Background Suppression with the Belle II Neural Network Trigger

Sebastian Skambraks

Max-Planck-Institut für Physik

Aug 21, 2017

Outline

Introduction
Belle II
Trigger
NeuroTrigger
Algorithm
Accuracy
Background
Simulation
Suppression

Neuro Team
S. Bähr, C. Kiesling, S. Neuhaus, S. Skambraks
Introduction - Belle II at SuperKEKB
located in Tsukuba, Japan at KEK

- asymmetric $e^+e^-$ collider
- $\Upsilon(4S)$ resonance
  - $B^0\bar{B}^0 / B^+B^-$
  - $\mathcal{L} = 8 \times 10^{35} \text{ cm}^{-2} \text{ s}^{-1}$
    - $(40 \times \text{KEKB})$
- average $p_T$: 500 MeV
- average track multiplicity: 11
Introduction - The Belle II Detector
Introduction - The Belle II Detector

Central Drift Chamber
56 layers
Input for L1 Track Trigger
Introduction - Belle II Background

Beam Background Tracks

- tracks generated at the beam-line & -wall with vertices \( z \neq 0 \) cm
- increase with luminosity
- main processes:
  - Touschek effect
  - radiative Bhabha back scatters
  - beam gas

⇒ need \( z \) vertex reconstruction at 1\(^\text{st} \) trigger level

NeuroTrigger Goals

- reject tracks from \( z \neq 0 \) cm
- single track \( z \)-vertex resolution < 2 cm
- latency < 1 \( \mu \)s
**Introduction - Interaction Region**

- scattering at material $\rightarrow$ background tracks
- two separate rings with different energies

**Bunch crossing:**

Nano-Beam scheme

\[
\begin{align*}
\sigma_{x,y,z} & \quad \text{beam size} \\
\phi & \quad \text{crossing angle} \\
d = \frac{\sigma_x}{\phi} & \quad \text{eff. bunch length} \\
\sigma_y & \approx 60 \text{ nm} \\
\phi & \approx 40 \text{ mrad} \\
d & \approx 0.2 \text{ mm}
\end{align*}
\]
Introduction - Belle II First Level Trigger

Requirements

- 30 kHz trigger rate
- 5 µs latency
⇒ deadtime-free pipelined operation
Introduction - Belle II First Level Trigger

Requirements

- 30 kHz trigger rate
- 5 µs latency

⇒ deadtime-free pipelined operation

CDC Trigger Pipeline

1. TSF
2. Finder
3. Tracker

Track Segment Finder
Hough Transformation (2D Tracks)
Neural Network (3D Tracks)

30 kHz
5 µs
GDL

Background Suppression with the Belle II Neural Network Trigger (Sebastian Skambraks)
Introduction - CDC Trigger

- 56 layers combined to 9 super layers (SL)
- 2336 track segments (TS) in 9 SL

<table>
<thead>
<tr>
<th>SL</th>
<th>angle (mrad)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>45.4 – 45.8</td>
</tr>
<tr>
<td>4</td>
<td>-55.3 – -64.3</td>
</tr>
<tr>
<td>6</td>
<td>63.1 – 70.0</td>
</tr>
<tr>
<td>8</td>
<td>-68.5 – -74.0</td>
</tr>
</tbody>
</table>

Stereo SL configuration
Introduction - CDC Trigger

- 56 layers combined to 9 super layers (SL)
- 2336 track segments (TS) in 9 SL

Stereo SL configuration:

<table>
<thead>
<tr>
<th>SL</th>
<th>angle (mrad)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>45.4 – 45.8</td>
</tr>
<tr>
<td>4</td>
<td>-55.3 – -64.3</td>
</tr>
<tr>
<td>6</td>
<td>63.1 – 70.0</td>
</tr>
<tr>
<td>8</td>
<td>-68.5 – -74.0</td>
</tr>
</tbody>
</table>

Track Segment

≈ 15 mm

≈ 1.2 m
≈ 2.4 m
≈ 16 cm

NeuroTrigger Input

- position, drift time and left/right information of TS priority wires
- 2D track estimates ($p_T, \varphi$)
Introduction - CDC Trigger

- axial layers
- stereo layers

xt – relation (nonlinear)
Introduction - CDC Trigger

- axial layers
- stereo layers
- $\Upsilon(4S)$ Event

xt – relation (nonlinear)
Introduction - CDC Trigger

- axial layers
- stereo layers
- γ(4S) Event
- background noise

xt – relation (nonlinear)
Introduction - CDC Trigger

- axial layers
- stereo layers
- \( \Upsilon(4S) \) Event
- background noise
- track segments (TS)

\[ xt - relation \] (nonlinear)

Background Suppression with the Belle II Neural Network Trigger (Sebastian Skambraks)
NeuroTrigger - Multi Layer Perceptron

Properties

- robust function approximator
- massively parallel processing
- short deterministic runtime
- neuron: \( y = \tanh(w_ix_i + w_0) \)
- network: \( z_k = f(w_{kj}f(w_{ji}x_i)) \)

Training

- minimize \( \sum_i (z_i^{\text{true}} - z_i^{\text{Net}})^2 \)
- RPROP (backpropagation)

input one TS Hit per SL per track
(position \( \varphi_{\text{rel}} \), \( \alpha \) and time \( t \))

output \( z, \theta \) estimate
NeuroTrigger - Input Representation

\[ \varphi_{rel} : \text{TS position relative to 2D track} \]
\[ \alpha : \frac{2D \text{ arc length to TS}}{r_{2D}} \]

- use track estimates provided by 2D finder
- 3 inputs per SL, values: \((t, \varphi_{rel}, \alpha)\)
- dedicated networks for missing hits
NeuroTrigger - Accuracy

- 5 networks total (for missing stereo hits)
- different bkg noise levels
- IP efficiency: flag IP events with $z \in [-6, 6]$ cm

![Graphs showing $\Delta z$ vs. $p_T$ and IP efficiency vs. $p_T$.]

**$\Delta z$ vs. $p_T$**

- $\Delta z$ [cm]
- Neuro 0% Bkg 1.3
- Neuro 50% Bkg 1.7
- Neuro 100% Bkg 2.0
- Neuro 200% Bkg 3.1
- Neuro 300% Bkg 4.4

**IP efficiency vs. $p_T$**

- $\epsilon$ [%]
- Neuro 0% Bkg 98.2
- Neuro 50% Bkg 94.5
- Neuro 100% Bkg 91.8
- Neuro 200% Bkg 82.8
- Neuro 300% Bkg 73.7
### Background Simulation

#### Luminosity

<table>
<thead>
<tr>
<th>Process</th>
<th>Background</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TwoPhoton</strong></td>
<td>$e^+ e^- \rightarrow e^+ e^- e^+ e^-$</td>
</tr>
<tr>
<td><strong>Bhabha S</strong></td>
<td>$e^+ e^- \rightarrow e^+ e^- \gamma$</td>
</tr>
<tr>
<td><strong>Bhabha M</strong></td>
<td>$e^+ e^- \rightarrow e^+ e^- \gamma$</td>
</tr>
<tr>
<td><strong>Bhabha L</strong></td>
<td>$e^+ e^- \rightarrow e^+ e^- \gamma$</td>
</tr>
<tr>
<td><strong>Touschek</strong></td>
<td>intra bunch scatt.</td>
</tr>
<tr>
<td><strong>Coulomb</strong></td>
<td>$e^\pm N \rightarrow e^\pm N$</td>
</tr>
<tr>
<td><strong>Brems</strong></td>
<td>$e^\pm N \rightarrow e^\pm N \gamma$</td>
</tr>
</tbody>
</table>

#### Machine

- **Two Photon:**
  - $e^+ e^- \rightarrow e^+ e^- e^+ e^-$
  - $e^+ e^- \gamma \gamma$

- **Bhabha S:**
  - $e^+ e^- \rightarrow e^+ e^- \gamma$

- **Bhabha M:**
  - $e^+ e^- \rightarrow e^+ e^- \gamma$

- **Bhabha L:**
  - $e^+ e^- \rightarrow e^+ e^- \gamma$

- **Touschek:**
  - intra bunch scatt.

- **Coulomb:**
  - $e^\pm N \rightarrow e^\pm N$

- **Brems:**
  - $e^\pm N \rightarrow e^\pm N \gamma$
Luminosity Background - Radiative Bhabha

largest contribution
- initial state radiation (ISR) (t-channel)

\[ e^+ e^- \rightarrow e^+ e^- \gamma \]

\[ \theta_e^+ [^\circ] \]

\[ 0.5 \ 1 \ 10 \ 0.5 \ 1 \ 10 \ \theta_e^- [^\circ] \]

3 Bhabha cases:
- small angle (BhabhaS)
- medium angle (BhabhaM)
- large angle (BhabhaL)

Bhabha cross section strongly depends on scattering angle (t-channel)
Machine Background

**Touschek**
elastic inner beam
- \( e^+ e^- \rightarrow e^+ e^- \)

**Coulomb**
elastic beam gas
- \( e^+ \rightarrow e^+ e^+ \)

**Brems**
inelastic beam gas
- \( e^+ \rightarrow e^+ \gamma \)

Simulation

1. calculate deviations of beam particles from the nominal beam orbit
2. track these particles until they get lost
3. loss positions are input to detector simulation
Background - Material Scattering

- initial particles from
  - beam pipe (machine background)
  - IP (luminosity background)
- scattered particles show detector structure and beam pipe

- backscatters: reducible with track trigger
- tracks from the IP: irreducible QED
  - prescaled with Bhabha veto (combine track & calorimeter information)

Initial Particles

Track Vertices

Background Suppression with the Belle II Neural Network Trigger (Sebastian Skambraks)
Background - Material Scattering

Initial Bkg Particles before Scattering

- primary particles from the bkg generators
- only events with a 2D trigger
- luminosity bkg only from the IP
Background - Material Scattering

Tracks seen in the Trigger

- particles after detector simulation
- bkg particles matched to 2D trigger tracks
  \[ \approx 80 \text{ kHz reducible} \quad (z \neq 0) \]
  \[ \approx 40 \text{ kHz irreducible} \quad (z = 0) \]
Background - Reconstruction

Neural Network Track Estimates

- 3D reconstructed bkg with the neural network
- neuro z range limited to $[-50, 50]$ cm
Background - Suppression

Z Cut

- cumulative bkg rate after a cut on the neural network $z$
- $z_{\text{cut}}$ is varied in 5 cm steps
Background Suppression

Z Cut

- cumulative bkg rate after a cut on the neural network $z$
- $z_{cut}$ is varied in 5 cm steps
Background - Suppression

Z Cut (Tracks not from IP)

- only tracks with $|z_{MC}| \geq 1$ cm
- cumulative bkg rate after a cut on the neural network $z$
- $z_{cut}$ is varied in 5 cm steps

Rate: 81.4 kHz
Conclusion

Background

- 2 background types: luminosity background (generated at the IP) and machine background (generated at the walls of the beam pipe)
- scattering of background tracks at material leads to spread in $z$
- $\approx 82$ kHz reducible background (tracks not from the IP) and $\approx 40$ kHz irreducible background (tracks from the IP)

Neural Network Trigger

- robust $z$-vertex estimation with the neural networks
- significant background reduction with $z$ cut
- allows to consider a single track trigger
Backup
IP Efficiency

- **efficiency $\varepsilon$:** efficiency to correctly flag tracks from the IP
- **fake rate FR:** rate of tracks wrongly flagged as IP tracks
- **split background data in**
  - “ip-tracks”: $z \in [-1, 1]$ cm
  - “displaced”: $z \notin [-1, 1]$ cm
- **vary $z_{\text{cut}}$ in 1 cm steps ($z_{\text{cut}} \in [1..50]$)**

<table>
<thead>
<tr>
<th>$z_{\text{cut}}$/cm</th>
<th>FR /kHz</th>
<th>$\varepsilon$ / %</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>5.6</td>
<td>89.5</td>
</tr>
<tr>
<td>16</td>
<td>10.9</td>
<td>98.5</td>
</tr>
<tr>
<td>22</td>
<td>15.6</td>
<td>99.3</td>
</tr>
<tr>
<td>28</td>
<td>20.3</td>
<td>99.9</td>
</tr>
</tbody>
</table>
Background - Suppression

**Multiplicity**

- Touschek
- Coulomb
- Brems
- Bhabha M
- Bhabha L
- Bhabha S
- Two Photon

**Single Track**

**Multi Track**

<table>
<thead>
<tr>
<th>rate [kHz]</th>
<th>1 track</th>
<th>≥ 2 tracks</th>
</tr>
</thead>
<tbody>
<tr>
<td>lumi. bkg</td>
<td>89.7</td>
<td>14.5</td>
</tr>
<tr>
<td>machine bkg</td>
<td>16.1</td>
<td>1.2</td>
</tr>
<tr>
<td>total</td>
<td>105.8</td>
<td>15.7</td>
</tr>
</tbody>
</table>
Introduction - The Belle II Detector
Introduction - The Belle II Detector

Central Drift Chamber
56 layers
Input for L1 Track Trigger
Introduction - The Belle II Detector

Calorimeter

$e^-$

$e^+$

$K_L$ and $\mu$ detector

Particle Identification
TwoPhoton

$t$ channel

$s$ channel
Luminosity Background

- primary vertex at the IP ($z = 0$)
- $e^+ e^-$ from the IP directly hit the CDC
- back scattered particles hit the CDC
Machine Background

- Touschek increase via nano beam scheme
- small beam pipe \((r \approx 1\,\text{cm})\), resulting in worse vacuum conditions
- beam gas scattering increased via bad vacuum in the beam pipe

![Graphs showing Touschek, Coulomb, and Brems distributions.](image-url)

**\(p_T\)**

![Graph showing \(p_T\) distribution with a rate of 17.3 kHz.](image-url)

**\(z\)**

![Graph showing \(z\) distribution with a rate of 17.3 kHz.](image-url)
Background - Track Properties

- \( \approx 106 \text{ kHz} \) single track background
- \( \approx 16 \text{ kHz} \) multi track background
- most scattered particles: protons (from nuclear spallation)

**Particles**

<table>
<thead>
<tr>
<th>( \pi^- )</th>
<th>( e^+ )</th>
<th>( e^- )</th>
<th>( \pi^+ )</th>
<th>( p )</th>
<th>other</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>40</td>
<td>20</td>
<td>20</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

**Multiplicity**

<table>
<thead>
<tr>
<th># tracks</th>
<th>( \text{kHz} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>90</td>
</tr>
<tr>
<td>( \geq 3 )</td>
<td>80</td>
</tr>
</tbody>
</table>

**Background Suppression with the Belle II Neural Network Trigger (Sebastian Skambraks)**
MC particles after the detector simulation matched to 2D trigger tracks

final state $e^-$

Background Suppression with the Belle II Neural Network Trigger (Sebastian Skambraks)
MC particles after the detector simulation matched to 2D trigger tracks

final state $e^+$

Background Suppression with the Belle II Neural Network Trigger (Sebastian Skambraks)
MC particles after the detector simulation matched to 2D trigger tracks

Final state $\pi^-$

- $p_T$ / GeV
- $\phi$ / rad
- $cos(\theta)$
- $z$ / cm

Frequency: 4.8 kHz
MC particles after the detector simulation matched to 2D trigger tracks

**final state $\pi^+$**

- **$p_T / \text{GeV}$**
  - rate: 3.7 kHz
  - $p_T / \text{GeV}$: 0.0, 0.2, 0.4, 0.6

- **$\phi / \text{rad}$**
  - rate: 3.7 kHz
  - $\phi / \text{rad}$: $-2, 0, 2$

- **$\cos(\theta)$**
  - rate: 3.7 kHz
  - $\cos(\theta)$: $-1.0, -0.5, 0.0, 0.5, 1.0$

- **$z / \text{cm}$**
  - rate: 3.7 kHz
  - $z / \text{cm}$: $-200, -100, 0, 100, 200$
MC particles after the detector simulation matched to 2D trigger tracks

**final state p**

![Graphs showing distributions of pT, phi, cos(theta), and z with rate indication of 72.2 kHz for different events.]
MC particles after the detector simulation matched to 2D trigger tracks

final state other
Generator particles scattering to final states

\( e^+ \)

\( e^- \)

\( \pi^- \)

\( \pi^+ \)

\( p \)