Empirical Pion and Kaon PDFs

Nobuo Sato
University of Connecticut
Pion and Kaon Structure at an Electron-Ion Collider, Argonne, 2017
Motivations

PHYSICAL REVIEW D 93, 054011 (2016)

Pion structure function from leading neutron electroproduction
and SU(2) flavor asymmetry

J. R. McKenney,¹,²,³ Nobuo Sato,² W. Melnitchouk,² and Chueng-Ryong Ji¹
¹North Carolina State University, Raleigh, North Carolina 27695, USA
²Jefferson Lab, Newport News, Virginia 23606, USA
³University of North Carolina, Chapel Hill, North Carolina 27599, USA
(Received 15 December 2015; published 7 March 2016)

- Study pion structure using pion exchange models and pQCD
- **Unknowns:**
  - pion splitting functions (UV regulator as a parameter)
  - small $x \ F_2^\pi$ (shape parameters)
- **Input:**
  - leading neutron cross sections from H1 and Zeus
  - $\bar{d} - \bar{u}$ asymmetry from E866
  - large $x$ pion PDFs (SMRS)
Motivations

PHYSICAL REVIEW D 93, 054011 (2016)

Pion structure function from leading neutron electroproduction and SU(2) flavor asymmetry

J. R. McKenney,\textsuperscript{1,2,3} Nobuo Sato,\textsuperscript{2} W. Melnitchouk,\textsuperscript{2} and Chueng-Ryong Ji\textsuperscript{1}

\textsuperscript{1}North Carolina State University, Raleigh, North Carolina 27695, USA
\textsuperscript{2}Jefferson Lab, Newport News, Virginia 23606, USA
\textsuperscript{3}University of North Carolina, Chapel Hill, North Carolina 27599, USA
(Received 15 December 2015; published 7 March 2016)
Feasibility to constrain $F_2^\pi$ at small $x$

Potential reduction of the uncertainty in the pion sea PDFs

Next step: fit pion PDFs instead of $F_2^\pi$

... but first we need to learn how to fit pion PDFs at large $x$
Data analysis framework:

The goal is to estimate:

\[ E[\mathcal{O}] = \int d^n a \ P(a|\text{data}) \ \mathcal{O}(a) \]

\[ V[\mathcal{O}] = \int d^n a \ P(a|\text{data}) \ [\mathcal{O}(a) - E[\mathcal{O}]]^2 \]

- \( \mathbf{a} = (N, a, b, \ldots) \) is a vector of parameters

  i.e. \( f(x, Q_0^2) = N x^a (1 - x)^b P(x) \)

- \( \mathcal{O}(\mathbf{a}) \) is an observable

  i.e. PDFs, PPDFs, FF, cross sections
Data analysis framework:

The goal is to estimate:

\[
E[\mathcal{O}] = \int d^m a \ P(a|data) \ \mathcal{O}(a)
\]

\[
V[\mathcal{O}] = \int d^m a \ P(a|data) \ [\mathcal{O}(a) - E[\mathcal{O}]]^2
\]

Maximum Likelihood

- Maximize \( P(a|data) \rightarrow a_0 \)

- \( E[\mathcal{O}] \approx \mathcal{O}(a_0) \)

- \( V[\mathcal{O}] \approx \text{Hessian, Lagrange multipliers} \)
Data analysis framework:

The goal is to estimate:

\[
E[O] = \int d^n a \ P(a|data) \ O(a)
\]

\[
V[O] = \int d^n a \ P(a|data) \ [O(a) - E[O]]^2
\]

Maximum Likelihood

- Maximize \( P(a|data) \rightarrow a_0 \)
- \( E[O] \approx O(a_0) \)
- \( V[O] \approx \text{Hessian, Lagrange multipliers} \)

Monte Carlo methods

- \( P(a|data) \rightarrow \{a_k\} \)
- \( E[O] \approx \frac{1}{N} \sum_k O(a_k) \)
- \( V[O] \approx \frac{1}{N} \sum_k [O(a_k) - E[O]]^2 \)
Data analysis framework:

The goal is to estimate:

\[ E[\mathcal{O}] = \int d^n a \ P(a|data) \ \mathcal{O}(a) \]

\[ V[\mathcal{O}] = \int d^n a \ P(a|data) \left[ \mathcal{O}(a) - E[\mathcal{O}] \right]^2 \]

Maximum Likelihood

- Maximize \( P(a|data) \rightarrow a_0 \)
- \( E[\mathcal{O}] \approx \mathcal{O}(a_0) \)
- \( V[\mathcal{O}] \approx \text{Hessian, Lagrange multipliers} \)

Monte Carlo methods

- \( P(a|data) \rightarrow \{a_k\} \)
- \( E[\mathcal{O}] \approx \frac{1}{N} \sum_k \mathcal{O}(a_k) \)
- \( V[\mathcal{O}] \approx \frac{1}{N} \sum_k [\mathcal{O}(a_k) - E[\mathcal{O}]]^2 \)
MC Methods

- MCMC or HMC
- bootstrap + cross validation + iterative convergence (JAM)
- Nested sampling (in this talk)
MC Methods

- MCMC or HMC
- bootstrap + cross validation + iterative convergence (JAM)
- **Nested sampling (in this talk)**

The idea

\[
Z = \int d^n a \; \mathcal{P}(a|\text{data})
\]

The algorithm returns \(\{a_k, w_k\}\)

\[
E[\mathcal{O}] \approx \frac{1}{N} \sum_k w_k \mathcal{O}(a_k)
\]

\[
V[\mathcal{O}] \approx \frac{1}{N} \sum_k w_k [\mathcal{O}(a_k) - E[\mathcal{O}]]^2
\]
MC Methods

- MCMC or HMC
- bootstrap + cross validation + iterative convergence (JAM)
- Nested sampling (in this talk)

The idea

\[ Z = \int d^n a \ P(a | \text{data}) \]

The algorithm returns \( \{a_k, w_k\} \)

\[ E[\mathcal{O}] \approx \frac{1}{N} \sum_k w_k \mathcal{O}(a_k) \]

\[ V[\mathcal{O}] \approx \frac{1}{N} \sum_k w_k [\mathcal{O}(a_k) - E[\mathcal{O}]]^2 \]
Details of the fitting machinery

- The code is written entirely in python (Standard in modern data analysis)

- Mellin space based DGLAP solver up to NNLO (Benchmarked against PEGASUS)

- $x$ space DY code at NLO using nCTEQ PDFs integrated within the Mellin space machinery (significant speed performance)

- Nested sampling software for the MC sampling (nestle)
The analysis

- **Data sets**
  - E615
  - NA10

- **Pion PDFs to be fitted**
  - \( q_v = \bar{u}_v = d_v \)
  - \( q_s = 2(u + \bar{d} + s) \)
  - \( g \)

- **Parametrization at \( Q_0^2 = 1 \text{GeV}^2 \)**
  - \( f(x) = N x^a (1 - x)^b \)
  - \( N_{qV} \) and \( N_g \) are fixed by sum rules
Data sets

![Graph showing data sets](image-url)
Nested sampling
Results

Fermilab – E615 PLab = 252GeV

CERN – NA10 PLab = 194GeV

CERN – NA10 PLab = 286GeV
Results

Fermilab – E615 PLab = 252GeV

CERN – NA10 PLab = 194GeV

CERN – NA10 PLab = 286GeV

$Q^2 = 1 \text{ GeV}^2$

$Q^2 = 10 \text{ GeV}^2$

$Q^2 = 100 \text{ GeV}^2$
Results

\[ 2 \int_0^1 dxxq_V(x) \]

\[ \int_0^1 dxxq_S(x) \]

\[ \int_0^1 dxxg(x) \]

<table>
<thead>
<tr>
<th></th>
<th>( Q^2 = 1 \text{ GeV}^2 )</th>
<th>( Q^2 = 10 \text{ GeV}^2 )</th>
<th>( Q^2 = 100 \text{ GeV}^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valence</td>
<td>0.566 ± 0.098</td>
<td>0.470 ± 0.081</td>
<td>0.413 ± 0.071</td>
</tr>
<tr>
<td>Sea</td>
<td>0.039 ± 0.095</td>
<td>0.036 ± 0.071</td>
<td>0.035 ± 0.058</td>
</tr>
<tr>
<td>Glue</td>
<td>0.202 ± 0.577</td>
<td>0.307 ± 0.422</td>
<td>0.356 ± 0.332</td>
</tr>
</tbody>
</table>
Outlook

■ new analysis of pions PDFs within the MC framework is on the way

■ inclusion of LN data from HERA will allow to constrain the sea PDF beyond SMRS

■ the Mellin machinery allows to extend the analysis to include threshold resummation

■ the fitting framework can be extended to perform EIC studies to test impact on pion PDFs