CMS Conference Report

January 12, 2006

Tracking and b and τ tagging in the CMS high level trigger

Marcel Vos, for the CMS collaboration

INFN - sezione di Pisa, Largo Pontecorvo 3, Pisa, Italy

Abstract

The CMS trigger stands for the daunting task of selecting rare signal processes amidst the 40 Million bunch crossings. While the tracker information is not available in the first level of the trigger, it plays a crucial role in the high level trigger. In this contribution, the τ and b identification performance in the trigger is discussed. Hadronic τ decay can efficiently be reconstructed and selected using tracker information. An inclusive b-jet trigger based on a lifetime tag allows the jet thresholds to be significantly reduced, thus providing a very much improved trigger efficiency for hadronic final with b-quarks, like hadronically decaying top pairs.

Presented at *Tracking in High Multiplicity Environments*, Zurich, October 3-7, 2005

Submitted to Nucl. Instr. Meth. A

1 Introduction

The beam energy and luminosity of the Large Hadron Collider will give rise to an unprecedented QCD production rate. The online selection of rare signal events among the background will be even more important than it has been in previous experiments.

The CMS trigger is composed of two distinct levels. The first stage, known as first level or L1, is entirely implemented in custom hardware. It performs a rapid (latency of the order of 3 μs) decision on the basis of information from the calorimeters and muon chambers. In the second stage, the High Level Trigger or HLT, the complete event information - including that from the tracker - has been assembled. HLT decision algorithms are implemented in software and run on a large filter farm.

The tracker information plays a crucial role in the High Level Trigger. The initial momentum measurement of the muon is refined by connecting to the measurements in the internal tracking volume. Photons and electrons are distinguished by reconstructing the tracks in the region indicated by the electromagnetic calorimeter deposit. Use of the tracker information yields the required rejection, while maintaining the p_T threshold for muons and electrons close to those required at the first level. A very complete description of the CMS high level trigger algorithms is found in reference [1].

While final states with one or more isolated leptons (electrons, muons) of moderate momentum are efficiently selected, the trigger for purely hadronic final states is considerably more challenging, as the large background rate drives up the thresholds. Hadronic final states with τ of b decays can be reconstructed and identified (“tagged”) on the basis of tracker information.

In this contribution, the online b- and τ-tagging algorithms implemented in the CMS High Level Trigger are described.

2 Track reconstruction in the HLT environnent

In the CMS HLT environment, algorithms of arbitrary complexity can in principle be implemented. The most severe constraint is posed by the available CPU time. Several techniques are employed in the CMS HLT to reconstruct high-quality tracks at a minimum computing load.

A considerable speed-up is obtained by a regional application of the algorithm. Typically, regions-of-interest are defined on the basis of the result of the preceding trigger level (a calorimetric cluster, the extrapolation of a track in the muon chambers, etc.) Seeds are reconstructed only in a multi-dimensional region (constraining the η ϕ direction of the tracks, the origin in the $R\phi$ and Rz planes, and the minimum transverse momentum). Trajectory building and fitting then proceeds in the same fashion as for global tracking.

The time performance of the regional approach allows the default offline algorithm - the combinatorial Kalman filter based track finder discussed in another section of these proceedings [2] - to be used in the later stages of the high level trigger. The track reconstruction thus obtained is of comparable quality as the offline reconstruction. Only at the very edges of the region parameters is a minor degradation of the efficiency observed (due to the limited precision of seeds). The fake rate can be controlled to the level of 1 % by an adequate choice of track quality requirements.

For the earlier stages of the high level trigger, speed becomes even more of a concern. An extremely fast reconstruction based on hit triplets in the pixel detector has been developed. The algorithm is described in detail in reference [4]. An overview of the performance is given in these proceedings [2]. Pixel-only reconstruction is sufficiently fast that global reconstruction of all tracks with transverse momentum greater than 1 GeV can be performed in the high level trigger [1)]

The simplified pattern recognition has to rely on three hits out of three pixel layers, thus posing a severe requirement on the single layer efficiency. After detailed simulations of all efficiency loss mechanisms [5], the effect on the efficiency is expected to be quite limited, but of course the robustness of the algorithm against defective components is much reduced. The fake rate is rather well controlled by the three-out-of-three requirement (to the level of 10 %, see reference [2]) and can be further improved by requiring compatibility with the primary vertex.

The track parameter resolution suffers from the small lever arm (the pixel barrel layers are located at radii of 4, 7

[1)] The CPU requirements measured include only the pure algorithmic part. Loading and unpacking times of the detector data, and cluster reconstruction have to be evaluated separately.

and 10 cm). For a 10 GeV track, the resolution of the full tracker is better than 1 %, whereas the pixel-only tracks have a resolution of approximately 25 %. This uncertainty affects the precision of the impact parameter estimate, leading to a signficant degradation of the pixel-only d_0 resolution with respect to the full-tracker performance. In the longitudinal plane, the resolution is much less affected. Finally, it has been shown that sufficiently accurate error estimates are possible with the simplified track fit of the pixel-only reconstruction [3].

In online conditions, the robustness of the algorithms against changing conditions becomes especially important. The b-tagging algorithms based on lifetime require the 3-dimensional position of the primary vertex to be accurately known. The beam is expected to be sufficiently narrow that in the transverse plane the beam position can be taken as the vertex position without significant loss of precision. The beam position is expected to be sufficiently stable during a fill of the LHC (O (10) hours), but has to be determined at the start of each fill. A study shows that 500 minimum bias events are sufficient to reach the required precision [3]. The z-coordinate (along the beam line) of the primary vertex has to be reconstructed for each event.

3 Online τ identification

The τ lepton offers a great physics perspective at the LHC, the most well-known being that of the detection of the MSSM Higgs bosons through its decay into pairs of τ. Whenever the τ decays into a muon or electron plus a neutrino, the event is triggered in the standard muon and electron streams. The large majority of τ, however, decay hadronically. Such τ jets are identified in the CMS trigger using a combination of calorimeter and tracker information.

The selection of τ jet events at the first trigger level (L1) is entirely based on the calorimeter response. As τ jets tend to be much more collimated than a quark or gluon jet of comparable transverse energy, τ jets are efficiently selected requiring the energy deposition in towers containing ECAL and HCAL cells to follow certain narrow patterns. An E_T threshold 93 GeV (66 GeV) for events with a single (double) τ-jets is applied.

In the High Level Trigger the reconstruction of the calorimetric information is repeated with the full granularity. A modest background rejection of a factor 3 can be achieved by an isolation criterion based on the comparison of the ECAL transverse energy deposit in a narrow ($\Delta R < 0.13$)and a large ($\Delta R < 0.4$) cone.

For confirmed candidates the tracker information in a narrow region around the candidate is accessed. The selection is based on the comparison of the number of tracks contained in a narrow signal cone (typically $\Delta R < 0.07$) and a much larger isolation cone (typically in the range $0.2 \Delta R < 0.4$). The direction of the signal and isolation cones are defined by that of the leading track within $\Delta R < 0.1$ of the direction of the jet.

For isolated τ jets, one expects one or three stiff tracks inside the signal cone and none in the isolation. The leading track in quark or gluon jets is generally much softer and tracks tend to spill over into the isolation cone. The selection applied in the HLT requires at least one track with transverse energy over a certain threshold

In the CMS HLT the τ tagging algorithm [7] may be applied on pixel-only tracks or on tracks reconstructed in the full tracker. The pixel-only algorithm is much faster. Tracks reconstructed in the full tracker have a superior transverse energy resolution and thus allow a tight requirement on the leading track (typically 20 GeV) to be defined. The full-tracker τ selection thus yields an efficiency of 45-50 % for τ jets from the decay of Higgs bosons with a mass from 200 to 500 GeV at a di-jet event rejection of a factor 1000.

The complementary nature of the pixel-only and full-tracker approaches allows to choose the one that fits best for application. Thus, the single τ jet trigger relies on the full-tracker reconstruction to reach the required rejection on a single object, while for the double τ jet trigger stream - with a much more relaxed selection on each of the objects - benefits from the superior time performance of the pixel-only algorithm.

4 Online b tagging

Many interesting fully hadronic final states contain one or several b jets. Examples are hadronic decays of bbH, SUSY decays, etc. The selection of such final states can greatly benefit from the experimental sensitivity for the presence of b jets. Efficient b-tagging allows the di-jet background to be significantly reduced. In the trigger, the additional rejection obtained from the b-tag requirement can be used to lower the thresholds on the N-jet triggers.

The copious $b\bar{b}$ production at the LHC limits the rejection that can be obtained from b-tagging to roughly a factor 20, a small factor compared to the rejection of 1000 needed to reduce the L1 to the HLT output rate. Rejections beyond this figure are in principle possible, but inevatibly yield very low signal efficiency. In the CMS HLT, b-tagging information is combined with pure jet trigger to create an inclusive b-jets trigger.

In the first stage, jets are reconstructed based on calorimeter information. A threshold is applied on the transverse energy of the first, third and fourth jet of 350, 15- and 55 GeV, respectively. This first step is intended to yield a fast rejection of a factor 50.

For the remaining events, the tracker data are accessed. To optimize the time performance, a fast preselection is applied using pixel-only tracks. Global reconstruction of triplets of pixel hits yields an accurate measurement of the primary vertex: for the relevant event topologies the efficiency to reconstruct and identify the primary vertex is better than 97 % [3, 4].

The b-tagging selection [6] requires at least two pixel-only tracks of good quality with a signed 3D impact parameter significance greater than 2.5 in a cone around the jet. This criterion is applied to the two leading jets in the event. For events in which at least one jet is tagged, seeds are generated in a region around the tagged jet. These seeds are followed through the full tracker using the (offline) Kalman-filter based track finder. On these high-quality tracks a much tighter b-tag is required: 3 tracks with a signed 3D impact parameter significance greater than 3. While this last step is rather time-consuming (300 ms/jet), it is only applied at a very limited rate (several tens of Hz), thus yielding an acceptable CPU requirement for the average HLT event.

The b-tagging performance of the pixel-only and full-tracker requirements is shown in figure 1. Clearly, the degraded resolution of pixel-only tracks with respect to that of the full tracker leads to a reduced tagging power. Indeed, for a tagging efficiency of 50 % on the leading jets in hadronic $t\bar{t}$ events the full-tracker algorithm easily limits the contamination of the tagged sample by light jets to 1-2 %, whereas for the pixel-only algorithm a purity of 7 % is obtained. For larger tagging efficiency, however, the pixel-only algorithm performs very well: for 70 % efficiency, a rejection of a factor 5 is found, quite similar to that obtained with the full tracker. Moreover, the tagging results are extremely correlated: application of both algorithms in series - with relatively loose cuts in the pixel-only algorithm and a tight selection in the final step - yields an efficiency that is identical within 2 % to that of the accurate full-tracker algorithm alone. Thus, a very fast rejection is achieved of 80 % of events, with virtually no loss in efficiency.

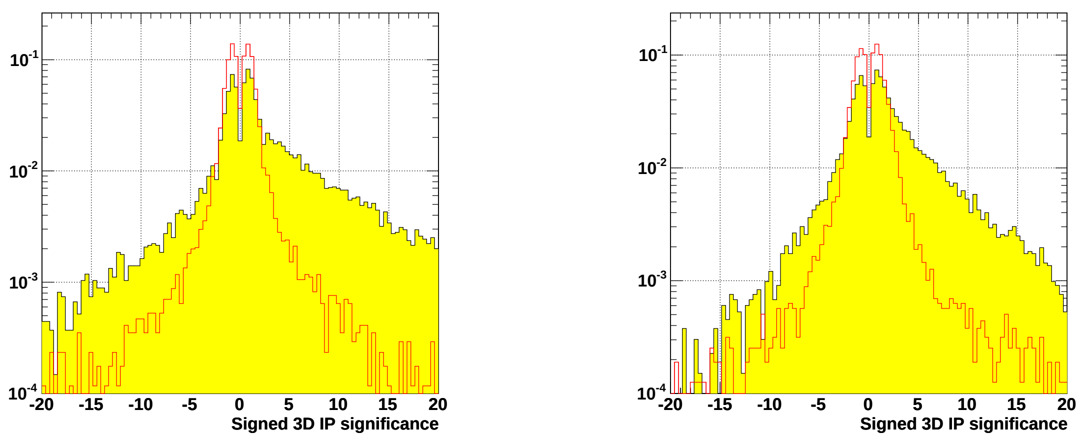

Figure 1: A comparison of the 3D impact parameter distribution for regional reconstruction (leftmost figure) and pixel-only reconstruction (rightmost figure). The 3D impact parameter distribution for tracks in b jets (filled histogram) and in light jets (neither b nor c, empty histogram) is shown. The results are obtained on reference (b-enriched) di-jet samples with jet E_T from 50-80 GeV. The jet flavour is determined from Monte Carlo truth information by associating the jet to the heaviest parton (after FSR) in an association cone of $\Delta R = 0.5$.

The chain described above yields a total trigger rate of 16.7 Hz. As the b-jet trigger is highly correlated to some of the other triggers (mostly the inclusive jets trigger), about half of these events trigger at least one other stream. The trigger efficiency (including the L1 efficiency) for fully hadronic $t\bar{t}$ events is 15 %, where the inclusive jets trigger without b-tagging information reaches only 4.3 %.

5 Conclusions

The CMS high level trigger relies heavily on tracker information to efficiently select a range of objects in the final state. Two algorithms for fast track reconstruction are available: a fast pixel-only algorithm allows for fast rejection in the early stages, while regional application of the combinatorial Kalman filter track finder yields a high-quality track collection in the later stages. Track-based τ and b-tagging algorithms have been implemented in the HLT. The tagging efficiency and purity for τ- and b-jets is evaluated in detail. Several physics channels - good examples are MSSM Higgs bosons decaying into a pair of τ leptons and fully hadronic top decays - thus attain a significant potential.

This work was supported by the European Union under contract number HPRN-CT-2002-00326

References

[1] The CMS collaboration. The Trigger and Data Acquisition project, volume II: Data Acquisition & High-Level Trigger, Technical Design Report, CERN/LHCC 2002-26

[2] T. Todorov, these proceedings.

[3] M.A. Vos and F. Palla, HLT b-tagging, CMS note in preparation

[4] S. Cucciarelli, M. Konecki, D. Kotliński, T. Todorov, Track parameter evaluation and primary vertex finding with the pixel detector, CMS Note 2003/026

[5] S. Cucciarelli, D. Kotlinski, T. Todorov, Position determination of the pixel hits, CMS Note 2002/049 D. Kotlinski, Pixel Detector data rates and inefficiencies, CMS Internal Note 2003/004 S. Cucciarelli, D. Kotlinski, Pixel hit reconstruction, CMS Internal Note 2004/014

[6] G. Segneri and F. Palla, Lifetime-based b-tagging with CMS, CMS Note 2002/046 A. Rizzi, F. Palla and G. Segneri, Track impact parameter based b-tagging, CMS note in preparation

[7] G. Bagliesi et al., τ identification in CMS, CMS note in preparation S. Gennai, τ identification in CMS, CMS Conference Report 2005/018