Application of Bayesian methods in effective field theory (EFT)

## Dick Furnstahl APS/DNP meeting, October, 2019



THE OHIO STATE UNIVERSITY



## Special thanks to my BUQEYE collaborators!







# Outline

- Overview of Bayesian methods for EFT
- Truncation errors and model checking
- Aspects of EFT parameter estimation
- Snapshots of Bayesian EFT applications
- Take-away points and advertisements

## **Quantum chromodynamics has many EFTs!**



## [From lain Stewart]

# Ingredients of EFTs that invite Bayesian statistics



- Choice of degrees of freedom and organization ("power counting")
- Well-defined order-by-order construction used to predict observables ⇒ *truncation error*
- Short distance physics is encoded in low-energy constants (LECs) that need to be *fitted*
- *Expectation* is that observables improve as powers of a small parameter (e.g., ratios of scales) that may not be well determined
- Need to *validate* the EFT and provide robust theory error bars

## Ingredients of EFTs that invite Bayesian statistics

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

 $\begin{aligned} \mathsf{pr}(\mathsf{EFT}_1 \mid \boldsymbol{y}_{\mathsf{exp}}, I) / \mathsf{pr}(\mathsf{EFT}_2 \mid \boldsymbol{y}_{\mathsf{exp}}, I) \\ \implies \text{ evidence ratio given data} \end{aligned}$ 

 $\mathsf{pr}(c_n(\boldsymbol{x}) \mid \boldsymbol{\vec{c_k}}, I) \implies \mathsf{pdf} \text{ of expansion coefficients}$ 

 $\begin{aligned} \mathsf{pr}(\boldsymbol{\theta} \mid \boldsymbol{y}_{\mathsf{exp}}, \boldsymbol{\Sigma}_{\mathsf{exp}}, I) &\propto \mathsf{pr}(\boldsymbol{y}_{\mathsf{exp}} \mid \boldsymbol{\theta}, \boldsymbol{\Sigma}_{\mathsf{exp}}, I) \, \mathsf{pr}(\boldsymbol{\theta} \mid I) \\ \implies \mathsf{pdf} \text{ of LECs given data } \boldsymbol{y}_{\mathsf{exp}}, \boldsymbol{\Sigma}_{\mathsf{exp}} \; (\mathsf{and} \; \boldsymbol{\Sigma}_{\mathsf{th}}!) \end{aligned}$ 

 $\mathsf{pr}(\Lambda_b \mid \boldsymbol{\theta}, I) \implies \mathsf{pdf} \text{ of EFT breakdown scale}$ 

 $\operatorname{pr}(\mu, \sigma \mid I) \implies \operatorname{pdf}$  of hyperparameters

Use rules of probability to manipulate Marginalize over nuisance parameters



provide robust theory error bars



NEW ENGLAND SECTION MEETING MARCH 16-17 BOSTON, MA

15 16

29 30

ZARE FALL AND

18

22 23 24 25 26 2.

Nonequilibrium interband phase textures induced by vortex splittin superconductors (A. S. Mosquera Polo et al., Phys. Rev. B 96, 0545

19 20

25

26

27

28

Bayesian truncation errors in chiral effective field theory: nucleon-nucleon observables [J. A. Melendez et *al.*, Phys. Rev. C **96**, 024003 (2017)].

29

30

31

OHIO-REGION SECTION MEETING MARCH 23-24 EAST LANSING, MI

APS APRIL MEETING APRIL 14-17 COLUMBUS, OH

DIVISION OF L	ASER SCIENCE	MEETING	(CLEC

## **Example: chiral EFT expansion for NN observables**

How do we apply Bayesian uncertainty quantification and model checking?



## **Example: chiral EFT expansion for NN observables**

How do we apply Bayesian uncertainty quantification and model checking?



## Generic example of EFT expansion (based on NN analysis)

- $V_{NN}(\vec{a}) = V_{LO} + V_{NLO} + \dots + V_{N^k LO} \Longrightarrow$  | Schrödinger Eq. |  $\Longrightarrow y_k(x; \vec{a})$
- One can change variables for convenience/insight.
- $\Delta y_n = y_{\text{ref}} c_n Q^n$

$$y_3 = y_{ref} \left[ \frac{C_0 Q^0 + C_1 Q^1 + C_2 Q^2 + C_3 Q^3}{C_3 Q^3} \right]$$



## **Generic example of EFT expansion (based on NN analysis)**

Coefficients from NN scattering look like the example on the previous slide!



## **Model discrepancy: EFT theory errors**

• Statistical model relating experiment and theory (all are functions of x):

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

- Experimental discrepancy: usually normally distributed, can be correlated
- Theoretical discrepancy: systematic, often difficult to quantify, often correlated
- EFTs: specific expectations for theory error
- Theory prediction at a certain order:  $[Q < 1: expansion parameter (e.g., p/\Lambda_b)]$

$$oldsymbol{y}_k = oldsymbol{y}_{ ext{ref}} \sum_{n=0}^k c_n Q^n \quad + oldsymbol{y}_{ ext{ref}} \sum_{n=k+1}^\infty c_n Q^n$$

Theory prediction at order k Omitted terms from truncation

Bayesian: treat the  $c_n$ s as random variables, depending on x; train on  $c_n$  for  $n \le k$ 

## **Gaussian process model of the coefficients**



## Model checking: Does our model refer to reality?

### Use metric to measure GP-ness to test model: Mahalanobis distance



This is what success looks like!

## Model checking: Does our model refer to reality?

### Use metric to measure GP-ness to test assumption: Mahalanobis distance



This is what failure looks like!

## Model checking: Does our model refer to reality?

EKM semi-local NN potential (R = 0.9 fm). Results for total *np* cross section.



Ok for low orders but a problem at the highest order.

## Model checking: Do our error bands have statistical meaning?

Credible interval diagnostic



## What is the breakdown scale of our EFT?

• The breakdown scale of the EFT  $\Lambda_b$ :

 $Q \approx \frac{f(p, m_{\pi})}{\Lambda_b}$ 

- Once energies approach  $\Lambda_b$ , the EFT no longer works.
- By promoting  $\Lambda_b$  to a random variable, its posterior can be produced as a byproduct of  $\delta y_{\text{th}}$  estimation.



Melendez et al., in preparation (2019)

# Symmetry energy in infinite matter (preliminary!)

Christian Drischler, Jordan Melendez, et al., in preparation

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



Many interesting questions to investigate and comparisons to make!

## **Parameter estimation: Exploring projected posteriors**

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

$$oldsymbol{y}_{ ext{exp}} = oldsymbol{y}_{ ext{th}} + \delta oldsymbol{y}_{ ext{exp}}$$

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

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

## **Projected posteriors as a diagnostic tool**

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



- 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!)



## **Projected posteriors as a diagnostic tool**

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



- 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

## **Projected posteriors as a diagnostic tool**

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

1S0 channel at N3LO



 $\widetilde{C}_{1S0} = 0.55^{+0.00}_{-0.00}$ 

- Set  $D_{1S0}^2$  to 0 and use only 3 parameters
- Description of data still good (and potential softer) ٠
- Reinert et al., EPJA (2018) improved potential ٠



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





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

Repeat same problem, vary givens



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

Repeat same problem, vary givens



## Parameter estimation including truncation errors

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

We can put a prior on the unknown higher-order coefficients:  ${f c}_n | ar c \sim {\cal N}(0,ar c^2)$ 

Just a sum of Gaussian random variables using this prior (which we have validated)

## Two extreme assumptions about correlations

1. Theory error at each kinematic point (Q) is completely uncorrelated:

$$[\delta y_{\rm th}]_i = [y_{\rm ref}]_i \sum_{n=k+1}^{\infty} (c_n]_i Q_i^n$$

2. Theory error at each kinematic point (Q) is fully correlated, with same coefficients:

$$[\delta y_{\rm th}]_i = [y_{\rm ref}]_i \sum_{n=k+1} c_n Q_i^n$$

## Parameter estimation including truncation errors

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

Full posterior pdf including theory error:

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

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:

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

# **Effect of including truncation errors**

Use partial-wave cross sections extracted from PWA phase shifts



# **Effect of including truncation errors in parameter estimation**

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:





## Correlated assumption

# **Effect of including truncation errors in parameter estimation**

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:



## But won't the Bayesian sampling calculations take too long?

#### Eigenvector Continuation as an Efficient and Accurate Emulator for Uncertainty Quantification

arXiv:1909.08446

S. König,<sup>1,2,\*</sup> A. Ekström,<sup>3,†</sup> K. Hebeler,<sup>1,2,‡</sup> D. Lee,<sup>4,§</sup> and A. Schwenk<sup>1,2,5,¶</sup>



Figure 1. Comparison of different emulators for the  ${}^{4}$ He ground-state energy using 12 training data points to explore a space where three LECs are varied. The left panel includes samples for both interpolation (solid symbols) and extrapolation (semi-transparent symbols). See main text on how these are defined. The right panel shows the same data restricted to interpolation samples (note the smaller axis range).

## But won't all of these Bayesian sampling calculations take too long?

Global sensitivity analysis of bulk properties of an atomic nucleus



arXiv: 1910.02922

Andreas Ekström<sup>1</sup> and Gaute Hagen<sup>2,3</sup>

"We have to use  $(16 + 1) \cdot 216 = 1,114,112$  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."

## **Snapshots of Bayesian methods for EFT**



80 83 30

T-MAY

1.00

1.00

1.00

1.14

1.00

STREET IN

Analysis: measuring cross-section

cross section's error bar

anisotropy can reduce low-energy



#### Bayesian model selection: Two competing EFT power countings

Pradeepa Premarathna and Gautam Rupak; arXiv:1906.04143



# Form factors of deuteron with *consistent* chiral EFT current operators

Hermann Krebs and Evgeny Epelbaum (from HK talk at INT, July 2019)



## http://www.lenpic.org/

- A major advantage of EFT: consistent current operators
- Here: new developments by HK and EE for chiral EFT
- Colored band is 68% Bayes credible interval for truncation error from convergence pattern



Truncation error band from Bayesian analysis: 68% DoB,  $\Lambda_b = 600 \, {\rm MeV}$  Furnstahl et al. 15

# EFT for nuclear vibrations with quantified uncertainties

E. A. Coello Pérez and Thomas Papenbrock, PRC (2015, 2016)

proton separation energy in lead 1.12 vibrational state in tin See also Coello Pérez, Menéndez, Schwenk, PRC (2018); Chen, Kaiser, Meissner, Meng, EPJA (2017), PRC (2018)

Collective and algebraic models have long been used to describe data in heavy nuclei. Some models "fail" to describe data. But: How does one really know in the absence of theoretical uncertainties?

Here: EFT truncation errors from convergence pattern.



"The systematic improvement inherent to EFT approaches is evident."

# EFT + Bayes for nuclear reactions: ${}^{3}\text{He}(\alpha, \gamma){}^{7}\text{Be}$

Xilin Zhang, Ken Nollett, Daniel Phillips, arXiv:1909.07287 See also PRC (2015) on <sup>7</sup>Be(p, γ)<sup>8</sup>B

6

 $\sim$ 

<sup>0.0</sup> (keV1

0.2

0.0

 $10^{-1}$ 

 $10^{-2}$ 

 $10^{-3}$ 

 $10^{-4}$ 

0.0

0.5

 $\Delta S/S|$ 

- "Supermodel": project existing models into EFT space->marginalize EFT params
- Some LECs for potential models are not preferred by data (analysis enabled by EFT); data prefer natural values in agreement with ab initio calculations
- Analysis: first realized that capture data constrains scattering parameters
- Analysis: measuring cross-section anisotropy can reduce low-energy cross section's error bar



## **Bayesian model selection: Two competing EFT power countings**

Pradeepa Premarathna and Gautam Rupak; arXiv:1906.04143

posterior PDF: 
$$P(\theta|D, I) = \frac{P(D|\theta, I)P(\theta|I)}{P(D|I)}$$
  
prior PDF:  $P(\theta|I)$  EFT expectations/bias  
likelihood PDF:  $P(D|\theta, I) \propto \exp(-\frac{1}{2}\chi^2), \quad \chi^2 = \sum_{i=1}^{N} \frac{[D_i - \mu_i(\theta)]^2}{\sigma_i^2}$   
evidence:  $P(D|I) = \int [d\theta]P(D|\theta, I)P(\theta|I)$   
— harder to calculate  
— not needed for parameter estimation  
— needed to compare models  
Models:  $\frac{P(M_1|D, I)}{P(M_2|D, I)} = \frac{P(M_1|I)}{P(M_2|I)} \times \frac{P(D|M_1, I)}{P(D|M_2, I)}$ 

Nested sampling:  $\ln \frac{P(M_A \mid D, H)}{P(M_B \mid D, H)} \approx 31.6 \pm 0.9$ 

## Application to ${}^{3}\text{He}(\alpha, \gamma){}^{7}\text{Be}$ in Halo EFT



 $\ln \frac{P(M_A \mid D^*, H)}{P(M_B \mid D^*, H)} \approx 0.2 \pm 0.8$ 

# **Take-away Points**

- Bayesian statistics is an ideal framework for EFT uncertainty quantification.
- Theory discrepancy model based on convergence pattern and naturalness priors; GP truncation error model accounts for correlations;  $\Lambda_b$  estimates.
- Model checking is an essential part of Bayesian UQ.
- Parameter estimation with LEC priors and model discrepancy minimizes dependence on how much data is used; posterior diagnostics point to physics issues.
- Need to sample for parameter estimation and the propagation of uncertainties: Eigenvector continuation is very promising solution to computational load!
- Model selection is a frontier for EFT Bayesian statistics.







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



## gsum: A Bayesian model of series convergence using Gaussian sums.

The gsum package provides convenient classes that allow one to analyze the convergence pattern of Effective Field Theory (EFT) observables. Specifically, this is a conjugacy-based implementation of the statistical model described in <u>arXiv:1904.10581</u>.

Correlated truncation errors in effective field theory: The code behind the manuscript. This Jupyter notebook provides the Python code to reproduce *all* of the plots from Melendez et al., Phys. Rev. C **100**, 044001 (2019), <u>arXiv:1904.10581</u>.

## **Journal of Physics G Special Focus Issue**

Focus on further enhancing the interaction between nuclear experiment and theory through information and statistics (ISNET 2.0)

#### **Guest Editors**

Dick Furnstahl Ohio State University David Ireland University of Glasgow Daniel Philips Ohio University

"Following on from the hugely successful first edition in 2015, Journal of Physics G: Nuclear and Particle Physics is delighted to announce a second focus issue inspired by the ISNET workshops (Information and Statistics in Nuclear Experiment and Theory."

> Two articles already published! Please consider contributing!

https://iopscience.iop.org/journal/0954-3899/page/ISNET2

## **TALENT:** Training in Advanced Low-Energy Nuclear Theory

The TALENT initiative aims at providing advanced training to graduate students and young researchers in low-energy nuclear theory. TALENT offers intensive three-week courses on a rotating set of topics.

## Bayes2019. Learning from Data: Bayesian Methods and Machine Learning,

at the University of York, UK, June 10-28, 2019. See the github site for many Jupyter notebooks and lecture notes: <u>https://nucleartalent.github.io/Bayes2019/</u>

## TALENT courses scheduled for 2020 (see <u>http://nucleartalent.org</u> for more info):

- Atomic Nuclei as Open Quantum Systems: Unifying Nuclear Structure and Reactions will be held at the INT in Seattle, WA from June 22 to July 10, 2020.
   Contact: Christian Forssén (christian.forssen@chalmers.se).
- **Density Functional Theory and Self-Consistent Methods** will be held at LBNL, in Berkeley, CA. The (tentative) dates are July 6 to July 24, 2020. **Contact:** Nicolas Schunck (schunck1@llnl.gov).
- Machine Learning and Data Analysis for Nuclear Physics will be held at the ECT\* in Trento, Italy. The (tentative) dates are June 29 to July 17, 2020. Contact: Morten Hjorth-Jensen (hjensen@msu.edu).