Tracking, alignment and flavor-tagging performance in ATLAS and CMS

LHCP 2020

The Eighth Annual Conference on Large Hadron Collider Physics

May 25-30, 2020

Paris France, Sorbonne Université

(IN2P3/CNRS, IRFU/CEA)

Topics:

Standard Model and Beyond, Higgs Boson, Flavour, Heavy-Ions

http://lhcp2020.fr

IFIC INSTITUT DE FÍSICA CORPUSCULAR

VNIVERSITAT DE VALÈNCIA

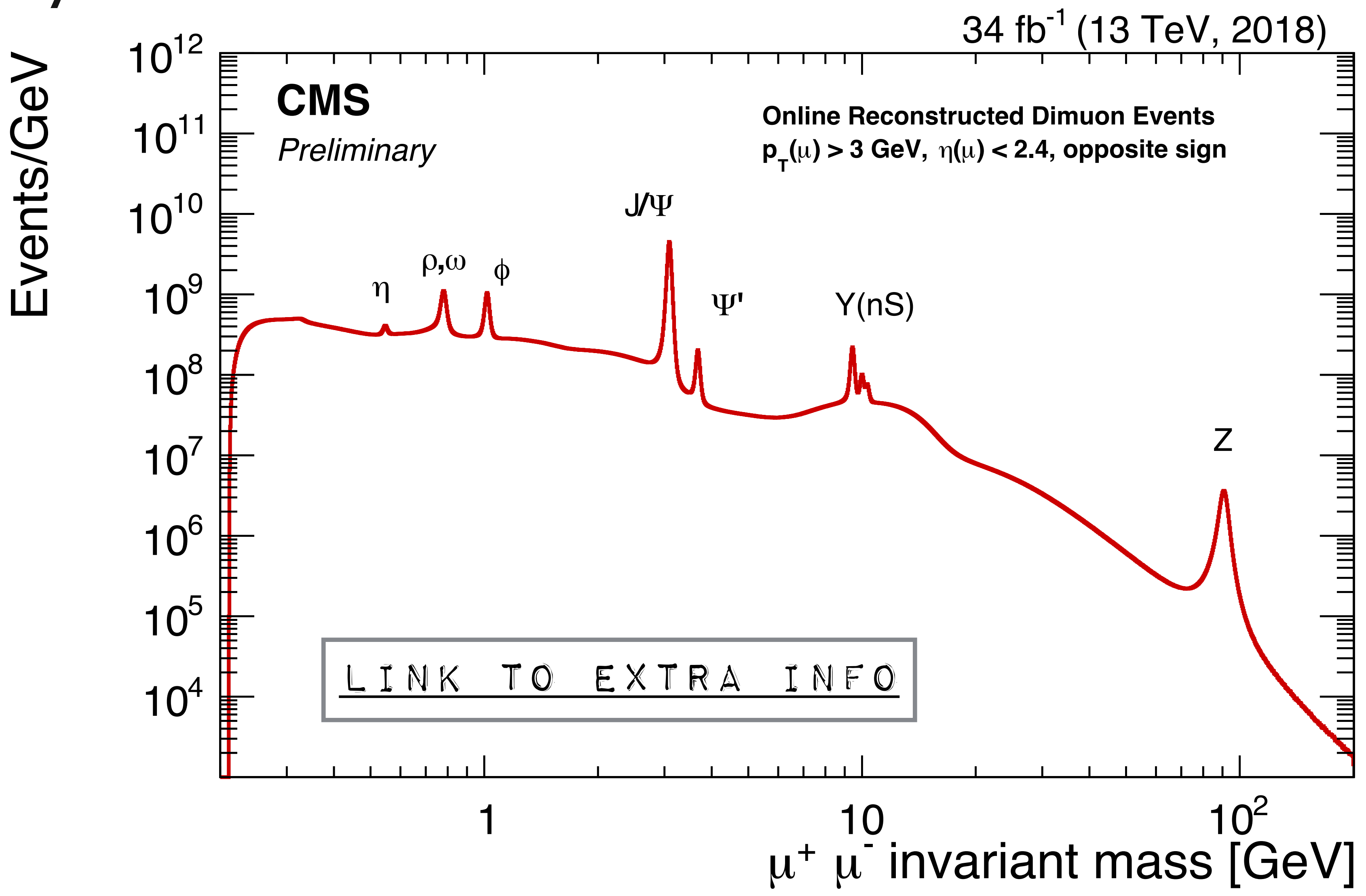
CSIC CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

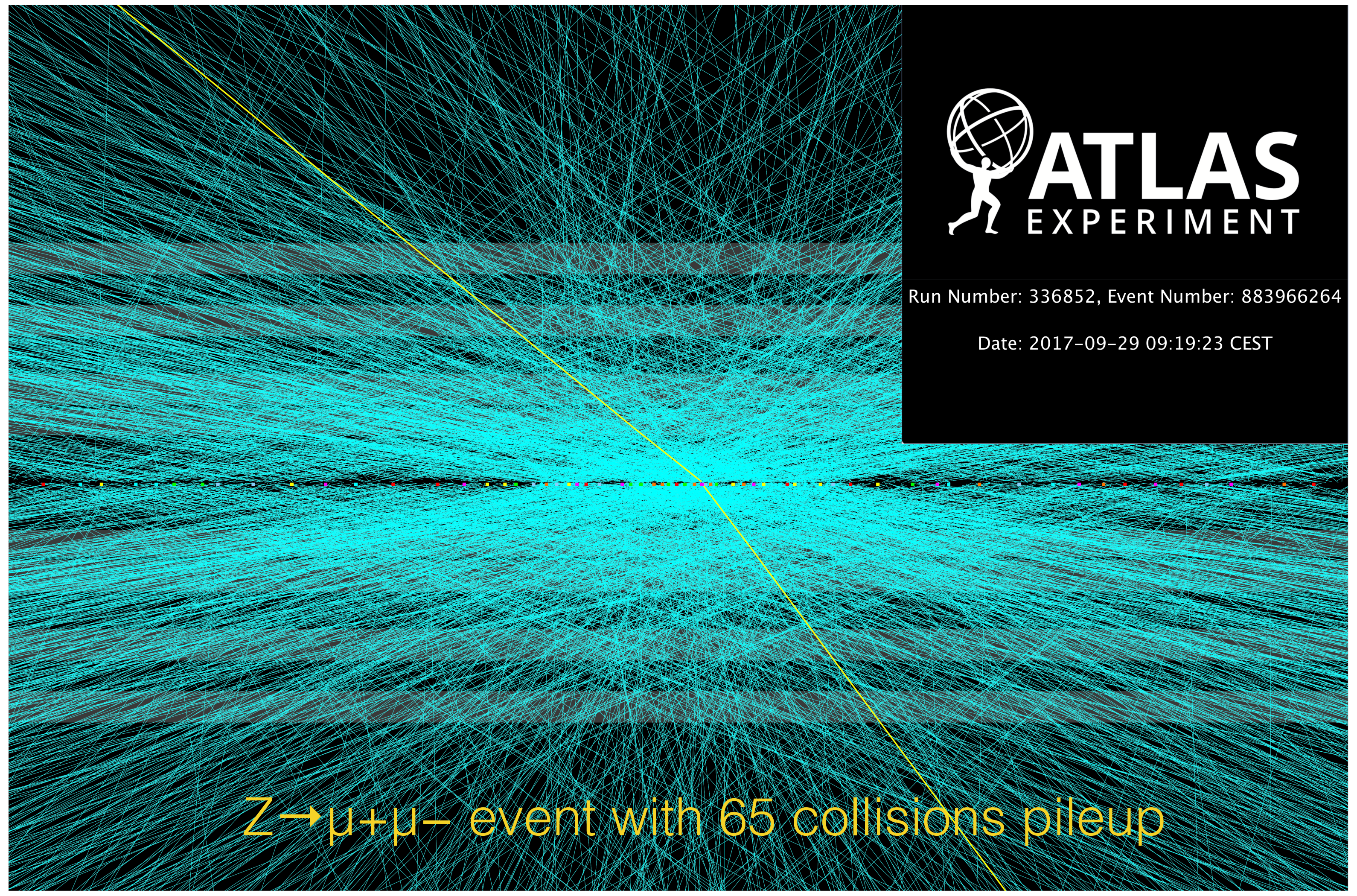
ATLAS IFIC VALÈNCIA

Introduction

- A good performance of the trackers is a key ingredient of the success of the physics program
  - An accurate determination of the charged particles properties is necessary for
    - Momentum & impact parameter
    - Invariant masses have to be determined with precision
    - Secondary vertices must be fully reconstructed: evaluate short lifetimes
    - Others: Kink reconstruction, despairing tracks…

- Challenges for the tracking systems of the LHC detectors
  - Momenta of particles ranging from MeV to TeV
  - High multiplicity of charged particles (up to 1000 for $\mathscr{L} \sim 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$)
    - Even higher for heavy ion collisions
  - Large background from secondary activities of the particles
  - Multiple Coulomb Scattering in detector frames, supports, cables, pipes.
  - Complex modular tracking systems
    - combining different detecting technologies, different resolutions
  - Varying detector resolutions
    - Radius, polar angle (θ) or pseudorapidity (η)
  - Very high event rates
    - large amount of data with demanding high requirements of:
      - CPU and storage
      - Tracking CPU budget

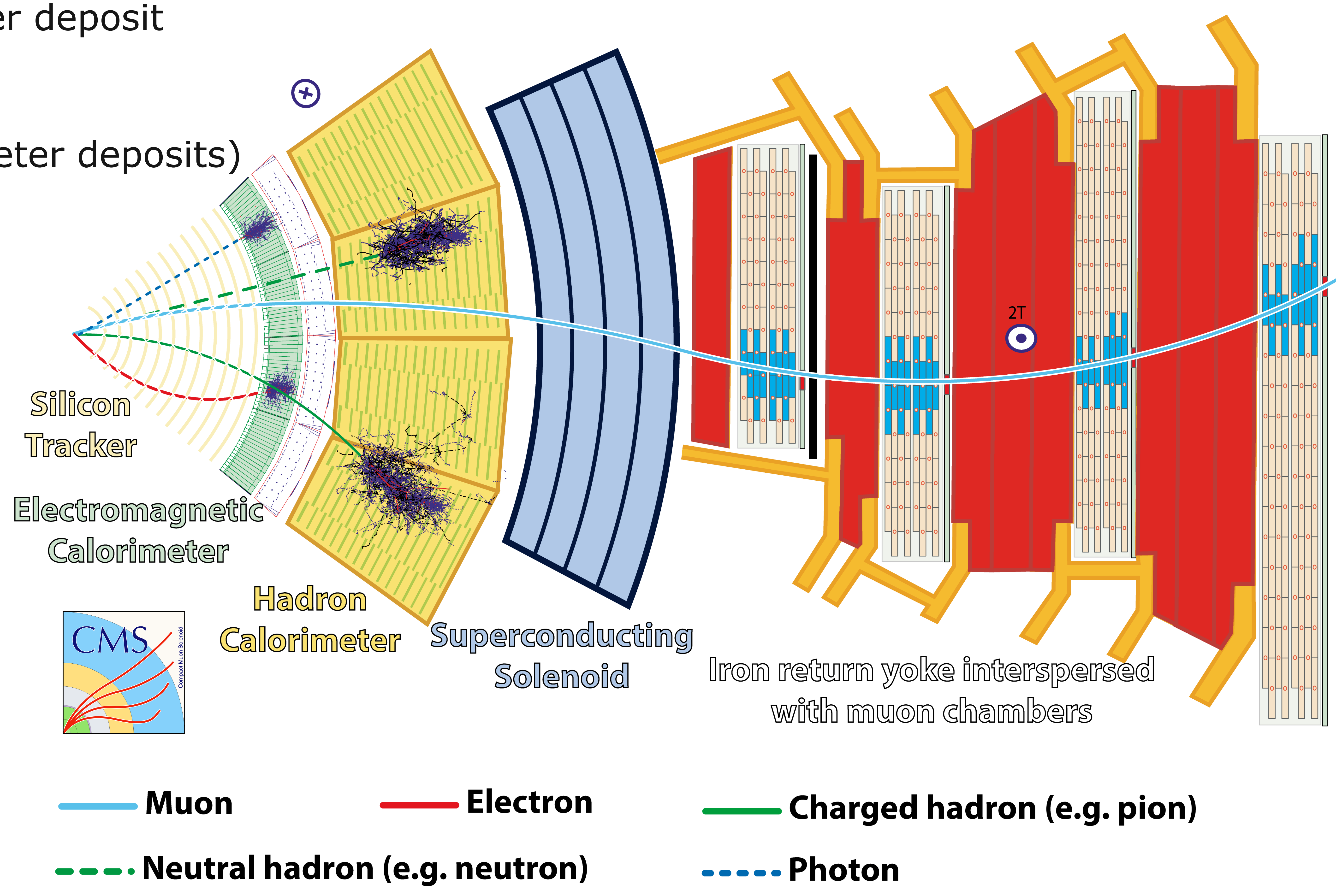




Z→μ+μ− event with 65 collisions pileup

Charged particles in the detector

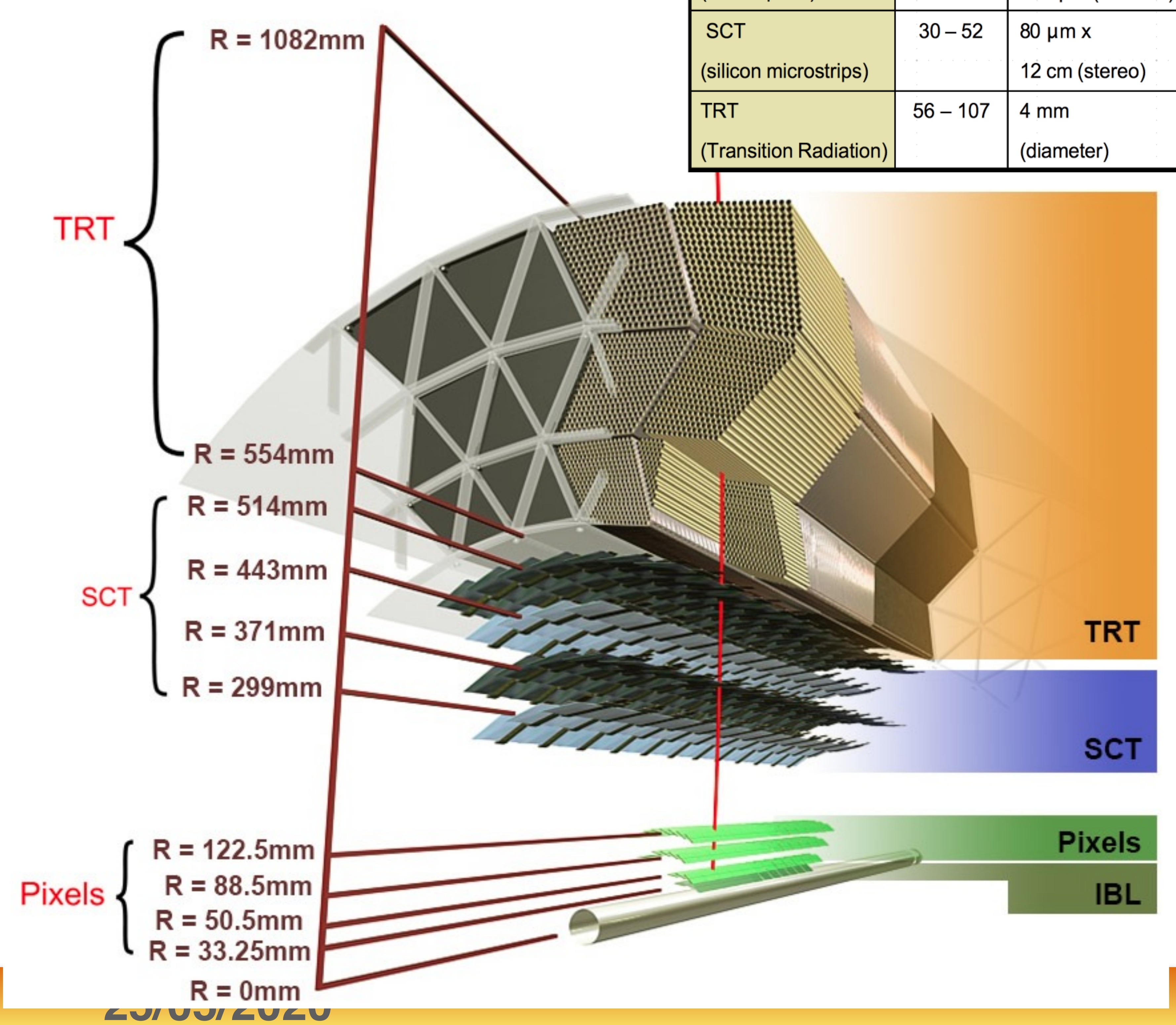
- Outgoing particles leave different signatures in various sub-detectors
  - Muons:
    - Combined track in inner tracker and muon spectrometer
  - Electrons:
    - Bremstrahlung corrected tracking and EM calorimeter deposit
  - Tracks in jets
    - Reconstructed in the inner tracker (EM/HAD calorimeter deposits)
    - High density of tracks
    - flavour tagging and hadronic taus identification
  - Photons:
    - No tracking
    - But track reconstruction for γ conversions



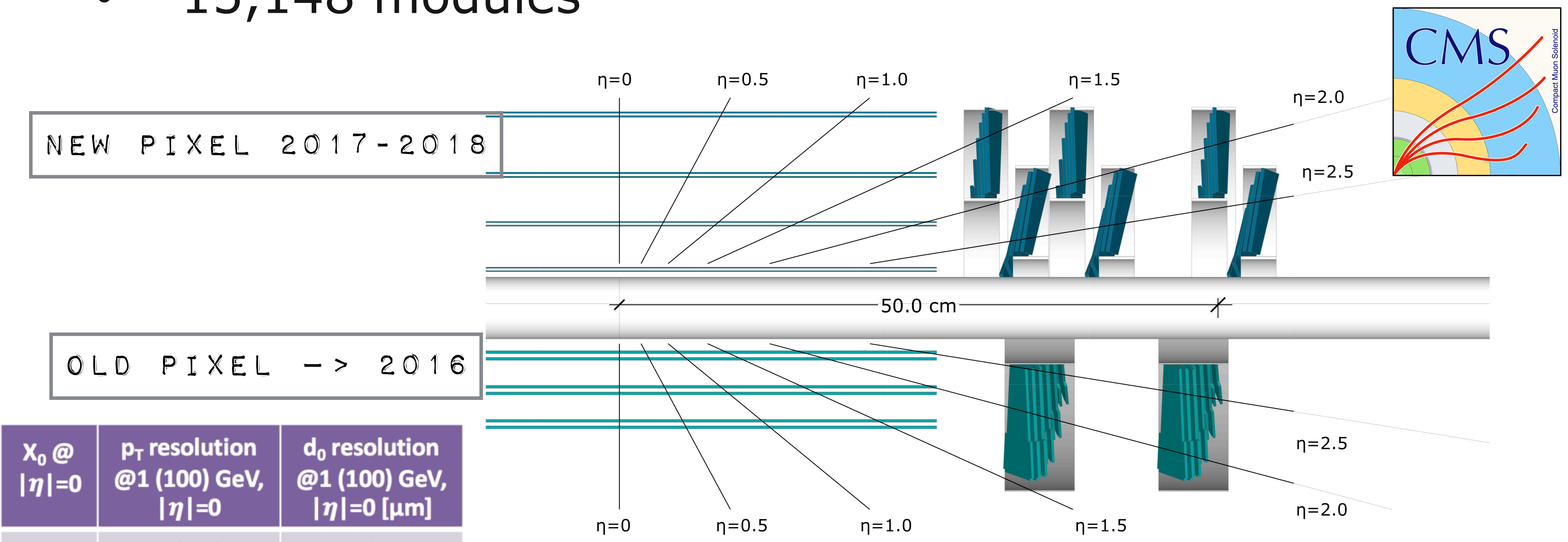
ATLAS & CMS Tracking systems

- ATLAS: Inner Detector (ID)
  - |η| < 2.5 & 2 T solenoid field
  - Pixel and IBL
    - 1744+280 modules
  - strips: SCT (double sided)
    - 4088 modules
  - Gas drift tubes: TRT (30 measurements)

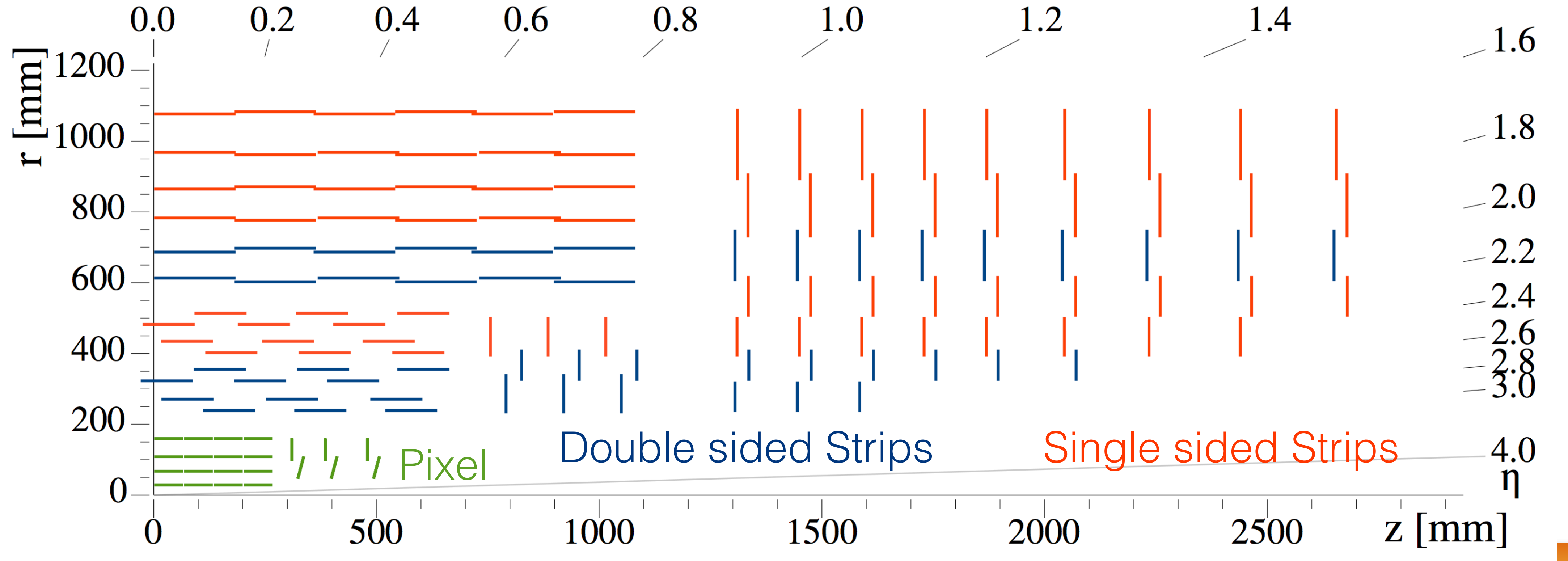
Subdetector	r (cm)	Elements size	Resolution (X * Y)	hits/track (average)	channels
Pixels & IBL (silicon pads)	5 – 12.5	50 µm x 400 µm (250 IBL)	10 µm x 115 µm	3 Pixels 1 IBL	92×10^6
SCT (silicon microstrips)	30 – 52	80 µm x 12 cm (stereo)	17 µm x 580 µm	8	6.3×10^6
TRT (Transition Radiation)	56 – 107	4 mm (diameter)	130 µm	30	3.5×10^5



- The CMS tracker has a lηl < 3 coverage
  - All silicon: |η| < 2.5 & 3.8 T solenoid field
  - Pixel (replaced in 2017. Phase 1)
    - 4 barrel layers & 3 end-cap disks
    - 1856 modules & 124M channels
  - Strips
    - 10 barrel layers (4 double sided) & 12 end-cap disks
    - 15,148 modules



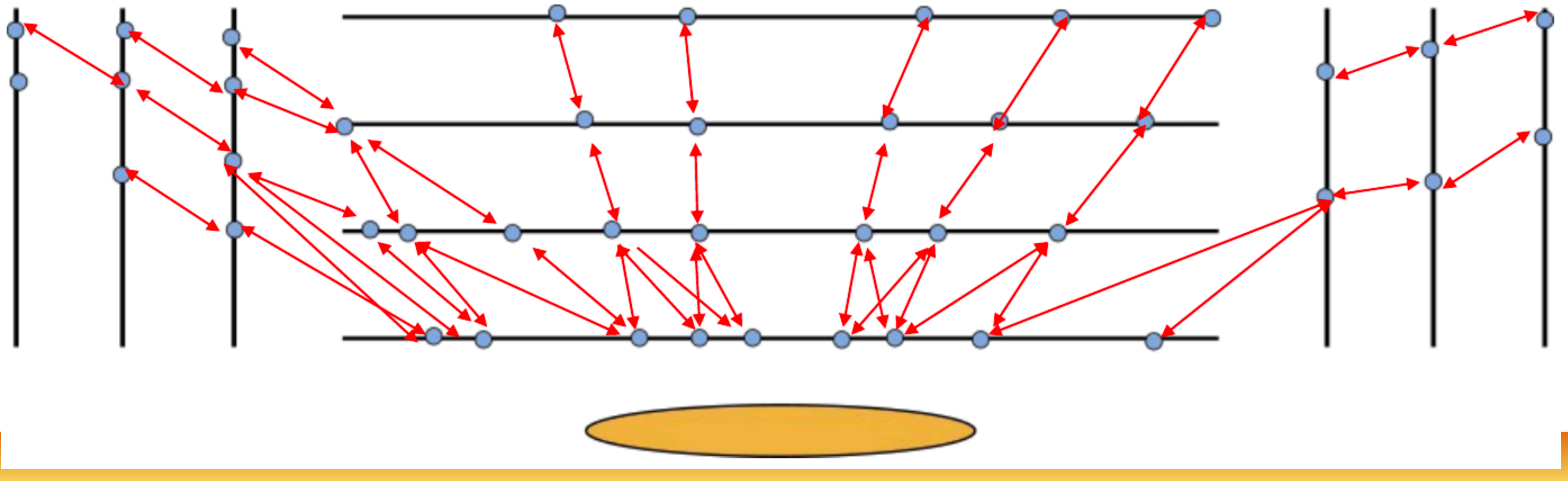
R_{inner} [cm]	R_{outer} [m]	$\lvert\eta\rvert$ coverage	B field [T]	X_0 @ $\lvert\eta\rvert=0$	p_T resolution @1 (100) GeV, $\lvert\eta\rvert=0$	d_0 resolution @1 (100) GeV, $\lvert\eta\rvert=0$ [µm]
3	1.1	3.0	3.8	0.4	0.7 (1.5)%	90 (20)



Track reconstruction

- From detector hits to tracks:
  - Track reconstruction refers to the process of using the hits to obtain estimates for the momentum and position parameters of the charged particles
- CMS: Combinatorial Track Finder (CTF)
  - Kalman filter for pattern recognition and track fitting
  - Track classification / selection is done using BDT trained for each step
  - Iterative process
    - initial iterations search for tracks that are easiest to find (large p_T & near interaction region)
    - After each iteration, hits associated with tracks are removed (reduce combinatorics)
    - Further iterations:
      - search for more difficult classes of tracks (low-p_T, or greatly displaced tracks)

- Cellular Automaton
  - Implemented to cope with the increase of luminosity
    - triplet & quadruplet pixel track seeds
  - Parallelism, keep high efficiency and low fake rate
  - Avoid $\mathcal{O}(n^3)$ combinatorics
    - p_T and interaction region
  - Documentation: F. Pantaleo



Seed generation

Pixel, strips or mix using only a few (2 or 3 or 4) hits. Initial estimates

Track finding

Kalman filter. Extrapolation searching for additional hits

Track fitting

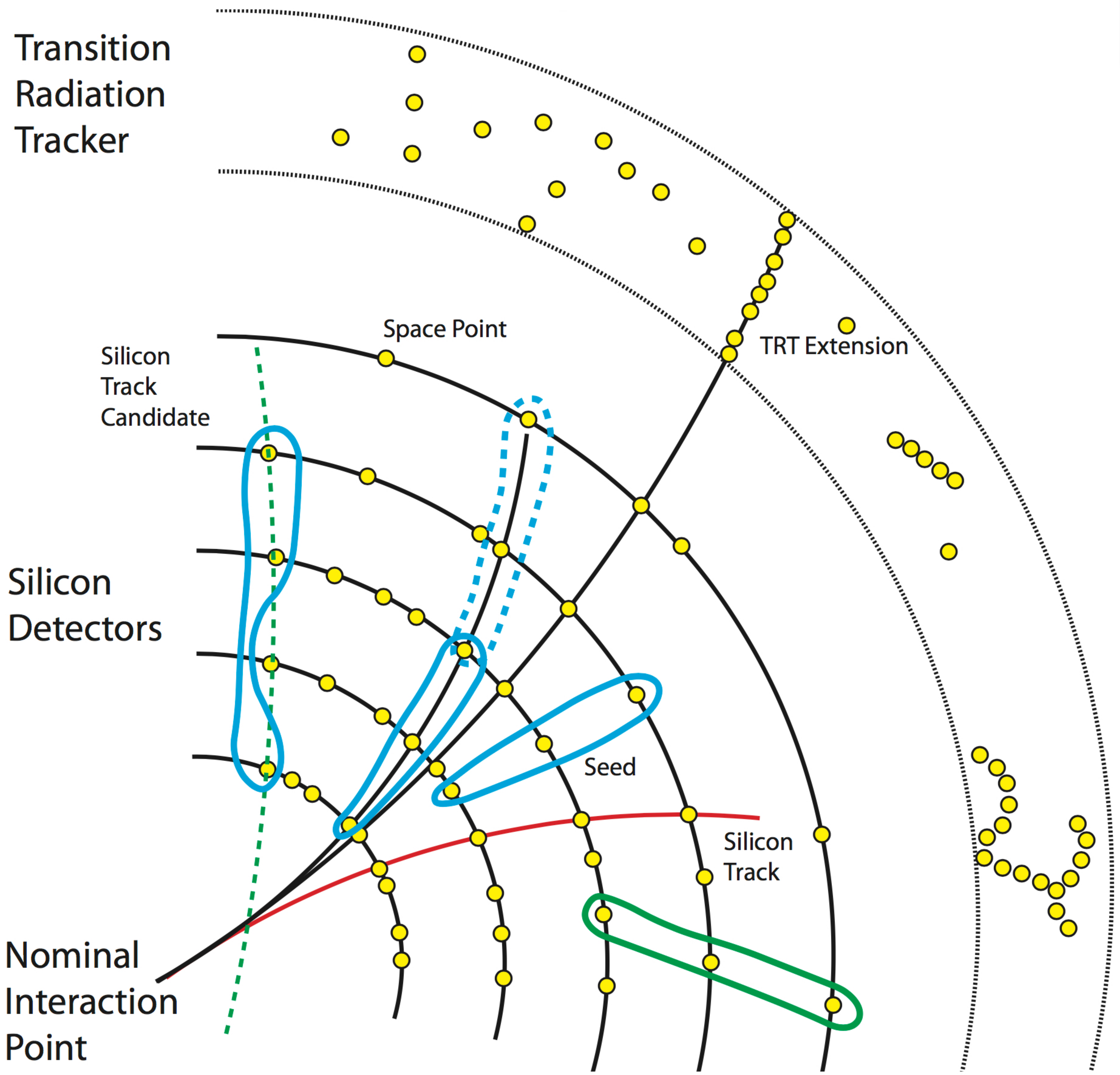
Kalman filter & smoother. Best estimate of track parameters

Track selection

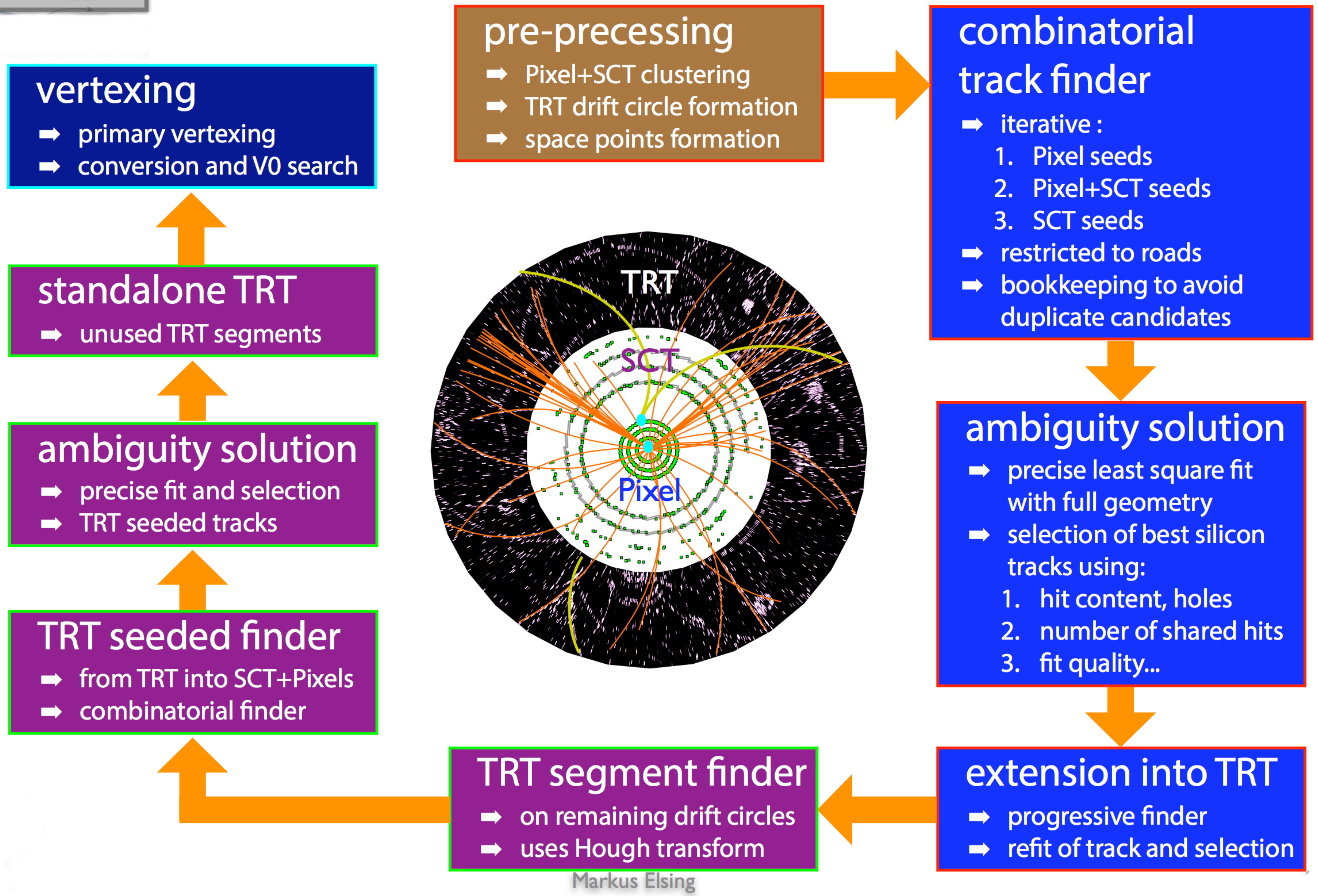
Track quality requirements & compatibility with Interaction Region

Track reconstruction

- From detector hits to tracks



ATLAS NewTracking Software Chain



pre-precessing
- ➡ Pixel+SCT clustering
- ➡ TRT drift circle formation
- ➡ space points formation

combinatorial track finder
- ➡ iterative :
  1. Pixel seeds
  2. Pixel+SCT seeds
  3. SCT seeds
- ➡ restricted to roads
- ➡ bookkeeping to avoid duplicate candidates

ambiguity solution
- ➡ precise least square fit with full geometry
- ➡ selection of best silicon tracks using:
  1. hit content, holes
  2. number of shared hits
  3. fit quality...

extension into TRT
- ➡ progressive finder
- ➡ refit of track and selection

TRT segment finder
- ➡ on remaining drift circles
- ➡ uses Hough transform

TRT seeded finder
- ➡ from TRT into SCT+Pixels
- ➡ combinatorial finder

ambiguity solution
- ➡ precise fit and selection
- ➡ TRT seeded tracks

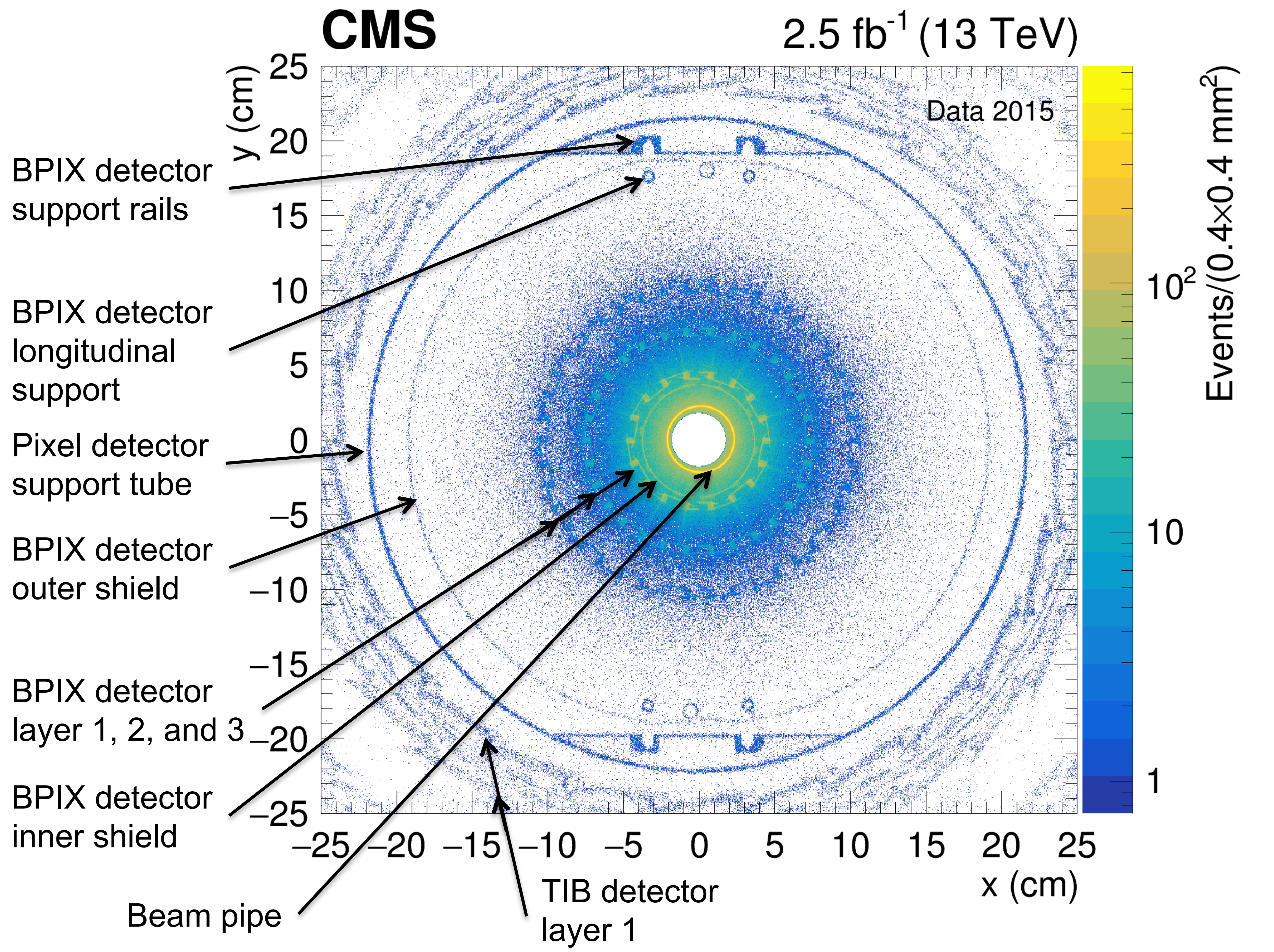
standalone TRT
- ➡ unused TRT segments

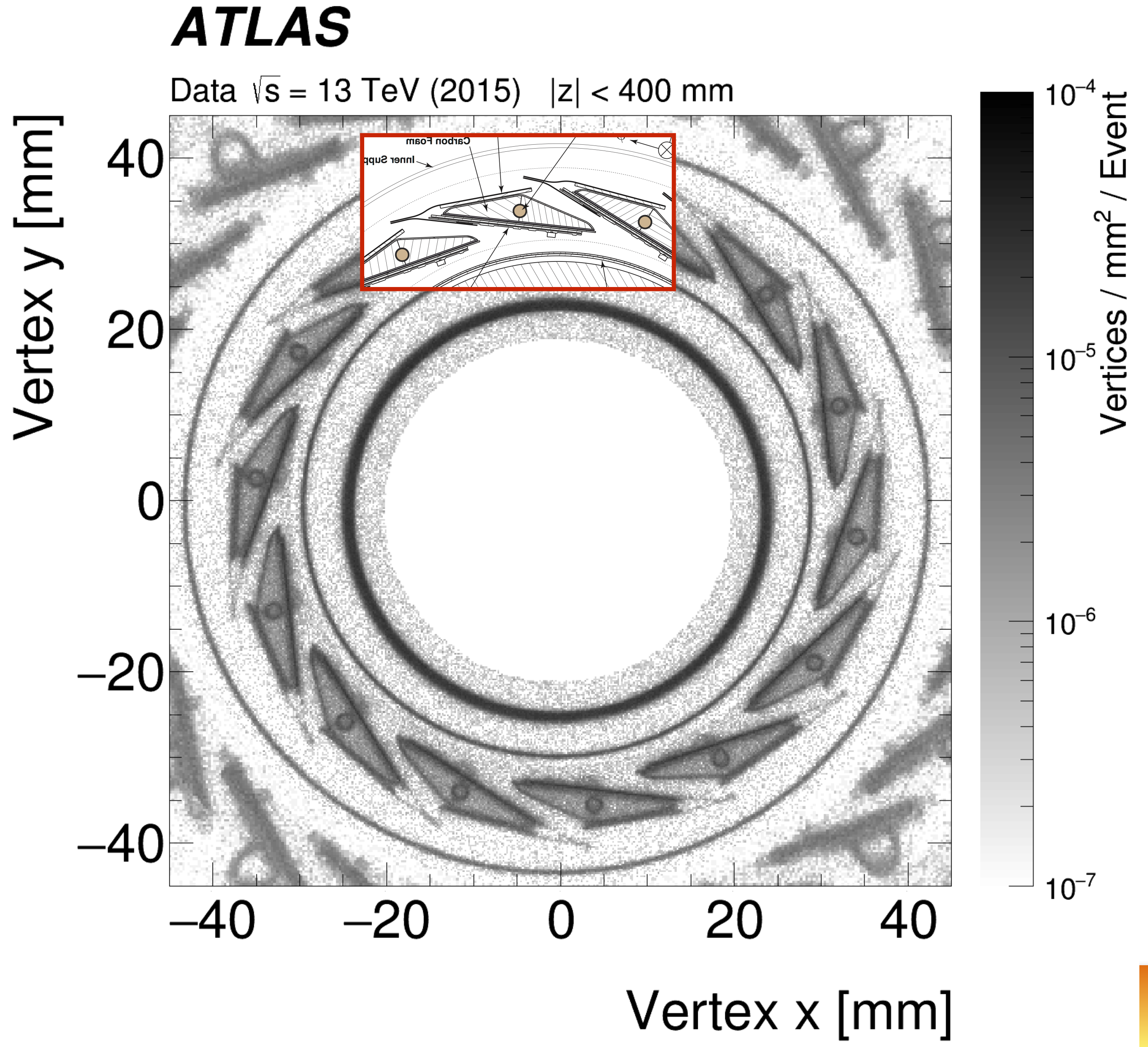
vertexing
- ➡ primary vertexing
- ➡ conversion and V0 search

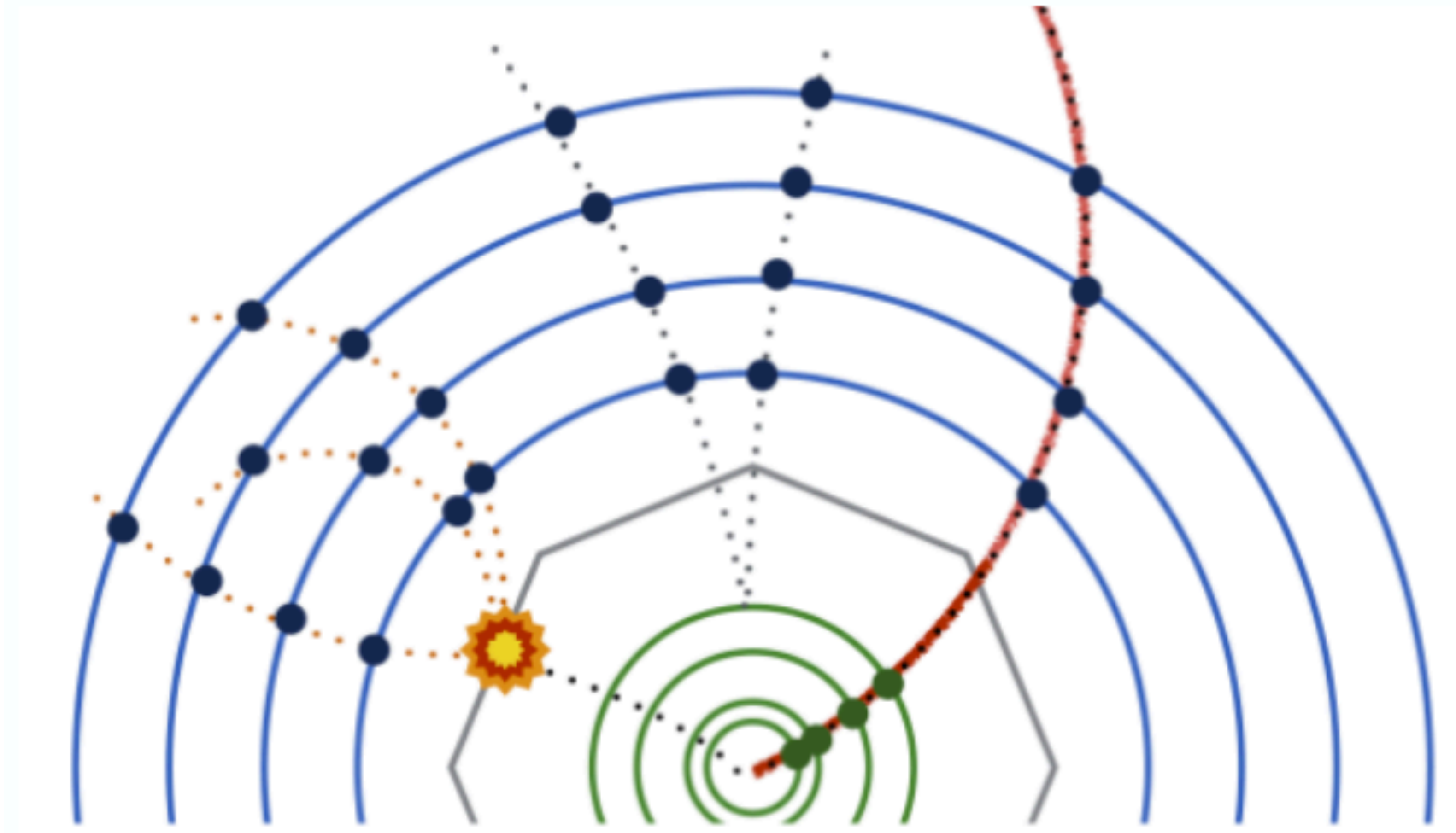
Markus Elsing

Material

- The precise knowledge of the material traversed by the particles is a key ingredient for a performant (precise) track reconstruction
  - The material affects the reconstruction of tracks through multiple scattering, energy loss, electron bremsstrahlung, photon conversions, and nuclear interactions
- Previous material knowledge from the detector design
  - Components & composition
  - Location of detector & ancillary elements
    - Beam pipe & supporting structures
    - Readout electronics
    - HV, LV cablesCooling pipes, etc
- Secondary interactions help to map the material
  - Hadronic interactions
  - Photon conversions



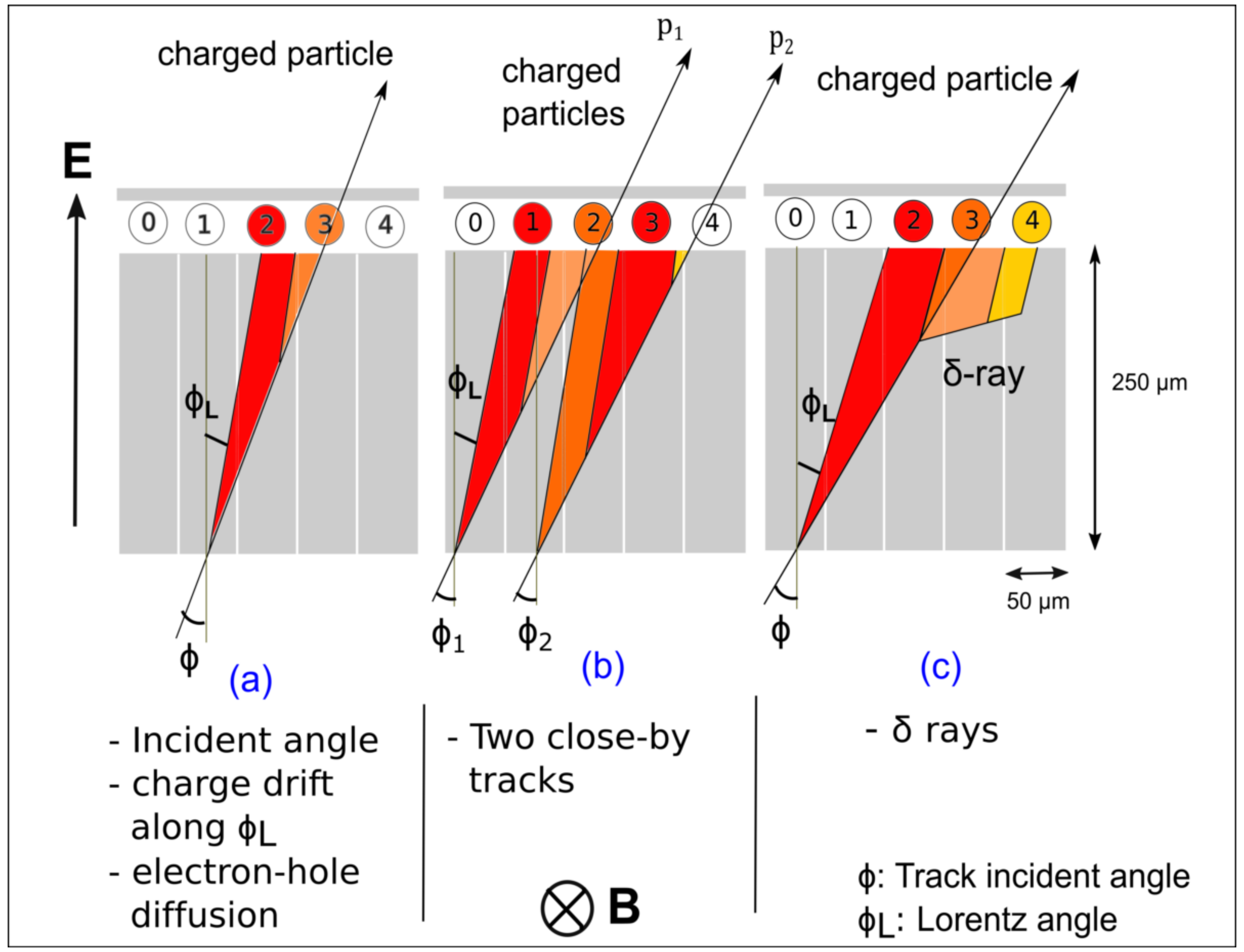


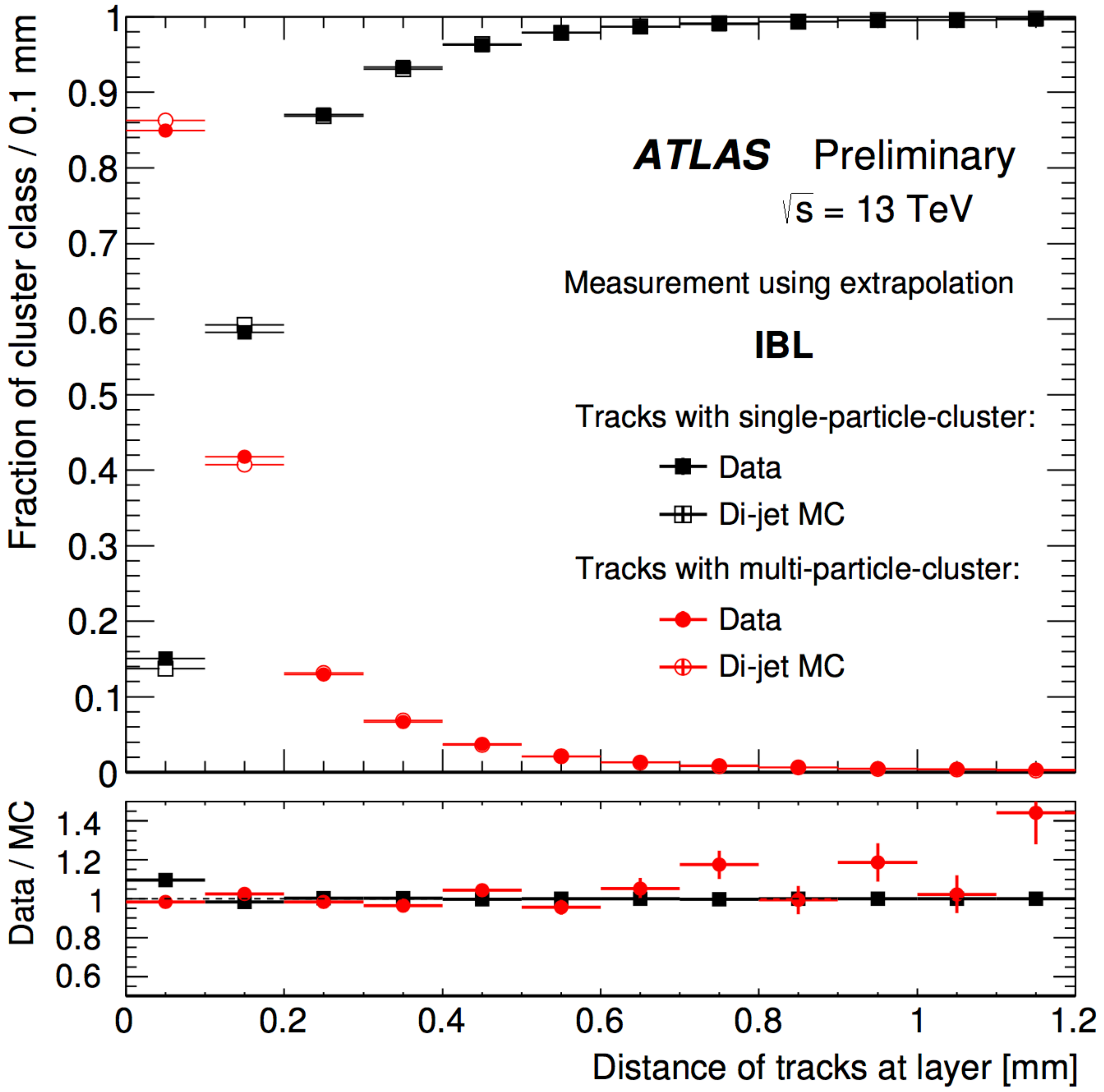


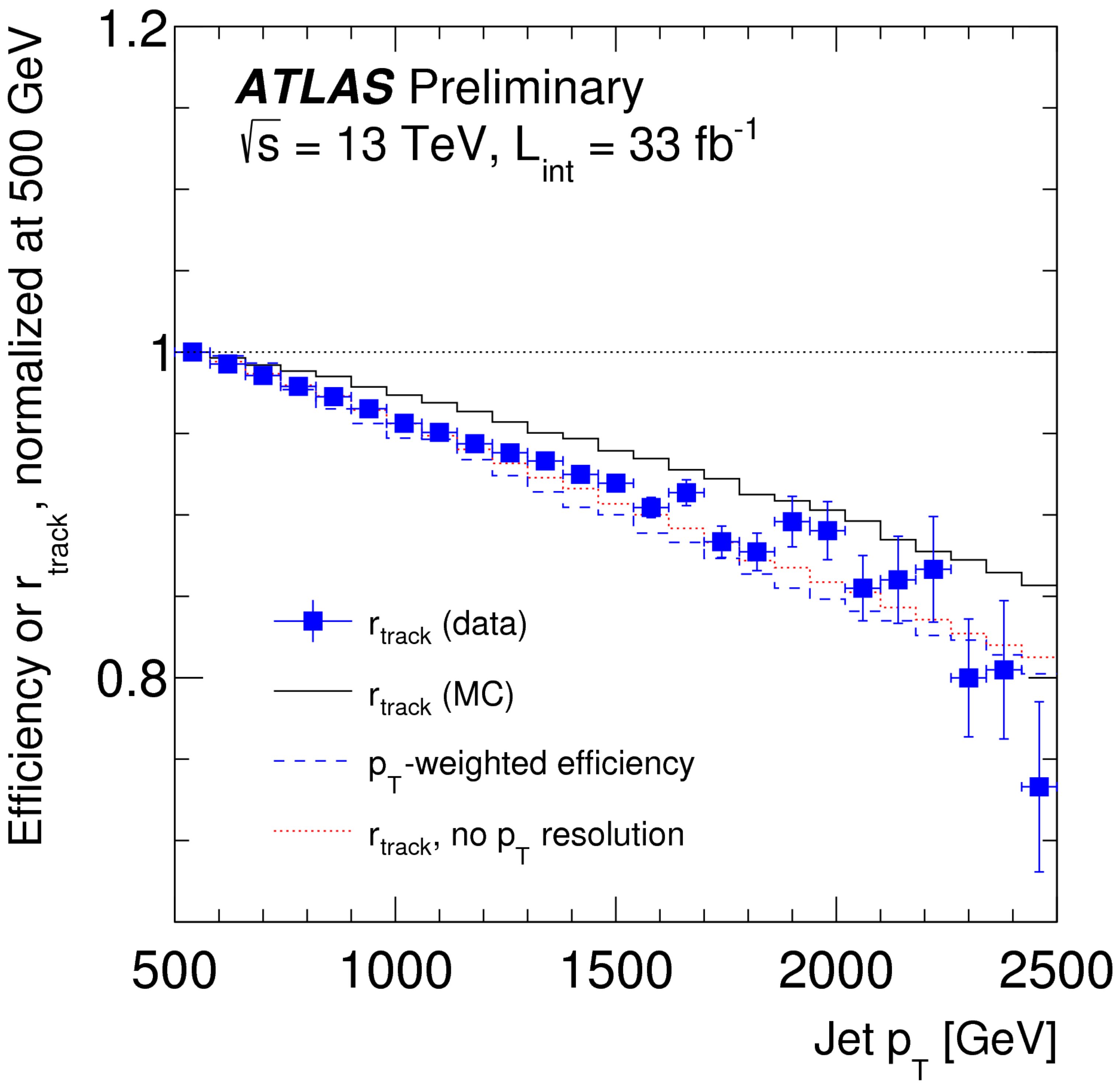
Tracking in dense environments

- The tracking inside jet core becomes inefficient in high transverse momentum jets
  - the collimated environment produces merged cluster from different tracks on the Pixel detector
  - Shared clusters → Separation of tracks inside jets can be smaller than pixel size
  - Tracks with many shared clusters → low quality & rejected

- Recover performance with a NN approach
  - separate clusters originating from single and multiple particles and to estimate hit positions within clusters

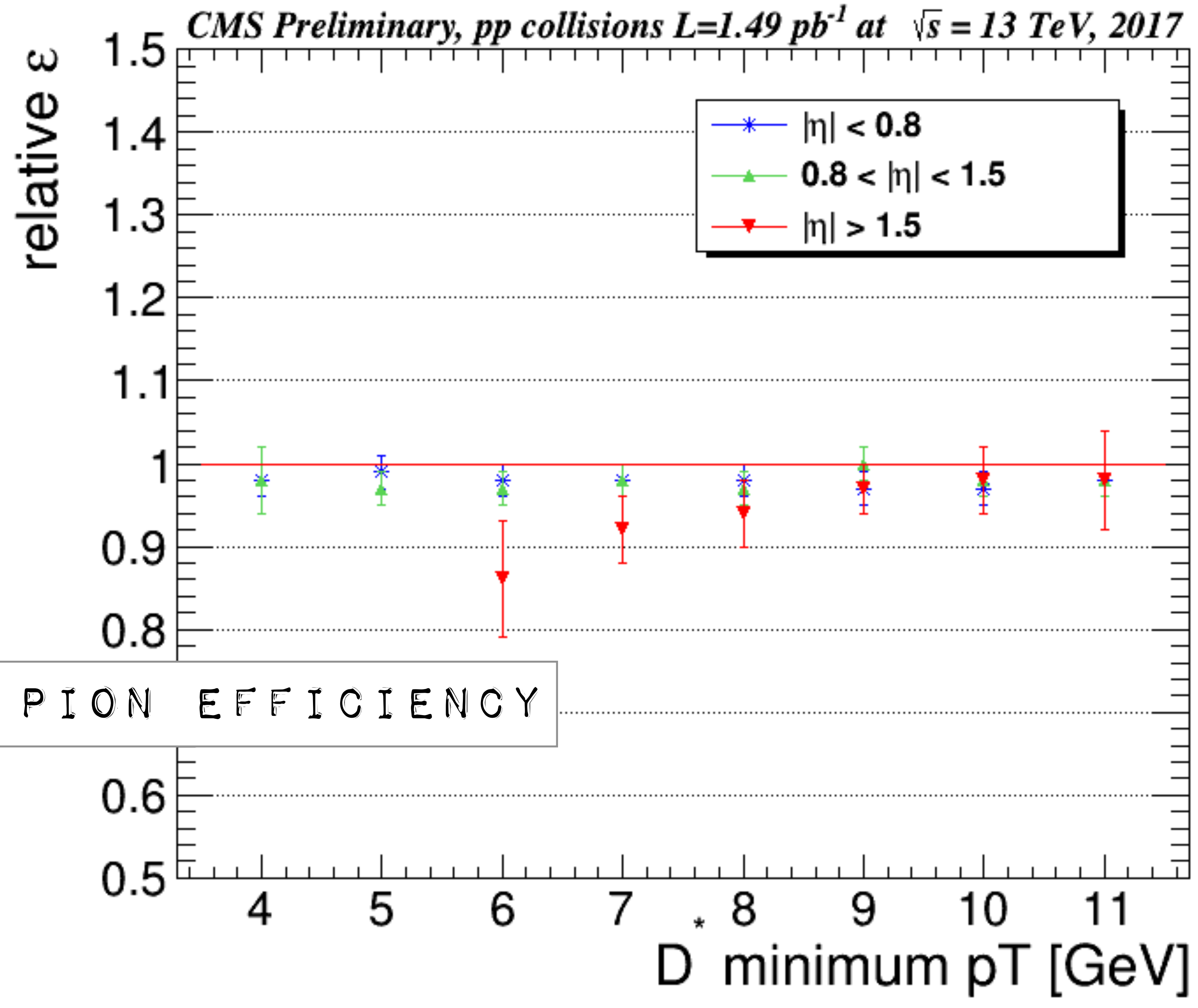




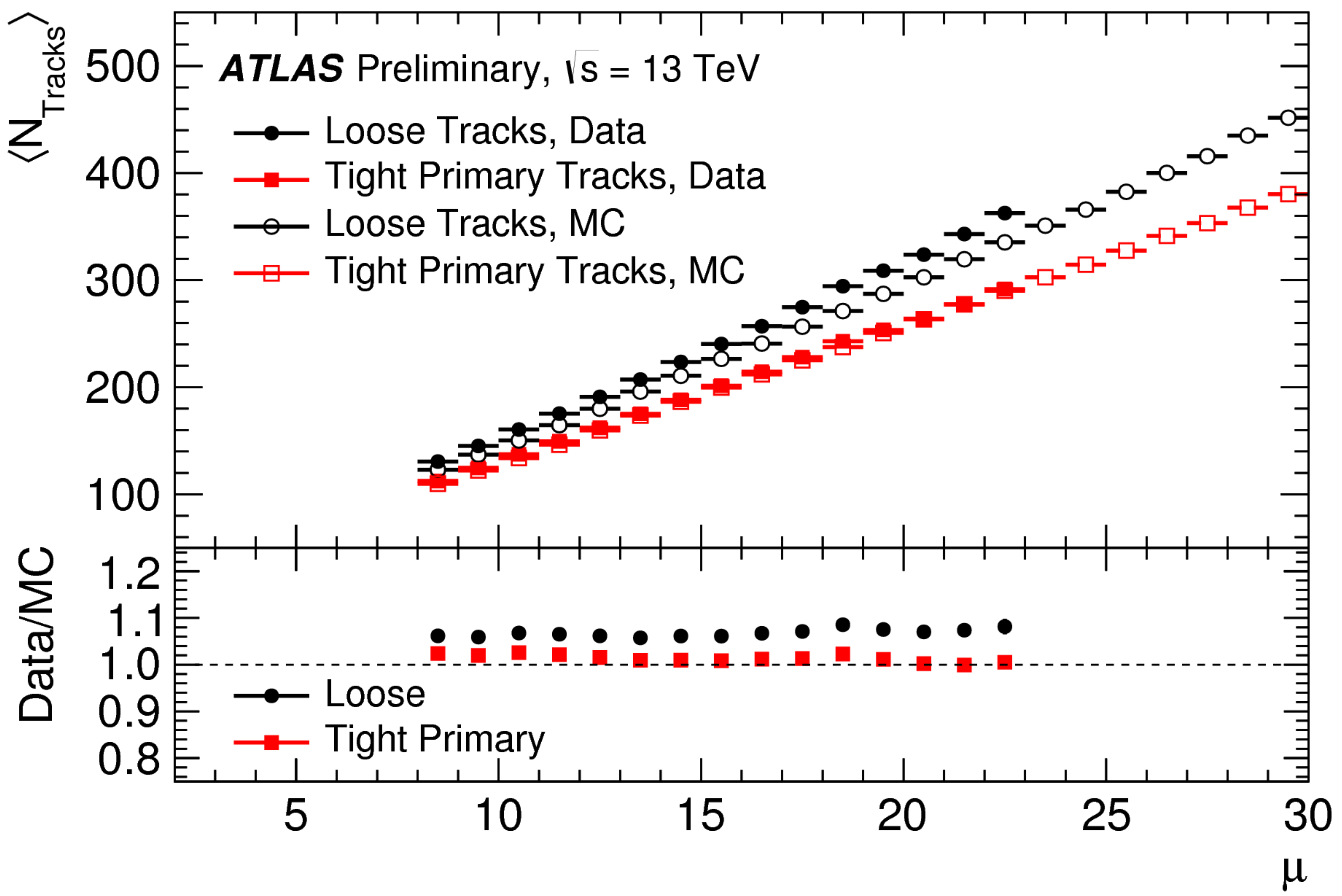


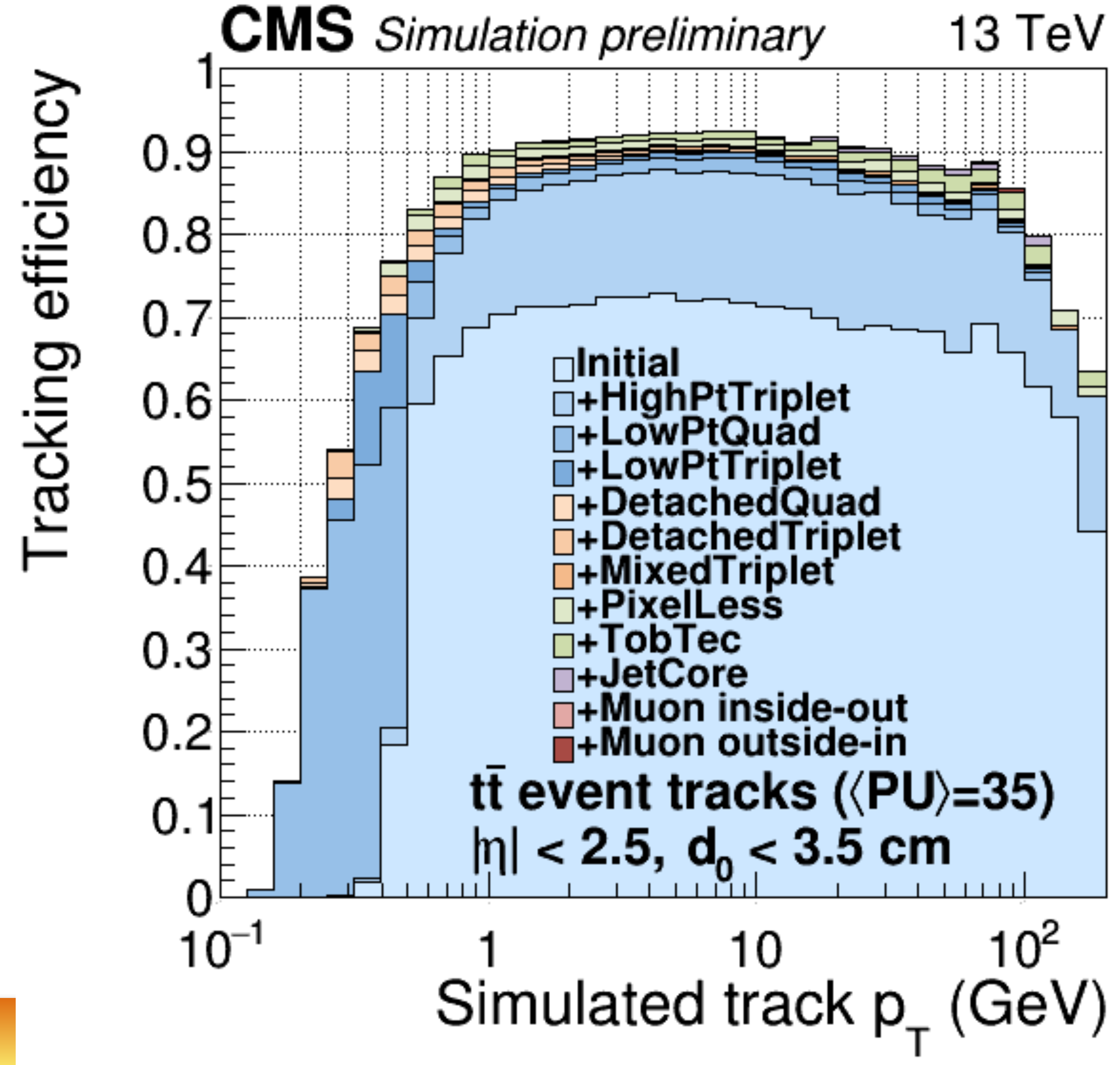
Tracking efficiency

- CMS uses a tag and probe method
  - Pions from D* →kππ decays
  - J/ψ & Z decays to μ+μ−
    - Tracker-only seeded collection
    - All-Tracks collection:
      - Tagged muons from the Muon System as seed for tracks in the inner tracker
- ATLAS
  - MC based
  - Efficiency study vs interactions per beam crossing
    - Loose & Tight tracks selection



CHARGED PION EFFICIENCY

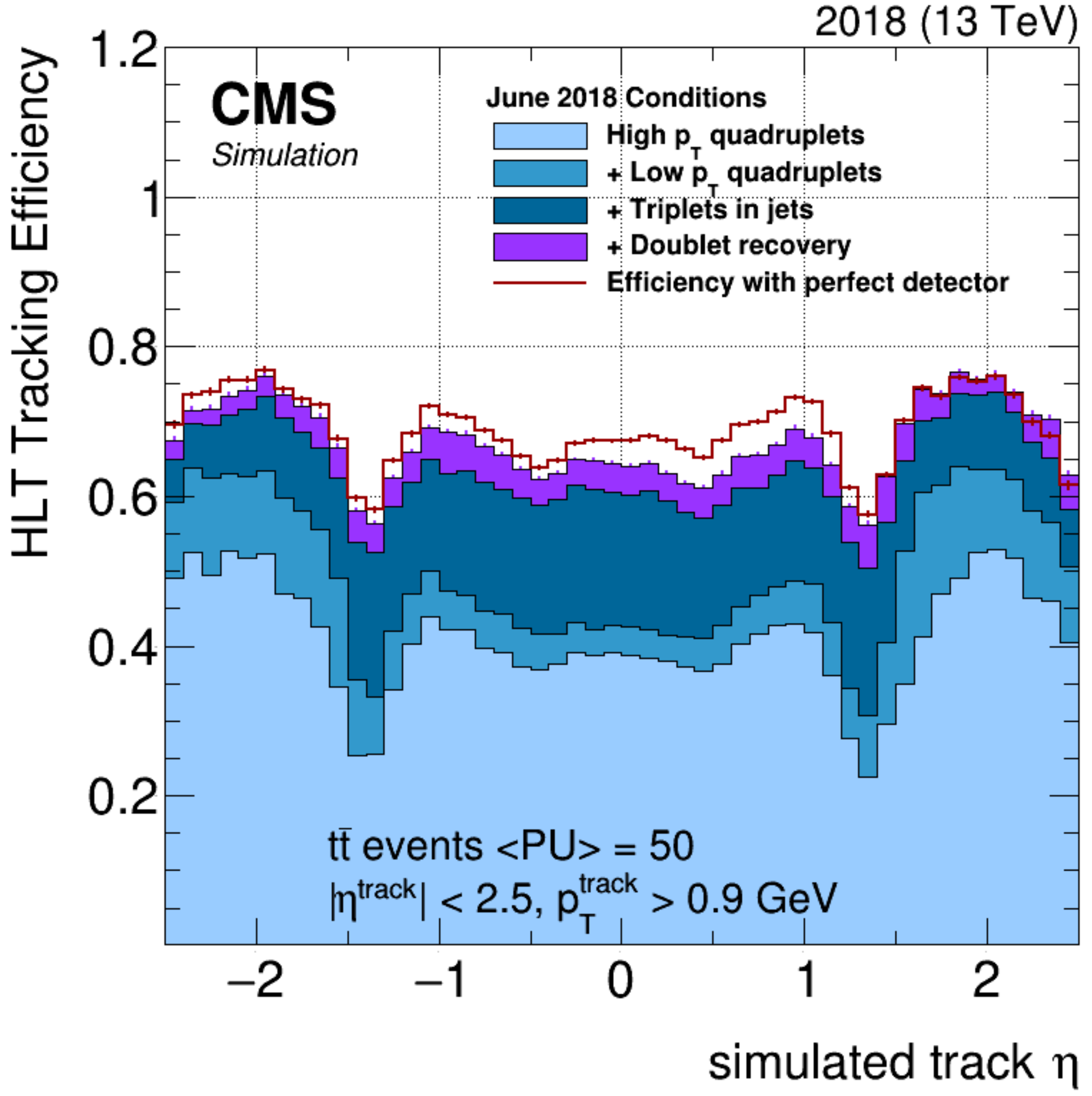


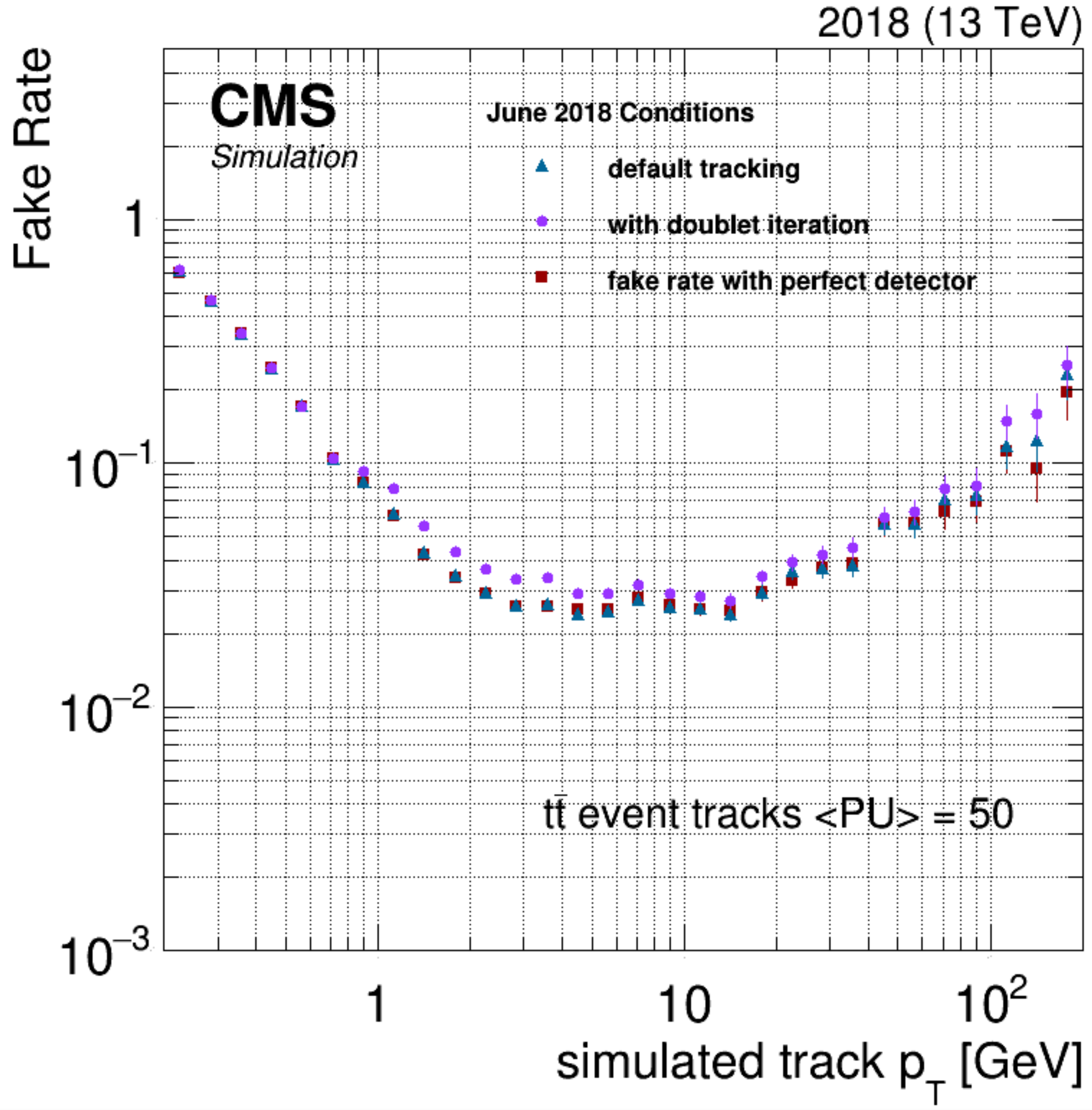


Tracking at HLT

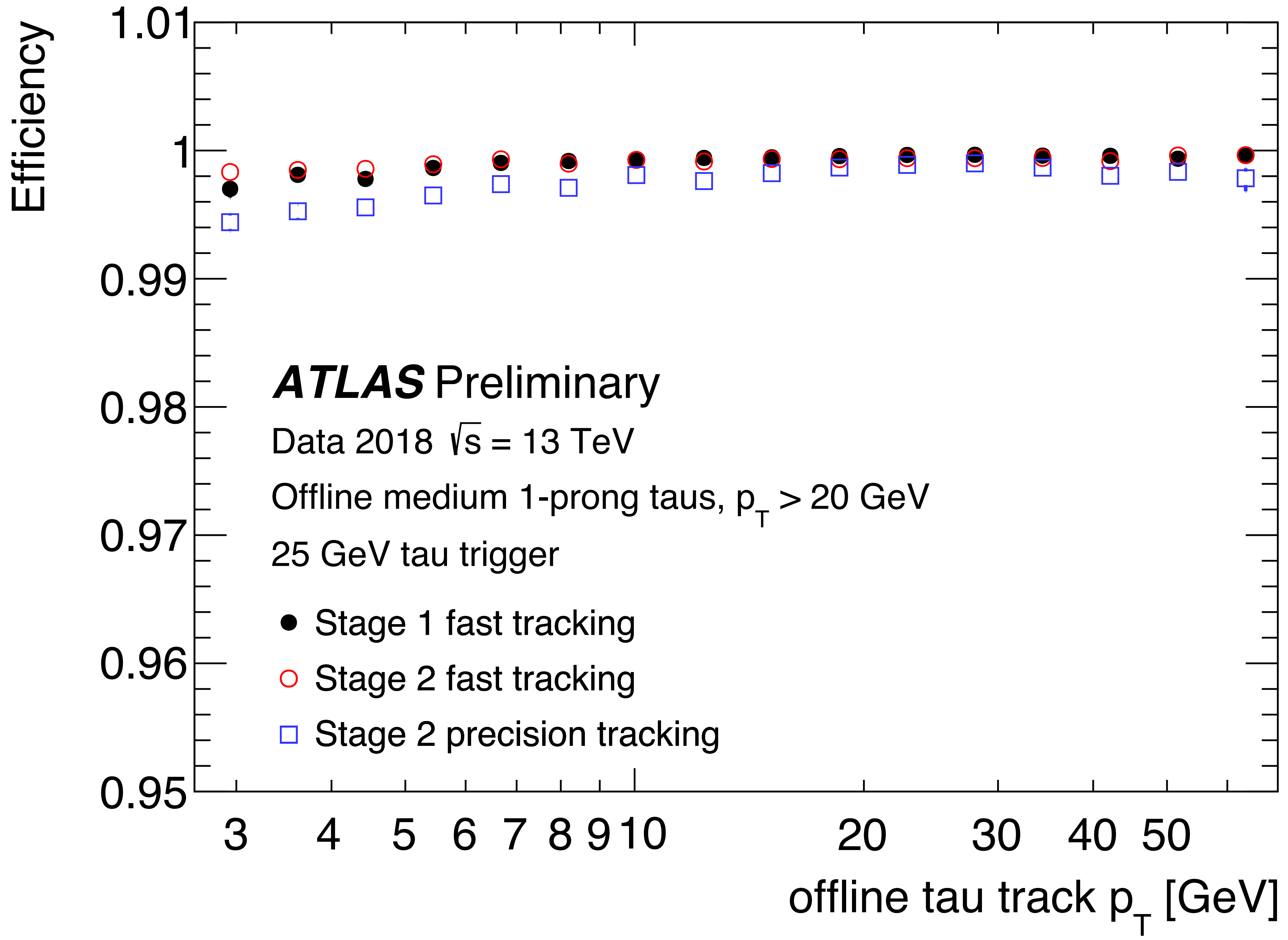
- Tight trigger reduction rate required by the experiments
- A software trigger system requires a trade-off between the complexity of the algorithms, the sustainable output rate, and the selection efficiency

- CMS: the tracking efficiency and fake rate are measured in simulated ttbar events
  - Mean number of additional interactions: 50





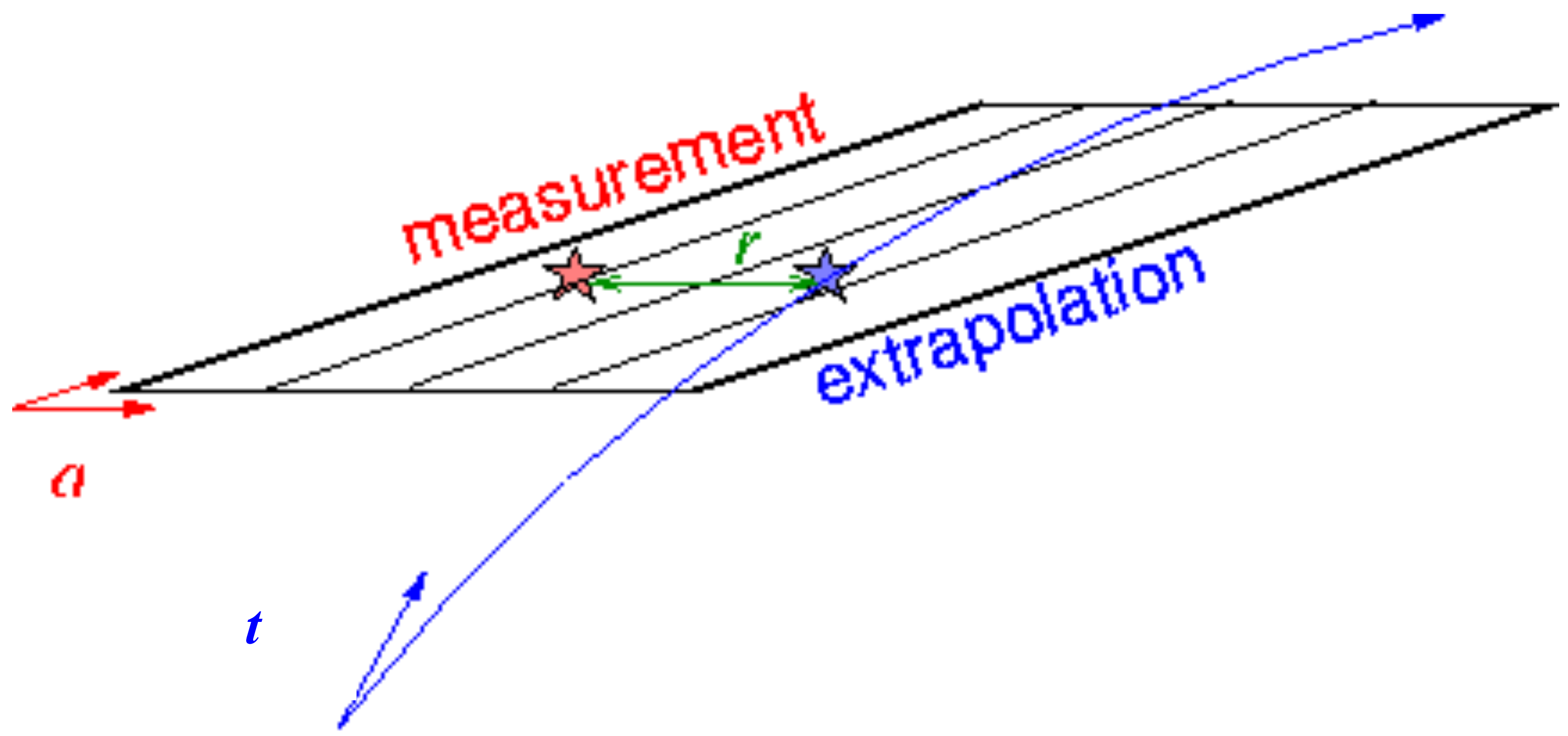
- ATLAS: two-stage fast tracking



Alignment

- Alignment is concerned with determining the actual geometry of the tracking system and following its eventual changes in time
- ATLAS and CMS use track based detector alignment
  - Track-hit residual minimisation
  - Global$\chi2$
  - MILLIPEDE and HipPy

$$\chi^2 = \sum_{\forall t} \left[\boldsymbol{r}^T(\boldsymbol{t},\boldsymbol{a})\; V^{-1}\; \boldsymbol{r}(\boldsymbol{t},\boldsymbol{a}) \right]$$



  - Minimization → solve linear system with many degrees of freedom

$$\frac{d\chi^2}{d\boldsymbol{a}} = 0 \quad \rightarrow \quad \overset{\text{N}_{\text{DoF}}\text{ x N}_{\text{DoF}}\text{ matrix}}{\sum_{\forall t} \left[\left(\frac{d\boldsymbol{r}}{d\boldsymbol{a}} \right)^T V^{-1} \left(\frac{d\boldsymbol{r}}{d\boldsymbol{a}} \right) \right]} \enclose{circle}{\delta \boldsymbol{a}} + \overset{\text{N}_{\text{DoF}}\text{ vector}}{\sum_{\forall t} \left(\frac{d\boldsymbol{r}}{d\boldsymbol{a}} \right)^T V^{-1}\, \boldsymbol{r}} = 0$$

Track parameters

$$\boldsymbol{t} = \{(d_0, z_0, \phi_0, \theta, q/p), (\theta_{scat}, \ldots)\}$$

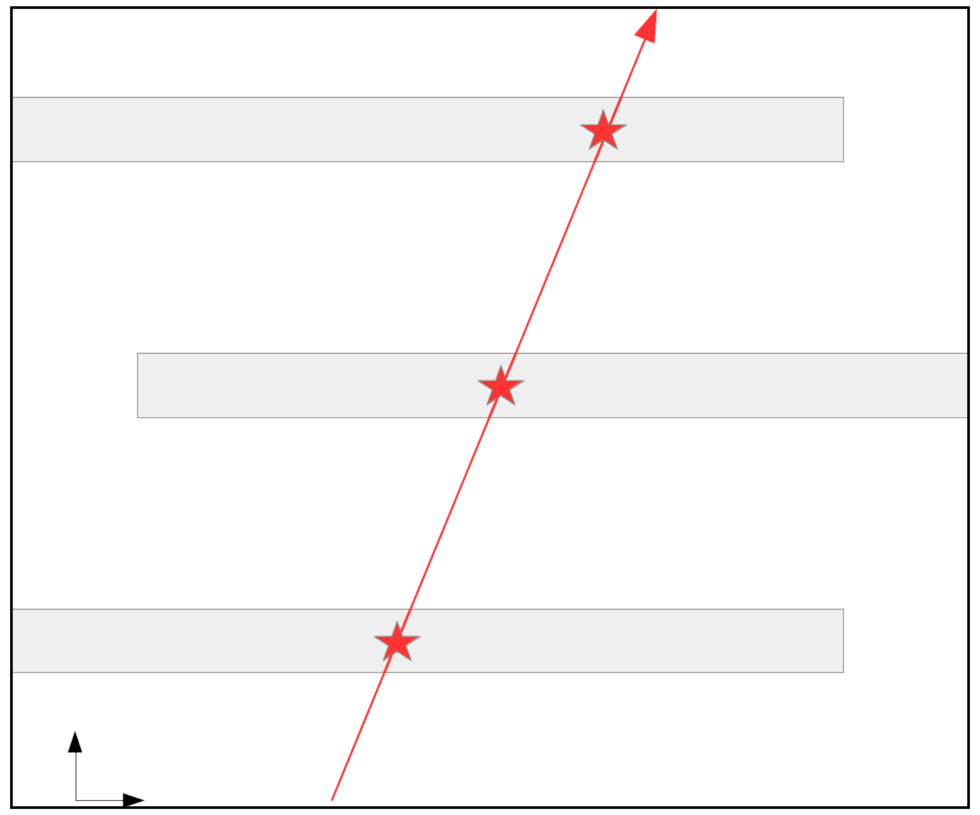
Alignment parameters

$$\boldsymbol{a} = (T_x, T_y, T_z, R_x, R_y, R_z) \times N_{struct}$$

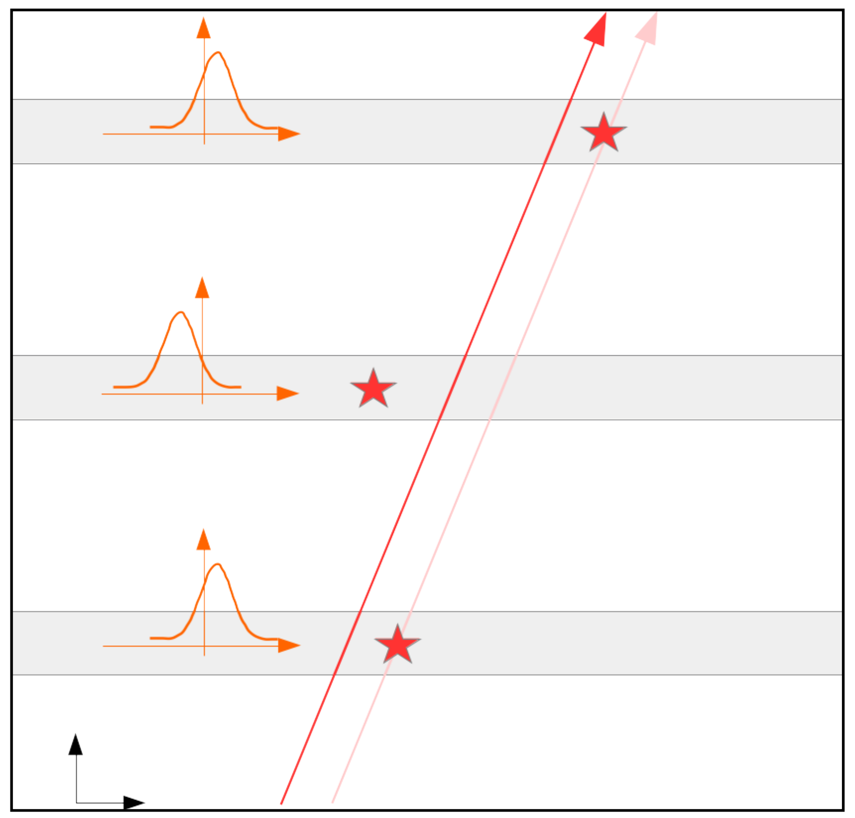
Hit covariance matrix

$$V = \begin{bmatrix} \sigma^2_{hit\ 1} & & \\ & \ddots & \\ & & \sigma^2_{hit\ n} \end{bmatrix}$$

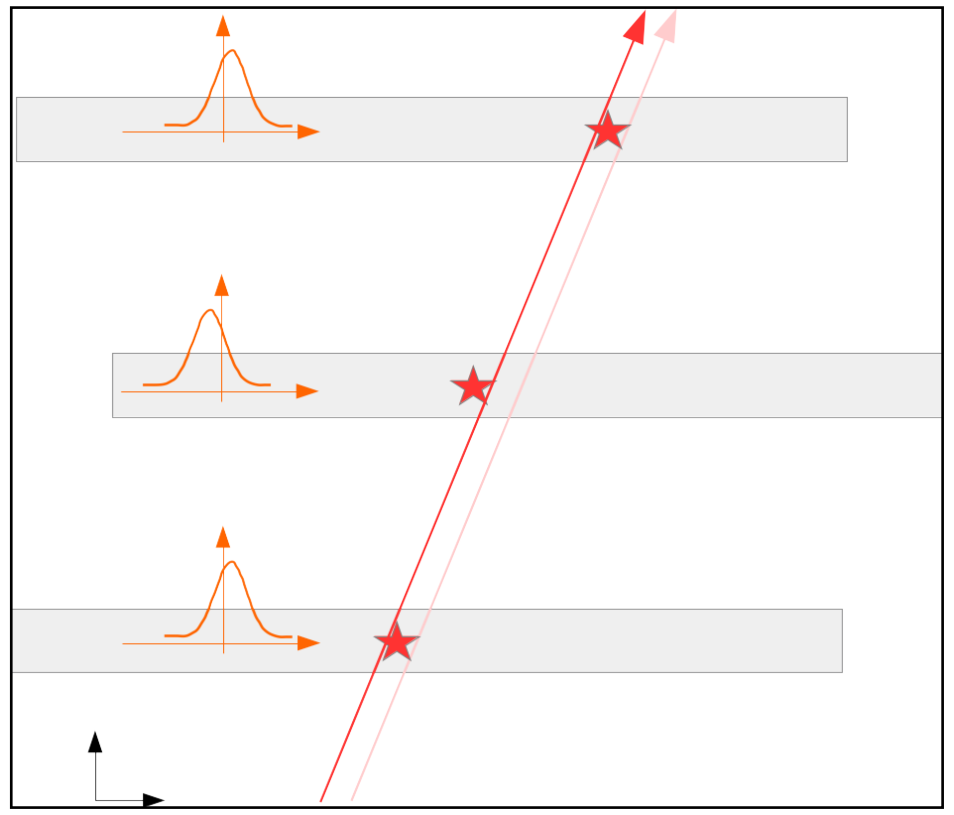
Real



Before alignment

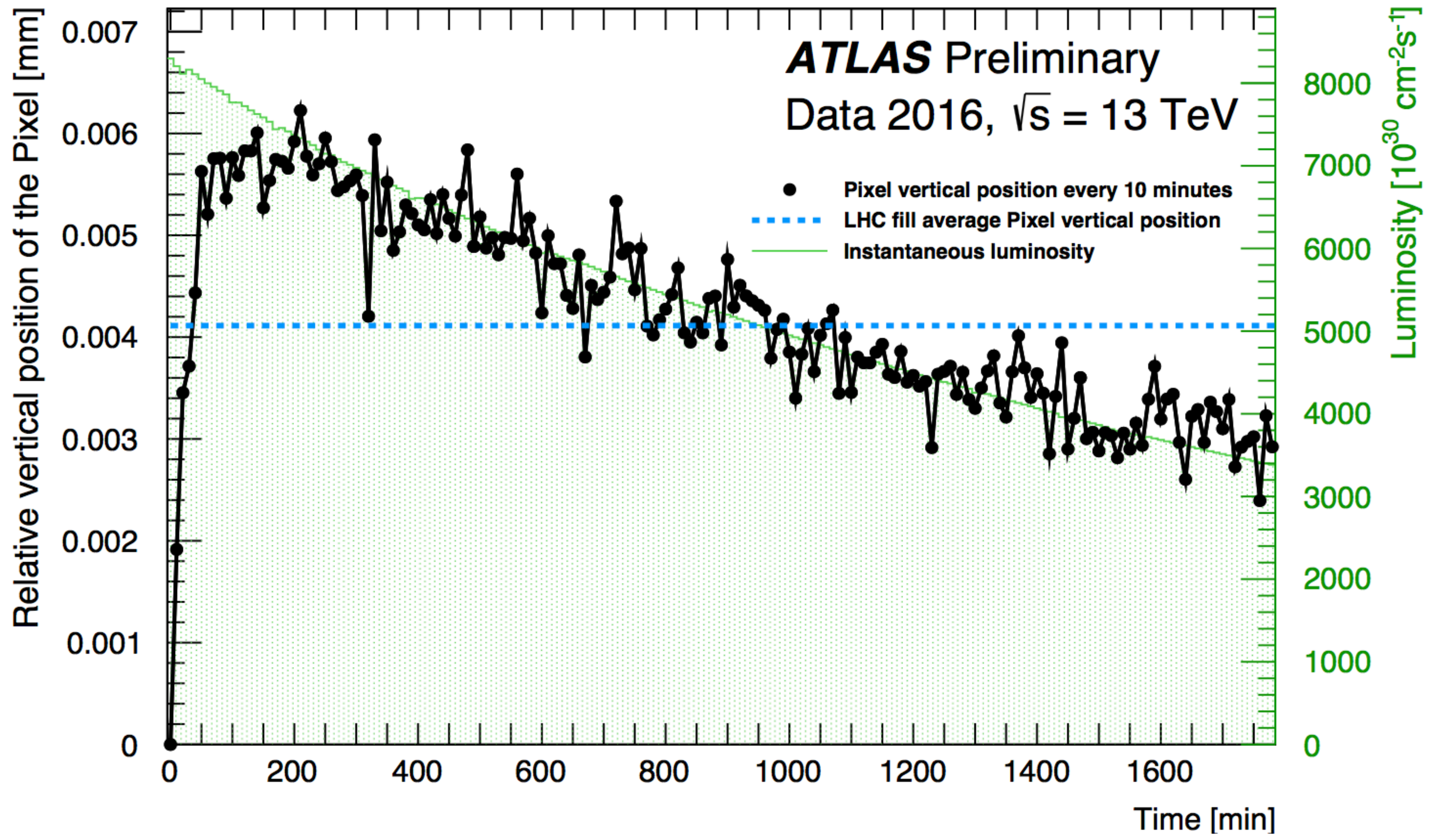


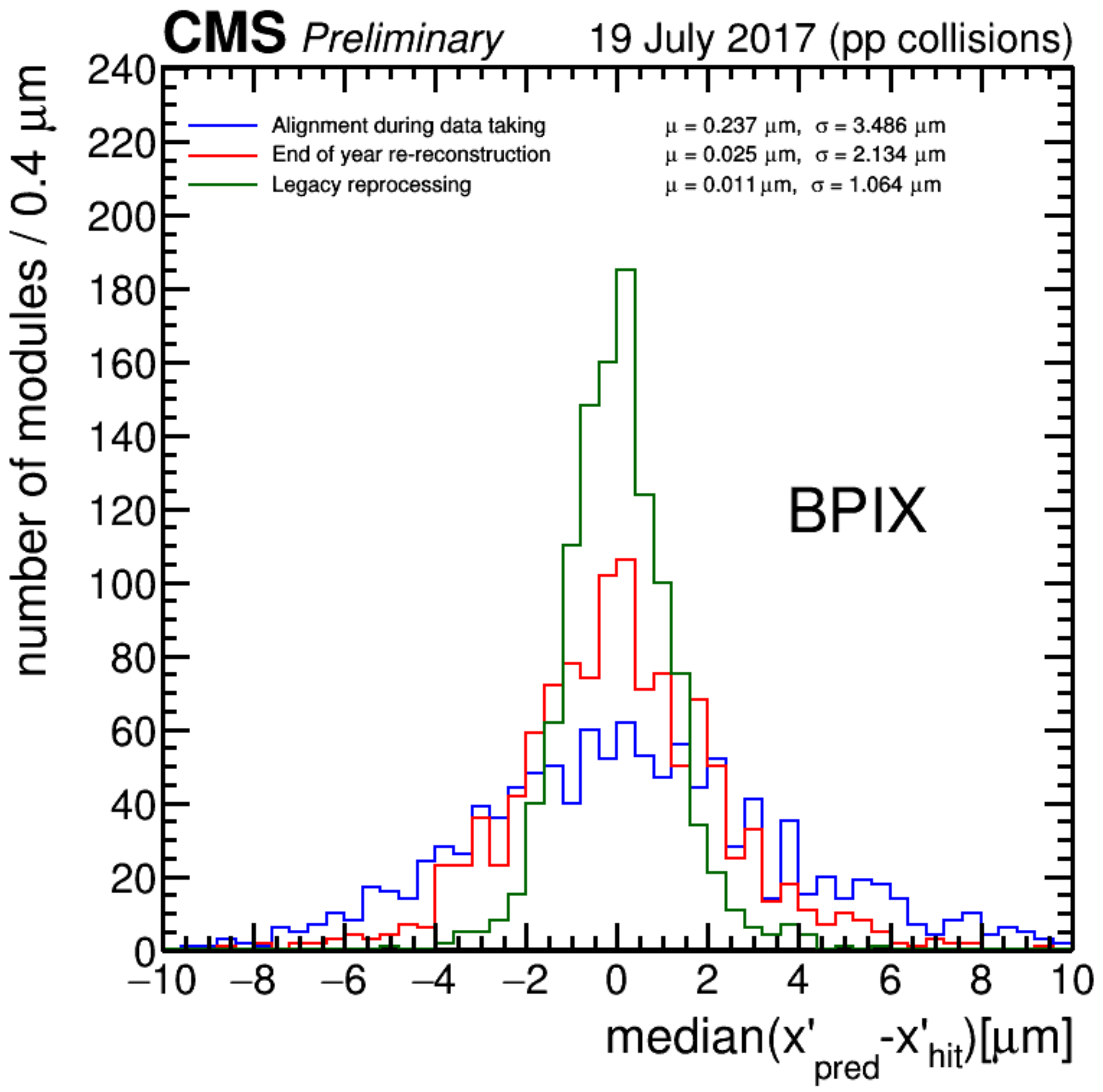
After alignment

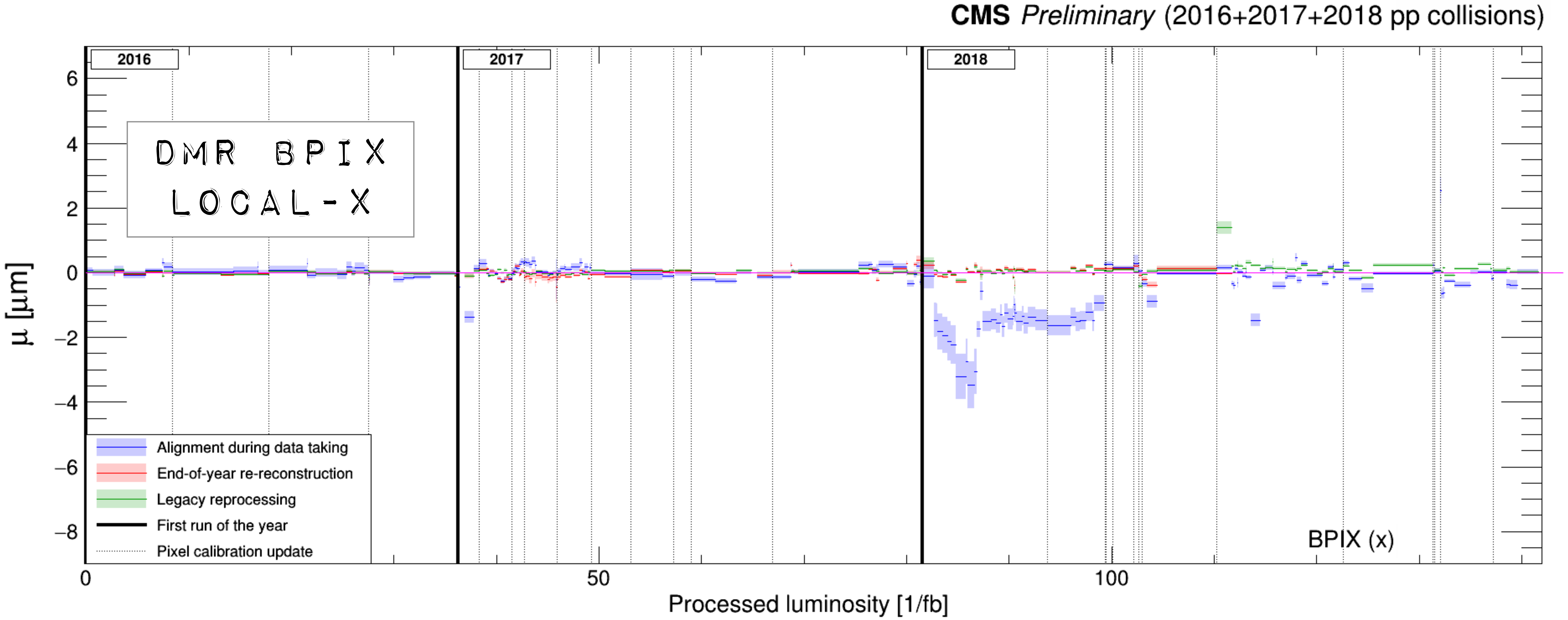


Alignment

- Hierarchical approach:
  - Proceed from large structures to individual modules
- Detector stability and time dependent movements
  - Short time scale movements and long term stability
  - Prompt alignment
  - Whole Run2 alignment → Legacy

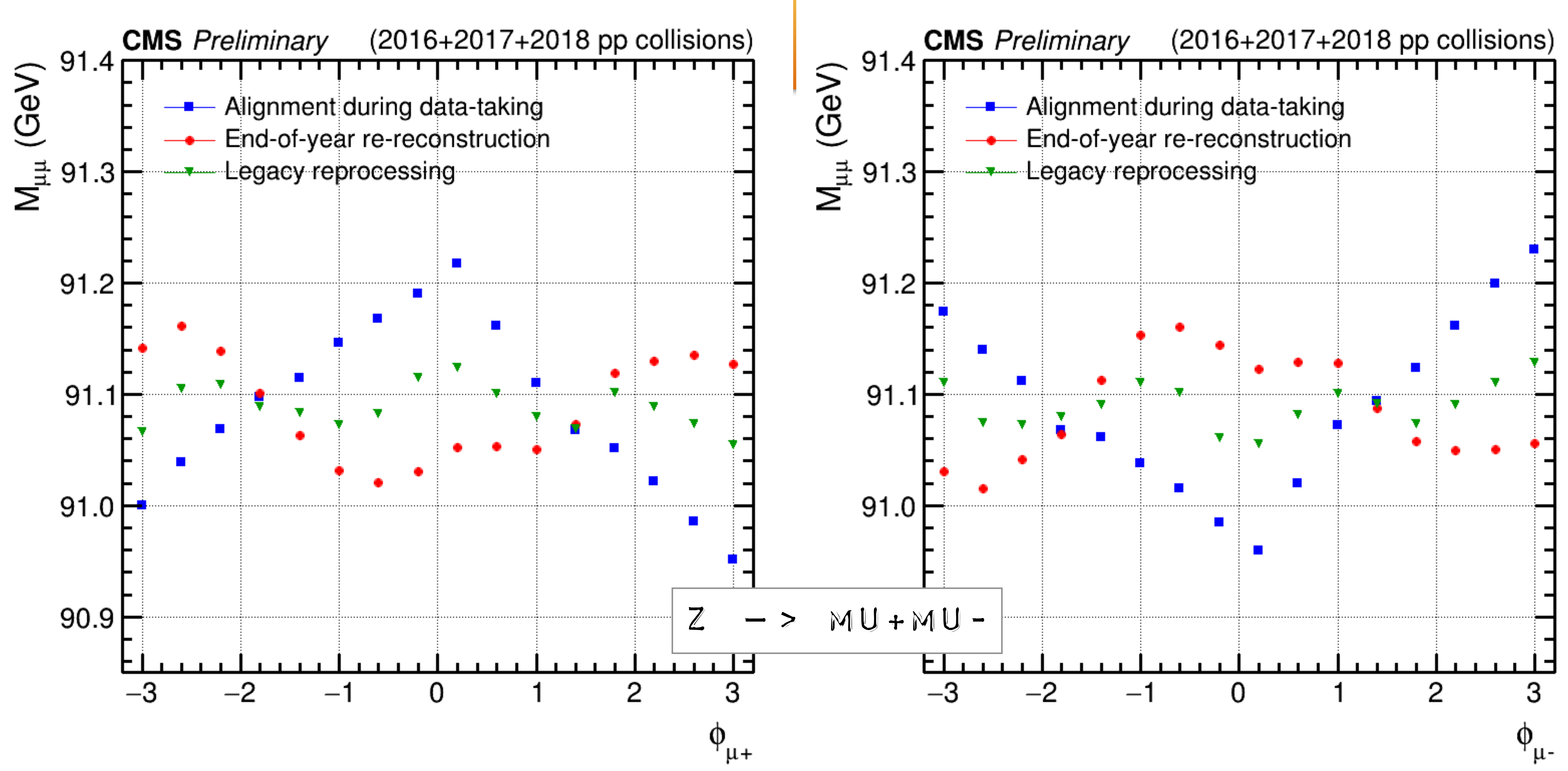




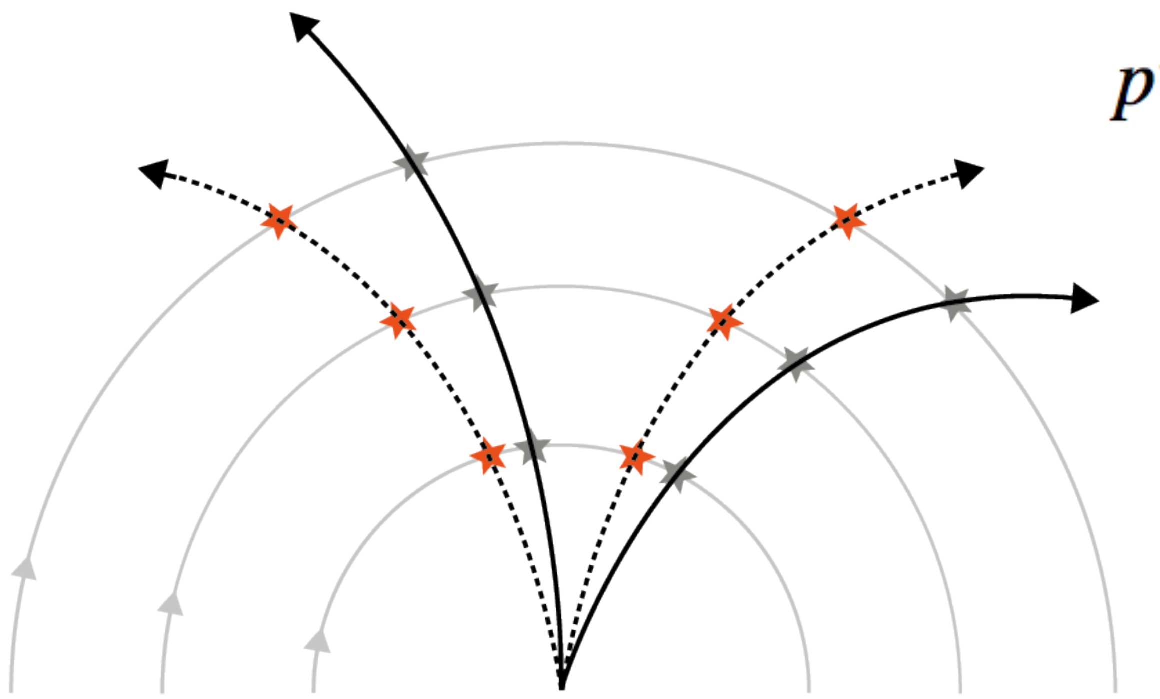


Alignment

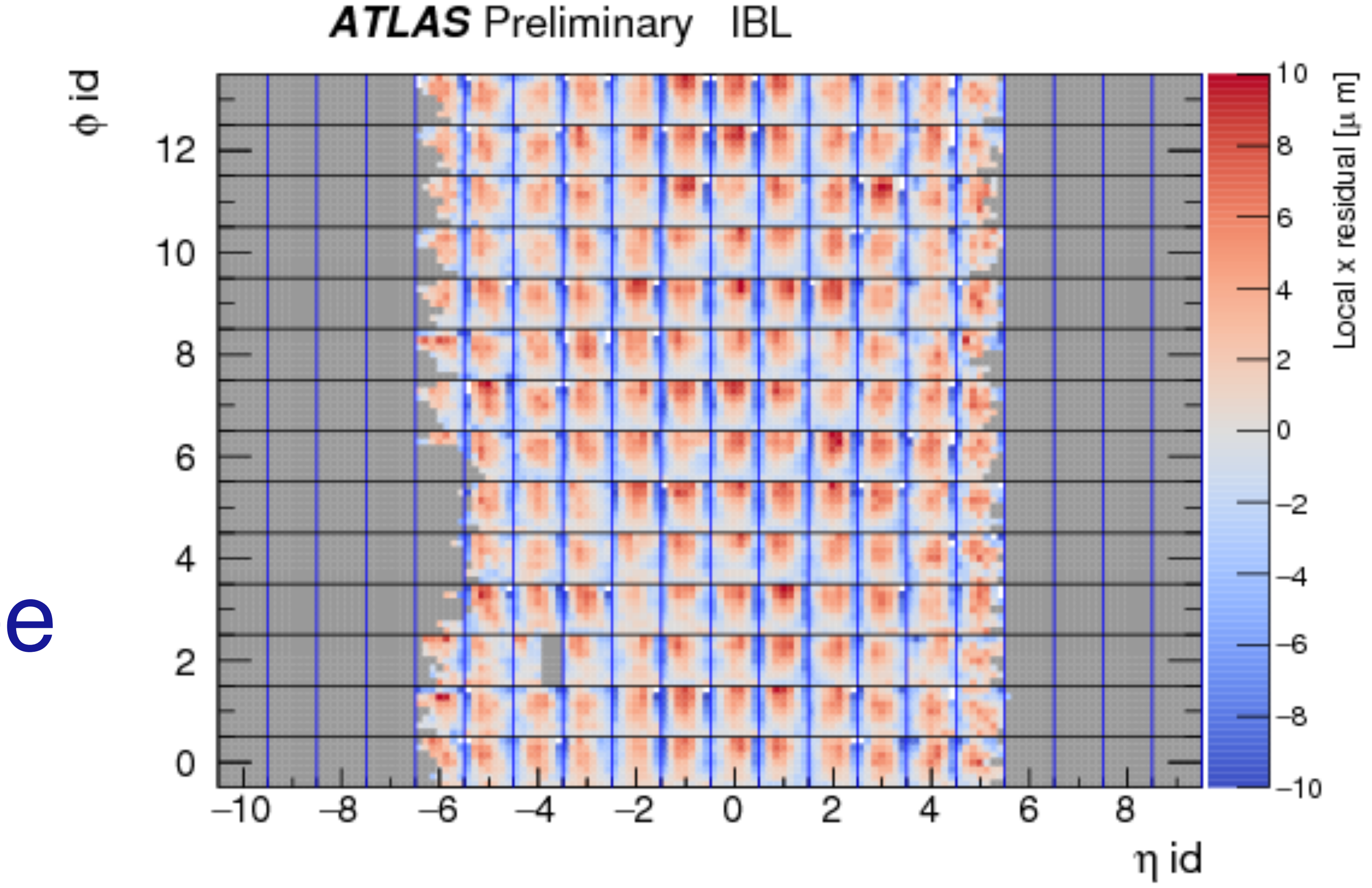
- Weak modes
  - Track based alignment has low sensitivity to misalignments that may leave the track χ2 almost invariant
  - This may introduced biases in track parameters → Sagitta, scale or impact parameter biases
  - dedicated alignment campaigns
  - Validate with Z and J/ψ → μ+μ−, E/p
  - Example: sagitta biases

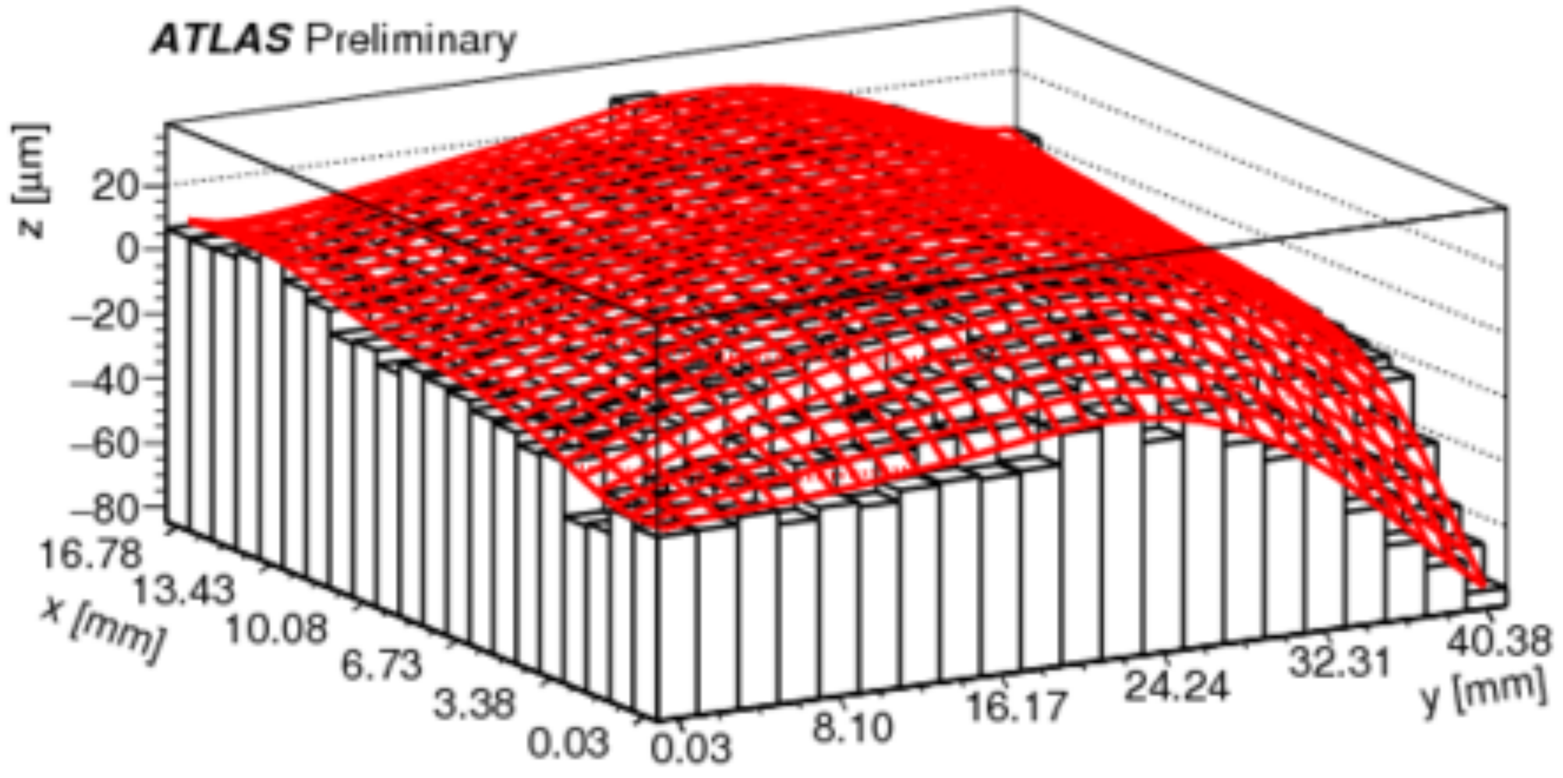


Z –> MU+MU-



$$p' = p\,(1 + q\,p_{\mathrm{T}}\,\delta_{\mathrm{sagitta}})^{-1}$$

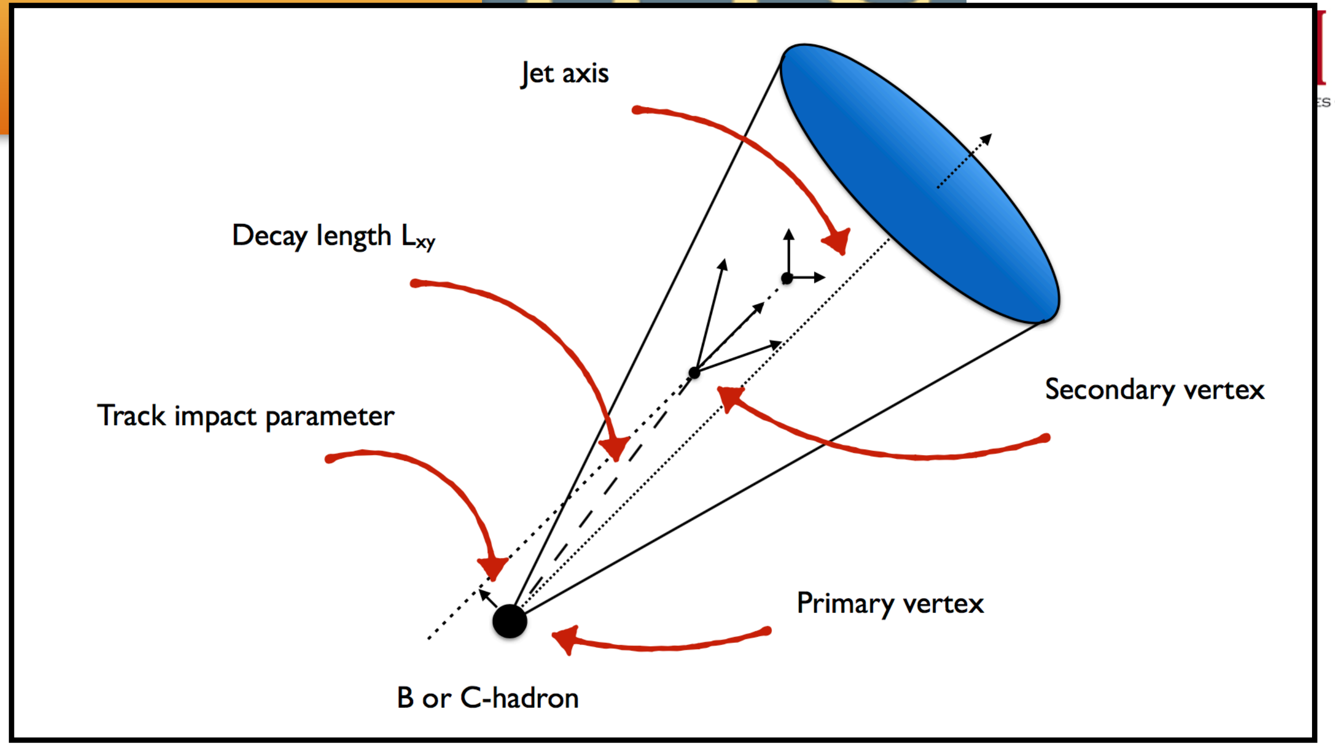




- Detailed evaluation of sensors shape
  - e.g. ATLAS IBL modules

Flavour tagging

- Identification of jets containing b- or c- hadrons is crucial for physics analysis (Higgs, top, SUSY…)
- Flavour-tagging relies on a variety of track related observables
  - Impact parameter, secondary vertices, Particle ID (muons), jets, …
  - Use Machine Learning techniques to combine them into a single classifier
    - High Level Taggers
    - Algorithms and techniques are in constant evolution
      - Also the output classifiers



Properties of jets containing b hadrons:

High mass (~5 GeV)

Long lifetime (1.5 ps)

Large $\gamma c\tau$ (few mm)

Larger number of charged particles

Leptonic decays ($b \rightarrow \mu X$ 20%)

Inputs	High level variables		High + low level variables
Network	NN based	DNN based	
		Dense	CNN→RNN→Dense
Output	b, other	b, c, l	b, bb, lepb, c, l, g
Tagger	CSV	DeepCSV	DeepJet

[Courtesy of Xavier Coubez] 2015 2016 2017 2018 2019

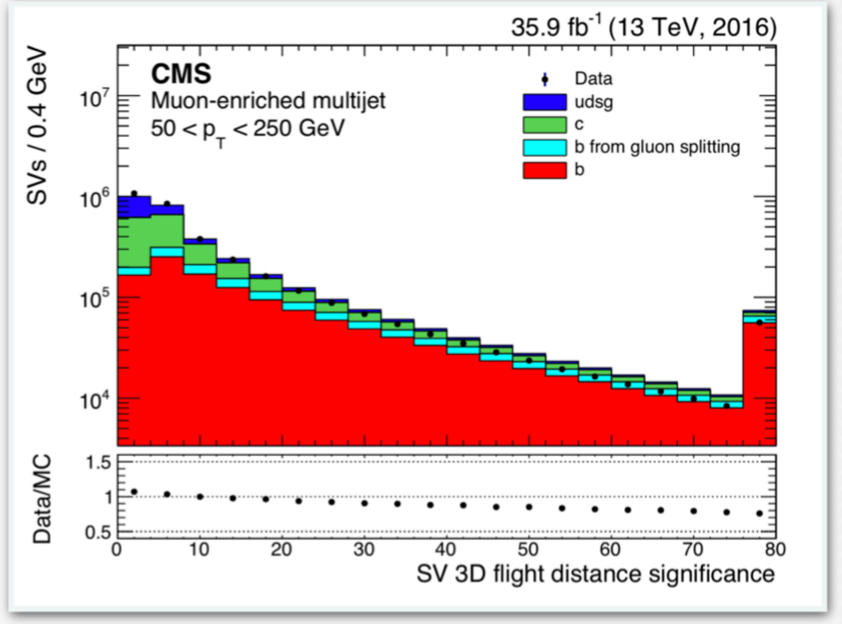
[Disclaimer: this is an example of CMS flavour tagger evolution. ATLAS flavour tagging also evolved with time]

Flavour tagging

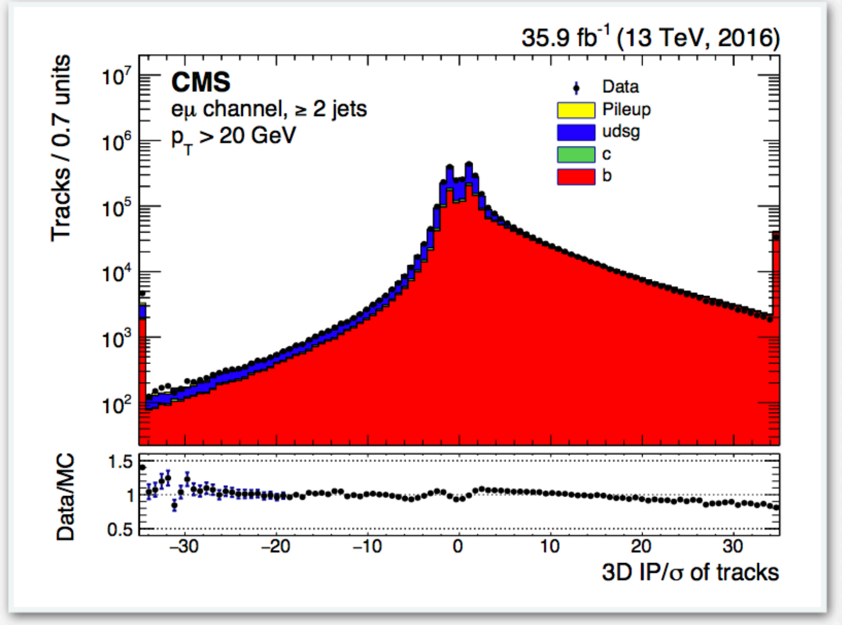
LINK TO DOCUMENTATION

AK4 taggers

[Courtesy of Xavier Coubez]



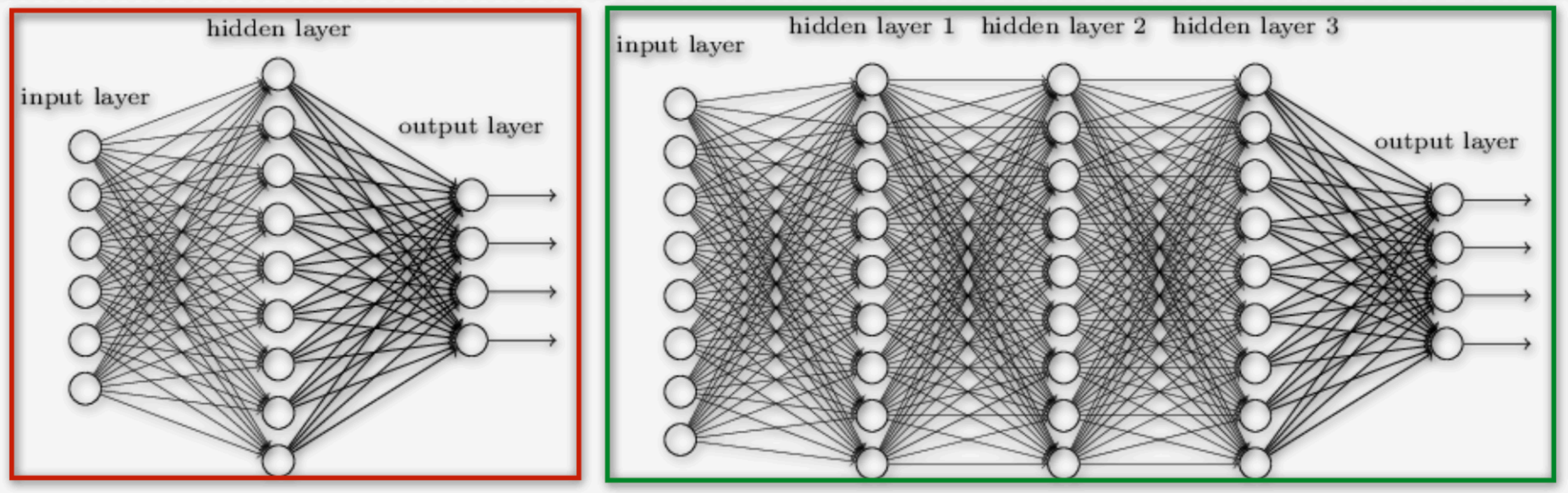
+



+…

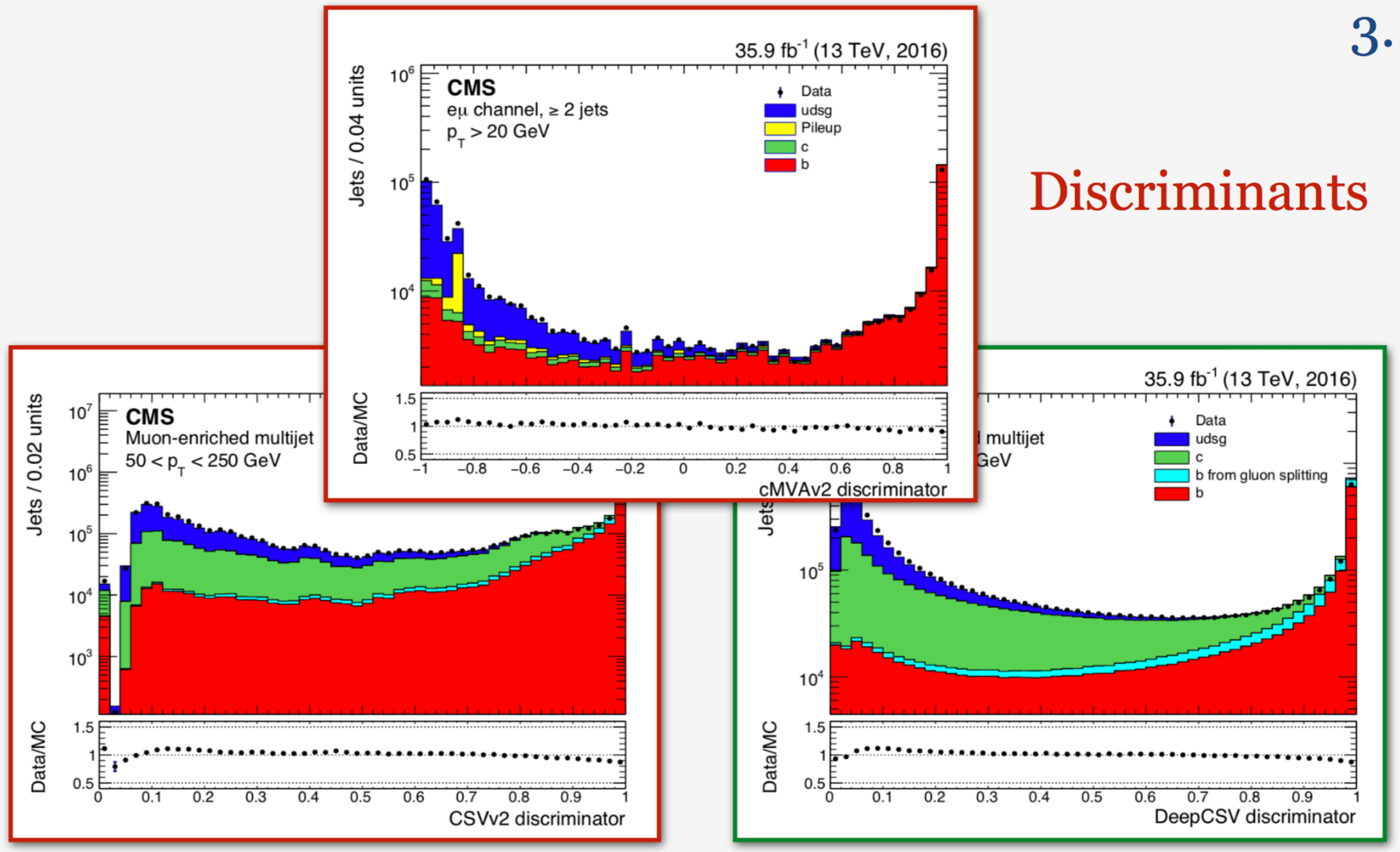
Input variables

1.



Neural network architectures

2.

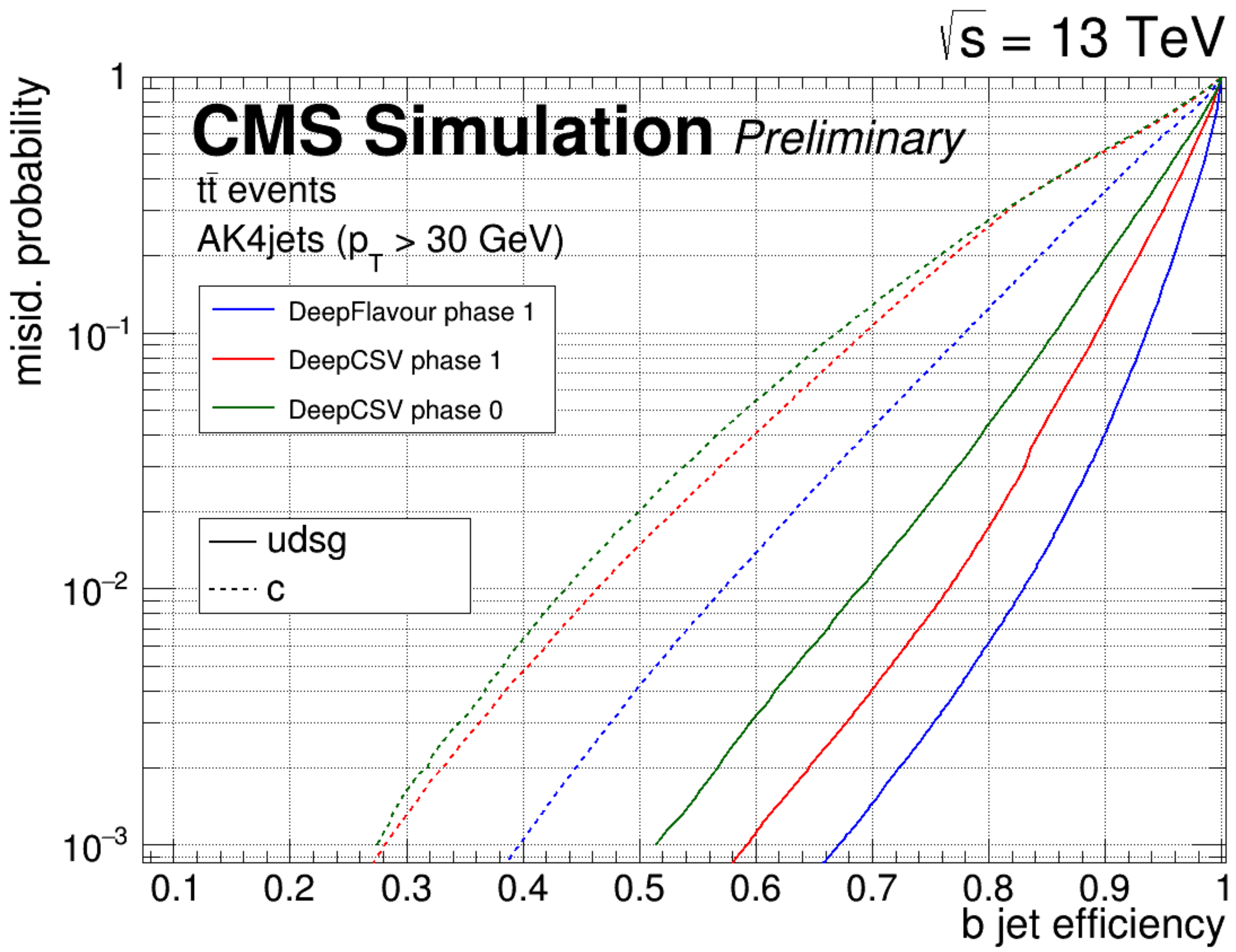


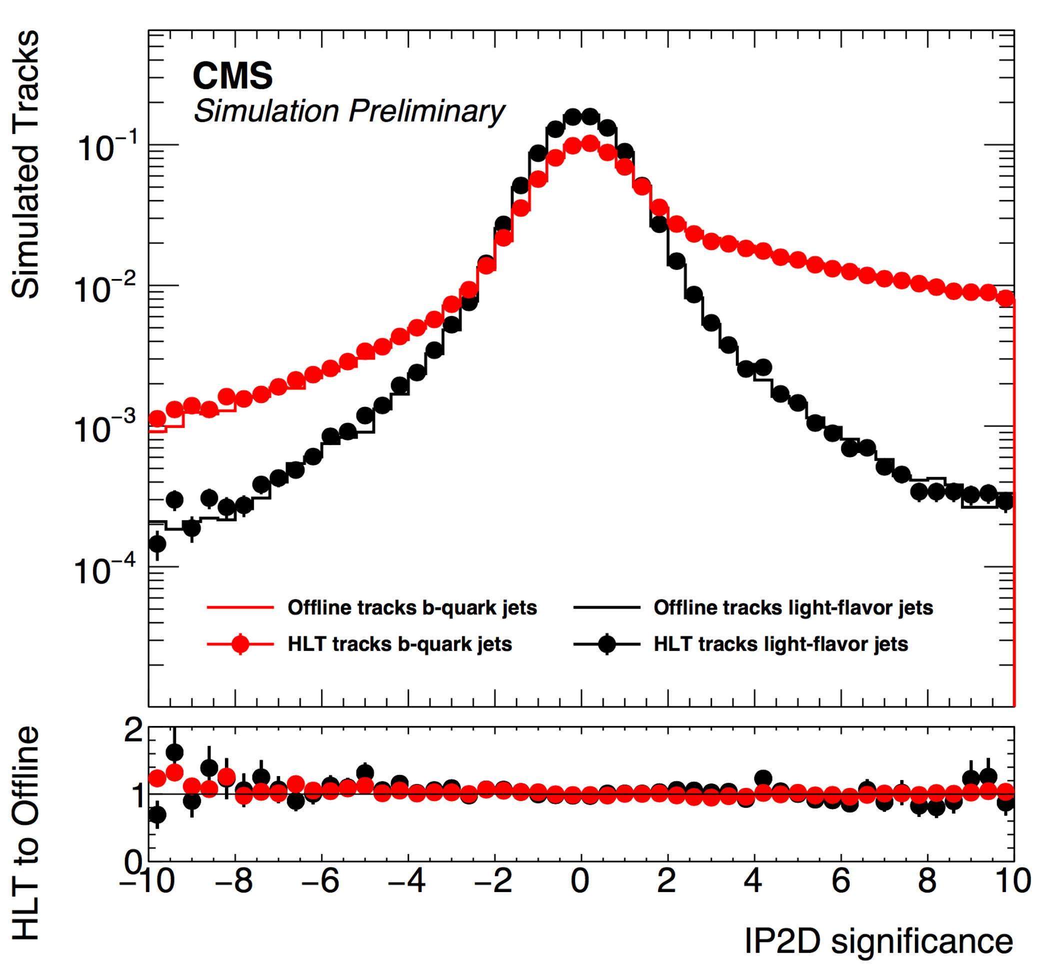
3.

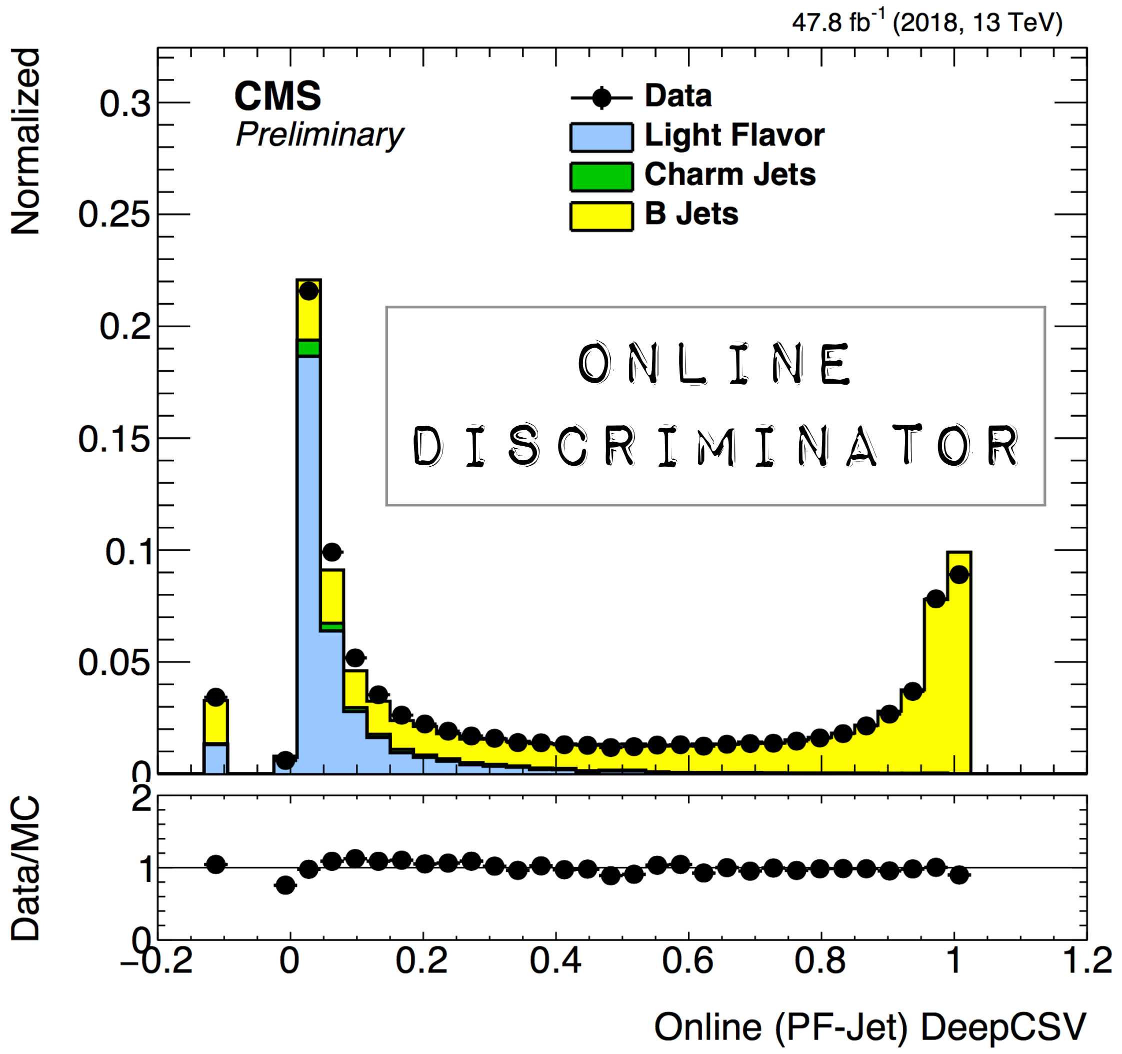
Discriminants

HLT

B-jet trigger performance



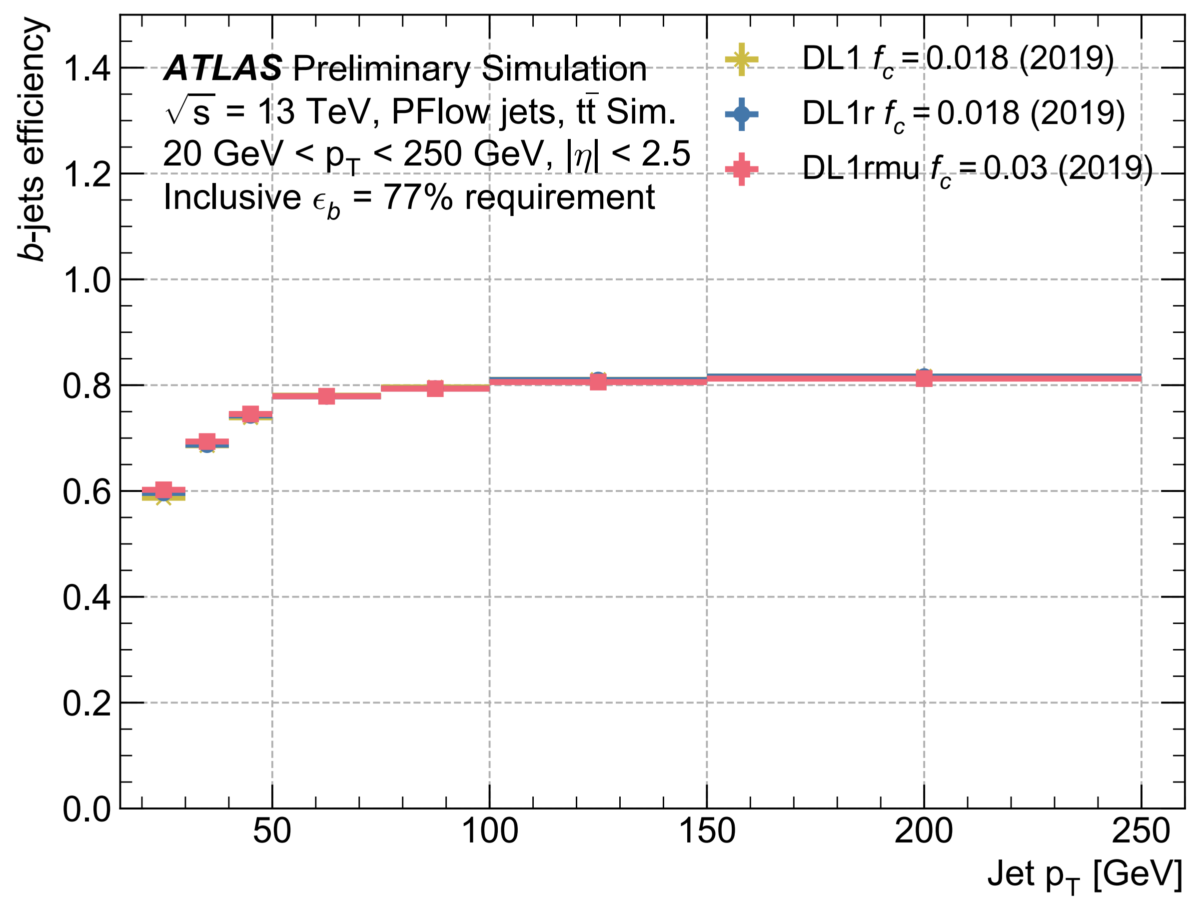


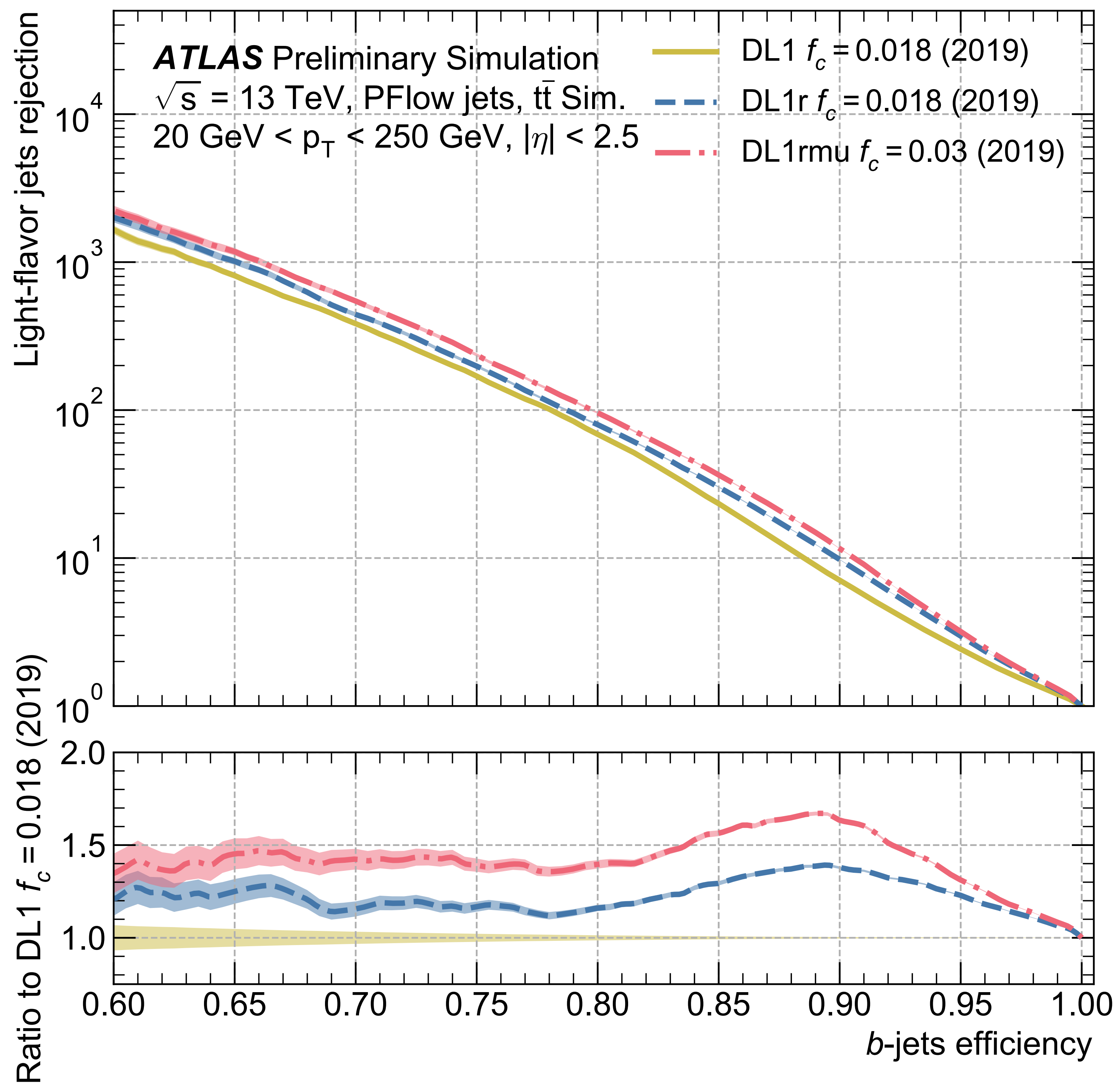


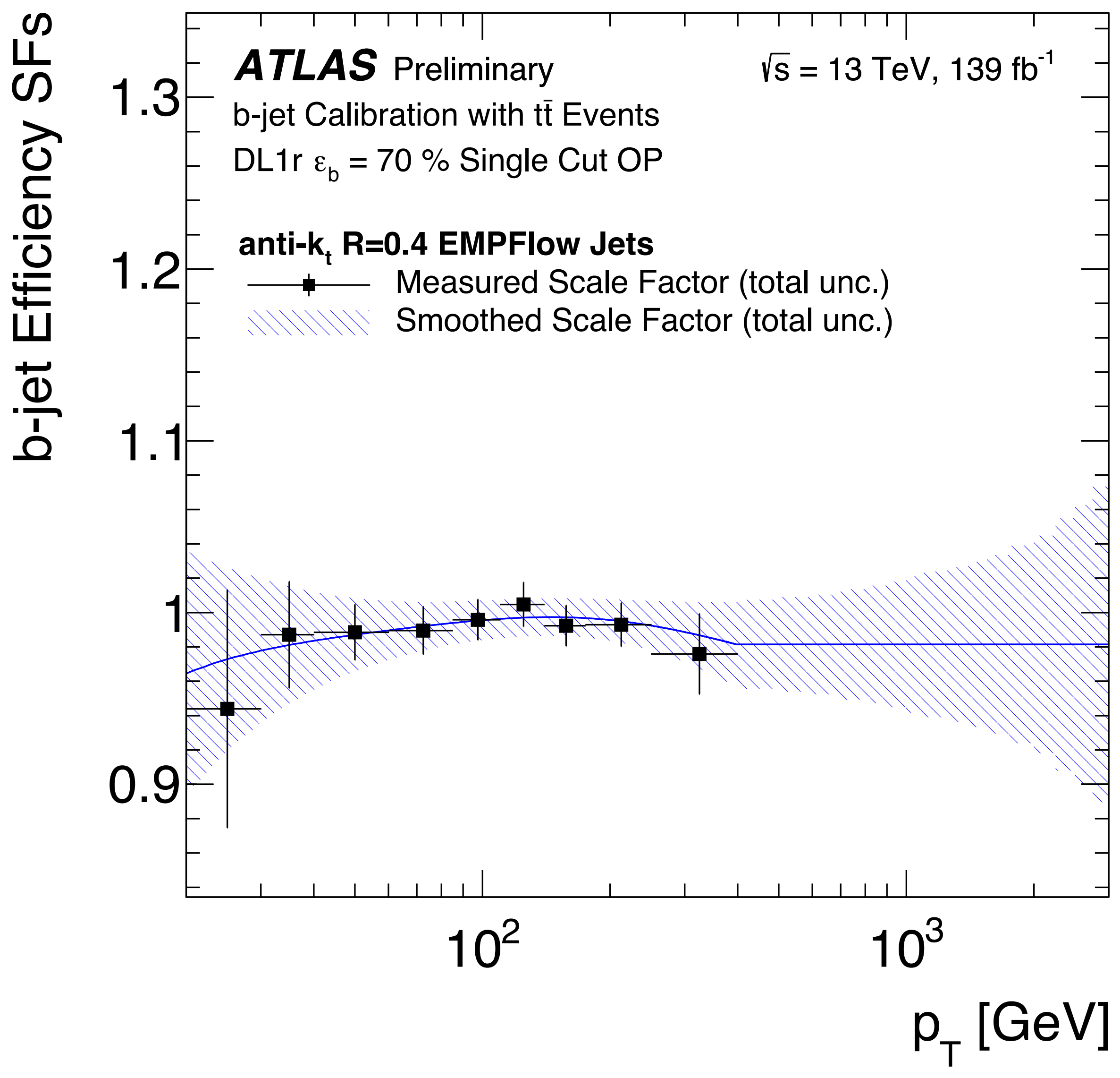
ONLINE DISCRIMINATOR

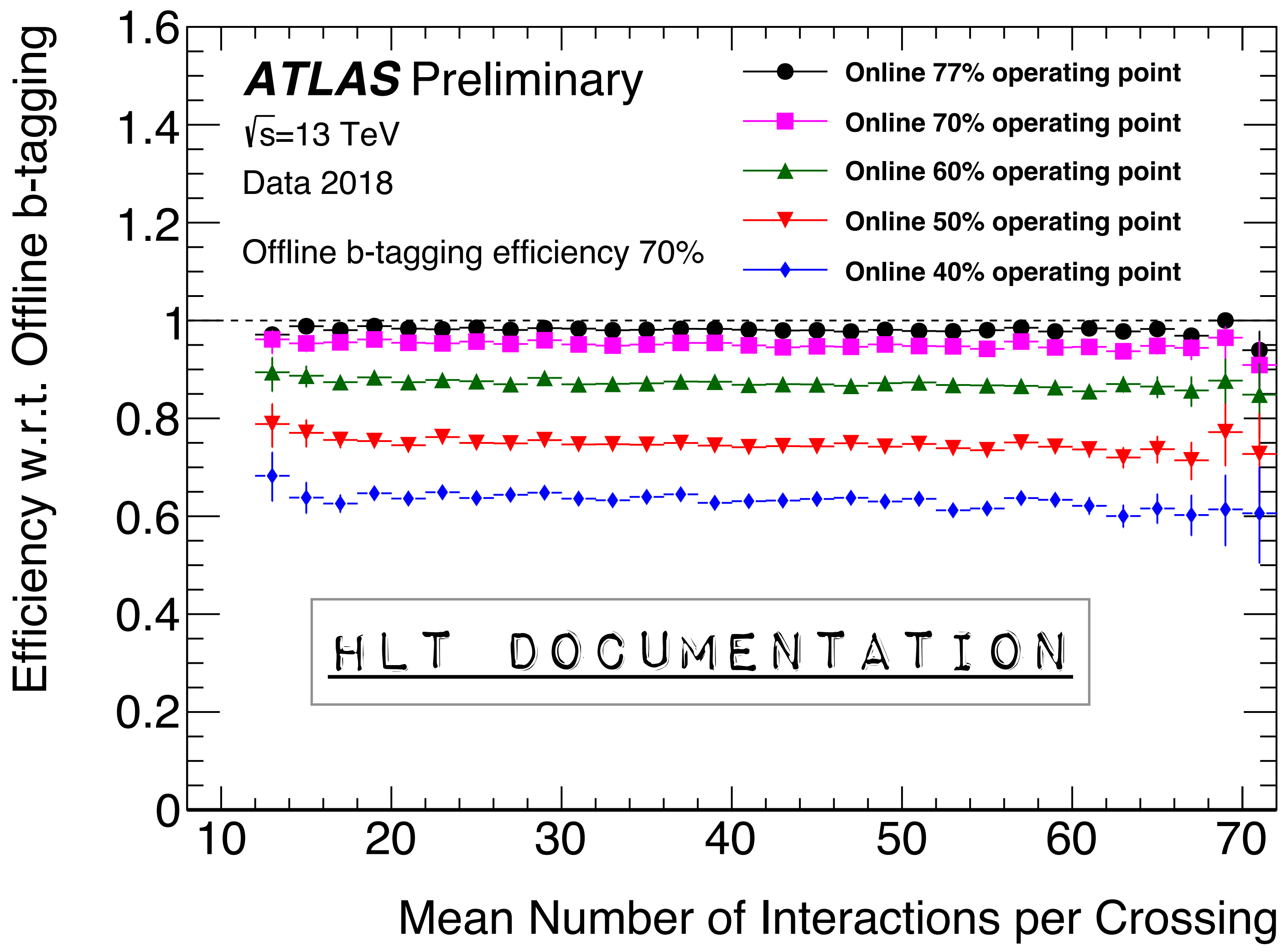
Flavour tagging

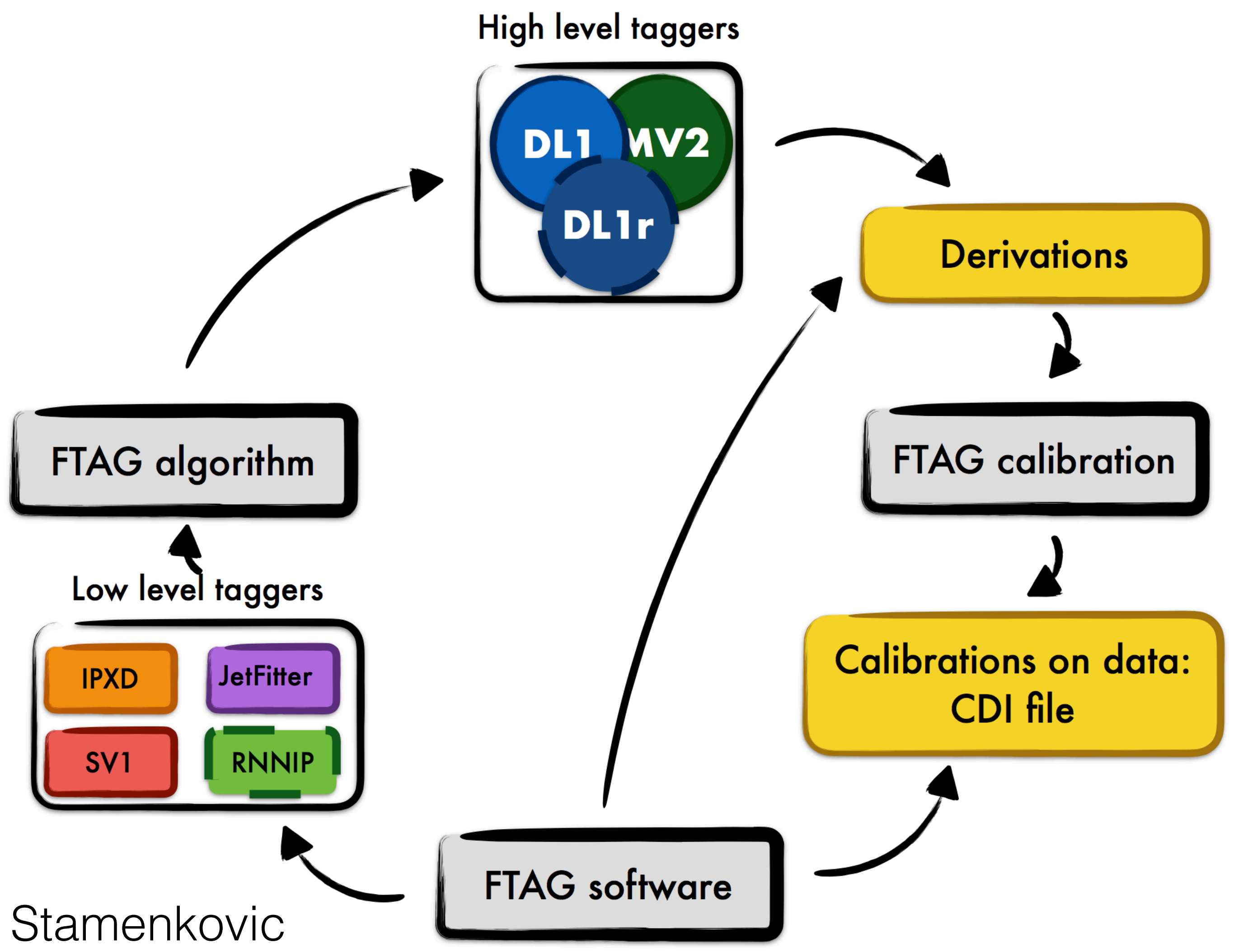
- High level taggers in ATLAS
  - MV2 is a BDT based algorithm
  - DL1 is a ML-based algorithm
  - DL1r evolution using RNN
  - Recent developments
    - achieving factors of ~1000 in light jet rejection and good performance across a much broader p_T range
  - Working points (e.g. ε_b = 77%)
  - Documentation:
    - Performance with 2019 calibration
    - Tagging efficiency, mis-tag rates…









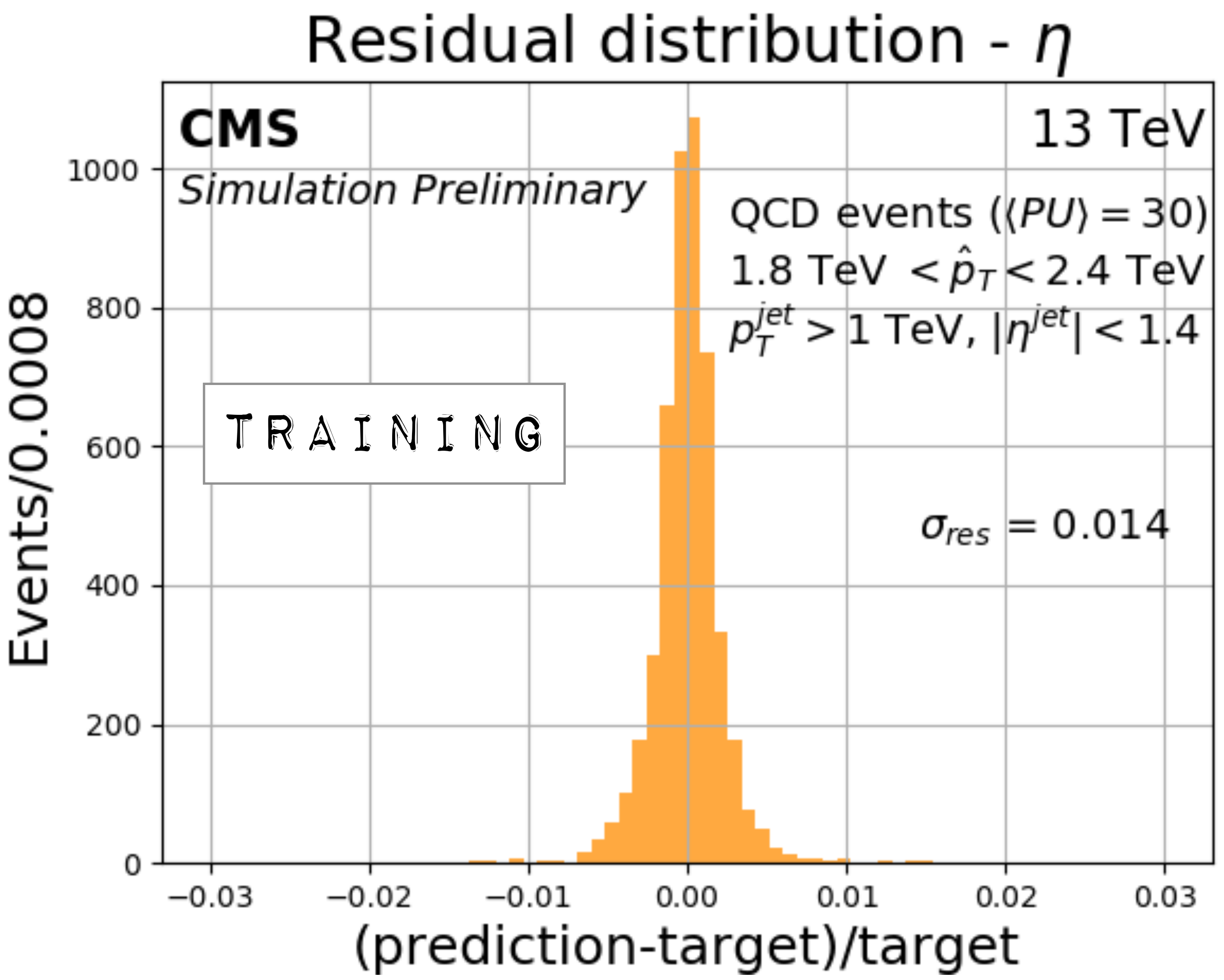


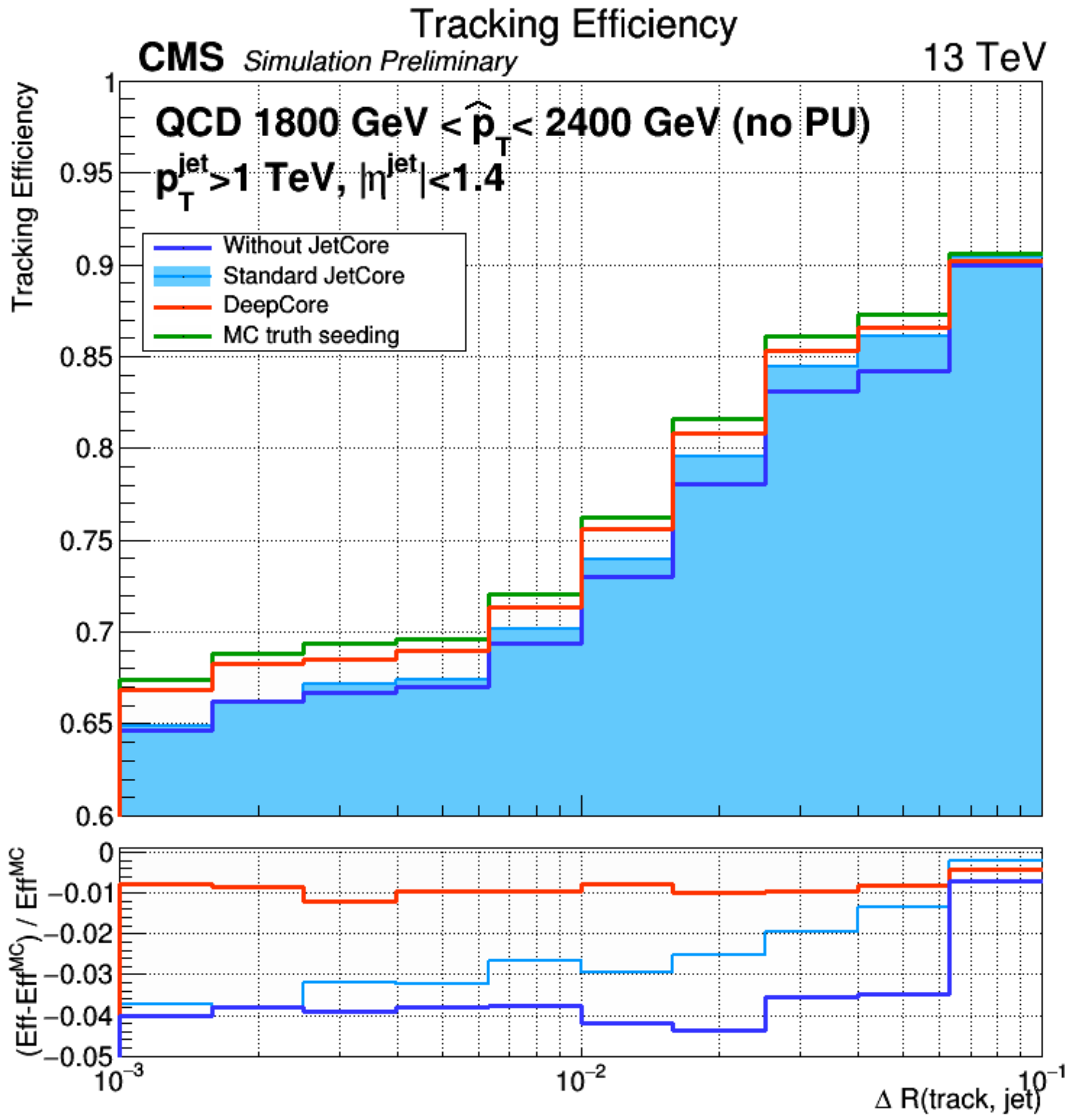
Marko Stamenkovic

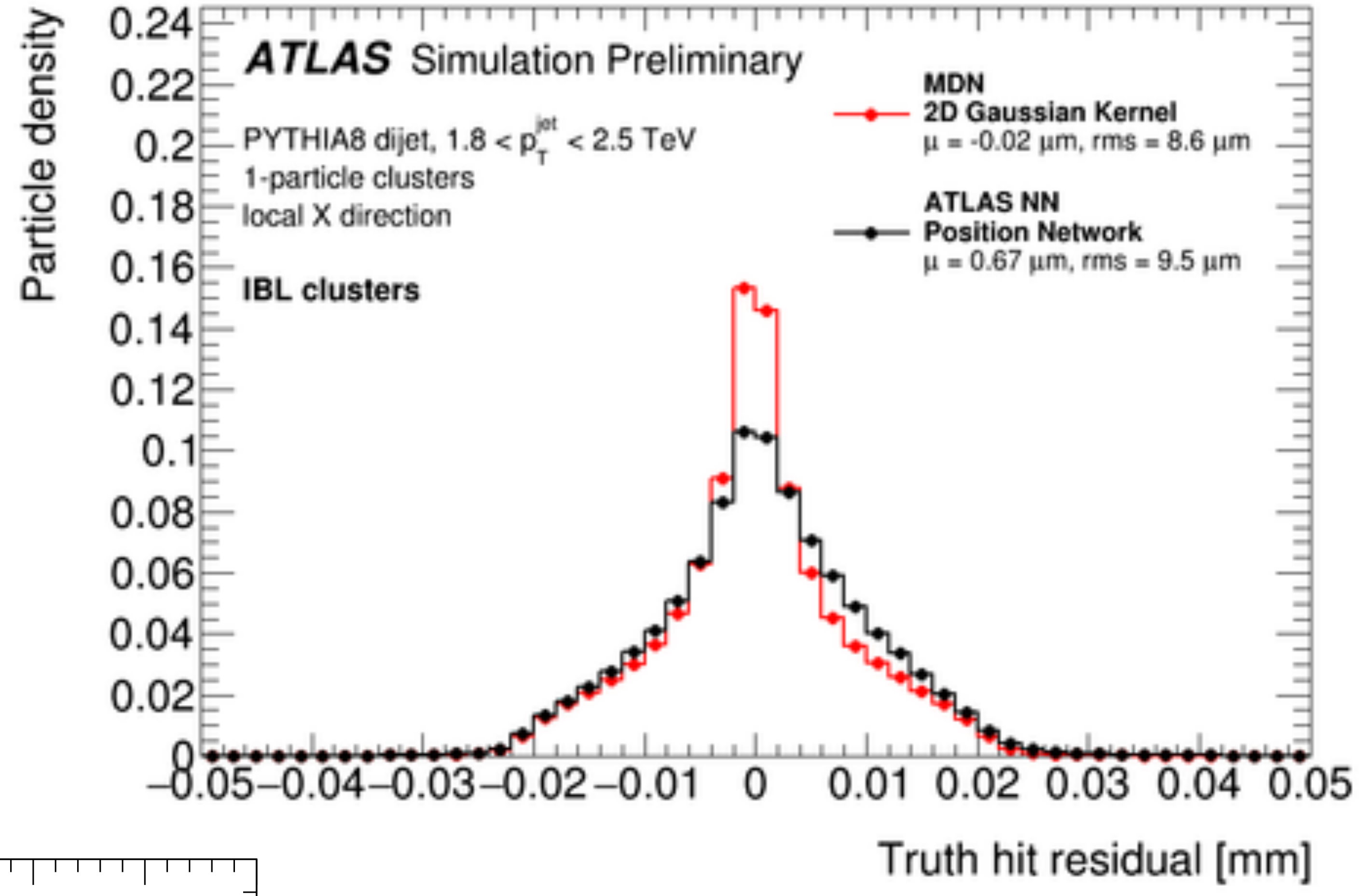
LHC Run 3 preparations: challenging pileup

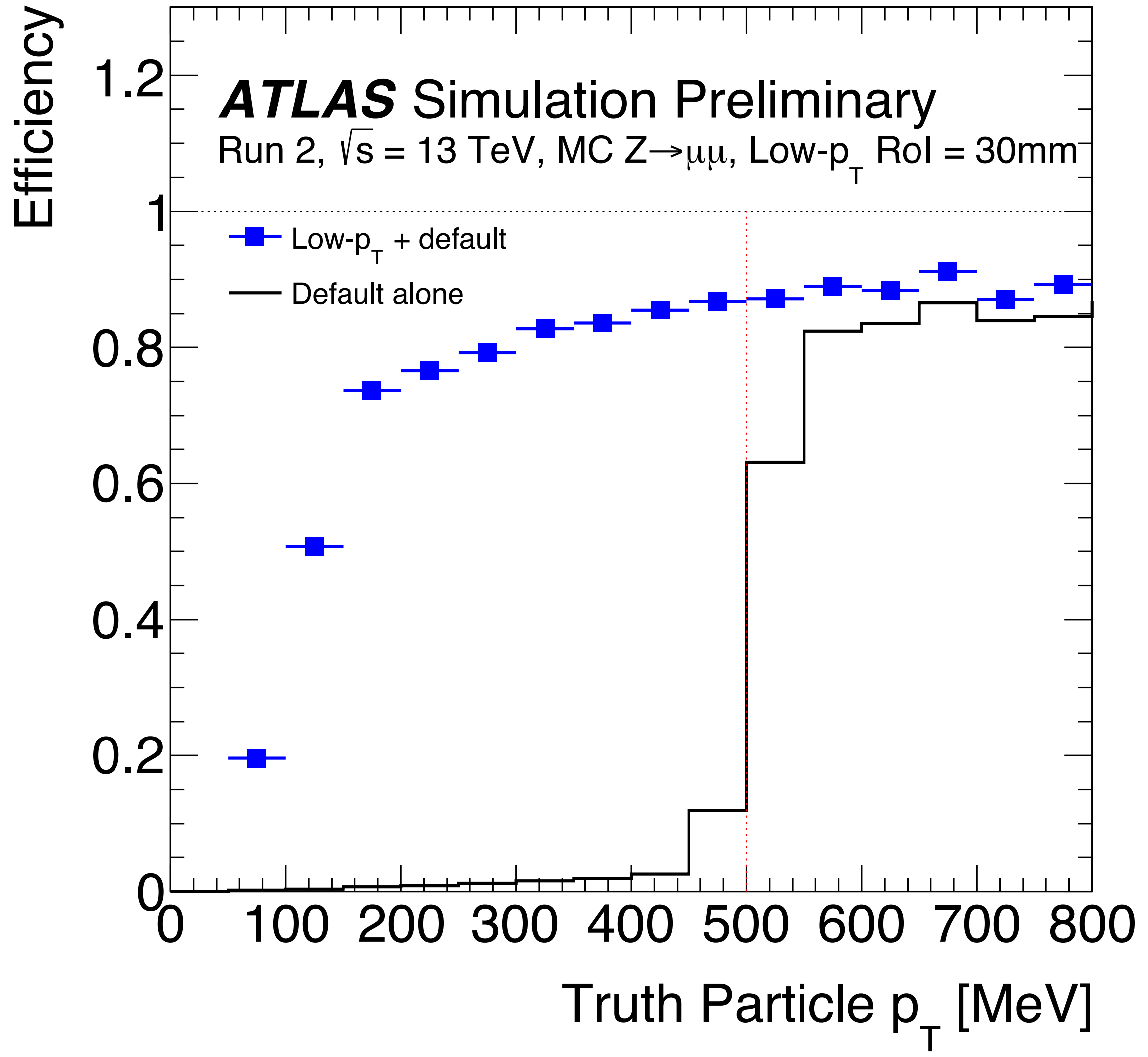
- CMS: new seeding for tracks in high p_T jets
  - Developing a Convolutional NN (CNN) DeepCore
    - Goal: to improve the tracking performance by skipping the pixel clustering
  - Cellular Automaton
    - Produce the track-seeds directly from the raw pixel information of the four layers in the jet core region.
  - Track Classification via DNN

- ATLAS
  - Tuning of the track selection aiming to keep efficiency high and low fake rate
  - Mixture Density Network (MDN) can estimate both hit position and associated uncertainty simultaneously
  - Low p_T tracking



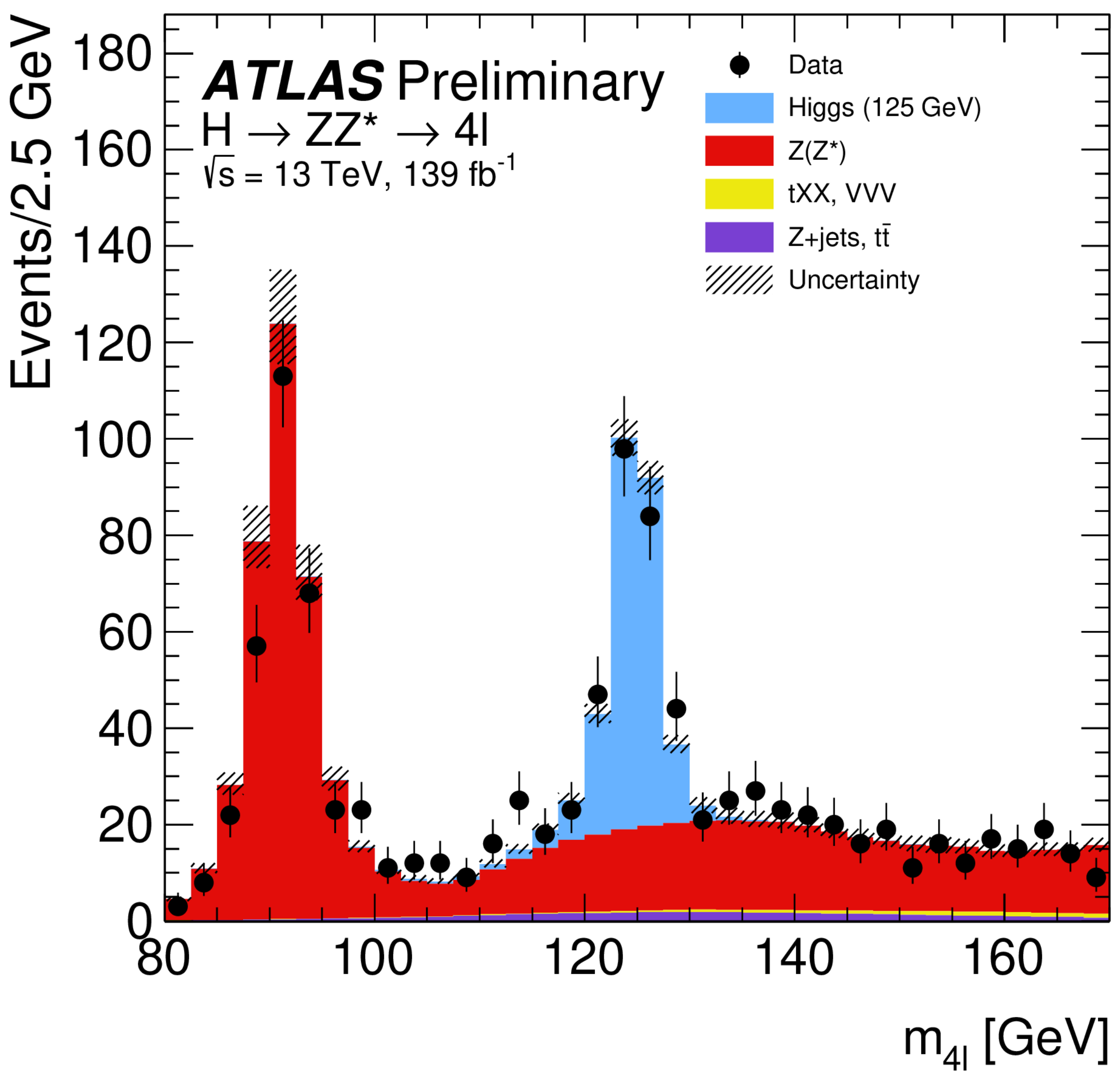


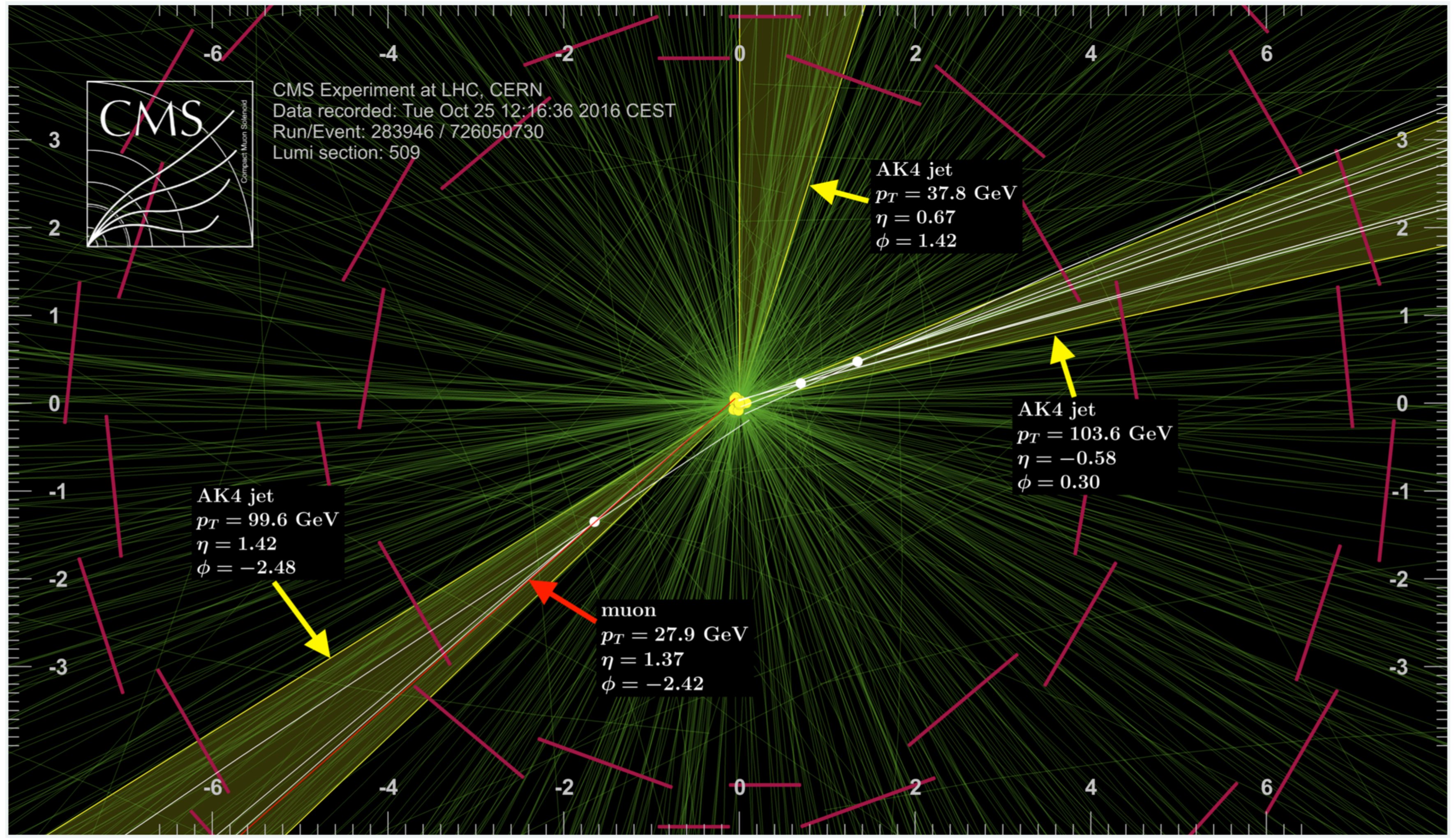




Summary

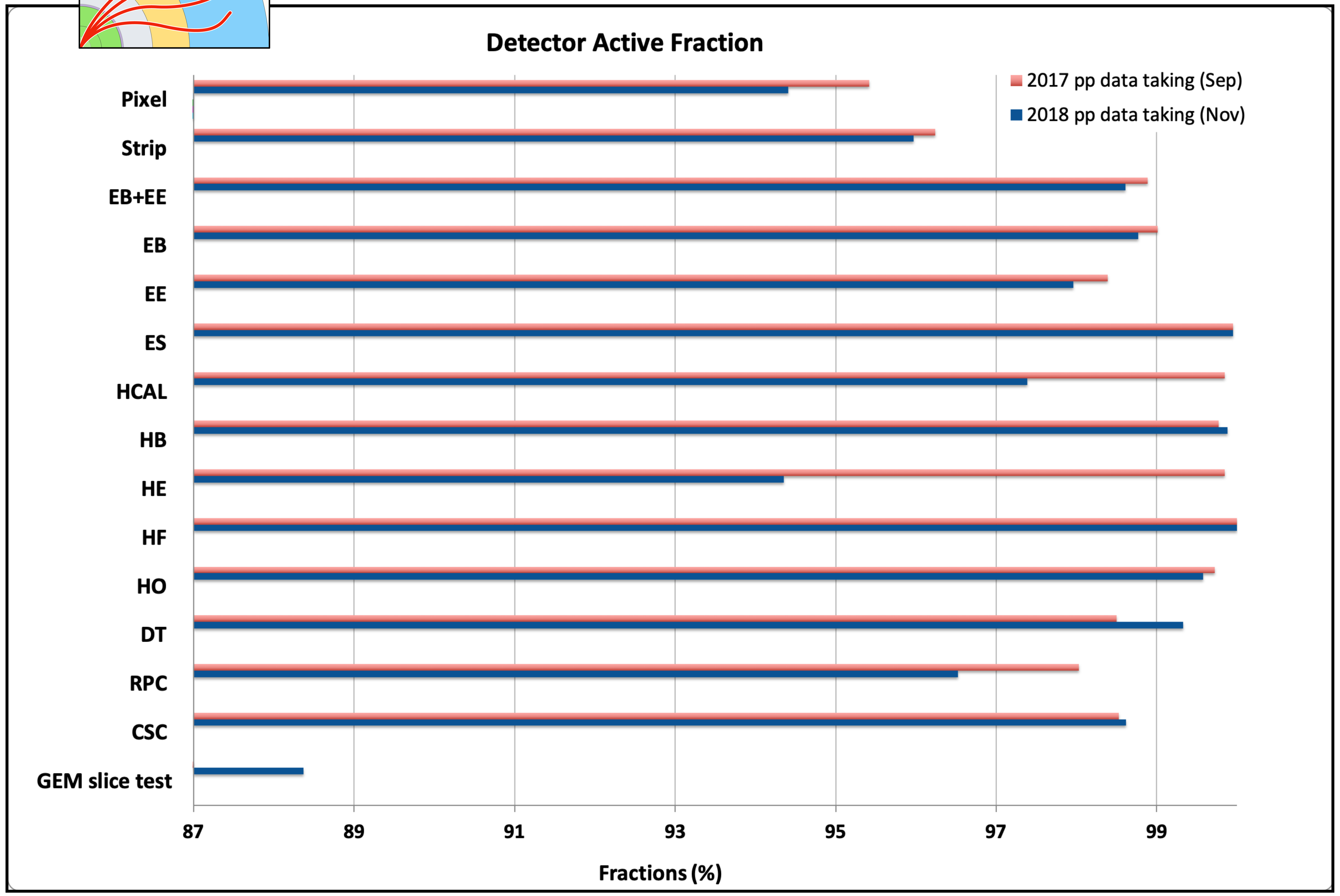
- ATLAS & CMS Trackers are a key ingredient for the success of the LHC physics program
- Very efficient operations of the ATLAS & CMS trackers
- Tracking algorithms need to provide high-quality tracks efficiently and with an efficient use of resources.
  - Despite challenging conditions at the LHC in Run2
- Tracking is in continuous development
  - Using novel tools (NN, ML…) for conditioning the hits and the track seeds
  - Tracking efficiency from simulation and data-driven techniques
- Detector alignment
  - Correcting short and long time scale movements
  - Crosschecks to avoid/mitigate weak modes (→ track parameter biases)
- Flavour tagging techniques in continuous development
  - Use ML techniques to combine the basic observables into classifier
- Run 3 preparations to cope with higher data rates
  - Pileup mitigation
- Thanks to all tracking, alignment and flavour tagging teams of ATLAS & CMS
- Apologise for leaving some topics out



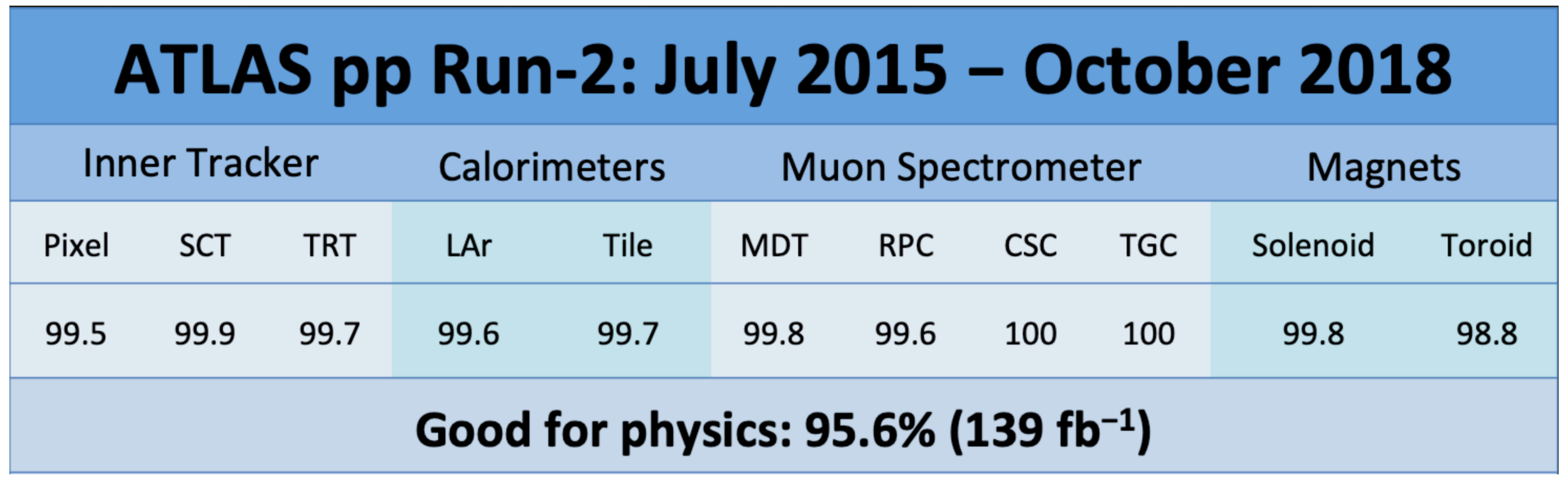


Operation

- Efficient operation of both trackers during LHC Run 2



Subdetector	Number of Channels	Approximate Operational Fraction
Pixels	92 M	95.7%
SCT Silicon Strips	6.3 M	98.6%
TRT Transition Radiation Tracker	350 k	97.2%
LAr EM Calorimeter	170 k	100 %
Tile Calorimeter	5200	99.5%
Hadronic End-Cap LAr Calorimeter	5600	99.7%
Forward LAr Calorimeter	3500	99.8%
LVL1 Calo Trigger	7160	99.9%
LVL1 Muon RPC Trigger	383 k	100%
LVL1 Muon TGC Trigger	320 k	99.9%
MDT Muon Drift Tubes	357 k	99.7%
CSC Cathode Strip Chambers	31 k	93.0%
RPC Barrel Muon Chambers	383 k	93.3%
TGC End-Cap Muon Chambers	320 k	98.9%
ALFA	10 k	99.9%
AFP	430 k	97.0%



ATLAS pp Run-2: July 2015 – October 2018

Inner Tracker			Calorimeters		Muon Spectrometer				Magnets	
Pixel	SCT	TRT	LAr	Tile	MDT	RPC	CSC	TGC	Solenoid	Toroid
99.5	99.9	99.7	99.6	99.7	99.8	99.6	100	100	99.8	98.8

Good for physics: 95.6% (139 fb^{-1})