Brief Overview: Pion and Kaon Structure Functions

- Pion and kaon structure functions play an important role in understanding the origin of mass.

  - Proton has three quarks and a mass of about 1 GeV. Take one of the quarks away to get a pion and get a mass that's much less than 1/3 of the proton mass – Why?

  - Do the same and check out the kaon mass: get about one half of the proton mass – strange?

  - Pion and kaon structure is different – can it explain mass?

- Data are sparse. We know essentially nothing about the contribution of sea quarks and gluons

- Some nomenclature: **F2**== a (dynamic) structure function accounting for the substructure of a particle and allows access to the **PDF**s=Parton Distribution Function

Objective: Pion and Kaon PDFs

Goal: Impact of projected F2 data on pion (kaon) PDFs?

- What needs to be done:
  - Projected F2 data for pion and kaon from the Sullivan process with flexible choice of x and Q^2 bins – simulation
  - With projected F2 data, see what uncertainties one gets for, e.g. the gluon PDFs.
  - Develop/improve upon statistical models, e.g. M. Alberg et al. – here, get the proton, pion, kaon PDFs from a detailed balance statistical approach

Introduction and Background

Motivation: quarks, gluons, hadrons…

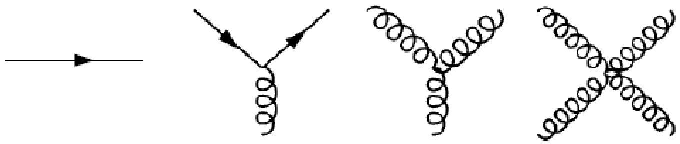
- The strong force is described in terms of coloured quarks and gluons

$$L_{QCD} = \bar{\psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi - g\left(\bar{\psi}\gamma^{\mu}T_{a}\psi\right)A^{a}_{\mu} - \frac{1}{4}G^{a}_{\mu\nu}G^{\mu\nu}_{a}$$

“color”

a quark

a gluon



- But, only colour-neutral hadrons can be detected – colour confinement

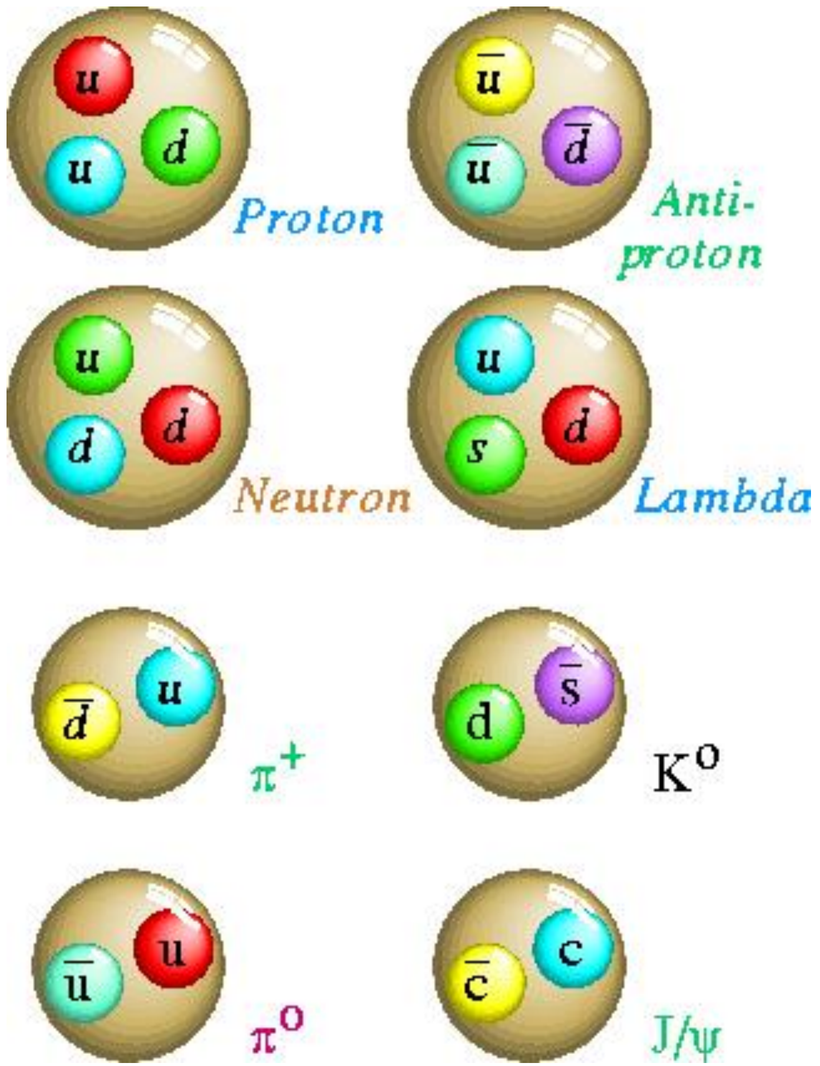
> ➢ How can one understand pions, kaons, protons or neutrons in terms of quarks and gluons?

Hadrons are made of quarks

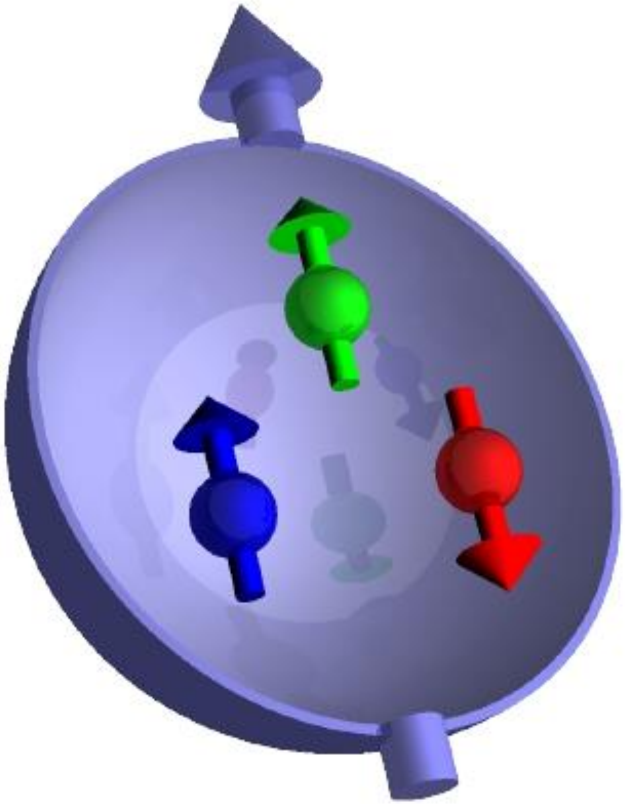
- 6 flavours (and 3 colours)
  - Up, down, strange
  - Charm, bottom, top

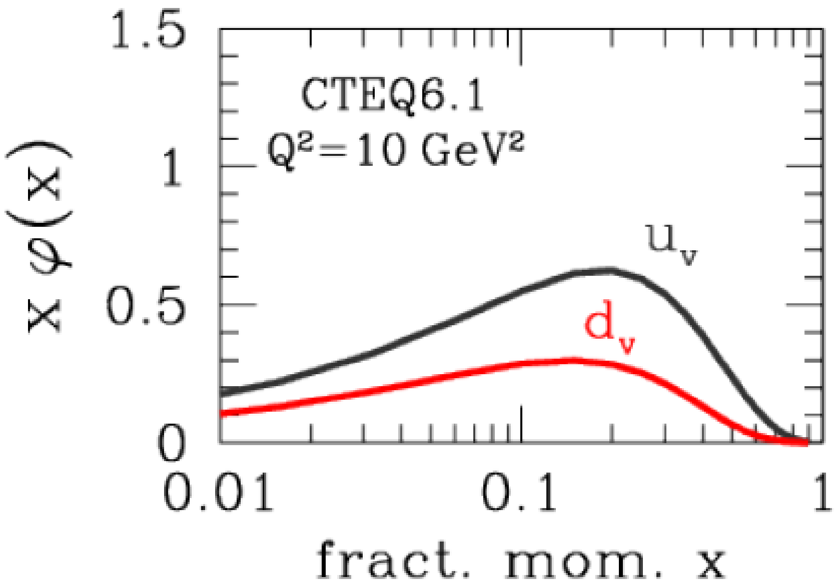
  - Spin 1/2
  - Isospin (u=1/2, d=-1/2)
  - Strangeness (s=1)

- Confined in colourless hadrons
  - Mesons – 2 quarks
  - Baryons – 3 quarks
  - Tetraquarks - ?
  - Pentaquarks - ???

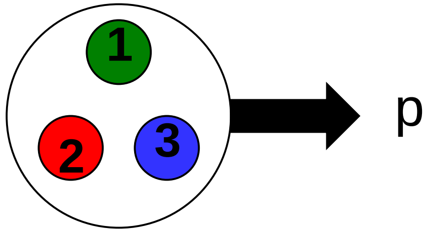


Nucleons are made of 3 quarks…

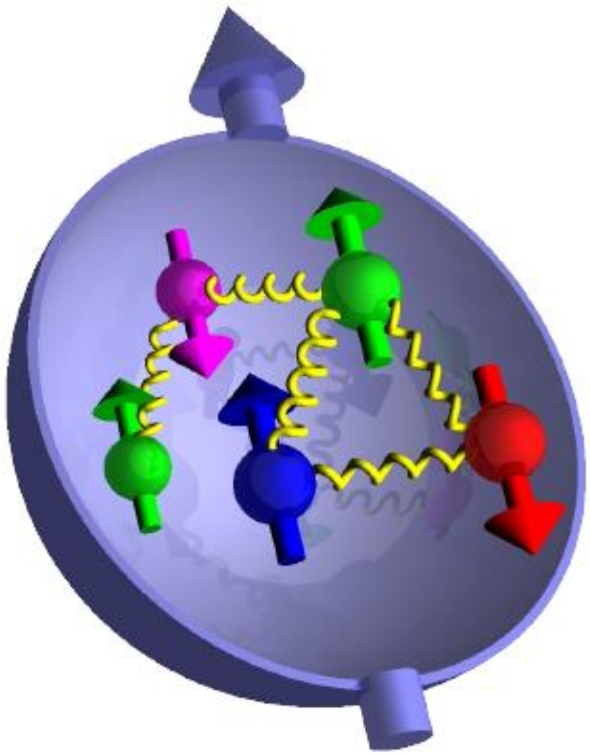


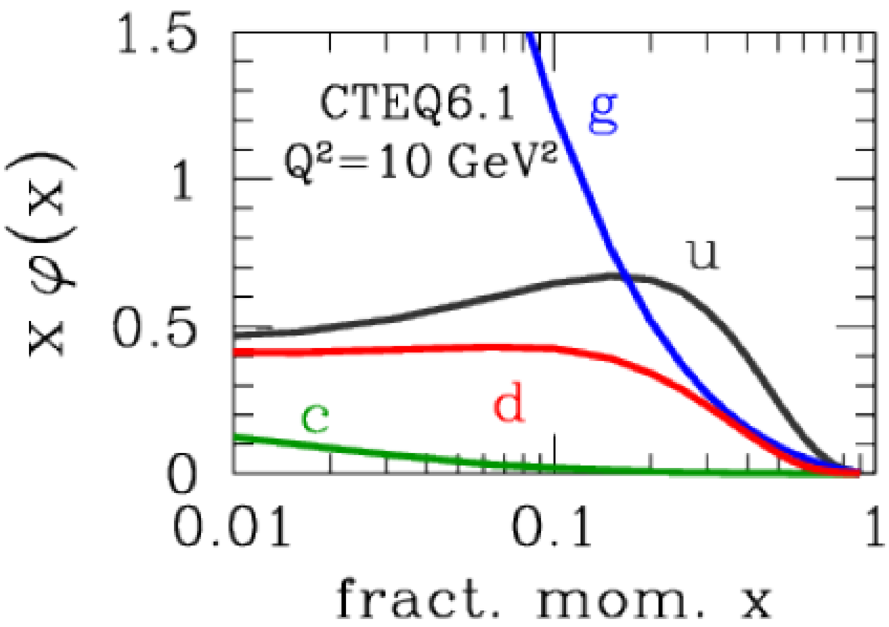


- x is the fraction of momentum carried by a quark in a nucleon momentum moving quickly to the right (here)

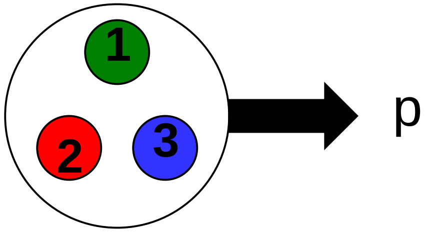


…and gluons, and sea quarks…

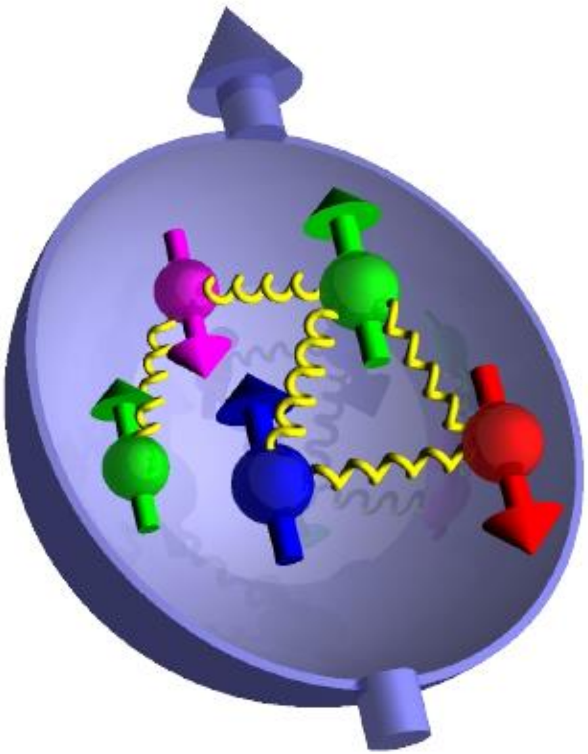


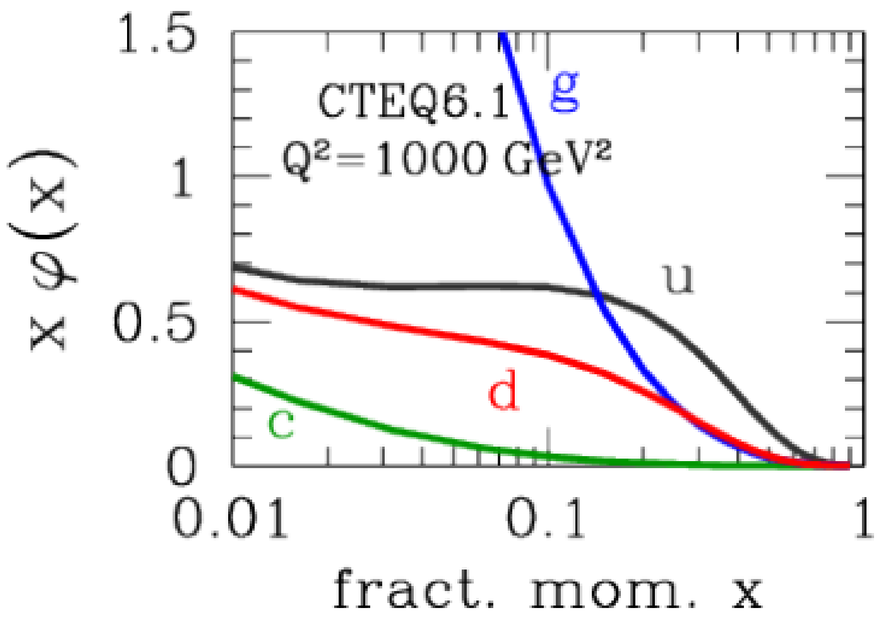


- x is the fraction of momentum carried by a quark in a nucleon momentum moving quickly to the right (here)

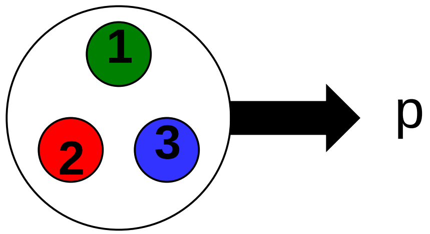


…and gluons, and sea quarks…

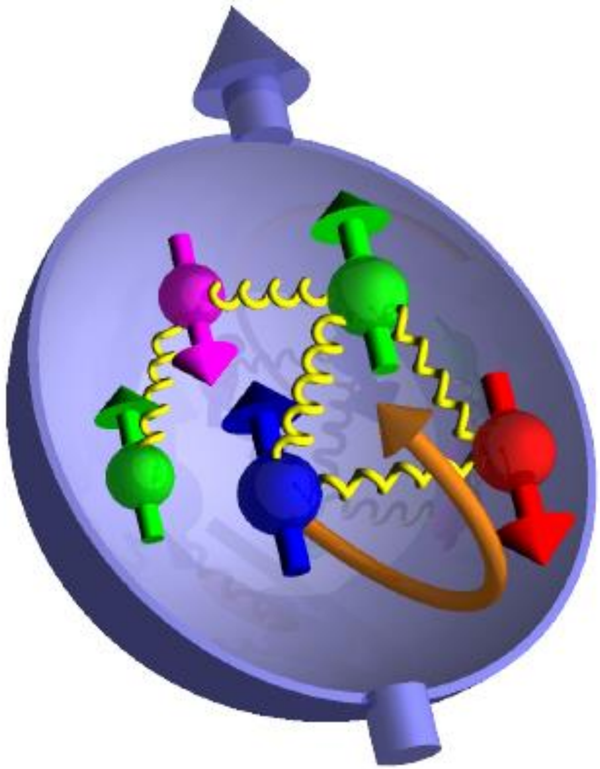


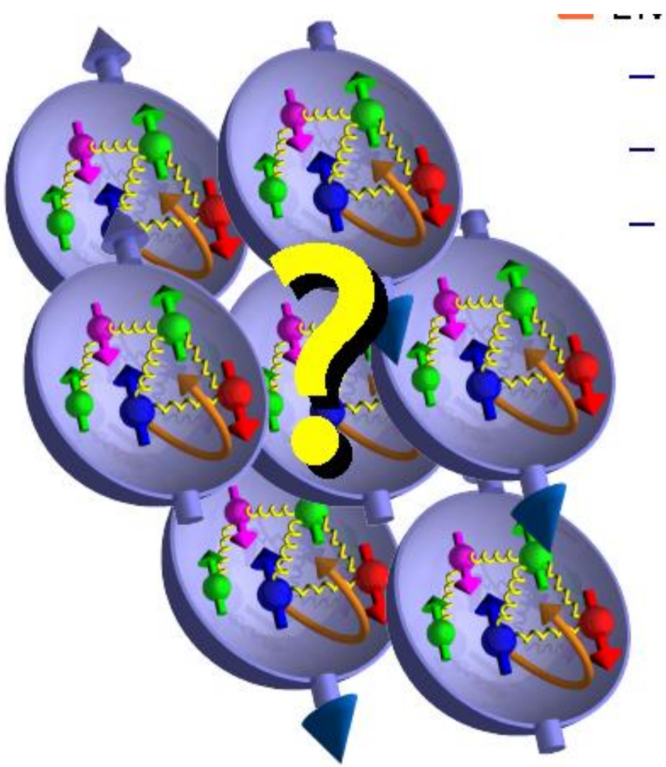


- x is the fraction of momentum carried by a quark in a nucleon momentum moving quickly to the right (here)

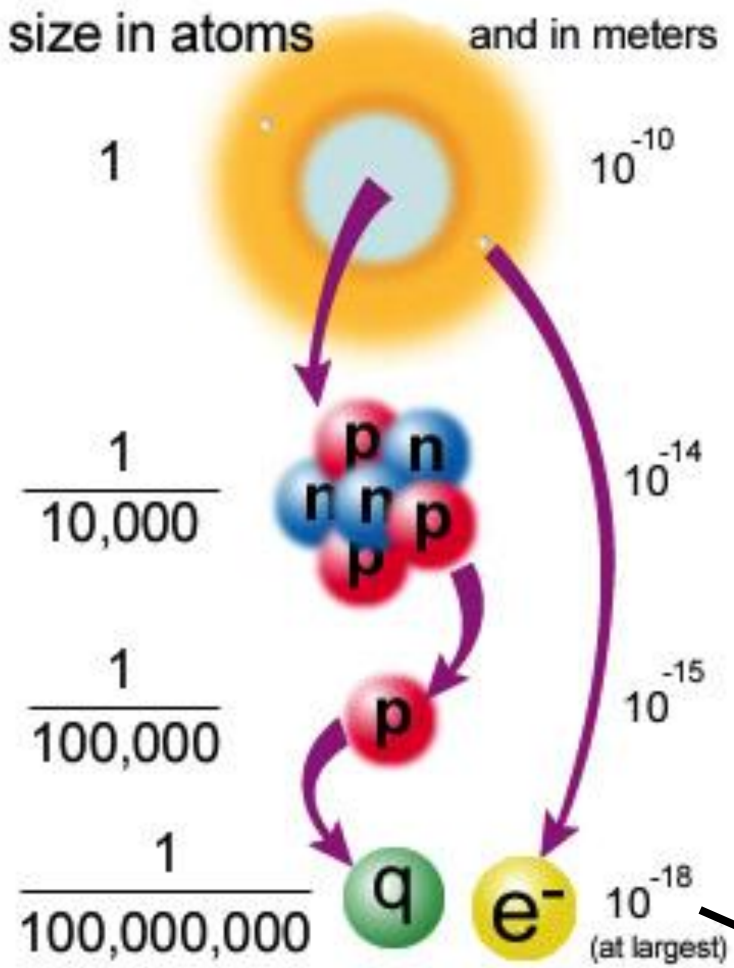


…spinning and orbiting around…and interacting





How to probe the nucleons / quarks?



Electron/lepton scattering experiments employ high momentum point-like leptons, + electromagnetic interactions, which are well understood, to probe hadronic structure (which isn't).

High energy electrons are a great tool for the job!

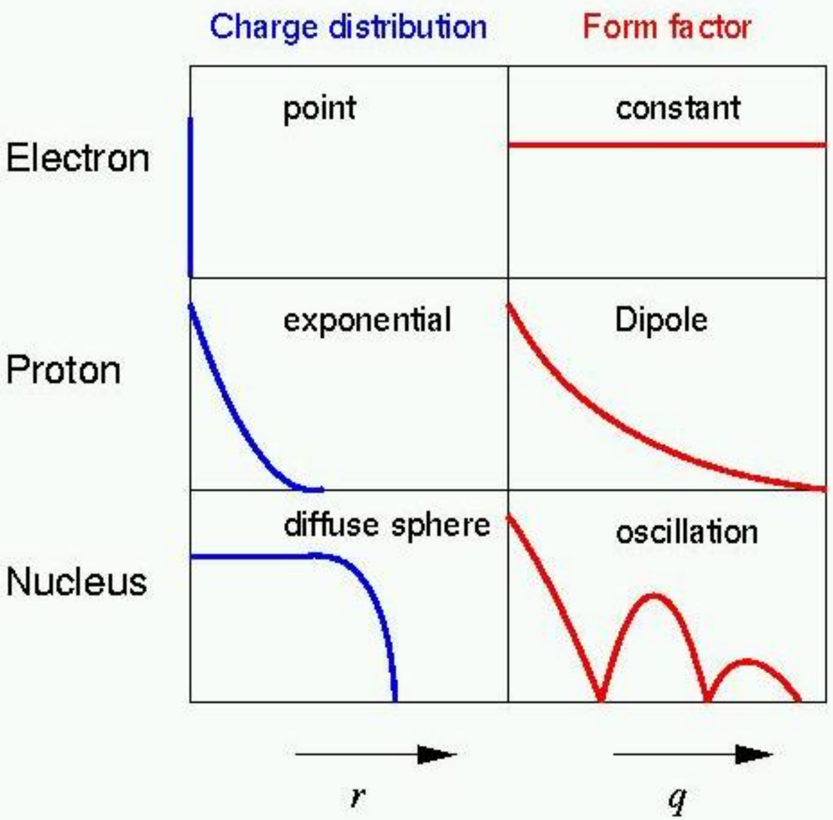
short distance -> large momentum (Uncertainty Principle!)

$$d_{probed} \propto \lambda\!\!\!^{-} = \frac{\hbar}{p} \approx 10^{-18}\ \mathrm{m}$$

Charge and Magnetic Moment Distributions

Probability of *elastic* interaction: $\frac{d\sigma}{d\Omega} \Big/ \left(\frac{d\sigma}{d\Omega}\right)_{point} = \left[\frac{G^2{}_E(Q^2) + \tau G^2{}_M(Q^2)}{1+\tau} + 2\tau G^2{}_M(Q^2)\tan^2\frac{\theta}{2}\right]$ $\qquad \tau = \frac{Q^2}{4M^2}$

- Form Factors are (in some limit) Fourier transforms of charge and magnetic moment distributions



- Elastic cross section

$$\left(\frac{d\sigma}{d\Omega}\right)_{\exp} = \left(\frac{d\sigma}{d\Omega}\right)_{\mathrm{Mott}} \left|F(q^2)\right|^2$$

- Form factor

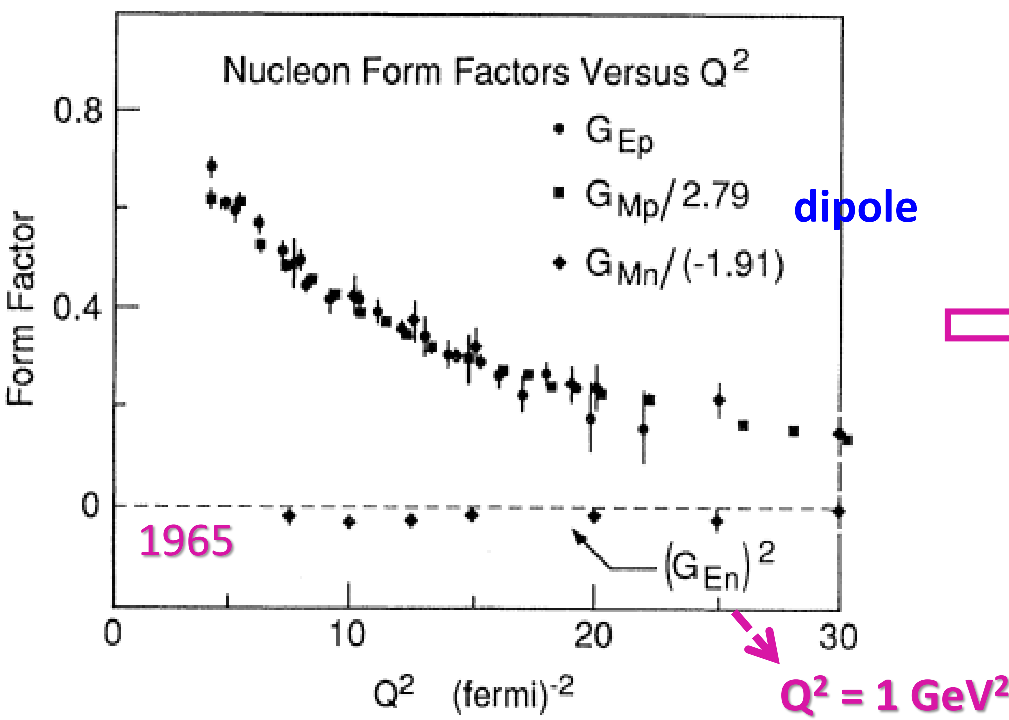
$$F(q^2) = \int e^{iqx/\hbar}\rho(x)d^3x$$

The form factor as a Fourier transformation of the charge distribution is a non-relativistic concept.

How Do the Charge and Magnetic Moment Distribute?

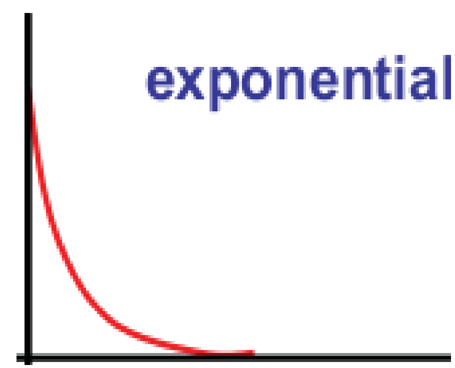
Probability of elastic interaction: $\frac{d\sigma}{d\Omega} \Big/ \left(\frac{d\sigma}{d\Omega}\right)_{point} = \left[\frac{G^2{}_E(Q^2) + \tau G^2{}_M(Q^2)}{1+\tau} + 2\tau G^2{}_M(Q^2) \tan^2\frac{\theta}{2}\right]$ $\quad \tau = \frac{Q^2}{4M^2}$

- The Q^2 dependence of form factors was measured…



⇨

Distribution of Charge or Magnetic Moment

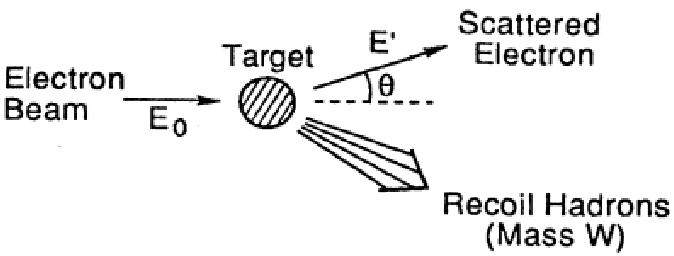


Caveat: The Form Factor as the Fourier transformation of a charge distribution is a non-relativistic concept.

Matter Puzzle: What's Inside the Proton?

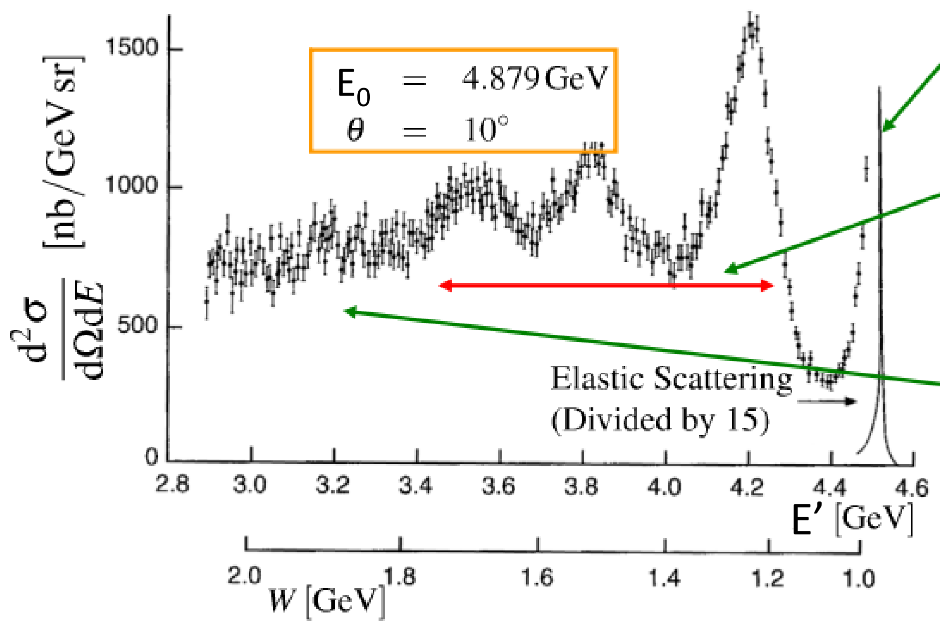
- Is the proton elementary?

To find out increase the probe's ability of resolving structure (decrease $\frac{\hbar}{Q}$)



$$E' = \frac{E_0 - \frac{W^2 - M^2}{2M}}{1 + \frac{2E_0}{M}\sin^2\frac{\theta}{2}} \qquad \nu = E_0 - E' \qquad y = \frac{\nu}{E_0} \qquad x = \frac{Q^2}{2M\upsilon}$$

$$W^2 = M^2 + 2M\upsilon - Q^2$$



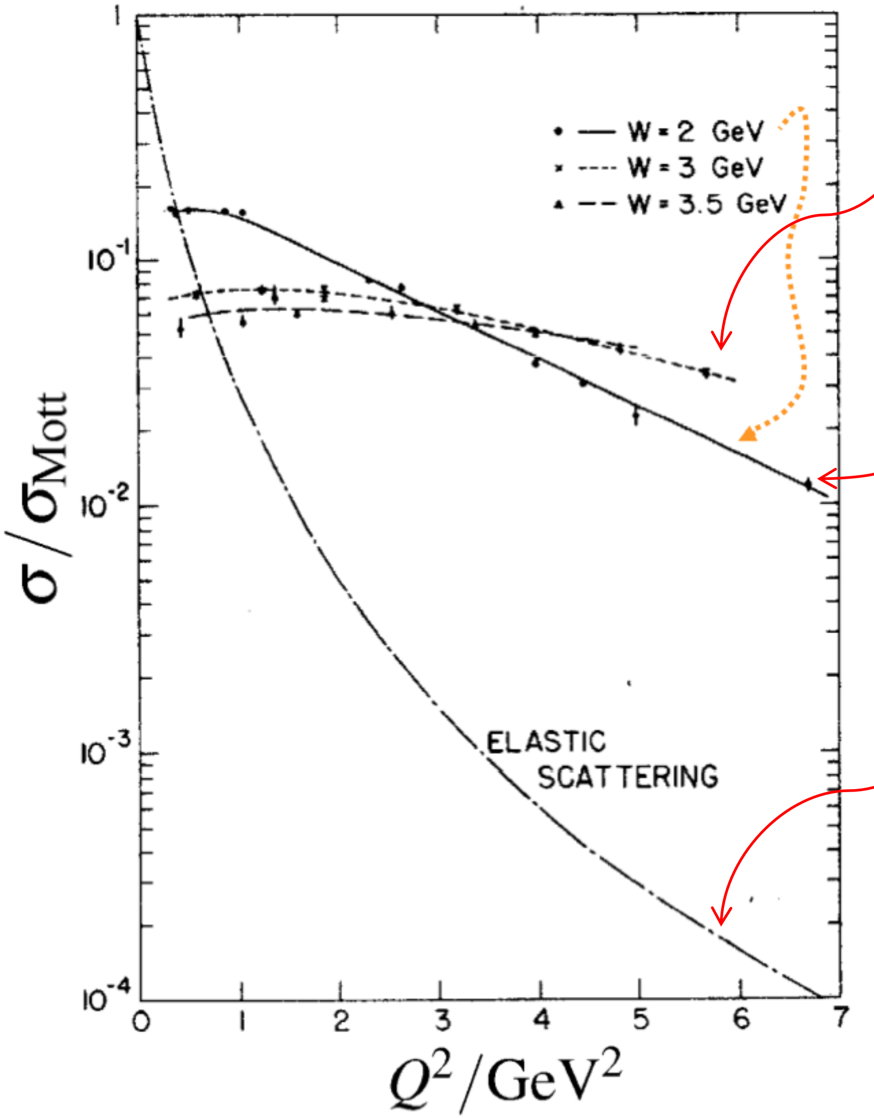
Elastic scattering: proton stays intact, $W = M$

Inelastic scattering: proton gets excited, produce excited states or proton's resonances, $W = M_{resonance}$

Deep inelastic scattering: proton breaks up and we end up with a many particle final state, W = *large*

Looking deep inside the Proton

M.Breidenbach et al.,
Phys. Rev. Lett. 23 (1969) 935



- **Deep Inelastic scattering** cross sections almost independent of Q^2!
  i.e. "Form factor" ➔1

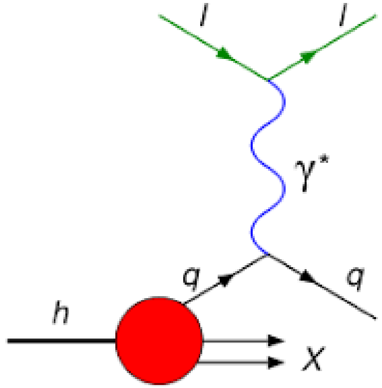
Scattering off point-like objects within the proton

- **Inelastic scattering** cross sections only weakly dependent on Q^2

- **Elastic scattering** falls of rapidly with Q^2 due to the proton not being point-like (i.e. form factors)

$$\frac{\sigma}{\sigma_{\mathrm{Mott}}} = \left(\frac{1}{(1+Q^2/0.71)^2}\right)^2 \propto Q^{-8}$$

Structure Functions in Deep Inelastic Electron-Nucleon Scattering



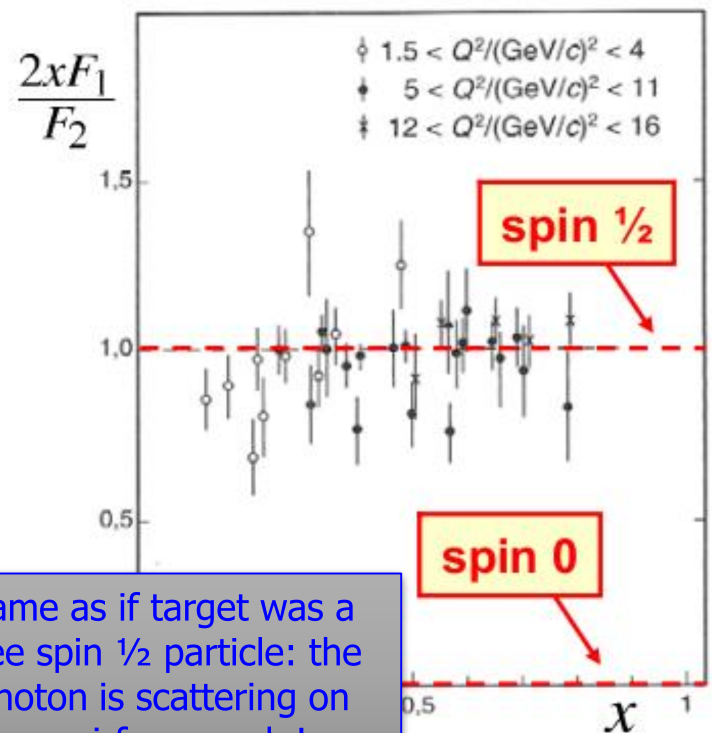
Probability of inelastic interaction:

$$\frac{d^2\sigma}{d\Omega dE'} = \frac{\alpha^2}{4E^2{}_0 \sin^4\frac{\theta}{2}} \cos^2\frac{\theta}{2}\left[\frac{1}{\nu}F_2(x,Q^2) + \frac{2}{M}F_1(x,Q^2)\tan^2\frac{\theta}{2}\right]$$

Unpolarized *"Structure Functions"* $F_1(x,Q^2)$ and $F_2(x,Q^2)$:

- Account for the sub-structure of the protons and neutrons
- x = fraction of nucleon momentum carried by struck quark
- Give access to *partonic structure* of the nucleon, i.e.

$$F_2^p = x\left[\frac{4}{9}(u+\overline{u}) + \frac{1}{9}(d+\overline{d}) + \frac{1}{9}(s+\overline{s})\right]$$

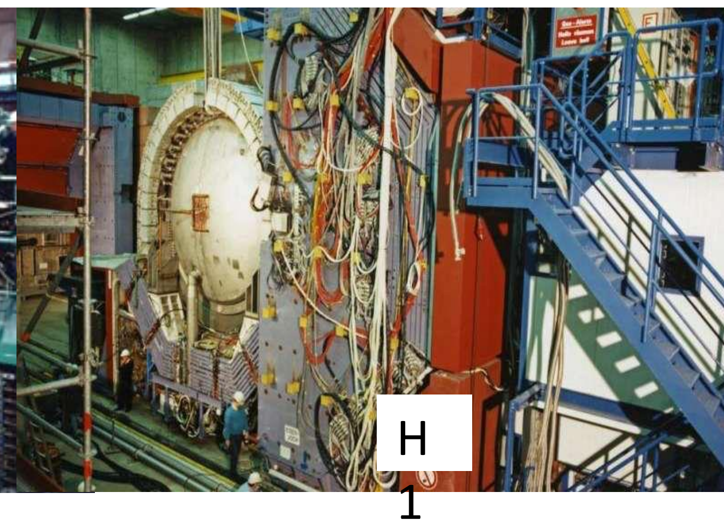


Same as if target was a free spin ½ particle: the photon is scattering on quasi-free quark !

- Comparing the DIS cross section formula with the Mott and Dirac elastic cross sections for particles of mass m = xM and spin 1/2
- *If point-like constituents were spin zero particles, we would expect F_1 to be zero*

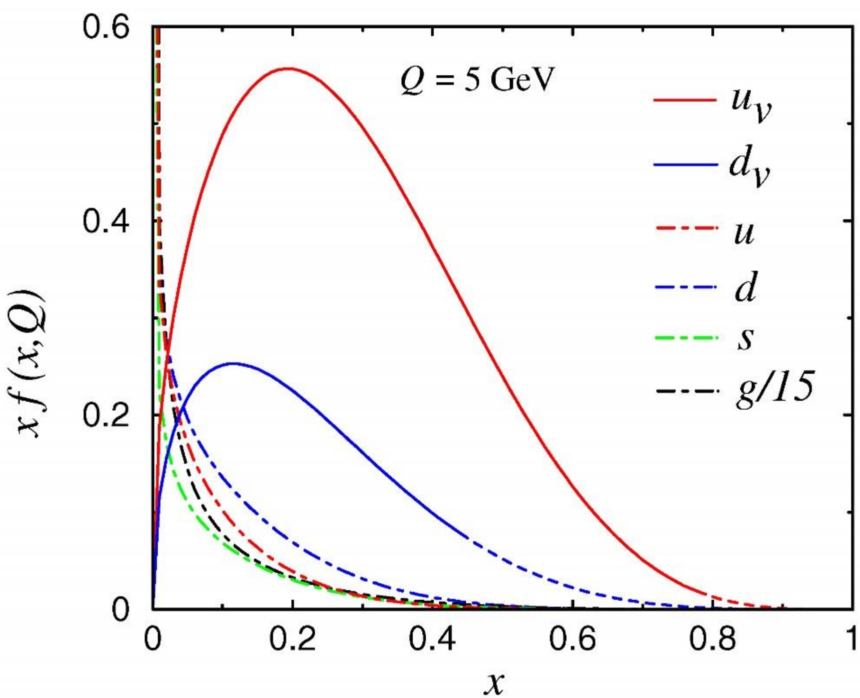
Fast forward….

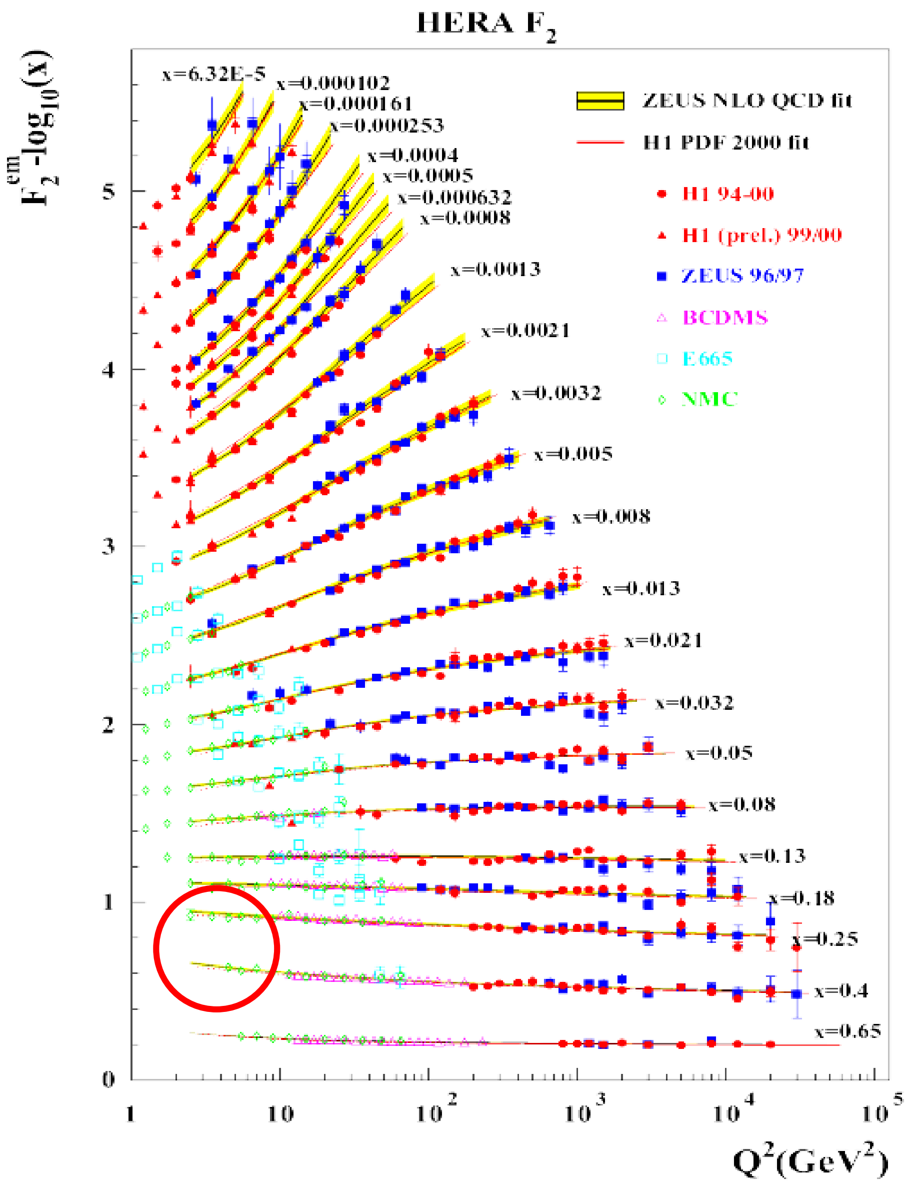
30+ years of charged lepton Deep Inelastic Scattering at *<u>multiple</u>* laboratories including SLAC (to ~2000), CERN 80-90s EMC, NMC, BCDMS..), DESY (90s – 21st century H1, ZEUS,...), and ***<u>more!</u>***



Q^2 Evolution of the F_2 *Proton* Structure Function

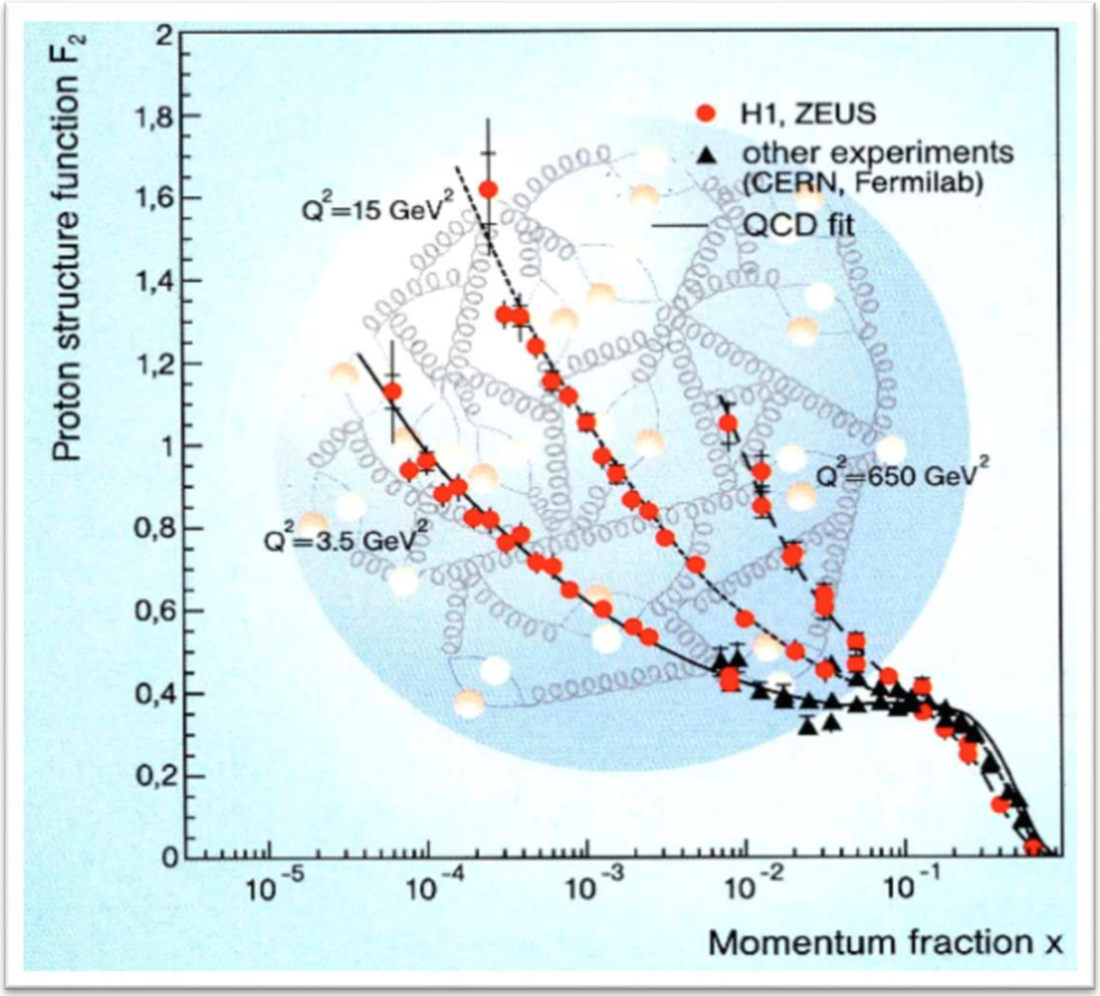
F_2^p Structure Function measured over impressive range of x and Q^2



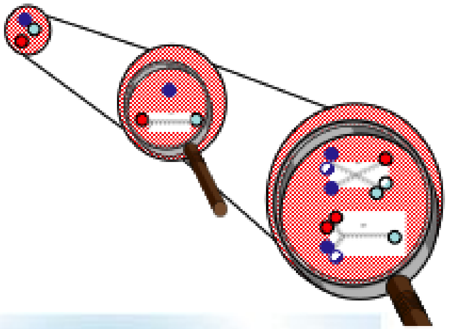


Allows extraction of *Parton Distribution Functions* f(x,Q^2) - think momentum distribution of quarks

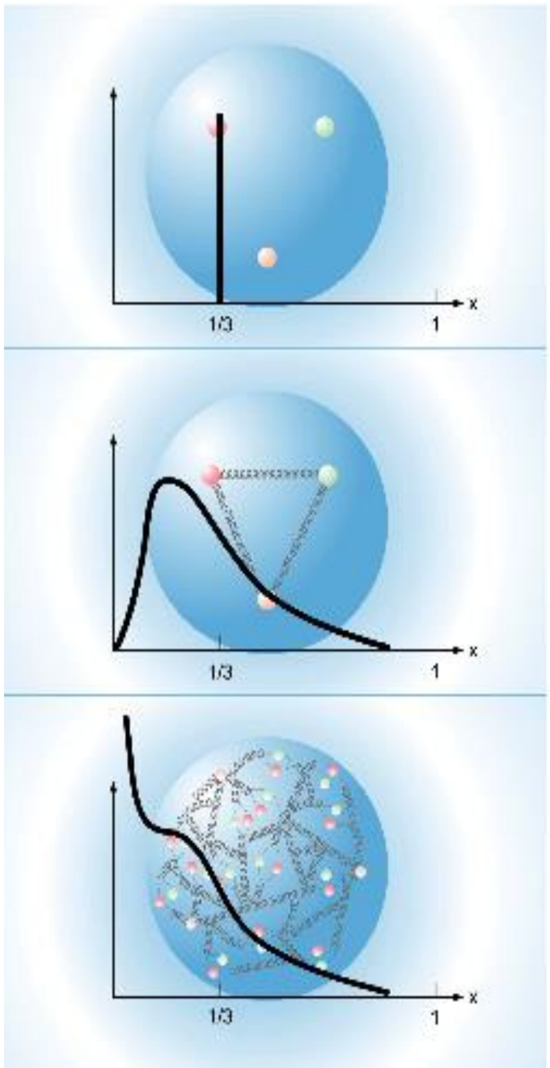
Scaling Violations

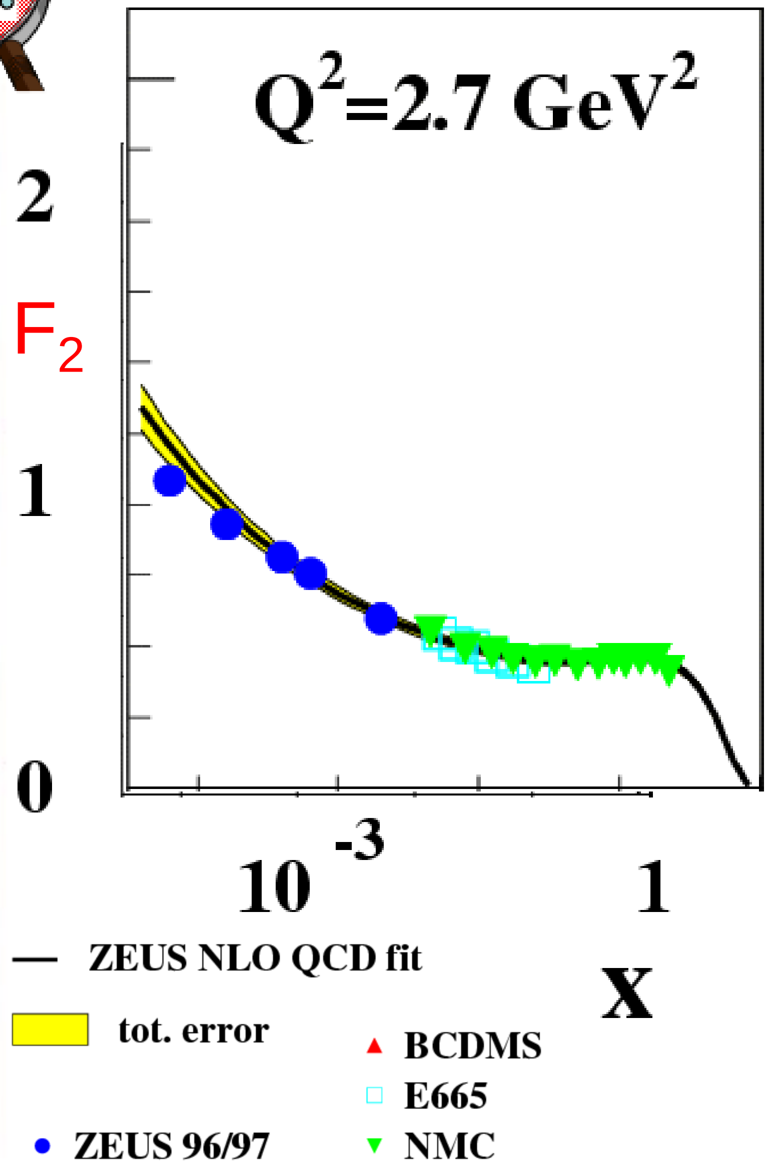


- Scaling violation is due to the fact that the quarks radiate gluons that can "materialize" as q-qbar pairs (sea quarks)
- Increasing Q^2 increases the resolution of the probe ($\sim\hbar/\sqrt{Q^2}$) and thus increases the probability of seeing these (abundant) low x partons
- The parton distribution functions (PDFs) can not be calculated from first principle of QCD but their Q^2 dependence is calculable in perturbative QCD using the DGLAP evolution equations



Proton Structure Function F_2





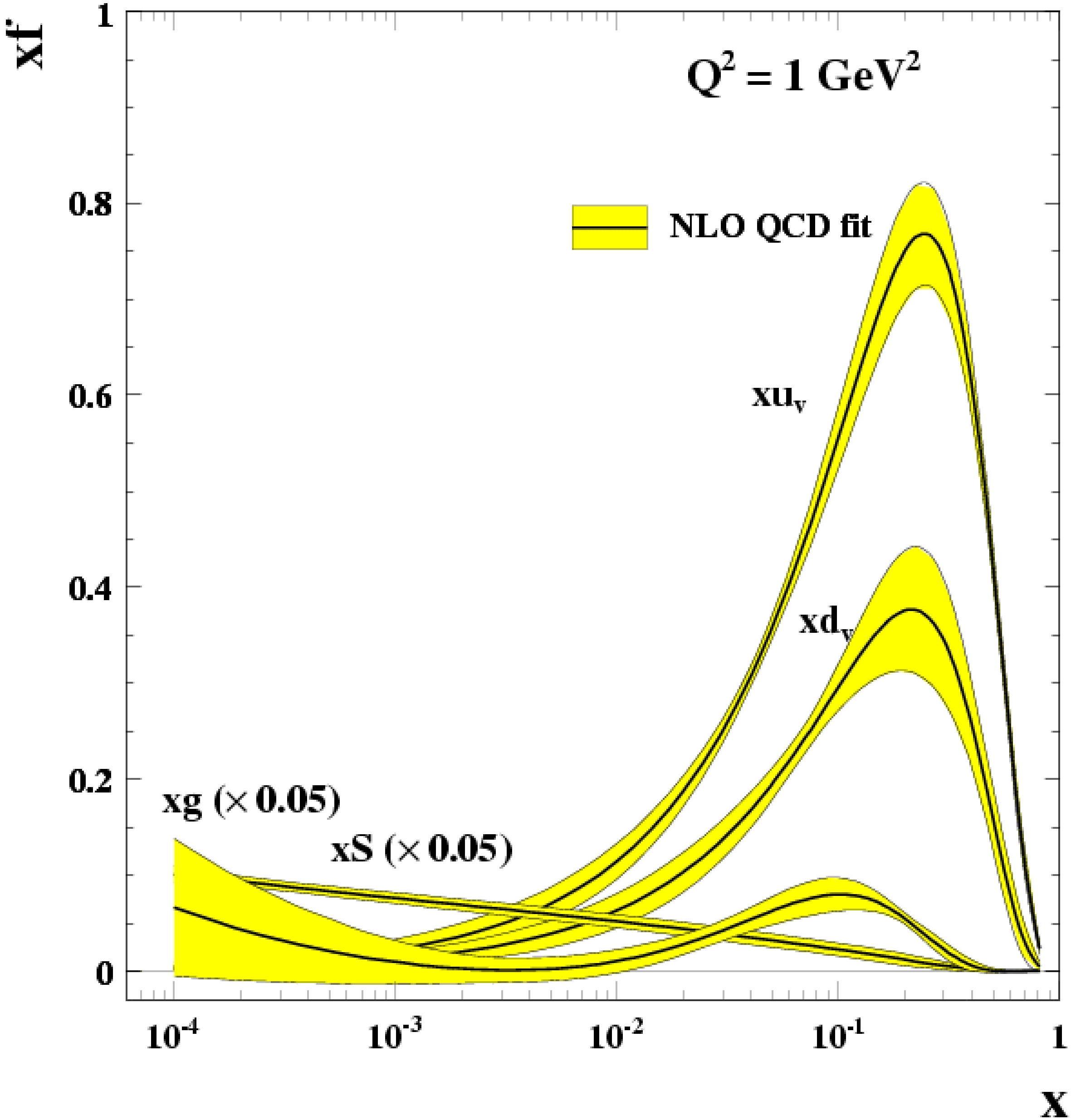
Q^2 dependence described by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution equations

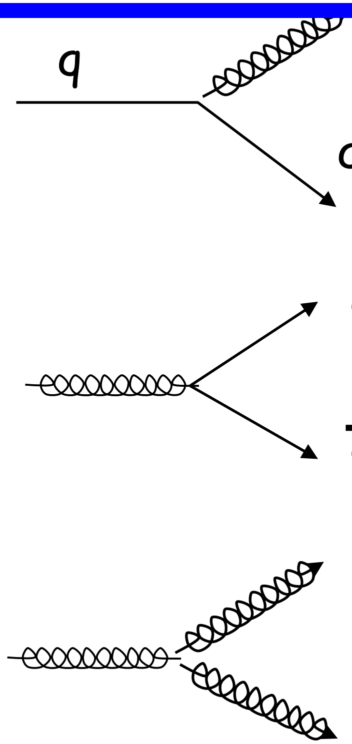
Increasing Q^2:

- High x decrease
- Low x increase

Allows extraction of Parton Distribution Functions $f(x,Q^2)$ through Q^2 evolution From high x, low Q to high Q, low x

Parton Distribution Functions and QCD Evolution





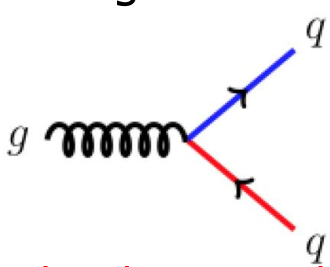
Valence quarks maximum around *x=0.2; f(x) →0* for x→1 and x→0

Sea quarks and gluons - contribute at low values of x

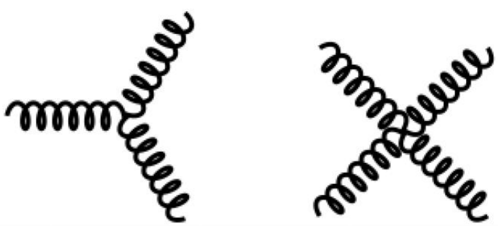
Quantum Chromo Dynamics

Gluons are the messengers for the quark-quark interactions

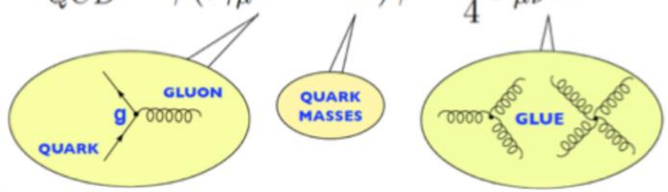
Quantum Chromo Dynamics (QCD) is the theory that governs their behaviour



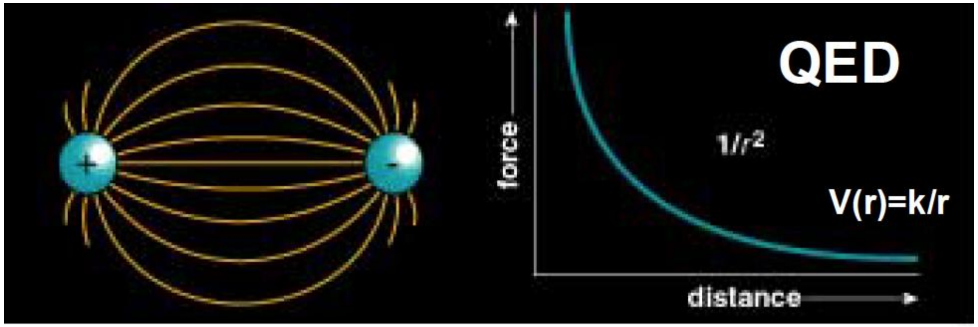
Gluons carry color charge, and we can draw 3- and 4- gluon diagrams *(self-interaction)*

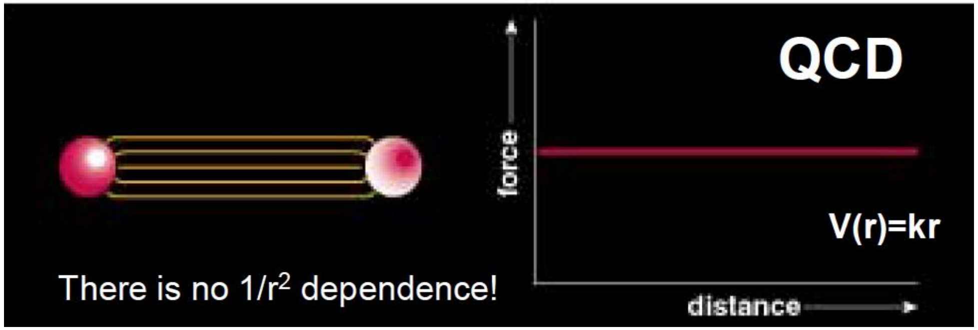


$$\mathcal{L}_{QCD} = \bar{\psi}(i\gamma_\mu D^\mu - m)\psi - \frac{1}{4}G_{\mu\nu}G^{\mu\nu}$$



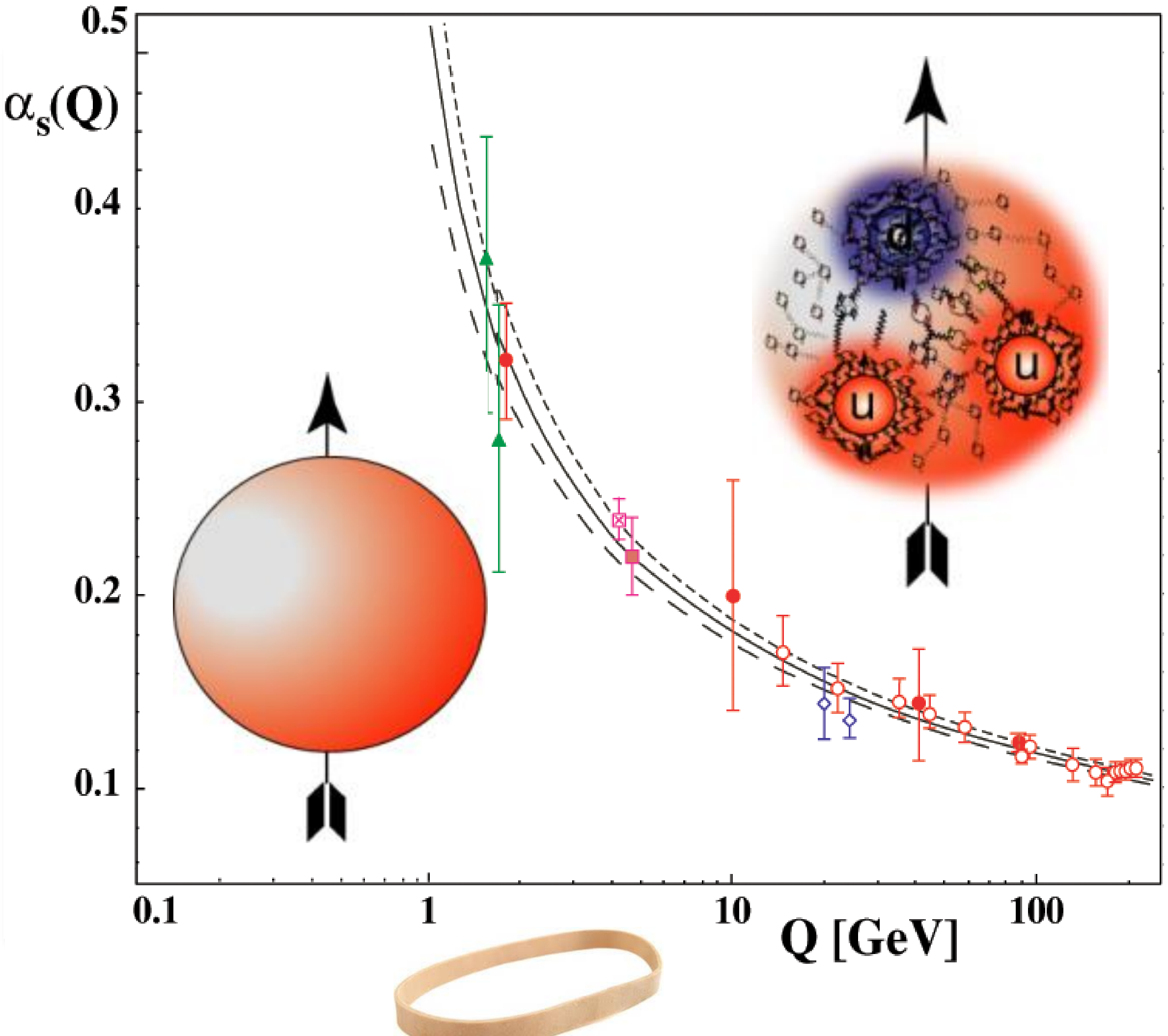
The strong force does not get weaker with large distances (opposite to the EM force) and blows up at distances around 10^{-15} m (the radius of the nucleon)





Quantum ChromoDynamics

2004 David Gross, David Politzer and Frank Wilczek



At short distances
quarks move as though they are free → **Asymptotic freedom**
Physics at short distance is understood through perturbation theory - $\alpha_s(m_Z)$= 0.1189(10)
Perturbative QCD tested up to 1% level

At longer distances
Confinement ensures that only hadronic final states are observed
Quarks can be removed from the proton, but cannot be isolated!!!
We never see a free quark
QCD still unsolved in non-perturbative region
Insights into soft phenomena exist through qualitative models and quantitative numerical (lattice) calculations

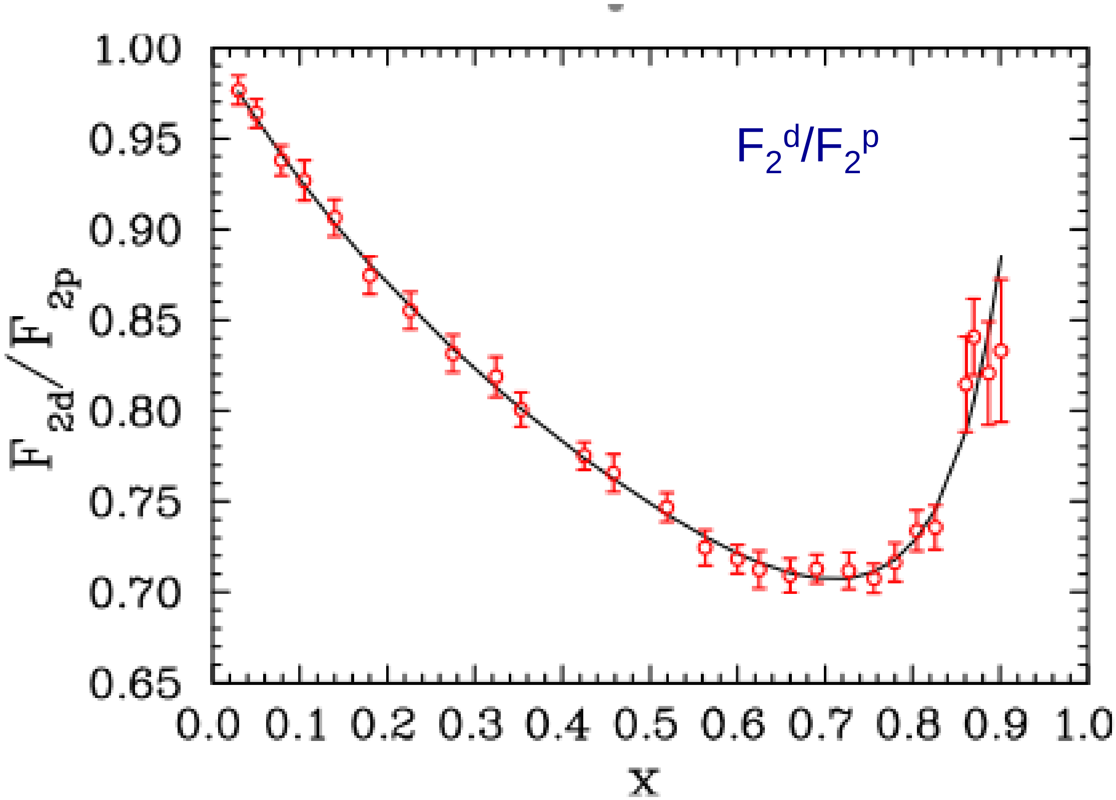
Puzzles

Important alert: the deuteron is also a nucleus!

Neutron structure is typically derived from deuterium target data by subtracting proton data

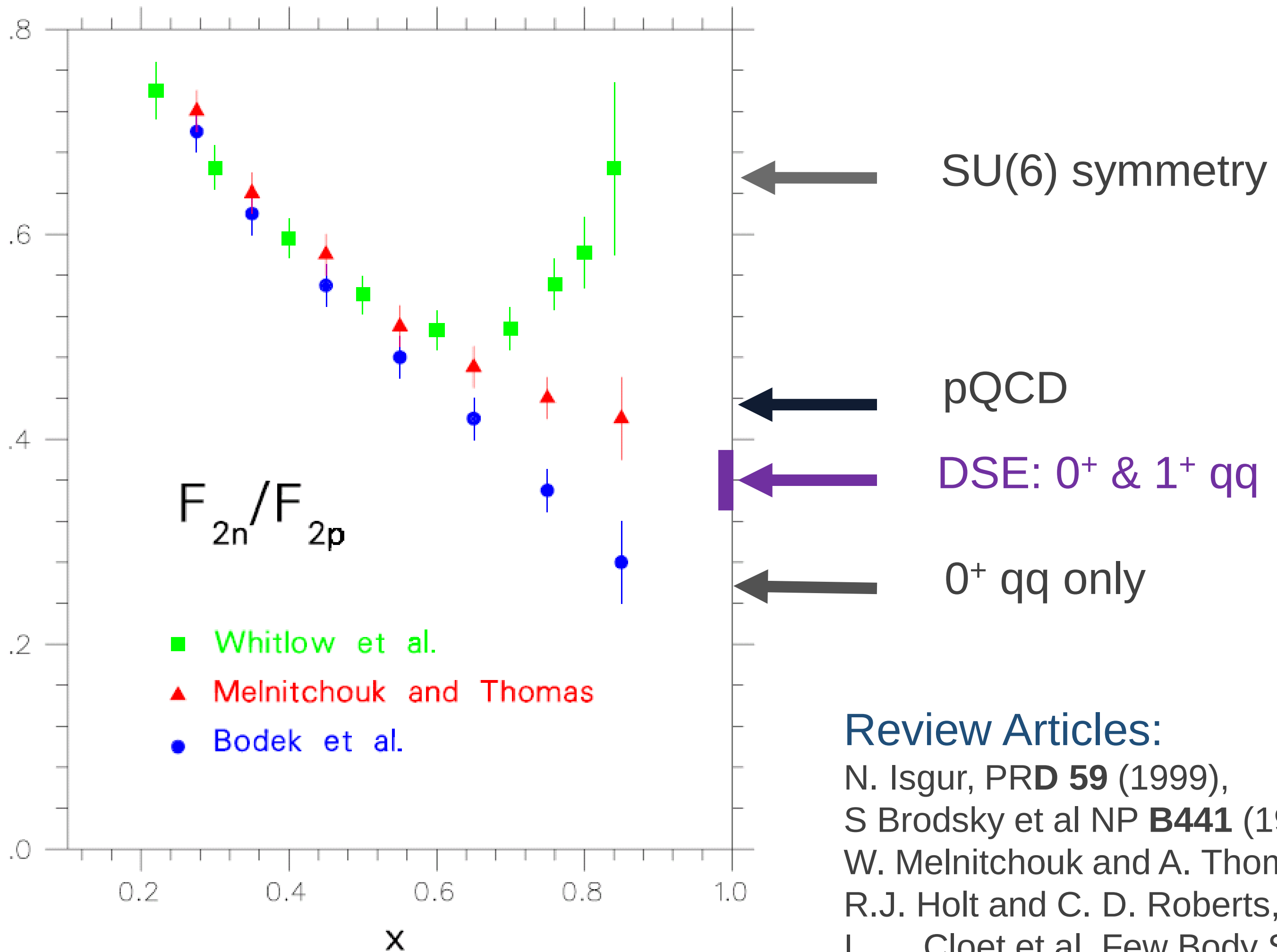
…..but…..

Large uncertainty in unfolding nuclear effects (Fermi motion, off-shell effects, deuteron wave function, coherent scattering, final state interactions, nucleon structure modification ("EMC"-effect),……….



F_2^n/F_2^p (and, hence, d/u) is essentially unknown at large x:

- Conflicting fundamental theory pictures
- Data inconclusive due to uncertainties in deuterium nuclear corrections

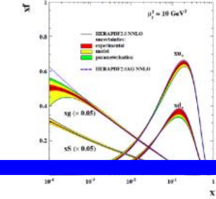


Review Articles:

N. Isgur, PR**D 59** (1999),
S Brodsky et al NP **B441** (1995),
W. Melnitchouk and A. Thomas PL **B377** (1996) 11,
R.J. Holt and C. D. Roberts, arXiv:1003.4666 [nucl-th],
I. Cloet et al, Few Body Syst. **46** (2009) 1.

Your textbook may be lying…

Large Uncertainties on Large x Valence pdfs

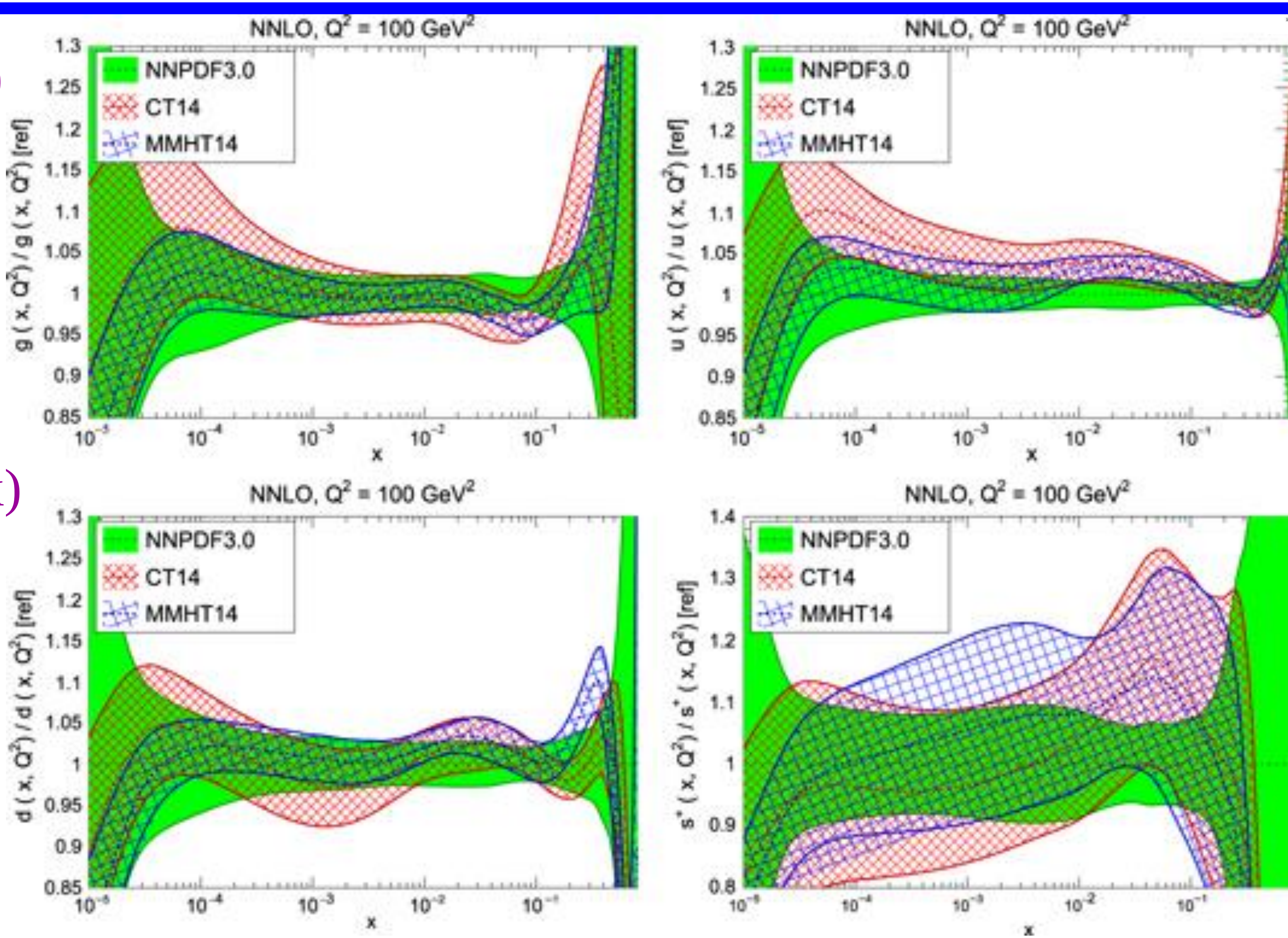


g(x)

u(x)

d(x)

s(x)



Lack of precision at large x hidden by fact that PDFs tend to zero

From Paul Newman

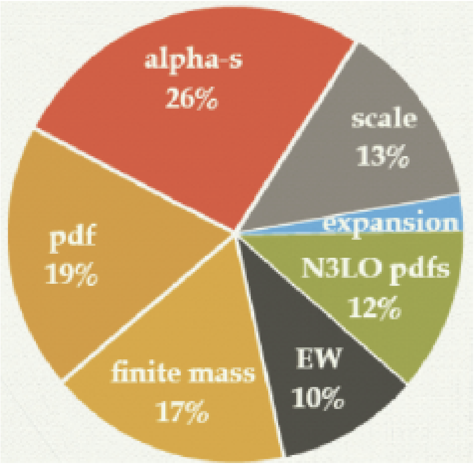
Higgs X-Section / Coupling PDF Uncertainties

Theoretical Uncertainties

After N^3LO calculation of gluon-fusion Higgs cross section at 13 TeV → much reduced scale uncertainty

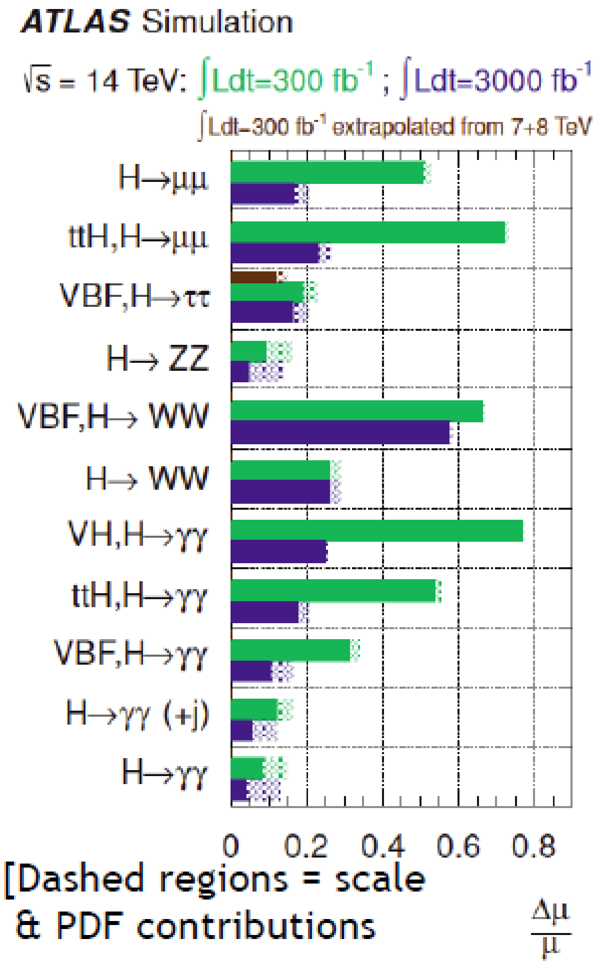
... largest sources of unertainty:
- PDFs [1.9%]
- α_s [2.6%]

with additional 1.2% uncertainty on non-availability of N^3LO PDFs



[Anastasiou et al [1503.06056], Dulat, CERN Dec '15]

Projected Experimental Uncertainties



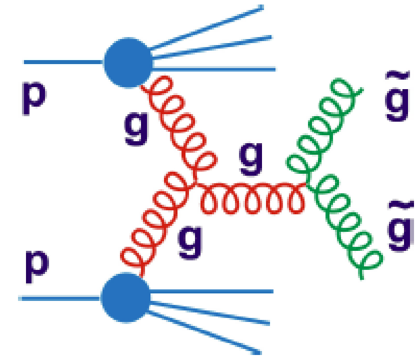
[Dashed regions = scale & PDF contributions

... much of Higgs sector becomes PDF limited in HL-LHC era ...
(though it's $x \sim 10^{-2}$, so not really today's topic)

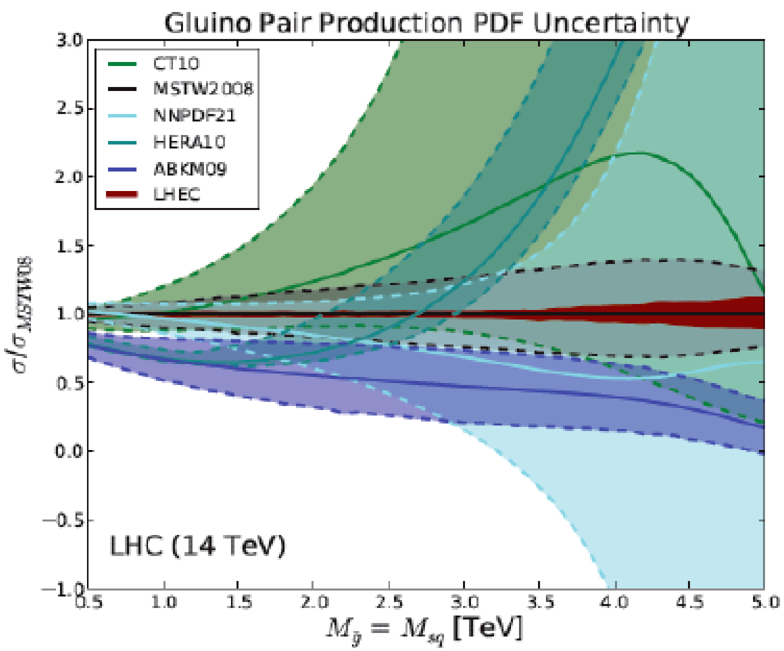
From Paul Newman

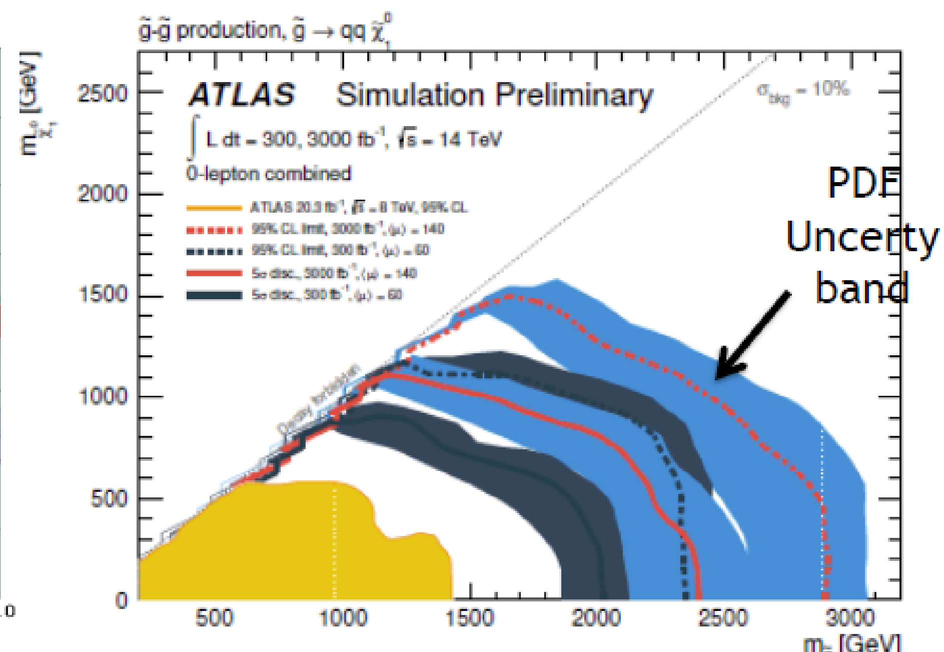
e.g. High Mass 2 Gluino Production

- Signature is excess @ large invariant mass
- Expected SM background (e.g. gg → gg) poorly known for s-hat > 1 TeV.

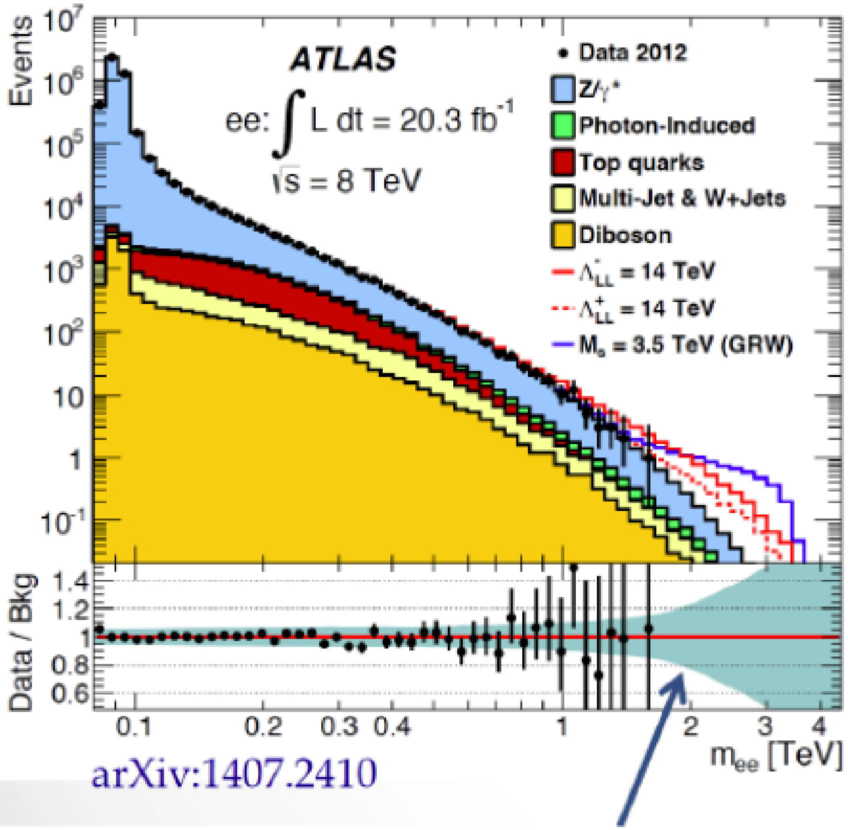


- Both signal & background uncertainties driven by error on gluon density … essentially unknown for masses much beyond 2 TeV

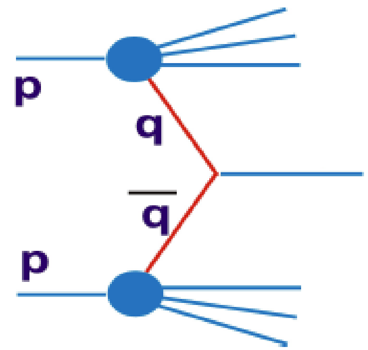




High x (Anti)-Quarks Matter Too …



arXiv:1407.2410



- BSM sensitivity through excess in high mass Drell-Yan limited by high x antiquark uncertainties as well as valence

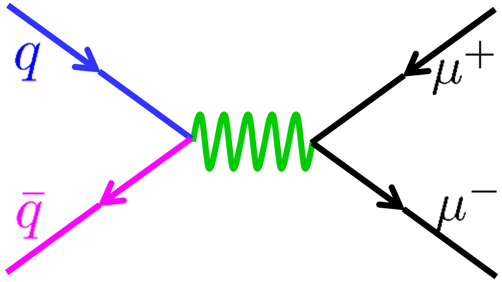
… bottom line is that much of the LHC search programme will become limited by the high x parton density uncertainties as we head towards the ultimate lumi of the LHC unless there is a transormation in precision in the meantime …

And then there is the

Pion and Kaon

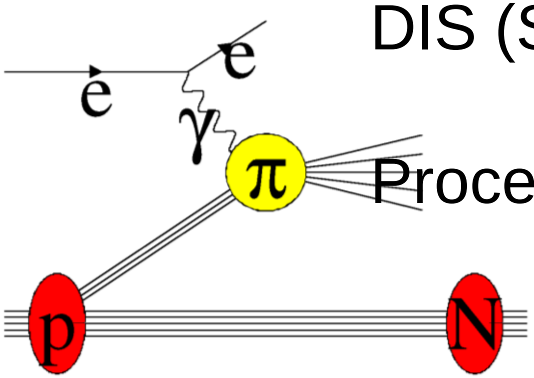
World Data on pion structure function F_2^{π}

Pion Drell-Yan

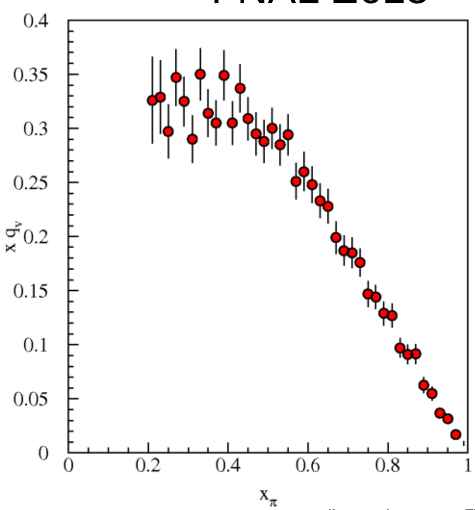


Data much more limited than nucleon…

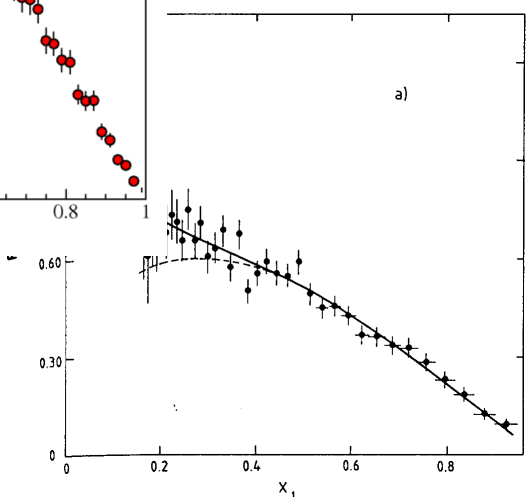
DIS (Sullivan Process)

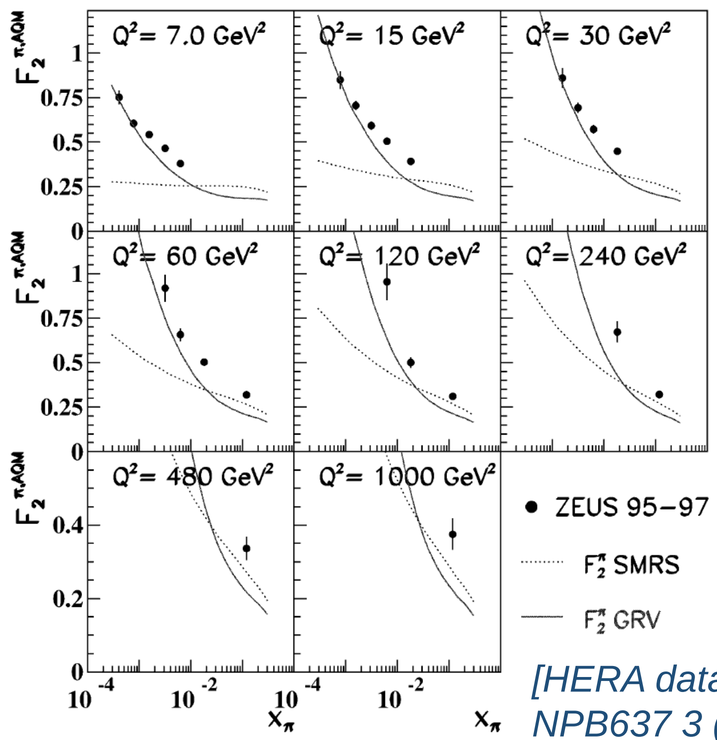


FNAL E615



CERN NA3





[HERA data [ZEUS, NPB637 3 (2002)]

Calculable Limits for Parton Distributions

- Calculable limits for ratios of PDFs at *x = 1*, same as predictive power of $x \rightarrow 1$ limits for spin-averaged and spin-dependent proton structure functions (asymmetries)

$$\boxed{\left.\frac{u_V^K(x)}{u_V^\pi(x)}\right|_{x\to 1} = 0.37\,, \quad \left.\frac{u_V^\pi(x)}{\bar{s}_V^K(x)}\right|_{x\to 1} = 0.29}$$

- On the other hand, inexorable growth in both pions' and kaons' gluon and sea-quark content at asymptotic Q^2 should only be driven by pQCD splitting mechanisms. Hence, also calculable limits for ratios of PDFs at *x = 0*, *e.g.*,

$$\boxed{\lim_{x\to 0} \frac{u^K(x;\zeta)}{u^\pi(x;\zeta)} \stackrel{\Lambda_{\rm QCD}/\zeta\simeq 0}{\rightarrow} 1}$$

The inexorable growth in both pions' and kaons' gluon content at asymptotic Q^2 provides connection to gluon saturation.

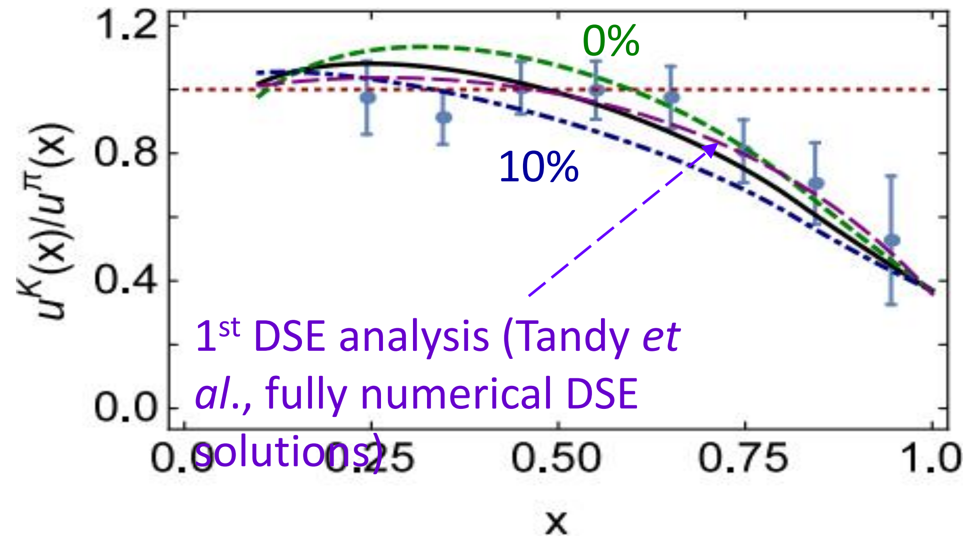
Gluon Content in Kaon and Pion

Based on Lattice QCD calculations and DSE calculations:

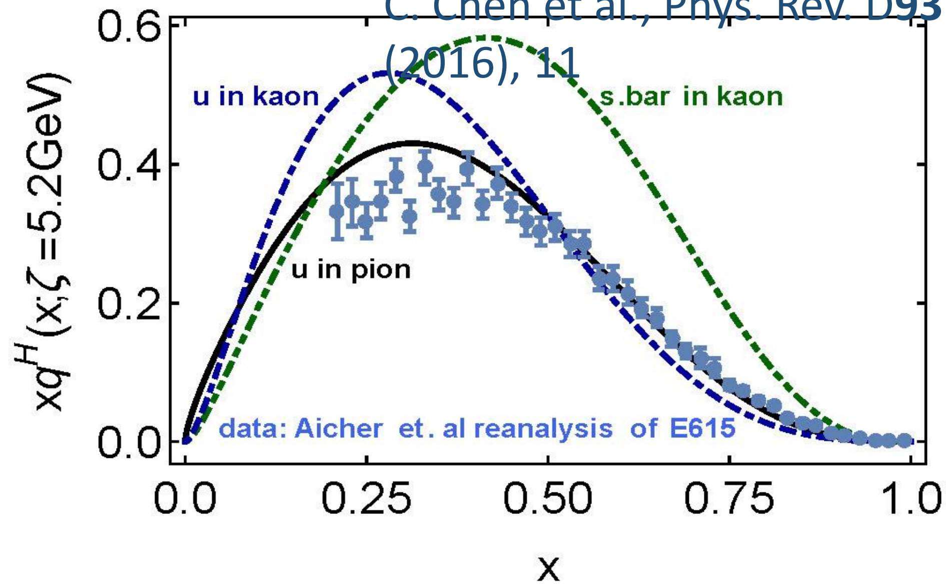
- ❑ Valence quarks carry 2/3 of the kaon's momentum at the light front, at the scale used for Lattice QCD calculations, or roughly 95% at the perturbative hadronic scale

- ❑ At the same scale, valence-quarks carry 52% of the pion's light-front momentum, or roughly 65% at the perturbative hadronic scale

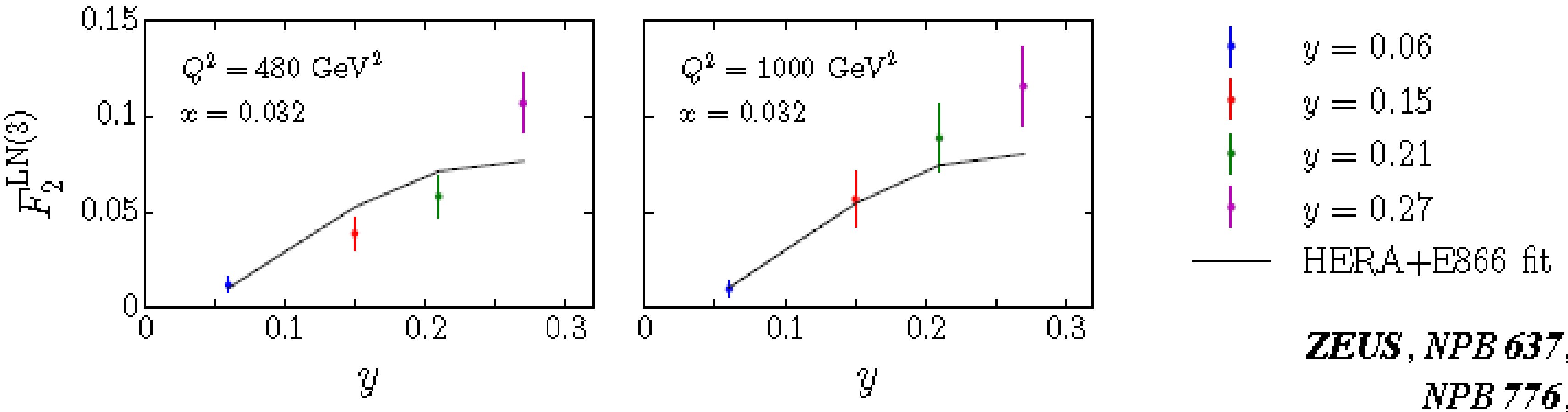
Thus, at a given scale, there is far less glue in the kaon than in the pion



C. Chen et al., Phys. Rev. D**93** (2016), 11



Combined Fit to HERA LN and E866 DY Data

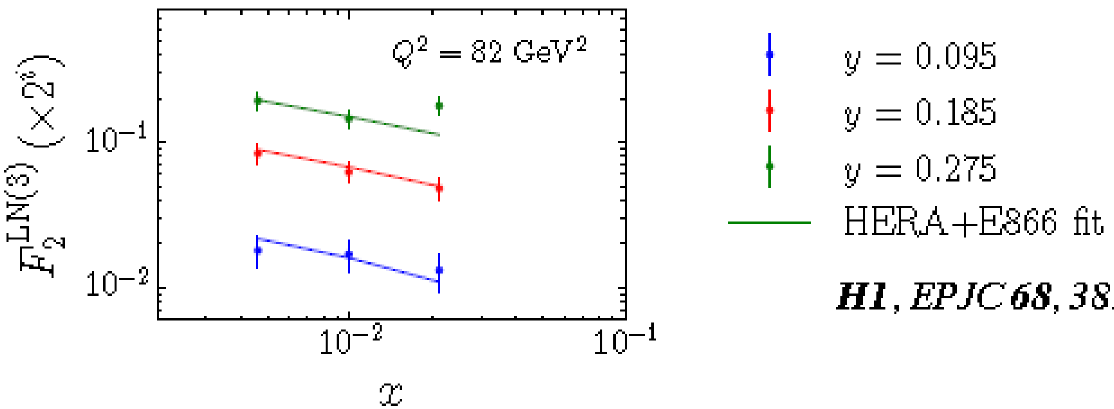


***ZEUS*, *NPB* 637, 3 (2002)**

***NPB* 776, 1 (2007)**

- ❑ Quality of fit depends on y-range fitted – to reduce model dependence fit up to y_{cut}=0.3 to which data can be described in term of π exchange

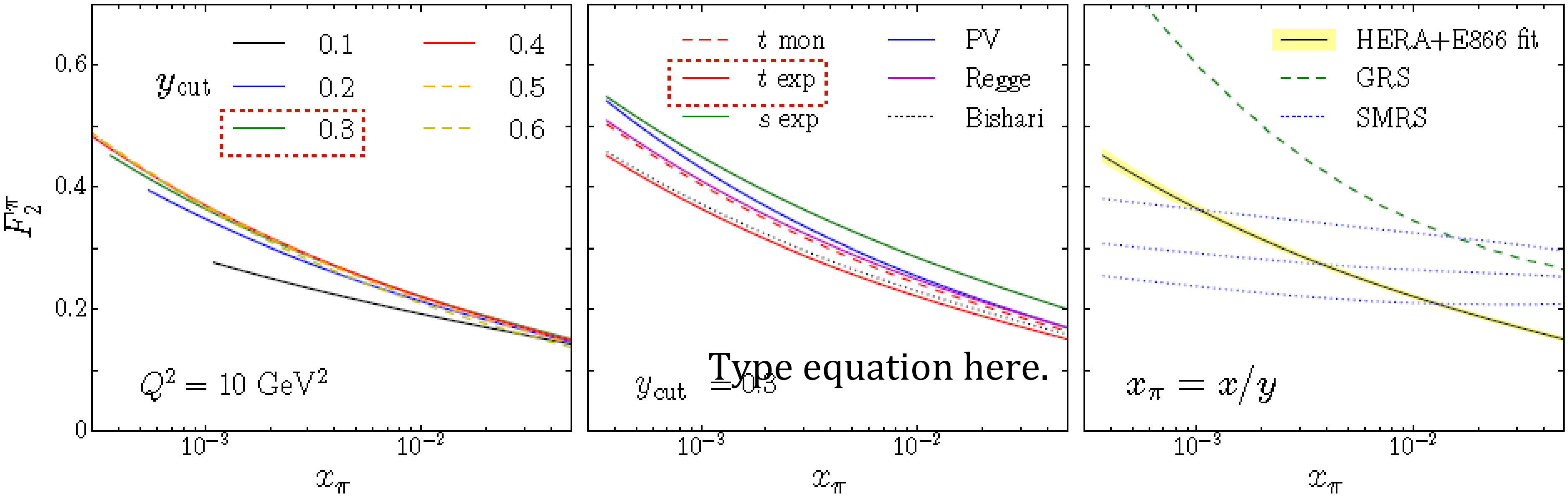
χ^2/dof=1.27 for 202 (187+15) points



***H1*, *EPJC* 68, 381 (2010)**

- ❑ Best fits for largest number of points by t-dependent exponential (and t-monopole) regulators

Extracted Pion Structure Function

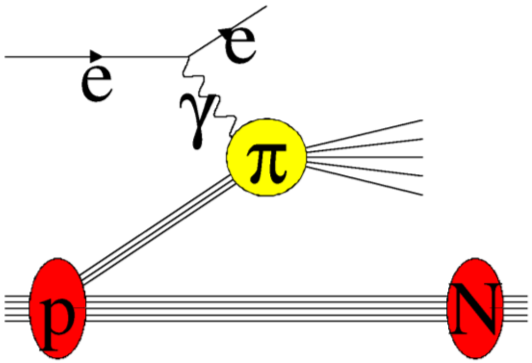


$$F_2^\pi = N\, x_\pi^a \left(1 - x_\pi\right)^b, \quad a = a_0 + a_1\eta$$

$$\eta \sim \log\left(\log Q^2\right)$$

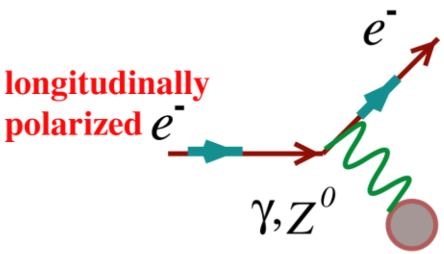
- Stable values of F_2^π at $4\mathrm{x}10^{-4}$ ~< x_π ~<0.03 from combined fit
- Shape similar to GRS fit to πN Drell-Yan data (for x_π >~0.2) but smaller magnitude

Electroweak Pion and Kaon Structure Functions



- The Sullivan Process will be sensitive to u and dbar for the pion, and likewise *u* and *sbar* for the kaon.
- Logarithmic scaling violations may give insight on the role of gluon pdfs
- Could we make further progress towards a flavour decomposition?

1) Using the Neutral-Current Parity-violating asymmetry A_{PV}



2) Determine xF_3 through neutral/charged-current interactions

$$F_2^{\gamma} = \sum_q e_q^2\, x\,(q+\bar{q})$$

In the parton model:

$$F_2^{\gamma Z} = 2\sum_q e_q\, g_V^q\, x\,(q+\bar{q})$$ *Use different couplings/weights*

$$x\,F_3^{\gamma Z} = 2\sum_q e_q\, g_A^q\, x\,(q-\bar{q})$$ *Use isovector response*

$$F_2^{W^+} = 2\,x\,(\bar{u}+d+s+\bar{c}) \quad F_3^{W^+} = 2\,(-\bar{u}+d+s-\bar{c}) \quad F_2^{W^-} = 2\,x\,(u+\bar{d}+\bar{s}+c) \quad F_3^{W^-} = 2\,(u-\bar{d}-\bar{s}+c)$$

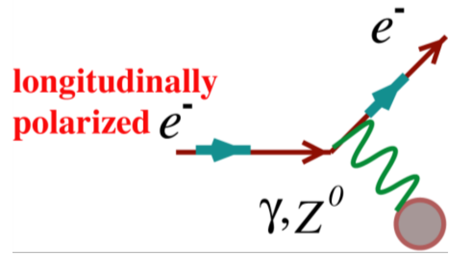
3) Or charged-current through comparison of electron versus positron interactions

$$A = \frac{\sigma_R^{CC,e^+} \pm \sigma_L^{CC,e^-}}{\sigma_R^{NC} + \sigma_L^{NC}} \qquad A = \frac{G_F^2\, Q^4}{32\,\pi^2\,\alpha_e^2}\left[\frac{F_2^{W^+} \pm F_2^{W^-}}{F_2^{\gamma}} - \frac{1-(1-y)^2}{1+(1-y)^2}\,\frac{xF_3^{W^+} \mp xF_3^{W^-}}{F_2^{\gamma}}\right]$$

Disentangling the Flavour-Dependence (I)

1) Using the Neutral-Current Parity-violating asymmetry A_{PV}

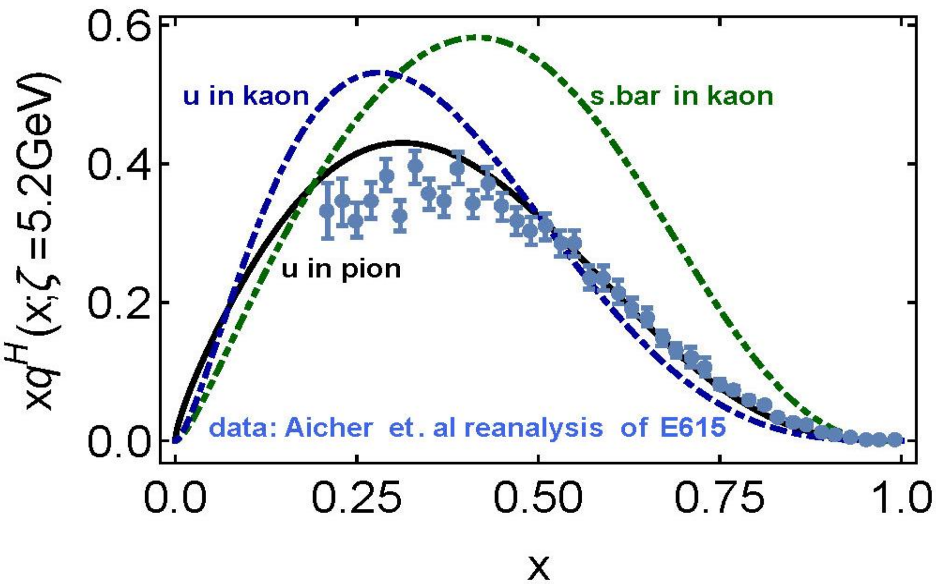
$$A_{PV} = \frac{\sigma_R - \sigma_L}{\sigma_R + \sigma_L}$$



$$a_{2\pi}(x) = \frac{2\sum_q e_q\, g_V^q\,(q+\bar{q})}{\sum_q e_q^2\,(q+\bar{q})} \simeq \frac{6\,u_\pi^+ + 3\,d_\pi^+}{4\,u_\pi^+ + d_\pi^+} - 4\sin^2\theta_W,$$

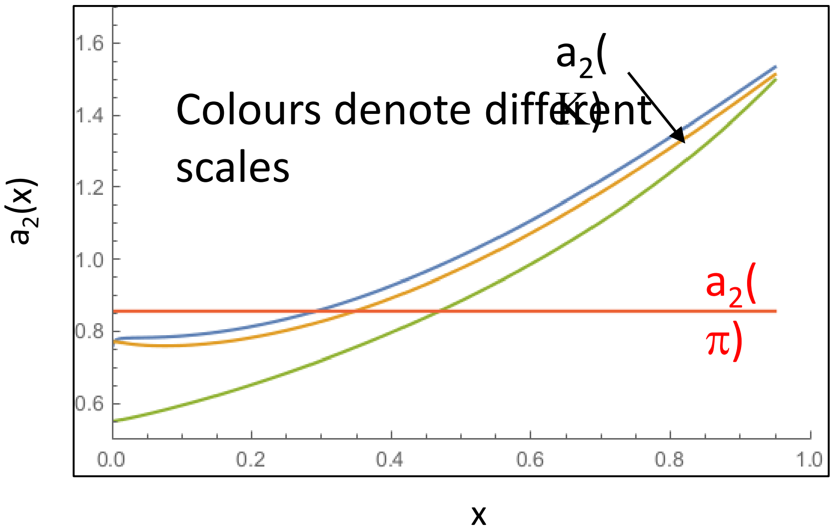
$$a_{2K}(x) = \frac{2\sum_q e_q\, g_V^q\,(q+\bar{q})}{\sum_q e_q^2\,(q+\bar{q})} \simeq \frac{6\,u_K^+ + 3\,s_K^+}{4\,u_K^+ + s_K^+} - 4\sin^2\theta_W.$$

C. Chen et al., Phys. Rev. D**93** (2016), 11



DSE-based parton distributions in pion and kaon

Calculation by C.D. Roberts et al.



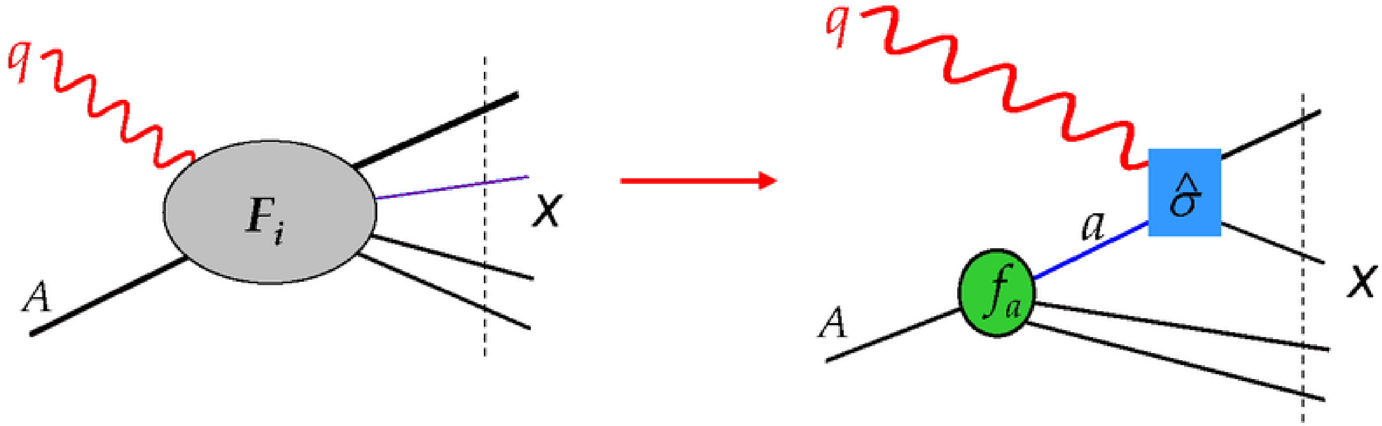
a_2 picks up different behaviour of u and sbar.
Flavour decomposition in kaon possible?

What are we missing?

- ✓ We discovered that (nearly) massless quarks and gluons make up the nucleon and that QCD governs their interactions.

- We had hoped to find out how quarks and gluons and their interactions give rise to the characteristics of the nucleons.
  - Spin
  - Mass
  - Bulk
- We also hoped that we would be able to find out how NN interactions work in terms of QCD.
  - How nuclear forces arise.
  - How nuclear characteristics come about
- We were able to do this kind of things with EM and atoms.
- *So what's going on…*

What longitudinal factorization did



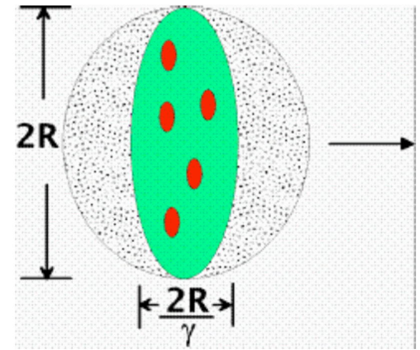
$$\lim_{Q^2 \to \text{large},\, x \text{ fixed}} F_i(x, Q^2) = f_a \otimes \hat{\sigma}$$

Function only of x (i.e. longitudinal momentum)
Our quarks and gluons as constituents of the proton only exist longitudinally.

Limits of Longitudinal Information

infinite momentum frame

What we know



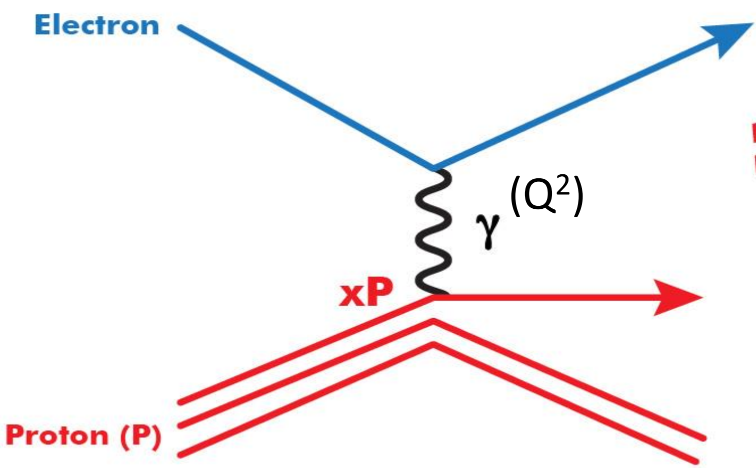
What is the quark and gluon structure of the proton?
-orbital motion?
-color charge distribution?
-how does the mass come about?
-origin of nucleon-nucleon interaction?

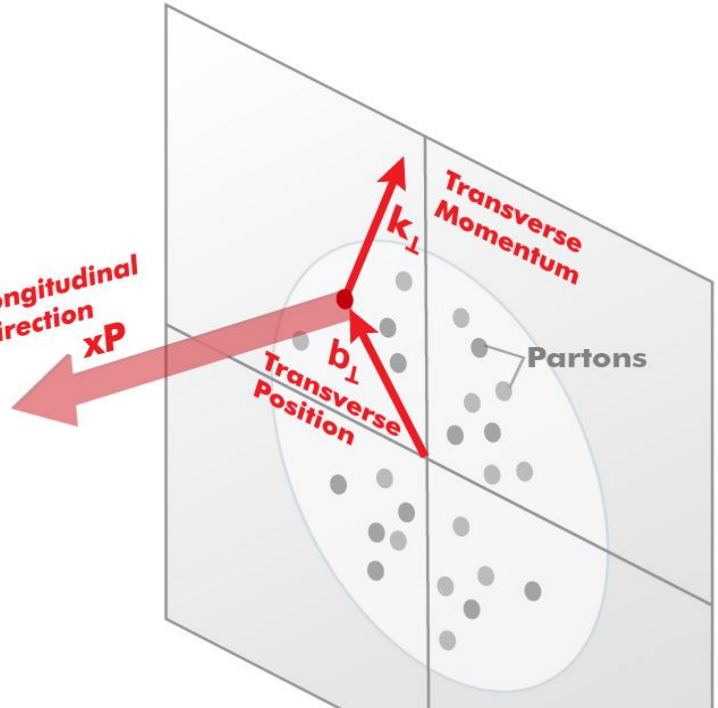
Parton frozen transversely. Framework does not incorporate any transverse information.

But this was the only way to define quark-gluon structure of proton in pQCD.

Progress in pQCD Theory (~1980-~2010)

Factorization II



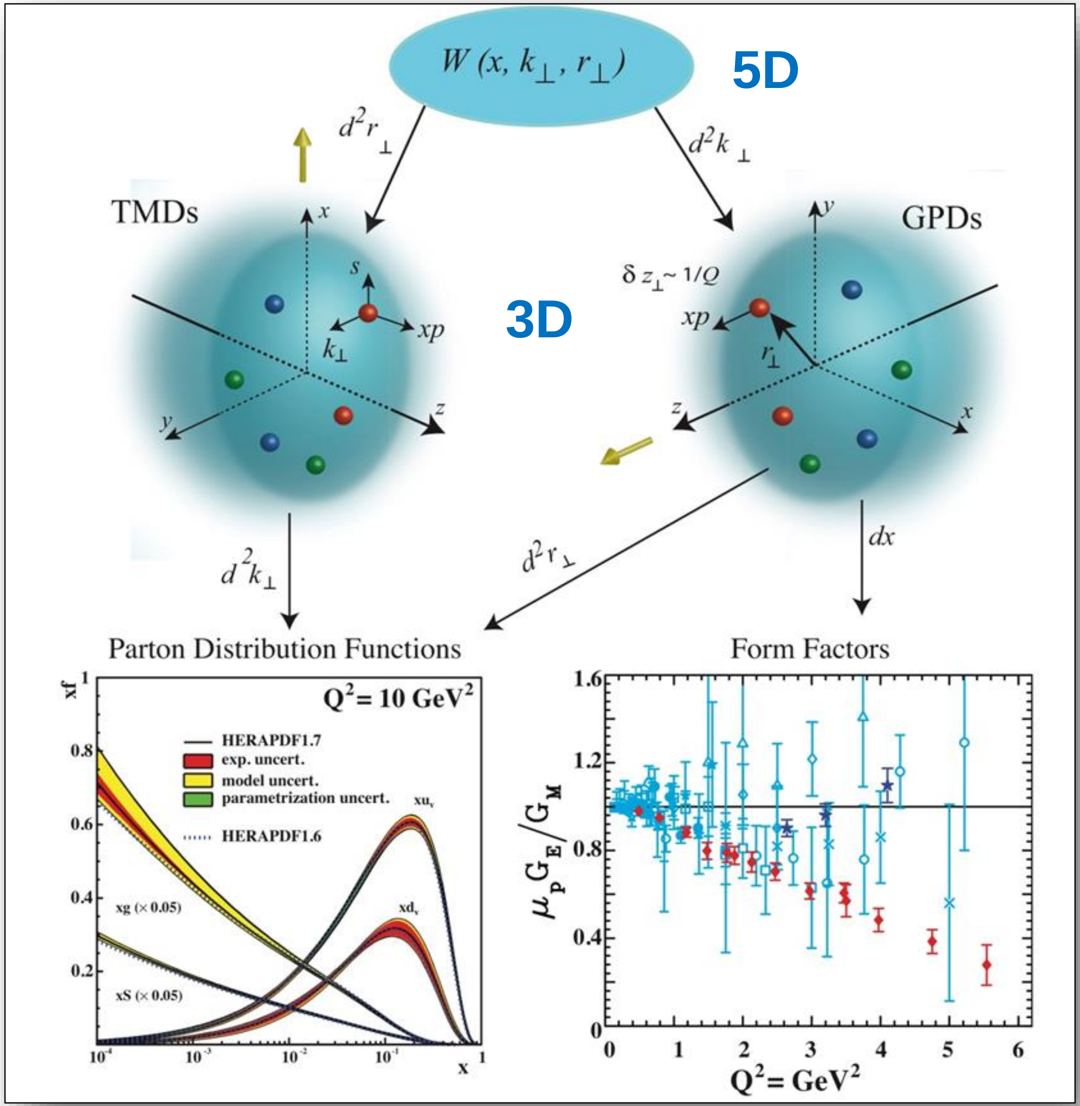


Parton Distribution Functions:
Longitudinal only—
No way to interpret nucleon partonic structure in rest frame

3D (Transverse) Structure
TMD's, GPD's—
Now we know what to measure to understand the 3D structure of nucleons

Transverse Momentum Dependent Distributions (TMD): k_t
Generalized Parton Distributions (GPD): b_t

New Paradigm for Nucleon Structure



- TMDs
  - Confined motion in a nucleon (semi-inclusive DIS)
- GPDs
  - Spatial imaging (exclusive DIS)
- Requires
  - High luminosity
  - Polarized beams and targets

3D Imaging of Quarks and Gluons

$W(x,b_T,k_T)$

Momentum space

$\int d^2b_T$

Coordinate space

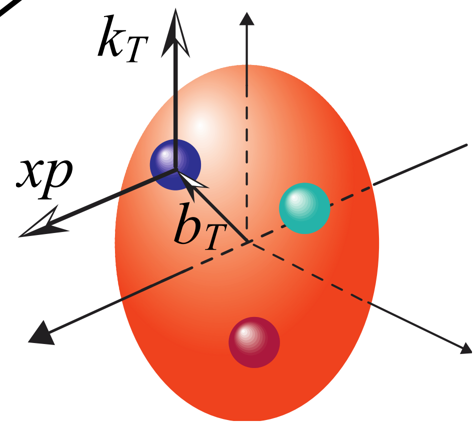
$\int d^2k_T$

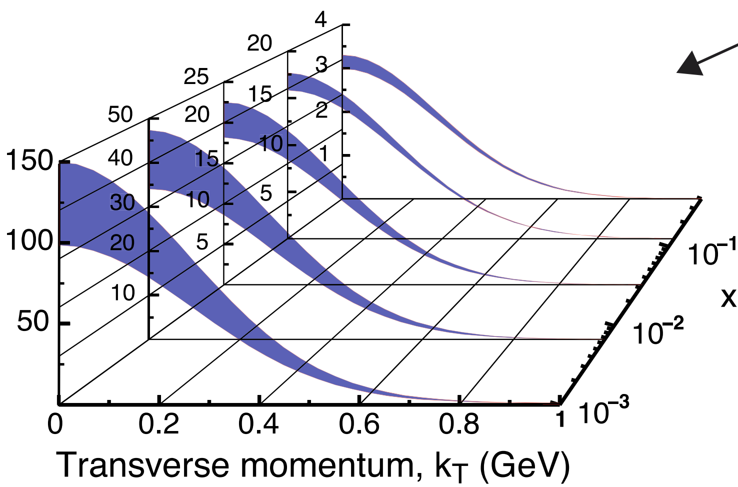
$f(x,k_T)$

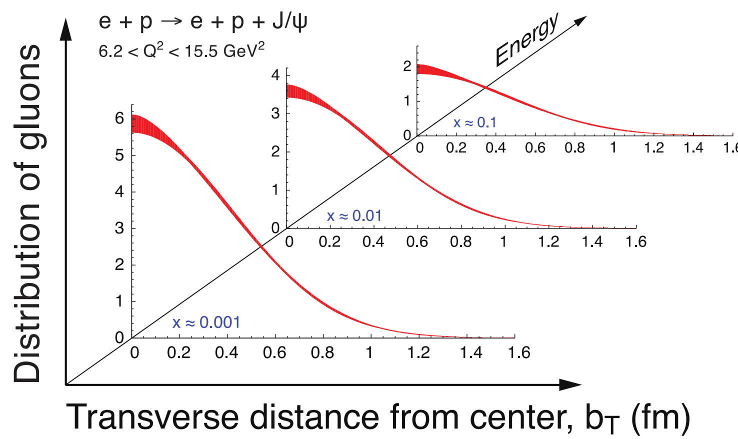
$f(x,b_T)$

Quarks

Gluons







Spin-dependent 3D momentum space images from semi-inclusive scattering

Spin-dependent 2D (transverse spatial) + 1D (longitudinal momentum) coordinate space images from exclusive scattering

How to extract PDFs from Data?

How to extract PDFs from data

- **Problem**:
  - we need a set of PDFs in order to calculate a particular hard-scattering process (say, at LHC)

- **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 (CJ10), MSTW2008, NNPDF2.1, ABM11, JR, HERAPDF1.5

Global PDF fits as a tool

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

How to extract PDFs from data

- Choice of 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

- 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* $W^{\pm}, \ Z^{0}, \ \text{DY lepton pairs}$
  - 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 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

Parameterizations

- 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.
- May have to make some arbitrary decisions on parameter values that are not well constrained by the data
- A smaller numbers of parameters is not always better - it is the description of the data that counts.

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

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

PDF uncertainties

- **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
    - ...

PDF uncertainties

Hessian method

- PDF parameters denoted by $\{a_\mu\}$, $\mu = 1, \ldots, d$
- As a byproduct of the fitting process, one obtains the Hessian $H_{\mu\nu}$

$$H_{\mu\nu} \equiv \frac{1}{2} \frac{\partial^2 \chi^2}{\partial a_\mu \, \partial a_\nu}$$

  which is evaluated at the minimum of χ^2.

- To estimate the error on some observable $X(a)$, taking into account only the experimental errors which entered into the calculation of χ^2 one uses the "Master Formula"

$$(\Delta X)^2 = T \sum_{\mu,\nu} \frac{\partial X}{\partial a_\mu} (H^{-1})_{\mu\nu} \frac{\partial X}{\partial a_\nu}$$

"tolerance"

PDF uncertainties

- **Tolerance $T = \Delta\chi$**
  - Open a textbook, T=$\Delta\chi$=1 means 67% confidence level
  - But Hessian method works only if
    - all data sets are statistically compatible
    - Exp. errors are Gaussian…
    - …and have not been underestimated (*e.g.*, by neglect of a source of systematics)
  - Correct this by a larger tolerance factor so that most data (90%, 67% of them) fall inside the PDF error band
    - CTEQ6.1 used T=10, MRST used T=5
    - Nowadays a bit more refined procedure are adopted

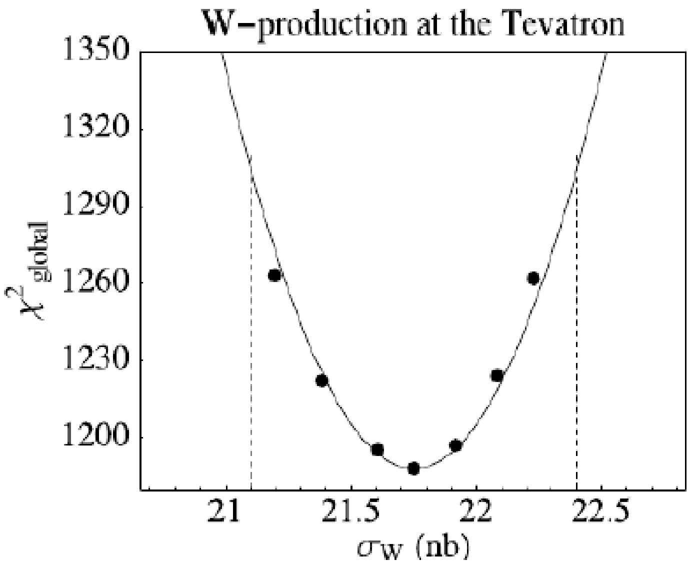
PDF uncertainties

Lagrange multipliers method

- Given an observable X, minimize a new function for fixed values of Lagrange multiplier λ

$$\Psi(\lambda, A) = \chi^2_{\rm g}(A) + \lambda(X(A) - X_0)$$

- Obtain a new set of parameters, A_{min}, and the pair $\{\chi^2_{\rm g}(\lambda),\, X(\lambda)\}$
- Repeating for many variables, one obtains $\chi^2_{\rm g}(X)$
- Chose a tolerance, read off the PDF error ΔX



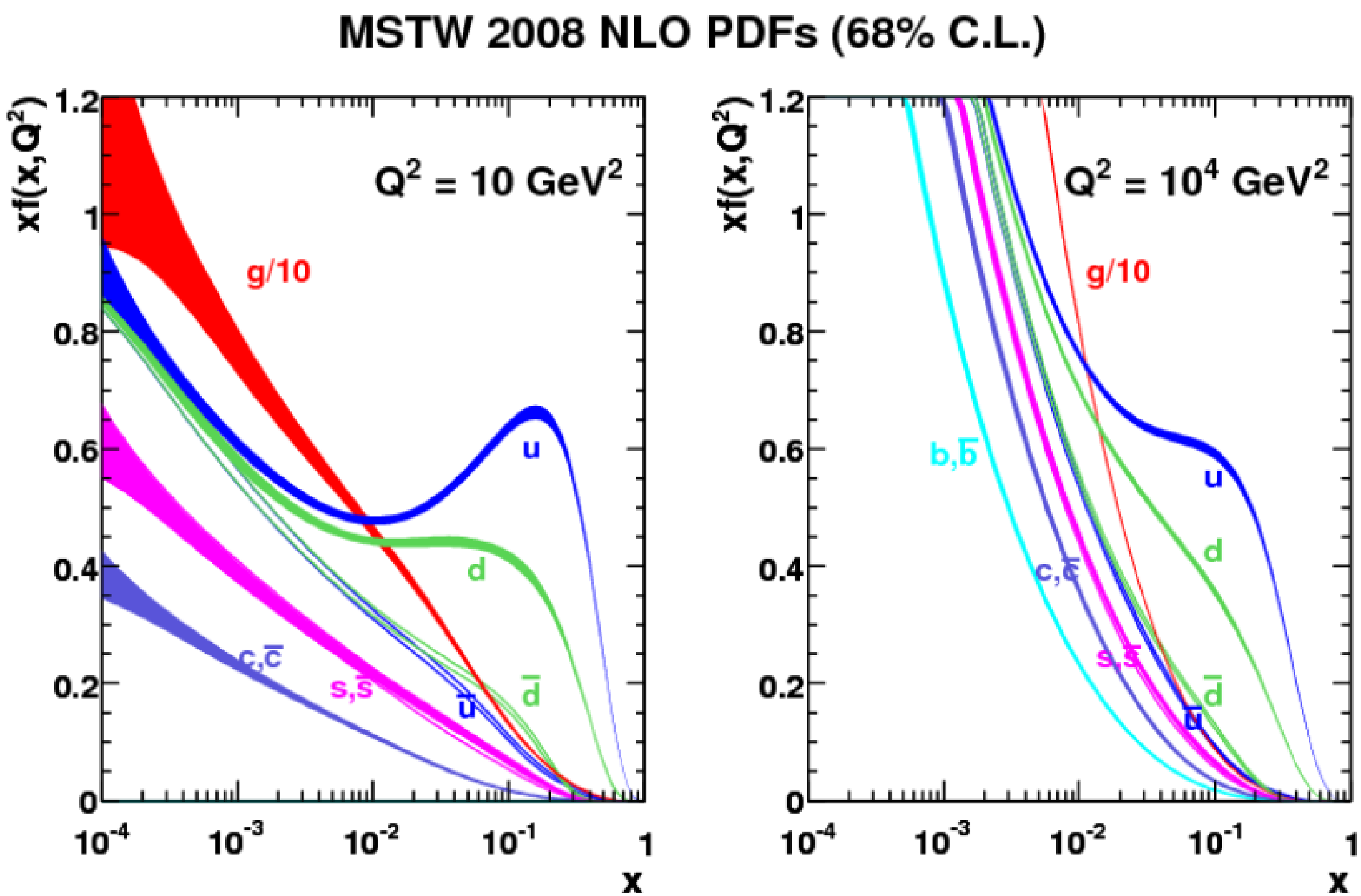
PDF uncertainties

Monte-Carlo method

- Generate many replicas of the chosen data set
- In each replica, randomize central data point within quoted errors
- Make a fit for each replica
- Obtain PDF errors from statistical analysis of all fit results

- This is adopted by the NNPDF collaboration, but is not limited to neural network based fits

Examples



MSTW 2008 NLO PDFs (68% C.L.)
xf(x,Q²)
Q² = 10 GeV²
g/10
u
d
d̄
ū
s,s̄
c,c̄
x
Q² = 10⁴ GeV²
b,b̄

Impact of new data, eic

- **Questions**
  - What are the requirements in terms of energy, luminosity?
  - What physics do we expect to learn?
  - "Is it worthwhile building that accelerator?"

- **For example:**
  - Is a DIS cross section measurement at the EIC going to improve the PDF measurements?

- **This we can anwer with a global fit:**
  - Generate pseudo-data
  - Include them in a global fit
  - Compare with old result