Estimation of the material budget from tracker data to improve the quality of the track fit

Moritz Nadler, Rudolf Frühwirth

Last corrections: July 18, 2011

Contents

1 Motivation and Introduction 2

2 Estimation Methods 2
- 2.1 Global linear estimator . . . 3
- 2.2 Forward-backward Kalman filter . . . 5

3 Simulation Experiments 7
- 3.1 The program . . . 7
- 3.2 Setup . . . 7
- 3.3 Results from Global linear estimator . . . 8
- 3.4 Results from forward-backward Kalman filter . . . 8

4 Validation with Real Data 12
- 4.1 Setup . . . 12
- 4.2 Results . . . 13

5 Conclusions and Outlook 14

We present two methods to improve the reconstruction of particle tracks by explicit estimates of the amount of material passed by the tracks. The first one is based on a linearized least-squares estimator, the second one on the combination of a forward and a backward Kalman filter. Our simulation studies show that it is possible to obtain a chi-square distribution that is indistinguishable from the one using the true material information. One of the methods has been validated using test beam data. Results of this validation are presented.

1 Motivation and Introduction

We investigate the possibility to extract information on the amount of material in a tracking detector by using the position measurements in the detector, and to improve the track reconstruction with this information. There are two main scenarios where this could be useful in track reconstruction: First, a detailed description of the materials in the detector is available, but the track reconstruction uses a simplified model. Such a simplified model might have some shortcomings that deteriorate the momentum resolution and the distribution of test quantities such as pulls and the total χ^2 of the track. Second, a device with unknown material budget is put in the way of the particles, for instance in a test beam. We have developed methods that are able to improve track reconstruction by explicit estimation of the material budget and to produce a satisfactory total χ^2-distribution of the reconstructed tracks.

As an assumption about the material budget has to be used in the reconstruction of the tracks, the method is necessarily iterative. If the unknown thickness of the material layers is considered as a latent variable, the iteration can be interpreted as an expectation-maximization (EM) algorithm [1]. In this interpretation the expectation step corresponds to the estimation of the thickness, and the maximization step to the estimation of the track parameters, using the most recent estimate of the thickness.

The methods proposed here use charged tracks that interact with the material of the detector only by multiple scattering and energy loss by ionization. The methods can and should be regarded as a complement to other approaches, which use pair production of electrons or strong interactions of hadrons for obtaining information about the distribution and amount of material present in a detector. The main restriction of our approach is the assumption that the material is concentrated in the measurement layers.

2 Estimation Methods

The basic idea of material estimation is to exploit the effects of the material on the particle trajectories. In the following we use a cylindrical detector model in a constant magnetic field. The track model is therefore a helix with five parameters to describe the trajectory between material layers. The parameters are $(\Phi, z, \theta, \beta, \kappa)$, where Φ is the azimuth angle of a point on the track, z is the Cartesian coordinate parallel to the beam, θ is the polar angle between the momentum vector and the z-axis, $\beta = \phi - \Phi$, where ϕ is the azimuthal angle of the momentum vector, and κ is the curvature of the projection of the helix on the x-y-plane. The estimation algorithm is of course independent of the particular track model and can easily be reformulated in a different coordinate system.

The mapping of the track parameters onto the measurements of one layer $\boldsymbol{m}_k$ is given by

$$\mathbf{H}_k = \begin{pmatrix} R_k & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}, \tag{1}$$

where R_k is the radius of the kth layer.

We use two material effects in this investigation:

(i) multiple scattering, as described by the Highland formula [2, 3]:

$$\begin{aligned}\sigma_{\mathrm{ms},\theta} &= \frac{0.015E}{p^2}\sqrt{\frac{X}{\sin(\theta)\cos(\beta)}}\left(1+0.038\ln\left(\frac{X}{\sin(\theta)\cos(\beta)}\right)\right),\\ \sigma_{\mathrm{ms},\beta} &= \frac{\sigma_{\mathrm{ms},\theta}}{\sin(\theta)},\end{aligned} \tag{2}$$

where X is the nominal thickness of the material traversed in units of radiation lengths, E the total particle energy and p the particle momentum; and
(ii) energy loss, as described by the Bethe-Bloch formula [4]:

$$\Delta E(E) = \frac{n_e e^4}{4\pi m_e \epsilon_0^2}\left(\frac{1}{1-\frac{m^2}{E^2}}\right)\left(\ln\left(\frac{2m_e(E^2-m^2)}{Im^2}\right)+\frac{m^2}{E^2}-1\right)\Delta x, \tag{3}$$

where n_e is electron volume density, e is elementary charge, ϵ_0 is the electric constant, m is the passing particle's mass, m_e is the electron mass and I the mean excitation potential. Eqs. (2) and (3) use SI units with the exception GeV for E, GeV/c for p and GeV/c^2 for m. Our method only estimates the amount of material. The type of material must be known to the algorithm.

The estimation of the layer thicknesses $\boldsymbol{X}$ from reconstructed track data is a non-linear problem because the track reconstruction depends on the knowledge of $\boldsymbol{X}$. A common approach to this kind of problem is to apply an iteration scheme that resembles an EM algorithm [1]. It starts with a prior guess of $\boldsymbol{X}_0$, then uses a linear estimator to get $\boldsymbol{X}_1$. This is reiterated until a convergence criterion is met. The linear estimator is derived by Taylor-expanding the non-linear model to first order.

Two linear estimators were developed to work in conjunction with the iteration scheme. The first one is a global linear estimator that estimates directly all $\boldsymbol{X}$, all scattering angles $\boldsymbol{\alpha}$ and all their correlations. It can be considered as an extended track fit with breakpoints [5, 6]. The second one is an indirect estimation of the material effects in every layer from energy loss and multiple scattering, by combining a forward and backward Kalman filter [7].

2.1 Global linear estimator

The global linear estimator uses a linearized model that models the position measurements as a function of the initial track parameters $\boldsymbol{p}_1$, the thicknesses $\boldsymbol{X}$, and the actual scattering angles $\boldsymbol{\alpha}$. It incorporates prior information about the scattering angles by using their mean value (zero) as a virtual measurement and their variance in the covariance matrix of the model. The model therefore reads:

$$\begin{pmatrix}\boldsymbol{m}\\ \boldsymbol{0}\end{pmatrix} = \begin{pmatrix}\mathbf{B} & \mathbf{A} & \mathbf{D}\\ \mathbf{0} & \mathbf{0} & \mathbf{1}\end{pmatrix}\begin{pmatrix}\boldsymbol{p}_1\\ \boldsymbol{X}\\ \boldsymbol{\alpha}\end{pmatrix} + \boldsymbol{\epsilon} \quad \text{with } \mathrm{Cov}(\boldsymbol{\epsilon}) = \mathbf{E} = \begin{pmatrix}\mathbf{V} & \mathbf{0}\\ \mathbf{0} & \mathbf{Q}\end{pmatrix}. \tag{4}$$

The left hand side of Eq. (4) are the position measurements $\boldsymbol{m}$, consisting of the $(R\Phi, z)$ measurements of the L detector layers, and the expectation values of the scattering angles $\boldsymbol{\alpha}$, which are all equal to zero. The covariance matrix $\mathbf{E}$ of the error vector $\boldsymbol{\epsilon}$ is diagonal because both $\mathbf{V}$,

which holds the variances of the $(R\Phi, z)$ measurements, and $\mathbf{Q}$, which holds the variances of the multiple scattering angles, are diagonal.

The parameters to be estimated are (i) the five initial track parameters $\boldsymbol{p}_1$, (ii) the layer thicknesses $\boldsymbol{X}$, and (iii) the scattering angles $\boldsymbol{\alpha}$ in all layers. By convention a measurement is taken at the beginning of the material of a layer therefore there is no measurement after the material of the last layer and the its material cannot be estimated. $\mathbf{B}$, $\mathbf{A}$ and $\mathbf{D}$ are the Jacobian matrices of $\boldsymbol{m}$ with respect to the estimated parameters:

$$\mathbf{B} = \frac{\partial \boldsymbol{m}}{\partial \boldsymbol{p}_1}, \quad \mathbf{A} = \frac{\partial \boldsymbol{m}}{\partial \boldsymbol{X}}, \quad \mathbf{D} = \frac{\partial \boldsymbol{m}}{\partial \boldsymbol{\alpha}}. \tag{5}$$

To get expressions for $\mathbf{B}$, $\mathbf{A}$, and $\mathbf{D}$ that can be implemented one has to write them as products of partial derivatives. $\mathbf{B}$ is assembled from L submatrices $\mathbf{b}_k \in \mathbb{R}^{2\times 5}$, $k = 1, \ldots, L$:

$$\mathbf{b}_k = \frac{\partial \boldsymbol{m}_k}{\partial \boldsymbol{p}_1} = \frac{\partial \boldsymbol{m}_k}{\partial \boldsymbol{p}_k} \prod_{i=k-1}^{1} \frac{\partial \boldsymbol{p}_{i+1}}{\partial \boldsymbol{p}_{\mathrm{ME},i}} \frac{\partial \boldsymbol{p}_{\mathrm{ME},i}}{\partial \boldsymbol{p}_i}, \tag{6}$$

where $\boldsymbol{p}_i$ is the track state in layer i before material effects and $\boldsymbol{p}_{\mathrm{ME},i}$ is the track state after material effects. $\partial \boldsymbol{m}_k / \partial \boldsymbol{p}_k$ is equal to $\mathbf{H}_k$, the mapping of the track state on the measurements; $\partial \boldsymbol{p}_{i+1} / \partial \boldsymbol{p}_{\mathrm{ME},i}$ is equal to $\mathbf{F}_i$, the propagation matrix extrapolating the track parameter from layer i to layer $i+1$; and

$$\frac{\partial \boldsymbol{p}_{\mathrm{ME},i}}{\partial \boldsymbol{p}_i} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \frac{\partial \kappa_{\mathrm{ME},i}}{\partial \kappa_i} \end{pmatrix}, \tag{7}$$

because κ is the only track parameter that is changed by the material effects. Note that the linear model is computed from a reference track that does not incorporate multiple scattering.

$\mathbf{A}$ is assembled from $L(L-1)$ submatrices $\boldsymbol{a}_{i,j} \in \mathbb{R}^{2\times 1}$, $i = 1, \ldots, L$, $j = 1, \ldots, L-1$:

$$\boldsymbol{a}_{i,j} = \frac{\partial \boldsymbol{m}_i}{\partial X_j} = \begin{cases} \dfrac{\partial \boldsymbol{m}_i}{\partial \boldsymbol{p}_i} \left(\displaystyle\prod_{k=i-1}^{j+1} \frac{\partial \boldsymbol{p}_{k+1}}{\partial \boldsymbol{p}_{\mathrm{ME},k}} \frac{\partial \boldsymbol{p}_{\mathrm{ME},k}}{\partial \boldsymbol{p}_k} \right) \dfrac{\partial \boldsymbol{p}_{j+1}}{\partial \boldsymbol{p}_{\mathrm{ME},j}} \dfrac{\partial \boldsymbol{p}_{\mathrm{ME},j}}{\partial X_j} & \text{if } i > j+1, \\ \dfrac{\partial \boldsymbol{m}_i}{\partial \boldsymbol{p}_i} \dfrac{\partial \boldsymbol{p}_i}{\partial \boldsymbol{p}_{\mathrm{ME},j}} \dfrac{\partial \boldsymbol{p}_{\mathrm{ME},j}}{\partial X_j} & \text{if } i = j+1. \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

$\partial \boldsymbol{m}_i / \partial \boldsymbol{p}_i$ and $\partial \boldsymbol{p}_{k+1} / \partial \boldsymbol{p}_{\mathrm{ME},k}$ are the same as in $\mathbf{B}$, and

$$\frac{\partial \boldsymbol{p}_{\mathrm{ME},j}}{\partial X_j} = \begin{pmatrix} 0 & 0 & 0 & 0 & \dfrac{\partial \kappa_{\mathrm{ME},j}}{\partial X_j} \end{pmatrix}^{\mathrm{T}}. \tag{9}$$

$\mathbf{D}$ is assembled from $L(L-1)$ submatrices $\mathbf{d}_{i,j} \in \mathbb{R}^{2\times 2}$, $i = 1,\ldots,L$, $j = 1,\ldots,L-1$:

$$\mathbf{d}_{i,j} = \frac{\partial \boldsymbol{m}_i}{\partial \boldsymbol{\alpha}_j} = \begin{cases} \dfrac{\partial \boldsymbol{m}_i}{\partial \boldsymbol{p}_i}\left(\prod_{k=i-1}^{j+1} \dfrac{\partial \boldsymbol{p}_{k+1}}{\partial \boldsymbol{p}_{\mathrm{ME},k}}\dfrac{\partial \boldsymbol{p}_{\mathrm{ME},k}}{\partial \boldsymbol{p}_k}\right)\dfrac{\partial \boldsymbol{p}_{j+1}}{\partial \boldsymbol{p}_{\mathrm{ME},j}}\dfrac{\partial \boldsymbol{p}_{\mathrm{ME},j}}{\partial \boldsymbol{\alpha}_j} & \forall\, i > j+1 \\ \dfrac{\partial \boldsymbol{m}_i}{\partial \boldsymbol{p}_i}\dfrac{\partial \boldsymbol{p}_i}{\partial \boldsymbol{p}_{\mathrm{ME},j}}\dfrac{\partial \boldsymbol{p}_{\mathrm{ME},j}}{\partial \boldsymbol{\alpha}_j} & \forall\, i = j+1 \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

$\partial \boldsymbol{m}_i/\partial \boldsymbol{p}_i$ and $\partial \boldsymbol{p}_{k+1}/\partial \boldsymbol{p}_{\mathrm{ME},k}$ are the same as in $\mathbf{B}$ and $\mathbf{A}$, and

$$\frac{\partial \boldsymbol{p}_{\mathrm{ME},j}}{\partial \boldsymbol{\alpha}_j} = \begin{pmatrix} 0 & 0 & 1 & 0 & -\kappa_j \dfrac{\cos(\theta_j)}{\sin(\theta_j)} \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}^{\mathrm{T}}. \tag{11}$$

To fully assemble $\mathbf{M}$ one therefore needs to calculate two Jacobians numerically per track and layer (Eqs. (7) and (9)), because all the matrices $\mathbf{F}_k$ are already calculated during the generation of the reference track.

Finally, the least-squares estimator is given by

$$\begin{pmatrix} \boldsymbol{p}_1 \\ \boldsymbol{X} \\ \boldsymbol{\alpha} \end{pmatrix} = \left(\mathbf{M}^{\mathrm{T}}\mathbf{E}^{-1}\mathbf{M}\right)^{-1}\mathbf{M}^{\mathrm{T}}\mathbf{E}^{-1}\begin{pmatrix} \boldsymbol{m} \\ \mathbf{0} \end{pmatrix} \quad \text{with } \mathbf{M} = \begin{pmatrix} \mathbf{B} & \mathbf{A} & \mathbf{D} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{pmatrix}. \tag{12}$$

2.2 Forward-backward Kalman filter

To estimate the layer thickness with a Kalman filter, first one has to estimate the track state in every layer before the material effects with the forward filter, and after the material effects with the backward filter. Then the differences $\Delta\beta$, $\Delta\theta$ and ΔE have to be calculated. It is important that in the layer where the current difference is computed the update step is done either by the forward or by the backward filter, so that every measurement is used exactly once when forming the difference. In our implementation of the estimator the forward filter includes the measurement of the current layer.

The differences in energy, θ and β can be used to get three largely independent estimates of X. In order to compute ΔE, one has to transform the two κ values to the momentum p and then to the energy E. Because a linear approximation of the highly non-linear transformation $p = h(\kappa) = \left|\frac{B_z}{\kappa \sin(\theta)}\right|$ introduces a bias in p, p was Taylor-expanded up to second order in κ:

$$p(\kappa) \approx h(\kappa_0) + h'(\kappa_0)(\kappa - \kappa_0) + \frac{1}{2}h''(\kappa_0)(\kappa - \kappa_0)^2. \tag{13}$$

If κ_0 is chosen as the local estimate of κ, which is the mean of its posterior distribution, taking the expectation in Eq. 13 gives the second-order transformation

$$p = \frac{B_z}{|\kappa|\sin(\theta)} + \frac{B_z \sigma_\kappa^2}{\sin(\theta)|\kappa^3|}. \tag{14}$$

The second-order variance can be computed by using the relation $\sigma_p^2 = \mathrm{E}(p^2) - \mathrm{E}(p)^2$. If Eq. 13 is substituted in this relation, in the Gaussian assumption the variance turns out to be

$$\sigma_p^2 = \left(\frac{B_z}{\sin(\theta)}\right)^2 \frac{\sigma_\kappa^2}{\kappa^4} + 2\left(\frac{B_z}{\sin(\theta)}\right)^2 \frac{\sigma_\kappa^4}{\kappa^6} + \left(\frac{B_z \cos(\theta)}{\kappa \sin^2(\theta)}\right)^2 \sigma_\theta^2, \tag{15}$$

where the last term is the contribution of θ to the variance of p. The final transformation from p to E uses standard linear error propagation:

$$E = \sqrt{p^2 + m^2}, \quad \sigma_E^2 = p\left(p^2 + m^2\right)^{-\frac{1}{2}} \sigma_p^2. \tag{16}$$

The variance of the difference ΔE is equal to $\sigma_{\Delta E}^2 = \sigma_{E,\mathrm{f}}^2 + \sigma_{E,\mathrm{b}}^2$, i.e., the sum of the variances of the forward and the backward filter estimates. At this point the Bethe-Bloch formula (Eq. (3)) is used to compute Δx, the layer thickness in meters. E, the particle energy before the material effects, is also needed in Eq. (3); therefore the correlation between ΔE and E has to be taken into account when calculating the variance on Δx:

$$\sigma_{\Delta x}^2 = \left(\frac{\mathrm{d}\Delta x}{\mathrm{d}\Delta E} \quad \frac{\mathrm{d}\Delta x}{\mathrm{d}E}\right) \mathrm{Cov}((\Delta E, E)) \begin{pmatrix} \frac{\mathrm{d}\Delta x}{\mathrm{d}\Delta E} \\ \frac{\mathrm{d}\Delta x}{\mathrm{d}E} \end{pmatrix}, \tag{17}$$

with

$$\mathrm{Cov}((\Delta E, E)) = \begin{pmatrix} \sigma_{\Delta E}^2 & \sigma_{E,\mathrm{f}}^2 \\ \sigma_{E,\mathrm{f}}^2 & \sigma_{E,\mathrm{f}}^2 \end{pmatrix}. \tag{18}$$

After a final transformation from Δx to X and $\sigma_{\Delta x}$ to σ_X, the estimated thickness from N tracks is computed by the weighted mean

$$\bar{X} = \frac{\sum_i^N \frac{X_i}{\sigma_{X,i}^2}}{\sum_i^N \frac{1}{\sigma_{X,i}^2}}.$$

For the multiple-scattering based estimation of X a maximum-likelihood estimator was developed. In each layer, the observed angle $\Delta\theta$ between the track segments estimated by the forward and the backward filter, respectively, is assumed to be normally distributed with mean 0 and a variance σ^2. This can be written as

$$f(\Delta\theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\Delta\theta^2}{2\sigma^2}\right), \tag{19}$$

where σ^2 is the sum of σ_{ms}^2, the multiple scattering variance given by the Highland formula (Eq. (2)), and σ_{meas}^2, the variance of the measurement error of $\Delta\theta$, which is just the sum of the variances of θ from the forward and the backward filter. The log-likelihood function therefore reads:

$$\ln L(X) = \ln \prod_i f_i = K + \sum_i \left[-\ln(\sigma_{\mathrm{ms},i}^2 + \sigma_{\mathrm{meas},i}^2) - \frac{\Delta\theta_i^2}{\sigma_{\mathrm{ms},i}^2 + \sigma_{\mathrm{meas},i}^2}\right], \tag{20}$$

where the product runs over all reconstructed tracks. $\sigma_{\mathrm{ms},i}^2$ is replaced by the Highland formula, and the log-likelihood function is maximized with respect to X. K includes all terms in $\ln L(X)$ that do not depend on X and therefore are irrelevant for the maximization. Solving the equation $\ln L - \ln L_{\max} + \frac{1}{2} = 0$ for X gives the standard deviation of the maximum-likelihood estimator.

The log-likelihood estimator of X that uses the $\Delta\beta$ is done by repeating this process and replacing $\Delta\theta$ with $\Delta\beta$ and σ_{ms}^2 with $\sigma_{\mathrm{ms},\beta}^2$.

3 Simulation Experiments

3.1 The program

The MATLAB [8] program has three main parts: (i) the "simulation" part to create the true particle trajectories or to generate a reference track from given starting values; (ii) the "estMat" part which iteratively estimates the thicknesses $\boldsymbol{X}$ using a forward-backward Kalman filter; and (iii) the "globalEst" part that implements the global linear regression on the thicknesses $\boldsymbol{X}$ and the scattering angles $\boldsymbol{\alpha}$.

The detector simulation starts with the five true track parameters in the innermost layer and propagates them through the detector. It is assumed that the material to be estimated is concentrated in the measurement layers. In each layer, the energy loss is calculated with the Bethe-Bloch formula (Eq. (3)), and random scattering angles are drawn from a normal distribution with a standard deviation according to the Highland formula (Eq. (2)). Then the track parameters are propagated to the next layer. As a simplification, the multiple scattering uses the momentum after the full energy loss. The position measurements are smeared by normally distributed errors.

Both the Kalman filter and the global regression model need a reference track. It is constructed by adding random errors to the true state vector in the first layer, and propagating the track through all layers, using the helix model with energy loss but without multiple scattering. Part of the reference track generation is the calculation of the propagation matrices $\mathbf{F}_k$ that are the Jacobians of the non-linear track propagator.

The iteration scheme in "estMat" is dampened in order to avoid large oscillations of the estimated thicknesses. In each iteration, the thickness used as an input for the next iteration is 0.34 times the current estimate plus 0.66 times the previous estimate. Other numerical values for the dampening were tried but this ones lead to a fast convergence without risking a breakdown of the iteration scheme.

3.2 Setup

As a test case for our material estimation methods a simplified model of the CMS silicon tracker was implemented. Table 1 shows the parameters of this model. Apart from the X values they were taken from [9]. The thicknesses were chosen to add up to the total amount of material stated in [9], i.e., 0.4 radiation lengths of silicon. One layer was assigned a different value in order to check if the estimation methods could recognize the difference.

The simulated events consist of single randomly chosen π^+ or π^- particles. The initial true track parameters are selected from uniform random distribution, the ranges of which are given in Table 2. The transverse momentum p_T was chosen as 1 GeV for all tracks, and the homogeneous magnetic field B_z was set to 4 T.

The results shown in this section use some ideal conditions: perfect alignment, exact knowledge of detector resolution and perfect particle identification. Additional tests to check the influence of wrong particle identification or underestimation of the detector resolution were also conducted.

Layer	R [m]	$\sigma_{R\Phi}$ [µm·rad]	σ_z [mm]	X [X_0]
1	0.044	20	0.02	0.03
2	0.073	20	0.02	0.03
3	0.102	20	0.02	0.03
4	0.225	23	0.23	0.03
5	0.339	23	0.23	0.03
6	0.4185	35	34	0.04
7	0.498	35	34	0.03
8	0.608	53	0.53	0.03
9	0.692	53	0.53	0.03
10	0.78	53	52	0.03
11	0.868	53	52	0.03
12	0.965	35	52	0.03
13	1.08	35	52	0.03

Table 1: Parameters used in the simplified CMS silicon tracker model. From left to right: layer number, radius R, resolution in $R\Phi$, resolution in z and the thickness of the layer as a fraction of the radiation length.

	Φ [rad]	z [m]	θ [rad]
min	0	−0.1	$\pi/4$
max	2π	0.1	$\pi/2$

Table 2: Range of the uniform distributions that are used to generate the initial track parameters.

3.3 Results from Global linear estimator

Table 3 shows the results of the estimation of $\boldsymbol{X}$ with the global estimator. No iteration scheme was used, and the true thicknesses were used as the initial values. Table 3 shows that the global estimator, while working in principle, requires a very large number of tracks—in the order of 10^7—to estimate $\boldsymbol{X}$ with sufficient accuracy.

The final estimate of $\boldsymbol{X}$ is a weighted mean over the estimates from all tracks. Table 3 shows the weighted means, their standard error and their standard scores with respect to the true values. The errors scale as expected with the inverse square-root of the number of tracks. With 10^5 tracks the errors are still so large that negative estimates of the thickness occur. With 10^7 tracks the errors are a few percent of the thickness, therefore the estimates can be considered as satisfactory.

An inspection of the joint covariance matrix of all estimates reveals very strong correlations between the thicknesses, up to 0.97. Adjacent layers always have negative correlation, so that there is strict alternation between positive and negative correlations.This explains the alternating signs of the standard scores in Table 3.

3.4 Results from forward-backward Kalman filter

Table 4 shows the result of the material estimation with a single step of the forward-backward Kalman with the true thicknesses as the initial values. The true values are reproduced within the statistical errors. Note that the estimator requires two well-defined track segments from both filters, so that there are no reliable estimates in the innermost two and outermost two layers. The

tracks	10^5			10^6			10^7		
layer	X	σ_X	z_X	X	σ_X	z_X	X	σ_X	z_X
3	424	90	1.37	328	29	0.971	286	9	−1.510
4	154	76	−1.93	295	24	−0.196	307	8	0.950
5	576	117	2.36	328	37	0.759	306	12	0.475
6	−81	157	−3.07	344	50	−1.130	397	17	−0.198
7	812	151	3.38	371	48	1.480	301	15	0.088
8	−175	158	−3.02	231	50	−1.380	306	16	0.368
9	819	181	2.87	370	57	1.220	298	18	−0.088
10	−128	182	−2.36	257	57	−0.732	304	18	0.260
11	642	173	1.98	346	55	0.850	309	17	0.531
12	48	145	−1.73	263	46	−0.807	296	15	−0.246

Table 3: Results of the estimation of X, its standard error (both in units of $10^{-4}X_0$), and standard scores z_X calculated with the global linear estimator for different track numbers. See Table 1 for the true values of X.

tracks	10^4				10^5			
layer	X	σ_X	z_X	r_X	X	σ_X	z_X	r_X
2	287	6	−2.01	−0.043	291	2	−4.40	−0.030
3	300	7	−0.01	−0.000	306	2	2.40	0.018
4	278	8	−2.81	−0.073	302	3	0.65	0.005
5	305	10	0.55	0.018	305	3	1.73	0.018
6	423	9	2.51	0.059	402	3	0.75	0.005
7	309	9	0.92	0.029	303	3	0.97	0.010
8	298	9	−0.19	−0.005	302	3	0.67	0.006
9	313	10	1.40	0.043	307	3	2.31	0.023
10	310	12	0.82	0.032	300	4	0.04	0.000

Table 4: Results of the estimation of X its σ_X (both in $10^{-4}X_0$), the standard score z_X, and the relative error $r_X = (X - X_{\text{tr}})/X_{\text{tr}}$ calculated with the forward-backward Kalman estimator for different track numbers. The track parameters used for the material estimation were reconstructed with the true $\boldsymbol{X}$ values. They can be found in Table 1

fact that there is no estimate of X in layer 11 is only due to the poor resolution in the outer layers. A test run with smaller resolution produced an estimate in layer 11 with the same quality as the other estimates.

The comparison of Table 4 with Table 3 shows that the estimator based on the Kalman filter needs much less tracks to achieve the same accuracy. This can be explained by noting that the scattering angles are estimated from two independent track segments, without injecting any prior information about their variance. Additionally only the direct estimation of $\boldsymbol{X}$ was used in the global regression model. Actually, 10^4 tracks of low momentum are sufficient to estimate the thickness to a relative precision of a few percent.

If the starting values of the thickness are different from the true ones, the iteration scheme with dampening has to be applied. A test was performed with starting thicknesses of 2% in all layers. The improvement of the track reconstruction in the first four iteration is illustrated by

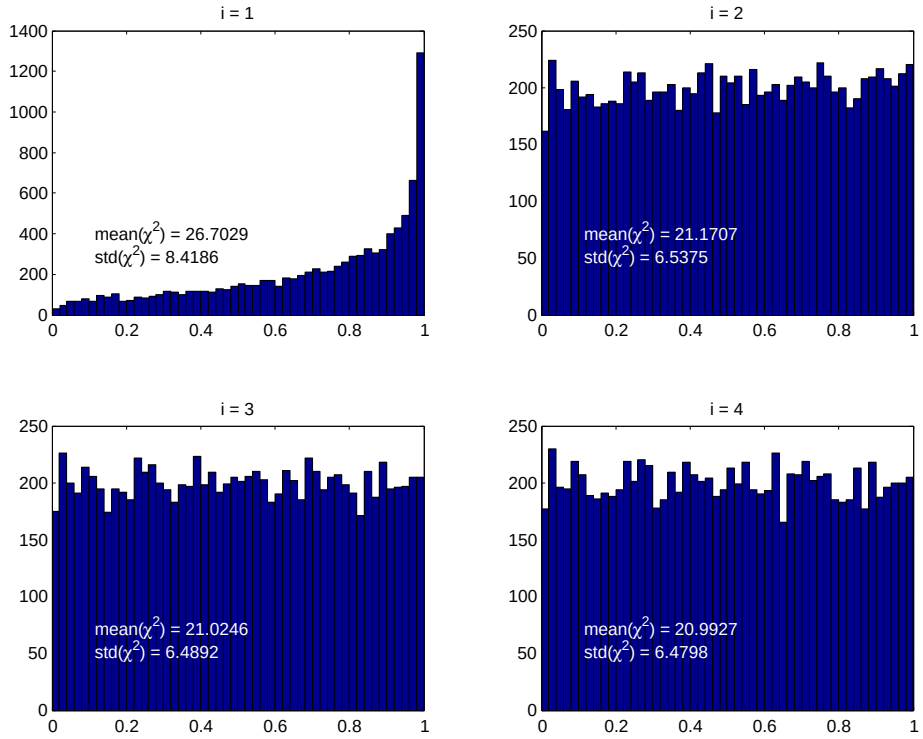

Figure 1: Distribution of the p-values of the total χ^2 of the forward filter, for the first four iterations. The p-values should be uniformly distributed.

the distributions of the p-values of the total χ^2 of the forward filter (Figure 1). In the course of four iterations, the distribution quickly converges to a nearly perfect uniform one. The estimated thicknesses are shown in Table 5. The four iterations took approximately 10 minutes on an Intel Core i3 CPU with 3.33 GHz.

The iteration is stopped as soon as the mean of the total χ^2 and the smoothed χ^2 values of the individual layers stabilize. We have observed that further iterations actually tend to degrade the smoothed χ^2-distributions and estimated thicknesses while the total χ^2-distribution and the total material budgets remain stable. The reason for this behavior is the presence of strong correlations between adjacent layers which are not taken into account in the Kalman filter approach.

Table 6 shows the weights used to combine the information from the three different X estimates. Two of them are based on the multiple scattering in θ and β, respectively, and the third one on the change of κ due to energy loss. The table shows the result of the 4th iteration. One can see that most of the information comes from the θ scattering in the inner layers and from β in the outer layers. This is a direct result of the good z resolution of the inner pixel detectors and the bad z resolution of the outer strip detectors.

The relative contribution of the estimation from energy loss to the weighted mean is very small

i \layer	2	3	4	5	6	7	8	9	10	total
0	200	200	200	200	200	200	200	200	200	2900
1	274	303	333	355	345	322	299	283	241	3965
2	287	287	303	365	364	330	318	313	256	4030
3	301	289	276	362	371	324	317	327	268	4043
4	311	303	259	357	377	319	310	332	277	4056
10	269	376	242	327	389	331	299	319	292	4056

Table 5: Estimated X values in $10^{-4} X_0$ for the first four and the 10th iteration. Because X cannot be estimated in layers 1, 11, 12 and 13, they are set to the true value. The row $i = 0$ contains the starting values (a-priori guesses) of X. See Table 1 for the true values. The sum of all X including the non-estimated ones is also shown.

layer	2	3	4	5	6	7	8	9	10
θ	0.988	0.890	0.629	0.068	0.063	0.070	0.000	0.000	0.000
β	0.012	0.109	0.361	0.893	0.897	0.887	0.963	0.967	0.967
κ	0.000	0.001	0.011	0.039	0.040	0.043	0.037	0.033	0.033

Table 6: Weights (normalised inverse variances) of the X estimates in all layers after the 4th iteration. The weights are used to combine information from multiple scattering in θ and β and from energy loss via the change in κ.

and never exceeds 5 %. The contribution is even less when using particles with higher momenta. The results shown here would therefore hardly change if the estimation from energy loss would be discarded.

In general the material estimation works better with low momentum tracks because the material effects are stronger. Higher momentum can to some extent be compensated by using more tracks for the estimation. Table 7 shows a comparison of (A) the estimated thicknesses of the standard setup, (B) the estimates from 10000 tracks with a $p_T = 10\,\text{GeV}$, and (C) estimates from 40000 tracks with $p_T = 10\,\text{GeV}$. The resulting total $\chi^2_{\text{tot,f}}$ is still perfect with 10000 tracks with $p_T = 10\,\text{GeV}$, having a mean of 20.9076 and an rms of 6.4097 after four iterations.

We also have tested the effect of wrong assumptions on the detector resolution and on the particle mass. If the detector resolution used in the track reconstruction is 10 % better than the one used to generate the measurements, the average $\chi^2_{\text{tot,f}}$ is 22.8238, and the rms is 7.0804, somewhat too large. The estimated thicknesses are shown in setup D. If the generated particles are muons and the Kalman filter uses the pion mass instead (setup E), the effect is very small: the average $(\chi^2_{\text{tot,f}})$ is 20.9645, and the rms is 6.4715.

For the results in this section, σ^2_θ was neglected in Eq. (15). This term increases σ^2_p and through error propagation σ^2_X, the variance of the X estimate from energy loss. This effect is visible mostly in the outer layers because of their lower z resolution. As a consequence, the estimate of X from energy loss has even smaller weight, and the results in this section are virtually the same if σ^2_θ is included in Eq. (15).

setup	\layer	2	3	4	5	6	7	8	9	10
A	X	333	330	225	348	389	310	297	341	293
	std(X)	6.73	7.25	8.23	9.64	9.14	9.59	9.03	9.46	11.8
B	X	502	320	172	345	229	387	470	72.7	720
	std(X)	51.5	52.7	39.5	28.2	28	46.5	49.2	855	191
C	X	274	344	165	371	322	388	213	440	257
	std(X)	25.6	26.1	19.7	13.8	13.3	22.7	25.9	40.5	101
D	X	332	362	214	369	387	341	300	340	306
	std(X)	6.74	7.23	8.3	9.63	9.17	9.55	9.03	9.46	11.7
E	X	331	329	222	347	387	309	296	341	288
	std(X)	6.69	7.21	8.19	9.59	9.09	9.55	8.99	9.41	11.7

Table 7: Estimated X values and their error in $10^{-4} X_0$ after four iterations, for different setups (see Table 1 for the true values). A is the standard setup. B uses p_T = 10 GeV instead of 1 GeV. C uses 40000 tracks with p_T = 10 GeV. In setup D the filter uses a resolution that is 10 % better than the one used in generating the measurements. In setup D the generated particles are muons while the filter assumes the pion mass for reconstruction. In comparison to Table 5 this table shows the direct estimation results before they get dampened with the previous estimation.

Layer	z [cm]	σ_x [µm]	σ_y [µm]	X [$10^{-3} X_0$]
1	0	4.3	4.3	7.33
2	15	4.3	4.3	7.33
3	30	4.3	4.3	7.33
4	39.27	114	14.4	?
5	44.17	111	11	?
6	50.17	111	12.8	?
7	68	4.3	4.3	7.33
8	83	4.3	4.3	7.33
9	98	4.3	4.3	7.33

Table 8: Parameters known about the test-beam setup. From left to right: layer number, distance in beam direction z, resolution σ in x, resolution σ in y and the thickness of the layer X

4 Validation with Real Data

4.1 Setup

Thanks to the generous assistance of the ATLAS 3D collaboration and the ATLAS IBL collaboration, we were able to validate the forward-backward Kalman filter estimator on real data, taken in a test beam with the EUDET telescope. The setup was similar to the one described in [10]. It consisted of three layers of EUDET pixel detectors with known thickness, 3 devices under test plus infrastructure with unknown material budget, and three more EUDET pixel layers of the same type as the first three layers. The devices under test were samples of 3D silicon pixel sensors. There was no magnetic field. The beam particles were pions with an energy of 120 GeV. Table 8 shows the parameters used to estimate the material budget of the three central layers. It should be stressed that the actual sensors are only a small fraction of the material budget, which is dominated by the infrastructure.

	real data				simulation			
i \layer	4	5	6	total	4	5	6	total
0	641	641	641	1923	641	641	641	1923
1	952	946	923	2821	947	945	923	2815
2	967	954	924	2845	961	957	933	2851
3	974	955	918	2847	963	958	932	2853
4	980	955	912	2847	964	958	930	2852

Table 9: Test beam data. Estimated X values (in $10^{-4} X_0$) for the first 4 iterations. $i = 0$ are the are the initial guesses of X. The sum of all X includes only the 3 estimated layers. Left: real data; right: simulated data.

The main difference to the "CMS" setup is the missing magnetic field. As a consequence, the tracks do not have a curvature κ, and only the effect of multiple scattering can be used to estimate the thickness. We recall, however, that the contribution of energy loss is almost negligible in any case. The reference track was computed using the first position measurement, the beam direction, and the known beam momentum.

To limit the amount of outliers in the measurements used for the material estimation a preliminary track reconstruction with several cuts was applied. The cuts were tuned to give a well shaped $\chi^2_{\text{tot,f}}$-distribution after four iterations of the material estimation. This resulted in a sample of 99966 tracks.

As an additional cross-check, a sample of 99966 simulated tracks was generated in the test-beam setup, using the estimated amount of material in the three central layers.

4.2 Results

Figure 2 shows the development of the p-values of the total χ^2 of the tracks. It is obvious that the initial guess of the material budget is too small. The estimates converge to around 0.95% of a radiation length per layer, including the entire infrastructure (see Table 9). The corresponding plots with simulated tracks look very similar.

By combining the forward and the backward filter estimates, individual (smoothed) chi-squares can be obtained for each layer. Their distribution is summarized in Table 10, both for real and simulated data. While the distributions from simulated data are perfect, the real data show some discrepancies. These could be due either to deviations of the actual resolution from the nominal one, or to the fact that the material is not really concentrated in the detector layers.

Unfortunately, it was not possible to validate the material estimates by comparing them to the true values because the latter were never determined by the experimenters. In this test run preliminary and bulky cooling and support structures accounted for most of the material budget, but they have been modified in the meantime. The starting values shown in Table 9 were rough guesses by a member of the collaboration [11]. We were assured that the result of our material estimation did indeed improve the quality indicators of the track reconstruction in their analysis framework.

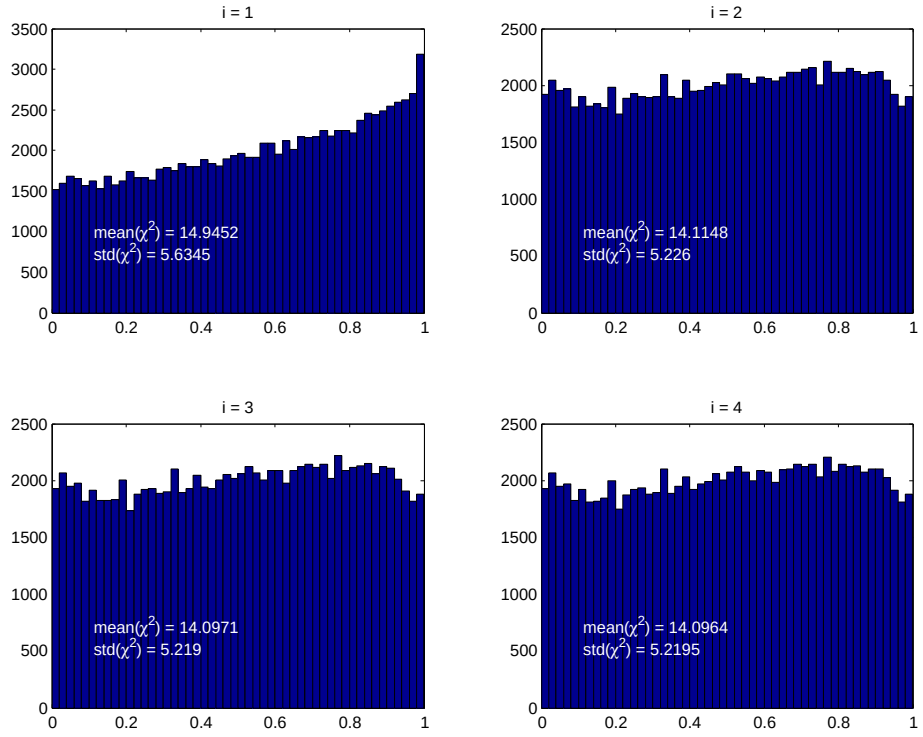

Figure 2: Test beam data. Distribution of the p-values of the total χ^2 of the forward filter, for the first four iterations. The p-values should be uniformly distributed.

layer	1	2	3	4	5	6	7	8	9
mean	1.888	1.611	2.086	2.026	2.290	2.020	2.307	1.702	1.900
rms	1.755	1.575	2.035	1.635	2.509	1.734	2.280	1.676	1.790
mean	2.079	2.007	2.004	2.002	2.012	1.993	1.991	1.996	1.997
rms	2.077	2.005	2.004	2.003	2.004	2.001	1.988	1.999	1.989

Table 10: Mean and rms of the smoothed χ^2-distributions in every layer, after four iterations. The two top rows are from real data, the two bottom rows are from simulated data.

5 Conclusions and Outlook

We have presented two methods for estimating the amount of material in a detector using charged tracks. Simulation studies show that only the forward-backward Kalman filter approach is feasible in terms of number of tracks required to obtain satisfactory estimate. We have been able to validate the latter method with test-beam data, in spite of a relatively high beam momentum of 120 GeV. Four iterations over the track sample are sufficient to obtain a reliable estimate of the total material budget and a nearly perfect χ^2-distribution.

We observe that the individual thicknesses do not always converge to their true values, although the total material budget does. In fact, there tends to be an alternating over- and underestimation of the thickness. The most likely reason for this is neglecting the correlations between the state vectors in different layers. These correlations can in principle be computed by error propagation in both the forward and the backward Kalman filter. Once they are known, they can be used to set up a global covariance matrix for the estimated multiple scattering angles $\Delta\theta$ and $\Delta\beta$ in all layers.

A restriction in our current approach is the assumption that the material is concentrated in the measurement layers. This is clearly not the case in many tracking detectors. A future development of our method will have to relax this assumption and deal with material between the measurement layers, possibly spread over a larger distance.

An unexpected outcome of our simulation studies is the fact that compared to multiple scattering energy loss by ionization contributes very little to the estimate of the material thickness. It could be worth studying whether energy loss of electrons by bremsstrahlung can make a significant contribution to the material estimate.

Acknowledgements

We thank the ATLAS 3D collaboration and the ATLAS IBL collaboration for the permission to use their test beam data for a validation of our estimation method. Special thanks are due to H. Gjersdal (Oslo) who prepared the test beam data for us and helped us in tuning the reconstruction.

References

[1] A.P. Dempster, N.M. Laird and D.B. Rubin, Maximum Likelihood from Incomplete Data via the EM Algorithm. *J. Royal Stat. Soc. B* **39**(1) (1977) 1–38.

[2] V.L. Highland, Some practical remarks on multiple scattering. *Nucl. Instrum. Meth.* **129** (1975) 497–499.

[3] K. Nakamura et al. (Particle Data Group), *J. Phys. G* **37** 075021 (2010), Section 27.3.

[4] K. Nakamura et al. (Particle Data Group), *J. Phys. G* **37** 075021 (2010), Section 27.2.

[5] R. Frühwirth et al., *Data analysis techniques for high-energy physics experiments*, 2nd ed. Cambridge University Press, Cambridge, UK, 2000.

[6] V. Blobel, A new fast track-fit algorithm based on broken lines. *Nucl. Instrum. Meth. Phys. Res. A* **566** (2006) 14–17.

[7] R. Frühwirth, Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth. Phys. Res. A* **262** (1987) 444–450.

[8] D.C. Hanselman and B.L. Littlefield, *Mastering Matlab 7*, Prentice Hall, 2005.

[9] The CMS Collaboration, The CMS experiment at the CERN LHC, *J. Instr.* **3** (2008) S08004. http://iopscience.iop.org/1748-0221/3/08/S08004

[10] P. Grenier et al., Test Beam Results of 3D Silicon Pixel Sensors for the ATLAS Upgrade. http://arxiv.org/abs/1101.4203

[11] H. Gjersdal, personal communication.