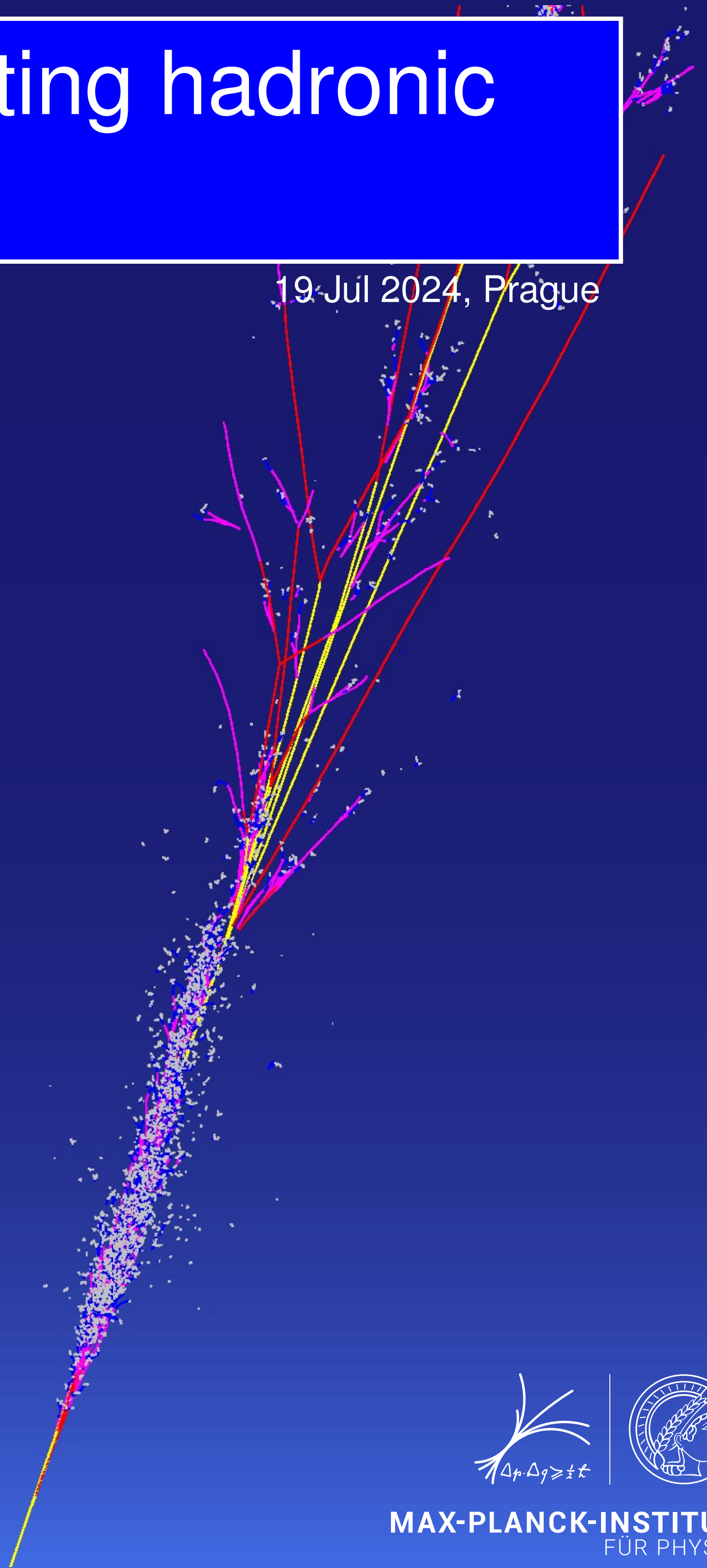
New techniques for reconstructing and calibrating hadronic objects with ATLAS

ICHEP 2024 Sven Menke, MPP München 19 Jul 2024, Prague

- Hadronic objects in ATLAS
  - Importance of Pile-Up
  - Topological clustering
- Time as a new discriminant
- Calibration methods
  - ML for cluster calibration
  - ML for jet energy calibration
  - ML for energy and mass for large ΔR-jets
- Performance of missing transverse momentum
- Conclusions



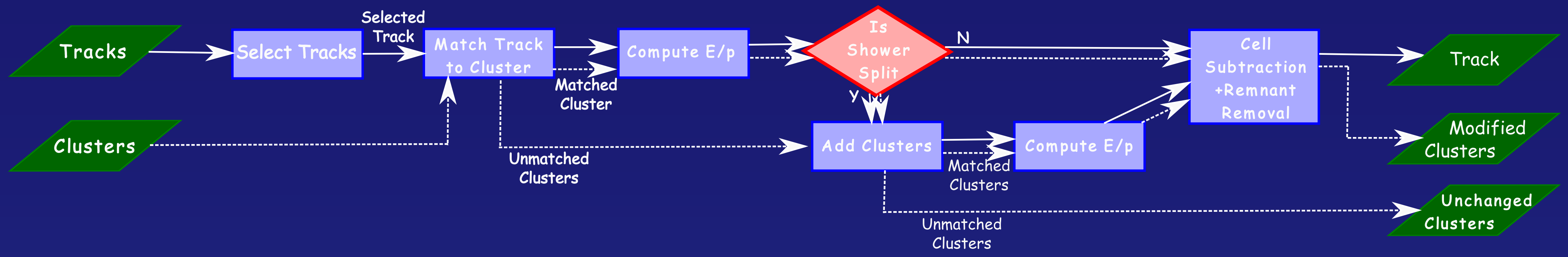
ATLAS EXPERIMENT

MAX-PLANCK-INSTITUT FÜR PHYSIK

Hadronic objects in ATLAS

- jets in ATLAS are made out of topological clusters (calorimeter) and charged particle tracks (inner detector)
- clusters and tracks are combined to form higher level objects (with 4-vectors) as input to jet-clustering
  - Particle Flow Objects (PFO) for small ΔR jets (see sketch)

Eur.Phys.J.C77(2017)466

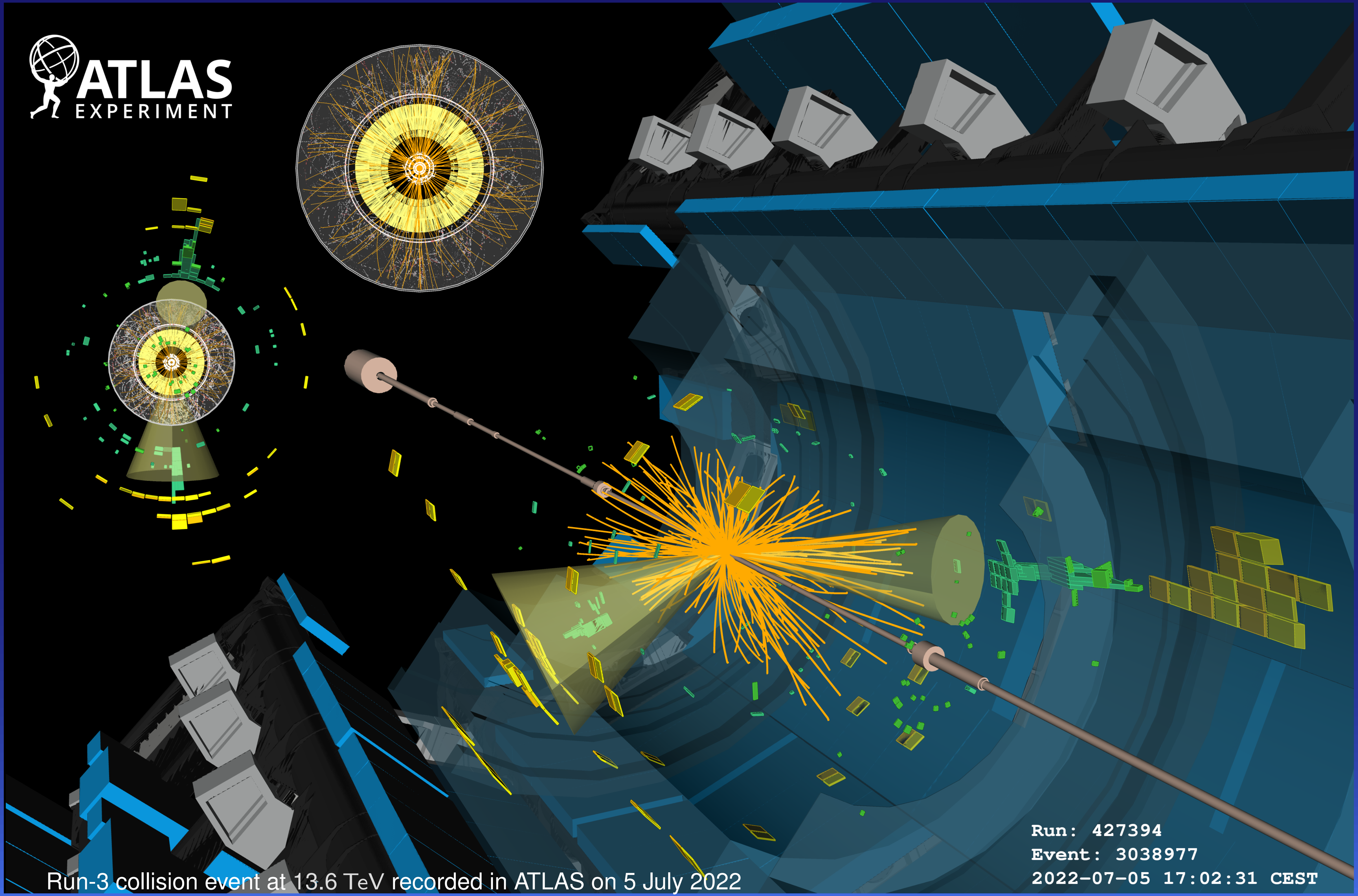


flow chart particle flow algorithm

  - Track-CaloClusters (TCC) for large ΔR jets (splitting the cluster energy to all matching tracks with track's $p_{\perp}$-fraction in matched tracks as weight, ATL-PHYS-PUB-2017-015)
  - Unified Flow Objects (UFO) combine PFO and TCC depending on environment to make best of both Eur.Phys.J.C81(2020)334)
- jet clustering is performed with FastJet – anti-k_t with $\Delta R = 0.4$ (small) or $\Delta R = 1.0$ (large)
- jets are then calibrated in several steps for energy ($p_{\perp}$), momentum direction and mass (for large ΔR)

Di-jet event in Run-3

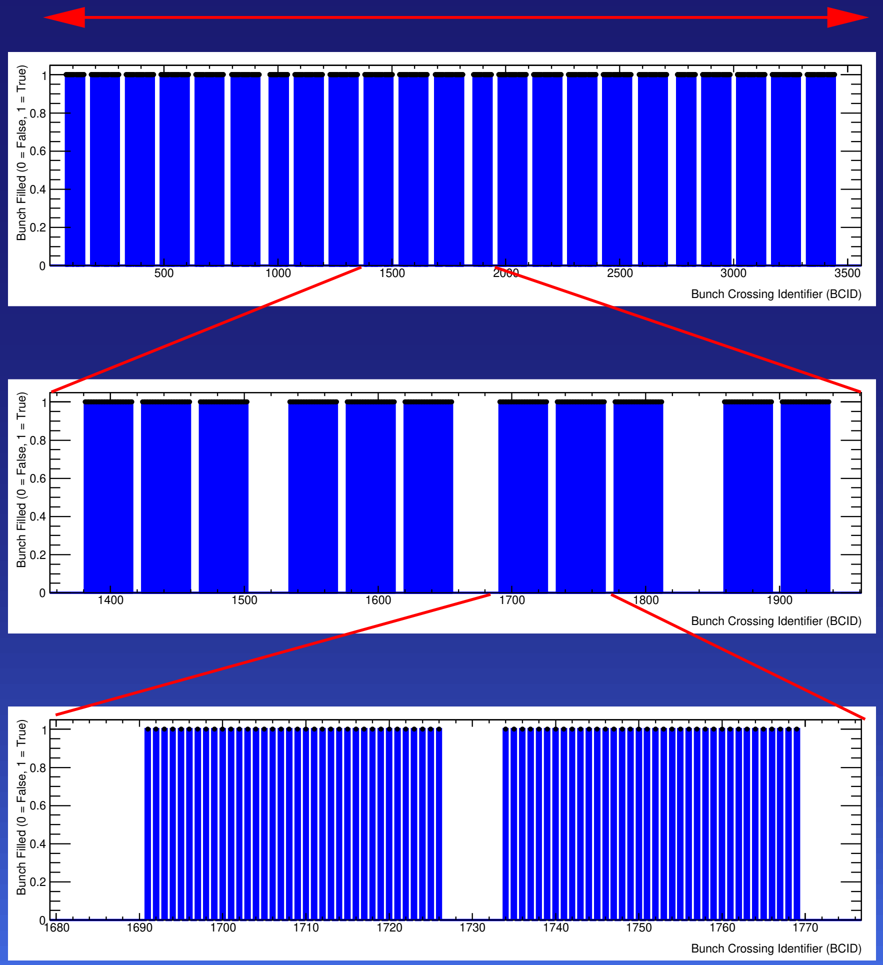
twiki.cern.ch/twiki/bin/view/AtlasPublic/EventDisplayRun3Collisions



Pile-Up impact on calorimeter signals in ATLAS

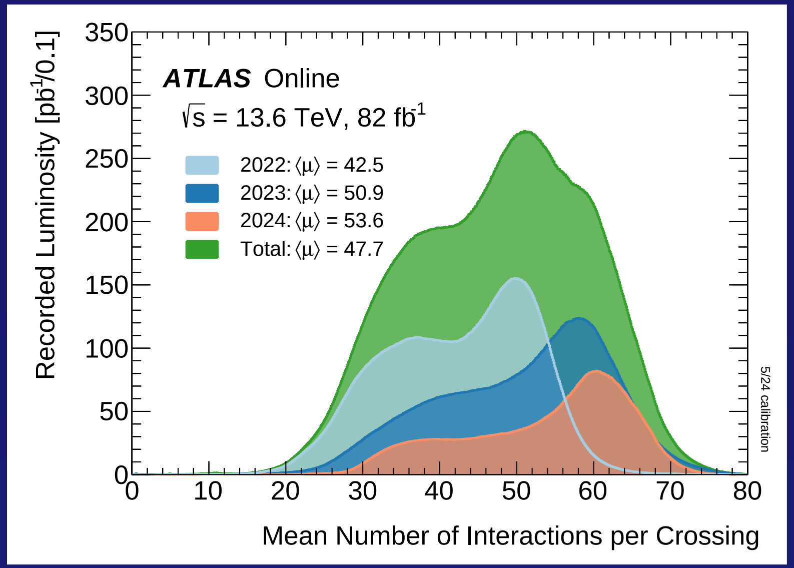
- Pile-Up characteristics
  - μ: average number of interactions per crossing ~ 50 in Run-3
  - Δt: bunch distance 24.95 ns
  - signal integration time for the LAr-calorimeters ~ 500 ns

26.7 km $\equiv$ 3564 bunch places each $\Delta t = 24.95$ ns



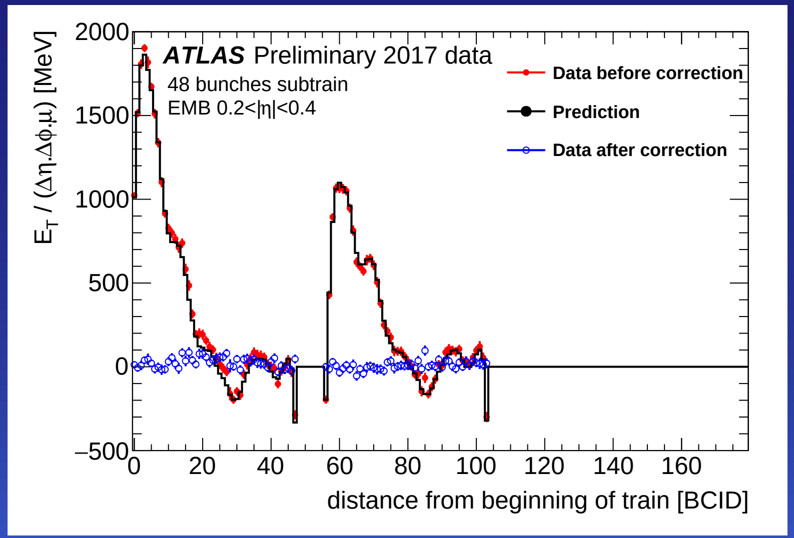
- typical LHC bunch structure in 2022-2024
  - 2340 colliding bunch pairs (2835 is theoretical max)
  - 23 trains ~ 800 ns apart (LHC injection)
  - 2-3 sub-trains with gaps of 200 ns (SPS injection)
  - 36 filled bunches per sub-train
    - Pile-Up impact depends on bunch crossing Number (BCID)
    - up to 20 colliding bunch pairs contribute to signal
- see **arXiv:2407.10819** and **Turning noise into data: using pileup for physics** poster by A. Pirttikoski

twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun3



interactions per crossing 2022-2024

twiki.cern.ch/twiki/bin/view/AtlasPublic/LArCaloPublicResults2015

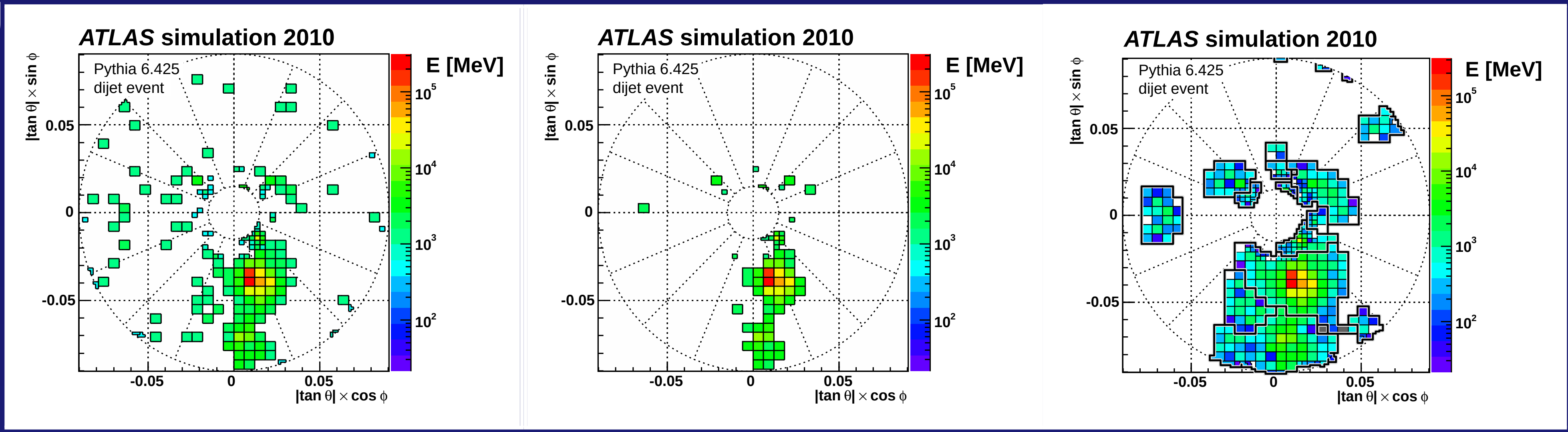


LAr baseline shift

Topological clustering

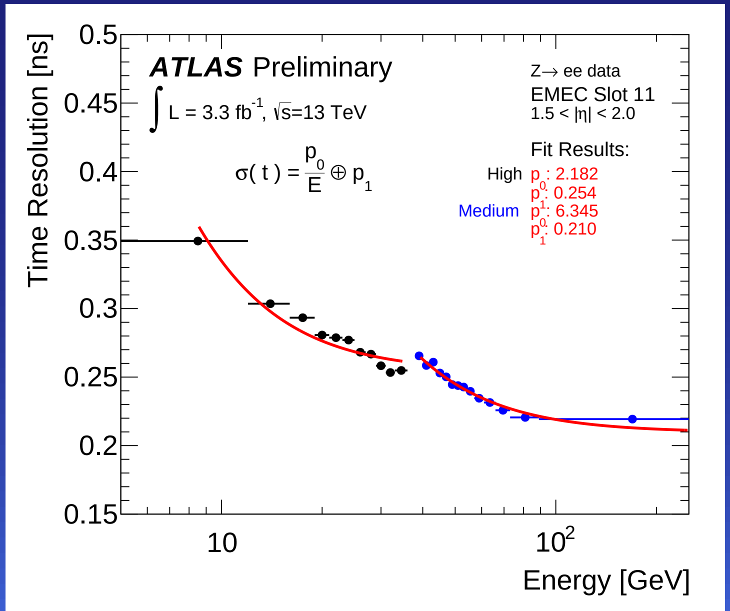
- jet constituents, $\tau^{\pm}$, $e^{\pm}$ and γ are made out of topological cell clusters (TopoClusters)
  - 3d energy blobs of neighbouring calorimeter cells around seeds with $|E| > 4\sigma$
  - direct seed neighbours with $|E| > 2\sigma$ become seeds too
  - proto-clusters are re-clustered around local energy maxima
  - σ is the expected noise $\equiv \sigma_{\text{elec}} \oplus \sigma_{\text{pile-up}}$

Eur.Phys.J.C77(2017)490



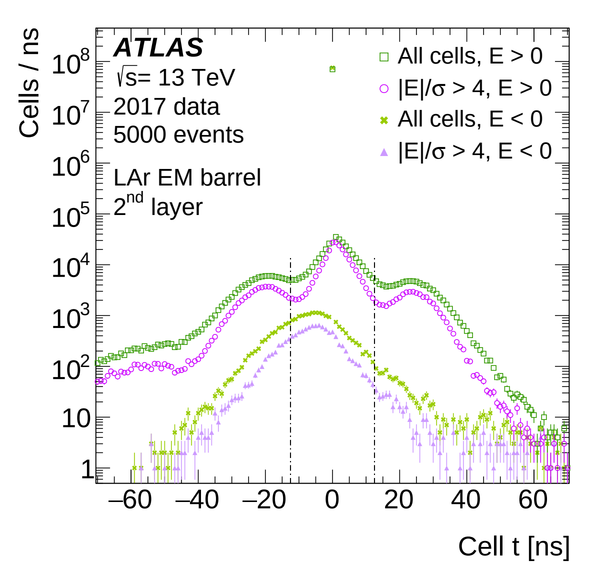
$|E| > 2\,\sigma_{\text{noise}}$ $|E| > 4\,\sigma_{\text{noise}}$ 4 / 2 / 0 TopoClusters

twiki.cern.ch/twiki/bin/view/AtlasPublic/LArCaloPublicResults2015



LAr EMEC time resolution

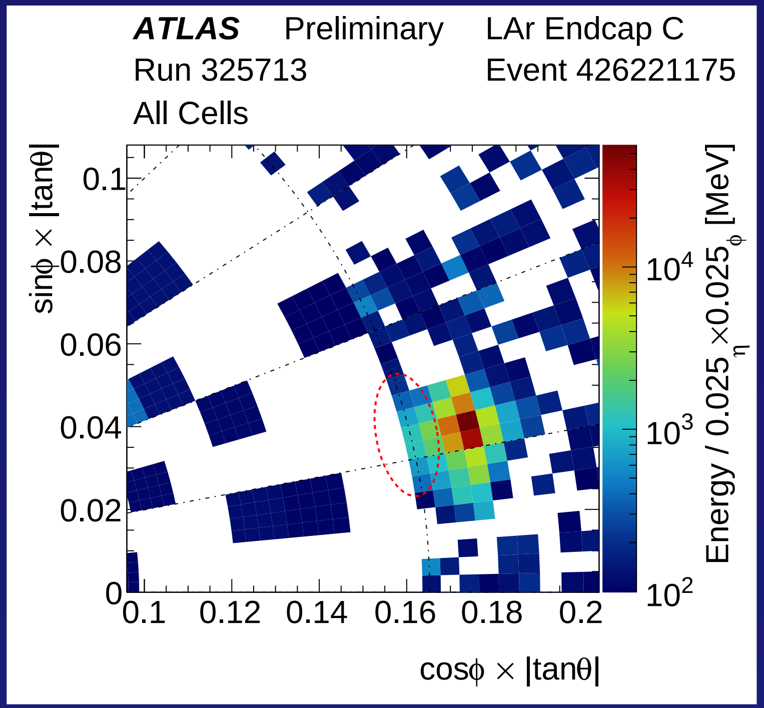
Eur.Phys.J.C84(2024)455



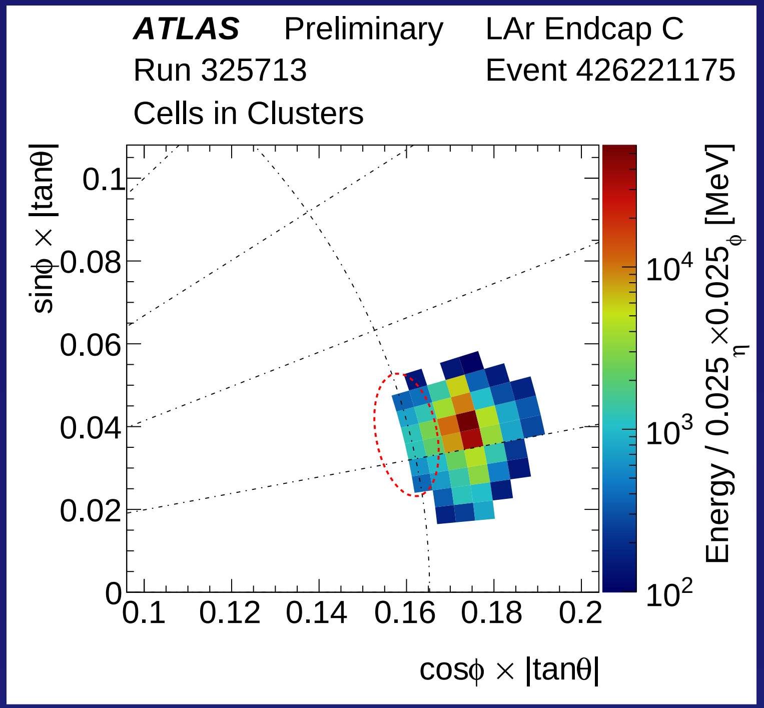
Cell Time in LAr EMB

- calorimeters have excellent time resolution!
  - intrinsic time resolution in LAr samplings is ~ 60 ps at high energies
  - time has always been reconstructed alongside energy since the beginning of data taking
  - added recently to the topological clustering algorithm as additional discriminator (cut at $|t| < 12.5$ ns) for any cell that has $|E| > 4\sigma$
  - but restrict the time cut to those cells with $E < 20\sigma$
    - to keep significant, positive energy deposits that are out-of-time (searches for exotic, long-lived particles)

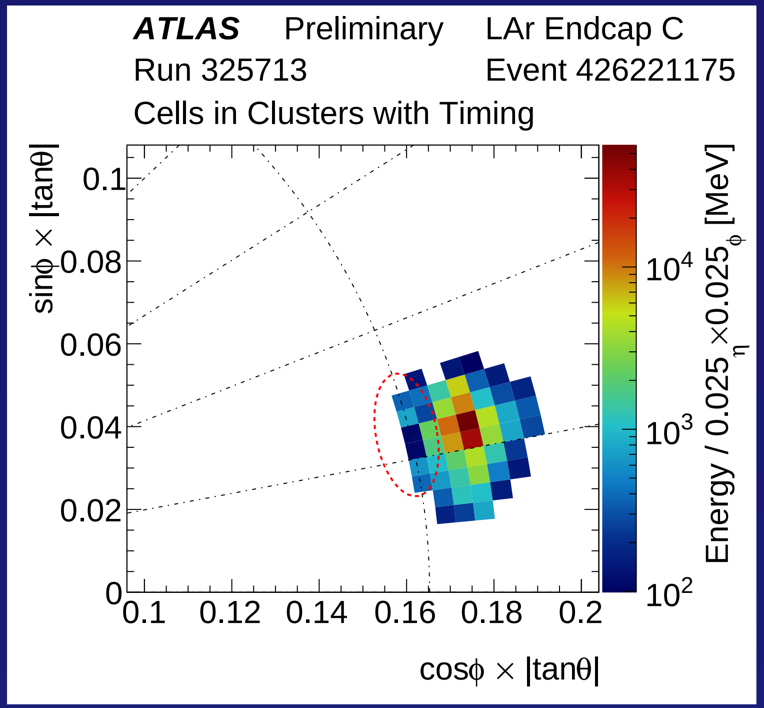
twiki.cern.cn/twiki/bin/view/AtlasPublic/LArCaloPublicResults2015



LAr cell energy sums above 100 MeV



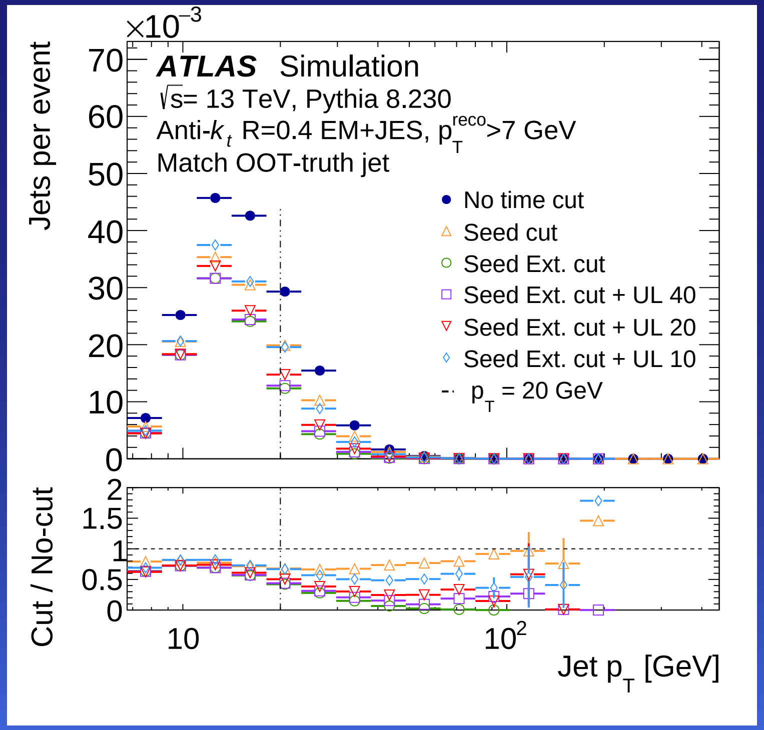
inside default clusters



inside clusters with time discrimination

- new default in Run-3

Eur.Phys.J.C84(2024)455



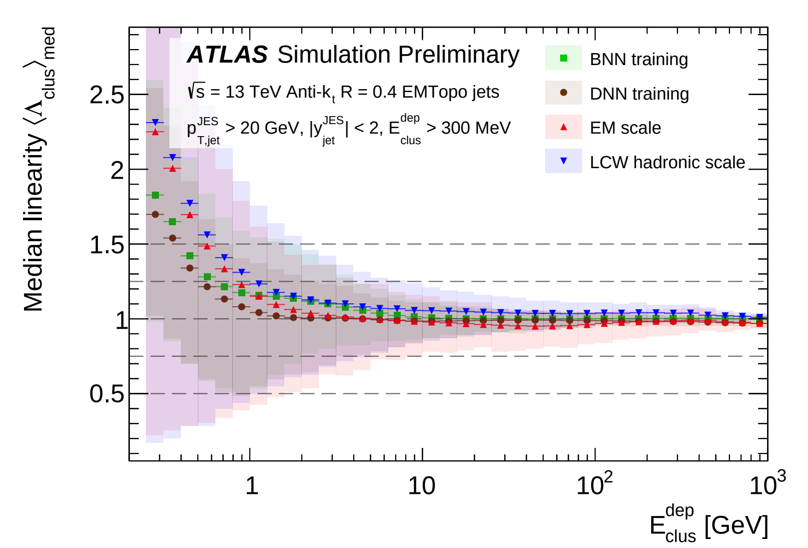
OOT Pile-Up jets vs. $p_\perp$

- new time discriminant further reduces residual out-of-time (OOT) Pile-Up that was not suppressed by default topological clustering
  - entire clusters are removed (reduces background)
  - and cells inside clusters are removed (improves signal, see plots above)
- removes OOT Pile-Up jets (see plot to the right)
  - by $\sim 50\%$ at $p_\perp \simeq 20$ GeV; by $\sim 80\%$ at $p_\perp \geq 50$ GeV
  - number of in-time jets remains unchanged
  - resolution improves by $\sim 5\%$
- removes fakes for $\tau^\pm$, $e^\pm / \gamma$

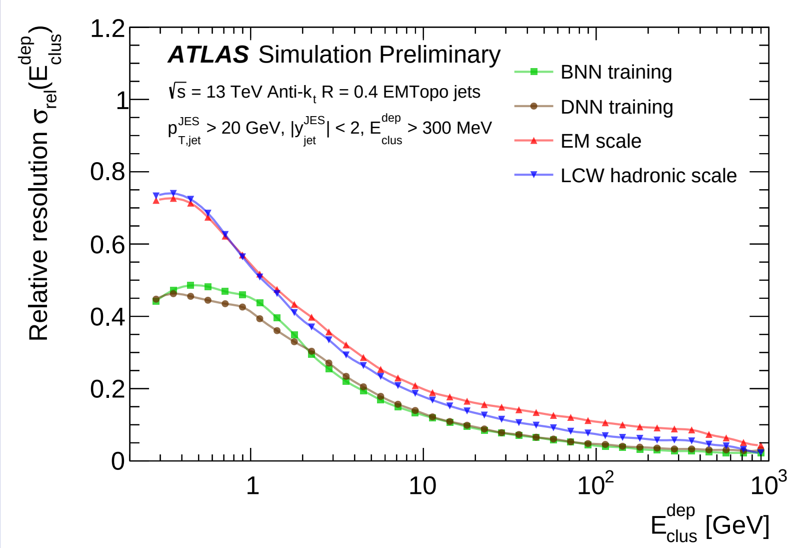
Calibration methods

- Local Hadronic Calibration (a.k.a. Local Cell Weighting, LCW, Eur.Phys.J.C77(2017)490)
  - 4 step procedure to bring the energy scale of clusters from the raw ”EM”-scale to the particle-level ”LC”-scale
  - Classification: compute EM-probability p^{EM} form shapes
  - Cell-Weighting: apply hadronic (HAD) and electromagnetic (EM) weights:
    - $w_{\mathrm{cell}} = (1 - p^{\mathrm{EM}})\, w_{\mathrm{HAD}} + p^{\mathrm{EM}}\, w_{\mathrm{EM}}$
  - for 3 different corrections:
    - corrections for hadronic non-compensation
    - corrections for out-of-cluster deposits
    - corrections for out-of-calorimeter (dead-material) deposits
- Jets (Eur.Phys.J.C81(2021)689)
  - can use either EM- or LC-scale objects (clusters or flow objects)
  - are corrected for Pile-Up (jet-area correction and residual Pile-Up correction)
  - get their energy corrected by MC-derived Jet-Energy-Scale correction
  - flavour dependency and resolution gets improved by Global-Calibration (MC-derived, keeping average energy scale constant)
  - data is corrected *in-situ* from measured $p_{\perp}$ balance of jets in multi-jet and $Z^{0}/\gamma +$ jet events to match MC

- Idea: Apply machine learning to Local Hadronic Calibration
  - to explore the applicability of neural networks to calorimetric calibration
  - so far done with the first of the three correction steps (non-compensation) and implicit classification
    - biggest difference to legacy LCW: Pile-Up is included
  - out-of-cluster and dead-material corrections still to come
- Input quantities for the NNs:
  - kinematics: (E^{EM}_{clus}, y^{EM}_{clus}) • significance: (E_{clus} / σ_{clus}) • time: (t_{clus}, $\mathrm{Var}_{clus}(t_{cell})$) • cluster moments: (depth, centroid, EM-fraction, energy density, lateral and longitudinal dispersion, compactness) • environment: (isolation, N_{PV}, μ)



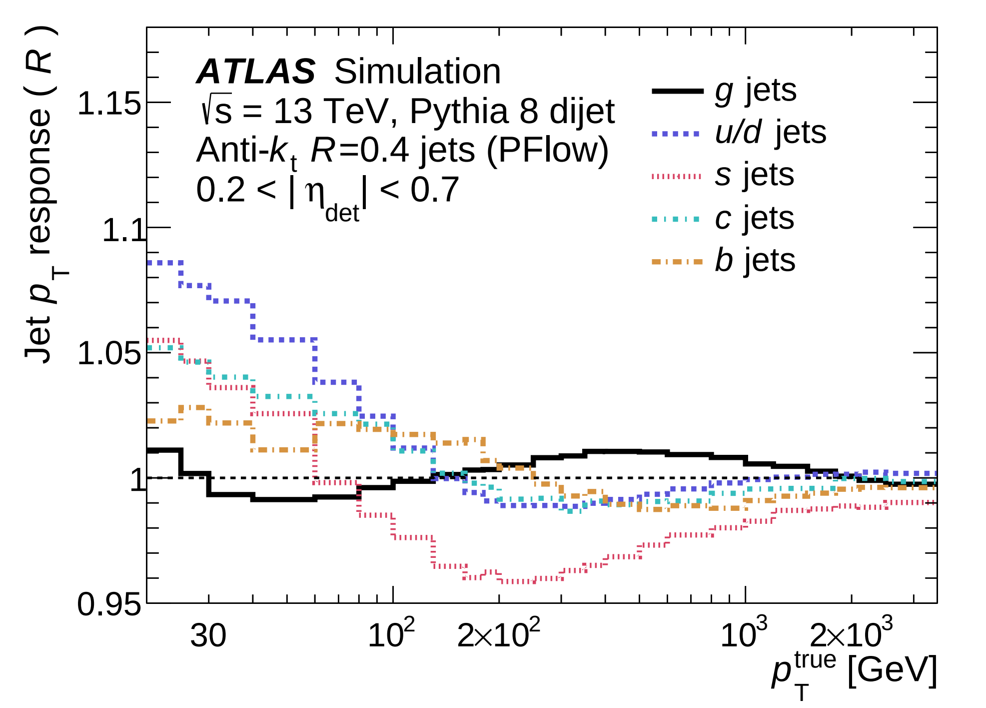
Median linearity of reconstructed E^{dep}_{clus}



Relative resolution of reconstructed E^{dep}_{clus}

- trained NNs:
  - DNN with leaky Gaussian kernel
  - BNN with regularised negative log-likelihood
- linearity (left) and resolution (right) of NNs compared to EM- and LC-scale on simulated clusters from di-jets with Pile-Up
  - NNs outperform legacy LCW (removal of Pile-Up)
    - but Pile-Up removal is not part of LCW ...
  - DNN slightly better than BNN
  - encouraging result to implement the other steps

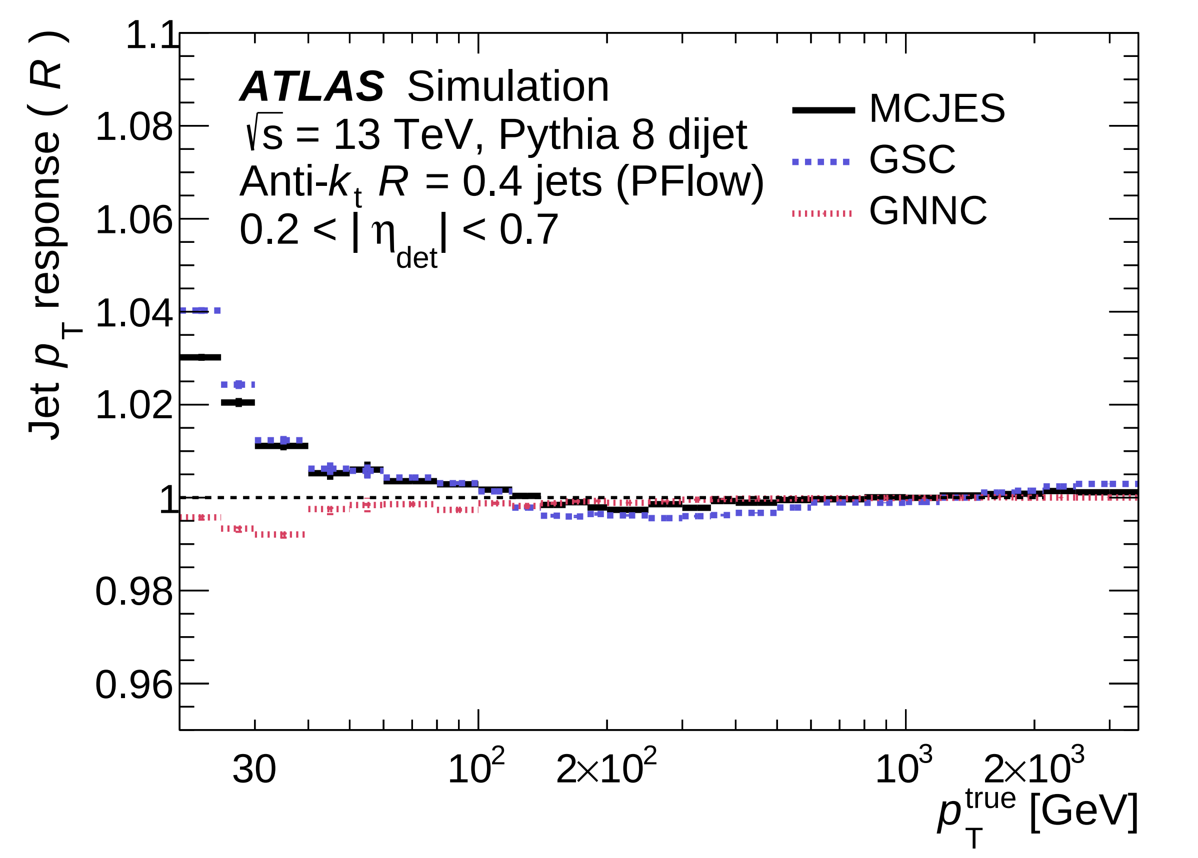
- Global-Calibration is applied to jets after setting the jet energy scale (MCJES) (based on MC simulations and energy E and pseudo-rapidity η of the jet)
- Global Sequential Calibration (GSC) (used for Run-2)
  - uses many kinematic observables in addition to $p_\perp$: • charged $p_\perp$ fraction f_{charged} • energy fractions in first Tile & third EM layer f_{Tile0} • f_{LAr3} • number of tracks N_{track} • $p_\perp$-weighted average track distance w_{track} • number of associated muon segments N_{segments}
  - since JES is kept unchanged, the six corrections can be applied and checked independently of each other ► requires uncorrelated observables



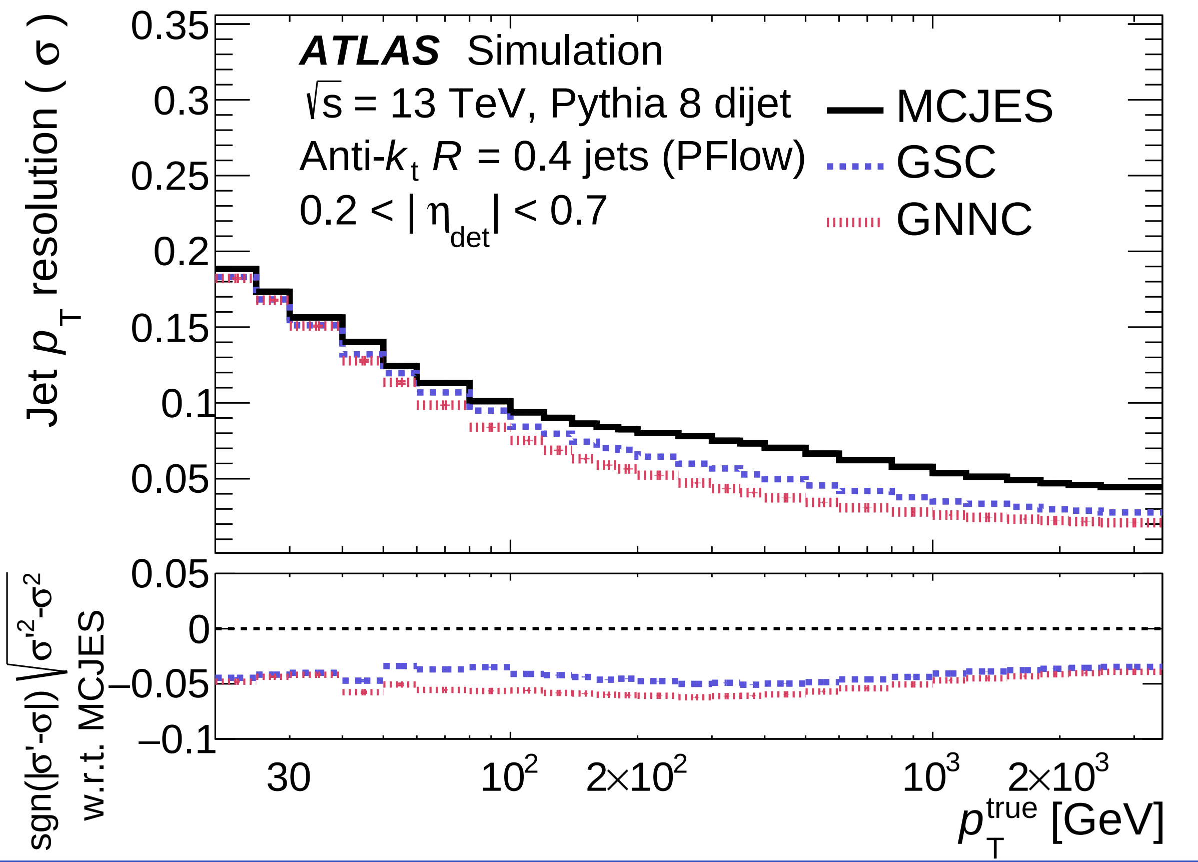
$p_\perp$ response after MCJES

- Global Neural Network Calibration (GNNC) (new, will be used for Run-3)
  - alternative to GSC
  - trains a DNN with jet observables for a simultaneous correction to $p_\perp$ and leaky Gaussian kernel loss-function
    ► allows the use of correlated variables; is allowed to change JES
  - in addition to the GSC observables it uses: • 12 more (i.e. all 14) layer energy fractions $f_{\text{LAr0-3,Tile0-2,HEC0-3,FCal0-2}}$ • number of clusters with 90% energy $N_{90\%}$ • η • Pile-Up variables μ, N_{PV}

- closure and resolution in $p_\perp$ compared after MCJES, MCJES+GSC and MCJES+GNNC (here for $0.2 < |\eta| < 0.7$, similar results in all other η-regions)
- small non-closure for GSC at low $p_\perp$ stems from MCJES (GSC keeps JES unchanged)
- GNNC does change JES and hence improves the MCJES closure at low $p_\perp$
- resolution improves by $15 - 25\%$ for GNNC compared to GSC



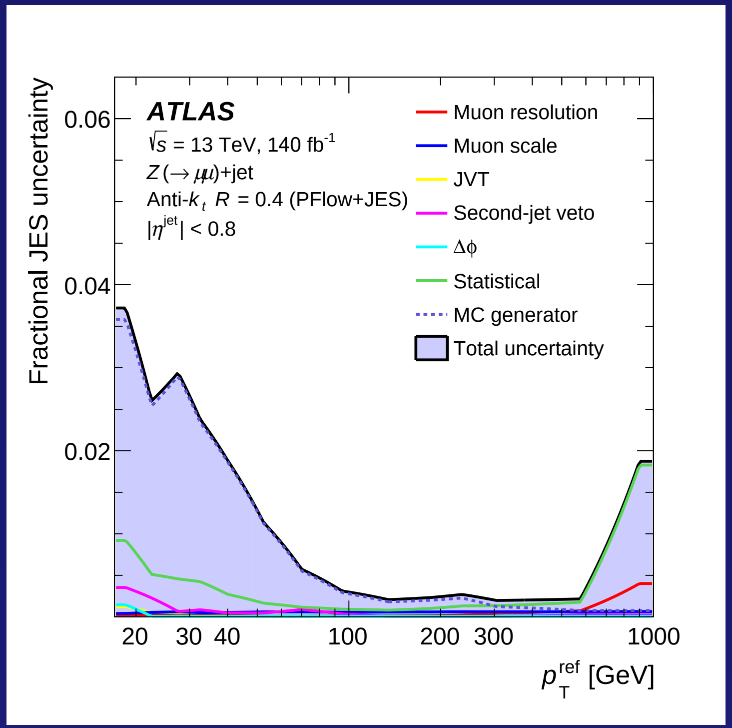
Jet $p_\perp$ response closure



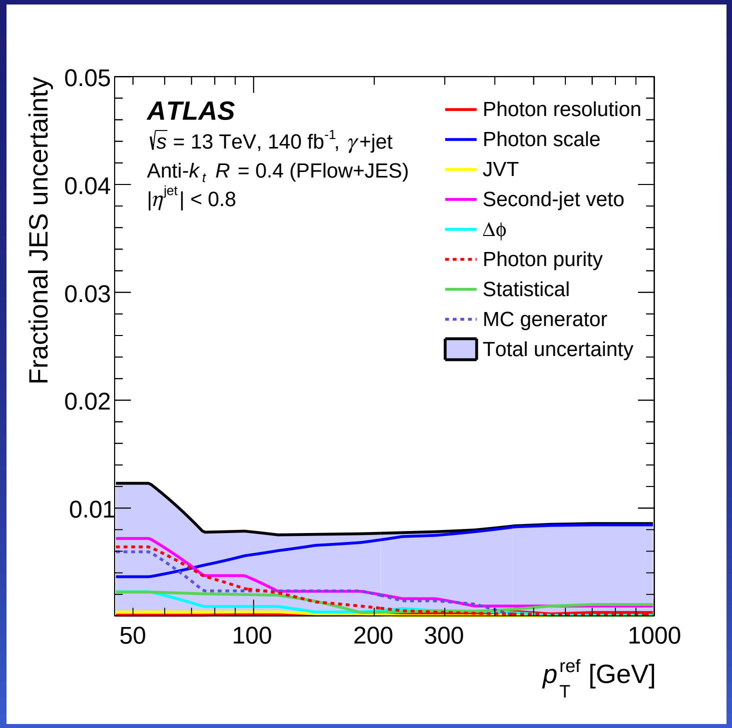
Jet $p_\perp$ resolution

Eur.Phys.J.C83(2023)761

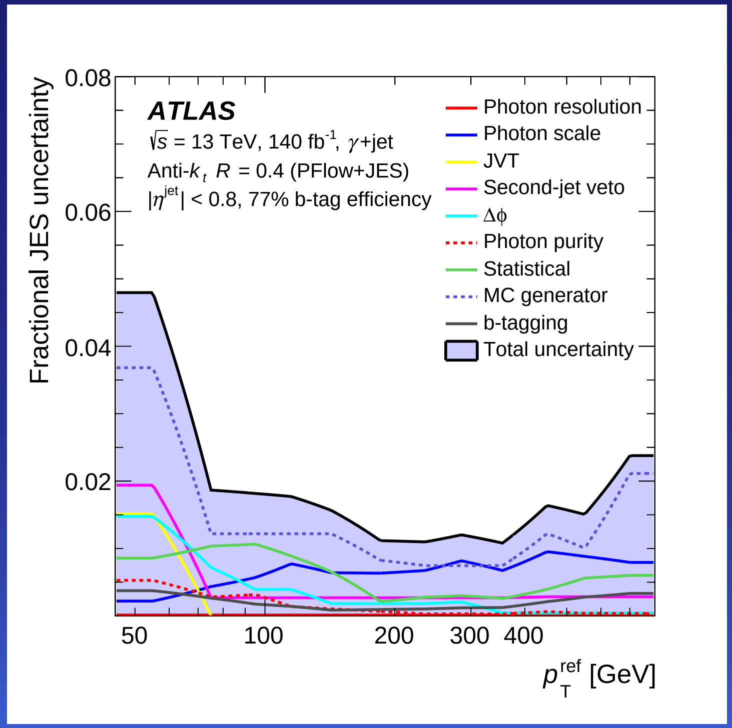
- $Z^0 + \text{jet}$ and $\gamma + \text{jet}$ data are compared *in-situ* to simulations to bring the final JES in data to simulation level (after MCJES+GNNC and *in-situ* η-intercalibration with multi-jet events)
- Missing-$E_\perp$ Projection Fraction (MPF) is used to calculate $p_\perp$-balance between Z^0 / γ and the full hadronic recoil
  - best for Pile-Up and lower $p_\perp$
- $O(1\%)$ precision is achieved over a large $p_\perp$-range



JES uncertainty vs. $p_\perp$: MPF method $Z^0(\to \mu^+\mu^-) + \text{jet}$



JES uncertainty vs. $p_\perp$: DB method $\gamma + \text{jet}$

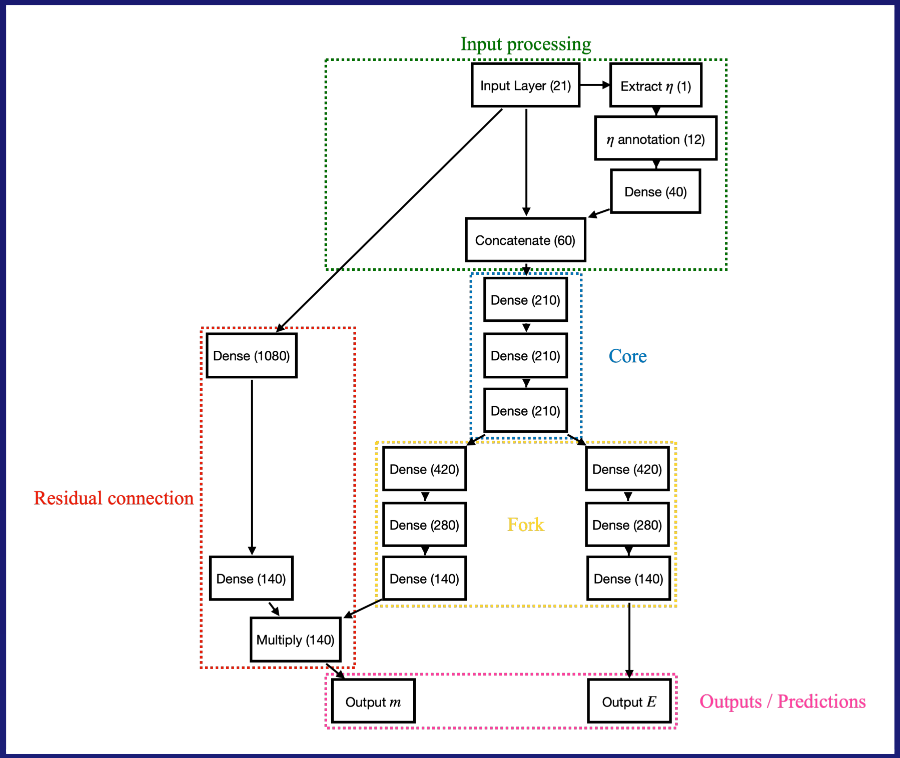


b-jet JES uncertainty vs. $p_\perp$: DB method $\gamma + \text{b-jet}$

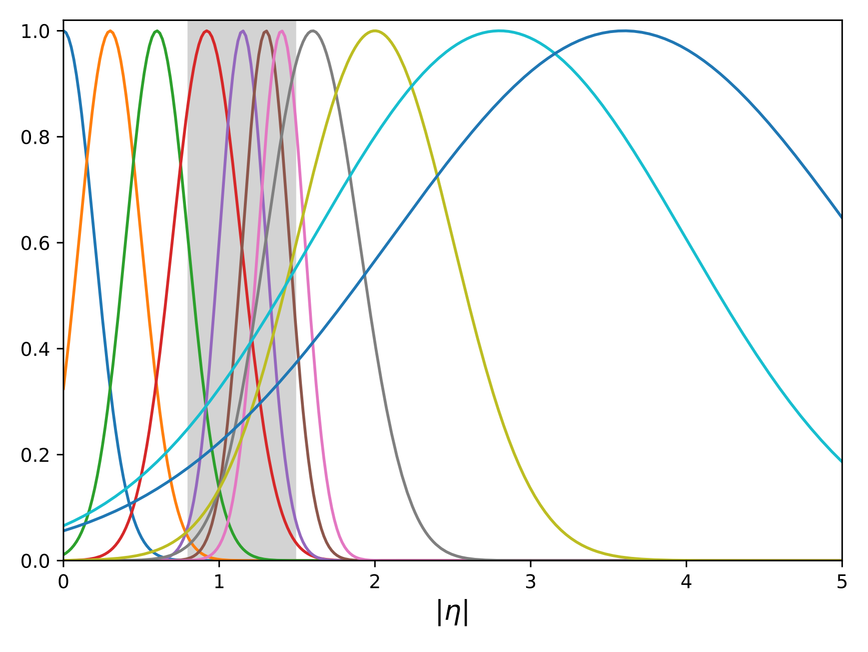
- Direct Balance (DB) is used in $\gamma + \text{jet}$ events to measure the balance of γ against one (possibly b-tagged) jet
  - good at $p_\perp > 100\,\text{GeV}$ and for single *b*-jets
- up to $O(1\%)$ precision on-top of general JES uncertainty

Calibration of E and m of large ΔR-jets

- large ΔR jets: good for boosted topologies of heavy resonances
  - the asymmetric response in energy and mass requires dedicated calibration for both
  - both remain highly correlated though and a combined calibration approach is hence desirable
- complex DNN with η annotation (adding 11 Gaussian η-dependent weights to input)
  - inputs: jet kinematics E, m, η, 8 jet substructure variables, 7 detector-level energy or $p_\perp$ fractions, Pile-Up environment N_{PV}, μ
  - initial training for both E and m
  - loss function is sum of negative log-likelihood predicting μ and σ of Gaussian distributions in E and m
  - then fork and optimise separately for E and m (can freeze the other)
  - residual connection for m improves the focus on most important inputs for m
- trained on 270 M jets from fully simulated di-jet events (based on Pythia8 and Geant4; other generators, physics for cross-checks)



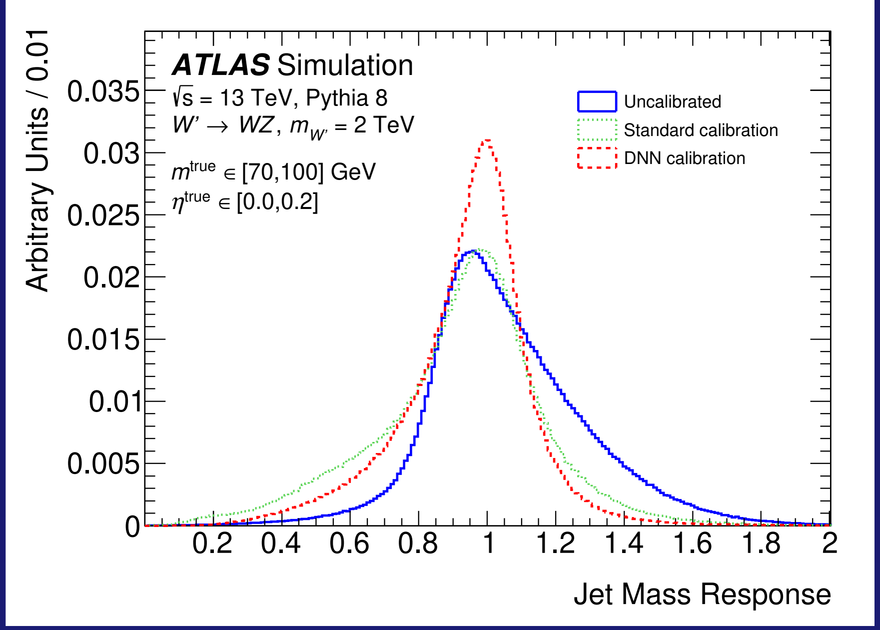
DNN architecture for E and m



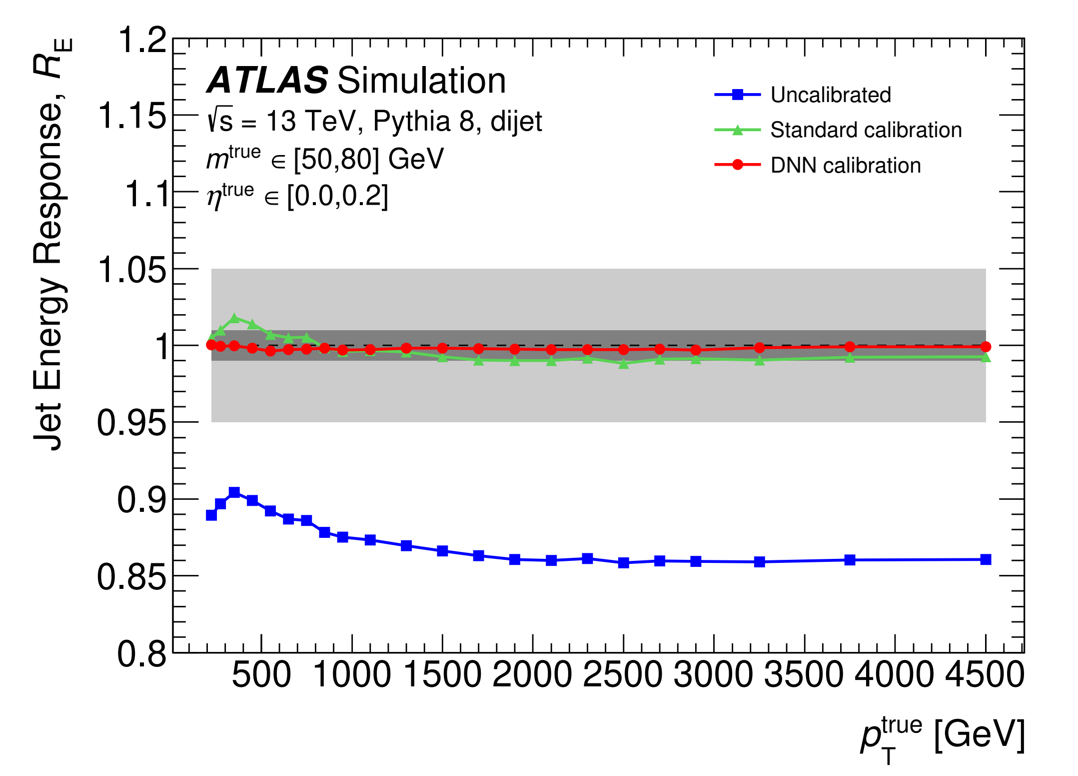
η-annotation functions

Calibration of *E* and *m* of large ΔR-jets

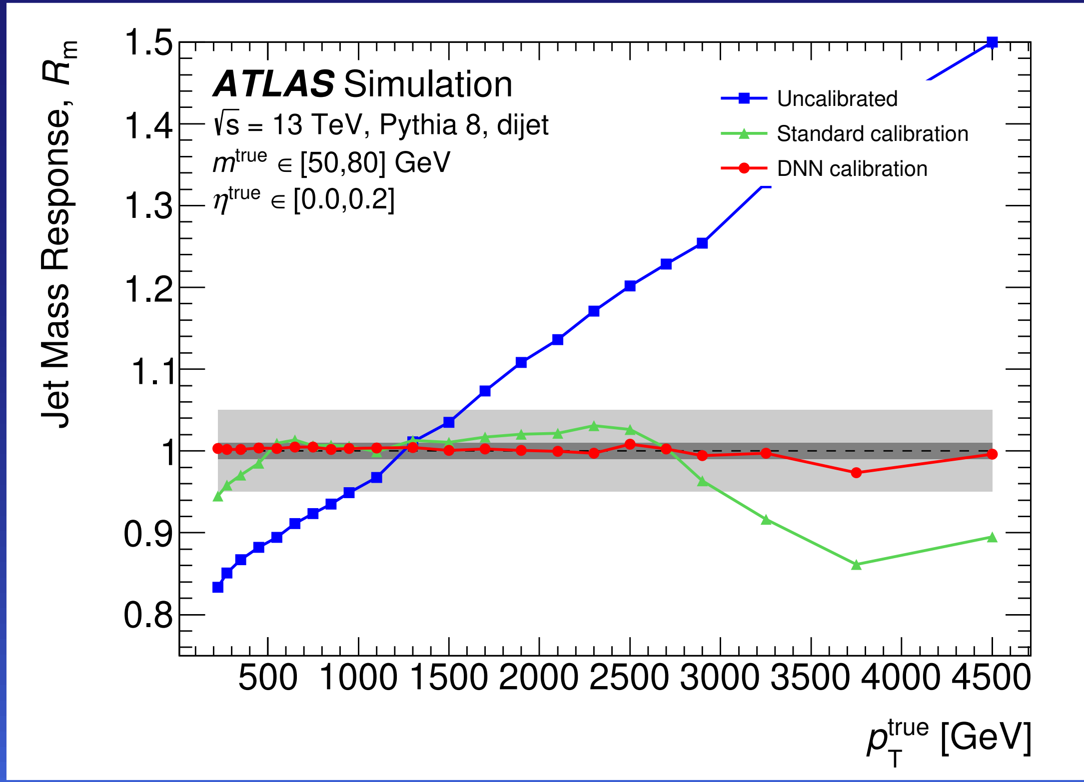
- comparison of the DNN calibration (red) with standard calibration (green) and no calibration (blue)
  - DNN outperforms standard calibration in energy- and mass-scale closure and resolution for both E and m
  - typical resolution improvement of $> 30\%$ for $p_{\perp} > 500$ GeV
  - robust against Pile-Up
  - performs also better on topologies not used in the training (boosted heavy bosons)



Mass response for boosted $W^{\pm}$ / Z^0



Energy response for different calibrations



Mass response for different calibrations

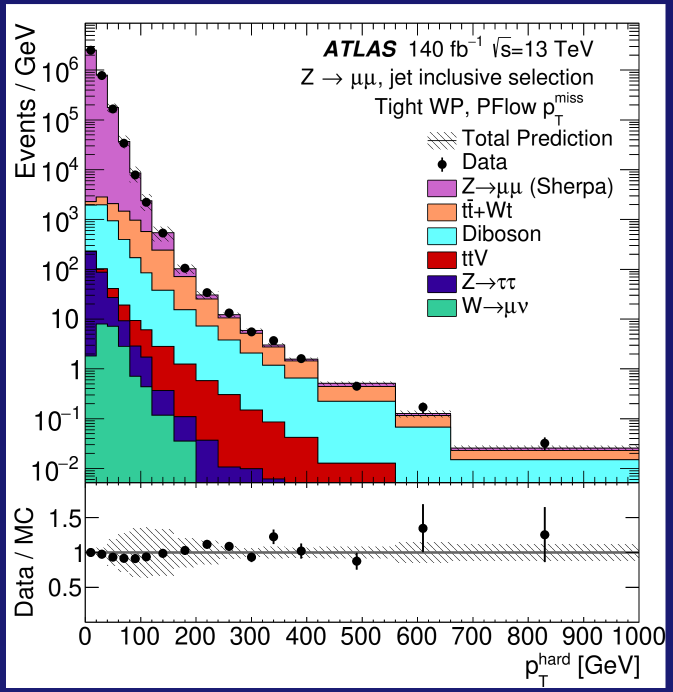
- 2D missing transverse momentum vector $\boldsymbol{p}_\perp^{\text{miss}}$ is derived from $\boldsymbol{p}_\perp^{\text{obj}} = (p_x^{\text{obj}}, p_y^{\text{obj}})$ from all "hard" objects (obj) and a remaining "soft" term from "unused" tracks $\boldsymbol{p}_\perp^{\text{track}}$:

$$\boldsymbol{p}_\perp^{\text{miss}} = -\boldsymbol{p}_\perp^{\text{hard}} - \boldsymbol{p}_\perp^{\text{soft}}, \text{ with } \quad \boldsymbol{p}_\perp^{\text{hard}} = \sum_{\text{obj}=e,\gamma,\tau,\mu,\text{jet}} \boldsymbol{p}_\perp^{\text{obj}} \quad \text{and} \quad \boldsymbol{p}_\perp^{\text{soft}} = \sum_{\text{unused tracks}} \boldsymbol{p}_\perp^{\text{track}}$$

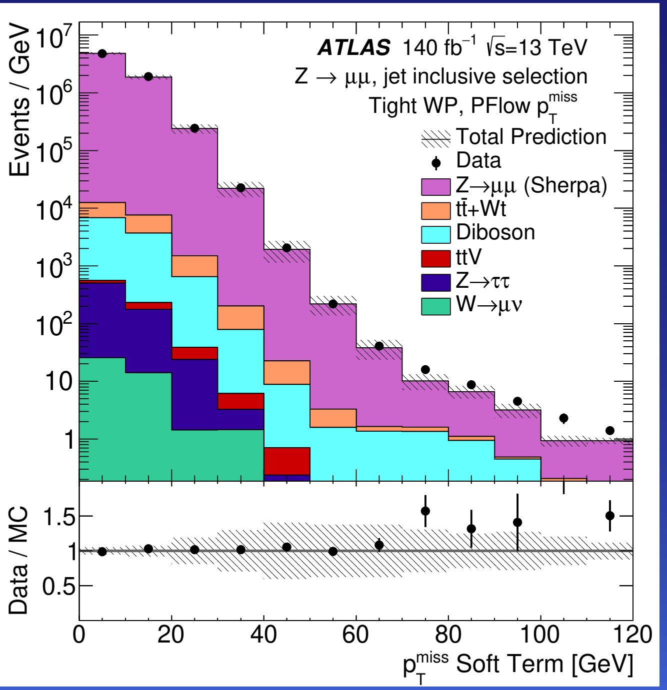
- scalar transverse momentum sum to evaluate the scale:

$$\sum p_\perp = \sum_{\text{obj}=e,\gamma,\tau,\mu,\text{jet}} p_\perp^{\text{obj}} + \sum_{\text{unused tracks}} p_\perp^{\text{track}}$$

- Run-2 performance updated with full Run-2 dataset for use of PFlow objects for jets
- evaluation in $Z^0 \to \mu\mu$ and $Z^0 \to ee$ events (no real $p_\perp^{\text{miss}}$ expected)
  - dominant systematic in $p_\perp^{\text{hard}}$ from JES (bump at ~ 100 GeV)
  - small excess in $p_\perp^{\text{soft}}$ tail in data from fake electrons



$p_\perp^{\text{hard}}$ in $Z^0 \to \mu\mu$



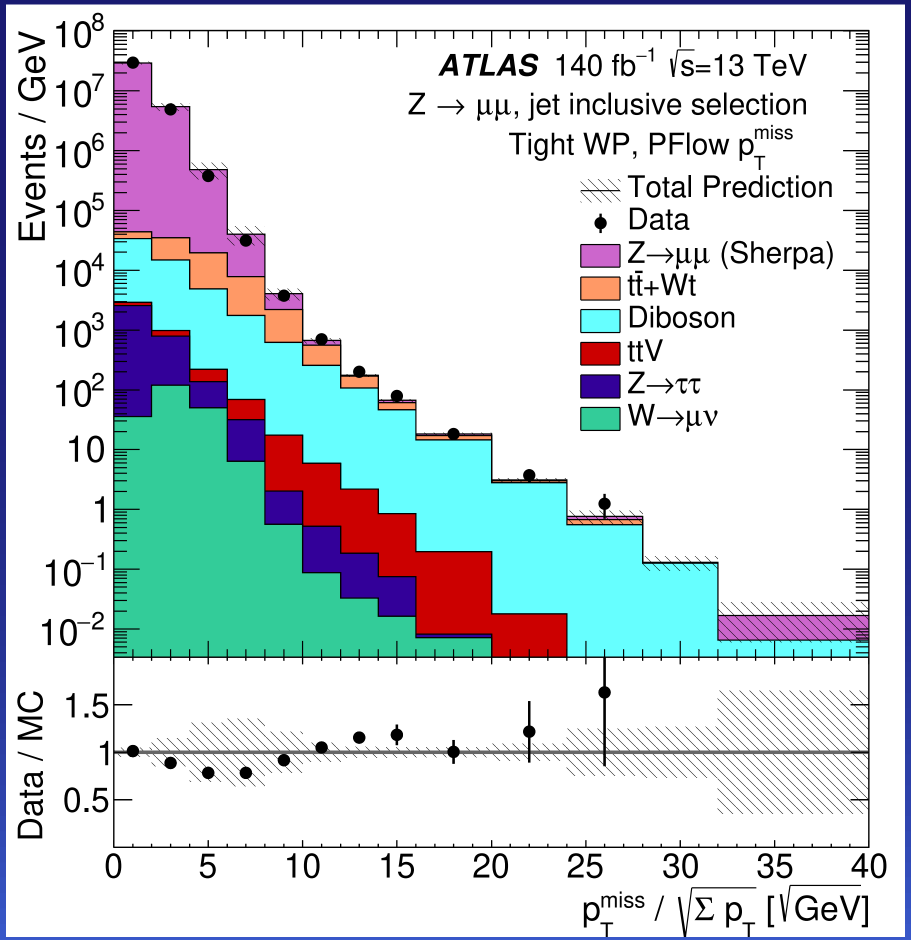
$p_\perp^{\text{soft}}$ in $Z^0 \to \mu\mu$

Missing transverse momentum

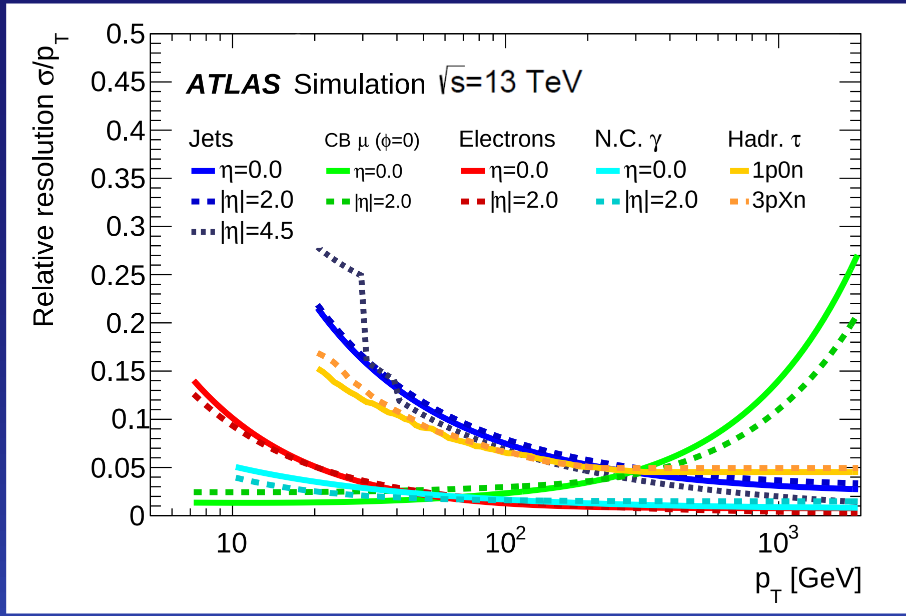
- event-based significance $S_{H_\perp} = p_\perp^{\text{miss}} / \sqrt{H_\perp}$ is based on $H_\perp = \sum_{\text{jet}} p_\perp^{\text{jet}}$, which is approximate only (assumes calorimeter-like resolution)

- missing transverse momentum significance evaluated on object-based uncertainties $\boldsymbol{V}$:

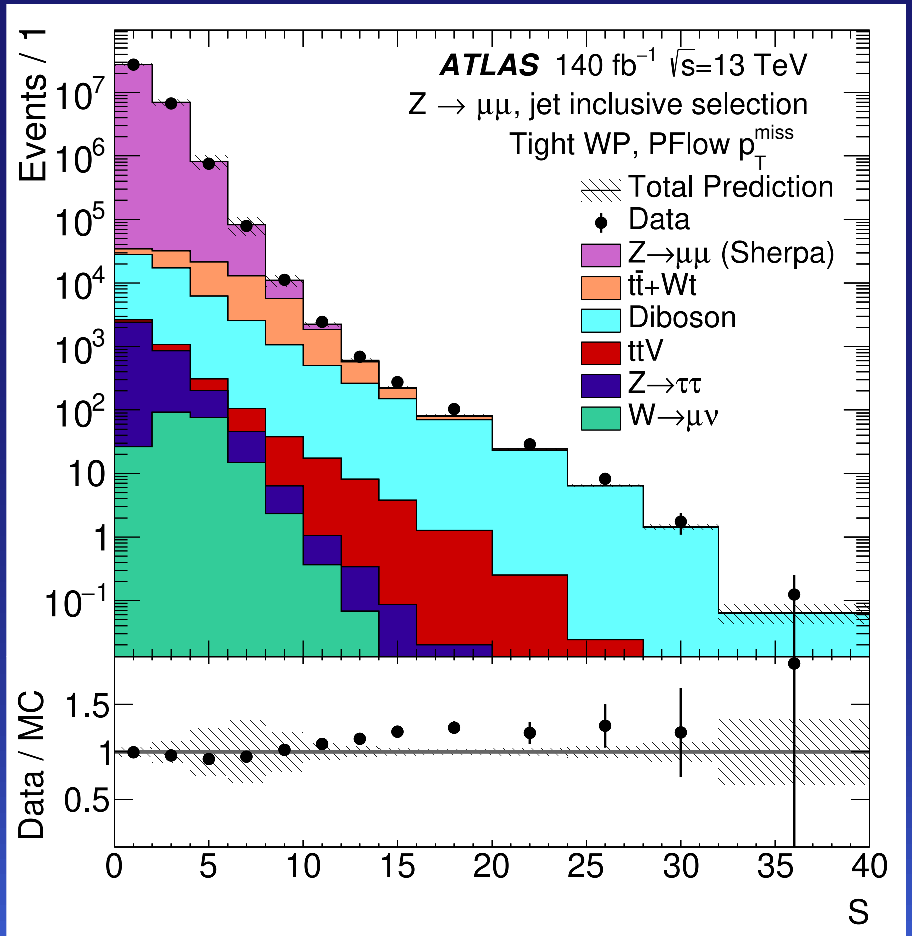
$$S = \sqrt{\boldsymbol{p}_\perp^{\text{miss}\,T} \left(\sum_{\text{obj}} \boldsymbol{V}^{\text{obj}} \right)^{-1} \boldsymbol{p}_\perp^{\text{miss}}}$$



$p_\perp^{\text{miss}} / \sqrt{\sum p_\perp}$ in $Z^0 \to \mu\mu$



relative resolutions of the objects entering $p_\perp^{\text{miss}}$



Object-based significance S in $Z^0 \to \mu\mu$

Conclusions

- reconstruction and calibration of hadronic objects in ATLAS is a very active field
  - Pile-Up remains the biggest challenge
  - time as a new discriminant in calorimetry helps reducing it
  - new ML-based techniques start to replace legacy calibration methods for energy and mass
- Run-2 performance results:
- jet calibration
  - $\mathcal{O}(1\%)$ precision reached for jet energy scale, $\mathcal{O}(15 - 30\%)$ improvements in resolution for energy and mass
  - additional b-jet energy scale uncertainty measured to $\mathcal{O}(1\%)$
- missing transverse momentum
  - benefits from reconstruction and calibration advancements – especially from jets
  - object-based significance sharpens the MET discrimination power
- Run-3 analyses benefit from these improvements