Combining Physics and Bayesian Statistics to Validate Models and Infer Their Parameters

Jordan Melendez[1] Dick Furnstahl[1] Daniel Phillips[2] Matt Pratola[1] Sarah Wesolowski[3]

August 17, 2019

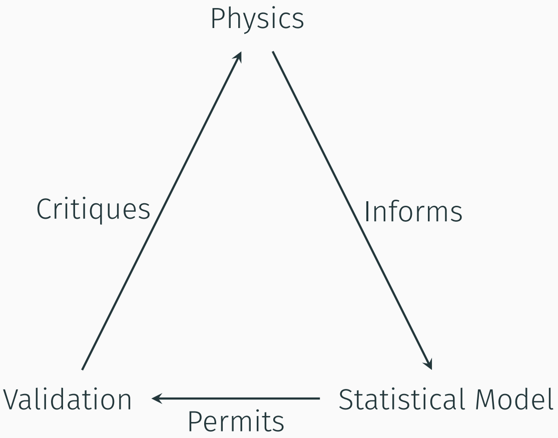
[1]The Ohio State University

[2]Ohio University

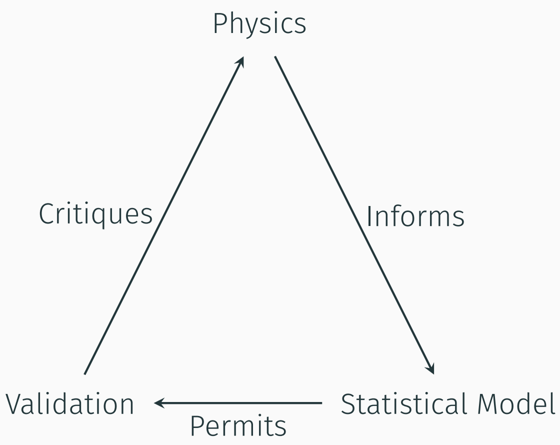
[3]Salisbury University

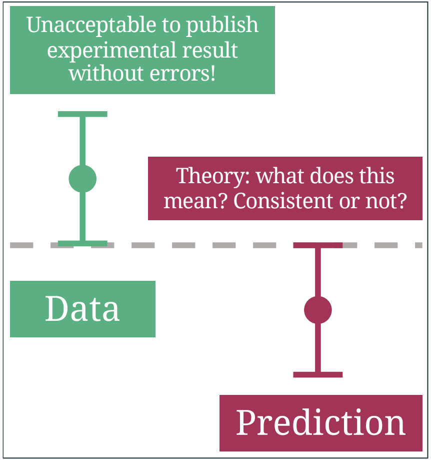
Based on:
- S. Wesolowski, R. J. Furnstahl, J. A. Melendez, D. Phillips, arXiv:1808.08211; &
- J. A. Melendez, S. Wesolowski, and R. J. Furnstahl Phys. Rev. C **96**, 024003, Editors' Suggestion

Outline


Physics
Critiques
Informs
Validation
Permits
Statistical Model

Outline


Physics
Critiques
Informs
Validation
Permits
Statistical Model


Unacceptable to publish experimental result without errors!
Theory: what does this mean? Consistent or not?
Data
Prediction

Why Bayesian?

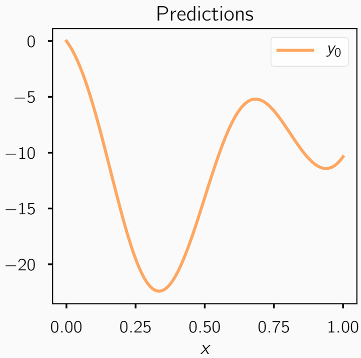
- More intuitive
  - Confidence intervals $\rightarrow$ credible intervals
  - p-values $\rightarrow$ posterior probabilities
- Conventional results often recovered as a special case
- Assumptions are made explicit
- Well suited for incorporating theoretical error

Physical Motivation

Chiral EFT in One Slide

- $V_{NN}(\vec{a}) = V_{LO} + V_{NLO} + \cdots + V_{N^kLO} \Longrightarrow \boxed{\text{Schrödinger Eq.}} \Longrightarrow y_k(x; \vec{a})$

$$\{y_0\}$$

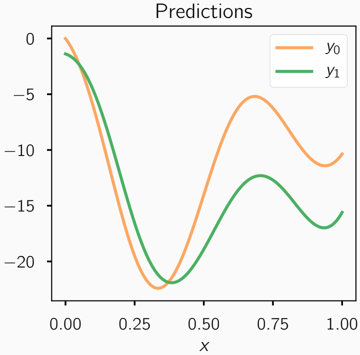


$$y_0 \to \text{LO}$$

Chiral EFT in One Slide

- $V_{NN}(\vec{a}) = V_{LO} + V_{NLO} + \cdots + V_{N^kLO} \Longrightarrow$ Schrödinger Eq. $\Longrightarrow y_k(x; \vec{a})$

$$\{y_0, y_1\}$$



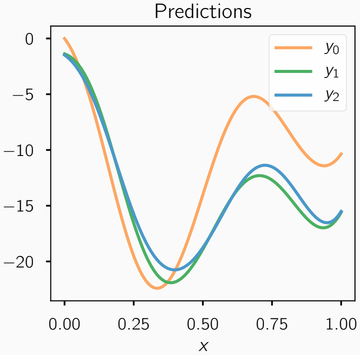
$y_0 \to$ LO

$y_1 \to$ NLO

Chiral EFT in One Slide

- $V_{NN}(\vec{a}) = V_{LO} + V_{NLO} + \cdots + V_{N^kLO} \Longrightarrow \boxed{\text{Schrödinger Eq.}} \Longrightarrow y_k(x; \vec{a})$

$$\{y_0, y_1, y_2\}$$



$y_0 \rightarrow$ LO

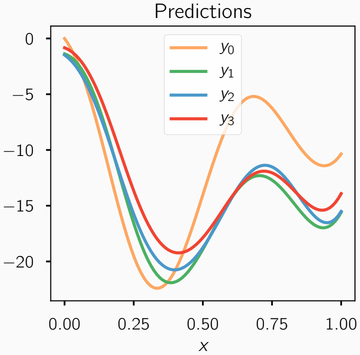
$y_1 \rightarrow$ NLO

$y_2 \rightarrow$ N^{2}LO

Chiral EFT in One Slide

- $V_{NN}(\vec{a}) = V_{LO} + V_{NLO} + \cdots + V_{N^kLO} \Longrightarrow$ Schrödinger Eq. $\Longrightarrow y_k(x; \vec{a})$

$$\{y_0, y_1, y_2, y_3\}$$



$y_0 \rightarrow$ LO

$y_1 \rightarrow$ NLO

$y_2 \rightarrow$ N^{2}LO

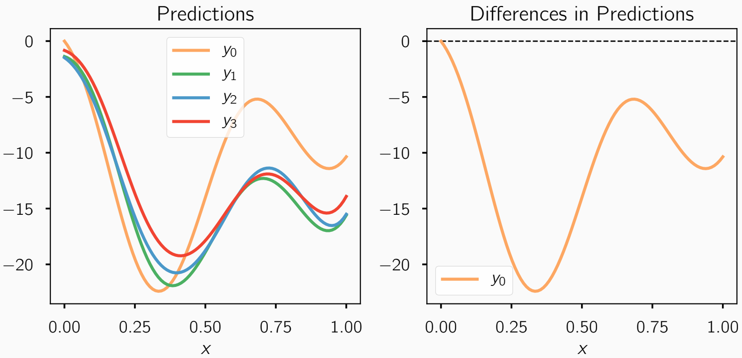
$\vdots$

$y_k \rightarrow$ N^{k}LO

Chiral EFT in One Slide

- $V_{NN}(\vec{a}) = V_{LO} + V_{NLO} + \cdots + V_{N^kLO} \Longrightarrow$ Schrödinger Eq. $\Longrightarrow y_k(x; \vec{a})$
- One can change variables for convenience/insight.

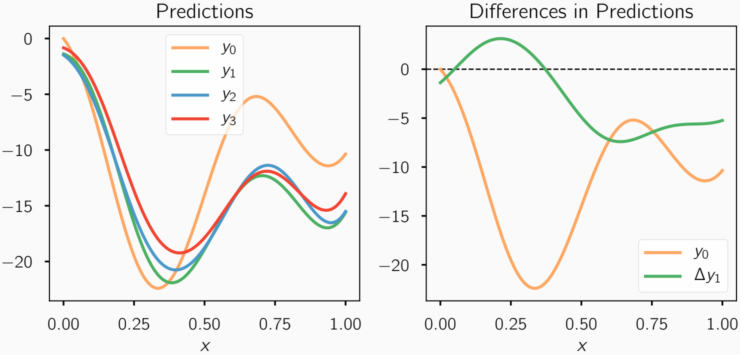
$$y_0 = y_0$$



Chiral EFT in One Slide

- $V_{NN}(\vec{a}) = V_{LO} + V_{NLO} + \cdots + V_{N^kLO} \Longrightarrow$ Schrödinger Eq. $\Longrightarrow y_k(x; \vec{a})$
- One can change variables for convenience/insight.

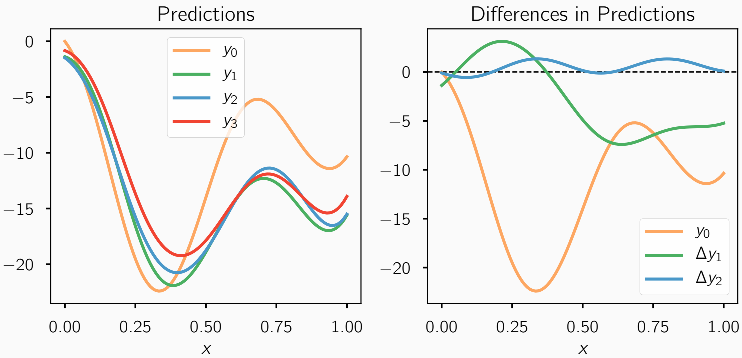
$$y_1 = y_0 + \Delta y_1$$



Chiral EFT in One Slide

- $V_{NN}(\vec{a}) = V_{LO} + V_{NLO} + \cdots + V_{N^kLO} \Longrightarrow$ Schrödinger Eq. $\Longrightarrow y_k(x; \vec{a})$
- One can change variables for convenience/insight.

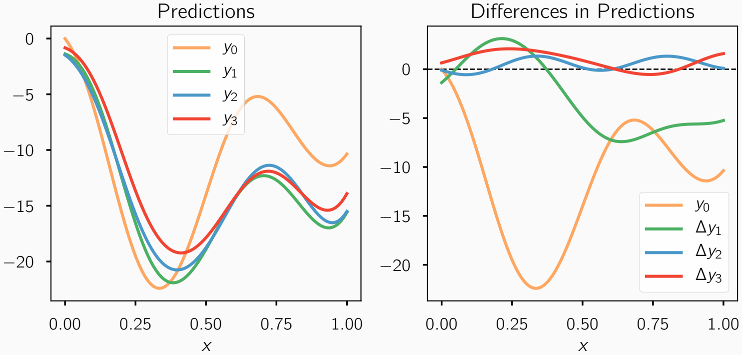
$$y_2 = y_0 + \Delta y_1 + \Delta y_2$$



Chiral EFT in One Slide

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

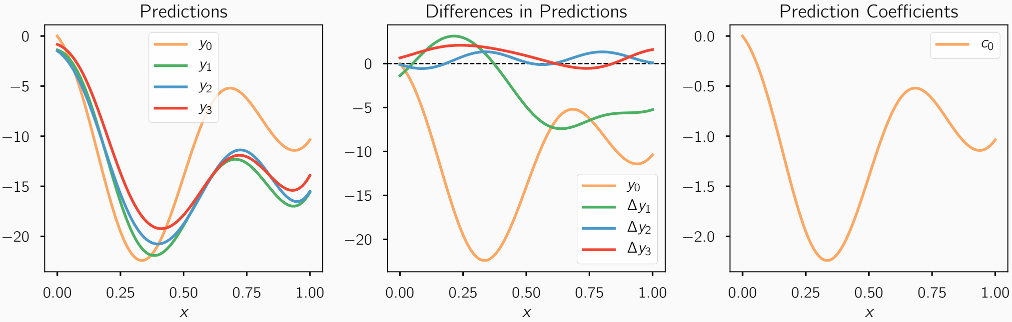
$$y_3 = y_0 + \Delta y_1 + \Delta y_2 + \Delta y_3$$



Chiral EFT in One Slide

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

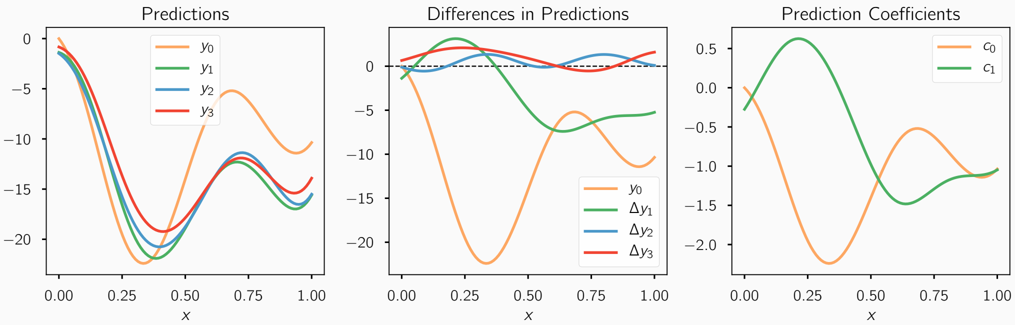
$$y_0 = y_{\text{ref}} \left[c_0 Q^0 \right]$$



Chiral EFT in One Slide

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

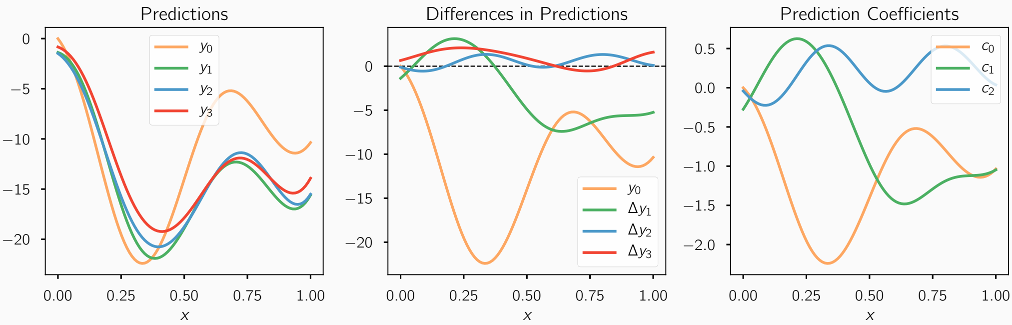
$$y_1 = y_{\text{ref}} \left[c_0 Q^0 + c_1 Q^1 \right]$$



Chiral EFT in One Slide

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

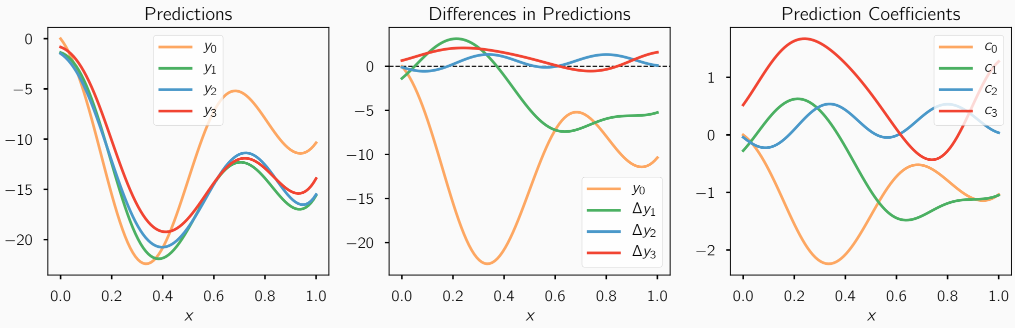
$$y_2 = y_{\text{ref}} \left[c_0 Q^0 + c_1 Q^1 + c_2 Q^2 \right]$$



Chiral EFT in One Slide

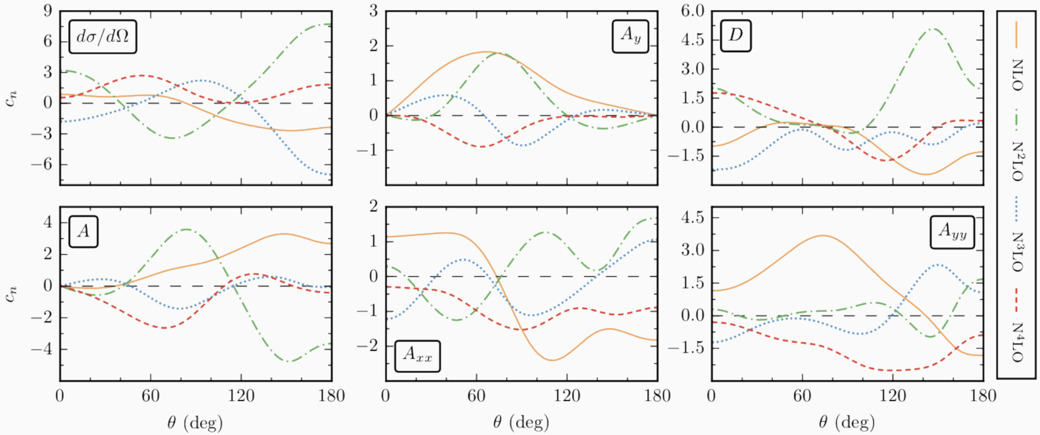
- $V_{NN}(\vec{a}) = V_{LO} + V_{NLO} + \cdots + V_{N^kLO} \Longrightarrow \boxed{\text{Schrödinger Eq.}} \Longrightarrow y_k(x;\vec{a})$
- One can change variables for convenience/insight.
- $\Delta y_n = y_{\text{ref}} c_n Q^n$

$$y_3 = y_{\text{ref}}\left[c_0 Q^0 + c_1 Q^1 + c_2 Q^2 + c_3 Q^3\right]$$



Real Life (Neutron-Proton Scattering)

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



Statistical Model

$$y_{\mathrm{exp}}(x) = y_{\mathrm{th}}(x, \vec{a}) + \delta y_{\mathrm{th}}(x) + \delta y_{\mathrm{exp}}$$

To theorists, magic

$$y_{\text{exp}}(x) = y_{\text{th}}(x, \vec{a}) + \delta y_{\text{th}}(x) + \delta y_{\text{exp}}$$

Parameters

Discrepancy

χ^2 fit

$$y_{\text{exp}}(x) = y_{\text{th}}(x, \vec{a}) \qquad + \delta y_{\text{exp}}$$

rigorous fit

$$y_{\text{exp}}(x) = y_{\text{th}}(x, \vec{a}) + \delta y_{\text{th}}(x) + \delta y_{\text{exp}}$$

$$\overbrace{y_{\text{th}}(x, \vec{a}) + \delta y_{\text{th}}(x)}^{\text{Full Prediction}}$$

$$y_{\text{exp}}(x) = y_{\text{th}}(x, \vec{a}) + \delta y_{\text{th}}(x) + \delta y_{\text{exp}}$$

$$y_{\text{exp}}(x) = y_{\text{th}}(x, \vec{a}) + \underbrace{\delta y_{\text{th}}(x)}_{} + \delta y_{\text{exp}}$$

How do we design this?

How does it affect fitting $\vec{a}$?

The Hierarchical Model

- Decompose prediction

$$y_k = y_0 + \sum_{n=1}^{k} \Delta y_n \quad \Bigg\}$$

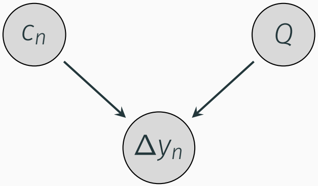
Data

The Hierarchical Model

- Decompose prediction

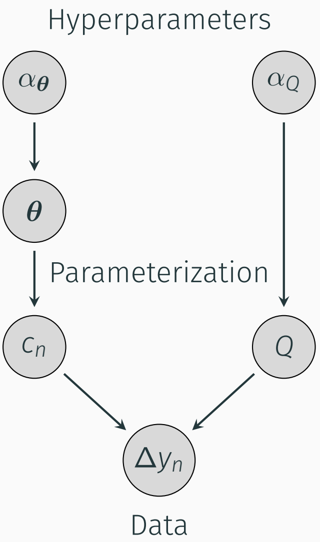
$$\left.\begin{aligned} y_k &= y_0 + \sum_{n=1}^{k} \Delta y_n \\ &= y_{\text{ref}} \sum_{n=0}^{k} c_n Q^n \end{aligned}\right\}$$

Parameterization



Data

The Hierarchical Model



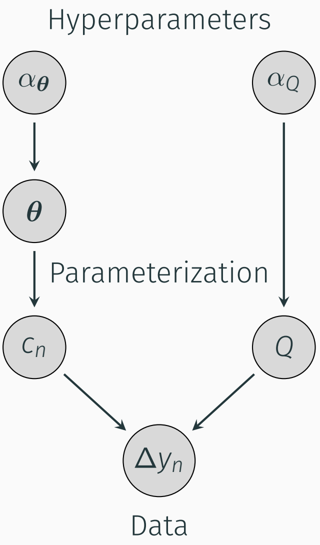
- Decompose prediction

$$\left.\begin{aligned} y_k &= y_0 + \sum_{n=1}^{k} \Delta y_n \\ &= y_{\text{ref}} \sum_{n=0}^{k} c_n Q^n \end{aligned}\right\} \implies \delta y_k = y_{\text{ref}} \sum_{n=k+1}^{\infty} c_n Q^n$$

- Promote c_n (and Q) to random variables

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

The Hierarchical Model



- Decompose prediction

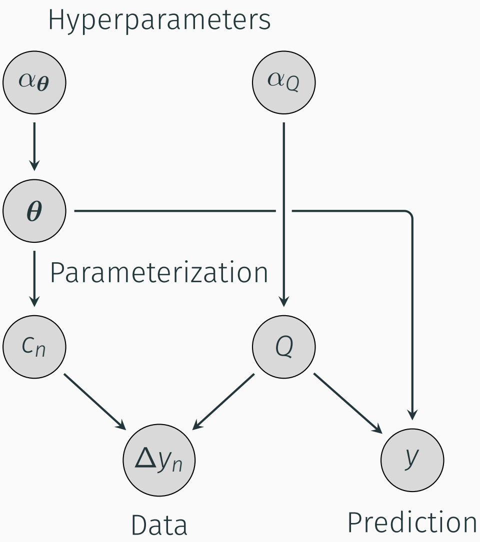
$$\left.\begin{aligned} y_k &= y_0 + \sum_{n=1}^{k} \Delta y_n \\ &= y_{\text{ref}} \sum_{n=0}^{k} c_n Q^n \end{aligned}\right\} \implies \delta y_k = y_{\text{ref}} \sum_{n=k+1}^{\infty} c_n Q^n$$

- Promote c_n (and Q) to random variables

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

- Train statistical parameters ($\boldsymbol{\theta}$ and Q) and LECs $\vec{a}$ on data

The Hierarchical Model



- Decompose prediction

$$\left.\begin{aligned} y_k &= y_0 + \sum_{n=1}^{k} \Delta y_n \\ &= y_{\text{ref}} \sum_{n=0}^{k} c_n Q^n \end{aligned}\right\} \implies \delta y_k = y_{\text{ref}} \sum_{n=k+1}^{\infty} c_n Q^n$$

- Promote c_n (and Q) to random variables

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

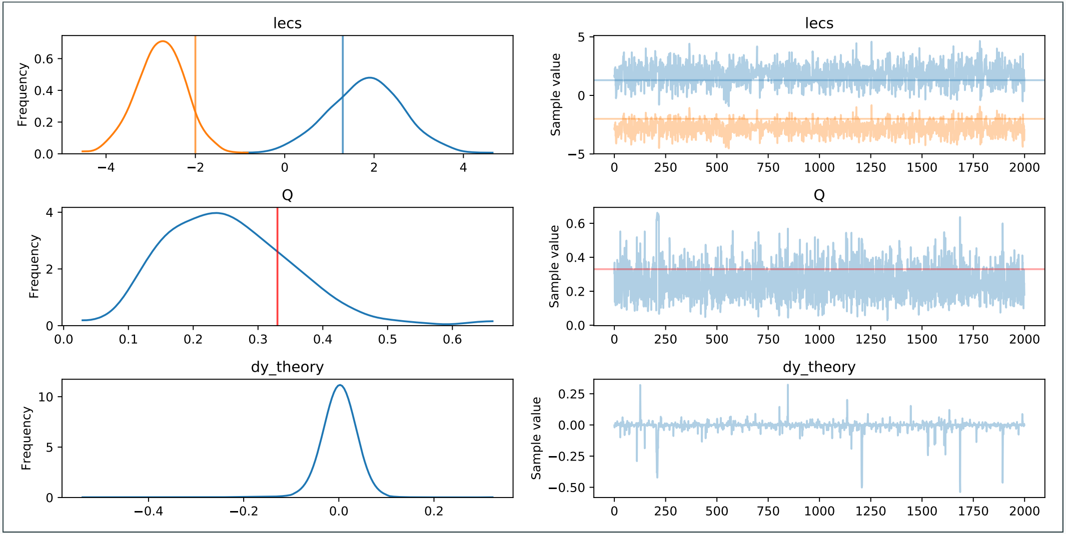
- Train statistical parameters ($\boldsymbol{\theta}$ and Q) and LECs $\vec{a}$ on data
- Compare $y_k + \delta y_k$ to experiment y_{exp}

```python
# Setup code goes up here ...
with Model() as model: # Build statistical model with PyMC3
    Q = Beta('Q', alpha=5, beta=15) # Priors
    sigma = Lognormal('sigma', mu=0, sd=1)
    c_n = Normal('c_n', mu=0, sd=sigma, shape=(n_high_ords, 1))
    dy_theory = Deterministic('dy_theory',
        tt.sum(c_n * Q**higher_orders, axis=0)) # ∑ c_n Q^n

    # Setup chiral EFT variables: y_k(x; a⃗) + δy_k(x) and prior pr(a⃗).
    lecs = MvNormal('lecs', mu=0, cov=np.identity(2), shape=2)
    y_theory = compute_observable(X, lecs) + dy_theory

    # Fit chiEFT parameters and statistical parameters
    MvNormal('y_exp', mu=y_theory, cov=exp_cov, observed=exp_data)
    MvNormal('c_n_obs', mu=0, cov=sigma**2 * R, observed=coeffs)
    trace = sample(draws=500)
```

Toy Results



Physics Results

So far we've

- added a model discrepancy term δy_{th}
- adopted a Bayesian approach

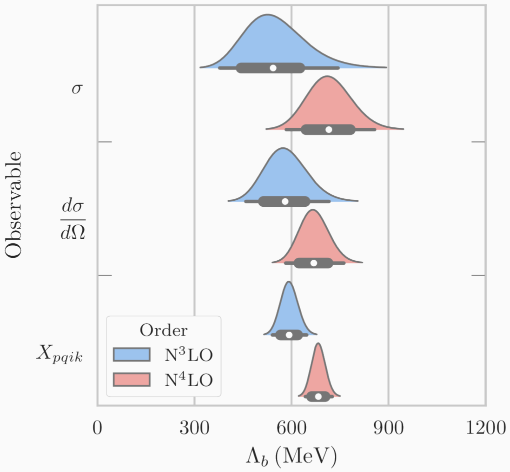
So what?

Rigorous Estimation of Unknown Quantities

- The breakdown scale of the EFT Λ_b:

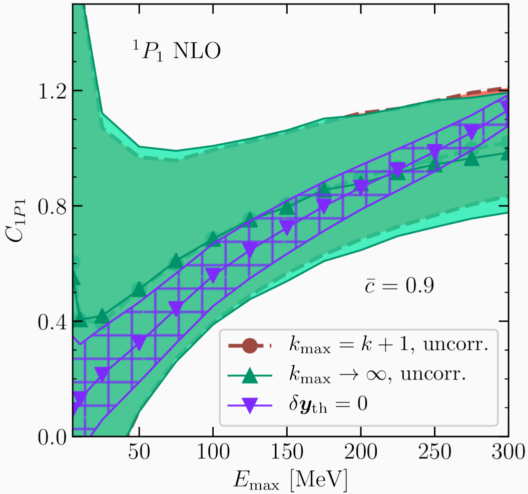
$$Q \approx \frac{f(p, m_\pi)}{\Lambda_b}$$

- Once energies approach Λ_b, the EFT no longer works.
- By promoting Λ_b to a random variable, its posterior can be produced as a byproduct of δy_{th} estimation.



Adding Theory Error Reduces Bias

- y axis: posterior median $\pm 1\sigma$
- x axis: max energy of data in fit

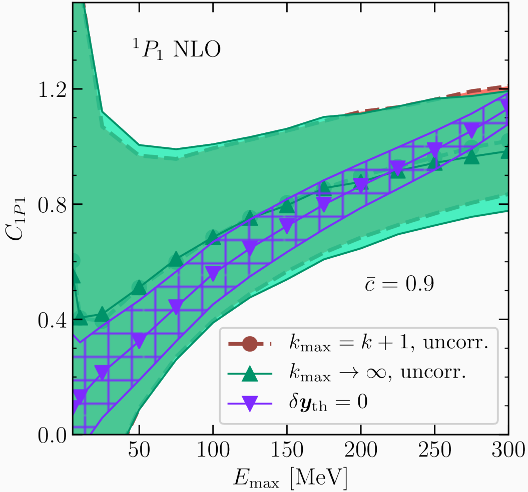


Adding Theory Error Reduces Bias

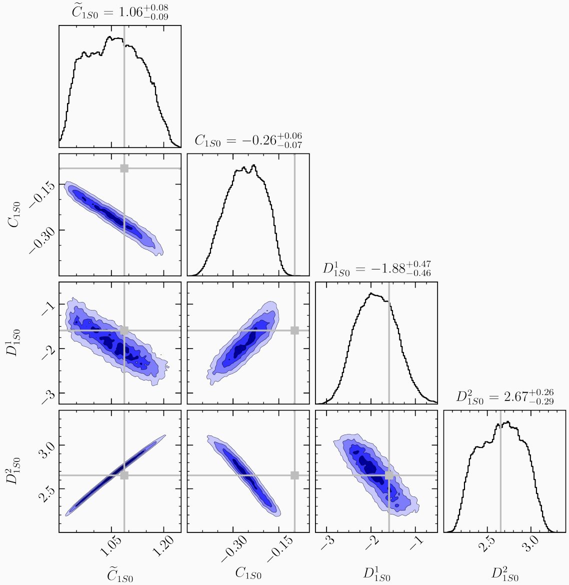
- y axis: posterior median $\pm 1\sigma$
- x axis: max energy of data in fit
- Q, and hence δy_{th}, grows with energy

$$\delta y_{\text{th}} = y_{\text{ref}} \sum_{n=k+1}^{k_{\max}} c_n Q^n$$

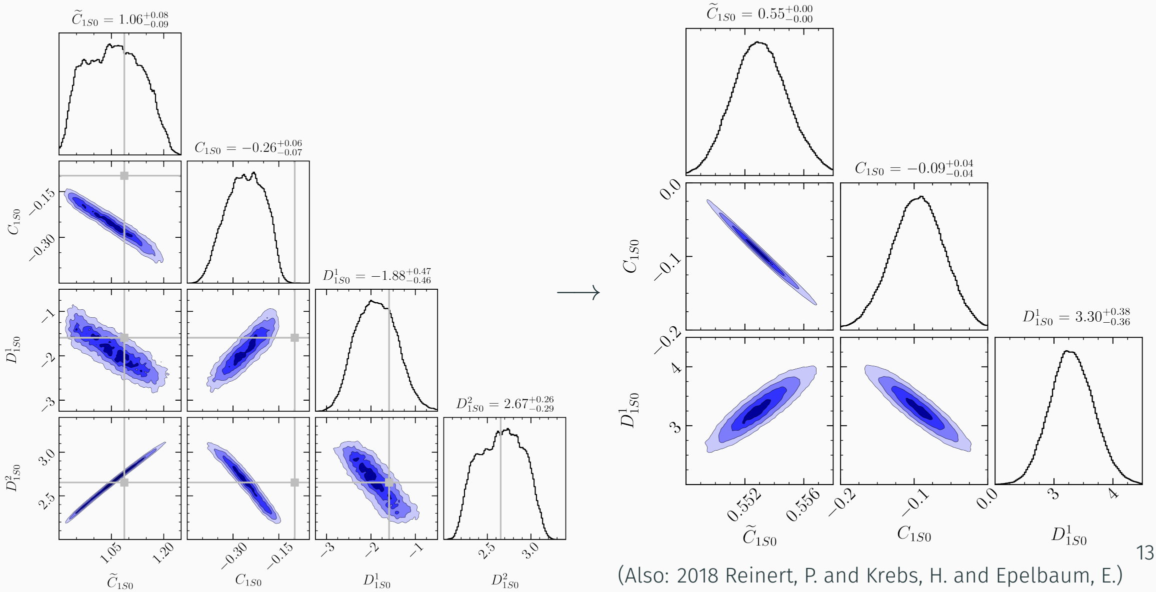
- This weights high energy data less!
- Stabilizes LEC fit as a function of E
- LEC uncertainty is more realistic



Discovering Physics (Redundant LECs)



Discovering Physics (Redundant LECs)



(Also: 2018 Reinert, P. and Krebs, H. and Epelbaum, E.)

Model Validation

Model Validation

As far as the laws of mathematics refer to reality, they are not certain, and as far as they are certain, they do not refer to reality.

— Albert Einstein

Model Validation

> *As far as the laws of mathematics refer to reality, they are not certain, and as far as they are certain, they do not refer to reality.*
>
> — *Albert Einstein*

Does our model refer to reality? How can we check?

Model Validation

As far as the laws of mathematics refer to reality, they are not certain, and as far as they are certain, they do not refer to reality.

— Albert Einstein

Does our model refer to reality? How can we check?

Custom Tests

1. Are the theory error bands actually working?

Bayesian Solution

1. Use the posterior predictive $\mathrm{pr}(y \mid M)$ to generate fake data

Model Validation

As far as the laws of mathematics refer to reality, they are not certain, and as far as they are certain, they do not refer to reality.

— *Albert Einstein*

Does our model refer to reality? How can we check?

Custom Tests

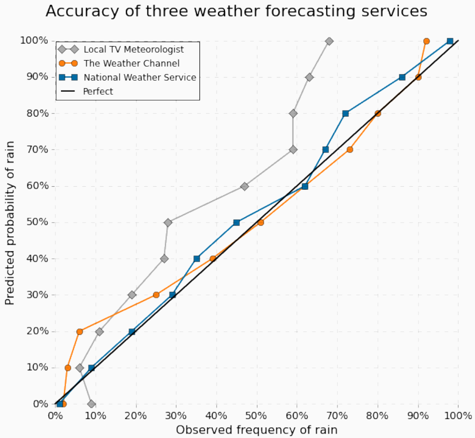
1. Are the theory error bands actually working?
2. Are the EFT parameters sufficient? Are they overfitting?

Bayesian Solution

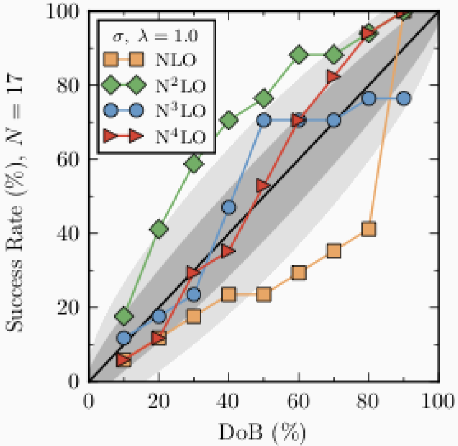
1. Use the posterior predictive $\mathrm{pr}(y \mid M)$ to generate fake data
2. Can assign probabilities to models

Credible Interval Diagnostics

Are our error bands working as advertised? Requires a reference distribution.



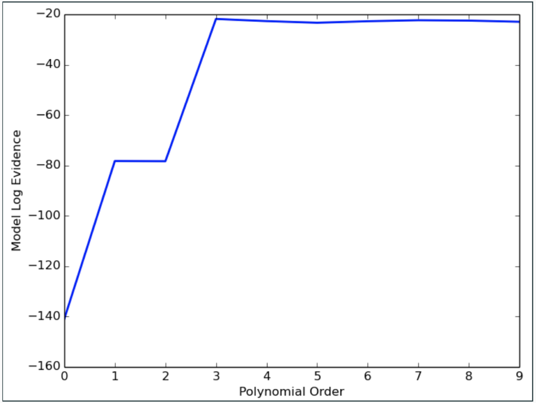
Source: "The Signal and the Noise" by Nate Silver | Author: Randy Olson (randalolson.com / @randal_olson)



Model Selection: Bayes Factors

- Which model is better supported by the data, M_0 or M_1?

$$\frac{\text{pr}(M_1 \mid \mathcal{D})}{\text{pr}(M_0 \mid \mathcal{D})} = \frac{\text{pr}(\mathcal{D} \mid M_1)}{\text{pr}(\mathcal{D} \mid M_0)} \frac{\text{pr}(M_1)}{\text{pr}(M_0)}$$

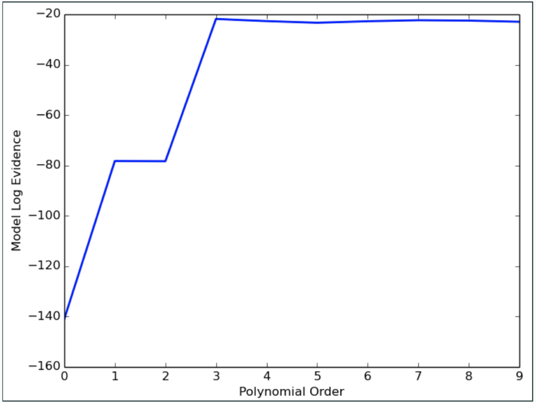


Model Selection: Bayes Factors

- Which model is better supported by the data, M_0 or M_1?

$$\frac{\mathrm{pr}(M_1 \mid \mathcal{D})}{\mathrm{pr}(M_0 \mid \mathcal{D})} = \frac{\mathrm{pr}(\mathcal{D} \mid M_1)}{\mathrm{pr}(\mathcal{D} \mid M_0)} \frac{\mathrm{pr}(M_1)}{\mathrm{pr}(M_0)}$$

- Requires computation of evidence $\mathrm{pr}(\mathcal{D} \mid M)$ which can be tricky
- But can be used to prevent overfitting, and to choose between competing models!



Takeaway Points

Building Statistical Models

Takeaway Points

Building Statistical Models

- List what you know and what you don't

$$y_{\text{exp}}(x) = y_{\text{th}}(x, \vec{a}) + \delta y_{\text{th}}(x) + \delta y_{\text{exp}}$$

Takeaway Points

Building Statistical Models

- List what you know and what you don't
- Build it in the language of random variables

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

Takeaway Points

Building Statistical Models

- List what you know and what you don't
- Build it in the language of random variables
- Check that it lives up to its own standards

Takeaway Points

Building Statistical Models

- List what you know and what you don't
- Build it in the language of random variables
- Check that it lives up to its own standards

Takeaway Points

Building Statistical Models

- List what you know and what you don't
- Build it in the language of random variables
- Check that it lives up to its own standards

What is Gained

- Full accounting of uncertainty (keeps experimentalists happy)

Takeaway Points

Building Statistical Models

- List what you know and what you don't
- Build it in the language of random variables
- Check that it lives up to its own standards

What is Gained

- Full accounting of uncertainty (keeps experimentalists happy)
- Can uncover physics and falsify models

Takeaway Points

Building Statistical Models

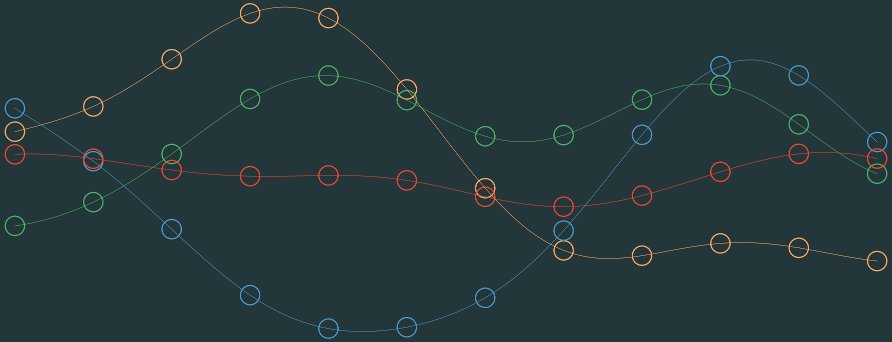
- List what you know and what you don't
- Build it in the language of random variables
- Check that it lives up to its own standards

What is Gained

- Full accounting of uncertainty (keeps experimentalists happy)
- Can uncover physics and falsify models
- Is easy due to new computational tools, like `PyMC3`

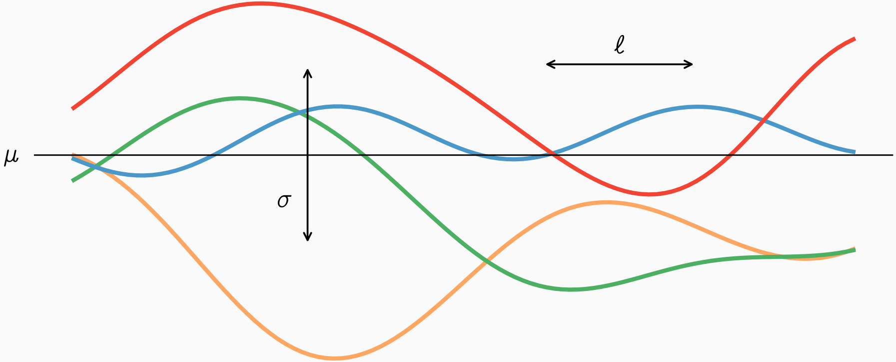
Thank you!

arxiv:1904.10581 arXiv:1808.08211 arXiv:1704.03308



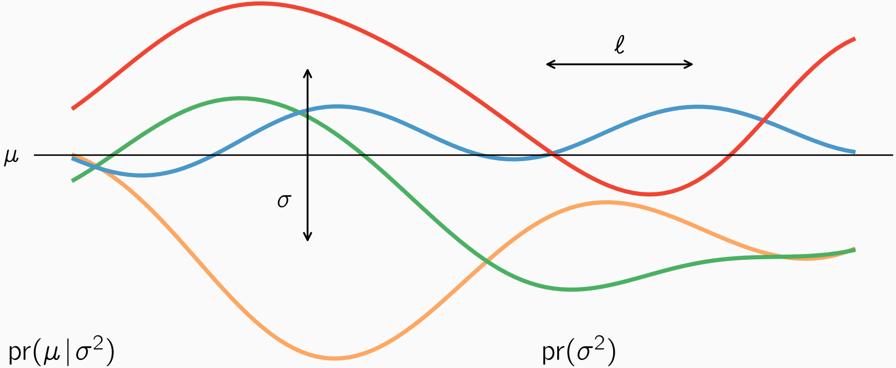
Gaussian Process Priors on Observable Coefficients

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



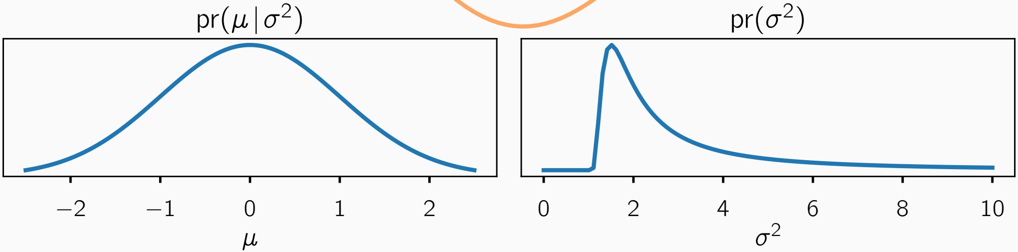
Gaussian Process Priors on Observable Coefficients

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



Conjugate priors:

$$\mu \mid \sigma^2 \sim \mathcal{N}(m, \sigma^2 V)$$
$$\sigma^2 \sim \mathrm{IG}(a, b)$$

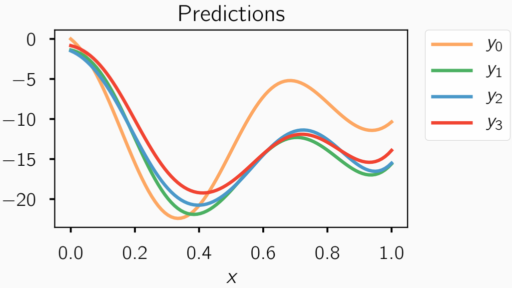


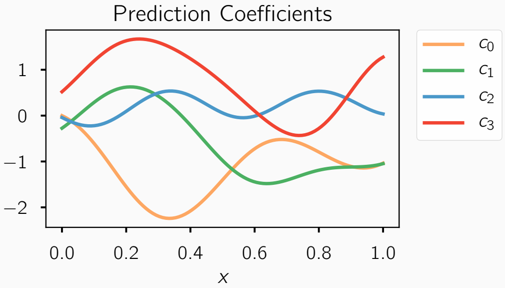
Model Building

Main equation

$$y_k = y_{\mathrm{ref}} \sum_{n=0}^{k} c_n Q^n$$

$$c_n \equiv \frac{y_n - y_{n-1}}{y_{\mathrm{ref}} Q^n}$$



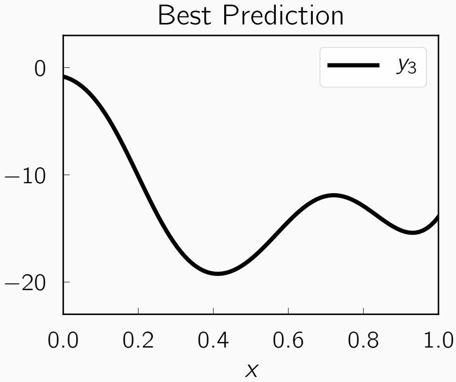


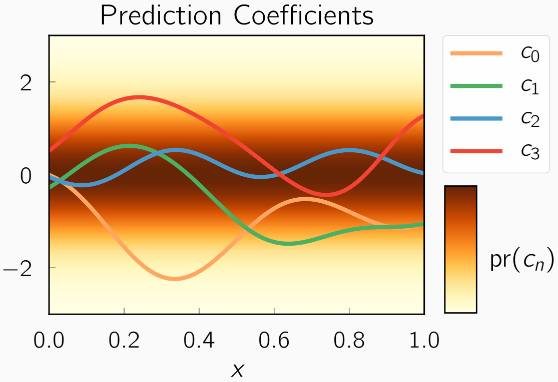
Model Building

Main equation

$$y_k = y_{\text{ref}} \sum_{n=0}^{k} c_n Q^n$$

$$c_n \equiv \frac{y_n - y_{n-1}}{y_{\text{ref}} Q^n}$$



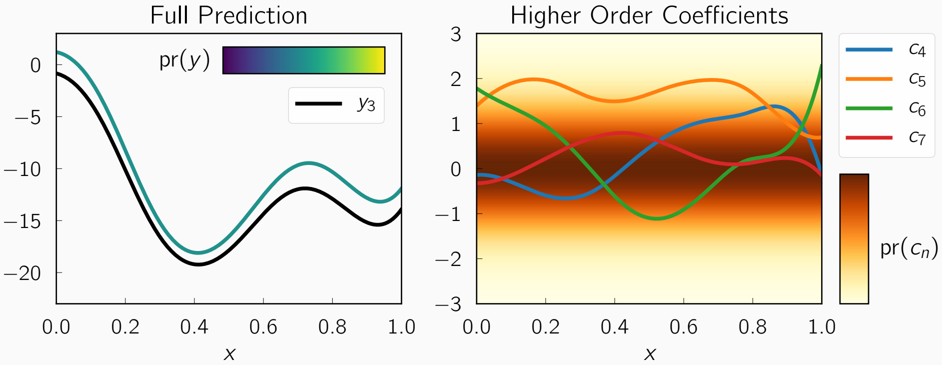


Model Building

Main equation

$$y = y_{\text{ref}} \sum_{n=0}^{\infty} c_n Q^n$$

$$c_n \equiv \frac{y_n - y_{n-1}}{y_{\text{ref}} Q^n}$$

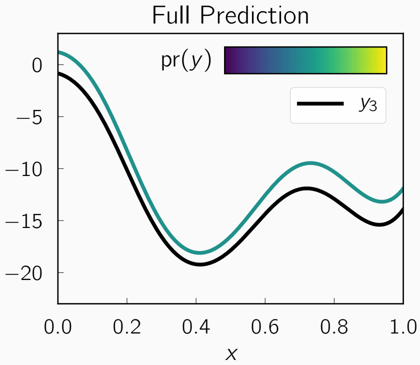


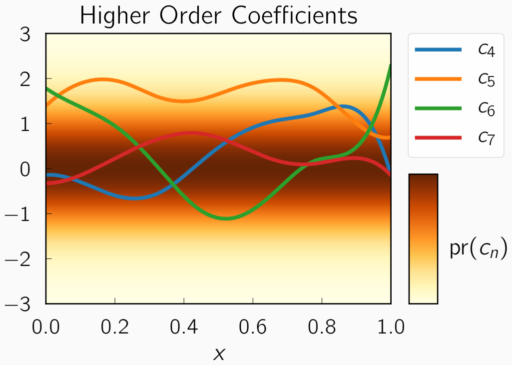
Model Building

Main equation

$$y = y_{\text{ref}} \sum_{n=0}^{\infty} c_n Q^n$$

$$c_n \equiv \frac{y_n - y_{n-1}}{y_{\text{ref}} Q^n}$$



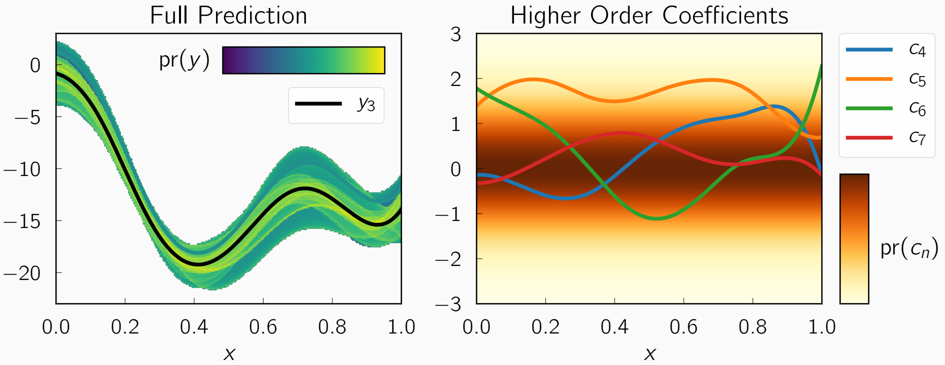


Model Building

Main equation

$$y = y_{\text{ref}} \sum_{n=0}^{\infty} c_n Q^n$$

$$c_n \equiv \frac{y_n - y_{n-1}}{y_{\text{ref}} Q^n}$$



Discrepancy Distribution

Remember the goal:

$$y_{\mathrm{exp}}(x) = y_{\mathrm{th}}(x, \vec{a}) + \delta y_{\mathrm{th}}(x) + \delta y_{\mathrm{exp}}$$

Discrepancy Distribution

Remember the goal:

$$y_{\mathrm{exp}}(x) = y_{\mathrm{th}}(x, \vec{a}) + \delta y_{\mathrm{th}}(x) + \delta y_{\mathrm{exp}}$$

Our convergence assumptions

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

$$\delta y_{\mathrm{th}}(x) = y_{\mathrm{ref}} \sum_{n=k+1}^{\infty} c_n Q^n$$

Discrepancy Distribution

Remember the goal:

$$y_{\mathrm{exp}}(x) = y_{\mathrm{th}}(x, \vec{a}) + \delta y_{\mathrm{th}}(x) + \delta y_{\mathrm{exp}}$$

Our convergence assumptions

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

$$\delta y_{\mathrm{th}}(x) = y_{\text{ref}} \sum_{n=k+1}^{\infty} c_n Q^n$$

Gaussian sum rules

$$a\mathcal{N}(\mu_1, \Sigma_1) + b\mathcal{N}(\mu_2, \Sigma_2) = \mathcal{N}(a\mu_1 + b\mu_2, a^2\Sigma_1 + b^2\Sigma_2)$$

Discrepancy Distribution

Remember the goal:

$$y_{\rm exp}(x) = y_{\rm th}(x, \vec{a}) + \delta y_{\rm th}(x) + \delta y_{\rm exp}$$

Our convergence assumptions

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

$$\delta y_{\rm th}(x) = y_{\rm ref} \sum_{n=k+1}^{\infty} c_n Q^n$$

Gaussian sum rules

$$a\mathcal{N}(\mu_1, \Sigma_1) + b\mathcal{N}(\mu_2, \Sigma_2) = \mathcal{N}(a\mu_1 + b\mu_2, a^2\Sigma_1 + b^2\Sigma_2)$$

Discrepancy Distribution

$$\mathrm{pr}(\delta y_{\rm th} \mid \boldsymbol{\theta}) = \mathcal{GP}\left(\mu_{\rm th}, \Sigma_{\rm th}\right) = \mathcal{GP}\left(\frac{\mu Q^{k+1}}{1-Q}, y_{\rm ref}^2 \frac{\sigma^2 Q^{2(k+1)}}{1-Q^2} R_\ell\right)$$

Implications for EFT Fitters

Standard χ^2

$$\sum_i \frac{[y_{\text{exp},i} - y_{\text{th},i}(\vec{a})]^2}{\sigma_{\text{exp}}^2} = \sum_i \frac{r(x_i, \vec{a})^2}{\sigma_{\text{exp}}^2}$$

Implications for EFT Fitters

Standard χ^2

$$\sum_i \frac{[y_{\exp,i} - y_{\mathrm{th},i}(\vec{a})]^2}{\sigma_{\exp}^2} = \sum_i \frac{r(x_i, \vec{a})^2}{\sigma_{\exp}^2}$$

χ^2 + Theory Error

$$\chi^2_{\mathrm{mod}}(\vec{a}) = \vec{r}^{\intercal}(\vec{a})(\Sigma_{\mathrm{th}} + \Sigma_{\exp})^{-1}\vec{r}(\vec{a})$$

Implications for EFT Fitters

Standard χ^2

$$\sum_i \frac{[y_{\mathrm{exp},i} - y_{\mathrm{th},i}(\vec{a})]^2}{\sigma^2_{\mathrm{exp}}} = \sum_i \frac{r(x_i, \vec{a})^2}{\sigma^2_{\mathrm{exp}}}$$

χ^2 + Theory Error

$$\chi^2_{\mathrm{mod}}(\vec{a}) = \vec{r}^{\intercal}(\vec{a})(\Sigma_{\mathrm{th}} + \Sigma_{\mathrm{exp}})^{-1}\vec{r}(\vec{a})$$

Prediction

$$y_{\mathrm{exp}}(x) \overset{?}{\approx} y_{\mathrm{th}}(x, \vec{a})$$

Implications for EFT Fitters

Standard χ^2

$$\sum_i \frac{[y_{\text{exp},i} - y_{\text{th},i}(\vec{a})]^2}{\sigma^2_{\text{exp}}} = \sum_i \frac{r(x_i, \vec{a})^2}{\sigma^2_{\text{exp}}}$$

χ^2 + Theory Error

$$\chi^2_{\text{mod}}(\vec{a}) = \vec{r}^{\intercal}(\vec{a})(\Sigma_{\text{th}} + \Sigma_{\text{exp}})^{-1}\vec{r}(\vec{a})$$

Prediction

$$y_{\text{exp}}(x) \overset{?}{\approx} y_{\text{th}}(x, \vec{a})$$

Prediction + Theory Error

$$\text{pr}(y_{\text{exp}}) = \mathcal{N}\left[y_{\text{th}}(x, \vec{a}) + \mu_{\text{th}}, \Sigma_{\text{th}}\right]$$

Implications for EFT Fitters

Standard χ^2

$$\sum_i \frac{[y_{\mathrm{exp},i} - y_{\mathrm{th},i}(\vec{a})]^2}{\sigma^2_{\mathrm{exp}}} = \sum_i \frac{r(x_i, \vec{a})^2}{\sigma^2_{\mathrm{exp}}}$$

χ^2 + Theory Error

$$\chi^2_{\mathrm{mod}}(\vec{a}) = \vec{r}^{\mathsf{T}}(\vec{a})(\Sigma_{\mathrm{th}} + \Sigma_{\mathrm{exp}})^{-1}\vec{r}(\vec{a})$$

Prediction

$$y_{\mathrm{exp}}(x) \overset{?}{\approx} y_{\mathrm{th}}(x, \vec{a})$$

Prediction + Theory Error

$$\mathrm{pr}(y_{\mathrm{exp}}) = \mathcal{N}\left[y_{\mathrm{th}}(x, \vec{a}) + \mu_{\mathrm{th}}, \Sigma_{\mathrm{th}}\right]$$

- Gaussian process correlations propagate via Σ_{th} matrix (computed once!)

Implications for EFT Fitters

Standard χ^2

$$\sum_i \frac{[y_{\text{exp},i} - y_{\text{th},i}(\vec{a})]^2}{\sigma^2_{\text{exp}}} = \sum_i \frac{r(x_i, \vec{a})^2}{\sigma^2_{\text{exp}}}$$

χ^2 + Theory Error

$$\chi^2_{\text{mod}}(\vec{a}) = \vec{r}^{\intercal}(\vec{a})(\Sigma_{\text{th}} + \Sigma_{\text{exp}})^{-1}\vec{r}(\vec{a})$$

Prediction

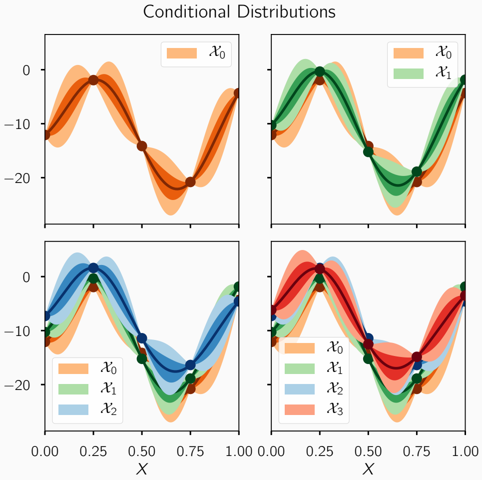
$$y_{\text{exp}}(x) \overset{?}{\approx} y_{\text{th}}(x, \vec{a})$$

Prediction + Theory Error

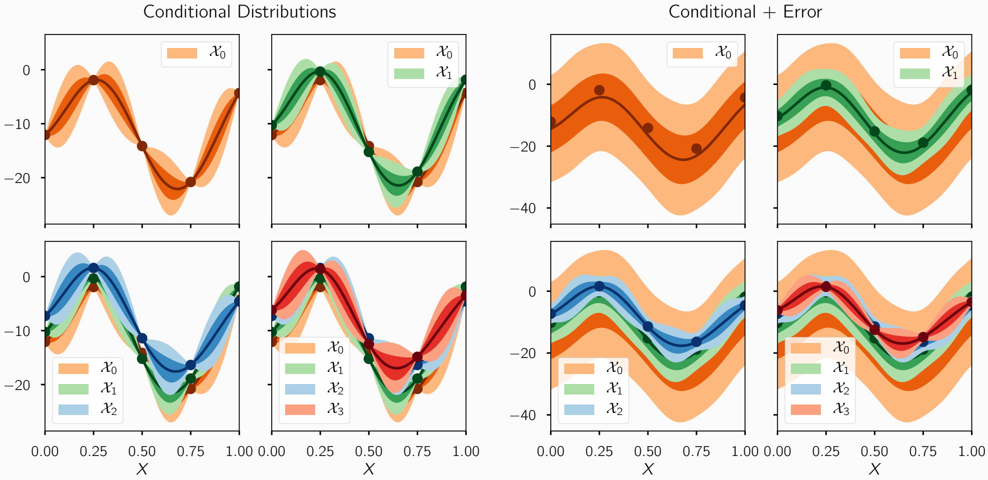
$$\text{pr}(y_{\text{exp}}) = \mathcal{N}\left[y_{\text{th}}(x, \vec{a}) + \mu_{\text{th}}, \Sigma_{\text{th}}\right]$$

- Gaussian process correlations propagate via Σ_{th} matrix (computed once!)
- Different correlation assumptions $\rightarrow$ different results!

Quantifying Truncation Uncertainty



Quantifying Truncation Uncertainty



What You Get for (Almost) Free: Evidence, Length Scale, & Breakdown Scale

This model permits mostly analytic calculation of evidence

$$\mathrm{pr}(\mathcal{D} \mid \ell, Q) = \frac{\Gamma(a^\star)}{\Gamma(a)} \frac{b^a}{(b^\star)^{a^\star}} \sqrt{\frac{|V^\star|}{|V|}} \frac{|2\pi R_\ell|^{-(k+1)/2}}{|Q|^{k(k+1)/2}}$$

What You Get for (Almost) Free: Evidence, Length Scale, & Breakdown Scale

This model permits mostly analytic calculation of evidence

$$\mathrm{pr}(\mathcal{D} \mid \ell, Q) = \frac{\Gamma(a^\star)}{\Gamma(a)} \frac{b^a}{(b^\star)^{a^\star}} \sqrt{\frac{|V^\star|}{|V|}} \frac{|2\pi R_\ell|^{-(k+1)/2}}{|Q|^{k(k+1)/2}}$$

Important for model comparison and for posteriors:

$$\mathrm{pr}(\ell \mid \mathcal{D}, Q) \propto \mathrm{pr}(\mathcal{D} \mid \ell, Q)\, \mathrm{pr}(\ell)$$

$$\mathrm{pr}(Q \mid \mathcal{D}, \ell) \propto \mathrm{pr}(\mathcal{D} \mid \ell, Q)\, \mathrm{pr}(Q)$$

What You Get for (Almost) Free: Evidence, Length Scale, & Breakdown Scale

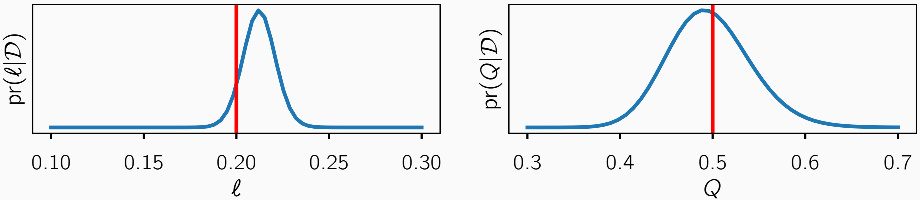
This model permits mostly analytic calculation of evidence

$$\mathrm{pr}(\mathcal{D} \mid \ell, Q) = \frac{\Gamma(a^\star)}{\Gamma(a)} \frac{b^a}{(b^\star)^{a^\star}} \sqrt{\frac{|V^\star|}{|V|}} \frac{|2\pi R_\ell|^{-(k+1)/2}}{|Q|^{k(k+1)/2}}$$

Important for model comparison and for posteriors:

$$\mathrm{pr}(\ell \mid \mathcal{D}, Q) \propto \mathrm{pr}(\mathcal{D} \mid \ell, Q)\,\mathrm{pr}(\ell)$$

$$\mathrm{pr}(Q \mid \mathcal{D}, \ell) \propto \mathrm{pr}(\mathcal{D} \mid \ell, Q)\,\mathrm{pr}(Q)$$



What You Get for (Almost) Free: Evidence, Length Scale, & Breakdown Scale

This model permits mostly analytic calculation of evidence

$$\mathrm{pr}(\mathcal{D} \,|\, \ell, Q) = \frac{\Gamma(a^\star)}{\Gamma(a)} \frac{b^a}{(b^\star)^{a^\star}} \sqrt{\frac{|V^\star|}{|V|}} \frac{|2\pi R_\ell|^{-(k+1)/2}}{|Q|^{k(k+1)/2}}$$

Important for model comparison and for posteriors:

$$\mathrm{pr}(\ell \,|\, \mathcal{D}, Q) \propto \mathrm{pr}(\mathcal{D} \,|\, \ell, Q)\, \mathrm{pr}(\ell)$$

$$\mathrm{pr}(Q \,|\, \mathcal{D}, \ell) \propto \mathrm{pr}(\mathcal{D} \,|\, \ell, Q)\, \mathrm{pr}(Q)$$

Here, $Q \propto \dfrac{1}{\Lambda_b}$