Parton Distribution Functions and the LHC

James Stirling
Cambridge University

• introduction: parton distribution functions
• the MSTW* project
• implications for LHC physics

*with Alan Martin, Robert Thorne, Graeme Watt
1

depth inelastic scattering
and
parton distributions
parton distribution functions

\[ f_{i/A}(x, Q^2) \]

• introduced by Feynman (1969) in the parton model, to explain Bjorken scaling in deep inelastic scattering data; interpretation as probability distributions

• according to the QCD factorisation theorem for inclusive hard scattering processes, universal distributions containing long-distance structure of hadrons; related to parton model distributions at leading order, but with logarithmic scaling violations

• key ingredients for Tevatron and LHC phenomenology
for example, in Deep Inelastic Scattering

\[
\frac{1}{x} F_{2}^{l\rho}(x, Q^2) = x \sum_{q} e_{q}^{2} \int_{x}^{1} \frac{dy}{y} q(y, Q^2) \left\{ \delta(1 - \frac{x}{y}) + \frac{\alpha_{S}(Q^2)}{2\pi} C_{q}(x/y) \right\} \\
+ x \sum_{q} e_{q}^{2} \frac{\alpha_{S}(Q^2)}{2\pi} \int_{x}^{1} \frac{dy}{y} g(y, Q^2) C_{g}(x/y) + \mathcal{O}(\alpha_{S}^{2}) \\
+ \mathcal{O}(1/Q^2) \text{ (higher twist, mass corrections)}
\]

where the scale dependence of the parton distributions is calculable in QCD perturbation theory

\[
\mu^{2} \frac{\partial}{\partial \mu^{2}} f_{i}(x, \mu^{2}) = \frac{\alpha_{S}(\mu^{2})}{2\pi} \sum_{j} \int_{x}^{1} \frac{dy}{y} f_{j}(y, \mu^{2}) P_{ij}(x/y, \alpha_{S}(\mu^{2}))
\]

… and \( f_{i}(x, \mu_{0}^{2}) \) determined from

- lattice QCD
- fits to data
MSTW2008NLO
$Q^2 = 10^1 - 10^5 \text{GeV}^2$
beyond leading order in pQCD:

\[ P(x, \alpha_s) = P^{(0)}(x) + \left( \frac{\alpha_s}{2\pi} \right) P^{(1)}(x) + \left( \frac{\alpha_s}{2\pi} \right)^2 P^{(2)}(x) + \ldots \]

the 2004 calculation of the complete set of \( P^{(2)} \) splitting functions by Moch, Vermaseren and Vogt completes the calculational tools for a consistent NNLO (massless) pQCD treatment of lepton-hadron and hadron-hadron collider hard-scattering cross sections

Note:
- \( f_i^{LO} \), \( f_i^{NLO} \) and \( f_i^{NNLO} \) are different quantities!
- beyond LO, they depend on the factorisation prescription (e.g. MS)
and in hadron-hadron collisions

\[ \sigma_{pp \to X + \ldots} = \sum_{ij} \int_0^1 dx_1 dx_2 f_i(x_1, \mu_F^2) f_j(x_2, \mu_F^2) \times \tilde{\sigma}_{ij \to X + \ldots} \left(x_1, x_2, \{\mu^\mu_i\}, \alpha_S(\mu_R^2), \frac{Q^2}{\mu_R^2}, \frac{Q^2}{\mu_F^2} \right) \]

where \( Q \) is the ‘hard scale’ (e.g. = \( M_X \)), usually \( \mu_F, \mu_R = cQ \), and \( \tilde{\sigma} \) is known to fixed order in pQCD (sometimes with additional resummed large log corrections, QED corrections, ....)
momentum fractions $x_1$ and $x_2$ determined by mass and rapidity of $X$
pdfs in 1993

<table>
<thead>
<tr>
<th></th>
<th>$\mu$-DIS</th>
<th>$\nu$-DIS</th>
<th>Prompt $\gamma$</th>
<th>D-Yan</th>
<th>$W, Z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRS '88</td>
<td>EMC + ...</td>
<td>CDHSW</td>
<td>AFS($+J/\psi$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DFLM '88</td>
<td>(EMC + ...)</td>
<td>CHARM + ...</td>
<td>-</td>
<td>E288 + ...</td>
<td>-</td>
</tr>
<tr>
<td>ABFOW '89</td>
<td>BCDMS</td>
<td>-</td>
<td>WA70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMRS '90</td>
<td>EMC</td>
<td>CDHSW</td>
<td>WA70</td>
<td>E605</td>
<td>-</td>
</tr>
<tr>
<td>MT '90</td>
<td>EMC</td>
<td>CDHSW</td>
<td>-</td>
<td>E288</td>
<td>-</td>
</tr>
<tr>
<td>KMRS '90</td>
<td>BCDMS</td>
<td>CDHSW</td>
<td>WA70</td>
<td>E605</td>
<td>-</td>
</tr>
</tbody>
</table>

- MRS (Apr '92)
  - MRS (Nov '92)
- (sets $D_0, D_{--}$)
- (sets $D_0, D_{--}$)

- CTEQ ('93)

MRS, DIS93, Durham
### pdfs in 2009

<table>
<thead>
<tr>
<th>Data set</th>
<th>$N_{\text{pts.}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1 MB 99 e$^+p$ NC</td>
<td>8</td>
</tr>
<tr>
<td>H1 MB 97 e$^+p$ NC</td>
<td>64</td>
</tr>
<tr>
<td>H1 low $Q^2$ 96–97 e$^+p$ NC</td>
<td>80</td>
</tr>
<tr>
<td>H1 high $Q^2$ 98–99 e$^−p$ NC</td>
<td>126</td>
</tr>
<tr>
<td>H1 high $Q^2$ 99–00 e$^+p$ NC</td>
<td>147</td>
</tr>
<tr>
<td>ZEUS SVX 95 e$^+p$ NC</td>
<td>30</td>
</tr>
<tr>
<td>ZEUS 96–97 e$^+p$ NC</td>
<td>144</td>
</tr>
<tr>
<td>ZEUS 98–99 e$^−p$ NC</td>
<td>92</td>
</tr>
<tr>
<td>ZEUS 99–00 e$^+p$ NC</td>
<td>90</td>
</tr>
<tr>
<td>H1 99–00 e$^+p$ CC</td>
<td>28</td>
</tr>
<tr>
<td>ZEUS 99–00 e$^−p$ CC</td>
<td>30</td>
</tr>
<tr>
<td>H1/ZEUS e$^±p$ $F^\text{charm}$</td>
<td>83</td>
</tr>
<tr>
<td>H1 99–00 e$^+p$ incl. jets</td>
<td>24</td>
</tr>
<tr>
<td>ZEUS 96–97 e$^+p$ incl. jets</td>
<td>30</td>
</tr>
<tr>
<td>ZEUS 98–00 e$^±p$ incl. jets</td>
<td>30</td>
</tr>
<tr>
<td>DØ II $p\bar{p}$ incl. jets</td>
<td>110</td>
</tr>
<tr>
<td>CDF II $p\bar{p}$ incl. jets</td>
<td>76</td>
</tr>
<tr>
<td>CDF II $W \rightarrow l\nu$ asym.</td>
<td>22</td>
</tr>
<tr>
<td>DØ II $W \rightarrow l\nu$ asym.</td>
<td>10</td>
</tr>
<tr>
<td>DØ II $Z$ rap.</td>
<td>28</td>
</tr>
<tr>
<td>CDF II $Z$ rap.</td>
<td>29</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data set</th>
<th>$N_{\text{pts.}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCDMS $\mu p$ $F_2$</td>
<td>163</td>
</tr>
<tr>
<td>BCDMS $\mu d$ $F_2$</td>
<td>151</td>
</tr>
<tr>
<td>NMC $\mu p$ $F_2$</td>
<td>123</td>
</tr>
<tr>
<td>NMC $\mu d$ $F_2$</td>
<td>123</td>
</tr>
<tr>
<td>NMC $\mu n/\mu p$</td>
<td>148</td>
</tr>
<tr>
<td>E665 $\mu p$ $F_2$</td>
<td>53</td>
</tr>
<tr>
<td>E665 $\mu d$ $F_2$</td>
<td>53</td>
</tr>
<tr>
<td>SLAC ep $F_2$</td>
<td>37</td>
</tr>
<tr>
<td>SLAC ed $F_2$</td>
<td>38</td>
</tr>
<tr>
<td>NMC/BCDMS/SLAC $F_L$</td>
<td>31</td>
</tr>
<tr>
<td>E866/NuSea pp DY</td>
<td>184</td>
</tr>
<tr>
<td>E866/nuSea pd/pp DY</td>
<td>15</td>
</tr>
<tr>
<td>NuTeV $\nu N$ $F_2$</td>
<td>53</td>
</tr>
<tr>
<td>CHORUS $\nu N$ $F_2$</td>
<td>42</td>
</tr>
<tr>
<td>NuTeV $\nu N$ $xF_3$</td>
<td>45</td>
</tr>
<tr>
<td>CHORUS $\nu N$ $xF_3$</td>
<td>33</td>
</tr>
<tr>
<td>CCFR $\nu N \rightarrow \mu \mu X$</td>
<td>86</td>
</tr>
<tr>
<td>NuTeV $\nu N \rightarrow \mu \mu X$</td>
<td>84</td>
</tr>
</tbody>
</table>

| All data sets                               | 2743              |

G2 = New w.r.t. MRST 2006 fit.
how pdfs are obtained

• choose a factorisation scheme (e.g. MSbar), an order in perturbation theory (LO, NLO, NNLO) and a ‘starting scale’ $Q_0$ where pQCD applies (e.g. 1-2 GeV)

• parametrise the quark and gluon distributions at $Q_0$, e.g.

$$f_i(x, Q_0^2) = A_i x^{a_i} [1 + b_i \sqrt{x} + c_i x] (1 - x)^{d_i}$$

• solve DGLAP equations to obtain the pdfs at any $x$ and scale $Q > Q_0$; fit data for parameters $\{A_i, a_i, \ldots, \alpha_S\}$

• approximate the exact solutions (e.g. interpolation grids, expansions in polynomials etc) for ease of use; thus the output ‘global fits’ are available ‘off the shelf”, e.g.

```
SUBROUTINE PDF (X, Q, U, UBAR, D, DBAR, ..., BBAR, GLU)
```

input | output
the pdf industry

• many groups now extracting pdfs from ‘global’ data analyses (MSTW, CTEQ, NNPDF, HERAPDF, ABKM, GJR, …)

• broad agreement, but differences due to
  — choice of data sets (including cuts and corrections)
  — treatment of data errors
  — treatment of heavy quarks \((s,c,b)\)
  — order of perturbation theory
  — parameterisation at \(Q_0\)
  — theoretical assumptions (if any) about:
    • flavour symmetries
    • \(x \to 0,1\) behaviour
    • …
recent global fits


GJR (Jimenez-Delgado, Reya): global (DIS + DY) NLO, NNLO FFNS, VFNS fits [arXiv:0909.1711]

HERAPDF (H1+ZEUS): NLO VFNS fit to DIS data [arXiv:0906.1108]

MSTW (Martin, Stirling, Thorne, Watt): global LO, NLO and NNLO VFNS fits [arXiv:0901.0002]

NNPDF (Ball et al): NLO VFNS fit to DIS data using neural net technology to avoid parameterisation dependence [arXiv:0906.1958]

+ older fits
2

MSTW*

*Alan Martin, WJS, Robert Thorne, Graeme Watt

the MRS/MRST/MSTW project

- since 1987 (with Alan Martin, Dick Roberts, Robert Thorne, Graeme Watt), to produce 'state-of-the-art' pdfs
- combine experimental data with theoretical formalism to perform 'global fits' to data to extract the pdfs in user-friendly form for the particle physics community
- currently widely used at HERA and the Fermilab Tevatron, and in physics simulations for the LHC
- currently, the only available NNLO pdf sets with rigorous treatment of heavy quark flavours
MSTW 2008

- supersedes MRST sets
- new data (see next slide)
- new theory/infrastructure
  - $\delta f_i$ from new dynamic tolerance method: 68%cl (1$\sigma$) and 90%cl (cf. MRST) sets available
  - new definition of $\alpha_S$ (no more $\Lambda_{QCD}$)
  - new GM-VFNS for $c, b$ (see Martin et al., arXiv:0706.0459)
  - new fitting codes: FEWZ, VRAP, fastNLO
  - new grids: denser, broader coverage
  - slightly extended parametrisation at $Q_0^2$: 34-4=30 free parameters including $\alpha_S$

code, text and figures available at:
http://projects.hepforge.org/mstwpdf/
data sets used in fit

<table>
<thead>
<tr>
<th>Data set</th>
<th>(N_{\text{pts.}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1 MB 99 e(^+)p NC</td>
<td>8</td>
</tr>
<tr>
<td>H1 MB 97 e(^+)p NC</td>
<td>64</td>
</tr>
<tr>
<td>H1 low (Q^2) 96–97 e(^+)p NC</td>
<td>80</td>
</tr>
<tr>
<td>H1 high (Q^2) 98–99 e(^-)p NC</td>
<td>126</td>
</tr>
<tr>
<td>H1 high (Q^2) 99–00 e(^+)p NC</td>
<td>147</td>
</tr>
<tr>
<td>ZEUS SVX 95 e(^+)p NC</td>
<td>30</td>
</tr>
<tr>
<td>ZEUS 96–97 e(^+)p NC</td>
<td>144</td>
</tr>
<tr>
<td>ZEUS 98–99 e(^-)p NC</td>
<td>92</td>
</tr>
<tr>
<td>ZEUS 99–00 e(^+)p NC</td>
<td>90</td>
</tr>
<tr>
<td>H1 99–00 e(^+)p CC</td>
<td>28</td>
</tr>
<tr>
<td>ZEUS 99–00 e(^+)p CC</td>
<td>30</td>
</tr>
<tr>
<td>H1/ZEUS e(^\pm)p (F^\text{charm})</td>
<td>83</td>
</tr>
<tr>
<td>H1 99–00 e(^+)p incl. jets</td>
<td>24</td>
</tr>
<tr>
<td>ZEUS 96–97 e(^+)p incl. jets</td>
<td>30</td>
</tr>
<tr>
<td>ZEUS 98–00 e(^\pm)p incl. jets</td>
<td>30</td>
</tr>
<tr>
<td>DØ II (p\bar{p}) incl. jets</td>
<td>110</td>
</tr>
<tr>
<td>CDF II (p\bar{p}) incl. jets</td>
<td>76</td>
</tr>
<tr>
<td>CDF II (W \to l\nu) asym.</td>
<td>22</td>
</tr>
<tr>
<td>DØ II (W \to l\nu) asym.</td>
<td>10</td>
</tr>
<tr>
<td>DØ II Z rap.</td>
<td>28</td>
</tr>
<tr>
<td>CDF II Z rap.</td>
<td>29</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data set</th>
<th>(N_{\text{pts.}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCDMS (\mu p) (F_2)</td>
<td>163</td>
</tr>
<tr>
<td>BCDMS (\mu d) (F_2)</td>
<td>151</td>
</tr>
<tr>
<td>NMC (\mu p) (F_2)</td>
<td>123</td>
</tr>
<tr>
<td>NMC (\mu d) (F_2)</td>
<td>123</td>
</tr>
<tr>
<td>NMC (\mu n/\mu p)</td>
<td>148</td>
</tr>
<tr>
<td>E665 (\mu p) (F_2)</td>
<td>53</td>
</tr>
<tr>
<td>E665 (\mu d) (F_2)</td>
<td>53</td>
</tr>
<tr>
<td>SLAC (ep) (F_2)</td>
<td>37</td>
</tr>
<tr>
<td>SLAC (ed) (F_2)</td>
<td>38</td>
</tr>
<tr>
<td>NMC/BCDMS/SLAC (F_L)</td>
<td>31</td>
</tr>
<tr>
<td>E866/NuSea (pp) DY</td>
<td>184</td>
</tr>
<tr>
<td>E866/NuSea (pd/pp) DY</td>
<td>15</td>
</tr>
<tr>
<td>NuTeV (\nu N) (F_2)</td>
<td>53</td>
</tr>
<tr>
<td>CHORUS (\nu N) (F_2)</td>
<td>42</td>
</tr>
<tr>
<td>NuTeV (\nu N\times F_3)</td>
<td>45</td>
</tr>
<tr>
<td>CHORUS (\nu N\times F_3)</td>
<td>33</td>
</tr>
<tr>
<td>CCFR (\nu N \to \mu\mu X)</td>
<td>86</td>
</tr>
<tr>
<td>NuTeV (\nu N \to \mu\mu X)</td>
<td>84</td>
</tr>
</tbody>
</table>

| All data sets                | 2743                  |

- Red = New w.r.t. MRST 2006 fit.
values of $\chi^2/N_{\text{pts}}$ for the data sets included in the MSTW2008 global fits

<table>
<thead>
<tr>
<th>Data set</th>
<th>LO</th>
<th>NLO</th>
<th>NNLO</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCDMS $ud$ $F_2$ [102]</td>
<td>162 / 142</td>
<td>190 / 151</td>
<td>188 / 151</td>
</tr>
<tr>
<td>NMC $pP$ $F_2$ [53]</td>
<td>137 / 115</td>
<td>121 / 123</td>
<td>115 / 123</td>
</tr>
<tr>
<td>NMC $ud$ $F_2$ [53]</td>
<td>126 / 115</td>
<td>102 / 125</td>
<td>93 / 123</td>
</tr>
<tr>
<td>NMC $uP$ [103]</td>
<td>131 / 127</td>
<td>130 / 128</td>
<td>135 / 148</td>
</tr>
<tr>
<td>E605 $pp$ $F_2$ [104]</td>
<td>56 / 53</td>
<td>57 / 53</td>
<td>63 / 53</td>
</tr>
<tr>
<td>E605 $ud$ $F_2$ [104]</td>
<td>49 / 53</td>
<td>53 / 53</td>
<td>63 / 53</td>
</tr>
<tr>
<td>SLAC $eP$ $F_2$ [105, 106]</td>
<td>24 / 18</td>
<td>30 / 37</td>
<td>31 / 37</td>
</tr>
<tr>
<td>SLAC $ed$ $F_2$ [105, 106]</td>
<td>12 / 18</td>
<td>30 / 38</td>
<td>26 / 38</td>
</tr>
<tr>
<td>NMC/BCDMS/SLAC $F_2$ [32-34]</td>
<td>28 / 24</td>
<td>38 / 31</td>
<td>32 / 31</td>
</tr>
<tr>
<td>ES66/NuSea $pp$ $DY$ [107]</td>
<td>239 / 184</td>
<td>228 / 184</td>
<td>257 / 184</td>
</tr>
<tr>
<td>ES66/NuSea $pd/pp$ $DY$ [108]</td>
<td>14 / 15</td>
<td>14 / 15</td>
<td>14 / 15</td>
</tr>
<tr>
<td>NuTeV $\nu N$ $F_2$ [37]</td>
<td>49 / 49</td>
<td>49 / 53</td>
<td>46 / 53</td>
</tr>
<tr>
<td>CHORUS $\nu N$ $F_2$ [38]</td>
<td>21 / 37</td>
<td>20 / 42</td>
<td>29 / 42</td>
</tr>
<tr>
<td>NuTeV $\nu N$ $xF_2$ [37]</td>
<td>92 / 45</td>
<td>40 / 45</td>
<td>34 / 45</td>
</tr>
<tr>
<td>CHORUS $\nu N$ $xF_2$ [38]</td>
<td>44 / 33</td>
<td>31 / 33</td>
<td>36 / 33</td>
</tr>
<tr>
<td>CCFR $\nu N \rightarrow \mu X$ [39]</td>
<td>63 / 86</td>
<td>66 / 86</td>
<td>69 / 86</td>
</tr>
<tr>
<td>NuTeV $\nu N \rightarrow \mu X$ [39]</td>
<td>44 / 40</td>
<td>39 / 40</td>
<td>45 / 40</td>
</tr>
<tr>
<td>H1 MB 90 $e^{+}\nu$ NC [31]</td>
<td>9 / 8</td>
<td>9 / 8</td>
<td>7 / 8</td>
</tr>
<tr>
<td>H1 MB 97 $e^{+}\nu$ NC [109]</td>
<td>46 / 64</td>
<td>42 / 64</td>
<td>51 / 64</td>
</tr>
<tr>
<td>H1 low $Q^2$ 96-97 $e^{+}\nu$ NC [109]</td>
<td>54 / 80</td>
<td>44 / 80</td>
<td>45 / 80</td>
</tr>
<tr>
<td>H1 high $Q^2$ 98-99 $e^{+}\nu$ NC [110]</td>
<td>134 / 126</td>
<td>122 / 126</td>
<td>124 / 126</td>
</tr>
<tr>
<td>H1 high $Q^2$ 99-00 $e^{+}\nu$ NC [36]</td>
<td>153 / 147</td>
<td>131 / 147</td>
<td>153 / 147</td>
</tr>
<tr>
<td>ZEUS SVX 95 $e^{+}\nu$ NC [111]</td>
<td>35 / 30</td>
<td>35 / 30</td>
<td>35 / 30</td>
</tr>
<tr>
<td>ZEUS 96-97 $e^{+}\nu$ NC [112]</td>
<td>118 / 144</td>
<td>86 / 144</td>
<td>86 / 144</td>
</tr>
<tr>
<td>ZEUS 98-99 $e^{+}\nu$ NC [113]</td>
<td>61 / 92</td>
<td>54 / 92</td>
<td>54 / 92</td>
</tr>
<tr>
<td>ZEUS 99-00 $e^{+}\nu$ NC [114]</td>
<td>75 / 90</td>
<td>63 / 90</td>
<td>65 / 90</td>
</tr>
<tr>
<td>H1 99-00 $e^{+}\nu$ CC [35]</td>
<td>28 / 28</td>
<td>29 / 28</td>
<td>29 / 28</td>
</tr>
<tr>
<td>ZEUS 99-00 $e^{+}\nu$ CC [36]</td>
<td>36 / 30</td>
<td>38 / 30</td>
<td>37 / 30</td>
</tr>
<tr>
<td>H1/ZEUS $eP$ $F_2^{\text{herm}}$ [41, 47]</td>
<td>110 / 83</td>
<td>107 / 83</td>
<td>95 / 83</td>
</tr>
<tr>
<td>H1 99-00 $e^{+}\nu$ incl. jets [59]</td>
<td>109 / 24</td>
<td>19 / 24</td>
<td>—</td>
</tr>
<tr>
<td>ZEUS 96-97 $e^{+}\nu$ incl. jets [57]</td>
<td>85 / 30</td>
<td>30 / 30</td>
<td>—</td>
</tr>
<tr>
<td>ZEUS 98-00 $e^{+}\nu$ incl. jets [58]</td>
<td>102 / 30</td>
<td>17 / 30</td>
<td>—</td>
</tr>
<tr>
<td>D0 II $pf$ incl. jets [56]</td>
<td>139 / 110</td>
<td>114 / 110</td>
<td>123 / 110</td>
</tr>
<tr>
<td>CDF II $pp$ incl. jets [54]</td>
<td>143 / 76</td>
<td>56 / 76</td>
<td>54 / 76</td>
</tr>
<tr>
<td>CDF II $W \rightarrow \ell\nu$ asym. [48]</td>
<td>50 / 22</td>
<td>29 / 22</td>
<td>30 / 22</td>
</tr>
<tr>
<td>D0 II $W \rightarrow \ell\nu$ asym. [49]</td>
<td>25 / 10</td>
<td>25 / 10</td>
<td>25 / 10</td>
</tr>
<tr>
<td>D0 II $Z$ rap. [32]</td>
<td>26 / 28</td>
<td>19 / 28</td>
<td>17 / 28</td>
</tr>
<tr>
<td>CDF II $Z$ rap. [52]</td>
<td>52 / 20</td>
<td>49 / 20</td>
<td>50 / 20</td>
</tr>
</tbody>
</table>

All data sets 3069 / 2598 2543 / 2699 2480 / 2615
MSTW input parametrisation

At input scale \( Q_0^2 = 1 \text{ GeV}^2 \):

\[
x_u = A_u x^{\eta_1} (1 - x)^{\eta_2} (1 + \epsilon_u \sqrt{x} + \gamma_u x)
\]

\[
x_d = A_d x^{\eta_3} (1 - x)^{\eta_4} (1 + \epsilon_d \sqrt{x} + \gamma_d x)
\]

\[
x_s = A_s x^{\delta_s} (1 - x)^{\eta_s} (1 + \epsilon_s \sqrt{x} + \gamma_s x)
\]

\[
x_{\bar{d}} - x_{\bar{u}} = A_\Delta x^{\eta_\Delta} (1 - x)^{\eta_{\Delta} + 2} (1 + \gamma_\Delta x + \delta_\Delta x^2)
\]

\[
x_g = A_g x^{\delta_g} (1 - x)^{\eta_g} (1 + \epsilon_g \sqrt{x} + \gamma_g x) + A_{g'} x^{\delta_{g'}} (1 - x)^{\eta_{g'}}
\]

\[
x_s + x_{\bar{s}} = A_+ x^{\delta_s} (1 - x)^{\eta_+} (1 + \epsilon_s \sqrt{x} + \gamma_s x)
\]

\[
x_s - x_{\bar{s}} = A_- x^{\delta-} (1 - x)^{\eta-} (1 - x/x_0)
\]

Note: 20 parameters allowed to go free for eigenvector PDF sets, cf. 15 for MRST sets
which data sets determine which partons?

<table>
<thead>
<tr>
<th>Process</th>
<th>Subprocess</th>
<th>Partons</th>
<th>$x$ range</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ell^\pm {p, n} \rightarrow \ell^\pm X$</td>
<td>$\gamma^* q \rightarrow q$</td>
<td>$q, \bar{q}, g$</td>
<td>$x \gtrsim 0.01$</td>
</tr>
<tr>
<td>$\ell^\pm n/p \rightarrow \ell^\pm X$</td>
<td>$\gamma^* d/u \rightarrow d/u$</td>
<td>$d/u$</td>
<td>$x \gtrsim 0.01$</td>
</tr>
<tr>
<td>$pp \rightarrow \mu^+ \mu^- X$</td>
<td>$u\bar{u}, d\bar{d} \rightarrow \gamma^*$</td>
<td>$\bar{q}$</td>
<td>$0.015 \lesssim x \lesssim 0.35$</td>
</tr>
<tr>
<td>$pn/\bar{p}p \rightarrow \mu^+ \mu^- X$</td>
<td>$(u\bar{d})/(u\bar{u}) \rightarrow \gamma^*$</td>
<td>$d/\bar{u}$</td>
<td>$0.015 \lesssim x \lesssim 0.35$</td>
</tr>
<tr>
<td>$\nu(\bar{\nu}) N \rightarrow \mu^- (\mu^+) X$</td>
<td>$W^* q \rightarrow q'$</td>
<td>$q, \bar{q}$</td>
<td>$0.01 \lesssim x \lesssim 0.5$</td>
</tr>
<tr>
<td>$\nu N \rightarrow \mu^- \mu^+ X$</td>
<td>$W^* s \rightarrow c$</td>
<td>$s$</td>
<td>$0.01 \lesssim x \lesssim 0.2$</td>
</tr>
<tr>
<td>$\bar{\nu} N \rightarrow \mu^+ \mu^- X$</td>
<td>$W^* \bar{s} \rightarrow \bar{c}$</td>
<td>$\bar{s}$</td>
<td>$0.01 \lesssim x \lesssim 0.2$</td>
</tr>
<tr>
<td>$e^\pm p \rightarrow e^\pm X$</td>
<td>$\gamma^* q \rightarrow q$</td>
<td>$g, q, \bar{q}$</td>
<td>$0.0001 \lesssim x \lesssim 0.1$</td>
</tr>
<tr>
<td>$e^+ p \rightarrow \bar{\nu} X$</td>
<td>$W^+ {d, s} \rightarrow {u, c}$</td>
<td>$d, s$</td>
<td>$x \gtrsim 0.01$</td>
</tr>
<tr>
<td>$e^\pm p \rightarrow e^\pm c\bar{c} X$</td>
<td>$\gamma^* c \rightarrow c, \gamma^* g \rightarrow c\bar{c}$</td>
<td>$c, g$</td>
<td>$0.0001 \lesssim x \lesssim 0.01$</td>
</tr>
<tr>
<td>$e^\pm p \rightarrow \text{jet} + X$</td>
<td>$\gamma^* g \rightarrow q\bar{q}$</td>
<td>$g$</td>
<td>$0.01 \lesssim x \lesssim 0.1$</td>
</tr>
<tr>
<td>$pp \rightarrow \text{jet} + X$</td>
<td>$gg, gg, qq \rightarrow 2j$</td>
<td>$g, q$</td>
<td>$0.01 \lesssim x \lesssim 0.5$</td>
</tr>
<tr>
<td>$pp \rightarrow (W^\pm \rightarrow \ell^\pm \nu) X$</td>
<td>$ud \rightarrow W, \bar{u}\bar{d} \rightarrow W$</td>
<td>$u, d, \bar{u}, \bar{d}$</td>
<td>$x \gtrsim 0.05$</td>
</tr>
<tr>
<td>$pp \rightarrow (Z \rightarrow \ell^+ \ell^-) X$</td>
<td>$uu, dd \rightarrow Z$</td>
<td>$d$</td>
<td>$x \gtrsim 0.05$</td>
</tr>
</tbody>
</table>
MSTW2008(NLO) vs. CTEQ6.6

Note:

CTEQ error bands comparable with MSTW 90%cl set (different definition of tolerance)

CTEQ light quarks and gluons slightly larger at small $x$ because of imposition of positivity on gluon at $Q_0^2$. 
the asymmetric sea

• the sea presumably arises when ‘primordial’ valence quarks emit gluons which in turn split into quark-antiquark pairs, with suppressed splitting into heavier quark pairs

• so we naively expect

\[ u \approx d > s > c > \ldots \]

• but why such a big d-u asymmetry? Meson cloud, Pauli exclusion, …?

The ratio of Drell-Yan cross sections for \( pp, pn \rightarrow \mu^+\mu^- + X \) provides a measure of the difference between the \( u \) and \( d \) sea quark distributions.
earliest pdf fits had SU(3) symmetry: $s(x, Q_0^2) = \bar{s}(x, Q_0^2) = \bar{u}(x, Q_0^2) = \bar{d}(x, Q_0^2)$

later relaxed to include (constant) strange suppression (cf. fragmentation):

$$s(x, Q_0^2) = \bar{s}(x, Q_0^2) = \frac{\kappa}{2} \left[ \bar{u}(x, Q_0^2) + \bar{d}(x, Q_0^2) \right]$$

with $\kappa = 0.4 - 0.5$

nowadays, dimuon production in $\nu N$ DIS (CCFR, NuTeV) allows ‘direct’ determination:

$$\frac{d\sigma}{dxdy} (\nu_\mu (\bar{\nu}_\mu) N \rightarrow \mu^+ \mu^- X) = B_c N A \frac{d\sigma}{dxdy} (\nu_\mu s (\bar{\nu}_\mu \bar{s}) \rightarrow c \mu^- (\bar{c} \mu^+) X)$$

in the range $0.01 < x < 0.4$

data seem to prefer $s(x, Q_0^2) - \bar{s}(x, Q_0^2) \neq 0$

theoretical explanation?!
\[ \frac{100\pi}{G_F M_N E_v} \frac{d\sigma}{dx dy}(\nu_N \rightarrow \mu^- \mu^- X) \text{ in GeV}^2 \]

MSTW 2008 NNLO PDF fit, \( \chi^2 = 13 \) for 21 DOF

- \( E_\nu = 88.26 \text{ GeV} \)
  - \( y = 0.324 \)

- \( E_\nu = 88.29 \text{ GeV} \)
  - \( y = 0.558 \)

- \( E_\nu = 88.29 \text{ GeV} \)
  - \( y = 0.771 \)

- \( E_\nu = 174.29 \text{ GeV} \)
  - \( y = 0.324 \)

- \( E_\nu = 174.29 \text{ GeV} \)
  - \( y = 0.558 \)

- \( E_\nu = 174.29 \text{ GeV} \)
  - \( y = 0.771 \)

- \( E_\nu = 247.00 \text{ GeV} \)
  - \( y = 0.324 \)

- \( E_\nu = 247.00 \text{ GeV} \)
  - \( y = 0.558 \)

- \( E_\nu = 247.00 \text{ GeV} \)
  - \( y = 0.771 \)
\[ \kappa(x, Q^2) = \frac{s(x, Q^2) + \bar{s}(x, Q^2)}{u(x, Q^2) + \bar{d}(x, Q^2)} \]
charm, bottom

considered sufficiently massive to allow pQCD treatment: \( g \to Q\bar{Q} \)

distinguish two regimes:
(i) \( Q^2 \sim m_H^2 \) include full \( m_H \) dependence to get correct threshold behaviour
(ii) \( Q^2 \gg m_H^2 \) treat as \( \sim \)massless partons to resum \( \alpha_S^n \log^n(Q^2/m_H^2) \) via DGLAP

**FFNS:** OK for (i) only  
**ZM-VFNS:** OK for (ii) only

consistent **GM** (=general mass)**-VFNS** now available (e.g. ACOT(\( \chi \)), Roberts-Thorne) which interpolates smoothly between the two regimes

**Note:** definition of these is tricky and non-unique (ambiguity in assignment of \( O(m_H^2/Q^2) \) contributions), and the implementation of improved treatment (e.g. in going from **MRST2006** to **MSTW 2008**) can have a big effect on light partons
charm and bottom structure functions

• MSTW 2008 uses \textit{fixed} values of $m_c = 1.4$ GeV and $m_b = 4.75$ GeV in a GM-VFNS

• can study the sensitivity of the fit to these values
MSTW 2008 uses \textit{fixed} values of $m_c = 1.4$ GeV and $m_b = 4.75$ GeV in a GM-VFNS --- can study the sensitivity of the fit to these values.
dependence on $m_c$ at NLO in 2008 fits (preliminary)

<table>
<thead>
<tr>
<th>$m_c$ (GeV)</th>
<th>$\chi^2_{global}$ 2699 pts</th>
<th>$\chi^2_{F_2}$ 83 pts</th>
<th>$\alpha_s(M^2_Z)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>2730</td>
<td>264</td>
<td>0.1181</td>
</tr>
<tr>
<td>1.2</td>
<td>2626</td>
<td>188</td>
<td>0.1187</td>
</tr>
<tr>
<td>1.3</td>
<td>2563</td>
<td>134</td>
<td>0.1194</td>
</tr>
<tr>
<td>1.4</td>
<td>2543</td>
<td>107</td>
<td>0.1202</td>
</tr>
<tr>
<td>1.5</td>
<td>2545</td>
<td>97</td>
<td>0.1208</td>
</tr>
<tr>
<td>1.6</td>
<td>2574</td>
<td>104</td>
<td>0.1214</td>
</tr>
<tr>
<td>1.7</td>
<td>2627</td>
<td>129</td>
<td>0.1221</td>
</tr>
</tbody>
</table>

- correlation between $m_c$ and $\alpha_S$
- for low $m_c$ overshoot low $Q^2$ medium $x$ data badly
- preferred value (1.4 GeV) towards lower end of pole mass determination
- (asymmetric) uncertainty from global fit of order $\pm 0.15$ GeV
- in contrast, only weak sensitivity to $m_b$
impact of Tevatron jet data on fits

• a distinguishing feature of pdf sets is whether they use (MRST/MSTW, CTEQ,...) or do not use (H1, ZEUS, Alekhin, NNPDF,...) Tevatron jet data in the fit: the impact is on the high-x gluon
(Note: Run II data requires slightly softer gluon than Run I data)

• the (still) missing ingredient is the full NNLO pQCD correction to the cross section, but not expected to have much impact in practice

D Runnable inclusive jet data (cone, R = 0.7)
MSTW 2008 NLO PDF fit (|y_R| = |y_F| = p_T, χ² = 114 for 110 pts.

Gluon distribution at Q^2 = 10^4 GeV^2

- MSTW 2008 NLO (68% C.L.)
- Fit with Tevatron Run I jet data
- Fit without any Tevatron jets
- ZEUS 2005 Jets NLO
dijet mass distribution from D0

Rominsky, DIS09
a note on $\alpha_S$

- world average value (PDG 2008):

$$\alpha_{S,NNLO}^{\overline{MS}}(M_Z^2) = 0.1176 \pm 0.002$$

- MSTW global fit value (minimum $\chi^2$):

$$\alpha_{S,NNLO}^{\overline{MS}}(M_Z^2) = 0.1171$$

- the pdf error sets are generated with $\alpha_S$ fixed at its ‘best fit’ value, therefore variation of (e.g. jets, top, etc at LHC) cross sections with $\alpha_S$ is not explicitly included in the ‘pdf error’

Note: $$\alpha_{S,NLO}^{\overline{MS}}(M_Z^2) = 0.1202$$
MSTW variable-$\alpha_S$ sets*

- allow $\alpha_S$ to vary in global fit
  \[ \alpha_{S, NNLO}^{MS,NNLO}(M_Z^2) = 0.1171^{+0.0014}_{-0.0014} \]

- for fixed $\alpha_S \pm \delta\alpha_S$, produce sets with ‘pdf errors’, as before

- note gluon $- \alpha_S$ anticorrelation at small $x$ and quark $- \alpha_S$ anticorrelation at large $x$

- use resulting sets to quantify combined ‘pdf + $\alpha_S$’ error on observables

pdf, $\alpha_S$ uncertainties in jet cross sections

Inclusive jet cross sections with MSTW 2008 NLO PDFs

Tevatron, $\sqrt{s} = 1.96$ TeV

$LHC, \sqrt{s} = 14$ TeV

$0.1 < y < 0.7$

$0.0 < y < 0.8$

fastNLO with $\mu_R = \mu_F = p_T$

$K_T$ algorithm with $D = 0.7$

PDF only

PDF + $\alpha_S$

$P_T$ (GeV)

$P_T$ (GeV)
Note: still dominated by scale variation uncertainty - see earlier
pdfs at LHC:
— impact on precision phenomenology
— can we constrain pdfs further?
where $X=W, Z, H, \text{high-}E_T \text{ jets, SUSY sparticles, black hole, ...}$, and $Q$ is the ‘hard scale’ (e.g. $= M_X$), usually $\mu_F = \mu_R = Q$, and $\hat{\sigma}$ is known ...

- to some fixed order in pQCD (and EW), e.g. high-$E_T$ jets

\[
\hat{\sigma} = A\alpha_S^2 + B\alpha_S^3
\]

- or ‘improved’ by some leading logarithm approximation (LL, NLL, ...) to all orders via resummation

---

**the QCD factorization theorem** for hard-scattering (short-distance) inclusive processes

\[
\sigma_X = \sum_{a,b} \int_0^1 dx_1 dx_2 f_a(x_1, \mu_F^2) f_b(x_2, \mu_F^2) \\
	\times \hat{\sigma}_{ab\to X} \left( x_1, x_2, \{p_1^\mu\}; \alpha_S(\mu_R^2), \alpha(\mu_R^2), \frac{Q^2}{\mu_R^2}, \frac{Q^2}{\mu_F^2} \right)
\]
momentum fractions $x_1$ and $x_2$ determined by mass and rapidity of $X$
pdfs at LHC – the issues

• high precision cross section predictions require accurate knowledge of pdfs: \( \delta\sigma_{\text{th}} = \delta\sigma_{\text{pdf}} + \ldots \)
  \( \rightarrow \) improved signal and background predictions
  \( \rightarrow \) easier to spot new physics

• ‘standard candle’ processes (e.g. \( \sigma_Z \)) to
  - check formalism (factorisation, DGLAP, …)
  - measure machine luminosity?

• learning more about pdfs from LHC measurements. e.g.
  - high-\( E_T \) jets \( \rightarrow \) gluon?
  - \( W^+, W^-, Z^0 \) \( \rightarrow \) quarks?
  - forward Drell-Yan \( \rightarrow \) small \( x \)?
  - …
how important is pdf precision?

**Example 1**: \( \sigma(M_H=120 \text{ GeV}) @ \text{LHC} \)

- \( \delta \sigma_{pdf} \approx \pm 2\% \), \( \delta \sigma_{ptNNLO} \approx \pm 10\% \)
- \( \delta \sigma_{ptNLL} \approx \pm 8\% \)

\[ \rightarrow \delta \sigma_{\text{theory}} \approx \pm 10\% \]


**Example 2**: \( \sigma(Z^0) @ \text{LHC} \)

- \( \delta \sigma_{pdf} \approx \pm 2\% \), \( \delta \sigma_{ptNNLO} \approx \pm 2\% \)

\[ \rightarrow \delta \sigma_{\text{theory}} \approx \pm 3\% \]
• Example 3: $\sigma(tt) \ @ \ LHC$

$\delta\sigma_{pdf} \approx \pm 2\%, \ \delta\sigma_{ptNNLO_{approx}} \approx \pm 3\%$

$\rightarrow \ \delta\sigma_{\text{theory}} \approx \pm 4\%$
pdf uncertainty on $\sigma(gg \rightarrow H)$

$\delta \sigma / \sigma$ typically ± 2-3% pdf uncertainty, except near edges of phase space
parton luminosity functions

• a quick and easy way to assess the mass and collider energy dependence of production cross sections

\[ \hat{\sigma}_{ab \rightarrow X} = C_X \delta(\hat{s} - M^2) \]
\[ \sigma_X = \int_0^1 dx_a dx_b f_a(x_a, M^2) f_b(x_b, M^2) C_X \delta(x_a x_b - \tau) \equiv C_X \left[ \frac{1}{s} \frac{\partial \mathcal{L}_{ab}}{\partial \tau} \right] \quad (\tau = M^2/s) \]
\[ \frac{\partial \mathcal{L}_{ab}}{\partial \tau} = \int_0^1 dx_a dx_b f_a(x_a, M^2) f_b(x_b, M^2) \delta(x_a x_b - \tau) \]

• i.e. all the mass and energy dependence is contained in the X-independent parton luminosity function in \[ \]
• useful combinations are \( ab = gg, \sum_q q\bar{q}, \ldots \)
• and also useful for assessing the uncertainty on cross sections due to uncertainties in the pdfs
\[ \begin{align*}
\frac{dL}{d\hat{s}} & = gg \\
\sum_i (gq_i + g\bar{q}_i + q_i g + \bar{q}_i g) \\
\sum_i (q_i \bar{q}_i + \bar{q}_i q_i)
\end{align*} \]

Huston, Campbell, S (2007)
parton luminosity (68%cl) uncertainties at LHC

\[ G = g + \frac{4}{9} \sum_q (q + \bar{q}) \]
LHC at 7 and 10 TeV

ratios of parton luminosities
at 7 TeV, 10 TeV and 14 TeV LHC

MSTW2008NLO

\[ \Sigma q\bar{q} \]

\[ gg \]

\[ M_X (\text{GeV}) \]

\[ 10^2 \text{ to } 10^3 \]
comparison of gg luminosities at LHC (7 TeV) with 90%cl pdf uncertainty bands
pdfs at LHC – the issues

• high precision cross section predictions require accurate knowledge of pdfs: \( \delta \sigma_{th} = \delta \sigma_{pdf} + \ldots \)
  → improved signal and background predictions
  → easier to spot new physics

• ‘standard candle’ processes (e.g. \( \sigma_Z \)) to
  – check formalism (factorisation, DGLAP, …)
  – measure machine luminosity?

• learning more about pdfs from LHC measurements. e.g.
  – high-\( E_T \) jets \( \rightarrow \) gluon?
  – \( W^+, W^-, Z^0 \) \( \rightarrow \) quarks?
  – forward DY \( \rightarrow \) small \( x \)?
  – …
standard candles: $\sigma(W,Z) \at \text{LHC}$

- cross sections (total and rapidity distributions) known to NNLO pQCD and NLO EW; perturbation series seems to be converging quickly

- EW parameters well measured at LEP

- samples pdfs where they are well measured (in $x$) in DIS

- … although the mix of quark flavours is different: $F_2$ and $\sigma(W,Z)$ probe different combinations of $u,d,s,c,b \rightarrow$ sea quark distributions important

- precise measurement of cross section ratios at LHC (e.g. $\sigma(W^+)/\sigma(W^-)$, $\sigma(W)/\sigma(Z)$) will allow these subtle effects to be explored further
at LHC, ~30% of W and Z total cross sections involves s,c,b quarks
Note: at NNLO, factorisation and renormalisation scale variation \(M/2\) → 2\(M\) gives an additional ± 2% change in the LHC cross sections
predictions for $\sigma(W,Z) @ LHC$ (Tevatron)

<table>
<thead>
<tr>
<th></th>
<th>$B_u \cdot \sigma_W$ (nb)</th>
<th>$B_\parallel \cdot \sigma_Z$ (nb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSTW 2008 NLO</td>
<td>21.17 (2.659)</td>
<td>2.001 (0.2426)</td>
</tr>
<tr>
<td>MSTW 2008 NNLO</td>
<td>21.72 (2.747)</td>
<td>2.051 (0.2507)</td>
</tr>
<tr>
<td>MRST 2006 NLO</td>
<td>21.21 (2.645)</td>
<td>2.018 (0.2426)</td>
</tr>
<tr>
<td>MRST 2006 NNLO</td>
<td>21.51 (2.759)</td>
<td>2.044 (0.2535)</td>
</tr>
<tr>
<td>MRST 2004 NLO</td>
<td>20.61 (2.632)</td>
<td>1.964 (0.2424)</td>
</tr>
<tr>
<td>MRST 2004 NNLO</td>
<td>20.23 (2.724)</td>
<td>1.917 (0.2519)</td>
</tr>
<tr>
<td>CTEQ6.6 NLO</td>
<td>21.58 (2.599)</td>
<td>2.043 (0.2393)</td>
</tr>
<tr>
<td>Alekhin 2002 NLO</td>
<td>21.32 (2.733)</td>
<td>2.001 (0.2543)</td>
</tr>
<tr>
<td>Alekhin 2002 NNLO</td>
<td>21.13 (2.805)</td>
<td>1.977 (0.2611)</td>
</tr>
</tbody>
</table>

MSTW
predictions for $\sigma(W,Z)$ @ Tevatron, LHC

- MRST/MSTW NNLO: 2008 ~ 2006 > 2004 mainly due to changes in treatment of \textit{charm}
- CTEQ: 6.6 ~ 6.5 > 6.1 due to changes in treatment of $s,c,b$
- NLO: CTEQ6.6 2\% higher than MSTW 2008 at LHC, because of slight differences in quark (u,d,s,c) pdfs, difference within quoted uncertainty
\[ R_\pm = \frac{\sigma(W^+ \to l^+ \nu)}{\sigma(W^- \to l^- \nu)} \]

\[
R_\pm \approx \frac{u(x_1)\bar{d}(x_2) + c(x_1)\bar{s}(x_2) + (1 \leftrightarrow 2)}{d(x_1)\bar{u}(x_2) + s(x_1)\bar{c}(x_2) + (1 \leftrightarrow 2)}
\]

\[ \delta \sigma_{\text{th}} \approx \delta \sigma_{\text{pdf}} \approx \pm 1\%, \]

\[ \delta \sigma_{\text{expt}} \approx ??? \]
$W \rightarrow l \nu_l$ charge asymmetry at LHC

$W^-$

$W^+$

Using MSTW2007 NLO (prel.) PDFs [arXiv:0706.0455]

Thin blue lines: LO FEWZ with $\Gamma_W = 0$

Thick red lines: NLO FEWZ with $\Gamma_W = 2.1$ GeV

$p_T > 10$ GeV

$10 < p_{T,t} < 25$ GeV

$25 < p_{T,t} < 35$ GeV

$35 < p_{T,t} < 45$ GeV
pdf uncertainty on \( d\sigma(dY)/dM dy \)
\( d\sigma(W^+)/dy, d\sigma(W)/dy, d\sigma(Z)/dy \)
at LHC using MSTW2008NLO (68%cl)

\( M = 8 \text{ GeV} \)
\( 16 \text{ GeV} \)
\( 24 \text{ GeV} \)

pdf uncertainty on
\( R_{WZ} = d\sigma(W)/d\sigma(Z) \), \( R_{WW} = d\sigma(W^+)/d\sigma(W) \)
\( A_{W} = (d\sigma(W^+)/d\sigma(W^+)-d\sigma(W^+)/d\sigma(W^+)) \)
at LHC using MSTW2008NLO (68%cl)
using the $W^{\pm}$ charge asymmetry at the LHC

- at the Tevatron $\sigma(W^+) = \sigma(W^-)$, whereas at LHC $\sigma(W^+) \sim 1.3\sigma(W^-)$

- can use this asymmetry to calibrate backgrounds to new physics, since typically $\sigma_{NP}(X \rightarrow W^+ + \ldots) = \sigma_{NP}(X \rightarrow W^- + \ldots)$

- example:
  
  \[ gg \rightarrow t\bar{t} \rightarrow W^+ W^- b\bar{b} \rightarrow W^\pm (\rightarrow l^\pm + \nu) + 4\text{jets} \]

  in this case
  
  \[ \sigma_{signal}(W^+ + 4\text{jets}) = \sigma_{signal}(W^- + 4\text{jets}) \]

  whereas...
  
  \[ \sigma_{QCD\text{bgd}}(W^+ + 4\text{jets}) \neq \sigma_{QCD\text{bgd}}(W^- + 4\text{jets}) \]

  which can in principle help distinguish signal and background
pdfs at LHC – the issues

• high precision cross section predictions require accurate knowledge of pdfs: \( \delta \sigma_{\text{th}} = \delta \sigma_{\text{pdf}} + \ldots \)
  → improved signal and background predictions
  → easier to spot new physics

• ‘standard candle’ processes (e.g. \( \sigma_Z \)) to
  — check formalism (factorisation, DGLAP, …)
  — measure machine luminosity?

• learning more about pdfs from LHC measurements. e.g.
  — high-\( E_T \) jets \( \rightarrow \) gluon?
  — \( W^+, W^-, Z^0 \) \( \rightarrow \) quarks?
  — forward DY \( \rightarrow \) small \( x \)?
  — …
impact of LHC measurements on pdfs

- the standard candles: central $\sigma(W,Z,tt,jets)$ as a probe and test of pdfs in the $x \sim 10^{-2\pm1}$, $Q^2 \sim 10^{4-6}$ GeV$^2$ range where most New Physics is expected (H, SUSY, ....)

- forward production of (relatively) low-mass states (e.g. $\gamma^*$,dijets,...) to access partons at $x<<1$ (and $x\sim1$)
Unique features

• pseudo-rapidity range 1.9 - 4.9
  — 1.9 - 2.5 complementary to ATLAS/CMS
  — > 2.5 unique to LHCb

• beam defocused at LHCb: 1 year of running = 2 fb⁻¹

• trigger on low momentum muons: \( p > 8 \text{ GeV}, p_T > 1 \text{ GeV} \)

  access to unique range of \((x,Q^2)\)
LHCb

→ detect forward, low $p_T$ muons from $q\bar{q} \rightarrow \mu^+ \mu^-$
Impact of 1 fb⁻¹ LHCb data for forward Z and $\gamma^*$ ($M = 14$ GeV) production on the gluon distribution uncertainty.
summary

• precision phenomenology at high-energy colliders such as the LHC requires an accurate knowledge of the distribution functions of partons in hadrons

• determining pdfs from global fits to data is now a major industry… the MSTW collaboration has released its latest (2008) LO, NLO, NNLO sets

• pdf uncertainty for ‘new physics’ cross sections not expected to be too important (few % level), apart from at very high mass

• ongoing high-precision studies of standard candle cross sections and ratios (→ new PDF4LHC benchmarking initiative)

• potential of LHCb to probe very small $x$ via low-mass Drell-Yan

• 7 TeV LHC is already interesting!
extra slides
**Tevatron parton kinematics**

\[ x_{1,2} = (M/1.96 \text{ TeV}) \exp(\pm y) \]

\[ Q = M \]

- \( M = 1 \text{ TeV} \)
- \( M = 100 \text{ GeV} \)
- \( M = 10 \text{ GeV} \)

**LHC parton kinematics**

\[ x_{1,2} = (M/14 \text{ TeV}) \exp(\pm y) \]

\[ Q = M \]

- \( M = 10 \text{ TeV} \)
- \( M = 1 \text{ TeV} \)
- \( M = 100 \text{ GeV} \)
- \( M = 10 \text{ GeV} \)

**Graphs:**

- **Tevatron**
  - HERA
  - Fixed target

- **LHC**
  - HERA
  - Fixed target
Higgs ($M_H = 120$ GeV) with MSTW 2008 NNLO PDFs

**Tevatron, $\sqrt{s} = 1.96$ TeV**

**LHC, $\sqrt{s} = 14$ TeV**

- $\alpha_s^2(M_Z^2)$
- $gg$ luminosity

68% C.L. uncertainties

**Gluon at $Q^2 = M_H^2 = (120$ GeV)$^2**

- MSTW 2008 NNLO (68% C.L.)
- Fix $\alpha_s$ at +68% C.L. limit
- Fix $\alpha_s$ at -68% C.L. limit

Ratio to MSTW 2008 NNLO

$y = 0$
- at LHC
- at Tevatron

$x$
MSTW 2008 NNLO ($\alpha_S$) PDF fit

$\alpha_S(M_Z^2)$ vs. dataset labels.

- 90% C.L.
- 68% C.L.
- Global min.
The graph shows the ratio of the Higgs production cross-sections calculated using different PDF sets and with and without a specific factor.

- **Tevatron**: Shows the ratio for production at the Tevatron collider.
- **LHC**: Shows the ratio for production at the LHC collider.

The cross-sections are expressed as a fraction of the reference cross-sections calculated using MSTW2008 PDFs. The solid blue line represents the LO gg → H cross-section with NNLO PDFs, while the dashed blue line represents the LO gg → H cross-section without the overall $\alpha_s^2$ factor. The solid red line represents the LO gg → H cross-section with NNLO PDFs, and the dashed red line represents the LO gg → H cross-section without the overall $\alpha_s^2$ factor.
\( \gamma^* \rightarrow \mu\mu \) selection

- 4-vector Pythia + LHCb acceptance

Dominant at Z but large backgrounds at low masses

8th April 2008
Ronan McNulty et al, at DIS08
using the $W^{+}$ charge asymmetry at the LHC

- at the Tevatron $\sigma(W^+) = \sigma(W^-)$, whereas at LHC $\sigma(W^+) \sim 1.3\sigma(W^-)$

- can use this asymmetry to calibrate backgrounds to new physics, since typically $\sigma_{NP}(X \rightarrow W^+ + \ldots) = \sigma_{NP}(X \rightarrow W^- + \ldots)$

- example:

\[
gg \rightarrow t\bar{t} \rightarrow W^+ W^- b\bar{b} \rightarrow W^{\pm}(\rightarrow l^{\pm} + \nu) + 4\text{jets}
\]

in this case

$\sigma_{\text{signal}}(W^+ + 4\text{jets}) = \sigma_{\text{signal}}(W^- + 4\text{jets})$

whereas...

$\sigma_{\text{QCDbkgd}}(W^+ + 4\text{jets}) \neq \sigma_{\text{QCDbkgd}}(W^- + 4\text{jets})$

which can in principle help distinguish signal and background
for $n_{\text{jet}} > 1$ dominant subprocess is:

\[
g + q^+ \rightarrow W^+ + q^- + ng \\
 g + q^- \rightarrow W^- + q^+ + ng
\]

<table>
<thead>
<tr>
<th></th>
<th>%qq</th>
<th>%qg</th>
<th>%gg</th>
</tr>
</thead>
<tbody>
<tr>
<td>W+0j</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>W+1j</td>
<td>75</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>W+2j</td>
<td>18</td>
<td>75</td>
<td>7</td>
</tr>
<tr>
<td>W+3j</td>
<td>18</td>
<td>72</td>
<td>10</td>
</tr>
</tbody>
</table>
$W^+/W^-$ ratio:

- very sensitive to $u/d$ pdf ratio
- varies with $y_W$
- depends slightly on $n_{\text{jet}}$ and $E_{Tj}(\text{min})$
- fairly independent of scale choice etc
$70 \text{ GeV} \times 7 \text{ TeV ep}$
LHC parton kinematics

\[ x_{1,2} = (M/14 \text{ TeV}) \exp(\pm y) \]

\[ Q = M \]

\[ M = 10 \text{ TeV} \]

\[ M = 1 \text{ TeV} \]

\[ M = 100 \text{ GeV} \]

\[ M = 10 \text{ GeV} \]

\[ y = 6, 4, 2, 0, 2, 4, 6 \]

\[ Q^2 \text{ (GeV}^2) \]

\[ x \]

LHeC

HERA

fixed target
MSTW2008 vs MRST2006
MSTW2008 vs Alekhin2002

- Up valence distribution at $Q^2 = 10^6 \text{ GeV}^2$
- Down valence distribution at $Q^2 = 10^6 \text{ GeV}^2$
- Up quark distribution at $Q^2 = 10^6 \text{ GeV}^2$
- Down quark distribution at $Q^2 = 10^6 \text{ GeV}^2$
- Strange quark distribution at $Q^2 = 10^6 \text{ GeV}^2$
- Gluon distribution at $Q^2 = 10^6 \text{ GeV}^2$
MSTW2008 vs NNPDF1.0

Up valence distribution at $Q^2 = 10^2$ GeV$^2$

- MSTW 2008 NLO (68% C.L.)
- NNPDF1.0 (1000 replicas)

Up antiquark distribution at $Q^2 = 10^4$ GeV$^2$

- MSTW 2008 NLO (68% C.L.)
- NNPDF1.0 (1000 replicas)
\[ R(W/Z) = \frac{\sigma(W)}{\sigma(Z)} @ \text{Tevatron & LHC} \]

CDF 2007: \[ R = 10.84 \pm 0.15 \text{ (stat)} \pm 0.14 \text{ (sys)} \]
scaling violations measured at HERA
pdf uncertainties

\[ \Delta \chi^2_{\text{global}} \equiv \chi^2_{\text{global}} - \chi^2_{\text{min}} = \sum_{i,j=1}^{n} H_{ij} (a_i - a_0^i) (a_j - a_0^j) \]

\[ \vec{a} - \vec{a}^0 = \sum_{k=1, n} z_k \vec{e}_k \text{ where } (H^{-1}) \cdot \vec{e}_k = \lambda_k \vec{e}_k, \quad \vec{e}_k \cdot \vec{e}_l = \lambda_k \delta_{kl} \]

then \[ \Delta \chi^2_{\text{global}} = \sum_{k=1, n} z_k^2 \leq T^2 \quad (T = \text{tolerance}) \]

this defines a set of n ‘error’ pdfs, spanning the allowed variation in the parameters, as determined by \( T \):

\[ \vec{a} (S_k^\pm) = \vec{a}^0 \pm T \vec{e}_k \]

rather than using a fixed value of \( T \) (cf. MRST, CTEQ), we determine the ‘dynamic’ tolerance for each eigenvector from the condition that all data sets should be described within their 68% or 90% or … confidence limit
similar, but not identical, to published CDF Run-2 midpoint jet data.
CDF Run II inclusive jet data, $\chi^2 = 119$ for 72 pts.

<table>
<thead>
<tr>
<th>CDFI [179] (33 pts.)</th>
<th>DØI [176] (90 pts.)</th>
<th>CDFII [54] (76 pts.)</th>
<th>DØII [56] (110 pts.)</th>
<th>$\Delta\chi^2_{\text{non-jet}}$ (2513 pts.)</th>
<th>$\alpha_S(M^2_Z)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>53</td>
<td>119</td>
<td>64</td>
<td>117</td>
<td>0</td>
<td>0.1197</td>
</tr>
<tr>
<td><strong>51</strong></td>
<td><strong>48</strong></td>
<td><strong>132</strong></td>
<td><strong>180</strong></td>
<td>9</td>
<td>0.1214</td>
</tr>
<tr>
<td>56</td>
<td>110</td>
<td><strong>56</strong></td>
<td><strong>114</strong></td>
<td>2</td>
<td>0.1202</td>
</tr>
<tr>
<td><strong>53</strong></td>
<td><strong>85</strong></td>
<td><strong>68</strong></td>
<td><strong>117</strong></td>
<td>1</td>
<td>0.1204</td>
</tr>
</tbody>
</table>

Table 8: $\chi^2$ values for Tevatron data on inclusive jet production from various global fits. The large bold values indicate the data sets explicitly included in each fit. The “MSTW 2008” fit includes only the Run II data. We also indicate the increase in the $\chi^2$ for all data sets other than Tevatron inclusive jets when these data are added to the fit. Finally, the value of $\alpha_S(M^2_Z)$ is given for each fit. All fits are carried out at NLO with a scale choice of $\mu_R = \mu_F = p_T$. 