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 Ji[1]
[1]*North Carolina State University, Raleigh, North Carolina 27695, USA*
[2]*Jefferson Lab, Newport News, Virginia 23606, USA*
[3]*University of North Carolina, Chapel Hill, North Carolina 27599, USA*
(Received 15 December 2015; published 7 March 2016)

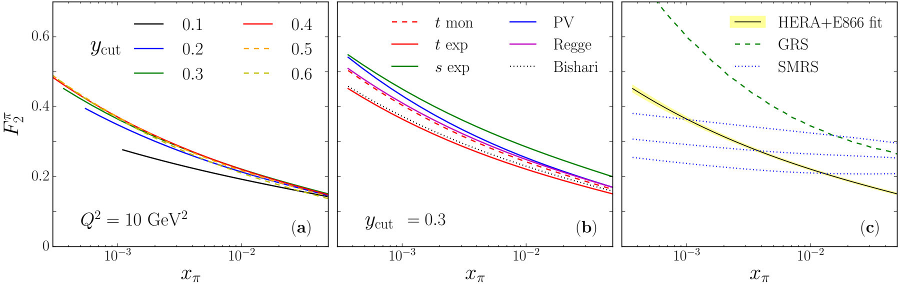
- Study pion structure using pion exchange models and pQCD
- **Unknowns**:
  - pion splitting functions (UV regulator as a parameter)
  - small x F_2^π (shape parameters)
- **Input**:
  - leading neutron cross sections from H1 and Zeus
  - $\bar{d} - \bar{u}$ asymmetry from E866
  - large x pion PDFs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

Motivations

PHYSICAL REVIEW D **93**, 054011 (2016)

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

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



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 Ji[1]
[1]*North Carolina State University, Raleigh, North Carolina 27695, USA*
[2]*Jefferson Lab, Newport News, Virginia 23606, USA*
[3]*University of North Carolina, Chapel Hill, North Carolina 27599, USA*
(Received 15 December 2015; published 7 March 2016)

- Feasibility to constrain F_2^π at small x
- Potential reduction of the uncertainty in the pion sea PDFs
- Next step: fit pion PDFs instead of F_2^π

... but first we need to learn how to fit pion PDFs at large x

Data analysis framework:

The goal is to estimate:

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

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

- $\boldsymbol{a} = (N, a, b, ...)$ is a vector of parameters

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

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

  i.e. PDFs, PPDFs, FF, cross sections

Data analysis framework:

The goal is to estimate:

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

Maximum Likelihood

- Maximize $\mathcal{P}(\boldsymbol{a}|data) \rightarrow \boldsymbol{a}_0$
- $\mathrm{E}[\mathcal{O}] \approx \mathcal{O}\left(\boldsymbol{a}_0\right)$
- $\mathrm{V}[\mathcal{O}] \approx$Hessian,
  Lagrange multipliers

Data analysis framework:

The goal is to estimate:

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

Maximum Likelihood

- Maximize $\mathcal{P}(\boldsymbol{a}|data) \rightarrow \boldsymbol{a}_0$
- $\mathrm{E}[\mathcal{O}] \approx \mathcal{O}\left(\boldsymbol{a}_0\right)$
- $\mathrm{V}[\mathcal{O}] \approx$Hessian, Lagrange multipliers

Monte Carlo methods

- $\mathcal{P}(\boldsymbol{a}|data) \rightarrow \{\boldsymbol{a}_k\}$
- $\mathrm{E}[\mathcal{O}] \approx \frac{1}{N}\sum_k \mathcal{O}(\boldsymbol{a}_k)$
- $\mathrm{V}[\mathcal{O}] \approx \frac{1}{N}\sum_k [\mathcal{O}(\boldsymbol{a}_k) - \mathrm{E}[\mathcal{O}]]^2$

Data analysis framework:

The goal is to estimate:

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

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

Maximum Likelihood

- Maximize $\mathcal{P}(\boldsymbol{a}|data) \rightarrow \boldsymbol{a}_0$
- $\mathrm{E}[\mathcal{O}] \approx \mathcal{O}\left(\boldsymbol{a}_0\right)$
- $\mathrm{V}[\mathcal{O}] \approx$Hessian, Lagrange multipliers

Monte Carlo methods

- $\mathcal{P}(\boldsymbol{a}|data) \rightarrow \{\boldsymbol{a}_k\}$
- $\mathrm{E}[\mathcal{O}] \approx \frac{1}{N} \sum_k \mathcal{O}(\boldsymbol{a}_k)$
- $\mathrm{V}[\mathcal{O}] \approx \frac{1}{N} \sum_k [\mathcal{O}(\boldsymbol{a}_k) - \mathrm{E}[\mathcal{O}]]^2$

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

$$Z = \int d^n a \;\; \mathcal{P}(\boldsymbol{a}|data)$$

The algorithm returns $\{\boldsymbol{a}_k, w_k\}$

$\mathrm{E}[\mathcal{O}] \approx \frac{1}{N} \sum_k w_k \mathcal{O}(\boldsymbol{a}_k)$

$\mathrm{V}[\mathcal{O}] \approx \frac{1}{N} \sum_k w_k [\mathcal{O}(\boldsymbol{a}_k) - \mathrm{E}[\mathcal{O}]]^2$

MC Methods

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

The idea

$$Z = \int d^n a \;\; \mathcal{P}(\boldsymbol{a}|data)$$

The algorithm returns $\{\boldsymbol{a}_k, w_k\}$

$$\mathrm{E}[\mathcal{O}] \approx \frac{1}{N} \sum_k w_k \mathcal{O}(\boldsymbol{a}_k)$$

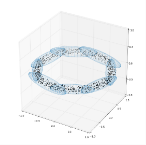
$$\mathrm{V}[\mathcal{O}] \approx \frac{1}{N} \sum_k w_k [\mathcal{O}(\boldsymbol{a}_k) - \mathrm{E}[\mathcal{O}]]^2$$

Nestle

/ˈnesəl/ *(rhymes with "wrestle")*

Pure Python, MIT-licensed implementation of nested sampling algorithms.

Nested Sampling is a computational approach for integrating posterior probability in order to compare models in Bayesian statistics. It is similar to Markov Chain Monte Carlo (MCMC) in that it generates samples that can be used to estimate the posterior probability distribution. Unlike MCMC, the nature of the sampling also allows one to calculate the integral of the distribution. It also happens to be a pretty good method for robustly finding global maxima.



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)

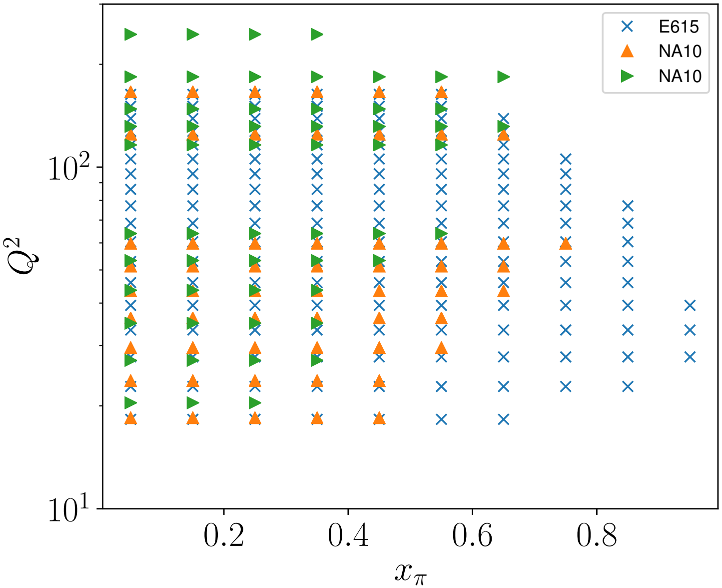
The analysis

- **Data sets**
  - E615
  - NA10

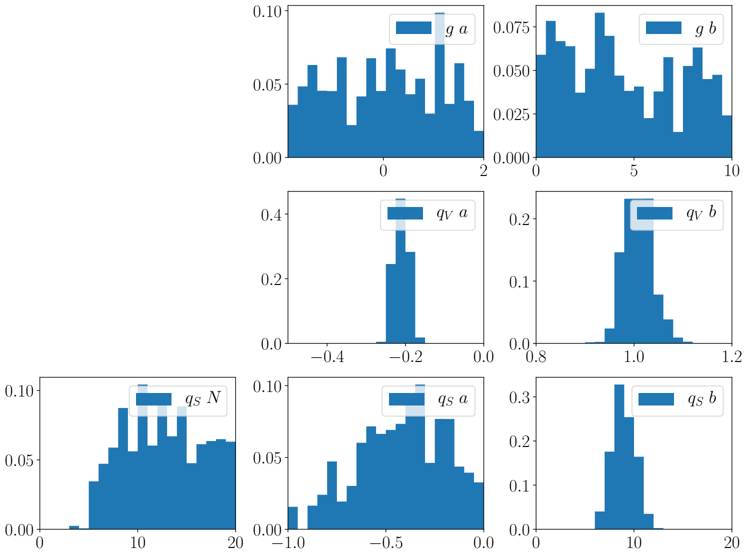
- **pion PDFs to be fitted**
  - $q_v = \bar{u}_v = d_v$
  - $q_s = 2(u + \bar{d} + s)$
  - g

- **parametrization at** $Q_0^2 = 1\mathrm{GeV}^2$
  - $f(x) = Nx^a(1-x)^b$
  - N_{q_V} and N_g are fixed by sum rules

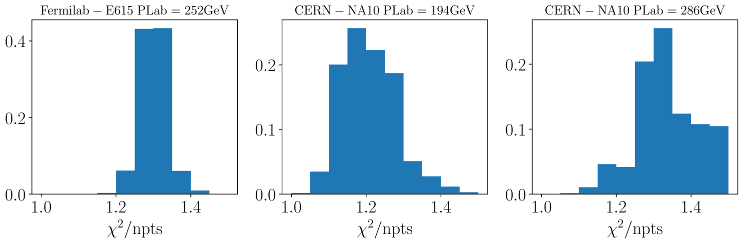
Data sets



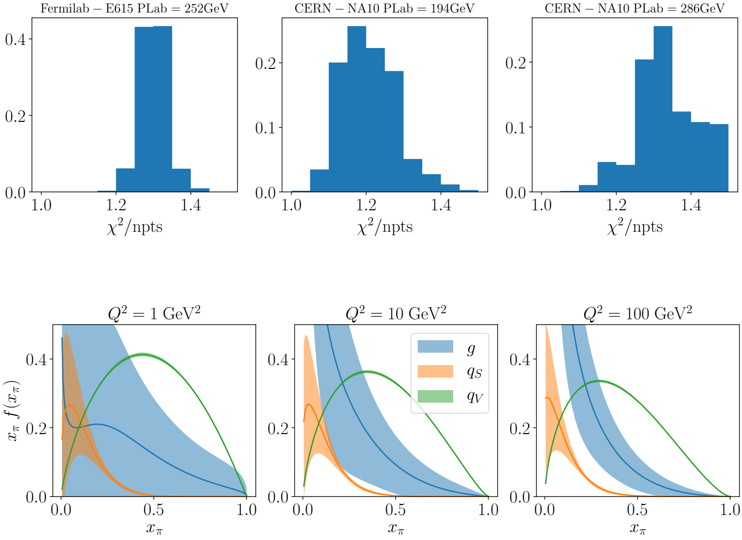
Nested sampling



Results

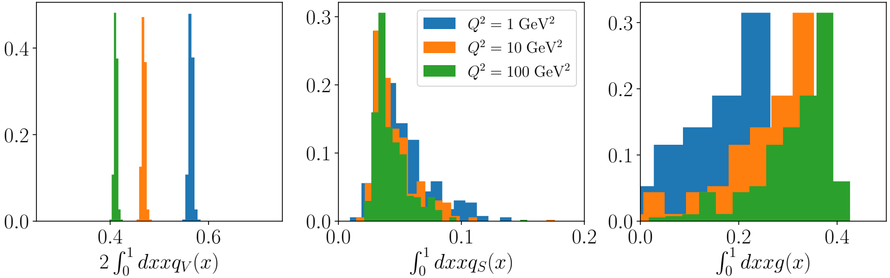

Fermilab − E615 PLab = 252GeV
CERN − NA10 PLab = 194GeV
CERN − NA10 PLab = 286GeV
0.0
0.2
0.4
0.1
1.0
1.2
1.4
χ^2/npts
χ^2/npts
χ^2/npts

Results



Fermilab – E615 PLab = 252GeV
CERN – NA10 PLab = 194GeV
CERN – NA10 PLab = 286GeV
χ^2/npts
$Q^2 = 1$ GeV^2
$Q^2 = 10$ GeV^2
$Q^2 = 100$ GeV^2
g
q_S
q_V
$x_\pi\ f(x_\pi)$
x_π

Results



–	$Q^2 = 1\ \mathrm{GeV}^2$	$Q^2 = 10\ \mathrm{GeV}^2$	$Q^2 = 100\ \mathrm{GeV}^2$
Valence	0.566 ± 0.098	0.470 ± 0.081	0.413 ± 0.071
Sea	0.039 ± 0.095	0.036 ± 0.071	0.035 ± 0.058
Glue	0.202 ± 0.577	0.307 ± 0.422	0.356 ± 0.332

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