PineAPPL: towards NLO EW corrections in NNPDF

arXiv:2008.12789

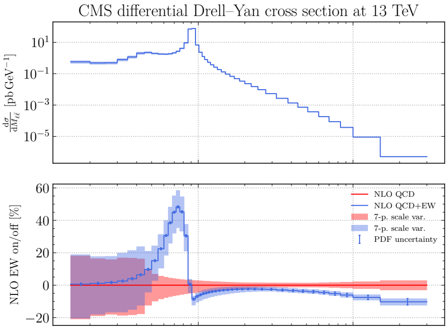
Christopher Schwan

Radcor & Loopfest 2021, 21 May 2021

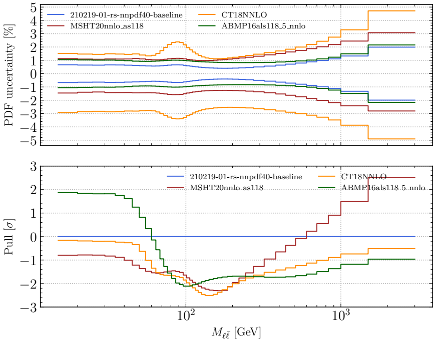
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 740006

NLO EW PDF fit with NNPDF: motivation

- NNPDF4.0 – see Maria's and Michael/Shayan/Tommaso's talk – aims for $\lesssim 1\,\%$ accuracy/precision
- Further improvement need previously neglected contributions: theory uncertainties, NLO EW, ...
- NLO EW: never been done in a global (all processes) and consistent way (correct data)



- Can include extreme phase-space regions with large EW corrections: large $M_{\ell\ell}$ in Drell–Yan, large p_T of Z/W bosons, ... → constrain high x
- These regions are going to be measured more precisely in the future, e.g. predictions for CMS DY $L = 2.8\,\mathrm{fb}^{-1}$ @ 13 TeV [CMS Collaboration]



NLO EW in a PDF fit: What has to be done?

✓ Theory: Need corrections in the form of interpolation grids: PineAPPL

- interfaced with MG5_aMC@NLO, see [S. Carrazza, E.R. Nocera, C.S., M. Zaro]; now released in v3.1.0: `https://launchpad.net/mg5amcnlo`
- WIP: `SHERPA`/`MCgrid` [E. Bothmann et al.]
- API available in `C`, `C++`, `Fortran`, `Python`, `Rust`

✓ Data: Needs careful selection

- no subtraction of FSR
- no photon-initiated subtraction
- proper observable definition

→ subset of NNPDF4.0's LHC data + ...

✓ Write/test runcards for all PDF processes and ...

✗ Run them (WIP)

✗ Implement changed data (WIP)

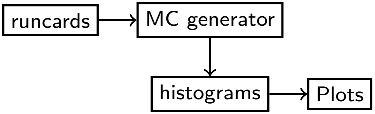
→ Run fit

What is PineAPPL?

PineAPPL

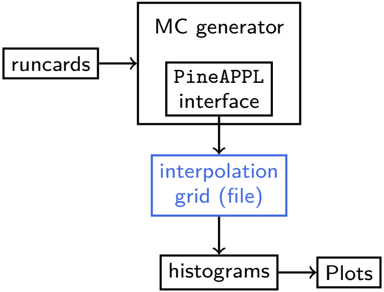
A tool/library for storing PDF-independent theoretical predictions in interpolation grids [S. Carazza, E.R. Nocera, C.S., M. Zaro]

Typical setup:



1. write runcards for your process/distributions
2. MC generates histograms with PDFs baked-in
3. changing PDFs is slow—need to rerun everything

With `PineAPPL`:



- grids can be convolved with arbitrary PDF sets in a matter of seconds

Interpolation techniques and applications

Not a new idea:

- APPLgrid [T. Carli et al.] or
- fastNLO [T. Kluge, K. Rabbertz, M. Wobisch]

Interpolate PDFs $f_a(x)$ with kernels $L_i(x), x \mapsto (0,1)$:

$$f_a(x) = \sum_{i=1}^{\infty} f_a^i L_i(x)$$

Convolution turns into a sum:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\mathcal{O}} = \sum_a \int \mathrm{d}x\, f_a(x) \frac{\mathrm{d}\sigma_a}{\mathrm{d}\mathcal{O}}(x) = \sum_{a,i} f_a^i G_i$$

Interpolation grid $\{G_i\}_{i=1}^{\infty}$ indep. of PDFs:

$$G_i = \int \mathrm{d}x\, L_i(x) \sigma_a(x)$$

Applications of interpolation grids

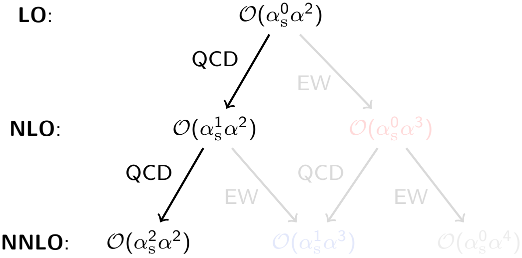
- input for PDF fits or
- study impact of PDFs uncertainties of observables

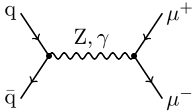
Why PineAPPL?

→ need support EW corrections (any powers of α and/or α_s)

- performance becomes very important: more bins, initial-states, scale variations ...
- powerful tooling (see later slides)

Which EW corrections will be included for $\mathrm{pp} \to \ell\bar{\ell}$?





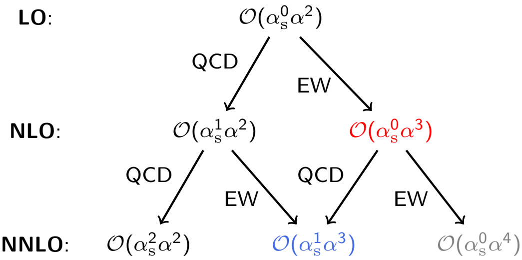
- NNLO QCD corrections included in PDF fits

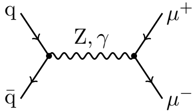
→ include also higher-order α contributions in NNPDF 4.x
  - for all processes
  - check impact of the corrections,
  - be more inclusive, etc.

→ `PineAPPL` supports all higher-orders

- $\mathcal{O}(\alpha_s \alpha^3)$ might become available at some point . . .
- lots of progress → talks by A. Behring, A. Sankar, N. Rana, L. Buonocore, A. von Manteuffel

Which EW corrections will be included for $\mathrm{pp} \to \ell\bar{\ell}$?





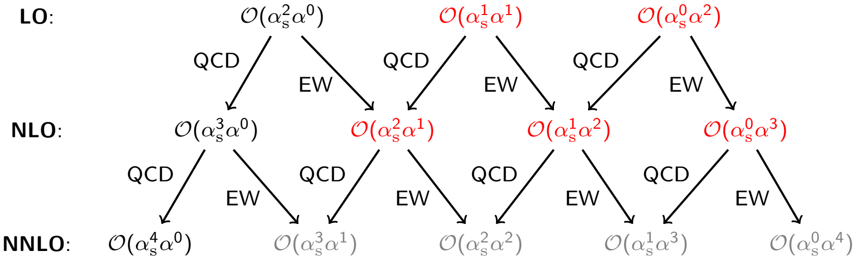
- NNLO QCD corrections included in PDF fits

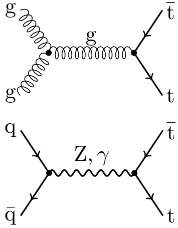
→ include also higher-order α contributions in NNPDF 4.x
  - for all processes
  - check impact of the corrections,
  - be more inclusive, etc.

→ `PineAPPL` supports all higher-orders

- $\mathcal{O}(\alpha_s\alpha^3)$ might become available at some point ...
- lots of progress → talks by A. Behring, A. Sankar, N. Rana, L. Buonocore, A. von Manteuffel

Which EW corrections will be included for $\mathrm{pp} \to \mathrm{t\bar{t}}$?



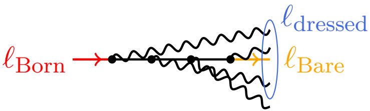


- NNLO QCD corrections included in PDF fits

→ include also higher-order α contributions in NNPDF 4.x
  - for all processes
  - check impact of the corrections,
  - be more inclusive, etc.

→ `PineAPPL` supports all higher-orders

Which datasets do we fit? Final-state radiation (FSR) subtraction



- pre-FSR data/Born leptons: leptons "before they radiate", calculated using shower inversion (PHOTOS), from

→ post-FSR data/dressed leptons: leptons with photons recombined around $\Delta R_{f\gamma}$, typically $\Delta R_{f\gamma} = 0.1$

- bare leptons: non-collinear safe
- dressing factors

$$C_{\text{dress}} = \frac{\mathrm{d}\sigma_{\text{post-FSR}}/\mathrm{d}\mathcal{O}}{\mathrm{d}\sigma_{\text{pre-FSR}}/\mathrm{d}\mathcal{O}}$$

  can be very large, up to 50 % in invariant mass distributions

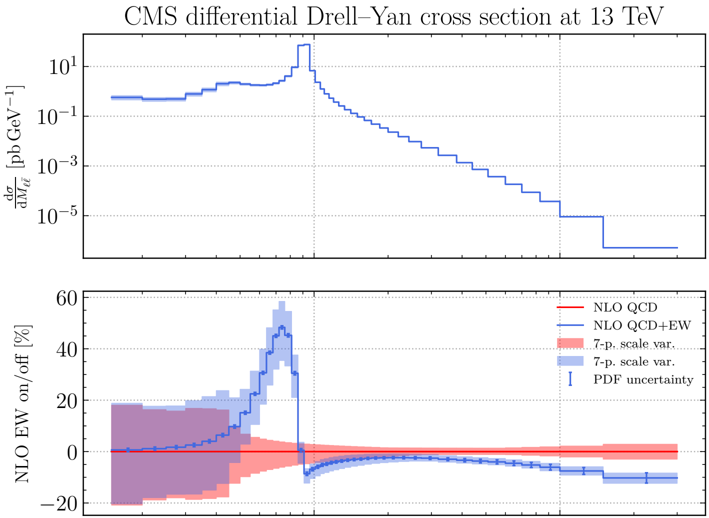
- pre-FSR data for comparisons with QCD-only theory predictions

→ post-FSR data for comparisons with EW corrections (up to one photon at NLO)

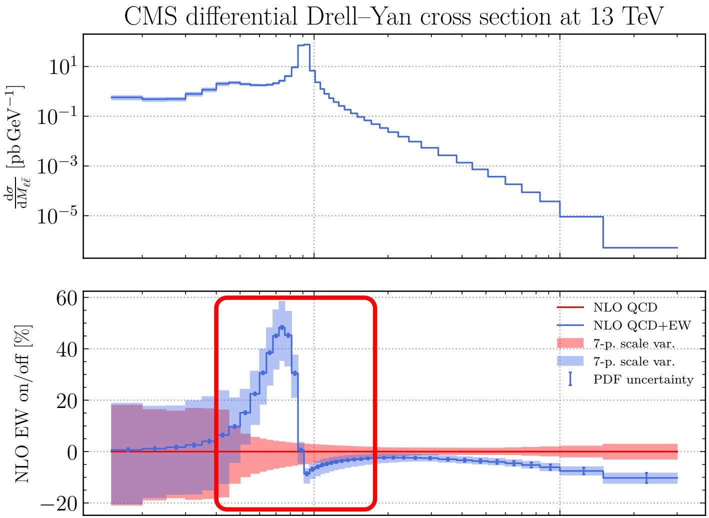
- For some analyses post-FSR data (preferred choice) not published: double counting issue with pre-FSR data!
- Often C_{dress} (+uncertainty) and pre-FSR dataset given ⇒ need to change systematic uncertainties!

→ NLO EW PDF dataset largely determinedy by whether dressed-lepton observables/post-FSR dataset is available

Effects of FSR in Drell–Yan

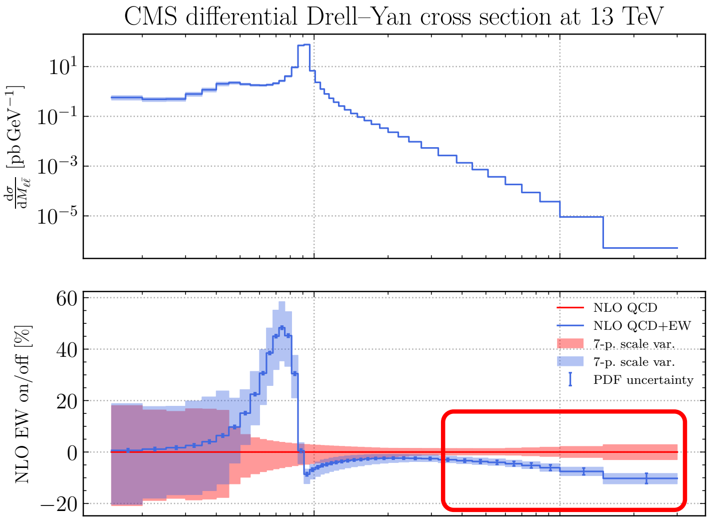


Effects of FSR in Drell–Yan



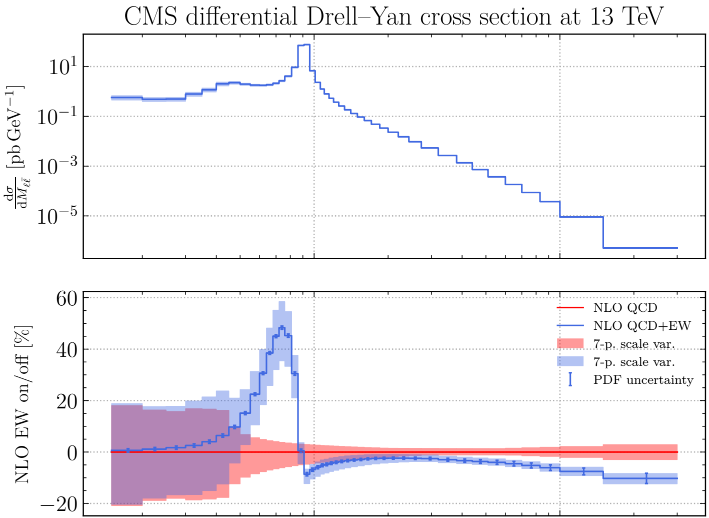
- large distortions of the NLO EW around the Z peak: QED FSR
- very large if bins are small
- subtracted in data in pre-FSR datasets
- scale-variation band much larger: interplay with theory uncertainties in PDF

Effects of FSR in Drell–Yan



- large distortions of the NLO EW around the Z peak: QED FSR
- very large if bins are small
- subtracted in data in pre-FSR datasets
- scale-variation band much larger: interplay with theory uncertainties in PDF
- not described by FSR: weak corrections for large $M_{\ell\ell}$ $\rightarrow$ need full EW corrections

Effects of FSR in Drell–Yan

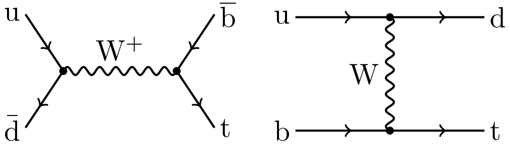


- large distortions of the NLO EW around the Z peak: QED FSR
- very large if bins are small
- subtracted in data in pre-FSR datasets
- scale-variation band much larger: interplay with theory uncertainties in PDF
- not described by FSR: weak corrections for large $M_{\ell\ell}$ $\rightarrow$ need full EW corrections
- Experiments: please publish (also) dressed-lepton observables/post-FSR datasets – we can't use it otherwise

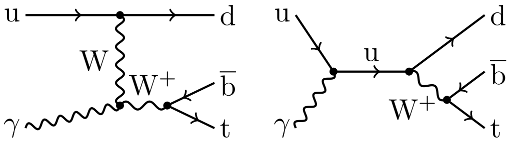
t-channel single-top production

Not properly/easily definable at NLO EW (see also [R. Frederix, D. Pagani, I. Tsinikos]):

- Analyses, e.g. [ATLAS collaboration], treat s-channels as irreducible background
- will be included in NNPDF4.0
- single-production at LO:



- but at NLO EW not (gauge-invariantly) separable:



$\rightarrow$ ignore these datasets

- probably not too important, but see [E.R. Nocera, M. Ubiali, C. Voisey]

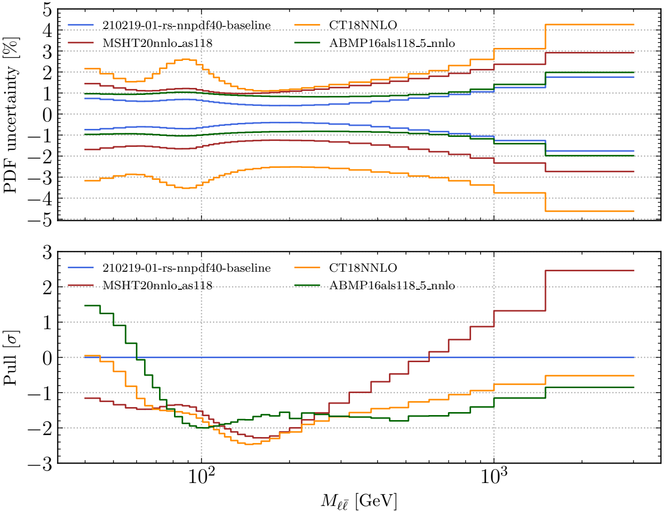
How can I use Madgraph5_aMC@NLO+PineAPPL?

1. Install Madgraph5_aMC@NLO: `https://launchpad.net/mg5amcnlo`
2. Install PineAPPL: `https://github.com/N3PDF/pineappl`
3. Example for generating the DY plot (next page) available at `https://github.com/N3PDF/pineappl/tree/master/examples/mg5amcnlo`

- only one line needs to be added in the mg5amc runcard (rhs)
- No two-phase generation of the grids needed
- Replaces the `aMCfast` [V. Bertone et al.] interface in Madgraph5_aMC@NLO v2.x
- `PineAPPL`'s CLI allows to easily produce convolutions and plots

```
launch processname
[..]
set ptl = 25.0
set etal = 2.5
set mll_sf = 116
set req_acc_FO 0.001
set pineappl True
done
quit
```

Example: DY plot




```
$ pineappl --silent-lhapdf plot \
    DY_14.pineappl.lz4 \
    210219-01-rs-nnpdf40-baseline \
    MSHT20nnlo_as118 \
    CT18NNLO \
    ABMP16als118_5_nnlo > plot.py
$ python3 plot.py
```

- PDF uncertainties for
  - NNPDF 4.0 candidate fit
  - MSHT20
  - CT18 (only main set)
  - ABMP16
- pull: weighted difference w.r.t. NNPDF 4.0 in units of σ

Summary

- NLO EW corrections for PDF fits: upgrading the precision of PDF processes
- `PineAPPL`: interpolation tool/library for storing PDF-independent theoretical predictions `https://github.com/N3PDF/pineappl`
- Built-in support in Madgraph5_aMC@NLO v3.1.0
- data issues: FSR subtraction, photon-initiated subtraction, ...

$\rightarrow$ effectively determines the datasets we can use for the fit

Outlook:

- We will publish grids for ATLAS/CMS/LHCb analyses (PDF processes) soon

$\rightarrow$ `PineAPPL` already public and Open Source: `https://n3pdf.github.io/pineappl`

- NLO EW PDF fit not too far away

Interpolation grids

For PDF fitting we need PDF independent predictions. Use Lagrange interpolation,

$$f_a(x_1, Q^2) f_b(x_2, Q^2) \approx \sum_{i,j,k} f_a(x_i, Q_k^2) f_b(x_j, Q_k^2) L_i(x_1) L_j(x_2) L_k(Q^2),$$

with Lagrange polynomials L_i over the 3D grid $\left\{(x_i, x_j, Q_k^2)\right\}_{i,j,k}$. Insert into master formula:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\mathcal{O}} = \sum_{a,b} \int_0^1 \mathrm{d}x_1 \int_0^1 \mathrm{d}x_2 \int_{Q^2_{\mathrm{min}}}^{Q^2_{\mathrm{max}}} \mathrm{d}Q^2\, f_a(x_1, Q^2) f_b(x_2, Q^2) \frac{\mathrm{d}\sigma_{ab}}{\mathrm{d}\mathcal{O}}(x_1, x_2, Q^2, \mathcal{O})$$

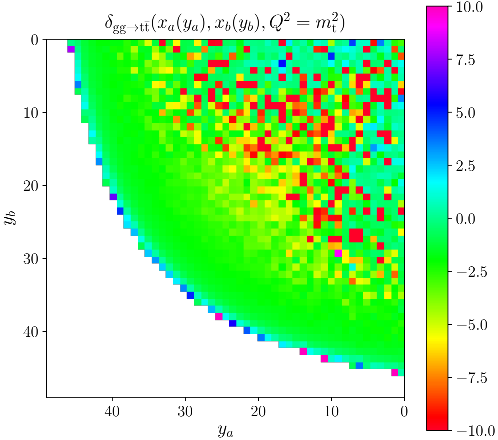
$$= \sum_{a,b} \sum_{i,j,k} \sum_{m,n} f_a(x_i, Q_k^2) f_b(x_j, Q_k^2) \alpha_{\mathrm{s}}^m(Q^2) \alpha^n \frac{\mathrm{d}\Sigma_{abijkmn}}{\mathrm{d}\mathcal{O}}$$

where

$$\frac{\mathrm{d}\Sigma_{abijkmn}}{\mathrm{d}\mathcal{O}} = \int_0^1 \mathrm{d}x_1 \int_0^1 \mathrm{d}x_2 \int_{Q^2_{\mathrm{min}}}^{Q^2_{\mathrm{max}}} \mathrm{d}Q^2\, L_i(x_1) L_j(x_2) L_k(Q^2) \frac{\mathrm{d}\sigma_{ab}^{(i,k)}}{\mathrm{d}\mathcal{O}}(x_1, x_2, Q^2, \mathcal{O})$$

$\rightarrow$ generate $\frac{\mathrm{d}\Sigma_{abijkmn}}{\mathrm{d}\mathcal{O}}$ once, perform PDF convolutions very quickly off-line

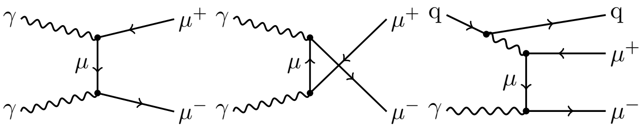
Example: $\Sigma_{\mathrm{gg}ij021}/\Sigma_{\mathrm{gg}ij020}$, $\mathcal{O}(\alpha_\mathrm{s}^2\alpha)/\mathcal{O}(\alpha_\mathrm{s}^2)$ for $\mathrm{gg} \to \mathrm{t\bar{t}}$ @ 8 TeV



- no interpolation in y_a, y_b, or Q^2
- correction for ixs roughly $-0.5\,\%$
- $y_{a/b}(x) = -\ln x_{a/b} + 5(1 - x_{a/b})$, $y(1) = 0$

- lower left corner → production threshold
- at threshold: Coulomb singularity
- $y_a \leftrightarrow y_b$ symmetry: initial-state symmetry of $\mathrm{gg} \to \mathrm{t\bar{t}}$
- negative correction for larger x_a, x_b

Subtraction of photon–photon contribution

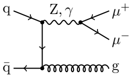


- For ATLAS and CMS it seems to be standard procedure to subtract double-photon induced contributions:

  The photon-induced process, $\gamma\gamma \to \ell\bar{\ell}$, is simulated at LO using Pythia 8 and the MRST2004qed PDF set.

- I am not sure why this is done
- This is a problem: proton contains photons, should be counted towards signal!
- Size of the LO contribution can become significant in large-invariant-mass bins (3 %) depending on the used PDF—up to twice as large for pre-LUXQED photon PDFs

Z transverse momentum



$$\mu = M_{\mathrm{Z}} \text{ vs. } \mu = \sqrt{M_{\mathrm{Z}}^2 + (p_{\mathrm{T}}^{\ell\bar{\ell}})^2}$$

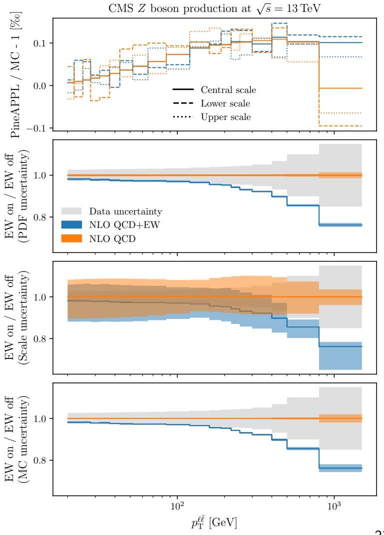
- FSR issues similar to DY
- no photon subtraction

static scale:

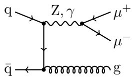
- accidental cancellation of NLO QCD correction → uncertainty band shrinks
- NLO EW are artificially enhanced because of normalisation

dynamic scale:

- scale variation is stabilised
- still significant EW corrections, comparable to data uncertainty



Z transverse momentum



$$\mu = M_{\mathrm{Z}} \text{ vs. } \mu = \sqrt{M_{\mathrm{Z}}^2 + (p_{\mathrm{T}}^{\ell\bar{\ell}})^2}$$

- FSR issues similar to DY
- no photon subtraction

static scale:

- accidental cancellation of NLO QCD correction → uncertainty band shrinks
- NLO EW are artificially enhanced because of normalisation

dynamic scale:

- scale variation is stabilised
- still significant EW corrections, comparable to data uncertainty