Automated calculation of jet fragmentation at NLO in QCD

combining Monte Carlo generators and fragmentation functions

Xiao-min Shen

Shanghai Jiao Tong University and DESY

in collaboration with ChongYang Liu, Bin Zhou, Jun Gao

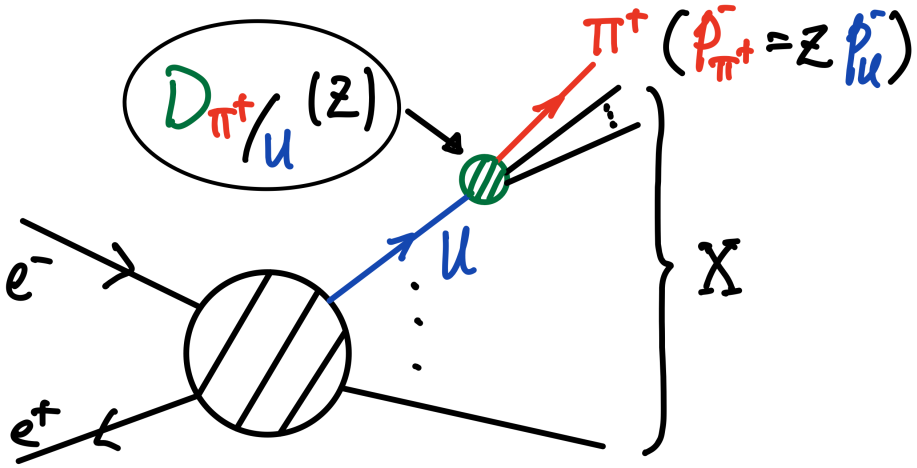
(arXiv:2305.14620)

EPS-HEP 2023, 25 August 2023

The Fragmentation Functions (FFs)

- counterpart of PDFs in the final state

- $D_{h/i}(z)$: intuitive probabilistic interpretation

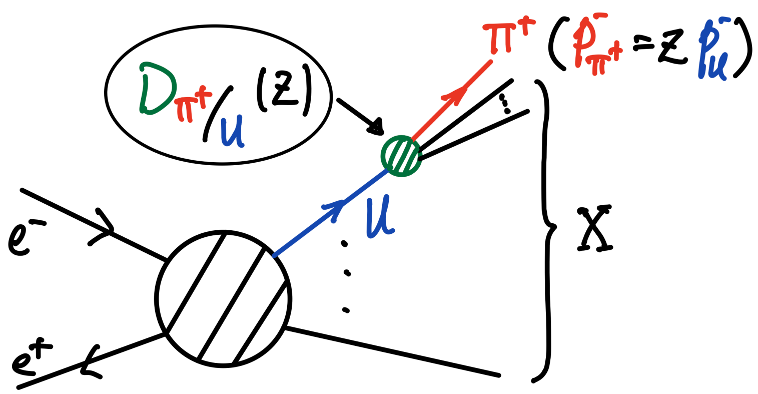


  - number density of collinear fragmentation of
  - massless unpolarized parton i
  - into an unpolarized hadron h with **momentum fraction** z

- sufficiently inclusive observables
  - $e^+e^- \to h+X$ (Single Inclusive Annihilation): $\hat{\sigma} \otimes \mathrm{FF}$
  - $e^-p \to e^-+h+X$ (Semi-Inclusive DIS): $\hat{\sigma} \otimes \mathrm{PDF} \otimes \mathrm{FF}$
  - $p\, p \to h+X$: $\hat{\sigma} \otimes \mathrm{PDF} \otimes \mathrm{PDF} \otimes \mathrm{FF}$

Example: Single Inclusive Annihilation

- energy fraction x_h of hadron h in SIA $e^+e^- \to h+X$:



$$\frac{d\sigma}{dx_h}=\sum_{i=q,\bar{q},g}\int\frac{dz}{z}C_i^0(x/z,\epsilon)D^0_{h/i}(z,\epsilon)\equiv\sum_i C_i^0(\epsilon)\otimes D^0_{h/i}(\epsilon)$$

$$C_q^0(x,\epsilon,\alpha_s) \sim \delta(1-x)+\frac{\alpha_s}{2\pi}\left(P^+_{qq}(x)\left[\frac{-1}{\epsilon}+\ln\frac{Q^2}{\mu_R^2}\right]+C_F f_q(x)\right)$$

$$C_g^0(x,\epsilon,\alpha_s) \sim 2\times\frac{\alpha_s}{2\pi}\left(P^+_{gq}(x)\left[\frac{-1}{\epsilon}+\ln\left(\frac{Q^2}{\mu_R^2}\right)\right]+C_F f_g(x)\right)$$

- divergence of the coeff. functions absorbed by the bare FFs

$$\frac{d\sigma}{dx_h}=\sum_i C_i^0(\epsilon)\otimes D^0_{h/i}(\epsilon)=\sum_i C_i(\alpha_s(\mu_R),\mu_D)\otimes D_{h/i}(\mu_D)$$

$$D^0_{h/i}(x,\epsilon)\equiv D_{h/i}(x,\mu_D)+\left(\frac{\mu_R^2}{\mu_D^2}\right)^{\epsilon}\frac{\alpha_s(\mu_R)}{2\pi}\frac{1}{\epsilon}P^{+(0)}_{ji}\otimes D_{h/j}(x,\mu_D)$$

- evolution equation

$$\frac{d}{d\ln\mu_D^2}D_{h/i}(z,\mu_D) = \frac{\alpha_s(\mu_R)}{2\pi}\sum_j\int_z^1\frac{dy}{y}P^+_{ji}(y)\,D_{h/j}\left(\frac{z}{y},\mu_D\right)$$

Motivation: Combine MC generators and FFs

- **analytical** calculation of the coefficient (structure) functions
  - available hard processes are **limited**, usually implemented **case-by-case**
  - analytical results with various **selection conditions** may not be available

- this work:
  - a new prescription to combine **general-purpose MC generators** and **FFs**
  - implemented with MG5_aMC@NLO
    - automated NLO calculation for various processes
    - cuts, jet reconstruction are possible

$\rightarrow$ deal with IR divergence in **4-dim spacetime**

Local subtraction with one identified hadron at NLO

- Local subtraction for IRC safe observable at NLO QCD

$$\frac{d\sigma}{dF} = \int d\,\mathrm{PS}_m \,[|M|^2_{B,m} + |M|^2_{V,m} + |\tilde{\mathcal{I}}|^2_m]\,\delta\,(\hat{F}(p_m; f_m) - F)$$
$$+ \int d\,\mathrm{PS}_{m+1}\,[|M|^2_{R,m+1}\,\delta\,(\hat{F}(p_{m+1}; f_{m+1}) - F) - |\mathcal{I}|^2_{m+1}\,\delta\,(\hat{F}(\tilde{p}_m; \tilde{f}_m) - F)]$$

- With one identified hadron, e.g. hadron p_T, naively

$$\frac{d\sigma}{dp_{T,h}} = \int dx \int d\,\mathrm{PS}_m \,[|M|^2_{B,m} + |M|^2_{V,m} + |\tilde{\mathcal{I}}|^2_m] \sum_{i=1}^{m} \delta\,(p_{T,h} - x\,p_{T,i})\, D^0_{h/i}(x)$$
$$+ \int dx \int d\mathrm{PS}_{m+1}\left[|M|^2_{R,m+1} \sum_{i=1}^{m+1} \delta\,(p_{T,h} - x\,p_{T,i}) D^0_{h/i}(x) - |\mathcal{I}|^2_{m+1} \sum_{\tilde{i}=1}^{m} \delta(p_{T,h} - x\,\tilde{p}_{T,\tilde{i}})\, D^0_{h/\tilde{i}}(x)\right]$$

  - **Problem: collinear divergences not locally cancelled**
    - additional subtraction terms, or
    - in this work: **slicing of phase space**

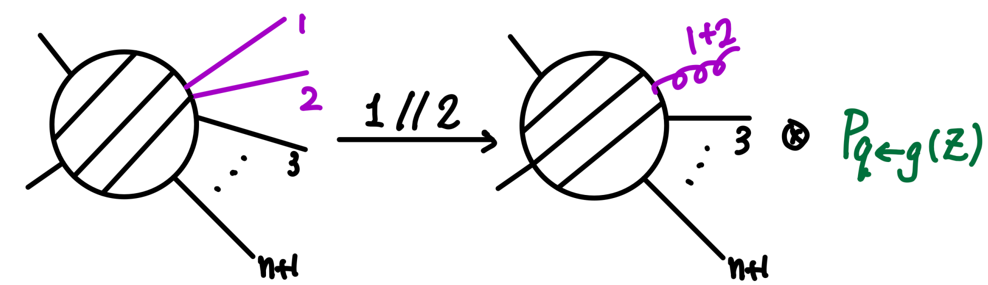
Framework: local subtraction + PS slicing

- slicing of phase space: $dPS_{m+1} = dPS_{m+1}\Theta(\Delta R_{kl} - \lambda) + dPS_{m+1}\Theta(\lambda - \Delta R_{kl})$

$$\sum_{\{kl\}} \int d\,x \int d\,\mathrm{PS}_{m+1}\, \Theta\left(\lambda - \Delta\, R_{kl}\right) \left[|M|^2_{R,m+1} \sum_{i=1}^{m+1} \delta\left(p_{T,h} - x\, p_{T,i}\right) D^0_{h/i}(x) \right.$$

$$\left. - |\mathcal{I}|^2_{m+1} \sum_{\tilde{i}=1}^{m} \delta\left(p_{T,h} - x\,\tilde{p}_{T,\tilde{i}}\right) D^0_{h/\tilde{i}}(x) \right] \equiv |\tilde{\mathcal{J}}|^2_m(\lambda)$$

  - in collinear region: the matrix elements factorize
  - factorize the phase space and integrate analytically



- arrive at a rather compact form

Born matrix element

$$|\tilde{\mathcal{J}}|^2_m(\lambda) \;=\; \int d\,x \int d\,\mathrm{PS}_m |M|^2_{B,m} \frac{\alpha_S(\mu_R)}{2\,\pi} \frac{(4\,\pi)^\epsilon}{\Gamma(1-\epsilon)} \sum_{i=1}^{m} \delta\left(p_{T,h} - x\, p_{T,i}\right)$$

$$\times \left[\left(-\frac{1}{\epsilon} + \ln \frac{\lambda^2\, p^2_{T,i}}{\mu_R^2} \right) \sum_j P^{+(0)}_{ji} \otimes D_{h/j}(x,\, \mu_D) + \tilde{D}_{h/i}(x,\, \mu_D) \right]$$

IR div. absorbed into FFs

finite part from $\frac{1}{\epsilon} \times O(\epsilon)$

FMNLO: interfaced to MG5_aMC@NLO

- The master formula for x_h distribution in SIA

$$\frac{d\,\sigma}{d\,x_h} = \sum_{i=1}^{m} \int \frac{d\,x_i}{x_i}\left[\frac{d\,\sigma_m^{(0)}}{d\,x_i} + \frac{d\,\tilde{\sigma}_m^{(1)}}{d\,x_i}\right] D_{h/i}(x_h/x_i,\,\mu_D) + \sum_{i=1}^{m+1} \int \frac{d\,x_i}{x_i}\left[\frac{d\,\tilde{\sigma}_{m+1}^{(1)}}{d\,x_i}\right] D_{h/i}(x_h/x_i,\,\mu_D)$$
$$+ \sum_{i=1}^{m} \int \frac{d\,x_i}{x_i}\left[\frac{\alpha_S(\mu_R)}{2\,\pi}\frac{d\,\sigma_m^{(0)}}{d\,x_i}\right] (\bar{D}_{h/i}(x_h/x_i,\,\mu_D) + \tilde{D}_{h/i}(x_h/x_i,\,\mu_D))$$

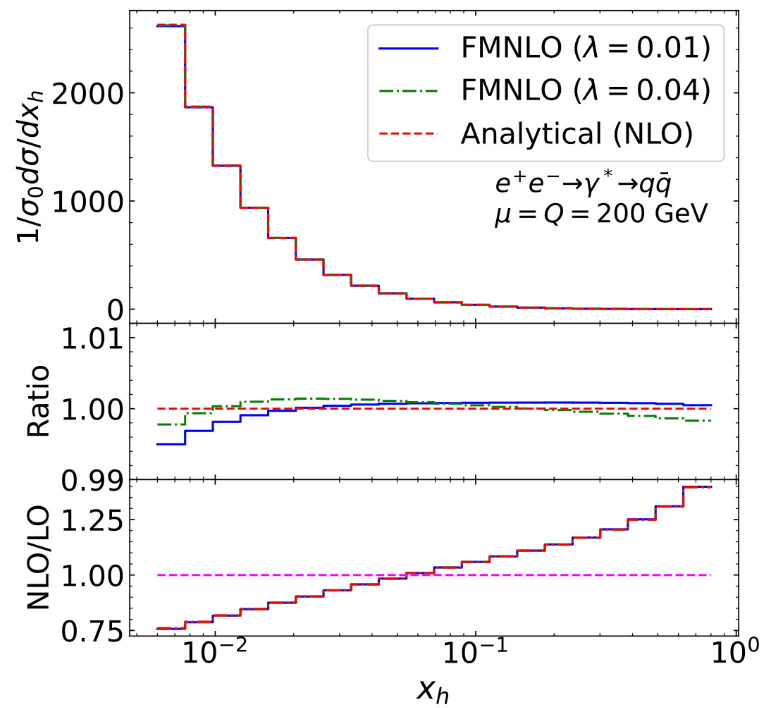
- **FMNLO**: our framework interfaced to MG5_aMC@NLO

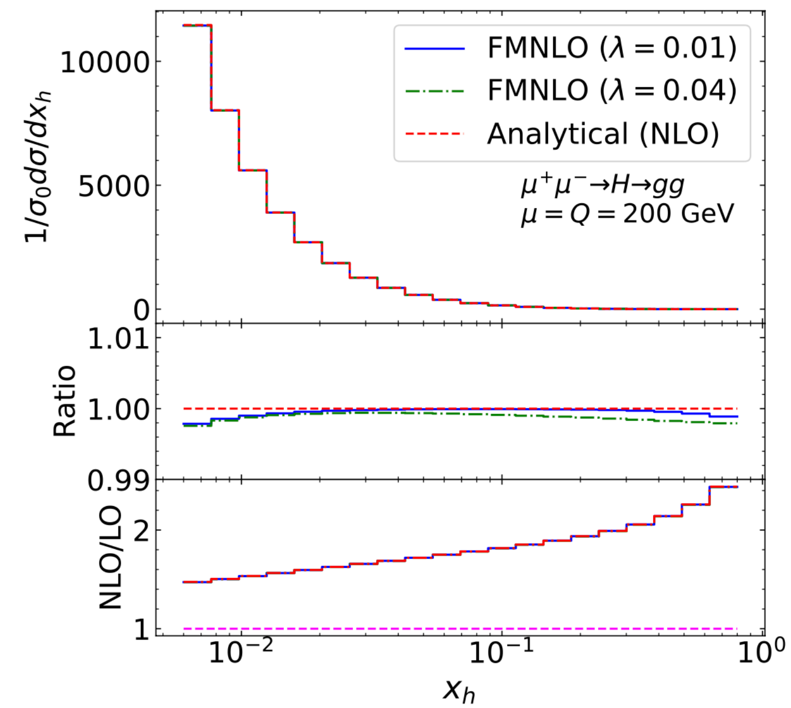
$$\frac{d\,\sigma}{d\,x_h} = \sum_{i=q,\bar{q},g} \sum_{k=1}^{N_x} \sum_{l=1}^{N_Q} [G(x_h)^i_{k,l}\, D^{k,l}_{h/i} + \bar{G}(x_h)^i_{k,l}\,(\bar{D}^{k,l}_{h/i} + \tilde{D}^{k,l}_{h/i})]$$

  - partonic ME: calculated once and stored using histograms in MG5_aMC@NLO
  - FFs are approximated by an interpolation on a 2D grid of (x, μ_D)
  - evolution and convolution: HOPPET

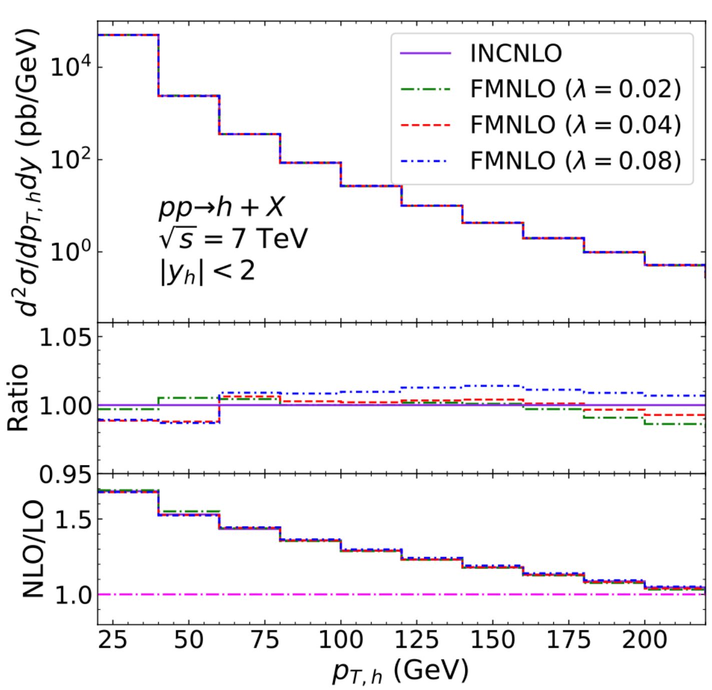
Validation of FMNLO

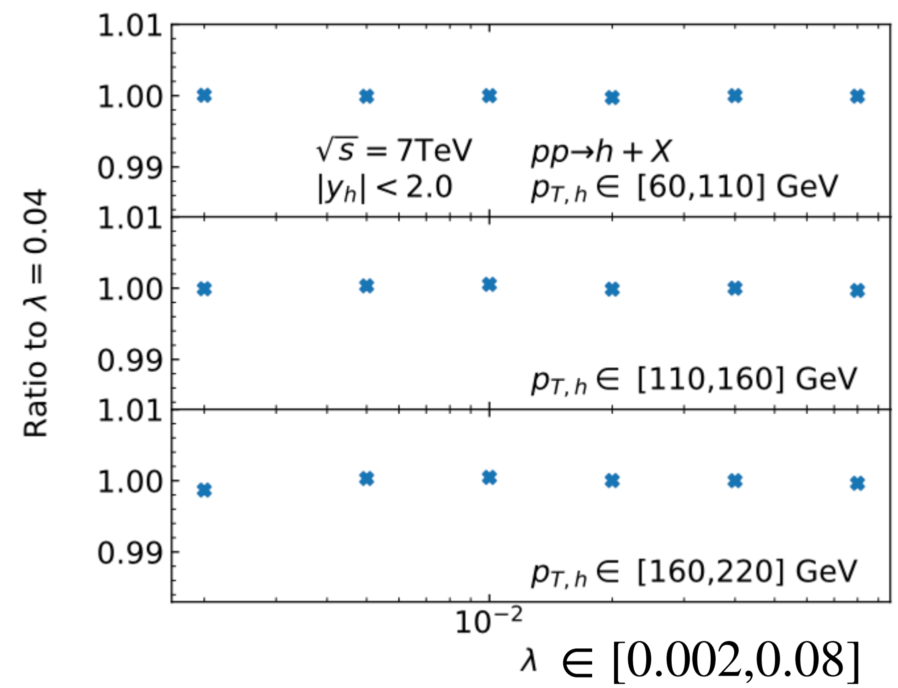
leptonic collision





pp collision





NLO fit of charged hadron FFs using LHC data

- measurements of **unidentified charged hadron ($\pi^{\pm}$, $K^{\pm}$, etc.)** at the LHC

Experiments	lum.	observables	N_{pt}	Range	tagging
CMS 5.02 TeV	27.4 pb^{-1}	$1/N_j\, dN_{trk}/d\xi_T^{\gamma}$	8(5)	$\xi_T^{\gamma} \in [0.5, 4.5]$	γ
ATLAS 5.02 TeV	25 pb^{-1}	$1/N_j\, dN_{trk}/dp_{T,h}$	10(7)	$p_{T,h} \in [1, 100]$GeV	γ
CMS 5.02 TeV	320 pb^{-1}	$1/N_Z\, dN_{trk}/dp_{T,h}$	14(11)	$p_{T,h} \in [1, 30]$GeV	Z
ATLAS 5.02 TeV	160 pb^{-1}	$1/N_Z\, d^2N_{trk}/dp_{T,h}\, d\Delta\phi$	15(9)	$p_{T,h} \in [1, 60]$GeV	Z
ATLAS 13 TeV	33 fb^{-1}	$1/N_j\, dN_{trk}/d\zeta$(central)	261(143)	$\zeta \in [0.002, 0.67]$	dijet
ATLAS 13 TeV	33 fb^{-1}	$1/N_j\, dN_{trk}/d\zeta$(forward)	261(143)	$\zeta \in [0.002, 0.67]$	dijet

  - exclude data points with small jet p_T or small hadron momentum fraction

- parametrization of the FFs to unidentified charge hadrons

$$x\, D_{h/i}(x, Q_0 = 5\text{GeV}) = a_{i,0}\, x^{\alpha_i}\, (1-x)^{\beta_i}\, (1 + a_{i,1}\, x + a_{i,2} x^2)$$

  - assume FFs equal for all (anti-)quarks
  - 10 free parameters in total

NLO fit of FFs using LHC data

- Search for the best fit with MINUIT

$$\chi^2(\text{FFs}) = \sum_{i,j=1}^{N_{\text{pt}}} (T_i - D_i)[\text{cov}^{-1}]_{ij}\,(T_j - D_j)$$

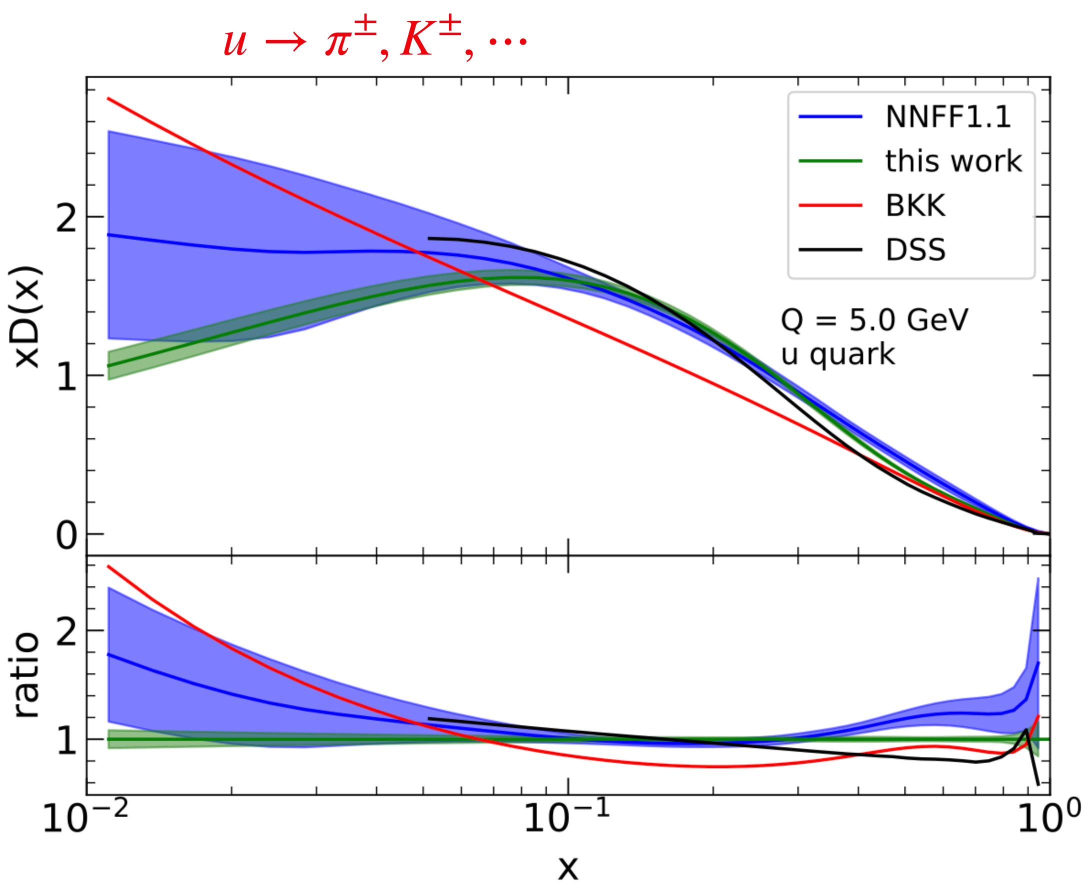
  - assuming theoretical uncertainties fully correlated within each subset of the data
  - nominal scale choices: $\mu_F = \mu_R = \sum m_T/2, \mu_D = \max\{p_T\}$

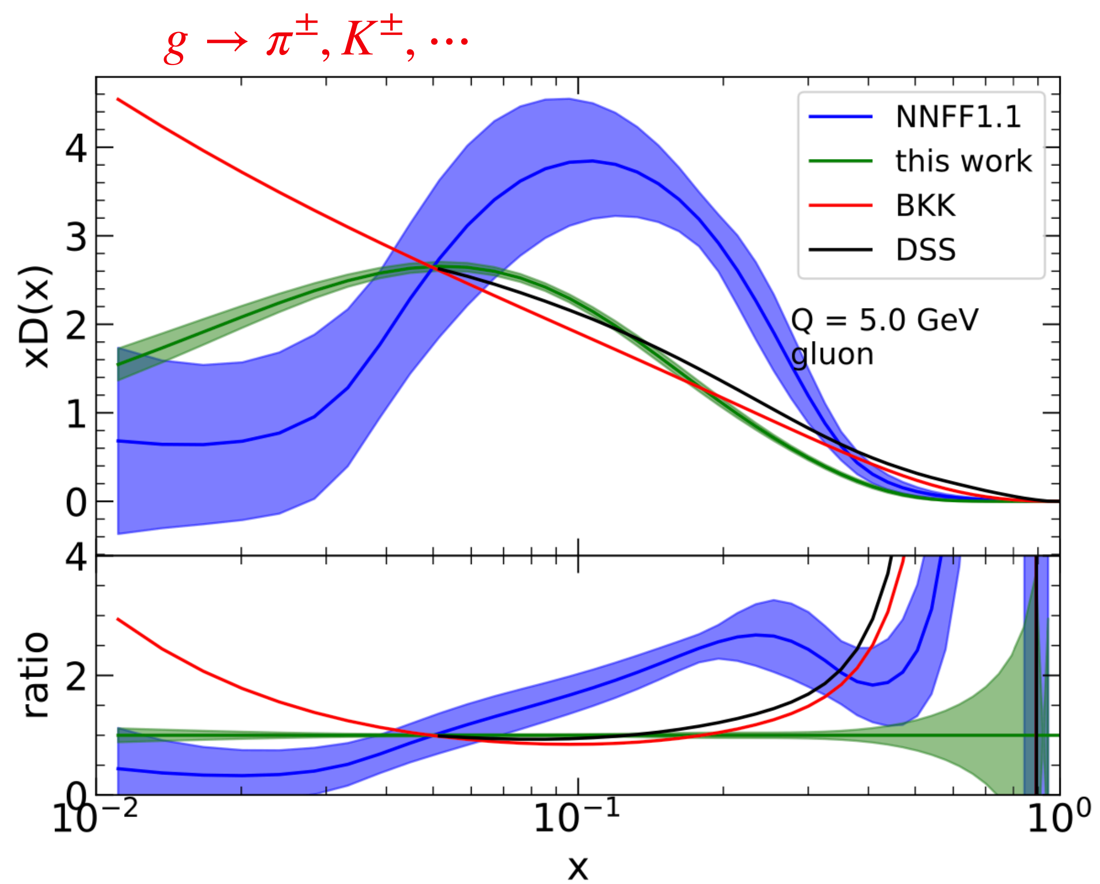
Experiments	N_{pt}	$\chi^2(/N_{pt})$, **NLO**	$\chi^2(/N_{pt})$, **NLO**$_{w/o\ th.}$	$\chi^2(/N_{pt})$, **LO**$_{w/o\ th.}$
CMS γ	5	11.3(2.27)	28.8(5.76)	48.5(9.71)
ATLAS γ	7	17.8(2.55)	18.8(2.68)	40.5(5.78)
CMS Z	11	16.2(1.47)	24.8(2.25)	906.9(82.4)
ATLAS Z	9	47.5(5.27)	48.1(5.34)	348.8(38.8)
ATLAS central jets	141	98.1(0.69)	112.9(0.79)	833.7(5.83)
ATLAS froward jets	141	76.4(0.53)	98.0(0.68)	855.6(5.98)
Total	318	267.4(0.84)	331.2(1.04)	3034.0(9.54)

  - good agreement to ATLAS dijet measurement
  - large χ^2 for ATLAS Z boson measurement

The quark and gluon FFs

- FFs to unidentified charged hadrons ($\text{parton} \to \pi^{\pm}, K^{\pm}, \cdots$)
  - error criterions $\Delta\chi^2 = 1$





- up quark FF : good agreement in $0.1 < x < 0.3$; large deviation in small x region
- gluon FF: notable disparities

Summary

- a new prescription combining **general-purpose MC generators** and **FFs** at **NLO QCD**
  - based on local subtraction + phase space slicing
  - has been realized with MG5_aMC@NLO → **FMNLO** is publicly available
  - suitable for frag. measurements with cuts, tagging, jet reconstruction
- a **preliminary** fit of unidentified charged hadrons FFs
  - include new LHC data: high energy/luminosity; with isolated photon or Z
- future improvements
  - include semi-inclusive DIS (SIDIS) at NLO
  - fit of all quark flavor FFs with more SIA, SIDIS, pp data

Summary

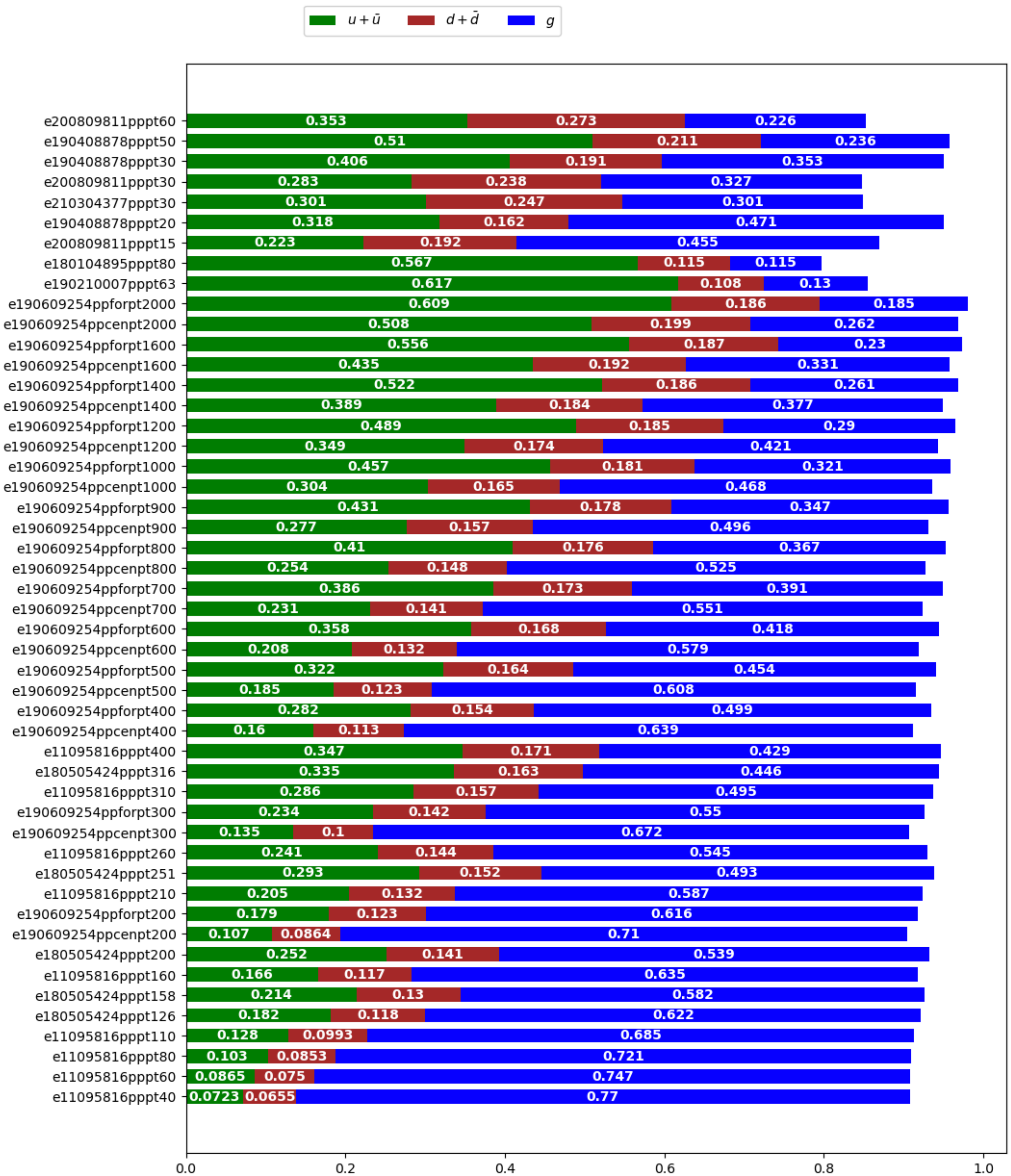
- a new prescription combining **general-purpose MC generators** and **FFs** at **NLO QCD**
  - based on local subtraction + phase space slicing
  - has been realized with MG5_aMC@NLO → **FMNLO** is publicly available
  - suitable for frag. measurements with cuts, tagging, jet reconstruction
- a **preliminary** fit of unidentified charged hadrons FFs
  - include new LHC data: high energy/luminosity; with isolated photon or Z
- future improvements
  - include semi-inclusive DIS (SIDIS) at NLO
  - fit of all quark flavor FFs with more SIA, SIDIS, pp data

Thank you!

backup slides

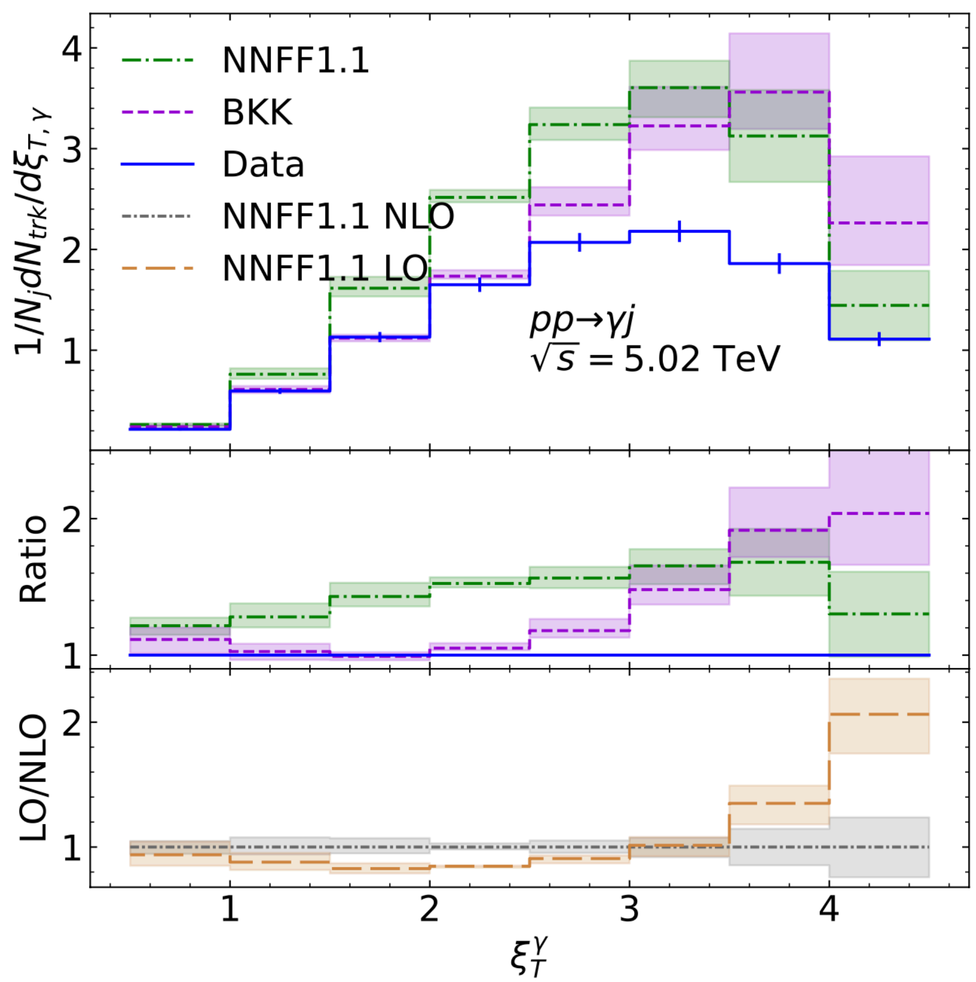
NLO fit of FFs for LHC data

quark	α	β	a_0	a_1	a_2	$\langle x \rangle$
best-fit	0.375	2.166	6.016	-2.292	2.083	0.586
unc.(scan)	$^{+0.03}_{-0.03}$	$^{+0.11}_{-0.12}$	$^{+0.55}_{-0.56}$	$^{+0.10}_{-0.10}$	$^{+0.18}_{-0.20}$	–
unc.(Hessian)	$^{+0.03}_{-0.03}$	$^{+0.09}_{-0.10}$	$^{+0.45}_{-0.44}$	$^{+0.08}_{-0.08}$	$^{+0.16}_{-0.16}$	$^{+0.007}_{-0.008}$
gluon	α	β	a_0	a_1	a_2	$\langle x \rangle$
best-fit	0.710	10.224	44.080	-3.527	11.786	0.510
unc.(scan)	$^{+0.09}_{-0.16}$	$^{+1.09}_{-0.91}$	$^{+19.54}_{-13.54}$	$^{+0.95}_{-0.85}$	$^{+3.54}_{-3.60}$	–
unc.(Hessian)	$^{+0.09}_{-0.10}$	$^{+0.91}_{-0.93}$	$^{+18.9}_{-14.1}$	$^{+0.92}_{-0.83}$	$^{+3.32}_{-3.52}$	$^{+0.011}_{-0.012}$

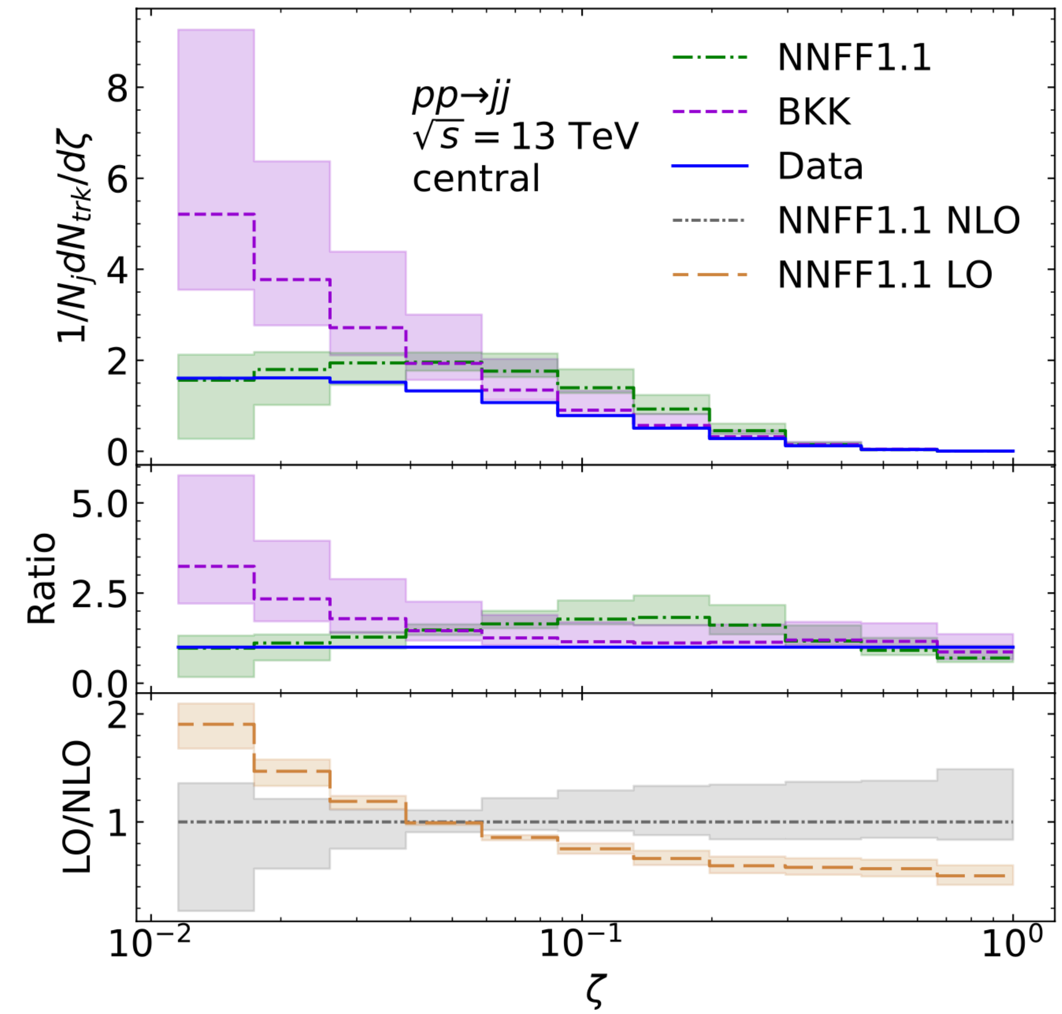


$u+\bar{u}$
$d+\bar{d}$
g
e200809811pppt60 0.353 0.273 0.226
e190408878pppt50 0.51 0.211 0.236
e190408878pppt30 0.406 0.191 0.353
e200809811pppt30 0.283 0.238 0.327
e210304377pppt30 0.301 0.247 0.301
e190408878pppt20 0.318 0.162 0.471
e200809811pppt15 0.223 0.192 0.455
e180104895pppt80 0.567 0.115 0.115
e190210007pppt63 0.617 0.108 0.13
e190609254ppforpt2000 0.609 0.186 0.185
e190609254ppcenpt2000 0.508 0.199 0.262
e190609254ppforpt1600 0.556 0.187 0.23
e190609254ppcenpt1600 0.435 0.192 0.331
e190609254ppforpt1400 0.522 0.186 0.261
e190609254ppcenpt1400 0.389 0.184 0.377
e190609254ppforpt1200 0.489 0.185 0.29
e190609254ppcenpt1200 0.349 0.174 0.421
e190609254ppforpt1000 0.457 0.181 0.321
e190609254ppcenpt1000 0.304 0.165 0.468
e190609254ppforpt900 0.431 0.178 0.347
e190609254ppcenpt900 0.277 0.157 0.496
e190609254ppforpt800 0.41 0.176 0.367
e190609254ppcenpt800 0.254 0.148 0.525
e190609254ppforpt700 0.386 0.173 0.391
e190609254ppcenpt700 0.231 0.141 0.551
e190609254ppforpt600 0.358 0.168 0.418
e190609254ppcenpt600 0.208 0.132 0.579
e190609254ppforpt500 0.322 0.164 0.454
e190609254ppcenpt500 0.185 0.123 0.608
e190609254ppforpt400 0.282 0.154 0.499
e190609254ppcenpt400 0.16 0.113 0.639
e11095816pppt400 0.347 0.171 0.429
e180505424pppt316 0.335 0.163 0.446
e11095816pppt310 0.286 0.157 0.495
e190609254ppforpt300 0.234 0.142 0.55
e190609254ppcenpt300 0.135 0.1 0.672
e11095816pppt260 0.241 0.144 0.545
e180505424pppt251 0.293 0.152 0.493
e11095816pppt210 0.205 0.132 0.587
e190609254ppforpt200 0.179 0.123 0.616
e190609254ppcenpt200 0.107 0.0864 0.71
e180505424pppt200 0.252 0.141 0.539
e11095816pppt160 0.166 0.117 0.635
e180505424pppt158 0.214 0.13 0.582
e180505424pppt126 0.182 0.118 0.622
e11095816pppt110 0.128 0.0993 0.685
e11095816pppt80 0.103 0.0853 0.721
e11095816pppt60 0.0865 0.075 0.747
e11095816pppt40 0.0723 0.0655 0.77
0.0
0.2
0.4
0.6
0.8
1.0

Applications: compared with LHC measurements



[CMS, 1801.04895]



[ATLAS, 1906.09254]

$$\xi_T^\gamma := \ln\left[\frac{\vec{p}_{T,\gamma}^{\,2}}{-\vec{p}_{T,\gamma}\cdot\vec{p}_{T,h}}\right] \xrightarrow{\text{LO}} -\ln z$$

$$\zeta := \frac{p_{T,h}}{p_{T,j}}$$