Bayesian Fitting
in
Effective Field Theory

Dick Furnstahl

Department of Physics
Ohio State University

February, 2006

Collaborators: D. Phillips (Ohio U.), U. van Kolck (Arizona),
R.G.E. Timmermans (Groningen, Nijmegen)
What Does The Typical Physicist Know About Bayesian Statistics?

- Nothing!
- But there are exceptions:
  - Experimentalists searching for weak signatures of particles
  - Some “Beyond the Standard Model” theorists
  - Some lattice gauge theorists doing “constrained curve fitting”
- Even where applied, Bayesian methods are controversial
- Not applied as yet in EFT or nuclear structure calculations
  - The problem described here is just one (simple) example where Bayesian methods might be useful
Constrained Curve Fitting Example (hep-lat/0110175)

- Monte Carlo estimates of a meson correlator $G(t)$ are generated at 24 time steps, $t = 0, 1, \cdots, 23$.
- Theory says the exact correlator has the form:

$$G_{th}(t; A_n, E_n) = \sum_{n=1}^{\infty} A_n e^{-E_n t}$$

- Challenge: Fit an infinite number of amplitudes $A_n$ and energies $E_n$ using only 24 $G(t)$’s.
- Standard procedure:
  - Keep only first few terms in $G_{th}$
  - Fit using only Monte Carlo data from $t \geq t_{\text{min}}$
  - But choosing $t_{\text{min}}$ is \textit{ad hoc}
Results From Standard Fits

- Left: 2-term fit to \( \sum_{n=1}^{\infty} A_n e^{-E_n t} \) for \( t \geq t_{\text{min}} \)
- Competition between large systematic errors for small \( t_{\text{min}} \)
  and large statistical errors for large \( t_{\text{max}} \)
- Right: Fit values for lowest two energies vs. # of terms
Unconstrained vs. Constrained Fits

- Goal: Fit all data using as many terms as we wish
- Plan: Add priors for reasonable $A_n$’s and $E_n$’s in $\sum_n A_n e^{-E_n t}$

Left: unconstrained. Right: constrained.
Overview: Problem(s) to Be Solved

Method: Effective Field Theory (EFT)

Request: Advice on Applying Bayesian Methods
Overview: Problem(s) to Be Solved

Method: Effective Field Theory (EFT)

Request: Advice on Applying Bayesian Methods
The Islands of Strong Interaction Physics

Figure 1: From QCD vacuum to heavy nuclei: the intellectual connection between the hadronic many-body problem (quark-gluon description of a nucleon) and the nucleonic many-body problem (nucleus as a system of Z protons and N neutrons). The bridges illustrate major physics challenges: the mechanism of quark confinement, the understanding of the bare nucleon-nucleon interaction in terms of the quark-gluon dynamics, and the understanding of the effective interactions in heavy nuclei in terms of the bare force.
The Big Picture (adapted from Richter @INPC2004)
Table of the Nuclides
Problems with Extrapolations

- Mass formulas and energy functionals do well where there is data, but elsewhere . . .

![Graph showing two-neutron separation energies for Sn isotopes](image)

- The position of the neutron drip line is uncertain.
- Unknown nuclear deformations or as yet uncharacterized phenomena, such as the presence of neutron halos or neutron skins, make theoretical predictions highly uncertain.
- Experiments for the Sn isotopes with N=80–100 will greatly narrow the choice of viable models.
Input to Many-Body Problem: Internucleon Force

- Reproduce data from scattering protons from neutrons, etc.
- Difficult problem with long history
- M.L. Goldberger, at the Midwestern Conference on Theoretical Physics, Purdue University, 1960:

  “There are few problems in nuclear theoretical physics which have attracted more attention than that of trying to determine the fundamental interaction between two nucleons. It is also true that scarcely ever has the world of physics owed so little to so many. . . . It is hard to believe that many of the authors are talking about the same problem or, in fact, that they know what the problem is.”
Successful Fits to Phase Shift Data Achieved

- Fit energies from 0 to 350 MeV, with goal of $\chi^2$/dof $\approx 1$
- Account only for measurement errors
- Table from Argonne $v_{18}$ paper:

<table>
<thead>
<tr>
<th>Bin (MeV)</th>
<th>$N_{pp}$</th>
<th>PWA93</th>
<th>$v_{18}$</th>
<th>$N_{np}$</th>
<th>PWA93</th>
<th>$v_{18}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0–0.5</td>
<td>134</td>
<td>134.5</td>
<td>136.3</td>
<td>10</td>
<td>9.7</td>
<td>11.8</td>
</tr>
<tr>
<td>0.5–2</td>
<td>63</td>
<td>39.7</td>
<td>41.1</td>
<td>5</td>
<td>3.8</td>
<td>7.4</td>
</tr>
<tr>
<td>2–8</td>
<td>48</td>
<td>45.0</td>
<td>36.0</td>
<td>55</td>
<td>52.4</td>
<td>51.0</td>
</tr>
<tr>
<td>8–17</td>
<td>108</td>
<td>103.0</td>
<td>111.6</td>
<td>182</td>
<td>168.3</td>
<td>164.8</td>
</tr>
<tr>
<td>17–35</td>
<td>59</td>
<td>63.1</td>
<td>72.2</td>
<td>293</td>
<td>226.6</td>
<td>234.9</td>
</tr>
<tr>
<td>35–75</td>
<td>243</td>
<td>213.4</td>
<td>251.5</td>
<td>328</td>
<td>335.2</td>
<td>339.3</td>
</tr>
<tr>
<td>75–125</td>
<td>167</td>
<td>169.5</td>
<td>171.5</td>
<td>232</td>
<td>237.1</td>
<td>231.3</td>
</tr>
<tr>
<td>125–183</td>
<td>343</td>
<td>379.7</td>
<td>415.7</td>
<td>333</td>
<td>336.8</td>
<td>363.5</td>
</tr>
<tr>
<td>183–290</td>
<td>239</td>
<td>285.9</td>
<td>304.8</td>
<td>517</td>
<td>494.6</td>
<td>574.0</td>
</tr>
<tr>
<td>290–350</td>
<td>383</td>
<td>360.7</td>
<td>421.3</td>
<td>571</td>
<td>599.0</td>
<td>708.0</td>
</tr>
<tr>
<td>0–350</td>
<td>1787</td>
<td>1794.5</td>
<td>1962.0</td>
<td>2526</td>
<td>2463.5</td>
<td>2685.8</td>
</tr>
</tbody>
</table>

- No theoretical errors, all data treated equally
Outline

Overview: Problem(s) to Be Solved

Method: Effective Field Theory (EFT)

Request: Advice on Applying Bayesian Methods
Resolution and the Pointillists

- George Seurat painted using closely spaced small dots (∼ 0.4 mm wide) of pure pigment

- Why do the dots blend together?
Resolution and the Pointillists

- George Seurat painted using closely spaced small dots (≈ 0.4 mm wide) of pure pigment

Why do the dots blend together?
Wavelength and Resolution
Wavelength and Resolution
Wavelength and Resolution
Wavelength and Resolution
Wavelength and Resolution
Wavelength and Resolution
Wavelength and Resolution
Wavelength and Resolution
Wavelength and Resolution
Principles of Effective Low-Energy Theories

\[ \lambda \ll R \]
Principles of Effective Low-Energy Theories

If system is probed at low energies, fine details not resolved
Principles of Effective Low-Energy Theories

- If system is probed at low energies, fine details not resolved
  - use low-energy variables for low-energy processes
  - short-distance structure can be replaced by something simpler without distorting low-energy observables
Effective Field Theory Ingredients

From “Crossing the Border” [nucl-th/0008064]

1. Use the most general $\mathcal{L}$ with low-energy dof’s consistent with the global and local symmetries of the underlying theory

2. Declaration of regularization and renormalization scheme

3. Well-defined power counting $\Rightarrow$ expansion parameters
Effective Field Theory Ingredients: Chiral NN

From “Crossing the Border” [nucl-th/0008064]

1 Use the most general $\mathcal{L}$ with low-energy dof’s consistent with the global and local symmetries of the underlying theory
   - $\mathcal{L}_{\text{eft}} = \mathcal{L}_{\pi\pi} + \mathcal{L}_{\pi N} + \mathcal{L}_{\text{NN}}$
   - chiral symmetry $\implies$ systematic long-distance pion physics

2 Declaration of regularization and renormalization scheme
   - momentum cutoff and “Weinberg counting”
   - use cutoff sensitivity as measure of uncertainties!

3 Well-defined power counting $\implies$ expansion parameters
   - use the separation of scales $\implies$ $\left\{ \mathbf{p}, \frac{m_\pi}{\Lambda_\chi} \right\}$ with $\Lambda_\chi \sim 1$ GeV
   - chiral symmetry $\implies$ $V_{NN} = \sum_{\nu=\nu_{\text{min}}}^{\infty} c_\nu Q^\nu$ with $\nu \geq 0$
   - naturalness: parameters are $\mathcal{O}(1)$ in appropriate units
Chiral Lagrangian

- Unified description of $\pi\pi$, $\pi N$, and $NN \cdots N$
- Lowest orders:

\[
\mathcal{L}^{(0)} = \frac{1}{2} \partial_\mu \pi \cdot \partial^\mu \pi \quad \text{and} \quad \frac{1}{2} M^2 \pi^2 + N^\dagger \left[ i \partial_0 + \frac{g_A}{2F} \tau_\sigma \cdot \nabla \pi - \frac{1}{4F^2} \tau \cdot (\pi \times \dot{\pi}) \right] N
\]

\[
- \frac{1}{2} C_S (N^\dagger N)(N^\dagger N) - \frac{1}{2} C_T (N^\dagger \sigma N)(N^\dagger \sigma N) + \ldots
\]

\[
\mathcal{L}^{(1)} = N^\dagger \left[ 4c_1 M^2 - \frac{2c_1}{F^2} M^2 \pi^2 + \frac{c_2}{F^2} \dot{\pi}^2 + \frac{c_3}{F^2} (\partial_\mu \pi \cdot \partial^\mu \pi)
\]

\[
- \frac{c_4}{2F^2} \epsilon_{ijk} \epsilon_{abc} \sigma_i \tau_a (\nabla_j \pi_b)(\nabla_k \pi_c) \right] N
\]

\[
- \frac{D}{4F} (N^\dagger N)(N^\dagger \sigma \tau N) \cdot \nabla \pi - \frac{1}{2} E (N^\dagger N)(N^\dagger \tau N) \cdot (N^\dagger \tau N) + \ldots
\]

- Infinite # of unknown parameters in hierarchy
**Chiral Effective Field Theory for Two Nucleons**

- Epelbaum, Meißner, et al.
- Also Entem, Machleidt
- $\mathcal{L}_{\pi N} + \text{match at low energy}$

<table>
<thead>
<tr>
<th>$Q^\nu$</th>
<th>$1\pi$</th>
<th>$2\pi$</th>
<th>$4N$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Chiral Effective Field Theory for Two Nucleons

- Epelbaum, Meißner, et al.
- Also Entem, Machleidt
- $\mathcal{L}_{\pi N} + \text{match at low energy}$

<table>
<thead>
<tr>
<th>$Q^\nu$</th>
<th>$1\pi$</th>
<th>$2\pi$</th>
<th>$4N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q^0$</td>
<td></td>
<td></td>
<td>(2)</td>
</tr>
</tbody>
</table>

Dick Furnstahl
Bayesian Fitting in EFT
Chiral Effective Field Theory for Two Nucleons

- Epelbaum, Meißner, et al.
- Also Entem, Machleidt
- $\mathcal{L}_{\pi N}$ + match at low energy

<table>
<thead>
<tr>
<th>$Q^\nu$</th>
<th>$1_\pi$</th>
<th>$2_\pi$</th>
<th>$4N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q^0$</td>
<td>$\pi$</td>
<td></td>
<td>(2)</td>
</tr>
<tr>
<td>$Q^1$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dick Furnstahl
Bayesian Fitting in EFT
Chiral Effective Field Theory for Two Nucleons

- Epelbaum, Meißner, et al.
- Also Entem, Machleidt
- $L_{\pi N}$ + match at low energy

<table>
<thead>
<tr>
<th>$Q^\nu$</th>
<th>$1_\pi$</th>
<th>$2_\pi$</th>
<th>$4N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q^0$</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /> (2)</td>
</tr>
<tr>
<td>$Q^1$</td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td></td>
</tr>
<tr>
<td>$Q^2$</td>
<td><img src="image6.png" alt="Image" /></td>
<td><img src="image7.png" alt="Image" /> (7)</td>
<td></td>
</tr>
</tbody>
</table>

Dick Furnstahl

Bayesian Fitting in EFT
Chiral Effective Field Theory for Two Nucleons

- Epelbaum, Meißner, et al.
- Also Entem, Machleidt
- $\mathcal{L}_{\pi N}$ + match at low energy

<table>
<thead>
<tr>
<th>$Q^\nu$</th>
<th>$1\pi$</th>
<th>$2\pi$</th>
<th>$4N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q^0$</td>
<td>$\pi$</td>
<td></td>
<td>(2)</td>
</tr>
<tr>
<td>$Q^1$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Q^2$</td>
<td>$\pi^*$</td>
<td>$\nabla^2$</td>
<td>(7)</td>
</tr>
<tr>
<td>$Q^3$</td>
<td>$\mathcal{L}_{\pi N}$</td>
<td>$\mathcal{L}_{\pi N}$</td>
<td></td>
</tr>
</tbody>
</table>

Dick Furnstahl
Bayesian Fitting in EFT
Chiral Effective Field Theory for Two Nucleons

- Epelbaum, Meißner, et al.
- Also Entem, Machleidt
- $\mathcal{L}_{\pi N}$ + match at low energy

<table>
<thead>
<tr>
<th>$Q^\nu$</th>
<th>$1_\pi$</th>
<th>$2_\pi$</th>
<th>$4N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q^0$</td>
<td>(\pi)</td>
<td></td>
<td>(\nabla^2) (2)</td>
</tr>
<tr>
<td>$Q^1$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Q^2$</td>
<td></td>
<td></td>
<td>(\nabla^2) (7)</td>
</tr>
<tr>
<td>$Q^3$</td>
<td>(\pi)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$Q^4$</td>
<td>many</td>
<td>many</td>
<td>(\nabla^4) (15)</td>
</tr>
</tbody>
</table>

Dick Furnstahl
Bayesian Fitting in EFT
Motivation For Applying Effective Field Theory

- Systematic calculations with theoretical error estimates
- Reliable, model independent extrapolation
- Analogy between EFT and basic numerical analysis
  - naive error analysis: pick a method and reduce the mesh size (e.g., increase grid points) until the error is “acceptable”
  - sophisticated error analysis: understand scaling and sources of error (e.g., algorithm vs. round-off errors)
  \[\Rightarrow\text{Does it work as well as it should?}\]
- representation dependence \[\Rightarrow\text{not all are equally effective!}\]
- extrapolation: completeness of an expansion basis
Error Plots in Numerical Analysis

Numerical Derivatives

\[ f'(x) = \frac{[f(x+h)-f(x)]}{h} + O(h) \]

relative error

mesh size h
Error Plots in Numerical Analysis

Numerical Derivatives

\[ f'(x) = \frac{f(x+h)-f(x)}{h} + O(h) \]

\[ f'(x) = \frac{f(x+h/2)-f(x-h/2)}{h} + O(h^2) \]

Relative error vs. mesh size h
Error Plots in Numerical Analysis

Numerical Derivatives

- $f'(x) = [f(x+h)-f(x)]/h + O(h)$
- $f'(x) = [f(x+h/2)-f(x-h/2)]/h + O(h^2)$
- Richardson extrapolation $O(h^4)$

Dick Furnstahl

Bayesian Fitting in EFT
Error Plots in Numerical Analysis

Numerical Integration

relative error $h_n$ to $h_{n-1}$

trapezoid rule $O(h^2)$

mesh size $h$

Dick Furnstahl
Bayesian Fitting in EFT
Error Plots in Numerical Analysis

Numerical Integration

- trapezoid rule $O(h^2)$
- Simpson’s rule $O(h^4)$

Relative error $h_n$ to $h_{n-1}$

Mesh size $h$

Dick Furnstahl
Bayesian Fitting in EFT
Error Plots in Numerical Analysis

Numerical Integration

- trapezoid rule $O(h^2)$
- Simpson’s rule $O(h^4)$
- Milne’s rule $O(h^6)$

Mesh size $h$ to $h_{n-1}$
The Representation Can Make A Difference!

- E.g., elliptic integral:

\[ \int_0^1 \sqrt{(1 - x^2)(2 - x)} \, dx \]
The Representation Can Make A Difference!

- E.g., elliptic integral:

\[ \int_0^1 \sqrt{(1 - x^2)(2 - x)} \, dx \]

- How do the numerical errors behave?

Numerical Integration

- trapezoid rule \( O(h^2) \)
- Simpson’s rule \( O(h^4) \)
- Milne’s rule \( O(h^6) \)

![Graph showing relative error vs mesh size for different integration rules.](image-url)
The Representation Can Make A Difference!

- E.g., elliptic integral:
  \[\int_{0}^{1} \sqrt{(1-x^2)(2-x)} \, dx\]

- How do the numerical errors behave?

- After transformation:
  \[\int_{0}^{\pi/2} \sin^2 y \sqrt{2-\cos y} \, dy\]
The Representation Can Make A Difference!

- E.g., elliptic integral:
  \[ \int_0^1 \sqrt{(1 - x^2)(2 - x)} \, dx \]

- How do the numerical errors behave?

- After transformation:
  \[ \int_0^{\pi/2} \sin^2 y \sqrt{2 - \cos y} \, dy \]

![Numerical Integration Graph]

- trapezoid rule \(O(h^2)\)
- Simpson’s rule \(O(h^4)\)
- Milne’s rule \(O(h^6)\)

Dick Furnstahl  Bayesian Fitting in EFT
The Representation Can Make A Difference!

- E.g., elliptic integral:
  \[ \int_0^1 \sqrt{(1 - x^2)(2 - x)} \, dx \]

- How do the numerical errors behave?

- After transformation:
  \[ \int_0^{\pi/2} \sin^2 y \sqrt{2 - \cos y} \, dy \]

### Numerical Integration

- Trapezoid rule $O(h^2)$
- Simpson’s rule $O(h^4)$
- Milne’s rule $O(h^6)$

Graph showing the relative error $h_n$ to $h_{n-1}$ for different mesh sizes $h$. The graph compares the numerical integration before and after transformation.
Error Plots in Effective Field Theory

G. P. Lepage, “How to Renormalize the Schrödinger Equation”

Errors in the $^1S_0$ phase shifts versus energy through orders $\Lambda^{-2}$ and $\Lambda^{-4}$.

Fit using data at low energy only.
Error Plots in Effective Field Theory

G. P. Lepage, “How to Renormalize the Schrödinger Equation”

Errors in the $^1S_0$ phase shifts versus energy for different values of the cutoff $\Lambda$. 

Dick Furnstahl  
Bayesian Fitting in EFT
Best Calculation: Theoretical Error A Posteriori

Dick Furnstahl
Bayesian Fitting in EFT
Naturalness of Coefficients (Epelbaum et al.)

- Georgi-Manohar naive dimensional analysis (NDA):

\[
\mathcal{L}_\chi^{\text{eft}} = c_{lmn} \left( \frac{N^\dagger(\cdots)N}{f_\pi^2 \Lambda_\chi} \right)^l \left( \frac{\pi}{f_\pi} \right)^m \left( \frac{\partial^\mu, m_\pi}{\Lambda_\chi} \right)^n f_\pi^2 \Lambda_\chi^2
\]

- \( f_\pi \sim 100 \text{ MeV} \) and \( \Lambda_\chi \sim 1000 \text{ MeV} \)

- check NLO, NNLO constants from \( \mathcal{L}_{NN} \) (cutoff 500...600 MeV):

<table>
<thead>
<tr>
<th>( f_\pi^2 )</th>
<th>( C_S )</th>
<th>( f_\pi^2 )</th>
<th>( C_T )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f_\pi^2 \Lambda_\chi^2 )</td>
<td>( C_1 )</td>
<td>( 4 f_\pi^2 \Lambda_\chi^2 )</td>
<td>( C_2 )</td>
</tr>
<tr>
<td>( f_\pi^2 \Lambda_\chi^2 )</td>
<td>( C_3 )</td>
<td>( 4 f_\pi^2 \Lambda_\chi^2 )</td>
<td>( C_4 )</td>
</tr>
<tr>
<td>( 2 f_\pi^2 \Lambda_\chi^2 )</td>
<td>( C_5 )</td>
<td>( f_\pi^2 \Lambda_\chi^2 )</td>
<td>( C_6 )</td>
</tr>
<tr>
<td>( 4 f_\pi^2 \Lambda_\chi^2 )</td>
<td>( C_7 )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- \( \frac{1}{3} \lesssim c_{lmn} \lesssim 3 \implies \text{natural!} \implies \) truncation error estimates
- \( f_\pi^2 C_T \) unnaturally small \( \implies \) SU(4) spin-isospin symmetry
Outline

Overview: Problem(s) to Be Solved

Method: Effective Field Theory (EFT)

Request: Advice on Applying Bayesian Methods
Summary of (Some of the) Questions

- How can we incorporate the expected behavior of the theoretical error based on naturalness and the EFT hierarchy?
- Frequently Asked Question: How much of the scattering data should be fit (i.e., up to what energy)?
  - Traditionalist says: fit all data up to 350 MeV with $\chi^2$/dof $\approx 1$
  - EFT practitioner says: only use the low-energy data
  - How do we use all the data, accounting for the expected better description at low energy?
- How do we calculate uncertainties in our best-fit parameters?
  - This is particularly important when using the fit to calculate elsewhere (e.g., nuclei)