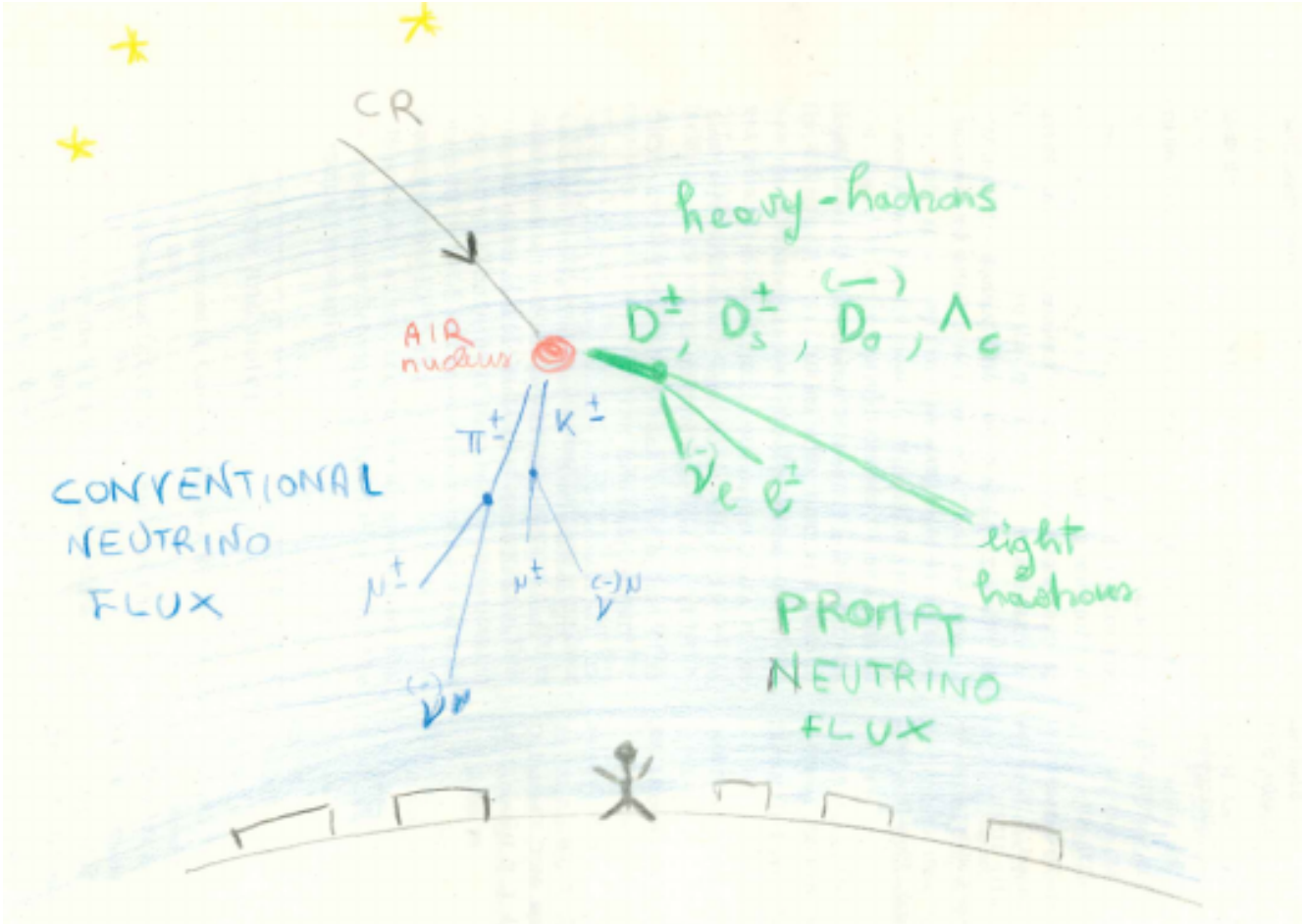
Constraints from heavy-quark production on proton PDFs

Oleksandr Zenaiev
(CERN)

Mainz
8.10.2019

Atmospheric ν **are background for** *astrophysical* ν**:**



Production of atmospheric ν:

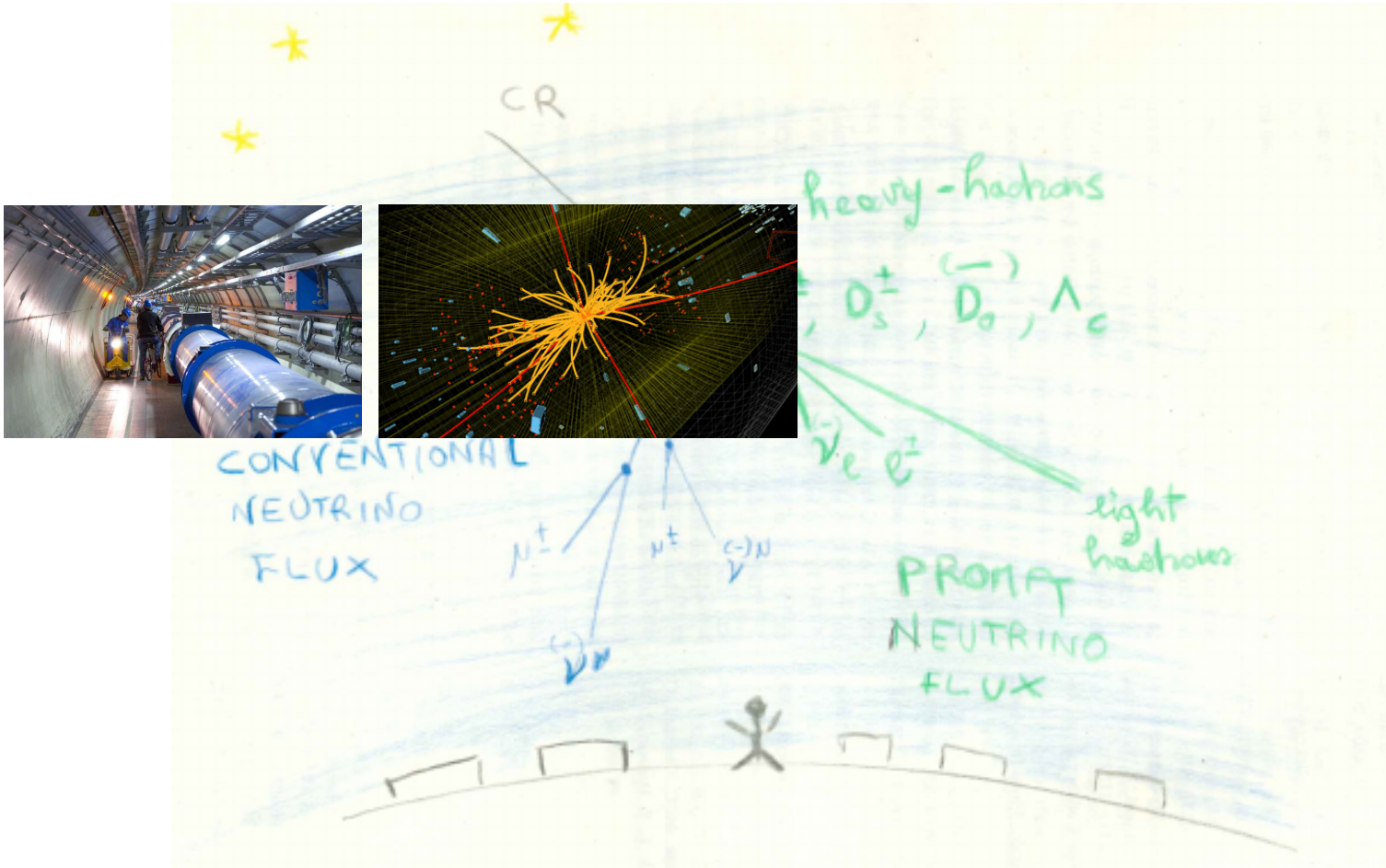
- cosmic rays (CR) + atmospheric nuclei
  $\rightarrow$ light and heavy hadrons $\rightarrow$ conventional and **prompt** ν fluxes
- spectra of conventional and **prompt** ν fluxes are different because of different hadroproduction cross sections and decay properties

Atmospheric ν **are background for** *astrophysical* ν**:**

Production of atmospheric ν:

- cosmic rays (CR) + atmospheric nuclei
  $\rightarrow$ light and heavy hadrons $\rightarrow$ conventional and **prompt** ν fluxes
- spectra of conventional and **prompt** ν fluxes are different because of different hadroproduction cross sections and decay properties

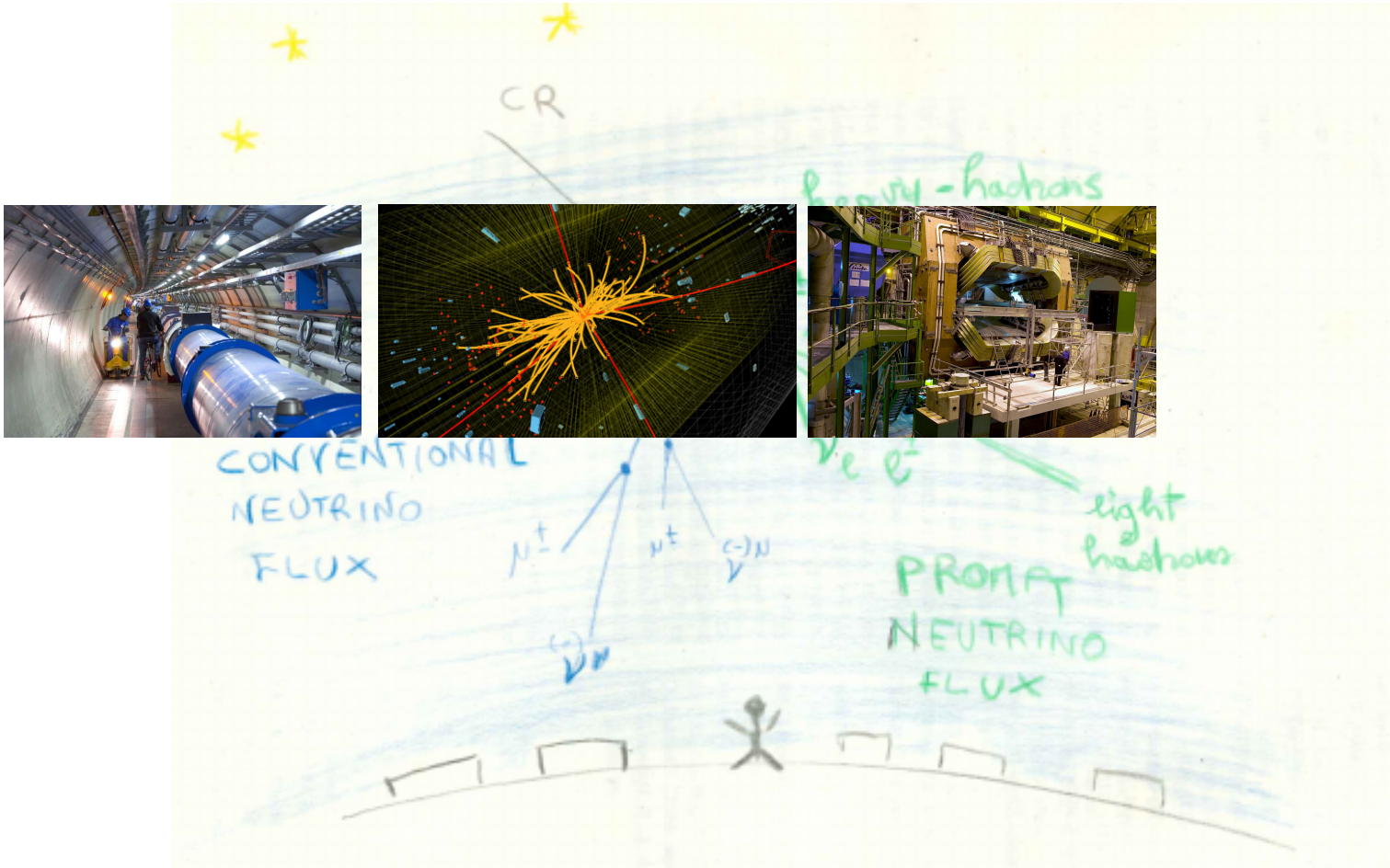
Atmospheric ν **are background for** *astrophysical* ν**:**



Production of atmospheric ν:

- cosmic rays (CR) + atmospheric nuclei
  $\rightarrow$ light and heavy hadrons $\rightarrow$ conventional and **prompt** ν fluxes
- spectra of conventional and **prompt** ν fluxes are different because of different hadroproduction cross sections and decay properties

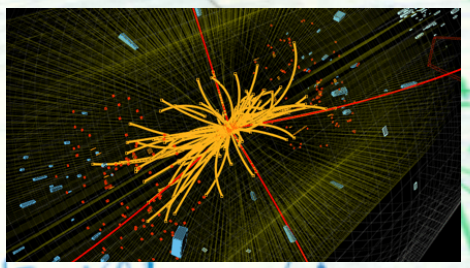
Atmospheric ν **are background for** *astrophysical* ν**:**

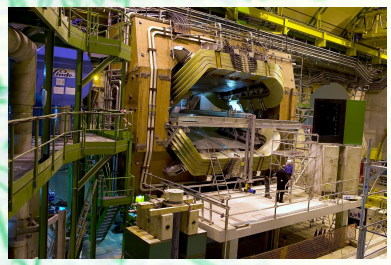


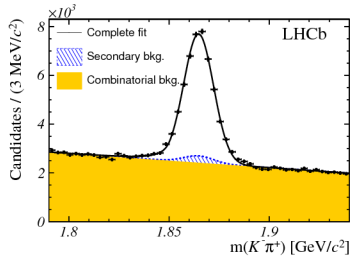
Production of atmospheric ν:

- cosmic rays (CR) + atmospheric nuclei
  $\rightarrow$ light and heavy hadrons $\rightarrow$ conventional and **prompt** ν fluxes
- spectra of conventional and **prompt** ν fluxes are different because of different hadroproduction cross sections and decay properties

Atmospheric ν **are background for** *astrophysical* ν**:**







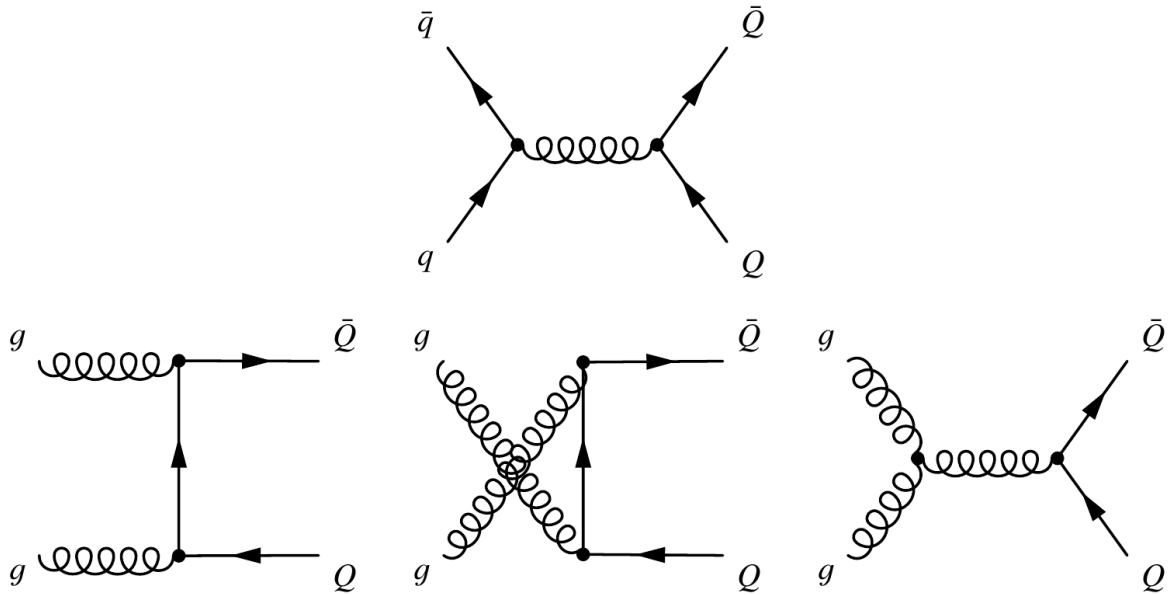
Production of atmospheric ν:

- cosmic rays (CR) + atmospheric nuclei
  $\rightarrow$ light and heavy hadrons $\rightarrow$ conventional and **prompt** ν fluxes
- spectra of conventional and **prompt** ν fluxes are different because of different hadroproduction cross sections and decay properties

Theoretical description of charm hadroproduction

Ingredients:

- parton distribution functions (PDFs)
- hard-scattering perturbative partonic cross sections
- fragmentation functions / parton shower + hadronization



$$\sigma_{N_1 N_2 \to H+X} = \Sigma_{ab} PDF_a^{N_1}(x_a, \mu_F) PDF_b^{N_2}(x_b, \mu_F) \otimes$$
$$\otimes \hat{\sigma}_{ab \to cX}(x_a, x_b, z, \mu_F, \mu_R, \alpha_S(\mu_R), m_c) \quad (+\text{fragmentation})$$

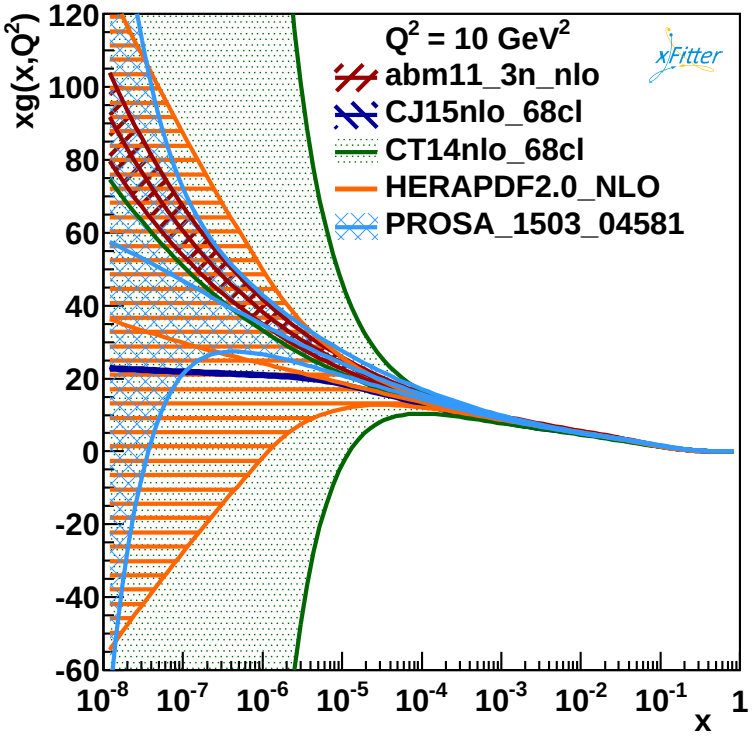
Theoretical description of charm hadroproduction

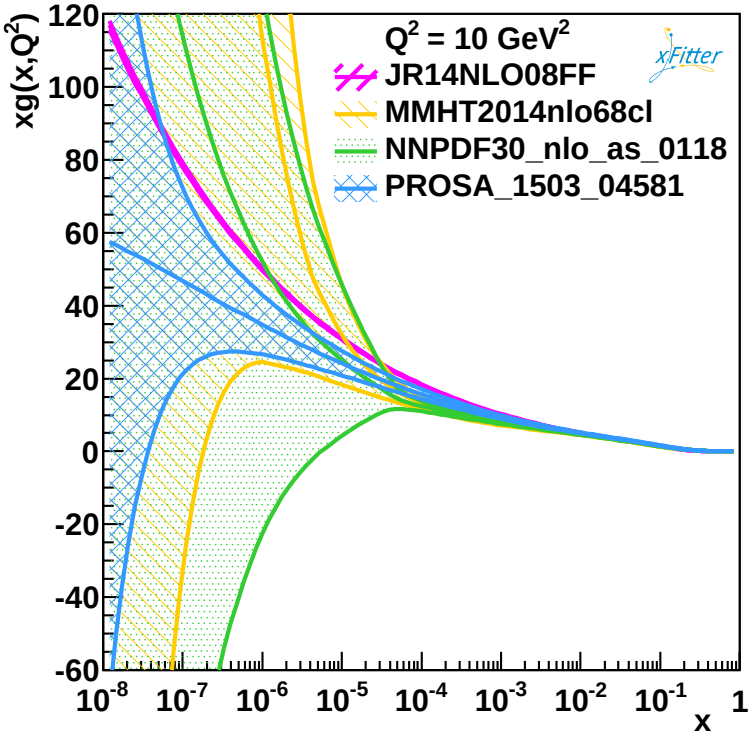
$$\sigma_{N_1 N_2 \to H+X} = \Sigma_{ab} PDF_a^{N_1}(x_a, \mu_F) PDF_b^{N_2}(x_b, \mu_F) \otimes$$
$$\otimes \hat{\sigma}_{ab \to cX}(x_a, x_b, z, \mu_F, \mu_R, \alpha_S(\mu_R), m_c) \quad (+\text{fragmentation})$$

Uncertainties in computation:

- proton/nuclear PDF uncertainties at small parton memonetum fraction x (can be very large for high $E \longleftrightarrow$ small parton x)
  $\to$ *main topic of this talk*
- perturbative hadroproduction cross section: NLO is state-of-the-art fixed-order prediction for differential charm production cross sections
  $\to$ various ways of treating heavy quark mass, fixed-flavour-number scheme (FFNS) vs. general-mass variable-flavour-number scheme (GM-VFNS)
- fragmentation $c \to D, \Lambda$

Gluon PDFs

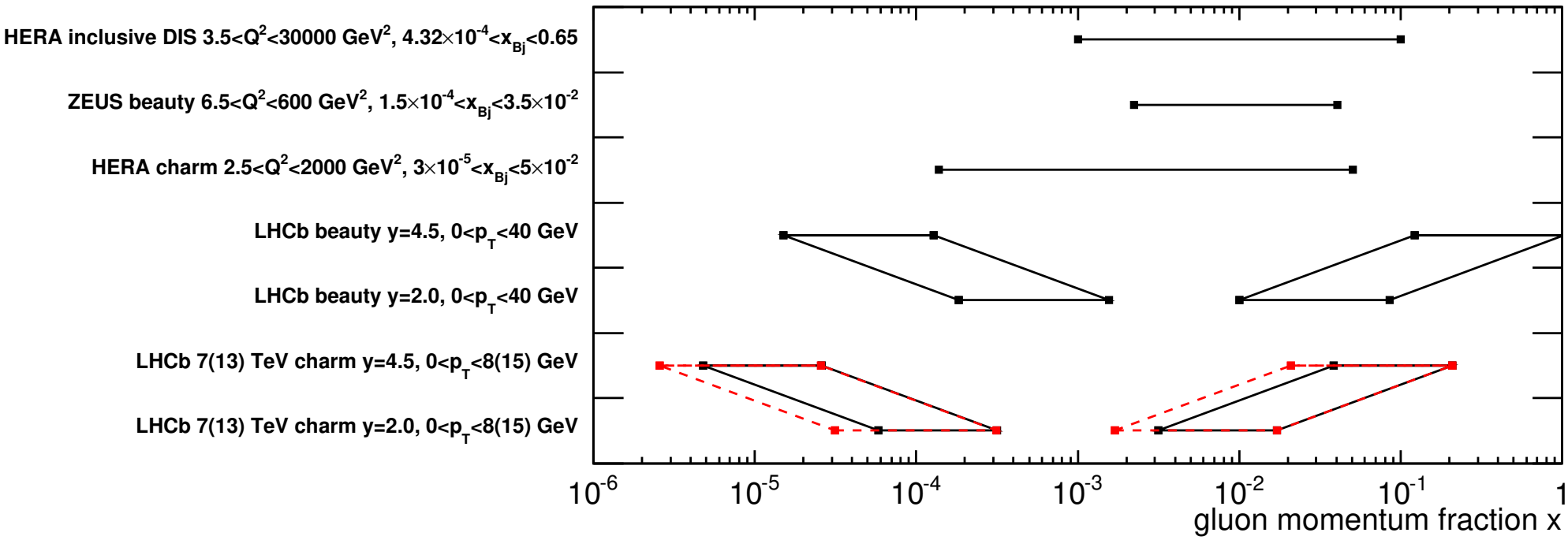




[1611.03815]

- PDFs are exctracted from fits to experimental data
- Production of prompt neutrinos requires as input gluon PDF at low x, previously not constrained by any data
  - depends on additional assumption (e.g. PDF parametrisation)
  - *gluon PDF uncertainties can be arbitrary large at low x!*
- PROSA fit was the first QCD analysis of heavy-flavour (HF) data from LHCb, which extended coverage of gluon PDF to $x \gtrsim 10^{-6}$ [1503.04581]
- Similar studies were later presented by R. Gauld et al. [1506.08025, 1610.09373]

Kinematics of HF hadroproduction at LHCb



- LHCb measured double-diff. $\frac{\mathrm{d}^2\sigma}{\mathrm{d}p_T\mathrm{d}y}$ for c and b at 7 TeV, $2.0 < y < 4.5$:
  - charm: $0 < p_T < 8$ GeV [NPB871 (2013) 1]
  - beauty: $0 < p_T < 40$ GeV [JHEP08 (2013) 117]
- At LO: $\boldsymbol{x_{1,2}} = \frac{\sqrt{\boldsymbol{p_T^2}+\boldsymbol{m_Q^2}}}{\boldsymbol{E_P}}\mathrm{e}^{\pm \boldsymbol{y}}$
- c @ 7 TeV data: $x \gtrsim \frac{1.4}{3500}\mathrm{e}^{-4.5} \approx 4 \times 10^{-6}$ (uncovered by HERA data)
- c @ 13 TeV: even lower x, however not in PROSA 2015 fit
- b @ 7 TeV: $x \lesssim 1$, however not yet studied carefully

PROSA PDF fit

PROSA fit done using **xFitter** (former HERAFitter): open source framework for QCD+EW analyses
www.xfitter.org

LHCb data fitted together with HERA data (basics of any PDF fit)

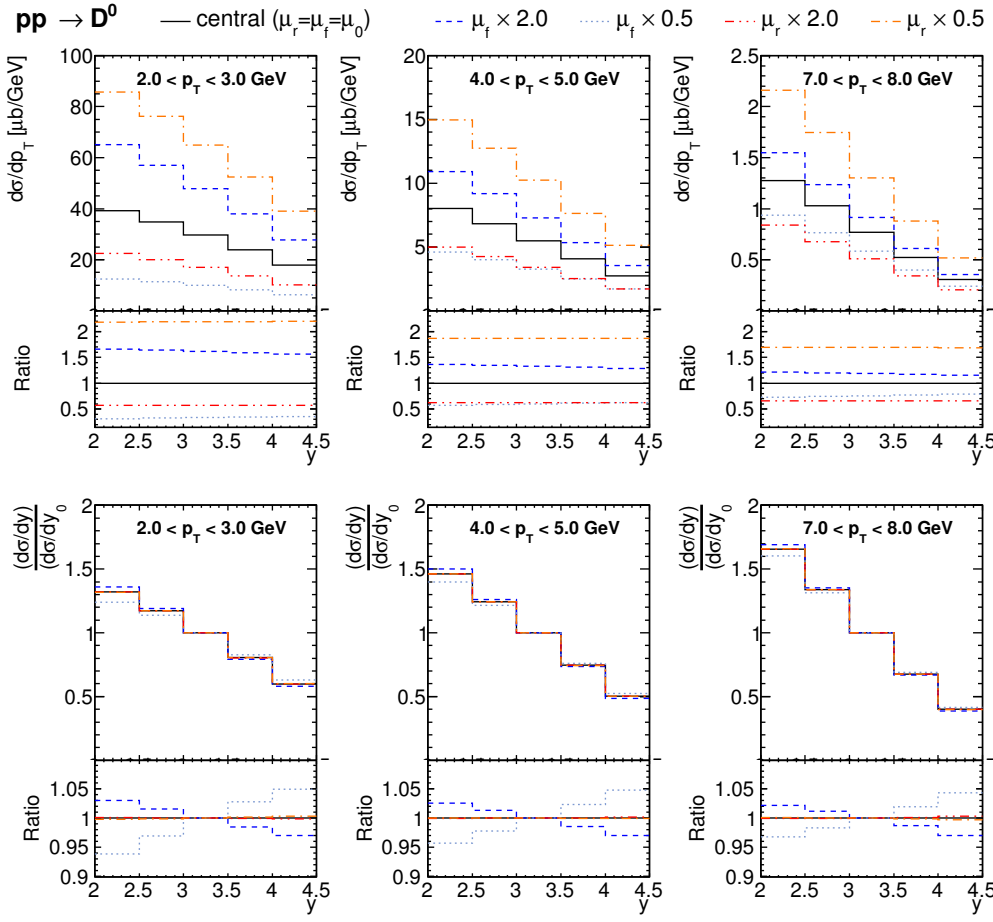
NLO predictions for LHCb data using Mangano-Nason-Ridolfi (MNR) calculations [NPB327 (1989) 49]] (implemented and now avail. in xFitter)

Setup followed NLO HERAPDF fits

Important ingredient was **normalisation of LHCb data**:

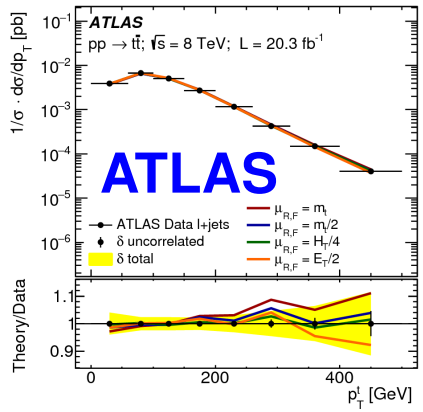
- only *y* shape used in the fit (correlated with *x* in PDFs)
- assumption: i.e. scale variations (i.e. missing higher-order corrections) do not depend on y
  - → y is determined mostly by the boost of parton-parton frame
- **order of magnitude** reduction of NLO scale uncertainties

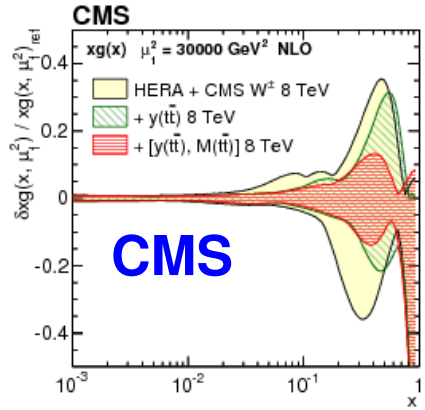
PROSA [1503.04581]

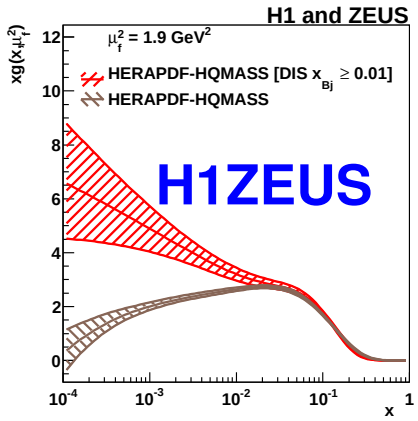


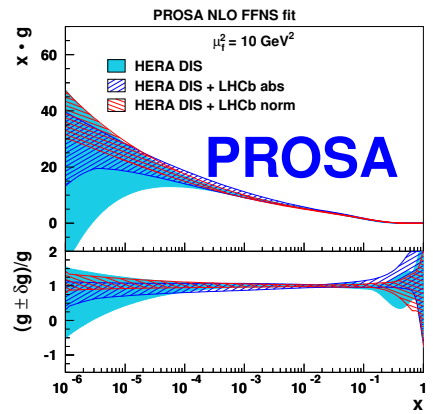
xFitter [xfitter.org]

- xFitter (HERAfitter before 2016) is open source QCD fit framework ready to extract PDFs, theory parameters, and assess impact of new data
  - various PDF parameterisations, theoretical processes and schemes, and data sets (> 50)
  - $\rightarrow$ unique such tool
- $\sim$ **80** publications using xFitter (widely used by LHC community)
- **9** publications by xFitter developers team
- tutorials to get familiar with xFitter
  - presented at CMS Data Analysis School (DESY, September 2018)
- latest stable release xFitter-2.0.1 OldFashioned
  - give it a try `www.xfitter.org`







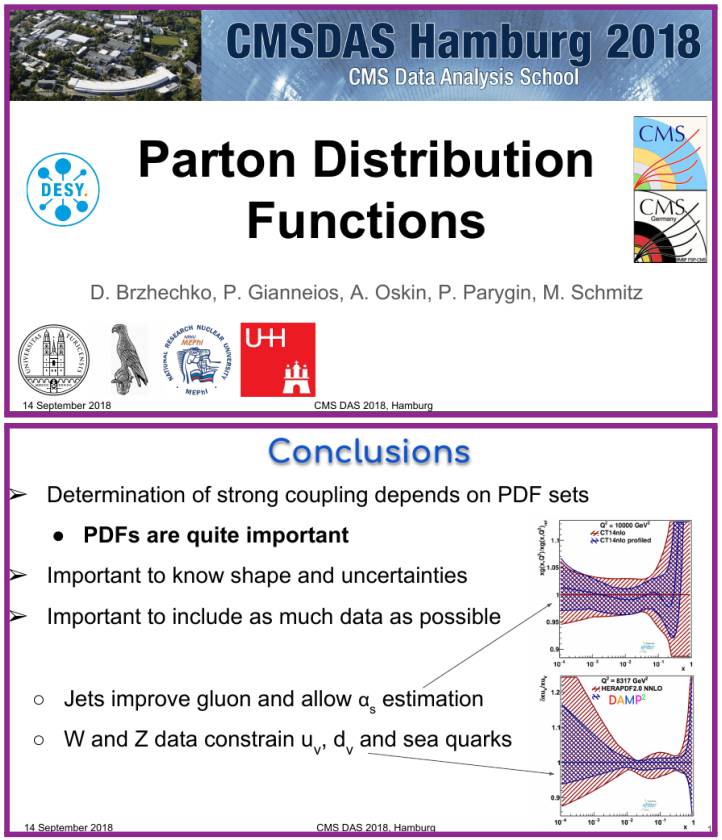


xFitter [xfitter.org]

- xFitter (HERAfitter before 2016) is open source QCD fit framework ready to extract PDFs, theory parameters, and assess impact of new data
  - various PDF parameterisations, theoretical processes and schemes, and data sets (> 50)

  $\rightarrow$ unique such tool
- $\sim$ **80** publications using xFitter (widely used by LHC community)
- **9** publications by xFitter developers team
- tutorials to get familiar with xFitter
  - presented at CMS Data Analysis School (DESY, September 2018)
- latest stable release xFitter-2.0.1 OldFashioned
  - give it a try `www.xfitter.org`

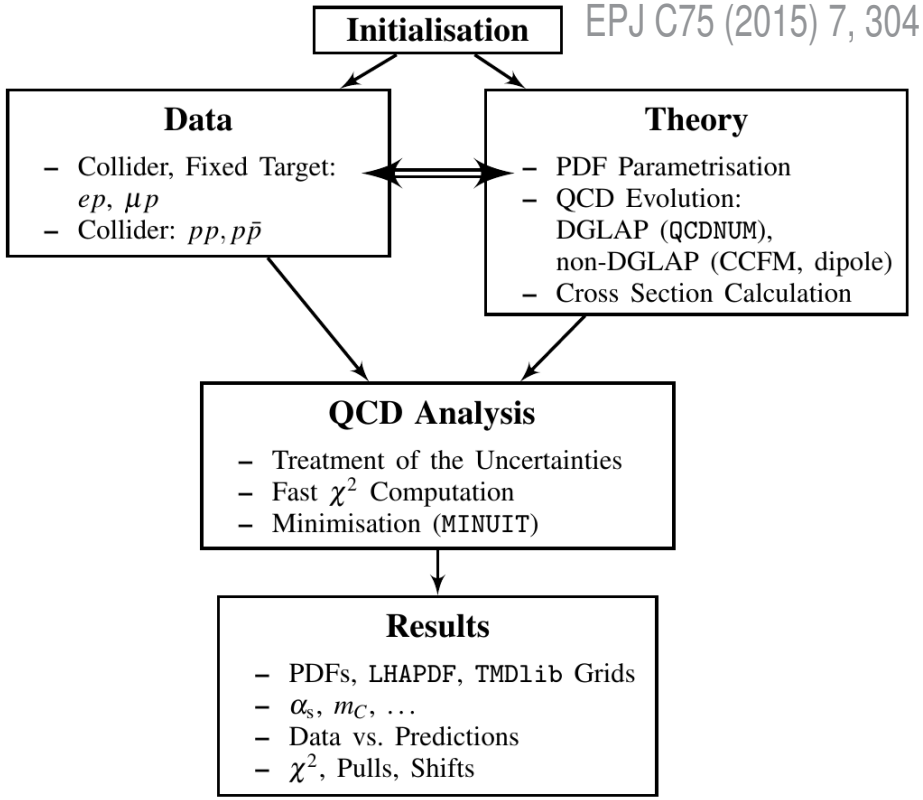
New "PDF set" produced in 2 days using xFitter:

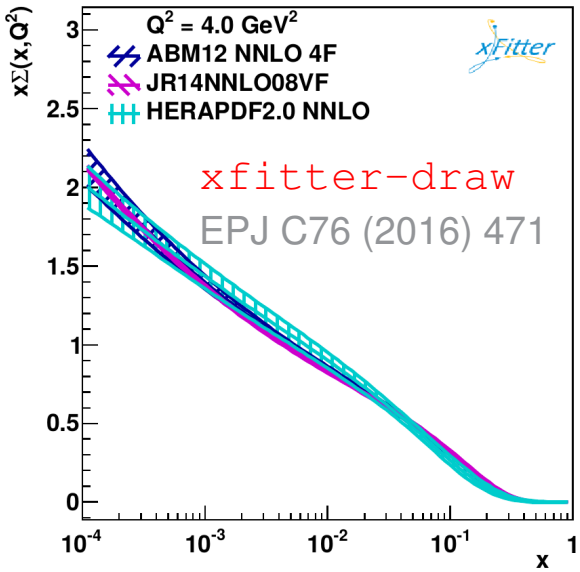


xFitter workflow

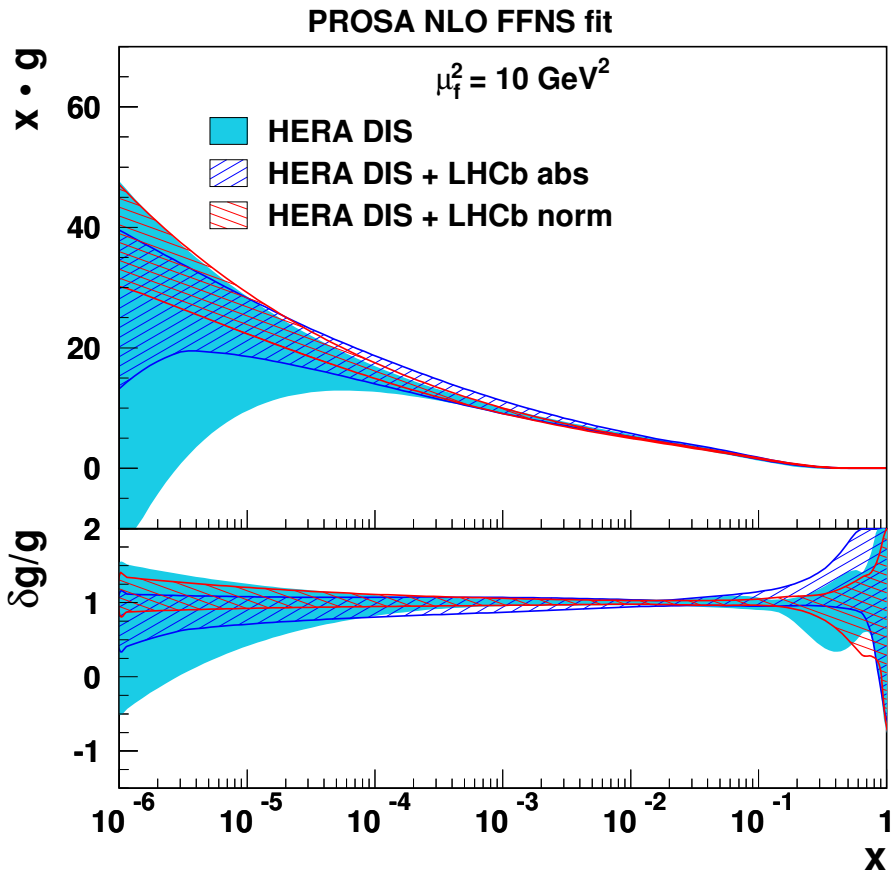
Main ingredients of a QCD analysis:

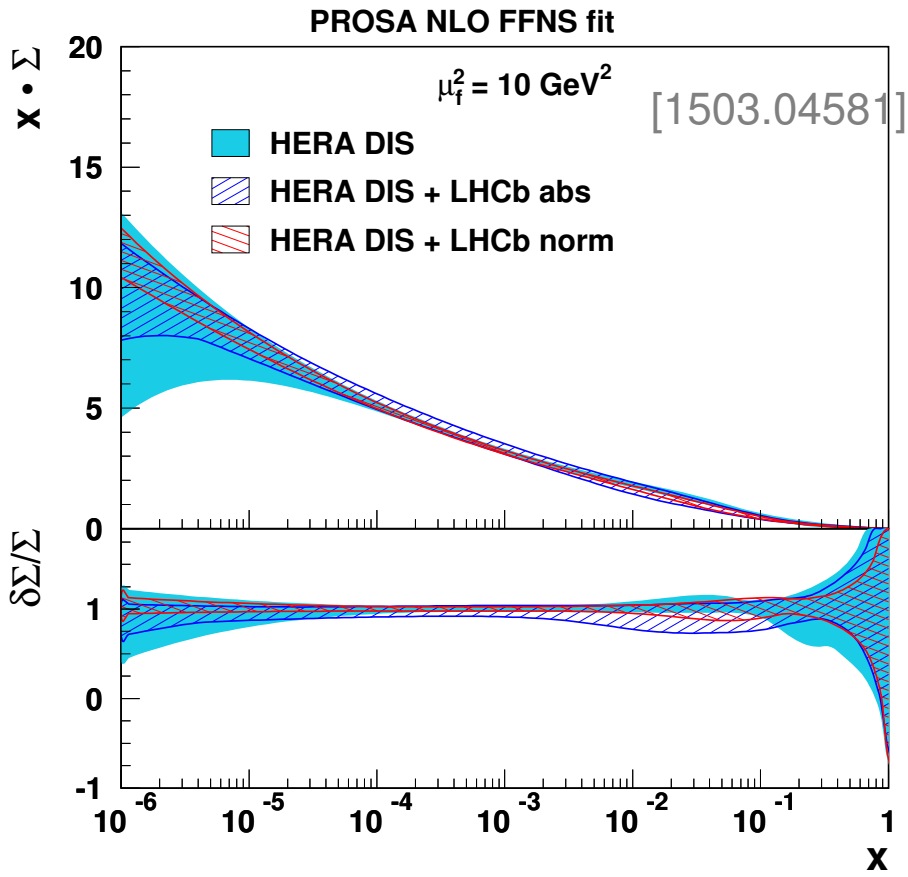
- Parametrise PDF at initial scale
  - polynomial, Chebyshev etc.
  - under development: fully flexible modular PDF decomposition and parametrisation
- Evolve to the scale of the process:
  - DGLAP (QCDNUM, APFEL) + QED
  - non-DGLAP (CCFM, dipole)
- Calculate hard scattering:
  - for deep inelastic scattering: various heavy-flavour schemes
  - fast techniques for N(N)LO calculations: fastNLO, ApplGrid, APFELgrid
  - external packages (Hathor etc.)
- Calculate χ^2 w.r.t data:
  - account for correlations
  - various treatment of uncertainties
- Analyse/minimise χ^2:
  - simple calculation of theoretical predictions
  - χ^2 minimisation using MINUIT (under development: interfacing more minimisers)
  - reweighting/profiling
- Plotting tools





PROSA PDF fit

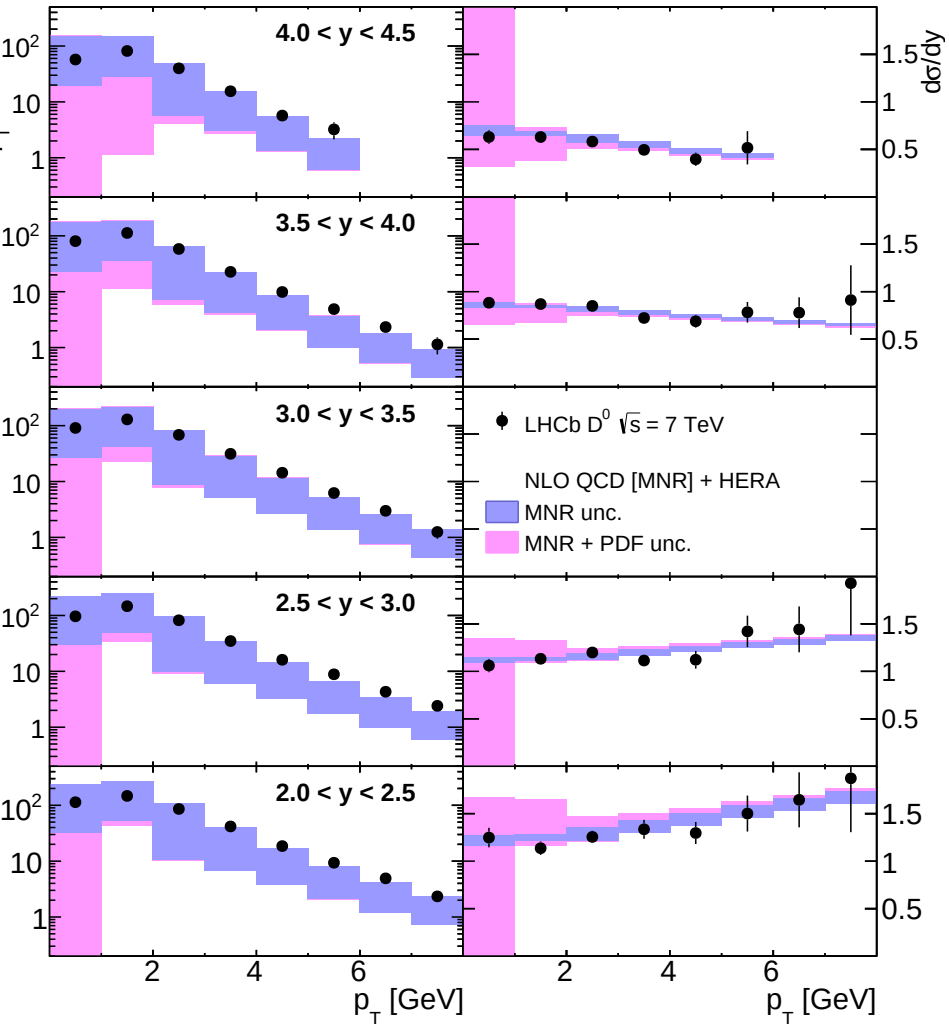




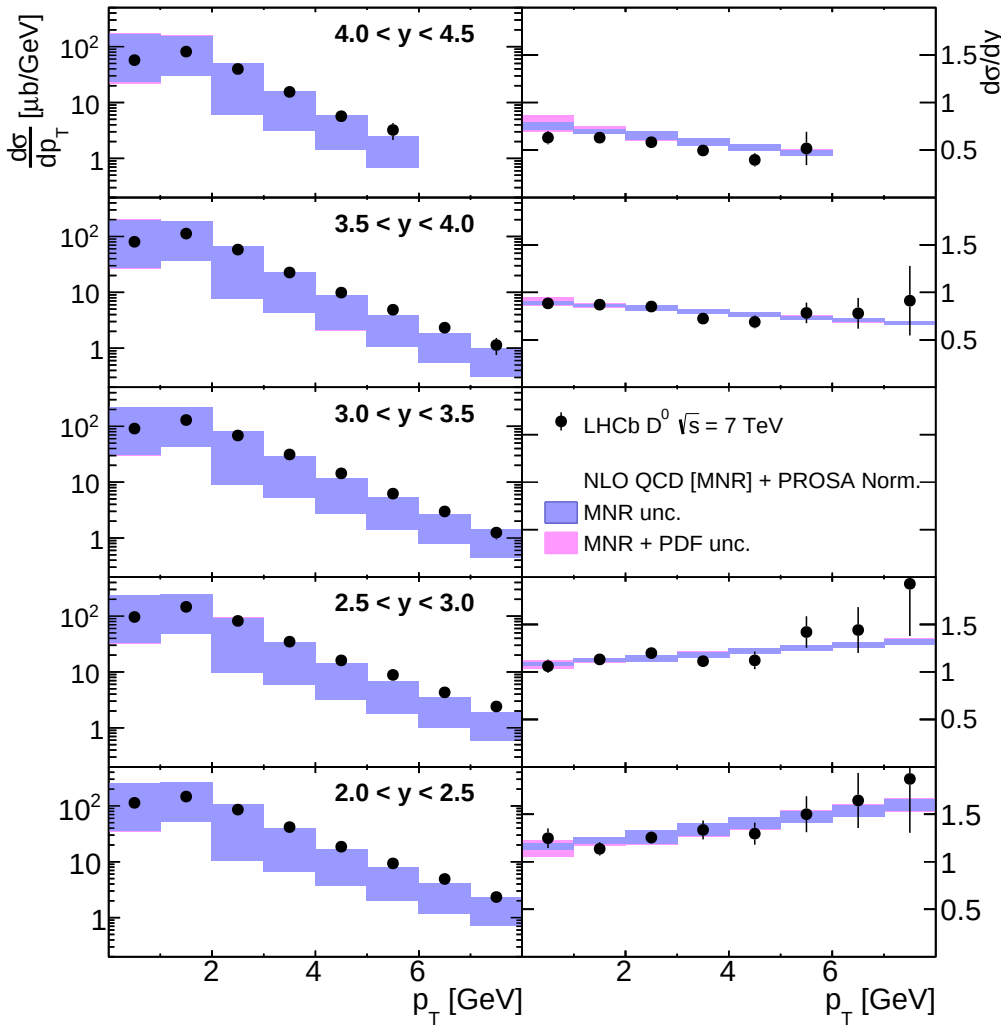
- Significant reduction of gluon and sea-quark [$\Sigma = 2(\bar{u} + \bar{d} + \bar{s})$] unc. at $x < 10^{-4}$ in fits with LHCb data (owing to correlation between g and Σ)
- Smaller PDF uncertainties when using **normalised** LHCb cross sections
- These PDFs are very useful for astrophysical ν calculations because of directly constrained low-x gluon distribution
- *But do PDFs determined using normalised HF cross sections still describe absolute HF cross sections?*

PROSA PDF fit: description of 7 TeV LHCb data

HERA only PDFs



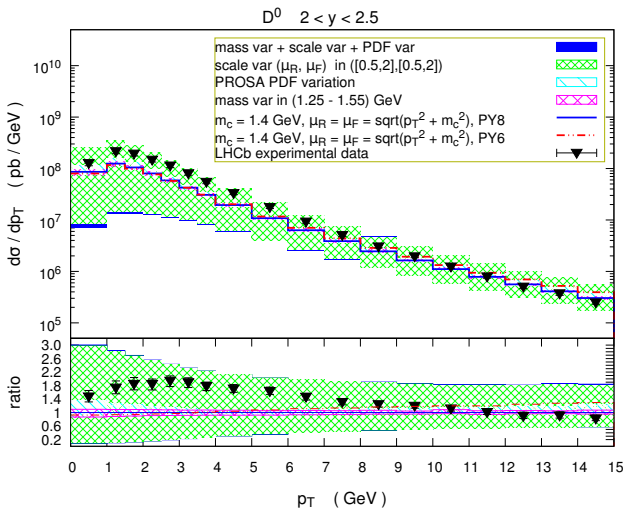
HERA + LHCb (normalised) PDFs

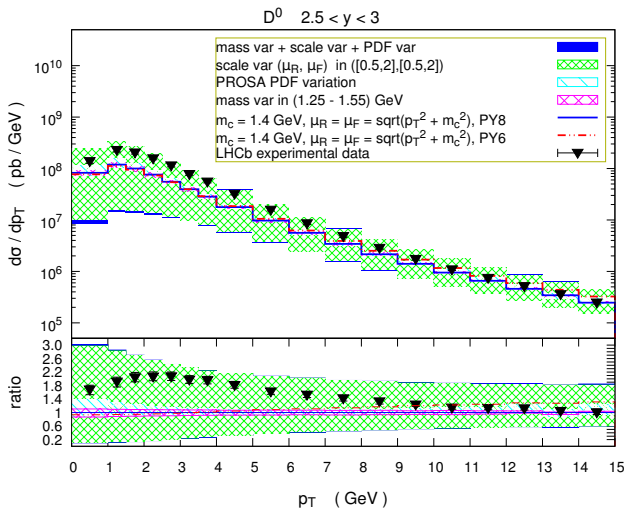


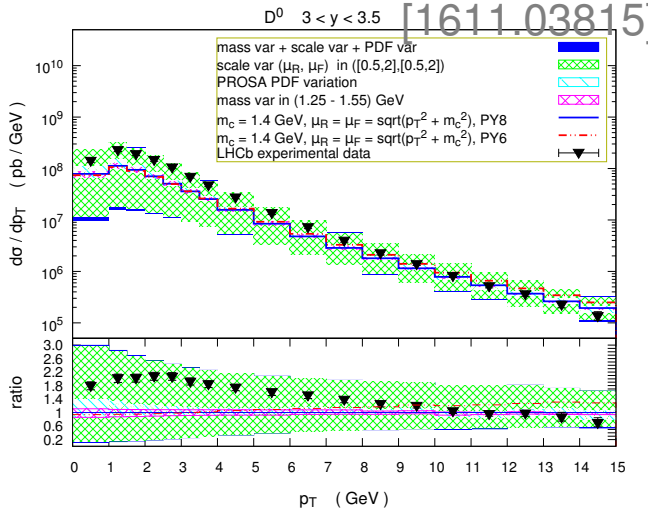
[1612.02371]

- Absolute cross sections are well described within large NLO scale unc.

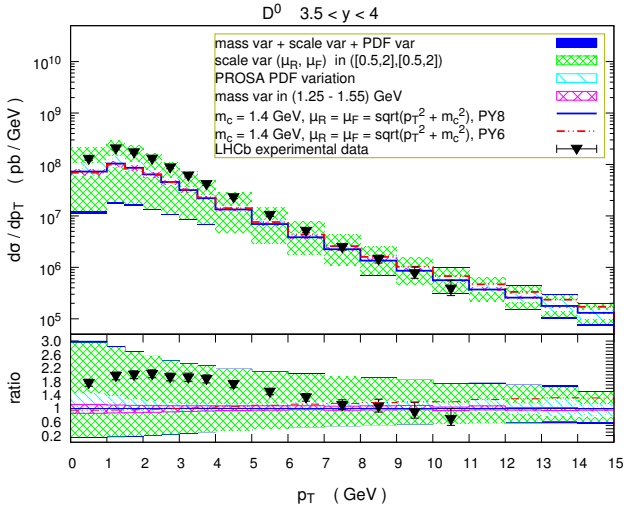
PROSA PDFs vs. 13 TeV LHCb data

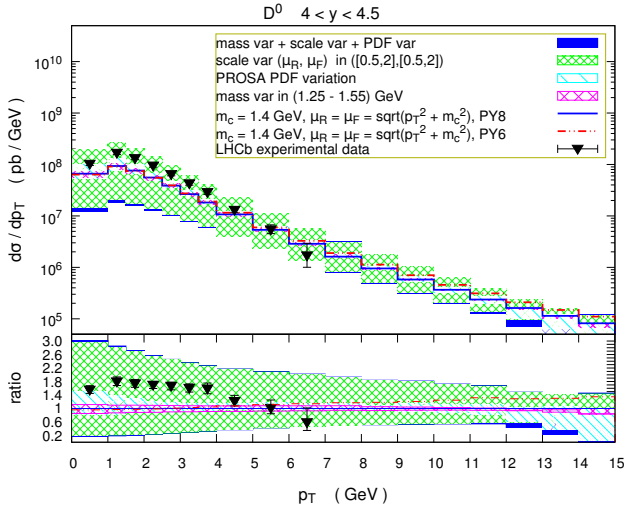






[1611.03815]





- PROSA PDFs, NLO + PS (POWHEGBOX + PYTHIA8)
- These data [JHEP 1603 (2016) 159 + errata] not used in PDF fit
- Good description within uncertainties

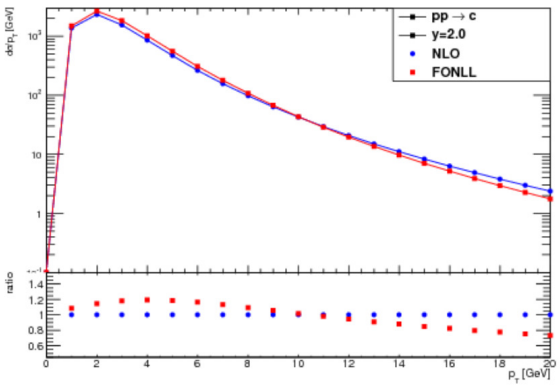
Theoretical uncertainties dominated by NLO scale $\gg$ experimental ones:

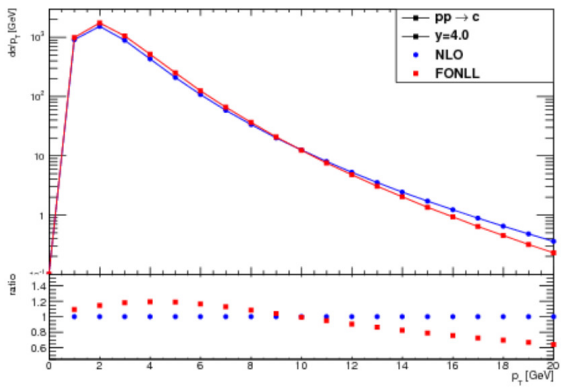
- limit further interpretation of these data (e.g. m_c determination)
- limit testing power: different modifications in theory input can lead to a “good agreement” with data
- **NNLO wanted!** (recently appeared for differential $t\bar{t}$)

NLO vs. FONLL

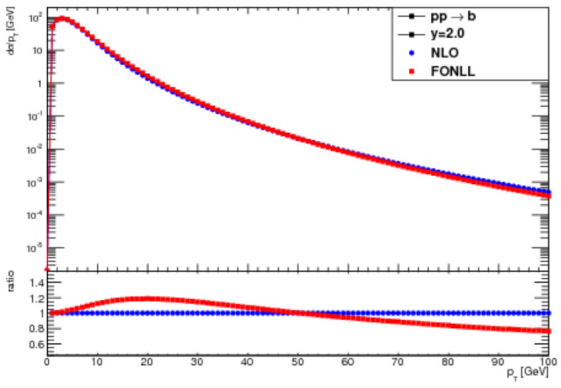
[DESY-THESIS-2015-012]

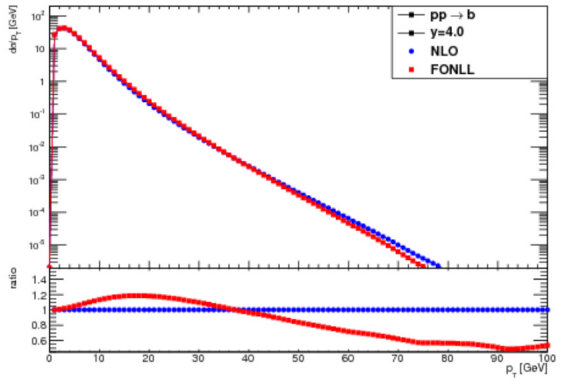
Charm:



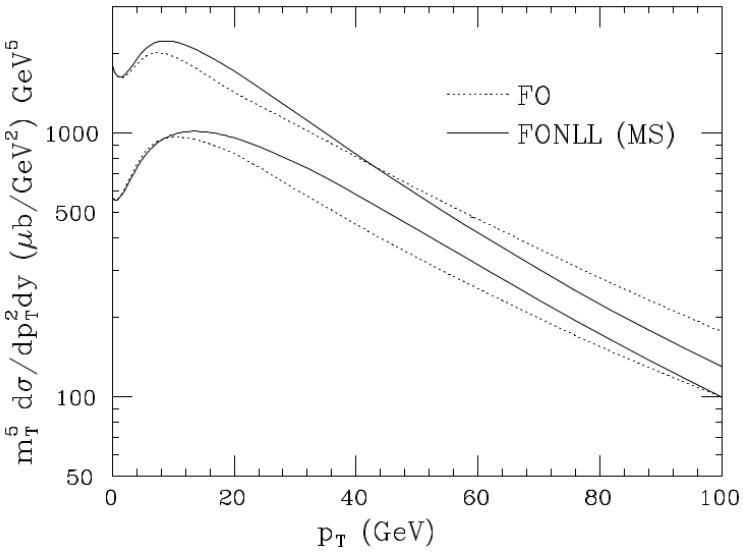


Beauty:





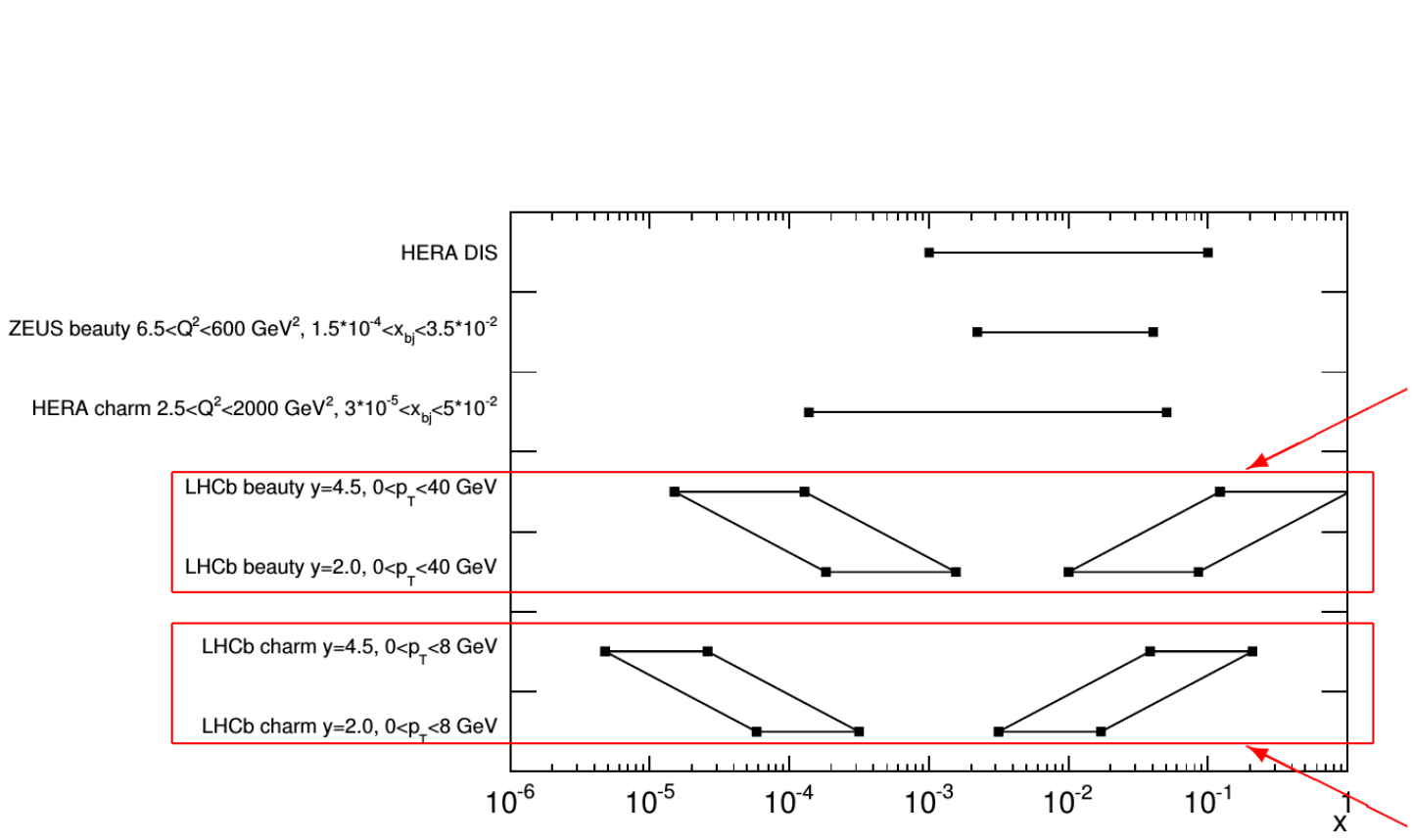
In LHCb kinematic region difference $<20\% \Rightarrow$ very well within scales uncertainties (NLL correction is small)

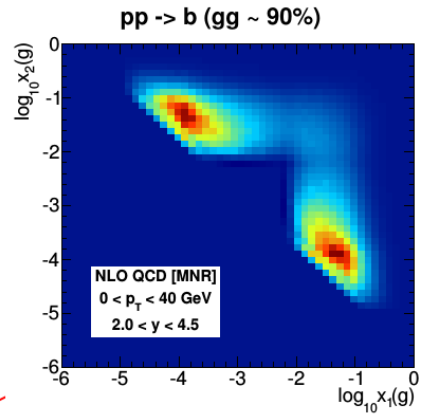


Cacciari et all., JHEP05 1998 007: "...In summary, we observed that our resummation procedure indicates the presence of a small enhancement in the intermediate p_T region, followed by a reduction of the cross section (and of the uncertainty band) at larger p_T..."

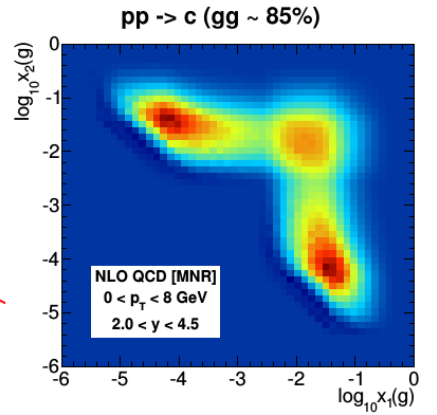
Kinematics: gluon x ranges

[DESY-THESIS-2015-012]





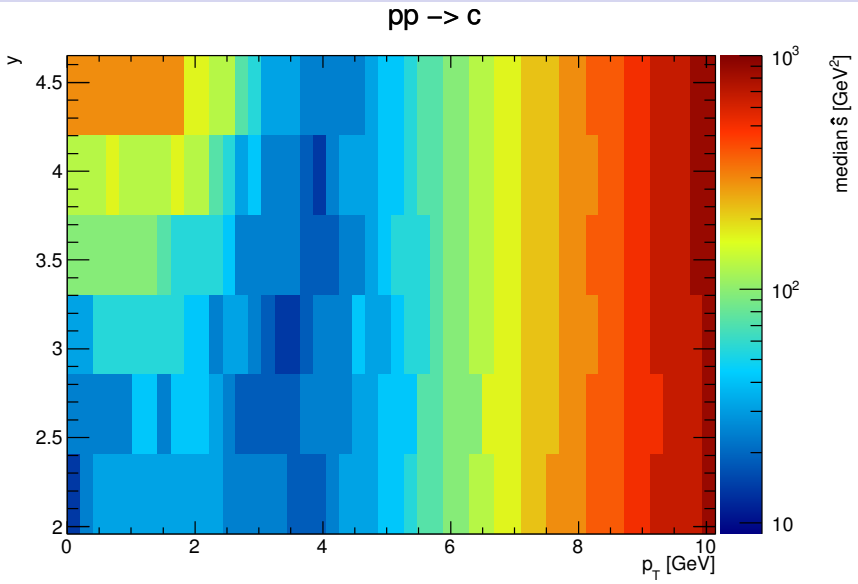
$(Q^2 \sim 20..2000\ GeV^2)$

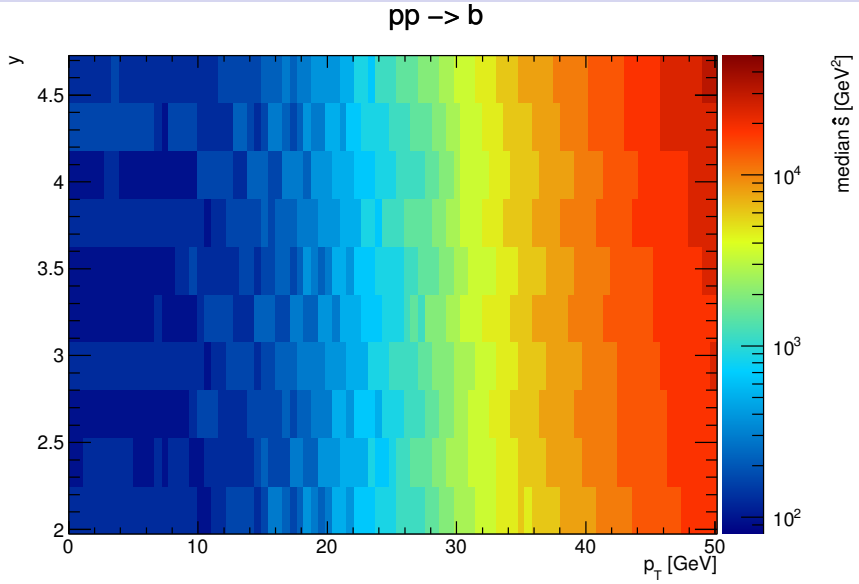


- LHCb data cover $x \sim 10^{-5}..10^{-3}$ (*small x*) and $x \sim 10^{-3}..10^{-1}$ (*medium x*)
- *Medium x* covered by HERA data $\Rightarrow$ expect improvement at *small x*

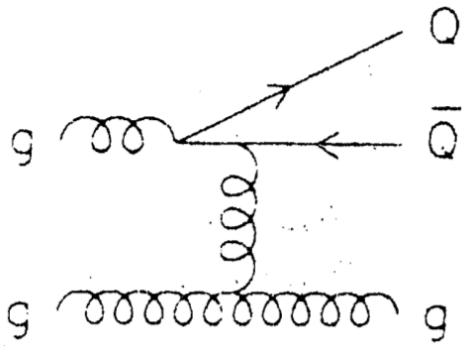
Kinematics: low p_T region

[DESY-THESIS-2015-012]





For low p_T charm ($0.0 \leq p_T \leq 2.5$) significant contribution ($\sim$ 50%) comes from $\hat{s}$ $10^2..10^5$ GeV^2: not dominated by low x g. Effect increases with increasing y. No such trend for beauty.



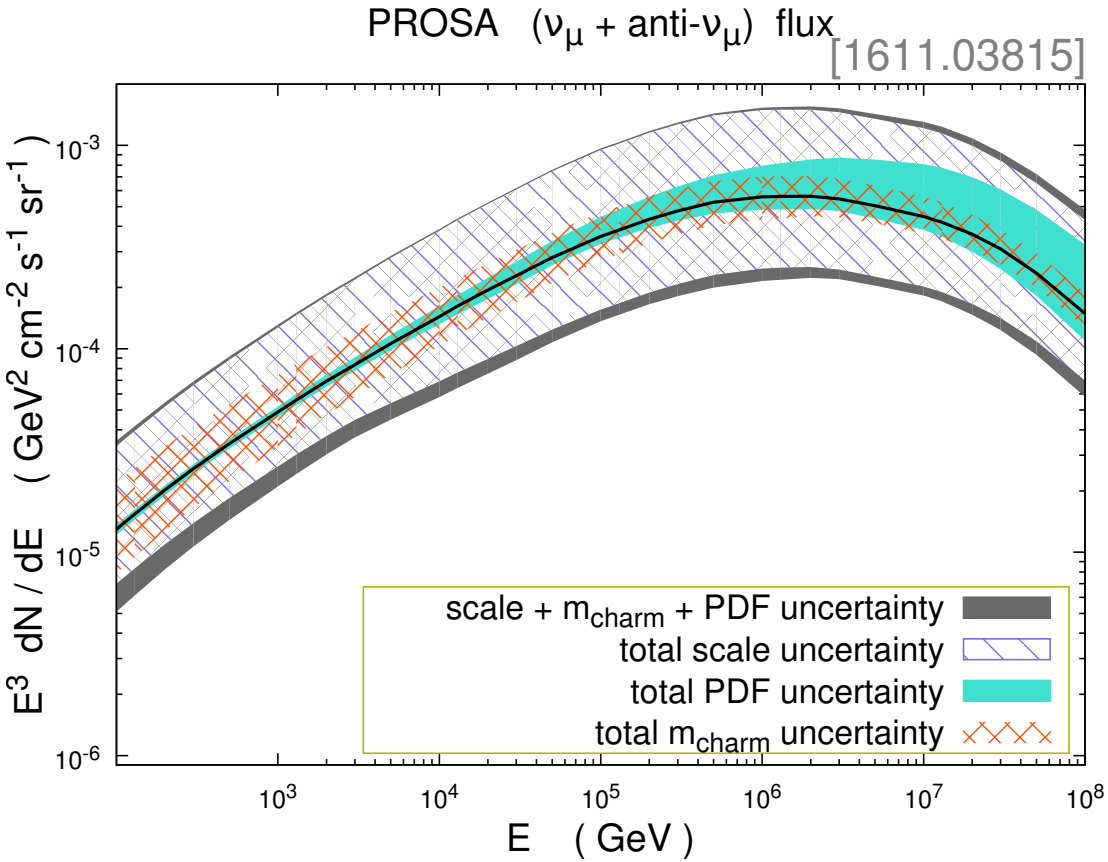
Not a problem of the FFNS calculations, rather an important correction to the LO kinematics:

$\sim$ 50% of low p_T forward data does not depend on low x gluons

Measurements of double charm production might be usefull.

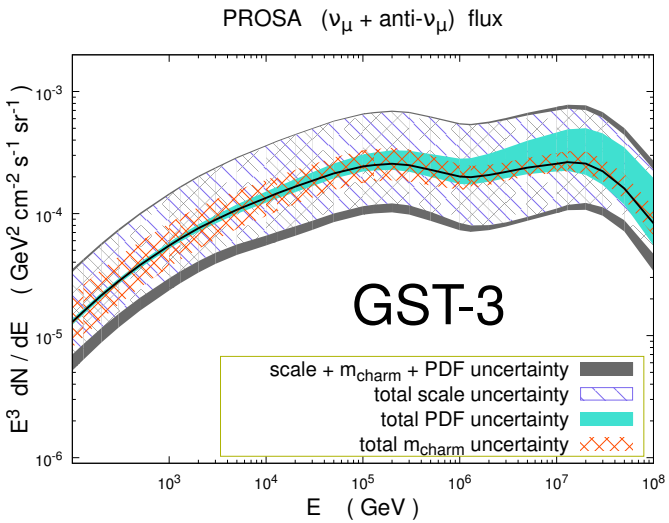
Prompt ν fluxes: NLO scale, m_c and PDF uncertainties

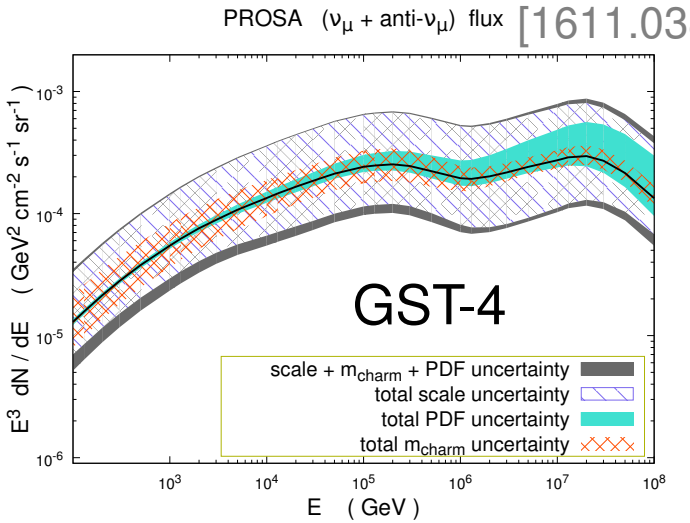
$$\frac{d\phi_j(E_j,X)}{dX} = -\frac{\phi_j(E_j,X)}{\lambda_j^{\rm int}(E_j)} - \frac{\phi_j(E_j,X)}{\lambda_j^{\rm dec}(E_j)} + \\ + \sum_{k\neq j} S_{\rm prod}^{k\to j}(E_j,X) + \sum_{k\neq j} S_{\rm decay}^{k\to j}(E_j,X) + S_{\rm reg}^{j\to j}(E_j,X)\,.$$



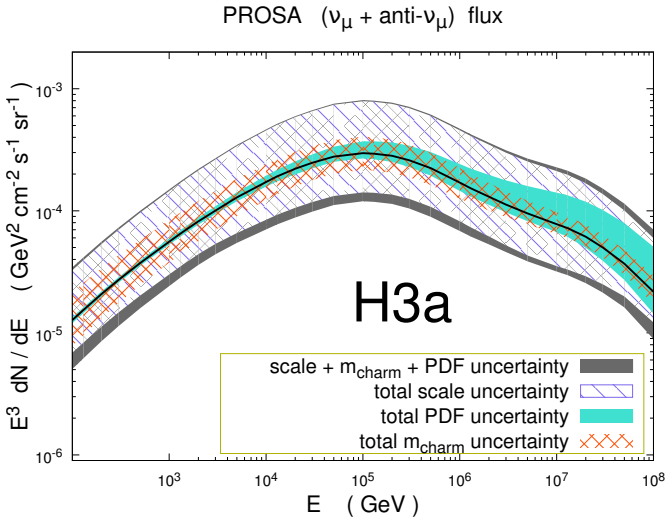
- Uncertainties for prompt ν flux are dominated by NLO scale uncertainties
- PDF uncertainties are under control owing to LHCb data in PROSA fit

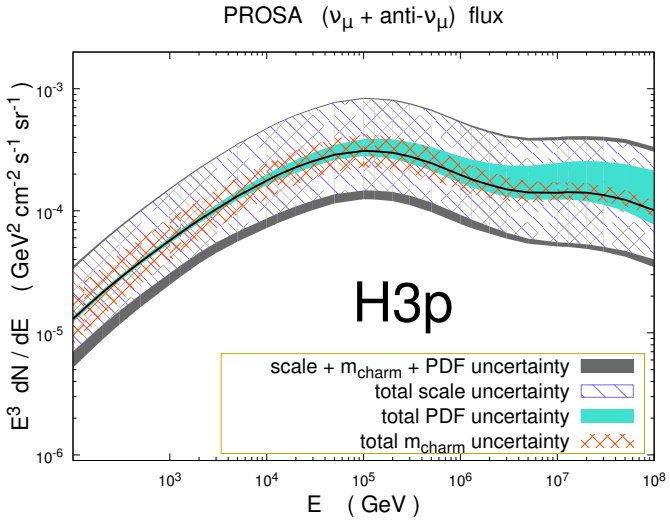
Comparison using as input different primary CR fluxes





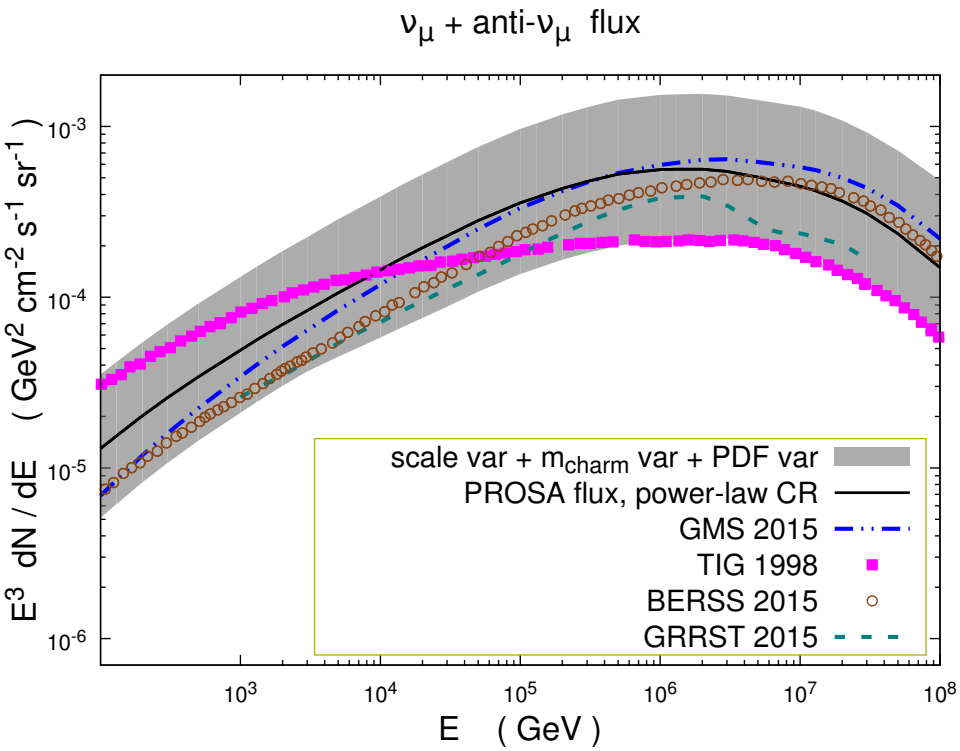
[1611.03815]

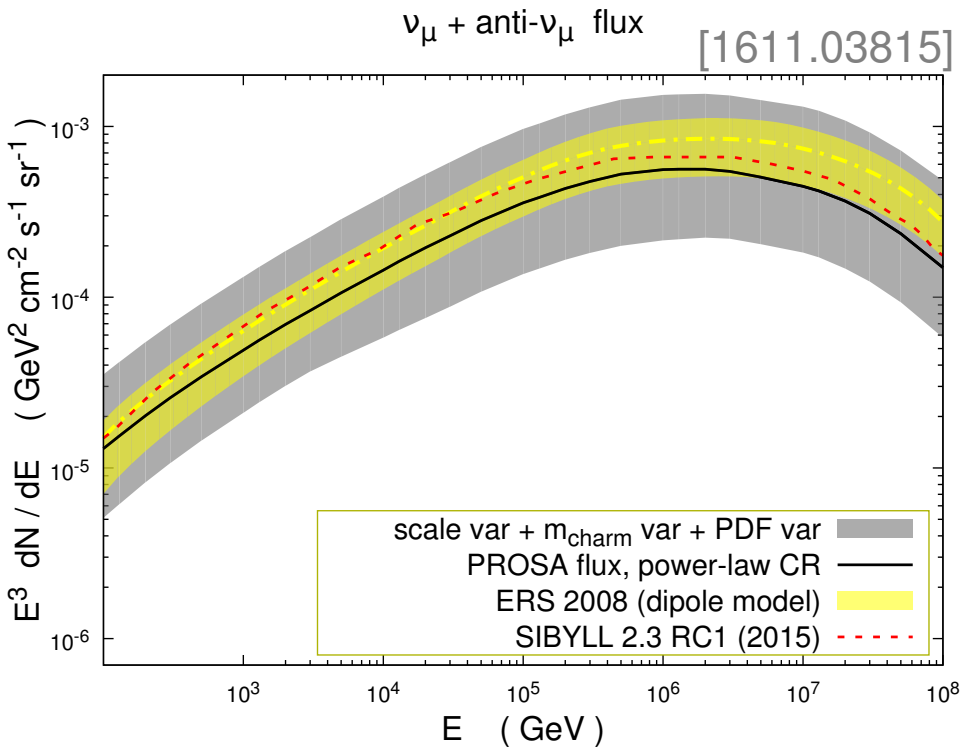




- Compared GST-3, GST-4, H3a, H3p primary fluxes
- PDF uncertainties are under control
- Dependence on primary CR becomes large at $E \gtrsim 10^7$ GeV

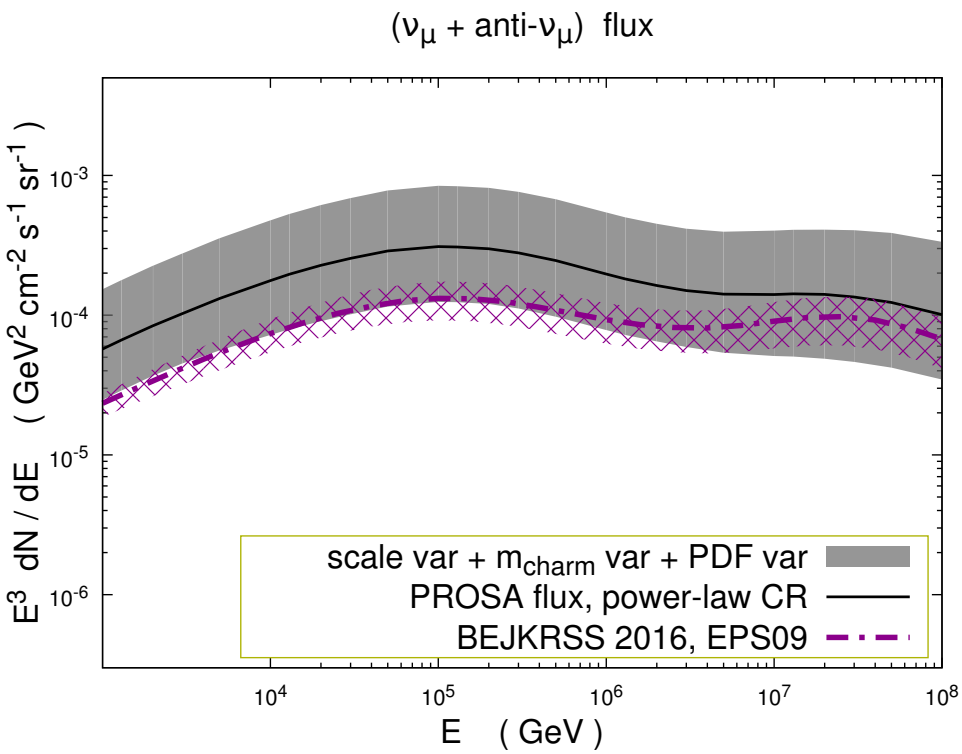
Prompt ν fluxes: comparison to other predictions

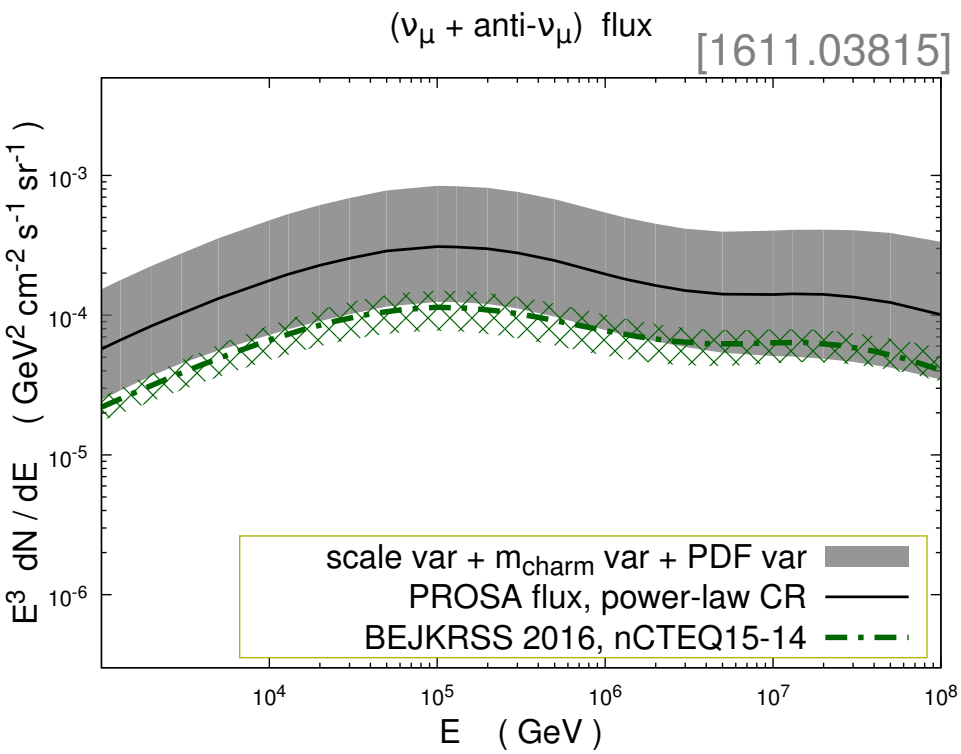




- Good agreement with other calculations based on perturbative QCD (left)
- Good agreement also with other phenomenological approaches (right)

Prompt ν fluxes: proton vs. nuclear PDFs

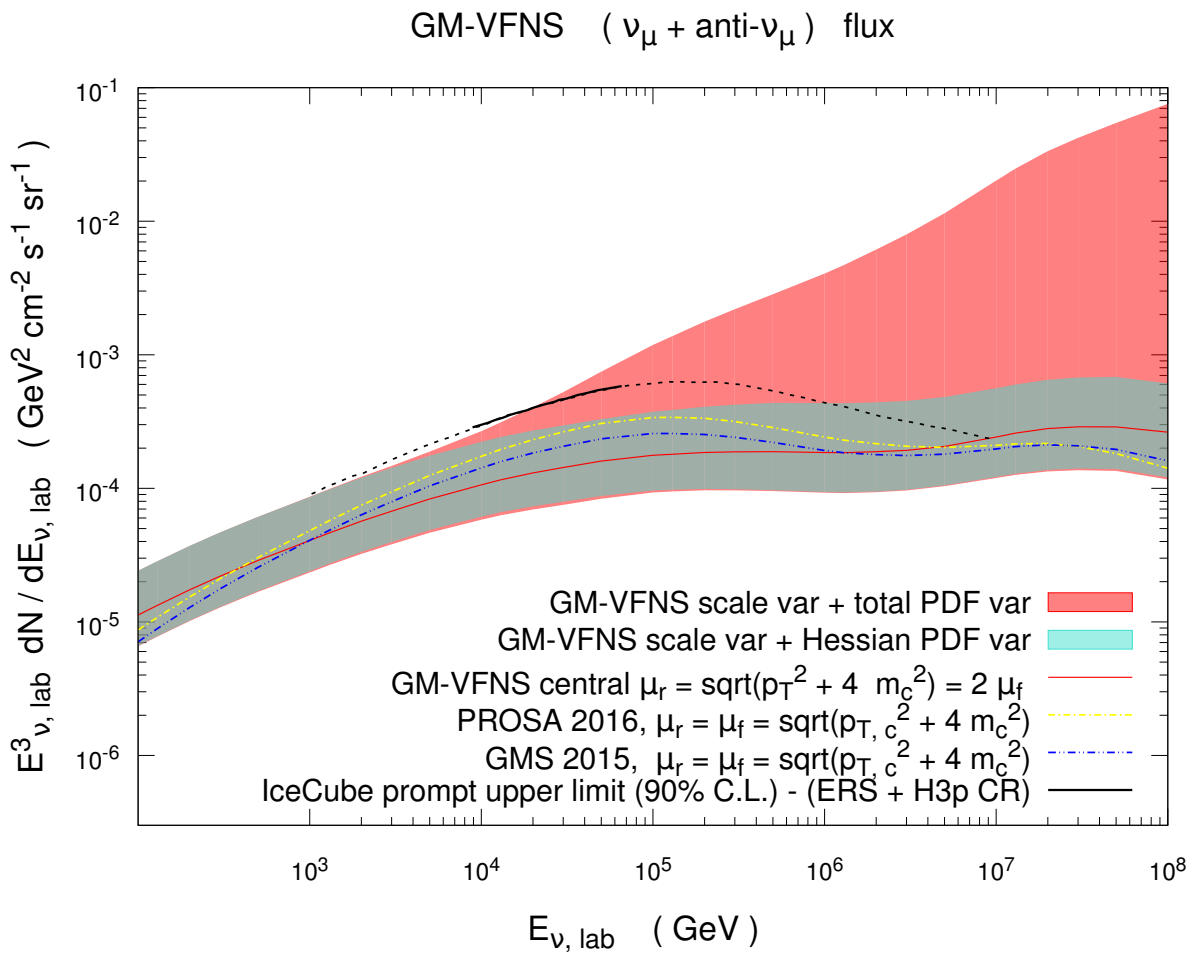




- Any uncertainties on nuclear PDFs are not included in these plots!
  (and scale variations were restricted in BEJKRSS)
- Still they are close/within PROSA uncertainty band
- *But how would another proton PDFs behave, which do not have LHCb data constraining low x gluons?*

Prompt ν fluxes: comparison vs GM-VFNS + CT14nlo

[M. Benzke, M.V. Garzelli, B. Kniehl, G. Kramer, S. Moch, G. Sigl, arXiv:1705.10386]

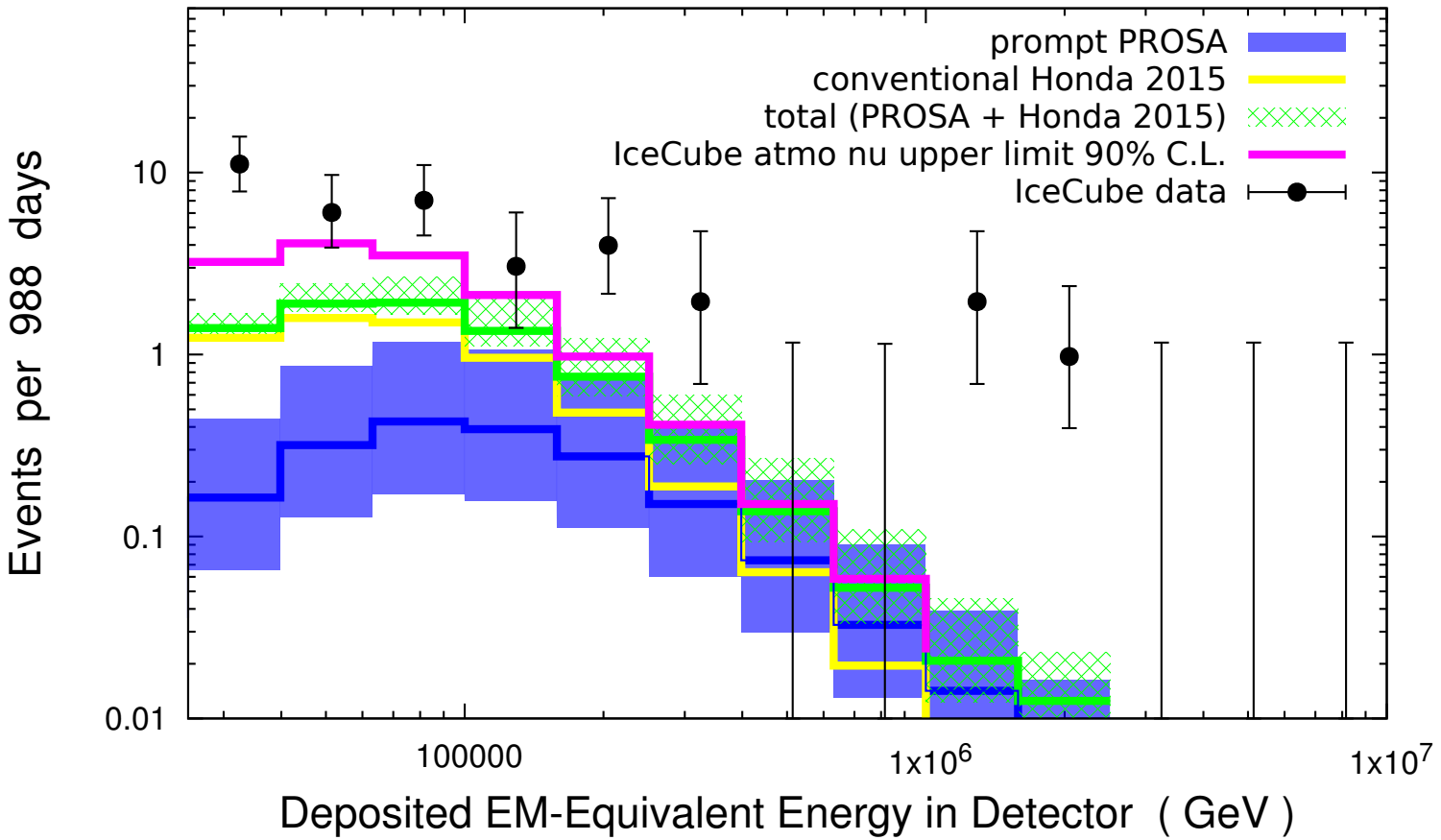


- Predictions calculated in GM-VFNS using CT14nlo PDFs
- Uncertainties of GM-VFNS are dominated by CT14nlo PDFs
- Their upper band is larger than the IceCube upper limit

  $\Rightarrow$ IceCube results are capable to constrain PDFs!

Total atmospheric ν flux

[1611.03815]



- Predictions for the number of prompt, conventional and total expected atmospheric neutrino events for the IceCube 988-day HESE analysis, as compared to the IceCube lepton data
- IceCube upper limit lies well inside PROSA uncertainty band at high E_ν
- This shows potential of astrophysical measurements to provide complementary information about charm hadroproduction

Update on PROSA 2015 fit

- We are working on updating PROSA 2015 PDF fit with new LHC data from LHCb and ALICE

  $\rightarrow$ in particular LHCb charm data at 13 TeV which extend low-x coverage
- As several precise measurements are already available (with experimental uncertainties $< 10\%$), one needs **information on correlations of systematic uncertainties** to fully exploit potential of LHC data

  $\rightarrow$ already emphasised by R. Gauld in JHEP 1705 (2017) 084
- This is especially important when calculating normalised cross sections from published absolute cross sections
- One needs also information on correlation of systematic uncertainties which affect different final states and/or measurements at different energies

Correlated uncertainties (description from ALICE paper 1901.07979)

	D^0			D^+		D^{*+}		D_s^+	
p_T (GeV/c)	0-0.5	2-2.5	10-12	2-2.5	10-12	2-2.5	10-12	2-3	8-12
Signal yield	9%	3%	2%	3%	3%	3%	1%	7%	3%
Tracking efficiency	3%	4%	5%	4.5%	7%	4%	5%	4.5%	7%
Selection efficiency	0	5%	3%	4%	3%	5%	1%	8%	5%
PID efficiency	0	0	0	0	0	0	0	2.5%	0
p_T shape in MC	0	0	0	1%	0	1%	0	1%	0
Feed-down	$^{+1.1}_{-1.3}$%	$^{+3.6}_{-4.3}$%	$^{+3.8}_{-5.3}$%	$^{+2.4}_{-2.8}$%	$^{+2.3}_{-3.1}$%	$^{+3.0}_{-3.5}$%	$^{+1.8}_{-2.5}$%	$^{+2.8}_{-3.3}$%	$^{+3.4}_{-4.5}$%
Branching ratio		1.0%		3.1%		1.3%		3.5%	
Normalisation	2.1%								

Table 1: Summary of relative systematic uncertainties on D^0, D^+, D^{*+}, and D_s^+ measurements in different p_T intervals.

"The systematic uncertainty on PID, tracking, and selection efficiencies are mainly correlated among the different pT intervals, while the raw-yield extraction uncertainty is mostly uncorrelated."

This information is not really sufficient:

- need to know contributions of different systematic uncertainties for each bin
  → taking lowest of quoted values for all bins as correlated unc. (e.g. 3% for D^0 track. eff.)
- need to know how to treat 'mostly uncorrelated' uncertainties
  → assumed uncorrelated (calculated by subtracting correlated sources from total syst. unc.)
- need to know how to correlate uncertainties between different final states and c.m.e.
  → only tracking and lumi uncertainty assumed correlated between different final states
  → all unc. assumed uncorrelated between different c.m.e. (except branching ratio unc.)
- see e.g. sufficient description given for HERA data 1506.06042, 1804.01019

Summary

- Constraining PDFs at low x is important for hadronic interactions at high energies, as probed in astroparticle physics
- PROSA PDF fit: low x gluon constrained by LHCb charm and beauty production data

  $\rightarrow$ other PDF sets would definetely benefit from these data
- Once PDF are under control, largest uncertainties on prompt ν fluxes are due to QCD scale variations at NLO

  $\rightarrow$ higher-order calculations are needed
- LHC measurements of charm and beauty production will be used in future PDF fits

  $\rightarrow$ correlation matrices must be published
- Astrophysical measurements provide data which are complementary to colliders in constraining QCD