Parton distributions in hadrons: nucleon and pion

Wally Melnitchouk
Outline

- **Aim**: understand internal quark-gluon structure of hadrons
- **Method**: extract parton distribution functions (PDFs) from global QCD analysis, using new Monte Carlo-based methods

**Recent highlights:**
- Constraints from Fermilab & JLab data on *unpolarized* PDFs at high $x$
- First combined analysis of *polarized* DIS + SIDIS + SIA data, with *simultaneous* extraction of PDFs & fragmentation functions
- First MC analysis of nucleon’s *transversity* PDFs + lattice QCD
- First extraction of *pion* PDFs from Drell-Yan and HERA leading neutron production data
Parton distributions in hadrons

Generic process: inclusive particle production $AB \to CX$

\[
\sigma_{AB \to CX}(p_A, p_B) = \sum_{a,b} \int dx_a \: dx_b \: f_{a/A}(x_a, \mu) \: f_{b/B}(x_b, \mu)
\]

“factorization”

\[
\times \sum_n \alpha_s^n(\mu) \: \hat{\sigma}_{ab \to CX}^{(n)}(x_a p_A, x_b p_B, Q/\mu)
\]

universal functions $f_{a/A}$ characterize internal structure of bound state $A$

Collins, Soper, Sterman (1980s)
Parton distributions in hadrons

Most information on parton distribution functions obtained from inclusive deep-inelastic scattering (DIS)

\[ \frac{d^2 \sigma}{d\Omega dE'} = \frac{4\alpha^2 E'^2 \cos^2 \frac{\theta}{2}}{Q^4} \left( 2 \tan^2 \frac{\theta}{2} \frac{F_1}{2M} + \frac{F_2}{\nu} \right) \]

\[ \nu = E - E' \]
\[ Q^2 = \bar{q}^2 - \nu^2 \]
\[ W^2 = M^2 + Q^2 \left( 1 - \frac{x}{x} \right) \]
\[ x = \frac{Q^2}{2M\nu} \]

At leading order (LO) in pQCD, structure functions given in terms of charge-weighted sums of PDFs

\[ F_2(x, Q^2) = x \sum_q e_q^2 \, q(x, Q^2) \]
Parton distributions in hadrons

Precision PDFs needed to
(1) understand basic structure of QCD bound states
(2) compute backgrounds in searches for BSM physics

\[ Q^2 \text{ evolution feeds} \]
low \( x \), high \( Q^2 \) (“LHC”) 
from high \( x \), low \( Q^2 \) (“JLab”)

Information on PDFs obtained from
(1) nonperturbative approaches (low-energy models, DSE, \( \chi \)EFT)
(2) lattice QCD
(3) global QCD analysis
Parton distributions in hadrons

Compute matrix element of leading twist operator at a low scale $\mu$; evolve PDF to higher $Q^2$ using DGLAP

$$q(x, \mu) = \frac{M}{(2\pi)^3} \sum_n \int d^3p_n |\langle n|\gamma^- \gamma^+ \psi(0)|N \rangle|^2 \delta \left( (1 - x)M - p_n^+ \right)$$

Jaffe (1975, 1983)
Parisi, Petronzio (1976)

$\rightarrow$ reasonable phenomenology, but evolution (questionable at low scales $\sim 0.1-0.5 \text{ GeV}^2$) “washes out” many differences between models
Parton distributions in hadrons

More recently, significant progress made towards computing $x$ dependence of PDFs directly on lattice

Lin et al. (2015)

$X. Ji (2013); \; Ma, Qiu (2014)$

Radyushkin (2017)

MSTW

CJ12

Lattice

\[
q(x, \mu) = \int_{-\infty}^{\infty} \frac{d\xi}{|\xi|} C(\xi, \mu/xP) \tilde{q}(x/\xi, \mu, Pz)
\]

Lin et al. (2015)

\[
\tilde{q}(x, P) = \int \frac{dz}{4\pi} e^{-ixPz} \langle P|\bar{\psi}(0, z)\gamma_z W(z) \psi(0, 0)|P \rangle
\]

“quasi-PDF”
Global PDF analysis

Universality of PDFs allows data from different processes (DIS, SIDIS, jet production, Drell-Yan …) to be analyzed simultaneously.

Several dedicated global efforts to extract PDFs using factorization theorems + pQCD at a given order in $\alpha_s$:

- CTEQ, MRS/MMHT, HERAPDF, DSSV, … use standard maximum likelihood methods ($\chi^2$ minimization)
- NNPDF, JAM use Monte Carlo methods (neural networks, nested sampling)

Typically PDF parametrizations are nonlinear functions of PDF parameters, e.g. $x f(x, \mu) = N x^\alpha (1 - x)^\beta P(x)$

where $P$ is a polynomial, neural net, …

- multiple local minima present in the $\chi^2$ function
- thoroughly scan over sufficiently large parameter space
Global PDF analysis

A major challenge has been to characterize PDF uncertainties, especially in the presence of tensions among data sets.

Previous attempts sought to address tensions in data sets by introducing:

- “tolerance” factors (artificially inflating PDF errors)
- “neural net” parametrization (instead of polynomial parametrization), together with MC techniques

However, to address the problem in a more statistically rigorous way, one requires going beyond the standard $\chi^2$ minimization paradigm.

- utilize modern techniques based on Bayesian statistics
Bayesian approach to global analysis

Analysis of data requires estimating expectation values $E$ and variances $V$ of “observables” $\mathcal{O}$ (functions of PDFs) which are functions of parameters

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

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

“Bayesian master formulas"

Using Bayes’ theorem, probability distribution $\mathcal{P}$ given by

\[
\mathcal{P}(\tilde{a} | \text{data}) = \frac{1}{Z} \mathcal{L}(\text{data} | \tilde{a}) \pi(\tilde{a})
\]

in terms of the likelihood function $\mathcal{L}$
Bayesian approach to global analysis

**Likelihood function**

\[
\mathcal{L}(\text{data}|\vec{a}) = \exp \left( -\frac{1}{2} \chi^2(\vec{a}) \right)
\]

is a Gaussian form in the data, with \( \chi^2 \) function

\[
\chi^2(\vec{a}) = \sum_i \left( \frac{\text{data}_i - \text{theory}_i(\vec{a})}{\delta(\text{data})} \right)^2
\]

with priors \( \pi(\vec{a}) \) and “evidence” \( Z \)

\[
Z = \int d^n a \, \mathcal{L}(\text{data}|\vec{a}) \, \pi(\vec{a})
\]

→ \( Z \) tests if e.g. an \( n \)-parameter fit is statistically different from \( (n+1) \)-parameter fit
Bayesian approach to global analysis

Standard method for evaluating $E$, $V$ via maximum likelihood

$\rightarrow$ maximize probability distribution

\[ P(\tilde{a}|\text{data}) \rightarrow \tilde{a}_0 \]

$\rightarrow$ if $\mathcal{O}$ is linear in parameters, and if probability is symmetric in all parameters

\[ E[\mathcal{O}(\tilde{a})] = \mathcal{O}(\tilde{a}_0), \quad V[\mathcal{O}(\tilde{a})] \rightarrow \text{Hessian} \quad H_{ij} = \frac{1}{2} \frac{\partial^2 \chi^2(\tilde{a})}{\partial a_i \partial a_j} \bigg|_{\tilde{a} = \tilde{a}_0} \]

In practice, since in general $E[f(\tilde{a})] \neq f(E[\tilde{a}])$, maximum likelihood method often fails

$\rightarrow$ need more robust (Monte Carlo) approach

\[ E[\mathcal{O}] \approx \frac{1}{N} \sum_k \mathcal{O}(\tilde{a}_k), \quad V[\mathcal{O}] \approx \frac{1}{N} \sum_k \left[ \mathcal{O}(\tilde{a}_k) - E[\mathcal{O}] \right]^2 \]
Monte Carlo methods

First group to use MC for global PDF analysis was NNPDF, using neural network to parametrize $P(x)$ in

$$f(x) = N x^\alpha (1 - x)^\beta P(x)$$

— $\alpha, \beta$ are fitted “preprocessing coefficients”

Iterative Monte Carlo (IMC), developed by JAM Collaboration, variant of NNPDF, tailored to non-neutral net parametrizations

$N. Sato et al. (2016)$

Markov Chain MC (MCMC) / Hybrid MC (HMC)
— recent “proof of principle” analysis, ideas from lattice QCD

$Gbedo, Mangin-Brinet (2017)$

Nested sampling (NS) — computes integrals in Bayesian master formulas (for $E$, $V$, $Z$) explicitly

$Skilling (2004)$
Unpolarized Nucleon PDFs
Unpolarized PDFs

Ubiquity of proton $F_2$ data (SLAC, BCDMS, NMC, HERA, JLab, …) provides strong constraints on $u$-quark PDF over large $x$ range.

Absence of free-neutron data and smaller $|e_q|$ of $d$ quarks limit precision of $d$-quark PDF, especially at high $x$.

→ nuclear effects in deuterium obscure free-neutron structure.
Unpolarized PDFs

Valence $d/u$ ratio at high $x$ of particular interest

→ testing ground for nucleon models in $x \to 1$ limit

- $d/u \to 1/2$
  SU(6) symmetry

- $d/u \to 0$
  $S = 0$ $qq$ dominance
  (color-hyperfine interaction)

- $d/u \to 1/5$
  $S_z = 0$ $qq$ dominance
  (perturbative gluon exchange)

- $d/u \to 0.18 - 0.28$
  DSE with $qq$ correlations

→ considerable uncertainty at high $x$ from deuterium corrections (no free neutrons!)

\[
F_2^d(x, Q^2) = \int_x dy \ f(y, \gamma) \ F_2^N(x/y, Q^2)
\]

\[
f(y, \gamma) = \int \frac{d^3p}{(2\pi)^3} |\psi_d(p)|^2 \delta\left(y - 1 - \frac{\gamma p_z}{M}\right) \times \frac{1}{\gamma^2} \left[1 + \frac{\gamma^2}{y^2} \left(1 + \frac{2\gamma}{M} + \frac{\gamma^2}{2M^2} (1 - 3\gamma^2)\right)\right]
\]
Unpolarized PDFs

Valence $d/u$ ratio at high $x$ of particular interest

→ significant reduction of PDF errors with new JLab tagged neutron & FNAL $W$-asymmetry data

$$d + \bar{u} \rightarrow W^- \rightarrow \ell^- + \bar{\nu}$$

→ extrapolated ratio at $x = 1$

$$d/u \rightarrow 0.09 \pm 0.03$$

does not match any model…

→ upcoming experiments at JLab (MARATHON, BONuS, SoLID) will determine $d/u$ up to $x \sim 0.85$
Nucleon Helicity PDFs
Proton spin structure

Question of how proton spin decomposed into its $q$ & $g$ constituents has engrossed community for $>30$ years

→ in nonrelativistic quark model, spin of proton is carried entirely by quarks
\[
\Delta \Sigma = \Delta u^+ + \Delta d^+ + \Delta s^+ = 1
\]
while early data suggested that
\[
\Delta \Sigma \approx 0 \quad \Delta s^+ \approx -(0.1 - 0.2)
\]

→ proton spin sum requires
\[
\frac{1}{2} = \frac{1}{2} \Delta \Sigma + \Delta G + L_q + L_g
\]

... does remaining spin come from large gluon polarization or orbital angular momentum?

→ stimulated many advances in theory, experiment & analysis
→ recent JAM global analyses, including JLab 6 GeV data
Proton spin structure

**Impact of JLab data**

- **inclusion of JLab data increases**
  # data points by factor $\sim 2$

- **reduced uncertainty in** $\Delta s^+$, $\Delta g$
  through $Q^2$ evolution

- **s-quark polarization negative**
  from inclusive DIS data
  (assuming SU(3) symmetry)
Polarization of quark sea?

Inclusive DIS data cannot distinguish between $q$ and $\bar{q}$

$\rightarrow$ semi-inclusive DIS sensitive to $\Delta q$ & $\Delta \bar{q}$

$\sim \sum_q e_q^2 \left[ \Delta q(x) D_q^h(z) + \Delta \bar{q}(x) D_{\bar{q}}^h(z) \right]$

$\rightarrow$ but need fragmentation functions!

Global analysis of DIS + SIDIS data gives different sign for strange quark polarization for different fragmentation functions!

$\rightarrow \Delta s > 0$ for “DSS” FFs \hspace{1cm} de Florian et al. (2007)

$\Delta s < 0$ for “HKNS” FFs \hspace{1cm} Hirai et al. (2007)

$\rightarrow$ need to understand origin of differences in fragmentation!
Polarization of quark sea?

First MC analysis of fragmentation functions

$e^+e^- \rightarrow hX$

single-inclusive annihilation (SIA)

Sato, Ethier, WM, Hirai, Kumano, Accardi (2016)

→ convergence after ~ 20 iterations
Polarization of quark sea?

Simultaneous determination of spin PDFs and FFs, fitting to DIS, SIA and polarized SIDIS (HERMES, COMPASS) data

- No assumption of SU(3) symmetry
- $\Delta s$ slightly $> 0$ at high $x$, consistent with zero
- $\Delta s - \Delta \bar{s}$ & $\Delta \bar{u} - \Delta \bar{d}$ consistent with zero
Simultaneous analysis

Polarized strangeness in previous, DIS-only analyses was negative at $x \sim 0.1$, induced by SU(3) and parametrization bias.

Weak sensitivity to $\Delta s^+$ from DIS data & evolution
- SU(3) pulls $\Delta s^+$ to generate moment $\sim -0.1$
- Negative peak at $x \sim 0.1$ induced by fixing $b \sim 6 - 8$

Less negative $\Delta s = -0.03(10)$ gives larger total helicity $\Delta \Sigma = 0.36(9)$
Nucleon Transversity PDFs
Transversity distributions

Extraction of transversity (TMD) PDF from SIDIS data + isovector moment $g_T = \int dx (h_1^u - h_1^d)$ from lattice QCD

→ Collins asymmetry

\[ A_{UT}^{\sin(\phi_h + \phi_s)} \propto \frac{h_1 \otimes H_1}{f_1 \otimes D_1} \]

→ significantly reduced uncertainties with lattice constraint

Lin, WM, Prokudin, Sato, Shows (2018)
Transversity distributions

Extraction of transversity (TMD) PDF from SIDIS data
+ isovector moment \( g_T = \int dx (h^u_1 - h^d_1) \) from lattice QCD

→ distributions do not look very Gaussian!

→ MC analysis gives \( g_T = 1.0 \pm 0.1 \)

→ maximum likelihood analysis would have given \( g_T \approx 0.5 \)
Pion PDFs
PDFs in the pion

PDFs in the pion (in principle) simpler to compute than baryons, but are more difficult to study experimentally

→ most information has come from pion-nucleus (tungsten) Drell-Yan data (CERN, Fermilab)

→ constrains valence PDFs at $x \gg 0$ (uncertainty from gluon resummation)

→ pion sea quark & gluon PDFs at small $x$ mostly unconstrained

Shi, Mezrag, Zong (2018)
PDFs in the pion

Recently a new (Monte Carlo-based) global analysis used chiral effective field theory to include also leading neutron electroproduction from HERA

\[
\frac{d^3 \sigma^{\text{LN}}}{dx \, dQ^2 \, dy} \sim 2f_{\pi/N}(y) \, F_2^{\pi}(x_\pi, Q^2)
\]

\[N \rightarrow N + \pi\]
splitting function
(computed from $\chi$EFT)

\[x_\pi = x/y\]

McKenney, Sato, WM, C.-R. Ji (2018)

\[
x_\pi f(x_\pi)
\]

Barry, Sato, WM, C.-R. Ji (2018)

→ first constraints on pion PDFs at low $x$
PDFs in the pion

Larger gluon fraction in the pion than without LN constraint

Tagged DIS experiment at JLab \((e n \rightarrow e' p X)\) will probe pion structure at intermediate \(x\) values (between DY and LN)

extension to hyperon final state will probe kaon structure
Outlook

New paradigm in global analysis — *simultaneous* determination of collinear distributions using MC sampling of parameter space → providing new insights into quark/gluon structure of hadrons

*Short-term*: “universal” QCD analysis of all observables sensitive to collinear (unpolarized & polarized) PDFs and FFs

*Longer-term*: technology developed will be applied to global analysis of transverse momentum dependent (TMD) distributions to map out full 3-d image of hadrons → vital interplay between theory & experiment at JLab
대단히 감사합니다.