Available on CMS information server **CMS CR 2005/021**

CMS Conference Report

29 Septemebr 2005

Track and Vertex Reconstruction with the CMS Detector at LHC

S. Cucciarelli

CERN, Geneva, Switzerland

Abstract

Because of the high charged particle multiplicity expected in proton-proton collisions at the LHC experiments, track and vertex reconstructions are challenging tasks. The main track and vertex reconstruction techniques used in the CMS experiment are described in the following. Both the track and vertex reconstruction algorithms can be decomposed into a pattern recognition step followed by the estimate of the track and vertex parameters. The main algorithms used in CMS are based on the least-squares methods, but depending on the specific application improvements of the simple least-squares algorithms are found useful to enhance the performance.

Presented at *Beauty 2005*, Assisi (Perugia) Italy, June 20-24, 2005

1 Introduction

The proton beams at LHC are expected to collide at a centre-of-mass energy of 14 TeV with bunch spacing of 25 ns. In the high luminosity period ($2{\cdot}10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$) about 20 minimum bias events per bunch are foreseen that corresponds to 5000 charged tracks per event. A fast response of the CMS Tracker Detector [1] is needed to resolve the bunch-crossing time together with high precision of the spatial measurements to resolve tracks close to each-other.

The reconstruction of tracks and vertices in the high density environment of the LHC collision is very challenging. Track reconstruction algorithms have to exploit the high resolution of tracker devises: tracks with transverse momentum (p_T) of 100 GeV/c have a sagitta of only 1.8 mm in the CMS Tracker Detector. Track parameters have to be precisely estimated, a good impact parameter (d_0) resolution is essential to be able to tag b-jets.

At LHC, especially at the high luminosity, charged particles will originate from many different primary and secondary vertices. In order to reconstruct and understand the whole event, vertices need to be found with high efficiency from the set of reconstructed charged tracks and their position precisely estimated.

2 Track Reconstruction

The main algorithm used in the CMS experiment to reconstruct charged tracks is the Kalman Filter [2] which is a recursive formulation of the least-squares method of fitting a set of measurements to a track model. The least-squares method is optimal when the model is linear and the random noise is Gaussian. For non-linear models or non-Gaussian noise the Kalman Filter is still the optimal linear estimator.

The CMS tracker provides 2-3 three-dimensional measurements with spatial resolution of about 10 μm in the innermost Pixel detector, and up to 14 two-dimensional meausurements with resolution of about 30 μm in the Silicon-Strip detector. The tracker radius is 1.1 m, its coverage extends up to $|\eta| = 2.4$. The magnetic field is 4 Tesla and is almost parallel to the beam axis.

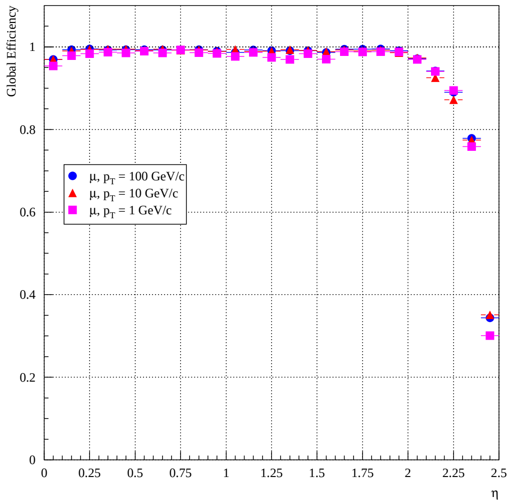
The combinatorial Kalman Filter is a recursive procedure which starts from an initial trajectory estimate and integrates pattern recognition and track fitting. In the default strategy for track reconstruction in CMS, the initial estimate of the trajectory (seed) consists on a track segment made of a set of compatible measurements in the pixel layers close to the beam pipe. The Kalman Filter proceeds by an alternating sequence of propagation and update steps. In the propagation step, the trajectory at the current surface is extrapolated to the next tracker layer. The covariance matrix is propagated as well taking into account multiple scattering and energy loss. In the update step, the extrapolated trajectory is combined with the observation. For each seed, a tree of possible track candidate is generated. The number of trajectories is limited according to their χ^2 and the number of missing hits. Since only the estimate in the last layer is based on the full track candidate information, the surviving track candidates are submitted to a Kalman smoothing procedure to obtain the optimal estimates at every measurement point.

The efficiency to reconstruct tracks with the combinatorial Kalman Filter is shown in Figures 1 and 2 as a function of the pseudorapidity η for single muons and pions respectively. Single muons are reconstructed with an efficiency around 100% up to η =2, while single pions are reconstructed with an efficiency of about 85% due to their nuclear interaction with the Tracker material (about 20% of pions with p_T =1 GeV/c do not reach the outer layer).

Transverse momentum resolution as a function of $|\eta|$ is shown in Figure 3, for single muons with different p_T values. The resolution of the transverse momentum is dominated by multiple scattering for p_T values below 10 GeV/c, by the detector spatial resolution for p_T values above 100 GeV/c, and it degrades for $|\eta| > 1.7$ as long as the track exits from the Tracker detector.

The Combinatorial Kalman Filter is also suitable at the High Level Trigger stage where the Tracker response has to be processed as fast as possible. In order to decrease the processing time, tracks are reconstructed using partial information from the Tracker Detector (limited number of hits). Tracks with 5-6 hits are reconstructed with a precision comparable to the offline reconstruction, while the processing time is reduced by a factor around 1.4 [3]. Figure 4 shows the transverse impact parameter resolution of tracks in b-jets as a function of the number of hits.

To better describe non-linearity, adaptive filters have been developed for specific applications. In particular, the Gaussian Sum Filter is suitable when the energy loss and bremsstrahlung are non-Gaussian and measurement errors have tails. As an application of the Gaussian-Sum Filter, the reconstruction of electron is presented.



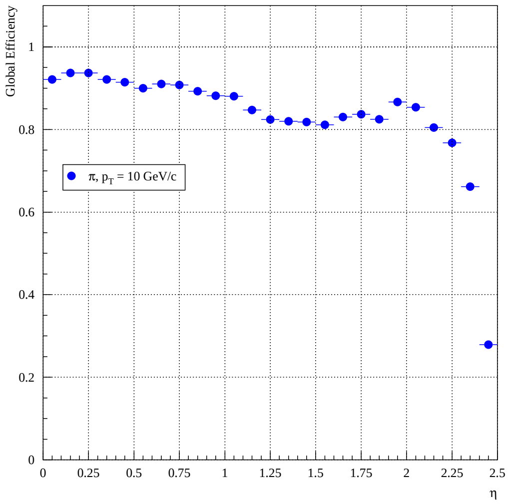


Figure 1: Track reconstruction efficiency for single muons with different p_T as a function of the pseudorapidity η.

Figure 2: Track reconstruction efficiency for single pions with $p_T = 10$ GeV/c as a function of the pseudorapidity η.

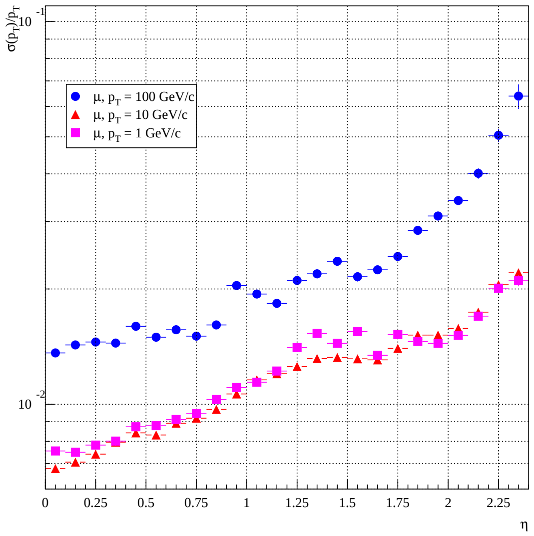


Figure 3: Relative transverse momentum resolution for single muons with different p_T values as a function of the pseudorapidity η.

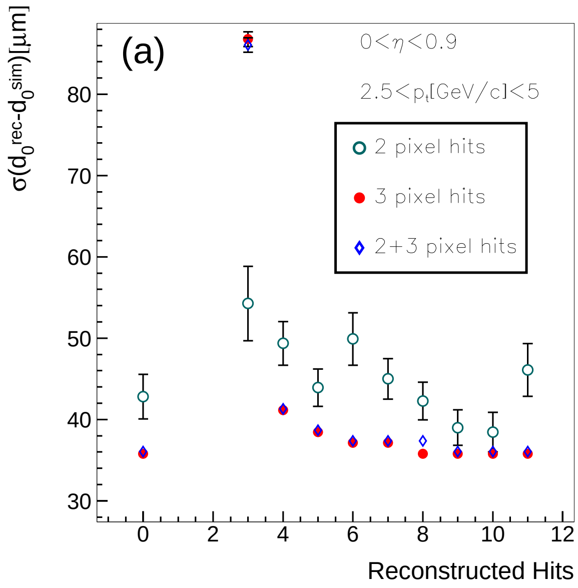


Figure 4: Transverse impact parameter resolution as a function of the number of hits used to reconstruct tracks in b-jets, 0 number of hits corresponds to full tracker information.

2.1 Electron reconstruction with the Gaussian-Sum Filter

The CMS Tracker detector is all made of silicon strips, thus it contains a relevant amount of material due in part to the active material and in part to services (electronics, cabling, etc.). The bremsstrahlung energy loss distribution of electrons propagating in matter is described by the Bethe-Heitler model [4] and it is highly non-Gaussian. With the Gaussian-Sum Filter (GSF), the energy loss distribution is modeled by a mixture of Gaussians instead of a single Gaussian as with the Kalman Filter. A detailed description of the electron reconstruction with the Gaussian-Sum Filter in the CMS Tracker can be found in reference [5].

The GSF is a non-linear generalization of the Kalman Filter where the distributions of all state vectors are a weighted sum of Gaussians. The GSF is implemented as a number of Kalman Filters running in parallel, where each Kalman Filter corresponds to one of the components of the Gaussian mixture. The weights of the different components are computed separately. As the Kalman Filter, the GSF alternates propagation and update steps. At each step the distribution of the state vectors are also allowed to be a mixture of Gaussians. The convolution of the mixture describing the predicted state with the mixture modeling the energy loss leads to an exponential explosion of the number of components. In order to limit the number of components, components are merged together according to a given distance definition and replaced by a single Gaussian.

Figure 5 shows the distribution of the momentum residuals of electrons with a p_T of 10 GeV/c reconstructed with the Kalman Filter and the GSF where a six-component mixture has been used to approximate the energy loss. The improvement of the GSF with respect to the Kalaman Filter is significant on the core of the residuals distribution, while tails are slightly reduced. The irreducible tails in the momentum residuals are due to the fact that the radiation in the innermost layer of the Tracker can not be detected. This effect can be partially compensated by including a vertex constraint. In addition, an improvement is expected if also the measured positions are modeled by a mixture of Gaussians.

3 Vertex Reconstruction

The performance on vertex reconstruction is clearly related to the quality of the track reconstruction. The vertex reconstruction typically involves two steps: the *vertex finding*, where clusters of tracks originating from the same vertex are individuated as vertex candidates, and the *vertex fitting*, where from a set of tracks the most compatible vertex position is computed and used to constrain track parameters at vertex. The reconstruction of primary vertices can also be performed at an early stage, without using the information from all the tracking system. In this case no vertex fitting is performed as tracks are not fit fully, but a helix approximation is made. The first estimation of the primary vertex of the trigger event is also used to constrain the track reconstruction in the full Tracker system.

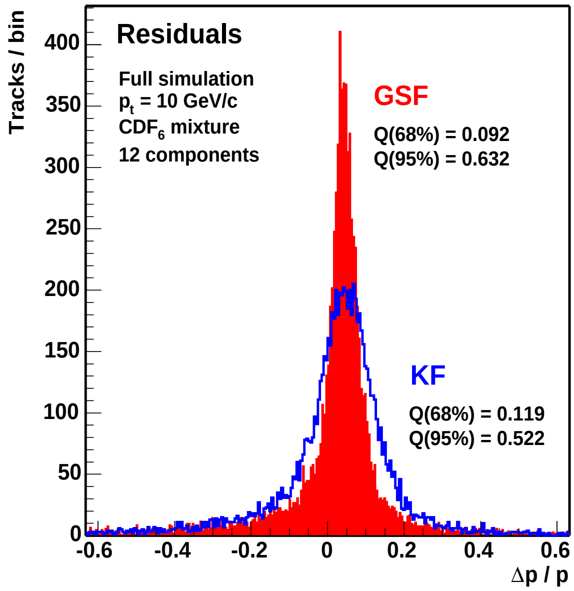


Figure 5: Residual distribution of the momentum for the Gaussian-Sum Filter (solid histogram) and the Kalman Filter (open histogram) for electrons with a p_T of 10 GeV/c.

3.1 Vertex Finding

A first primary-vertex measurement of the trigger event is needed already at High Level Trigger (HLT) stage, before the full track reconstruction is performed. The first estimation of the longitudinal component of the primary vertex (PV) is obtained with the Pixel detector response only. A fast tracking is performed in order to find sets of three hits compatible with a track to be used as inputs to the vertex finding. A description of the vertex finding algorithms with pixel and its performance can be found in reference [6]. Among the reconstructed vertex candidates, the trigger primary vertex is the one with the highest value of the track p_T sum. Reconstructed trigger primary vertices are found in a 500 μm windows around the simulated primary vertex with an efficiency of 95 to 100% for event topologies with high charged track multiplicity. Primary vertex efficiencies of events with few charged tracks in the final state, like H $\rightarrow \gamma\gamma$ and $\mathrm{B}_0 \rightarrow J/\Psi$, drop to 80%. Dedicated algorithims to find the trigger PV out of the vertex candidates are under study.

The offline vertex finding process is accomplished in two steps: first of all, primary vertices are reconstructed, identifying the one which triggered the event, and subsequently the reconstruction of displaced vertices from high lifetime particles, like b and τ, is performed.

The main algorithm to search for primary and secondary vertex implemented in CMS adopts a divisive approach. Tracks with less than 5% compatibility to the vertex candidates are discarded and the search for secondary vertices follows among the discarded tracks. The efficiency to find the trigger primary vertex in $\mathrm{b\bar{b}}$ events is 95% with no Pile-Up and decreases to 92% at low luminosity. The algorithm can be used also within a region of interest, as an example the efficiency to find primary vertices around a b-jet is about 80%.

The efficiency to find a secondary vertex in a b-jet with a p_T of 100 GeV/c is around 48% with a purity above 50%, for secondary vertices sufficiently distinct from the beam line (at least two tracks with a transverse impact parameter significance greater than three). That allows an efficiency to detect b-jets of about 50% with a mistagging rate close to 1% for jets from light quarks and gluon jets.

3.2 Vertex Fitting

Vertex fitting normally relies in the least-squares minimization formulated as a Kalman Filter. This method can be shown to be optimal when the model is linear, the measurement uncertainties are Gaussian and the vertex candidate is not contaminated by mis-measured or mis-associated tracks (*outliers*). Since none of these conditions will hold for the real data in LHC, "robustified" methods have been implemented and studied, as the *Trimming* and the *Adaptive* Vertex Fitters [7].

The Trimming Vertex Fitter discards the M out of N tracks which are the least compatible with the vertex candidate. The ratio M/N is called the trimming fraction and is an input parameter of the algorithm. An adequate value for the trimming fraction is 20%. The Adaptive Fitter is implemented as a re-weighted least-squares method. Tracks are down-weighted by a factor which is a sigmoidal function of the reduced track-vertex distance $\chi_{track} = (x_{track} - x_{vertex})/\sigma_{track}$. The Adaptive algorithms has two advantages with respect the Trimming algorithm: the weights can be fractional (soft assignment) and they are adapted in the course of iterations. In this case there is no need to set *a priori* the fraction of tracks to be rejected. Both these algorithms can be formulated as iterative re-weighted least-squares methods: they iterate from the initial vertex estimate until the vertex position converges.

Table 1: Performance comparison in terms of average χ^2 probability, resolution and pulls of vertices estimated with different vertex filters, when no outlier (left) and one outlier (right) is included in the four-track vertex.

	No track outlier			One track outlier		
Filter	Average P(χ^2)	Resolution [μm]	Pulls	Average P(χ^2)	Resolution [μm]	Pulls
Kalman Filter	0.32	70	1.41	0.18	115	1.61
GSF	0.47	54	1.00	0.37	63	1.11
Adaptive Filter	0.30	59	1.08	0.18	92	1.34
Adaptive-GSF	0.27	54	0.94	0.24	64	1.00

In order to have a better treatment of non-Gaussian tails of measurement errors, a Gaussian-Sum Vertex Fitter has been also developed and implemented [8]. In the context of vertex reconstruction, the measurements are the estimated track parameters. Their error distributions are modeled by a mixture of Gaussians. The distribution of the estimated vertex state is thus also distributed according to a mixture of Gaussians. In the fit an iterative procedure is applied: the estimate of the vertex is updated with one track at the time and each component of the vertex state mixture is updated with each component of the track parameter mixture. The weights of the component are calculated independently. As in the Adaptive filter the computation of the vertex position is independent of the computation of the track weights, the Kalman Filter can be replaced by the Gaussian-Sum Vertex Filter. The Gaussian-Sum Vertex Fitter indeed can be combined with the Adaptive Fitter in oder to both down-weight outliers and use the full mixture of track parameter errors.

A comparison of the different vertex fitters has been performed with a simplified simulation. No track reconstruction is done and track parameter are smeared according to a two-component Gaussian mixture. In Table 1, the average χ^2 probability, spatial resolution and pulls are listed for the different filters without outliers and with one outliers (mis-measured tracks) artificially introduced in the vertex. The non-linear fitters show in general better results, particularly in presence of one outliers.

4 Conclusions

The CMS experiment has very robust and versatile track reconstruction algorithms which are able to operate in a very challenging environment. The Kalman Filter is shown to give high reconstruction efficiency and good track parameter resolutions even in a difficult environment as proton-proton collisions at LHC.It can be succesfully used at the High Level Trigger stage with partial Tracker information (tracks reconstructed with five to six hits). More sofisticated methods are available for specific application. In particular the Gaussian-Sum Filter, used to reconstruct electrons, shows a clear improvement of the momentum resolution with respect to the Kalman Filter.

Several algorithms to reconstruct vertices have been studied and implemented in CMS providing both good efficiency in identifying vertex candidates and high precision in evaluating the best estimate of the position. Adaptive algorithms are shown to be more performant. Especially the combination of the Adaptive approach with Gaussian-Sum-Vertex Fitter allows to down-weight outliers and better treat non-Gaussain tails in the track parameter errors.

References

[1] The CMS Collaboration, The Tracker Project Technical Design Report, CERN/LHCC 98-6.

[2] R. Frühwirt, *Nucl. Instrum. Methods* **A 262** 444 (1987).

[3] The CMS Collaboration, The Trigger and Data Acquisition Project, CERN/LHCC 2002-26.

[4] H. Bethe and W. Heitler, *Proc. R. Soc.* **A 146**, 83 (1934).

[5] W. Adam, R. Frühwirt, A. Strandle and T. Todorov, CMS NOTE 2005/001.

[6] S. Cucciarelli, M. Konecki, D. Kotlinski, T. Tododrov, CMS NOTE 2003/026.

[7] R. Frühwirt, J. D'Hont, P. Vanlaer, W. Waltenberger, *IEEE Transactions on Nuclear Science* **Vol. 51**, 5 (2004).

[8] R. Frühwirt, T. Speer, CMS NOTE 2005/005.