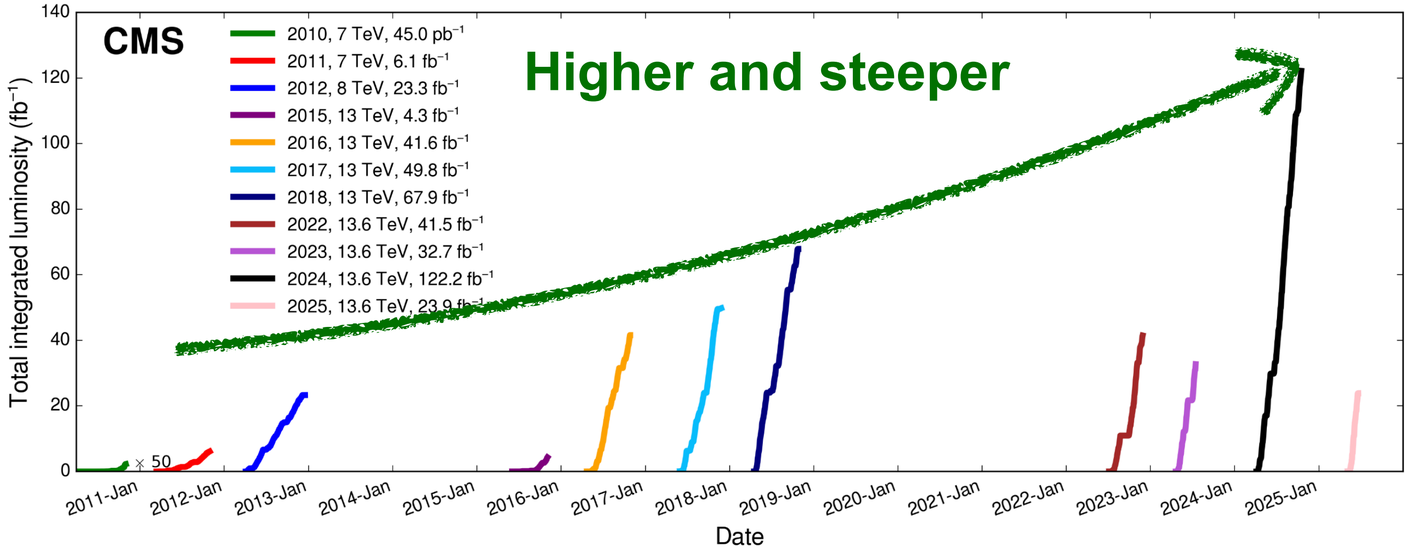
Physics Object Performance in CMS

Raffaella Tramontano
on behalf of the CMS Collaboration
EPS 2025 - Marseille
10-07-2025

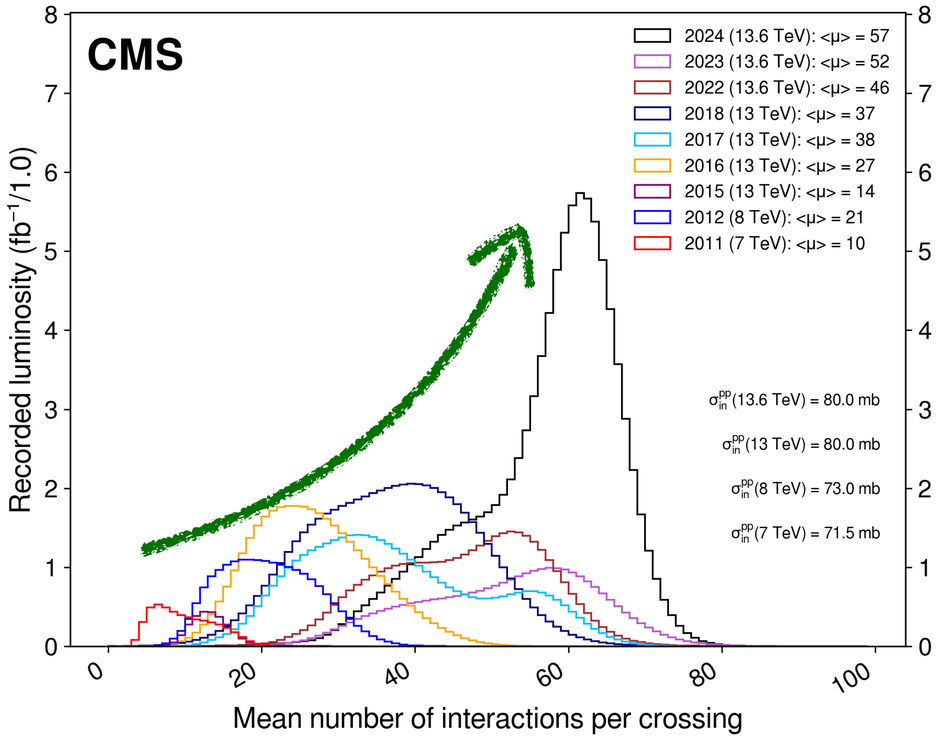
LHC Run3: records and challenges

CMS Lumi Public Results



More collisions, for longer!

How do CMS adapts to face the challenge?



- **higher instantaneous luminosities**
  - Increase number of interactions per bunch crossing - pile up

- **Lumi leveling throughout fill:**
  - Larger datasets in same recording time

Objects recipes

Reconstruction calibrations

- Radiation damage
- Macroscopic detector issues
- (Dead/noisy/masked channels)

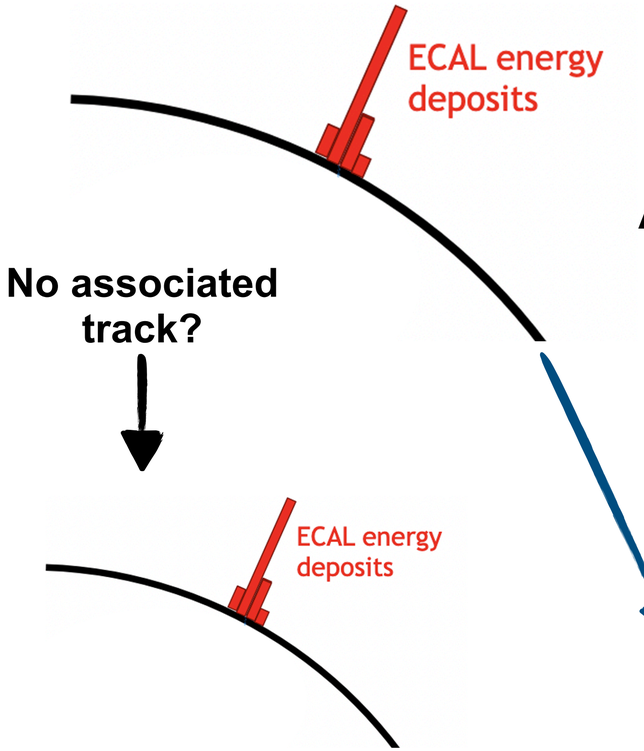
Identification

Non-standard phase spaces
High end ML techniques
in depth object characterization

Simulation Corrections

Material, radiation damage evolution
Known mismodeled variables
PU scenario dinamically changing

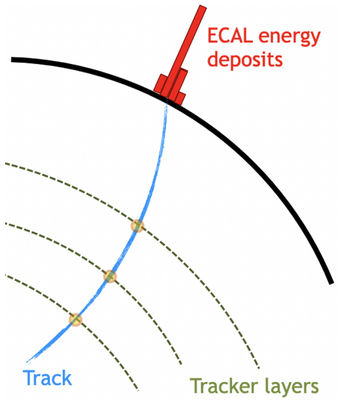
Electrons and photons



No associated track?

A photon!

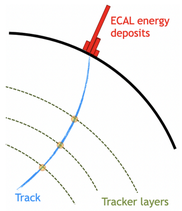
Associated track?



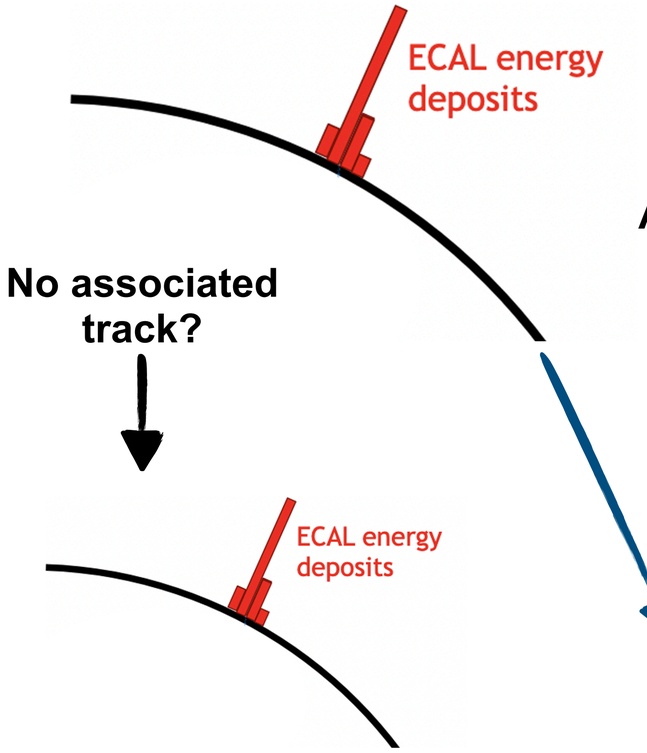
An electron!

Does the electron Brehm?

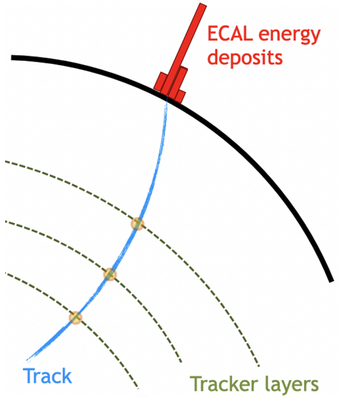
What if the photon converts?



Electrons and photons

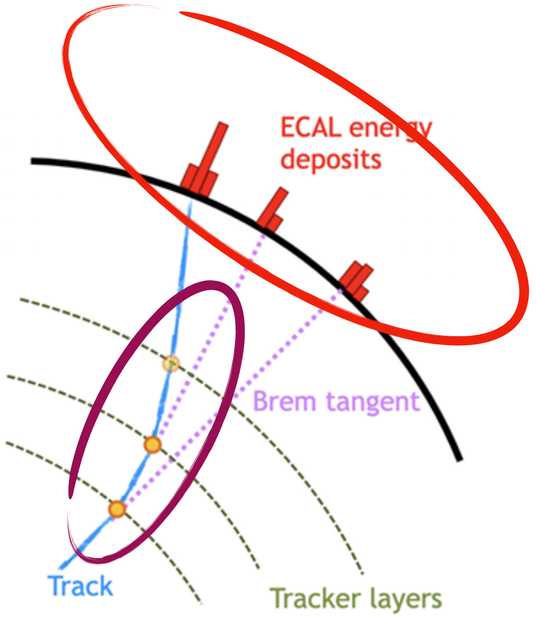


Associated track?



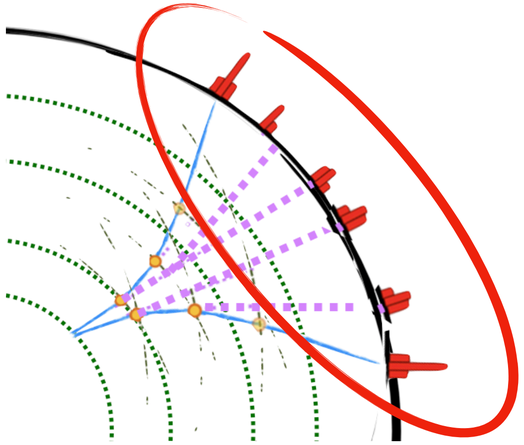
An electron!

Does the electron Brehm?



Dedicated tracking algorithm (GSF)
Recover from
Brehm kinks

What if the photon converts?

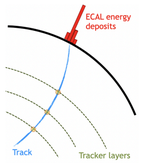


Dedicated clustering algorithms
patch all the energy deposits
together

Energy deposit regression + E-p combination
pT and E best estimates

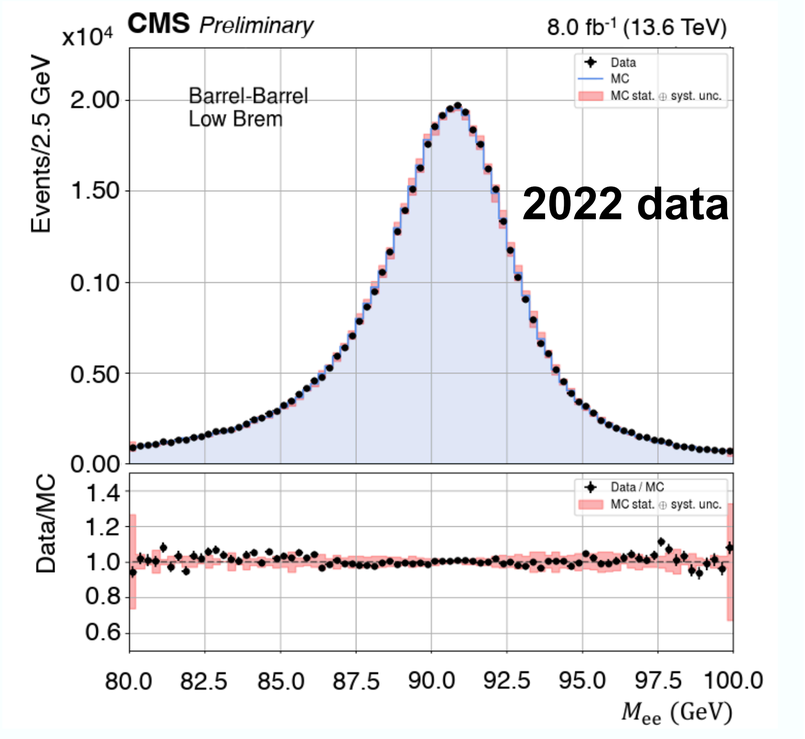
Simulation corrections

DP-2024-052



Scale and smearing corrections

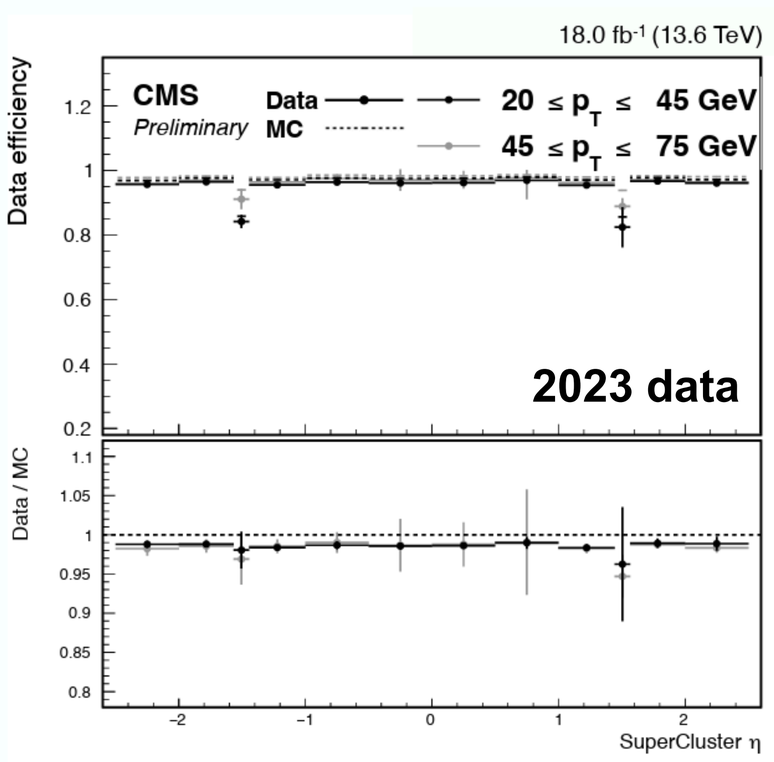
Residual energy corrections computed Z -> e+ e- standard candle



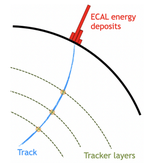
ECAL energy measurement:
shifted in DATA to match the SIM peak position
smeared in SIM to match DATA resolution

Reconstruction efficiency Scale Factors (electrons)

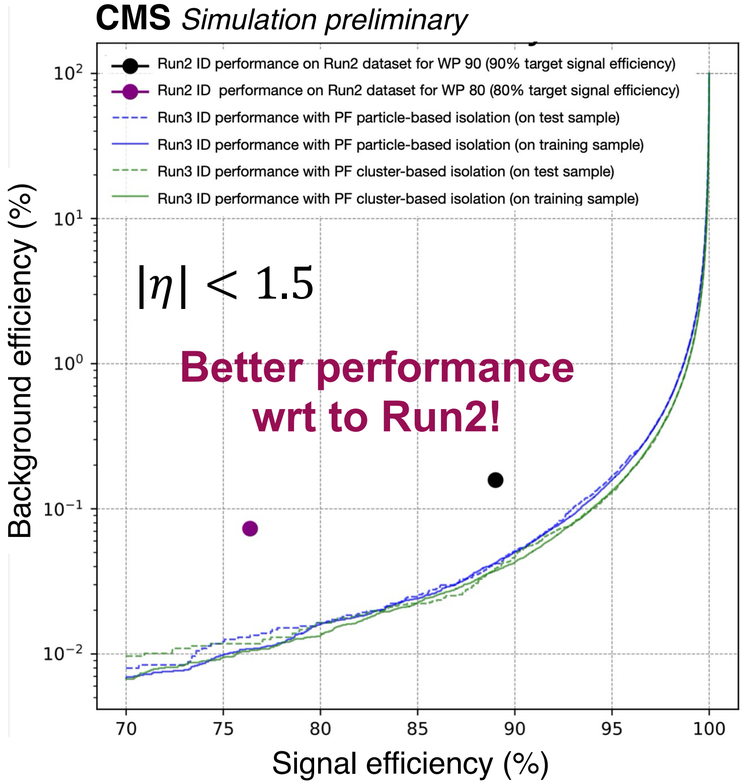
How many times does a Z->e+ e- energy deposit in ECAL has a corresponding associated track?



New identification algorithms



Run3 standard electron ID

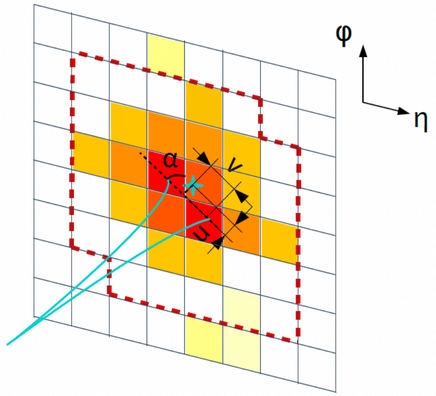


DP-2024-052

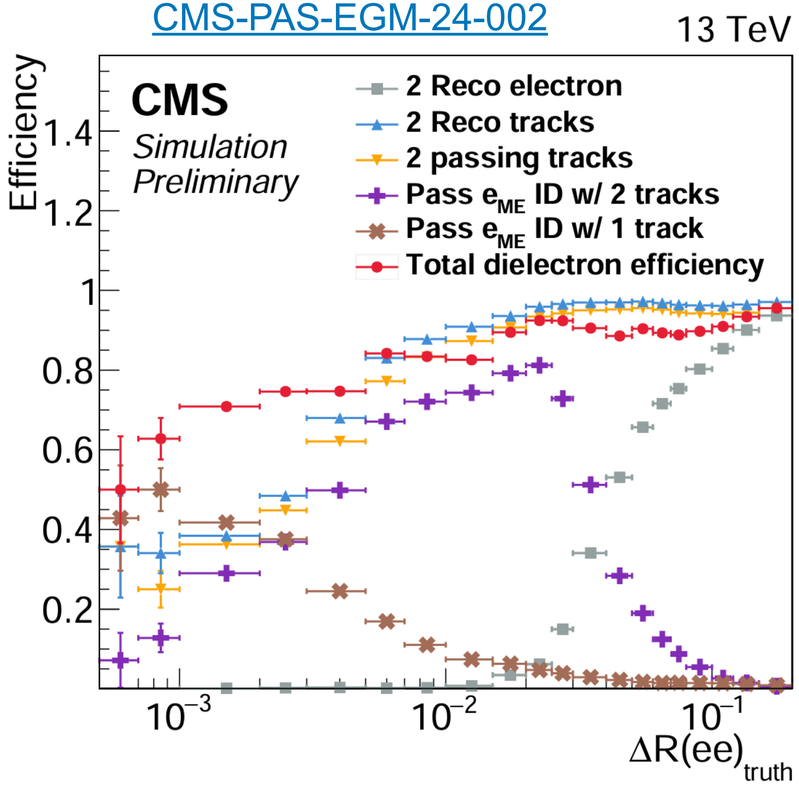
Highly boosted di-electron pairs identification

NEW!

1-2 tracks + single merged ECAL cluster



CMS-PAS-EGM-24-002



Di-track

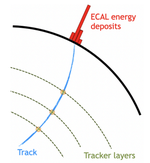
Standard di-electron

2 trk merged ID

1 trk merged ID

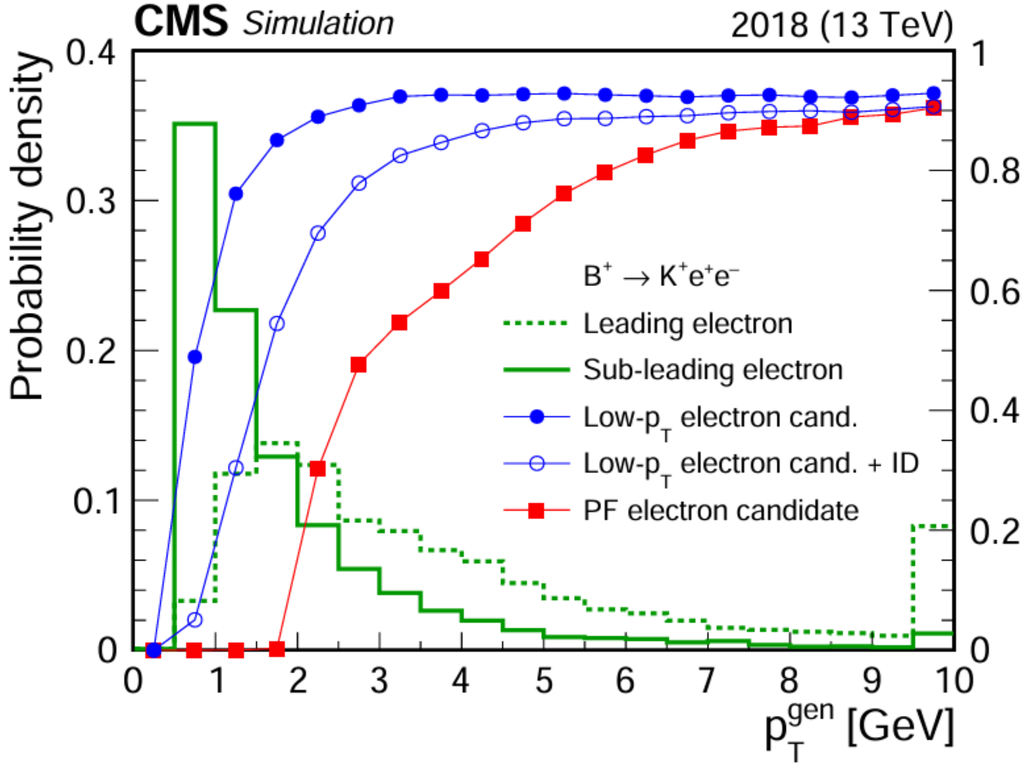
EGM novelties:

arXiv:2403.16134



Novel phase spaces:
Low Pt electron collection

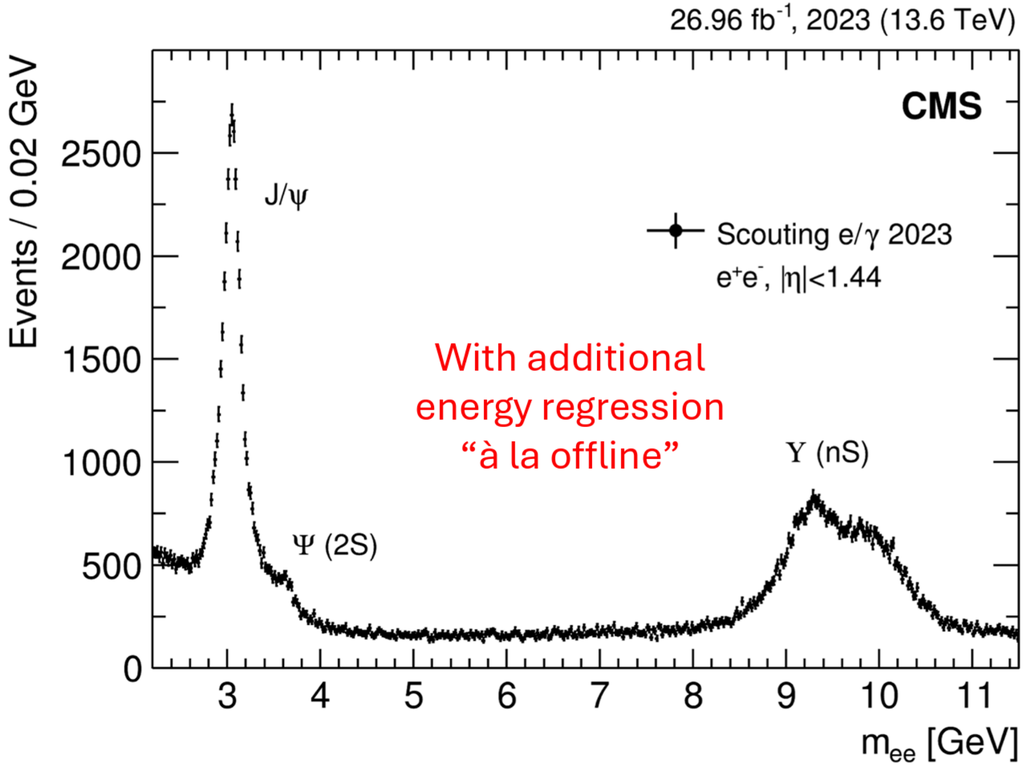
GSF track with looser track seed
Enhanced reconstruction efficiency in [1-10] GeV pT range
Fight rising background via dedicated ID



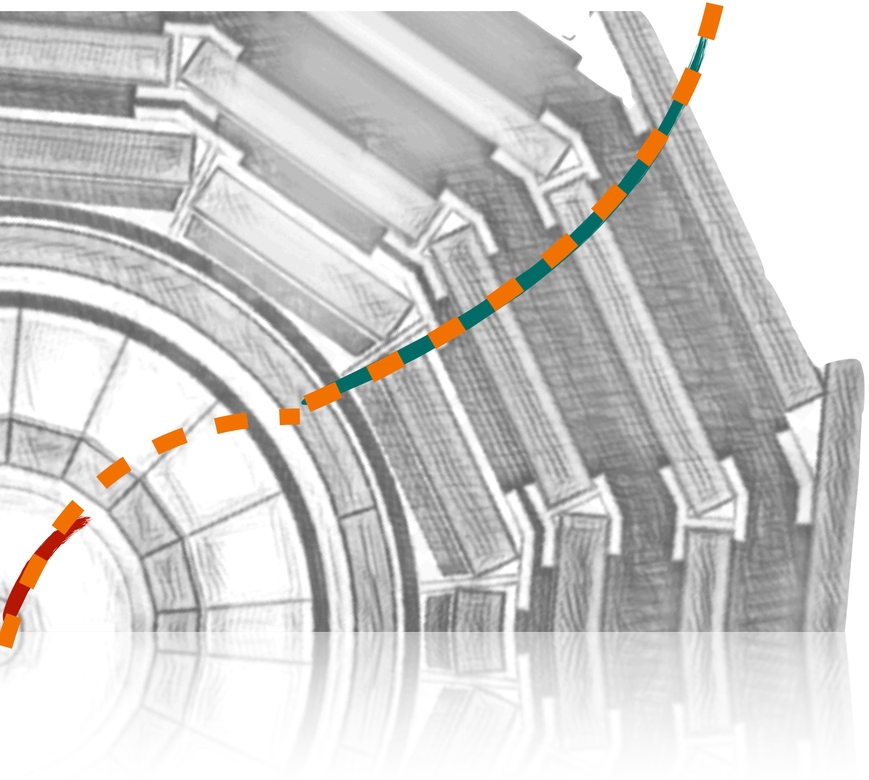
From Run3 available centrally in all CMS datasets!

Novel data rates:
Scouting electrons

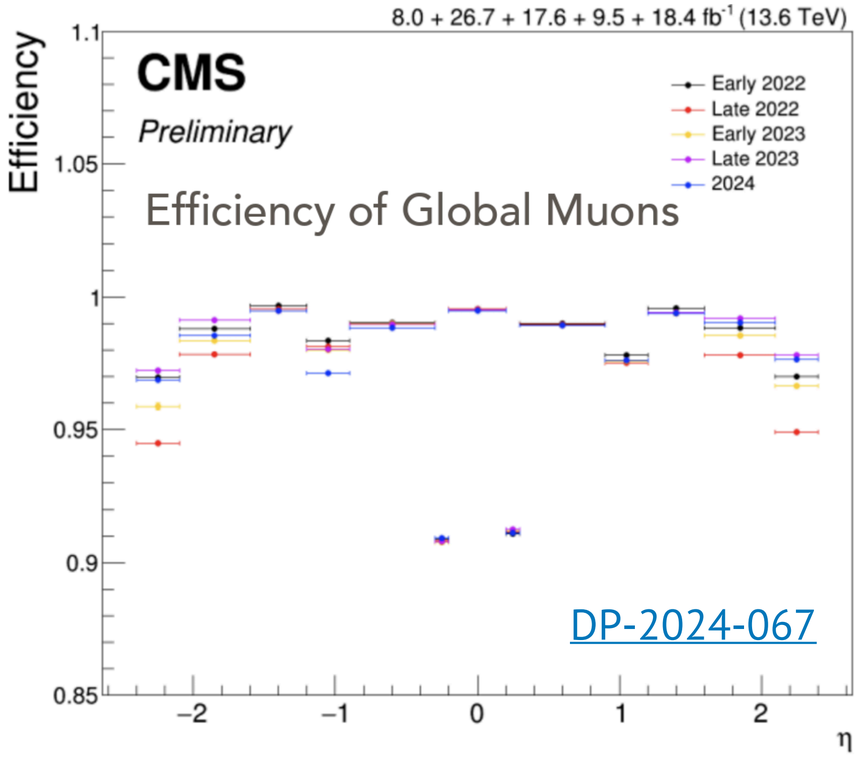
Limited event content information
Prompt HLT recostruction
Smaller event size = lower HLT thresholds!



Muon reconstruction



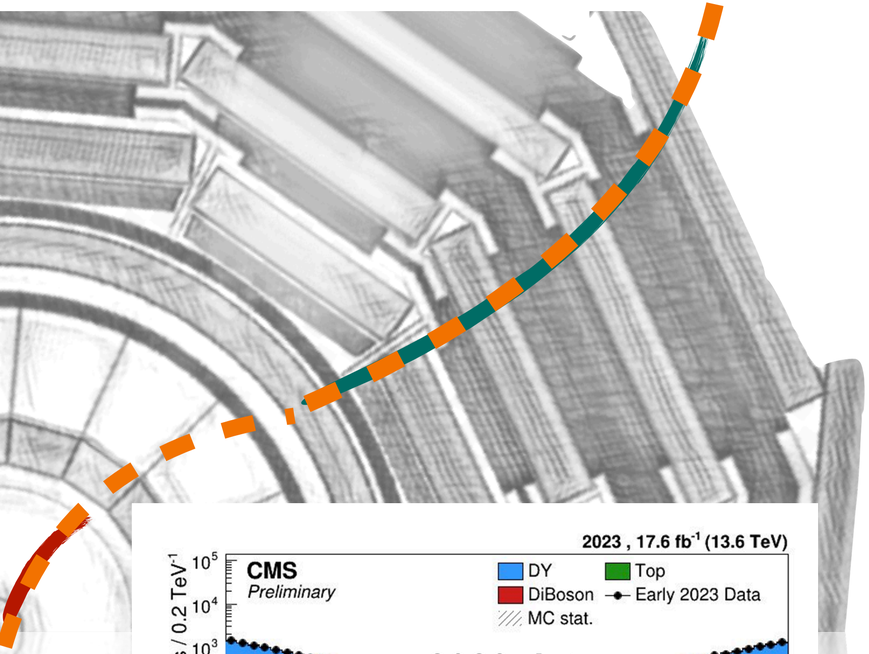
Global muon reconstruction efficiency > 95%!

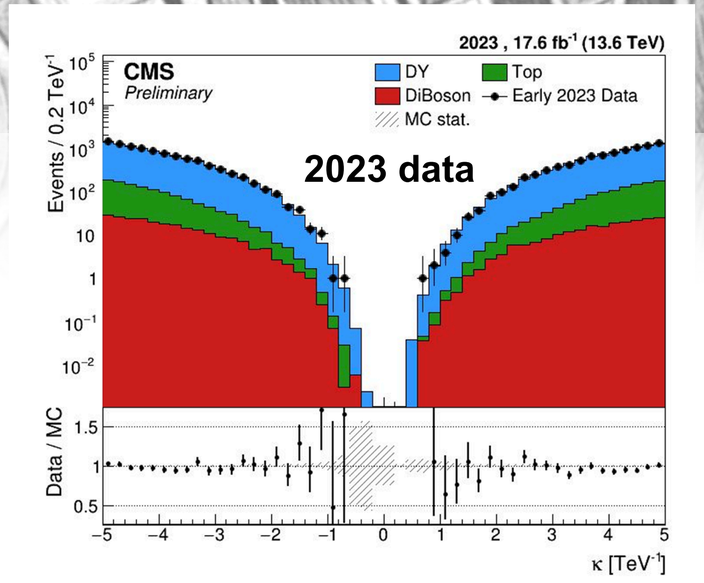


- **Tracker Muon:** Track + loose match to Muon system detectors
- **Standalone Muon:** Kalman Filter in the muon system only
  - LLP proxy
- **Global Muon:** outside in standalone-to-tracker match

Muon scale and smearing corrections

DP-2024-065





High pT: q/Pt method only

Medium pT: two steps procedure

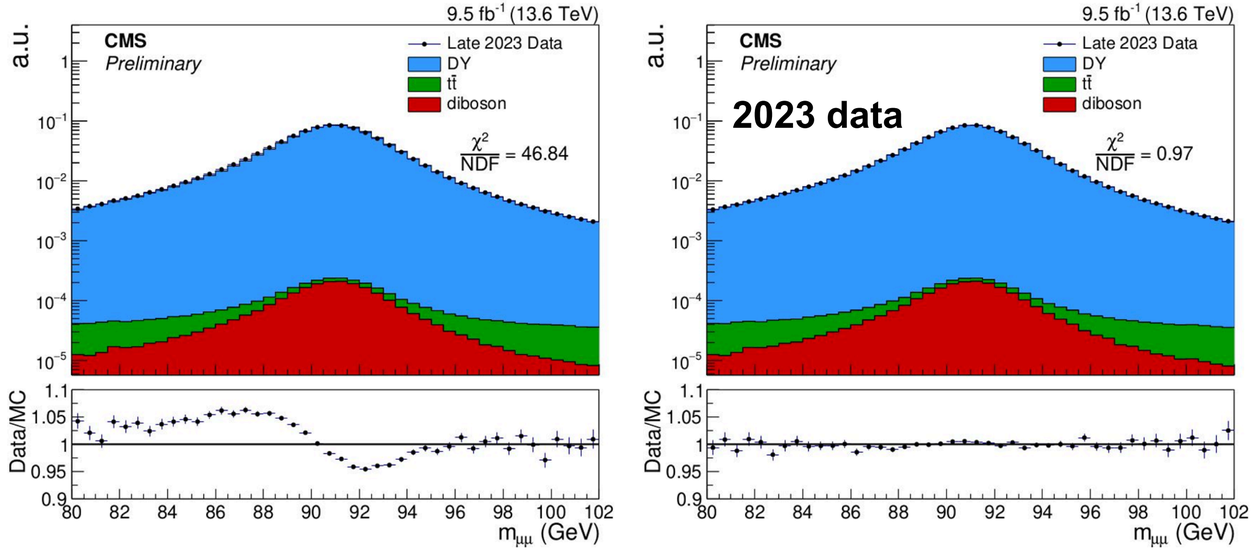
1. 1/pT correction per muon

- Assumes same behavior for μ^+ μ^- in
  - Kinematics
  - Sim vs data

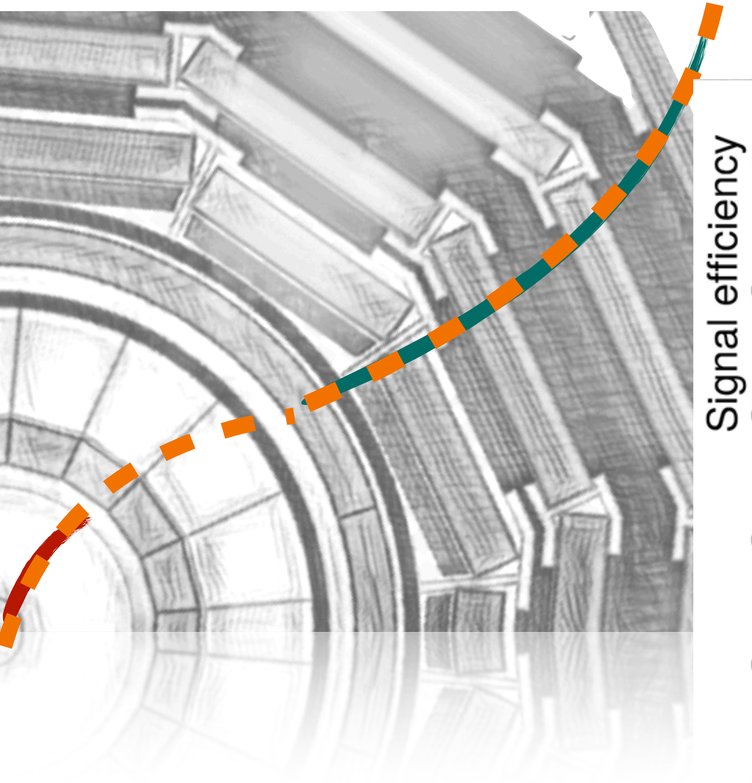
a. $\langle 1/p_T \rangle_{gen,\,neg} = M \langle 1/p_T \rangle_{reco,\,neg} - A$

b. $\langle 1/p_T \rangle_{gen,\,pos} = M \langle 1/p_T \rangle_{reco,\,pos} + A$

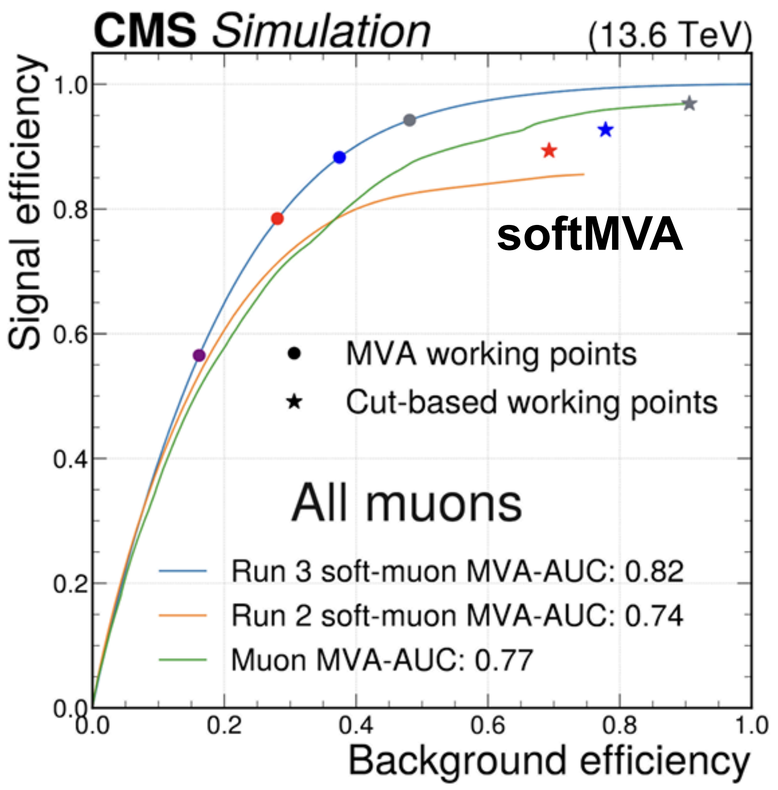
2. $Z \rightarrow \mu\mu$ peak position and resolution

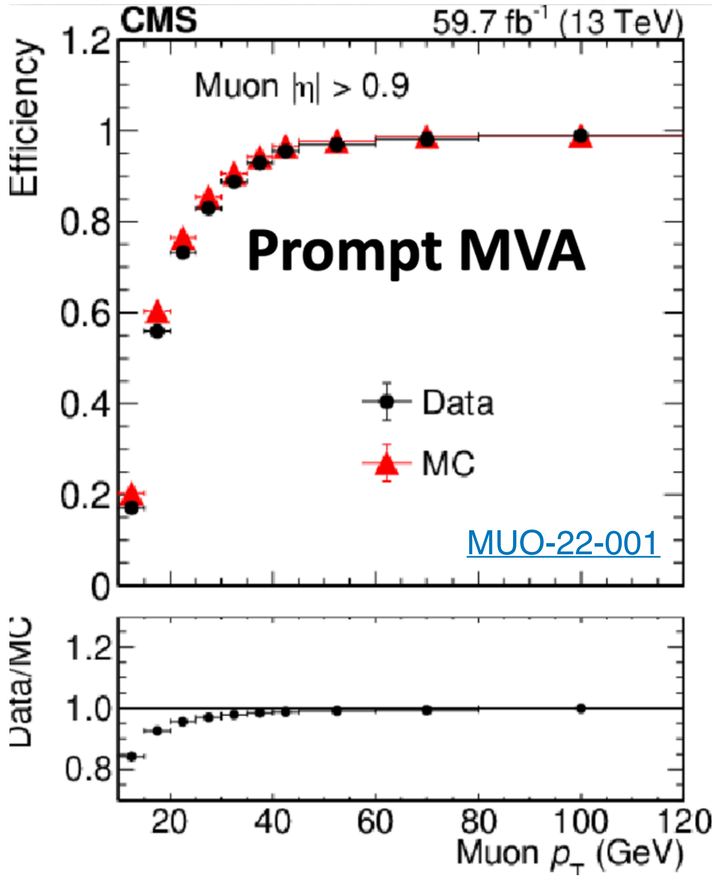


Same muons, different phase spaces:



arXiv:2412.17590





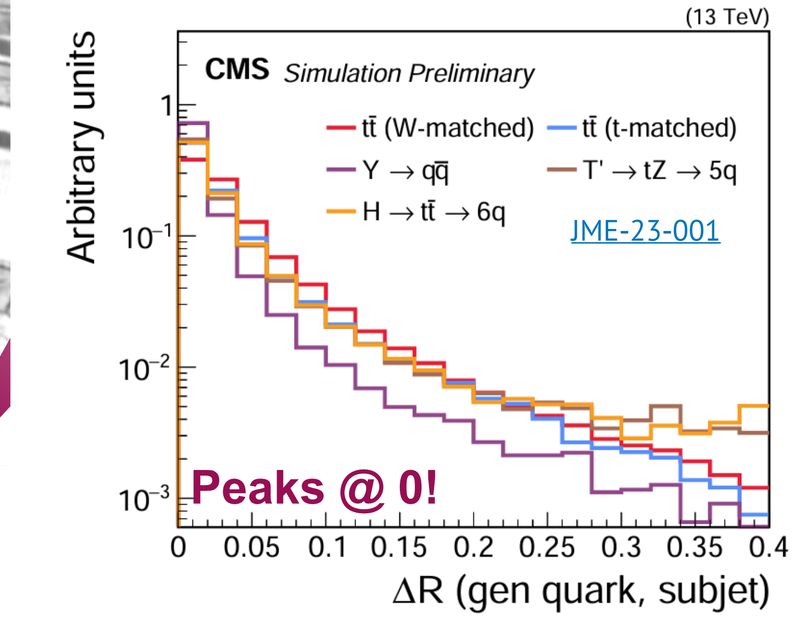
Less strict on prompt requirements
Optimized for B-meson/low-energy physics
Retrained for Run3 with better performance!

Specific for prompt SM processes:
W,Z H, τ decays,

Jets - Run3 standards

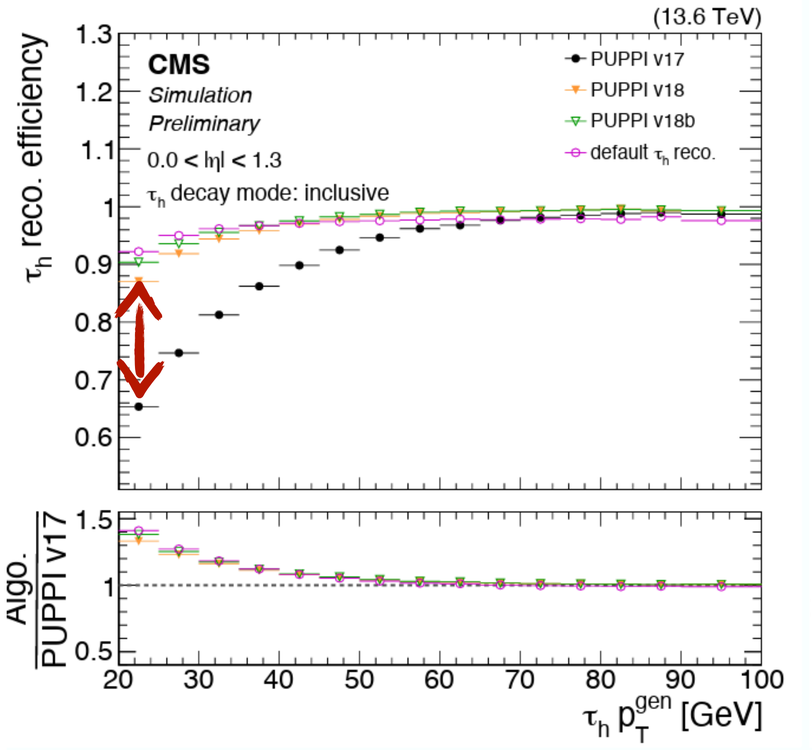
DP2024_043

Jet substructure information



- Via declustering: exclusive-kT algorithm
  - Build **quark-subjet pairs** in merged jets
- Jet is decomposed in subjets
  - Corrected **independently**

baseline algorithm: PUPPI



Hadronic taus often yield tracks not used in vtx fit
V17 not tuned to protect this tracks

New version v18 keeps non vertexed tracks from pT>4 GeV

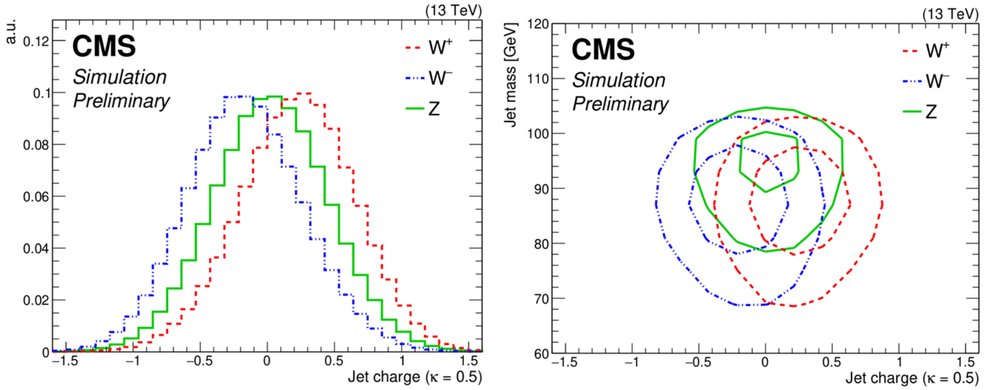
Hadronically decaying vector bosons ID

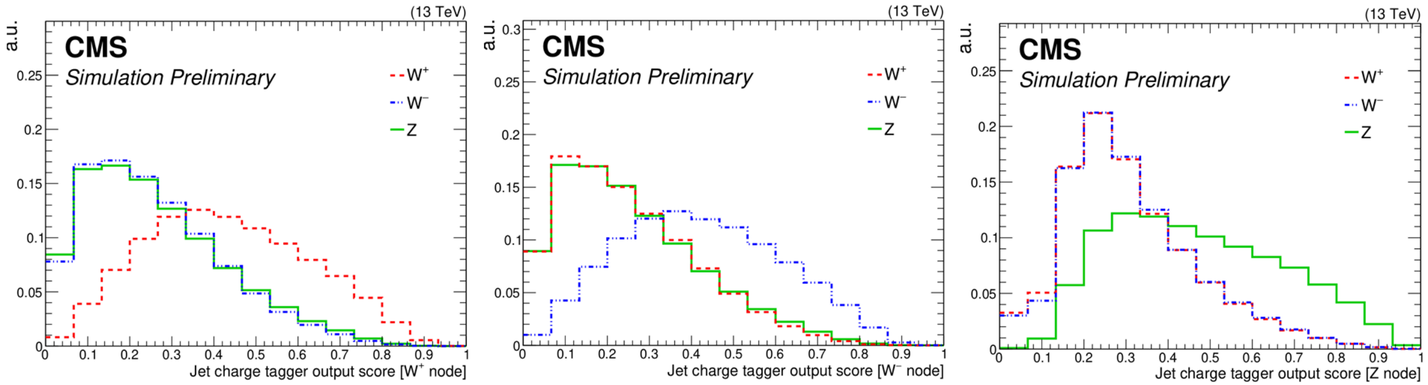
DP 2024-24

Jet charge taggers:

- Dynamic Graph Convolutional Neural Network (DGCNN) based on pNet
- On AK8 jets
- Binary (W+ vs W-) or ternary (W+,W-,Z) classification

Classification via jet variables



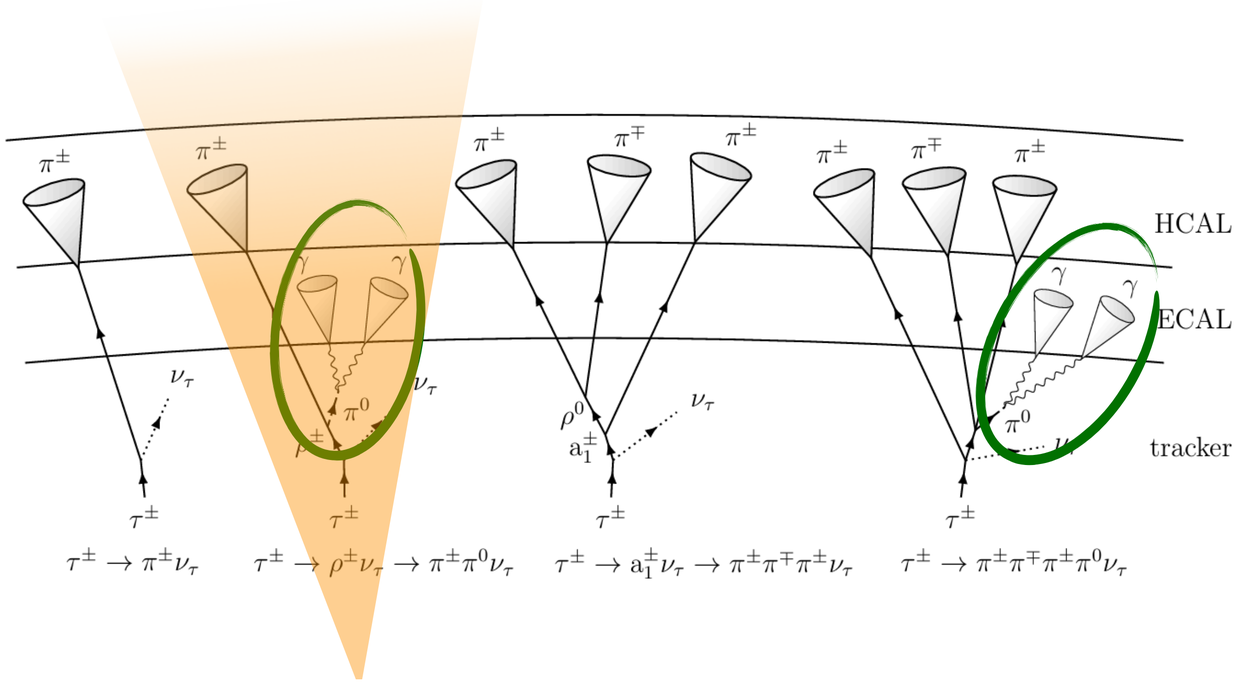


Classification via jet charge tagger

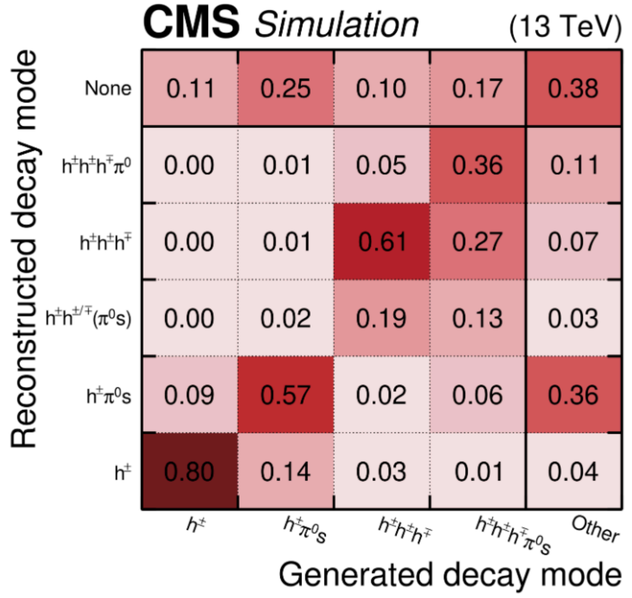
Tau reconstruction

JINST 13 (2018) 10, P10005

- Tau seeding region defined by **jets**
- Charged hadrons (HCAL + tracks)
- **HPS** (Hadron + strips) reconstructs $\pi^0 \rightarrow \gamma\gamma$
As ECAL energy strips



- Several combinations of charged and neutral pi0
- To build tau candidate
  - Selections in charge, mass compatibility **per decay channel**
- Leading pT canditate after selections: **Tau candidate**

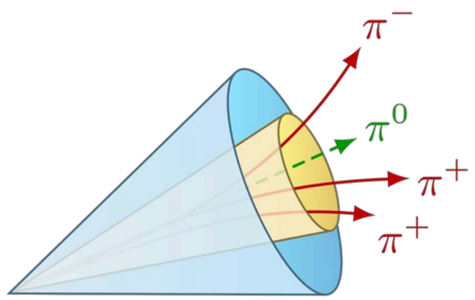


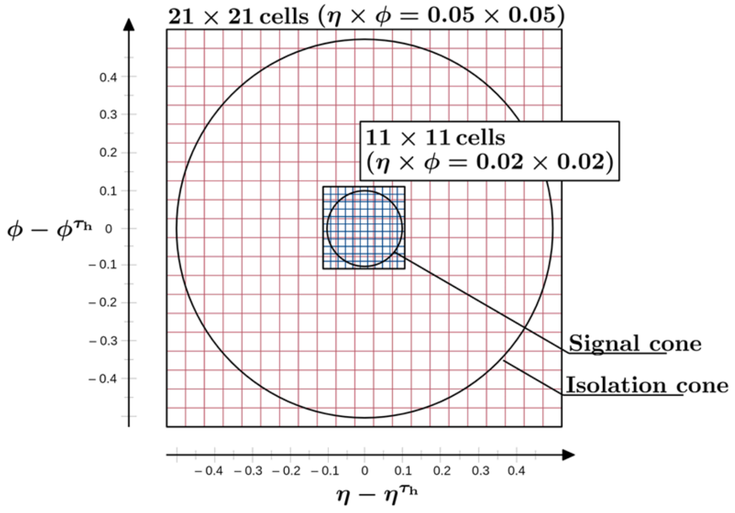
More on developments in tau's reco and ID in Paola's poster

Deep Tau with domain adaptation

CMS-PAS-TAU-24-001

More in Stepan's poster





CNN with:

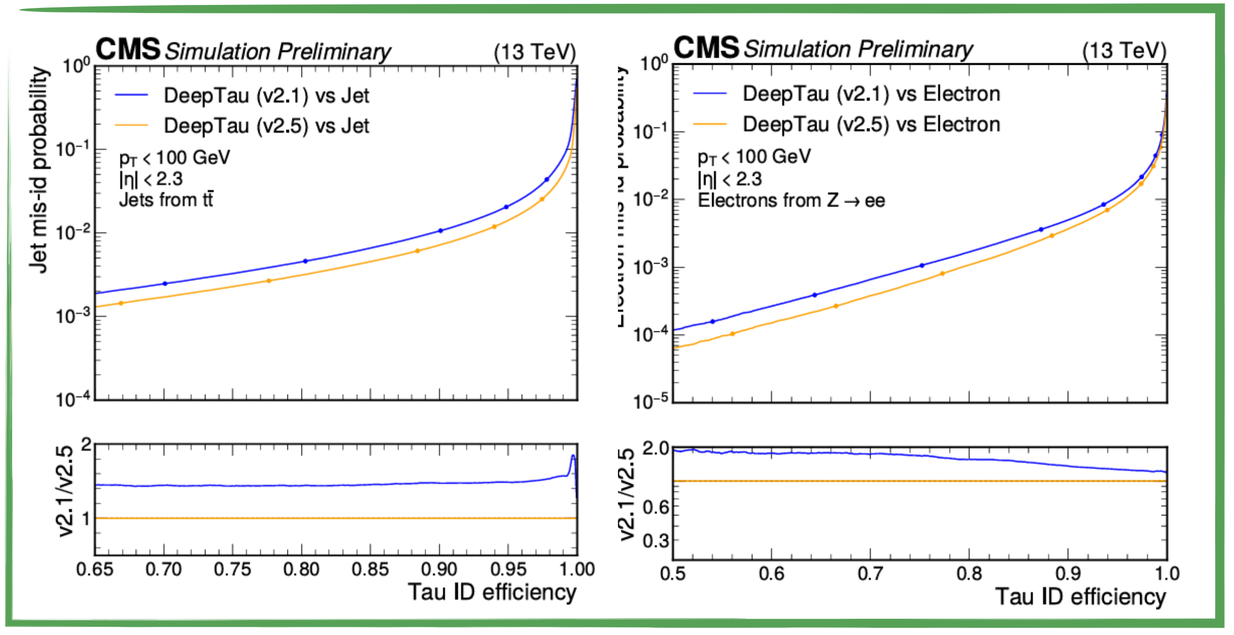
- Particle level input split in **two $\eta - \phi$ separate grids**
- High level tau candidate features

- **Mixed SM processes** for training

Train 1 epoch on mix of SIM + Data in control regions

Domain adaptation: **Run3 NEW!**

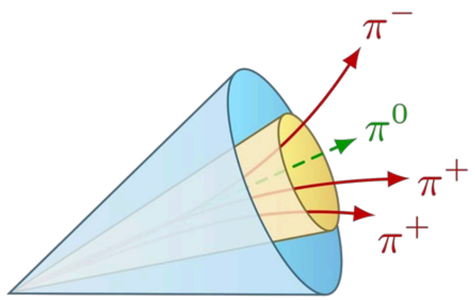
- **Target:** cure unwanted sim to data discrimination
  - Yielded by mismodelled variables + DNN train on SIM only
  - **Maximise** classification performance
  - **Minimize** discrimination between data and SIM

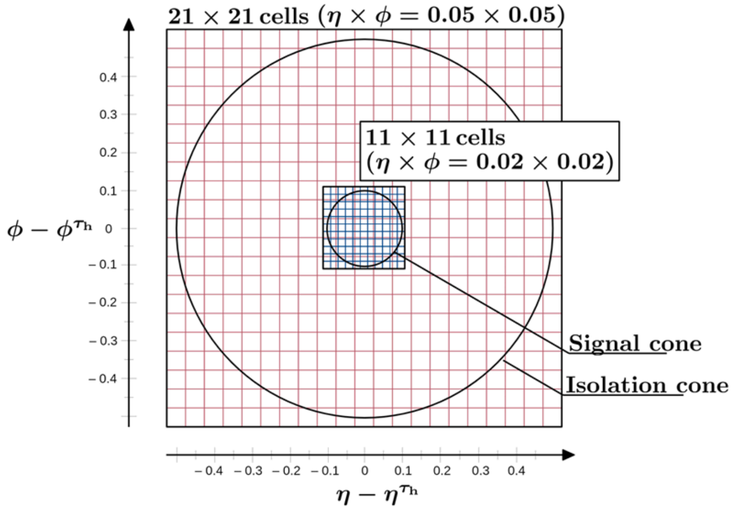


Deep Tau with domain adaptation

CMS-PAS-TAU-24-001

More in Stepan's talk





CNN with:

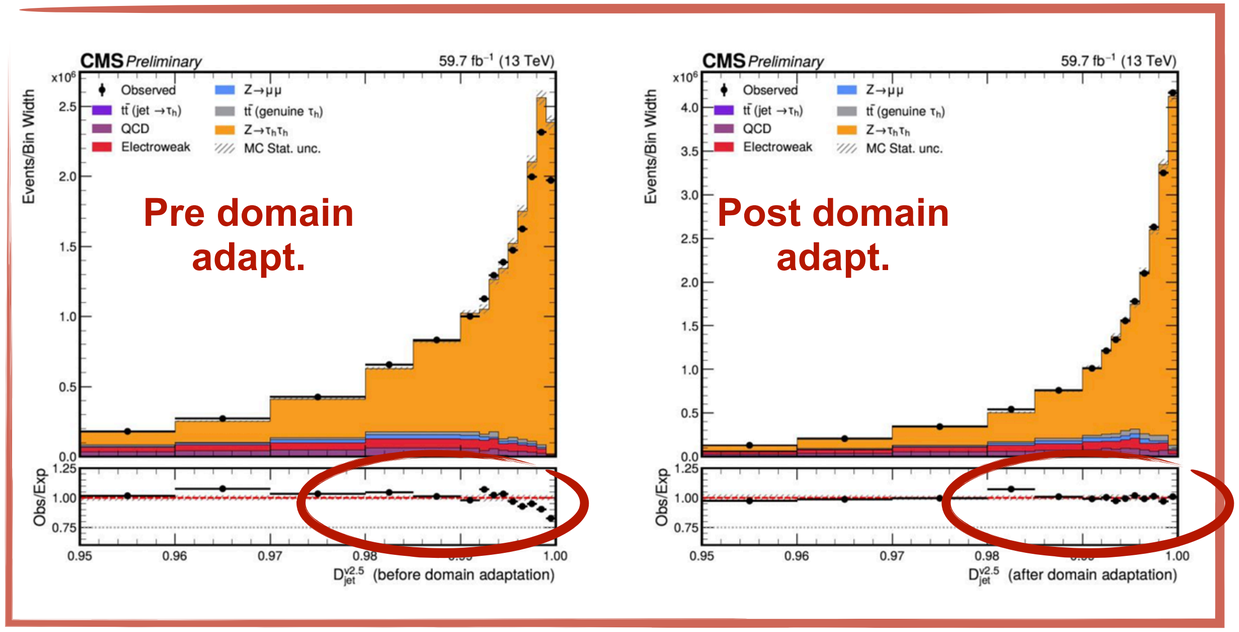
- Particle level input split in **two $\eta - \phi$ separate grids**
- High level tau candidate features

- **Mixed SM processes** for training

Train 1 epoch on mix of SIM + Data in control regions

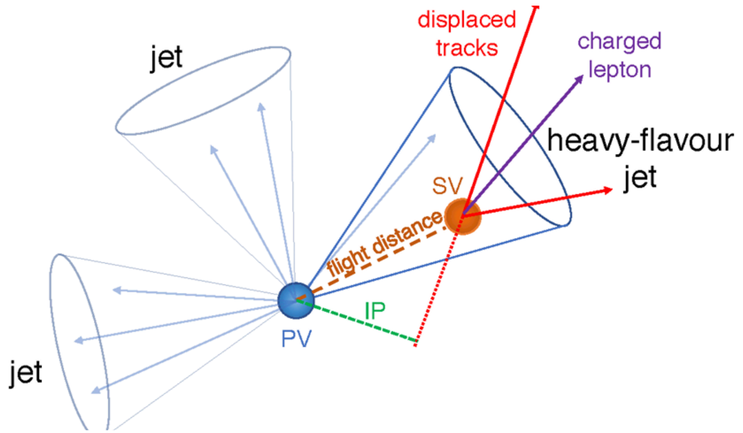
Domain adaptation: **Run3 NEW!**

- **Target:** cure unwanted sim to data discrimination
  - Yielded by mismodelled variables
  - + DNN train on SIM only
  - **Maximise** classification performance
  - **Minimize** discrimination between data and SIM



Deeper and neater - jet flavor tagging

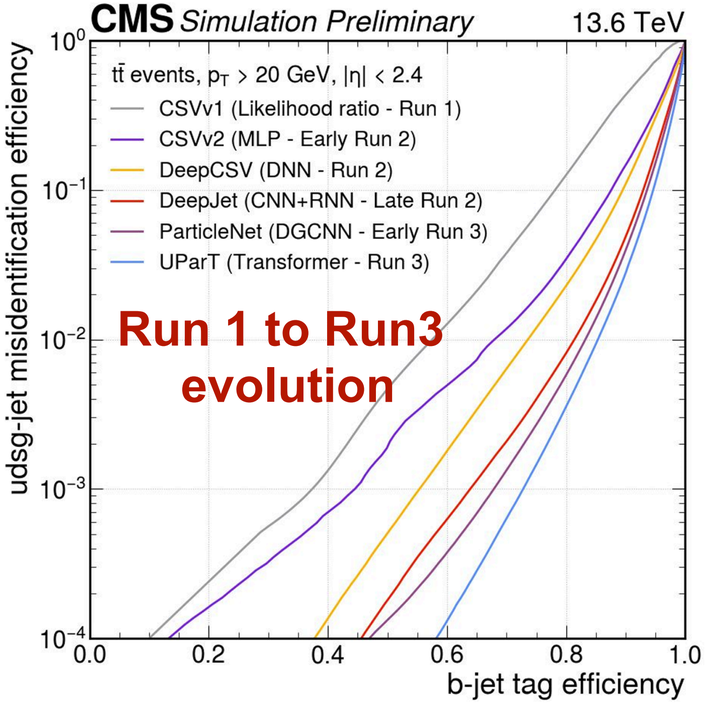
DP2024_066

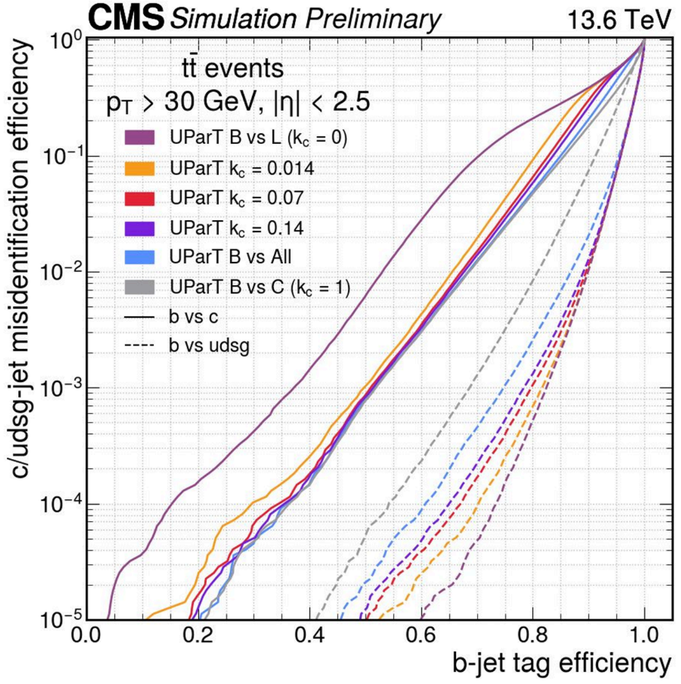


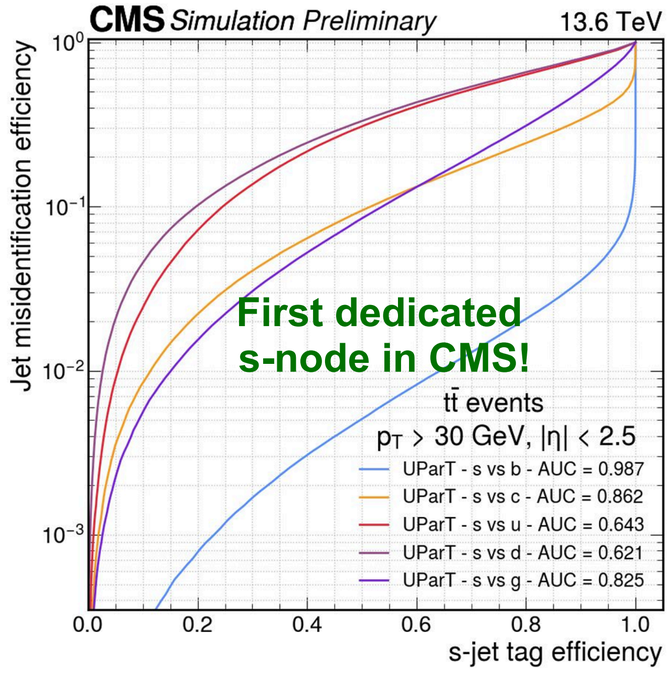
Multiple ML tasks via loss customization

$$L = \underbrace{CatEntropy(x, x_{truth})}_{\text{Classification}} + \underbrace{\gamma_{regr} \cdot log(cosh(y - y_{truth}))}_{\text{Regression}} + \underbrace{\gamma_{quantile} \cdot [p_{0.16}(z - z_{truth}) + p_{0.84}(z - z_{truth})]}_{\text{Quantile regression (resolution estimation)}}$$

UParT: the transformer paradigma





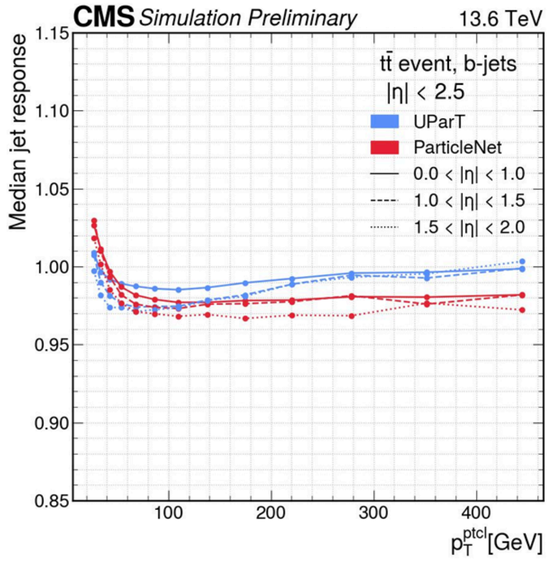


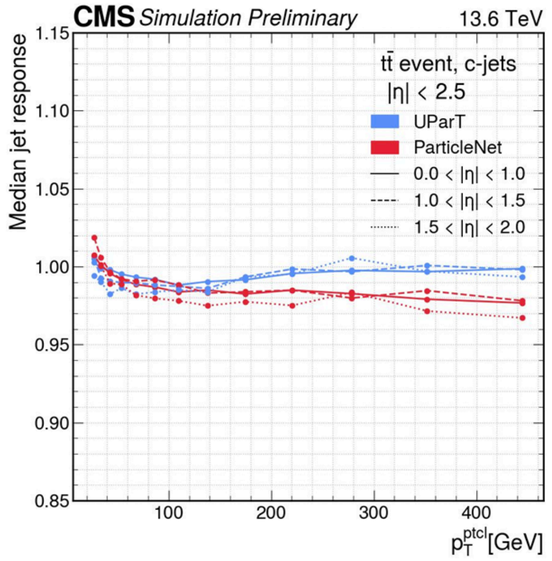
jet constituents x SV pairwise interaction novel info for discrimination

…and the other tasks

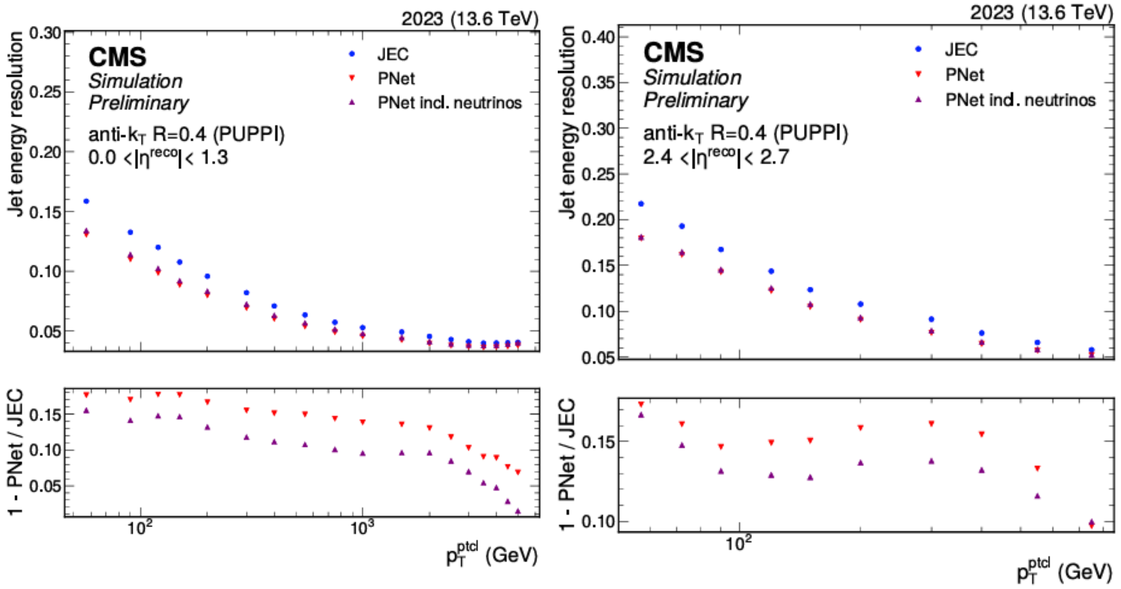
DP2024_066

Jet energy regression





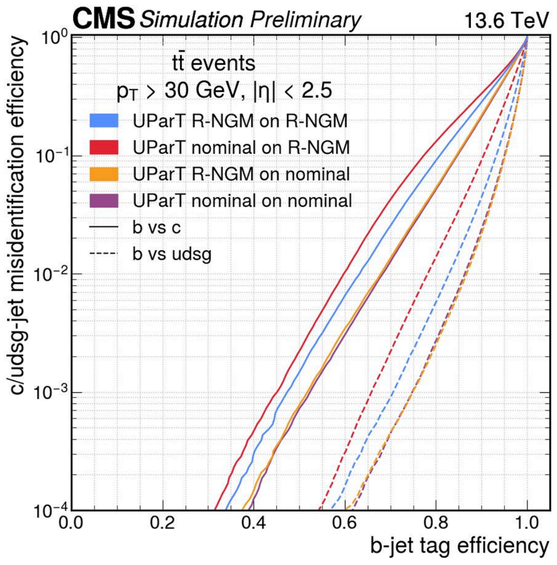
Jet energy resolution

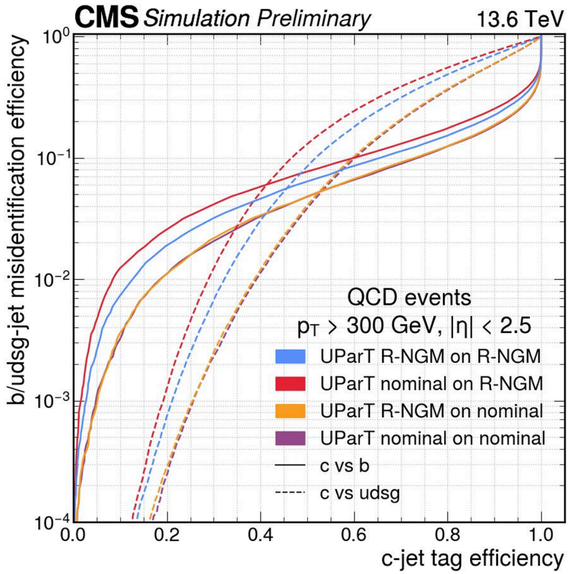


DP2024_064

See Donato's talk for more

Enhanced robustness with Rectified Normal Gradient Method





Conclusions

- CMS objects treatment keeps evolving:
  - **Novel** detector conditions
  - **Unprecendented** radiation levels
  - **Unprecedented** data taking conditions
  - **On-edge** ML development to toggle general and topical tasks
- Presented a **not at all exhaustive review** of main techniques and developments
- Run3 problem solving: **bridge towards impressive run conditions for HL-LHC**