Reparameterizing track errors in DC06

Vladimir Gligorov, Eduardo Rodrigues

University of Glasgow

T-Rec Meeting

26 May 2008

A big thanks to Hugo and Jose for their help!
OVERVIEW

- Motivation
- Fitting procedure
- DC06 results
  - Best fit polynomials & Residuals
  - Impact on HLT selections
- Future work
WHY ARE TRACK ERRORS IMPORTANT?

Without a correct calculation of track errors, using significance cuts in the trigger introduces inefficiencies.

No time for a full fit in HLT2

- Must parameterize the track errors

**The DC04 parametrization no longer works in DC06.**

---

T-Rec meeting, CERN, 29 September 2008
FITTING PROCEDURE
ERRORS AS A FUNCTION OF PT

- Follow the same procedure as used in Hugo’s DC04 note*
- The track errors are binned as a function of $1/p_T$
- A polynomial is fitted to these errors, and used in the HLT to calculate entries in the 5x5 track covariance matrix
  - Errors in x and y are assumed uncorrelated

* Ref: Hugo Ruiz, LHCb 2005-012
CALCULATING THE ERRORS (DETAILS)

- Extrapolate the track to the same Z point as the primary vertex

- Calculate the X (or Y) deviation of the track from the PV (binned in $p_T$)

- In each bin, use an iterative procedure to estimate the core Gaussian width of the X(Y) deviation from the PV

  - Errors on the PV position are considered negligible, hence ignored
CALCULATING THE ERRORS (MORE DETAILS)

- For each bin, iterate as follows
  1. Compute RMS; reject tracks > 8xRMS from the mean
  2. Compute RMS; reject tracks > 7xRMS from the mean
  3. Compute RMS; perform Gaussian fit in ±1xRMS region; Reject tracks > 6σ from the mean
  4. Repeat 3, rejecting tracks > 5σ from the mean
  5. Repeat 3, rejecting tracks > 4σ from the mean
  6. Perform a final Gaussian fit in the ±3xRMS region
DC06 results
DATA SAMPLE

- DaVinci v19r12
- 20,000 L0 stripped minbias events
- Select events with 1PV only
- Make all particles as pions using StdNoPIDsPions
  - Only Long tracks used!
- No MC truth information used
ERRORS AS A FUNCTION OF $1/p_T$

![Graph showing the relationship between $\delta(\Delta_x)$ (\textmu m) and $1/p_T$ (GeV$^{-1}$).]
WITH THE OLD PARAMETRIZATION...
BEST FIT POLYNOMIALS
EXPLANATION

- Plots have been made with second to sixth degree polynomials

- Show the sixth degree fits here, the rest are in the backups

- Note that the first few bins are the most important since they contain the high $p_T$ signal-like tracks

  - The first bin contains all tracks with $p_T > 10$ GeV and is hence especially important
FITTING WITH A 6TH DEGREE POLY
This plot shows the residual between the fitted and measured error for each bin of $p_T$, calculated at the midpoint of the bin.
This plot shows the width of the pull in bins of $p_T$. The pull in any one bin is computed by dividing the measured $\Delta_x$ for every track in that bin by the parameterized error. For perfect agreement the widths should be equal to 1.
HOW WELL IS THE ITERATIVE PROCEDURE ACTUALLY WORKING?
This plot shows the percentage of tracks removed by the iterative procedure before the final $3\sigma$ fit.
**Final 3σ Gaussian Fit Pulls**

First bin of $p_T$

- $\chi^2 / \text{ndf} = 137.7 / 61$
- Constant: $192 \pm 3.7$
- Mean: $0.01115 \pm 0.01376$
- Width: $0.8531 \pm 0.0128$

- Refitted in 1σ region

Last bin of $p_T$

- $\chi^2 / \text{ndf} = 302.2 / 61$
- Constant: $148.3 \pm 3.6$
- Mean: $0.01128 \pm 0.01634$
- Width: $1.02 \pm 0.02$

- Refitted in 1σ region

T-Rec meeting, CERN, 29 September 2008
Even after iteratively rejecting tracks, we do not get single gaussians, especially at low $p_T$.

Are we overestimating the errors (maybe)? How do you quote a single width for something which is a double gaussian anyway?
**EFFECT OF NEW PARAM. ON HLT2 SEL.**

Test the two parametrizations on a sample of $B_s \rightarrow D_s K$ events, selected by requiring all final state particles to have a $p_T > 600$ MeV

<table>
<thead>
<tr>
<th>With DC06 Error parametrization</th>
<th>With DC04 Error parametrization</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Applied Cut</strong></td>
<td><strong>HLT2 Eff. wrt. Offline</strong></td>
</tr>
<tr>
<td>IPS &gt; 3 (all final state)</td>
<td>85%</td>
</tr>
<tr>
<td>B IPS &lt; 4</td>
<td>98%</td>
</tr>
<tr>
<td>B flight significance &gt; 8</td>
<td>80%</td>
</tr>
<tr>
<td>IPS &gt; 3 (all final state)</td>
<td>91%</td>
</tr>
<tr>
<td>B IPS &lt; 4</td>
<td>100%</td>
</tr>
<tr>
<td>B flight significance &gt; 8</td>
<td>87%</td>
</tr>
</tbody>
</table>

Unfortunately, the DC04 parametrization was underestimating the errors, so the relative efficiencies actually get worse, not better... on the other hand, the minbias rejection should get better as well.
DEPENDANCE ON OTHER VARIABLES
Errors as a function of $\eta$
Errors as a function of $\Phi$
RESIDUAL DEPENDENCIES
WHEN FITTING AGAINST $p_T$
This plot shows the width of the pull in bins of $\eta$ when the $p_T$ parametrization is assumed. For perfect agreement the widths should be equal to one.
This plot shows the width of the pull in bins of $\phi$ when the $p_T$ parametrization is assumed. For perfect agreement the widths should be equal to one.
Taking Residual Dependencies into Account

- As noted in the DC04 study, there are substantial correlations between the different residual correction.

- Hence if we want an improvement on the $p_T$ only correction, we would need a look-up table.

- In first instance, consider the variables $p_T$, phi, and eta; each split into 80 bins.
  - The table then has 512000 entries (half if you assume perfect symmetry in phi).
  - Presents certain logistical difficulties... can it be implemented like magnetic field map?
  - Do we need it? Would need more than 20,000 tracks to do the fit for the table.
CONCLUSIONS AND FUTURE WORK
CONCLUSIONS AND FUTURE WORK

- The fitter is ready for release and public use
  - How do we envisage this to be used on real data? As a monitoring algorithm?
  - Do we want or need a look up table?
- Does the iterative procedure need refining?
  - Should switch to an unbinned fit for use with real data… sadly we now have a lot of time to work on this.
- Note detailing this work on the way soon
BACKUP
FITTING WITH A QUADRATIC

\[ \sigma(\Delta x) (\mu m) \]

\[ \text{Track } 1/p_T \ (\text{GeV})^{-1} \]

T-Rec meeting, CERN, 29 September 2008
Fitting with a quadratic (zoom)
This plot shows the residual between the fitted and measured error for each bin of $p_T$, calculated at the midpoint of the bin.
This plot shows the width of the pull in bins of $p_T$. The pull in any one bin is computed by dividing the measured $\Delta_x$ for every track in that bin by the parameterized error. For perfect agreement the widths should be equal to 1.
FITTING WITH A CUBIC
FITTING WITH A CUBIC (ZOOM)
RESIDUALS WITH CUBIC FIT

This plot shows the residual between the fitted and measured error for each bin of $p_T$, calculated at the midpoint of the bin.
This plot shows the width of the pull in bins of $p_T$. The pull in any one bin is computed by dividing the measured $\Delta x$ for every track in that bin by the parameterized error. For perfect agreement the widths should be equal to 1.
FITTING WITH A QUARTIC

\[ \sigma(\Delta_x) \text{ (\(\mu m\))} \]

\[ \text{Track } 1/p_T \text{ (GeV)}^{-1} \]
FITTING WITH A QUARTIC (ZOOM)
This plot shows the residual between the fitted and measured error for each bin of $p_T$, calculated at the midpoint of the bin.
This plot shows the width of the pull in bins of $p_T$. The pull in any one bin is computed by dividing the measured $\Delta x$ for every track in that bin by the parameterized error. For perfect agreement the widths should be equal to 1.
Fitting with a Quintic
FITTING WITH A QUINTIC (ZOOM)
This plot shows the residual between the fitted and measured error for each bin of $p_T$, calculated at the midpoint of the bin.
This plot shows the width of the pull in bins of $p_T$. The pull in any one bin is computed by dividing the measured $\Delta x$ for every track in that bin by the parameterized error. For perfect agreement the widths should be equal to 1.
INTERMEDIATE $1\sigma$ GAUSSIAN FIT

- Prob: 0.0001166
- Constant: 49.28 ± 1.11
- Mean: 0.001675 ± 0.002415
- Width: 0.1303 ± 0.0026

$\Delta x$ (mm)
DC04 results

* Ref: Hugo Ruiz, LHCb 2005-012
The parametrization

Figure 4: The best-fitting 4\textsuperscript{th}-degree polynomial in two different ranges of 1/p_T.
Figure 3: Residuals of the parameterization of the uncertainties
Tracks left out of the fit

- Because of the iterative procedure, some tracks are left out of the fit

- The proportion varies with $p_T$

Figure 5: Left: fraction of tracks left out of the final fit due to the iterative process for fitting the $\sigma$ in a $\pm 3\sigma$ region. Right: size of the region where the final fit is performed.