A z-Vertex Trigger for Belle II

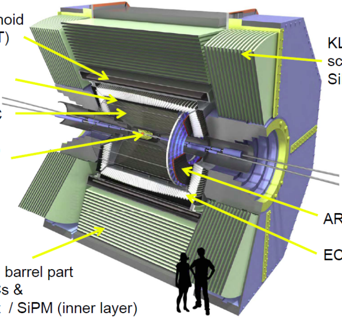
Sebastian Skambraks

Technische Universität München

May 27, 2014

Outline

Belle II Detector



Neuro Team

F. Abudinen (LMU), Y. Chen (TUM), M. Feindt (KIT), R. Frühwirth (HEPHY), M. Heck (KIT), C. Kiesling (MPI), A. Knoll (TUM), S. Neuhaus (TUM), S. Paul (TUM), T. Röder (TUM), J. Schieck (HEPHY), S. Skambraks (TUM)

Introduction

Goals

- build a z-vertex track trigger for the Belle II experiment
- achieve high precision (spatial resolution $\approx 1\,\mathrm{cm}$)
- get a fast decision ($< 1\,\mu\mathrm{s}$)

Methods

- CDC Track Segment data as input [IDs & clock cycle (2 ns timing)]
- current approaches:
  - MLP - Multi Layer Perceptron
  - Cascade prediction

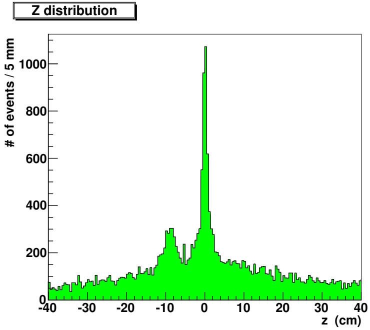


Figure: Offline z distribution in the Belle Experiment[a].

a) T. Abe et al., *Belle II Technical Design Report*, KEK-REPORT-2010-1, arXiv:1011.0352v1 [physics.ins-det] (2010).

Main approach

- sectorize input in (p_T, ϕ, θ)-Sectors
- many MLPs (one for each Sector)
- Hits are TS-IDs & drift times (2 ns resolution from TSF clock)

Look Up Table - Bayes

$$P(Sector|Hits) = P(Hits|Sector) \cdot \frac{P(Sector)}{P(Hits)} \tag{1}$$

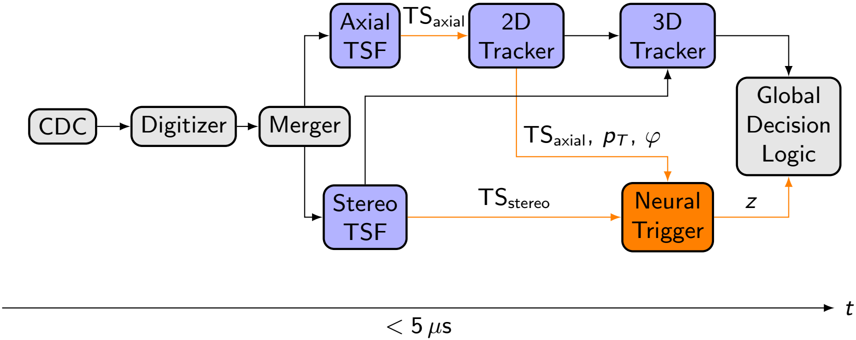
- modeled as nD array
- used to predict Sector & used to generate NN input

MLP

$$z(Hits, p_T, \phi, \theta) = NN_{p_T,\phi,\theta}(Hits) \tag{2}$$

- output float value interpreted as scaled z-position

Signal flow in the CDC Trigger

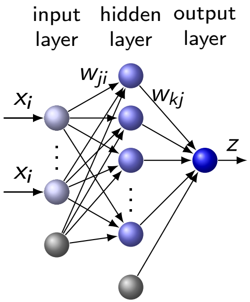


→ The neural network trigger will be implemented on a Virtex 7 FPGA

MLP - Multi Layer Perceptron

Motivation

- supervised machine learning
- universal real valued function approximation[b]
- short deterministic runtime



Basic Setup

- one neuron: $y = \tanh(\sum_{i=1} w_i \cdot x_i + w_0)$
- 3 layers, fully forward connected, hidden layer: $N_{hidden} = 3 \cdot N_{input}$
- 1 output node, output value interpreted as scaled z-vertex position
- train size $\approx Ndof \cdot 10$
- training with rprop algorithm (backpropagation)

b) R. Hecht-Nielsen, Theory of the Backpropagation Neural Network, IEEE first Annual Int. Conf. on Neural Networks, 593 - 605, 1989

Sectorization & TS selection

Sectorize input in p_T, ϕ, θ

all 2336 Track Segments (TS) as input is too much for one MLP

- constraint in (ϕ, θ) to shoot always in the same direction
- constraint in p_T to have the same curvature for the tracks

$\longrightarrow$ subset of relevant TS for each $(\phi,\ \theta,\ p_T)$ region

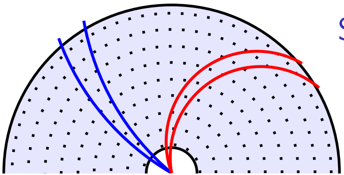


Figure: Two different $p_T - \phi$ sectors.

Selection criteria

- require >p% of the events to be covered by TS selection (chosen on a per superlayer basis)

$\rightarrow$ per sector $\approx$ 20 TS are used as input

“Relevant” Track Segments

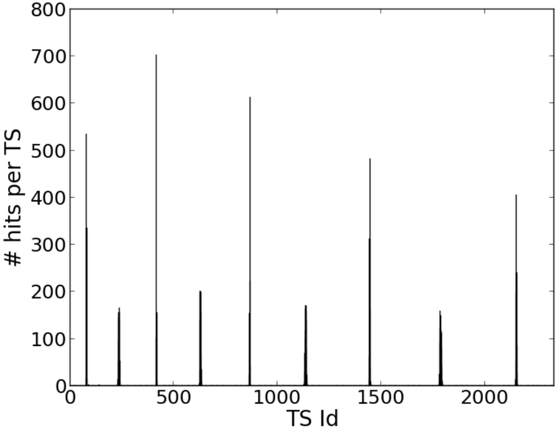


Figure: Hit distribution for events within a $p_T - \phi$ sector.

Track Parameters

$pt \in [1.43, 1.67]$ GeV	$\phi \in [180, 181]^\circ$	$\theta \in [35, 123]^\circ$	$z \in [-50, 50]$ cm

"Expert" MLP - Capabilities

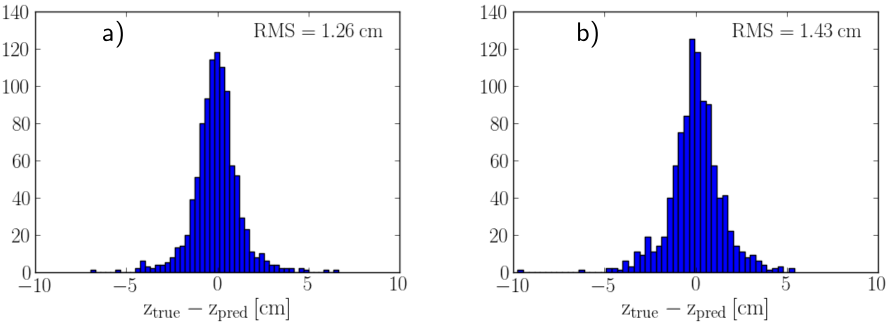


Figure: z-vertex prediction with an "expert" MLP in two p_T regions with $\phi \in [180, 181]^\circ$, $\theta \in [56, 62]^\circ$ and $z \in [-10, 10]$ cm.
a) $p_T \in [1.6, 1.67]$ GeV. b) $p_T \in [4.2, 4.8]$ GeV.

! high accuracy on the z-vertex within a small sector
→ construct one MLP for each sector ("expert")
→ find optimal solution to determine the sectors (preprocessing)

Preprocessing

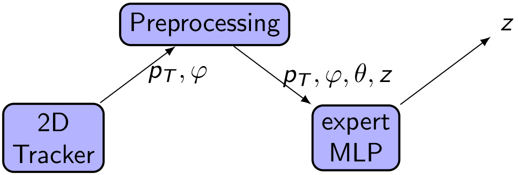


Figure: The $p_T - \phi$ from the 2D Tracker is enriched with $\theta - z$ information by the preprocessing to allow the selection of an "expert" MLP.

- $\mathcal{O}(10^6)$ sectors in total
- preprocessing determines correct sector
- "expert" MLP provides z-vertex value
- MLP is main candidate for preprocessing
- several rounds in the preprocessing might be required (ideal: 1 round)

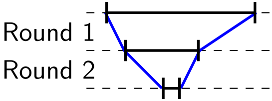


Figure: Incrementally shrink the sector size in the preprocessing.

θ Sector Finding with a MLP

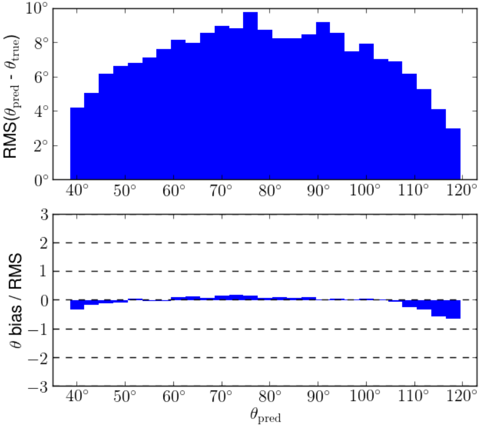


Figure: RMS of θ prediction in θ-bins.

- bins are assigned using the prediction θ_{pred}

$\rightarrow$ RMS is θ dependent: RMS(θ)

$\rightarrow$ MLP suitable for θ prediction

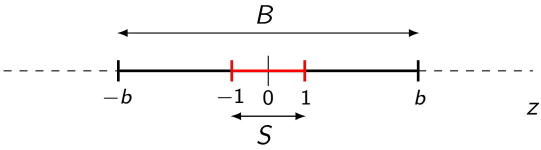
Track Parameters

$pt \in [1.43, 1.67]$ GeV	$\phi \in [180, 181]^\circ$	$\theta \in [35, 123]^\circ$	$z \in [-50, 50]$ cm

z Sector Finding with a MLP - Efficiency Analysis

a high efficiency for Interaction Region (IR) events is required
→ for the experiment, define two z-bins around IR ($z = 0\,\text{cm}$):

small $S = [-1, 1]\,\text{cm}$
big $B = [-b, b]\,\text{cm}$



b is varied in the experiment

Contingency Table

	$z_{\text{true}} \in S$	$z_{\text{true}} \notin S$
$z_{\text{pred}} \in B$	TP	FP
$z_{\text{pred}} \notin B$	FN	TN

Table: The prediction z_{pred} and true value z_{true} are evaluated in two different bins.

Efficiency

- efficiency

$$EFF = \frac{TP}{TP + FN} \tag{3}$$

- false positive rate

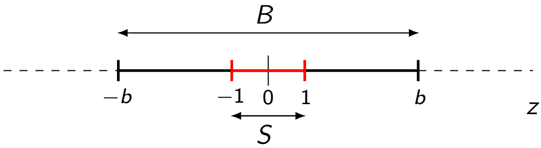
$$FPR = \frac{FP}{FP + TN} \tag{4}$$

z Sector Finding with a MLP - Efficiency Analysis

a high efficiency for Interaction Region (IR) events is required
$\rightarrow$ for the experiment, define two z-bins around IR ($z = 0\,\text{cm}$):

small $S = [-1, 1]\,\text{cm}$

big $B = [-b, b]\,\text{cm}$



b is varied in the experiment

Contingency Table

	$z_{\text{true}} \in S$	$z_{\text{true}} \notin S$
$z_{\text{pred}} \in B$	TP	FP
$z_{\text{pred}} \notin B$	FN	TN

Table: The prediction z_{pred} and true value z_{true} are evaluated in two different bins.

Efficiency

- efficiency

$$EFF = \frac{TP}{TP + FN} \quad (3)$$

- false positive rate

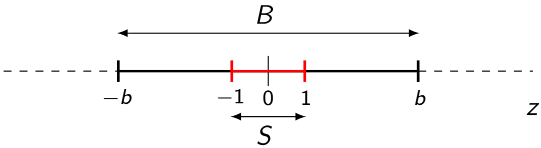
$$FPR = \frac{FP}{FP + TN} \quad (4)$$

z Sector Finding with a MLP - Efficiency Analysis

a high efficiency for Interaction Region (IR) events is required
$\rightarrow$ for the experiment, define two z-bins around IR ($z = 0\,$cm):

small $S = [-1, 1]\,$cm

big $B = [-b, b]\,$cm



b is varied in the experiment

Contingency Table

	$z_{\text{true}} \in S$	$z_{\text{true}} \notin S$
$z_{\text{pred}} \in B$	TP	FP
$z_{\text{pred}} \notin B$	FN	TN

Table: The prediction z_{pred} and true value z_{true} are evaluated in two different bins.

Efficiency

- efficiency

$$EFF = \frac{TP}{TP + FN} \qquad (3)$$

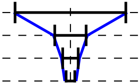
- false positive rate

$$FPR = \frac{FP}{FP + TN} \qquad (4)$$

3 Round z Prediction Chain Experiment

In each round b is set such that $EFF > 99\%$

$\rightarrow$ use more specialized predictor in next round



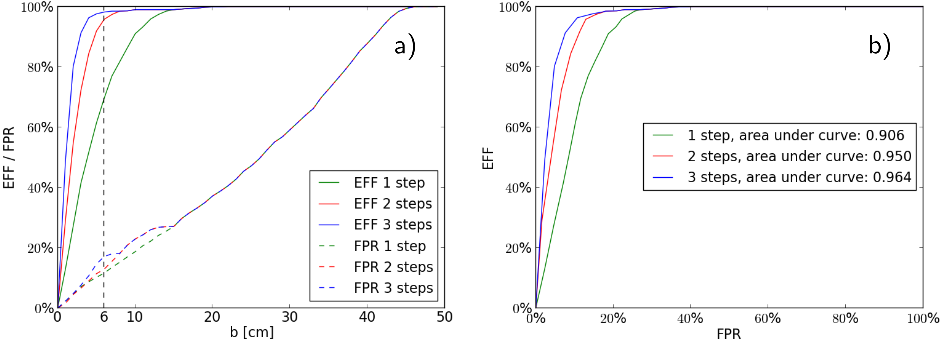


Figure: a) *EFF* and *FPR* of all 3 rounds vs. the half bin size b.
b) *EFF* vs. *FPR* as overall measure of the classification.

Track Parameters

$pt \in [1.43, 1.67]$ GeV	$\phi \in [180, 181]^\circ$	$\theta \in [35, 123]^\circ$	$z \in [-50, 50]$ cm

Conclusion

“expert” MLP

- “experts” are specialized to phase space sectors
- high z-resolution with $RMS < 2\,\text{cm}$
- sectorized solution requires $\mathcal{O}(10^6)$ “expert” MLP

Preprocessing

- preprocessing necessary for the sector finding
- MLP is good candidate for preprocessing
  (can predict θ and z in given $p_T - \phi$ sectors)

Outlook

- need generalization experiment to full detector acceptance region
- for hyper parameters and efficiency in the preprocessing further optimization is required
- Virtex 7 FPGA implementation anticipated

Thank you for your attention!