow$ model checking!
  Furnstahl et al, PRC (2015) and Melendez et al. PRC (2017)

- When can we improve on traditional methods?
  - χ^2 optimization procedures vs. Bayesian posteriors.
  - error propagation with covariance matrices
  - adding errors in quadrature

The Bayes Way

EFTs are special because they have a convergence *pattern*:

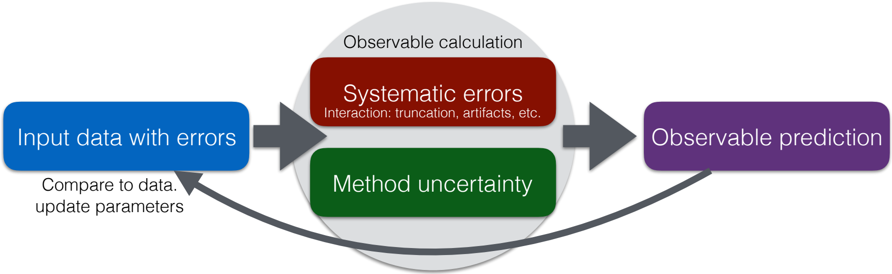
$$X(p) = X_0 \sum_{n=0}^{k} c_n\, Q^n, \qquad \Delta_k = \sum_{n=k+1}^{k_{\mathrm{max}}} c_n Q^n$$

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

- Bayes lets us consistently incorporate EFT details: priors:
  naturalness of LECs/observable expansion, truncation errors, breakdown scale.
- We can tell if EFTs are working $\rightarrow$ model checking!
  Furnstahl et al, PRC (2015) and Melendez et al. PRC (2017)
- When can we improve on traditional methods?
  - χ^2 optimization procedures vs. Bayesian posteriors.
  - error propagation with covariance matrices
  - adding errors in quadrature
- Parameter estimation framework for LEC estimation.
  Furnstahl et. al, J. Phys. G (2015) and sw et al., J. Phys. G (2016)
- Can combine LEC + truncation error consistently.
  Coming soon: sw, Furnstahl, and Phillips

Sources of error in parameter estimation

What goes into the parameter estimation procedure?



- Parameter estimation procedure entangles sources of error.
- Data + priors $\rightarrow$ sampling $\rightarrow$ LEC posterior distribution
- Focus on semi-local (coordinate-space) interaction of Epelbaum, Krebs, and Meißner (EKM) for case studies.
  NN contact terms in partial waves.
  Epelbaum, Krebs, and Meißner, Eur. Phys. J. A **51** (2015)
  Epelbaum, Krebs, and Meißner, PRL **115** (2015)

How does truncation error enter into the parameter estimation procedure?

χ^2-likelihood depends on observable calculation

$$\chi^2(\mathbf{a}) = \sum_{i=1}^{N} \left(\frac{d_i - X^k(p_i, \mathbf{a})}{\sigma_i^2} \right)$$

- Full observable $X(p) = X^k(p; \mathbf{a}) + X_0(p) \sum_{k+1}^{k_{\max}} c_n Q^n$

How does truncation error enter into the parameter estimation procedure?

χ^2-likelihood depends on observable calculation

$$\chi^2(\mathbf{a}) = \sum_{i=1}^{N} \left(\frac{d_i - X^k(p_i, \mathbf{a})}{\sigma_i^2} \right)$$

- Full observable $X(p) = X^k(p; \mathbf{a}) + X_0(p) \sum_{k+1}^{k_{\max}} \mathbf{c_n} Q^n$

Marginalize to introduce higher-order c_n's into likelihood

$$\mathrm{pr}(D|\mathbf{a}, I) = \int \mathrm{d}c_{k+1} \cdots \mathrm{d}c_{k_{\max}}$$

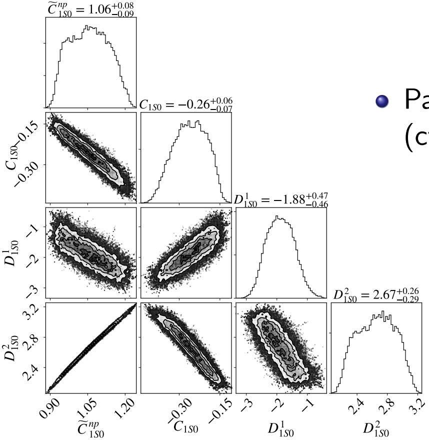
Stump *et al* Phys. Rev. D **65** (2001)

$$\mathrm{pr}(D|c_{k+1}, \cdots, c_{k_{\max}}, \mathbf{a}, I) \times \mathrm{pr}(c_{k+1}, \cdots, c_{k_{\max}}|I)$$

The final posterior $\mathrm{pr}(\mathbf{a}|D, I) \propto$ likelihood $\times$ prior

NN problem:
Numerous, high-precision data: what more do we learn?

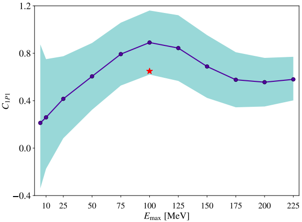
posterior for $^1S_0$ terms at N^{3}LO



- Parameter posteriors: *s*-wave redundancy. (cf. Christian and Hermann's talks)

Learning physics from Bayesian posteriors

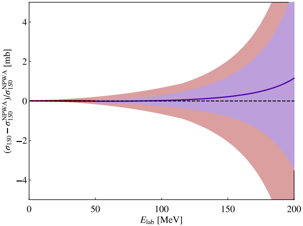
NN problem:
Numerous, high-precision data: what more do we learn?



- Parameter posteriors: *s*-wave redundancy. (cf. Christian and Hermann's talks)
- What range of data to use? $E_{\max}$ plots.

Learning physics from Bayesian posteriors

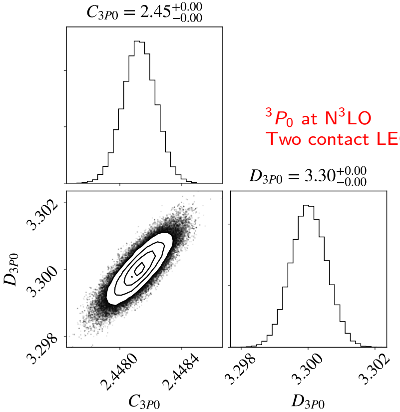
NN problem:
Numerous, high-precision data: what more do we learn?



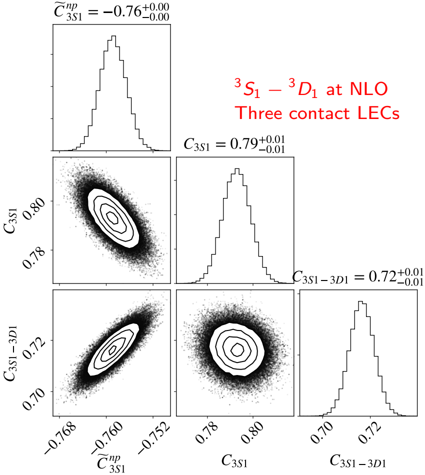
- Parameter posteriors: *s*-wave redundancy. (cf. Christian and Hermann's talks)
- What range of data to use? $E_{\max}$ plots.
- Formalism: combine truncation and parameter errors to make predictions.
- Briefly demonstrate these without getting into too much formalism.

Case study 1: parameter posterior for NN contact LECs

- Prior input information: naturalness and truncation errors.
- Framework outputs (LEC) parameter posteriors with uncertainties consistently included:



$^3P_0$ at N^{3}LO
Two contact LECs



$^3S_1$ – $^3D_1$ at NLO
Three contact LECs

Case study 1: redundancy in the s-waves

The operators in the s-wave sat N^{3}LO (Q^4) may be rewritten as:

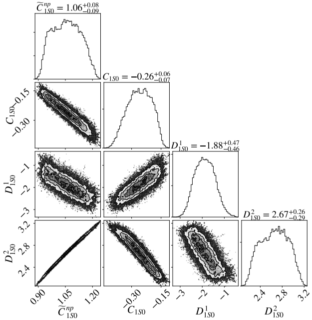
$$
\begin{aligned}
D^1_{(1S0)} p^2 p'^2 + D^2_{(1S0)} (p^4 + p'^4) \\
&= \frac{1}{4}(D^1_{(1S0)} + 2D^2_{(1S0)})(p^2 + p'^2)^2 \\
&\quad - \frac{1}{4}(D^1_{(1S0)} - 2D^2_{(1S0)})(p^2 - p'^2)^2 \\
&= (D^1_{(1S0)} + 2D^2_{(1S0)}) p^2 p'^2 + D^2_{(1S0)} (p^2 - p'^2)^2
\end{aligned}
$$

Case study 1: redundancy in the s-waves

The operators in the s-wave sat $\mathrm{N}^3\mathrm{LO}$ (Q^4) may be rewritten as:

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

The LEC posterior is non-Gaussian behavior with large correlations:



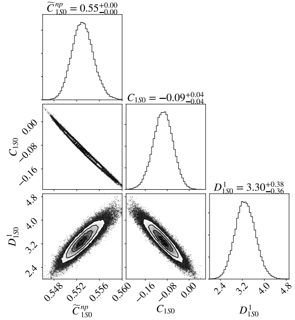
Case study 1: redundancy in the s-waves

The operators in the s-wave sat N^{3}LO (Q^4) may be rewritten as:

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

But becomes more Gaussian with restriction of parameters $D^2_{1S0} = 0$:

and equivalent description of data!



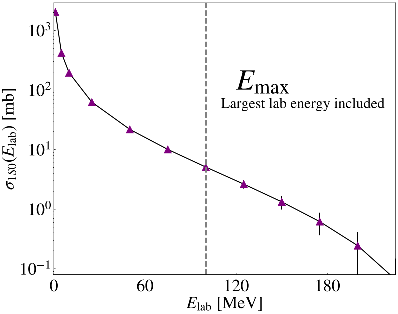
Case study 2: what range of data to use?

EFT convergence

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

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

As we go to higher E_{max}, more terms contribute in the expansion.



Case study 2: what range of data to use?

EFT convergence

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

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

As we go to higher $E_{\max}$, more terms contribute in the expansion.

- Including truncation error in fit procedure avoids overfitting.
- $E_{\max}$ plots: how high to make $k_{\max}$ to absorb UV physics?
- Bayesian model selection makes quantitative statements about how many terms are constrained by data.
- $E_{\max}$ plots serve as simpler proxy to model selection.

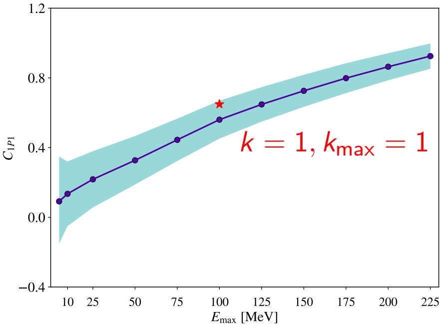
Case study 2: what range of data to use?

EFT convergence

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

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

As we go to higher $E_{\max}$, more terms contribute in the expansion.



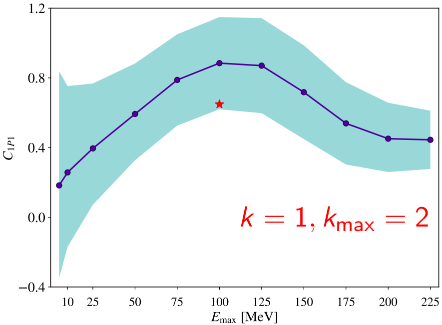
Case study 2: what range of data to use?

EFT convergence

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

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

With enough terms, LEC saturates as function of E_{max}.



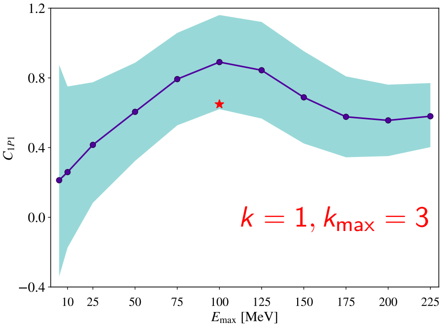
Case study 2: what range of data to use?

EFT convergence

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

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

With enough terms, LEC saturates as function of $E_{\max}$.



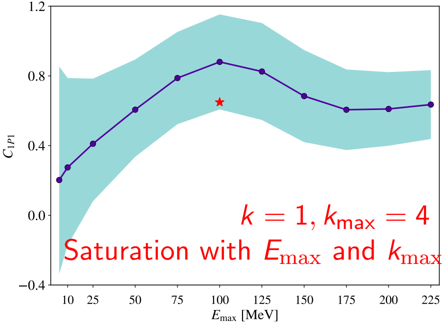
Case study 2: what range of data to use?

EFT convergence

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

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

With enough terms, LEC saturates as function of $E_{\max}$.

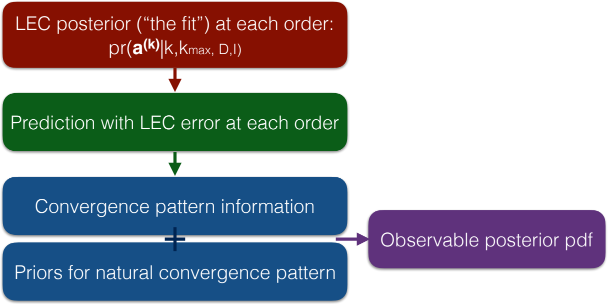


Case study 3: combining LEC and truncation uncertainty

How to make predictions with combined uncertainties?

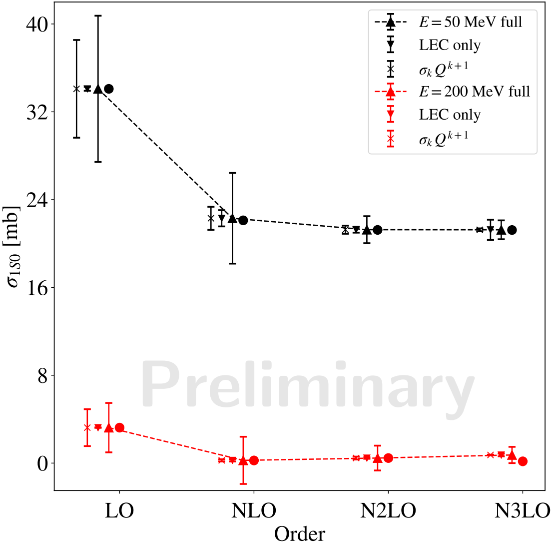
Using Bayesian methods, can derive a posterior for observables:

$$\mathrm{pr}(X(p)|D, k, k_{\mathrm{max}}, I)$$



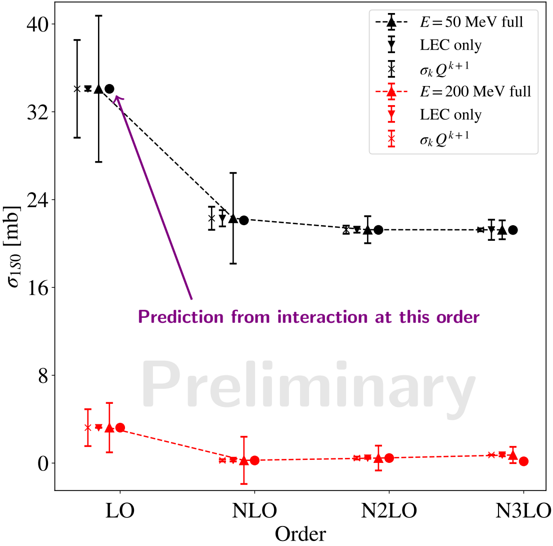
Case study 3: combining LEC and truncation uncertainty

How to make predictions with combined uncertainties?



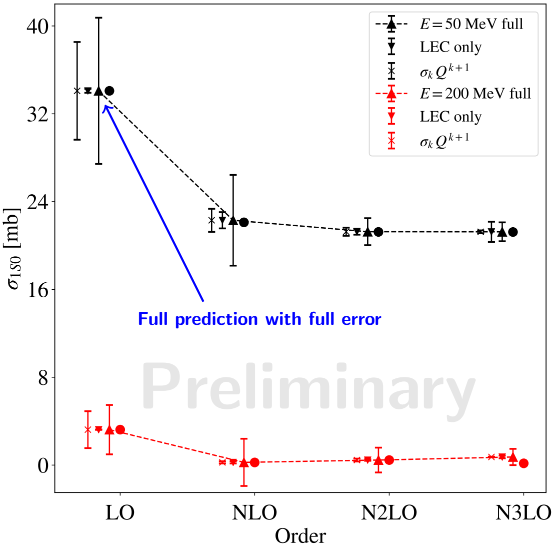
Case study 3: combining LEC and truncation uncertainty

How to make predictions with combined uncertainties?



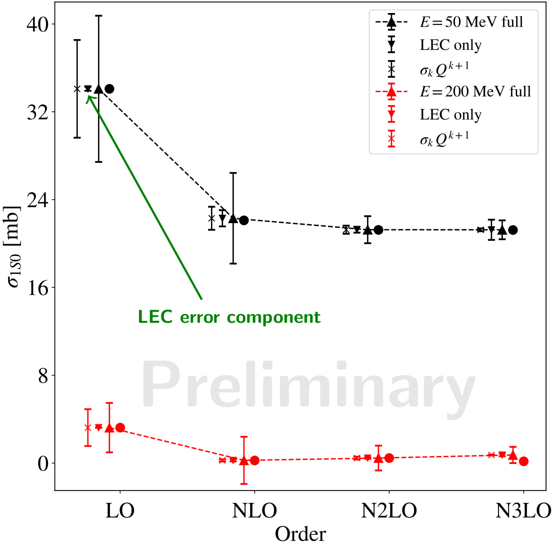
Case study 3: combining LEC and truncation uncertainty

How to make predictions with combined uncertainties?



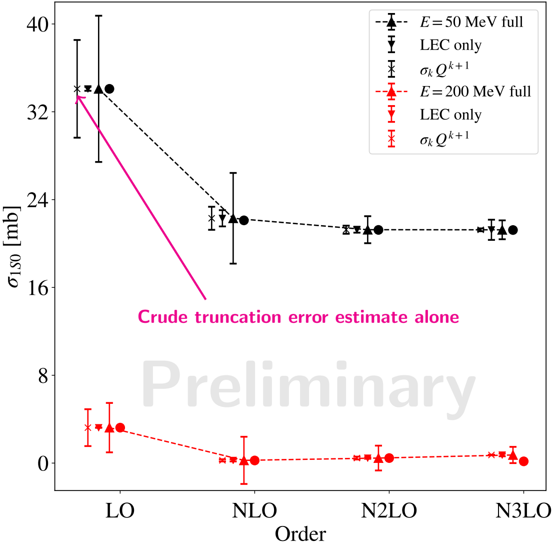
Case study 3: combining LEC and truncation uncertainty

How to make predictions with combined uncertainties?



Case study 3: combining LEC and truncation uncertainty

How to make predictions with combined uncertainties?

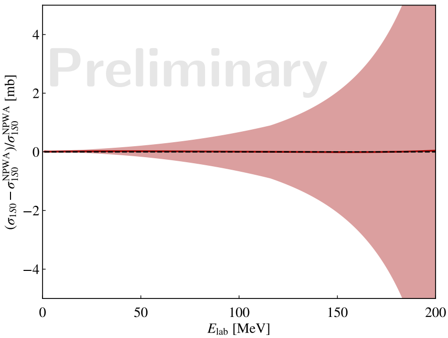


Combining LEC and truncation uncertainty: results

Plot relative uncertainty

$$\sigma_{\text{res}} = (\sigma_{\text{pred.}} - \sigma_{\text{NPWA}})/\sigma_{\text{NPWA}}$$

Note: this is singlet s component only!



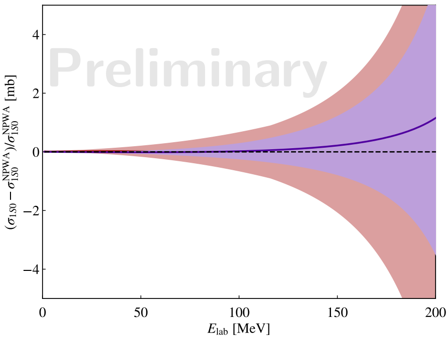
- Leading-order (Q^0) estimate reflects only prior input (not shown here)
- Next-to-leading order (Q^2) starts to know about convergence

Combining LEC and truncation uncertainty: results

Plot relative uncertainty

$$\sigma_{\text{res}} = (\sigma_{\text{pred.}} - \sigma_{\text{NPWA}})/\sigma_{\text{NPWA}}$$

Note: this is singlet s component only!



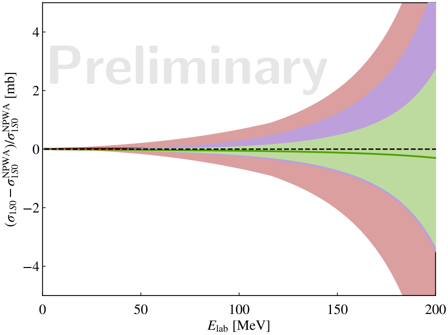
- Leading-order (Q^0) estimate reflects only prior input (not shown here)
- Next-to-leading order (Q^2) starts to know about convergence
- N^2LO (Q^3) includes even more convergence information

Combining LEC and truncation uncertainty: results

Plot relative uncertainty

$$\sigma_{\text{res}} = (\sigma_{\text{pred.}} - \sigma_{\text{NPWA}})/\sigma_{\text{NPWA}}$$

Note: this is singlet s component only!



- Leading-order (Q^0) estimate reflects only prior input (not shown here)
- Next-to-leading order (Q^2) starts to know about convergence
- N^2LO (Q^3) includes even more convergence information
- Finally, N^3LO (Q^4) has converged even further.

Summary and outlook

Summary

- Bayes $\rightarrow$ consistent analysis of error in *ab initio* calculations
- Here: EFT truncation error + LEC error/correlations.
- Model checking and validation possible!
- Can extract physics insight based on data
(cf. *s*-wave redundancy)

Outlook

- Extend to 3N and problems with few data.
- Bayesian model selection for deciding between formulations.
- Working on model selection for nd scattering in pionless EFT.
- Make code accessible for interaction practitioners.