



ELSEVIER

Computer Physics Communications 96 (1996) 189–208

Computer Physics
Communications

Vertex reconstruction and track bundling at the LEP collider using robust algorithms

R. Frühwirth, P. Kubinec¹, W. Mitaroff, M. Regler

Institut für Hochenergiephysik der Österreichischen Akademie der Wissenschaften, Nikolsdorfer Gasse 18, A-1050 Vienna, Austria

Received 15 January 1996

Abstract

This paper presents the application of a robust estimator to the problem of finding and estimating secondary vertices. After a brief review of the standard Kalman filter and its extension to an M -estimator, both estimators are studied on simulated samples of charm and bottom decays. It is shown that the M -estimator is superior to the Kalman filter. The implementation is shown to work also on real data from the LEP collider.

1. Introduction

The reconstruction of vertices by least-squares methods has a long tradition since the time of bubble chamber experiments. In those days the geometric vertex fit was a necessary prerequisite to a complete kinematics fit, which was one of the primary means of particle identification. Secondary vertices and their associated tracks could be recognized visually.

Today, the situation has changed enormously. In a collider experiment, the primary vertex is inside the beam tube, whereas tracks are reconstructed outside and have to be extrapolated back in order to reconstruct the primary and secondary decay vertices. Precision is gained by high-resolution vertex detectors and a correct treatment of multiple scattering. The identification of secondary tracks and subsequently of secondary vertices is now a problem which is no longer amenable to visual inspection, but has to be solved by automatic pattern recognition. In addition, track multiplicity is larger than it used to be and will grow to truly enormous figures in colliders like LHC. The algorithms used for vertex reconstruction have to be adapted accordingly.

It had been recognized already in 1973 by one of the authors (M.R.) that one can take advantage of the particular structure of the vertex fit when it is formulated as a linearized regression problem [1]. Later it was shown that there exists an algorithm which computes the estimates of the vertex position and the track parameters in $O(n)$ steps, and the full covariance matrix of the estimates in $O(n^2)$ steps, n being the track multiplicity [2]. In a general regression problem, $O(n^3)$ operations would be needed. Finally it was recognized

¹ Permanent address: Dept. of Nuclear Physics, Comenius University, Bratislava, Slovak Republic (supported by BMFWF, project no. OWP-50).

that this algorithm is a special case of the Kalman filter and the smoother [3]. It also was shown how the χ^2 -statistics associated with each step of the filter and smoother can be used to discriminate between primary and secondary tracks.

The Kalman filter and the smoother possess all the well-known optimal properties of linear least-squares estimators: estimates are unbiased, consistent, and have minimum variance among all linear estimates. In an exact linear model with Gaussian errors the estimate is efficient, i.e. has the smallest possible variance among all unbiased estimates.

Beside these nice features, least-squares estimators also have some less desirable properties, the most prominent of which is a lack of robustness. Due to the quadratic form of the objective function, outlying measurements have a large influence on the estimate, leading to serious distortions. Measurements are called outliers either if they do not conform to the model used in the regression, or if their errors are unusually large due to excessive tails of the error distributions. In a fit of the primary vertex, secondary tracks have to be considered as outliers, since they violate the assumption that all tracks in the fit originate at a common point in space. The fitted vertex position is distorted by secondary tracks to a certain degree. As a consequence, the probability of recognizing them is impaired. It therefore seems worthwhile to investigate whether this situation can be remedied by the use of robust estimators. These tend to be computationally expensive, so many of them are not suitable to the analysis of a large number of events. The M -estimator [4] is the obvious candidate, as it can be implemented as an iterated weighted least-squares estimator. A brief review of the Kalman filter and of the M -estimator is given in Section 2.

We have made a detailed comparison of least-squares and M -estimators on samples of simulated events containing secondary decay vertices of short-lived charm and bottom particles. The results are presented in Section 3. We also demonstrate that the algorithm proposed by the authors works well with real data collected by the DELPHI experiment at the LEP collider (Section 4).

2. Review of efficient vertex evaluation techniques

Suppose that there are n tracks of particles in a magnetic field which are to be fitted to a common vertex. For each reconstructed track, we have 5 fitted track parameters \mathbf{p}_k and the corresponding covariance matrix $\mathbf{V}_k = \mathbf{G}_k^{-1}$, defined at some reference surface. The fitted track parameters contain the full information from measurements by the tracking detectors, and are now in turn considered as (virtual) measurements for the purpose of the vertex fit. It is assumed that the tracks are uncorrelated, i.e. that the joint $(5n \times 5n)$ covariance matrix of all \mathbf{p}_k is (5×5) block-diagonal.

The aim of the vertex fit is the estimation of the vertex position \mathbf{x} and of the momentum vectors \mathbf{q}_k of all tracks at the common vertex. In addition, it can be used for the pattern recognition problem of discriminating between “good” tracks (originating from this vertex) and “bad” tracks (belonging to another vertex or to background).

2.1. The Kalman filter method

For an introduction to filter theory the reader is referred to a textbook (e.g. [5]). The terminology and notation used below is the same as in [3,6]. The parameters to be estimated are called *state vector* in filter language. Initially, the state vector consists only of the prior information about the vertex position, \mathbf{x}_0 , and its covariance matrix \mathbf{C}_0 . If no such information is available, \mathbf{x}_0 is arbitrary and \mathbf{C}_0^{-1} is set to zero. For each track, the state vector is augmented by the 3-momentum vector \mathbf{q}_k . Without process noise, the *system equation* is simply the identity

$$\mathbf{x}_k = \mathbf{x}_{k-1} = \mathbf{x}.$$

The functional dependence of the track parameters on the state vector is determined by the equation of motion (track model), and is described by the *measurement equation*, which in general is nonlinear,

$$\mathbf{p}_k = \mathbf{h}_k(\mathbf{x}, \mathbf{q}_k) + \boldsymbol{\epsilon}_k, \quad \text{cov}(\boldsymbol{\epsilon}_k) = \mathbf{V}_k.$$

If there is multiple scattering between the vertex and the reference surface, it has to be included as additional noise in \mathbf{V}_k .

A linear regression model is obtained by approximating \mathbf{h}_k by a first order Taylor expansion at some point $(\mathbf{x}_e, \mathbf{q}_{k,e})$,

$$\mathbf{h}_k(\mathbf{x}, \mathbf{q}_k) \approx \mathbf{h}_k(\mathbf{x}_e, \mathbf{q}_{k,e}) + \mathbf{A}_k(\mathbf{x} - \mathbf{x}_e) + \mathbf{B}_k(\mathbf{q}_k - \mathbf{q}_{k,e}) = \mathbf{c}_{k,e} + \mathbf{A}_k\mathbf{x} + \mathbf{B}_k\mathbf{q}_k,$$

with $\mathbf{A}_k = [\partial \mathbf{h}_k / \partial \mathbf{x}]_e$ and $\mathbf{B}_k = [\partial \mathbf{h}_k / \partial \mathbf{q}_k]_e$ being the (5×3) matrices of derivatives at $(\mathbf{x}_e, \mathbf{q}_{k,e})$.

The *Kalman filter* describes how to add a new track k to a vertex already fitted with $k - 1$ tracks,

$$\tilde{\mathbf{x}}_k = \mathbf{C}_k [\mathbf{C}_{k-1}^{-1} \tilde{\mathbf{x}}_{k-1} + \mathbf{A}_k^T \mathbf{G}_k^B (\mathbf{p}_k - \mathbf{c}_{k,e})],$$

$$\tilde{\mathbf{q}}_k = \mathbf{W}_k \mathbf{B}_k^T \mathbf{G}_k (\mathbf{p}_k - \mathbf{c}_{k,e} - \mathbf{A}_k \tilde{\mathbf{x}}_k),$$

$$\text{cov}(\tilde{\mathbf{x}}_k) = \mathbf{C}_k = (\mathbf{C}_{k-1}^{-1} + \mathbf{A}_k^T \mathbf{G}_k^B \mathbf{A}_k)^{-1},$$

$$\text{cov}(\tilde{\mathbf{q}}_k) = \mathbf{D}_k = \mathbf{W}_k + \mathbf{W}_k \mathbf{B}_k^T \mathbf{G}_k \mathbf{A}_k \mathbf{C}_k \mathbf{A}_k^T \mathbf{G}_k \mathbf{B}_k \mathbf{W}_k = \mathbf{W}_k + \mathbf{E}_k^T \mathbf{C}_k^{-1} \mathbf{E}_k,$$

$$\text{cov}(\tilde{\mathbf{x}}_k, \tilde{\mathbf{q}}_k) = \mathbf{E}_k = -\mathbf{C}_k \mathbf{A}_k^T \mathbf{G}_k \mathbf{B}_k \mathbf{W}_k,$$

with

$$\mathbf{W}_k = (\mathbf{B}_k^T \mathbf{G}_k \mathbf{B}_k)^{-1},$$

$$\mathbf{G}_k^B = \mathbf{G}_k - \mathbf{G}_k \mathbf{B}_k \mathbf{W}_k \mathbf{B}_k^T \mathbf{G}_k.$$

The chi-square of the filter step has two degrees of freedom. It is given by

$$\chi_{k,F}^2 = (\tilde{\mathbf{x}}_k - \tilde{\mathbf{x}}_{k-1})^T \mathbf{C}_{k-1}^{-1} (\tilde{\mathbf{x}}_k - \tilde{\mathbf{x}}_{k-1}) + \mathbf{r}_k^T \mathbf{G}_k \mathbf{r}_k,$$

$$\mathbf{r}_k = \mathbf{p}_k - \tilde{\mathbf{p}}_k, \quad \tilde{\mathbf{p}}_k = \mathbf{c}_{k,e} + \mathbf{A}_k \tilde{\mathbf{x}}_k + \mathbf{B}_k \tilde{\mathbf{q}}_k.$$

The total chi-square of the fit is equal to the sum of the chi-squares $\chi_{k,F}^2$ of all filter steps (which are independent random variables),

$$\chi_k^2 = \chi_{k-1}^2 + \chi_{k,F}^2.$$

It has $2k$ degrees of freedom if there is prior information on the vertex position, and $2k - 3$ degrees of freedom otherwise.

The choice of the expansion point $(\mathbf{x}_e, \mathbf{q}_{k,e})$ is in principle arbitrary; however, the approximation error of the linear expansion should be small compared to the measurement errors $\boldsymbol{\epsilon}_k$. If necessary, the function \mathbf{h}_k can be re-expanded at the new point $(\tilde{\mathbf{x}}_k, \tilde{\mathbf{q}}_k)$ and the filter can be recomputed until convergence has been reached.

Due to the simple form of the system equation the final estimate of the vertex position can actually be computed in one go (“global fit”),

$$\tilde{\mathbf{x}}_n = \mathbf{C}_n \left[\mathbf{C}_0^{-1} \mathbf{x}_0 + \sum_{k=1}^n \mathbf{A}_k^T \mathbf{G}_k^B (\mathbf{p}_k - \mathbf{c}_{k,e}) \right],$$

$$\text{cov}(\tilde{\mathbf{x}}_n) = \mathbf{C}_n = \left(\mathbf{C}_0^{-1} + \sum_{k=1}^n \mathbf{A}_k^T \mathbf{G}_k^B \mathbf{A}_k \right)^{-1}.$$

As there is no process noise, the *smoother* requires no more than recomputing the momentum vectors and covariance matrices with the final estimate of the vertex position,

$$\begin{aligned}\tilde{q}_k^n &= W_k B_k^T G_k (p_k - c_{k,e} - A_k \tilde{x}_n), \\ \text{cov}(\tilde{q}_k^n) &= D_k^n = W_k + W_k B_k^T G_k A_k C_n A_k^T G_k B_k W_k = W_k + E_k^{nT} C_n^{-1} E_k^n, \\ \text{cov}(\tilde{x}_n, \tilde{q}_k^n) &= E_k^n = -C_n A_k^T G_k B_k W_k.\end{aligned}$$

If required, the full covariance matrix and the total chi-square can be computed after smoothing,

$$\begin{aligned}\text{cov}(\tilde{q}_k^n, \tilde{q}_j^n) &= W_k B_k^T G_k A_k C_n A_j^T G_j B_j W_j = E_k^{nT} C_n^{-1} E_j^n, \quad k \neq j, \\ \chi_n^2 &= (x_0 - \tilde{x}_n)^T C_0^{-1} (x_0 - \tilde{x}_n) + \sum_{k=1}^n r_k^{nT} G_k r_k^n, \\ r_k^n &= p_k - \tilde{p}_k^n, \quad \tilde{p}_k^n = c_{k,e} + A_k \tilde{x}_n + B_k \tilde{q}_k^n,\end{aligned}$$

where r_k^n are the smoothed residuals of track k . Normally, the full covariance matrix is needed only if the vertex fit is followed by a kinematic fit. Even in this case, matrix elements of tracks not taking part in the kinematic fit need not be computed.

Computation of the full covariance matrix requires $O(n^2)$ arithmetic operations, whereas all other formulas can be computed by $O(n)$ operations.

The chi-square of the filter can be used to test whether a track is compatible with the fitted vertex. This test is, however, not symmetric: track k is tested against a vertex fitted with tracks 1 to $k-1$. A symmetric test can be constructed by using the chi-square of the smoother, which has two degrees of freedom as well. To this end we compute the estimate \tilde{x}_k^{n*} which results from removing track k from the fitted vertex by an *inverse Kalman filter* using $-G_k$ instead of G_k ,

$$\begin{aligned}\tilde{x}_k^{n*} &= C_k^{n*} [C_n^{-1} \tilde{x}_n - A_k^T G_k^B (p_k - c_{k,e})], \\ \text{cov}(\tilde{x}_k^{n*}) &= C_k^{n*} = (C_n^{-1} - A_k^T G_k^B A_k)^{-1}.\end{aligned}$$

The smoothed chi-square of track k can now be computed from r_k^n and \tilde{x}_k^{n*} ,

$$\chi_{k,S}^2 = (\tilde{x}_n - \tilde{x}_k^{n*})^T (C_k^{n*})^{-1} (\tilde{x}_n - \tilde{x}_k^{n*}) + r_k^{nT} G_k r_k^n.$$

In contrast to the filter, the smoothed chi-squares $\chi_{k,S}^2$ are no longer independent. Clearly, the chi-square of the smoother can be used as a test criterium for outlying tracks, in particular secondary tracks. Its usefulness is, however, rather limited, particularly when there are several such tracks. In that case, the estimate \tilde{x}_k^{n*} is biased by the remaining outliers, and the power of the test decreases. This will be demonstrated in Section 3.4.

2.2. The M -estimator method

As first suggested by one of the authors (R.F.), the estimation of the vertex can be made less sensitive to outliers by using a robustified version of the Kalman filter, based on the M -estimator [4]. The basic idea of the M -estimator is a modification of the objective function such that outliers have less influence on the estimate. This is achieved by downweighting the measurements with large residuals. Let us consider a homogeneous linear regression problem with m parameters x , a $(n \times m)$ model matrix A and n measurements y ,

$$y = Ax + \epsilon, \quad \text{cov}(\epsilon) = V.$$

We assume for the time being that V is diagonal with elements $(\sigma_1^2, \dots, \sigma_n^2)$. The objective function \mathcal{L} can be written as

$$\mathcal{L}(\mathbf{x}) = \sum_{i=1}^n (r_i/\sigma_i)^2,$$

with the residuals

$$r_i = y_i - \sum_{j=1}^m a_{ij}x_j.$$

Minimization of this objective function with respect to \mathbf{x} yields the least-squares estimate of \mathbf{x} . The M -estimator is based on a generalized objective function

$$\mathcal{M}(\mathbf{x}) = \sum_{i=1}^n \Psi(r_i/\sigma_i),$$

where $\Psi(t)$ is a suitable function, usually specified in terms of its derivative $\psi(t) = d\Psi/dt$. Clearly, there is a large variety of possible M -estimators. If we choose $\psi(t) = t$ we obtain the least-squares or L_2 -estimator. A choice of $\psi(t) = \text{sgn}(t)$ yields the L_1 -estimator.

A frequent choice of ψ is the following one:

$$\psi(t) = \begin{cases} t, & |t| \leq R, \\ R \text{sgn}(t), & |t| > R. \end{cases}$$

R is called the *constant of robustness*. It is usually chosen between 1 and 3. The objective function \mathcal{M} is minimized with respect to \mathbf{x} by setting the derivatives equal to 0,

$$\partial \mathcal{M} / \partial x_j = \sum_{i=1}^n (r_i/\sigma_i) (a_{ij}w_i/\sigma_i) = 0, \quad j = 1, \dots, m,$$

with

$$w_i = \frac{\psi(r_i/\sigma_i)}{r_i/\sigma_i} = \begin{cases} 1, & |r_i| \leq R\sigma_i, \\ R\sigma_i/|r_i|, & |r_i| > R\sigma_i. \end{cases}$$

The M -estimator is formally identical to a least-squares estimator with modified weights w_i/σ_i^2 , thus down-weighting the contribution of large residuals to the objective function \mathcal{M} . The estimate cannot be computed explicitly, however, as the weights depend on the unknown residuals r_i . Therefore, one has to resort to an iterative procedure. The initial estimate is computed by ordinary least-squares, corresponding to $w_i = 1$ for all i . The estimate of iteration $k - 1$ is used to compute residuals which are then used to compute the weights for iteration k . The iteration is stopped as soon as there is no significant change in the minimum value of the objective function \mathcal{M} . It should be noted that for some choices of the function ψ the M -estimator is not necessarily unique, the objective function having several local minima.

If the M -estimator is to be applied to the vertex fit, the weight matrices $\mathbf{G}_k = \mathbf{V}_k^{-1}$ of the virtual measurements (track parameters), which are in general non-diagonal, have to be transformed to diagonal form. This can be done, *separately* for each track k , by using an orthogonal transformation \mathbf{U}_k (because they are symmetric),

$$\mathbf{G}_k' = \mathbf{U}_k \mathbf{G}_k \mathbf{U}_k^T, \quad k = 1, \dots, n,$$

where the rows of \mathbf{U}_k are orthonormal eigenvectors of \mathbf{G}_k , and \mathbf{G}_k' is a diagonal matrix containing the (positive) eigenvalues of \mathbf{G}_k . The virtual measurements and the matrices of derivatives have to be transformed accordingly,

$$\mathbf{p}_k' = \mathbf{U}_k \mathbf{p}_k, \quad \mathbf{A}_k' = \mathbf{U}_k \mathbf{A}_k, \quad \mathbf{B}_k' = \mathbf{U}_k \mathbf{B}_k, \quad k = 1, \dots, n.$$

The first iteration of the M -estimator is just the global least-squares fit as described in the preceding subsection; in subsequent iterations the weights are modified as described above, yielding a new weight matrix \mathbf{G}_k'' . The global least-squares formalism is then applied again.

In case of several equally well-measured vertices, the M -estimator may exhibit a multi-stable behaviour. This can be avoided by including prior information (\mathbf{x}_0 , \mathbf{C}_0) about the position of the vertex wanted – e.g. the beam interaction profile for a primary vertex.

A statistical analysis of the M -estimator is considerably more difficult than the corresponding analysis of the linear least-squares estimator; in fact, little can be said about the small sample properties of the M -estimator. There are, however, some asymptotic results available. It can be shown that under certain regularity and uniqueness conditions the M -estimator is consistent and asymptotically normal. For mathematical details the reader is referred to the book of Huber [4].

The test of compatibility of a track with the fitted vertex can be computed according to one of three criteria:

- The smoothed chi-square $\chi_{k,S}^2$, using the original weight matrix \mathbf{G}_k' ;
- The smoothed chi-square $\chi_{k,S}'^2$, using the final (downweighted) weight matrix \mathbf{G}_k'' ;
- The product $\pi_k = \prod_{i=1}^5 w_i^k$ of the final weight corrections of track k .

The “chi-squares” $\chi_{k,S}^2$ and $\chi_{k,S}'^2$ are “large” if track k is an outlier; they are no longer actually χ^2 -distributed even if there is no outlier. The product π_k is “small” if track k is an outlier; its distribution is unknown. Therefore the critical values of the tests have to be determined by simulation studies.

After removing outlying tracks, the vertex is refitted with ordinary least-squares and the original weight matrices yielding a correct chi-square and covariance matrix. In practice, however, outlier removal by one of the tests above is never 100% efficient; type I errors (loss of “good” tracks) and type II errors (inclusion of “bad” tracks in the fit) will bias the total chi-square χ_n^2 in opposite directions, the net effect of which is difficult to assess, but should normally be small.

3. Monte Carlo study of charm and bottom events

In this section the efficiencies of testing track association to the correct vertex are systematically compared for simple impact parameter tests, the Kalman filter (Section 2.1) and the M -estimator (Section 2.2) methods, respectively. A clean but realistic environment is obtained by decays of charm or bottom events generated at the Z^0 pole. Mean lifetimes are $\bar{\tau} \approx 124, 140, 317$ and $388 \mu\text{m}/c$ for D^0, D_S^\pm, D^\pm and B mesons, respectively, yielding decay lengths $L = \beta\gamma c\tau$ with $\beta\gamma = P/m$. This is followed by a simulation of the DELPHI detector and the reconstruction of tracks by the usual DELPHI data analysis.

3.1. Event generation and detector simulation

Both data samples are generated by the program JETSET [7] using the Lund fragmentation model. For the charm sample, 100 events $e^-e^+ \rightarrow Z^0 \rightarrow c\bar{c}(g) \rightarrow \text{hadrons}$ are simulated. The bottom sample consists of 200 selected events $e^-e^+ \rightarrow Z^0 \rightarrow b\bar{b}(g) \rightarrow \text{hadrons}$ which fragment into B or B^* mesons only. Decays of the hadrons are generated by the detector simulation program (see below).

This study is based on the DELPHI detector layout of 1990 [8]. The track-sensitive detector modules are, from the inside outwards, in the “barrel region”: Vertex Detector (VD), Inner Detector (ID), Time Projection Chamber (TPC) and Outer Detector (OD); and in the “forward region”: two Forward Drift Chambers (FCA and FCB). The magnetic field is homogeneous with $\mathbf{B} = (0, 0, 1.2)$ T in the central tracking region.

The global coordinate system has its origin at the centre of the detector, the z -axis in direction of the e^- beam, and the x -axis pointing to the centre of the LEP ring. Notation for space points in cylindrical coordinates is (R, Φ, z) , and for vectors in polar coordinates (P, ϑ, φ) , with $0 \leq \vartheta \leq \pi$ and $0 \leq (\Phi; \varphi) < 2\pi$.

The detector simulation is performed by the standard program FASTSIM [9] which avoids the generation of detailed digitizings (raw data). Instead, it simulates the results as expected by a subsequent local pattern recognition in the individual detector modules (track elements, TE). Information on the simulated vertices (SP) and their tracks (ST) is kept and passed to the analysis program.

The VD's two cylindrical layers of Si-microstrips at radii $R_1 = 8.6$ cm, $R_2 = 11.3$ cm measure $R\Phi$ with an accuracy $\sigma_{R\Phi} = 5 \mu\text{m}$ for single tracks respectively $\sigma'_{R\Phi} = 30 \mu\text{m}$ for two tracks at a distance $\Delta_{R\Phi} < 100 \mu\text{m}$. In a modification made for this study, a double-sided layer is simulated at R_1 , measuring also z with an accuracy $\sigma_z = 30 \mu\text{m}$. The simulated TE consists therefore of 3 parameters $\mathbf{p} = (R\Phi, z, \varphi)$ at $R = R_1$ with $\text{var}(R\Phi) = \sigma_{R\Phi}^2$ and $\text{var}(z) = \sigma_z^2$; $\text{var}(\varphi) = 2\sigma_{R\Phi}^2/(R_2 - R_1)^2$ and the correlation $\rho(R\Phi, \varphi) = -1/\sqrt{2}$ are obtained by a straight-line “fit” of the $R\Phi$ measurements at R_1 and R_2 .

3.2. Track reconstruction

As explained above, the standard data analysis program DELANA [10] skips the local pattern recognition and starts with the simulated TE data as “virtual measurements” of the individual detector modules. First, a track search (global pattern recognition) is performed, linking the TEs into track strings (TS).

At this stage a modification is introduced for the VD: its TE is linked to the ID's TE using a-priori information from simulation, thus making sure to enter the “correct” VD measurements into the subsequent track fit. This is essential for our study.

The next stage performs the track fit by a Kalman filter algorithm [2,3] on all TEs of a TS list, using a helix track model with correct treatment of multiple scattering and the possibility of removing an “outlier” TE. The fitted tracks (TK) are extrapolated to a common reference surface at $R_{BEA} = 7$ cm, i.e. 0.8 cm inside the AL beam tube (TKX). The 5-dimensional parameter vector is defined by $\mathbf{p} = (R\Phi, z, \vartheta, \varphi, h/P)$ at R_{BEA} , with P being the track's momentum, $h = \text{sgn}(\partial\varphi/\partial s) = -\text{sgn}(QB_z)$, and Q and B_z being the track's charge and the z -component of the magnetic field, respectively. The corresponding covariance matrix is $\text{cov}(\mathbf{p}) = \mathbf{V}$.

Pointers contained in the data help to identify corresponding TKX-ST pairs if the track is not the product of a secondary reaction in the detector. Then an extrapolation of the original track (ST) from its true vertex position (SP) to the reference surface yields the “true” track parameters \mathbf{p}^{true} .

Inconsistencies between FASTSIM and DELANA in the treatment of detector measurement errors and/or multiple scattering and inefficiencies in the track search are clearly exhibited by plotting the normalized deviations $(\mathbf{p} - \mathbf{p}^{\text{true}})_i / \sqrt{(\mathbf{V})_{ii}}$ for the 5 parameter components which differ strongly from a (0,1) normal distribution, and the track fit's total χ^2 probability which shows a pronounced U-like shape.

Therefore, we keep only the track fit's covariance matrices \mathbf{V} , and replace all fitted track parameters \mathbf{p} by parameters simulated in accordance with \mathbf{V} ,

$$\mathbf{p} \rightarrow \mathbf{p} = \mathbf{p}^{\text{true}} + \mathbf{S}\mathbf{n}, \quad \mathbf{S}\mathbf{S}^T = \mathbf{V},$$

where \mathbf{n} is a (0,1) normally distributed, uncorrelated random 5-vector, and \mathbf{S} is a Cholesky factor of \mathbf{V} . The precision of the track fit may be derived from the distribution of the 5 components of $\Delta\mathbf{p} = \mathbf{p} - \mathbf{p}^{\text{true}}$, the r.m.s. of which are given in Table 1.

For comparison, also the impact parameters of all tracks with respect to the true vertex position (SP), either in space or projected onto the xy -plane, have been plotted, and their mean values are shown in Table 2.

The poor resolution in $R\Phi$ respectively the projected I/P, despite an accuracy of the VD of $\sigma_{R\Phi} = 5 \mu\text{m}$, might seem a paradox. In fact, this measurement can only marginally improve the resolution of φ and P^{-1} : because of multiple scattering in the walls of ID and TPC, their direction and momentum information gets

Table 1

The r.m.s. of the 5 components of $\Delta p = p - p^{\text{true}}$

Parameter	$c\bar{c}$ sample	$B\bar{B}$ sample
$\text{rms}(\Delta R\Phi)$	$140\ \mu\text{m}$	$138\ \mu\text{m}$
$\text{rms}(\Delta z)$	$1810\ \mu\text{m}$	$1885\ \mu\text{m}$
$\text{rms}(\Delta\vartheta)$	$6.1\ \text{mrad}$	$5.8\ \text{mrad}$
$\text{rms}(\Delta\varphi)$	$7.5\ \text{mrad}$	$6.9\ \text{mrad}$
$\text{rms}(\Delta P^{-1})$	$0.018\ \text{GeV}^{-1}$	$0.016\ \text{GeV}^{-1}$

effectively decoupled from the VD's position information, thus permitting a “zig-zag” fit; this has been proved by simulation of a layout with the ID removed [11]. Then inward error propagation, together with additional multiple scattering in the VD (0.48% r.l. per layer) and in the Al beam tube (1.12% r.l.), blow up the corresponding covariance matrix elements. (However, an improvement is achieved since 1991 by a new DELPHI detector layout including a smaller beam tube made of Be, and a closer third layer of the VD.)

3.3. Vertex evaluation

All parameter vectors p_k and corresponding error matrices V_k ($k = 1, \dots, n$), defined at R_{BEA} inside the beam tube, are now regarded as “virtual measurements” for the vertex fit. The parameters are defined as above, except for a re-definition of $h \rightarrow h = \text{sgn}(Q)$, reversing the sign of the 5th parameter if $B_z > 0$.

The parameters to be fitted are similarly defined as $x = (x, y, z)$ and $q_k = (\vartheta, \varphi, h/P)$. The expansion point chosen for the first approximate primary vertex fit (see below) is $x_e = x_0$ (the centre of the simulated interaction profile) and $q_{k,e}$ as obtained from inward tracking at the point of closest approach in the xy -projection. The interaction profile is not used as a measurement for the fit, so $C_0^{-1} = 0$.

The program for vertex reconstruction (FV) has been developed by one of the authors (W.M.), using the Kalman filter respectively M -estimator algorithms of Section 2. It is run as a module embedded within DELANA, performing the following steps:

- (1) *First approximate primary vertex fit*, using all fitted tracks as input, either
 - (a) by the *Kalman filter* method (Section 2.1), or
 - (b) by the *M-estimator* method (Section 2.2) with a constant of robustness $R = 1.5$ and convergence criterium $|\mathcal{M}(\tilde{x}^{\text{old}}) - \mathcal{M}(\tilde{x}^{\text{new}})| \leq 10$, in order to downweight the influence of outlier tracks (e.g. those originating from secondary vertices).
- (2) *Tests of track association* to this primary vertex, using the following criteria:
 - (a) the *Kalman filter's* smoothed $\chi^2_{k,S}$ probability, obtained by removal of the track being tested;
 - (b) the *M-estimator's* smoothed $\chi^2_{k,S}$ “probability”, formally calculated with the original weight matrix G_k' ;
 - (c) the *M-estimator's* smoothed $\chi^2_{k,S}'$ “probability”, formally calculated with the final (downweighted) weight matrix G_k'' ;
 - (d) the *M-estimator's* product π_k of the final weight corrections for one track.

Table 2

The mean values of the impact parameters in space or projected

Impact parameter	$c\bar{c}$ sample	$B\bar{B}$ sample
mean(I/P space)	$1435\ \mu\text{m}$	$1350\ \mu\text{m}$
mean(I/P proj.)	$348\ \mu\text{m}$	$320\ \mu\text{m}$

Table 3
Corresponding cut values on different test criteria

Test criterion (see above)	Cut value for $\alpha = 10\%$ losses	
	$c\bar{c}$ sample	$B\bar{B}$ sample
Kalman filter (a)	0.07	0.018
M -estimator (b)	0.087	0.063
(c)	0.13	0.107
(d)	0.82	0.78
I/P in space	4400 μm	4000 μm
I/P projected	1500 μm	1250 μm

Further steps of FV (*Secondary vertex search* by chi-square tests and *Final primary and secondary vertex fits* by the Kalman filter method) are used for studies of the accuracy of the final primary and secondary vertex fits with respect to the simulated true vertex parameters and of the accuracy of reconstructing the simulated interaction profile, which are presented elsewhere [11]. Here we present only results which have been obtained from steps 1 and 2 alone.

3.4. Results

For a systematic comparison of the efficiency of different test criteria, a quantitative measure is needed. This may be obtained as follows:

Starting with either the Kalman filter (step 1a) or the M -estimator (step 1b), the distribution of the test criterium being investigated (step 2a, respectively step 2b, c or d) can be plotted *separately* for primary vertex (PV) and secondary vertex (SV) tracks, which are identified (for this purpose only) by simulation information. The plots for the first three test criteria are shown (*left-hand side*) for PV tracks and (*right-hand side*) for SV tracks in Fig. 1a–c and Fig. 2a–c for the $c\bar{c}$ and $B\bar{B}$ samples, respectively.

For comparison with usual track association tests, similar plots (not shown) can be made for the impact parameters with respect to the simulated PV position.

Whatever the test criterium, choosing a cut value to exclude SV tracks from the final primary vertex fit will always be a compromise between loss of some “good” PV tracks (type I errors) and contamination with some “bad” SV tracks (type II errors). Choosing this value for each test criterium at a fixed quantile α for type I errors, we can directly evaluate from the SV plots (like Figs. 1 and 2, right-hand side) the percentage β of tracks contaminating with type II errors. Therefore, $1 - \beta$ is the *power* of the test, which should be large, and β should be low.

Fixing the first quantile at $\alpha = 10\%$ losses, the corresponding cut values on different test criteria follow from the PV plots (like Figs. 1 and 2, left-hand side) as shown in Table 3.

It is remarkable that the cut values for the M -estimator in (b) and (c) are closer to the theoretical value equal to $\alpha = 0.1$ than the cut value for the Kalman filter – which is a least-squares estimator – in (a), and that they depend much less on the event type. This is a consequence of the M -estimator’s robustifying properties.

Note that in mathematical statistics the terminology “test statistic” and “critical value” is preferred in place of “test criterium” and “cut value”, respectively.

Since β is expected to depend strongly on the true distance δ between the primary and secondary vertex, the distribution of the test criterium for SV tracks can be plotted separately for five δ -intervals with limits at 0, 0.1, 0.25, 0.5, 1 and 5 cm (not shown); the centres of these intervals are δ_i . The graphs $\beta(\delta_i)$ for six different test criteria (including those not shown in Figs. 1 and 2) are presented in Figs. 3a and 3b for the $c\bar{c}$ and $B\bar{B}$ samples, respectively.

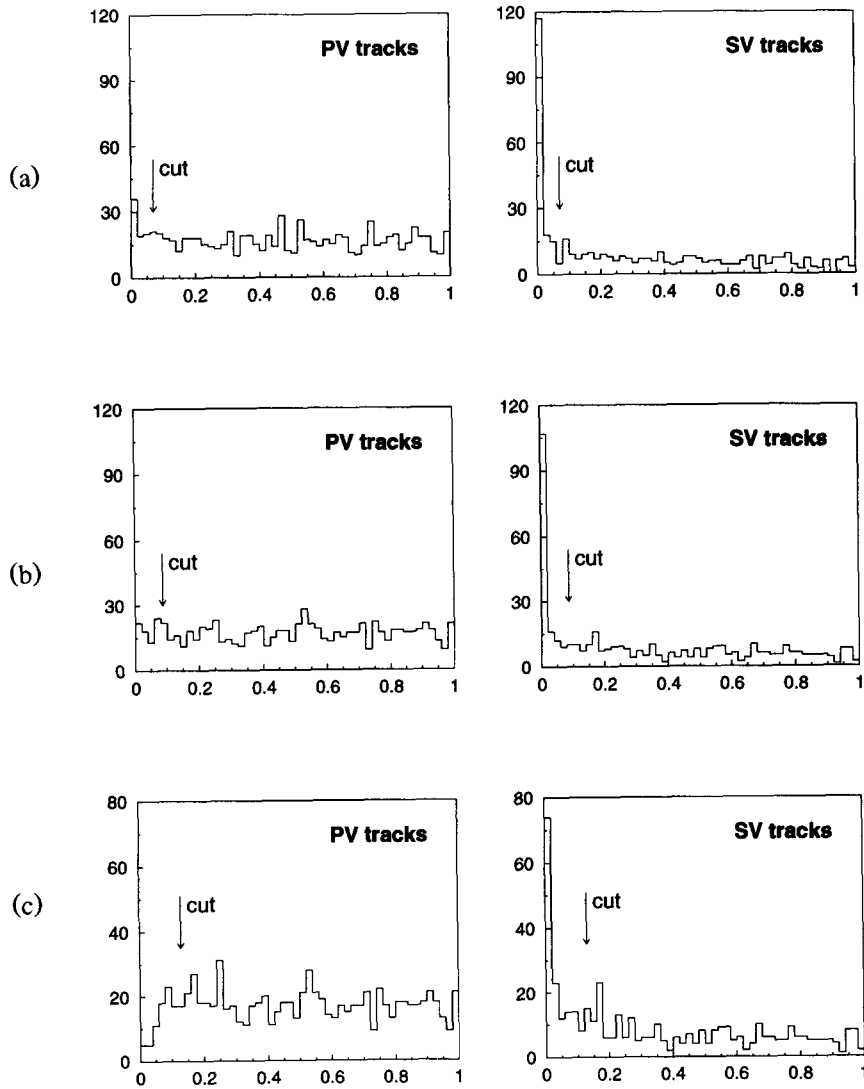


Fig. 1. Distribution of test criteria for primary and secondary vertex tracks ($c\bar{c}$). (a) Kalman filter, smoothed $\chi^2_{k,S}$ probability. (b) M -estimator, $\chi^2_{k,S}$ “probability” with original G_k' . (c) M -estimator, $\chi^2_{k,S}$ “probability” with final G_k'' .

Main results are

- An overall comparison between the two samples shows all six test criteria being less efficient for $c\bar{c}$ events (Fig. 3a) than for $B\bar{B}$ events (Fig. 3b) for $\delta > 5$ mm. This is caused by the kinematical differences of the two samples, as explained below, together with the effects of multiple scattering in the VD and AI beam tube.
- The Kalman filter and M -estimator tests (a)–(d) are clearly superior to the usual impact parameter tests, which use only part of the information available.
- The M -estimator tests (b)–(d) do not significantly differ from the Kalman filter test (a) in the case of $c\bar{c}$, but are clearly superior in the case of $B\bar{B}$. This is due to the kinematical differences between the two samples:
 - Charm events have a characteristic topology of two opposite secondary vertices with jet-like 2- or 3-particle decays. Therefore, a first approximate primary vertex fit by least-squares is only slightly distorted by SV tracks, and the M -estimator cannot give substantial improvement. This has been proven by comparing the

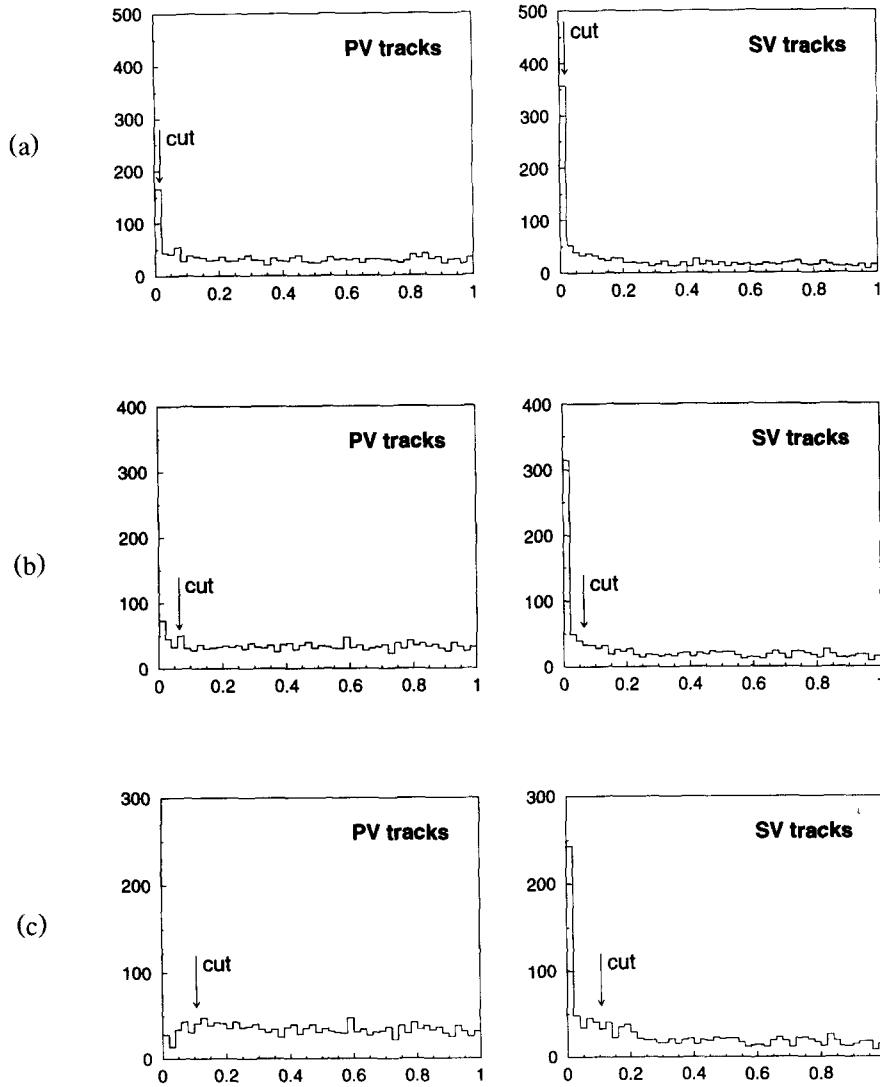


Fig. 2. Distribution of test criteria for primary and secondary vertex tracks ($B\bar{B}$). (a) Kalman filter, smoothed $\chi^2_{k,S}$ probability. (b) M -estimator, $\chi^2_{k,S}$ “probability” with original G_k' . (c) M -estimator, $\chi^2_{k,S}$ “probability” with final G_k'' .

total smoothed χ^2_n probability distributions for that fit, showing almost no difference when including either all tracks (standard method) or true PV tracks only (test option).

- Bottom mesons, because of their higher masses, have secondary vertices with lower momenta and wider decay angles than charm mesons, and have also cascade decays. In such events the first approximate primary vertex fit by least-squares is significantly distorted by outlier tracks. This is clearly seen by the spike in the first bin of Fig. 2a (left-hand side), and is the cause of the Kalman filter’s low cut value of 1.8% for a loss of $\alpha = 10\%$ (see Table 3 above). Thus, the M -estimator is now able to exhibit its superiority.
- Using the M -estimator’s smoothed $\chi^2_{k,S}$ probability (b) as a test criterium for the $B\bar{B}$ sample and accepting $\alpha = 10\%$ losses of PV tracks results in a contamination of $\beta \leq 45.1\%$ (16.2%) with SV tracks for decay lengths $\delta > 5\text{ mm}$ (10mm). Further improvements are to be expected with DELPHI’s new VD layout as mentioned at the end of Section 3.2.

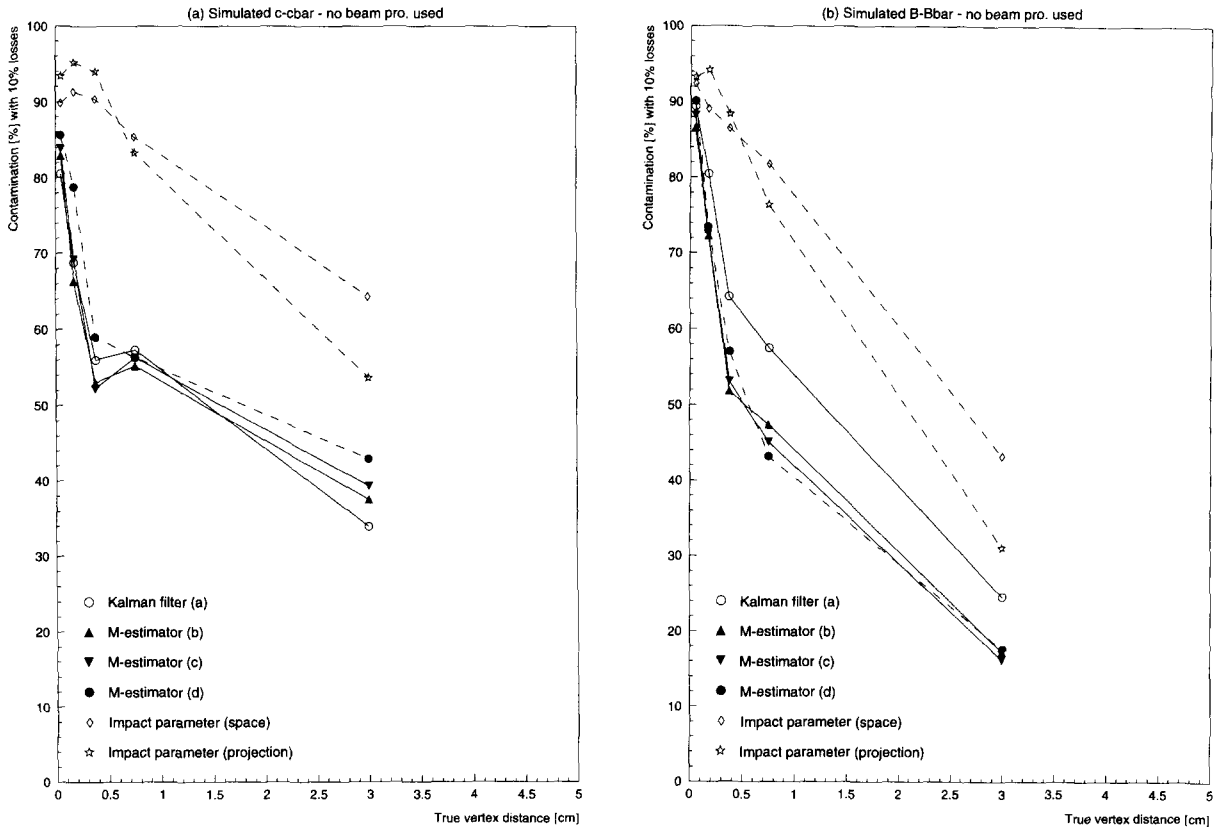


Fig. 3. Percentage of secondary vertex tracks contaminating the primary vertex. (a) $c\bar{c}$ sample, (b) $B\bar{B}$ sample.

In conclusion, the M -estimator is shown to be the best method for testing track association to the primary vertex and is particularly powerful for the identification of complex decays of heavy-flavoured hadrons at LEP.

4. Study of real bottom decays into $J/\psi \rightarrow \mu^+ \mu^-$

A first demonstration of our vertex evaluation algorithms has been performed with real event data, taken in 1990 and 1991 with the DELPHI detector at LEP, by reconstructing “long-lived” 2-prong decays of K_S^0 , Λ^0 and $\bar{\Lambda}^0$ [12]. Further studies, using real event data of 1992, have been performed with the aim of reconstructing “short-lived” decays of B mesons into J/ψ and $\psi(2S)$. Physics results of those are published elsewhere [13,14].

Here we present a study similar to that of Section 3, based upon reconstructed 2-prong decays $J/\psi \rightarrow \mu^+ \mu^-$. At LEP energies, these are produced predominantly by $e^- e^+ \rightarrow Z^0 \rightarrow b\bar{b}(g)$, followed by a colour-suppressed decay of one b -hadron, $B \rightarrow J/\psi X_S$.

4.1. Detector layout and event reconstruction

The DELPHI detector’s layout of 1992 [15] is an improved one, notably by a smaller beam tube at $R_{in} = 5.3$ cm, made of Be (0.4% r.l.) in the central region $|z| < 28.5$ cm, and a better Vertex Detector (VD) with an additional closer third layer at $R_0 = 6.3$ cm; all layers are single-sided, measuring only $R\phi$ with an

accuracy of $\sigma_{R\phi} = 11 \mu\text{m}$ (including alignment errors) for single-hits respectively with $\sigma'_{R\phi} \approx 150 \mu\text{m}$ for multi-hits.

LEP raw data of the 1992 run at the Z^0 peak were processed by the full data analysis program DELANA [10] version 92-C. Standard pattern recognition was used, both for the detector-local digitizings and for the global track search. The VD measurements are included in a second stage by a special DELANA module [16], and the tracks are re-fitted. In the case of simulated data (see below), no use was made of any a-priori information. This will, of course, influence the track reconstruction efficiency.

The following post-DST processing [17] was done: decoding of the particle identification; selection of hadronic events with at least two associated μ^\pm tracks (flagged by the muon chambers using standard “loose cuts”); and transformation of the track fit parameters from DST’s perigee coordinates to our standard coordinates $\mathbf{p} = (R\Phi, z, \vartheta, \varphi, h/P)$ at $R_{BEA} = 5 \text{ cm}$. These data correspond to the TKX of Section 3.2.

In order to correct for underestimated errors, the covariance matrices $V_k = \text{cov}(\mathbf{p}_k)$ are scaled by factors ≥ 1 . These have been determined by tuning carefully the total $\text{prob}(\chi_n^2, ndf)$ distribution to be flat and the reduced smoothed residuals (“pulls”),

$$(\mathbf{p}_k - \tilde{\mathbf{p}}_k^n)_i / \sqrt{(V_k - \mathbf{P}_k^n)_{ii}}, \quad i = 1, \dots, 5,$$

$$\mathbf{P}_k^n = \text{cov}(\tilde{\mathbf{p}}_k^n) = \mathbf{A}_k \mathbf{C}_n \mathbf{A}_k^T + \mathbf{A}_k \mathbf{E}_k^n \mathbf{B}_k^T + (\mathbf{A}_k \mathbf{E}_k^n \mathbf{B}_k^T)^T + \mathbf{B}_k \mathbf{D}_k^n \mathbf{B}_k^T,$$

to be (0,1) normally distributed, from a subsample of “good” primary vertex fits of tracks $k = 1, \dots, n$. As a result, the following 5 scaling factors are used: (1.55, 1.15, 1.3, 1.55, 1.2) in case of real data, and (1.3, 1, 1, 1.3, 1) in case of simulated data.

Our vertex evaluation program (FV) is run stand-alone with the input data described above. It performs the following steps:

- (1) *First approximate primary vertex fit*, either
 - (a) by the *Kalman filter* method (Section 2.1), using fitted tracks with a projected impact parameter (w.r. to the centre of the beam spot) $\leq 5 \text{ cm}$; or
 - (b) by the robust *M-estimator* method (Section 2.2) with $R = 1.5$ and the convergence criterium $|\mathcal{M}(\tilde{\mathbf{x}}^{\text{old}}) - \mathcal{M}(\tilde{\mathbf{x}}^{\text{new}})| \leq 10$ (yielding on average 4 iterations), using all fitted tracks as input.
- (2) *Tests of track association* to this primary vertex, using as criteria:
 - (a) the *Kalman filter*’s smoothed $\chi_{k,S}^2$ probability, obtained by removal respectively addition of the track being tested;
 - (b) the *M-estimator*’s smoothed $\chi_{k,S}^2$ “probability”, formally calculated with the original weight matrix \mathbf{G}_k' ;
 - (c) the *M-estimator*’s smoothed $\chi_{k,S}^{2'}$ “probability”, formally calculated with the final (downweighted) weight matrix \mathbf{G}_k'' ;
 - (d) the *M-estimator*’s product π_k of the final weight corrections for one track.
- (3) *Secondary vertex search* by *combinatorial bundling* of all tracks (submodule FVBUN2 version 65), using the test criterium (c) from above, and optimized by one of the authors (P.K.) for 2-prongs with (μ^+, μ^-) and located inside of R_{BEA} [14].
- (4) *Final primary and secondary vertex fits* by the Kalman filter method:
 - (a) *final primary vertex fit* with tracks not associated to the secondary vertex if $\text{prob}(\chi_{k,S}^2, 2) > 2.5\%$ (Kalman filter) or $\text{prob}(\chi_{k,S}^{2'}, 2) > 8\%$ (*M-estimator*),
 - (b) *secondary vertex fit* with the tracks as associated by FVBUN2, finally cutting on a total $\text{prob}(\chi_n^2, ndf) > 1\%$.

As an option, the interaction profile can be included in the first and final primary vertex fits (steps 1 and 4a). The corresponding information $\mathbf{x}_0, \mathbf{C}_0 = \text{cov}(\mathbf{x}_0)$ is either taken from a “beam spot data base” or is a-priori known, for real or simulated data, respectively. Further processing is done interactively with the program PAW [18].

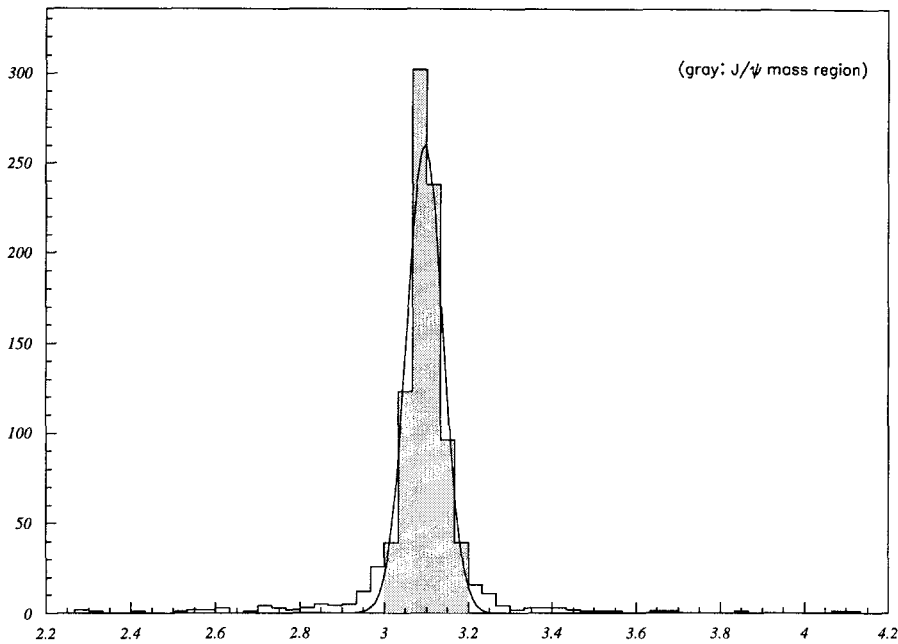


Fig. 4. Simulated data – distribution of the effective mass $(\mu^+\mu^-)/\text{GeV}$.

Our strategy is based on regarding as *secondary vertex tracks* only μ^+ and μ^- originating from reconstructed 2-prong secondary vertices with effective mass in the J/ψ region. Since an efficient particle identification by the muon chambers is essential for this analysis, only μ -candidate tracks with momenta $P \geq 3 \text{ GeV}$ are used.

On the other hand, determination of the cut values for a fixed α -quantile (see Section 3.4) requires a clean sample of *primary vertex tracks*, which can only be obtained from Monte Carlo data by using a-priori information from the event generator. Details are given below.

4.2. Analysis of simulated data

Ca. 25 000 Monte Carlo events $Z^0 \rightarrow b\bar{b}, B \rightarrow J/\psi X_s$ were generated by JETSET [19] and full detector simulation DELSIM [20] for the 1992 set-up. The full data analysis DELANA [10] was the same as that used for real data. Applying the standard post-DST selection criteria [14] gives 2291 simulated events with $J/\psi \rightarrow \mu^+\mu^-$ decays.

After vertex evaluation FV (bundling criteria see [14]), practically no background shows up in the effective mass $m_{\text{eff}}(\mu^+\mu^-)$ of reconstructed 2-prong secondary vertices (Fig. 4). A Gaussian fit to the J/ψ peak yields a mass resolution of $\sigma_m = (43 \pm 1.8) \text{ MeV}$. In the mass interval $3.0 < m_{\text{eff}}(\mu^+\mu^-) < 3.2 \text{ GeV}$ (gray band), 837 events were reconstructed. Looking at a-priori simulation information, we find only 3 secondary vertices (one of them in the mass interval defined above) which do not originate from a “true” J/ψ decay.

Regarding the reconstructed primary vertices of the sample of 837 events defined above, “true” primary vertex tracks can be identified by using a-priori simulation information. This sample is tested for association to the first approximate primary vertex by the criteria (a) of the Kalman filter (removals only) and (b)–(d) of the M -estimator (FV step 2); in addition, the impact parameters with respect to the reconstructed final primary vertex, both in space and in the xy -projection, are also used as test criteria if $\leq 2 \text{ cm}$ (in space) respectively $\leq 0.1 \text{ cm}$ (projection).

Plotting each test criterium for the basic sample of “good” primary vertex tracks, we can proceed as in Section 3.4: determining a corresponding lower or upper cut value (for criteria (a)–(d) or impact parameters,

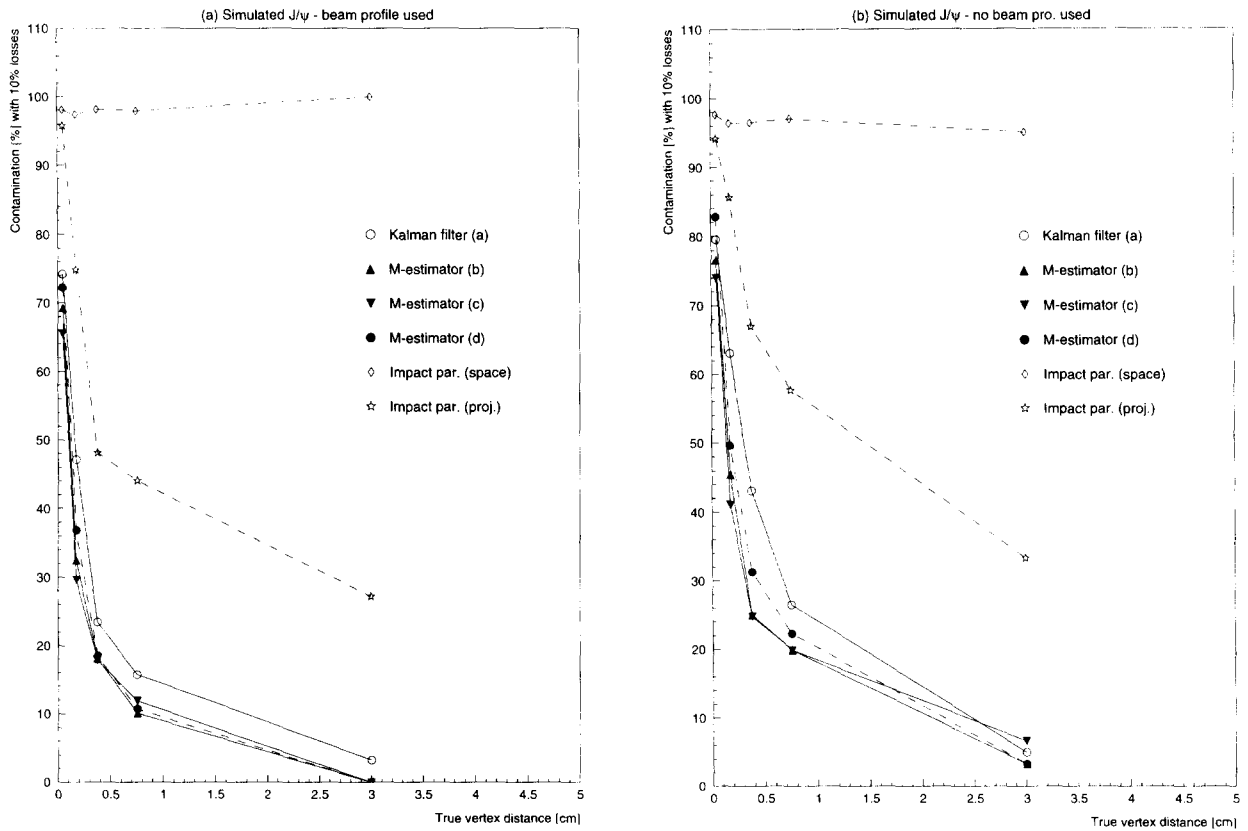


Fig. 5. Percentage of secondary vertex tracks contaminating the primary vertex (simulated data) – (a),(c) beam profile used, (b),(d) not used.

respectively), which will result in a fixed quantile $\alpha = 10\%$ losses (type I errors). These six cut values will be used further on.

Returning to the 837 reconstructed 2-prong secondary vertices of $J/\psi \rightarrow \mu^+ \mu^-$ decays: the true distance δ between simulated primary and secondary vertex is determined to fall into one of five intervals with limits at 0, 0.1, 0.25, 0.5, 1 and 5 cm (their centres being at δ_i). The sample of secondary vertex tracks μ^\pm is tested for primary vertex association by the Kalman filter, the M -estimator and the impact parameters, as above.

Each of the six test criteria is now plotted separately for the five δ -intervals. The cut values determined above are applied, thus allowing to calculate the percentages β of secondary vertex tracks contaminating the primary vertex track sample (type II errors).

The graphs $\beta(\delta_i)$ for the six test criteria are shown in Figs. 5a and 5b (for the beam profile included or not included, respectively, in the primary vertex fits). The same graphs are repeated in Figs. 5c and 5d, with the last two δ -intervals concatenated. Discussion is postponed to Section 4.4.

4.3. Analysis of real LEP data

The 1992 data contain (after DELANA [10]) 720 360 reconstructed “hadronic” Z^0 events, of which 28 866 events with at least two associated μ^\pm are selected by standard criteria, and passed to our vertex evaluation program FV (bundling vers. 65) [14].

The effective mass $m_{\text{eff}}(\mu^+ \mu^-)$ of reconstructed 2-prong secondary vertices is shown in Fig. 6. A fit with

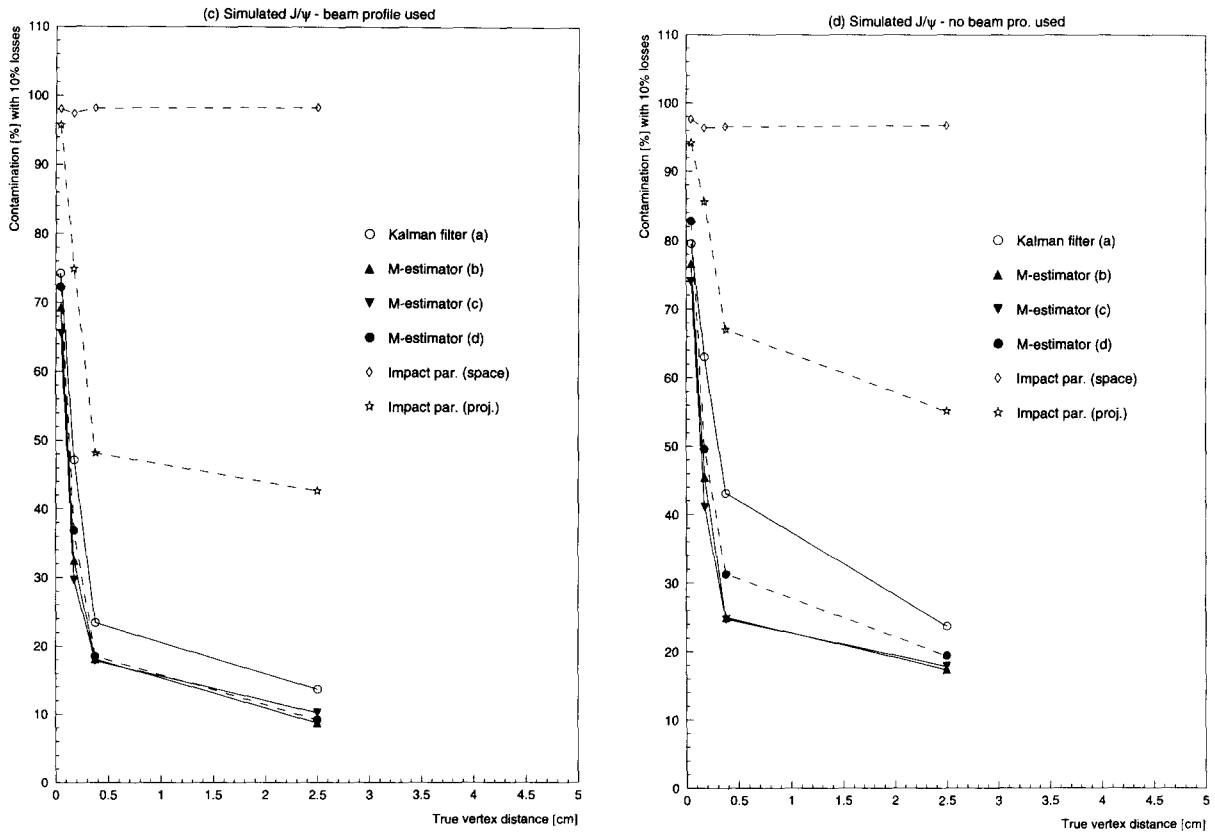


Fig. 5 — continued.

Gaussian peak plus exponential background yields a mass resolution of $\sigma_m = (47 \pm 7.2) \text{ MeV}$, and the deviation from the PDG value [21] is less than $0.3 \sigma_m$. In the mass interval $3.0 < m_{\text{eff}}(\mu^+ \mu^-) < 3.2 \text{ GeV}$ (gray band), 87 J/ψ event candidates are found; the estimated background is calculated to be 19.5 events.

The sample of 87 reconstructed 2-prong secondary vertices of candidate $J/\psi \rightarrow \mu^+ \mu^-$ decays is tested for association of the μ^\pm to the first approximate primary vertex by the criteria (a) of the Kalman filter (removals only) and (b)–(d) of the M -estimator (FV step 2); in addition, the impact parameters with respect to the reconstructed final primary vertex, both in space and in the xy -projection, are also used as test criteria if $\leq 2 \text{ cm}$ (in space) respectively $\leq 0.1 \text{ cm}$ (projection), as in the case of simulated data.

The δ -interval is now determined by the distance δ between reconstructed primary and secondary vertex. As before, each of the six test criteria is plotted separately for the five δ -intervals. The lower respectively upper cut values, which have been determined by a clean sample of simulated primary vertex tracks (see Section 4.2), are applied as before. Neglecting a possible bias from the background (which is not known), this allows to calculate the percentages β of secondary vertex tracks contaminating the primary vertex track sample (type II errors).

The graphs $\beta(\delta_i)$ for the six test criteria are shown in Figs. 7a and 7b (for the beam profile included or not included, respectively, in the primary vertex fits), with the first two and last two δ -intervals concatenated. Note, however, that due to the rather small statistics, the errors are much larger than before.

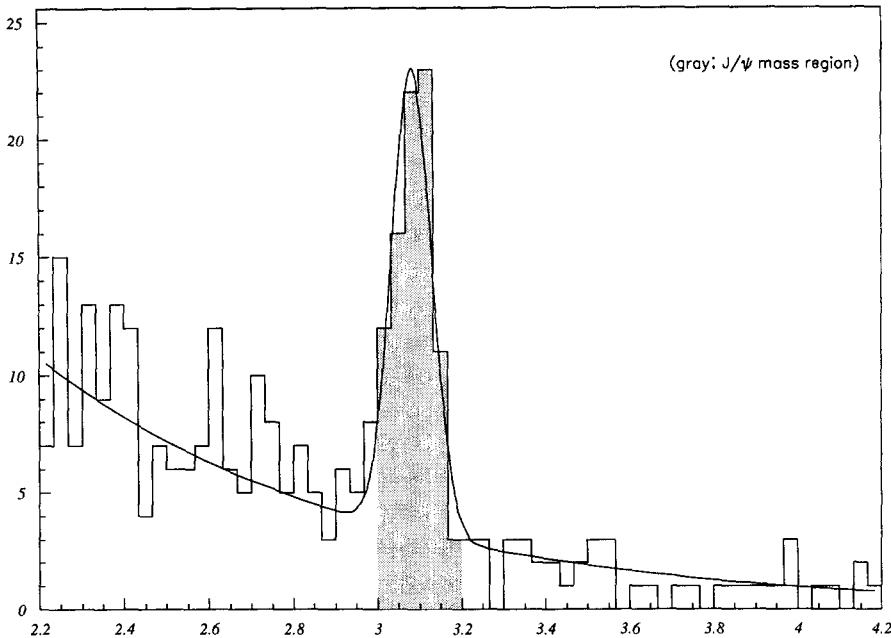


Fig. 6. Real data – distribution of the effective mass $(\mu^+\mu^-)/\text{GeV}$.

4.4. Discussion

The results of this study (Figs. 5, 7) are consistent with each other. Its general features are also consistent with those of the earlier simulation study of Section 3 (Fig. 3), with one exception: the impact parameter *in space* is no longer a good test criterium, giving a percentage of contamination $\beta > 90\%$ even for big decay lengths.

This is, of course, due to our virtual modification of the VD set-up of 1990 (inner layer being double-sided with a z -accuracy of $\sigma_z = 30\ \mu\text{m}$, see Section 3.1); no such modification has been made for the VD set-up of 1992 (only $R\Phi$, no z measured).

Comparing the earlier simulation of $B\bar{B}$ (Fig. 3b) with the corresponding present simulation of $B \rightarrow J/\psi \rightarrow \mu^+\mu^-$ without beam profile (Fig. 5b): the *projected* impact parameter's β is about the same. But the Kalman filter's and M -estimator's test criteria are now significantly better, with lower β over the whole range of decay lengths.

This is due to the very clean sample of selected J/ψ decays, the effect of which overcompensates a possibly worse track reconstruction efficiency caused by pattern recognition (see Section 4.1).

The ranking of the test criteria has remained the same: projected impact parameter worse than Kalman filter (a) worse than M -estimator (b)–(d); the gap between the latter two has become smaller, due to the fact that both have improved. But when concatenating the last two bins (Fig. 5d), the ranking becomes more significant again: Kalman filter (a) worse than M -estimator's product-of-weights (d) worse than M -estimator's χ^2 (b),(c). The splitting of the last two, which is a new feature, is significant for decay lengths in the range $0.1 < \delta < 1\ \text{cm}$.

The effect of using the beam profile shows up in a comparison with Fig. 5a or Fig. 5c: all test criteria have improved over the whole range of decay lengths, as expected. A closer look at Fig. 5c shows the splitting among the M -estimator's test criteria has disappeared; and the gap between the Kalman filter and the M -estimator has narrowed.

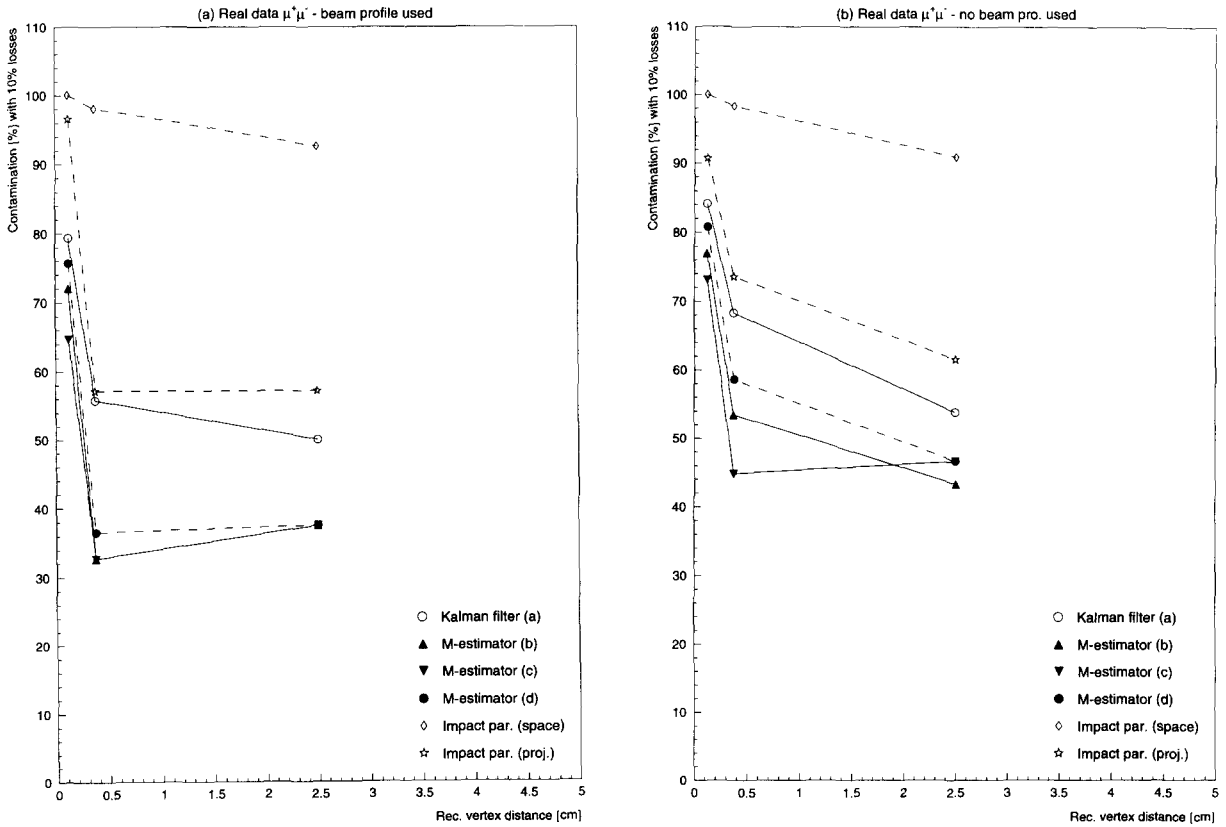


Fig. 7. Percentage of secondary vertex tracks contaminating the primary vertex (real data) – (a) beam profile used, (b) not used.

This last effect is exactly as expected: inclusion of the beam profile counterbalances distortions caused by multiple outlier tracks. In other words, inclusion of the beam profile has a “robustifying” effect by itself. Nevertheless, the M -estimator can still give an additional improvement.

For real data (Fig. 7), two points have to be kept in mind: the data sample contains a lot of background, and it has a much smaller statistics; therefore, the errors on the graphs are definitely much bigger than before. Note that the first two bins and the last two bins have been concatenated now.

The general features are nevertheless consistent with those of simulated data: ranking of the test criteria, effect of inclusion of the beam profile, etc. All test criteria behave significantly worse, while showing some “saturation effect” for decay lengths $\delta > 0.25$ cm, e.g. for the M -estimator at $\beta \approx 35\%$ (with beam spot) respectively 45% (without beam spot). But the gap between Kalman filter (a) and M -estimator (b)–(d) has become very significant again; this is, surprisingly, also the case when including the beam profile into the primary vertex fits.

In conclusion, the M -estimator’s smoothed χ^2 , although distorted towards smaller values, has again been proven to be the best criterium for testing track association to the primary vertex, even in the “dirty” environment of complex real data. Inclusion of the beam profile, however, adds to the power of the test and should be used whenever feasible.

5. Summary and outlook

The detection of secondary tracks and eventually of secondary decay vertices requires an estimate of the primary vertex position which has as little bias as possible. In other words, the estimation of the primary vertex position has to be robustified. In this note we have investigated the M -estimator which achieves robustification by downweighting the influence of outlying secondary tracks. It can be implemented as an iterated weighted least-squares estimator, thereby requiring only minimal modifications to the traditional Kalman filter or global least-squares estimator.

The separation of secondary tracks proceeds by measuring the distance of each track from the fitted primary vertex, leading to some sort of χ^2 -statistic. We have shown that the test statistics based on the M -estimator give consistently better performance than the test statistic based on the least-squares estimator, both with simulated and with real data. Tests based on impact parameters turn out not to be competitive. It should, however, be noted that the improvement due to the M -estimator depends on the event type; for instance, our simulation study shows that it is negligible for $c\bar{c}$ -events, but quite pronounced for $B\bar{B}$ -events.

The computational load of the M -estimator is directly proportional to the number of iterations required until convergence. At startup, the extra requirement of diagonalization can be performed *separately* on the (5×5) weight matrices $G_k, k = 1, \dots, n$ (see Section 2.2), which is a great benefit in case of many tracks. Thus, on average the M -estimator is slower than the Kalman filter. This is, however, not a serious problem, as the time spent in the vertex fit proper is usually only a tiny proportion of the total time required to analyze the event. As an example, the CPU time spent by the vertex evaluation program FV for real LEP data (see Section 4) is increased by less than 10%. There is therefore no reason not to take advantage of the capabilities of the M -estimator.

The vertex reconstruction program FV (presented in Section 3.3 and 4.1) is being recoded by its author (W.M.) with the aim of a detector-independent general-purpose package, VERA (“Vertex Evaluation by Robust Algorithms”), to be submitted to the CPC Program Library.

Acknowledgements

The authors wish to thank D. Liko for many useful discussions, and the DELPHI Physics Analysis Team for the data used in Section 4. One author (P.K.) gratefully acknowledges support from the Austrian Ministry of Science and Research.

References

- [1] M. Metcalf, M. Regler and C. Broll, A Split Field Magnet Geometry Fit Program – NICOLE, CERN 73-2, Geneva (1973).
- [2] P. Billoir, R. Frühwirth and M. Regler, Nucl. Instr. Meth. A 241 (1985) 115.
- [3] R. Frühwirth, Nucl. Instr. Meth. A 262 (1987) 444.
- [4] P.J. Huber, Robust Statistics (Wiley, New York, 1981).
- [5] A. Jazwinski, Stochastic Processes and Filtering Theory (Academic Press, New York, 1970).
- [6] M. Regler and R. Frühwirth, Reconstruction of Charged Tracks, Proc. 5th Advanced Study Inst. on Techniques and Concepts in High Energy Physics, St. Croix, Virgin Islands, July 1988, ed. T. Ferbel, NATO ASI Series B204 (Plenum Press, New York, 1990).
- [7] T. Sjöstrand et al., JETSET Vers. 6.3 – The Lund Monte Carlo Programs, CERN Program Library Pool W5035 etc., Geneva (1987).
- [8] P. Aarnio et al., DELPHI collaboration, Nucl. Instr. Meth. A 303 (1991) 233.
- [9] J. Cuevas et al., Fast Simulation for DELPHI (FASTSIM) Reference Manual. DELPHI 87-27/PROG-72/Rev, Geneva (1988).
- [10] A. Baroncelli et al., DELPHI Data Analysis Program (DELANA) User's Guide, DELPHI 89-44/PROG-137, Geneva (1989).
- [11] W. Mitaroff, Zur Bestimmung der Zerfallsvertices schwerer Teilchen im DELPHI-Experiment am LEP-Speicherring, PhD Thesis, Univ. of Technology, HEPHY-PUB-555/91, Vienna (1991).
- [12] P. Kubinec, D. Liko and W. Mitaroff, DELPHI Vertex Evaluation Studies Using K_S^0 and A^0/\bar{A}^0 , Proc. Int. Conf. Hadron Structure 92, Stará Lesná, Czecho-Slovakia, eds. D. Bruncko and J. Urbán (Šafárik University, Košice, 1992).

- [13] P. Kubinec, D. Liko and W. Mitaroff, Exclusive Reconstruction of Secondary Vertices with J/ψ and $\psi(2S)$ at LEP, HEPHY-PUB-625/95, Vienna (1995).
- [14] P. Kubinec, The Search for Secondary Vertices in the DELPHI Experiment, PhD Thesis, Comenius University, Bratislava (1995).
- [15] P. Abreu et al., DELPHI collaboration, Performance of the DELPHI Detector, Nucl. Instr. Meth. A, to be submitted.
- [16] C. Troncon, M. Caccia and A. Zalewska, VDANA – The Pattern Recognition Program of the Vertex Detector, DELPHI 89-49/PROG-139/TRACK-52, Geneva (1989).
- [17] V. Perevozchikov and N. Smirnov, PHDST Package Description – User's Manuals, DELPHI 92-118/PROG-189/Rev.3, Geneva (1994).
- [18] R. Brun, O. Couet, C. Vandoni and P. Zanarini, PAW – Physics Analysis Workstation, CERN Program Library Long Writeup Q121, Geneva (1995).
- [19] T. Sjöstrand, PYTHIA Vers. 5.7 and JETSET Vers. 7.4 – Physics and Manual, CERN TH-7112/93/Rev and CERN Program Library Pool W5035/W5044, Geneva (1994).
- [20] A. de Angelis et al., DELSIM – DELPHI Event Generation and Detector Simulation, DELPHI 89-67/PROG-142 and DELPHI 89-68/PROG-143, Geneva (1989).
- [21] L. Montanet et al., Particle Data Group, Review of Particle Properties, Phys. Rev. D 50 (1994) 1173.