First global Monte Carlo analysis of pion PDFs

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

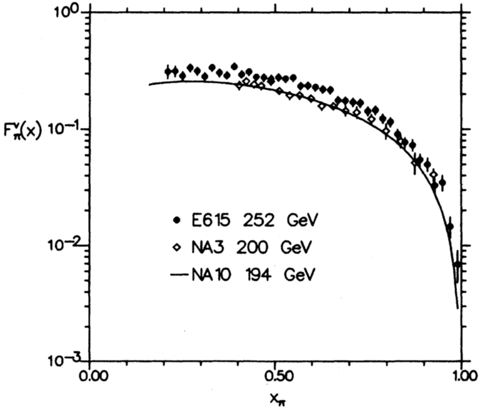
Nobuo Sato
TDIS collaboration meeting,
Jefferson Lab
February 22, 2018

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

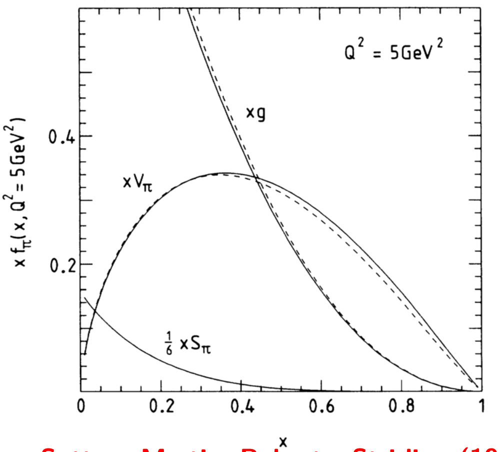
History



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

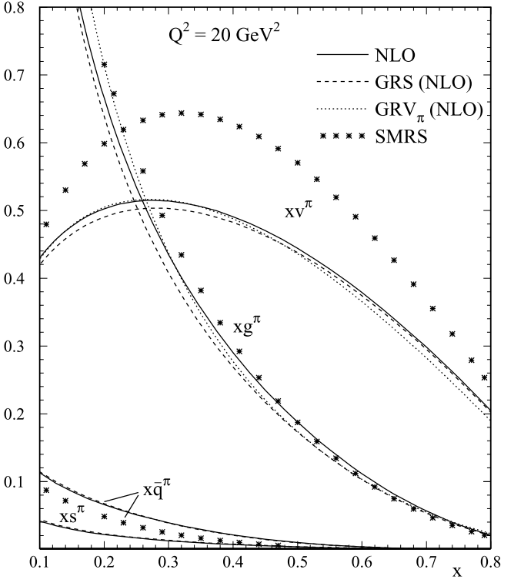
History



Sutton, Martin, Roberts, Stirling (1992)

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

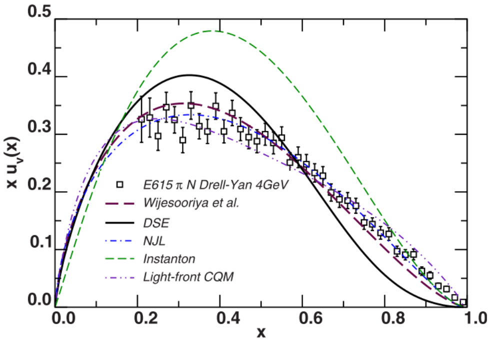
History



Gluck, Reya, Schienbein (1999)

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

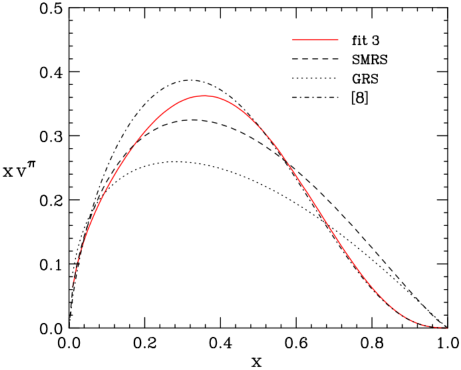
History



Holt, Roberts (2010)

- Various model calculations exist and they predict different large x_π for the valence distribution compared to global analysis

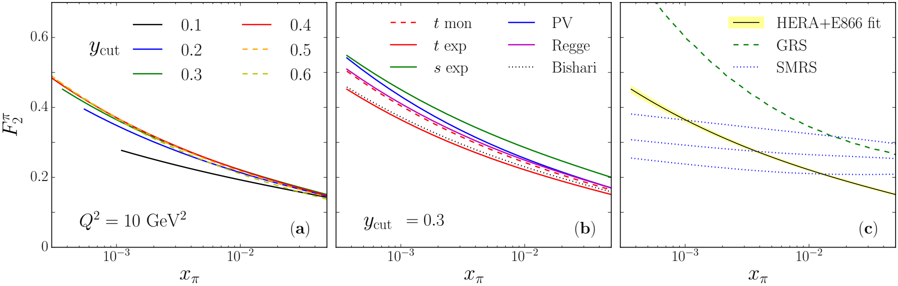
History



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

History



McKenney, NS, Melnitchouk, Ji (2015)

+ Data: Leading Neutron (HERA) (+ E866)
+ Fits of F_2^π
+ Approximate DGLAP
+ Feasibility study for a global analysis using LN data

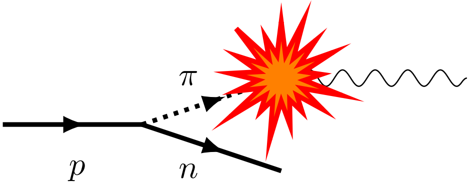
How to probe pion structure

- $\pi + A \to l\bar{l} + X$ (Drell-Yan)
  - Measurements at Fermilab and CERN
  - Constrains the medium to large x_π
  - Will be measured at COMPASS
- $\pi + A \to \gamma + X$ (prompt photons)
  - Measurements at Fermilab
  - Constraints on the gluon distribution at large x_π
  - Requires knowdlege of photon fragmentation function
  - Will be measured at COMPASS
- $e + p \to e' + n + X$ (SIDIS)
  - Leading neutron (LN) measurements at HERA
    $\to$ target fragmentation region
  - Constraints on small x_π
  - Process dominated by one pion exchange at forward angles
- ... and TDIS (JLab12, EIC)

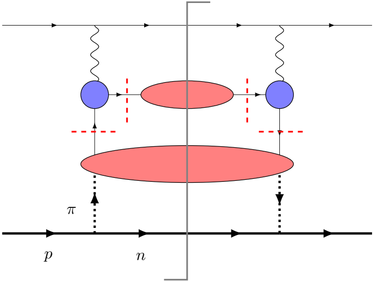
In this analysis
$\to$ only DY and SIDIS

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. $(p \to \pi + n) \to (e + \pi \to e' + X)$
  - In the collinear configuration the pion is nearly on-shell
  - $(p \to \pi + n)$ is described in χEFT
  - $(e + \pi \to e' + X)$ is described in pQCD $\to$ DIS

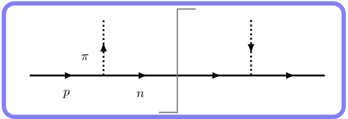


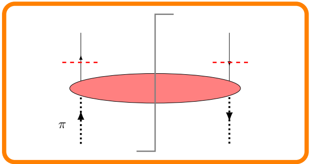
One pion exchange (OPE)



pQCD

χEFT





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

χEFT setup

- The splitting function ($y = k^+/p^+ = x/x_\pi$)

$$f_{p\to\pi+n}(y) = \frac{g_A^2 M^2}{(4\pi f_\pi)^2} \int dk_\perp^2 \frac{y(k_\perp^2 + y^2 M^2)}{(1-y)^2 D_{\pi N}^2} |F|^2$$

- UV regulators used in the literature

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

- where
  - $g_A = 1.267$ nucleon axial charge
  - $f_\pi = 93\mathrm{MeV}$ pion decay constant
  - $D_{\pi N} \equiv t - m_\pi^2 = -\frac{1}{1-y}[k_\perp^2 + y^2 M^2 + (1-y) m_\pi^2]$

pQCD setup

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

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

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

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

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

Likelihood analysis setup

- **likelihood analysis using Bayesian stat.**
+ Bayes theorem:

$$\mathcal{P}(f|data) = \frac{1}{Z}\mathcal{L}(data|f)\pi(f)$$

+ The likelihood function *Gaussian likelihood*

$$\mathcal{L}(data|f) = \exp\left[-\frac{1}{2}\sum_i \left(\frac{d_i - \mathrm{model}_i(f)}{\delta d_i}\right)^2\right]$$

+ The prior function to restrict unphysical regions of f. e.g.

$$\pi(f) = \begin{cases} 1 & \mathrm{condition}(f) == \mathrm{True} \\ 0 & \mathrm{condition}(f) == \mathrm{False} \end{cases}$$

Likelihood analysis setup

- In practice f needs to be parametrized e.g

$$f(x) = Nx^a(1-x)^b(1+c\sqrt{x}+dx+...)$$
$$f(x) = Nx^a(1-x)^b\mathrm{NN}(x;\{w_i\})$$
$$f(x) = \mathrm{NN}(x;\{w_i\}) - \mathrm{NN}(1;\{w_i\})$$

- The pdf for f becomes

$$\boldsymbol{a} = (N,a,b,c,d,...)$$
$$\mathcal{P}(\boldsymbol{a}|d) = \frac{1}{Z}\mathcal{L}(d|\boldsymbol{a})\pi(\boldsymbol{a})$$
$$\mathcal{L}(d|\boldsymbol{a}) = \exp\left[-\frac{1}{2}\sum_i\left(\frac{d_i - \mathrm{model}_i(\boldsymbol{a})}{\delta d_i}\right)^2\right]$$
$$\pi(\boldsymbol{a}) = \prod_i \theta(a_i - a_i^{min})\theta(a_i^{max} - a_i)$$

$$\boxed{\begin{array}{c} \mathcal{P}(f|d) = \frac{1}{Z}\mathcal{L}(d|f)\pi(f) \\ \downarrow \\ \mathcal{P}(\boldsymbol{a}|d) = \frac{1}{Z}\mathcal{L}(d|\boldsymbol{a})\pi(\boldsymbol{a}) \end{array}}$$

Likelihood analysis setup

- Having the pdf for f we can compute

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

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

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

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

Nested resampling

- **The basic idea**: compute

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

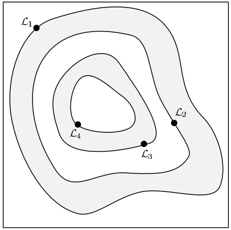
→ The algorithm traverses ordered isolikelihood contours in the variable X such that X follows the progression $X_i = t_i X_{i-1}$

→ The variable t_i is estimated statistically

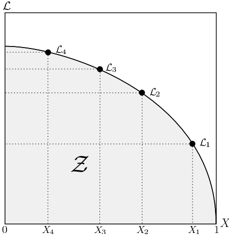
→ The algorithm can be optimized iteration to iteration. One can sample only in the regions where the likelihood is larger → "importance sampling"

→ The nested sampling is parallelizable

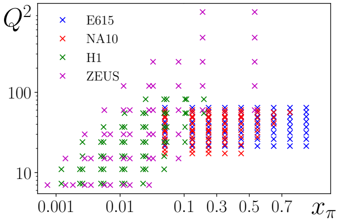
$\mathcal{L}(\text{data}|\boldsymbol{a})$ in $\boldsymbol{a}$ space

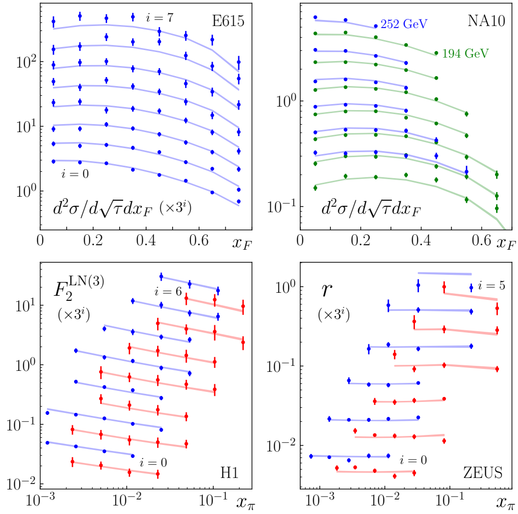


$\mathcal{L}(X)$ in X space



Results





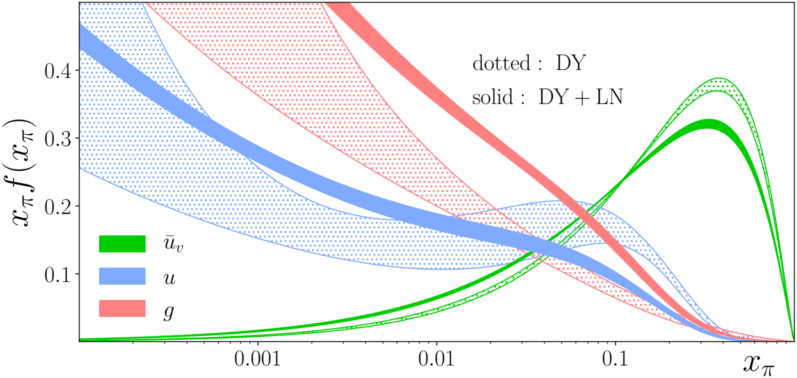
+ Our new analysis extends previous pion PDF analysis down to $x \sim 0.001$

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

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

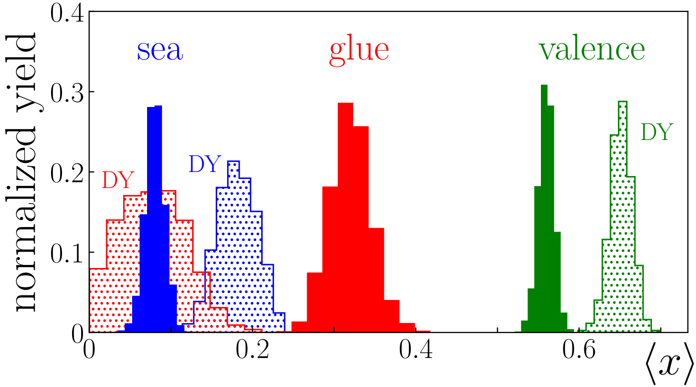
$$r(x, Q^2, y) = \frac{d^3\sigma^{LN}/dxdQ^2dy}{d^2\sigma^{inc}/dxdQ^2}\Delta y$$

Results



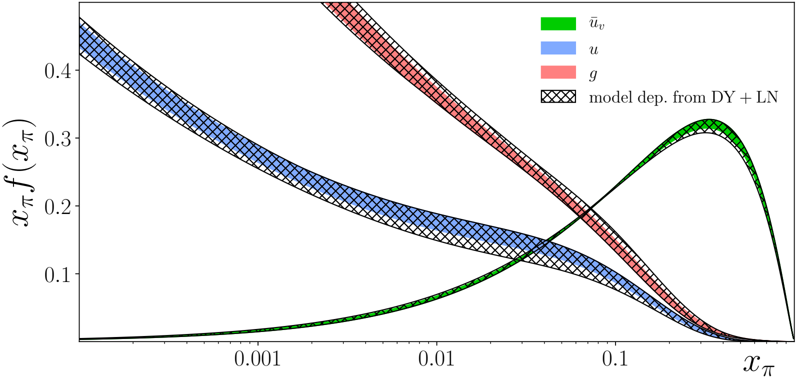
- Significant reduction of the uncertainties on the glue and sea at small x_π
- Non-overlapping uncertainties show some tensions between the LN and DY data. Future TDIS (JLab12/EIC) is needed to establish a more reliable pion PDFs in the intermediate x_π region.

Results



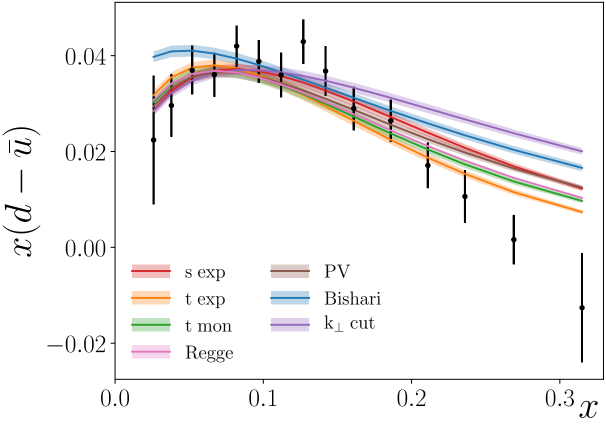
- Constraints from HERA significantly increase $\langle x_\pi^g \rangle$.
  The role of the glue is more important than suggested by DY alone
- In contrast, the strength of the sea is reduced
- Due to momentum sum rule $\left\langle x_\pi^{\text{valence}} \right\rangle$ decreases

Results



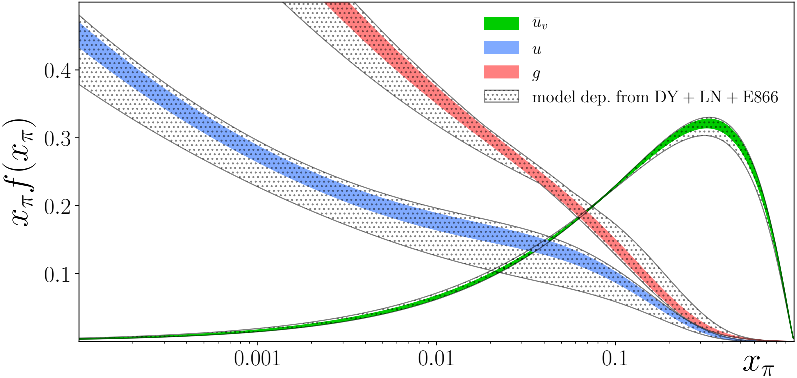
- Uncertainty from model dependence of DY+LN is comparable to current experimental uncertainties
- DY+LN cannot discriminate among the OPE models since the LN data is only sensitive to region where $|F| \approx 1$ for all models

Results



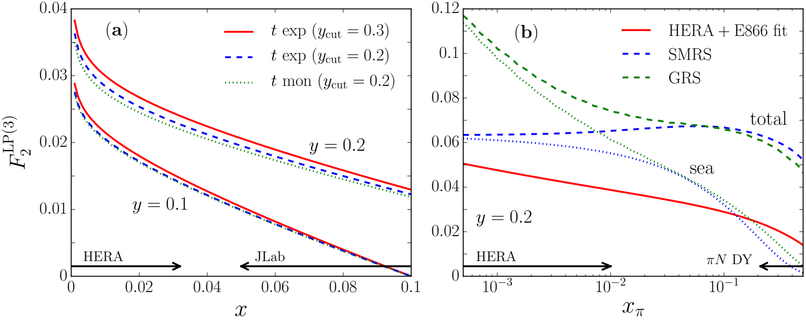
- Since LN+DY alone is not capable to discriminate among the OPE models the predictions for $\bar{d}-\bar{u}$ are not accurate
- We performed an additional analysis of LN+DY+E866
  $\rightarrow$ good description of E866 data except for large x

Results



- 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_π
- Resolve tensions between LN and DY data
  $\rightarrow$ to improve accuracy of pion PDFs
- Constraints on kaon PDFs $\rightarrow$ Virtually no data

Summary and outlook

- **New analysis of pion PDFs**
  - New combined analysis of pion PDFs from small to large x_π 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