First simultaneous extraction of (spin) PDFs & FFs from global QCD analysis

Wally Melnitchouk

Nobuo Sato  Jacob Ethier  Alberto Accardi

Connecticut / JLab  William & Mary  Hampton / JLab

JLab Angular Momentum Collaboration
http://www.jlab.org/JAM

- arXiv:1705.05889
- PRD 94, 114004 (2016)
- PRD 93, 074005 (2016)
Outline

- Aim to understand quark & gluon spin structure of proton
- New Iterative Monte Carlo (IMC) methodology, with Bayesian determination of PDF errors
  → single-fitting technology problematic for some PDFs

First global QCD analysis of polarised inclusive DIS including (6 GeV) JLab data

First MC extraction of fragmentation functions (FFs) from $e^+e^-$ single-inclusive annihilation (SIA) data

First combined analysis of polarised DIS + SIDIS + SIA data, with simultaneous extraction of PDFs and FFs
  → resolves strange quark polarisation puzzle
Proton spin structure

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

→ in nonrelativistic quark model, spin of proton is carried entirely by quarks

$$
\Delta \Sigma = \Delta u^+ + \Delta d^+ + \Delta s^+ = 1 \\
\Delta q^+ \equiv \Delta q + \Delta \bar{q}
$$

→ experimentally, from hadron weak decays have

$$
\Delta u^+ - \Delta d^+ = g_A = 1.269(3) \quad [SU(2) \text{ symmetry}] \\
\Delta u^+ + \Delta d^+ - \Delta s^+ = a_8 = 0.586(31) \quad [SU(3) \text{ symmetry}]
$$

→ third combination of flavours measured in polarised DIS

$$
\int_0^1 dx \ g_1 = \frac{1}{18} (4\Delta u^+ + \Delta d^+ + \Delta s^+)
$$
Proton spin structure

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

→ solving system of equations, early indications were that

$$\Delta \Sigma \approx 0! \quad \Delta s^+ \approx -(0.1 - 0.2)$$  

EMC (1988)

→ since 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 polarisation?
— orbital angular momentum?

→刺激 many advances in theory, experiment & data analysis
Data analysis methodology

Analysis of data requires estimating expectation values and variances of observables \( \mathcal{O} (= \text{PDFs, FFs}) \)

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

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

→ probability distribution

\[
\mathcal{P}(\vec{a}|\text{data}) \propto \mathcal{L}(\text{data}|\vec{a}) \, \pi(\vec{a})
\]

Bayes’ theorem

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

→ likelihood function

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

5
Data analysis methodology

- Standard method for evaluating $E$, $V$ is “maximum likelihood”
  - maximize probability distribution
    $$\mathcal{P}(\tilde{a}| \text{data}) \rightarrow \tilde{a}_0$$
  - if $\mathcal{O}$ linear in parameters, and if probability is symmetric in all parameters
    $$E[\mathcal{O}(\tilde{a})] = \mathcal{O}(\tilde{a}_0), \quad V[\mathcal{O}(\tilde{a})] \rightarrow \text{Hessian}$$

- In practice, since in general $E[f(\tilde{a})] \neq f(E[\tilde{a}])$, maximum likelihood method will sometimes fail
  - need more versatile approach (e.g. Monte Carlo)

$$E[\mathcal{O}] \approx \frac{1}{N} \sum_k \mathcal{O}(\tilde{a}_k), \quad V[\mathcal{O}] \approx \frac{1}{N} \sum_k [\mathcal{O}(\tilde{a}_k) - E[\mathcal{O}]]^2$$
Data analysis methodology

- Use traditional functional form for input distribution shape
\[ xf(x) = Nx^a (1 - x)^b (1 + c\sqrt{x} + dx + \cdots) \]

but sample significantly larger parameter space than possible in single-fit analyses

- no assumptions for exponents
- cross-validation to avoid overfitting
- iterate until convergence criteria satisfied
- unambiguous determination of PDF uncertainties
Inclusive DIS global analysis

Fit experimental asymmetries (longitudinal & transverse) rather than derived structure functions

\[ A_{\parallel} = \frac{\sigma^{\uparrow\downarrow} - \sigma^{\uparrow\uparrow}}{\sigma^{\uparrow\downarrow} + \sigma^{\uparrow\uparrow}} = D(A_1 + \eta A_2) \]
\[ A_{\perp} = \frac{\sigma^{\uparrow\Rightarrow} - \sigma^{\uparrow\Leftarrow}}{\sigma^{\uparrow\Rightarrow} + \sigma^{\uparrow\Leftarrow}} = d(A_2 - \zeta A_1) \]

where
\[ A_1 = (g_1 - \gamma^2 g_2) \frac{2x}{(1 + \gamma^2)F_2 - F_L} \]
\[ A_2 = \gamma (g_1 + g_2) \frac{2x}{(1 + \gamma^2)F_2 - F_L} \]

\[ \gamma^2 = \frac{4M^2x^2}{Q^2} \quad x = \frac{Q^2}{2M(E_i - E_f)} \quad F_2^{\text{LO}} \sum_q e_q^2 x(q + \bar{q}) \]

→ remove assumptions about \( R = \sigma_L/\sigma_T \) ratio
Inclusive DIS global analysis

■ Inclusive DIS data constrain $\Delta u^+$ & $\Delta d^+$ distributions
  $\rightarrow$ mostly insensitive to polarised strangeness and glue

■ Assume $g_1, g_2$ can be described as sum of twist $\tau = 2$ and higher twist (generally $1/Q$ suppressed) terms

$$g_1 = g_1^{\tau 2(TMC)} + g_1^{\tau 3(TMC)} + g_1^{\tau 4}$$

$$g_2 = g_2^{\tau 2(TMC)} + g_2^{\tau 3(TMC)}$$

includes target mass corrections (TMCs), proportional to

$$\rho^2 = 1 + \frac{4M^2x^2}{Q^2}$$
Inclusive DIS global analysis

- Leading twist contributions
  
  \[ g_{1,\tau_2}^{p,\text{LO}} = \frac{1}{2} \left( \frac{4}{9} \Delta u^+ + \frac{1}{9} \Delta d^+ + \frac{1}{9} \Delta s^+ + \cdots \right) \]

- Work in moment space (faster numerical computations)
  
  \[ \rightarrow \quad n\text{-th moment of function} \quad \Delta q^{(n)} = \int_0^1 dx \, x^{n-1} \Delta q(x) \]

- Structure function moments at next-to-leading order (NLO)
  
  \[ g_{1,\tau_2}^{(n)} = \frac{1}{2} \sum_q e_q^2 \left( \Delta C_{qq}^{(n)} \Delta q^{(n)} + \Delta C_{g}^{(n)} \Delta g^{(n)} \right) \]
  
  \[ g_{2,\tau_2}^{(n)} = -\frac{n - 1}{n} g_{1,\tau_2}^{(n)} \]

  hard scattering (Wilson) coefficients
Inclusive DIS global analysis

Higher twist corrections

→ twist-3 part of $g_1$ related to twist-3 part of $g_2$

$$g_1^{\tau 3}(x) = (\rho^2 - 1) \left[ g_2^{\tau 3}(x) - 2 \int_x^1 \frac{dz}{z} g_2^{\tau 3}(z) \right]$$

→ twist-3 part of $g_2$ parametrised via twist-3 PDFs $D_q$

$$g_2^{\tau 3}(x) = D(x) - \int_x^1 \frac{dz}{z} D(z), \quad D(x) = \sum_q e_q^2 \ D_q(x)$$

→ phenomenological twist-4 part

$$g_1^{\tau 4}(x) = \frac{H(x)}{Q^2}$$
### Convergence criteria

- Convergence after \(~ 5\text{–}6\) iterations

### Sensitivity to kinematic cuts

- Stability for \(W^2 > 4\ \text{GeV}^2\)
- and \(Q^2 > 1\ \text{GeV}^2\)
Inclusive DIS global analysis

Previous world data

- JAM15
- no HT
- syst (+)
- syst (−)

- EMC
  - $Q^2 \in [3.5, 29.5]$  
- SMC − 98
  - $Q^2 \in [1.3, 58.0]$  

- E80/E130
  - $Q^2 \in [3.3, 5.4]$  

- E80/E130
  - $Q^2 \in [5.3, 9.9]$  

- E143
  - $Q^2 \in [1.0, 1.4]$  

- E155
  - $Q^2 \in [1.0, 5.9]$  

- HERMES
  - $Q^2 \in [1.1, 3.1]$  

- HERMES
  - $Q^2 \in [1.2, 9.6]$  

- HERMES
  - $Q^2 \in [2.6, 14.3]$  

- HERMES
  - $Q^2 \in [2.0, 29.8]$  

- HERMES
  - $Q^2 \in [1.1, 53.1]$  

- COMPASS
  - $Q^2 \in [1.1, 62.1]$  

- COMPASS
  - $Q^2 \in [1.1, 96.1]$  

- COMPASS
  - $Q^2 \in [2.6, 14.3]$  

- COMPASS
Impact of JLab data

JLab eg1-dvcs (CLAS) data

→ signal of higher twist!
Impact of JLab data

→ twist-3 PDFs large!
→ same sign as twist-2
→ twist-4 term negligible

→ matrix element
\[ d_2 = 2g_1^{(3)} + 3g_2^{(3)} \]
related to
"color polarisability"
or "transverse force"
acting on quarks
Polarization of quark sea?

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

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

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

  → but need fragmentation functions!

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

  → $\Delta s > 0$ for "DSS" parametrisation \textit{de Florian et al., PRD75, 094009 (2007)}

  $\Delta s < 0$ for "HKNS" parametrisation \textit{Hirai et al., PRD75, 114010 (2007)}

  → need to understand origin of differences in fragmentation!
**IMC analysis of fragmentation functions**

**Analyse single-inclusive $e^+e^-$ annihilation data for pion & kaon production from DESY, CERN, SLAC & KEK from $Q \sim 10$ GeV to $Z$-boson pole**

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Ref.</th>
<th>Observable</th>
<th>$Q$ (GeV)</th>
<th>$N_{\text{dat}}$</th>
<th>norm.</th>
<th>$\chi^2$</th>
<th>$N_{\text{dat}}$</th>
<th>norm.</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARGUS</td>
<td>[26]</td>
<td>Inclusive</td>
<td>9.98</td>
<td>35</td>
<td>1.024 (1.058)</td>
<td>51.1 (55.8)</td>
<td>15</td>
<td>1.007</td>
<td>8.5</td>
</tr>
<tr>
<td>Belle</td>
<td>[38,39]</td>
<td>Inclusive</td>
<td>10.52</td>
<td>78</td>
<td>0.900 (0.919)</td>
<td>37.6 (21.7)</td>
<td>78</td>
<td>0.988</td>
<td>10.9</td>
</tr>
<tr>
<td>BABAR</td>
<td>[40]</td>
<td>Inclusive</td>
<td>10.54</td>
<td>39</td>
<td>0.993 (0.948)</td>
<td>31.6 (70.7)</td>
<td>30</td>
<td>0.992</td>
<td>4.9</td>
</tr>
<tr>
<td>TASSO</td>
<td>[23–25]</td>
<td>Inclusive</td>
<td>12.44</td>
<td>29</td>
<td>(*)</td>
<td>37.0 (38.8)</td>
<td>18</td>
<td>(*)</td>
<td>14.3</td>
</tr>
<tr>
<td>TPC</td>
<td>[27–29]</td>
<td>Inclusive</td>
<td>29.00</td>
<td>18</td>
<td>1</td>
<td>36.3 (57.8)</td>
<td>16</td>
<td>1</td>
<td>47.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$uds$ tag</td>
<td>29.00</td>
<td>6</td>
<td>1</td>
<td>3.7 (4.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$b$ tag</td>
<td>29.00</td>
<td>6</td>
<td>1</td>
<td>8.7 (8.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$c$ tag</td>
<td>29.00</td>
<td>6</td>
<td>1</td>
<td>3.3 (3.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HRS</td>
<td>[30]</td>
<td>Inclusive</td>
<td>29.00</td>
<td>2</td>
<td>1</td>
<td>4.2 (6.2)</td>
<td>3</td>
<td>1</td>
<td>0.3</td>
</tr>
<tr>
<td>TOPAZ</td>
<td>[37]</td>
<td>Inclusive</td>
<td>58.00</td>
<td>4</td>
<td>1</td>
<td>4.8 (6.3)</td>
<td>3</td>
<td>1</td>
<td>0.9</td>
</tr>
<tr>
<td>OPAL</td>
<td>[32,33]</td>
<td>Inclusive</td>
<td>91.20</td>
<td>22</td>
<td>1</td>
<td>33.3 (37.2)</td>
<td>10</td>
<td>1</td>
<td>6.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$u$ tag</td>
<td>91.20</td>
<td>5</td>
<td>1.203 (1.203)</td>
<td>6.6 (8.1)</td>
<td>5</td>
<td>1.185</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d$ tag</td>
<td>91.20</td>
<td>5</td>
<td>1.204 (1.203)</td>
<td>6.1 (7.6)</td>
<td>5</td>
<td>1.075</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$s$ tag</td>
<td>91.20</td>
<td>5</td>
<td>1.126 (1.200)</td>
<td>14.4 (11.0)</td>
<td>5</td>
<td>1.173</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$c$ tag</td>
<td>91.20</td>
<td>5</td>
<td>1.174 (1.323)</td>
<td>10.7 (6.1)</td>
<td>5</td>
<td>1.169</td>
<td>13.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$b$ tag</td>
<td>91.20</td>
<td>5</td>
<td>1.218 (1.209)</td>
<td>34.2 (36.6)</td>
<td>4</td>
<td>1.177</td>
<td>10.9</td>
</tr>
<tr>
<td>ALEPH</td>
<td>[34]</td>
<td>Inclusive</td>
<td>91.20</td>
<td>22</td>
<td>0.987 (0.989)</td>
<td>15.6 (20.4)</td>
<td>18</td>
<td>1.008</td>
<td>6.1</td>
</tr>
<tr>
<td>DELPHI</td>
<td>[35,36]</td>
<td>Inclusive</td>
<td>91.20</td>
<td>17</td>
<td>1</td>
<td>21.0 (20.2)</td>
<td>27</td>
<td>1</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$uds$ tag</td>
<td>91.20</td>
<td>17</td>
<td>1</td>
<td>13.3 (13.4)</td>
<td>17</td>
<td>1</td>
<td>22.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$b$ tag</td>
<td>91.20</td>
<td>17</td>
<td>1</td>
<td>41.9 (42.9)</td>
<td>17</td>
<td>1</td>
<td>9.1</td>
</tr>
<tr>
<td>SLD</td>
<td>[31]</td>
<td>Inclusive</td>
<td>91.28</td>
<td>29</td>
<td>1.002 (1.004)</td>
<td>27.3 (36.3)</td>
<td>29</td>
<td>0.994</td>
<td>14.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$uds$ tag</td>
<td>91.28</td>
<td>29</td>
<td>1.003 (1.004)</td>
<td>51.7 (55.6)</td>
<td>29</td>
<td>0.994</td>
<td>42.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$c$ tag</td>
<td>91.28</td>
<td>29</td>
<td>0.998 (1.001)</td>
<td>30.2 (40.4)</td>
<td>29</td>
<td>1.000</td>
<td>31.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$b$ tag</td>
<td>91.28</td>
<td>29</td>
<td>1.005 (1.005)</td>
<td>74.6 (61.9)</td>
<td>28</td>
<td>0.992</td>
<td>134.1</td>
</tr>
<tr>
<td>Total:</td>
<td></td>
<td></td>
<td>459</td>
<td></td>
<td>599.3 (671.2)</td>
<td>391</td>
<td></td>
<td>395.0</td>
<td></td>
</tr>
</tbody>
</table>

$\chi^2/N_{\text{dat}} = 1.31 (1.46)$

$\chi^2/N_{\text{dat}} = 1.01$
IMC analysis of fragmentation functions

\[
\begin{align*}
\text{pion production} \\
\text{kaon production}
\end{align*}
\]

→ generally good agreement
→ some issues at high $z$
→ smaller $\chi^2$ for $K$ cf. $\pi$

due to larger uncertainties
IMC analysis of fragmentation functions

→ favoured FFs well constrained; unfavoured not as well...
→ nontrivial shape of \( s \rightarrow K \) fragmentation
→ very hard \( g \rightarrow K \) fragmentation?  \( \text{(robust feature!)} \)
IMC analysis of fragmentation functions

→ qualitatively similar behavior as in previous analyses

→ JAM strange FF closer to DSS parametrization
  — impact on $\Delta s^+$ extraction??
Simultaneous analysis

Using same IMC tools, new analysis attempted first ever simultaneous determination of spin PDFs and FFs, fitting to DIS, SIA and polarised 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

Using same IMC tools, new analysis attempted first ever simultaneous determination of spin PDFs and FFs, fitting to DIS, SIA and polarised SIDIS (HERMES, COMPASS) data

<table>
<thead>
<tr>
<th>process</th>
<th>target</th>
<th>$N_{\text{dat}}$</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIS</td>
<td>$p, d, ^3\text{He}$</td>
<td>854</td>
<td>854.8</td>
</tr>
<tr>
<td>SIA ($\pi^\pm, K^\pm$)</td>
<td></td>
<td>850</td>
<td>997.1</td>
</tr>
<tr>
<td>SIDIS ($\pi^\pm$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HERMES [15]</td>
<td>$d$</td>
<td>18</td>
<td>28.1</td>
</tr>
<tr>
<td>COMPASS [16]</td>
<td>$d$</td>
<td>20</td>
<td>8.0</td>
</tr>
<tr>
<td>COMPASS [17]</td>
<td>$p$</td>
<td>24</td>
<td>18.2</td>
</tr>
<tr>
<td>SIDIS ($K^\pm$)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HERMES [15]</td>
<td>$d$</td>
<td>27</td>
<td>18.3</td>
</tr>
<tr>
<td>COMPASS [16]</td>
<td>$d$</td>
<td>20</td>
<td>18.7</td>
</tr>
<tr>
<td>COMPASS [17]</td>
<td>$p$</td>
<td>24</td>
<td>12.3</td>
</tr>
<tr>
<td>Total:</td>
<td></td>
<td>1855</td>
<td>1969.7</td>
</tr>
</tbody>
</table>

→ SIDIS data dwarfed by DIS & SIA

→ smaller $\chi^2$ for $K$ cf. $\pi$

SIDIS due to larger errors
Simultaneous analysis

- Slightly positive $\Delta s^+$ at $x \approx 0.1$ attributed directly to deuteron $K^-$ production data.

![Graph showing $A_{1d}^{K^-}$ and $A_{1p}^{\pi^-}$](image)

- Significantly worse fit using negative JAM15 $\Delta s^+$ (also for COMPASS and $p$ data).
- No such effect observed for pion data sensitive to $\Delta \bar{u}$ & $\Delta \bar{d}$.
Simultaneous analysis

- Polarised strangeness in previous, DIS-only analyses (e.g. JAM15) was negative at $x \sim 0.1$, induced by $SU(3)$ and parametrisation bias.

- Negligible 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 $b$ parameter to $\sim 6 - 8$. 

Diagram showing $x \Delta s^+$ distribution with and without $SU(3)$ corrections.
Simultaneous analysis

- Fragmentation functions similar to previous SIA-only analysis

- Some constraint from SIDIS on unfavoured FFs (e.g. $s \rightarrow K^+$) but uncertainties still large

- New result more consistent with DSS at moderate $z$

- Qualitatively similar to NJL-Jet model calculation

Matevosyan et al., PRD 83, 114010 (2011)
Simultaneous analysis

- Statistical distribution of lowest moments (axial charges)

→ triplet charge $g_A$ consistent with $SU(2)$ value

→ hint of $SU(3)$ breaking in octet charge $a_8$ [cf. Bass, Thomas, PLB 684, 216 (2010)]

→ less negative $\Delta s = -0.03(10)$ gives larger total helicity $\Delta \Sigma = 0.36(9)$
Outlook

- New paradigm in global QCD analysis
  — *simultaneous* determination of collinear distributions using **Monte Carlo** sampling of parameter space

- First combined analysis of polarised DIS + SIDIS + SIA, with consistent extraction of spin PDFs and FFs without assuming **SU(3)** symmetry
  → resolves strange quark polarization puzzle: $\Delta s = -0.03(10)$
  (less negative than traditional **SU(3)** value $\Delta s \sim -0.1$)

- Near-term future: “universal” QCD analysis of all observables sensitive to collinear (unpolarised & polarised) PDFs and FFs

- Longer-term: apply IMC technology to global QCD analysis of transverse momentum dependent (TMD) PDFs and FFs