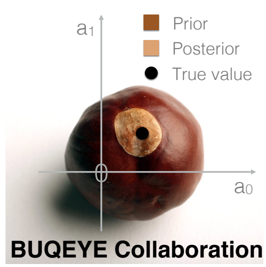
Moving forward with Bayesian parameter estimation for chiral EFT

Sarah Wesolowski

Progress in *ab initio* techniques in nuclear physics, TRIUMF 2019

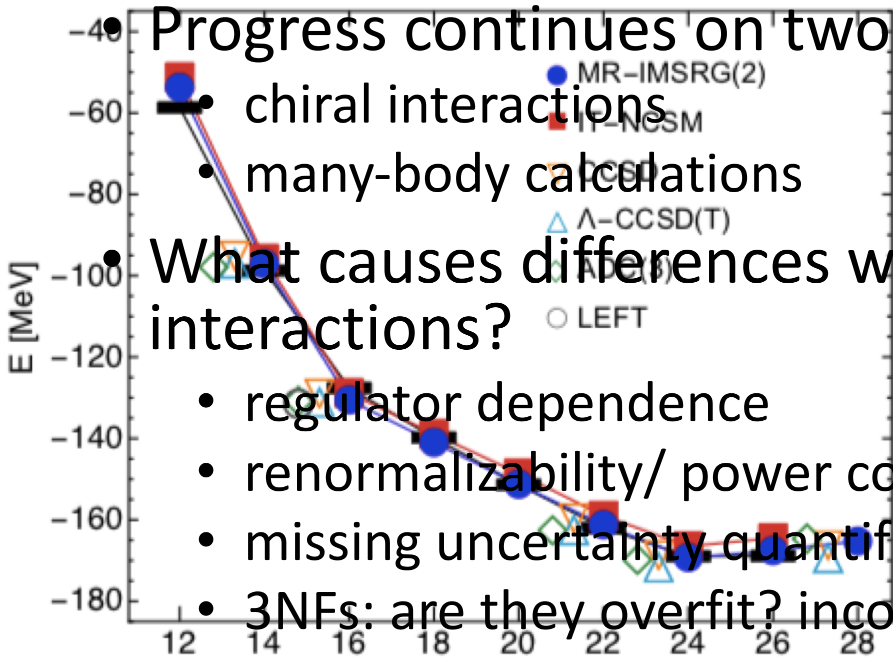


Dick Furnstahl (OSU)
Jordan Melendez (OSU)
Matt Pratola (OSU statistics)
Daniel Phillips (OU)

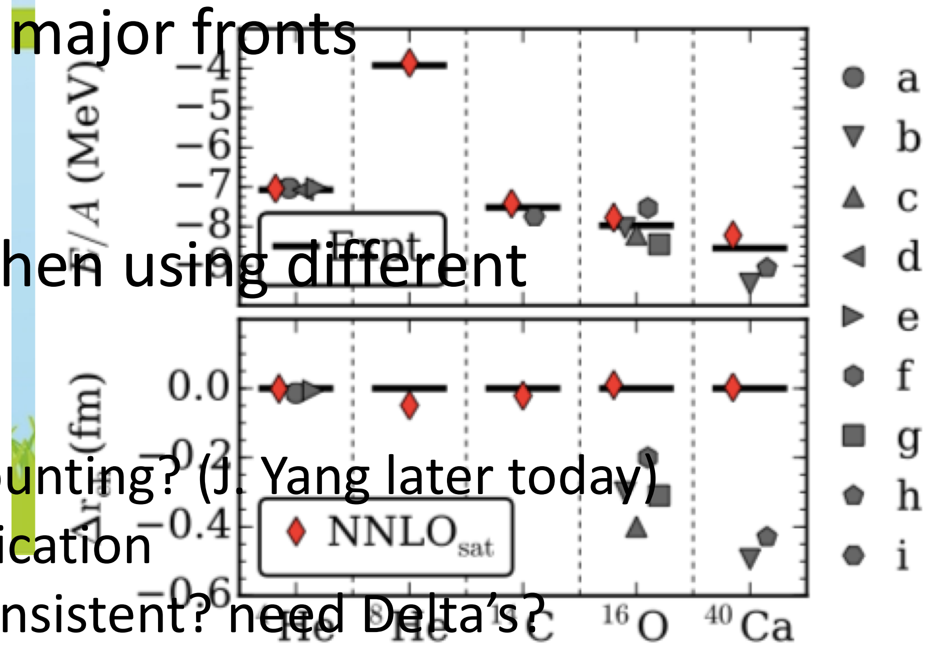
Supported in part by NSF, DOE, and the SciDAC NUCLEI project

Precision calculations go to the zoo

- A zoo of modern chiral effective field theory (EFT) interactions is now available (cf. Robert's talk)
- Progress continues on two major fronts
  - chiral interactions
  - many-body calculations
- What causes differences when using different interactions?
  - regulator dependence
  - renormalizability/ power counting? (J. Yang later today)
  - missing uncertainty quantification
  - 3NFs: are they overfit? inconsistent? need Delta's?
  - you can probably think of more...



Hergert et al., Journal of Physics: Conference Series (2018)



Ekstrom et al. PRC 91 051301 (2015)

Zoo of precise many-body methods

Zoo of interactions (maybe only 2 really...)

Recent progress in fitting interactions

- Two-body: nucleon-nucleon (NN) sector
  - LENPIC few/many-body NN calculations
    Binder et al. PRC **98**, 014002 (2018)
  - Ongoing work at Chalmers group beyond N^2LO sim/sep
    Carlsson et al., PRX **6** 011019 (2016)
  - Local chiral potentials with explicit Deltas
    Piarulli at al., PRC 91, 024003 (2015)
  - BUQEYE: recent Bayesian exploration (in partial waves)
    Wesolowski et al., JPG 46, 045102 (2019)
- πN sector: coefficients from Roy-Steiner analysis
  Hoferichter at al., Phys. Rev. Lett. **115**, 192301 (2015)
- Three-body: NNN sector
  - LENPIC fits of c_D and c_E + few/many-body calculations
    Kai, Hermann, and Robert's talks;, Epelbaum et al., arxiv:1807.02848
  - More work with light nuclei in QMC
    Baroni et al., PRC **98** (2018) no.4, 044003
    Piarulli et al., PRL **120** (2018) no.5, 052503
    Lynn et al., PRL 116 (2016) 062501
  - BUQEYE+Chalmers to contrain c_D and c_E (Bayesian methods)

Case studies in NN parameter estimation

- Question: what is the value added of using Bayesian methods, especially in the NN sector?
- Answer: can explore issues in a controlled setting
    - Prior assumptions such as naturalness constraints

For *all* the details see:

Wesolowski et al., J Phys G **46**, 045102 (2019)

    - Impact and inclusion of truncation errors
- Observable calculations are quick, as opposed to few-body sector, so understand in NN sector first
- Ultimate products: making progress on UQ + understanding and testing physics issues

Constraining and testing an EFT: the user's guide

1. Outline EFT and all sources of uncertainty
2. Be as explicit (i.e. Bayesian) as possible
   (That doesn't technically mean changing your procedure!)
3. Check your results. Validation is key.

For even more details see:
Wesolowski et al., J. Phys. G **43**, 074001 (2016)

Setup: sources of uncertainty

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

- Method/numerical errors (often well-controlled)
- Theoretical errors (don't just exist in EFT)
  - EFTs: truncation uncertainty
  - EFTs: Correlations between observables (3NFs)
- Experimental errors, correlated or uncorrelated

Put together everything you know about each source: in Bayes language these are "priors"

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

Parameter estimation

- Framework for parameter estimation
  Furnstahl et al., JPG 42 (2015) no.3, 034028
  Wesolowski et al., JPG 43 (2016) no.7, 074001
  Wesolowski et al., JPG (2019) in press
- Describing observables with theory:

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

- Reduces to χ^2 fitting without theory error
- But we know theoretical error exists, especially for effective field theories

$$y_{\mathrm{th}} = y_{\mathrm{ref}} \sum_{n=0}^{k} c_n Q^n \qquad \delta y_{\mathrm{th}} = y_{\mathrm{ref}} \sum_{n=k+1}^{\infty} c_n Q^n$$

For chiral EFT: $Q = \max\{p, m_\pi\}/\Lambda_b$

Parameter estimation

- What we want to calculate:

$$\mathrm{pr}(\vec{a}_k \mid \boldsymbol{y}_{\mathrm{exp}}, \Sigma_{\mathrm{exp}}, \Sigma_{\mathrm{th}}, I)$$

Any other background e.g., EFT naturalness

EFT low-energy constants at order k in the theory

Experimental values and uncertainties

Theory uncertainty

- Posterior pdf with naturalness and truncation error:

$$\mathrm{pr}(\vec{a}_k \mid \boldsymbol{y}_{\mathrm{exp}}, \Sigma_{\mathrm{exp}}, \Sigma_{\mathrm{th}}) \propto \mathrm{pr}(\boldsymbol{y}_{\mathrm{exp}} \mid \boldsymbol{y}_{\mathrm{th}}, \Sigma_{\mathrm{exp}}, \Sigma_{\mathrm{th}})\, \mathrm{pr}(\vec{a}_k \mid \bar{a})$$

Likelihood

Prior

- Naturalness: encoded in $\bar{a}$ "hyperparameter"
- Truncation error: Σ_{th}, theory error assumptions

Exploring projected posteriors

- Starting slow:

$$\boldsymbol{y}_{\mathrm{exp}} = \boldsymbol{y}_{\mathrm{th}} + \boxed{\delta \boldsymbol{y}_{\mathrm{exp}}}$$

- Work in a regime where theory error is very small
- High enough order where truncation error is small
- Regular least-squares likelihood with Gaussian prior

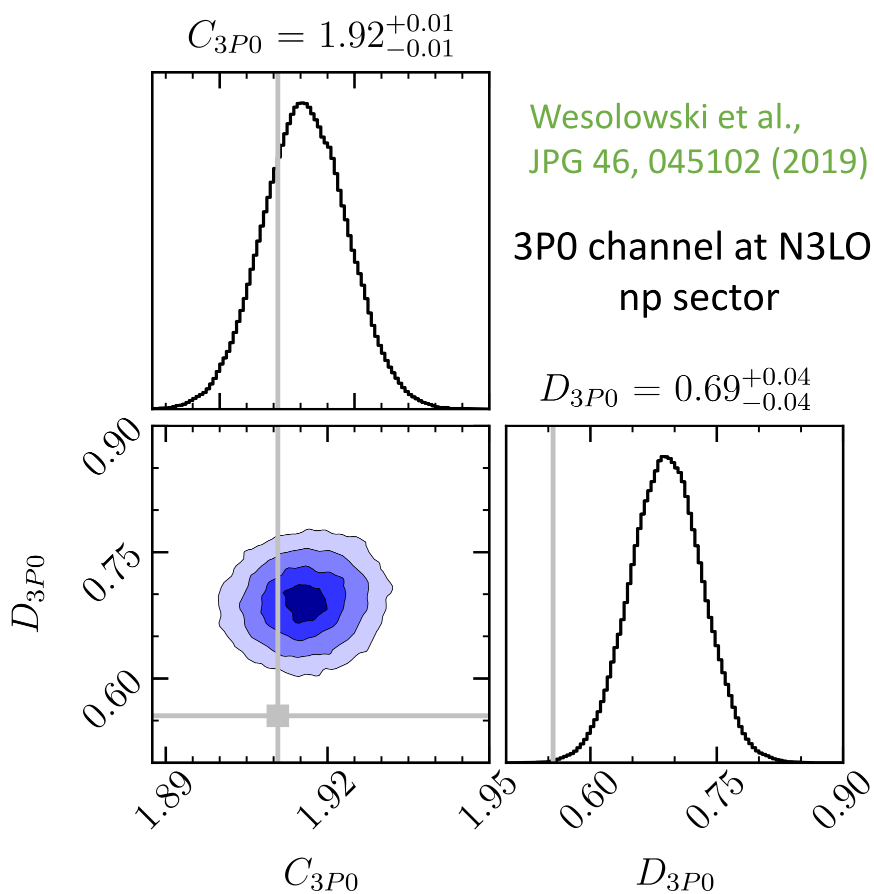
$$\mathrm{pr}(\vec{a}_k | \boldsymbol{y}_{\mathrm{exp}}, \Sigma_{\mathrm{exp}}) \propto \mathrm{pr}(\boldsymbol{y}_{\mathrm{exp}} | \vec{a}_k, \Sigma_{\mathrm{exp}})\ \mathrm{pr}(\vec{a}_k)$$

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

$$\boldsymbol{r} \equiv \boldsymbol{y}_{\mathrm{exp}} - \boldsymbol{y}_{\mathrm{th}}$$

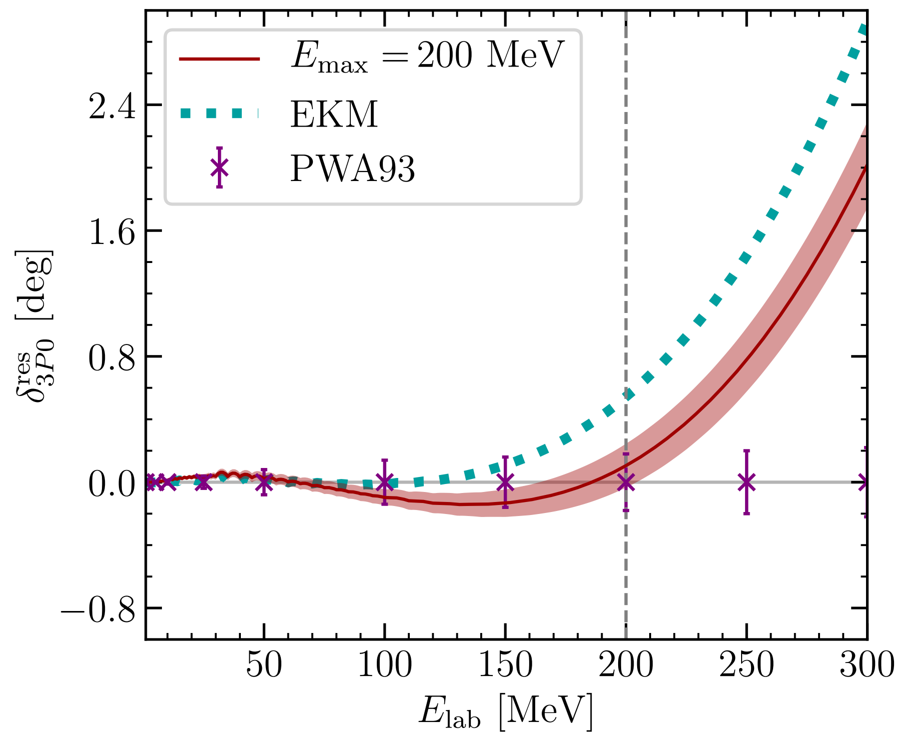
Exploring projected posteriors

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



Wesolowski et al., JPG 46, 045102 (2019)

3P0 channel at N3LO
np sector

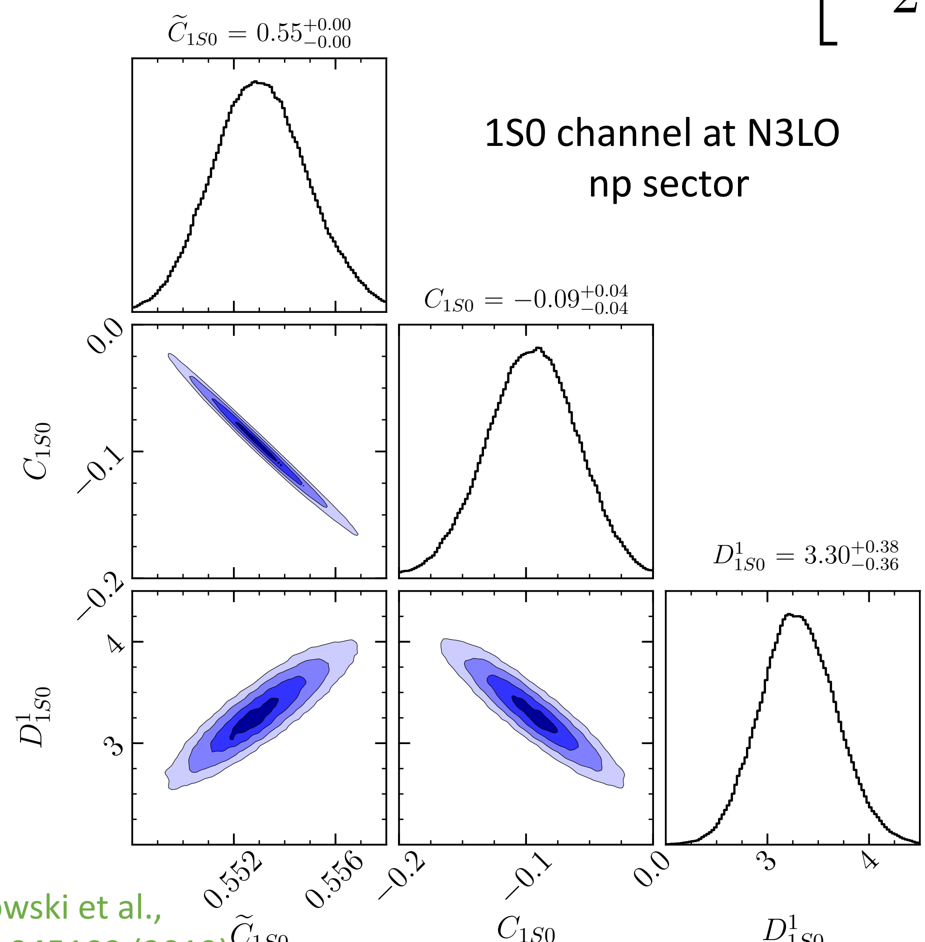


$$\delta^{\mathrm{res}}_{\mathrm{p.w.}} = \delta^{\mathrm{pred}}_{\mathrm{p.w.}} - \delta^{\mathrm{PWA93}}_{\mathrm{p.w.}}$$

Semilocal interactions from: Epelbaum et al., EPJ A 51 (2015) 53, Epelbaum et al., PRL **115** (2015) 122301

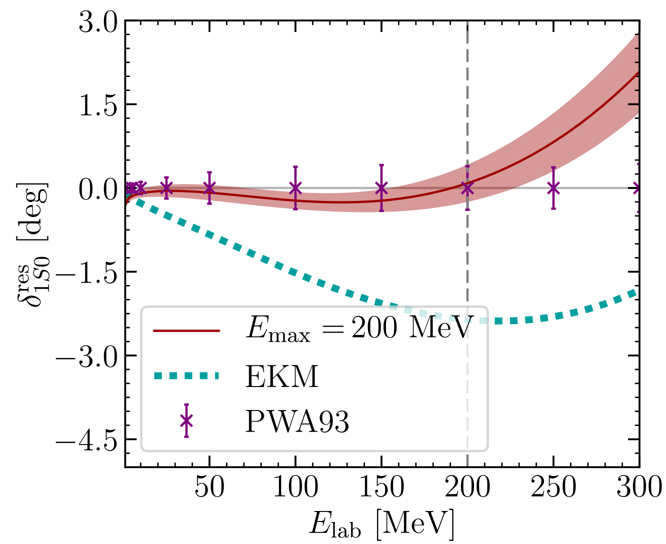
Exploring projected posteriors

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



1S0 channel at N3LO
np sector

- Set D^2_{1S0} to 0 and use only 3 parameters
- Description of data still good
- Recent explorations have confirmed this also

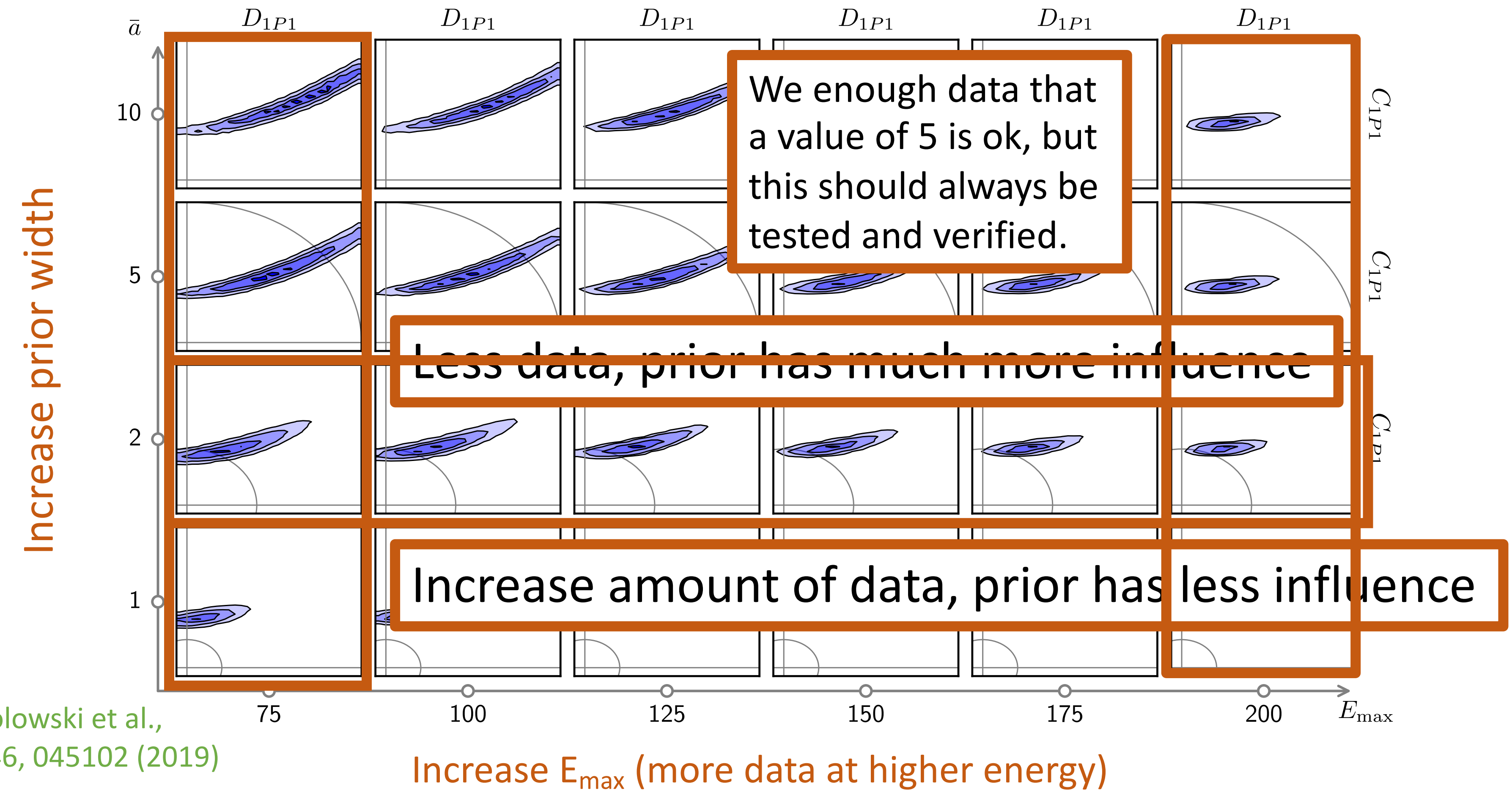


Wesolowski et al.,
JPG 46, 045102 (2019)

Effect of the prior

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

Repeat same problem, vary givens



Wesolowski et al., JPG 46, 045102 (2019)

Including truncation errors

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

Theoretical uncertainty takes this form, and we can put a prior on the unknown higher-order coefficients c_n

$$\delta \boldsymbol{y}_{\rm th} = \boldsymbol{y}_{\rm ref} \sum^{\infty} c_n Q^n \qquad \mathbf{c}_n | \bar{c} \sim \mathcal{N}(0, \bar{c}^2)$$

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

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

$| \, \bar{a})$

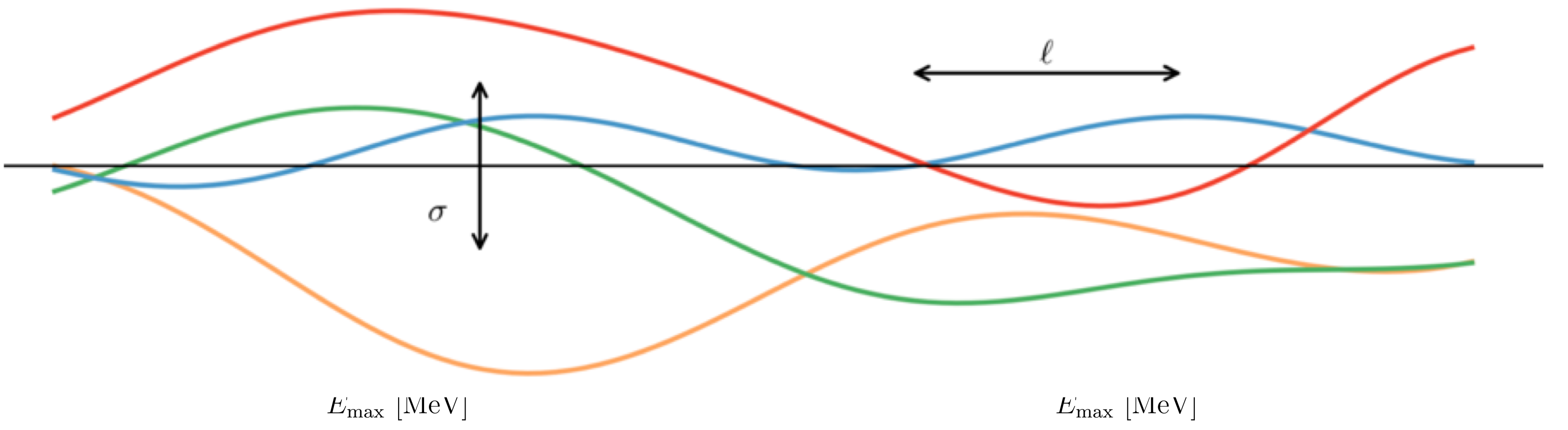
Form of theory covariance is quite simple and is can be added to the experimental covariance!

Wesolowski et al., JPG 46, 045102 (2019)

Effect of including truncation errors

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

Hypothesis: (Daniel's talk)
Coefficients can be modeled as Gaussian processes



Uncorrelated assumption

Correlated assumption

arxiv:1808:08211

Strategy and moving forward for 3N

- Nice, simple form of posterior

  (note: can marginalize over hyperparameters to avoid tight assumptions)
- Sample using Markov Chain Monte Carlo (MCMC)
- Fast for NN, but few-body not as cheap to compute
- Goal: understand impact of uncertainties from 3N
- Strategy: fix NN and πN and their uncertainties first

  Using NNLOsep (Carlsson et al., PRX **6** 011019 (2016))
- Theory uncertainty from fixed LECs and truncation error

$$\Sigma_{\rm th} = \Sigma_k + \Sigma_{\rm LECs}$$

- Then estimate c_D and c_E , simpler 2D problem

Summary

- Important to understand full information content of 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
- 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
- Moving forward… 3NF

Moving forward

- Combined, full error propagation. Put truncation error framework together with LEC error propagation
- c_D and c_E estimation, UQ for 3BFs, fuller analysis and comparison with recent work
- Incorporation of GP model for truncation errors into parameter estimation framework
- Additional open issues:
    - Testing power counting
    - Expansion parameter studies

Backup: Case studies in NN parameter estimation

- Case study 1: Exploring projected posteriors
  - Correlations between parameters
  - Are there multiple modes?
  - Parameter redundancy (occurs for three cases at N3LO)
  - Gaussianity: do covariance approximations work?
- Case study 2: "E_{max} plots"
  - An EFT at order Q^k is missing terms $\sim Q^{k+1}$
  - We account for these missing terms
  - Higher-order effects in data should not affect parameter estimates at lower orders in the EFT
  - E_{max} plots visualize this expectation: should "level off"
- Use "SCS" interactions of Epelbaum, Krebs, and Meißner

  Epelbaum et al., EPJ A 51 (2015) no.5, 53
  Epelbaum et al., PRL **115** (2015) no.12, 122301

Backup: Full posterior pdf with theory error

- Using particular assumptions:

  see Wesolowski et al., JPG (2019) in press [arXiv:1808.08211] for details

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

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

- Theory and experimental errors add!
- Naturalness prior imposes a penalty on large LECs
- Form of theory error:

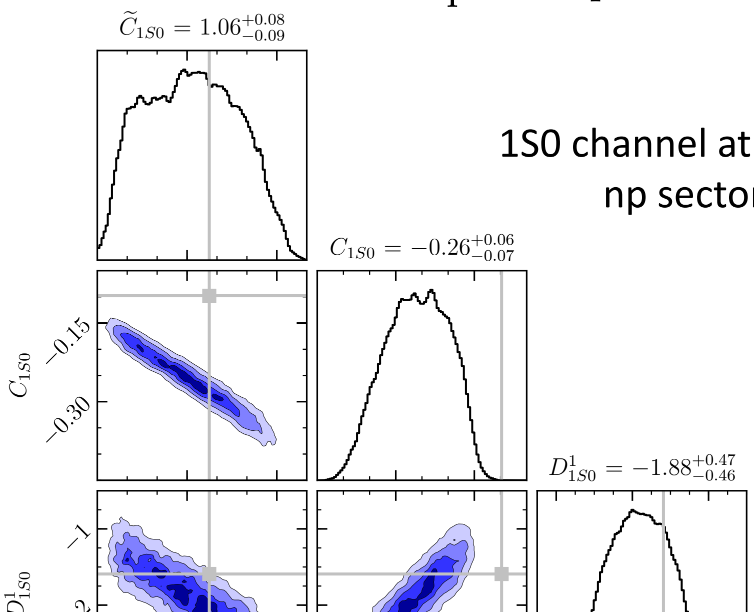
$$(\Sigma_{\rm th,uncorr.})_{ij} = (\boldsymbol{y}_{\rm ref})_i^2\, \bar{c}^2 \sum_{n=k+1}^{k_{\rm max}} Q_i^{2n}\delta_{ij} \xrightarrow[k_{\rm max}\to\infty]{} \frac{(\boldsymbol{y}_{\rm ref})_i^2\, \bar{c}^2\, Q_i^{2k+2}}{1-Q_i^2}\,\delta_{ij}$$

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

Reality: somewhere in between these two assumptions. Use Gaussian processes to model correlations (Daniel's talk next!)

Exploring projected posteriors

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



1S0 channel at N3LO
np sector

- Irregular structure of posterior
- Nothing necessarily wrong

- But here: a physics issues actually explains the structure

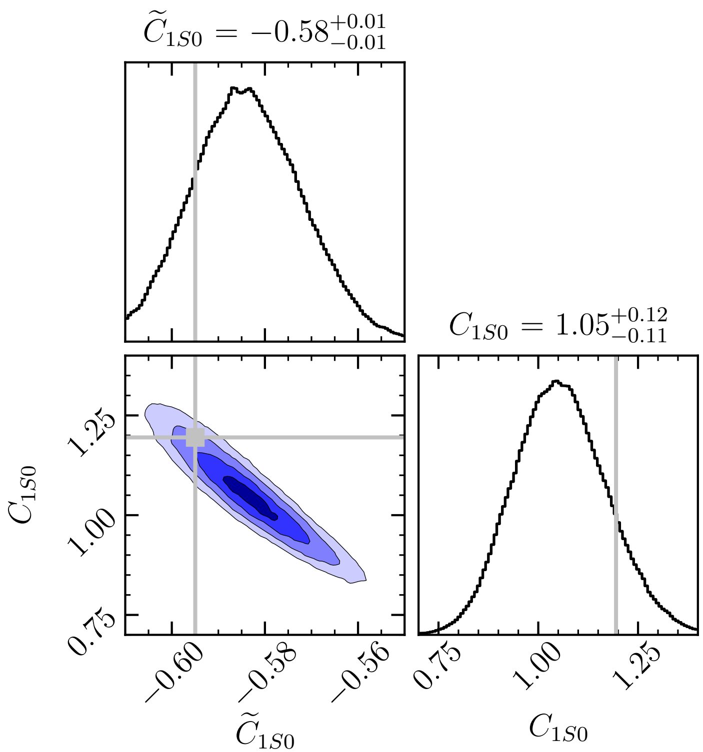
- Parameter redundancy at N3LO, one can be eliminated

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

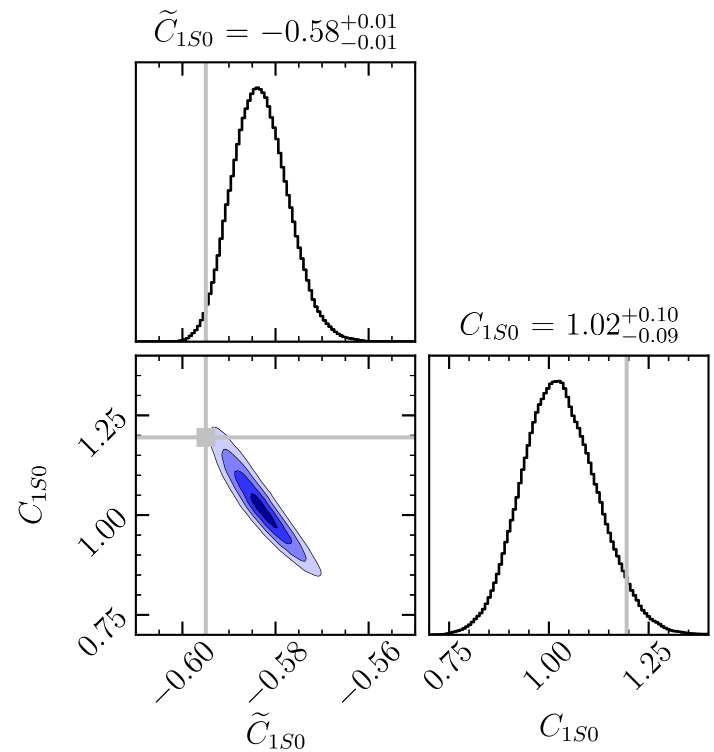
arxiv:1808:08211

Effect of including truncation errors

Use partial-wave cross sections extracted from PWA phase shifts



Uncorrelated assumption



Correlated assumption

arxiv:1808:08211