4D Vertexing at CMS

Jyoti Babbar
University of Trieste and INFN, Italy
On behalf on CMS Collaboration

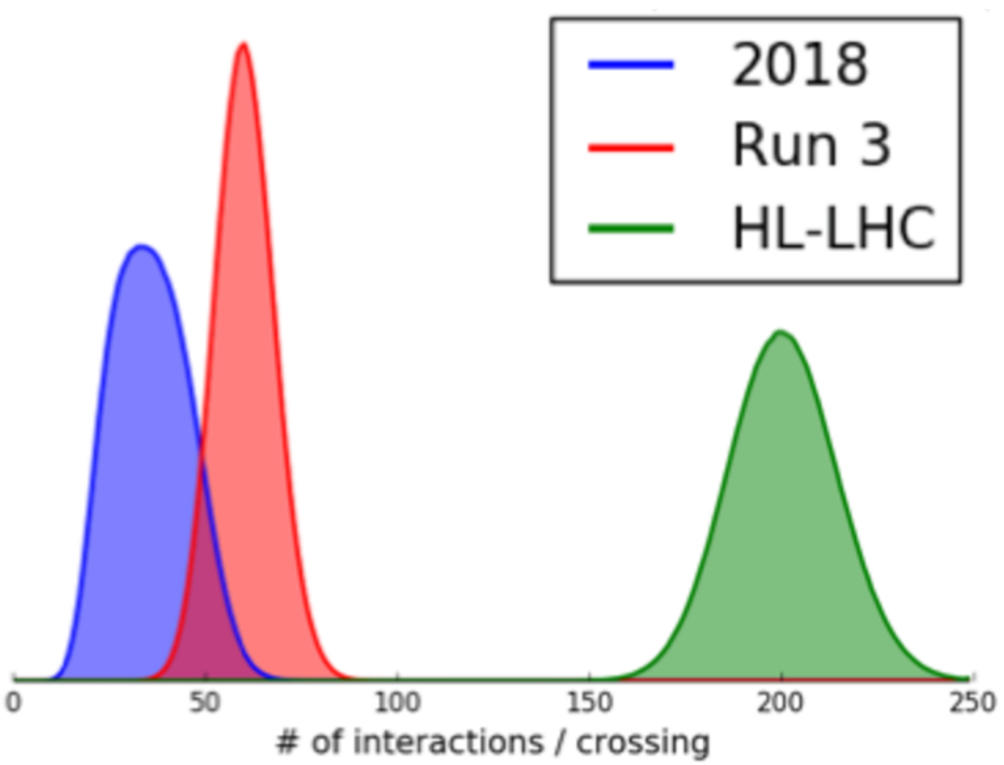
Istituto Nazionale di Fisica Nucleare

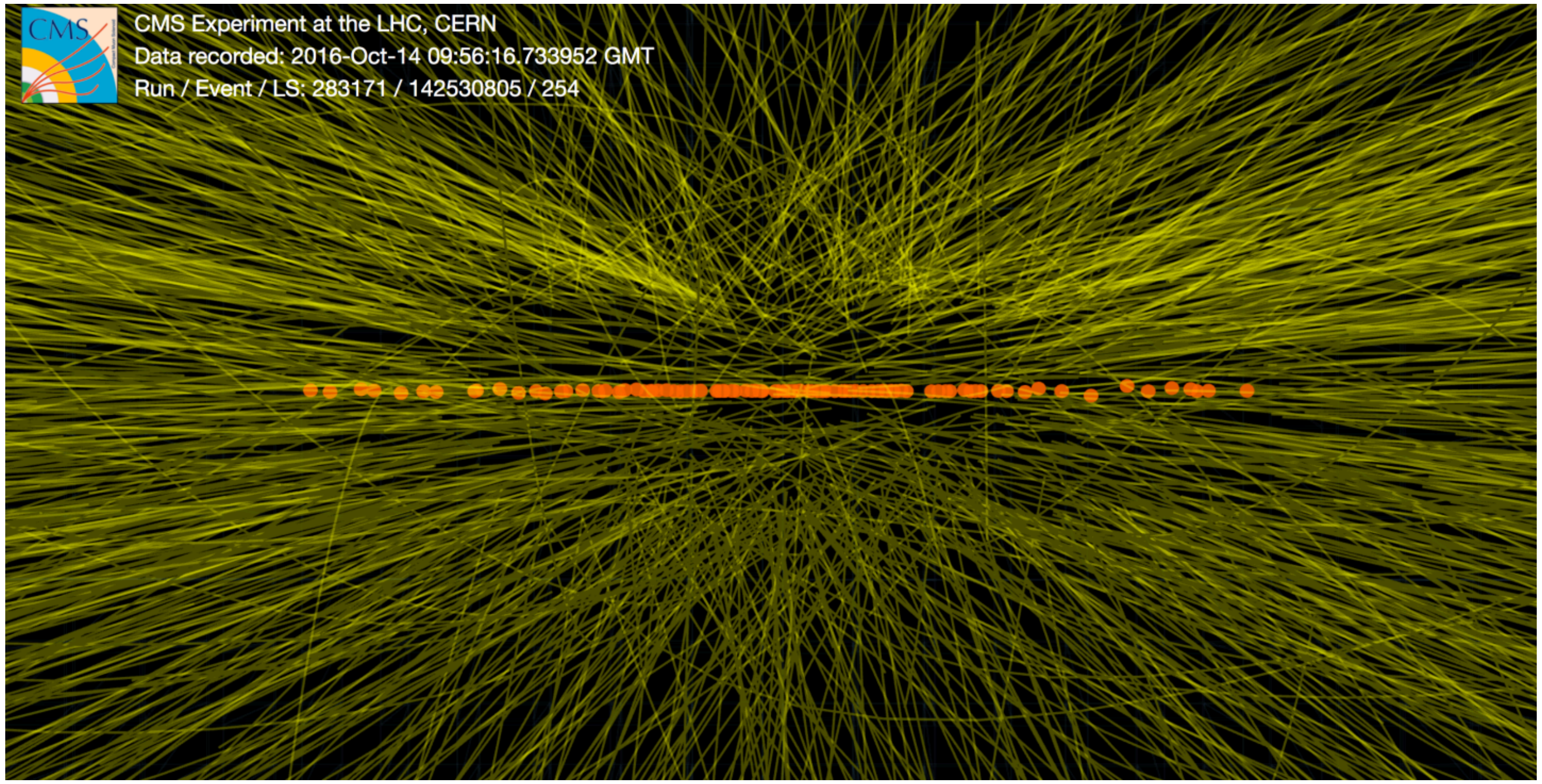
Tracking 2025 Workshop : Track Reconstruction in Particle Physics Experiments
Huizhou (China)
21-23 July

- HL-LHC
- CMS Upgrades
- Mip Timing Detector
- 4D vertex reconstruction
- Update of the 4D algorithm
- Performance study of:
  - vertex time resolution
  - reconstruction efficiency
  - pileup rejection

HL-LHC

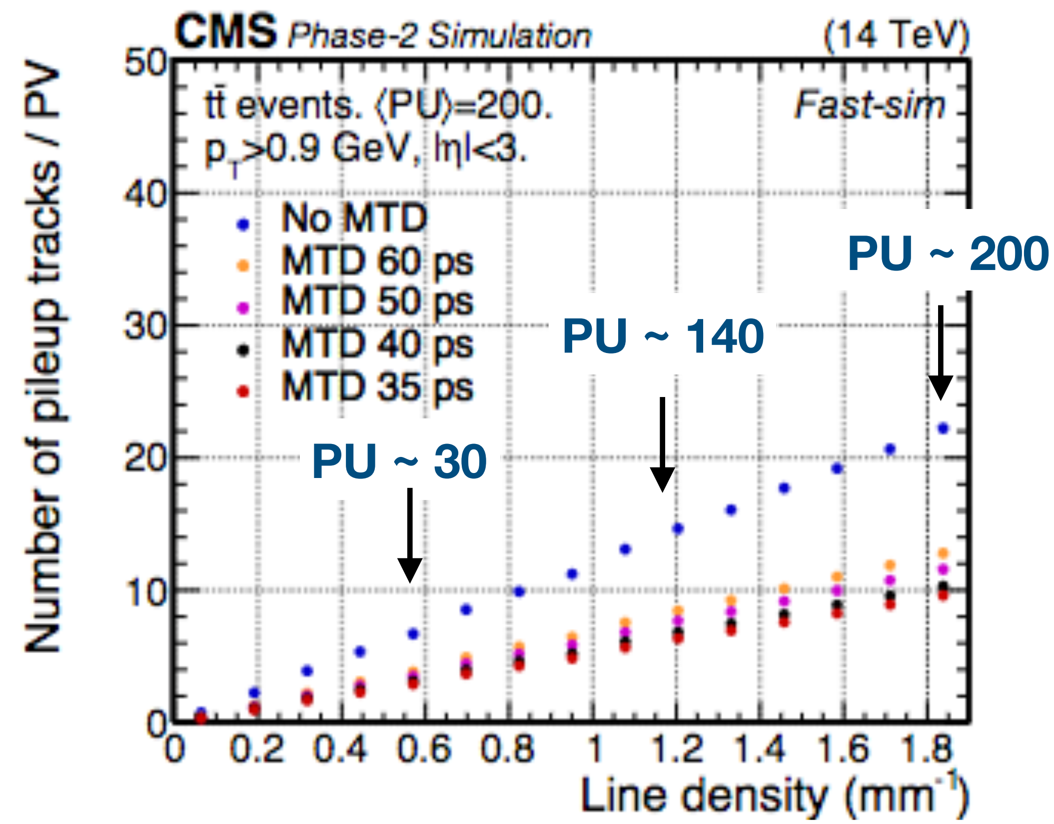
- High-luminosity LHC era (HL-LHC) starting in ~2030
  - x3-4 instantaneous luminosity
  - up to ~140-200 pileup (PU) interactions (~3x with respect to current situation)
  - x10 integrated luminosity
- Crucial to isolate interaction of interest and mitigate effects of PU on object reconstruction
- Current global event reconstruction relies on track-vertex association in space

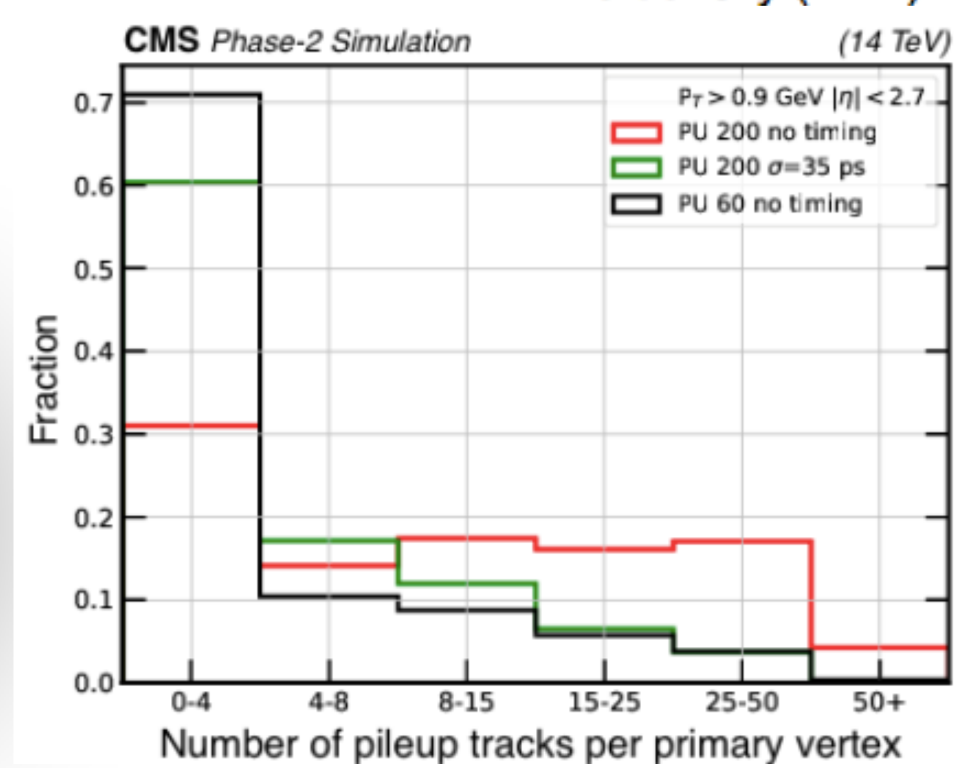




Experimental Challenges at HL-LHC

- Reconstruction depends on track-vertex assignments that become ambiguous when track resolution is comparable to vertex separation
- Vertex merging and the incorrect association of tracks with vertices distorts the final state kinematics.
- The efficiencies to correctly identify jets, leptons, and photons are affected; every object is degraded!
- Degraded reconstruction results in loss of sensitivity, undermining physics objectives motivating HL-LHC.
- Challenge: keep current performance during HL-LHC phase

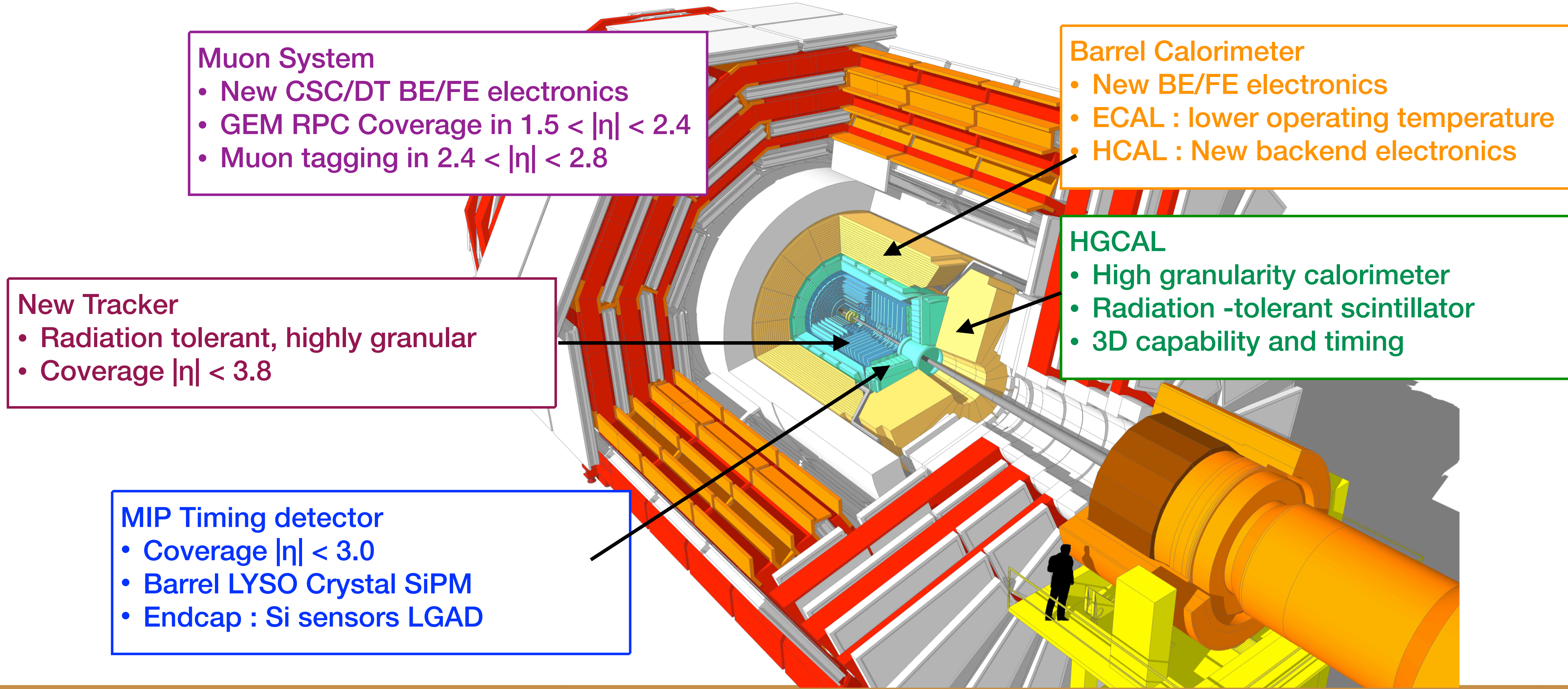




Timing significantly reduces the "effective" vertex line density —> 200 PU equivalent to current LHC PU (~50 PU)

Upgrade CMS detector to mitigate pileup and radiation damage

- New tracker and calorimeters with enhanced granularity and radiation tolerance.
- High-granularity calorimeters and tracker designed to operate in extreme pileup conditions
- **Minimum Ionizing Particle Timing Detector (MTD)** proposed for the CMS experiment Phase2 upgrade —- for precision timing of minimum ionizing particles (MIPs)

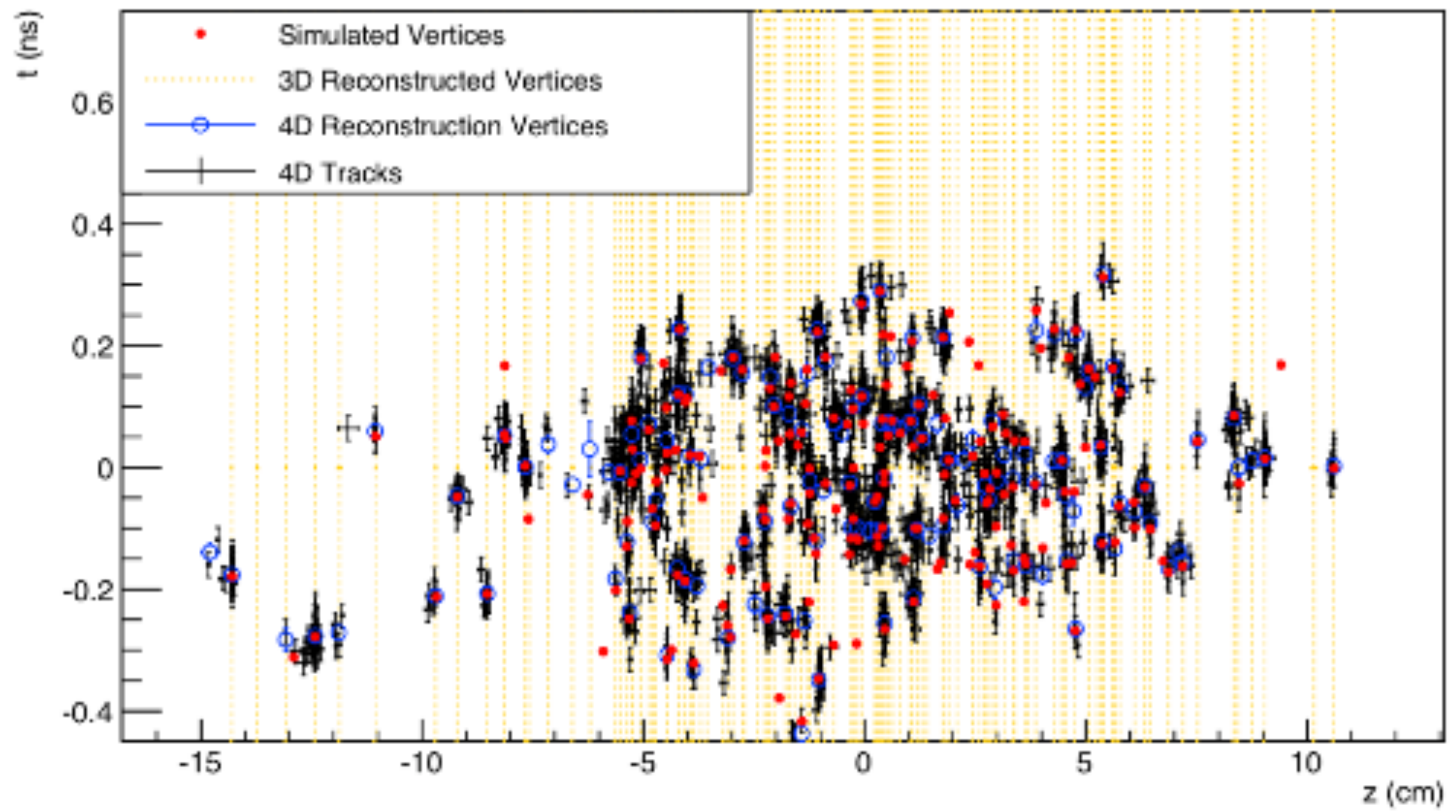


Why do we need a MIP Timing Detector in CMS?

- The MTD will provide timing information for MIPs with a 30-40 ps resolution
- This precision is significantly smaller than the O(**200 ps**) time spread of proton-proton collisions at the HL-LHC

Time-tagged charged tracks enable

- time compatibility check for track-vertex association
- Enhanced discrimination of vertices in space **and** time ("4D vertexing")



Vertices overlap in position, but separate in time

[CMS MTD Techinal Design Report]

MIP Timing Detector

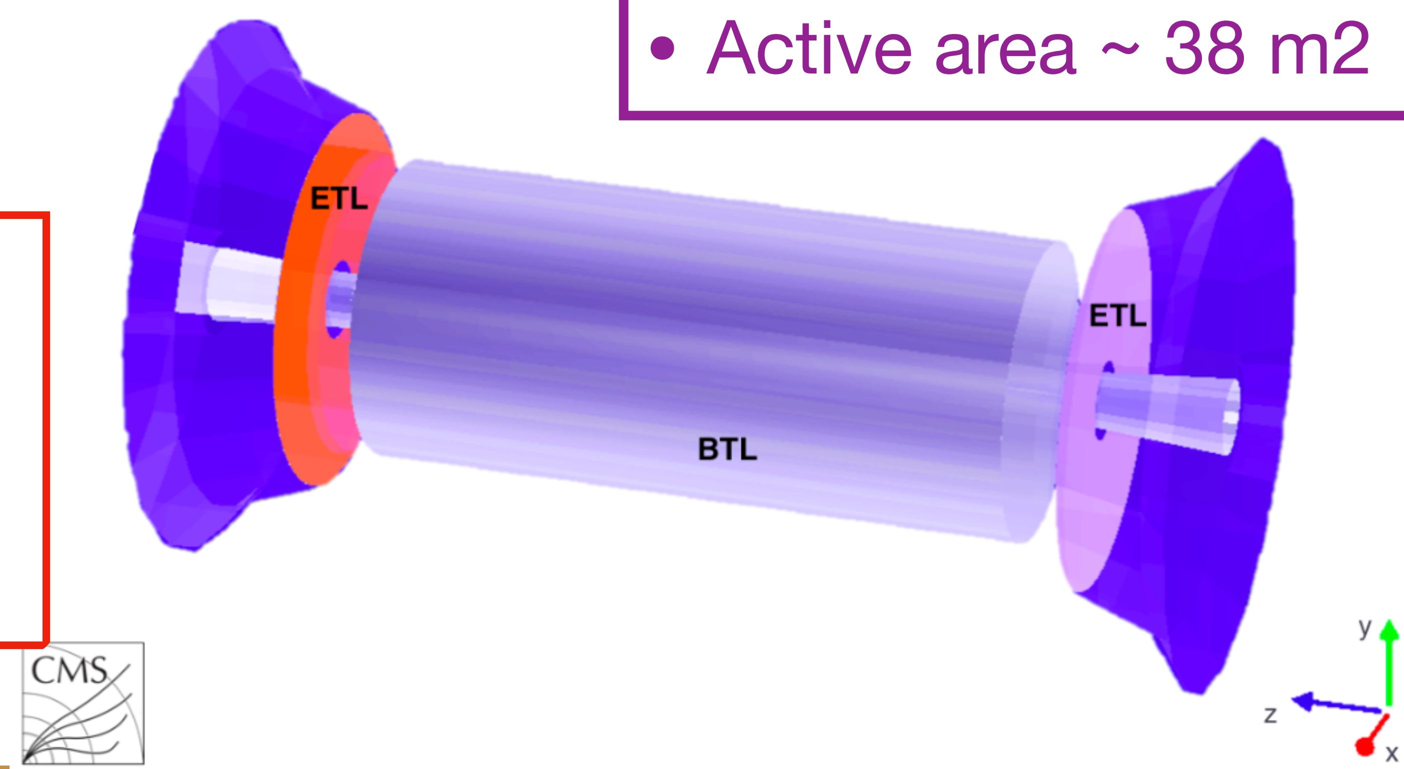
- Thin layer between tracker and calorimeters
- Almost hermetic (|η|<3)
- Different regions adopt different technologies, suited to the level of radiation dose:
  - Barrel Timing Layer (BTL)—arrays of LYSO crystal bars readout by SiPMs
  - Endcap Timing Layer (ETL)—Low Gain Avalanche Detector (LGAD) module

BTL

- LYSO Bars + SiPM readout
- |η| < 1.45 & pT > 0.7 GeV
- Active area ~ 38 m2

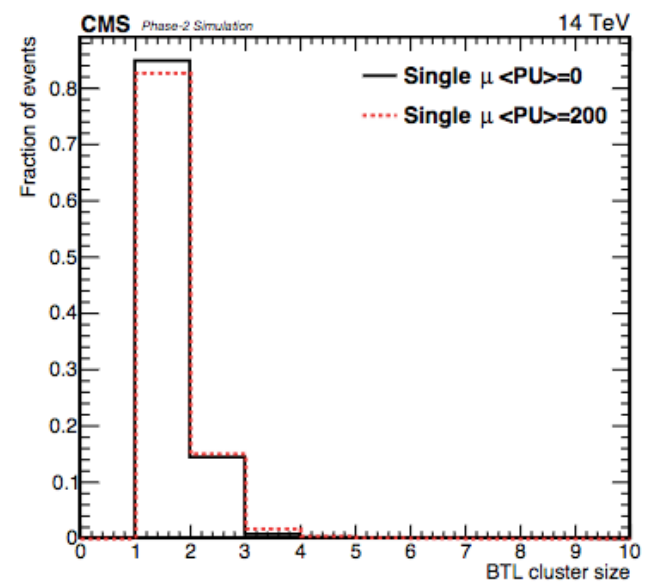
ETL

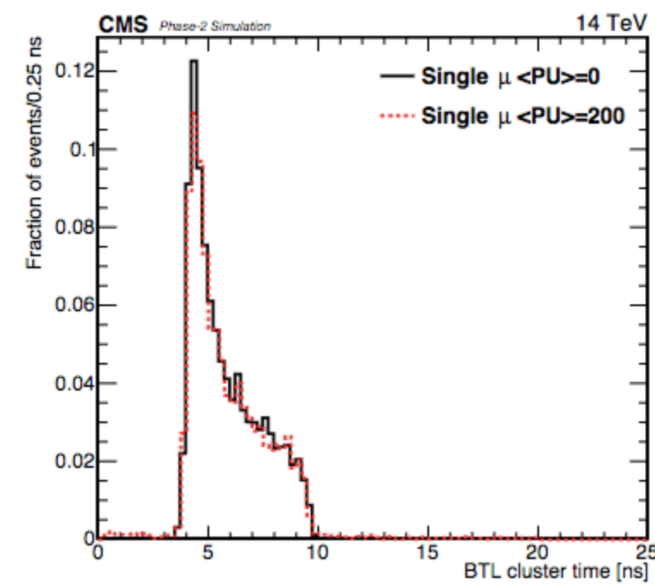
- Si with internal gain (LGADs)
- 1.6< |η| < 3.0
- Active area ~ 14 m2



Track Reconstruction With Timing

- **Local reconstruction:**
  - In BTL: single crystal measurements
  - $t_{av} = (t_L + t_R)/2$ (left and right SiPMs)
  - In ETL: pixel measurement
- **Topological clustering of adjacent MTD hits:**
  - In BTL: frequent multiple hits, especially at high η
  - In ETL: mostly single-hit clusters
  - Cluster time: weighted average of single-hit times



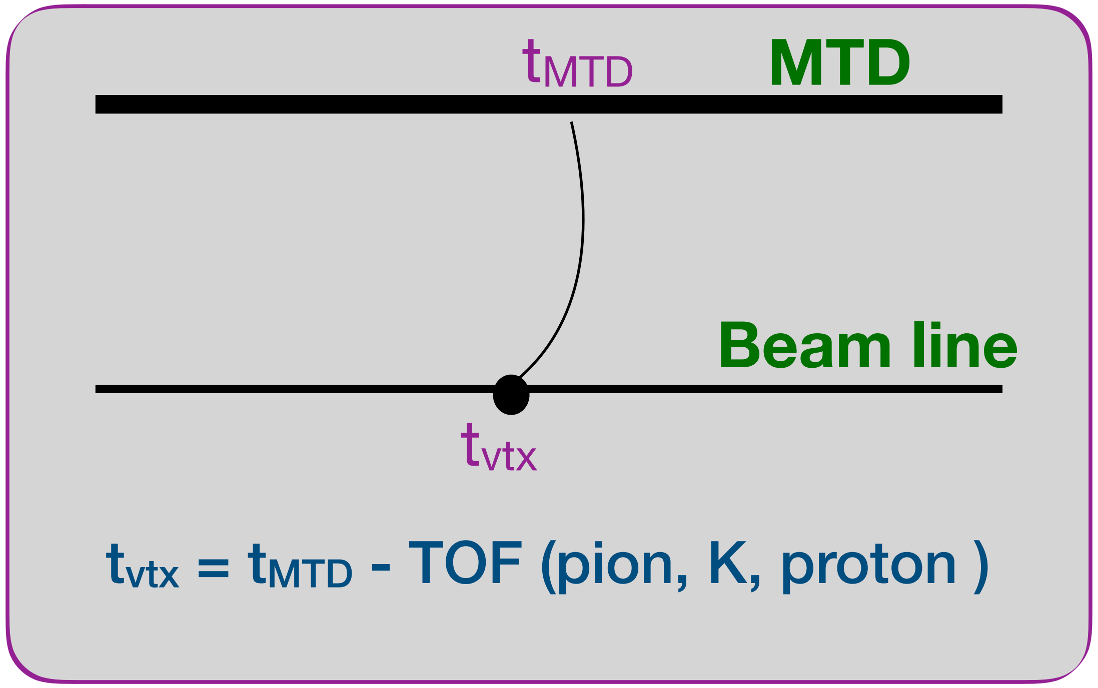


CMS MTD Techinal Design Report

- **Propagation of tracker tracks to MTD and matching with clusters**
  - Spatial matching based on χ^2 of track extrapolation
  - Time compatibility with beamspot constraint
  - To suppress background

Track Reconstruction With Timing

- Tracker + MTD track:
  - Re-fit track parameters including additional spatial measurement
  - Compute path length
- Given path length and momentum, velocity depends on mass hypothesis, a-priori unknown
- Propagate the track to its point of closest approach to beamline under various mass hypotheses (π, K, p)
  - Dedicated to primary vertex reconstruction
  - Similar approach may be used for secondary vertices exploiting their known position



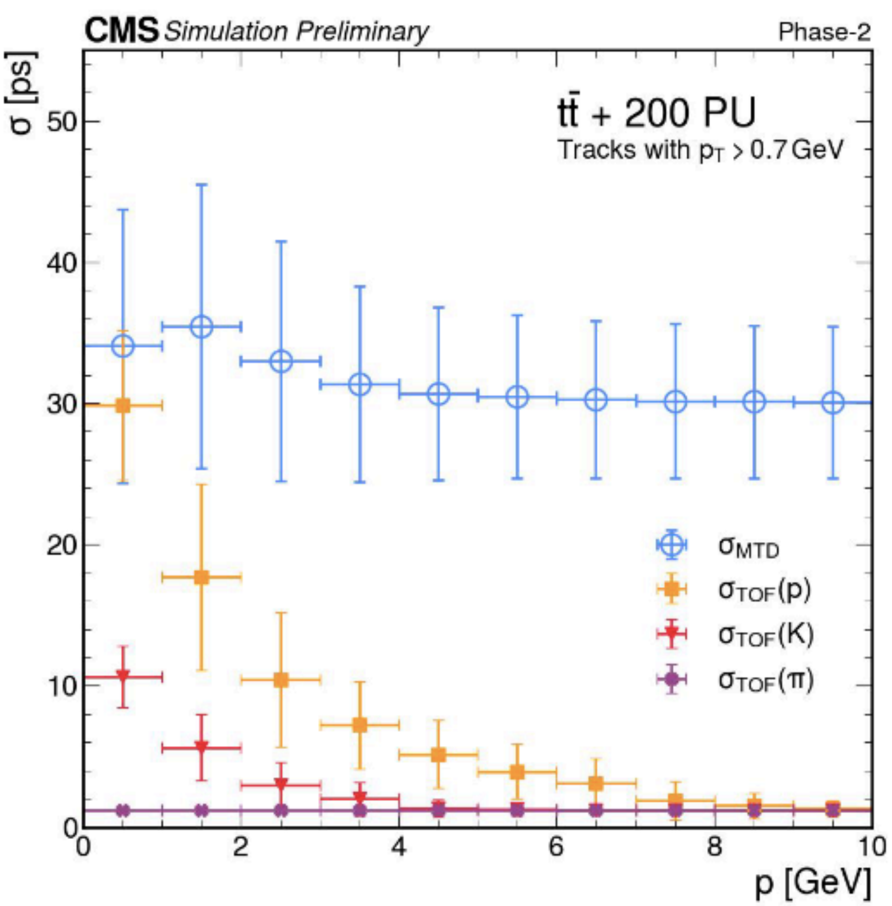
Track Time Uncertainty

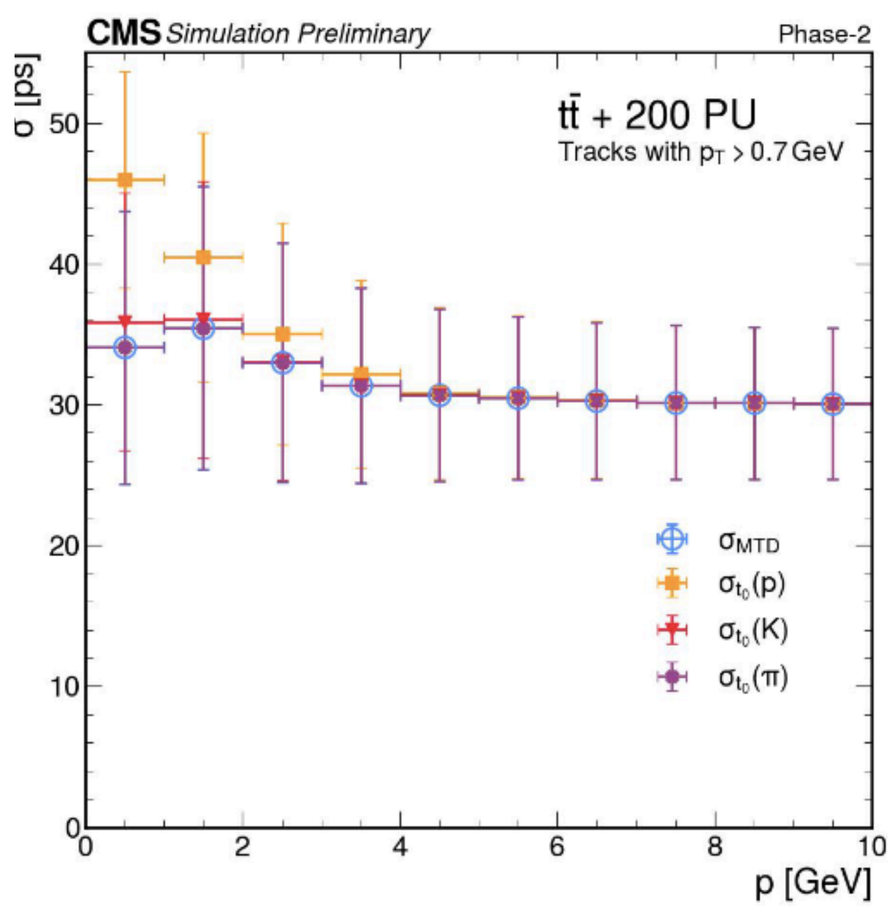
- Total track time uncertainty

$$\sigma(t_{vtx}) = \sigma(t_{MTD}) \oplus \sigma(t_{ToF}) \oplus \Delta(TOF_p - TOF_\pi)$$

- $\sigma(t_{ToF})$ propagates the uncertainty on particle's velocity derived from the reconstructed track momentum, that is particle hypothesis dependent

- CMS-DP/2024-048 studies shows that $\sigma(t_{ToF})$ has negligible impact





CMS-DP/2024-048

$\sigma(t_{MTD})$: estimated MTD time uncertainty
$\sigma(t_{TOF})$: TOF uncertainty
$\Delta(TOF_p - TOF_\pi)$: inflated uncertainty, due to π approximation

4D Vertex Reconstruction

4D vertex reconstruction and particle identification (**PID**) go hand in hand

- Measure track time @ MTD and momentum, velocity depends on the mas hypothesis
- So far vertex reconstruction - **legacy 4D** - in 2 steps

1st step

- Cluster vertex with pion hypotheses taking inflated uncertainties

$\sigma(t_{vtx}) = \sigma(t_{MTD}) \oplus \sigma(t_{TOF}) \oplus \Delta(TOF_p - TOF_\pi)$

- Calculate vertex time and PID

2nd step

- Calculate vertex using updated track times, and remove the inflated uncertainties

$\sigma(t_{vtx}) = \sigma(t_{MTD}) \oplus \sigma(t_{TOF})$

- Calculate vertex time and PID

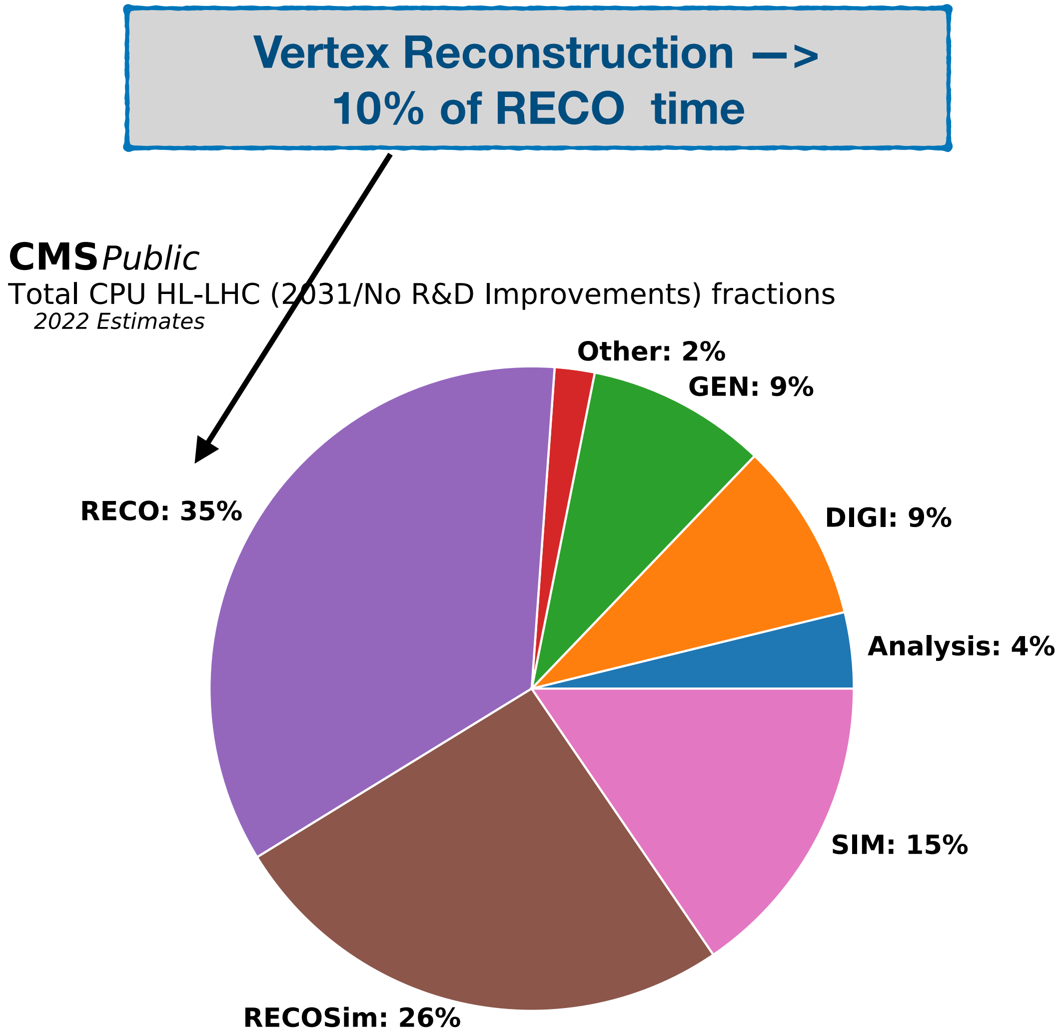
The updated 4D algorithm

The legacy 4D algorithm is sub-optimal:

- CPU-time consuming: in 1st step, inflated uncertainty dominates over MTD uncertainty at low momenta,

The **updated 4D** algorithm [CMS-DP/2024-085] replaces the 1st step of **legacy 4D** with 3D vertices:

- Possible thanks to **time** computation available for 3D vertices as well (**3Dt)**
- **Reduce** the vertex reconstruction **CPU-time** by 30% without loss in performance



CMS Offline Computing-Public Results

Vertex Time

In **legacy 4D**: compute time using only the mass hypothesis assigned after PID with a simple **weighted average**

$$t_{vtx} = \frac{\Sigma_i \frac{1}{\sigma^2_{t,i}} \cdot t_i}{\Sigma_i \frac{1}{\sigma^2_{t,i}}}$$

New: time computed with a deterministic annealing (DA) time algorithm using all 3 mass hypotheses, minimizing the cost function:

track time for **π**, K, p

$$F = -T \sum_{\text{tracks},i} w_{0,i} \log \left(Z_0 + \alpha_\pi e^{-\frac{(t_i(\pi) - t_v)^2}{2\sigma^2_{t_i}}} + \alpha_K e^{-\frac{(t_i(K) - t_v)^2}{2\sigma^2_{t_i}}} + \alpha_p e^{-\frac{(t_i(p) - t_v)^2}{2\sigma^2_{t_i}}} \right)$$

track weight from adaptive vertex fit

track time uncertainty for **π**, K, p

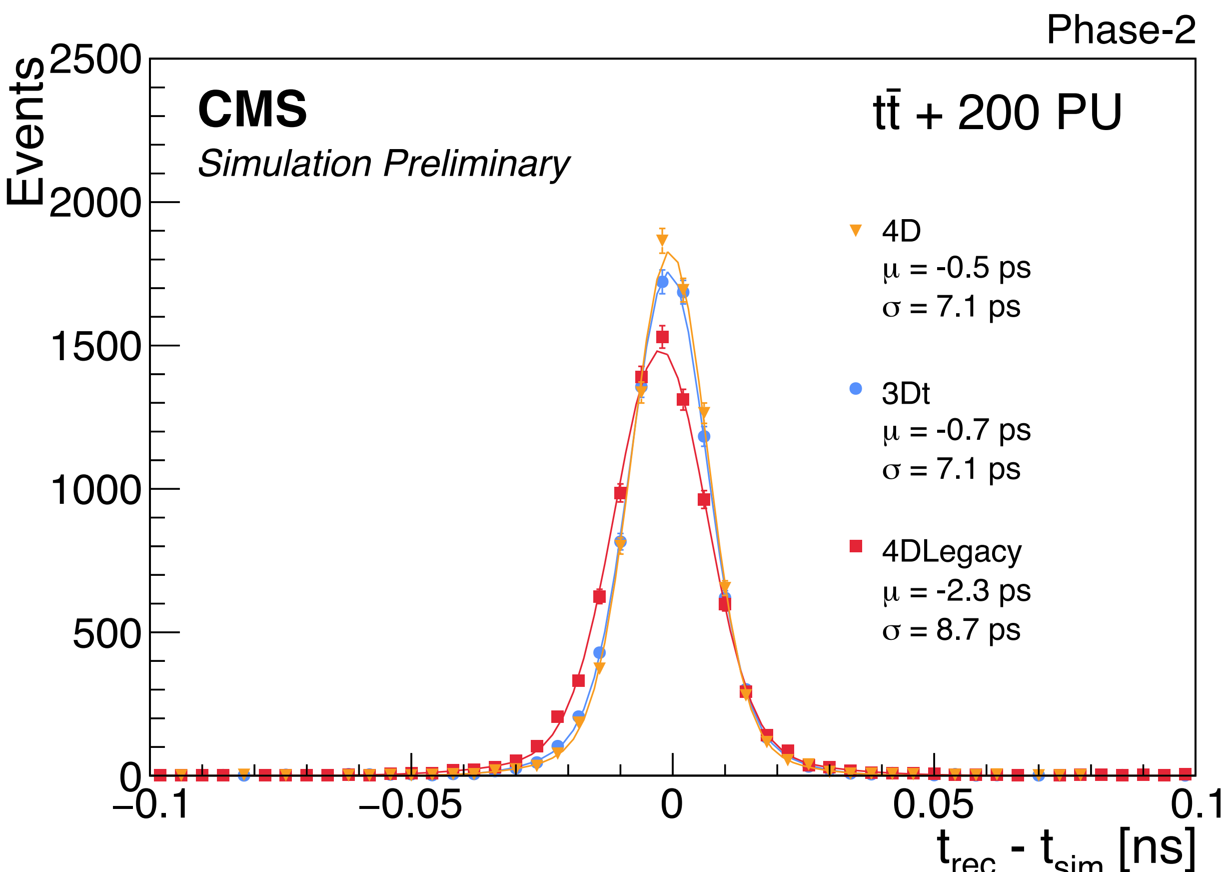
- **This algorithm can be applied to a reconstructed vertex regardless of the use of time in its clustering and fitting:**

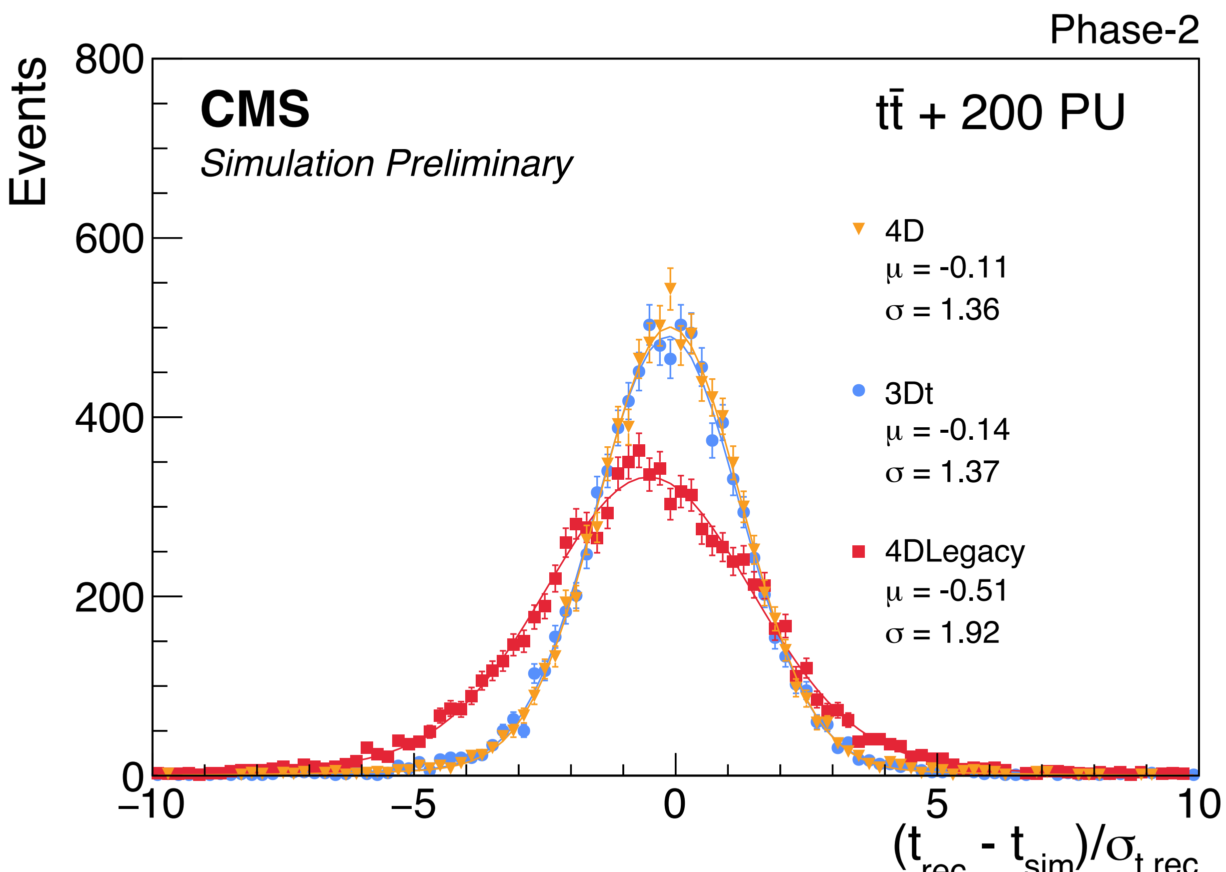
a prior probability for **π**, K, p (0.7, 0.2, 0.1)

  - **3Dt** vertex with the DA time calculation
  - updated **4D** with the DA time calculation in 2nd step

Signal Vertex Time

- The vertex time resolution and pull for signal vertices, the distributions are fitted with a double Gaussian: the parameters shown refer to the narrowest one
- The **3Dt and updated 4D algorithms** yield **notable improvements** in both time resolution and pull, compared to the **legacy 4D vertexing** approach
- The legacy 4D method exhibits a systematic negative bias, primarily due to incorrect mass hypothesis assignments (such as kaon or proton); this is effectively corrected in the updated algorithms.





CMS-DP/2024-085

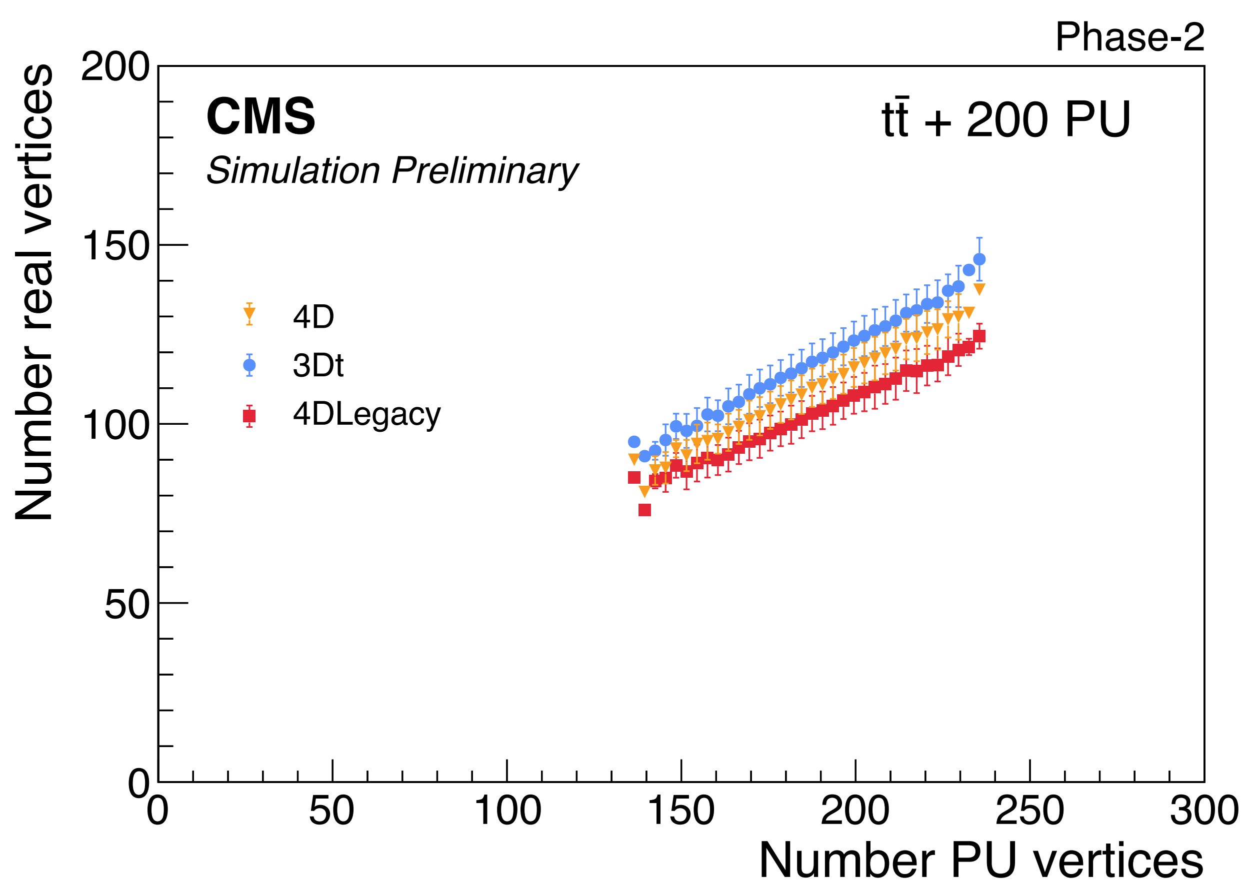
Finanziato dall'Unione europea NextGenerationEU | Ministero dell'Università e della Ricerca | Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA | ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

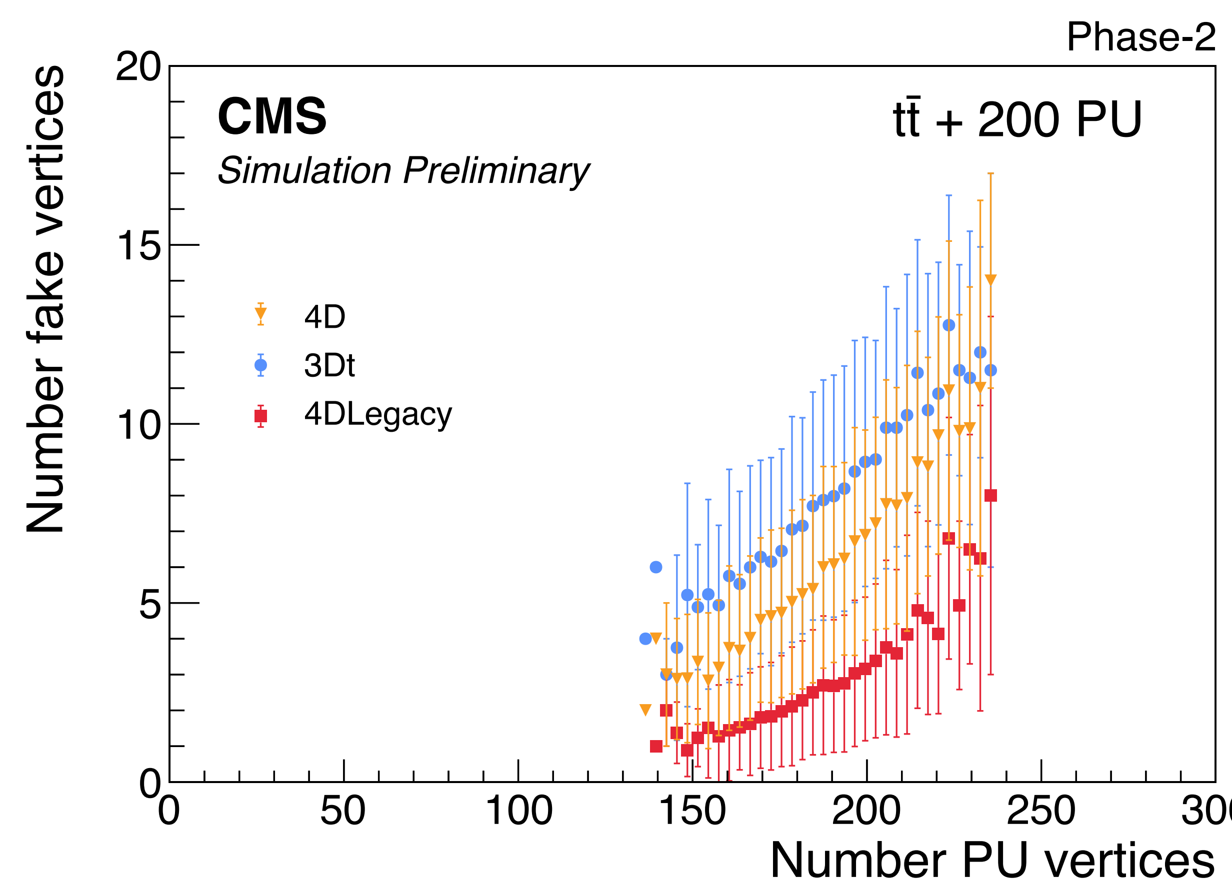
Number of Vertices

The number of reconstructed vertices as a function of the number of PU vertices for **real** and **fake** vertices

- Classification based on **matching to MC truth** - both true tracks and vertices (details in backup)
- The **3Dt** reconstructs more real vertices, but also more fakes
- The updated **4D** algorithm shows a **higher number of vertices** than the legacy, with a performance in between the **4D legacy** and the **3Dt** algorithms

Error bars indicate the **RMS** of the distribution, not the **uncertainty on the mean**



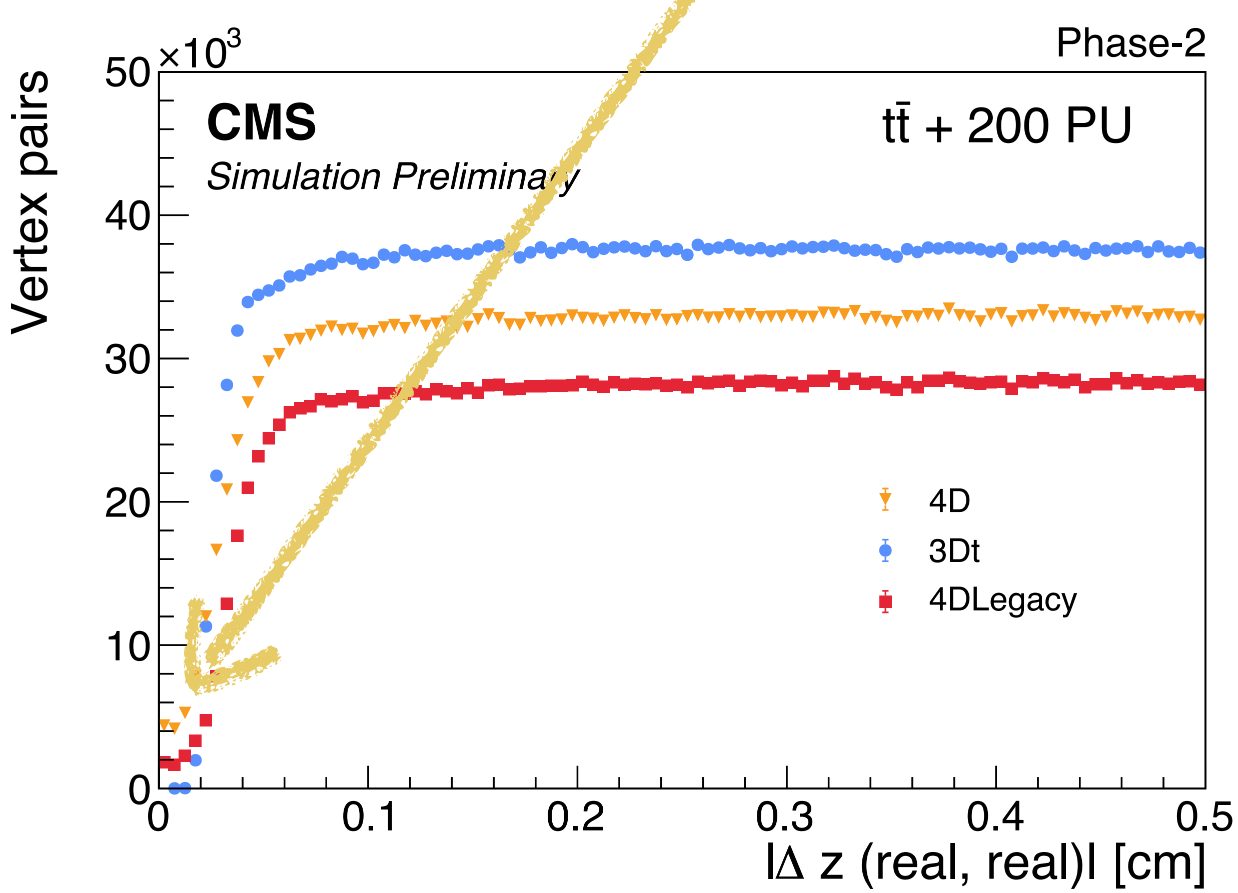


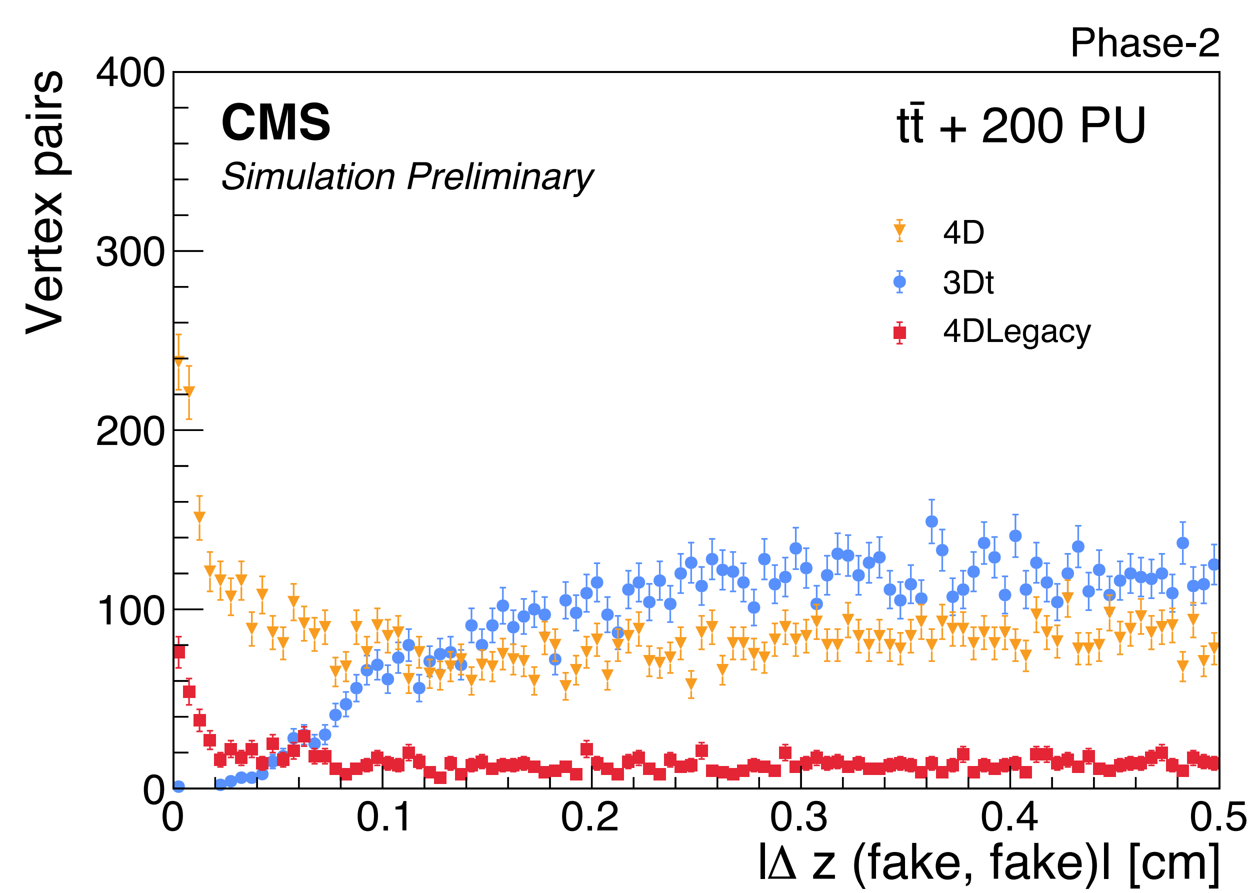
CMS-DP/2024-085

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

Distance in Z

- Distance between pairs of reco vertices: the updated **4D** algorithm shows more real-real vertex pairs close in z than **legacy 4D**, but also more fakes
- The **3Dt** algorithm is not designed to reconstruct vertices with separation less than ~0.3 mm
- The **improvement** in the new algorithm is visible especially for **real vertex pairs** with **Δz close to 0**
- Advantage in the use of **timing**: vertices that overlap in space can be separated in time

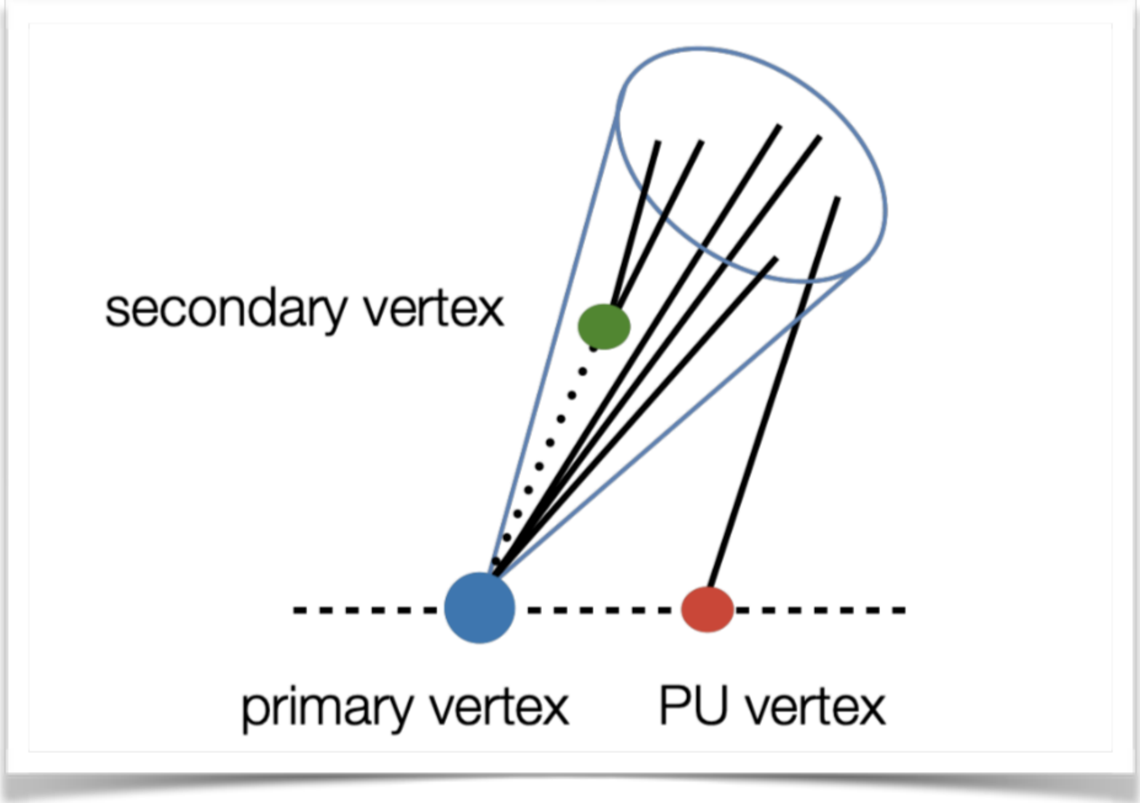




CMS-DP/2024-085

Pileup Contamination

- Compare vertex algorithms in terms of **PU rejection**:
  - Essential for accurate object reconstruction under high pileup conditions at the HL-LHC
  - primary goal of MTD
- Performance is monitored using observables derived from both tracks and jets
- **Tracks** associated to a reconstructed vertex are classified based on MC truth matching as:
  - Track from primary vertex
  - Track from secondary vertex
  - PU track
  - Fake track, not matched to any true particle

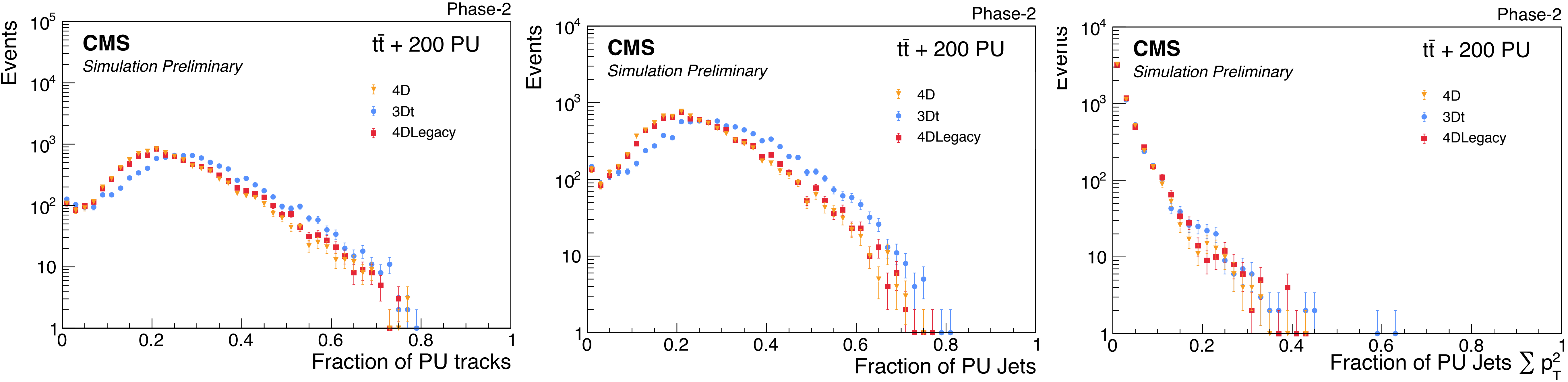


- **Jets** are built by clustering reconstructed charged tracks originating from the same vertex
- The relative **contribution of PU** to jet-based quantities is estimated by clustering jets without the PU tracks and recomputing the observables

Pileup Contamination for the Leading Vertex

Impact of PU on track multiplicity, jet multiplicity and sum of $p_T{}^2$ of jets

- The **4D vertexing algorithms** show a general improvement, reducing PU contamination by approximately **10–15%** compared to the **3Dt approach**
- In contrast, some variables—such as **the sum of jet $\mathbf{p_T{}^2}$**, which is used in vertex sorting—**are less sensitive to the vertex reconstruction algorithm**



CMS-DP/2024-085

Conclusions

- Precision timing from the MTD is essential for mitigating pileup (PU) at the HL-LHC.
- A set of tools has been developed for evaluating the performance of different vertex reconstruction algorithms in terms of:
  - Vertex time resolution
  - Number of reconstructed **true** and **fake** vertices
  - **PU rejection**
- These tools provide a benchmark for future algorithm development and exploration of advanced techniques
- The optimization of the 4D vertex reconstruction is presented
- **4D vertexing enables temporal separation of spatially overlapping vertices, enhancing pileup suppression**

Thank You!

Backup

References

[1] CMS Collaboration, "A MIP Timing Detector for the CMS Phase-2 Upgrade", technical report, CERN,Geneva, 2019.

[2] CMS Collaboration, "Improved use of MTD time in vertex reconstruction", CMS DP-2024-048 (2024).

[3] CMS Collaboration, "Update of the vertex reconstruction using track time from MTD", CMS DP-2024-085 (2024)

Vertex association to MC truth

To evaluate the performance of the vertex reconstruction, an algorithm has been developed to **match reconstructed vertices** to **MC truth**, based on the common origin of tracks in the reconstructed and simulated vertices. The reconstructed tracks are matched to the true simulated charged particles, and the simulated vertices from which they originate define the set of true primary vertices in the event. The matching algorithm is based on the sum of weights:

$$W_{\mathrm{os}} = \frac{w_{trk}}{\sigma^2_{z,trk}} \frac{1}{\mathrm{erf}(\sigma_t/\sigma_T)}$$

where w_{trk} is the weight assigned by the adaptive vertex fit, $\sigma_{z,trk}$ the track resolution, σ_t the track time uncertainty and σ_T the time width of the beamspot. The time dependent part is present only for tracks with time information.

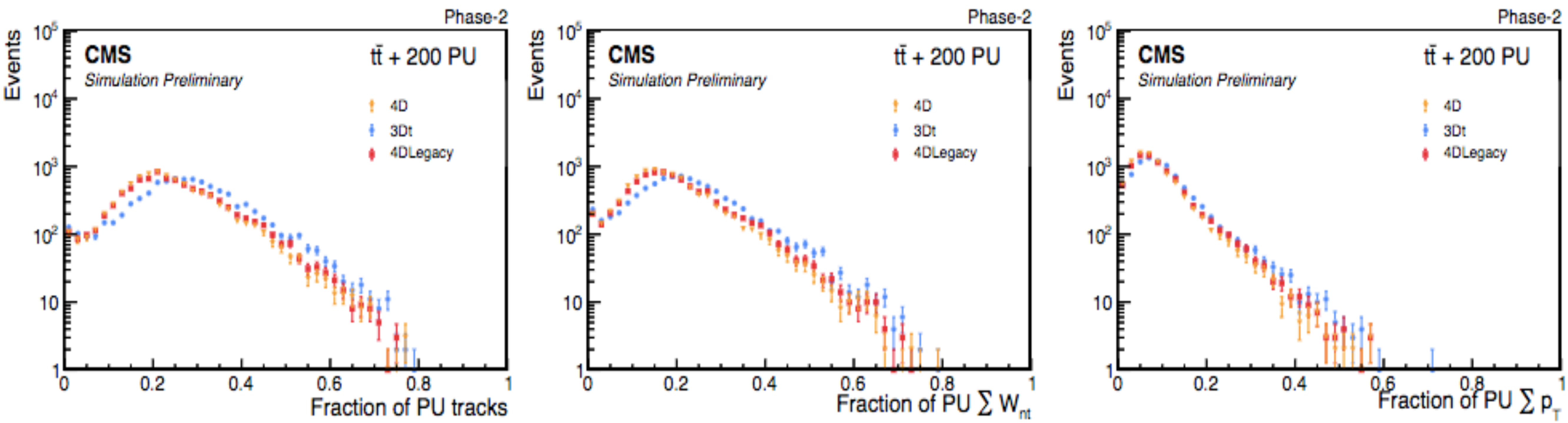
A **one-to-one matching** is performed: for a given **reconstructed vertex**, the **dominating simulated vertex** is the one with **largest sum of** W_{os}. Whenever possible, the dominating simulated vertex is matched to the reconstructed one. If a simulated vertex dominates more than one reconstructed vertex, the match is made between that simulated vertex and the reconstructed one that receives the largest weight among the dominated reconstructed vertices. The algorithm proceeds in an iterative manner for all the reconstructed vertices. Depending on the outcome of the algorithm, vertices are classified as:

- **real**: a good matching is found within the maximum allowed number of iterations (set to 8),
- **fake**: no matching can be found, meaning that there is no simulated vertex dominating the reconstructed vertex that does not dominate other vertices more.

5

CMS-DP/2024-085

Pileup contamination in vertices

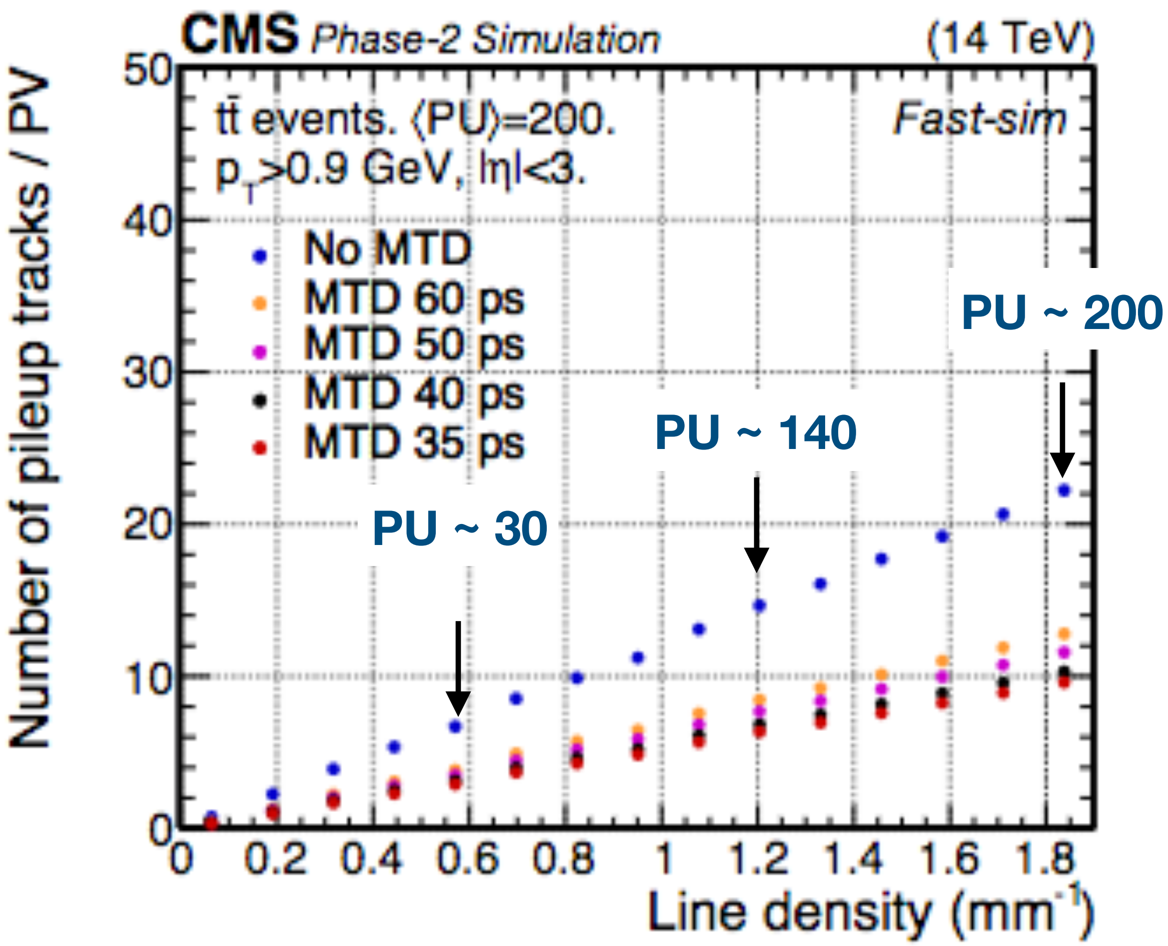


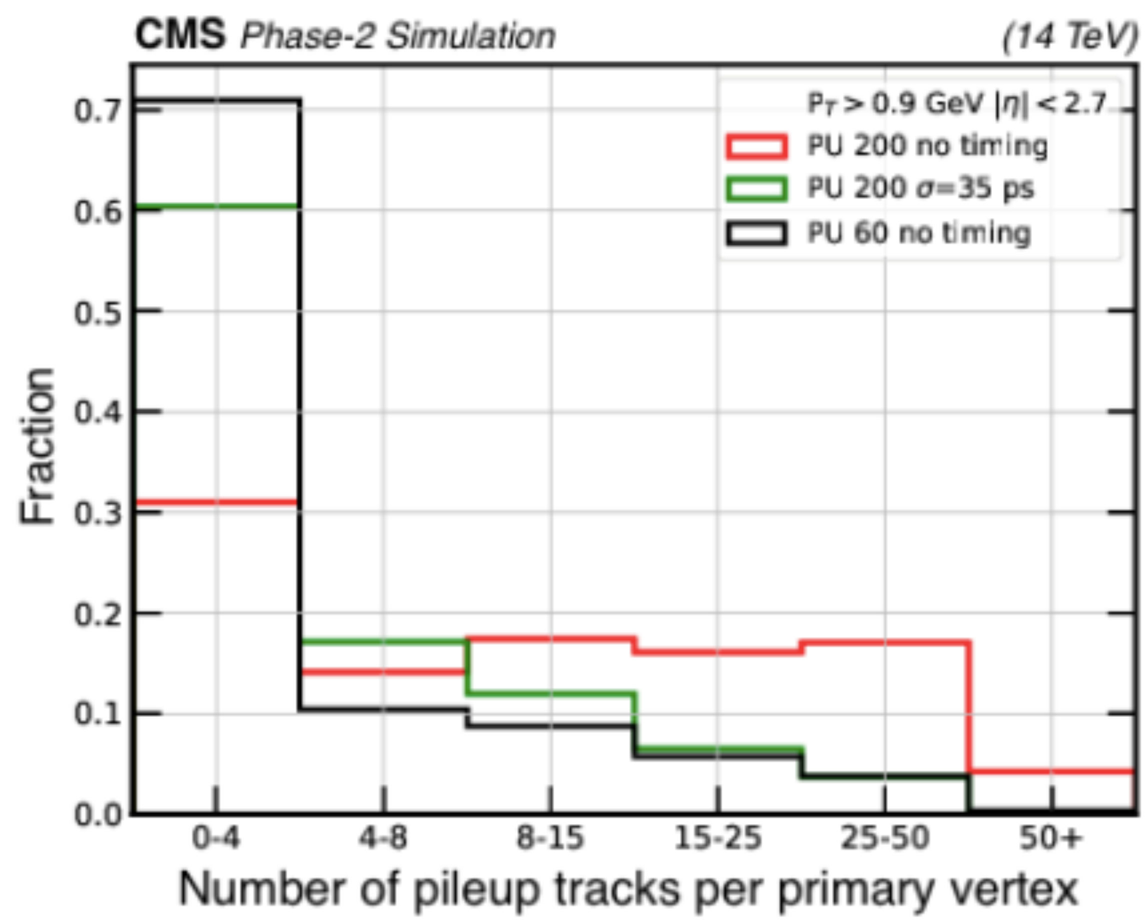
The relative contribution of PU to track multiplicity (left), $\sum_{trk} W_{\rm nt}^{trk} = \sum_{trk} w_{trk} \times \min(p_{T,trk}, 1)$ (middle) - where w_{trk} is the weight assigned by the adaptive vertex fit and p_T is the track's transverse momentum in GeV - and $\sum_{trk} p_{T,trk}$ (right) for the LV of the event, for the updated 4D (orange triangles), the 3Dt (blue circles) and the 4DLegacy (red squares) vertex reconstruction algorithms. The sum of $W_{\rm nt}^{trk}$ corresponds to the simple sum of weights of tracks, where low momentum tracks are downgraded, and is an auxiliary weight in the matching algorithm. A PU reduction of about 10/15% is generally observed in the updated 4D and 4D Legacy vertex algorithms with respect to the 3Dt one. 10

CMS-DP/2024-085

Precision Timing at CMS in HL-LHC

- Timing significantly reduces the "effective" vertex line density
  - 200 PU equivalent to current LHC PU (~50 PU)





CMS MTD Techinal Design Report