Journal of Physics: Conference Series **396** (2012) 022037 doi:10.1088/1742-6596/396/2/022037

Robust track fitting in the Belle II inner tracking detector

Moritz Nadler, Rudolf Frühwirth

E-mail: moritz.nadler@oeaw.ac.at, rudolf.fruehwirth@oeaw.ac.at

Abstract. Track fitting in the new inner tracker of the Belle II experiment uses the GENFIT package. In the latter both a standard Kalman filter and a robust extension, the deterministic annealing filter (DAF), are implemented. This contribution presents the results of a simulation experiment which examines the performance of the DAF in the inner tracker, in terms of outlier detection ability and of the impact of different kinds of background on the quality of the fitted tracks.

1. Introduction and motivation

The Belle II experiment [1] is the successor to the Belle experiment and is currently under construction at the KEK laboratory in Japan. The electron-positron collider KEKB, where Belle was taking data, delivered an instantaneous luminosity of about $2.11 \cdot 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. KEKB is in the process of being upgraded to SuperKEKB, aiming at an instantaneous luminosity of $8 \cdot 10^{35}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, an increase by a factor of forty. Many subsystems of Belle therefore have to be upgraded, too. The central tracking system will be replaced entirely in order to be able to cope with this increase. Beyond that, the hardware improvements should give the possibility of reconstructing tracks with much lower transverse momentum than before, and also better vertex resolution. This is required because of the smaller energy asymmetry of the upgraded accelerator and the resulting smaller boost. The illustration in Figure 1 shows the size and the main components of the Belle II detector.

The new inner tracking detector of Belle II has six sensor layers, two of them pixel layers and four of them double-sided Si strip layers. The radius of the outermost layer is about 14 cm. A computer generated image of the future device is shown in Figure 2. The inner tracker is surrounded by the central drift chamber, which is instrumental in measuring high-momentum tracks. Low-momentum tracks ($p_\mathrm{T} \lesssim 100\,\mathrm{MeV}/c$) are seen only in the six layers of the inner tracker, and the track reconstruction algorithms have to be designed such that they can be reconstructed with high efficiency. The main difficulties in this task are the strong material effects these tracks will suffer from and the expected high background level in the innermost layers. For this reason, a pattern recognition method using only the four Si strip layers of the inner tracker is currently under development. The track finding is presented in a separate article in these proceedings [2]. Here we discuss only the track fitting in the inner tracker. Evaluating and improving robust track fit algorithms by means of simulation studies is a natural first step in dealing with these challenges from the algorithmic point of view.

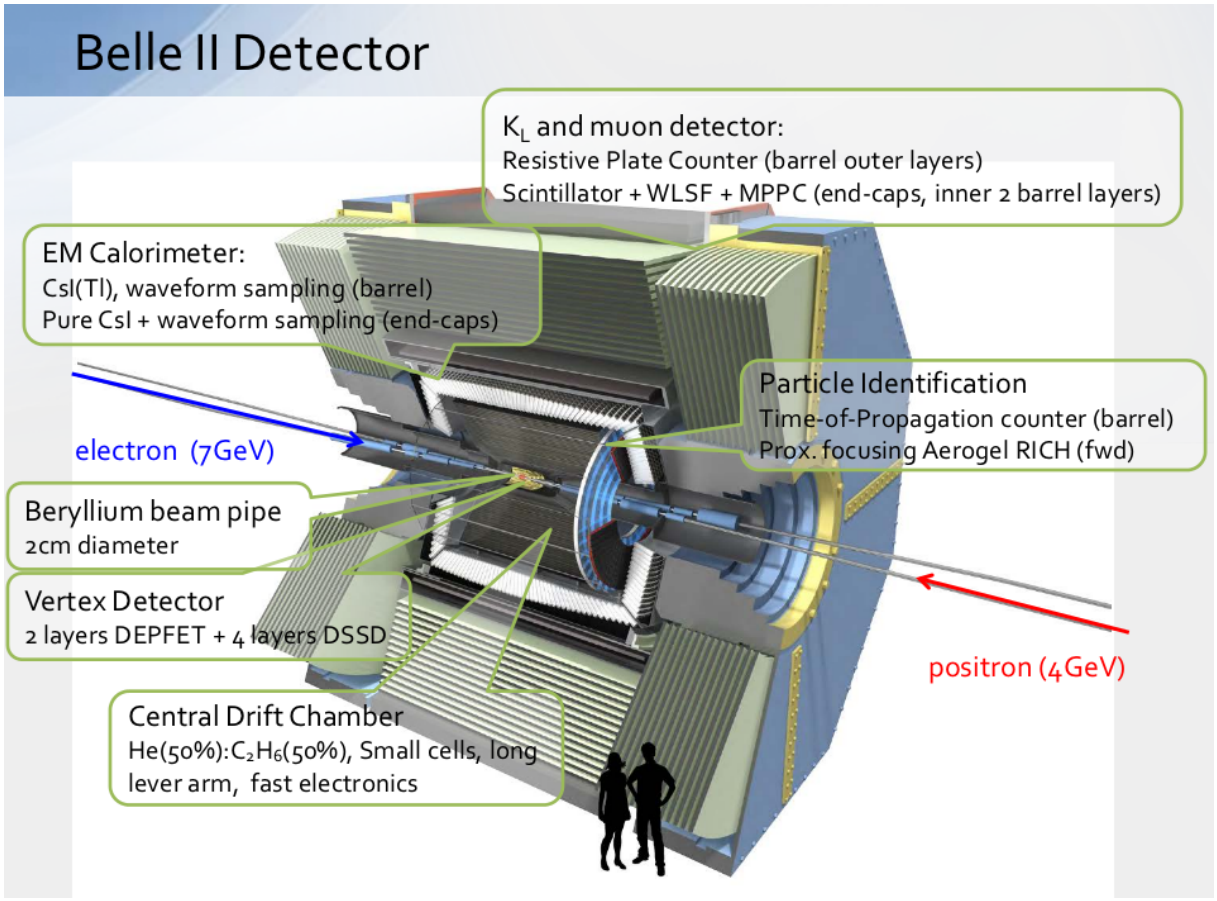

Figure 1. The main components of the Belle II detector

2. The software stack and tracking algorithms

2.1. The Belle II framework

The Belle Analysis Software Framework 2 (basf2) was started as a new software project. It incorporates ideas and concepts from other high energy physics frameworks, but does not use any legacy code. Modules for the framework can be written in C++ or Python. Its steering files are Python source files. The framework uses Geant4 [3] to simulate particle trajectories, Root [4] for data I/O (both developed at CERN) and Genfit [5] as its tracking toolkit. Genfit was developed at TU Munich and presented at CHEP 2012 in a poster session and a "lightning talk" in the plenary. Currently two track fit methods are implemented in Genfit, the Kalman filter and the Deterministic Annealing Filter (DAF) [6].

2.2. The fitting algorithms

The Kalman filter is an iterative least square estimator and is the standard method of estimating track parameters $\boldsymbol{x}$ and their covariance matrix $\boldsymbol{C}$ from tracking detectors in high energy physics. The algorithms starts with some prior information of $\boldsymbol{x}$ and $\boldsymbol{C}$ that normally comes as a rough estimate from the pattern recognition. The filter consists of an alternating prediction/update scheme. The prediction is obtained by propagating $\boldsymbol{x}$ onto the next measurement plane, using the equation of motion of a charged particle in a magnetic field, and linear error propagation of $\boldsymbol{C}$. In the propagation step the changes to $\boldsymbol{x}$ and $\boldsymbol{C}$ because of interaction with material are calculated, according to some appropriate model. The update is the correction of the prediction

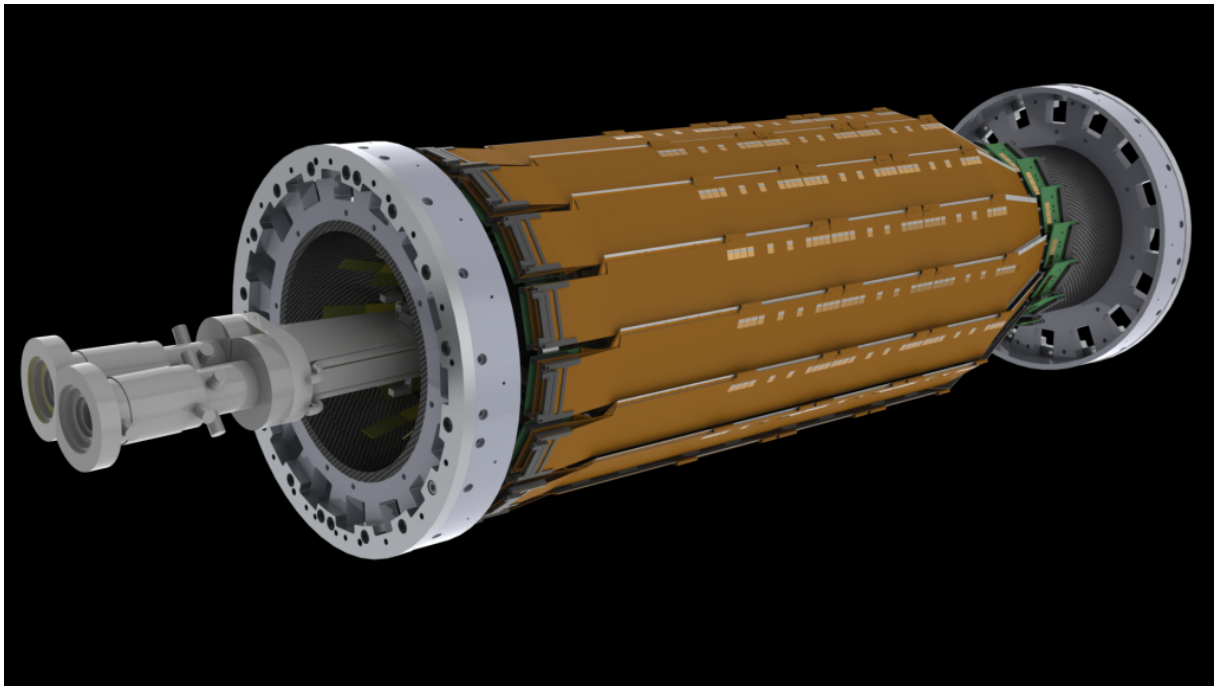

Figure 2. Computer generated image of the inner tracker. The beam pipe and the outer most Si sensor layer are visible. Diameter of the inner tracker $\approx 30\,\mathrm{cm}$, length $\approx 65\,\mathrm{cm}$

with the new measurement information, which is tantamount to a weighted average of the predicted track parameters and the measured position.

The DAF is an iterated Kalman filter with weights and annealing. It is designed for robust track fitting in the presence of outliers and background hits. In one iteration of the DAF the Kalman filter runs twice over all measurements, once in the forward and once in the backward direction. Then the information of the filters is combined in a so-called smoothing step. After that a weight is calculated for each measurement, using the residuals of the measurement and the smoothed track parameters $\boldsymbol{x}$. The weight can be interpreted as the probability of the measurement belonging to the track. This is sometimes referred to as "soft assignment". In the update steps of the following iteration the weight matrix (inverse covariance matrix) of every measurement is multiplied by the corresponding weight, so that measurements with small weight have less influence on the estimated track parameters.

The annealing increases the probability of finding the globally optimal solution even when the prior information (the track seed) is bad or the first few hits entering the Kalman filter are outliers. It is controlled by a temperature parameter. The DAF normally starts at a high temperature and lowers it after every iteration. If the final temperature is small, the weights tend to 0 or 1, and the "soft" assignment of the hit to the track becomes "hard".

3. Methods and setup

3.1. General approach

In order to study the background detection performance of the DAF in the environment of the Belle II inner tracker, first two track samples at two different particle energies were simulated. These track samples were then reconstructed using optimal conditions with the standard Kalman filter. Then artificial background was added to the track samples, and they were reconstructed again with the DAF. The parameters of the background generation were systematically changed,

Table 1. Parameters of the Belle II inner track as currently implemented in basf2. From left to right: layer number, radius R, resolution in u, resolution in v (u and v are orthogonal to each other, v is parallel to the beam direction) and the thickness t of the silicon sensors.

Layer	R [mm]	σ_u [µm]	σ_v [µm]	t [mm]
1	14	14	16	0.42
2	22	14	25	0.42
3	38	15	45	0.32
4	80	22	69	0.32
5	105	22	69	0.32
6	135	22	69	0.32

and the tracking results were compared to the results of the clean sample, using the impact parameters and momentum residuals as the comparison quantity.

To measure the background detection performance, two ratios were computed: the ratio of correctly identified background hits to the total number of background hits, and the ratio of non-background hits falsely identified as background hits to the total number of non-background hits. As the weights can take any value between 0 and 1, 0.5 was chosen as the threshold of background identification: a weight below or equal to 0.5 was interpreted as "found to be background" by the DAF, and a weight above 0.5 was interpreted as "found to be non-background" by the DAF.

3.2. Setup

Table 1 lists the characteristic properties of the inner tracker as currently used in basf2. The first two layers are pixel sensors (PXD), and the remaining four are double-sided strip sensors (SVD). The measurement resolutions given in Table 1 are simply the pitch size divided by $\sqrt{12}$. The final resolutions are not yet known but are expected to be somewhat better. Figure 3 is a schematic drawing of the the inner tracker illustrating the layout including the slanted parts in the three outermost layers, designed to minimize the material crossed by tracks in the forward direction.

The tracks were simulated and reconstructed with the following settings for particle gun, simulation and fitter algorithm:

- momentum fixed at $1\,\mathrm{GeV}/c$ or $100\,\mathrm{MeV}/c$, $\Phi \in [0°, 360°]$ uniformly distributed, $\theta = 90°$;
- either μ^+ or μ^- with equal probability;
- no other material/detectors than the SVD and PXD detectors;
- homogeneous magnetic field cut off after 15 cm, therefore no curling tracks;
- only tracks with exactly one hit in every layer;
- truth information used for "pattern recognition" and as starting values for track fit;
- other condition also optimal, e.g. only Gaussian smearing, detector resolution identical in digitizer and Genfit;
- probability cut of DAF: 0.001, annealing scheme of DAF: 1, 1, 1 (no annealing);
- about 3 800 Tracks per sample.

The magnetic field cut prevents tracks from re-entering the inner tracker. While Genfit is capable of fitting curling tracks, their efficiency is not yet optimal and they were therefore excluded. The selection of tracks with exactly one hit per layer ensures that all tracks have the same amount of information present in every layer, so that they are more easily comparable when background hits are placed in specific layers.

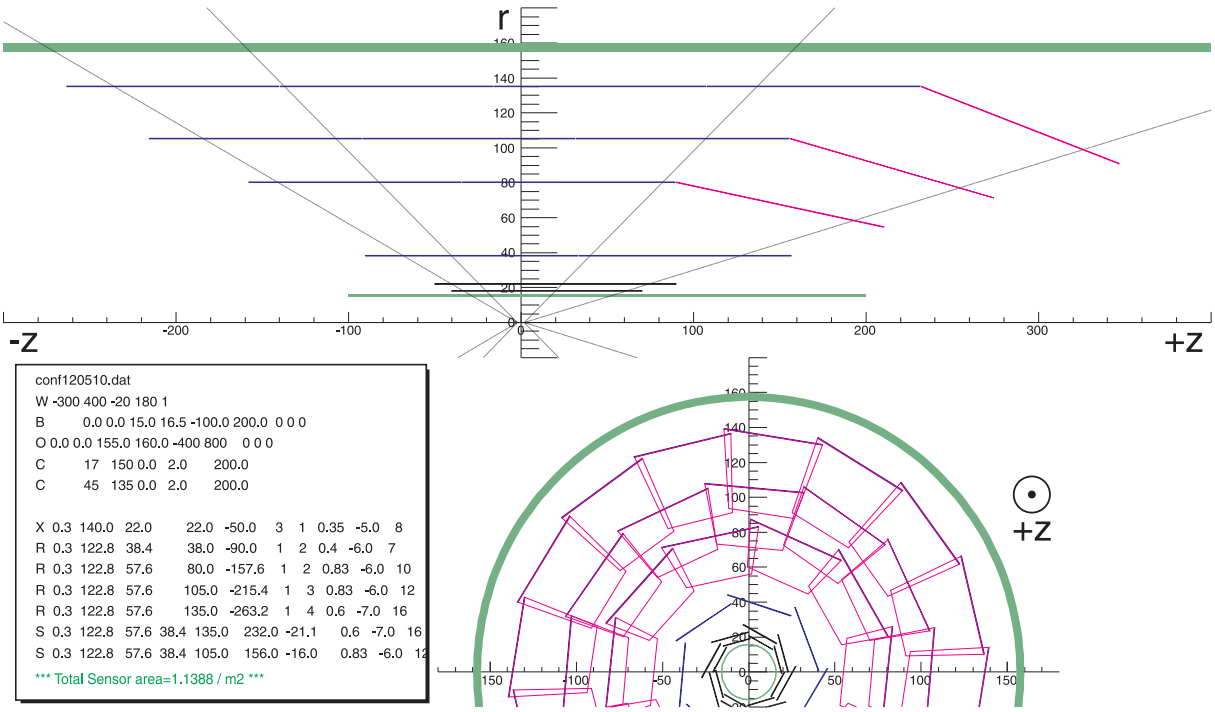

Figure 3. Schematic drawing of PXD and SVD sensor layers. Scale in mm.

3.3. Measurement and background generation

The measurements were generated by smearing the simulated true hits produced by Geant4 with a Gaussian error, using the resolutions in Table 1 as standard deviations. For the purpose of this study an artificial background and outlier generation module was implemented in basf2. In general, the module generates an additional background hit in the vicinity of a simulated hit. The input parameters of the module were set to create the following kind of background:

$$u_{\mathrm{BG}} = u_{\mathrm{true}} + a\sigma_u \sin(\phi),$$
$$v_{\mathrm{BG}} = v_{\mathrm{true}} + a\sigma_v \cos(\phi),$$

where u and v are the local coordinates of a hit in the sensor plane, and $\phi \in [0°, 360°]$ is randomly drawn from a uniform distribution. For each track sample a was set to one of three fixed values: 3, 4, or 5. Two different kind of background contaminated sample were generated: one additional background hit per track (1BG), giving a total of seven hits per track; and 6 additional background hits per track (6BG), giving a total of twelve hits per track. It has to be noted that the options and parameters described in this section where not chosen to achieve maximal realism, but to validate the Genfit implementation of the DAF and the code interfacing it with basf2.

4. Results

Figure 4 shows the ability of the DAF to detect background hits in the 1BG samples of 1 GeV/c tracks, when there is only one additional background hit per track. Figure 5 shows the rejection of good hits in the same track samples. In these figures a single data point corresponds to one fitted track sample, so that six different fitted track samples are shown in each of the figures, for each value of a. The rejection rate of good hits is shown only in the layer where the background hit was added.

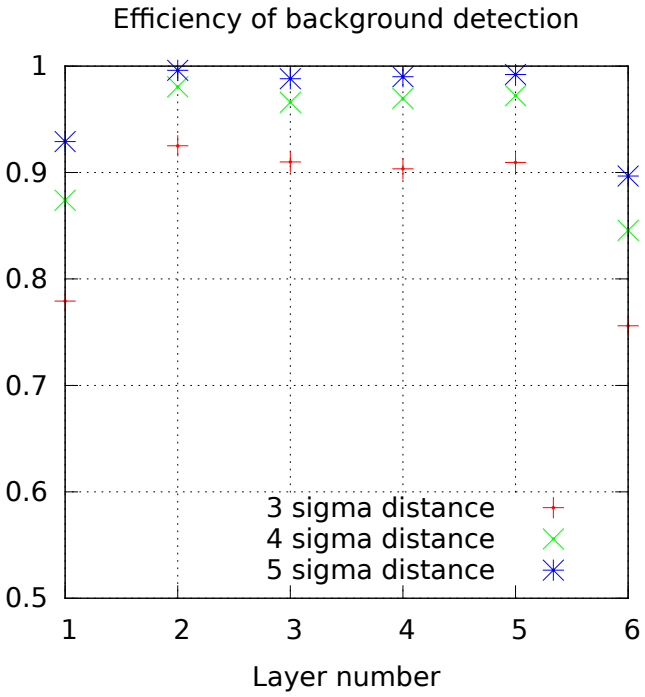

Figure 4. Results from the 1BG 1 GeV/c track samples. Each data point stands for an entire track sample.

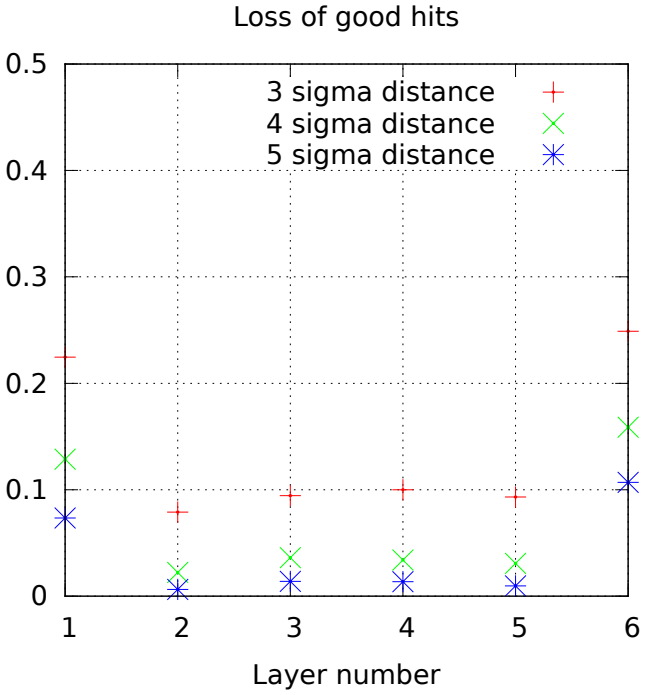

Figure 5. Results from the 1BG 1 GeV/c track samples. Each data point stands for an entire track sample.

Table 2. Standard deviations of the impact parameters and momentum residuals for the no BG and 1BG 1 GeV/c track samples. Background hits have $3\sigma_{\text{meas}}$ distance to true hit. Other distances ($a = 4$ and $a = 5$) give nearly identical result and are therefore not shown.

parameter	no BG	L1	L2	L3	L4	L5	L6
x [µm]	17	20	17	18	18	17	17
y [µm]	18	20	18	18	18	18	18
z [µm]	23	27	23	23	23	23	23
p_x [MeV/c]	30.2	30.6	29.7	31.1	31.3	30.2	30.0
p_y [MeV/c]	29.7	30.2	29.7	31.6	31.0	29.9	30.4
p_z [MeV/c]	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table 3. Standard deviations of the impact parameters and momentum residuals for the no BG and 6BG 1 GeV/c track samples, for different distances ($a = 3, 4, 5$).

parameter	no BG	$a = 3$	$a = 4$	$a = 5$
x [µm]	17	22	24	24
y [µm]	18	22	24	25
z [µm]	23	28	30	31
p_x [MeV/c]	30.2	34.0	35.7	36.1
p_y [MeV/c]	29.7	34.2	36.0	36.8
p_z [MeV/c]	1.0	1.1	1.1	1.1

As one can see in the two figures, the background detection of the DAF works very well in the 1 GeV/c case. Only if the background hit is in the last or first layer, the background detection drops below 90 %. In addition one can see that background hits with a larger distance from the real track are detected more easily.

Table 2 shows a comparison of the standard deviation of the impact parameters and momentum residuals at the track origin. A track sample without background (no BG) is compared to track samples with a single background hit in a specific layer (L1–L6). One can see that the distribution of the momentum residuals is not affected by the presence of a background hit. The impact parameters show a small deterioration only when the background hit is in the first layer. This is the case for all tested background to true hit distances ($a = 3, 4, 5$).

We now consider the 6BG samples with six background hits per track. Table 3 shows the standard deviations of impact parameters and momentum residuals. In contrast to Table 2, the results of the three different background distances are shown. The largest increase of any impact parameter standard deviation in Table 3 relative to the-non background case is 40 %. The largest increase for the momentum residuals is 24 %.

One can see in Figure 6 that in comparison to the 1BG samples the background detection efficiency has dropped slightly, and the impact parameters and momentum residuals in Table 3 are somewhat worse. A slight deterioration with rising distance a can be observed. This can be explained by the fact that the first layer has more influence on the impact parameters than the other five, and that in this layer the background detection efficiency drop is much more pronounced than in the other layers. Figure 7 shows that the misclassification of good hits as background is also somewhat more frequent.

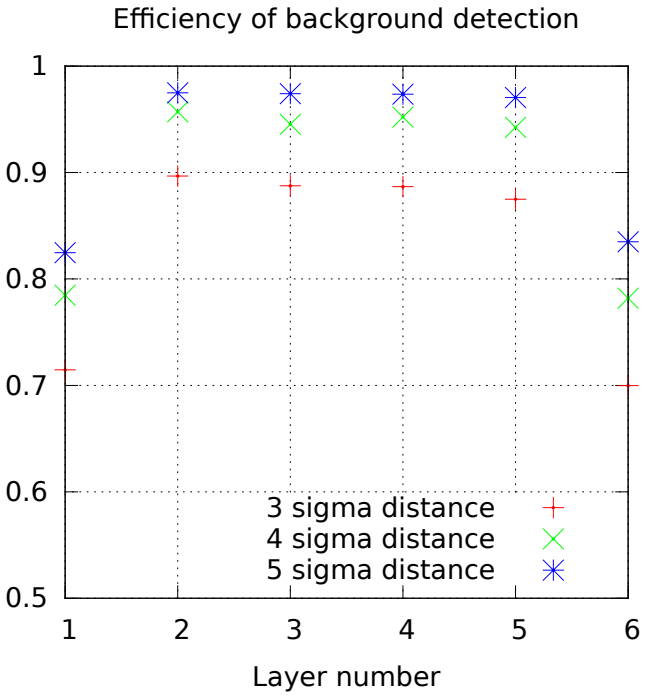

Figure 6. Results from the 6BG 1 GeV/c track samples. Each colour stands for an entire track sample.

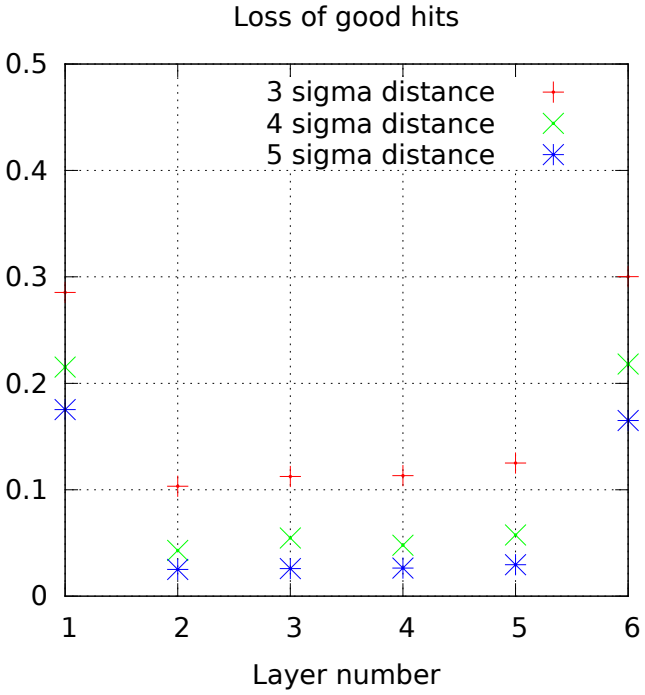

Figure 7. Results from the 6BG 1 GeV/c track samples. Each colour stands for an entire track sample.

Table 4. Standard deviations of the impact parameters and momentum residuals for the no BG and 1BG 100 MeV/c track samples. Background hits have $3\sigma_{\text{meas}}$ distance to true hit. Other distances ($a = 4$ and $a = 5$) give nearly identical result and are therefore not shown.

parameter	no BG	L1	L2	L3	L4	L5	L6
x [µm]	71	78	73	72	76	75	72
y [µm]	71	80	73	75	71	73	72
z [µm]	99	100	104	100	100	99	98
p_x [MeV/c]	3.4	3.3	3.2	3.3	3.2	3.3	3.3
p_y [MeV/c]	3.1	3.2	3.2	3.2	3.2	3.2	3.2
p_z [MeV/c]	0.7	0.8	0.7	0.8	0.7	0.8	0.8

Table 5. Standard deviations of the impact parameters and momentum residuals for the no BG and 6BG 100 MeV/c track samples, for different distances ($a = 3, 4, 5$).

parameter	no BG	$a = 3$	$a = 4$	$a = 5$
x [µm]	71	77	81	86
y [µm]	71	76	80	84
z [µm]	99	103	111	119
p_x [MeV/c]	3.4	3.4	3.4	3.3
p_y [MeV/c]	3.1	3.2	3.2	3.3
p_z [MeV/c]	0.7	0.7	0.8	0.8

Figures 8–11 and Tables 4–5 show the same information as Figures 4–7 and Tables 2–3, but for the 100 MeV/c track samples. One can see that both the background detection efficiency and the loss of good hits are worse than in the 1 GeV/c case. The corresponding residual tables, however, show that the deterioration of the reconstructed track parameters is even somewhat smaller. In the 1BG samples, both the impact parameters and the momentum residuals are affected very little, whereas in the 6BG samples the impact parameters are affected more strongly than the momentum residuals, but less than in the 1 GeV/c samples, which have a much better baseline resolution of the impact parameters.

The variation of the background detection efficiency between layers (see Figure 8) is now much larger. It is best in layer 2, because layers 1, 2 and 3 have the best measurement resolutions. On the other hand, in layer 6 the background detection is only around 60 % and therefore close to random guessing (50 %). The main reason for this large difference to the 1 GeV/c samples is multiple scattering. The random scattering angles are now so large that they produce shifts in the following layer that are in the order of 0.2 mm with respect to the undisturbed track, which is the same order of magnitude as the distance to the artificial background. Therefore the DAF has problems distinguishing these two cases.

It is important to note that in the 100 MeV/c samples the seven tracks (about 2‰) with the worst reconstruction results had to be deleted from the samples before the standard deviations were calculated. Because of numerical problems in the fit algorithm these tracks had very large residuals and therefore introduced a large bias into the estimation of the standard deviations.[1]

[1] As the authors only recognized this after the conference, the numbers shown in Table 4 and Table 5 are different from the ones shown in the conference talk.

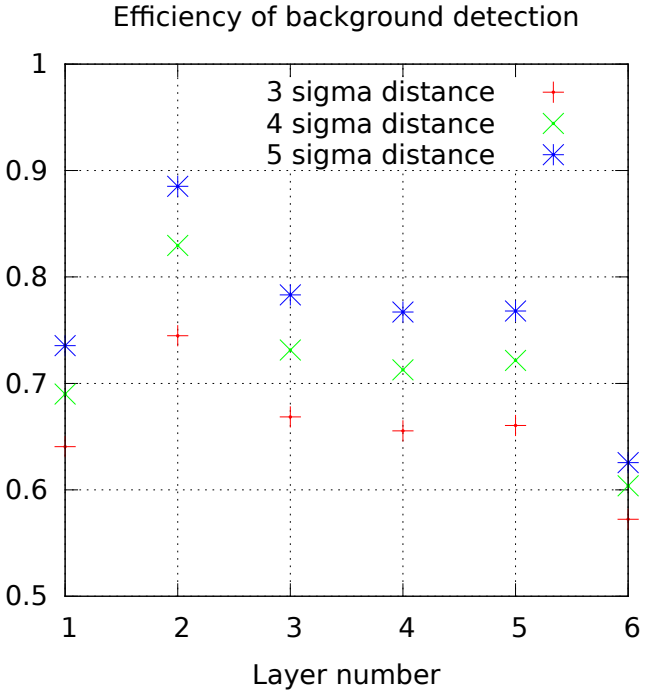

Figure 8. Results from the 1BG 100 MeV/c track samples. Each data point stands for an entire track sample.

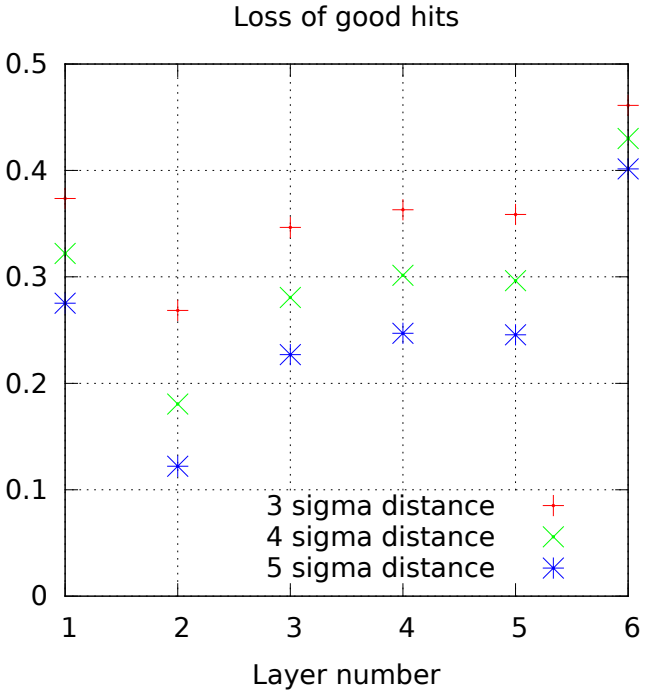

Figure 9. Results from the 1BG 100 MeV/c track samples. Each data point stands for an entire track sample.

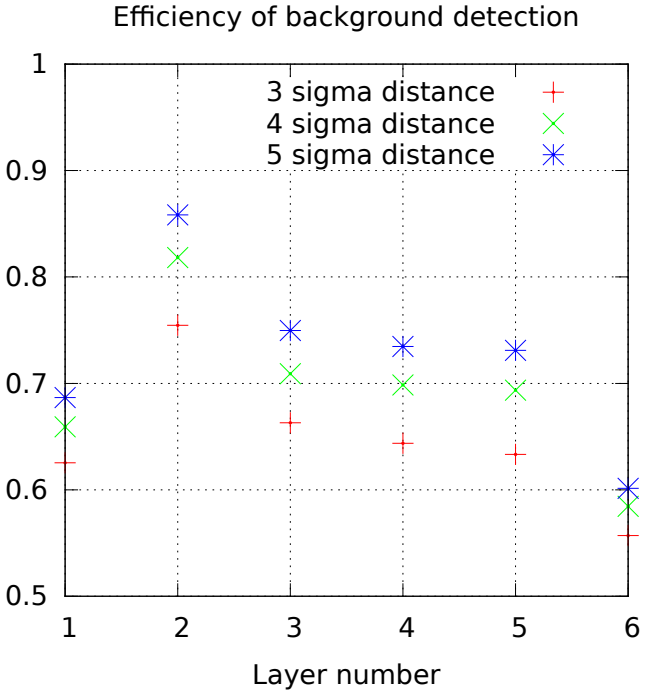

Figure 10. Results from the 6BG 100 MeV/c track samples. Each colour stands for an entire track sample.

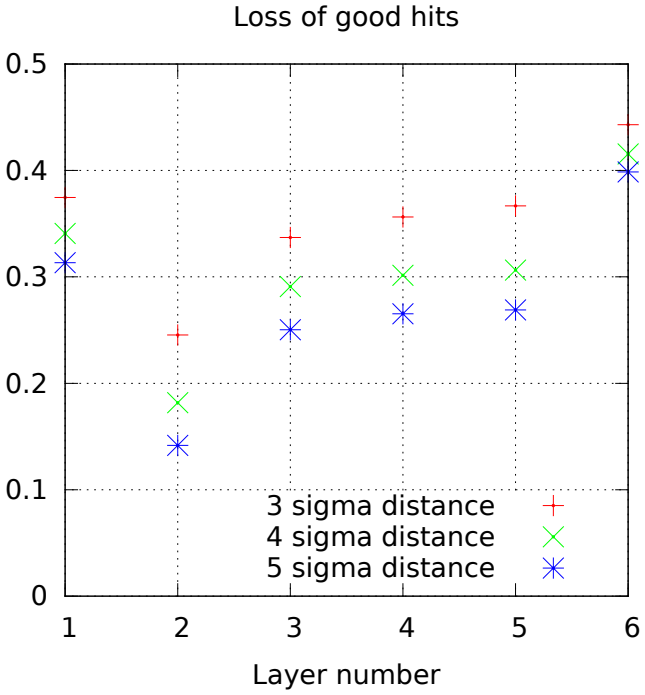

Figure 11. Results from the 6BG 100 MeV/c track samples. Each colour stands for an entire track sample.

5. Summary and Outlook

We have studied the performance of the DAF on artificially generated background hits. At moderate track energies background detection works very well. A single background hit per track in any layer besides L1 has virtually no effect on the impact parameters and the momentum estimate. If the track candidate contains as much background as real hits, the standard deviations of the impact parameters increase by 20% to 40%. The main reason is again the poorer performance of the background detection in L1. The momentum residuals are affected less, and their standard deviations increase by only 12% to 24%.

While the probability of background detection at low energies is worse, the effect on the impact parameters and the momentum residuals is smaller, because of much larger multiple scattering, and, in the case of the impact parameters, because the baseline resolution is already worse than in the $1\,\mathrm{GeV}/c$ sample.

It is of course desirable to make the tests described here more realistic. The two most important steps in this direction are:

(i) Using realistic digitization and clustering in the hit simulation, as well as simulation of ghost hits in the double-sided strip sensors. This is currently under test.

(ii) Using the SVD only pattern recognition. The background contaminated hit samples should not be fed directly to the DAF as it was done in this study, but to the pattern recognition, because in reality the track fitting will only have to deal with whatever the pattern recognition selects.

Several other aspect are interesting and will be investigated in the future, such as the possibility to improve background detection in the inner tracker with additional information from the central drift chamber or from the regions where sensors overlap in the PXD or SVD. The effect of annealing to temperatures well below $T = 1$ will also be studied.

Acknowledgements

The authors thank C. Schwanda (HEPHY Vienna) for interesting discussions and valuable feedback.

References

[1] Abe T *et al.* October 2010 *Belle II Technical Design Report* Belle II collaboration, KEK http://arXiv.org/abs/1011.0352v1
[2] Lettenbichler J, Frühwirth R and Nadler M to appear in 2012 *Proceedings of CHEP 2012* Journal of Physics: Conference Series
[3] Allison J *et al.* 2006 *IEEE Trans.Nucl.Sci.* **53** 270
[4] CERN Geneva *ROOT Reference Guide* URL: http://root.cern.ch/drupal/content/reference-guide
[5] Höppner C, Neubert S, Ketzer B and Paul S 2010 *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **620** 518
[6] Frühwirth R and Strandlie A 1999 *Computer Physics Communications* **120** 197