

CMS event reconstruction status in Run 2

Andrea Perrotta^{1,*} (on behalf of the CMS Collaboration)

¹INFN Sezione di Bologna, 40127 Bologna, Italy

Abstract. LHC Run2 began in April 2015 with the restart of the collisions in the CERN Large Hadron Collider. In the perspective of the offline event reconstruction, the most relevant detector updates appeared in 2017: the upgrade of the Pixel Detector, with the insertion of an additional layer closer to the beams, and the improved photodetectors and readout chips for the Endcap Hadron Calorimeter, which allow a finer longitudinal segmentation. The long shutdown between Run1 and Run2 was instrumental in the optimization of the reconstruction code and for the introduction of new algorithms to mitigate sensitivity to increased pileup, especially to out-of-time contribution with the advent of 25 ns separation between collisions compared to 50 ns in Run1. Such an optimization continued in the following years, when the reconstruction code of CMS evolved together with the improving of the performance of the LHC. The current status of the reconstruction software of the CMS experiment is described here, with emphasis on some of the recently integrated developments.

1 Introduction

The offline event reconstruction for the CMS experiment is based on the concept of Particle Flow (PF) [1]. Particles traversing the sensitive areas of the detector release an electric signal which is digitized by the acquisition system, and then locally reconstructed in the different subdetector units (Inner Tracker, Electromagnetic and Hadronic calorimeters, Muon detectors) to provide the building blocks for the subsequent overall event description. Those PF blocks are then considered globally, taking also possible superpositions of the signals into account, to reconstruct the particle-like final state objects which will be used in the analysis. The whole detector information is also exploited to provide particle identification. Therefore, the final objects which are made available for the analyses are particle flow objects like Jet, Photons, Electrons, charged and neutral Hadrons, Muons, Taus, as well as higher level observables like Vertices, b -Tag, Missing E_T .

Since the beginning of LHC Run2 in April 2015, the subdetectors of the CMS experiment started getting updated in view of the higher luminosities and harsher radiation environment as expected for the prosecution of Run2, or even for Run3 (expected to start after the second long LHC shutdown in 2020). The two CMS detector updates most relevant for the offline reconstruction were both integrated in 2017: the replacement of the Pixel Detector with an upgraded one which has an additional layer closer to the beams, which allows a more precise determination of the impact and other track parameters, and the replacement of photodetectors and readout chips in the Endcap Hadron Calorimeter with improved ones which provide a

*e-mail: andrea.perrotta@bo.infn.it

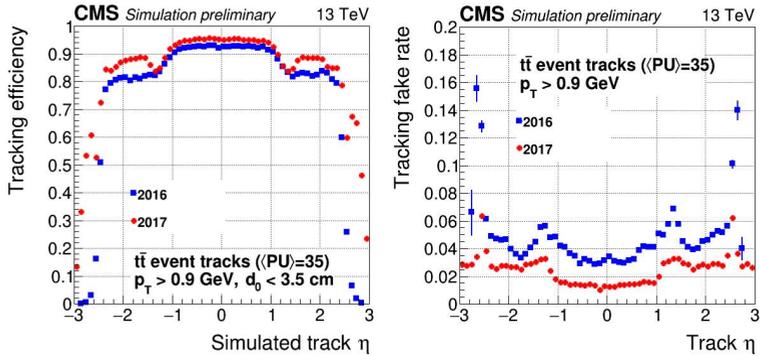


Figure 1. CMS Tracking efficiency (left) and fake rate (right) as function of pseudorapidity, comparing the performance of the tracking setups and algorithms of 2016 and 2017 [6].

finer longitudinal segmentation. The long shutdown between Run1 and Run2 was instrumental in the optimization of the reconstruction code and for the introduction of new algorithms to mitigate sensitivity to the increased pileup, especially to the out-of-time contribution which became quite relevant when the time separation between collisions was reduced to 25 ns from the previous 50 ns of Run1. Such an optimization continued in the following years, when the reconstruction code of CMS evolved together with the improving of the performance of the LHC.

In this contribution, the current status of the offline reconstruction software of the CMS experiment is described, together with its results and achievements in terms of physics objects. Emphasis will be given on some of the recently integrated developments. The computational performance of the offline reconstruction code of CMS is addressed in another contribution to this conference [2].

2 Tracking

The precise determination of the trajectory parameters of charged particles is instrumental for the characterization of the Particle Flow objects and the global event reconstruction. CMS upgraded its Pixel Detector in 2017 [3]: the insertion of an additional layer improves track impact parameter resolution and vertexing performance. In addition, the material budget in the endcaps was also reduced. Iterative tracking is adopted in CMS to deal with the large combinatorics [4]. It starts with very tight criteria for track seeds and reconstructs tracks by taking all available hits into account. Then, already used hits are cleared, seed criteria are loosened, and track reconstruction is repeated on the remaining unassociated hits.

Several improvements to the iterative tracking algorithm were possible thanks to the new Pixel detector. Among those improvements, the finding of the pixel seeds based on the Cellular Automaton (CA) technique [5] can be considered one of the most relevant: hit pairs are formed between detector layers; pair compatibility with respect to the interaction point is checked; hit triplets or quadruplets used for seeding are formed from compatible pairs. The improvements in the tracking efficiencies and in the fake rate obtained with the new detector setup and improved algorithms are shown in Fig. 1.

Timing performance improves with respect to the 2016 tracking algorithm thanks to the use of the CA seeding; a smaller fake rate also reduces the time spent in pattern recognition.

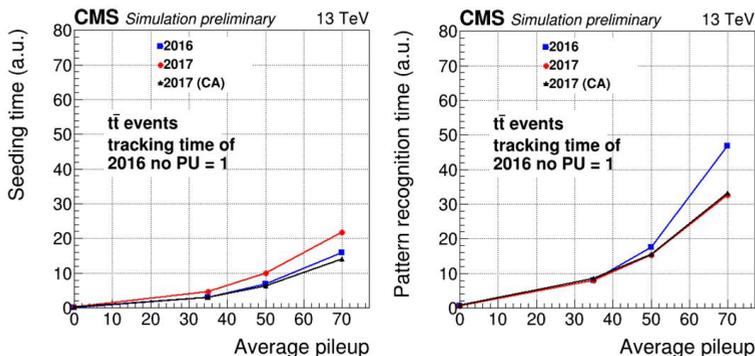


Figure 2. Comparisons of the time spent for the seeding (left) and pattern recognition (right) in the tracking algorithms used in 2016, and 2017 with and without Cellular Automaton used for the seeding [6].

This is even more relevant in the reconstruction of events with large pileup, as shown in Fig. 2.

3 Muons

Muon reconstruction and its performance during Run2 are described in [7]. Segments are reconstructed from local hits in the CMS muon detectors, and combined to build stand alone muon tracks. Inner tracks are also reconstructed in the Silicon Tracker. Global Muons are then obtained by merging stand alone and inner tracks (combined fit performed, p_T re-evaluated “outside-in”). Tracker Muons are identified “inside-out” by matching inner tracks with track segments locally reconstructed in the muon detectors. Ad-hoc algorithms specialized for high- p_T muons exist, with dedicated muon-ID and momentum refits which take the possible showering into account. Isolation deposits inside pre-defined isolation cones around the muons are also computed, either based on low-level detector quantities like calorimetric cells or on fully reconstructed particle flow objects.

4 b-Tagging

There are a few algorithms currently used for b -tag in CMS [8]. They are:

- CSVv2 (Combined Secondary Vertex): using an artificial neural network (NN) to combine track and vertex infos.
- DeepCSV: feeding a deep NN with the same variables as used in CSV, but with more tracks included.
- DeepFlavour: a deep NN algorithm based on 16 properties of up to 25 charged and 6 properties of 25 neutral PF jet constituents, as well as 17 properties from up to 4 secondary vertices associated with the jet.
- Double- b tagger: a dedicated algorithm for the identification of the decay of a boosted object to a b quark pair.

The insertion of the new Pixel detector and the implementation of the DeepCSV and DeepFlavour methods allow a major breakthrough in the b -tagging performance in CMS, as

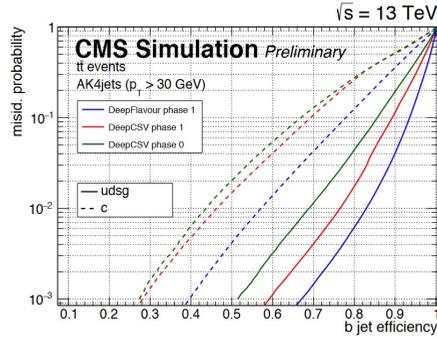


Figure 3. Comparisons of the b -tagging performance for the Phase0 DeepCSV algorithm, the Phase1 version of the same, and the Phase1 DeepFlavour algorithm [9].

can be seen in Fig. 3 for the simulation of the data collected in 2018, where the b -jet tagging efficiency is given as function of the mis-tagging probability for the previous DeepCSV used in Phase0 (i.e. Run 1) and for the new version of DeepCSV and for DeepFlavour, as used in Phase1 (i.e. Run 2).

5 Taus

The “hadrons-plus-strips” (HPS) algorithm is used to reconstruct and identify the hadronic decay products of tau leptons in various decay modes [10]. It works as follows:

- extract PF constituents (electrons, photons, muons, hadrons) from the reconstructed jets;
- reconstruct π^0 (as “Strips”);
- test all possible decay modes and assign the one with highest p_T and charge/strip multiplicity that pass the mass cut-window.

The probability for hadronic jets to be mis-tagged as tau jets is reduced by the use of isolation discriminators. Fig. 4 shows how the tau reconstruction efficiency and identification correlate with the misidentification probability while using a cut-based discriminator or an MVA-based discriminator, where the better performance of MVA based discriminator is evident.

6 Electrons and Photons

An ECAL- driven algorithm is adopted to reconstruct electrons and photons in CMS [12]. It starts from ECAL superclusters (group of one or more clusters of energy deposits in the Electromagnetic Calorimeter). Then for electrons, it uses superclusters to seed silicon tracker tracks; the Gaussian Sum Filter (GSF) tracking algorithm takes into account the momentum loss due to bremsstrahlung. Shower shape variables help separating single-core e/γ objects from multi-core π^0/η ones. Cut based or MVA based identifications are also implemented.

Improvements coming from the new reconstruction and identification algorithms and from the refined conditions are visible already in the re-reconstruction of the 2016 CMS data, as can be seen in Fig. 5 for the efficiency of the GSF tracking in the 2016 data when using the former reconstruction, or the re-reconstruction which implements the updated algorithm.

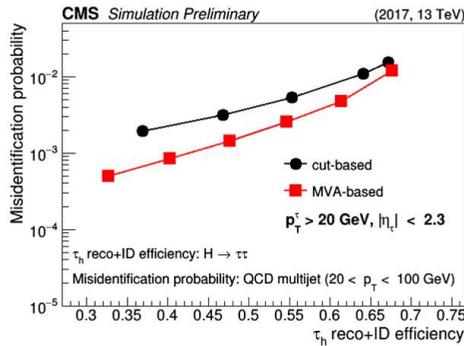


Figure 4. Comparisons of the tau reconstruction and identification performance for the cut based discriminator and the MVA based one [11].

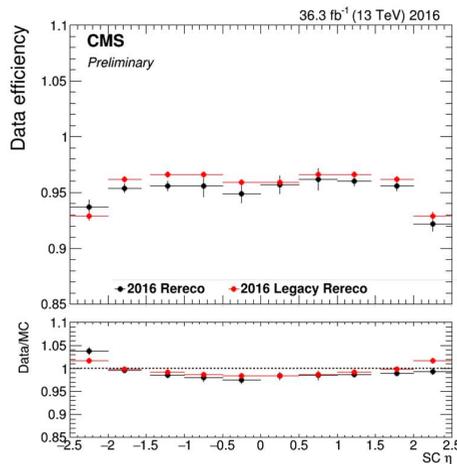


Figure 5. Improvements obtained in the GSF reconstruction efficiency as function of the pseudorapidity of the electron when using the new electron reconstruction on the 2016 data [13].

Electron/photon reconstruction and identification since 2017 take also advantage of the addition of the fourth silicon tracker layer, and the reduced overall material budget of the pixel detector. For a fixed electron ID efficiency, Fig. 6 shows the fake rates measured in Drell-Yan plus Jets events in 2016 and 2017 simulations. The improvement obtained with the new detector setup and reconstruction algorithm is evident, in particular for low p_T electrons in events with large pile-up. A big effort is also put on following the evolving data-taking and detector status and conditions.

7 Hadron Calorimeter

In early 2017, the readout electronics for one 20° wedge of the Hadronic Calorimeter (HCAL) endcap (HE) was upgraded with silicon photomultipliers (SiPMs) and QIE11 digitizers [14]. Based on the excellent experience, in 2018 the whole HE detector (36 wedges) has been upgraded. Besides the operational advantages of SiPMs with respect to HPDs (noise, gain

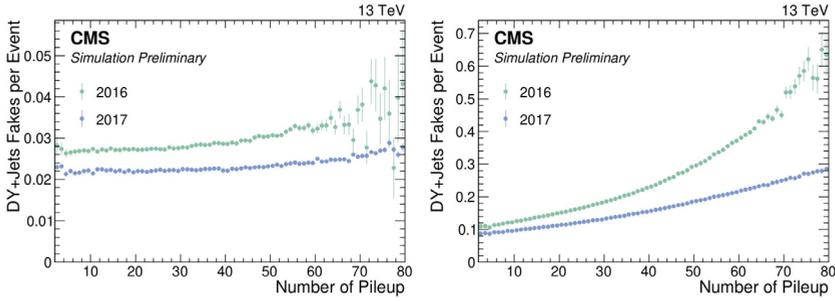


Figure 6. Fake rates measured in simulated Drell-Yan plus Jets events as function of the number of overlapped pile-up events, for high- p_T electrons (left) and low- p_T ones (right) in 2016 and 2017 CMS data [13].

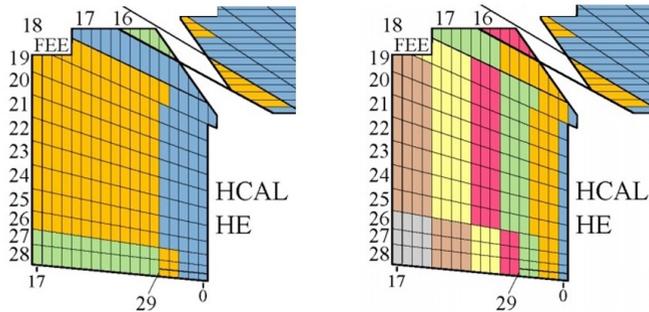


Figure 7. HCAL endcaps in 2016 (left) and the same with the current design of the readout (right).

stability) a finer longitudinal segmentation is now available for local reconstruction. This is visible in Fig. 7, where the past and current layout of a cross-section of the HE are displayed for comparison.

HB/HE out-of-time pileup mitigation in the HCAL local reconstruction is achieved by fits of up to three pulses in the bunch crossing of interest, the previous and following ones, and a flat baseline component. A non-negative least squares algorithm is used for optimization. The idea behind the fit is shown in Fig. 8 for both the barrel and the endcap HCAL modules: due to the signals added by pileup events collected in the nearby LHC bunch crossing, the in-time pulse overlaps with several out-of-time ones, whose contribution can be disentangled by fitting the expected shapes centered in the central and nearby bunch crossing times.

8 Conclusions

Before and during LHC Run2 the CMS detector was partially upgraded in order to better comply with the high radiation and instantaneous luminosity of the new run, and to improve at the same time its performance in terms of reconstruction and identification of physics objects. The offline event reconstruction of CMS, aimed at a global event description based on the Particle Flow concept, is taking advantage of all those detector upgrades. Moreover, new techniques and improved algorithms were also implemented to improve the performance

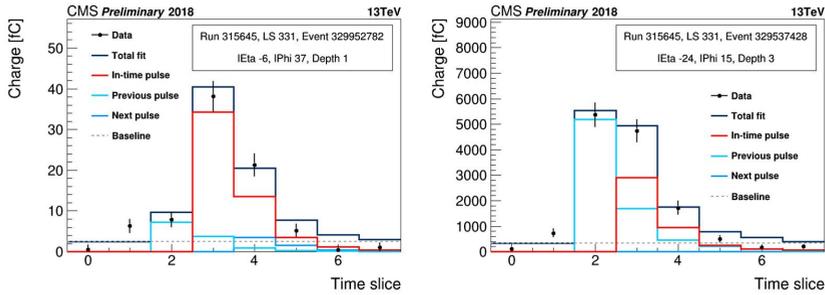


Figure 8. Charge distribution in the different time slices for the HCAL barrel (left) and endcap (right), showing separately the contribution of the signal coming from the different bunch crossings [15]. Difference in charge scale are due to differences in photo-detection efficiency and amplification.

of the reconstruction, both in terms of physics outputs and computational outcomes (cpu timing, memory consumption, usage of storage resources).

References

- [1] CMS Collaboration, *Particle-flow reconstruction and global event description with the CMS detector*, JINST **12**, P10003 (2017)
- [2] S. Krutelyov, *Performance in offline event reconstruction landscape at CMS*, these proceedings
- [3] CMS Collaboration, *CMS Technical Design Report for the Pixel Detector Upgrade*, CMS-TDR-011, CERN-LHCC-2012-016 (2012)
- [4] CMS Collaboration, *Description and performance of track and primary-vertex reconstruction with the CMS tracker*, JINST **9**, P10009 (2014)
- [5] D Funke et al, *Parallel track reconstruction in CMS using the cellular automaton approach*, J. Phys.: Conf. Ser. **513** 052010 (2014)
- [6] CMS Collaboration, *2017 Tracking performance plots*, CMS -DP -2017/015, <https://cds.cern.ch/record/2290524> (2017)
- [7] CMS collaboration, *Performance of the CMS muon detector reconstruction with proton-proton collisions at 13 TeV*, JINST **13**, P06015 (2018)
- [8] CMS Collaboration, *Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV*, JINST **13**, P05011 (2018)
- [9] CMS Collaboration, *Performance of b tagging algorithms in proton-proton collisions at 13 TeV with Phase 1 CMS detector*, CMS-DP-2018/033, <https://cds.cern.ch/record/2627468> (2018)
- [10] CMS Collaboration, *Reconstruction and identification of τ lepton decays to hadrons and ν_τ at CMS*, JINST **11**, P01019 (2016)
- [11] CMS Collaboration, *Tau Identification Performance in 2017 Data at $\sqrt{s}=13$ TeV*, CMS-DP-2018/026, <https://cds.cern.ch/record/2622155> (2018)
- [12] CMS Collaboration, *Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV*, JINST **10**, P06005 (2015)
- [13] CMS Collaboration, *Electron and Photon performance in CMS with the full 2017 data sample and additional 2016 highlights for the CALOR 2018 Conference*, CMS-DP-2018/017, <https://cds.cern.ch/record/2320638> (2018)

- [14] CMS Collaboration, *CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter*, CMS-TDR-010, CERN-LHCC-2012-015 (2012)
- [15] CMS Collaboration, *HCAL Out Of Time Pileup Subtraction and Energy Reconstruction*, CMS-DP-2018/018, <https://cds.cern.ch/record/2320408> (2018)