

Tracking, alignment and b-tagging performance and prospects in CMS

Jan-Frederik Schulte^{*†}

Purdue University

E-mail: jshulte@cern.ch

Charged particle tracks and track-based higher level observables are a key ingredient of the event reconstruction for the CMS experiment. In 2017, the CMS pixel detector was upgraded, promising significantly improved performance, but also requiring a concerted commissioning effort and improved algorithms to realize this potential from the start of the data taking in that year. Described are the alignment procedure and tracking algorithms used in 2017, as well as the potential tracking performance with future upgrades. Additionally, recent and future b-tagging developments and the resulting performance improvements are discussed.

Sixth Annual Conference on Large Hadron Collider Physics (LHCP2018)

4-9 June 2018

Bologna, Italy

^{*}Speaker.

[†]On behalf of the CMS Collaboration.

1. Introduction

A precise and efficient reconstruction of the particles produced in the proton-proton collisions is crucial for both measurements and searches for new physics at the CERN LHC. A key ingredient to this is the reconstruction of the trajectory of charged particles in the detector (“tracks”). The all-silicon tracking detector of the CMS experiment [1] allows for excellent precision in the track reconstruction and therefore efficient tagging of heavy-flavor jets. During the extended shutdown of the LHC between the 2016 and 2017 data taking periods, a new pixel detector was installed [2], adding a fourth detector layer in the central barrel part (BPix) as well as third discs in the forward parts of the detector (FPix). The geometry of the old and new pixel detectors are compared in Figure 1. While promising a significantly improved performance of the tracking in CMS, as well as of related higher level objects such as b -tagged jets, the installation of this new detector component posed a significant challenge for the detector alignment and track reconstruction algorithms to ensure optimal performance from the start of the data taking. In the following, the performance achieved with the new detector is described and an outlook on further developments in these areas in the future is presented.

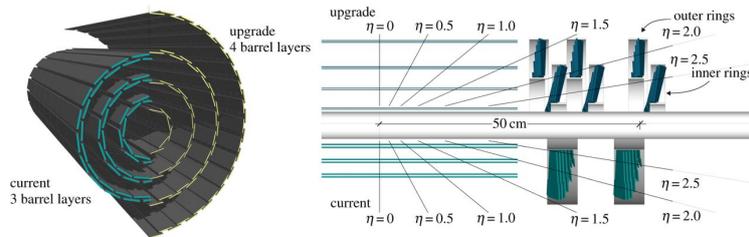


Figure 1: Comparison of the geometries of the initial and upgraded CMS pixel detectors.

2. Alignment

The resolution of the track-parameter measurements is limited by multiple scattering and the hit-position resolution. The latter is determined by the intrinsic position resolution of the detector modules as well as the uncertainty on the position, rotation, and surface deformation of the sensors. To achieve optimal performance, these uncertainties are reduced to a level well below the intrinsic resolution of 10-30 μm using a track-based alignment procedure. Two complementary approaches are used to perform least-squares fits to the track-hit residuals of a large sample of tracks to obtain the $\mathcal{O}(10^5)$ alignment parameters. MILLEPEDE II [3, 4, 5] performs a global fit of all parameters simultaneously, while HIPPIE fits track and alignment parameters separately, taking the correlations into account iteratively. A mix of collision and cosmic-ray data is used to control global systematic distortions of the detector. For this purpose $Z \rightarrow \mu\mu$ data is included, which allows imposing an invariant mass constraint.

For the data taking in 2017, a first alignment of the new pixel detector had to be performed based solely on cosmic-ray data. Using a sample of about 40000 tracks recorded with 0 T magnetic field, first the high-level mechanical structures of the detector and in a second step the individual

modules have been aligned. The resulting track-hit residuals in local x' direction in the BPix is shown on the left side of Figure 2. If the detector is well aligned, these residuals are expected to follow a narrow Gaussian distribution centred around 0. Before any alignment, the distribution is basically flat. After the module-level alignment, the distribution is well centred at 0 with a width of $83.4 \mu\text{m}$ [6].

Due to the limited amount of time between the ramp up of the CMS magnet to full 3.8 T field and the beginning of the data taking, only few cosmic rays could be recorded at full field. Still, a new alignment could be performed which significantly improved performance. This can be seen in the middle plot in Figure 2, showing that in the FPix, the track-hit residual resolution improved from 86.4 to $14.1 \mu\text{m}$ [7]. The final alignment during the commissioning phase of the new pixel detector was obtained including collision data, using about 35 million tracks in total, including several million $Z \rightarrow \mu\mu$ events. The right plot in Figure 2 shows that with this alignment, the width of the residual distribution in the FPix improved to $0.3 \mu\text{m}$, reaching a value well below the intrinsic resolution of the modules [7].

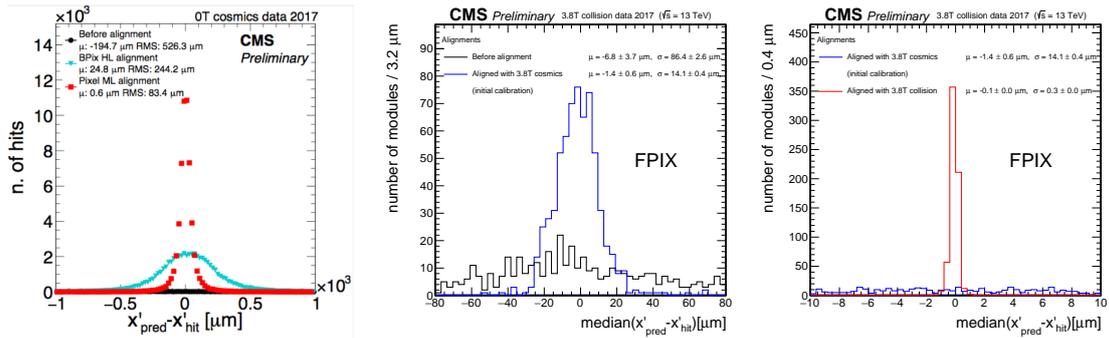


Figure 2: Track-hit residual distributions after different stages of the alignment of the pixel detector using 2017 data [6, 7].

The quality of the alignment can be validated using a variety of observables. Two examples are shown in Figure 3. The left side shows the mean track impact parameter in the x - y plane as a function of the track η . A significant bias can be observed for the endcaps when using only cosmic tracks for the alignment, which gets significantly reduced when including collision data. Similarly, as shown on the right of Figure 3, deformations in the detector introduce a bias in the measured Z boson mass in dimuon events when using only cosmic tracks in the alignment, which can be cured when making use of $Z \rightarrow \mu\mu$ events.

3. Tracking

Tracks are reconstructed in an iterative procedure, initially applying very strict requirements on the track seeds in the pixel detector, which are subsequently relaxed when the detector has been

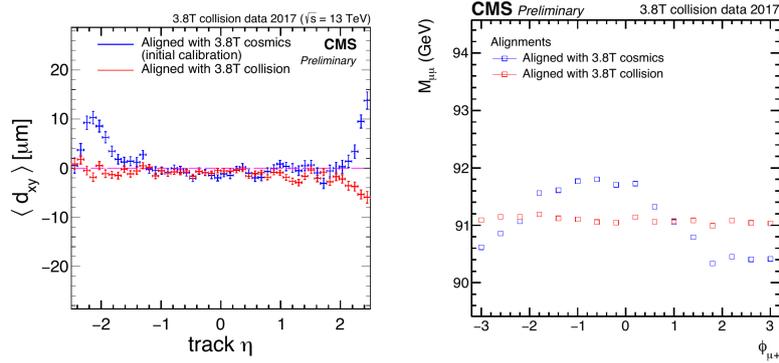


Figure 3: Mean track impact parameter in the x-y plane versus track η (left) and mean reconstructed Z boson mass as a function of ϕ of the positively charged muon (right) [6].

cleared of the hits already used in tracks found in previous iterations, thus reducing the combinatorial challenge [8]. The expected tracking efficiency based on MC studies achieved after updating the tracking algorithms to make use of the upgraded pixel detector is shown in Figure 4 [9].

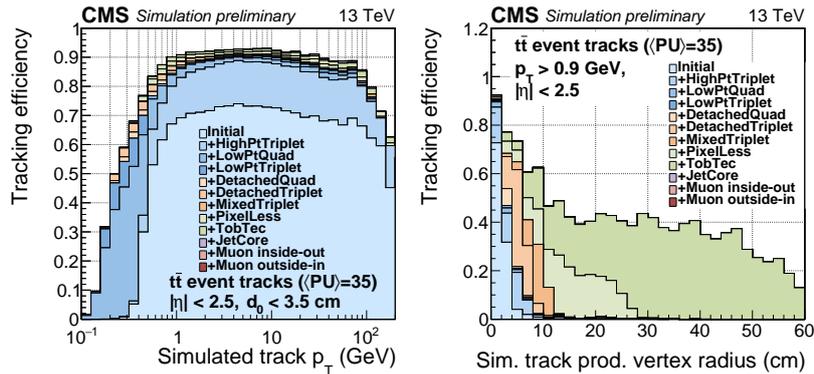


Figure 4: Tracking efficiency in MC as a function of track p_T (left) and production vertex radius (right) [9].

The most significant algorithmic improvement is the introduction of a “Cellular Automaton” (CA) algorithm to create track seeds, where hit pairs are found between detector layers. Track seeds are then created by forming hit triplets and quadruplets by finding pairs that share hits, after checking the compatibility of the seed with the interaction point. A sketch of the algorithm is shown in Figure 5.

A comparison of the tracking performance achieved with the CA seeding and the upgraded pixel detector in 2017 with the performance in 2016 is shown in Figure 6. A significant increase in tracking efficiency is observed, especially in the forward regions of the detector. At the same time, the track fake rate is significantly reduced. One major concern when designing the track reconstruction algorithms for the 2017 data taking was the time needed to perform the reconstruction. With the higher number of pixel detector modules, the time to find track seeds increases, especially for high numbers of additional proton-proton interactions (pileup) in the event, when using the same algorithm, as can be seen on the left of Figure 7. Utilizing the CA for track seeding, this time

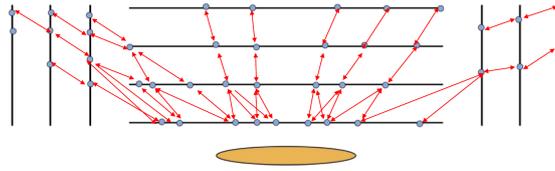


Figure 5: Schematic of the Cellular Automaton track seed algorithm.

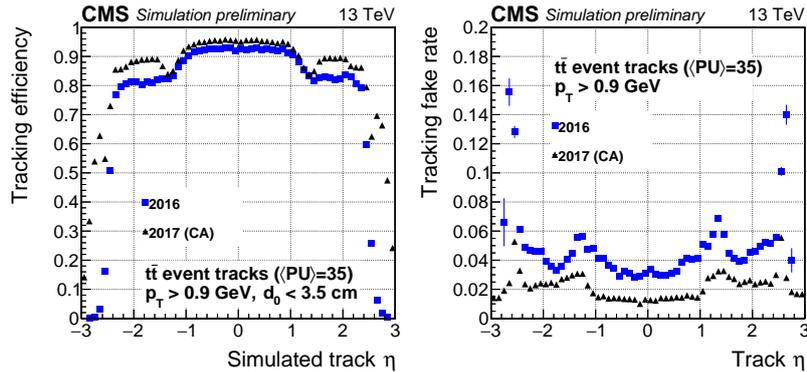


Figure 6: Comparison of tracking performance in 2016 and 2017 simulation. Shown are the tracking efficiency (left) and fake rate (right), both as a function of track η [9].

increase is mitigated and the seeding is even slightly faster than the 2016 setup. For the pattern recognition, the computing time is reduced compared to 2016 due to the reduced fake rate in the seeding, as shown on the right side of that Figure.

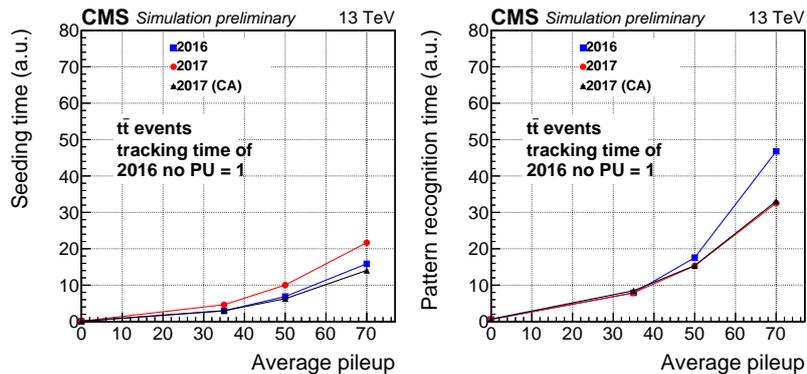


Figure 7: Average seeding (left) and pattern recognition (right) time in simulation as a function of pileup for both the 2016 and 2017 detector [9].

A further upgrade of the full CMS tracking detector, known as the Phase-2 upgrade, is foreseen in 2023. The pseudorapidity coverage will be extended to $|\eta| = 4$. With the upgraded detector, a similar tracking efficiency as with the current detector is achieved at much higher pileup of 140, as

shown on the left of Figure 8 [10]. Additionally, a significant improvement in the measurement of track parameters is expected, as illustrated for the impact parameter and the track p_T resolution in the middle and right plot of that Figure.

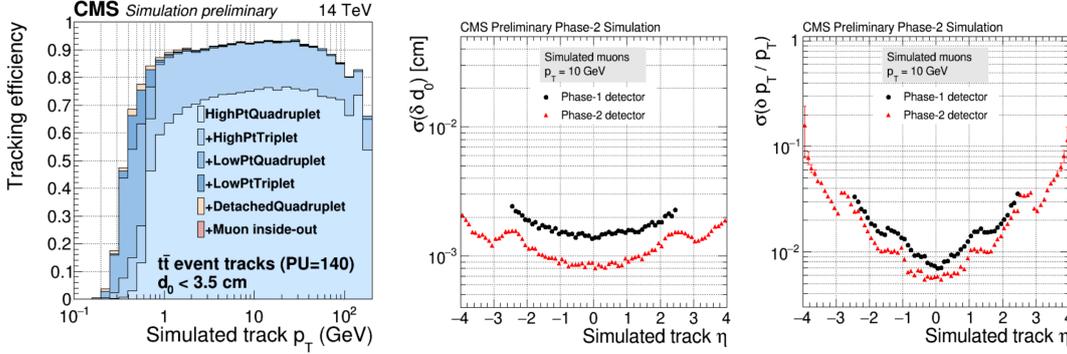


Figure 8: Projected tracking performance in simulation with the Phase-2 tracker upgrade [10].

4. B-Tagging

The identification of jets originating from b quarks (b -tagging) relies heavily on a large variety of track-based observables, such as track impact parameters and secondary vertices. Machine learning techniques, such as neural networks, are employed to combine this information into a single classifier. These algorithms are constantly improved upon to maximize b -tagging performance. In 2017, CMS introduced the DeepCSV tagger, which employs a deep neural network. Its performance is compared to the so far used CSVv2 algorithm in Figure 9 (left). For a mis-tag probability of 1%, the b -tagging efficiency is increased by 4% [11]. The b -tagging performance also improves due to the better tracking performance with the upgraded pixel detector in 2017. The right plot in Figure 9 shows the improved performance of the DeepCSV, which gains 7-8% efficiency at 1% mis-tag rate [12]. Going forward, a more complex neural network, taking into account neutral par-

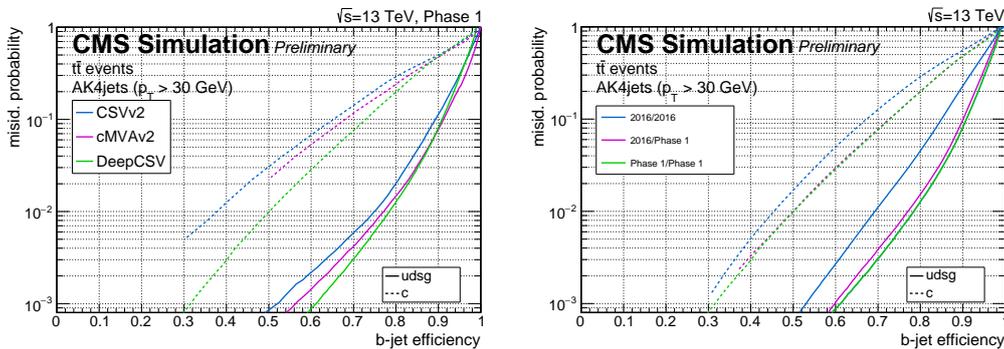


Figure 9: Simulated b -jet efficiency versus misidentification probability comparing different taggers (left) and comparing different pixel detectors for the DeepCSV tagger (right) [11, 12].

title candidates for the first time, will further improve the performance. The DeepFlavour algorithm performance is compared to the DeepCSV tagger in Figure 10. For jets with $p_T > 30$ GeV, the gain in efficiency is 4% at 1% mis-tag rate, while it reaches 20% for p_T between 600 and 1000 GeV [12].

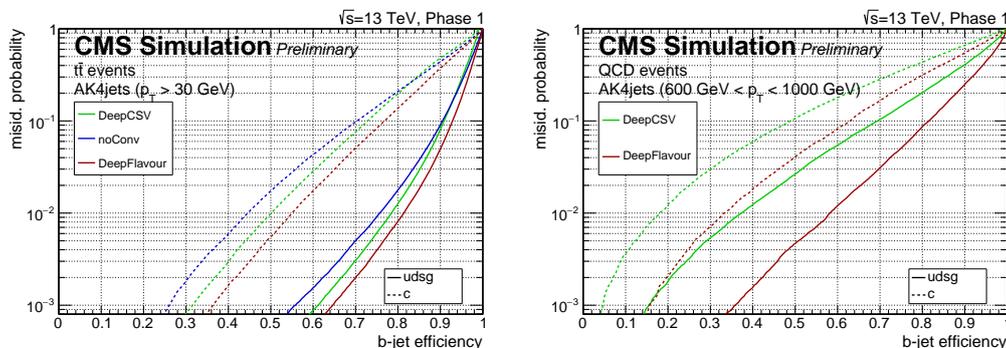


Figure 10: B-jet efficiency versus misidentification probability comparing the DeepCSV with the DeepFlavour tagger for jet $p_T > 30$ GeV in simulated $t\bar{t}$ events (left) and between 600 and 1000 GeV in simulated QCD events (right) [12].

5. Summary

The upgrade of the CMS pixel detector in 2017 posed a significant challenge for the detector alignment and tracking algorithms to deliver the improvement performance from the beginning of the data taking. After several alignment iterations with increasingly large datasets, very good performance was achieved. Upgrades to the track seeding algorithms allowed to exploit upgraded detector within the same CPU time budget, resulting in higher efficiency and lower fake rate. Tagging of jets originating from b-jet was improved both by the detector upgrade and significantly improved algorithms. Further upgrades to the detector and algorithms will further improve the performance of both tracking and b-tagging in the future.

References

- [1] CMS Collaboration, “The CMS Experiment at the CERN LHC”, 2008 JINST 3 S08004.
- [2] CMS Collaboration, “CMS Technical Design Report for the Pixel Detector Upgrade”, CMS-TDR-011, CERN-LHCC-2012-016 (2012).
- [3] V. Blobel, “Software alignment for tracking detectors”, Nucl. Instrum. Meth. A566 (2006) 5. doi:10.1016/j.nima.2006.05.157
- [4] CMS Collaboration, “Alignment of the CMS tracker with LHC and cosmic ray data”, JINST 9 (2014) P06009. doi:10.1088/1748-0221/9/06/P06009.
- [5] CMS Collaboration, “Alignment of the CMS Silicon Tracker during Commissioning with Cosmic Rays”, JINST 5 (2010) T03009. doi:10.1088/1748-0221/5/03/T03009
- [6] CMS Collaboration, “CMS Tracker Alignment Performance Results Start-Up 2017”, CMS-DP-2017-055, (2017).

- [7] CMS Collaboration, “Tracker Alignment Performance Plots after Commissioning”, [CMS-DP-2017-054](#), (2017).
- [8] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker”, JINST 9 (2014) P10009. doi:10.1088/1748-0221/9/10/P10009
- [9] CMS Collaboration, “CMS Tracking POG Performance Plots For 2017 with PhaseI pixel detector”, <https://twiki.cern.ch/twiki/bin/view/CMSPublic/TrackingPOGPerformance2017MC>, (2017).
- [10] CMS Collaboration, “Tracking Performance of the CMS Phase-2 Detector”, [CMS-DP-2017-010](#), (2017).
- [11] CMS Collaboration, “Heavy flavor identification at CMS with deep neural networks”, [CMS-DP-2017-005](#), (2017).
- [12] CMS Collaboration, “CMS Phase 1 heavy flavour identification performance and developments”, [CMS-DP-2017-013](#), (2017).