First global Monte Carlo analysis of pion PDFs

In Collaboration with:
P. Barry (NCSU), W. Melnitchouk (JLab), C.R. Ji (NCSU)

Nobuo Sato
TDIS collaboration meeting,
Jefferson Lab
February 22, 2018
Motivations

- **Parton distribution function in pions (and kaons)**
  + Much less known than proton PDFs
  + Easier to compute in lattice QCD and models than proton due to its simpler $q, \bar{q}$ structure
  + Predictions for $x \to 1$ behavior

- **Light quark asymmetry in the proton:** $\bar{d} > \bar{u}$
  + A nonperturbative phenomena i.e pQCD $\to \bar{d} \approx \bar{u}$
  + Origin of the asymmetry is still a puzzle.
  + E866 sign change at large $x$? $\to$ to be confirmed by Sea Quest
  + Possible explanations:
    - Chiral symmetry breaking $\to$ unique signature of pion loops
    - Pauli exclusion principle
Conway et al. (1989)

- Data: DY
- Fits of valence with some assumptions for the sea
- LO analysis (K factors)
- Approximate DGLAP
- Single fits and no errors
History

+ Data: DY, prompt photons
+ Fits of valence and glue with assumptions for the sea
+ NLO analysis
+ Single fits and no errors

Sutton, Martin, Roberts, Strirling (1992)
History

- Data: DY, prompt photons
- Fits of valence, sea and glue
- NLO analysis
- Single fits and no errors
- Inconsistent with SMRS

Glück, Reya, Schienbein (1999)
Various model calculations exist and they predict different large $x_\pi$ for the valence distribution compared to global analysis.

Holt, Roberts (2010)
Aicher, Schafer, Vogelsang (2010)

+ Data: DY
+ Fits of valence using GRS
  - sea and glue
+ NLO+NLL analysis
+ Single fits and no errors
+ The shape of valence agrees with DSE
The graphs show the behavior of $F^{\pi}_2$ as a function of $y_{\text{cut}}$ and $Q^2$ for different $x_{\pi}$ values.

**McKenney, NS, Melnitchouk, Ji (2015)**

+ Data: Leading Neutron (HERA) (+ E866)
+ Fits of $F^{\pi}_2$
+ Approximate DGLAP
+ Feasibility study for a global analysis using LN data
How to probe pion structure

- \( \pi + A \rightarrow l\bar{l} + X \) (Drell-Yan)
  - Measurements at Fermilab and CERN
  - Constrains the medium to large \( x_\pi \)
  - Will be measured at COMPASS

- \( \pi + A \rightarrow \gamma + X \) (prompt photons)
  - Measurements at Fermilab
  - Constraints on the gluon distribution at large \( x_\pi \)
  - Requires knowledge of photon fragmentation function
  - Will be measured at COMPASS

- \( e + p \rightarrow e' + n + X \) (SIDIS)
  - Leading neutron (LN) measurements at HERA
    - Target fragmentation region
  - Constraints on small \( x_\pi \)
  - Process dominated by one pion exchange at forward angles

- ... and TDIS (JLab12, EIC)
One pion exchange (OPE)

- LN at HERA
  + SIDIS in the target fragmentation region
  + Events from neutrons in the forward region
  + The reaction is dominated by single pion exchange i.e.
    \[(p \rightarrow \pi + n) \rightarrow (e + \pi \rightarrow e' + X)\]
  + In the collinear configuration the pion is nearly on-shell
  + \((p \rightarrow \pi + n)\) is described in \(\chi\)EFT
  + \((e + \pi \rightarrow e' + X)\) is described in pQCD \(\rightarrow\) DIS

\[
\begin{align*}
p & \rightarrow \pi + n \\
\rightarrow (e + \pi \rightarrow e' + X)
\end{align*}
\]
One pion exchange (OPE)

$$\frac{d\sigma}{dx dQ^2 dy} \sim f_{p \to \pi + n} (y) \times \sum_q \int_{x/y}^1 \frac{d\xi}{\xi} C(\xi) q \left( \frac{x/y}{\xi}, Q^2 \right)$$
The splitting function \( y = k^+/p^+ = x/x_\pi \)

\[
f_{p \to \pi + n}(y) = \frac{g_A^2 M^2}{(4\pi f_\pi)^2} \int dk^2_\perp \frac{y(k^2_\perp + y^2 M^2)}{(1 - y)^2 D^2_{\pi N}} |F|^2
\]

UV regulators used in the literature

\[
F = \begin{cases} 
[1 - \frac{(t-m^2_\pi)^2}{(t-\Lambda^2)^2}]^{1/2} & \text{Pauli-Villars} \\
(\Lambda^2 - m^2_\pi)/(\Lambda^2 - t) & \text{t-dependent monopole} \\
\exp[(t - m^2_\pi)/\Lambda^2] & \text{t-dependent exponential} \\
\exp[(M^2 - s)/\Lambda^2] & \text{s-dependent exponential} \\
y^{-\alpha}(t) \exp[(t - m^2_\pi)/\Lambda^2] & \text{Regge exponential,}
\end{cases}
\]

where

+ \( g_A = 1.267 \) nucleon axial charge
+ \( f_\pi = 93\text{MeV} \) pion decay constant
+ \( D_{\pi N} \equiv t - m^2_\pi = -\frac{1}{1-y}[k^2_\perp + y^2 M^2 + (1 - y)m^2_\pi] \)
pQCD setup

- $\pi^- + W \rightarrow l\bar{l} + X$ (Drell-Yan)

$$\frac{d\sigma}{dx_F dQ^2} = \sum_{a,b} \int d\xi d\zeta C_{a,b}(\xi, \zeta) f_{a/\pi^-}(\xi) f_{b/W}(\zeta)$$

- $e + p \rightarrow e' + n + X$ (LN)

$$\frac{d\sigma}{dx dQ^2 dy} \sim f_{p\rightarrow\pi+n}(y) \times \sum_q \int_{x/y}^{1} \frac{d\xi}{\xi} C(\xi) f_{q/\pi^+} \left( \frac{x/y}{\xi} \right)$$

- The hard coeffs are computed at NLO in pQCD
- We parametrize PDFs in $\pi^-$
  - Valence: $\bar{u}_v = d_v$
  - Sea: $u = d = s = \bar{s}$
  - Gluon: $g$
- For DY we use $W$ PDFs from EPS16
Likelihood analysis setup

- **likelihood analysis using Bayesian stat.**
  
  **Bayes theorem:**
  
  \[ P(f|data) = \frac{1}{Z} L(data|f) \pi(f) \]

  **The likelihood function** *Gaussian likelihood*
  
  \[ L(data|f) = \exp \left[ -\frac{1}{2} \sum_i \left( \frac{d_i - \text{model}_i(f)}{\delta d_i} \right)^2 \right] \]

  **The prior function** to restrict unphysical regions of \( f \). e.g.
  
  \[ \pi(f) = \begin{cases} 
    1 & \text{condition}(f) == \text{True} \\
    0 & \text{condition}(f) == \text{False} 
  \end{cases} \]
Likelihood analysis setup

- In practice $f$ needs to be parametrized e.g

\[
    f(x) = N x^a (1 - x)^b (1 + c \sqrt{x} + dx + \ldots)
\]

\[
    f(x) = N x^a (1 - x)^b \text{NN}(x; \{w_i\})
\]

\[
    f(x) = \text{NN}(x; \{w_i\}) - \text{NN}(1; \{w_i\})
\]

- The pdf for $f$ becomes

$$\mathbf{a} = (N, a, b, c, d, \ldots)$$

$$\mathcal{P}(\mathbf{a}|\mathbf{d}) = \frac{1}{Z} \mathcal{L}(\mathbf{d}|\mathbf{a})\pi(\mathbf{a})$$

$$\mathcal{L}(\mathbf{d}|\mathbf{a}) = \exp \left[ -\frac{1}{2} \sum_i \left( \frac{d_i - \text{model}_i(\mathbf{a})}{\delta d_i} \right)^2 \right]$$

$$\pi(\mathbf{a}) = \prod_i \theta(a_i - a_i^{\text{min}})\theta(a_i^{\text{max}} - a_i)$$

\[
    \mathcal{P}(f|\mathbf{d}) = \frac{1}{Z} \mathcal{L}(\mathbf{d}|f)\pi(f)
\]

\[
    \mathcal{P}(\mathbf{a}|\mathbf{d}) = \frac{1}{Z} \mathcal{L}(\mathbf{d}|\mathbf{a})\pi(\mathbf{a})
\]
Likelihood analysis setup

- Having the pdf for $f$ we can compute

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

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

- $\mathcal{O}$ is any function of $a$. e.g

$$\mathcal{O}(a) = f(x; a)$$

$$\mathcal{O}(a) = \int_x^1 \frac{d\xi}{\xi} C(\xi) f \left( \frac{x}{\xi}; a \right)$$

- How do we compute $E[\mathcal{O}]$, $V[\mathcal{O}]$?
  - Maximum likelihood
  - (Hessian, Lagrange multipliers)
  - Monte Carlo sampling
The basic idea: compute

\[ Z = \int L(\text{data}|a)\pi(a)d^n a = \int_0^1 L(X)dX \]

→ The algorithm traverses ordered isolikelihood contours in the variable \( X \) such that \( X \) follows the progression \( X_i = t_i X_{i-1} \)
→ The variable \( t_i \) is estimated statistically
→ The algorithm can be optimized iteration to iteration. One can sample only in the regions where the likelihood is larger → “importance sampling”
→ The nested sampling is parallelizable
Our new analysis extends previous pion PDF analysis down to $x \sim 0.001$

The OPE+pQCD can describe the HERA data simultaneously with the DY data

$$F_2^{LN(3)}(x, Q^2, y) = 2 f_{p \rightarrow \pi+n}(y) F_2^{\pi}(x_\pi, Q^2)$$

$$r(x, Q^2, y) = \frac{d^3\sigma^{LN}/dx\!dQ^2\!dy}{d^2\sigma^{inc}/dx\!dQ^2\!\Delta y}$$
Results

- Significant reduction of the uncertainties on the glue and sea at small $x_\pi$
- Non-overlapping uncertainties show some tensions between the LN and DY data. Future TDIS (JLab12/EIC) is needed to establish a more reliable pion PDFs in the intermediate $x_\pi$ region.
Results

Constraints from HERA significantly increase $\langle x_g^g \rangle$. The role of the glue is more important than suggested by DY alone.

In contrast, the strength of the sea is reduced.

Due to momentum sum rule $\langle x_{\pi}^{\text{valence}} \rangle$ decreases.
Results

- Uncertainty from model dependence of DY+LN is comparable to current experimental uncertainties.
- DY+LN cannot discriminate among the OPE models since the LN data is only sensitive to region where $|F| \approx 1$ for all models.
Since LN+DY alone is not capable to discriminate among the OPE models the predictions for $\bar{d} - \bar{u}$ are not accurate

We performed an additional analysis of LN+DY+E866  
→ good description of E866 data except for large $x$
Inclusion of E866 data does not change the pion PDFs significantly

The model dependence increases in order to simultaneously describe the E866 data

The results support the role of pion clouds as the origin of the sea asymmetry in the nucleon
Relevance of TDIS measurements

- Constraints at medium to large $x_\pi$
- Resolve tensions between LN and DY data → to improve accuracy of pion PDFs
- Constraints on kaon PDFs → Virtually no data
Summary and outlook

- **New analysis of pion PDFs**
  + New combined analysis of pion PDFs from small to large $x_\pi$ using LN+DY data
  + Momentum fractions carried by gluons are larger than in DY only analysis
  + To describe the LN data, momentum fraction of the valence distribution decreases
  + LN+DY alone at present cannot discriminate between the OPE models
  + Inclusion of E866 data does not change the pion PDFs significantly
  + The results support the role of pion cloud as the origin of the sea flavor asymmetry

- **Role of upcoming TDIS data**
  + Resolve some tensions between LN and DY
  + The data will improve the accuracy of the pion PDFs in the valence region