CMS Conference Report

February 12, 2004

A Gaussian-sum Filter for Vertex Reconstruction

R. Frühwirth

Institute of High Energy Physics, Vienna, Austria

T. Speer

University of Zürich, Switzerland

Abstract

A vertex reconstruction algorithm was developed based on the Gaussian-sum Filter and implemented in the framework of the CMS reconstruction program. While linear least-square estimators are optimal in case all observation errors are Gaussian distributed, a GSF offers a better treatment of non-Gaussian distributions of track parameter errors when these are modeled by Gaussian mixtures. Results are compared to the Kalman filter.

Presented at *IX International Workshop on Advanced Computing and Analysis Techniques in Physics Research*,
December 1-5, 2003, Tsukuba, Japan

1 Introduction

The most often used algorithm for vertex reconstruction is the Kalman Filter [1]. It is mathematically equivalent to a global least-squares minimization, which is the optimal estimator when the model is linear and all random noise is Gaussian. For non-linear models or non-Gaussian noise, it is still the optimal linear estimator.

One method that takes non-Gaussian distributions of measurement errors better into account is the Gaussian-sum filter (GSF). In this method, the distributions of both the measurements errors and the estimated quantities are modeled by mixtures of Gaussians. The main component of the mixtures would describe the core of the distributions and the tails would be described by one or several additional Gaussians. It is a non-linear estimator, where the weights of the components depend on the measurements.

This approach has first been developed and tested for track reconstruction [2], where it has been seen that in the presence of large tails, the GSF has smaller variance than the Kalman filter. It is particularly useful for electron reconstruction, as the Bethe-Heitler distribution of bremsstrahlung energy loss is highly non-Gaussian and can be modeled by a Gaussian mixture. This has been successfully implemented in the CMS reconstruction software [3], where a clear improvement of the track parameter resolution is seen.

2 The Gaussian-sum filter for vertex reconstruction

For the fit, as in the Kalman filter, an iterative procedure is applied, where the estimate of the vertex is updated with one track at the time. When one track is added to the vertex, each component of the vertex state mixture is updated with each component of the track measurement mixture by a Kalman filter, effectively doing an exhaustive combination of the components of the two mixtures. The number of components of the estimated vertex rises thus exponentially, as at each step it is multiplied by the number of components modeling the new track. The GSF is a weighted sum of several Kalman filters run in parallel, and only the weights of the components are calculated separately.

Simulation results

To validate the algorithm, a simplified simulation in a fully controlled environment has been used, where a single four-track vertex is generated per event. No track reconstruction is done, and track parameters are smeared according to a two-component Gaussian mixture model. The component modeling the non-Gaussian tails of the distribution (called *wide component*) has a standard deviation ten times larger than that of the core component (the *narrow component*). Their relative weights are of 90% for the narrow component and 10% for the wide component, and the standard deviations of the impact parameters are of 100 μm and 1000 μm respectively.

For a fit with the Kalman Filter, where only one component is used, the variance of the track parameters used is that of the dominant component¹⁾. The distributions of the vertex-coordinate residuals and pulls (figure 1) have a Gaussian core with tails, and 32% of the fits have a χ^2 -probability below 0.01.

When using the GSF, the two components are used, each with the correct weight and variance. In these first tests, the number of components kept is not limited. The distributions of the residuals (figure 2) have fewer outliers and the core, when fitted with a Gaussian, has a smaller standard deviation (called hereafter the *resolution*). These outliers are due to vertices with several track-outliers. The distributions of the pulls show no tails and are nearly perfectly Gaussian with a standard deviation nearly at 1. This

¹⁾ In case of a high percentage of tails, it is arguable what the best choice for that variance would be. The position of the estimated vertex, and hence the residuals, is nevertheless independent of the scaling factor of the variance used during the fit, whereas the uncertainty of the estimated vertex, and thus the pulls and the χ^2 , is no longer independent of that scaling factor.

Figure 1: Residual (left) and pull (middle) of the x -coordinate of the reconstructed vertex and χ^2 -probability (right) of the vertex fit using the Kalman Filter.

Figure 2: Residual (left) and pull (middle) of the x -coordinate of the reconstructed vertex and χ^2 -probability (right) of the vertex fit using the GSF, without limitations of the number of components.

indicates that the errors on the track-outliers are correctly taken into account and that a high weight is assigned to the correct component.

The distribution of the χ^2 -probability shows a dip at 0, large values of χ^2 occurring less frequently than expected. This can be explained by the fact that the tails of the narrow core component are well within the range of the wide component. Observations in the tails of the core are therefore interpreted as coming from the wide component, and their contribution to the χ^2 is accordingly small. Large values of χ^2 can occur only if all observations are in the tails of the wide component, and this has a very low probability with only four tracks.

These tests have been repeated with different two-component Gaussian mixture models, varying both the ratio of the standard deviations and the relative weight of the second component. To assess the improvement of the GSF with respect to the Kalman Filter, the half-widths of the symmetric intervals covering 50% and 90% of the residual distribution (called *50%* and *90% coverages*) of the x -coordinate and the *relative efficiency* are used. The *relative efficiency* is defined as the ratio of the generalized variances of the fits with the Kalman Filter and the GSF.

The relative efficiency and the 50% and 90% coverages are shown in figure 3 for the same ratio of the standard deviations of 10 as used in the first example. As can be expected, vertices fitted with the GSF always show an improvement over those fitted with the Kalman Filter. The coverages show the stability of the GSF with respect to the Kalman Filter. The highest relative efficiency is reached when the distance between the two-component Gaussian mixture and the single Gaussian is the largest, as measured by the *Kullback-Leibler Distance* [4]. With a higher proportion of outliers, the relative efficiency will decrease, as the tails of the track parameter errors start to dominate.

Figure 3: 50% (right) and 90% coverage (middle) and relative efficiency of vertex fits for tracks with two components as functions of the relative weight of the second component.

3 Limitation of the number of components

As the number of components of the vertex state mixture increases exponentially, it has to be limited to an acceptable number. This is achieved by clustering (collapsing) components which are close, according to a defined distance measurement, pair-wise until the desired number of components is reached. Two distance measurements are used, the *Kullback-Leibler Distance* [4] and the *Mahalanobis Distance* [5].

Simulation results

To test the sensitivity of the GSF with regard to the maximum number of components kept, the track parameters are smeared with a higher number of outliers, by increasing the relative weight of the wide component to 20%. Obviously, vertices fitted with the Kalman Filter are further degraded, with a higher number of outliers in the residual and pull distributions and a high number of fits (53%) with a χ^2 -probability below 0.01. The same vertices fitted with the GSF without limiting the number of components show only a modest increase of the number of outliers and of the RMS of the core of the residual distribution. For a four-track fit, the final estimate of the vertex has 16 components.

When limiting the maximum number of components kept during the fit to two, the distribution of the residuals is only slightly degraded, with a modest increase of the width and of the tails. Only the distribution of the χ^2 -probability changes noticeably. When increasing the maximum number of components to four, the results are nearly identical to those without limitation. In Table 1, the main features of the fits are shown for different maximum number of components, for either of the two distance measurements. There is no significant difference between the two distance measurements, and good results can be achieved with even a small number of components.

4 Conclusion

A Gaussian-sum filter for vertex reconstruction has been implemented in the CMS reconstruction software. It has been validated with a simplified simulation, with different distributions of track parameter errors. While these distributions are not those obtained with a real track reconstruction, the results with the Kalman filter are very close to those observed with fully simulated data, and indicate that similar results could be obtained with the GSF.

These tests demonstrate the validity of the algorithm and its efficiency in the presence of non-Gaussian noise. An improvement of the resolution and error estimate of the fitted vertex and of the χ^2 of the fit with respect of the Kalman Filter when the track parameters errors have non-Gaussian tails are obtained. In addition, when using electron tracks reconstructed with the GSF, the full mixture can be used, instead of only a single collapsed state. The GSF shows little sensitivity to the number of components kept during the fit, such that only a small number of these can be kept without degrading the fit too much.

Table 1: Comparison of the average χ^2 , resolution, pulls, 50% and 90% coverages and relative efficiency of the distributions of the estimated vertices, for different maximum number components using either the Kullback-Leibler Distance (KLD) or the Mahalanobis Distance (MD).

Nbr Comp.	Avg. χ^2	Resolution [μm]	RMS of pull	50% Cov. [μm]	90% Cov. [μm]	Rel. Eff.
Kalman Filter	21.35	150.5	1.93	99.1	548	-
Gaussian-sum filter						
No limitation	0.99	84.8	0.90	59.2	205	2.71
2 - KLD	0.91	90.5	0.88	64.7	253	1.93
3 - KLD	0.94	90.5	0.88	63.5	237	2.22
4 - KLD	0.95	85.4	0.89	59.6	215	2.53
6 - KLD	0.96	84.9	0.89	59.8	214	2.53
8 - KLD	0.96	83.9	0.90	59.2	205	2.71
2 - MD	0.93	92.2	0.84	64.0	246	2.00
3 - MD	0.95	89.7	0.85	63.1	228	2.16
4 - MD	0.95	84.9	0.89	59.6	211	2.55
6 - MD	0.97	84.6	0.89	59.8	205	2.56
8 - MD	0.97	83.9	0.90	59.2	205	2.70

This allows the filter to be used without incurring a high performance penalty.

References

- [1] R. Frühwirth, R. Kubinec, W. Mitaroff and M. Regler, *Comp. Phys. Comm.* **96**, 189 (1996)
- [2] R. Frühwirth, *Comp. Phys. Comm.* **100**, 1 (1997)
- [3] W. Adam, R. Frühwirth, A. Strandlie, T. Todorov, eConf **C0303241** (2003) TULT009
- [4] G. Kitagawa, *Comput. Math. Appl.* **18**, 503 (1989)
- [5] K.V. Mardia, J.T. Kent, J.M. Bibby, *Multivariate Analysis*, Academic Press, London, UK, 1979