Parton Distribution Functions
- lecture 1 -

Alberto Accardi
Hampton U. and Jefferson Lab

CTEQ summer school 2013

Plan of the lectures

- **Lecture 1 – Global PDF fits**
  - Why parton distributions?
  - Global PDF fits
  - Observables
  - The art of fitting

- **Lecture 2 – PDF uncertainties and applications**
  - Statistical and theretical uncertainties
  - Modern PDF sets
  - LHC applications
  - Large-*x* connections

Resources

Articles

- J.Rojo, “Parton distributions in the Higgs boson era “, arXiv:1305.3476
- A.Accardi, “The CJ12 parton distribution functions”, DIS2013 proceedings
  (see school's website)

Reviews

- P.Jimenez-Delgado, W.Melnitchouk, J.F.Owens,
  “Momentum and helicity distributions in the nucleon”, arXiv:1306.6515
- Forte, Watt, “Progress in partonic structure of proton”, arXiv:1301.6754

Lectures

- J.F. Owens' lectures, “Intro to parton model and pQCD”, 2013 summer school
- J.F. Owens, “PDF and global fitting”, 2007 summer school
- W.K.Tung, “pQCD and parton structure of the nucleon”, CTEQ website

Resources

- **Textbooks (mostly on DIS...)**
  - Halzen, Martin, “Quarks and leptons,” John Wiley and sons, 1984
  - Lenz et al. (Eds.), “Lectures on QCD. Applications,” Springer, 1997
    - esp. lectures by Levy, Rith, Jaffe
  - Devenish, Cooper-Sarkar, “Deep Inelastic Scattering,” Oxford U.P., 2004
  - Feynman, “Photon-hadron interactions,” Addison Wesley, 1972

...and a special thank to Jeff Owens,
who has been teaching me this business...

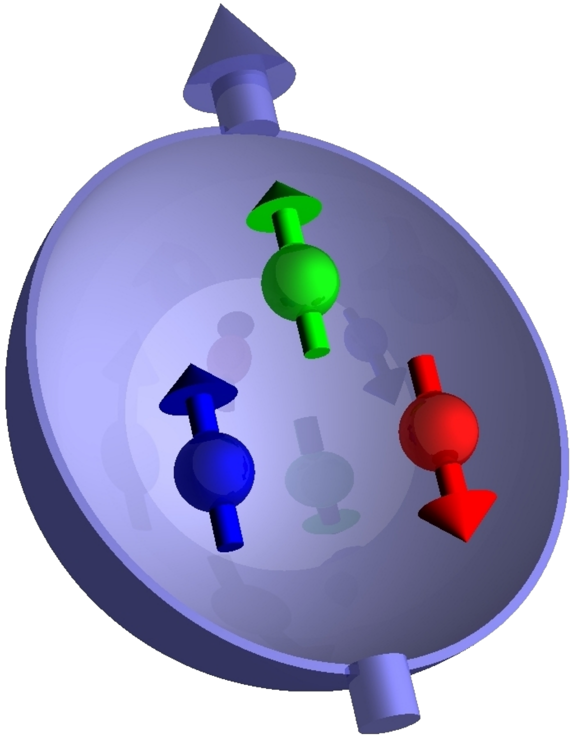
Lecture 1 - Global PDF fits

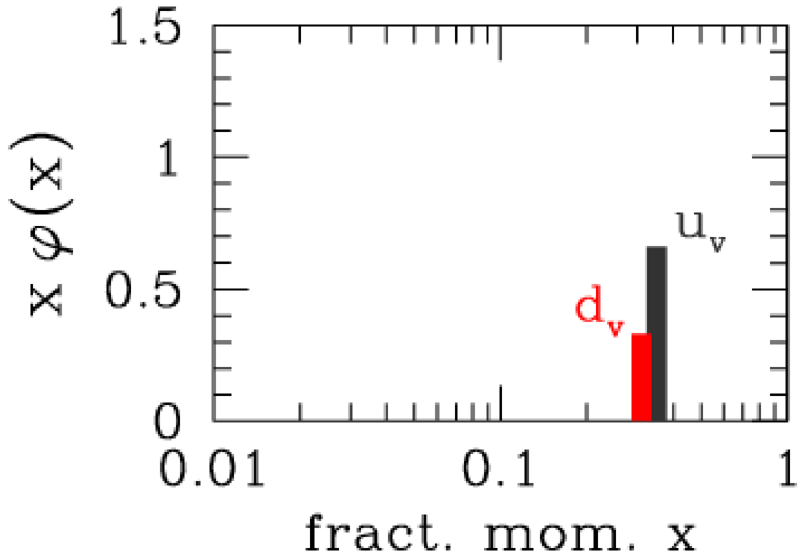
- **Why parton distributions?**
  - High-energy, hadronic and nuclear physics
- **Global PDF fits**
  - The basic ideas
- **Observables**
  - Sensitivity to specific quarks and gluons
- **Fitting**
  - A selection of fine details
  - PDF uncertainties: experimental, theoretical
  - Some recent PDFs

Why parton distributions?

PDFs and hadron structure

- Fundamental description of the structure of hadrons
  - Nucleons made of 3 quarks

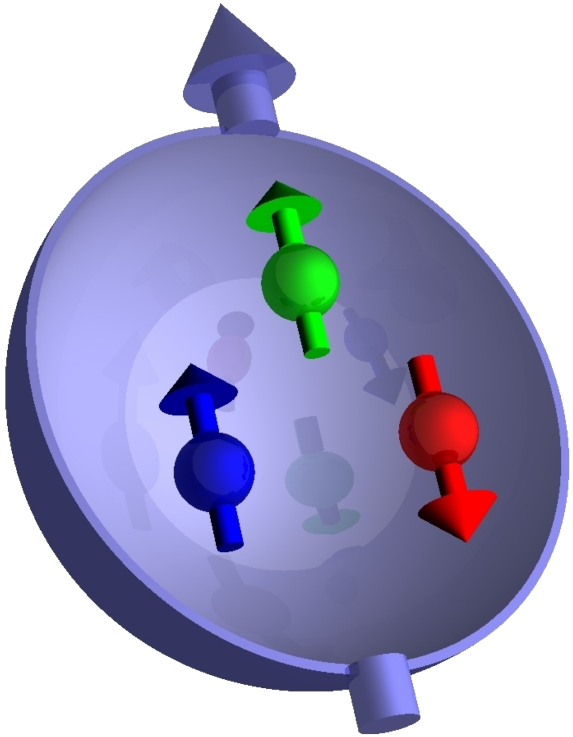


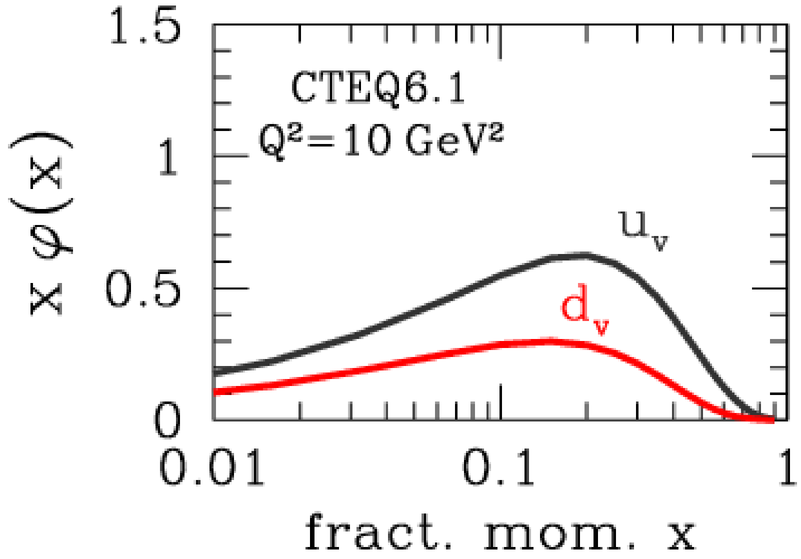


Fractional momentum: $x = \frac{p^+_{\text{parton}}}{p^+_{\text{nucleon}}}$ $p^\pm = \frac{1}{\sqrt{2}}(p_0 \pm p_3)$

PDFs and hadron structure

- Fundamental description of the structure of hadrons
  - Nucleons made of 3 **confined** quarks

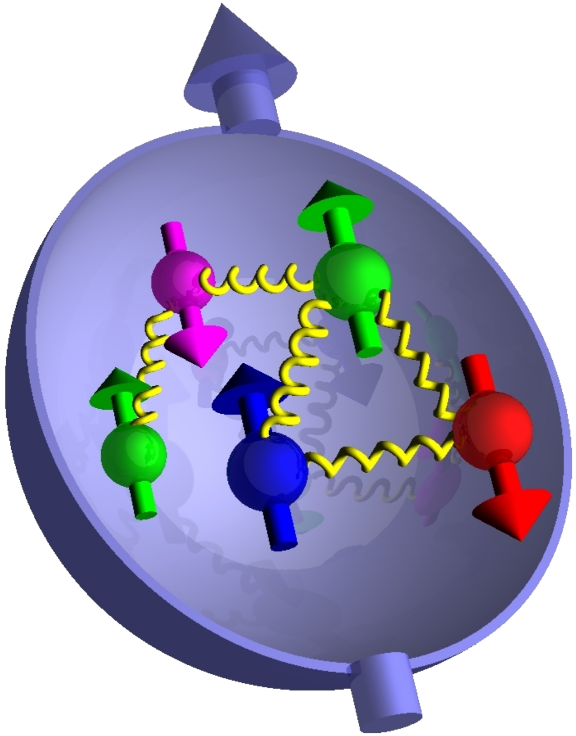


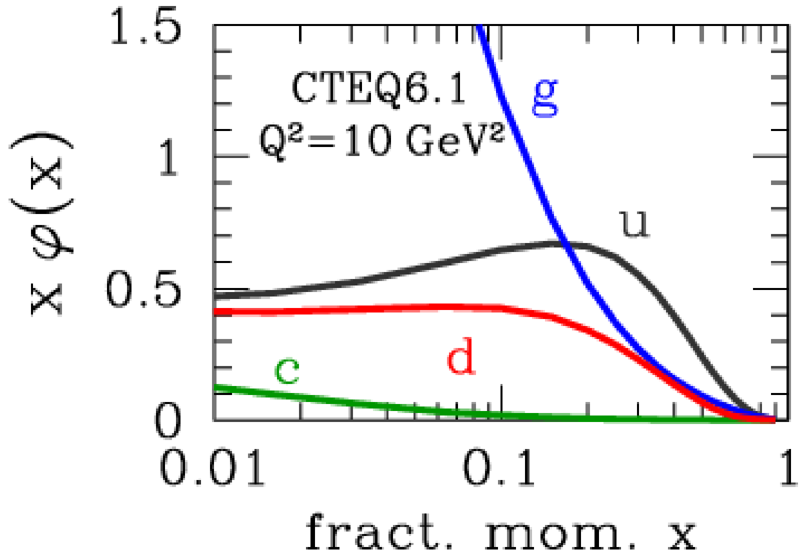


Fractional momentum: $$x = \frac{p^+_{\text{parton}}}{p^+_{\text{nucleon}}} \qquad p^{\pm} = \frac{1}{\sqrt{2}}(p_0 \pm p_3)$$

PDFs and hadron structure

- Fundamental description of the structure of hadrons
  - Nucleons made of 3 confined quarks, **and sea quarks and gluons**



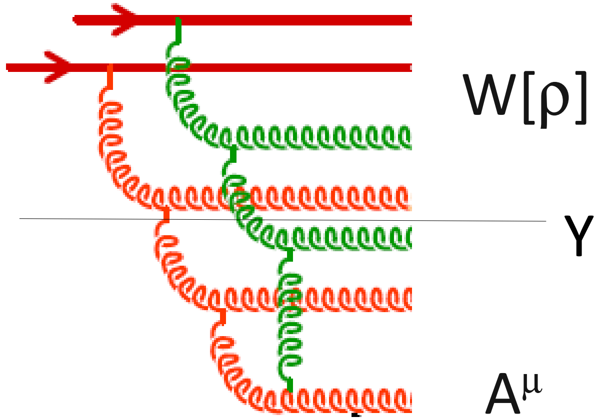


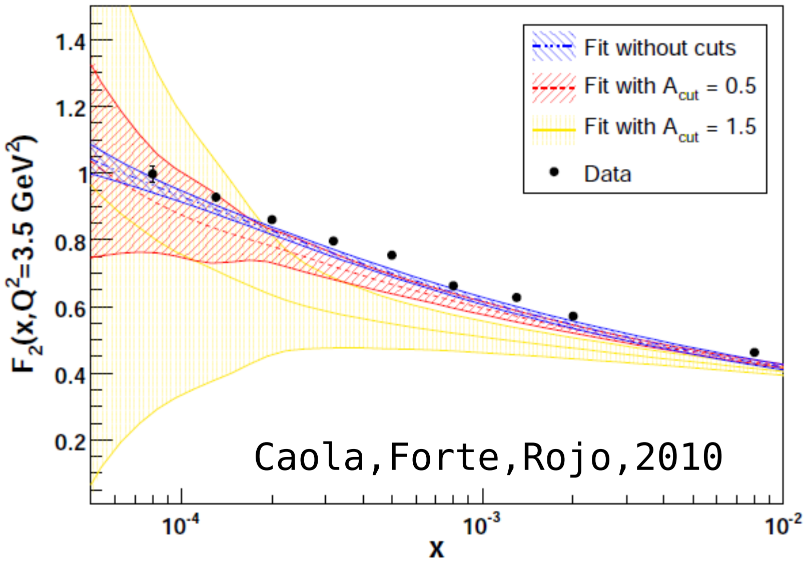
Fractional momentum: $x = \frac{p^+_{\rm parton}}{p^+_{\rm nucleon}}$ $\quad p^\pm = \frac{1}{\sqrt{2}}(p_0 \pm p_3)$

PDFs and hadron structure

- **Non-perturbative regime – dynamics of quark confinement**
  - SU(6) spin-flavor symmetry: $d/u \xrightarrow[x\to 1]{} 1/2$
  - Broken SU(6) : hard gluon exchange: $d/u \xrightarrow[x\to 1]{} 1/5$
  - Broken SU(6) : scalar diquark dominance: $d/u \xrightarrow[x\to 1]{} 0$

- **Perturbative regime – gluon saturation**





PDFs and High-Energy Physics

- Essential in perturbative calculations of hard scattering processes
  - Factorization & universality

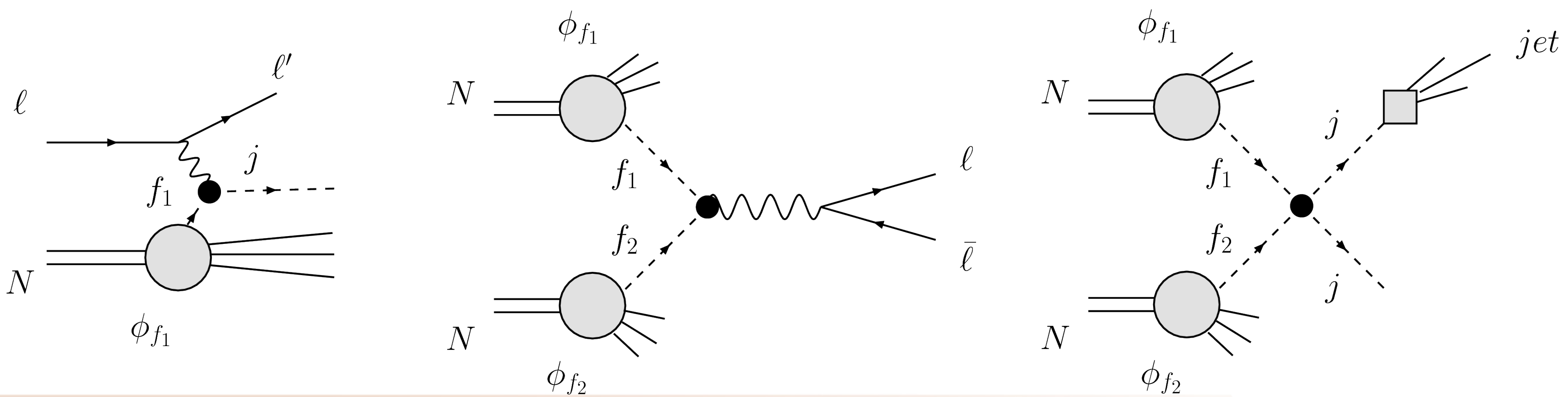
[J.Owens' lectures]

"Hard scale"
(large momentum transfer)^2

$$d\sigma = \sum_{f_1,f_2,i,j} \phi_{f_1}(Q^2) \otimes \hat{\sigma}^{f_1 f_2 \to ij}(Q^2) \otimes \phi_{f_2}(Q^2)$$

Parton Distribution Fns
(non-perturbative)

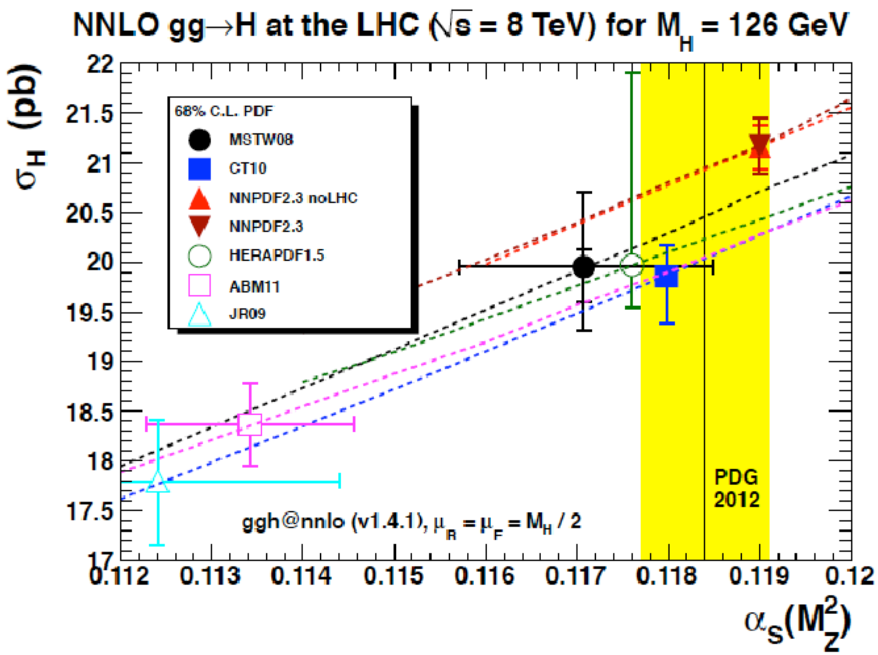
Partonic cross section
(calculable in pQCD)



PDFs and High-Energy Physics

- For example, **at the LHC**
  - Key ingredient for Higgs discovery
  - But PDF uncertainties fundamentally limit cross section calculations

- **Higgs cross section uncertainty**
  - Hence limitation on measurement of Higgs coupling
  - Cause: spread of PDF fits at medium-low *x*

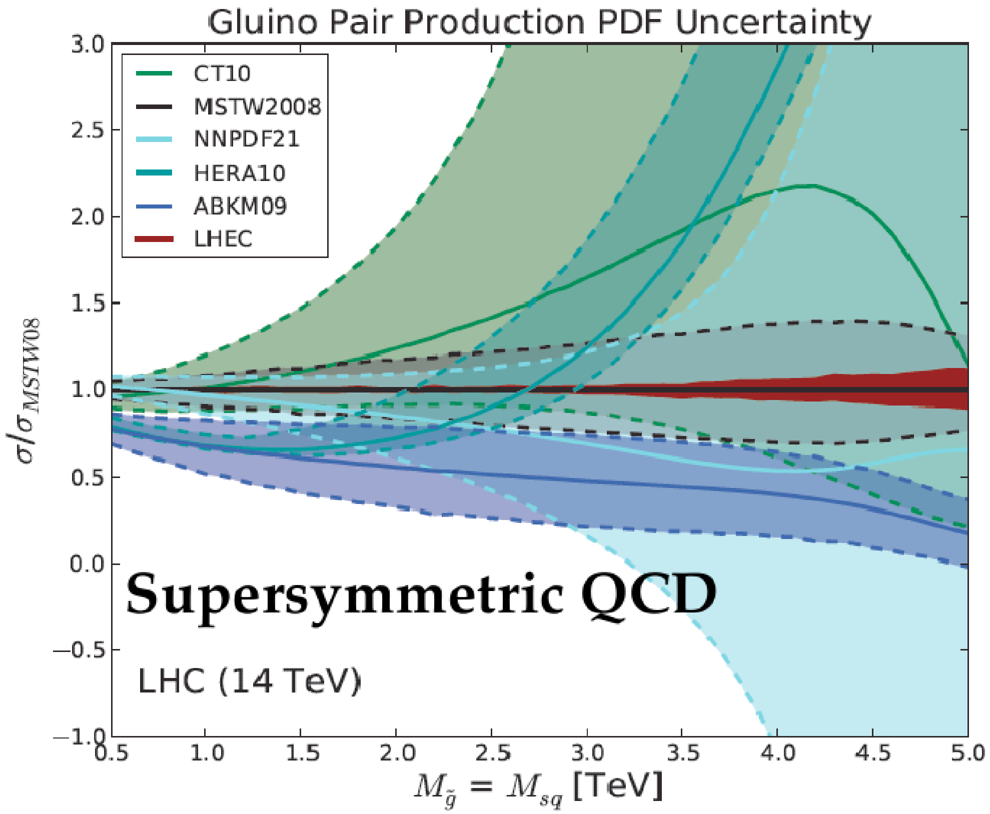


PDFs and High-Energy Physics

- **New heavy particle searches**
  - Large statistical uncertainties on large-*x* PDFs

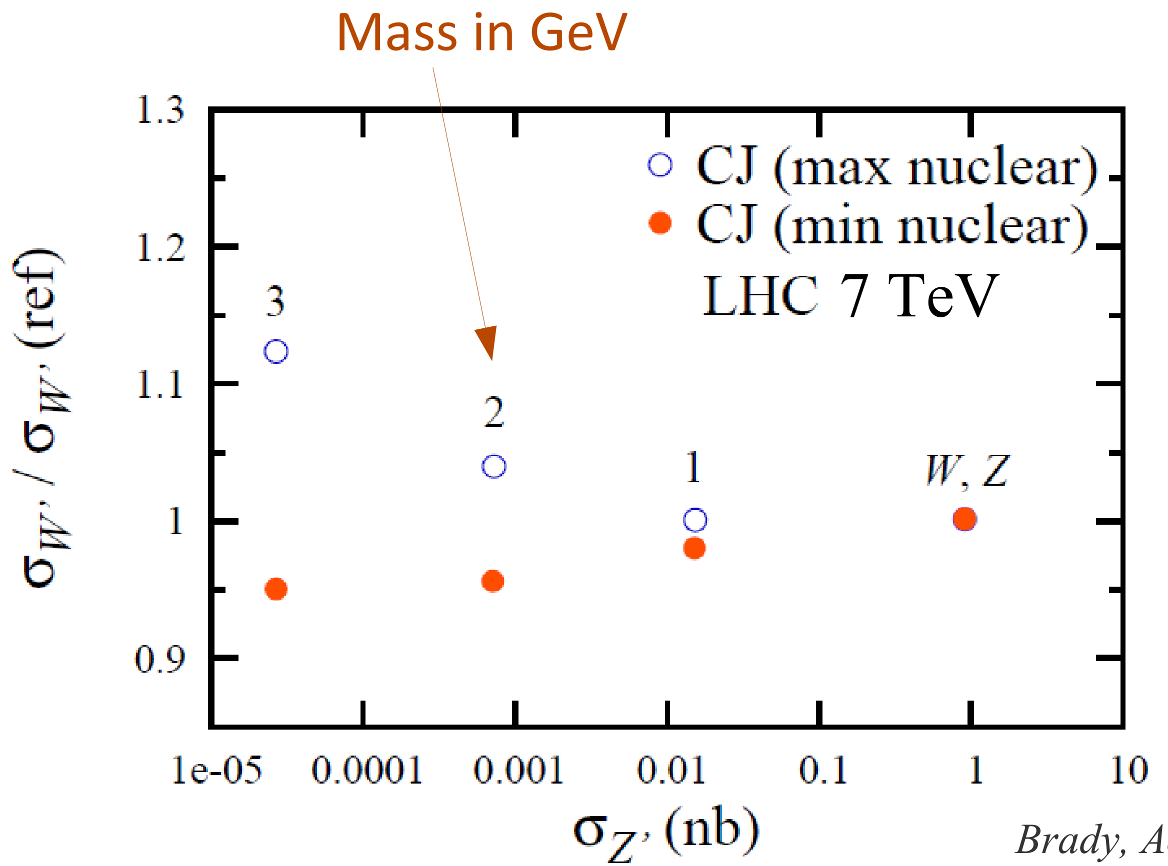
$$x \approx \frac{M}{\sqrt{s}} e^y$$

- ***Gluino searches***
  - Large statistical uncertainties in high-*x* gluons



PDFs and High-Energy Physics

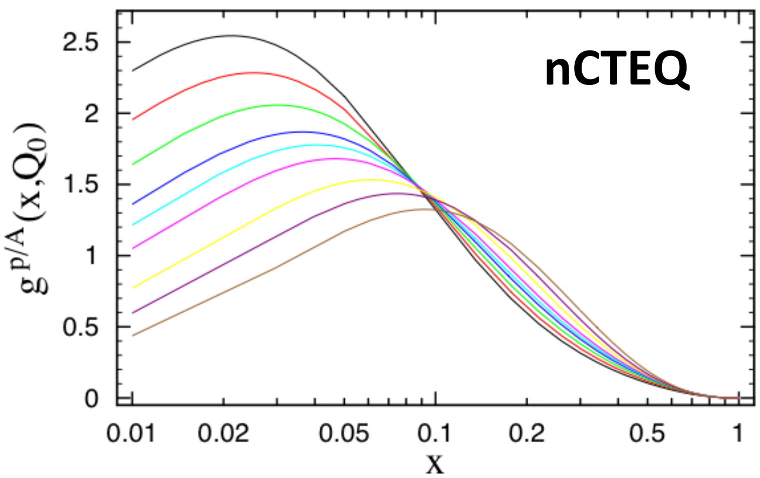
- ***W' and Z' total cross sections***
  - Large statistical and theoretical uncertainties on high-*x* *d* quarks
  - In the figure, "nuclear uncertainties" only
  - PDF uncertainties are comparable



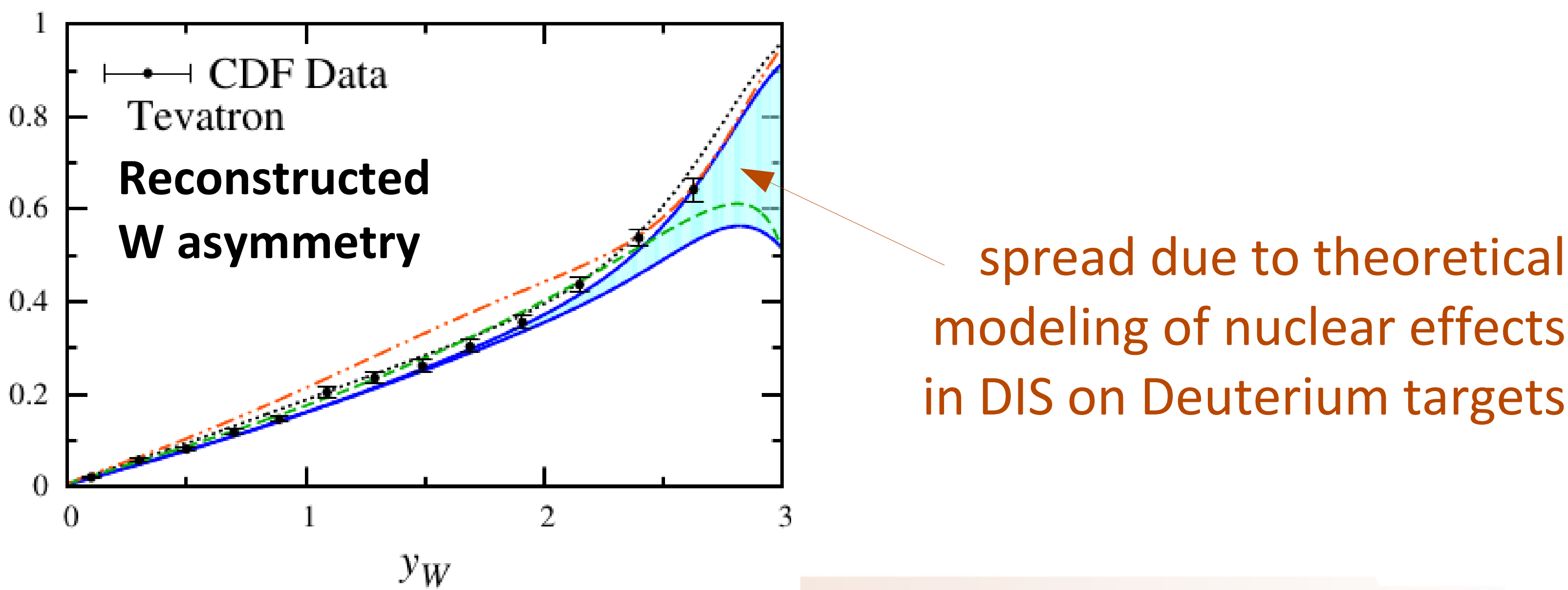
Brady, Accardi, Melnitchouk, Owens, JHEP 1206 (2012) 019

PDFs and nuclear physics

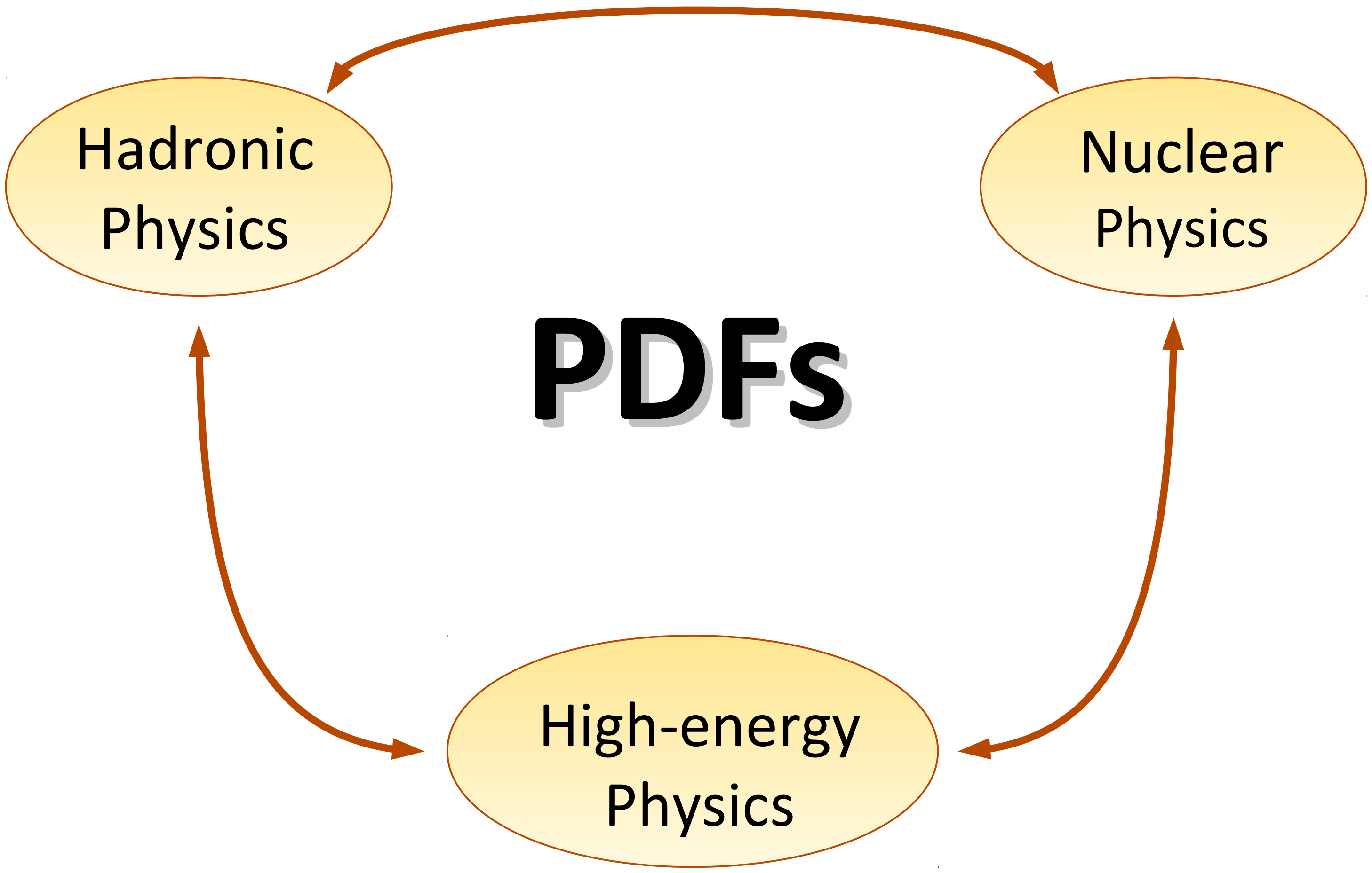
- Precise PDFs from proton target data
  → baseline for "nuclear PDFs"



- Compare proton target data with theory-corrected nuclear target data
  → test of nuclear theory models



These lectures' common thread



Global PDF fits

Fundamentals

❑ pQCD factorization

Factorization scale

Renormalization scale

$$d\sigma = \sum_{f_1,f_2,i,j} \phi_{f_1}(Q^2) \otimes \hat{\sigma}^{f_1 f_2 \to ij}(Q^2) \otimes \phi_{f_2}(Q^2)$$

Parton Distribution Fns
(non-perturbative)

Partonic cross section
(calculable in pQCD)

❑ Universality

- PDFs can be used to compute "any" hard scattering process
- In fact, the proof is not general but needs to be done process-by-process

Fundamentals

- **DGLAP evolution**:
  - you only need to know the PDFs at one scale Q_0

$$\frac{\partial q(x,t)}{\partial t} = \frac{\alpha_s(t)}{2\pi}\int_x^1 \frac{dy}{y}\left[P_{qq}(y)q(\frac{x}{y},t) + P_{qg}(y)g(\frac{x}{y},t)\right]$$
$$\frac{\partial g(x,t)}{\partial t} = \frac{\alpha_s(t)}{2\pi}\int_x^1 \frac{dy}{y}\left[P_{gq}(y)q(\frac{x}{y},t) + P_{gg}(y)g(\frac{x}{y},t)\right]$$

  - Coupled set of equations whose solutions show how the PDFs change with variations in the scale Q
  - Splitting functions have perturbative expansions α_s

$$t = \ln M_f^2/\mu^2$$

$$P_{ij}(y) = P_{ij}^0(y) + \frac{\alpha_s}{2\pi}P_{ij}^1(y) + \cdots$$

Fundamentals

- **Sum rules**
  - Charge conservation

$$2 = \int_0^1 dx \left[\underbrace{u(x) - \bar{u}(x)}_{= u_V(x)}\right] \qquad 1 = \int_0^1 dx \left[\underbrace{d(x) - \bar{d}(x)}_{= d_V(x)}\right]$$

$$0 = \int_0^1 dx \left[s(x) - \bar{s}(x)\right] = \int_0^1 dx \left[c(x) - \bar{c}(x)\right]$$

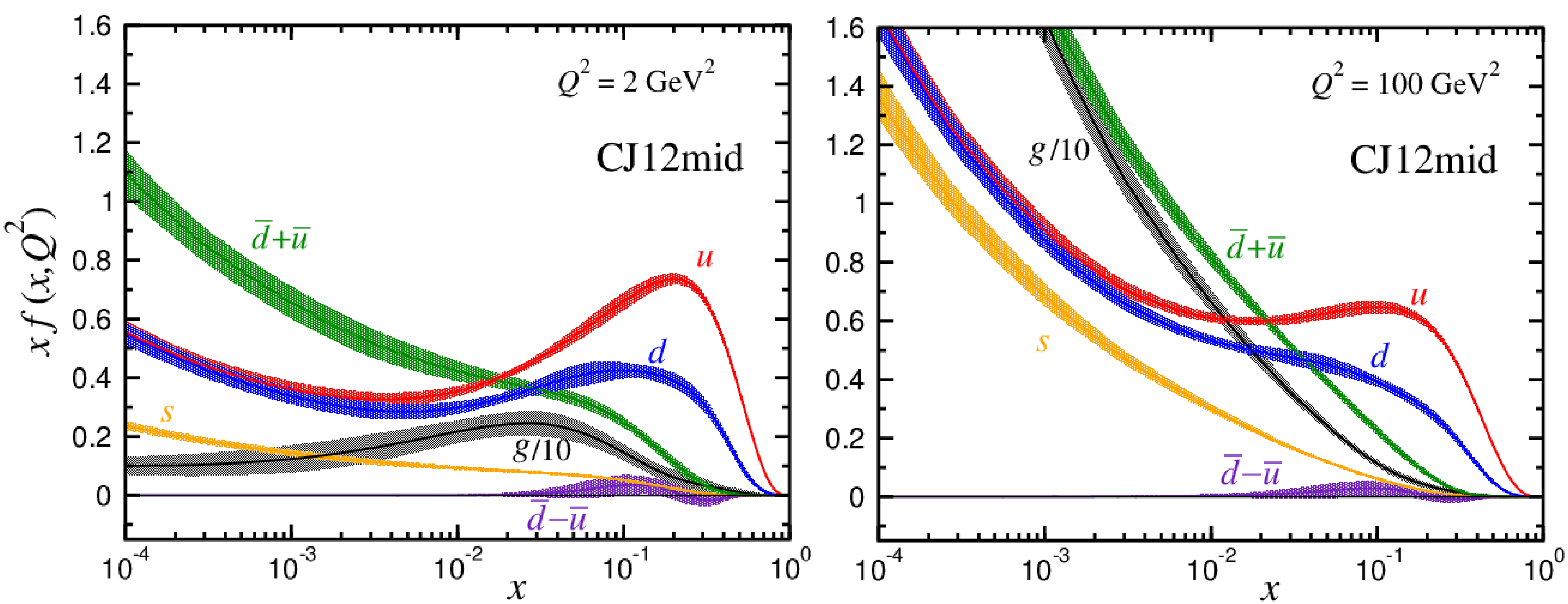
NOTE: does not mean $c(x) = \bar{c}(x)$!!

  - Momentum conservation

$$1 = \sum_{i=q,g} \int_0^1 dx\, x\, f_i(x)$$

Useful PDF properties - 1

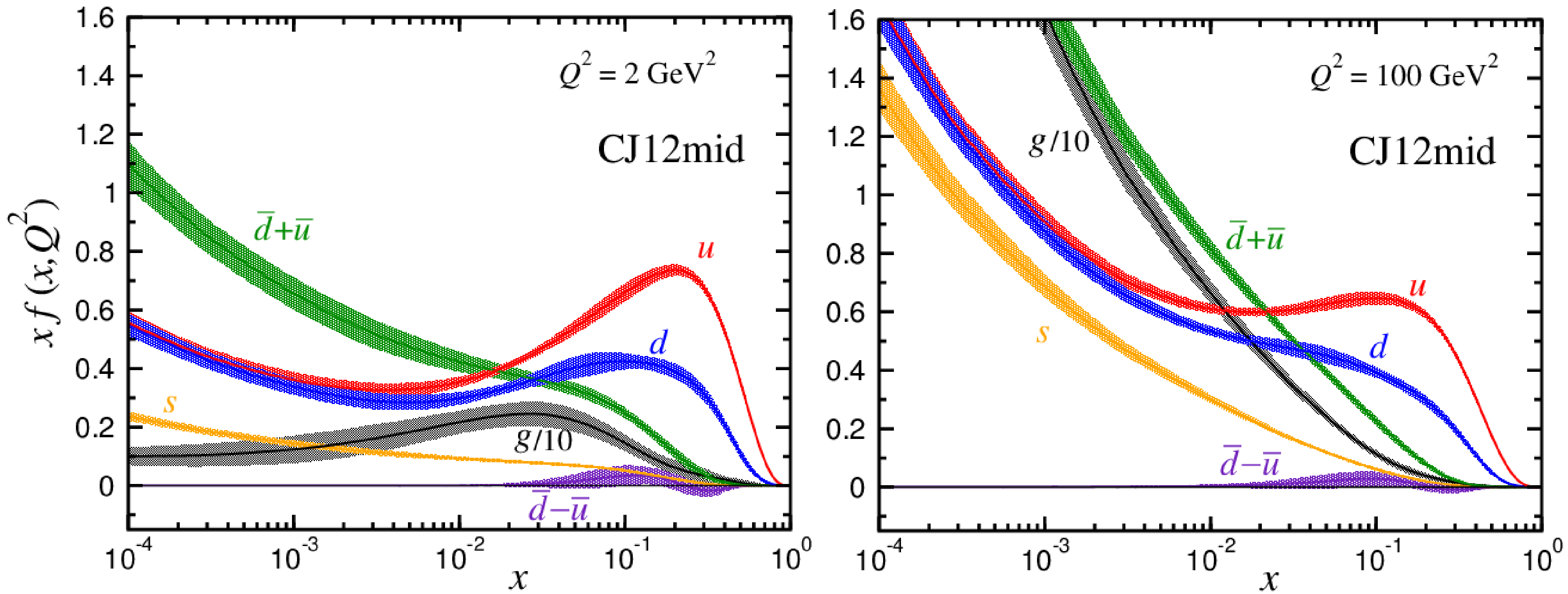
- The gluon dominates at low *x* and falls steeply as *x* increases
- Symmetric sea quarks: anti-*q* and *q* comparable at low *x* (and anti-*q* fall off in *x* even faster than the gluons)
- *u* and *d* dominate at large *x* with $u > d$; at low x, $u \approx d$



Owens, AA, Melnitchouk, PRD 87, 094012 (2013)

Useful PDF properties - 2

- Gluon radiation – QCD evolution in Q^2
  - Gluon radiation causes parton momentum loss:
    - At large *x*, quarks and gluons shift to the left: PDFs get steeper
  - Gluons create *q*, anti-*q* pairs, and *g*, *g* pairs:
    - At small *x*, quark and gluon PDFs increase, get steeper



Owens, AA, Melnitchouk, PRD 87, 094012 (2013)

Global PDF fits

- **Problem**:
  - we need a set of PDFs in order to calculate a particular hard-scattering process

- **Solution**:
  - Choose a data set for a set of different hard scattering processes
  - Generate PDFs using a parametrized functional form at initial scale Q_0; evolve them from Q_0 to any Q using DGLAP evolution equations
  - Use the PDF to compute the chosen hard scatterings
  - Repeatedly vary the parameters and evolve the PDFs again
  - Obtain an optimal fit to a set of data.

- **Modern PDF sets**: CTEQ-TEA (CT10), CTEQ-JLab (CJ12), MSTW2008, NNPDF2.1, ABM11, JR09, HERAPDF1.5

Global PDF fits

data

- DIS: p, d
- p+p(pbar) $\to$ l^+l^-, $W^\pm$, Z
- p+p(pbar) $\to$ jets, γ+jet

theory

- pQCD at NLO
- Factorization & universality
- Large-x, low-Q^2, nuclear corr.

fits

- Parametrize PDF at Q_0, evolve to Q
- Minimize χ^2

PDFs

$F_2(n)$

W, Z / W',Z', Higgs

(or any other "hard" observable)

Global PDF fits as a tool

- **Test new theoretical ideas**
  - *e.g.*, are sea-quarks asymmetric? Is there any "intrinsic" charm?
- **Phenomenology explorations**
  - *e.g.*, can CDF / HERA "excesses" be at all due to glue/quark underestimate at large x? Are there new particles at the LHC?
- **Test / constrain models**
  - *e.g.*, by extrapolating d/u at $x=1$
  - Possibly, constrain nuclear corrections
- **Limitations**
  - existing data
  - experimental uncertainty
  - theoretical uncertainty

As a user you should be aware of these

The art and science of global fitting - key points

- Choice of observables and data sets
- Choice of kinematic cuts to perform calculations with confidence
- Parametrized functional form for input PDFs at Q_0
- Definition of "optimal fit"
  - typically by a suitable choice of χ^2 function
- Truncation of the perturbative series:
  - LO; NLO (state-of-the-art)
  - NNLO (fully available for DIS, DY – partially for other processes)
- Treatment of errors
  - Experimental: statistical and systematic
  - Theoretical

Observables

Observables

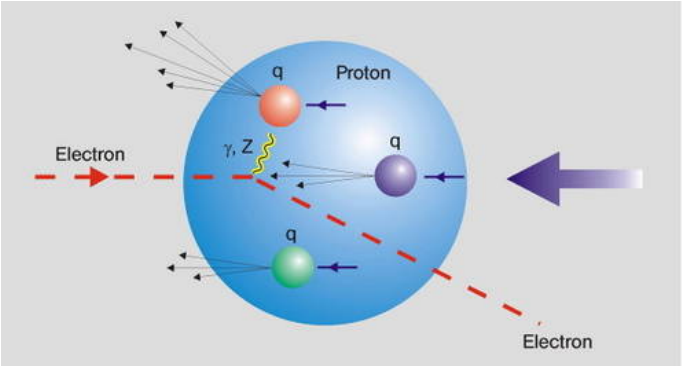
- Each observables involves a different linear combination, or product of PDFs: a diverse enough set of observables is needed for parton flavor separation
  - Some redundancy needed to cross-check data sets

- Typical data sets used in global fits
  - Inclusive DIS $\ell^{\pm} + p,\ \ell^{-} + D^{*}$
  - Vector boson production in *p+p, p+D* $W^{\pm},\ Z^{0},\ \mathrm{DY\ lepton\ pairs}$
  - Hadronic jets, *p+p or p+pbar:* inclusive jets, γ + jet
  - neutrino DIS: $\nu + A^{*}$

  * use of nuclear targets require consideration of nuclear corrections to measure the proton / neutron PDFs; typically these induce large theoretical uncertainty, the more so for heavy nuclei. Fixed target DY is so far an exception: the probed *x* values in the nucleus are small enough to neglect corrections.

- Need to establish a strategy to get to the particular PDFs one is interested in
  - Different groups make different choices

Deep Inelastic Scattering (DIS)



$$\ell^{\pm} + p \to \ell^{\pm} + X$$ NC: gamma, Z

$$\left.\begin{array}{l} \ell^{+}(\ell^{-}) + p \to \bar{\nu}(\nu) + X \\ \nu(\bar{\nu}) + A \to \ell^{-}(\ell+) + X \end{array}\right\}$$ CC: W^{+}, W^{-}

$$\frac{d\sigma^{NC}}{dx dQ^2} \propto Y_{+} F_2^{NC}(x, Q^2) \mp Y_{-} x F_3^{NC}(x, Q^2) - y^2 F_L^{NC}(x, Q^2)$$

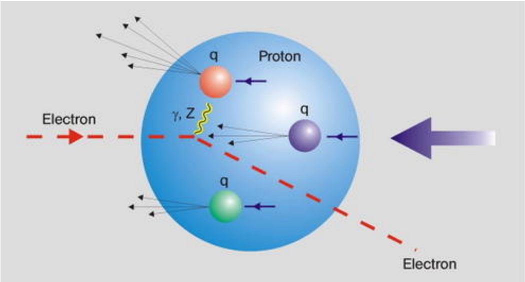
$$Q^2 = -p^2_{\gamma,Z,W} \quad x = \frac{2p \cdot q_{\gamma,Z,W}}{Q^2} \qquad y = \frac{q_{\gamma,Z,W} \cdot p}{k \cdot p}$$

- Electromagnetic probe at small Q2, electroweak couplings at large Q2

$$F_2^{NC} \propto F_2^{\gamma} + (g_V \pm \lambda g_A) \frac{G_F M_Z^2}{4\pi\alpha} \frac{Q^2}{Q^2 + M_Z^2} F_2^{\gamma Z}$$

$$x F_3^{NC} \propto (g_V \pm \lambda g_A) \frac{Q^2}{Q^2 + M_Z^2} x F_3^{\gamma Z}$$

Deep Inelastic Scattering (DIS)



- Leading Order DIS is a direct probe of quark and antiquarks

- **Proton target** at $Q^2 \ll M_Z^2$

$$F_2^\gamma(x,Q^2) \propto x \sum_i e_i^2 \left(q_i(x,Q^2) + \bar{q}_i(x,Q^2)\right)$$

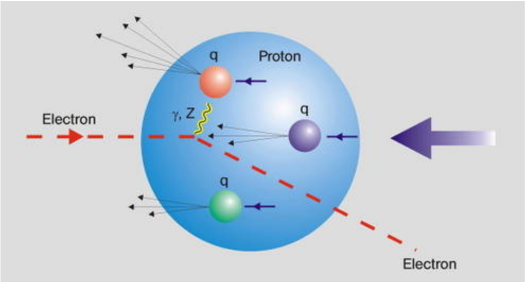
  - Each flavor is weighted by its charge squared
  - Gluon does not enter at LO (and $F_L = 0$)
  - Quarks and antiquarks enter togther

- **Deuterium target** = (interacting) proton + neutron

$$F_2^\gamma(x,Q^2) \propto x \sum_i \left(q_i(x,Q^2) + \bar{q}_i(x,Q^2)\right)$$

  - Different combination, allows *d* vs. *u* quark separation
  - Assumes isospin symmetry, up= dn and dp= un , no other nuclear fx
  - But corrections for binding and Fermi motion not small at *x* > 0.3

Deep Inelastic Scattering (DIS)



- **Neutrino scattering**: measure both F_2 and F_3

$$F_2^{\nu}(x,Q^2) = x\sum_i e_i^2\left(q_i(x,Q^2) + \bar{q}_i(x,Q^2)\right)$$

$$xF_3^{\nu}(x,Q^2) = x\sum_i e_i^2\left(q_i(x,Q^2) - \bar{q}_i(x,Q^2)\right)$$

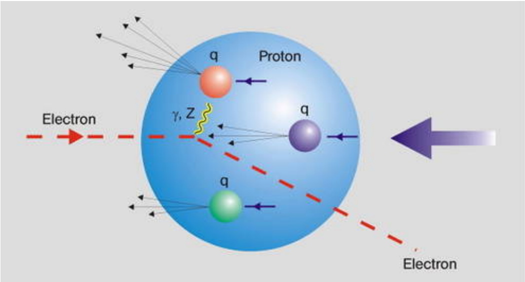
  - Separation of quarks and antiquarks
  - BUT: few data on proton targets (WA21/22)
  - Needs heavy nuclear targets, theoretical corrections for nuclear effects

- Same can be accomplished in **CC lepton DIS on protons** (e.g. at HERA)

$$F_2^{W^+} = x(\bar{u} + d + s + \bar{c})$$

$$F_2^{W^-} = x(\bar{u} - d - s + \bar{c})$$

Deep Inelastic Scattering (DIS)



- **γ-Z interference** allows further quark separation

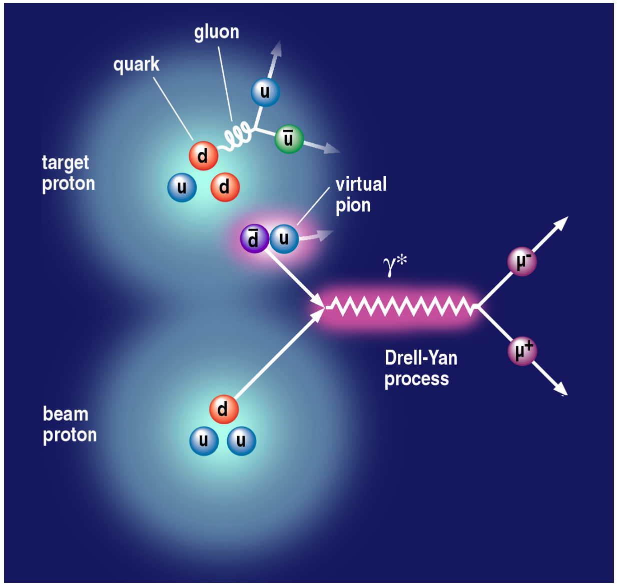
$$F_2^{\gamma Z} = x\sum_i B_i\left(q_i + \bar{q}_i\right)$$
$$xF_3^{\gamma Z} = x\sum_i D_i\left(q_i - \bar{q}_i\right)$$

  - Can be measured by comparing positive vs. negative helicity leptons and positrons vs. electrons

- In principle DIS on protons allows full quark flavor separation, but:
  - Data scarce for more "exotic" structure functions
  - Would require either a neutrino factory (far in the future) or an Electron-Ion Collider – EIC or LHeC (possibly in the 2020's)

- **We also need nuclear targets and/or hadron-hadron scattering**

Lepton pair production in hadronic collisions



$$p + p\,(\bar{p}) \rightarrow \gamma, Z, W \rightarrow \ell + \bar{\ell} + X$$

$$Q^2 = (p_\ell + p_{\bar{\ell}})^2$$

$$x_{1,2} = \frac{Q}{\sqrt{s}} \exp(\pm y)$$

$$\gamma \qquad \frac{\mathrm{d}\sigma}{\mathrm{d}Q^2 \mathrm{d}y} = \frac{4\pi\alpha^2}{9Q^2 s} \sum_i e_i^2 L^{ii}(x_1, x_2)$$

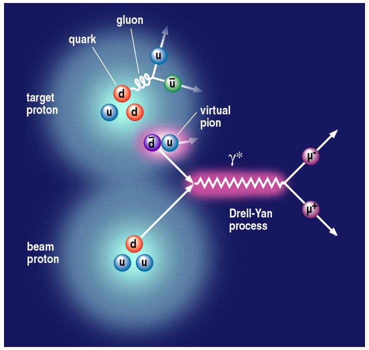
$$W \qquad \frac{\mathrm{d}\sigma}{\mathrm{d}y} = \frac{\pi G_F M_V^2 \sqrt{2}}{3s} \sum_{i,j} |V_{ij}^{\mathrm{CKM}}| L^{ij}(x_1, x_2)$$

$$Z \qquad \frac{\mathrm{d}\sigma}{\mathrm{d}y} = \frac{\pi G_F M_V^2 \sqrt{2}}{3s} \sum_i \left(V_i^2 + A_i^2\right) L^{ii}(x_1, x_2)$$

$$L^{ij}\,(x_1, x_2) \equiv q_{i/A}(x_1, Q^2)\, \bar{q}_{j/B}(x_2, Q^2) + \{A \leftrightarrow B\}$$

$q(x_2)$ in $p\bar{p}$ collisions

“Drell-Yan” pair production



- Away from Z,W resonance → mediated by photons

- At large (but not too much) y, and low Q^2

$$\sigma^{pp} \propto 4u(x_1)\bar{u}(x_2) + d(x_1)\bar{d}(x_2) + \ldots$$

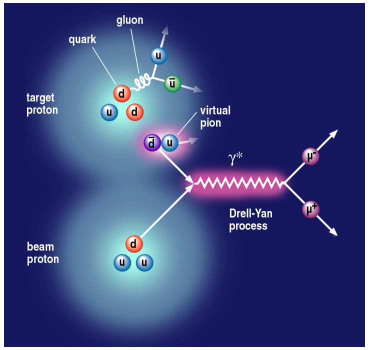
$$\sigma^{pD} \propto 4\big[4u(x_1) + d(x_1)\big]\,\big[\bar{u}(x_2) + \bar{d}(x_1)\big] + \ldots$$

  – σ^{pD}/σ^{pp} sensitive to the $\bar{d}/\bar{u}$ ratio

- At large *y* (large x_1, small x_2) same combinations as DIS on *p* and *D*
  – No new information

- Large Q^2 range at LHC: additional handles available
  – Evolution at moderate *x* different for strange and light quarks
  – at very large Q^2 charm quarks not negligible

W,Z production



- $Z \to l^+ + l^-$
  - *Z* kinematics reconstructed from charged lepton pair
  - Weak coupling helps with flavor separation

- $W \to l + \nu$

$$d\sigma^{W^+}/dy \approx u(x_1)d(x_2) + \bar{d}(x_2)\bar{u}(x_2) + \ldots$$

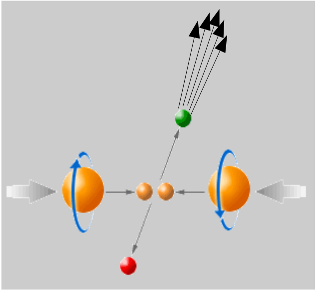
$$d\sigma^{W^-}/dy \approx u(x_1)d(x_2) + \bar{d}(x_2)\bar{u}(x_2) + \ldots$$

At large *y, i.e.*, small x_2: $d\sigma^{W^-}/d\sigma^{W^+} \approx d/u$

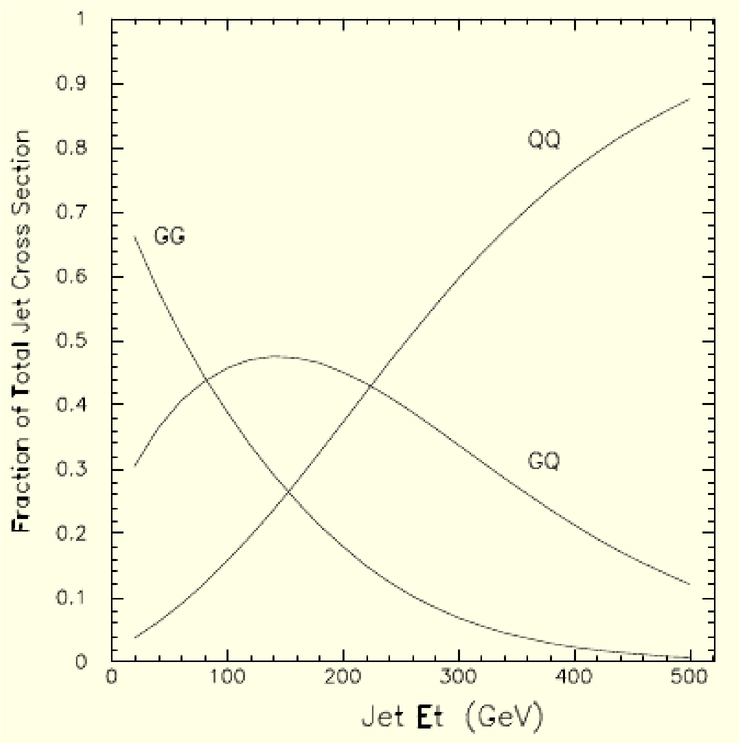
(Alternatively, charge asymmetry $(W^+ - W^-)/(W^+ + W^-)$)

  - But: missing energy, reconstruction of W kinematics is a challenge
  - Lepton decay limits reach to $x < 0.5$

Hadronic production of jets

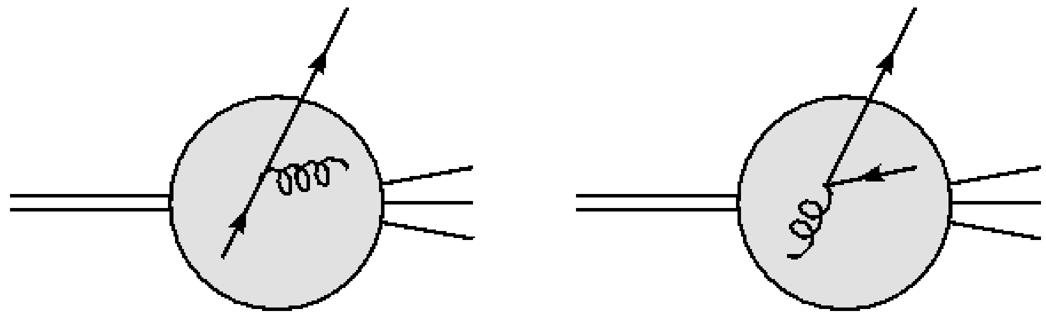


- The *qq* subprocesses do dominate the high-E_T region
  - But enough contribution from the gluons that data can be used to constrain the large-*x* gluon behavior
  - Combined with the low-*x* data and the momentum sum rule one has strong constraints on the gluon distribution



Gluons from DIS

- 2 methods
  - Scaling violations in F_2 $$G(x) \approx \frac{d}{d\log Q^2} F_2(x,Q^2)$$

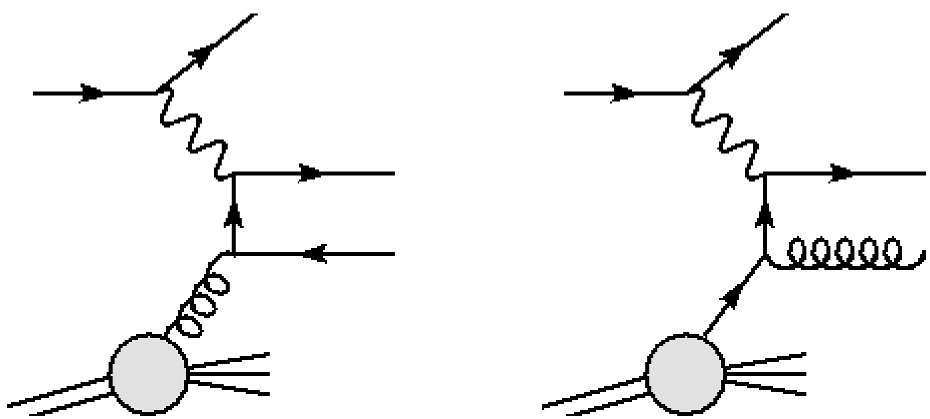


  - Longitudinal structure function F_L

$$F_L(x) = \gamma F_2 - 2xF_1(x) = 0 \text{ at LO}$$

$$\gamma = \sqrt{1 + 4x^2M^2/Q^2}$$

- Gluons subdominant in F_2
- But same order as quarks in F_L



Gluons from DIS

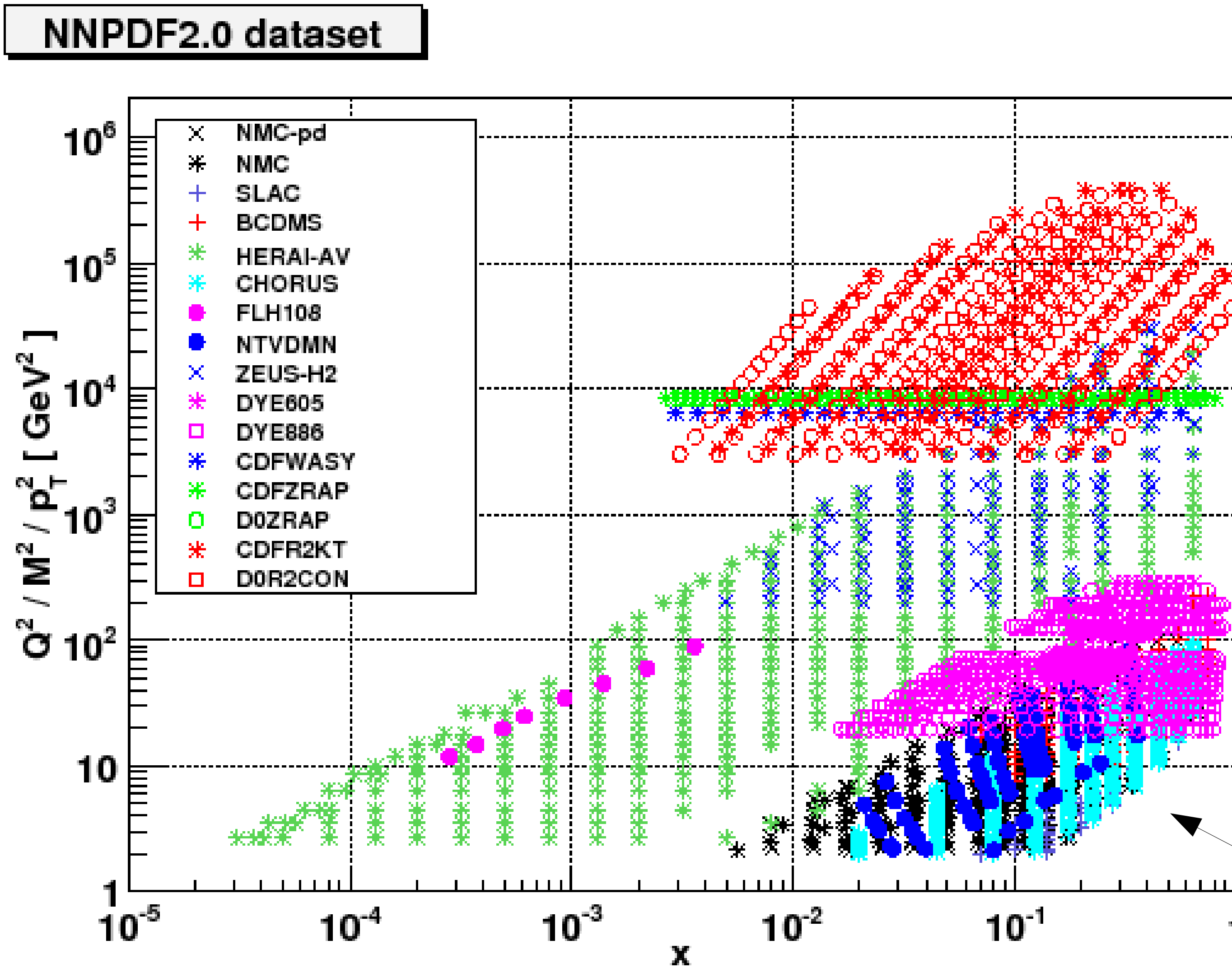
- Caveats
  - Experimentally separating F2 and FL requires measurements at different √s
  - Typically lower statistics than for F2
  - Systematic error analysis tricky

$$\frac{d\sigma^{NC}}{dx dQ^2} \propto F_2^{NC}(x,Q^2) - \frac{y^2}{1+(1-y)^2} F_L^{NC}(x,Q^2)$$

$$y = Q^2/(xs)$$

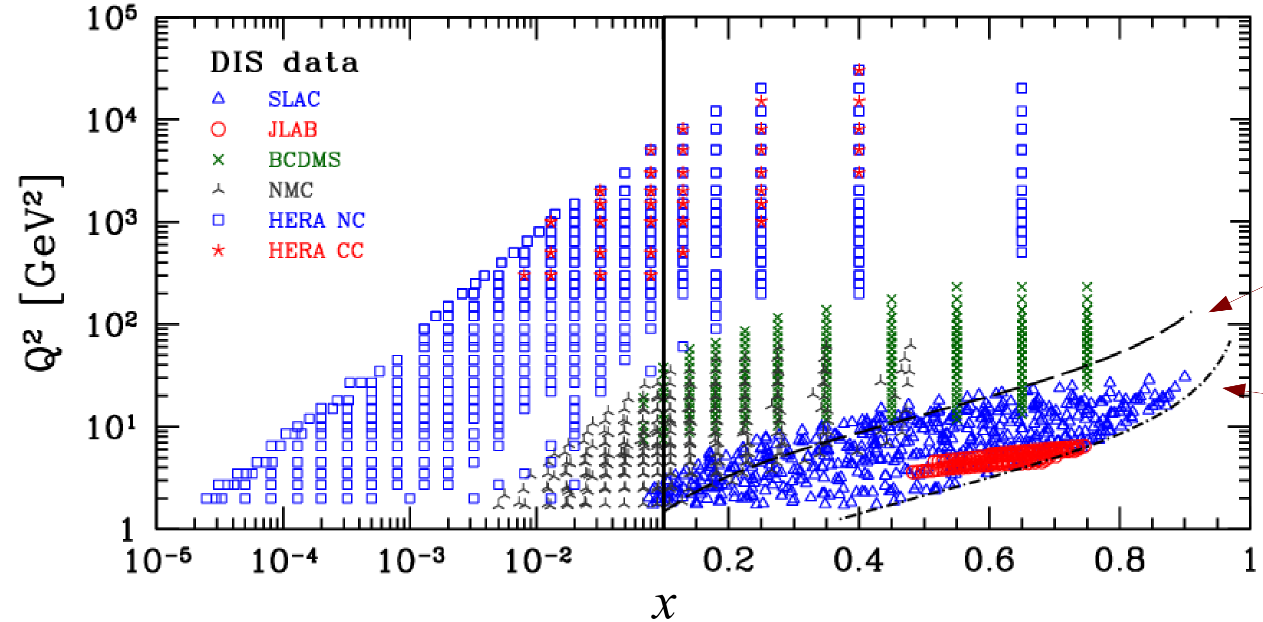
- In practice global fits can directly fit the cross section
  - effective F2 / FL separation

Data coverage in *x* and *Q* - NNPDF 2.0



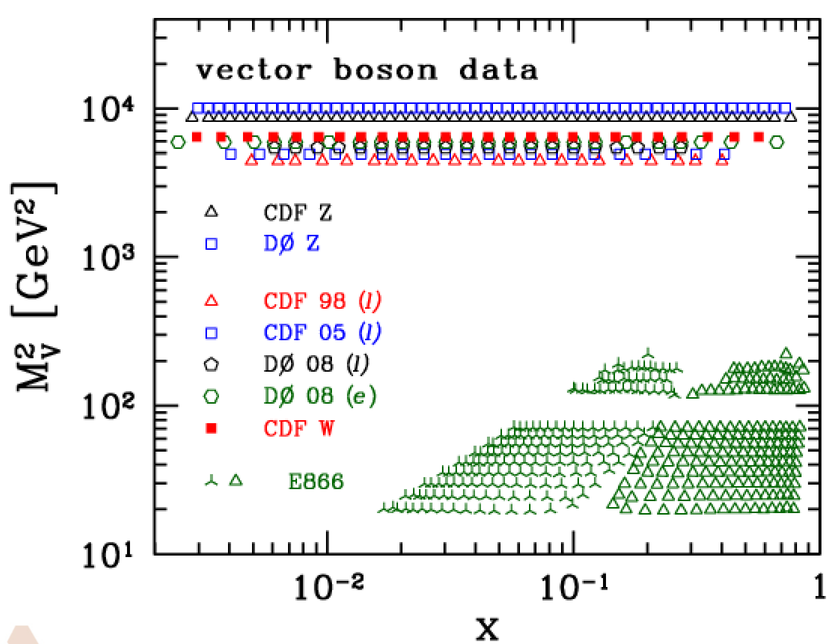
Note the empty triangle

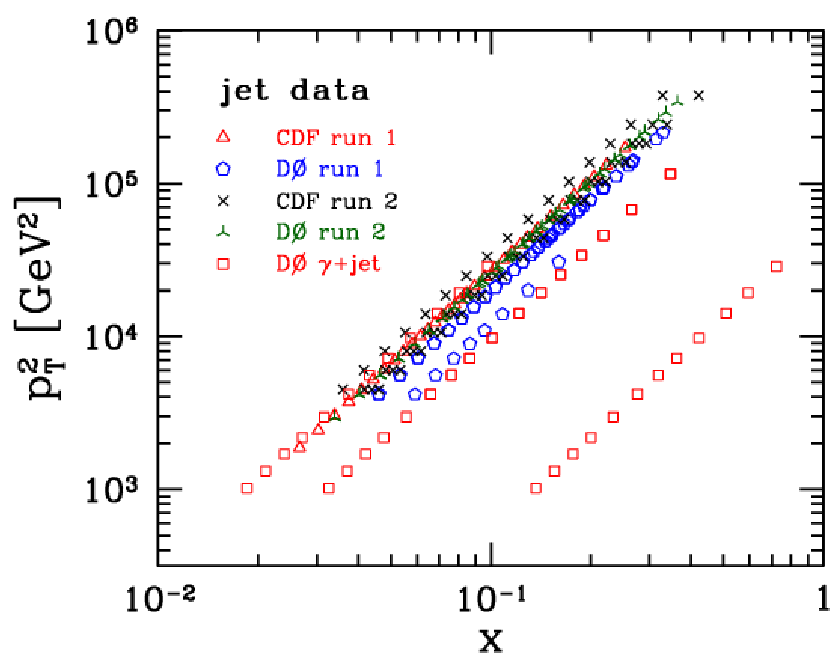
Data coverage in x and Q - CJ12 (by category)



standard cut
$W^2 \gtrsim 14\ \mathrm{GeV}^2$

CJ12, ABM11
$W^2 \gtrsim 3\ \mathrm{GeV}^2$





The art of fitting

Parametrization at Q_0

- In the beginning... first fits based on

$$f_i(x) = N_i x^{\alpha_i} (1-x)^{\beta_i}$$

- Estimate β by counting rules: $\beta = 2n_s - 1$ with n_s = spectator quarks no.
  - Valence quarks (qqq): $n_s = 2$, $\beta = 3$
  - Gluons ($qqqg$): $n_s = 3$, $\beta = 5$
  - Antiquarks ($qqqq\bar{q}$): $n_s = 4$, $\beta = 7$

- Estimate α from Regge arguments, behavior of gluon radiation:
  - Gluons and antiquarks: $\alpha \approx -1$
  - Valence quarks: $\alpha \approx -1/2$

- Overall normalization fixed by sum rules (momentum, charge conservation, ...)

Parametrization at Q_0

- With the large variety of precise data available today, needs more flexibility: multiply by a suitable function of *x*.
- Examples for *u*, *d* quarks and gluons
  - **CTEQ6.1** – *u,d,g*: $N\, x^{a_1} (1-x)^{a_2} e^{a_3 x} \left[1 + e^{a_4 x}\right]^{a_5}$
  - **MSTW2008** – *u,d*: $N\, x^{a_1} (1-x)^{a_2} \left[1 + a_3 \sqrt{x} + a_4 x\right]$

    g: $N\, x^{a_1} (1-x)^{a_2} \left[1 + a_3 \sqrt{x} + a_4 x\right] + N' x^{b_1} (1-x)^{b_2}$
- Caveats:
  - Choice of functional form or no. of free parameters can bias the results
  - Theoretical prejudices often built-in (*e.g.*, CTEQ gluons can't go negative, d/u ratio forced to either 0 or ∞ as $x \to 1$)
- **NNPDF**: obtains "unbiased fits" by a neural-network parametrization, using a very large linear basis of functional forms

Parametrization at Q_0

- Other points to keep in mind
  - One should increase the number of parameters and the flexibility of the parametrization until the data are well described
  - Adding more parameters past that point simply results in ambiguities, false minima, unconstrained parameters, etc.
    - But in Neural Network based fits this is turned into a virtue!
  - May have to make some arbitrary decisions on parameter values that are not well constrained by the data

“Optimal” fit

- Needs a numerical measure of how good a fit is
  - choose a suitable χ^2 function
  - vary parameters iteratively until χ^2 minimized

- Simplest choice

$$\chi^2 = \sum_i \frac{(D_i - T_i)^2}{\sigma_i^2}$$

D = exp.data
σ = uncorrelated exp. errors
T = calculation

  - OK for 1 data set
  - And if data is statistically limited (errors not “too small”)

- But nowadays we have
  - Several data sets for many observables
  - Correlated and uncorrelated errors
  - Overall normalization errors (due to, say, luminosity uncertainties)

"Optimal" fit

Normalization errors

- assign a χ^2 penalty for normalization errors (different choices possible)
- Fit optimal normalization f_N, compare to quoted one

$$\chi^2 = \sum_i \frac{(f_N D_i - T_i)^2}{\sigma_i^2} + \left[\frac{1 - f_N}{\sigma_N^{norm}}\right]^2$$

MSTW use a power 4

Point-to-point systematic errors

$$\chi^2 = \sum_i \frac{(D_i - \sum_{j=1}^k \beta_{ij} s_j - T_i)^2}{\sigma_i^2} + \sum_{j=1}^k s_j^2$$

- The data points D_i are shifted by an amount reflecting the systematic errors β with the shifts given the the s_j parameters
- There is a quadratic penalty term for non-zero values of the shifts s
- The minimum w.r.t. s_j are obtained analytically

"Optimal" fit

- **Minimization of biases in treatment of normalizations**
  - treat all errors on the same footing

the covariance matrix for each experiment is computed from the knowledge of statistical, systematic and normalization uncertainties as follows:

$$(\mathrm{cov}_{t_0})_{IJ} = \left(\sum_{l=1}^{N_c} \sigma_{I,l}\sigma_{J,l} + \delta_{IJ}\sigma_{I,s}^2 \right) F_I F_J + \left(\sum_{n=1}^{N_a} \sigma_{I,n}\sigma_{J,n} + \sum_{n=1}^{N_r} \sigma_{I,n}\sigma_{J,n} \right) F_I^{(0)} F_J^{(0)} \ , \quad (1)$$

where I and J run over the experimental points, F_I and F_J are the measured central values for the observables I and J, and $F_I^{(0)}$, $F_J^{(0)}$ are the corresponding observables as determined from some previous fit.

[Ball et al., Nucl.Phys.B838:136,2010]

- **Want to emphasize a given data set?** use

$$\chi^2 = \sum_k w_k \, \chi_k^2 + \sum_k w_{N,k} \left[\frac{1 - f_N}{\sigma_N^{norm}} \right]^2$$

  - the weights w_k and $w_{N,k}$ can be chosen to emphasize the contribution of a given experiment or normalization to the total χ^2

“Optimal” fit

- **Neural Network based fits** [Ball et al., Nucl.Phys.B838:136,2010]

  – Too many parameters for conventional χ^2 minimization (would fit everything, including statistical fluctuation...)

  – Solution:
    - Generate replicas of data set by randomly varying central values within their experimental uncertainties
    - Divide these pseudo-data sets into “training” and “control”
    - Reduce χ^2 in training set until the control χ^2 starts to increase

  – Method can in principle be used also in conventional global fits

Order of perturbation theory

- Lowest order in α_s (LO) - easy to do, but
  - Hard scattering subprocesses do not depend on the factorization scale
  - May be missing large higher order corrections
- Next-to-leading-order (NLO) - more complicated, but
  - Less dependent on scale choices since the PDFs and hard scattering subprocesses both contain scale dependences which (partially) cancel
  - Some higher order corrections are now included
- Next-to-next-to-leading-order (NNLO) - better, but
  - Splitting functions are known so NNLO evolution can be done
  - Some hard scattering subprocesses are known to NNLO (DIS, DY) but not high-ET jets (yet), and many other important for, say, LHC
- NLO remains the state-of-the-art, but full NNLO analyses are coming

Order of perturbation theory

- LO PDFs can be interpreted as probability distribution in *x*

$$q_{LO}(x) = \int \frac{dz^-}{2\pi} e^{ixp^+z^-} \langle p|\overline{\psi}(z^- n)\frac{\gamma^+}{2}\psi(0)|p\rangle$$

- At NLO, PDFs are defined to absorb IR and collinear divergences in the hard-scattering diagrams: they are no longer probabilities.

$$q_{NLO} = q_{LO} + \{\text{divergences}\}$$

  - Then, get rid of infinities by extracting PDFs from data (in this sense it is analogous to UV renormalization)
  - Note: a "divergence" is defined differently in different subtraction schemes (most commonly "modified minimal subtraction" $\overline{\text{MS}}$, or DIS)

- NLO PDFs are not "better" than LO PDFs – they are different objects:
  - **you should use LO PDFs in LO calculations, NLO PDFs in NLO calcul'ns**
  - **...and the same subtraction scheme, choice of scale**

PDF uncertainties - preview of Lecture 2

- **Experimental**:
  - uncertainties in measured data propagate into the fitted PDFs
  - can be quantified adapting statistical methods: "PDF error bands"
  - These PDF errors need to be interpreted with care
- **Theoretical**:
  - Several sources, cannot be quantified easily
    - Choice of data sets, kinematic cuts
    - Parametrization bias
    - Choice of χ^2 function
    - Truncation of pQCD series, heavy-quark scheme, scale choice
    - Higher-twist, target mass effects
    - Nuclear corrections
    - ...

Lecture 1 - recap

data

- DIS: p, d
- p+p(pbar) $\to$ l^+l^-, $W^\pm$, Z
- p+p(pbar) $\to$ jets, γ+jet

theory

- pQCD at NLO
- Factorization & universality
- Large-x, low-Q^2, nuclear corr.

fits

- Parametrize PDF at Q_0, evolve to Q
- Minimize χ^2

PDFs

F_2(n)

W, Z / W',Z', Higgs

(or any other “hard” observable)