Abstract. Vertexing is an important part of event reconstruction in CMS. Several non-linear reconstruction algorithms have been developed to improve the robustness of vertex fitting. An overview of the algorithms used in the CMS experiment is given, their properties are discussed and their performance on simulated data is shown.

1. Introduction
The Compact Muon Solenoid (CMS) is a multi-purpose detector at the Large Hadron Collider (LHC). The detector is designed for full event reconstruction, including tracking and vertexing using a full silicon tracker [1]. The identification of vertices plays an important role in event reconstruction. Precise coordinates of the primary event vertices are needed to assign tracks to collisions and to determine the event kinematics. Secondary vertices are tools for identifying long-lived particles like heavy flavor hadrons and τ leptons. Multiple overlapping events with high track density and particle interactions in the tracking volume are the main challenges for vertex reconstruction.

2. Vertex finding and fitting
A distinction is sometimes made between vertex finding and vertex fitting. Vertex finding is the task of identifying vertices within a given set of tracks, such as the tracks of a jet in case of flavor tagging or the full event in case of primary vertex finding. Vertex fitting on the other hand is the determination the vertex position assuming it is formed by a given set of tracks. The goodness of fit may be used to accept or discard a vertex hypothesis. In addition the vertex fit is often used to improve the measurement of track parameters at the vertex. For the more advanced algorithms described below, the distinction between fitting and finding is not strict, as they do a little bit of both in order to robustify the vertex parameter estimation.

In the ideal case of Gaussian track parameter errors without outliers, a least squares fit or equivalently the Kalman filter, which is the basis of all fitting algorithms in CMS, provides the best possible vertex estimate. The fit converges in all practical cases and the upper tail probability $P(\chi^2)$ is a useful measure of the goodness of fit. Every measurement enters the fit linearly and therefore result is very sensitive to outliers, which are frequently encountered in vertex reconstruction. An outlier can be a track that does not belong to the vertex it is fitted to or a track with badly measured parameters, typically as a consequence of interactions inside the tracking volume. Robustified versions of the Kalman Filter handle both kinds of outliers by rejecting or down-weighting some of the tracks. The Gaussian Sum filter[5] treats tails by modeling the non-Gaussian resolution functions more precisely.
Robustifications of the conventional Kalman fit use the compatibility of each track with the fitted vertex to identify outliers. The Trimmed Kalman Vertex Fitter (TKF) is an iterative procedure in which outlier candidates are removed one by one. After each iteration the track with the highest contribution to the vertex fit $\chi^2$ is removed if the corresponding $\chi^2$ probability is less than 5%. The final fit is a conventional Kalman fit with hard track assignments, however, the fit $\chi^2$ is biased by the outlier rejection and does not follow the usual $\chi^2$ distribution. With three dimensional tracks it is possible to reject all but one track in which case the fit fails completely. In practice this only happens at the permille level. Because the TKF produces a track assignment in addition to the vertex itself, it is used in vertex finding algorithms[3].

A slightly different approach is taken by the Adaptive Vertex Fitter (AVF), in which tracks are not completely included or rejected but the assignment is soft, i.e. each track receives a weight that is a function of its $\chi^2$ contribution to the vertex. After each iteration the weights are adapted to the new vertex position. Deterministic annealing ensures that local minima are avoided in this process [4].

The properties of the different fitters have been studied in simulated CMS data using events that underwent detailed detector simulation and full reconstruction [2]. As an example some results for the primary vertex in associated Higgs production ($t\bar{t}H$) events are shown in figure 1. These vertices contain on average 44 reconstructed tracks. The conventional Kalman Filter always converges but 80% of the fits yield an unacceptable $\chi^2$ probability ($< 1\%$). This reflects a significant fraction of outlier tracks 1. The TKF and the AVF remove 4.2 and 3.9 tracks,

1 In this case the all outliers are due to resolution tails because true secondary tracks have been excluded from the fit.
respectively, where the latter number is the difference between the sum of all track weights and
the number of tracks. While the effect of non-Gaussian tails is clearly visible in the vertices
fitted by the conventional Kalman Vertex Fitter, the robustified Fitters yield strongly reduced
tails and 30% improved coordinate resolution. Similar comparisons with other samples show
that the robust fitters generally give superior results compared to the conventional fitter when
the vertex multiplicity is high, while for low multiplicity (e.g. in $B_s \rightarrow J/\psi \Phi$) improvements are
marginal. The robust fitters need about a factor 5 more CPU time.

3. Gaussian Sum Filters

Gaussian Sum Filters have been successfully applied in CMS for the reconstruction of
electron tracks, which receive substantial resolution tails due to Bremstrahlung. The track
reconstruction provides not only the track parameters and a covariance matrix, but a set
covariances with appropriate weights that model the non-Gaussian resolution. A vertex fitter
making use of such tracks has been studied in [5] where clear improvements of resolution and
tails of the reconstructed vertex coordinates been demonstrated. The Fitter works as a set of
Kalman Filters running in parallel. The combinatorial growth of the number of components as
more tracks are added to the vertex would lead to prohibitive computation times. However, it
has been shown that the number of components can be restricted without severe degradation of
the result. The Gaussian Sum Fitter has recently been implemented in CMS.

4. Primary Vertex Reconstruction

Primary vertex reconstruction starts by selecting tracks that are compatible with beam line
and clustering those tracks according to their $z$-coordinate. For offline reconstruction, a vertex
is fitted from the tracks within a cluster. Although a large number of inelastic collisions can be
present in each bunch crossing, there is typically only one signal vertex containing the event of
interest. When taking the vertex with the largest sum of squared track transverse momenta,
the efficiency for reconstructing and selecting the correct vertex reaches 99% for $t\bar{t}H$ events.
For low multiplicity events like $H \rightarrow \gamma \gamma$ it can go down to 76%.

While the offline reconstruction uses fully reconstructed tracks, the online vertex
reconstruction [6] is based on track candidates found in the pixel detector (hit triplets). This
allows a fast determination of the primary vertex $z$-coordinate for high level triggers. It is
reconstructed without a fit with high efficiency and a resolution of 50$\mu$m or better.

5. Conclusions

Vertex reconstruction in the CMS experiment has to deal with a large number of tracks and
a significant fraction of outliers. While the conventional Kalman fitter yields good results
for low multiplicity vertices, robustified versions like the Adaptive Vertex Fit are superior for
high multiplicity. A Gaussian Sum fitter has been developed for vertices containing electrons
resolution tails from radiation. All results presented here have been obtained within the ORCA
software framework of CMS. In the meantime the relevant algorithms have been ported to the
new CMSSW framework and validated. Conventional and more advanced algorithms have been
studied in detail and are well suited for the analysis of LHC data expected to be taken starting
2008.

References

[1] CMS Collaboration, The Tracker System Project TDR, CERN-LHCC/98-6 and CMS Collaboration,
Addendum to the CMS Tracker TDR, CERN/LHCC 2000-16
[2] CMS Collaboration, CMS Physics TDR Vol. 1, CERN-LHCC/2006-001

2 within 500 $\mu$m in $z$ of the true signal vertex
[3] Th. Müller, C. Piasecki, G. Quast, C. Weiser, CMS NOTE 2006/027
[4] R. Frühwirt, W. Waltenberger, P. Vanaler, CMS NOTE 2007/008
[5] T. Speer, R. Frühwirt 2006 *Comp. Phys. Comm.* 174 935
[6] S. Cucciarelli, M. Konecki, D. Kotlinski, T. Todorov, CMS Note 2006/026