PROCEEDINGS OF SCIENCE

CMS b-Tagging and Tracking Commissioning in Cosmics

Jean-Roch Vlimant[*†]
University Of California, Santa Barbara
E-mail: vlimant@cern.ch

The construction of the CMS experiment was completed in 2008, and began taking data, recording particles coming from cosmic showers, in 2008 and 2009. We describe how charged particles are reconstructed with the CMS inner tracking system. Performance from simulation is compared to that observed in cosmic data, showing that startup conditions are very close to the ideal design. Jet flavor tagging at CMS is presented, along with simulated performance, expected performance at startup and plans for commissioning of jet b-tagging.

XXth Hadron Collider Physics Symposium
November 16 – 20, 2009
Evian, France

[*] Speaker.
[†] For the CMS collaboration. Special thanks to Kevin Burkett.

PoS(HCP2009)021

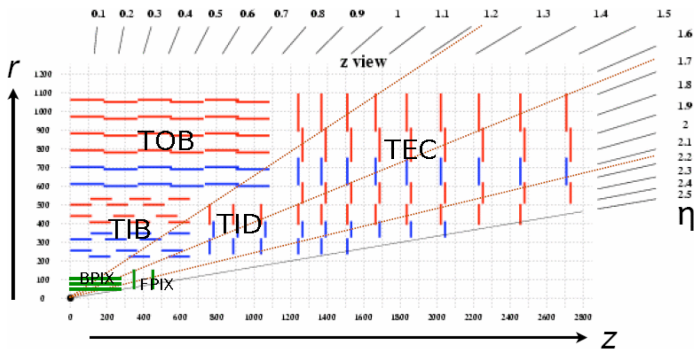


Figure 1: Layout of the layers of silicon sensors in the CMS inner tracking system. Single sided modules (red), stereo modules (blue) and pixel modules (green) are grouped in the different tracker sub-systems. Tracker inner barrel (TIB), tracker outer barrel (TOB), tracker inner disk (TID), tracker endcap (TEC) are labeled on this quadrant representation.

1. The CMS Inner Tracker

The CMS experiment is an all-purpose detector [1] which consists of different sub-detector not detailed in this paper. The CMS inner tracking system consists of a silicon strip tracker and a silicon pixel detector (see fig. 1) placed within a superconducting solenoid. The magnetic field operates at 3.8 Tesla over the tracking volume. The pixel detector has 3 barrel layers and 2 end cap disks on each side, with a total of about 66 million channels from pixels of size $100\times150\ \mu m^2$. It provides precise 3D measurement points close to the interaction point, useful for reconstruction of primary and secondary decay vertices, as well as very low pT track reconstruction. The strip tracker has 10 barrel layers and 12 end cap disks on each side, with a total of about 9.3 million channels from silicon strips with pitch ranges from 80 to 150 μm. 4 of the barrel layers and 3 rings of the end cap disks have double sided modules mounted with a stereo angle to provide 3D measurement points. The rest of the modules provide 2D measurement points. The inner tracking system [1] therefore provides a few (up to about 30 with double sided and module overlap) but precise measurement points. An inconvenience of the CMS silicon tracker is the rather large, for a tracking device, amount of dead material involved for support, cooling and cabling [1].

2. The CMS Tracking Software

The CMS collaboration has developed several tracking algorithms, some for general track reconstruction and some for the reconstruction of tracks in specific cases (not detailed here). The main algorithm in the CMS reconstruction is the combinatorial track finder (CTF) used in iterative steps. For simple events like cosmic events, with mainly a muon track in the detector, a dedicated algorithm [2] has been designed to perform brute force single track pattern recognition (called cosmic TF). Results from this algorithm are compared to results from the main algorithm in section 4.2.

Combinatorial Track Finder. The CTF algorithm consists of 3 stages • **Seeding.** Pair or triplet of hits are used in combinatorics, using compatibility with the interaction region or primary

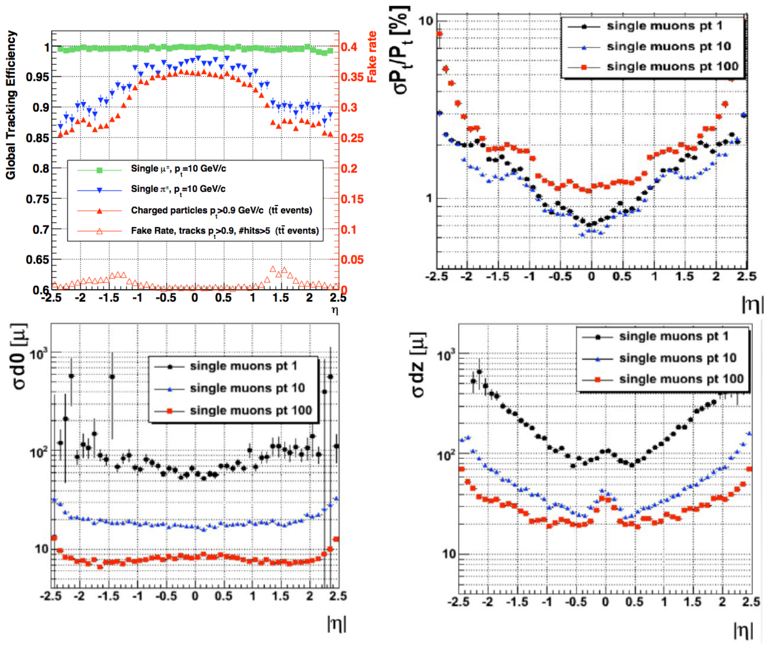


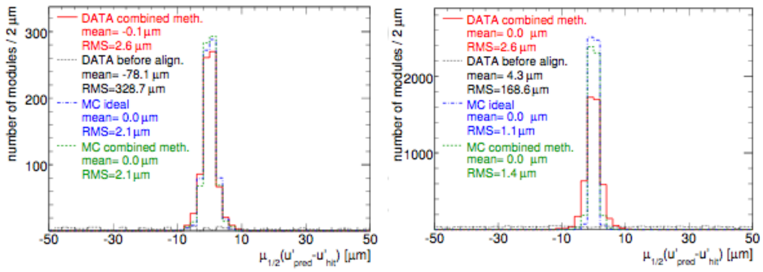
Figure 2: CMS track reconstruction efficiency and fake rate (top left) for different types of particles and transverse momentum resolution (top right) for muon tracks of 1, 10, 100 GeV of transverse momentum. Both plots are as a function of pseudo rapidity. Transverse (bottom left) and longitudinal (bottom right) impact parameter resolution for muon track of 1, 10, 100 GeV of transverse momentum, as a function of pseudo rapidity.

vertex depending on the iterative step. • **Pattern recognition.** From the initial track parameters of the seeds, hits are searched for in reachable layers, and added, using the Kalman update, to the trajectory candidate based on the chi squared compatibility between the hit and the predicted track positions. • **Final fit.** The collection of hits found in pattern recognition is refitted to provide the best track parameter estimation at each measurement point.

Iterative Tracking. In order to limit the combinatorics, while increasing the track parameter phase space reach, after completion of the first iteration of CTF (the 3 stages described above), hits belonging to high quality tracks are removed from the hits available to be used in finding tracks in the next steps. Additional iterations of the CTF are then performed (6 in total).

3. Expected Performance from Simulation

The tracking efficiency, the resolution on transverse momentum, and on transverse and longitudinal impact parameters are measured from simulation of various proton collisions final states (see figure 2). The efficiency to reconstruct muon tracks (10 GeV pT) is very close to 100%, while it is a bit lower for pions (10 GeV pT) because of a larger chance of nuclear interaction resulting in too few hits belonging to the track. The track efficiency for pions decreases in the end-cap because of the large amount of material within the pixel detector volume. The probability to reconstruct a fake track in a top pair production event (used as a benchmark for crowded events) is low but increases in the region of transition between barrel and disk where the amount of crossed material



RMS of DMR (µm)		# of aligned modules
BPIX (x)	2.5	761/768
BPIX (y)	4	
FPIX (x)	13	539/672
FPIX (y)	13	
TIB	3	2555/2724
TOB	3	5102/5208
TID	4	808/816
TEC	8	6346/6400

Figure 3: Distribution of the mean of residuals on the expected hit position in the transverse local coordinate for pixel barrel (left) and tracker outer barrel (middle). Table of the RMS of the distribution of median of residuals and number of modules for all sub-parts of the inner tracker (right).

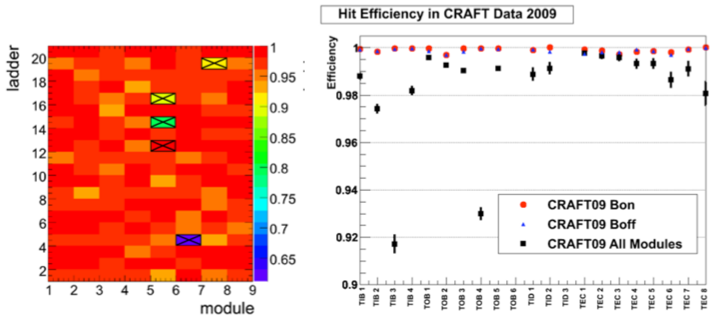


Figure 4: Single hit efficiency in the pixel (left) detector as a function of detector coordinate in the first barrel layer (crossed modules are known to have hardware problems) and in the strip (right) detector as a function of the layer for data taken with the magnetic field off and on, and adding modules with known issues.

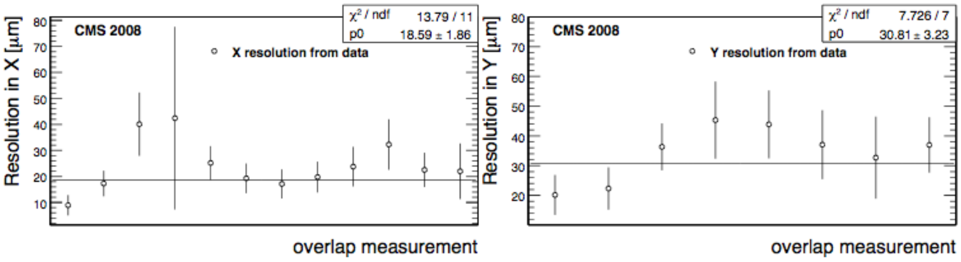
is maximum, increasing the chance of nuclear interaction and conversions, and hence increasing the combinatorics. The transverse momentum of muons is measured to the percent level up to 100 GeV. This resolution degrades gradually due to the change of strip pitch in the TEC, and degrades even more in the forward region because of the smaller lever arm when the track exits through the last layer of the TEC. The good coverage of the pixel detector allows measurement of a transverse impact parameter with a resolution of 10 to 100 microns, depending on transverse momentum. The longitudinal impact parameter is measured with slightly worse resolution, 20 to 600 microns.

4. Observed Performance with Cosmic Data

In 2008 and 2009, the CMS experiment has been taking cosmic data for use in commissioning [4]. Two campaigns of stable data taking with the solenoid powered on were performed, during each of which about 10^7 tracks passing through the tracker were recorded.

4.1 Alignment and single hit performance

The alignment of the CMS inner tracker was performed using cosmic data [5] and compared with alignment in simulated data (see fig 3). The precision of the alignment obtained is very



Sensor	Pitch (μm)	Resolution (μm)	Track angle $0^\circ - 10^\circ$	$10^\circ - 20^\circ$	$20^\circ - 30^\circ$	$30^\circ - 40^\circ$
TIB 1-2	80	Measurement	17.2 ± 1.9	14.3 ± 2.3	17.4 ± 3.2	25.7 ± 6.0
		MC Prediction	16.6 ± 0.5	11.8 ± 0.5	12.4 ± 0.6	17.9 ± 1.5
TIB 3-4	120	Measurement	27.7 ± 3.6	18.5 ± 3.1	16.1 ± 3.1	24.1 ± 6.7
		MC Prediction	26.8 ± 0.7	19.4 ± 0.8	17.2 ± 0.3	21.4 ± 2.0
TOB 1-4	183	Measurement	39.6 ± 5.7	28.0 ± 5.8	24.8 ± 6.5	32.8 ± 8.3
		MC Prediction	39.4 ± 1.3	27.8 ± 1.2	26.5 ± 0.3	32.5 ± 2.1
TOB 5-6	122	Measurement	23.2 ± 3.6	19.5 ± 3.6	20.9 ± 6.1	29.3 ± 9.7
		MC Prediction	23.8 ± 0.9	18.0 ± 0.5	19.2 ± 1.2	25.4 ± 1.6

Figure 5: Barrel pixel single hit resolution observed in cosmic data in the direction transverse (top left) and along (top right) to the beam pipe. The resolution expected from simulation is 22 μm and 28 μm in X and Y respectively. At the bottom, table of barrel strip single hit resolution observed in cosmic data and simulation, as a function of the angle with respect to the normal to the module surface. The resolution is optimum for median normal angles for which charge sharing is optimum [check that again please].

close to the ideal geometry of the tracker. Single hit efficiency is evaluated from cosmic data [6, 2] by propagating tracks to all reachable modules of the tracker, and checking that there is a corresponding hit on the module. To avoid edge effects, the propagated state is required to be well within the module boundaries. The single hit efficiency (see figure 4) is around 97% and 99% in the pixel and strip detectors, respectively. The inefficiencies are traceable to modules with hardware issues, for which repairs are underway. Single hit resolution was evaluated in the pixel [6] and strip trackers [2] using the "double difference method". In regions where two hits per layer are expected, the RMS of the difference between the difference of hit position and the difference of track predicted position contains a component from the track position uncertainty and a component from the single hit position resolution. Single hit position resolution is extracted assuming the track position uncertainty from the track fit itself. Results are reported in figure 5, showing good agreement with simulation. Performance is close to design [1].

4.2 Track Reconstruction Efficiency and Parameter Resolution

Several methods were used to estimate the track reconstruction efficiency in CMS using cosmic data [2], also comparing results from two different algorithms for cross checks. Results are presented in figure 6, showing overall good agreement with simulation and very good track reconstruction efficiency. • **Muon Matching Method:** Track reconstruction can be performed independently in the inner tracking system and in the muon system. The stand alone muon track reconstruction in the muon system is used to probe the reconstruction in the inner tracking system. • **Top-Bottom Tracking Method:** Cosmic track reconstruction is performed independently in both hemispheres of the tracker, and tracks from the top are used to probe the presence of corresponding

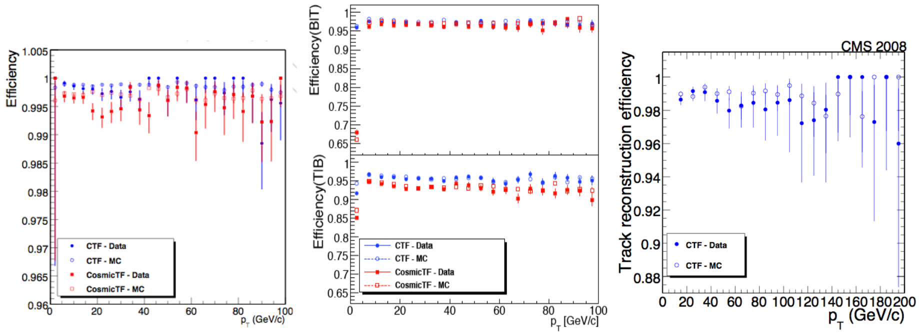


Figure 6: Cosmic track reconstruction efficiency, observed in cosmic data and simulation, as a function of track transverse momentum, for the CTF and the single track pattern algorithm, using the muon matching (left), top-bottom (middle) and collision like reconstruction (right) methods. Efficiency for the top-bottom method is given separately for the bottom (middle top) and the top (middle bottom).

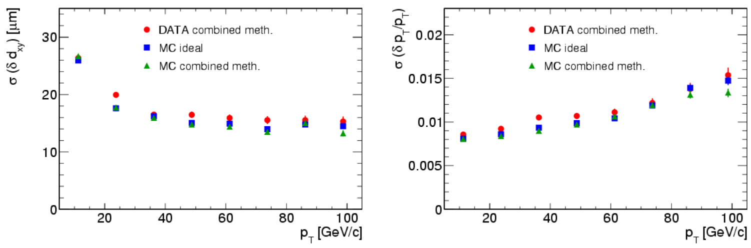


Figure 7: Transverse impact parameter (left) and momentum (right) resolution on cosmic track reconstruction as a function of transverse momentum resolution observed in cosmic data and simulation.

tracks in the bottom (and vice-versa). • **Collision Like Reconstruction Method:** Collision like tracking can be run in cosmic events and reconstruct the "upper" and "lower" parts of a cosmic track crossing the tracker as two independent tracks. Each track is used to probe the presence of a corresponding track on the other side.

The track of a cosmic muon crossing the tracker is split into two tracks at the point of closest approach to the beam line. The two tracks obtained therefore measure the same track parameters, and the width of residual (divided by $\sqrt{2}$) gives an estimation of the resolution of track reconstruction (see figure 7). The resolution on transverse momentum and transverse impact parameter in cosmic data is observed to be in good agreement with simulation and very close to the design resolution.

5. Jet b-Tagging at CMS

Except for the top quark, quarks and gluons produced at vertices hadronise into colorless jets of hadronic particles. The flavor of the initial quark at the origin of a jet can be identified with characteristics of the jet using tagging algorithms. Due to its intrinsic lifetime and how it decays semi-leptonically, b-jets can be distinguished from light quark jets. Identification of b-jets is fundamental to many physics analyses for improving the signal to noise ratio.

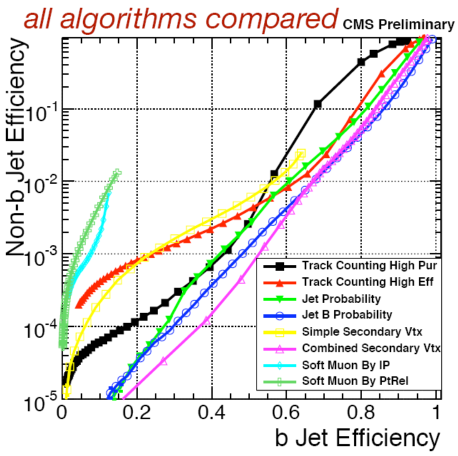


Figure 8: Probability of mistagging a light-quark jet as a b-jet versus the efficiency of tagging a b-jet as such, plotted for several b-tagging algorithms in the CMS reconstruction (not all described in the text, see [7] for details).

5.1 CMS Tagging Algorithms

The algorithms described below are common to the high energy physics community, but have been refined and extended for use in CMS [7]. They are described in order of complexity. • **Track Counting:** Requiring n tracks within a jet to have large impact parameter significance is a robust way to tag a b-jet. With two tracks one gets high efficiency, while three tracks provides high purity. • **Track Probability:** Calculating the combined probability that each track within the jet is compatible with the primary vertex provides a handle on discrimination of b-jets from other jets. • **Soft Lepton Tagging:** The presence of a lepton within a jet is an indication of the b-flavor of a jet. The momentum of the lepton relative to the jet can be used alone or combined with other information into a neural net to discriminate light flavor jets from b-jets. • **Secondary Vertex Tagging:** Secondary vertices within jets can be reconstructed. The distance significance between the primary and secondary vertices is a simple method of b-tagging. Combining information from the secondary vertex into a likelihood ratio is a bit more complex but provides better performance (see figure 8).

5.2 Expected Performance and Commissioning with Collision Data

The performance of b-tagging expected from simulation [7] is summarized in figure 8. Simple and robust algorithms like track counting have good expected performance that is just a bit worse than complex methods like secondary vertex tagging. As b-tagging methods are highly dependent on efficient tracking with good resolution, it is expected that the observed b-tagging performance will agree with that from simulation.

As soon as enough collision data with jet activity is recorded, the performance of b-tagging algorithms can be measured with several methods. • **Negative Tag Sample** [8]: A sign can be given to the transverse impact parameter of a track within a jet with respect to the jet axis. Tracks with negatively signed impact parameters are likely to come from resolution effects on tracks from the primary vertex of the jet. Therefore, a sample of jets constructed with negative impact parameter is enriched with light-flavor jets, and provides a simple way to estimate the mis-tag rate. • **Muon-in-Jet sample** [9]: A sample of jets constructed with the request of a muon being present is enriched with b-jets. A template fit applied based on the transverse momentum of the muon relative to the jet (p_T^{rel}), before and after the application of the tagger, allows a determination of the b-tagging efficiency. The so-called "system 8" method uses two stages of b-tagging with loose and tight requirements to solve a system of equations providing an estimation of tagging efficiency. • **Top Pair Production Sample** [10]: Due to the decays properties of the top quark, a sample selected on the signature of at least a leptonic decay of a W boson, with the addition of a b-tagged jet and a kinematic fit of the final state objects to top pair decay products, is enriched by at least one unbiased b-jet which can be used to estimate the b-tagging efficiency.

6. Conclusions

Track reconstruction with the CMS inner tracking system has been presented. The tracker performance observed during cosmic data taking has been shown to be close to the design performance, and in good agreement with simulation. CMS has a variety of algorithms to determine the b-flavor of the quark from which a jet originates. Methods to evaluate b-tagging performance with LHC collision data have been presented. The performance of b-tagging with collision data is expected to be close to that of the design performance.

References

[1] The CMS Collaboration, The CMS experiment at the CERN LHC, JINST 3 (2008) S08004.489.

[2] The CMS Collaboration, The CMS Silicon Strip Tracker Operation and Performance With Cosmic Rays in 3.8 T Magnetic Field, submitted to JINST (2009).

[3] The CMS Collaboration, Track Reconstruction in the CMS Tracker, CMS Note 2005/14.

[4] The CMS Collaboration, The CMS CRAFT Exercise, submitted to JINST (2009).

[5] The CMS Collaboration, Alignment of the CMS Silicon Tracker During Commissioning with Cosmic Ray Particles, JINST 2009.

[6] The CMS Collaboration, Performance of the CMS Pixel Tracker with Cosmic Rays, submitted to JINST (2009).

[7] The CMS Collaboration, Algorithms for b Jet identification in CMS, BTV-09-001.

[8] The CMS Collaboration, Evaluation of udsg Mistags for b-tagging using Negative Tags, BTV-07-002.

[9] The CMS Collaboration, Performance Measurement of b tagging Algorithms Using data containing muons within Jets, BTV-07-001.

[10] S. Lowette, J. DÕHondt, J.Heyninck, P. Vanlaer, Offline Calibration of b-Jet Identification efficiencies, CMS Note 2006/013.

PoS(HCP2009)021