# Adaptive Multi-track fitting

*R. Frühwirth*[1], *A. Strandlie*[2,3]

[1] Institute for High Energy Physics, Vienna, Austria
[2] Department of Technology, Gjøvik College, Norway
[3] Department of Physics, University of Oslo, Norway

**Abstract**

We address the problem of estimating the parameters of very close or overlapping tracks. In order to achieve optimal performance the estimation process has to be carried out concurrently with the task of assigning the observations to the track candidates. We present a new method based on the Deterministic Annealing Filter (DAF) which implements a global competition. The method is studied on simulated tracks in the ATLAS Transition Radiation Tracker, in which there is a large background of mirror hits in addition to the genuine observations. We show that the new method is superior to a sequential application of the DAF to the track candidates, and also superior to a simple competition either between hits or between tracks. An even better performance can be achieved if prior knowledge about the relation between hits and their mirror hits is taken into account.

Keywords: Adaptive track fitting, assignment problem

## 1 Introduction

Track reconstruction in the LHC detectors will be faced by the problem of finding and fitting high-momentum tracks in the presence of a large background of noise, backscattering, and low-momentum tracks. In some cases the problem is aggravated by the presence of mirror hits, for example in the Transition Radiation Tracker (TRT) of the ATLAS detector [1]. It has been shown previously [2] that for tracks which are isolated fairly well an iterated non-linear filter with annealing (the DAF) is a powerful tool for estimating the track parameters concurrently with solving the assignment problem, i.e. deciding which hits are "signal" and which are "noise".

The assignment problem is even more difficult if there are several tracks close to each other, again on a high background. In this case it is conceivable that a global decision rule performs better than pure competition between hits or pure competition between tracks. In the following we present such a global competition scheme and show that it performs according to this expectation. It is hardly surprising that using additional information on which hits are mirror images of each other results in a further improvement of the performance of the decision rule.

## 2 Multi-track filters

Suppose that in a given layer $k$ of the detector there are $m_k$ track candidates (templates) and $n_k$ hits which are compatible with the candidates. In order to keep the notation as simple as possible we keep $k$ fixed and drop the subscript $k$, so that we have $m$ templates and $n$ hits. We can describe the state (position, direction, curvature) of template $j$ by a state vector $\boldsymbol{x}_j$ and the observed values of hit $i$ by an observation vector $\boldsymbol{y}_i$. A state and an observation which belong to the same track are coupled by the measurement equation

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{\varepsilon},$$

where $\varepsilon$ is the observation error, which is assumed to be normal with zero mean and covariance matrix $\boldsymbol{V}$.

As a preparation for computing the assignment probability $p_{ij}$ between hit $i$ and template $j$ we set up the following matrix $\Phi$:

$$(\Phi)_{ij} = \varphi_{ij} = \varphi(\boldsymbol{y}_i; \boldsymbol{H}\boldsymbol{x}_j, \boldsymbol{V}),$$

where $\varphi(\cdot; \boldsymbol{\mu}, \boldsymbol{V})$ is the multivariate normal probability density with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{V}$. We now define four methods for computing the assignment probabilities.

**Method 1: Competition between hits.** There is competition between all hits for each track, but there is no competition between the tracks. This procedure is equivalent to the DAF [2]. The assignment probabilities are computed by dividing $\varphi_{ij}$ by its associated column sum plus a constant (normalization by columns):

$$p_{ij} = \frac{\varphi_{ij}}{\sum_k \varphi_{kj} + c}.$$

**Method 2: Competition between tracks.** There is competition between all tracks for each hit, but there is no competition between the hits. It is equivalent to the original Elastic Arms algorithm [3]. The assignment probabilities are computed by dividing $\varphi_{ij}$ by its associated row sum plus a constant (normalization by rows):

$$p_{ij} = \frac{\varphi_{ij}}{\sum_l \varphi_{il} + c}.$$

**Method 3: Global competition.** There is competition between all entries which are incompatible. This method is proposed here for the first time. The assignment probabilities are computed by dividing $\varphi_{ij}$ by the sum of all elements in the same row and column plus a constant (normalization by columns and rows):

$$p_{ij} = \frac{\varphi_{ij}}{\sum_k \varphi_{kj} + \sum_l \varphi_{il} - \varphi_{ij} + c}.$$

**Method 4: Competition between tracks and between mirror hits.** This is a refinement of Methods 2 by adding competition between a hit and its mirror hit. It is based on Lindströms algorithm [4]. The assignment probabilities are computed separately for each of pair of hit and mirror hit. If $(i_1, i_2)$ is such a pair, the assignment probabilities are computed according to

$$p_{i_k j} = \frac{\varphi_{i_k j}}{\sum_l \sum_\alpha \varphi_{i_\alpha l} + c}.$$

The normalization constant is therefore the sum of all elements in the respective submatrix plus a constant. If for some reason a hit has no mirror hit it is treated according to method 2. In a detector without mirror hits methods 2 and 4 coincide.

The constant $c$ effectively defines a cut beyond which the assignment probability quickly drops to 0. All observations with non-zero assignment probability are combined to a single observation by a weighted mean, the weights being proportional to the respective assignment probabilities. This combined observation is then used in the updating step of the filter. The filter is iterated until the assignment probabilities settle to their final values. In order to avoid suboptimal solutions (local minima) the iteration is carried out with annealing, similar to the Determistic Annealing Filter described in [2].

## 3 Simulation experiments in the ATLAS TRT

We now present results from simulation experiments with tracks in the ATLAS Inner Detector Transition Radiation Tracker (TRT). As the TRT is made of drift tubes each hit has a corresponding mirror hit. For details of the detector the reader is invited to consult the design report [1].

We have simulated a sample of $980$ pairs of tracks all coming from the origin, giving $1960$ tracks in total. A pair is simulated by first assigning to both tracks the same transverse momentum vector $\boldsymbol{p}_T$, the magnitude of this vector being larger than or equal to $1$ GeV/c. Each component of the momentum vector is then given a random perturbation of in the order of $1\ \%$ of its magnitude. This is done independently for both tracks. The outcome of this procedure gives everything from completely separated tracks to completely overlapping tracks.

Our multi-track algorithms require an initial guess of the parameters of the tracks in the pair, and we have in this work tested out two different initialization procedures. The first one is simply to initialize the track parameters by their true values. The second one makes a least-squares fit in the $R\Phi$-projection to all points in the track pair. This yields the "centre of gravity"-track (cog). The two initial tracks are the cog-track plus/minus one standard deviation of the track parameters. This procedure is called the "cog$+/-$" initialization.

The annealing schedule has to be chosen very carefully, in order to prevent the two templates from coming too close during the annealing, as they will have a tendency to remain close during the rest of the annealing process. Therefore the starting temperature cannot be too high, and the annealing cannot be too slow. We also have observed that the convergence of the algorithms at the final temperature is slower when treating multiple tracks. The number of iterations at the final temperature has therefore been increased to $15$.

We now compare the results of the four different methods presented above. The precision of the estimates will be assessed by the generalized variance $V$ of the residuals of the estimated track parameters with respect to the truth values. We first make an evaluation of the algorithms with the mirror hits turned off (Table I). The generalized variances are given relative to a single-track fit with correct assignment. All multi-track algorithms are initialized with the true parameter values. Methods $2$ and $4$ are the best and give equivalent results, as expected. Method $3$ behaves slightly worse than methods $2$ and $4$. Obviously the power of the general competition really arises in situations where there are noise hits in addition to track hits in a layer.

We now study the methods with the mirror hits turned on plus some variable amount of noise. Table II(a) shows the generalized variances of the different algorithms as a function of the noise level, relative to the above mentioned single-track fit. The templates have been initialized to the true track parameters. Again, method $1$ is consistently the least precise of all. Method $2$ is now also significantly less precise than methods $3$ and $4$. This is due to the lack of competition between the different hits in a layer. Methods $3$ and $4$ are comparable but, nevertheless, not totally equivalent in behaviour. At low noise levels, where mirror hits are the dominant source of noise, method $4$ is the best. However, method $3$ is the best at higher noise levels, as it has been specifically designed to cope with the problem of the presence of noise hits in addition to track hits.

Table II(b) shows results which are analogous to the ones from Table II(a), but now the cog$+/-$ method has been used for the initialization. Method $1$ is again the worst, followed by method $2$. Method $3$ now consistently exhibits a larger spread of the track parameters than method

| Method | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $V_{\text{rel}}$ | 250 | 1.94 | 2.53 | 1.94 |

**Table I:** The relative generalized variance of the four methods.

4. This means that method 3 for our jet sample obviously has a tendency to fall into a local minimum of the energy function more often than method 4. However, by using a more robust measure of the spread — the product $M$ of the medians of the squared track parameter residuals — it is possible to obtain a picture similar to the one using the truth initialization. In Table III we show ratios of quantity $M$ (method 4 divided by method 3) as a function of the noise level. The median products are comparable. Moreover, the qualitative feature of method 3 increasing its relative precision with increasing noise level is now apparent also with the cog$+/-$ initialization when invoking the robust quantities. Nevertheless, method 3 obviously strongly depends on starting up from reasonably good initial tracks in order to achieve its optimal performance.

| Noise level | Method 1 | Method 2 | Method 3 | Method 4 |
|---|---|---|---|---|
| 0 | 281 | 36.2 | 4.52 | 2.84 |
| 10 % | 270 | 58.7 | 5.35 | 4.35 |
| 20 % | 388 | 101 | 6.26 | 7.06 |
| 30 % | 358 | 185 | 7.19 | 9.51 |

(a)

| Noise level | Method 1 | Method 2 | Method 3 | Method 4 |
|---|---|---|---|---|
| 0 | 7.50E5 | 1.84E3 | 72.1 | 11.9 |
| 10 % | 7.53E5 | 2.70E3 | 102 | 21.9 |
| 20 % | 9.68E5 | 3.53E3 | 314 | 40.5 |
| 30 % | 9.80E5 | 3.73E3 | 249 | 43.3 |

(b)

**Table II:** The relative generalized variance with mirror hits and noise for four different algorithms. The algorithms have been initialized by the true parameter values (a) and by the cog$+/-$ method (b), respectively.

| Initialization method | Noise level 0 | 10 % | 20 % | 30 % |
|---|---|---|---|---|
| true | 0.95 | 1.12 | 1.26 | 1.22 |
| cog$+/-$ | 0.85 | 0.96 | 1.04 | 1.01 |

**Table III:** Ratio of median products $M$ (method 4 divided by method 3) as a function of the noise level and of the initialization method.

## 4 Conclusions and outlook

We have introduced a new decision scheme for implementing a global competition between hits and tracks. It has been shown to work better than existing schemes in the presence of mirror hits and high noise levels. We intend to pursue this approach in detectors without mirror hits, in particular the CMS Inner Tracker.

## References


1 *ATLAS Inner Detector Technical Design Report.* CERN/LHCC 97-16, CERN, Geneva, 1997.

2 R. Frühwirth and A. Strandlie, Track fitting with ambiguities and noise: a study of elastic tracking and non-linear filters. *Computer Physics Communications* 120 (1999) 197.

3 M. Ohlsson, C. Peterson and A. Yuille, Track finding with deformable templates — the elastic arms approach. *Computer Physics Communications* 71 (1992) 77.

4 M. Lindström, Track reconstruction in the ATLAS detector using elastic arms. *Nuclear Instruments and Methods* A357 (1995) 129.