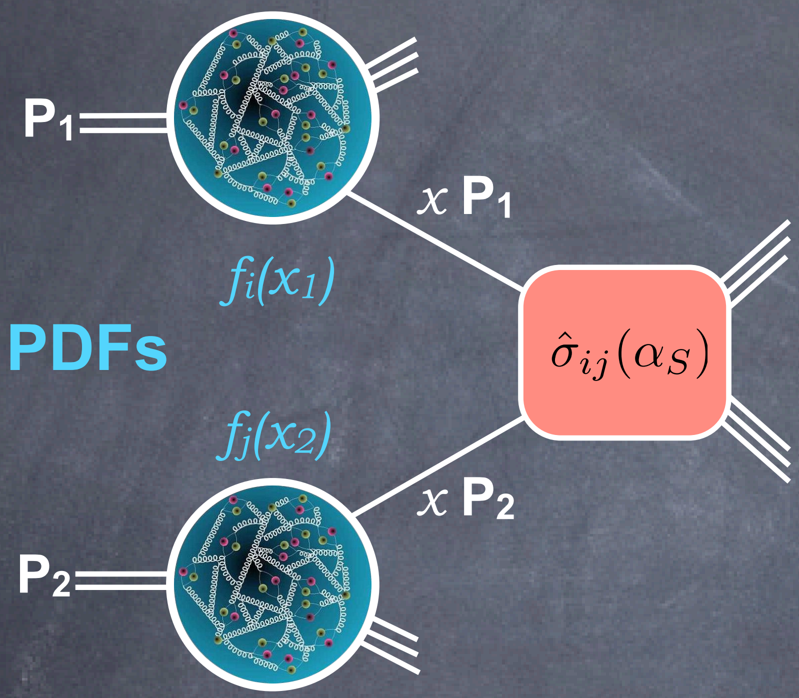
PROTON STRUCTURE IN THE LHC ERA: Impact of CMS Measurements on Parton Distribution Functions

Katerina Lipka
on behalf of the CMS Experiment

ICPPA, Moscow 2017

PARTICLE PRODUCTION IN *PP* COLLISIONS

proton structure **hard interaction**



Partons: quarks & gluons

Q^2 : typical energy scale in the process

x : partonic fraction of the proton momentum

Rate = (structure of 2 protons) $\otimes$ σ_{ij}

Parton Distribution Functions

$f_i\,(Q^2, x)$

provided by theory | **determined experimentally**

at the very edge of theory and experiment, correlated with fundamental QCD parameters

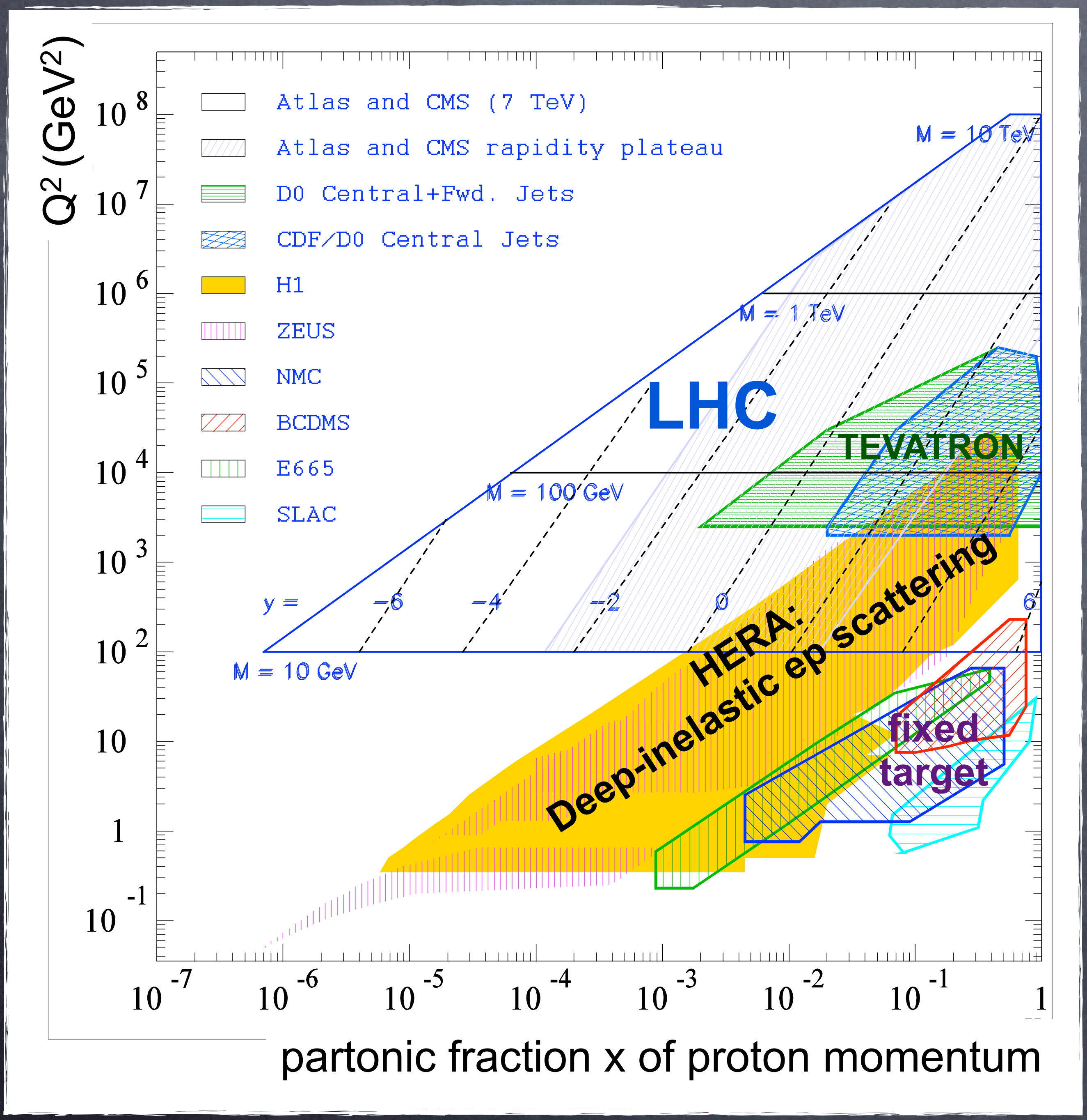
Improvement of PDFs precision demands theory & experiment collaboration and implies a variety of measurements and theory calculations

PDF DETERMINATION IN A QCD ANALYSIS

Parton Distribution Functions

$f_i(Q^2, x)$

provided by theory ← → determined experimentally

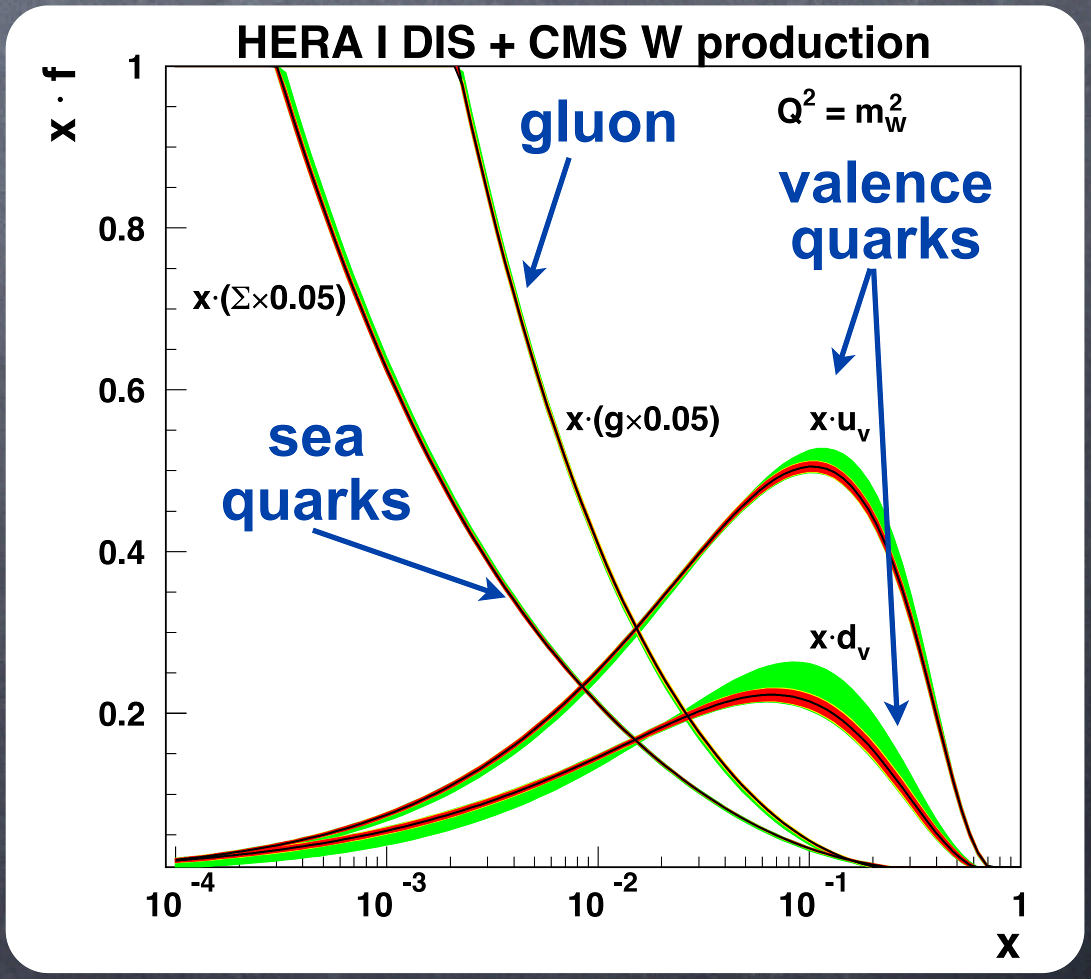


Example of PDF determination:

- parameterize x-shape at a scale Q^2_0:

$$f(x) = Ax^B (1-x)^C (1+Dx+Ex^2)$$

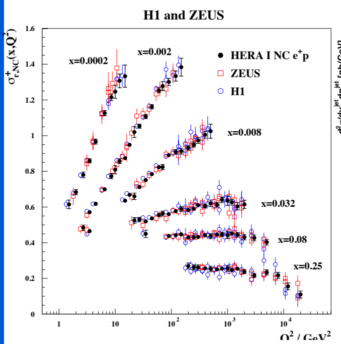
- evolve these PDFs to $Q^2 > Q^2_0$
  (e.g. using DGLAP evolution equations)
- construct expected cross sections
- χ^2- fit to the experimental data

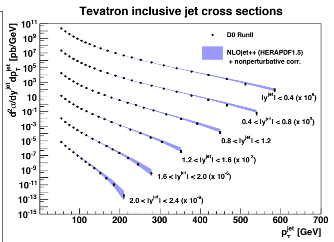


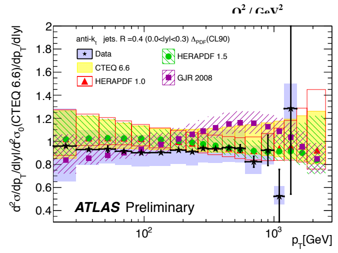
TOOL FOR PDF DETERMINATION

Unique tool to test impact of the measurements on e.g. PDFs **during data analysis**

experimental input







experiments:
HERA, Tevatron, LHC, fixed target

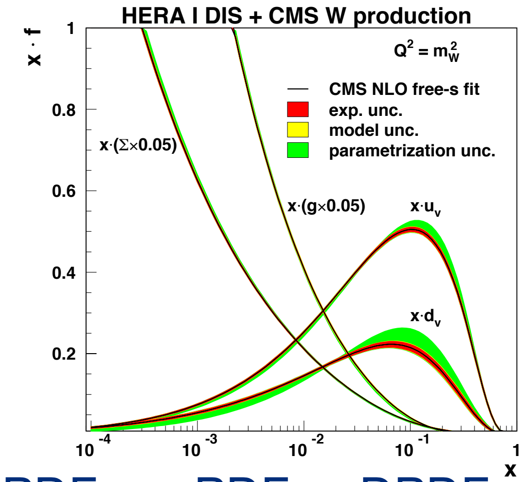
processes:
NC, CC DIS, jets, diffraction, heavy quarks (c,b,t)
Drell-Yan, W production

theoretical calculations/tools

Heavy quark schemes:	MSTW, CTEQ, ABM
Jets, W, Z production:	fastNLO, Applgrid
Top production	NNLO (Hathor)
QCD Evolution	DGLAP (QCDNUM)
	k_T factorisation
Alternative tools	NNPDF reweighting
Other models	Dipole model

+ Different error treatment models
+ Tools for data combination (HERAverager)

XFitter



PDF or uPDF or DPDF

$\alpha_S(M_Z), m_c, m_b, m_t, f_s, ..$

Theory predictions

Benchmarking

Comparison of schemes

open-source QCD framework

developed by experimentalists and theorists

contribution from HERA, ATLAS and CMS

https://www.xfitter.org/xFitter/

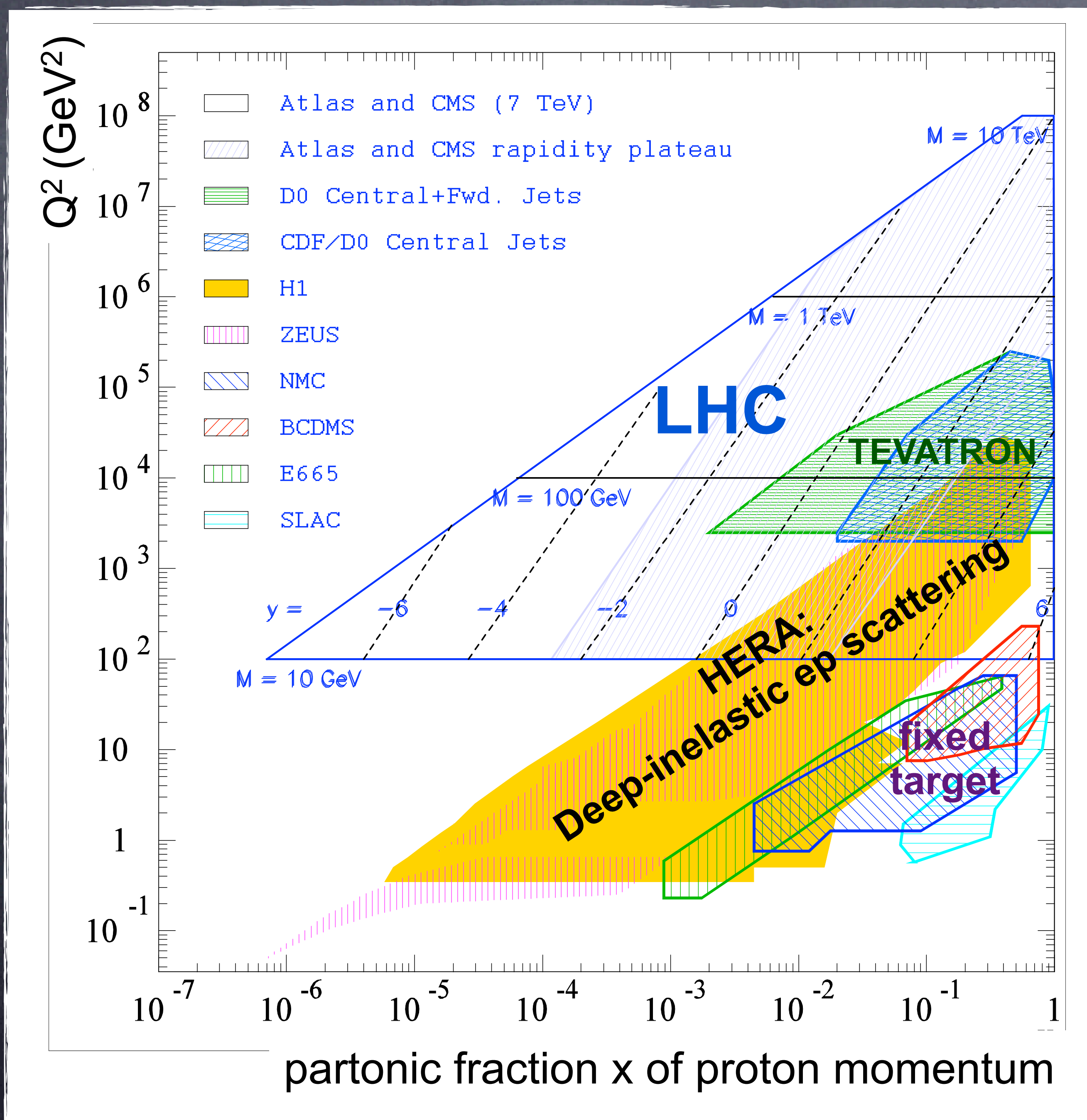
CMS results on PDFs shown in this talk are obtained using XFitter

PDF CONSTRAINTS FROM LHC

Parton Distribution Functions

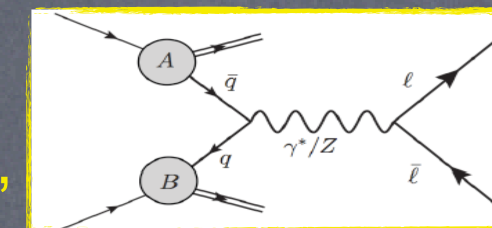
$f_i\,(Q^2, x)$

provided by theory ↙ ↘ determined experimentally

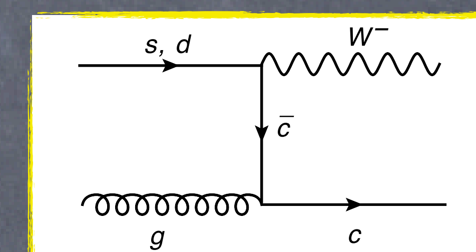


Impact of the LHC measurements:

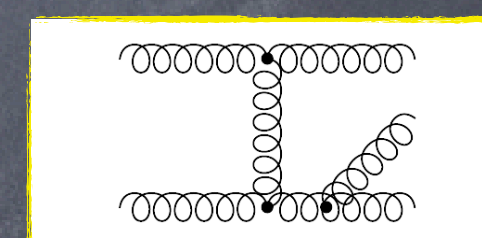
- DY: light quarks, flavor separation, gluon



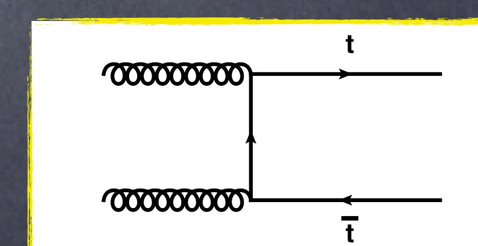
- V+HQ: s-quark, intrinsic charm



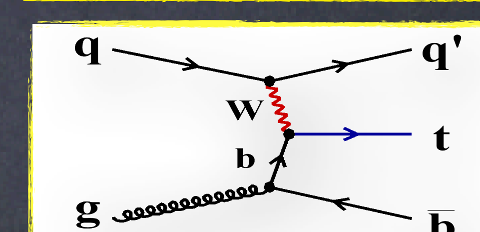
- jets: gluon, α_S medium-high x



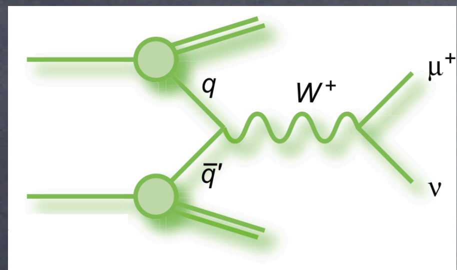
- top-pairs: gluon high x



- single top: u, d, b

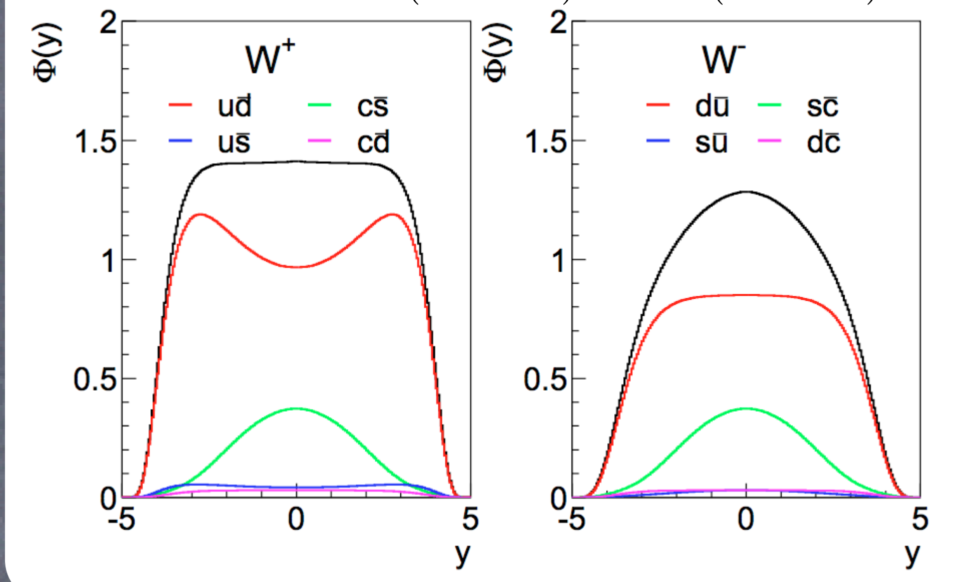


PROBING PDFs WITH W-BOSON PRODUCTION



production of W^+ or W^- probes different quark flavors

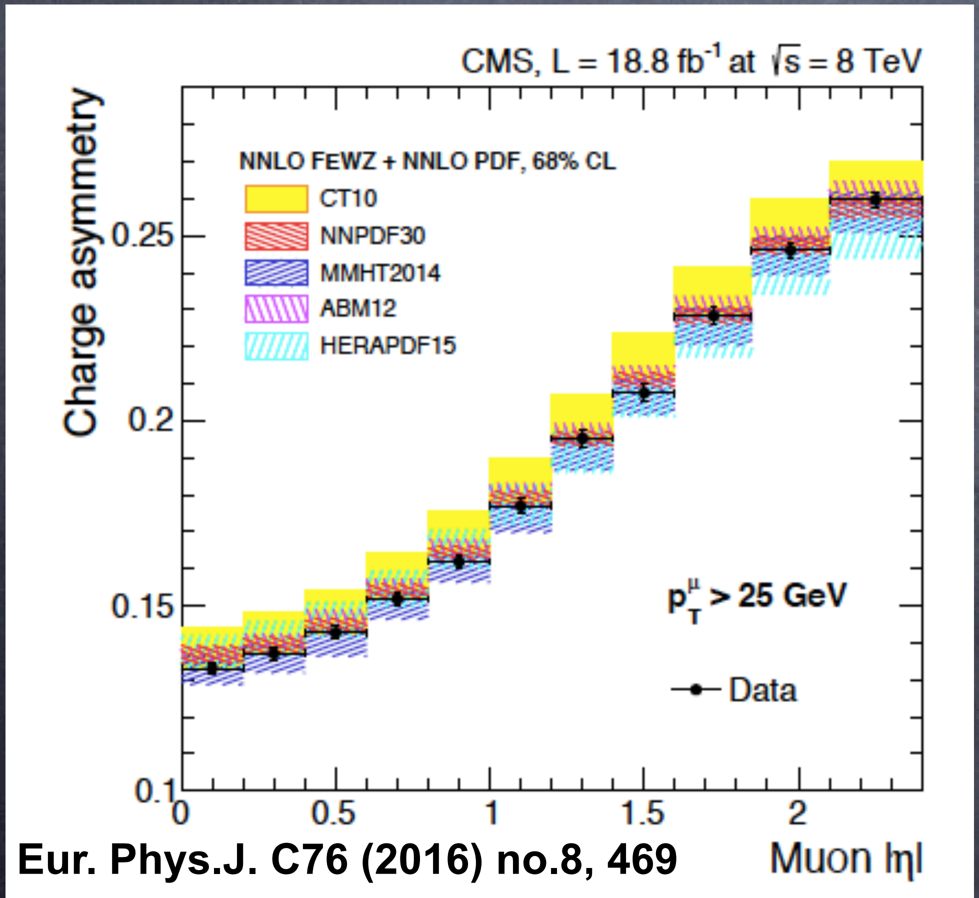
$$W^+ \approx 0.95(u\bar{d} + c\bar{s}) + 0.05(u\bar{s} + c\bar{d})$$
$$W^- \approx 0.95(d\bar{u} + s\bar{c}) + 0.05(d\bar{c} + s\bar{u})$$



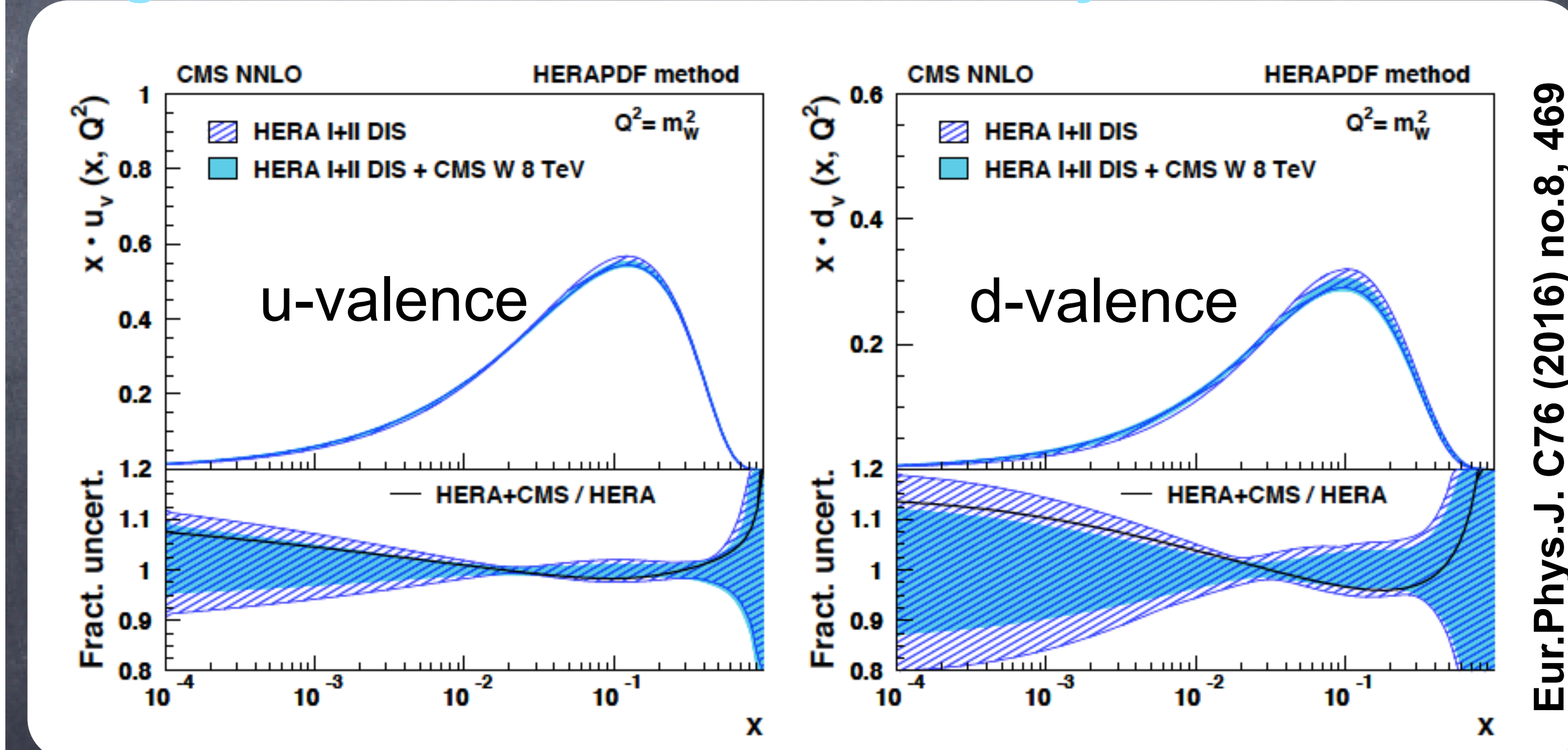
Lepton-charge asymmetry constrains valence

$$A_W = \frac{W^+ - W^-}{W^+ + W^-} \approx \frac{u_v - d_v}{u_v + d_v + 2u_{sea}}$$

CMS measurement used in a QCD analysis at NNLO

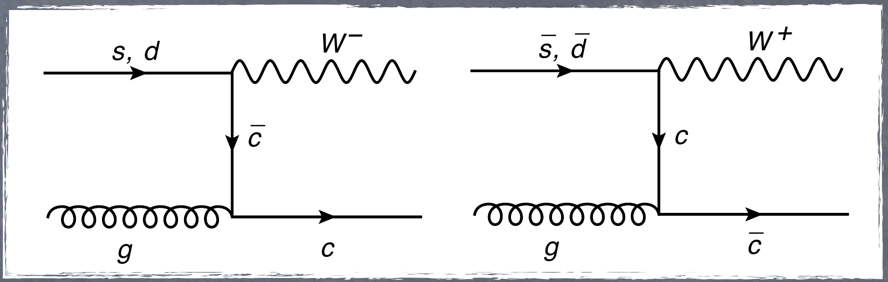


Significant reduction of uncertainty in the valence

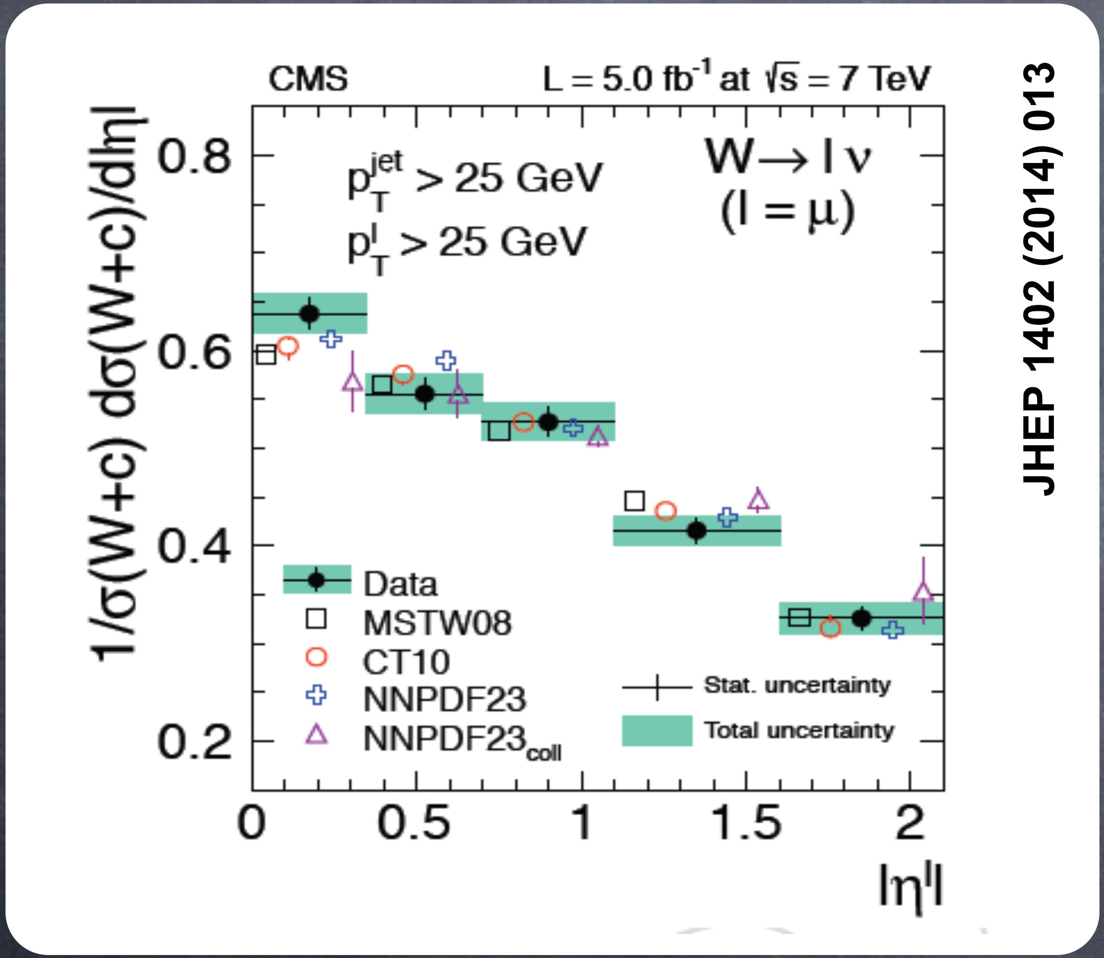


FLAVOUR DECOMPOSITION: W+CHARM

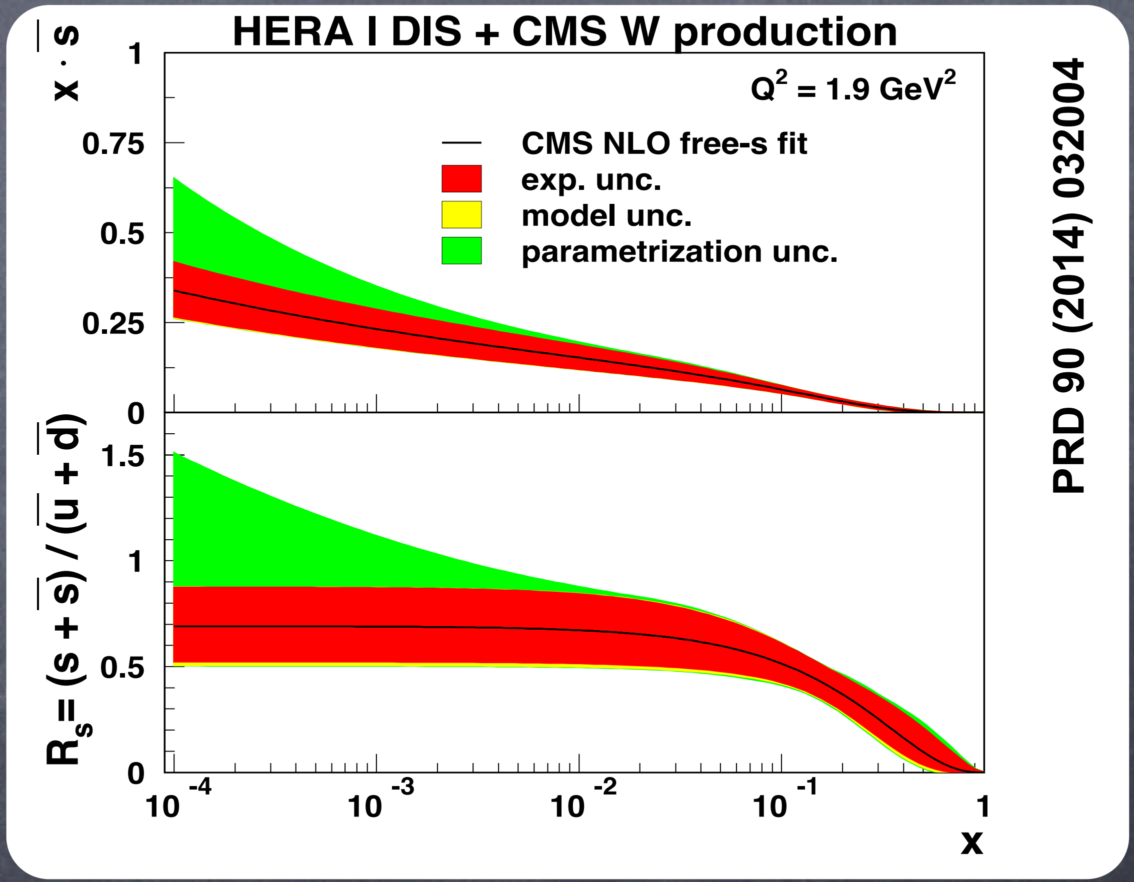
In pp collisions, production process of W+c probes strange quark directly at LO



Measure W+c-hadron production



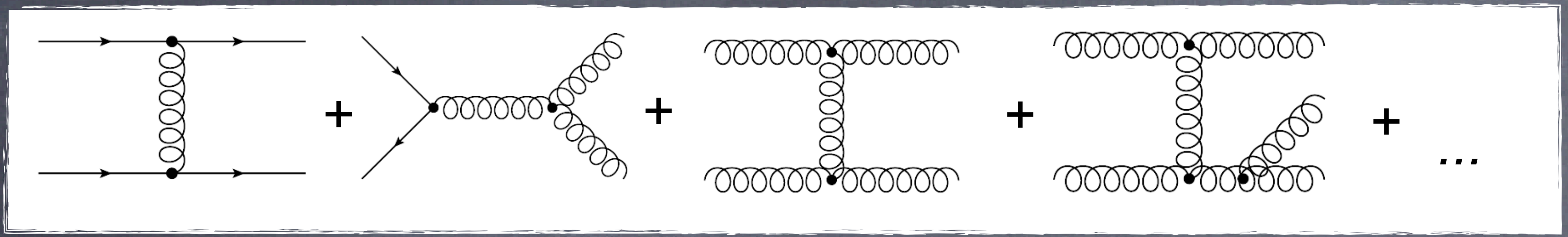
First direct determination of s-quark distribution at a hadron collider



Strangeness suppression factor $\kappa_s = 0.52^{+0.12}_{-0.10}(exp.)^{+0.05}_{-0.06}(mod.)^{+0.13}_{-0.10}(par.)$

in good agreement with neutrino experiments ***[Nucl.Phys. B876 (2013) 339, κ_s =0.59 ± 0.019]***

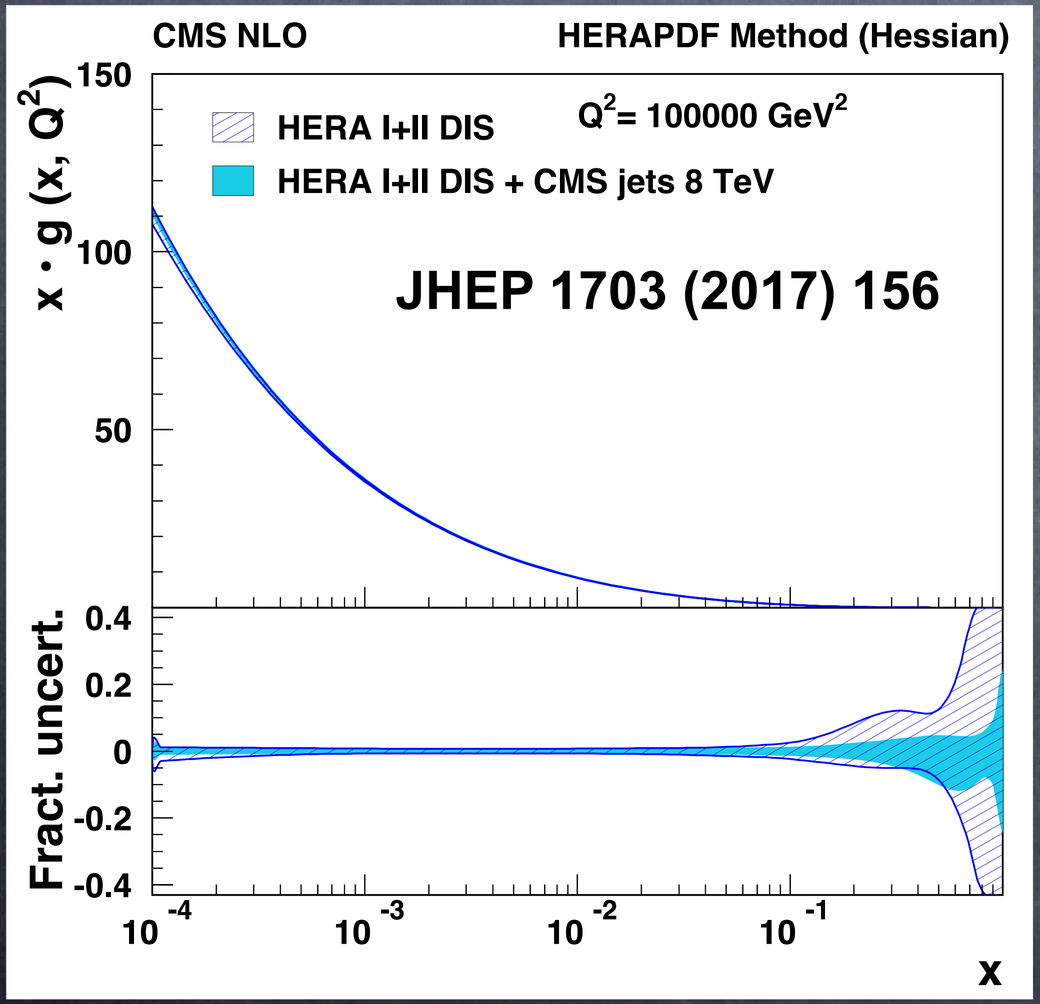
GLUON DISTRIBUTION: JET PRODUCTION



Jet production in pp collisions directly sensitive to PDFs and α_S

CMS 8 TeV, $\mathcal{L}$ = 19.7 fb^{-1} inclusive jet production,

2-differential cross sections vs jet p_T and y used in a QCD analysis at NLO:



Significant impact on the gluon distribution: reduced uncertainty at high x

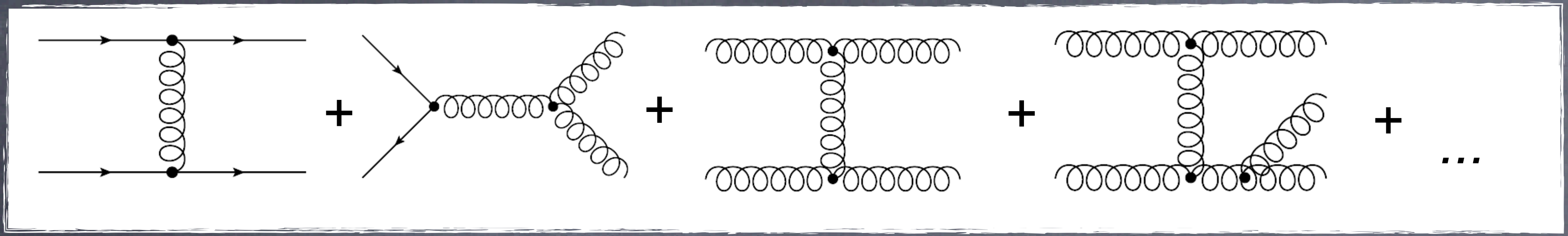
Strong Coupling, determined simultaneously with PDFs:

$$\alpha_s(M_Z) = 0.1185^{+0.0019}_{-0.0026}(PDF)^{+0.0022}_{-0.0018}(scale)$$

consistent with world average

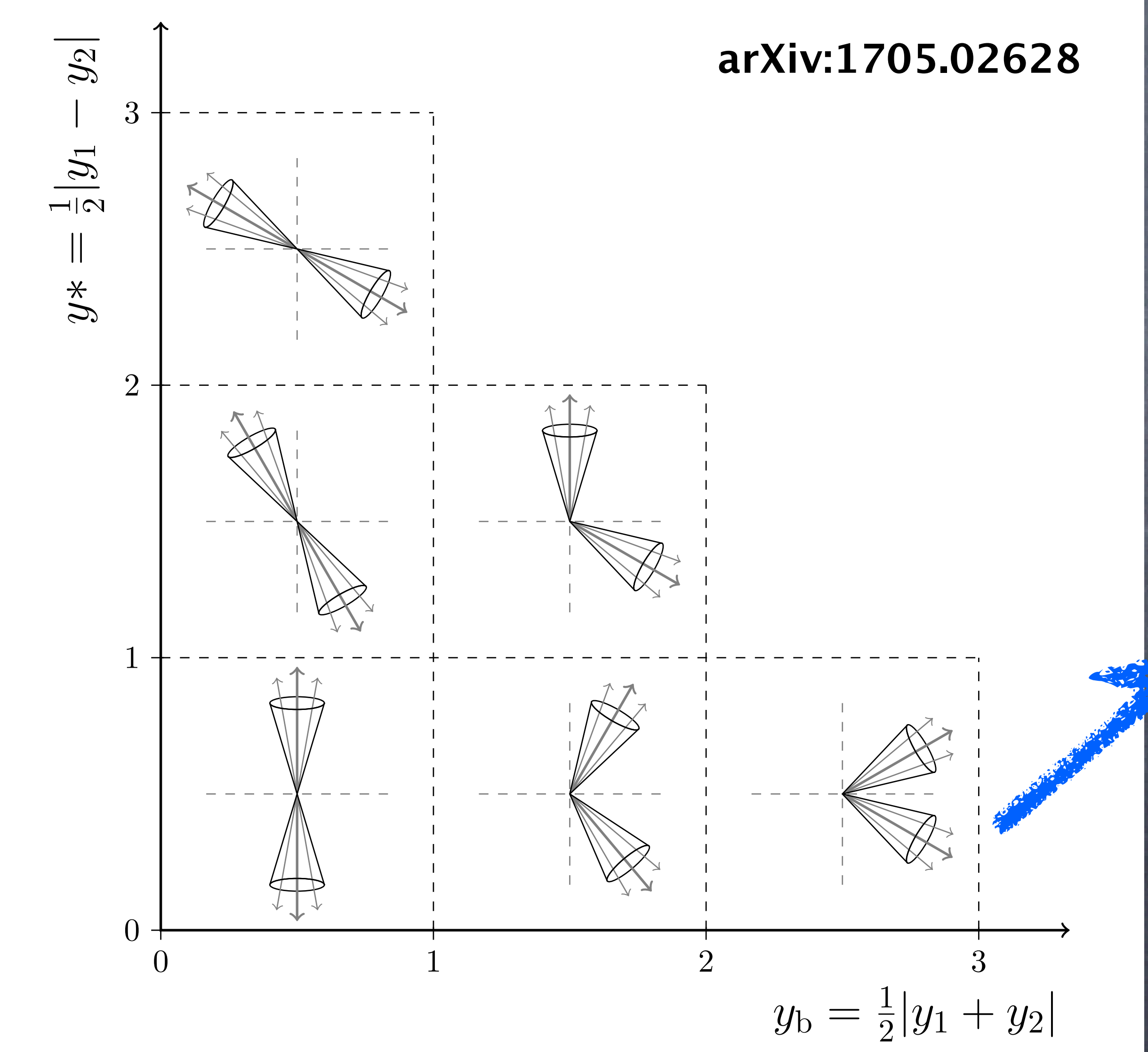
(dominant uncertainty from variations of the scales)

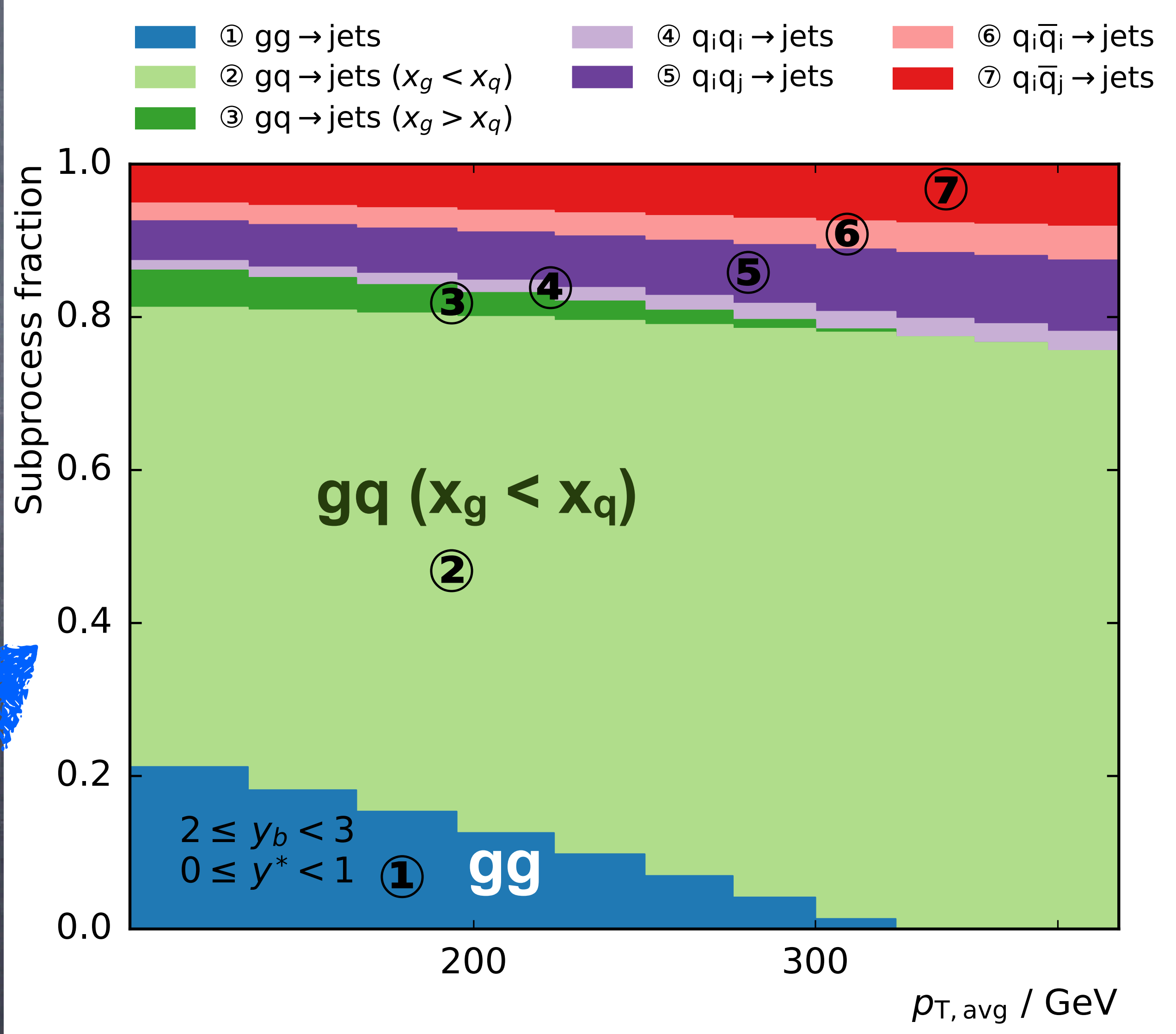
GLUON DISTRIBUTION: JET PRODUCTION



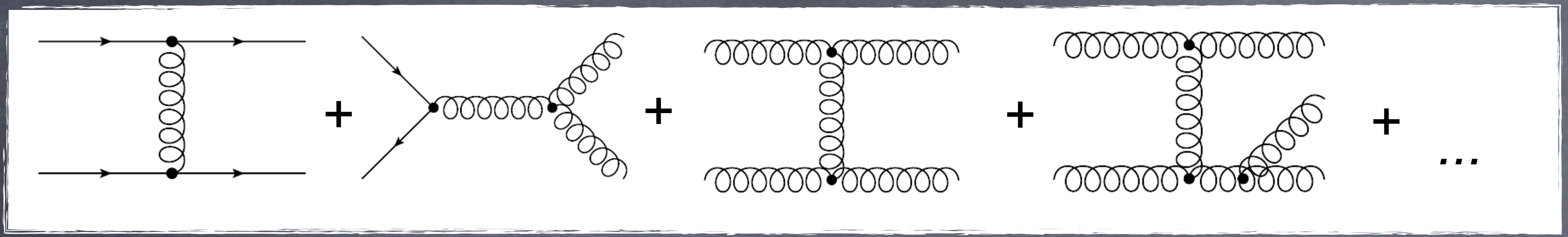
Jet production in pp collisions directly sensitive to PDFs and α_S

CMS 8 TeV, $\mathcal{L}$ = 19.7 fb^{-1} dijet production: 3-differential cross sections vs of jet <p$_T$>, rapidity separation and boost: probe x_1 and x_2 using different event topologies



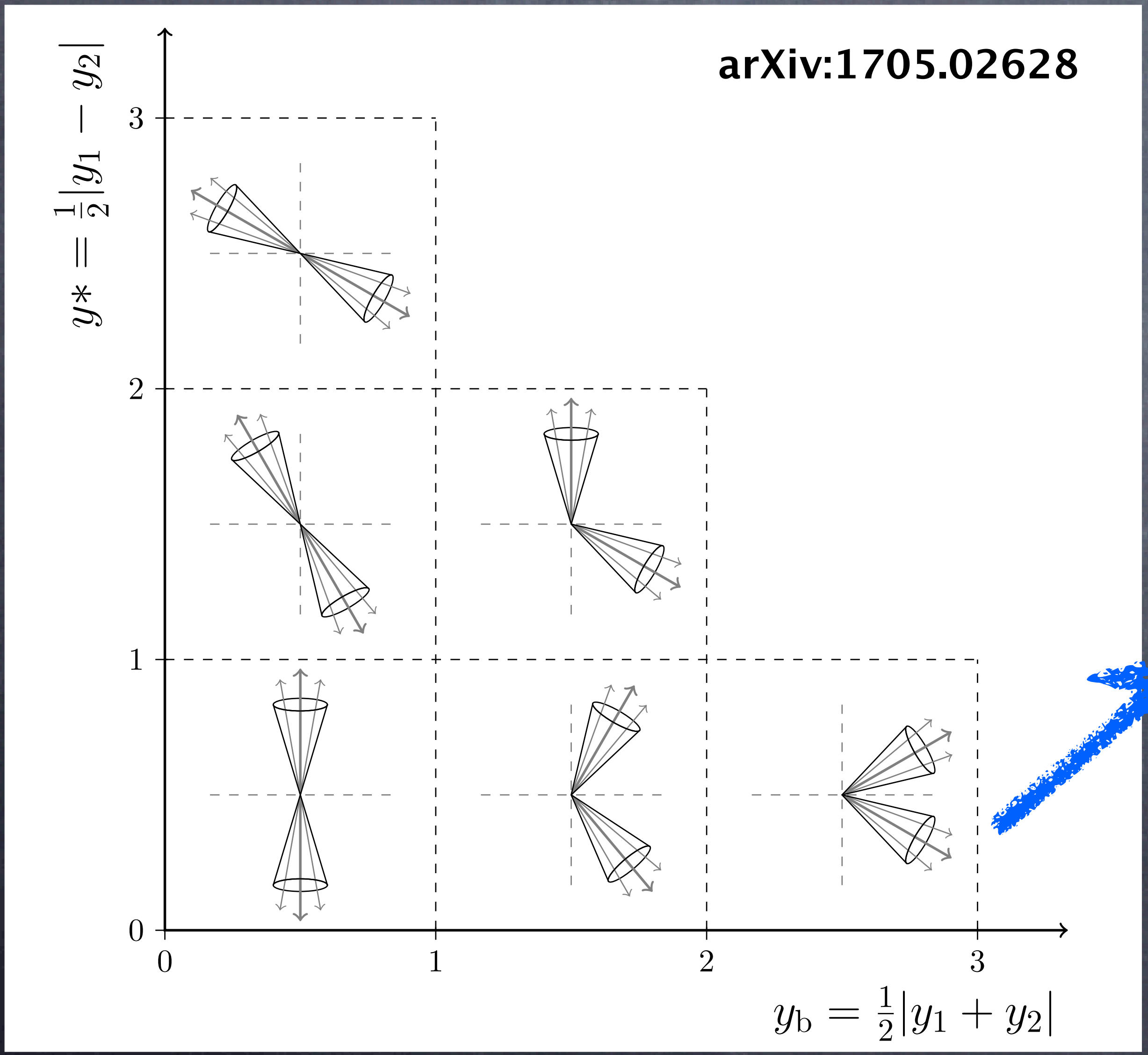


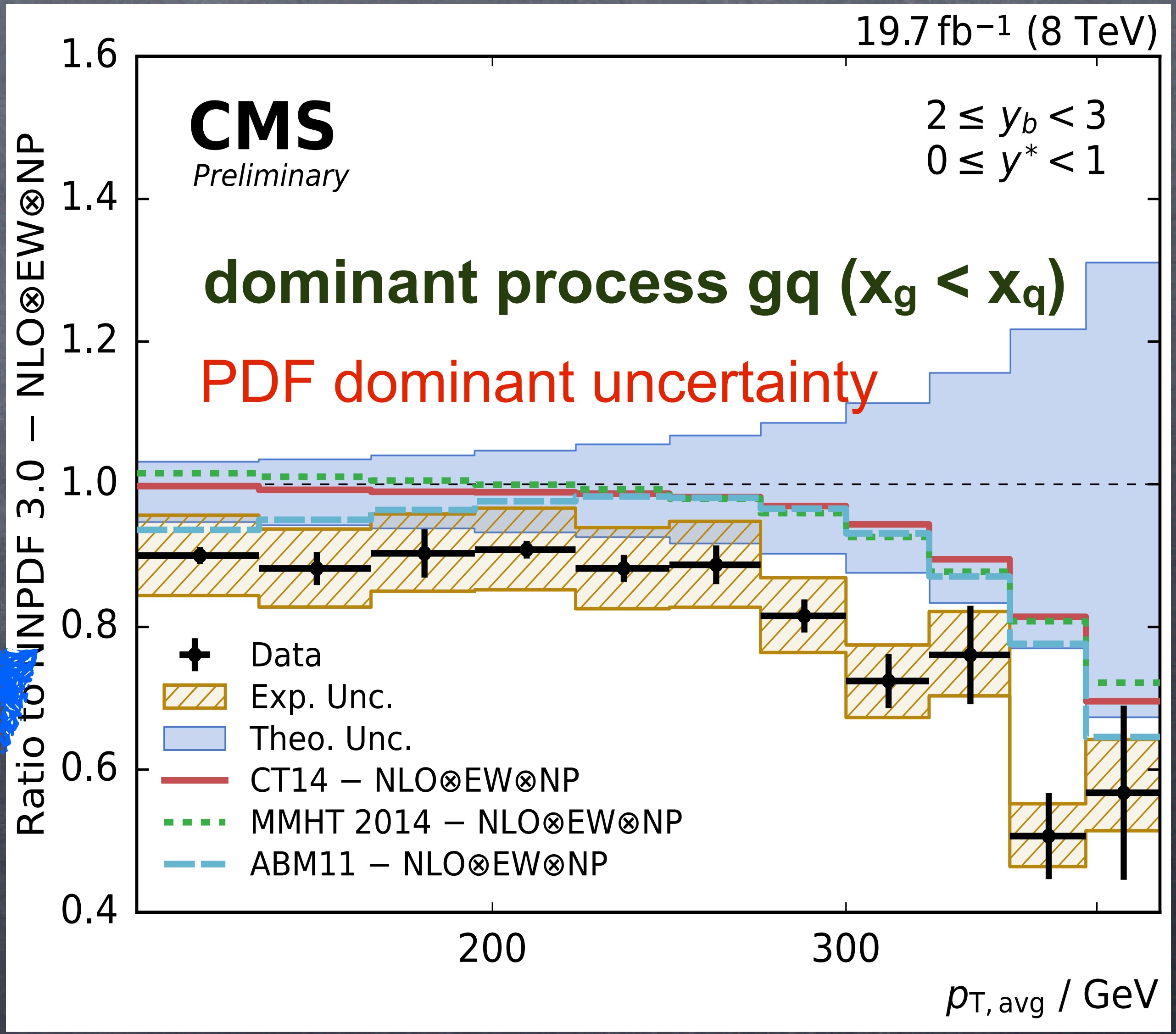
GLUON DISTRIBUTION: JET PRODUCTION



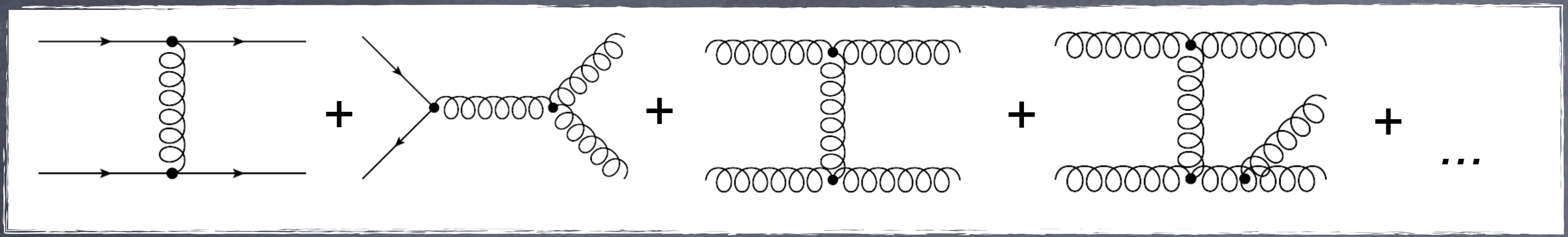
Jet production in pp collisions directly sensitive to PDFs and α_S

CMS 8 TeV, $\mathcal{L}$ = 19.7 fb^{-1} dijet production: 3-differential cross sections vs of jet <p_T>, rapidity separation and boost: probe x_1 and x_2 using different event topologies





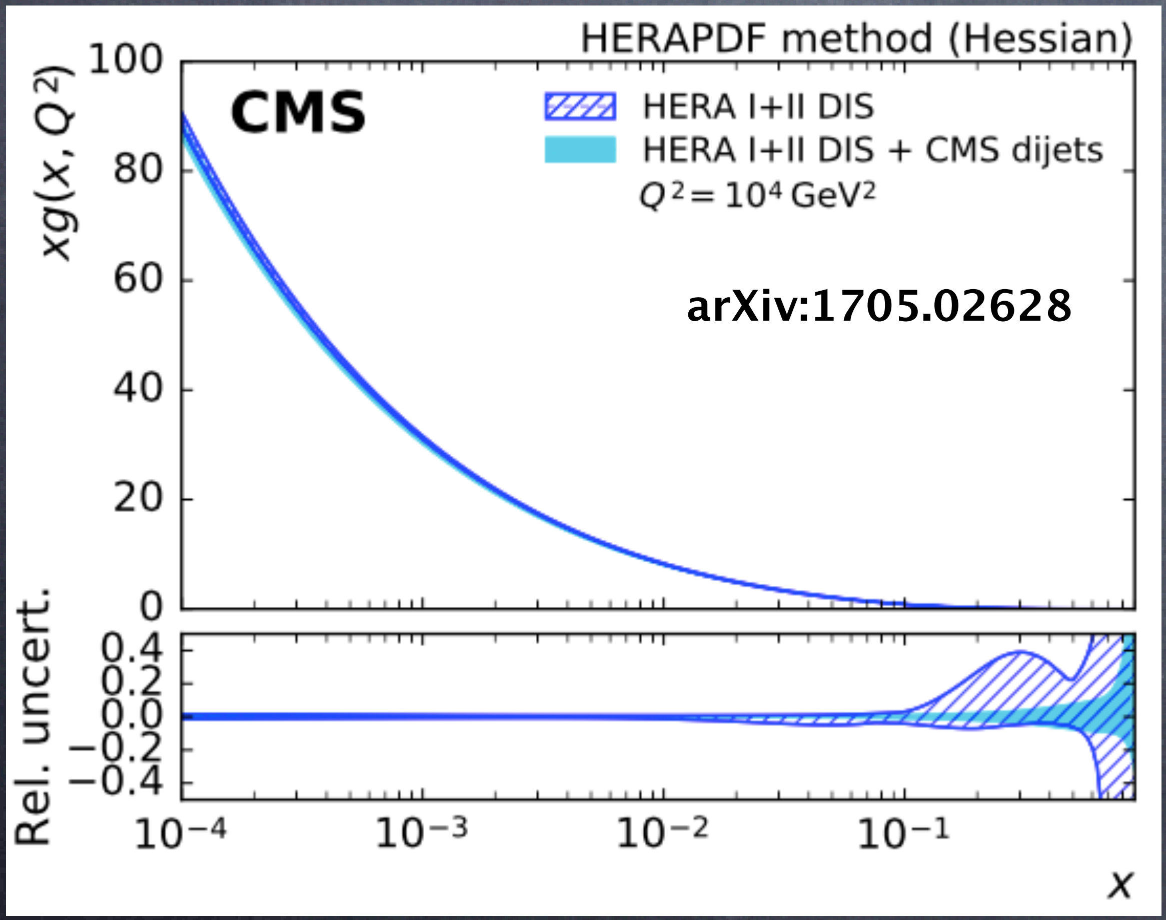
GLUON DISTRIBUTION: JET PRODUCTION



Jet production in pp collisions directly sensitive to PDFs and α_S

CMS 8 TeV, $\mathcal{L}$ = 19.7 fb^{-1} dijet production: 3-differential cross sections vs of jet <p_T>, rapidity separation and boost: probe x_1 and x_2 using different event topologies

By using dijet cross section in the QCD analysis in addition to HERA data...



Change in the gluon shape and reduced uncertainty at high x
similar as observed with inclusive jet data

Strong coupling determined simultaneously with PDFs:

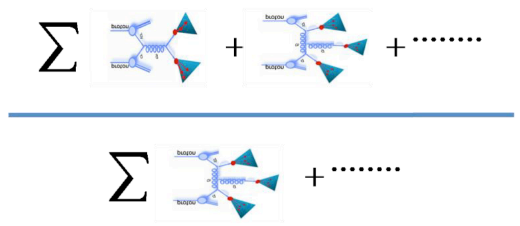
$$\alpha_s(M_Z) = 0.1199^{+0.0015}_{-0.0016}(PDF)^{+0.0026}_{-0.0016}(scale)$$

JETS @ CMS: GLUON AND STRONG COUPLING

CMS 8 TeV, $\mathcal{L}$ = 19.7 fb^{-1} multi-jet production CMS-PAS-SMP-16-008

Ratio of 3/2 inclusive jet cross sections

$$R_{32} = \frac{\sigma_3}{\sigma_2} = \frac{\sigma_{pp\to n\ jets+X;\, n\geq 3}}{\sigma_{pp\to n\ jets+X;\, n\geq 2}} =$$

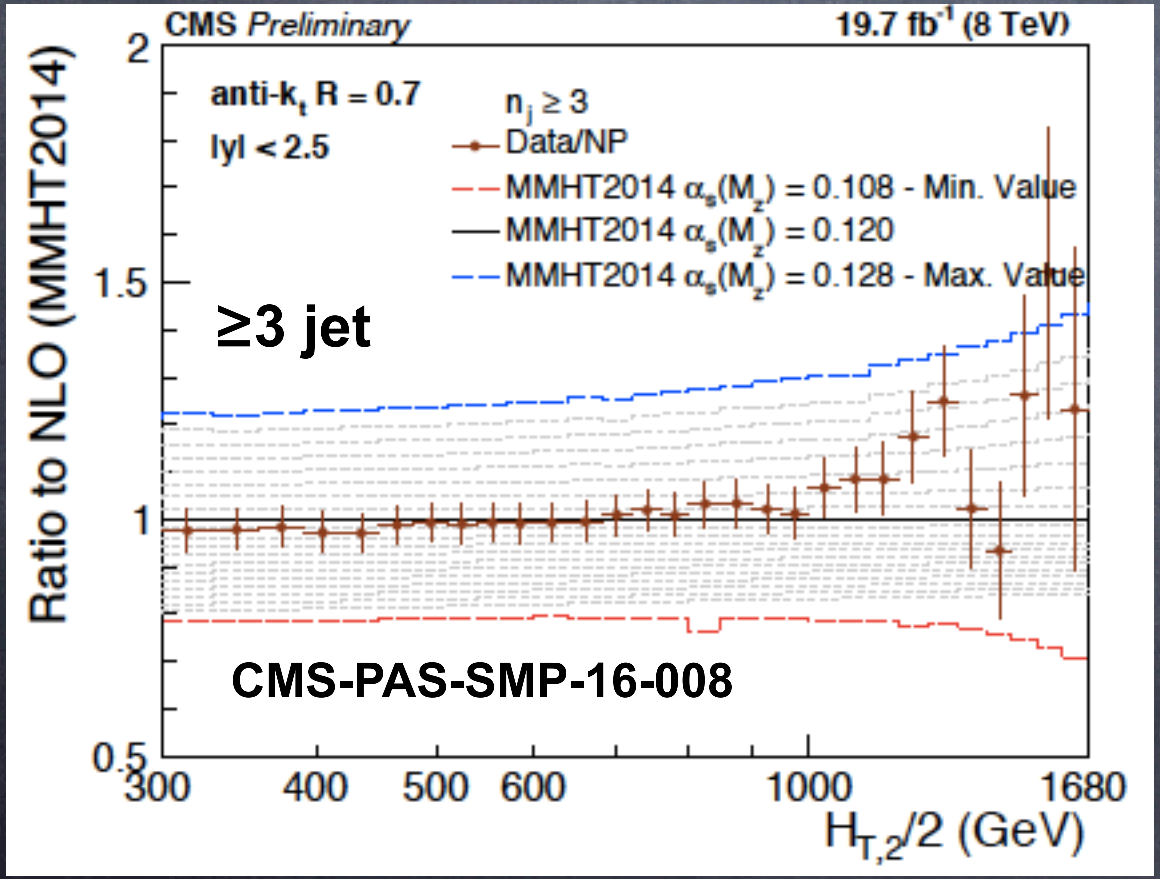


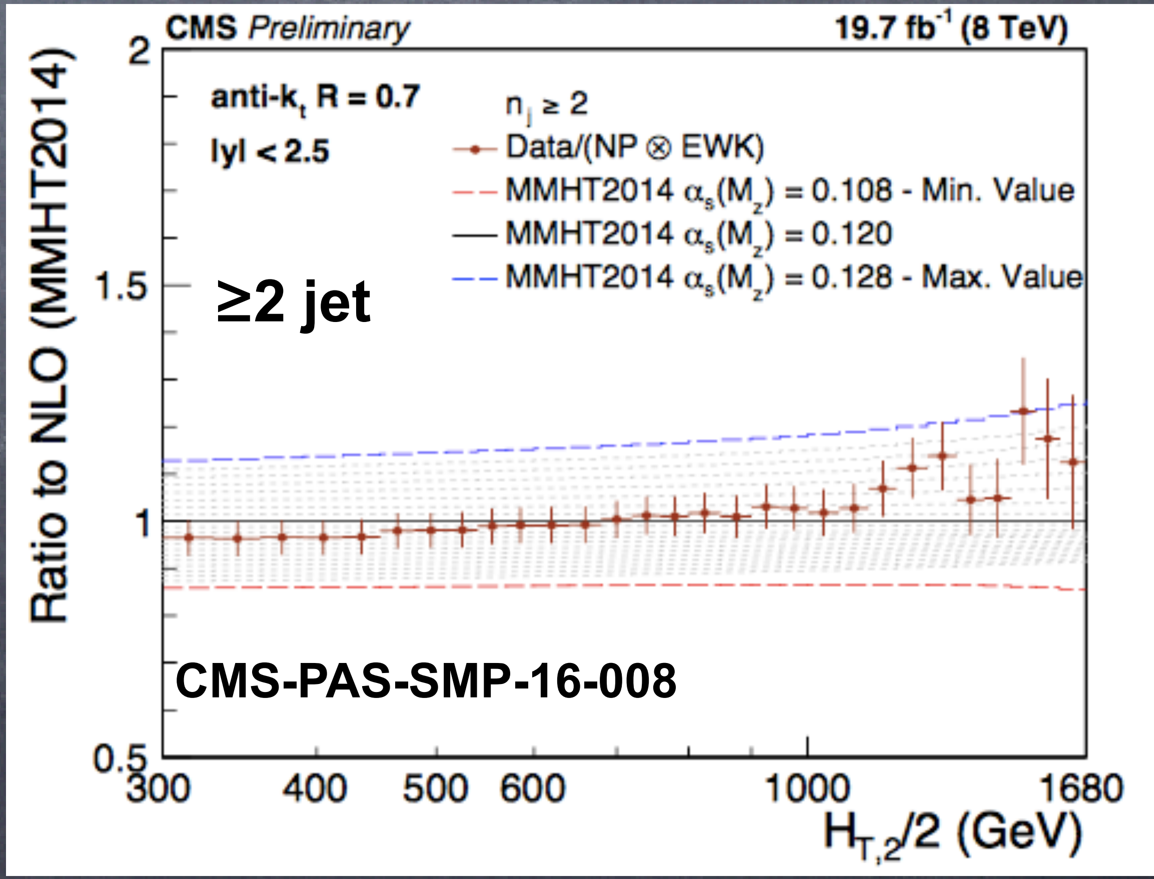
$\sim\alpha_s$

Theory: NLOJet++ via FastNLO, corrected for MPI, NP and EWK (2-jet)

scales $\mu_r=\mu_f=H_{T,2}/2=\frac{1}{2}(p_{T1}+p_{T2})$, varied independently by a factor of 2

NLO PDF sets studied: MSTW08, CT10, ABM11(N_F=5), NNPDF2.3 and 3.0 MMHT14, CT14

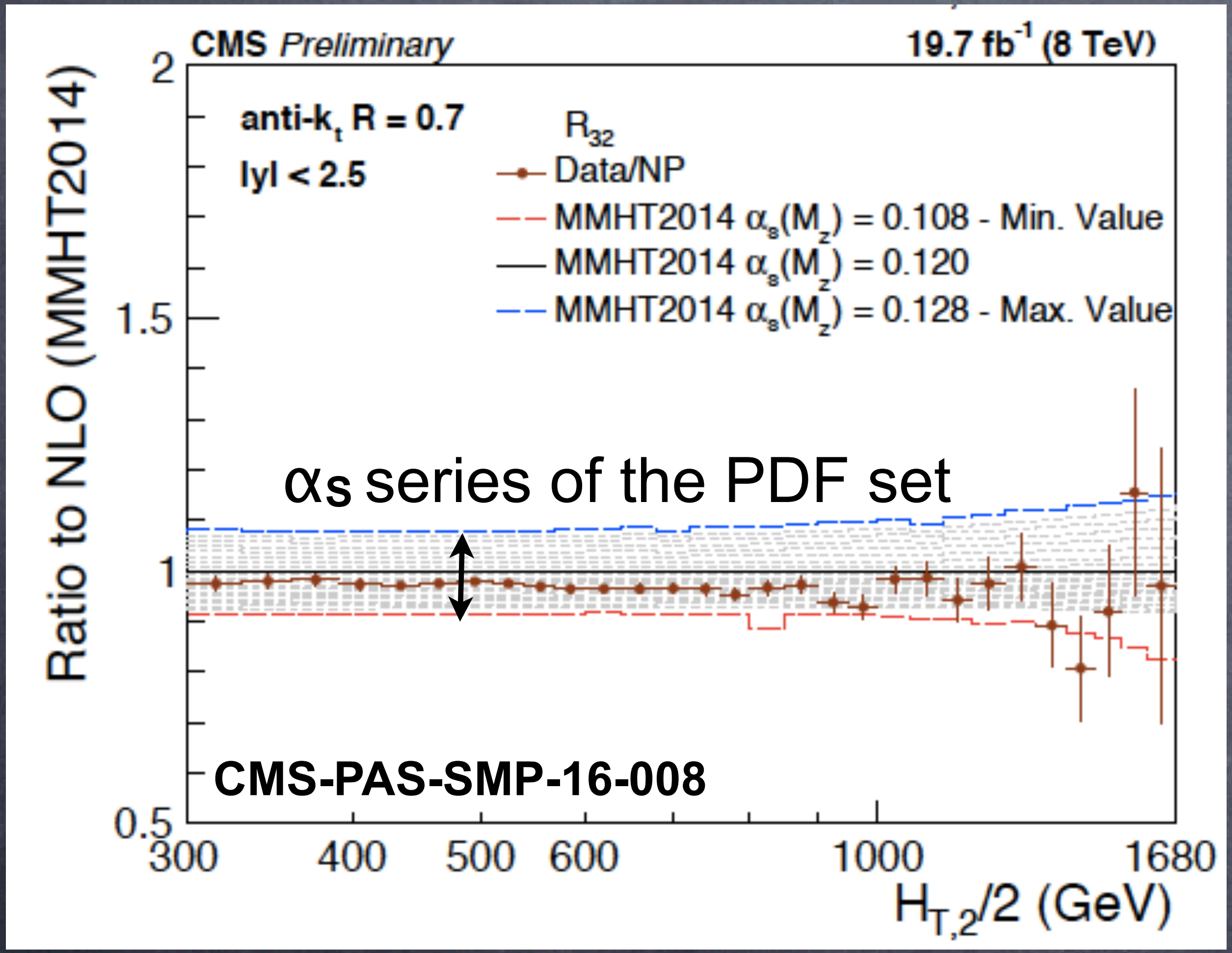




JETS @ CMS: GLUON AND STRONG COUPLING

Advantage of R_{32} : partial or full cancellation or reduction of experimental uncertainties, theory uncertainties due to NP effects, PDFs, scale choice, EWK corrections

α_S determined by minimizing χ^2 between the measurement and the theory



MMHT14: $\chi^2/\mathrm{n_{dof}}$ =24/28

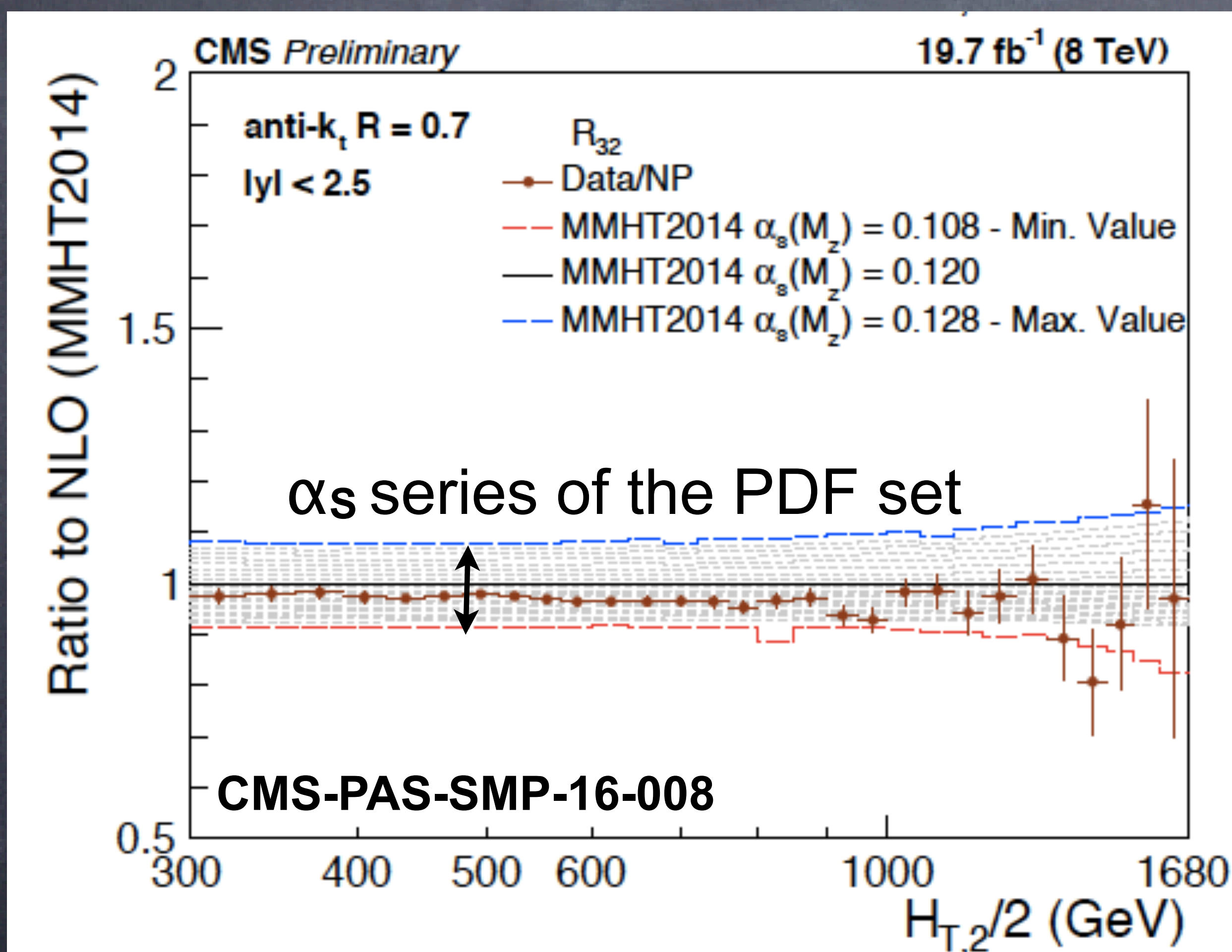
$$\alpha_S(M_Z) = 0.1142 \pm 0.0010(exp) \pm 0.0013(PDF) \pm 0.0014(NP)^{+0.0049}_{-0.0006}(scale)$$

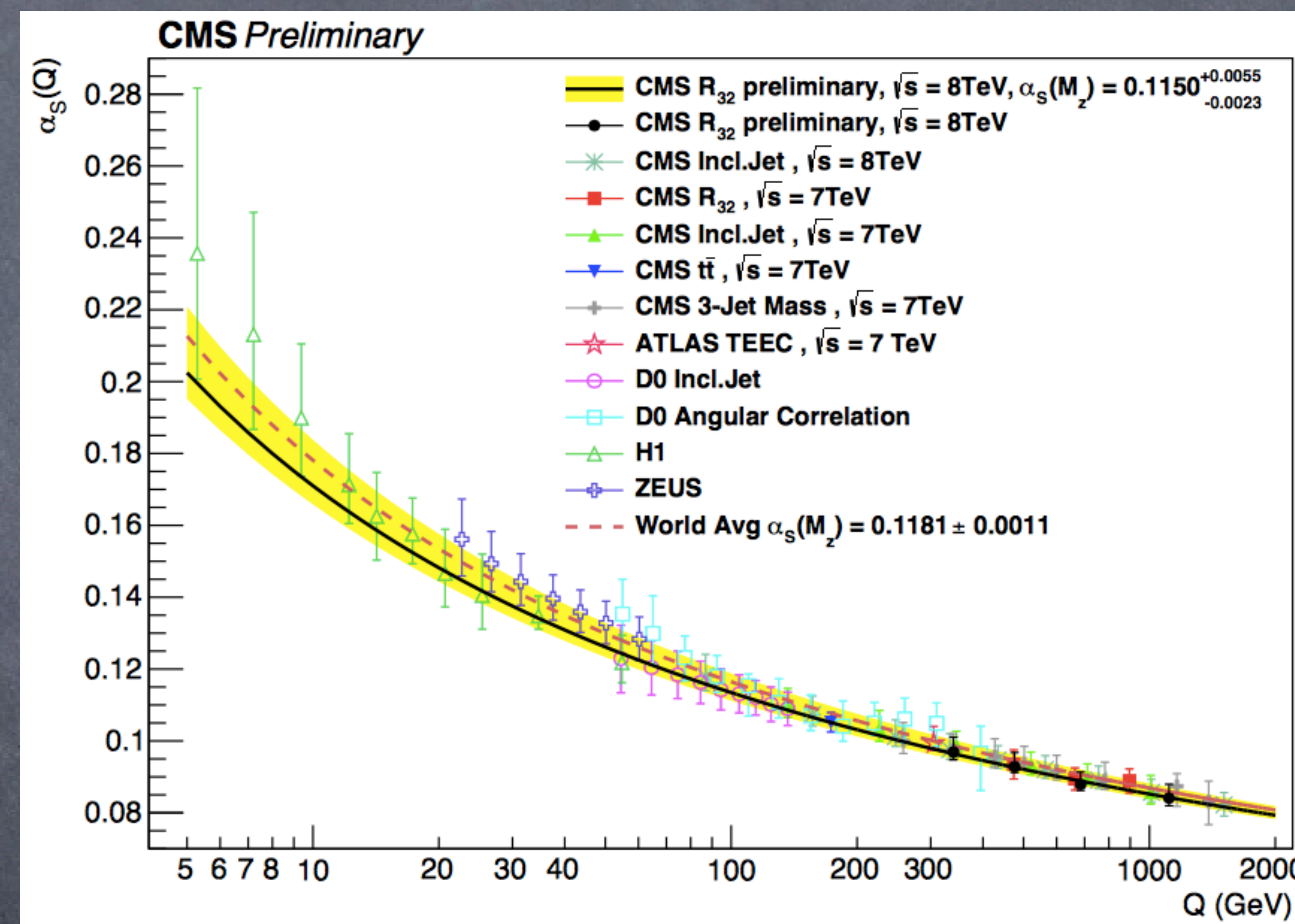
JETS @ CMS: GLUON AND STRONG COUPLING

Advantage of R_{32} : partial or full cancellation or reduction of experimental uncertainties, theory uncertainties due to NP effects, PDFs, scale choice, EWK corrections

α_S determined by minimizing χ^2 between the measurement and the theory

$\alpha_S(M_Z)$ value for each $H_{T,2}/2$ range → $\alpha_S(Q)$





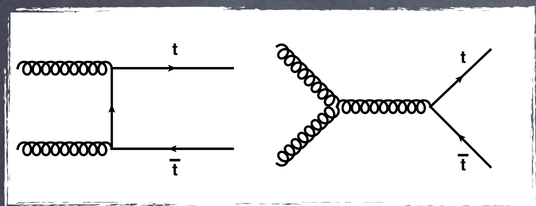
MMHT14: χ^2/n_{dof} =24/28

$$\alpha_S(M_Z) = 0.1142 \pm 0.0010(exp) \pm 0.0013(PDF)$$
$$\pm 0.0014(NP)^{+0.0049}_{-0.0006}(scale)$$

Evolution performed for $N_f = 5$ at 2-loops

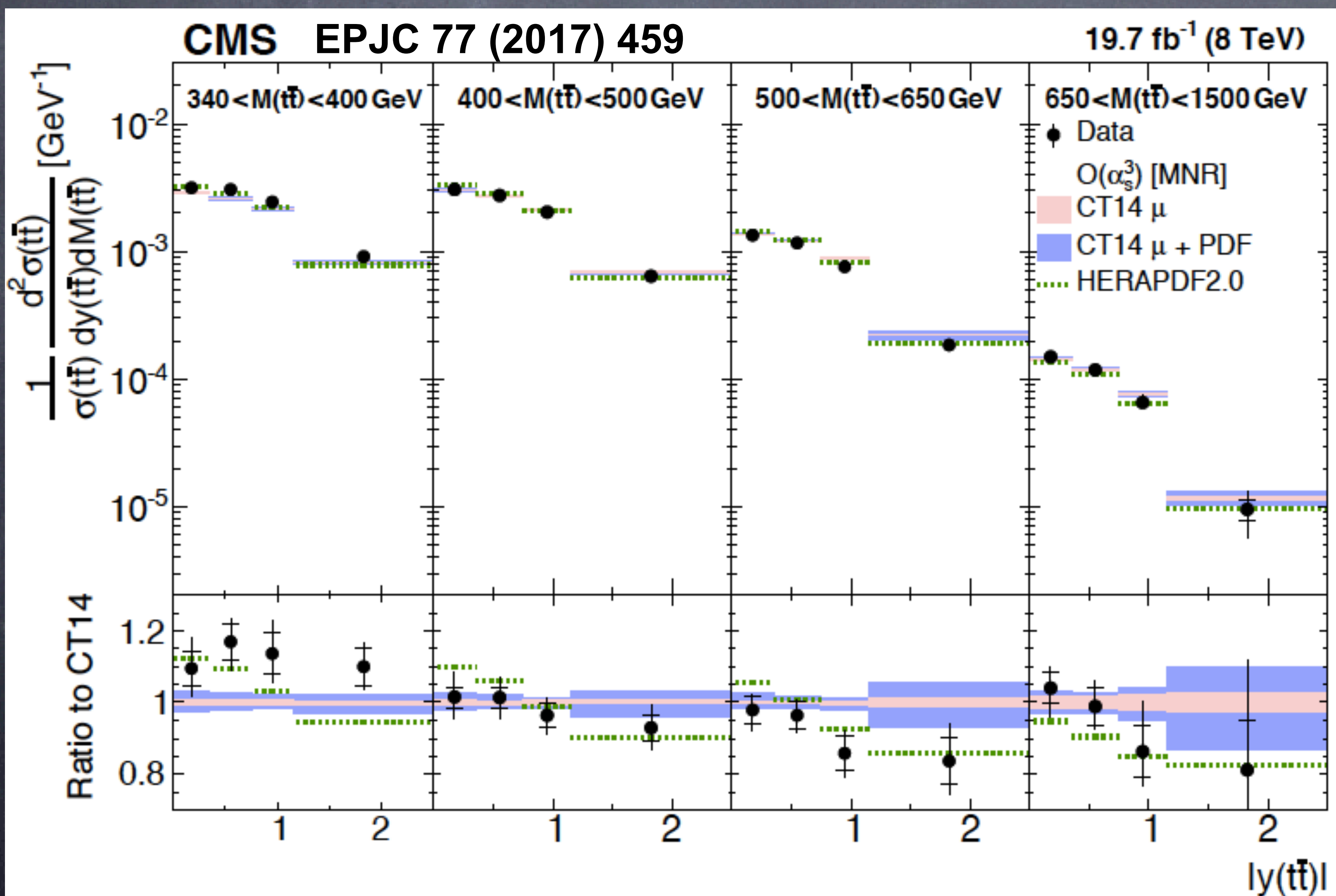
$t\bar{t}$ @ CMS: GLUON DISTRIBUTION AT HIGH X

In pp collisions top-quark pairs are produced via gg fusion probing gluon at high x



CMS 8 TeV, $\mathcal{L}$ = 19.7 fb^{-1} :

2d-differential $t\bar{t}$ cross sections, EPJC 77 (2017) 459



M(tt) and y(tt)

most sensitive to PDFs

at LO:

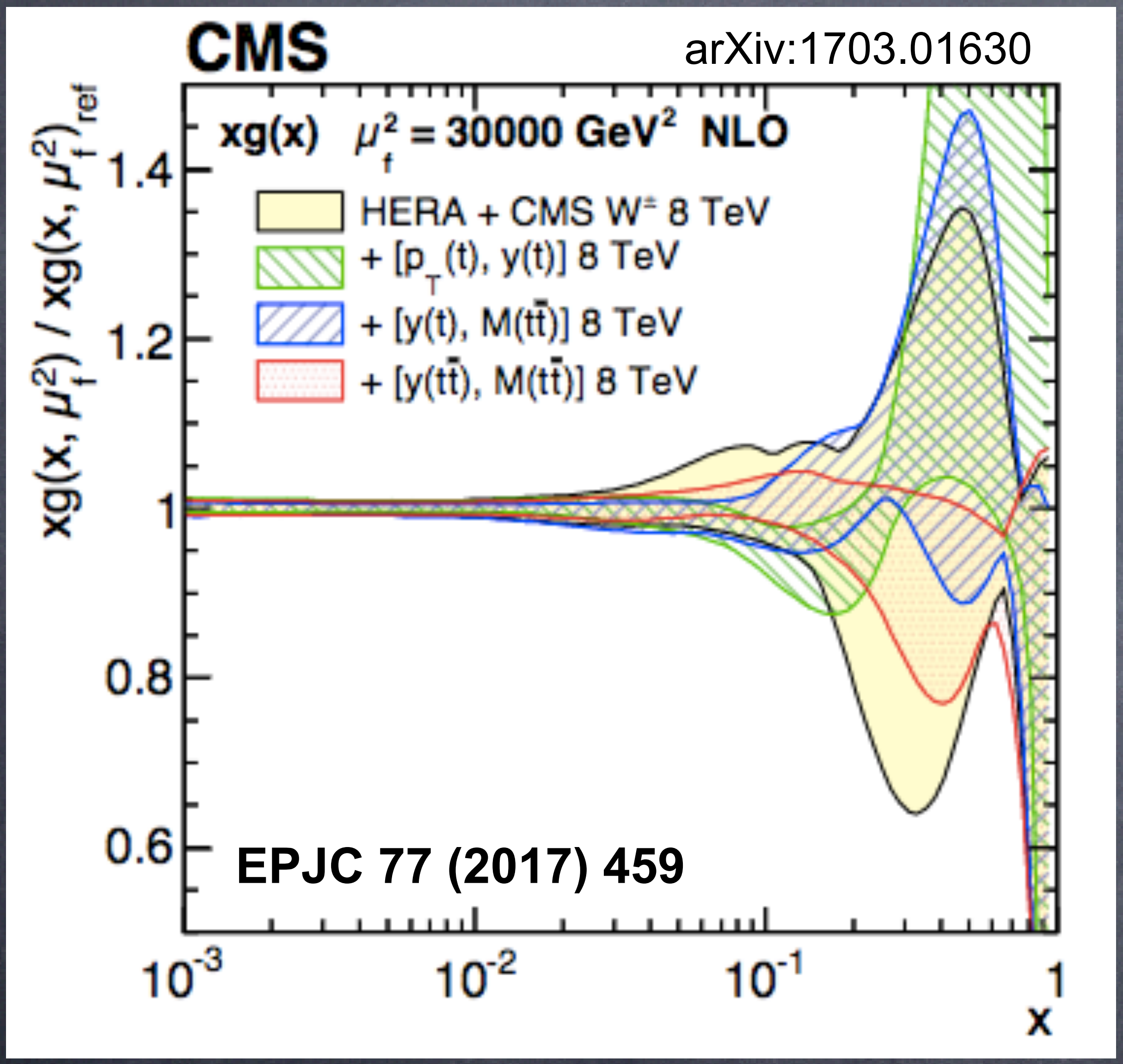
$$x_{1,2} = \frac{M(t\bar{t})}{\sqrt{s}} e^{\pm y(t\bar{t})}$$

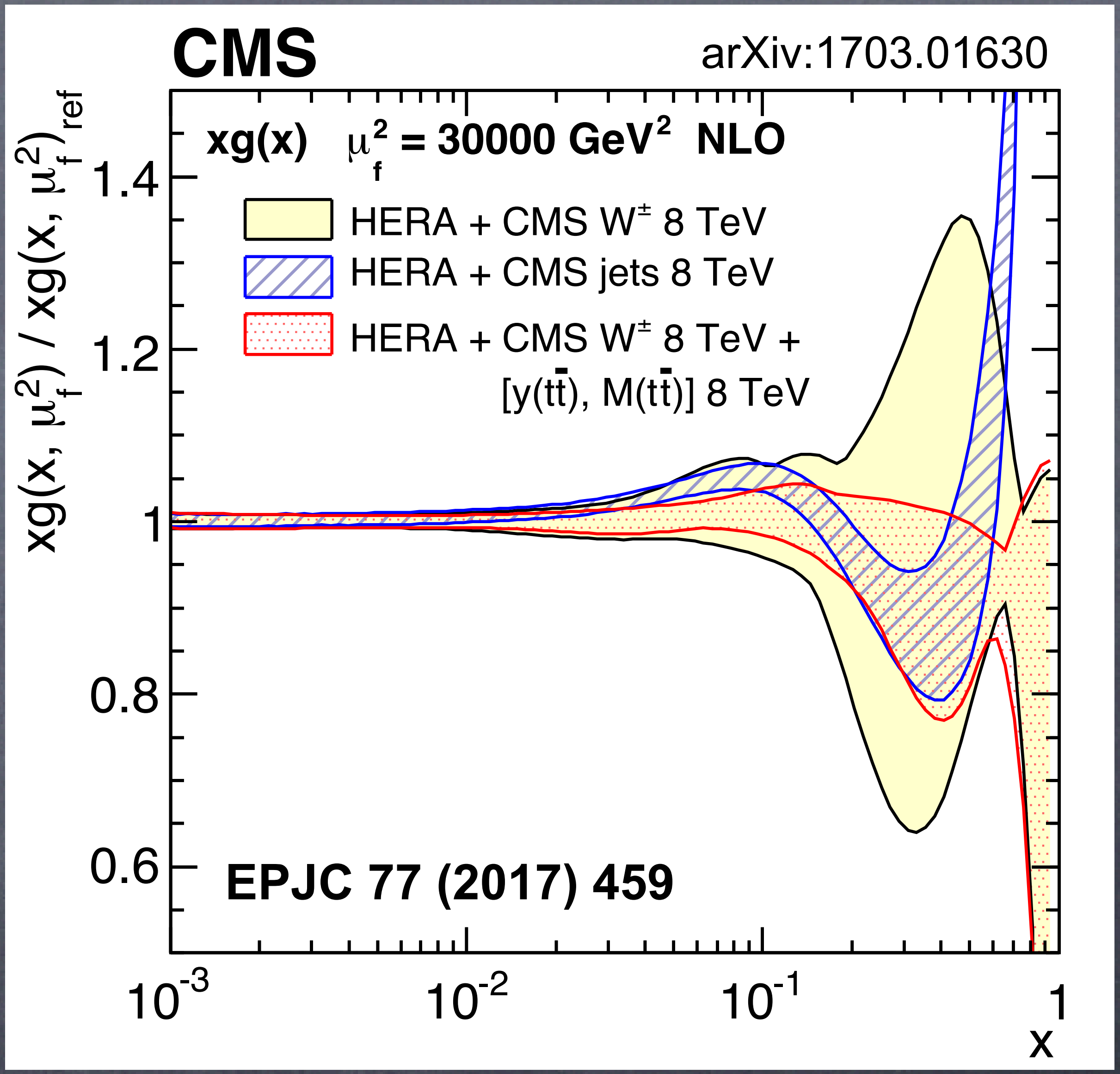
	HERA2	CT14
χ^2 (dof = 15)	29	16

$t\bar{t}$ @ CMS: GLUON DISTRIBUTION AT HIGH X

1-d and 2-d differential cross sections for different observables studied

Results compared to those obtained by using inclusive jets @ 8 TeV



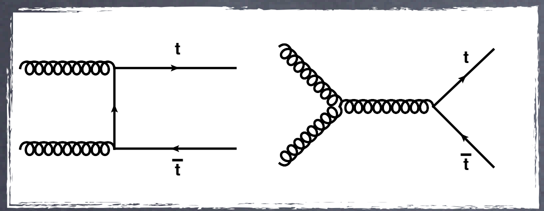


strongest constraints achieved by using 2d distributions in M_{tt} and y_{tt}

Recommend to use both data sets for further improvement of g(x) at high x

$t\bar{t}$ @ CMS: GLUON DISTRIBUTION AT HIGH X

In pp collisions top-quark pairs are produced via gg fusion probing gluon at high x

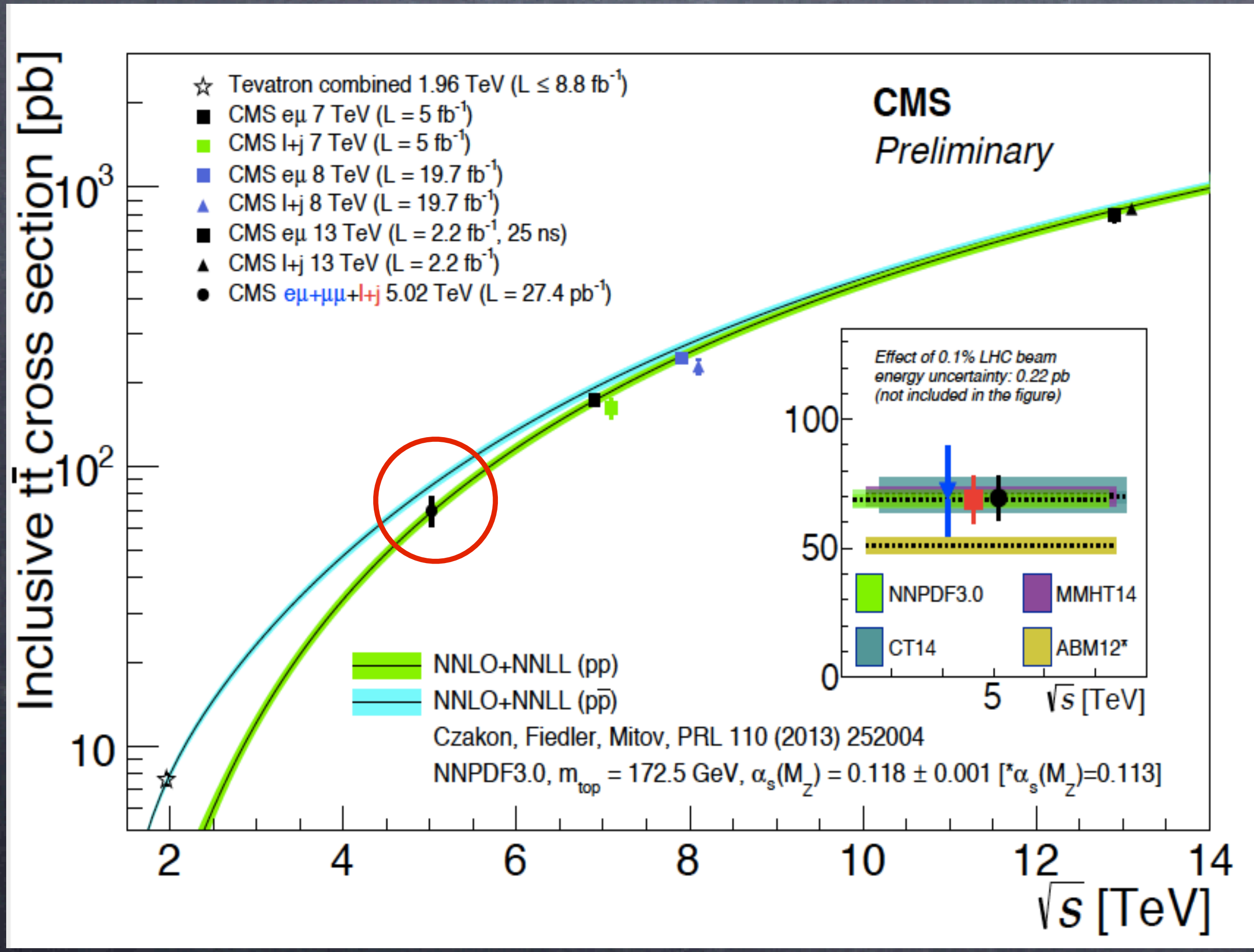


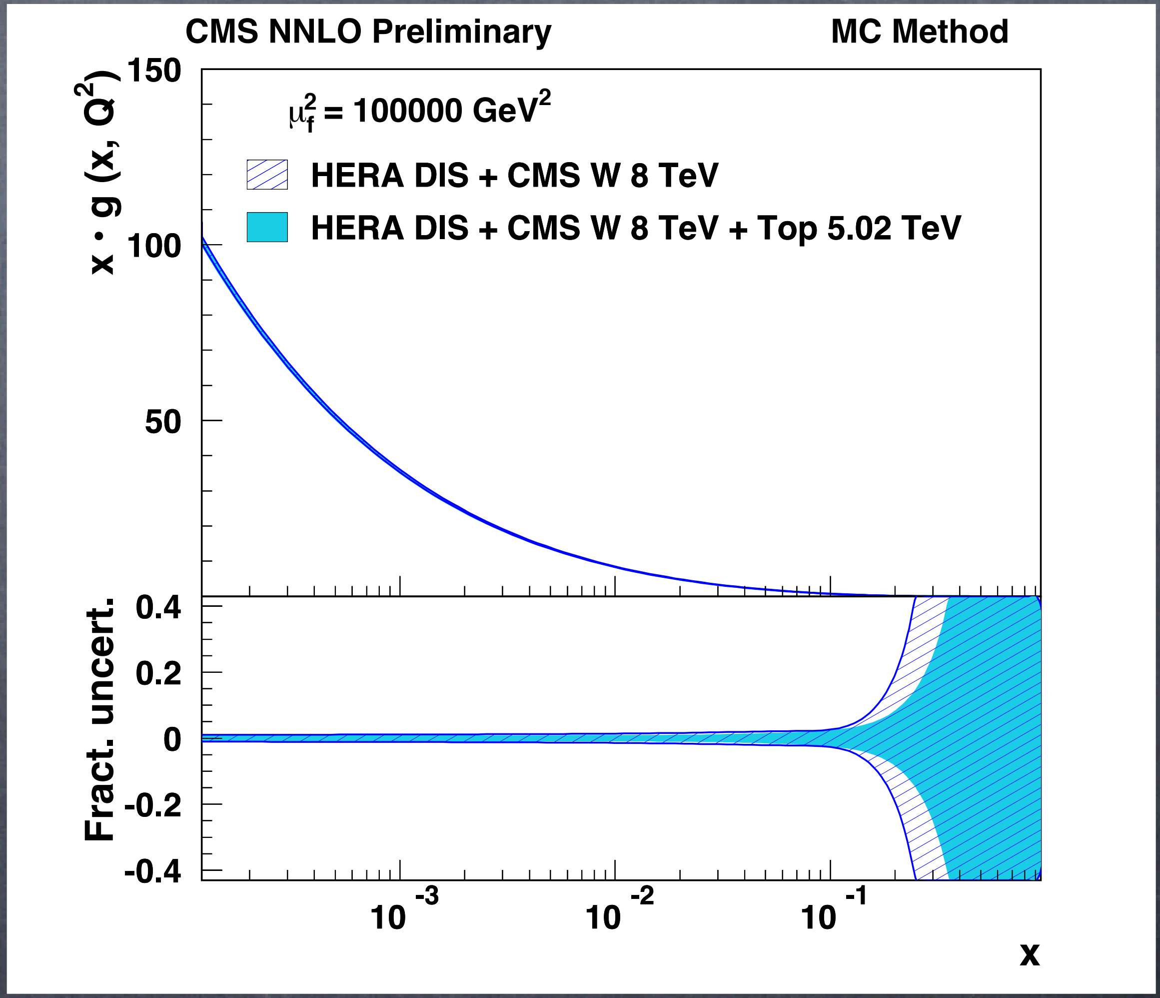
CMS 5.02 TeV, $\mathcal{L}$ = 27.4 pb^{-1} CMS-PAS-TOP-16-023

QCD analysis at NNLO

theory: HATHOR, m_t=172.5 GeV

new kinematic range probed

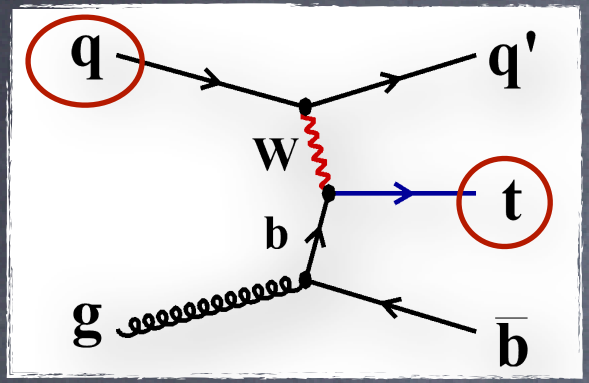




modest effect on g(x) at high x

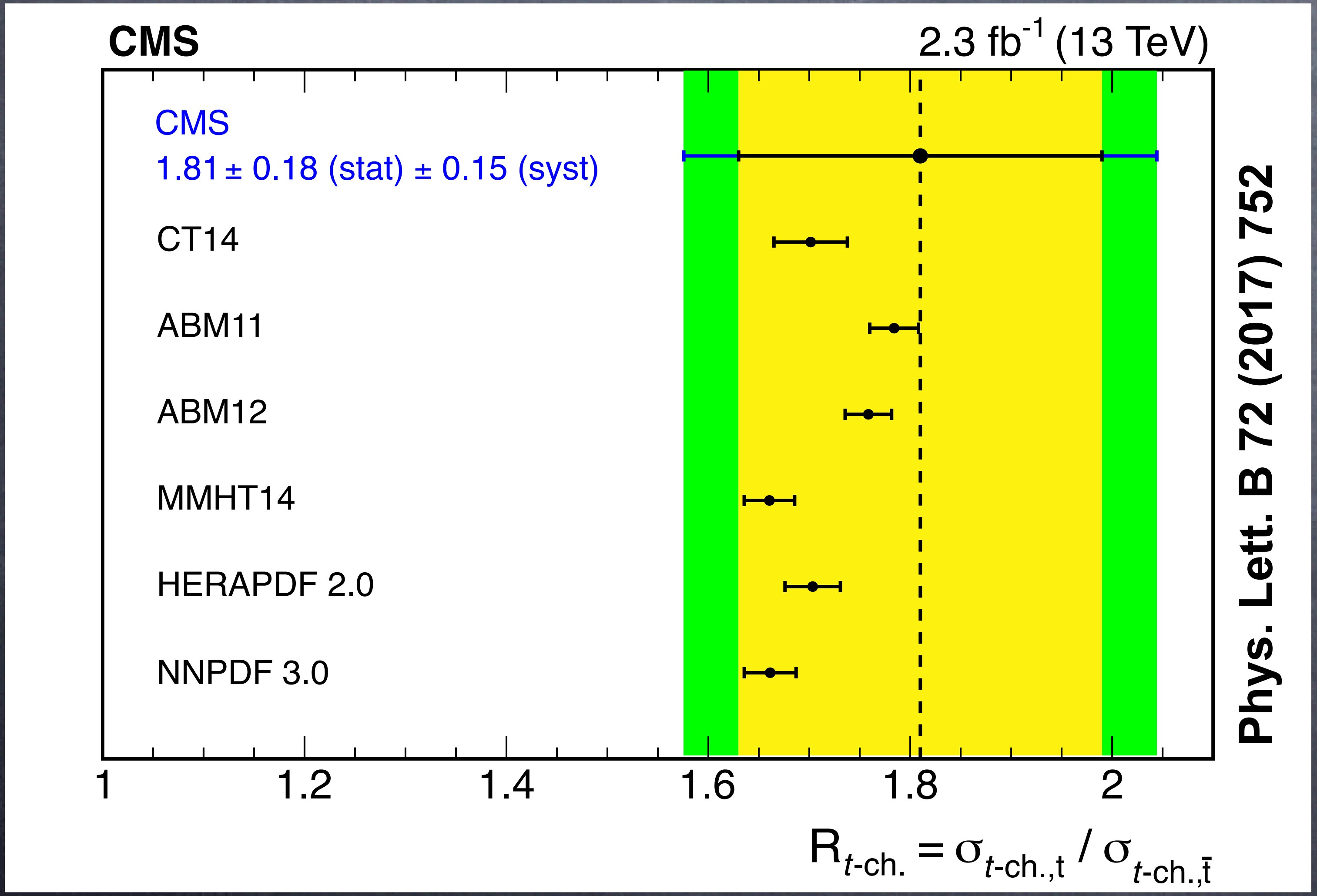
t and $\bar{t}$ @ CMS: PROBING THE LIGHT QUARKS

t-channel single top-quark production in pp collisions @ LHC



Probe the struck **light quark** through **charge** of top-quark measurement of $\sigma_t / \sigma_{\bar{t}}$ ratio R_t at **CMS 13 TeV (2.3 fb^{-1})**

Phys. Lett. B 72 (2017) 752



Dominant systematic uncertainty:
- Jet Energy Scale and Calibration
- Signal Modeling

Theory via POWHEG 4FS
Uncertainties account for variation of the scales and m_t

SUMMARY

LHC Run I CMS data used for improvement of PDF accuracy

• Inclusive and associated with Charm, W production constraints valence and strange sea of the proton

• jet data: gluon at medium & high x, strong coupling
→ getting even more interesting with available NNLO calculation

• Top-pair production has high potential to improve accuracy of g(x) at high x
→ remains important to constrain strong coupling & top quark mass

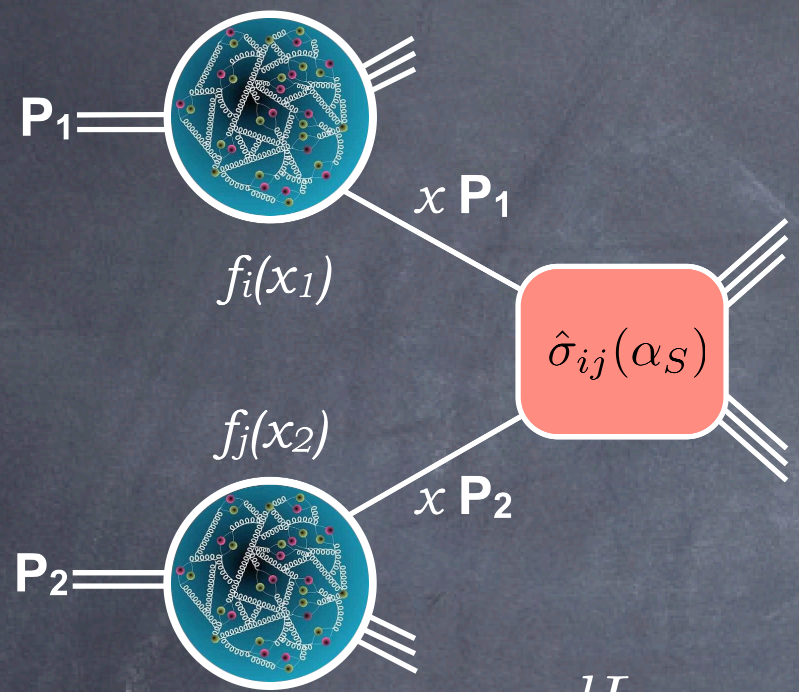
LHC Run II CMS data is forthcoming

Run I has shown high potential of the LHC
to improve the understanding of the proton structure,
more data are still to come to be used in precision QCD analyses

BACK UP

PARTICLE PRODUCTION IN *PP* COLLISIONS

proton structure hard interaction



Factorization:

proton structure ⊗ sub-process ME

$$\sigma(s) = \sum_{i,j} \int_{\tau_0}^{1} \boxed{\frac{d\tau}{\tau} \cdot \frac{dL_{ij}(\mu_F^2)}{d\tau}} \cdot \hat{s} \cdot \boxed{\hat{\sigma}_{ij}}$$

Ingredients for SM predictions for pp@LHC:

- partonic cross section calculable in pQCD
- parton luminosity:

$$\tau \cdot \frac{dL_{ij}}{d\tau} \propto \int_0^1 dx_1 dx_2 (x_1 f_i(x_1, \mu_F^2) \cdot x_2 f_j(x_2, \mu_F^2)) + (1 \leftrightarrow 2)\delta(\tau - x_1 x_2)$$

Parton Distribution Functions (PDFs)

universal functions of partonic fraction *x* of proton momentum and energy scale *Q* of the process

Precision of PDFs essential for interpretation of the LHC measurements
Precise LHC data are used to improve PDF constraints

CMS QCD ANALYSIS FRAMEWORK

QCD analyses at NLO / NNLO,

parton evolution in Q^2 via DGLAP as implemented in QCDNUM

Data in the QCD analysis:

- HERA I+II combined inclusive DIS data, Charged and Neutral Current *[JHEP 1001:109 (2010)]*
- Different CMS data sets (details in the next slides)

Experimental uncertainties: originate from uncertainties of the data, criterion $\Delta\chi^2=1$ is applied

Model input:

- Theory calculations at NLO/NNLO appropriate for each data set
- Starting scale of PDF evolution $Q^2_0 = 1.9$ GeV^2
- Heavy quark treatment: general mass variable flavor number scheme by Thorne-Roberts (TR)
- Heavy quark masses: $m_c = 1.4$ GeV, $m_b = 4.75$ GeV.

Model uncertainties: originate from variations of model input parameters:

1.35 GeV < m_c < 1.65 GeV, 4.3 GeV < m_b < 5 GeV, 3.5 GeV^2 < Q^2_{min} < 5 GeV^2

fraction of strange quarks in the sea f_s=0.31±0.08

CMS QCD ANALYSIS: PARAMETRISATION

Basic parametrization at the starting scale Q^2_0=1.9 GeV^2 (13+ free parameters):

$$
\begin{aligned}
xg(x) &= A_g x^{B_g} \cdot (1-x)^{C_g} - A'_g x^{B'_g} \cdot (1-x)^{C'_g}, \\
xu_v(x) &= A_{u_v} x^{B_{u_v}} \cdot (1-x)^{C_{u_v}} \cdot (1 + E_{u_v} x^2), \\
xd_v(x) &= A_{d_v} x^{B_{d_v}} \cdot (1-x)^{C_{d_v}}, \\
x\overline{U}(x) &= A_{\overline{U}} x^{B_{\overline{U}}} \cdot (1-x)^{C_{\overline{U}}}, \\
x\overline{D}(x) &= A_{\overline{D}} x^{B_{\overline{D}}} \cdot (1-x)^{C_{\overline{D}}}.
\end{aligned}
$$

$$
\begin{aligned}
x\,\overline{U} &= x\,\bar{u} \\
x\,\overline{D} &= x\,\bar{d} + x\,\bar{s} \\
B_{\overline{U}} &= B_{\overline{D}} \\
A_{\overline{U}} &= A_{\overline{D}}(1-f_s) \\
f_s &= \bar{s}/(\bar{d}+\bar{s}) \equiv 0.31 \pm 0.08
\end{aligned}
$$

Normalization parameters A are determined by QCD sum rules

B: define low-x behaviour, C: high-x shape

Parametrization uncertainties:

originate from variations on assumed parametrization, in which additional parameters are added one by-one in the functional form of the parametrization;
additional variation of 1. 5<Q^2_0<2.5 GeV^2

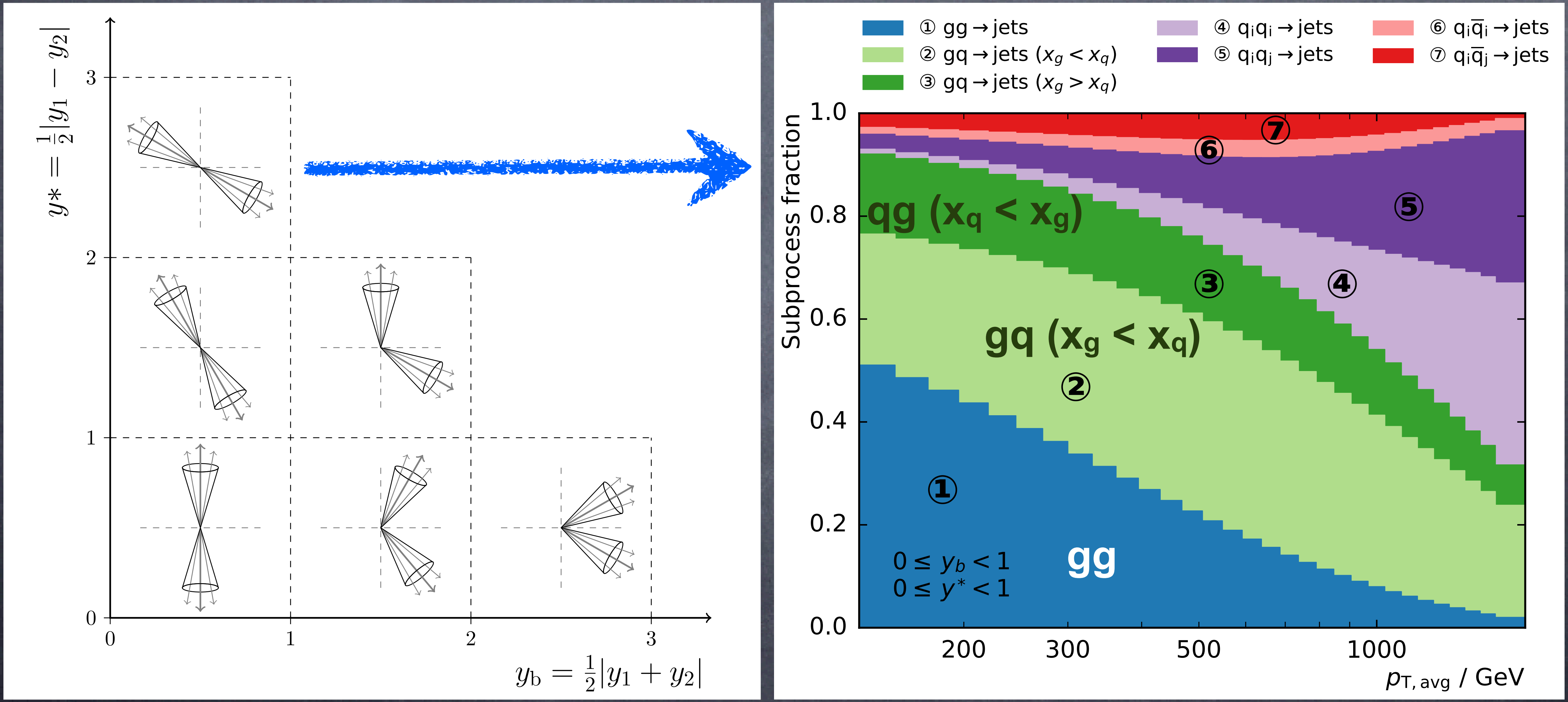
Largest difference of resulting PDFs to the central result (envelope) is assigned as uncertainty

JETS @ CMS: GLUON AND STRONG COUPLING

CMS 8 TeV, $\mathcal{L}$ = 19.7 fb^{-1} dijet production: CMS-PAS-SMP-16-011

3-differential cross sections vs of jet average p_T, rapidity separation and boost

Probing x_1 and x_2 using different event topologies

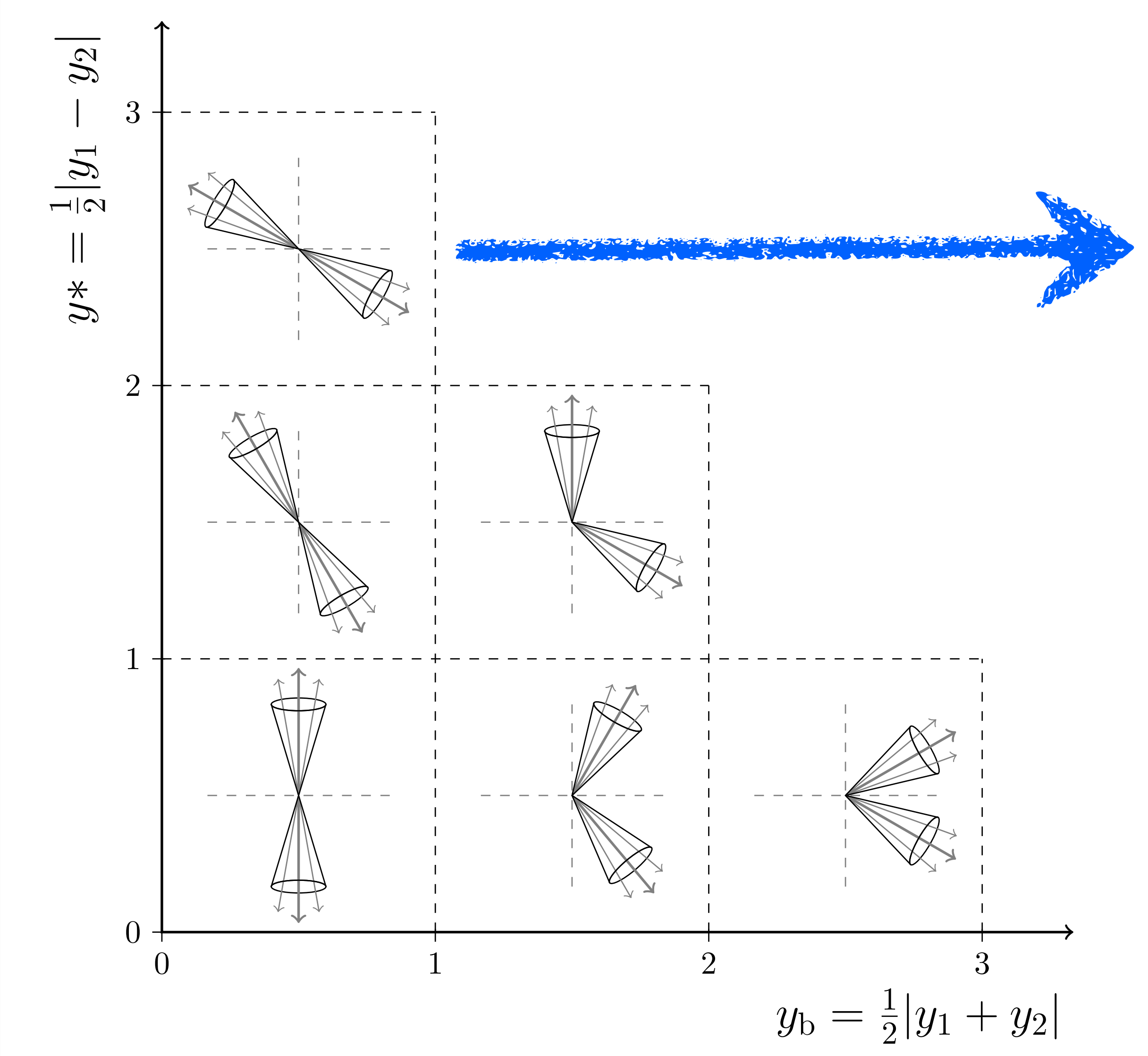


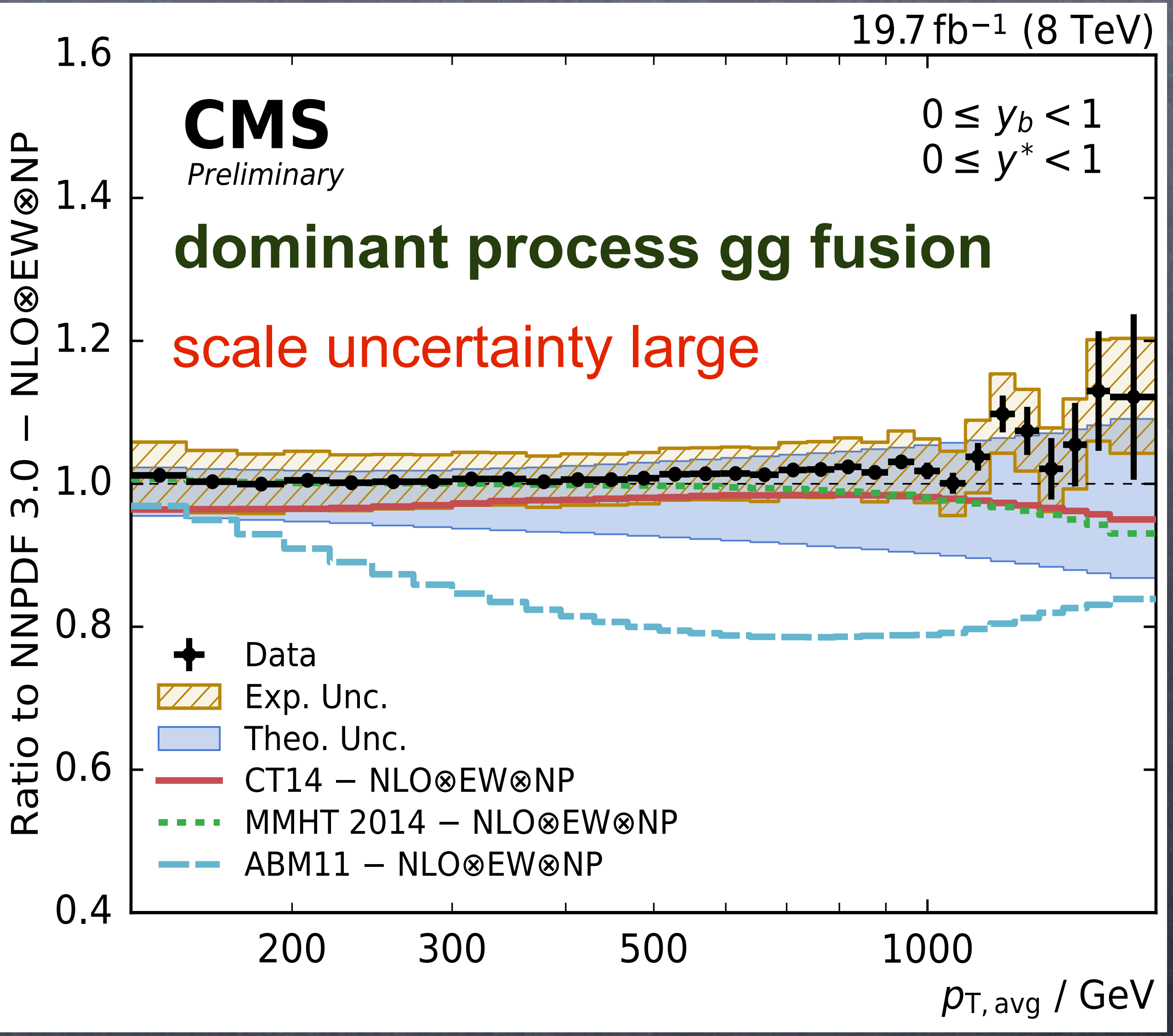
JETS @ CMS: GLUON AND STRONG COUPLING

CMS 8 TeV, $\mathcal{L}$ = 19.7 fb^{-1} dijet production: CMS-PAS-SMP-16-011

3-differential cross sections vs of jet average p_T, rapidity separation and boost

Probing x_1 and x_2 using different event topologies





JETS @ CMS: GLUON AND STRONG COUPLING

QCD analysis: XFitter 1.2. 2, baseline data HERA inclusive DIS [EPJ C 75 (2015) 580]

Theory via NLOJet++ via fastNLO, scale $\mu_r = \mu_f = p_{T,max} \cdot e^{0.3y^*}$

Q^2_0=1.9 GeV^2 :

$$xg(x) = A_g x^{B_g}(1-x)^{C_g} - A'_g x^{B'_g}(1-x)^{C'_g},$$
$$xu_v(x) = A_{u_v} x^{B_{u_v}}(1-x)^{C_{u_v}}(1 + D_{u_v}x + E_{u_v}x^2),$$
$$xd_v(x) = A_{d_v} x^{B_{d_v}}(1-x)^{C_{d_v}}(1 + D_{d_v}x),$$
$$x\overline{U}(x) = A_{\overline{U}} x^{B_{\overline{U}}}(1-x)^{C_{\overline{U}}}(1 + D_{\overline{U}}x),$$
$$x\overline{D}(x) = A_{\overline{D}} x^{B_{\overline{D}}}(1-x)^{C_{\overline{D}}},$$

$x\overline{U}(x) = x\bar{u}(x)$, and $x\overline{D}(x) = x\bar{d}(x) + x\bar{s}(x)$

$B_{\overline{U}} = B_{\overline{D}}$ and $A_{\overline{U}} = A_{\overline{D}}(1 - f_s)$

$Bd_v \neq Bu_v$

$\Rightarrow$ 16-parameter fit

Data are consistent very good fit quality for the CMS jet data

		HERA data		HERA & CMS data	
data set	n_{data}	χ^2_p	χ^2_p/n_{data}	χ^2_p	χ^2_p/n_{data}
NC HERA-I+II e^+p $E_p = 920$ GeV	332	382.44	1.15	406.45	1.22
NC HERA-I+II e^+p $E_p = 820$ GeV	63	60.62	0.96	61.01	0.97
NC HERA-I+II e^+p $E_p = 575$ GeV	234	196.40	0.84	197.56	0.84
NC HERA-I+II e^+p $E_p = 460$ GeV	187	204.42	1.09	205.50	1.10
NC HERA-I+II e^-p	159	217.27	1.37	219.17	1.38
CC HERA-I+II e^+p	39	43.26	1.11	42.29	1.08
CC HERA-I+II e^-p	42	49.11	1.17	55.35	1.32
CMS Triple-Differential Dijets	122	—	—	111.13	0.91
data set(s)	n_{dof}	χ^2	χ^2/n_{dof}	χ^2	χ^2/n_{dof}
HERA data	1040	1211.00	1.16	—	—
HERA & CMS data	1162	—	—	1372.52	1.18

$t\bar{t}$ @ CMS: GLUON DISTRIBUTION AT HIGH X

QCD analysis: XFitter 1.2.2,

baseline data: HERA inclusive DIS [EPJ C75 (2015) 580], CMS W± [EPJ C76 (2016) 469]

Theory for $t\bar{t}$ MCFM via ApplGrid, scales $\mu_{r,f} = \sqrt{m_t^2 + [p_T(t)^2 + p_T(\bar{t})^2]/2}$

Q^2_0=1.9 GeV^2 :

$$xg(x) = A_g x^{B_g} (1-x)^{C_g} (1 + E_g x^2 + F_g x^3) - A'_g x^{B'_g} (1-x)^{C'_g},$$
$$xu_v(x) = A_{u_v} x^{B_{u_v}} (1-x)^{C_{u_v}} (1 + D_{u_v} x + E_{u_v} x^2),$$
$$xd_v(x) = A_{d_v} x^{B_{d_v}} (1-x)^{C_{d_v}},$$
$$x\overline{U}(x) = A_{\overline{U}} x^{B_{\overline{U}}} (1-x)^{C_{\overline{U}}} (1 + D_{\overline{U}} x + F_{\overline{U}} x^3),$$
$$x\overline{D}(x) = A_{\overline{D}} x^{B_{\overline{D}}} (1-x)^{C_{\overline{D}}},$$

$x\overline{U}(x) = x\bar{u}(x)$, and $x\overline{D}(x) = x\bar{d}(x) + x\bar{s}(x)$

$B_{\overline{U}} = B_{\overline{D}}$ and $A_{\overline{U}} = A_{\overline{D}}(1 - f_s)$

⇒ 18-parameter fit

Data are consistent very good fit quality for the CMS data

Data sets	χ^2/dof			
	Nominal fit	$+[p_T(t), y(t)]$	$+[y(t), M(t\bar{t})]$	$+[y(t\bar{t}), M(t\bar{t})]$
CMS double-differential $t\bar{t}$		10/15	7.4/15	7.6/15
HERA CC e^-p, $E_p = 920$ GeV	57/42	56/42	56/42	57/42
HERA CC e^+p, $E_p = 920$ GeV	44/39	44/39	44/39	43/39
HERA NC e^-p, $E_p = 920$ GeV	219/159	219/159	219/159	218/159
HERA NC e^+p, $E_p = 920$ GeV	440/377	437/377	439/377	441/377
HERA NC e^+p, $E_p = 820$ GeV	69/70	68/70	68/70	69/70
HERA NC e^+p, $E_p = 575$ GeV	221/254	220/254	221/254	221/254
HERA NC e^+p, $E_p = 460$ GeV	219/204	219/204	219/204	219/204
CMS $W^\pm$ asymmetry	4.7/11	4.6/11	4.8/11	4.9/11
Correlated χ^2	82	87	91	89
Log-penalty χ^2	−2.5	+2.6	−2.2	−3.3
Total χ^2/dof	1352/1138	1368/1153	1368/1153	1366/1153