Towards NNFF2.0

Tanishq Sharma

DIS2026

MICHIGAN STATE
UNIVERSITY

Outline

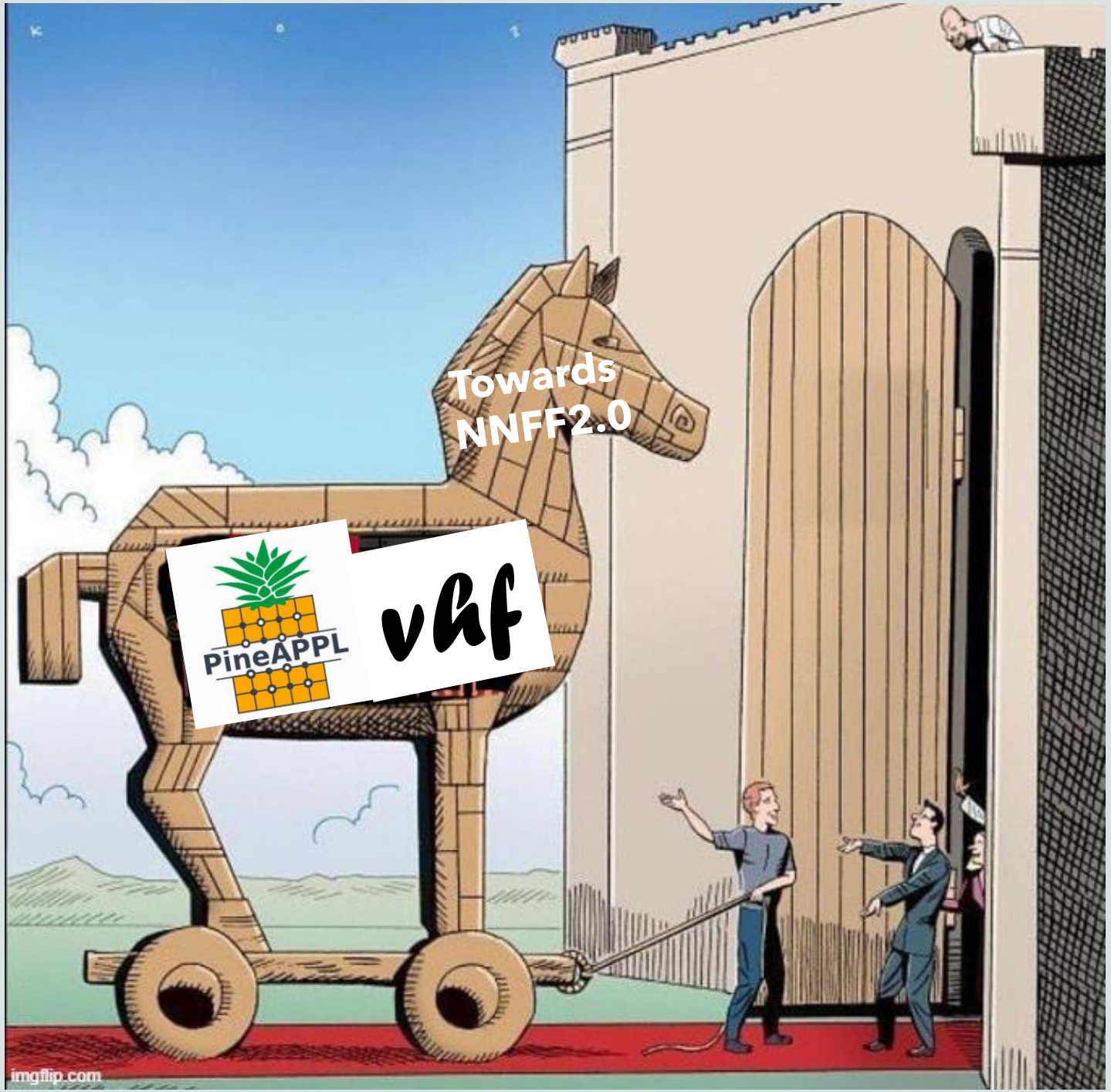
- Overview of the landscape of fragmentations functions
- Development of software tools
- Expected features of NNFF2.0

Overview of the FF landscape and motivation for NNFF2.0

- Several sets exist: NNFF, MAPFF, JAM, BDSSV, …
- Neural network based: NNFF (SIA only), MAPFF (aNNLO)
- There have been many new developments on the theory side
- Upcoming EIC will have a physics program dedicated to a detailed understanding of hadronic structure
- This motivates not only a new and updated fragmentation function determination but also development of software tools that leverage all the latest developments for such endeavors

Development of software tools

- Implementation of time-like splitting functions in EKO
- A new and robust new version of PineAPPL
- Development of Virtual Hadron factory (vhf): a tool for computation of theory predictions of SIA, SIDIS and DIS


Towards
NNFF2.0
PineAPPL
vhf
imgflip.com

PineAPPL: an interpolation grid library

$$\frac{d\sigma}{d\mathcal{O}} = \int_x^1 \frac{d\hat{x}}{\hat{x}} f\left(\frac{x}{\hat{x}}, \mathrm{Q}^2\right) C(\hat{x}, \mathrm{Q}^2)$$

based on piece-wise Lagrange polynomials

pre-compute over an interpolation basis, thus creating an array of weights which can be contracted with an array of PDF values at the interpolation nodes to obtain the final value of the observable

PineAPPL: the original version

- An interpolation grid is a 3d object of bins, orders and channels where each element in this 3d space is a sub-grid of weights
- Orders given by: $\alpha_s^k \alpha^l \log^m\left(\frac{\mu_R^2}{Q^2}\right) \log^n\left(\frac{\mu_F^2}{Q^2}\right)$, such that each combination of k,l,m and n is a single order
- Channels given by partonic flavors, such that $\sum_i \sum_j p_i * p_j$
- The structure of sub-grids fixed to 3 dimensions: Q^2, x_1, x_2
- This allowed for computation of interpolation grids for LHC processes with electroweak corrections, and flexible scale variations
- This is great for PDFs but FFs need more…

PineAPPL v1.0.0

- Global analysis of fragmentation function requires single inclusive annihilation, semi-inclusive deep inelastic scattering and hadron production in proton-proton collisions
- These processes require subgrids of dimension two, three and four respectively
- Furthermore, factorization of hard cross sections and fragmentation functions introduces a new scale: the fragmentation scale
- This requires changing the structure of grid (to introduce the fragmentation scale) and changing the structure of subgrids (to allow for flexible dimensionality)

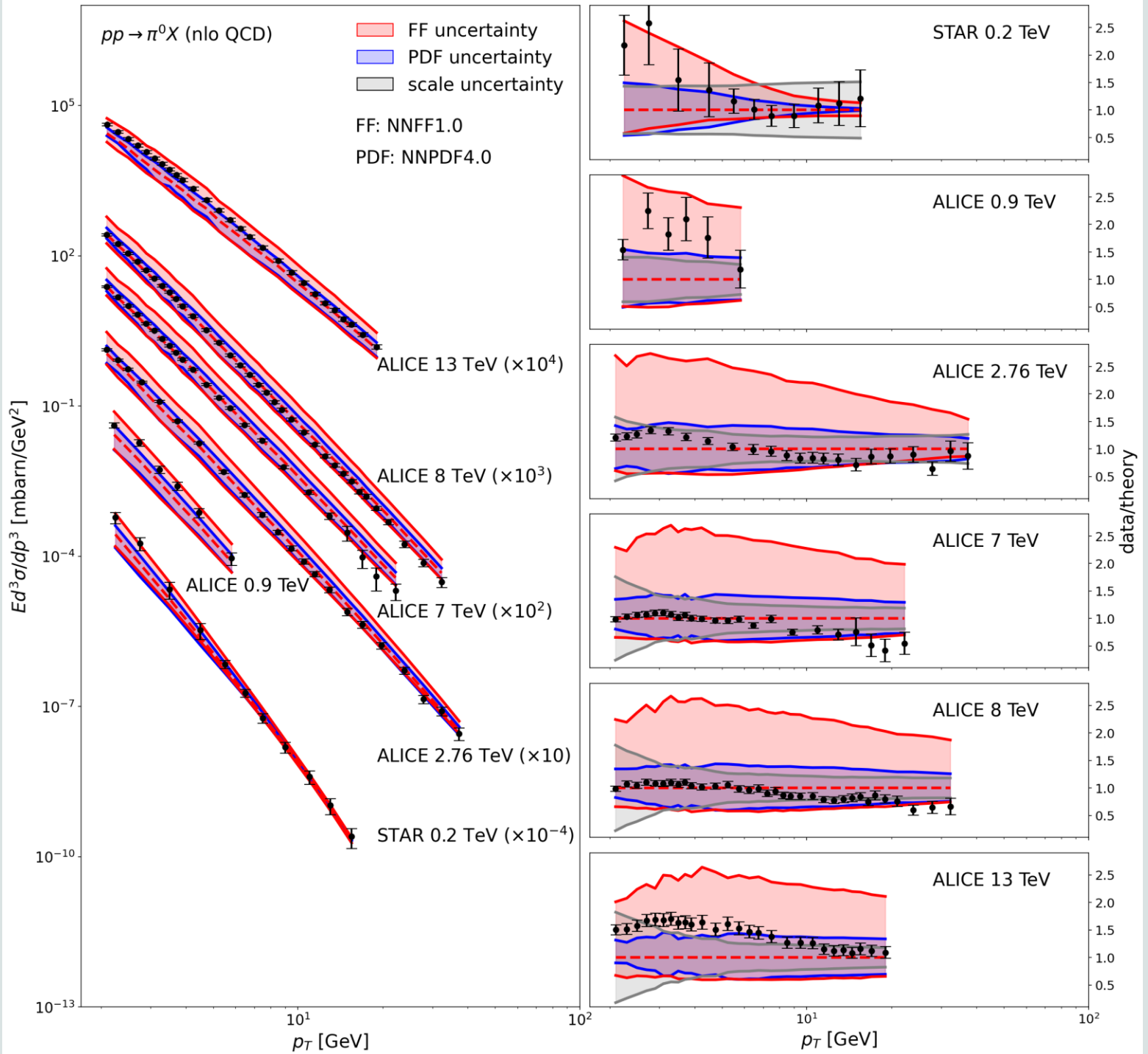
PineAPPL v1.0.0

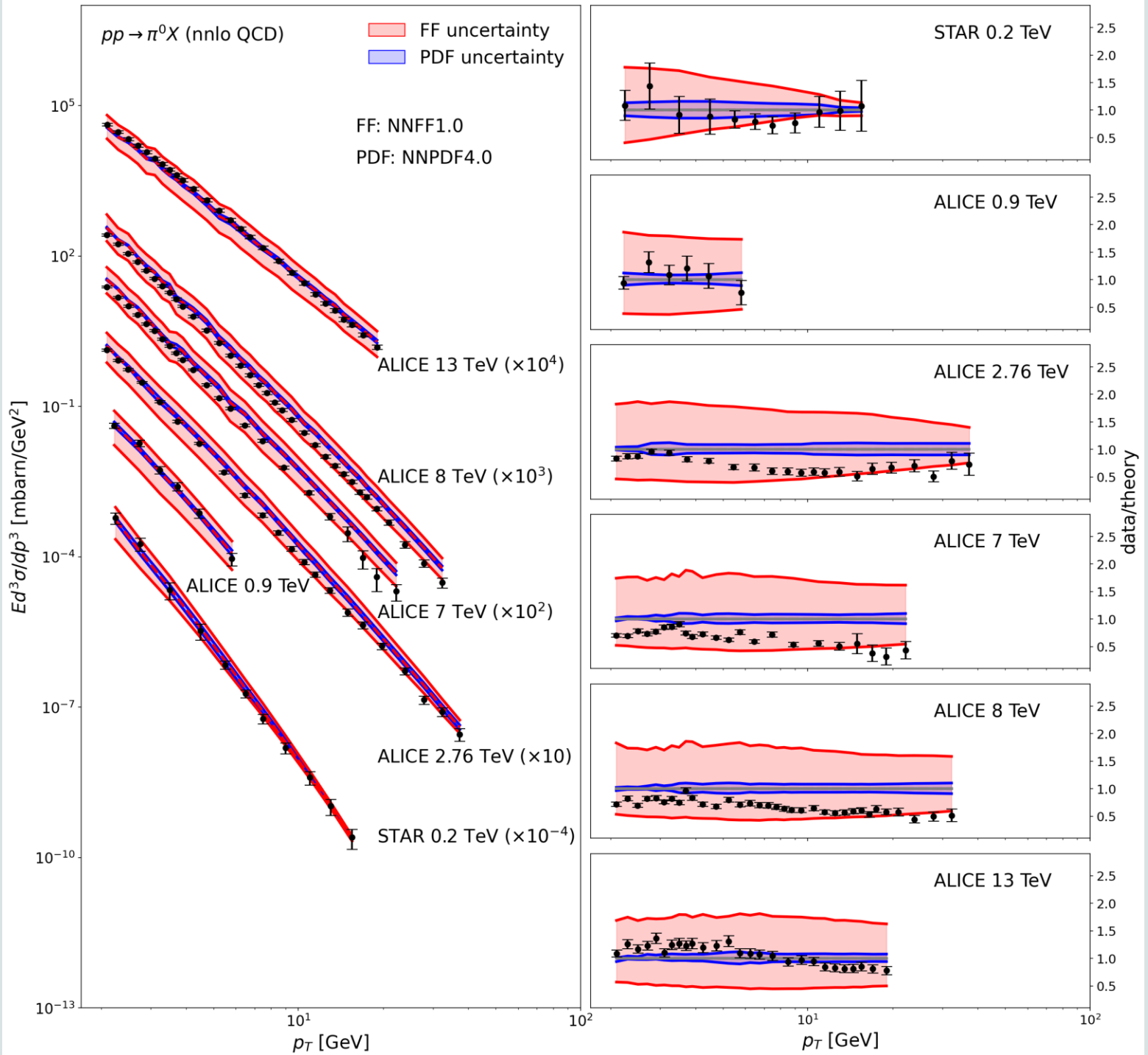
- We overhauled the entire library to make PineAPPL support arbitrary number of convolutions and fragmentation scale, making it highly flexible and future ready (think EIC)
- Expect a paper soon…

PineAPPL v1: fast and flexible theory predictions for present and future colliders

Tomáš Ježo[1], Emanuele R. Nocera[2], Tanjona R. Rabemananjara[3,4], Christopher Schwan[5,6], Tanishq Sharma[2,3,4], and Jan Wissmann[1]

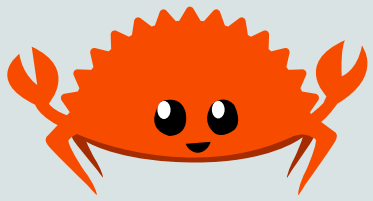
Note: tentative title^

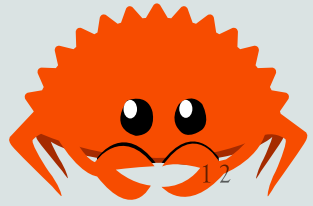

$pp \to \pi^0 X$ (nlo QCD)
FF uncertainty
PDF uncertainty
scale uncertainty
FF: NNFF1.0
PDF: NNPDF4.0
ALICE 13 TeV (×10^4)
ALICE 8 TeV (×10^3)
ALICE 0.9 TeV
ALICE 7 TeV (×10^2)
ALICE 2.76 TeV (×10)
STAR 0.2 TeV (×10^-4)
$Ed^3\sigma/dp^3$ [mbarn/GeV^2]
p_T [GeV]
STAR 0.2 TeV
ALICE 0.9 TeV
ALICE 2.76 TeV
ALICE 7 TeV
ALICE 8 TeV
ALICE 13 TeV
data/theory
p_T [GeV]


$pp \to \pi^0 X$ (nnlo QCD)
FF uncertainty
PDF uncertainty
FF: NNFF1.0
PDF: NNPDF4.0
ALICE 13 TeV (×10^4)
ALICE 8 TeV (×10^3)
ALICE 0.9 TeV
ALICE 7 TeV (×10^2)
ALICE 2.76 TeV (×10)
STAR 0.2 TeV (×10^−4)
$Ed^3\sigma/dp^3$ [mbarn/GeV^2]
p_T [GeV]
STAR 0.2 TeV
ALICE 0.9 TeV
ALICE 2.76 TeV
ALICE 7 TeV
ALICE 8 TeV
ALICE 13 TeV
data/theory
p_T [GeV]

vhf: Virtual Hadron Factory

- Built out of a need for a code to be able to produce theoretical predictions (actual calculations as well as PineAPPL grids) for SIA and SIDIS
- Written in Rust: highly efficient, parallelizable and future proof code
- Interfaced to PineAPPL and NeoPDF
- Uses a run card approach (as opposed to API approach) thus making the usage very straightforward with no learning curve
- Expect a detailed paper, in the near future …



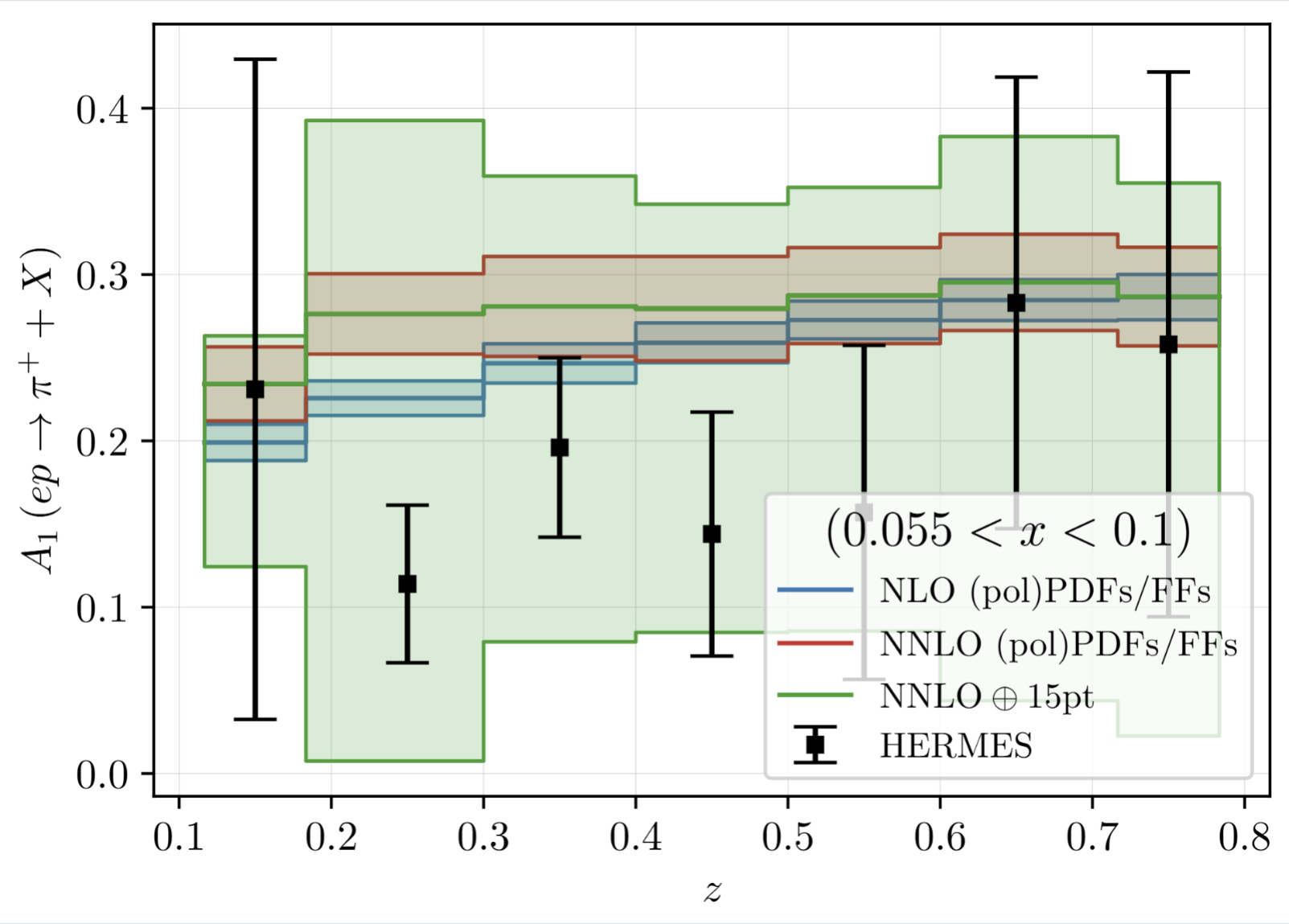


vhf > examples > test_run > ! runcard.yaml

```yaml
# Metadata and Run Settings
run_name: "HERA_F2_Test_Run"
output_type: "InterpolationGrid"
qcd_perturbative_order: 2  # 0=LO, 1=NLO, etc.
scale_variations: true
unit_prefix: "pb"

# Theory Settings
fns: "ZMVFNS"

# Process Definition
process:
  process_type: "DIS_NC"
  observable: "F2"
  lepton: 11           # Electron
  hadron: 2212         # Proton
  electroweak_interaction_type: "em"

# Binning Configuration
bins:
  xz_binning_type: "point"
  q_binning_type: "point"
  data:
    # KinVar::dis { x, q }
    - x: { val: 0.01 }
      q: { val: 10.0 }

    - x: { val: 0.01 }
      q: { val: 20.0 }

    - x: { val: 0.05 }
      q: { val: 50.0 }

    - x: { val: 0.1 }
      q: { val: 100.0 }

```

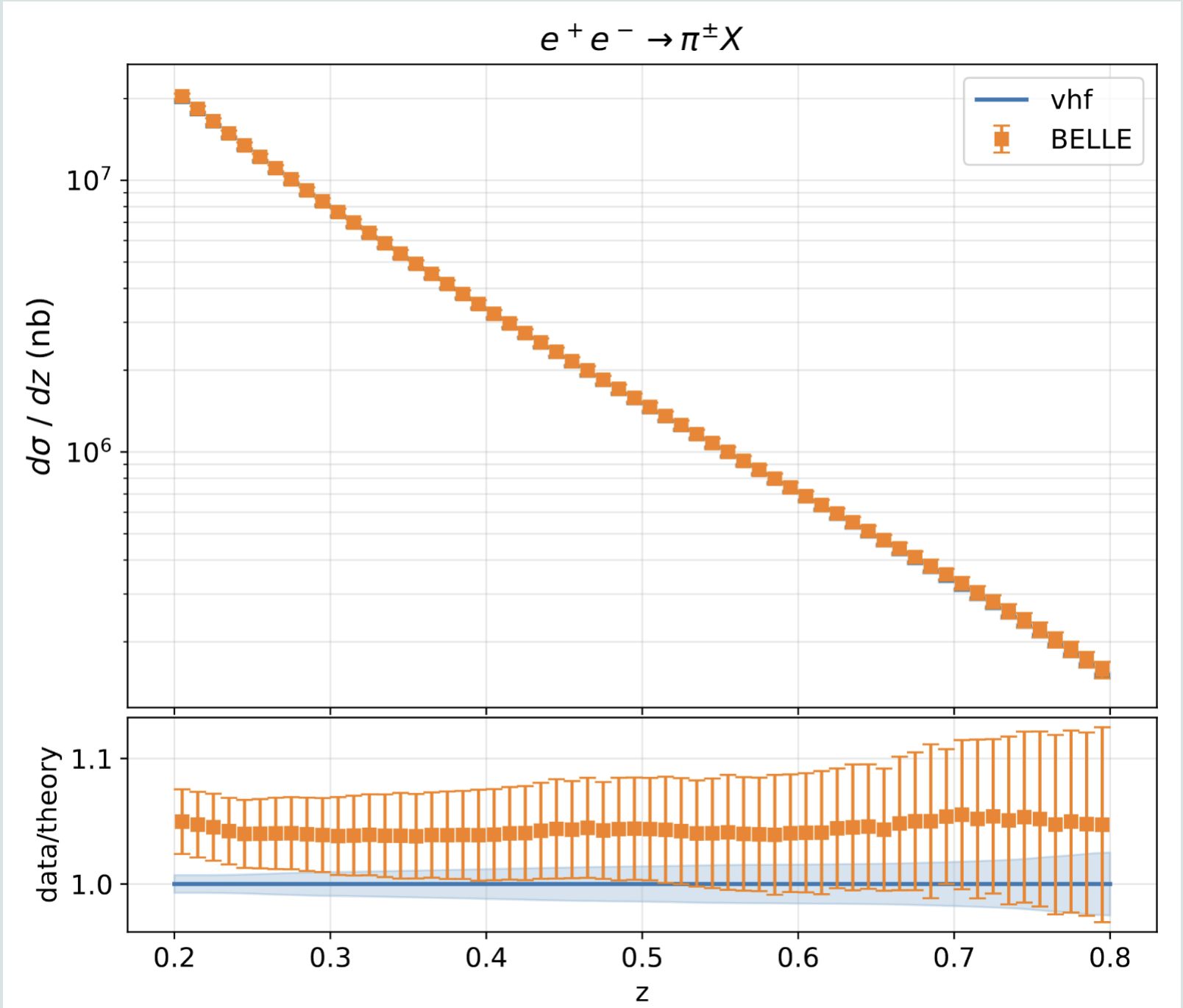
vhf: Virtual Hadron factory

- NC SIA structure functions: F_T, F_L & F_A
- NC/CC DIS structure functions: F_L, F_2, F_3, g_1, g_4 & g_L
- NC/CC SIDIS structure functions: F_T, F_L, F_3, g_1, g_T & g_L
- Allows for computation of structure functions, differential cross sections, and multiplicities
- ZM-VFNS is the only available scheme at the moment, but DIS to ultimately be extended to GM-VFNS to allow for its usage for stand alone DIS theory predictions too



PDF: NNPDF4.0
pPDF: NNPDFpol2.0
FF: MAPFF1.0

*Note: error bands only include 1 sigma uncertainty of pPDF


$e^+e^- \to \pi^\pm X$
vhf
BELLE
10^7
10^6
$d\sigma$ / dz (nb)
data/theory
1.1
1.0
0.2
0.3
0.4
0.5
0.6
0.7
0.8
z

FF: NNFF1.0

Current status of vhf

- The core library if fully complete
- This includes a self contained and independent implementation of many ingredients: Lagrange interpolation, Nielsen's generalized polylogs and harmonic polylogs, adaptive quadrature integrator, and more…
- The current focus is on automation, optimization and code refactoring for clarity and transparency
- The tool will be made publicly available for open use (though exact nature to be determined: container-based distribution vs open-source code?)

NNFF2.0: what to expect?

- A complete NNLO global analysis of pion, kaon (and possibly proton) collinear fragmentation functions including SIA, SIDIS and pp->HX experimental data
- Inclusion of missing higher order uncertainties (through scale variation)
- Neural network architecture determined by hyper-parameter optimization
- Most importantly: this will bring us one step closer to a future goal of a simultaneous determination of PDFs and FFs
- This will also allow the next generation of NNPDFpol PDF set to include SIDIS data in the pPDF extraction

Lots of work done, but lots more to do, so keep an eye out for updates towards this endeavor… 😊