# **P**arton **D**istribution **F**unctions (Part Deux)

Graeme Watt

CERN PH-TH

CERN Theory Group Retreat
*Les Houches Centre de Physique*
4th November 2011

# Introduction: Parton Distribution Functions

http://www.math.columbia.edu/∼woit/wordpress/?p=3715

"String theory has always been much more popular than phenomenology. Given the choice of working on the theory that unifies all physics and explains the big bang or computing grungy things about parton distribution functions, it's easy to see why this is." [Peter Woit, comment on his blog "*Not Even Wrong*", 19th May 2011]

# Introduction: **P**arton **D**istribution **F**unctions

http://www.math.columbia.edu/~woit/wordpress/?p=3715

"String theory has always been much more popular than phenomenology. Given the choice of working on the theory that unifies all physics and explains the big bang or computing grungy things about parton distribution functions, it's easy to see why this is." [Peter Woit, comment on his blog "*Not Even Wrong*", 19th May 2011]

## This talk:

Computing grungy
things about **P**arton **D**istribution **F**unctions.

# Introduction: **P**arton **D**istribution **F**unctions

http://www.math.columbia.edu/~woit/wordpress/?p=3715

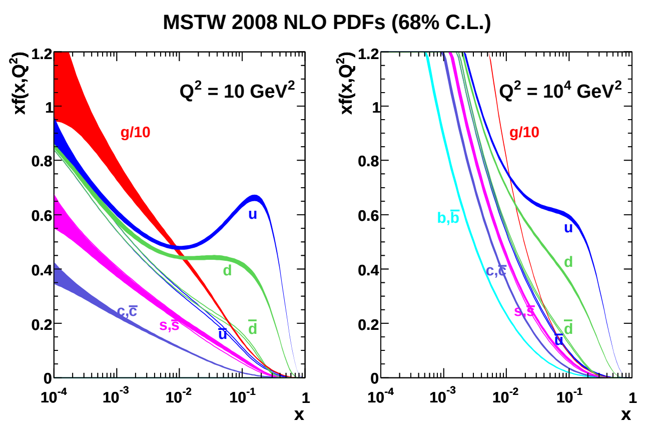
"String theory has always been much more popular than phenomenology. Given the choice of working on the theory that unifies all physics and explains the big bang or computing grungy things about parton distribution functions, it's easy to see why this is." [Peter Woit, comment on his blog "*Not Even Wrong*", 19th May 2011]

## This talk:

Computing ~~grungy~~ necessary and interesting things about **P**arton **D**istribution **F**unctions.

# MSTW 2008 PDFs [http://projects.hepforge.org/mstwpdf/]

A. D. **M**artin, W. J. **S**tirling, R. S. **T**horne, G. **W**att [arXiv:0901.0002]




- **Error bands** shown are obtained from propagation of experimental uncertainties on the fitted data points.

# Benchmark comparisons of PDFs from different groups

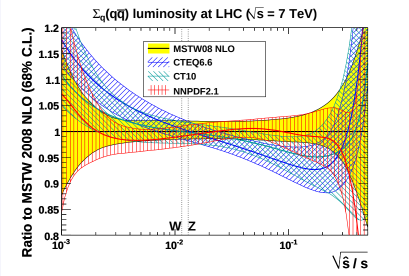
- Various fitting groups currently produce PDF sets: **MSTW**, CTEQ-TEA, NNPDF, HERAPDF, ABKM, (G)JR.
- Results often do not agree within the quoted uncertainties.
- Quantifying and understanding differences *between* groups is *as* (if not *more*) important as continued improvements *within* groups.
- Systematic comparisons performed over last couple of years [G.W., JHEP **09** (2011) 069, arXiv:1106.5788].
- Apparent major differences from classes of data fitted:

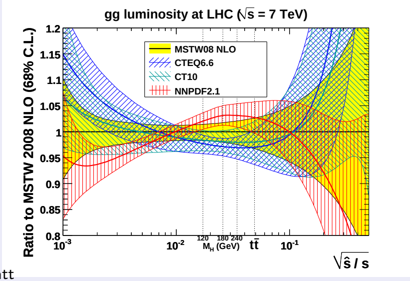
| | **MSTW08** | **CT10** | **NNPDF2.1** | HERAPDF1.5 | ABKM09 | JR09 |
|---|---|---|---|---|---|---|
| HERA DIS | ✔ | ✔ | ✔ | ✔ | ✔ | ✔ |
| Fixed-target DIS | ✔ | ✔ | ✔ | ✗ | ✔ | ✔ |
| Fixed-target DY | ✔ | ✔ | ✔ | ✗ | ✔ | ✔ |
| Tevatron $W$,$Z$ | ✔ | ✔ | ✔ | ✗ | ✗ | ✗ |
| Tevatron jets | ✔ | ✔ | ✔ | ✗ | ✗ | ✗ |

- "Global" $\equiv$ includes all five main categories of data.
- Next generation of global fits will also include **LHC** data.

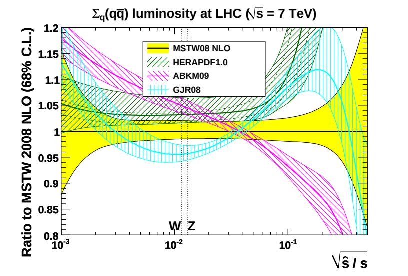
# Partonic luminosities [G.W., JHEP **09** (2011) 069, arXiv:1106.5788]

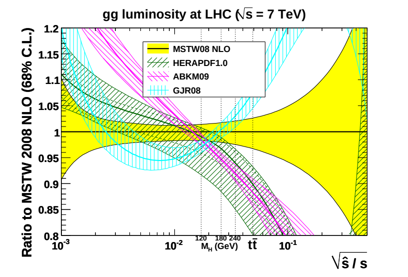
## MSTW vs. other global fits




## MSTW vs. non-global fits

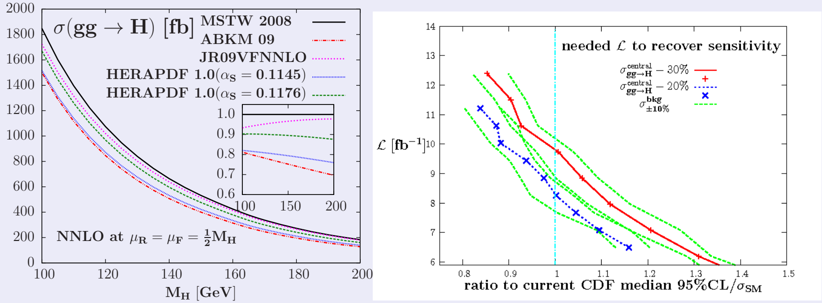


# Tevatron Higgs exclusion limits: a critical appraisal

- $\sigma(gg \to H)$ depends on gluon distribution at $x \sim M_H/\sqrt{s}$.
- Tevatron and LHC exclusion based on **MSTW 2008** PDFs.
- PDF+$\alpha_S$ uncertainty at 90% C.L. is $\approx$ 10% at Tevatron.
- But some **non-global** PDF sets can give much *lower* $\sigma(gg \to H)$.

[J. Baglio, A. Djouadi, S. Ferrag, R. M. Godbole, arXiv:1101.1832]

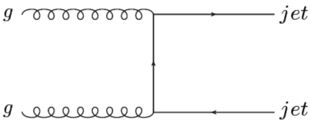


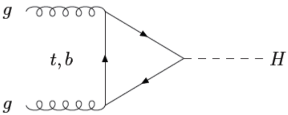
# Jets as a discriminator of the high-$x$ gluon distribution

**JETS**

PV: Any PDF should reproduce jet data if being used for Higgs

Closest observable to Higgs in terms of Luminosity, kinematics and power of coupling!



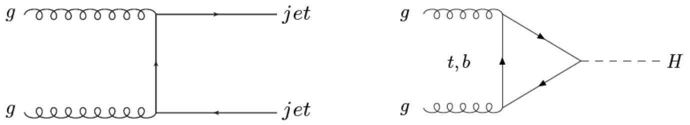

[D. de Florian, talk at "Higgs Hunting 2011", Orsay, France, 28th July 2011]

# Jets as a discriminator of the high-$x$ gluon distribution

***JETS***

PV: Any PDF should reproduce jet data if being used for Higgs

Closest observable to Higgs in terms of Luminosity, kinematics and power of coupling!




[D. de Florian, talk at "Higgs Hunting 2011", Orsay, France, 28th July 2011]

- Do PDFs from non-global fits describe Tevatron jet data?
  [R. S. Thorne and G.W., JHEP **08** (2011) 100, `arXiv:1106.5789`]
- Values of $\chi^2/N_{\rm pts.}$ for CDF Run II inclusive jet data:

| NNLO PDF | $\alpha_S(M_Z^2)$ | $\mu = p_T/2$ | $\mu = p_T$ | $\mu = 2p_T$ |
|---|---|---|---|---|
| **MSTW08** | 0.1171 | 1.39 | **0.69** | 0.97 |
| **ABKM09** | 0.1135 | 2.55 | **2.76** | 3.41 |

# Experimental error propagation in global PDF fits

## Hessian method (used by MSTW and CTEQ groups)

- Based on *covariance* (inverse *Hessian*) matrix of ≈20 fit parameters: diagonalise to produce *eigenvector* PDF sets.
- Varying values of tolerance $T = \sqrt{\Delta\chi^2}$ used to accommodate minor data inconsistencies (average $T \approx 3$ in MSTW08 fit).

## Monte Carlo sampling using data replicas (used by NNPDF)

- Generate replica data sets with shifted central values:

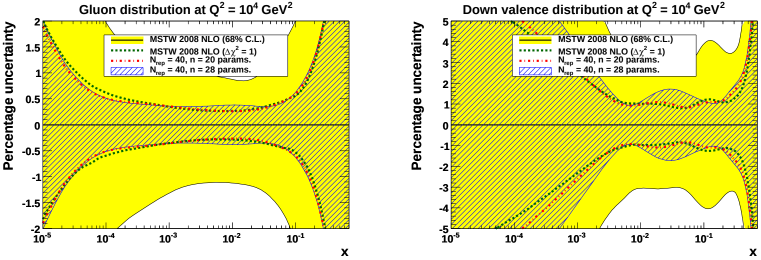
$$D_{m,i} \to \left( D_{m,i} + R_{m,i}^{\mathrm{uncorr.}}\, \sigma_{m,i}^{\mathrm{uncorr.}} + \sum_{k=1}^{N_{\mathrm{corr.}}} R_{m,k}^{\mathrm{corr.}}\, \sigma_{m,k,i}^{\mathrm{corr.}} \right) \cdot \left(1 + R_m^{\mathcal{N}}\, \sigma_m^{\mathcal{N}}\right)$$

- Calculate average and s.d. over $N_{\mathrm{rep}} \sim \mathcal{O}(100)$ PDF sets.
- Equivalent to Hessian method with tolerance $T = 1$.
- But more useful when fitting weakly-constrained parameters.

**Puzzle:** NNPDF *w/o* tolerance ≈ MSTW/CTEQ *with* tolerance.

# Hessian versus Monte Carlo sampling in MSTW fit

- Compare Hessian **with** or without tolerance (20 parameters) to Monte Carlo sampling (either 20 or 28 parameters).

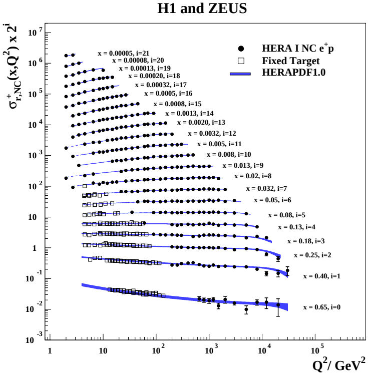



- MC sampling with 20 parameters $\approx$ Hessian with $\Delta\chi^2 = 1$.
- Moving from 20 $\to$ 28 parameters using MC sampling gives indication of parameterisation bias, i.e. moving from 4 $\to$ 7 parameters for input $xg$ and 3 $\to$ 4 parameters for input $xd_v$.
- **Conclusion:** parameterisation bias is likely to be small.

# Summary of recent work

- Detailed PDF comparisons of benchmark LHC cross sections
  ⇒ largest differences from global versus non-global fits
  [G.W., JHEP **09** (2011) 069, arXiv:1106.5788].
- PDF dependence of Higgs cross sections at Tevatron and LHC
  ⇒ importance of Tevatron jets to pin down high-$x$ gluon
  [R. S. Thorne and G.W., JHEP **08** (2011) 100, arXiv:1106.5789].
- *Monte Carlo sampling* in context of **MSTW** fit
  ⇒ parameterisation bias is likely to be small [work in progress].
- **Future:** making use of LHC data to constrain PDFs.

# Scaling violations of DIS structure functions




$$\boxed{F_2^{\mathrm{LO}}(x,Q^2) = \sum_q e_q^2\, x f_{q/p}(x,Q^2)}$$

Differentiate and insert DGLAP:

$$\Rightarrow \frac{\partial F_2^{\mathrm{LO}}(x,Q^2)}{\partial \ln Q^2} = \frac{\alpha_S}{2\pi}\left[ P_{q\leftarrow q}^{\mathrm{LO}} \otimes F_2^{\mathrm{LO}} + \sum_q e_q^2\, P_{q\leftarrow g}^{\mathrm{LO}} \otimes f_{g/p} \right]$$

- $F_2$ falls with $Q^2$ at large $x$ and rises with $Q^2$ at small $x$.
- $\Rightarrow$ Measure *scaling violations* to constrain small-$x$ gluon.
- **But** no direct constraint on large-$x$ gluon from inclusive DIS: only momentum sum rule and *imposed* parameterisation constraint (non-global fits typically only have 2 or 3 gluon parameters).

# Input parameterisation in MSTW 2008 NLO fit

## Input parameterisation ($Q_0^2 = 1$ GeV$^2$) in MSTW 2008 fit

$$xu_v = A_u\, x^{\eta_1}(1-x)^{\eta_2}(1+\epsilon_u\sqrt{x}+\gamma_u\, x)$$
$$xd_v = A_d\, x^{\eta_3}(1-x)^{\eta_4}(1+\epsilon_d\sqrt{x}+\gamma_d\, x)$$
$$xS = A_S\, x^{\delta_S}(1-x)^{\eta_S}(1+\epsilon_S\sqrt{x}+\gamma_S\, x)$$
$$x(\bar{d}-\bar{u}) = A_\Delta\, x^{\eta_\Delta}(1-x)^{\eta_S+2}(1+\gamma_\Delta\, x+\delta_\Delta\, x^2)$$
$$xg = A_g\, x^{\delta_g}(1-x)^{\eta_g}(1+\epsilon_g\sqrt{x}+\gamma_g\, x)+A_{g'}\, x^{\delta_{g'}}(1-x)^{\eta_{g'}}$$
$$x(s+\bar{s}) = A_+\, x^{\delta_S}\,(1-x)^{\eta_+}(1+\epsilon_S\sqrt{x}+\gamma_S\, x)$$
$$x(s-\bar{s}) = A_-\, x^{0.2}(1-x)^{\eta_-}(1-x/x_0)$$

- $A_u$, $A_d$, $A_g$ and $x_0$ are determined from sum rules.
- 28 parameters allowed to go free to find overall best fit.
  **But** restrict to 20 parameters for Hessian error propagation.
- MC sampling allows all 28 parameters free $\Rightarrow$ check potential bias.