Recent Developments on the CT14 Global Analysis of Q.C.D.

Daniel Stump

Department of Physics and Astronomy
Michigan State University

The CTEQ-TEA working group
Sayipjamal Dulat, Jon Pumplin, Carl Schmidt, Daniel Stump, C.-P. Yuan: *Michigan State* ;
Joey Huston: *Michigan State* ;
Tie-Jiun Hou, Pavel Nadolsky: *SMU* ;
Jun Gao: *Argonne* ; Marco Guzzi: *Manchester* ;

Part 1 "CTEQ-TEA PDFs and HERA I+II Combined Data"

T.-J. Hou, S. Dulat, *et al,* *[paper is in preparation]*

History: the CT10 (2010) PDFs and CT14 (2015) PDFs

CT10 (2010)	CT14 (2015)
HERA I combined data other short-distance processes (CERN, Fermilab, Tevatron) no LHC	HERA I combined data other short-distance processes (CERN, Fermilab, Tevatron) updated LHC inclusive jet production W and Z production

A new global analysis ≡ $CT14_{HERA2}$

- Make these changes w.r.t. CT14:
  - replace HERA I (N_{pts} = 579) by HERA I+II (N_{pts} = 1120) ★
  - delete NMC $F_{2p}(x,Q)$ (N_{pts} = 201)
  - replace prelim. CMS inclusive jet data by the up-dated table
  - add one more parameter to the strange quark PDF, $s(x,Q_0)$

Compare our results to:

- HERA : H. Abramowicz *et al*, EurPhyJ C75, 580 (2015)
- MMHT : L. A. Harland Lang et al, EurPhyJ C76, 186 (2016)
- NNPDF : J. Rojo, hep-ph 1508.07731 (2015)

Notations

- from 36 experiments we have

  D_i = central data values (i = 1 … N)

  σ_i = s.d. of uncorrelated errors (")

  β_{ji} = s.d. of correlated systematic errors

  ($j = 1 \ldots N_{sy}$; $i = 1 \ldots N$)

- from NNLO (or NLO QCD) we have

  T_i = theory value

  $= T_i (\{\alpha_\nu ; \nu = 1 \ldots 28\})$

- fit theory and data by χ^2 minimization,

$$\chi^2_{\text{global}} (\{\alpha\}) = \sum_{\text{expt}} \{ \chi^2_{\text{expt}} \}$$

$$\chi^2_{\text{expt}} = \min_{\{r_j\}} \left[\sum_i (D_i - \sum_j r_j \beta_{ji} - T_i)^2 / \sigma_i^2 + \sum_j r_j^2 \right] = \chi^2_{\text{red}} + R^2$$

PDF parameters

treating systematic errors as nuisance parameters

Comparing PDF results (CT14 and CT14$_{HERA2}$) to data (HERA1 and HERA2)

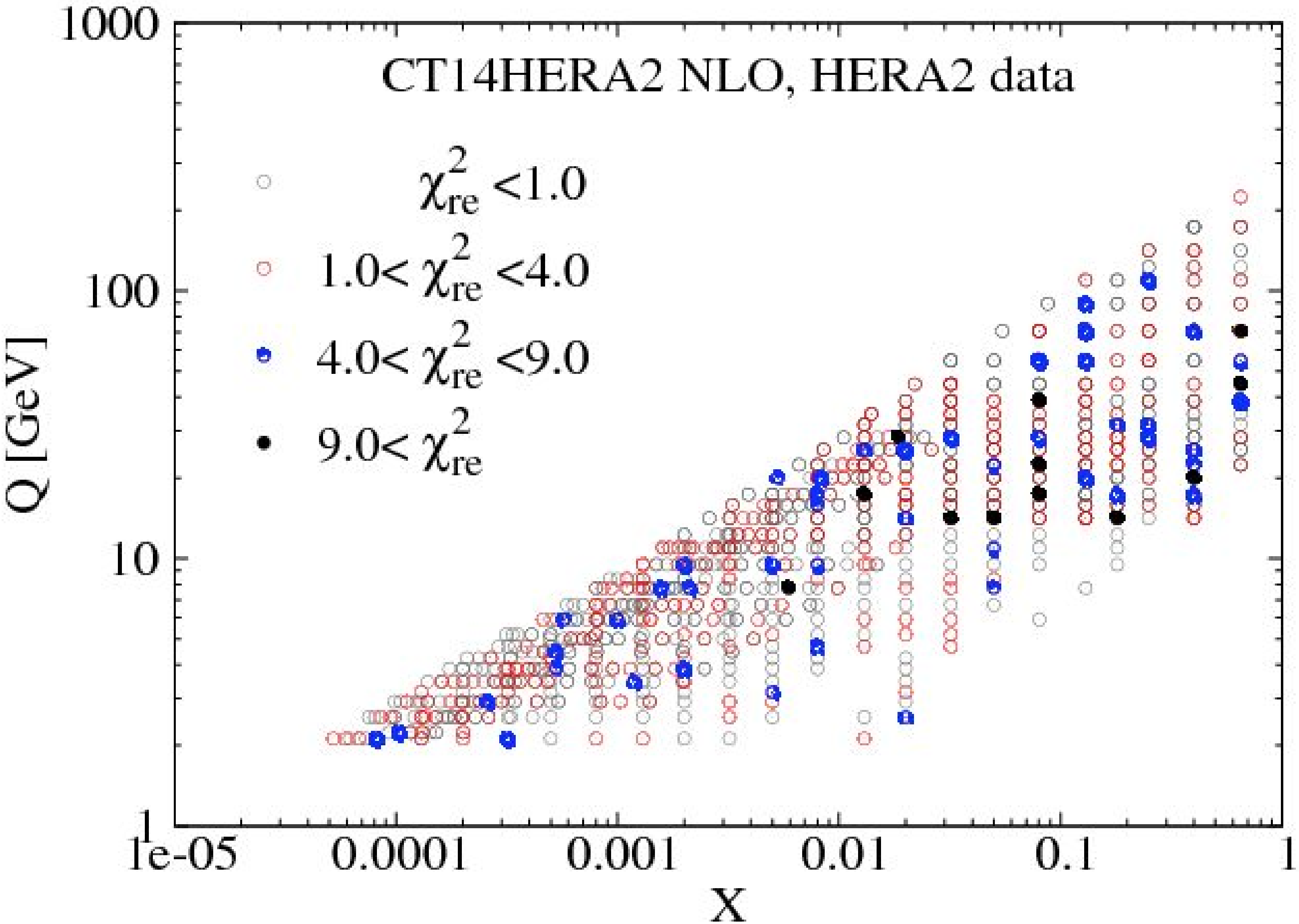
[HERA2 means the HERA I + II combined data (1120 points)]

PDFs	χ^2_{HERA1} $/N_1$	χ^2_{HERA2} $/N_2$	χ^2_{HERA2} $/N_2$
CT14 (NNLO)	591 /579 (fit)	1469 /1120 (not fit)	= 1.31
CT14$_{HERA2}$(NNLO)	610 /579 (not fit)	1402 /1120 (fit)	= 1.25

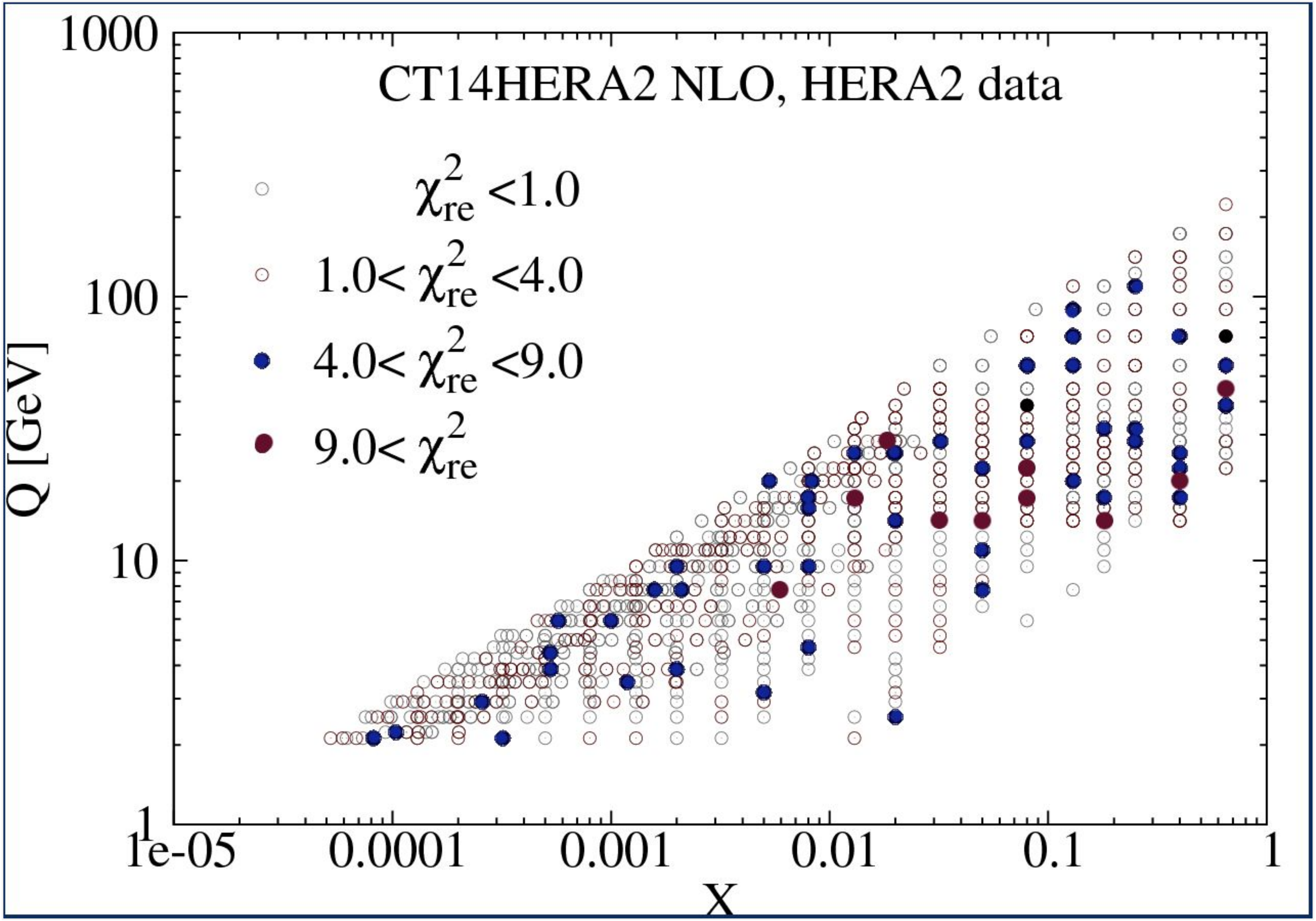
$[\, \chi^2/N \,]_{HERA2}$ is large
even when HERA2 is included in the global fit.

Why?

Reduced χ^2 's (for single data points) in the xQ plane



Reduced χ^2 's (for single data points) in the x-Q plane



Separate the four HERA2 DIS processes;
(Q_{cut} = 2 GeV)

	N_{pts}	$\chi^2_{red.}$ / N_{pts}
NC e^+p	880	1.11
CC e^+p	39	1.10
NC e^-p	159	1.45
CC e^-p	42	1.52
totals		
[reduced χ^2] /N	1120	1.17
χ^2 / N	1120	1.25
R^2 / N	1120	0.08

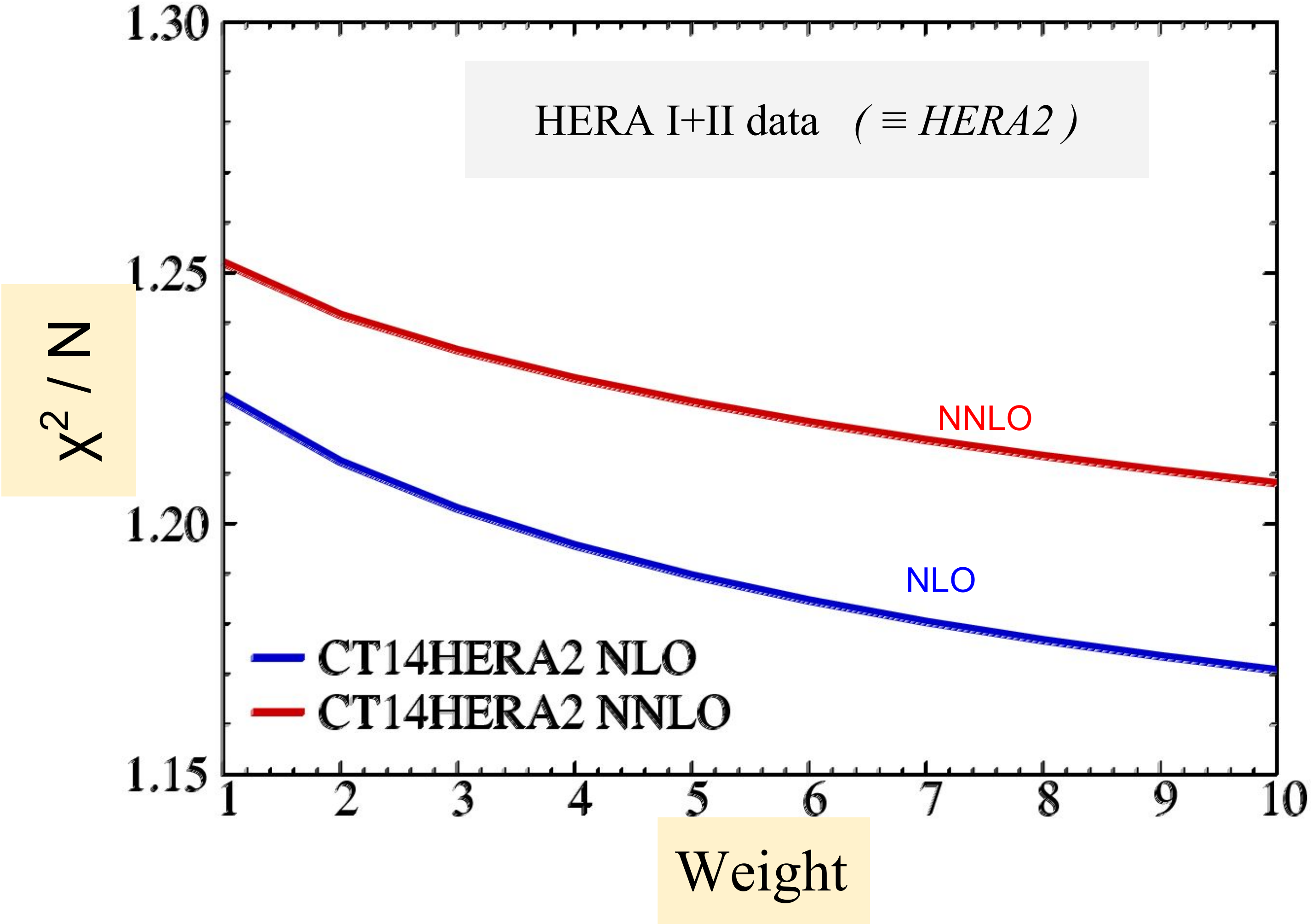
reduced χ^2 values

χ^2 = [reduced χ^2] + R^2

The quadratic penalty for 162 systematic errors = 87.5

We also studied the impact of different Q^2 kinematic cuts.

χ^2_{HERA2} /N_2 versus the ***weight*** assigned to the *HERA2* data in χ^2_{GLOBAL}



CT14$_{HERA2}$ PDFs compared to CT14

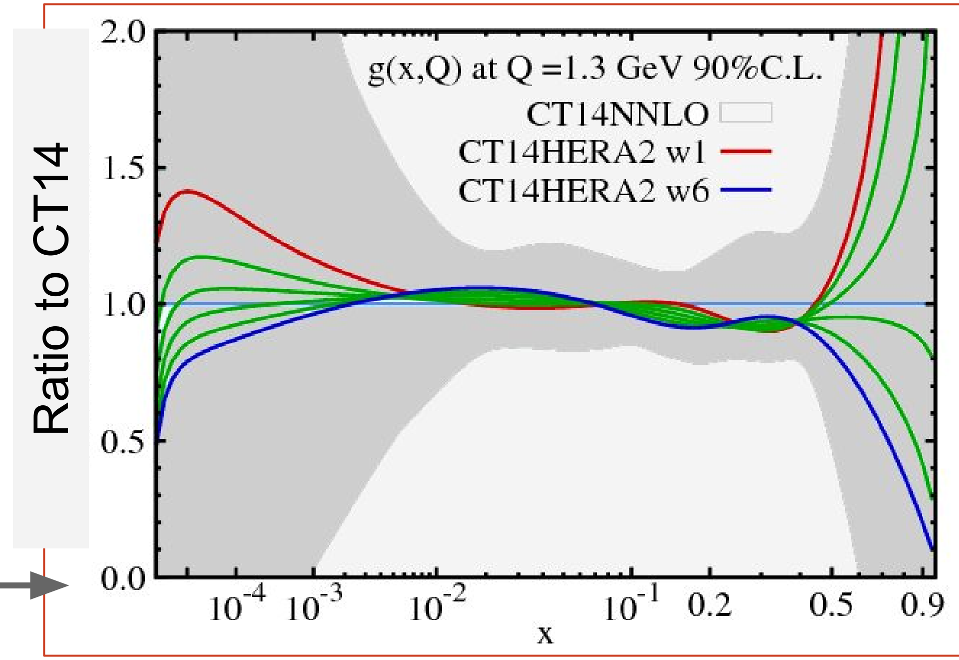
CT14HERA2

- Ratio to the standard CT14 PDF;
- six choices of **weight** applied to the HERA2 data set in the global fit (*nominal=1* to *heaviest=6*)

- CT14 Hessian error band (shaded)

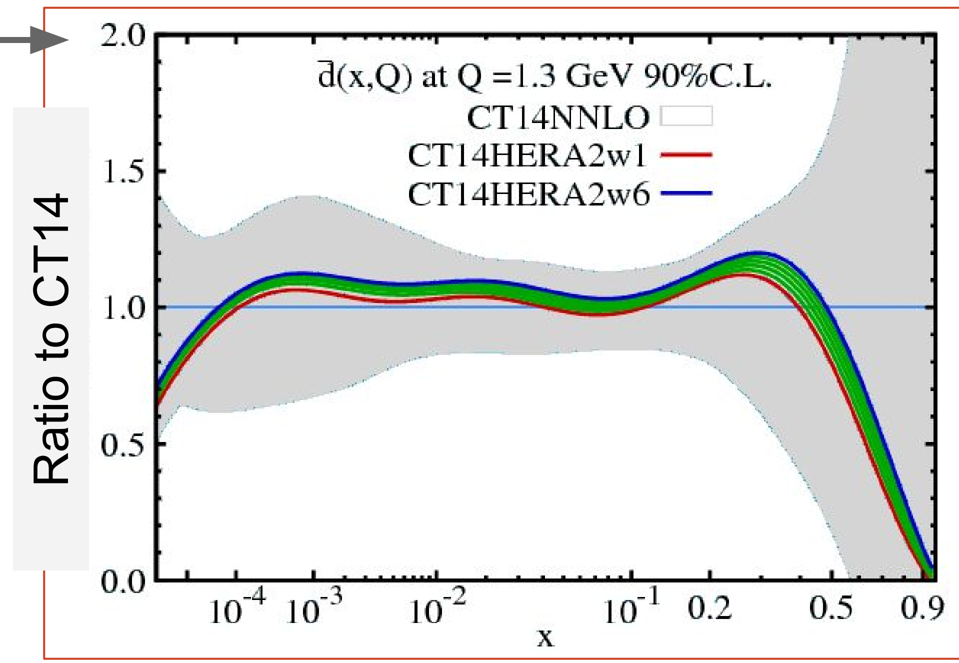
Comparing PDFs

gluon



d antiquark

CT14HERA2 is slightly larger but always in the error band



Impact of the *HERA2* data:

- ➢ skews the gluon pdf *vs.* x;
- ➢ pushes the d-antiquark up *vs.* x

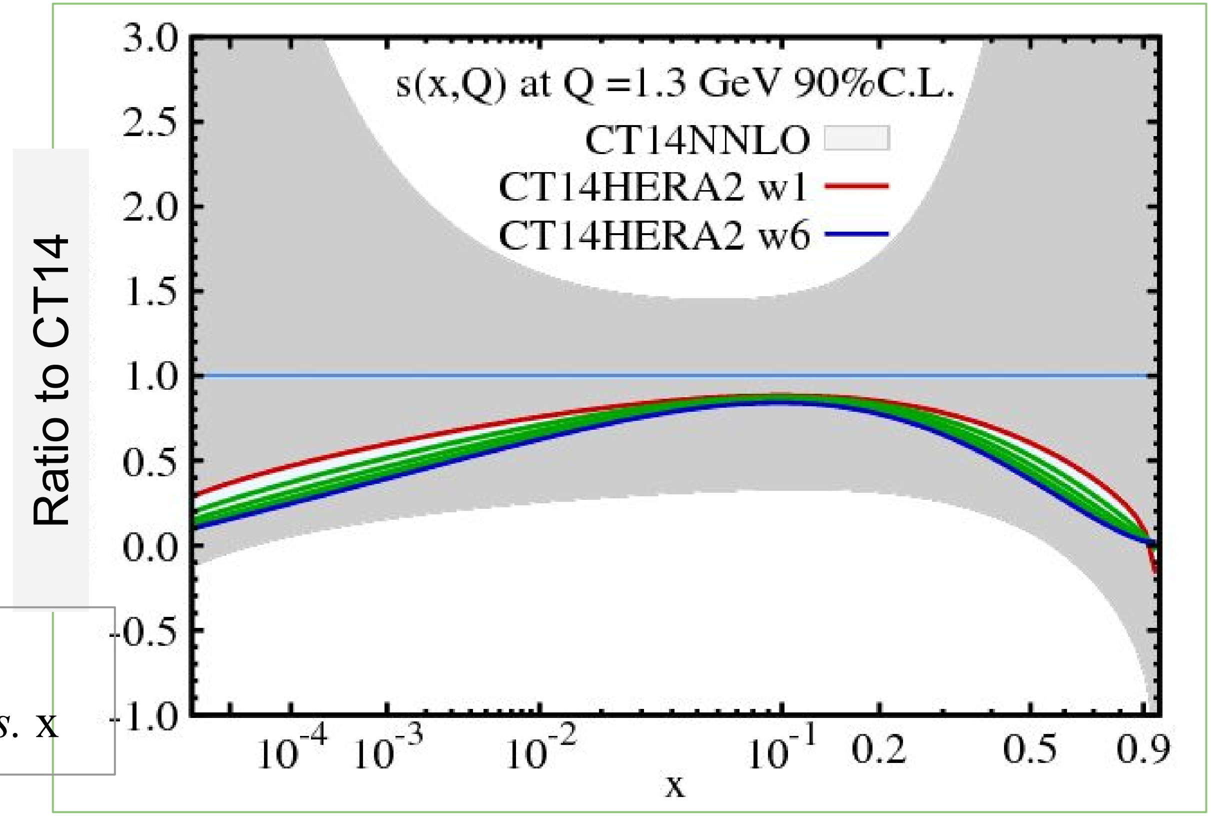
CT14HERA2

- Ratio to the standard CT14 PDF;
- six choices of **weight** applied to the HERA2 data set in the global fit (*nominal=1* to *heaviest=6*)

- CT14 Hessian error band (shaded)

Comparing PDFs

s quark

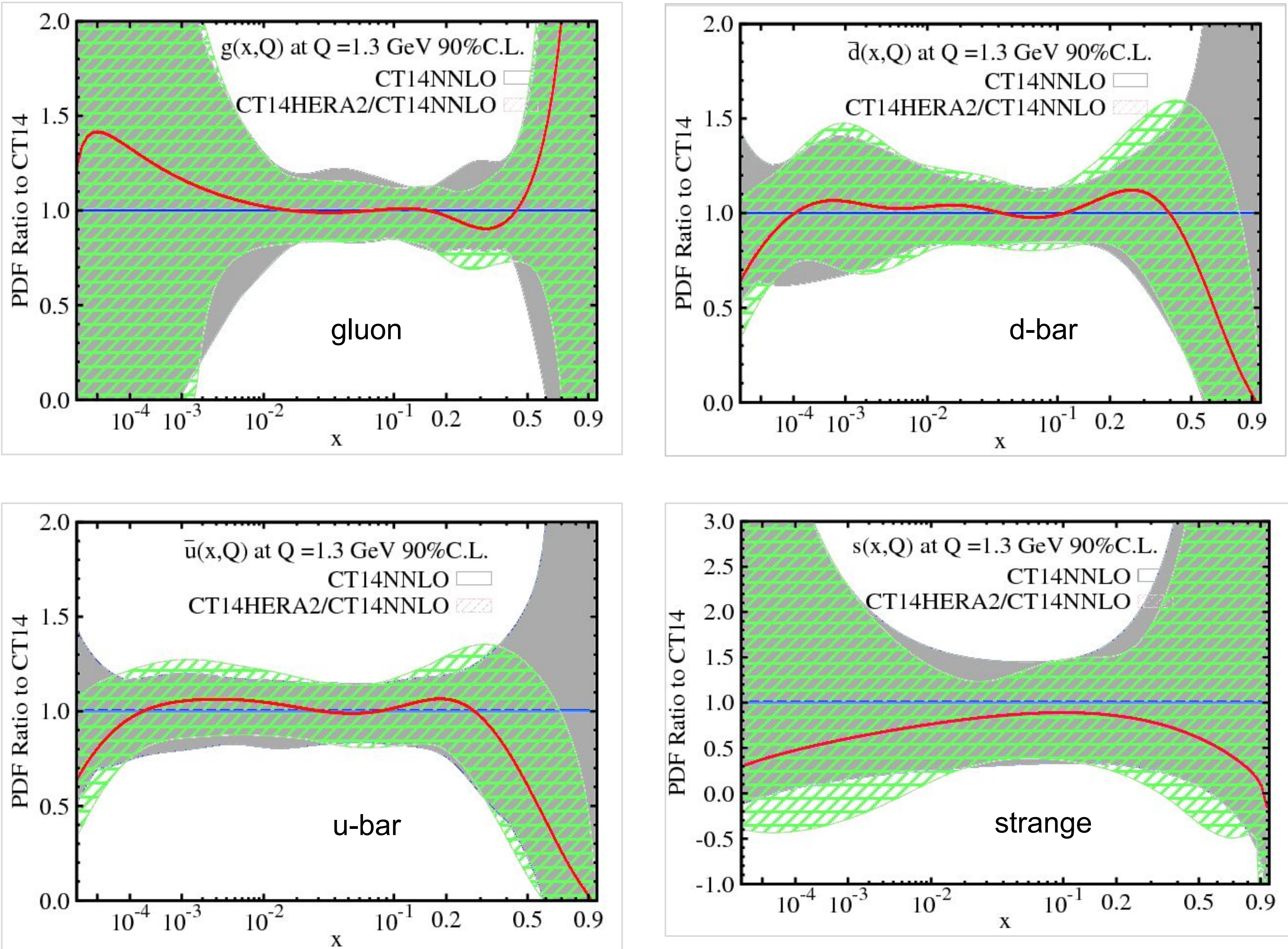


Impact of the *HERA2* data:

➢ pushes the s-quark down *vs.* x

Comparing $CT14_{HERA2}$ and CT14 ; plotting ratios f_{HERA2} / f ; $CT14_{HERA2}$ and CT14 error bands

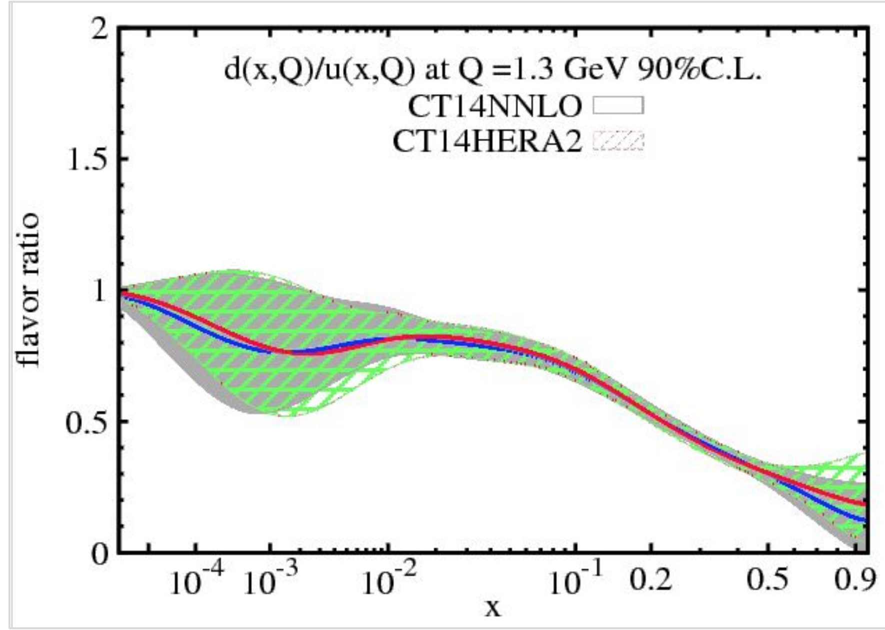
Comparing PDF uncertainty bands

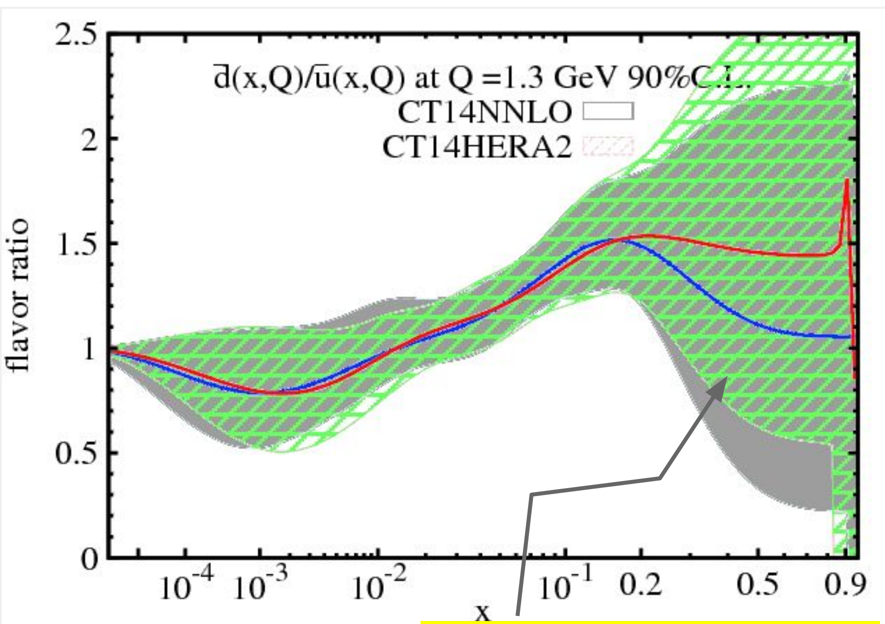


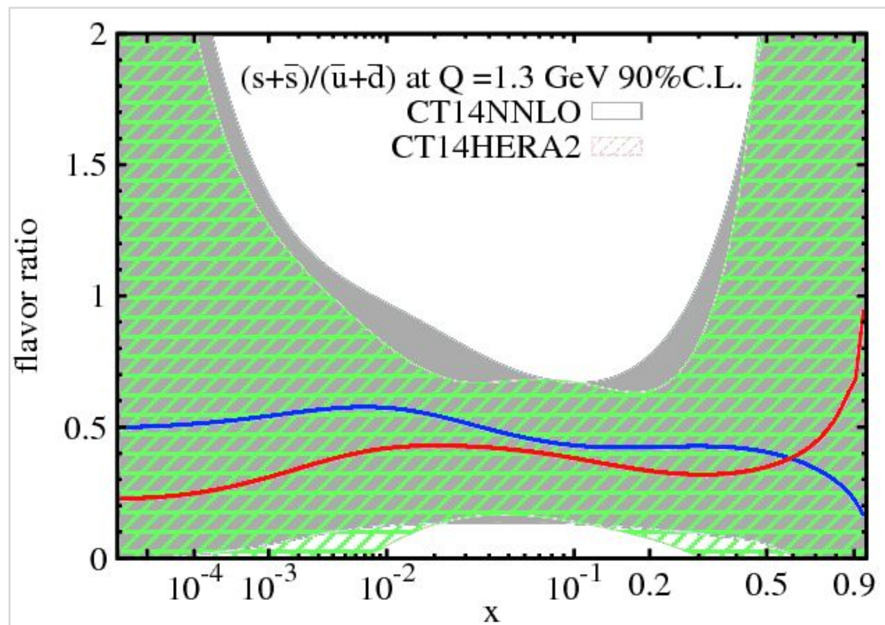
Flavor ratios at Q = 1.3 GeV

Blue = CT14 Red = $CT14_{HERA2}$

	d/u
db/ub	(s+sb)/(db+ub)







A curious result?
dbar/ubar > 1 at large x

Comparing cross sections

W$^{\pm}$ and **Z**$^{\mathbf{0}}$ production at the LHC

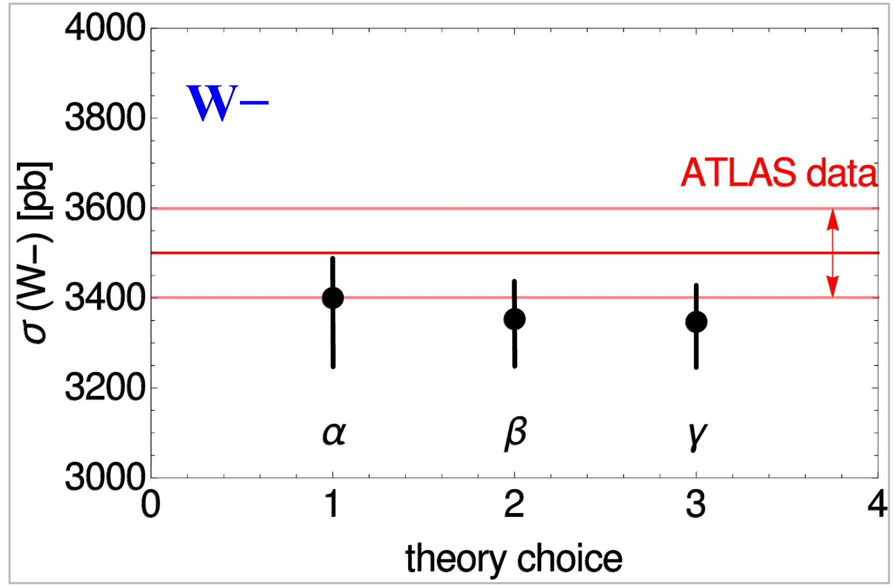
ATLAS fiducial cross section

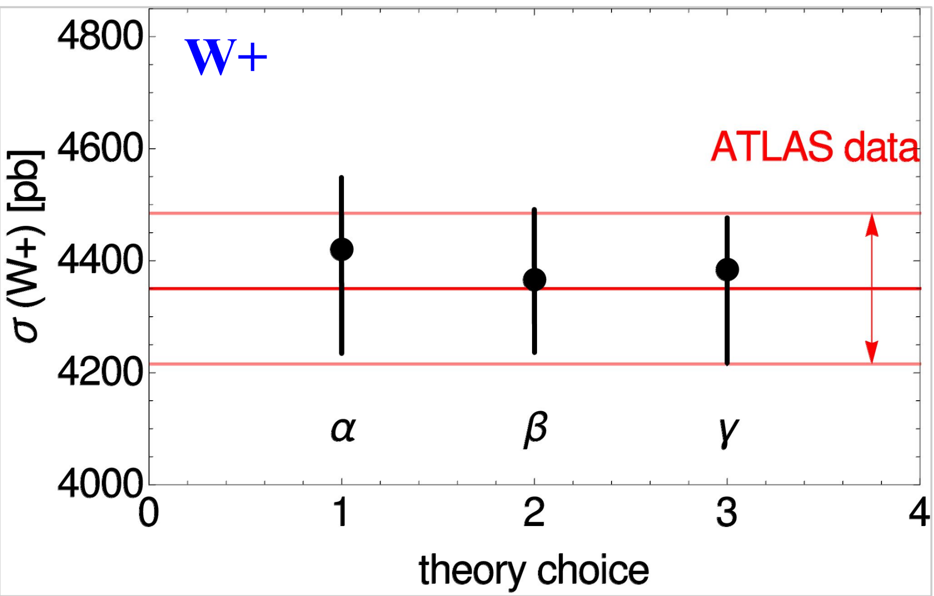
(∃ back-up slide on the CMS cross sections)

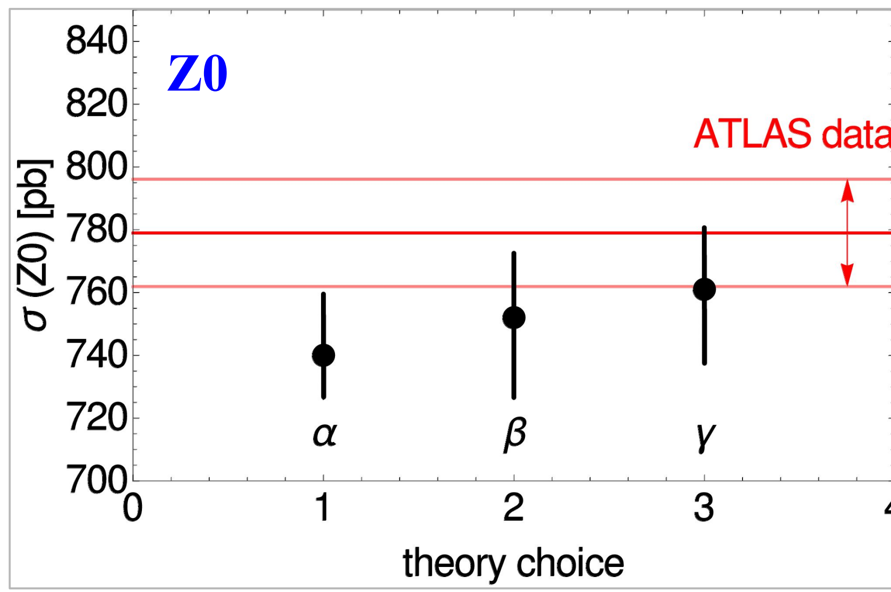
Theory calculations:

α = ATLAS calculation (DYNNLO)

β = CT14 ; γ = CT14$_{\text{HERA2}}$ (ResBos)







Part 1: Final conclusions

- There are some interesting but small changes in the PDFs, in going from CT14 to $\mathrm{CT14}_{\mathrm{HERA2}}$, *esp.* $\bar{u}$, $\bar{d}$, and s ;
- the changes are smaller than the current PDF uncertainties ;
- so we still recommend CT14 as the preferred PDFs for LHC Run 2 ;
- availability of $\mathrm{CT14}_{\mathrm{HERA2}}$.

Part 1: das Ende

Part 2

"Reconstruction of Monte Carlo Replicas from Hessian parton distributions"; Tie-Jiun Hou, P. Nadolsky, et al; arXiv:1607.06066 [hep-ph]

A Quick Review of the Hessian method

Parton DFs $f_v(x, Q_0) = F_v(x, \{\alpha\})$ parametrization $\{\alpha_i\,;\, i = 1 \ldots D\}$.

<u>*Figure of Merit*</u> $\chi^2(a) \approx \chi^2(0) + \sum_{ij=1}^{D} H_{ij}\, a_i\, a_j$ *(a=displacement from min)*

number of eigenvectors = D;
separate the + and − directions.

Result : 1 "central set" of PDFs and 2×D "error sets";
the LHAPDF format.

The prediction for an observable $X(f)$ is

$$\text{prediction} = X_{\text{central}}\ {}^{+\,\delta X_{\text{up}}}_{-\,\delta X_{\text{dn}}}$$

(possible asymmetric errors; contradicts the Gaussian hypothesis)

where

$$\delta X_{up} = \left\{ \sum_{i=1}^{D} [max(X_{+i} - X_0, X_{-i} - X_0, 0)]^2 \right\}^{1/2}$$

$$\delta X_{dn} = \left\{ \sum_{i=1}^{D} [max(X_0 - X_{+i}, X_0 - X_{-i}, 0)]^2 \right\}^{1/2}$$

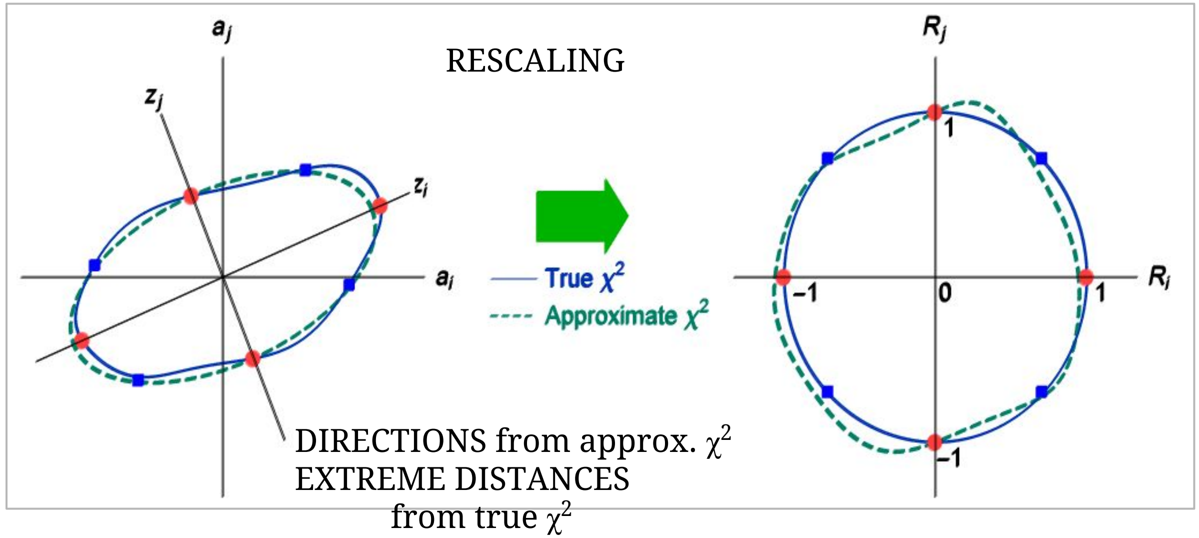
B **"Replicas"** Now generate 1,000 sets of PDFs, stochastically

$$\{\ f_v^{(k)}\ (x, Q_0)\ ;\ k = 1, 2, 3, \ldots, 1000\ \}$$

$\rightarrow$ $F_v\ (x, \{\alpha\}_k)$ where $\{\alpha\}_k$ is a random variate in D dimensions.

That's the basic idea,

but there are some developments …



- ❑ rescale from $\{a_1 \ldots a_D\}$ to $\{r_1 \ldots r_D\}$;

- ❑ $dP = (2\pi)^{-D/2} \exp[-½\, \mathbf{r} \cdot \mathbf{r}]\, d^D r$;
- ❑ Deal with the possibility that the Gaussian hypothesis is not valid; *e.g.*, what about the asymmetric errors?
- ❑ ***Ultimate goal*** : the *mean* and *standard deviation* of an ensemble of *X(f)*-values calculated with the replicas, should agree with the *central value* and *uncertainty* calculated with the (1 + 2D) Hessian PDFs.

I need to skip over some subtleties, for lack of time.

Hou, Nadolsky, *et al*, arXiv:1607.06066 [hep-ph]

Also, our results should be compared to

G. Watt and R. S. Thorne, JHEP **08**, 052 (2012); arXiv:1205.4024

We use the same basic method,
but with some different computational details:
"shift mean to best fit" , "asymmetry" ,
"positivity" , "Taylor series displacements"

$\Rightarrow$ Results …
(do replica results agree with Hessian?)

Hessian PDFs and Replica PDFs (linear method) and Replica PDFs (log method)

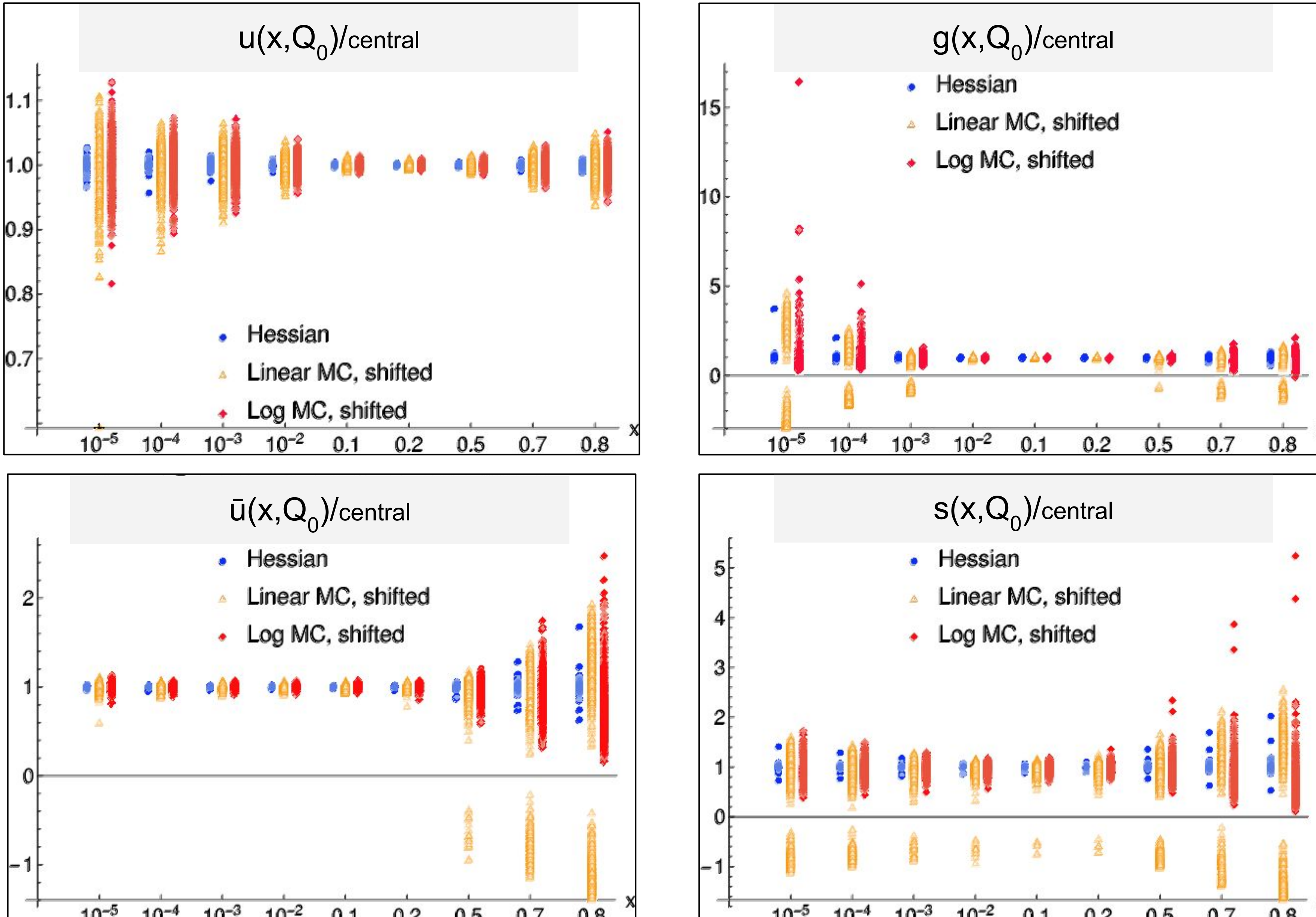
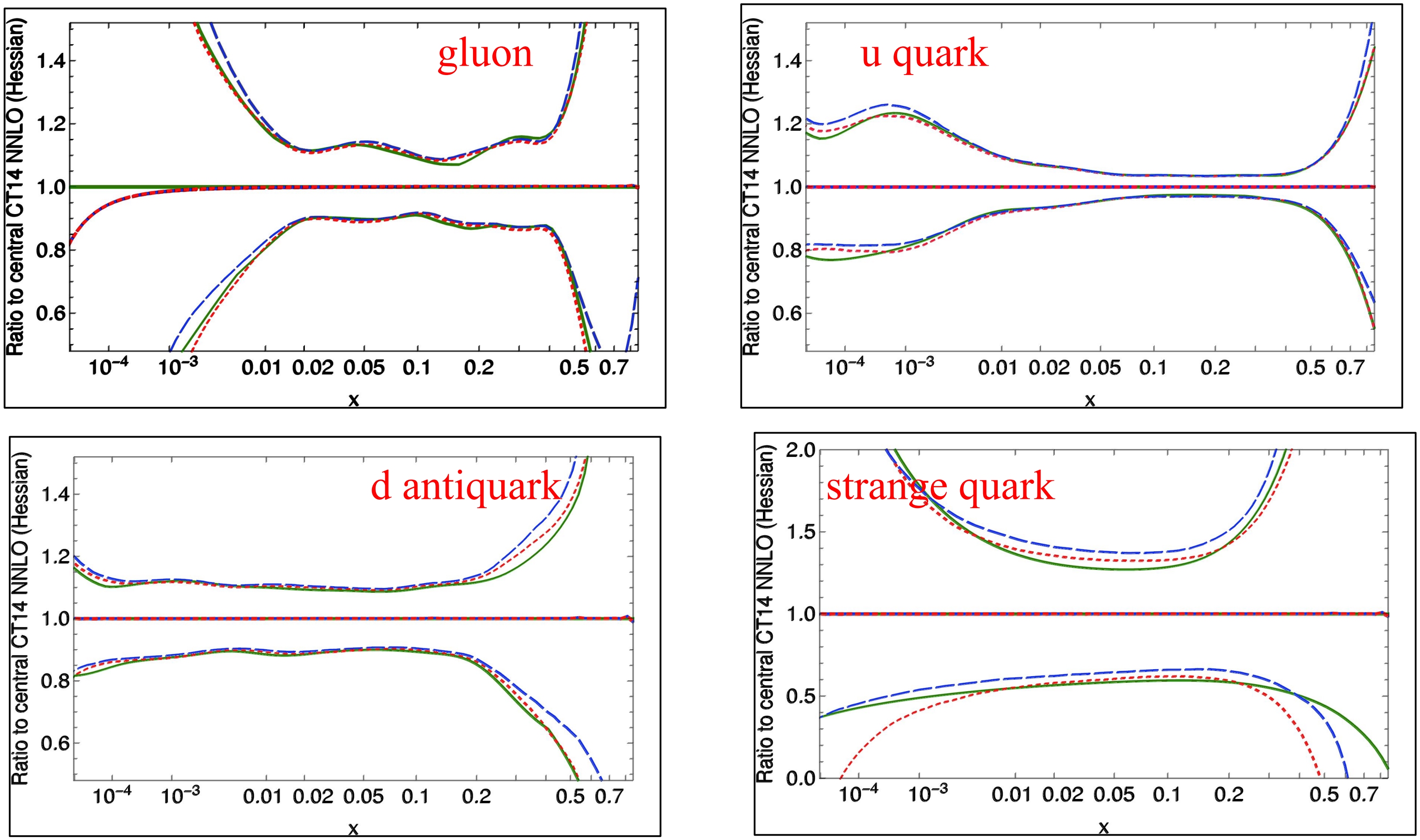


fig 2

Comparing PDF uncertainties;
i.e., repl.mean and SD ***versus*** Hessian

CT14 NNLO ***ASYMMETRIC uncertainties***
solid=Hessian ; dotted = MC1 ; dashed = MC2

fig 3



So indeed the S.D. of replicas is approximately equal to the Hessian uncertainty.
MC1 = linear MC (sampling f) ; MC2 = log MC (sampling $\ln|f|$)

Comparing PDF uncertainties;
i.e., repl.mean and SD ***versus*** Hessian

CT14 NNLO ***SYMMETRIC uncertainties***
solid=Hessian ; dotted = MC1 ; dashed = MC2

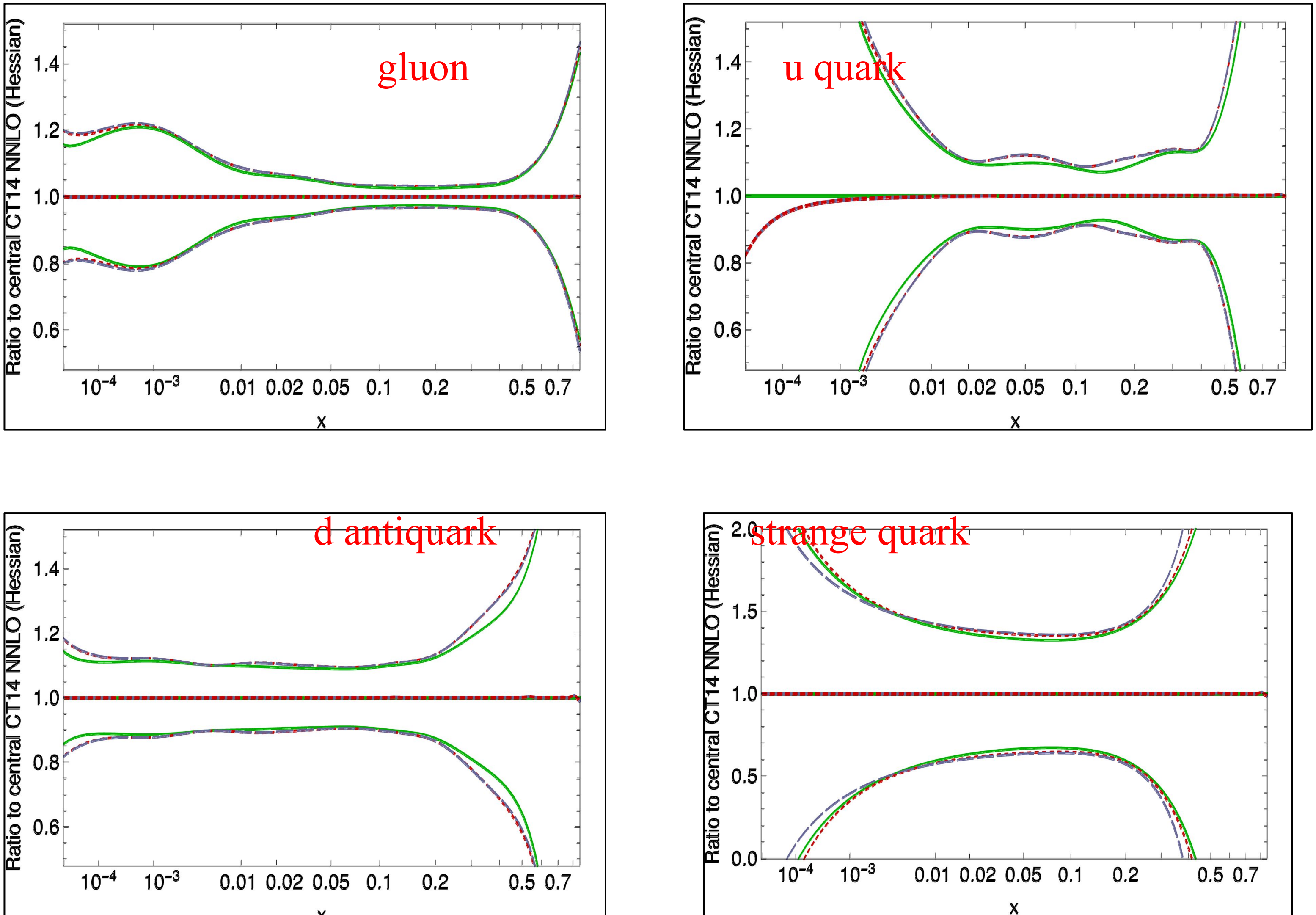


fig 4

Luminosity Functions :

$$L_{ab}(s, M^2, y_{cut}) = \frac{1}{1+\delta_{ab}}\left[\int_{\frac{M}{\sqrt{s}}e^{-y_{cut}}}^{\frac{M}{\sqrt{s}}e^{y_{cut}}} \frac{d\xi}{\xi}\, f_a(\xi, M)\, f_b\left(\frac{M}{\xi\sqrt{s}}, M\right) + (a \leftrightarrow b)\right]$$

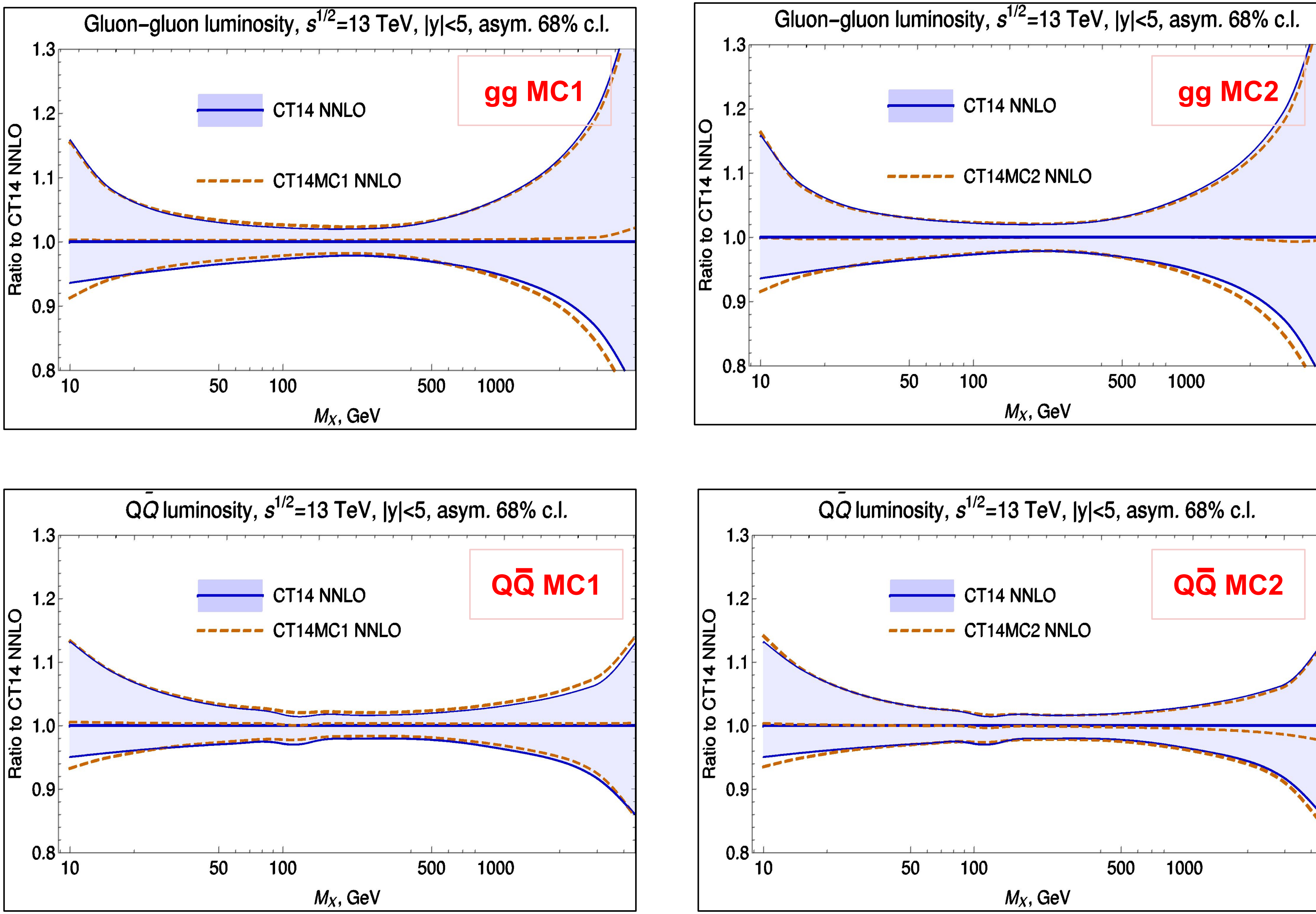
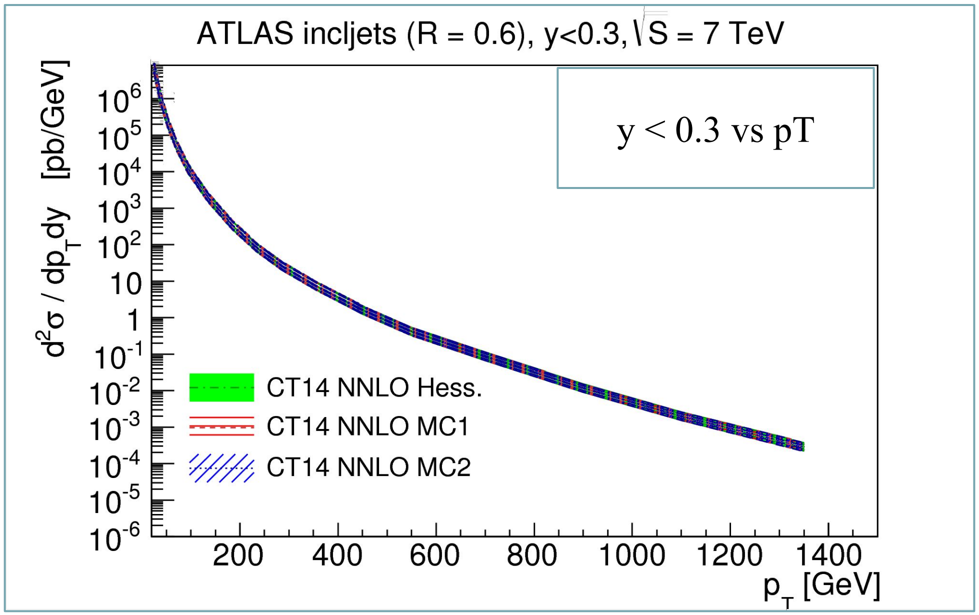


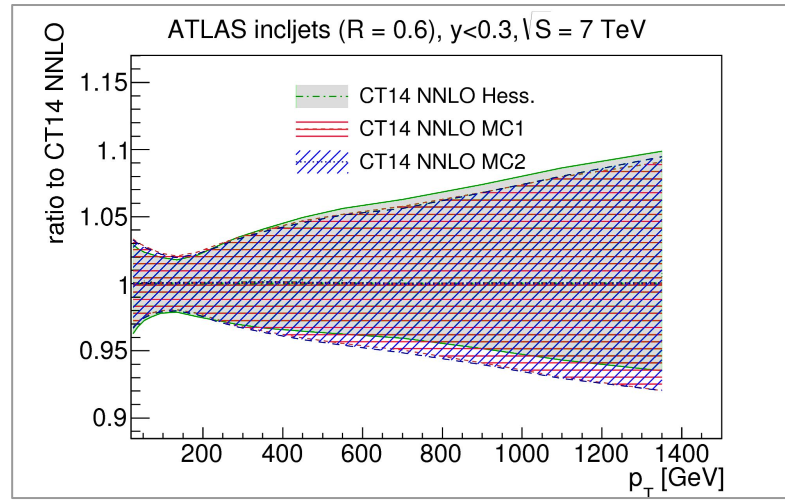
fig 7

ATLAS inclusive jet product'n @ 7 TeV

An example of a cross section calculation, comparing

- best fit with Hessian uncertainties
- mean and standard dev. of replicas
  - MC1 and MC2



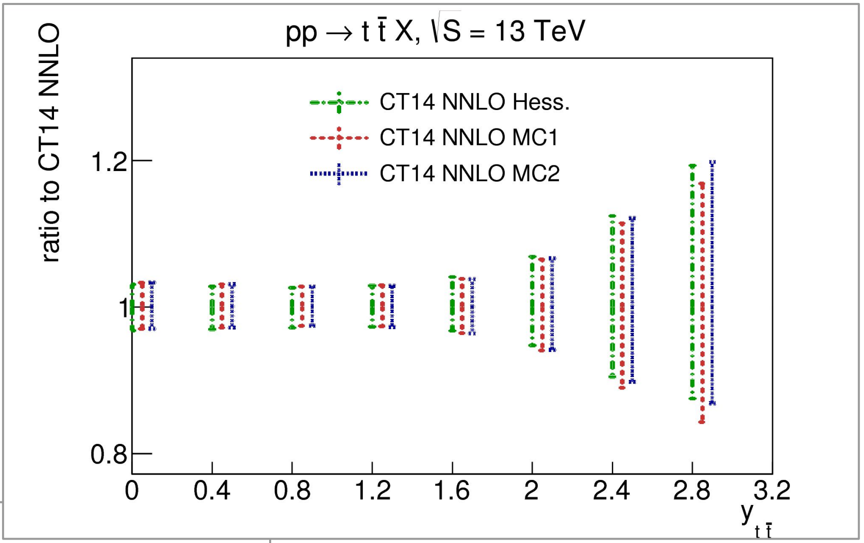


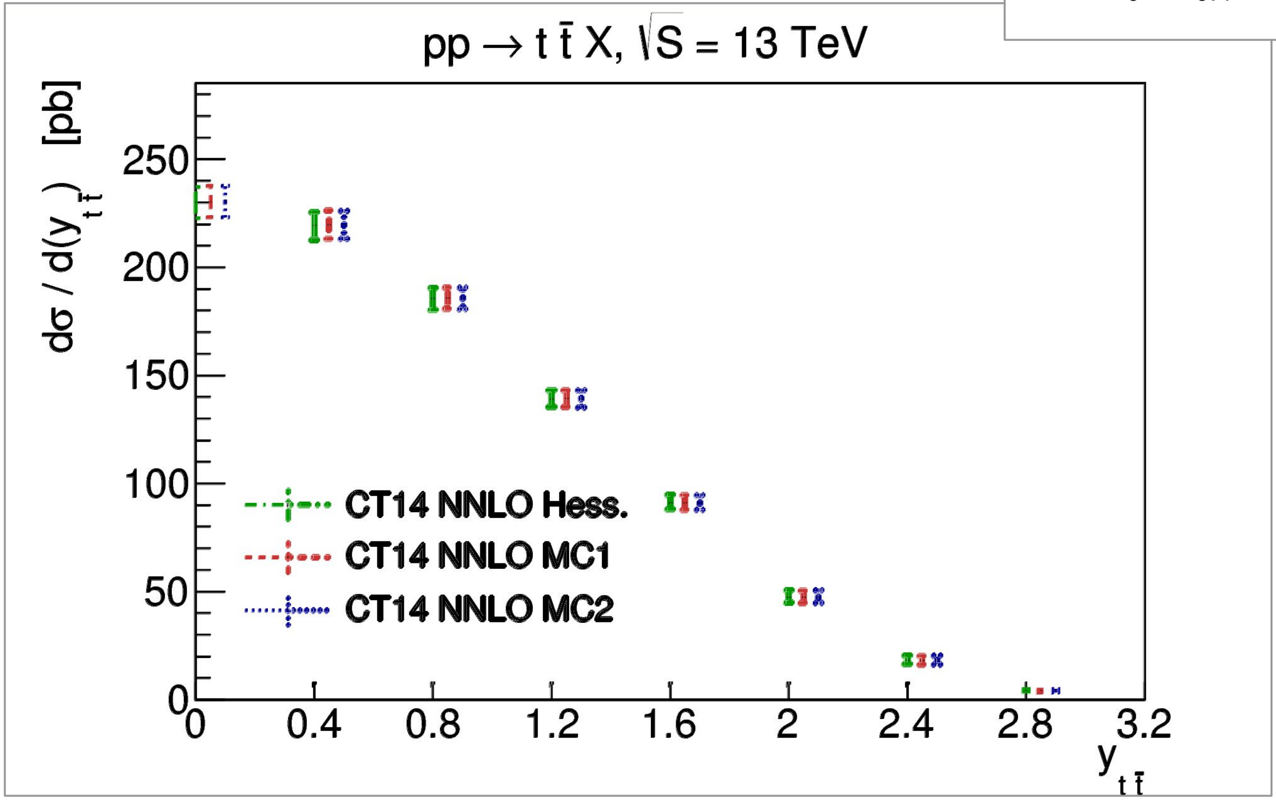
- The ***replica results*** closely approximate the ***Hessian results.***

<u>Inclusive top-antitop ($t\bar{t}$) production</u>

An example of a cross section calculation, comparing

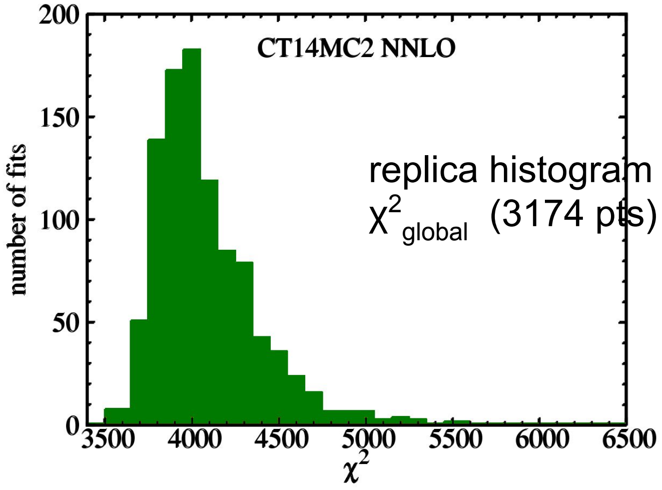
- best fit with Hessian uncertainties
- mean and standard dev. of replicas
  - MC1 and MC2

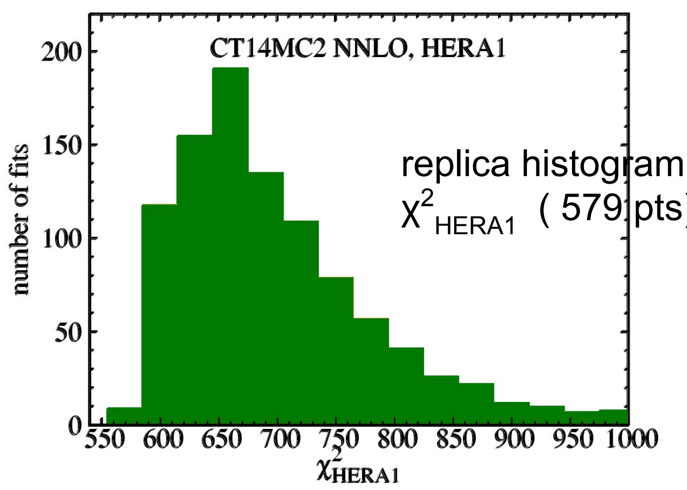




- The ***replica results*** closely approximate the ***Hessian results.***

<u>Only a large ensemble of MC replicas is meaningful.</u>





Most replicas are poor fits to the data; but the mean & SD do agree with the Hessian uncertainties.

- ❑ *What can we do with "replicas" ?*

replica re-weighting, unweighting, compression, meta PDFs, ...

- ❑ *Availability of the CT14 MC PDFs*

  http://hep.pa.msu.edu/cteq/public/
  http://lhapdf.hepforge.org
  http://metapdf.hepforge.org/mcgen

Part 2 : das Ende