Properties of robust vertex fitting algorithms at high luminosity HEP experiments

J. D'Hondt[a], R. Frühwirth[b], S. Tavernier[a], P. Vanlaer[c] and W. Waltenberger[b]

(a) *Vrije Universiteit Brussel*
Inter-university Institute for High Energies (IIHE-VUB)
Pleinlaan 2, B-1050 Brussels, Belgium

(b) *Universität Wien*
Institut für Hochenergiephysik
Nikolsdorfer Gasse 18, A-1050 Vienna, Austria

(c) *Université Libre de Bruxelles*
Inter-university Institute for High Energies (IIHE-ULB)
Boulevard du Triomphe, B-1050 Brussels, Belgium

IEEE-NSS Conference 19-25 October, Portland, Oregon, US

email: jdhondt@hep.iihe.ac.be

Abstract

Robust vertex fitting algorithms are expected to improve the knowledge of the vertex position and of its uncertainty, in the presence of mis-measured or mis-associated tracks. Such contaminations are likely to happen in high luminosity HEP experiments. We have performed a simulation study of the sensitivity of two types of robust algorithms : a trimmed least-squares estimator and an adaptive estimator. Their statistical properties are studied as a function of the source and the level of contamination, and compared to the results obtained with classical least-squares estimators.

1 Introduction

Vertex fitting usually relies on Least-Squares minimization techniques. This technique is statistically efficient and unbiased as long as the vertex measurements, i.e. the tracks, have Gaussian and perfectly known uncertainties, and as long as all tracks do originate from the fitted vertex.

However, in experiments or in detailed simulations, track parameter pulls very often exhibit some non-Gaussian tails. These are due either to non-Gaussian multiple Coulomb scattering and detector resolutions, or to imperfect modelling of the detector material and resolutions leading to an imperfect estimation of the track parameter error matrix. In addition, at high luminosity hadron colliders like the LHC, interesting physics (Higgs, SUSY, top,...) can manifest itself through short-lived particles like B-mesons or τ-leptons in the final state. The decay vertex of such particles is difficult to separate from the primary vertex, leading to mis-associations of tracks to vertices.

In estimation theory such contaminations are called outliers. Let us denote tracks with a wrong covariance matrix as Type-1 outliers, and tracks originating from a different vertex as Type-2 outliers. The Least Squares estimator ceases to be unbiased as soon as the measurement sample is contaminated. Robust estimators, unsensitive to outlying measurements, were proposed for vertex fitting in for example CMS [1].

2 Vertex fit algorithms

Fitting the vertex position consists in finding the vertex coordinates $\vec{r}$ which minimize a function F of the reduced distances between the tracks and the fitted vertex. When all tracks have known Gaussian uncertainties and actually originate from the vertex, the function F which yields an efficient and unbiased estimate for $\vec{r}$ is the sum of the squared reduced distances. This is the well-known Least Sum of Squares, or Least Squares (LS) technique. In this case

$$F = \sum_i (\vec{r} - g(\vec{p_i}))^T \mathbf{G}^T \mathbf{C}_i^{-1} \mathbf{G} (\vec{r} - g(\vec{p_i})); \quad i = 1 \rightarrow N_{\text{tracks}}$$

is a χ^2 variable with $2N_{\text{tracks}} - 3$ degrees of freedom, where $\vec{p_i}$ and $\mathbf{C}_i$ denote the parameters and covariance matrix of track i.

The function g is the measurement function which transforms the vector of track parameters $\vec{p_i}$ into a two-dimensional constraint on the vertex coordinates. The matrix $\mathbf{G}$ is the Jacobian matrix $\frac{\partial \vec{p}}{\partial \vec{r}}$ and the matrix product $\mathbf{G}^T \mathbf{C}_i^{-1} \mathbf{G}$ represents the weight of the constraint $g(\vec{p_i})$ in the vertex fit. In the theory of Optimal Estimation this is known as the Kalman formalism. A more complete formulation of Kalman vertex fitting can be found in [2].

To simplify the discussion let us formulate the problem in one dimension. The function to minimize reads in the Least Squares case:

$$F = \sum_i \frac{(r - r_i)^2}{\sigma_i^2}; \quad i = 1 \rightarrow N_{\text{tracks}} .$$

The robustness of an estimator can be defined as the sensitivity of the fit statistics (goodness-of-fit, resolution, uncertainty estimation and bias on the fitted parameters) to the presence of outliers.

This study considered two robust estimators:

- The Least Trimmed Sum of Squares (LTS) algorithm discards the M out of N tracks which have the largest distance relative to the vertex. The ratio $k = M/N$ is called the trimming fraction and is a free input parameter of the algorithm. The default trimming fraction is 20%.
- The Adaptive algorithm downweights distant tracks by a factor which is a sigmoidal function of the reduced distance to the vertex:

$$w_i(d_i) = \frac{e^{\frac{-d_i^2}{2T}}}{e^{\frac{-d_i^2}{2T}} + e^{\frac{-d_{\text{cutoff}}^2}{2T}}} \leq 1; \qquad (1)$$

where $d_i = (r - r_i)/\sigma_i$ is the reduced distance of track i, d_{cutoff} is the distance where the weight function drops to 0.5 and T is the parameter controlling the sharpness of the drop.

Both robust algorithms are iterative and can actually be formulated as iterative reweighted Least Squares algorithms, with the function to minimize at iteration k:

$$F^k = \sum_i w_i^k \cdot (d_i^k)^2; \quad i = 1 \rightarrow N_{\text{tracks}} ,$$

with weights $w_i^k = w_i^k(d_i^{k-1})$ computed according to the vertex position at iteration $k-1$. In the case of the LTS algorithm, the weights assume discrete values 0 or 1.

3 Toy Model : the VertexGun Monte Carlo

The interpretation of the envisaged results concerning the vertex fitting algorithms is simplified by using a new developed Monte Carlo track generator rather than complex real physics events. This simplified setup is illustrated in Figure 1. The statistical properties of the different vertex estimators are studied by introducing a known contamination of outliers (Type-1 or Type-2) in the track sample. For each simulated vertex both the number of tracks and the fraction of inlying to outlying tracks are configurable.

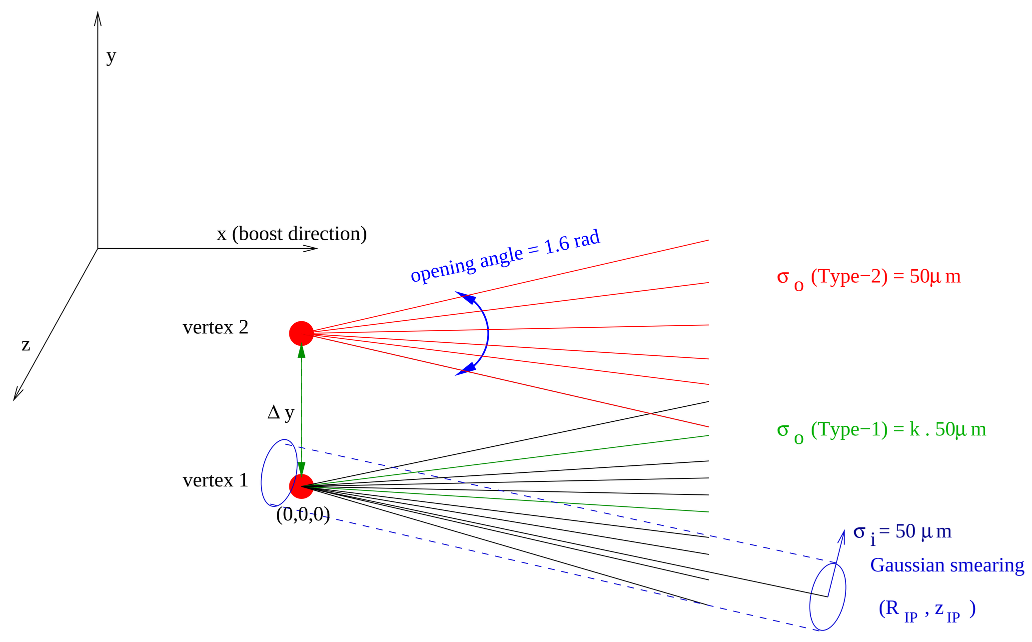

Figure 1: Illustration of the VertexGun parameters.

The tracks are uniformly generated with a momentum between 1 and 30 GeV/c within a cone with an opening angle 2θ equal to 1.6 rad in the direction around the x-axis (i.e. the boost direction). The 4 Tesla magnetic field is oriented along the z-axis. The effect of the track reconstruction is simulated by smearing the track parameters with a 5-dimensional Gaussian distribution. The smearing of the impact parameters ($R_{\text{IP}}, z_{\text{IP}}$) for inliers is done with a fixed spread of 50 μm.

- Type-1 outliers are simulated as the inliers but by smearing their impact parameters with a spread of $k_{\text{Type}-1} \cdot 50\mu$m, while the same covariance matrix is used for both inliers and outliers in the computation of the vertex position. This simulates outliers which are thought to be $k_{\text{Type}-1}$ more precise than they truly are. The outliers of Type-1 can perturb the pull distribution of the vertex estimator and they can reduce its resolution.

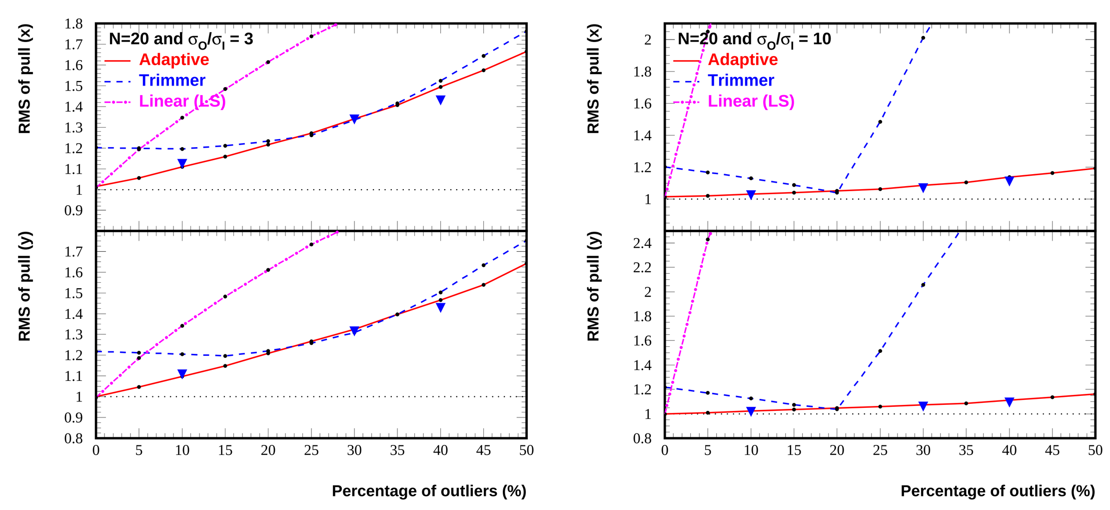

Figure 2: The RMS of the pull distribution ($(r_{\text{fit}} - r_{\text{sim}})/\sigma_{\text{fit}}$) of different vertex estimators as a function of the fraction of Type-1 outliers in the sample of in total 20 tracks. This both in the direction longitudinal and transversal to the global boost of the tracks. The triangles indicate the results obtained at other values of the otherwise fixed trimming fraction, being 10%, 30% and 40%. On the left the coordinates of the outliers are smeared 3 times broader than the inliers ($k_{\text{Type}-1} = 3$), while on the right they are smeared 10 times broader($k_{\text{Type}-1} = 10$).

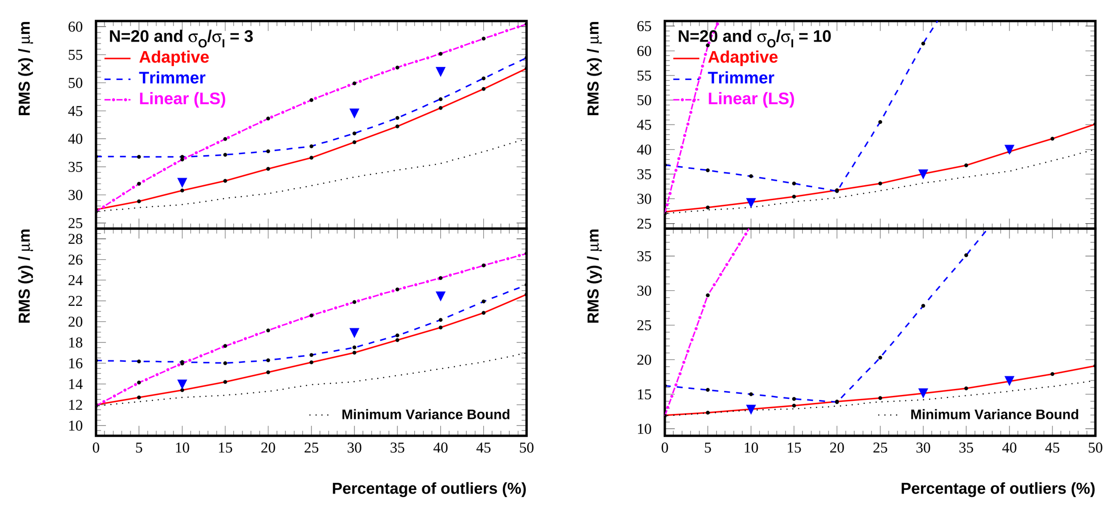

Figure 3: The RMS of the residuals distribution ($r_{\text{fit}} - r_{\text{sim}}$) of different vertex estimators as a function of the fraction of Type-1 outliers in the sample of in total 20 tracks. This both in the direction longitudinal and transversal to the global boost of the tracks. The triangles indicate the results obtained at other values of the otherwise fixed trimming fraction, being 10%, 30% and 40%. On the left the coordinates of the outliers are smeared 3 times broader than the inliers ($k_{\text{Type}-1} = 3$), while on the right they are smeared 10 times broader ($k_{\text{Type}-1} = 10$). The results are compared to the MVB obtained with no outliers and using the LS vertex fitter.

- Type-2 outliers are simulated by generating a secondary vertex separated with a distance Δy from the primary one along the y-axis. This is a worst case scenario since outliers from the same secondary vertex attract the fitted vertex towards their vertex, introducing a bias on the estimation of the primary vertex. Such outliers are called leverage points as they bias the LS estimate by a quantity which is proportional to Δy. Also the resolution of the vertex position estimator can be degraded by the presence of Type-2 outliers.

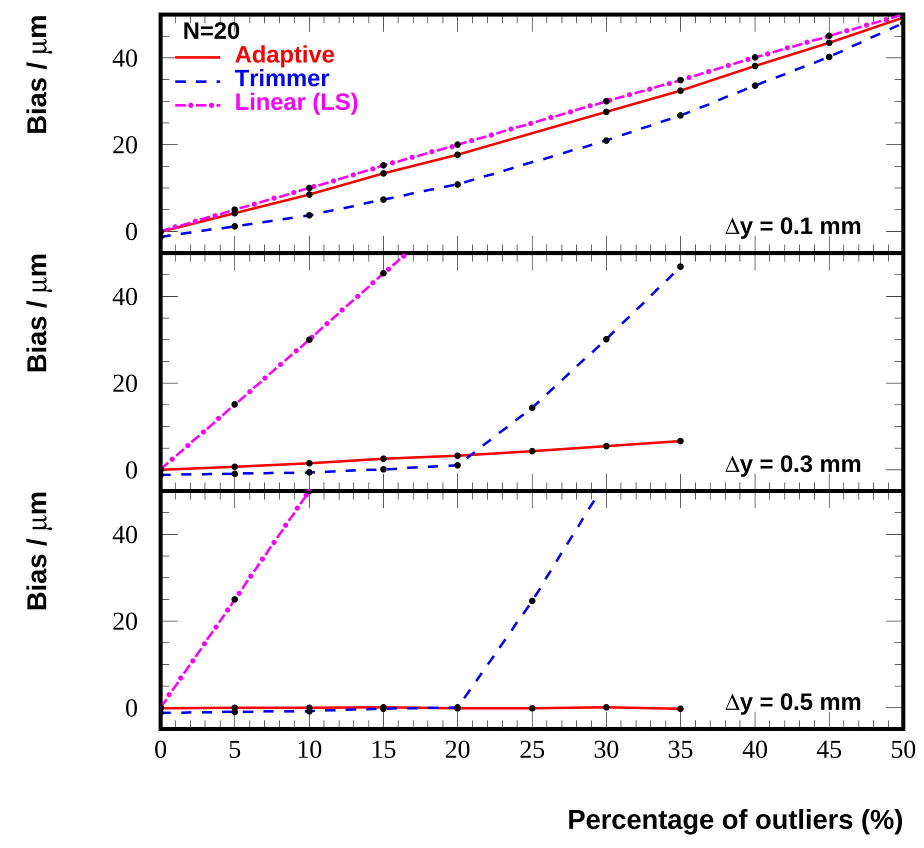

Figure 4: The average bias on the vertex position estimated with different algorithms as a function of the fraction of Type-2 outliers in the sample of in total 20 tracks. This both in the direction longitudinal and transversal to the global boost of the tracks and for several separations Δy between the primary (inliers) and secondary (Type-2 outliers) source of tracks.

In each case 10000 events are simulated in which the vertex position and uncertainty on the position are estimated. For a perfect fit algorithm the observed χ^2_{obs} should follow a uniform probability distribution P($\chi^2_{\text{ndf}} \geq \chi^2_{\text{obs}}$) between 0 and 1. The average of this distribution over the 10000 events should therefore be 0.5. The pull distribution should follow a normal Gaussian distribution centred at zero and with unity width (i.e. Root Mean Square or RMS = 1). The bias is estimated from the shift in the $\vec{r}_{\text{fit}} - \vec{r}_{\text{gen}}$ distribution (the distribution of the residuals) away from zero. The resolution is estimated as the width of a Gaussian fit on this distribution. The fit with a Gaussian function is motivated by the Gaussian smearing which is applied on the track parameters. The obtained resolution can be compared with the Minimum Variance Bound (MVB) resolution which reflects the optimal information present in the events.

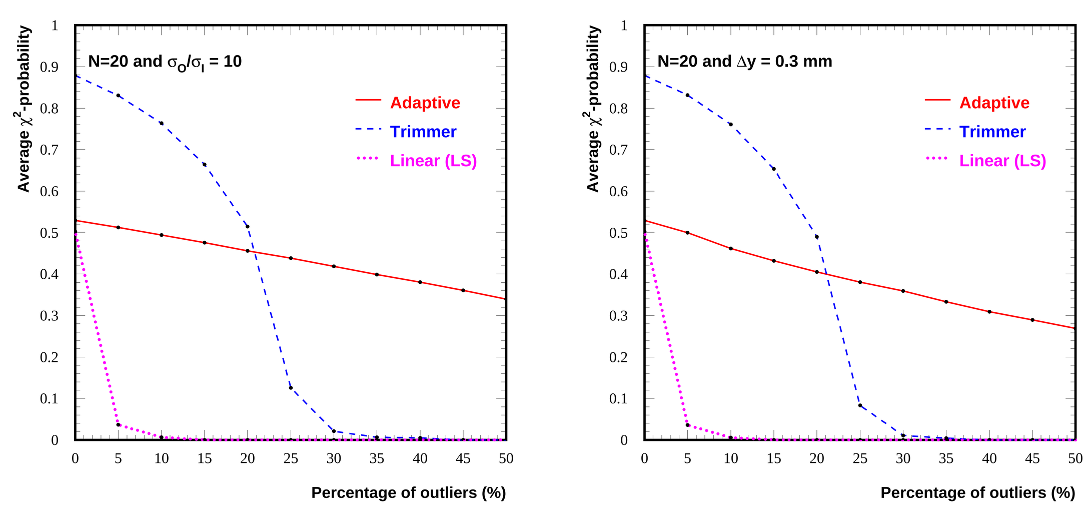

Figure 5: The average χ^2 probability of different vertex fitters over all events as a function of the fraction of Type-1 (left) and Type-2 (right) outliers.

4 Conclusion

It was found that the most robust algorithm is provided by the adaptive fitter, while the LTS fitter performes equally well only when the trimming fraction is set equal to the apriori unknown number of outliers associated to the vertex. The improvement of using robust fitters becomes more pronounced when the track parameters of the outliers are more different from the inliers as it becomes easier to reject them from the track sample. This was observed for track multiplicities associated to the vertex ranging from 3 to 20.

The above study was performed to understand the major properties and features of the different vertex fitting algorithms. In a real physics case at for example the LHC we will observe much more complicated events. Hence one should execute a certain vertex finder which associates tracks to possibly different vertices, before starting the fit to estimate the vertex position. In a following paper the statistical properties of the vertex estimators will be studied at the level of detector simulated tracks in realistic $pp \rightarrow q\bar{q}$ events.

References

[1] R. Frühwirth et al., *New developments in vertex reconstruction for CMS*, Nucl. Instrum. Methods **A502**(2003) 699.

[2] R. Frühwirth et al., *Vertex reconstruction and track bundling at the LEP collider using robust algorithms*, Computer Physics Communications 96 (1996) 189.