Bachelor's Colloquium

Electron track reconstruction studies and improvement for LHCb's real-time analysis trigger

Furkan Cetin

Heidelberg University

cetin@physi.uni-heidelberg.de

April 26, 2024

Motivation

tracking at LHCb:

- tracks of charged particles are reconstructed with efficiencies well above 90%
- **but**: electron tracking underperforms significantly compared to other particles

→ due to bremsstrahlung

- interesting for *e.g.* LFU, FCNC[1]

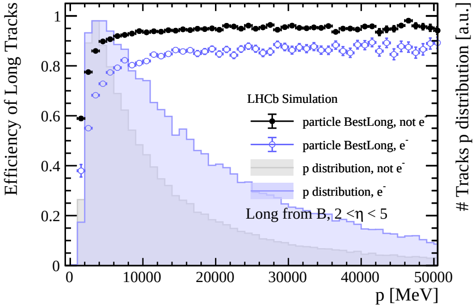


Figure: Long track efficiency of electrons and other particles.

[1] doi: 10.1088/1742-6596/1271/1/012009

The LHCb Experiment

- major experiment at LHC, located at CERN
- tracking system: Vertex Locator (VELO), Upstream Tracker (UT), magnet, Scintillating Fibre (SciFi) tracker

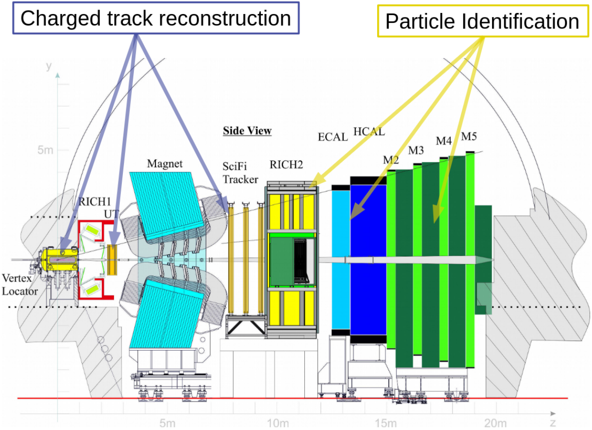


Figure: Side view of the LHCb detector in the non-bending y-z plane [1].

Electrons at LHCb

electrons lose energy while traversing the LHCb detector:

- energy loss dominated by bremsstrahlung

→ typically lose 30% of energy upstream

- bremsstrahlung obstructs reconstruction

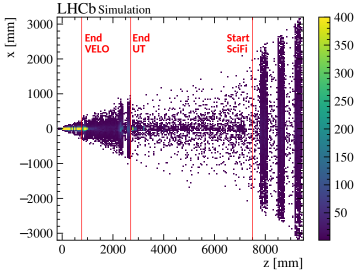


Figure: Bremsstrahlung vertices.

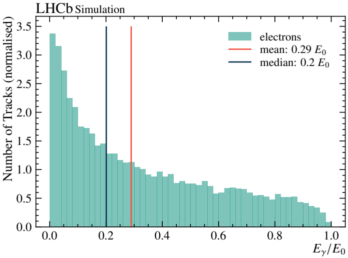


Figure: Energy loss of electrons in the LHCb detector upstream of the magnet.

Tracking of Charged Particles

- **Long tracks** offer the highest relative momentum resolution with 0.5%

→ main type used in physics analyses

→ reconstructed via two methods: Forward tracking and **Track Matching**

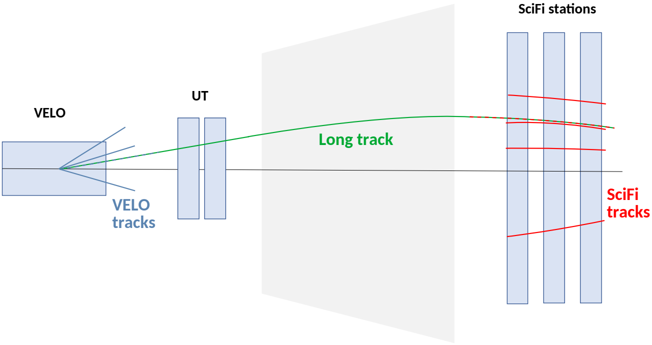


Figure: Sketch of LHCb track types in the x-z plane.

The Matching Algorithm

Basic idea:

- combine VELO and SciFi tracks of same particle to create Long tracks
- quantify the level of agreement, *i.e.* a match, between reconstructed VELO and SciFi tracks
- evaluate a neural network to obtain a combined estimate of being correctly matched

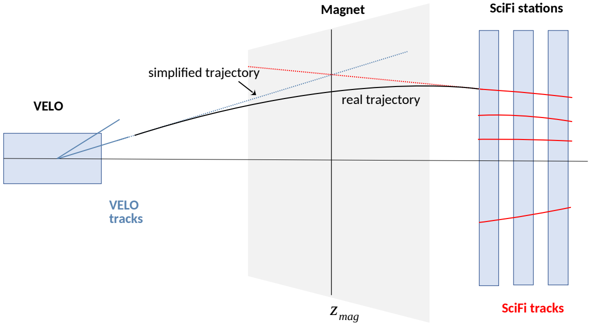


Figure: Basic idea of the Matching algorithm in the x-z plane.

Track Reconstruction in the Real-Time Analysis Trigger

LHCb's software trigger is split into two stages:

HLT1: partial event reconstruction

HLT2: full event reconstruction

→ **Matching**

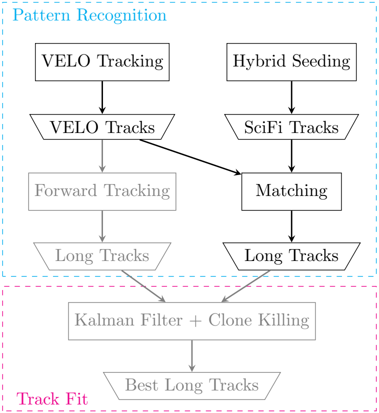


Figure: Dataflow in the Long-track reconstruction in HLT2.

Improving Electron Reconstruction

currently, all track-reconstruction tuning explicitly excludes electrons

solution:

- run dedicated electron reconstruction sequence in HLT2
- expand the Matching algorithm

  → must be computationally cheap to be run in the trigger
- tune parameterisations to electrons
- upper limit for the efficiency around 93%, determined by the VELO and SciFi track efficiency

Track Preselection

electron track identification:

- tracks can be extrapolated to the ECAL and matched to EM-showers:

→ VELO tracks to the EM-showers of photons emitted by electrons in the VELO

→ SciFi tracks to the EM-showers of the electrons directly

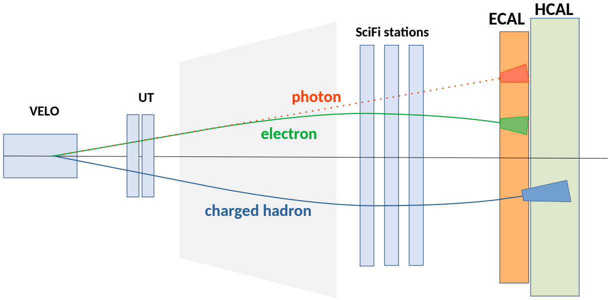


Figure: Different particle signatures in the LHCb detector.

Parameter Tuning

tune all approximations and estimates used during the reconstruction to electrons, *e.g.* the z_{mag} parameter in the optical model

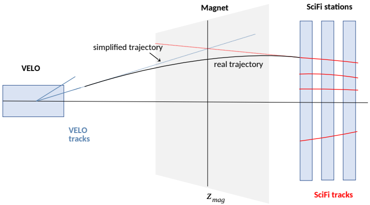


Figure: Magnet kick position z_{mag}.

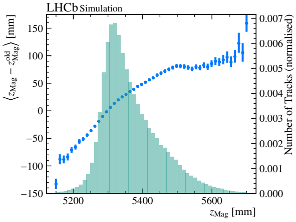


Figure: Old z_{mag} parameterisation for electrons.

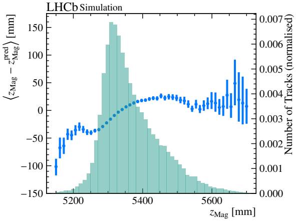


Figure: New z_{mag} parameterisation for electrons.

Input Variables I

- six input variables
- most important:

$D_x = |x_V(z_{\text{mag}}) - x_T(z_{\text{mag}})|$

$D_y = |y_V(z_{\text{mag}}) - y_T(z_{\text{mag}})|$

→ used to define a combined quality measure, mimicking a χ^2:

$$\chi^2_{\text{match}} = \frac{D_x^2}{\sigma_x^2} + \frac{D_y^2}{\sigma_y^2} + \frac{|\Delta t_y|^2}{\sigma_{t_y}^2}$$

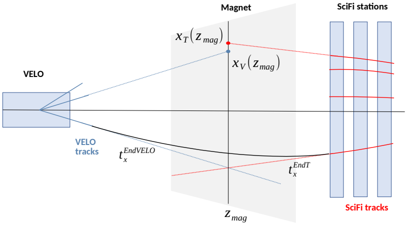


Figure: Bending plane.

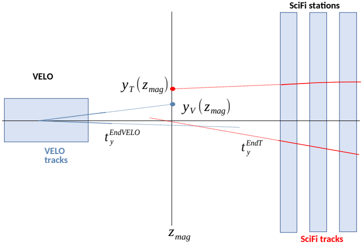


Figure: Non-bending plane.

Input Variables II

- six input variables
- most important:

$$D_x = |x_V(z_{mag}) - x_T(z_{mag})|$$

$$D_y = |y_V(z_{mag}) - y_T(z_{mag})|$$

→ used to define a combined quality measure, mimicking a χ^2:

$$\chi^2_{\text{match}} = \frac{D_x^2}{\sigma_x^2} + \frac{D_y^2}{\sigma_y^2} + \frac{|\Delta t_y|^2}{\sigma_{t_y}^2}$$

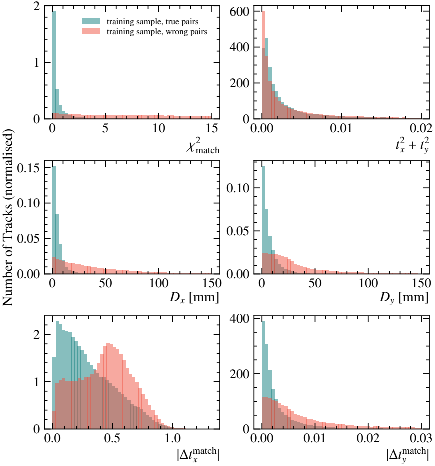


Figure: Matching MLP input variables.

Test on Electron Tracks

- new Matching algorithm tested on electron tracks
- efficiencies above 90%
- low ghost rate

Table: Integrated Matching efficiencies using perfect track selection.

Category	Old $\varepsilon[\%]$	New $\varepsilon[\%]$
Long	64.02	78.33
Long from B	83.46	87.90
$p > 5\,\mathrm{GeV}$	86.22	90.05
$p > 3\,\mathrm{GeV}\ p_T > 0.5\,\mathrm{GeV}$	87.72	90.86
Ghost rate	17.44%	13.96%

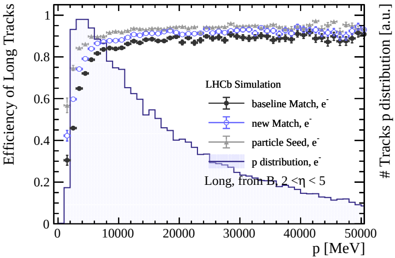


Figure: Efficiency using perfect selection.

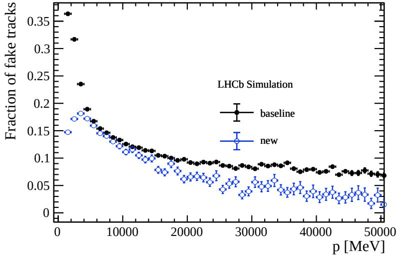


Figure: Ghost rate.

Improved Electron Reconstruction Efficiency

- new Matching algorithm + ECAL filter
- efficiency increases by 3%
- higher ghost rate

→ reduce with ECAL filter

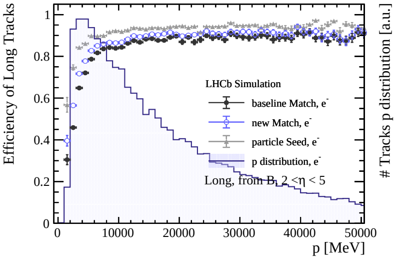


Figure: Matching efficiency using ECAL filter.

Table: Integrated Matching efficiencies using ECAL filtered tracks.

Category	Old $\varepsilon[\%]$	New $\varepsilon[\%]$
Long	64.02	76.41
Long from B	83.46	86.40
$p > 5\,\mathrm{GeV}$	86.22	88.55
$p > 3\,\mathrm{GeV}\ p_T > 0.5\,\mathrm{GeV}$	87.72	89.39
Ghost rate	17.44%	26.37%

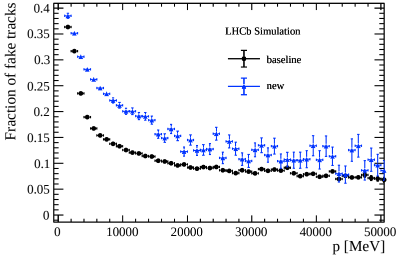


Figure: Ghost rate.

Conclusion

summary:

- significant improvement in electron reconstruction
  - ECAL filter to control ghosts
  - improved Matching for electrons

→ efficiencies around 90%

next steps:

- ECAL filter ready for SciFi tracks, but not for VELO tracks
- use the electron track-finding for real data
- optimise track fit
  - → improve ghost rate

References I

[1] P. A. Günther, "Track reconstruction development and commissioning for LHCb's run 3 real-time analysis trigger."

[2] A. Hoecker et al., "TMVA - toolkit for multivariate data analysis."

[3] V. Gligorov et al., "RTA and DPA dataflow diagrams for run 1, run 2, and the upgraded LHCb detector,"

[4] A. Scarabotto, "Track matching for LHCb upgrade 2."

[5] C. A. Beteta et al., "Calibration and performance of the LHCb calorimeters in run 1 and 2 at the LHC."

BACKUP

Multi-layer Perceptron (MLP)

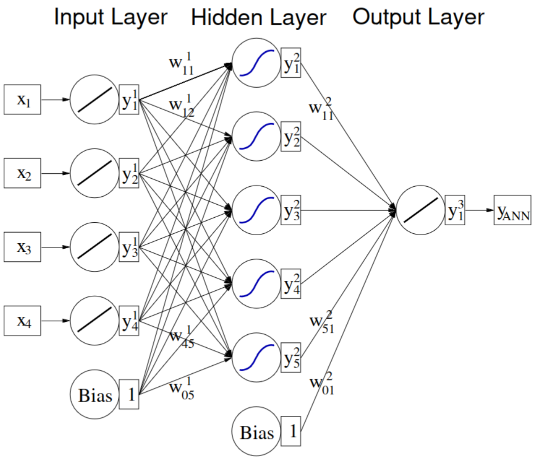


Figure: MLP with one hidden layer [2].

Multi-layer Perceptron (MLP)

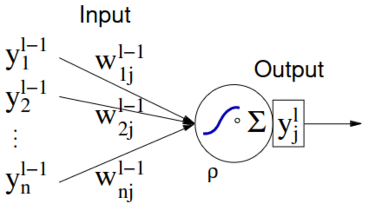


Figure: Single neuron j in layer l with n input connections [2].

Variable	Definition
x_i	inputs
w_i	synaptic weights
b	bias weight
y_i	predicted values
d_i	target values
n	# input connections
η	learning rate

neuron activation function

$$\rho = \mathrm{ReLU} = \max\left(0, b + \sum_{i=1}^{n} y_i w_i\right)$$

error function

$$E = \mathrm{CE}(y_i, d_i) = -\frac{1}{n}\sum_{i=1}^{n}(d_i \log(y_i) + (1-d_i)\log(1-y_i))$$

→ gradient descent of weight ($\eta \equiv 0.02$)

$$w_i' = w_i + \Delta w_i = w_i - \eta \frac{\partial E(w_i)}{\partial \mathrm{w_i}}$$

Magnetic field at LHCb

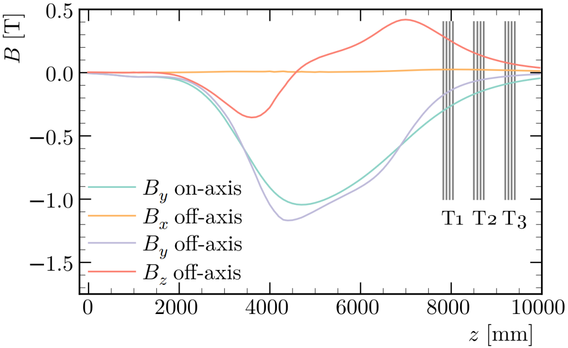
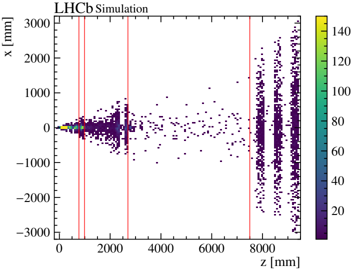
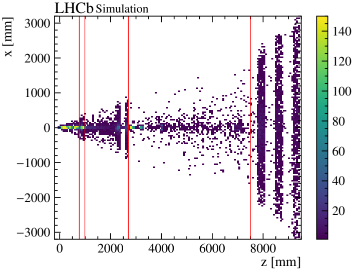


Figure: Magnetic field strength components as functions of z [1].

Bremsstrahlung



a: Found.



b: Lost.

Figure: Distributions of bremsstrahlung vertices.

Performance Metrics

two main performance indicators for track reconstruction with pattern recognition:

1. *track finding efficiency* $\varepsilon = \frac{N_{\text{reconstructible \& MC-matched}}}{N_{\text{reconstructible}}}$
2. *fake track fraction* or *ghost rate* $r_{\text{fake}} = \frac{N_{\text{fake}}}{N_{\text{total}}}$

   where N is the number of tracks in the corresponding category

reconstructible:

- must pass through at least three VELO layers, leaving hits in each
- must leave a hit in at least one x and one stereo layer in each of the SciFi stations

MC-matched:

- more than 70% of the hits must originate from the same simulated particle

RTA Trigger

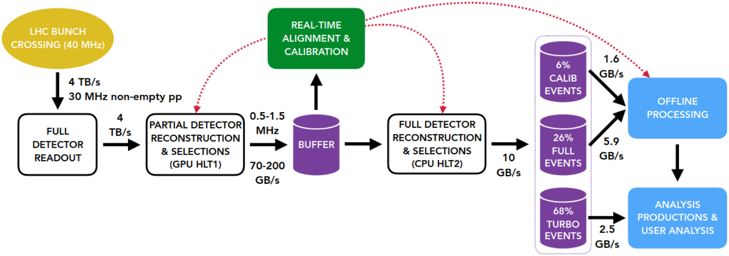
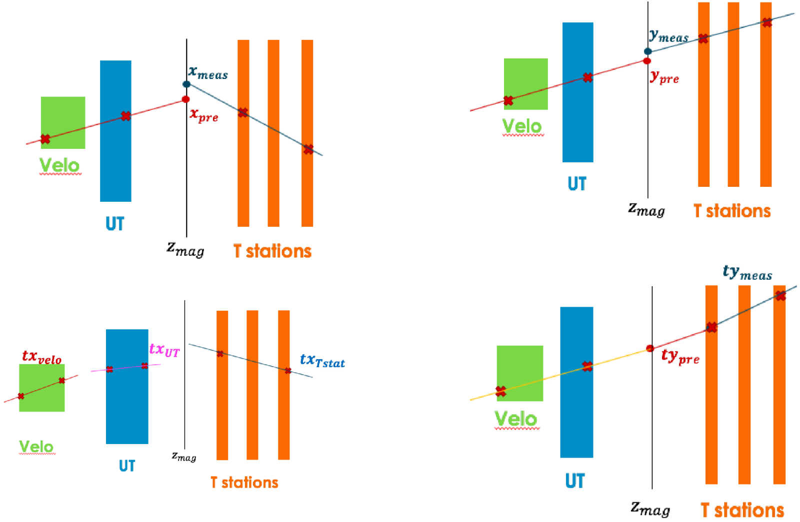


Figure: LHCb software trigger dataflow [3].

Matching Variables

input variables of the Matching algorithm [4]:



Matching Variables

$$t_x^2 + t_y^2 = \left(\frac{\mathrm{d}x}{\mathrm{d}z}\right)^2 + \left(\frac{\mathrm{d}y}{\mathrm{d}z}\right)^2$$

$$|\Delta t_x| = |t_x^{\mathrm{EndT}} - t_x^{\mathrm{EndVelo}}|$$

$$|\Delta t_y| = |t_y^{\mathrm{EndT}} - t_y^{\mathrm{EndVelo}}|$$

$$D_x^{\mathsf{match}} = \left| x_T - x_V + t_x^{\mathsf{EndT}}(z_{\mathsf{mag}}^{\mathsf{match}} - z_T) - t_x(z_{\mathsf{mag}}^{\mathsf{match}} - z_V) \right|$$

$$D_y^{\mathsf{match}} = \left| y_T - y_V + t_y^{\mathsf{EndT}}(z_{\mathsf{match}} - z_T) - t_y(z_{\mathsf{match}} - z_V) - y_{\mathsf{corr}}^{\mathsf{match}} \right|$$

$$\chi^2_{\mathsf{match}} = \frac{D_x^2}{\sigma_x^2} + \frac{D_y^2}{\sigma_y^2} + \frac{|\Delta t_y|^2}{\sigma_{t_y}^2}$$

Input Variables

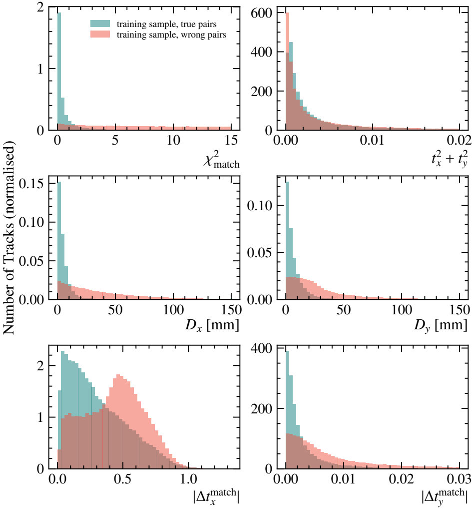


Figure: Matching MLP input variables.

some changes in the preselection cuts yield better performance

Table: Matching MLP input variables for electron Matching.

Variable	Preselection
χ^2_{match}	< 15
D_x	$< 300\ \mathrm{mm}$
D_y	$< 300\ \mathrm{mm}$
$\lvert\Delta t_x^{\mathrm{match}}\rvert$	< 2.0
$\lvert\Delta t_y^{\mathrm{match}}\rvert$	< 0.15
$t_x^2 + t_y^2$	

Electron Matching Tuning

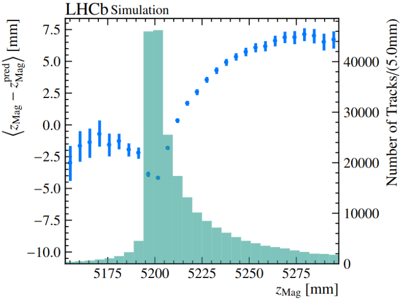


Figure: Old $z_{\rm mag}$ parameterisation for not-electrons [1].

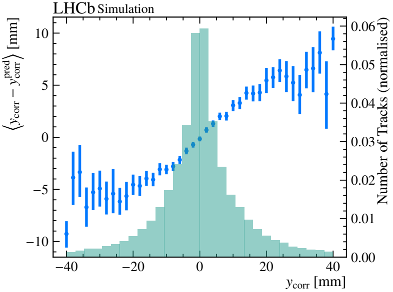


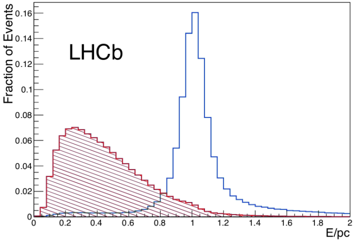
Figure: New $y_{\rm corr}$ parameterisation for electrons.

$$z_{\rm mag}^{\rm match} = c_0 + c_3 t_x^2 + c_7 t_y^2 + |\Delta t_x| \left(c_1 + c_{11} t_y^2 \right) + \Delta t_x^2 \left(c_5 + c_{10} |x_T| \right) + \Delta t_x t_x \left(c_4 + c_9 |x_T| \right) + |x_T| \left(c_2 + c_8 t_x^2 + c_6 |x_T| + c_{12} |x_T|^2 \right)$$

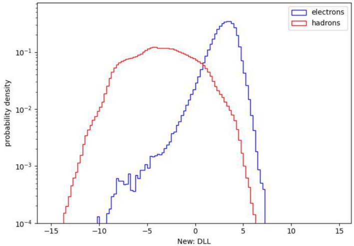
$$y_{\rm corr}^{\rm match} = t_y \left(c_2 |\Delta t_x|^2 + c_3 |\Delta t_y|^2 \right) + \Delta t_y \left(c_0 + c_1 \Delta t_x \right)$$

ECAL Filter

SciFi track selection using the already implemented ECAL filter



a: E/p using Run 1 data.



b: DLL using simulated events.

Figure: Distributions of variables in the ECAL for electrons (blue) and hadrons (red). Figure (a) adopted from Ref. [5], and (b) from (M. V. Veghel, personal communication, January 2024).

Electron Matching Response

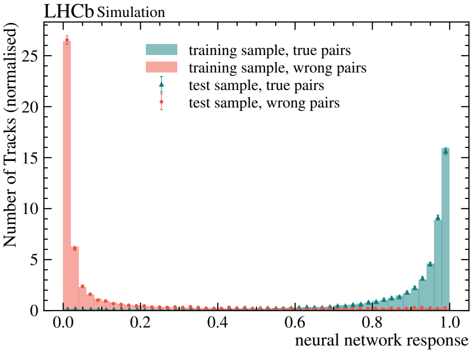


Figure: Output of the electron Matching MLP.

BestLong Efficiency

Table: Integrated Best Long efficiency of electrons and other particles.

BestLong	$e^{\pm}$ $\varepsilon[\%]$	not-$e^{\pm}$ $\varepsilon[\%]$
From B	82.04	91.08
$p > 5\,\mathrm{GeV}$ from B	84.15	94.00
Ghost rate	10.26%	

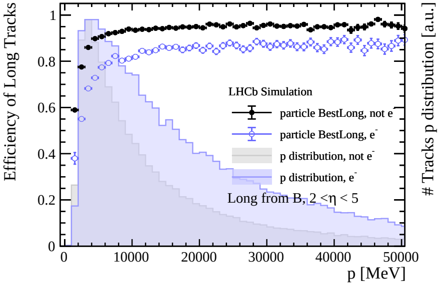
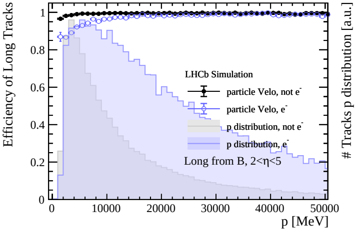
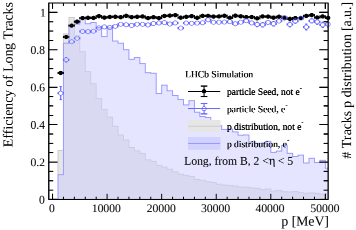


Figure: Long track efficiency of electrons and other particles, after Kalman filter.

VELO and T Track Efficiencies



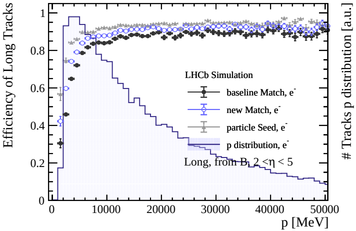
a: VELO Tracking efficiency.



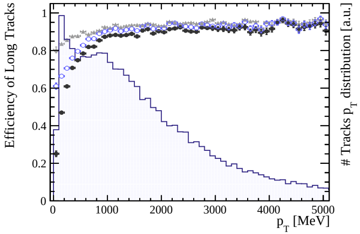
b: Hybrid Seeding efficiency.

Figure: Reconstruction efficiency of (a) the VELO Tracking, and (b) the Hybrid Seeding algorithm.

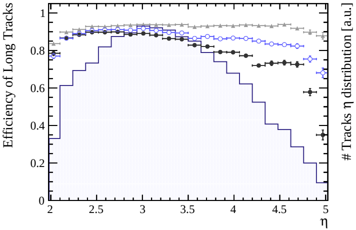
New Matching Efficiencies - Long from B



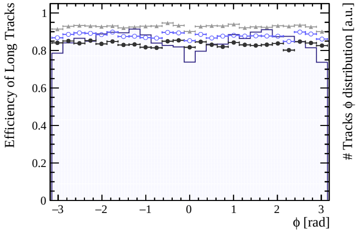
a: Efficiency *vs.* momentum.



b: Efficiency *vs.* transverse momentum.



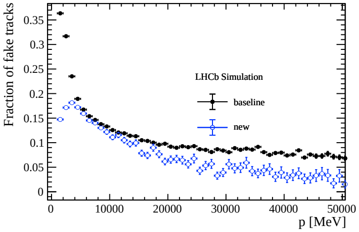
c: Efficiency *vs.* pseudorapidity.



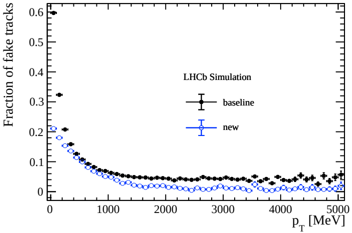
d: Efficiency *vs.* azimuthal angle.

Figure: Track finding efficiency of the electron Matching algorithm.

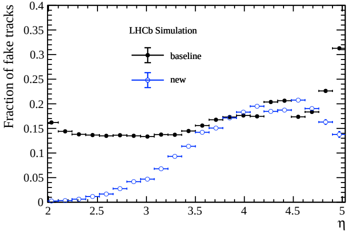
New Matching Fake Track Fraction - Long from B



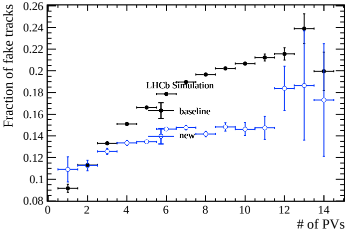
a: Fake track fraction *vs.* momentum.



b: Fake track fraction *vs.* transverse momentum.

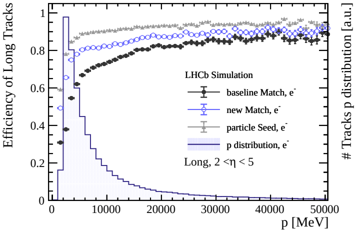


c: Fake track fraction *vs.* pseudorapidity.

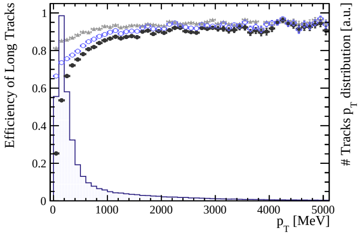


d: Fake track fraction *vs.* number of primary vertices.

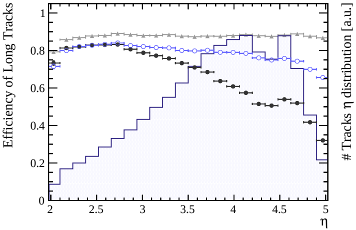
New Matching Efficiencies - Long



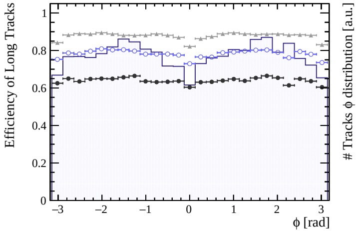
a: Efficiency *vs.* momentum.



b: Efficiency *vs.* transverse momentum.



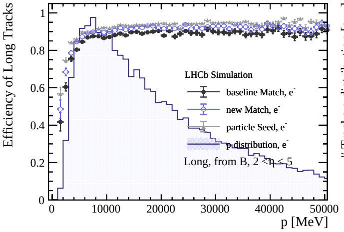
c: Efficiency *vs.* pseudorapidity.



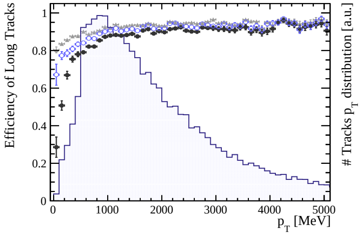
d: Efficiency *vs.* azimuthal angle.

Figure: Track finding efficiency of the electron Matching algorithm.

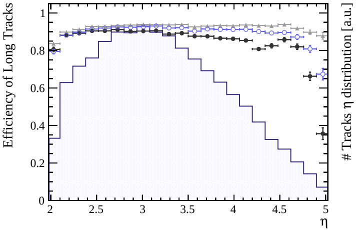
New Matching Efficiencies - Long from B $p > 3\,\mathrm{GeV}$ $p_T > 0.5\,\mathrm{GeV}$



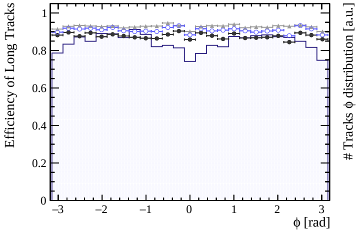
a: Efficiency *vs.* momentum.



b: Efficiency *vs.* transverse momentum.



c: Efficiency *vs.* pseudorapidity.



d: Efficiency *vs.* azimuthal angle.

Figure: Track finding efficiency of the electron Matching algorithm.