Increasing track reconstruction efficiency in dense environments at ATLAS

On behalf of the ATLAS collaboration
CTD Valencia
02-05 April 2019
Outline

• Introduction
• Tracking in ATLAS ➔ Focus on dense environments
• Tracking for high $p_T$ B-hadrons:
  • Defining a region of interest
  • Tackling the efficiency loss
  • Improving track quality
• Tracking for high $p_T$ $\pi$s:
  • Fixing merged tracks
• Conclusion
The long lifetime of certain particles (B hadrons, \(\tau\)s, etc...) leads them to having characteristic properties that allow us to identify them:

- Large positive impact parameter
- Displaced secondary vertex
- Weak decay chain (PV \(\rightarrow\) b \(\rightarrow\) c)

The identification of jets originating from B-hadrons (B-tagging) is crucial for many interesting physics signature at the Large Hadron Collider (LHC):

- Top quarks decay into W bosons and b-quarks about 100% of the time
- The Standard Model Higgs boson predominantly decays into b-anti b-quark pairs
- Many searches for new physics, e.g. supersymmetry, involve final states with b-quarks

TRACKING IS ESSENTIAL
ATLAS Inner Detector

• The Inner Detector (ID) measures the trajectories of charged particles originating from the collision point

• Three detector technologies:
  • Silicon pixels (in order):
    ➔ IBL
    ➔ B Layer (BL)
    ➔ Pix 2
    ➔ Pix 3
  • Silicon strips (SCT)
  • Transition radiation (TRT)
Tracking in ATLAS

**PRE-PROCESSING**

**COMBINATORIAL TRACK FINDING**

**EXTENSION TO TRT**

Input tracks

Fit tracks fulfilling minimum requirements

Score track

Create stripped down version of track candidate

Rank tracks

Evaluate minimum requirements:
- n shared cluster
- n holes
- n clusters
- problematic pixel cluster

Score track

if bad score

if min req. not fulfilled

Rejected tracks

Accept track

if min req. fulfilled

Output tracks

Rejected tracks

if min req. fulfilled

if bad score

AMBIGUITY PROCESSOR

Input tracks

Fit tracks fulfilling minimum requirements

Score track

Create stripped down version of track candidate

Rank tracks

Evaluate minimum requirements:
- n shared cluster
- n holes
- n clusters
- problematic pixel cluster

Score track

if bad score

if min req. not fulfilled

Rejected tracks

Accept track

if min req. fulfilled

Output tracks
Tracking in Dense Environments: why are tracks rejected?

- In boosted environment
  - separation between particles ~ cluster size
  - reconstructed as merged cluster
  - shared clusters penalised (reduce fakes/duplicates)
- ATLAS has a neural network (NN) to split shared cluster → tackling remaining inefficiencies
Anatomy of a B jet

At high $p_T$ Bs can be so **displaced** that they start decaying after the first layers of the ID

🔍

Our current tracking algorithm isn’t optimised for this scenario

🔍

We **lose efficiency** in track reconstruction ➔

lose efficiency in B tagging

🔍

B-hadrons have ~4 tracks: whatever gain on per track efficiency results in ~4 fold gain in B tagging
Consequences for B-tagging at high $p_T$?

Tracking is a key part in this loss (2 main effects):

- **Loss of hits** $\rightarrow$ lower track reconstruction efficiency
- **Track quality** $\rightarrow$ degrades impact parameter measurement
Efficiency loss

This loss for tracks coming from high pT Bs is the results of tight cuts

We loosen the following cuts

- nominal maximum number of shared hits: $N_{\text{hits}}^{\text{shared}} < 3 \rightarrow N_{\text{hits}}^{\text{shared}} < 7$
- nominal minimum number of non shared hits: $N_{\text{hits}}^{\text{non-shared}} > 5 \rightarrow N_{\text{hits}}^{\text{non-shared}} > 1$
- nominal minimum number of Si hits to allow splitting: $N_{\text{hits}}^{\text{Si}} > 8 \rightarrow N_{\text{Si}}^{\text{hits}} > 6$
Regaining efficiency

Recover efficiency
• ~15% for B-hadrons decaying after Pix2 and ~10% after BL
• 4 tracks per B → increase in B-tagging should be larger
Defining a Region Of Interest (ROI)

We want to improve our tracking only for tracks coming from B otherwise \(\rightarrow\) **time consuming** and **fake tracks**

**IDEALLY**
Want to apply changes in the tracking only to B tracks

**IN REALITY**
Can limit the changes to high energy calorimeter clusters corresponding to high pT jets

**THREE PARAMETERS**

- \(P_t^{\text{cut}}\) on track candidates
- \(E_{\text{cut}}\)
- \(dR = \sqrt{\Delta \phi^2 + \Delta \eta^2}\)

between cluster and track < 0.1
Defining a Region Of Interest (ROI)

We want to target as many tracks from high $p_T$ B as we can.
Wrong hits assignment

Picking up wrong inner layer hits means pulling the track which results in poor fit quality

We try to remove recursively the hits (starting from IBL up to Pix2) and we refit the track each time

if the quality improves (20% improvement) we keep the stripped down version of the track
The wrong hits assignment problem

Tracks with wrong hits
• Bad fit quality
• Large pull distribution ➜ lose information on impact parameter

Good tracks = no wrong hits
Results for tracks from B hadrons

Showing the % of tracks from B with a wrong IBL and/or BL hit

~ 50% improvement for $p_T > 400$ GeV

~ 20/30% improvement for $p_T > 400$ GeV
Results for tracks from B hadrons

Overall effect if we consider all pixel hits we have an improvement of ~10%
Only apply the refit procedure up to Pix2

ATLAS Simulation Preliminary
$s=13$ TeV
Impact on number of tracks per jet

Increase in number of tracks in jet < 2%, convoluted effect of fake rate and increase in efficiency
Outline

• Introduction
• Tracking in ATLAS ➔ Focus on dense environments
• Tracking for high $p_T$ B-hadrons:
  • Defining a region of interest
  • Tackling the efficiency loss
  • Improving track quality
• Tracking for high $p_T$ $\pi$s:
  • Fixing merged tracks
• Conclusion
High $p_T \tau$s

- Highly collinear particles $\rightarrow$ many shared hits
- Two tracks reconstructed as one $\rightarrow$ merged tracks
- Problematic for high $p_T \tau$s: analysis level has specific selections for 3-prong $\tau$s $\rightarrow$ missing tracks = $\tau$s not identified
- Hard to tackle with simple cut-based approach
- Special MC sample: no pileup, particle gen. with $\eta = 0$, single $\tau$ to 3 $\pi$
### Building the BDT

<table>
<thead>
<tr>
<th>Variable</th>
<th>Num. of Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Track $p_T$</td>
<td>1</td>
</tr>
<tr>
<td>Track $\eta$</td>
<td>1</td>
</tr>
<tr>
<td>Track $\phi$</td>
<td>1</td>
</tr>
<tr>
<td>Num. clusters on each pixel layer</td>
<td>4</td>
</tr>
<tr>
<td>Highest charge deposited in a cluster on each pixel layer</td>
<td>4</td>
</tr>
<tr>
<td>$dE/dx$ in each pixel layer</td>
<td>4</td>
</tr>
<tr>
<td>$</td>
<td>W_c - W_o</td>
</tr>
<tr>
<td>Boolean for whether a hit on each pixel layer is flagged as split</td>
<td>4</td>
</tr>
<tr>
<td>$</td>
<td>L_c - L_o</td>
</tr>
<tr>
<td>Num. clusters on each SCT layer</td>
<td>4</td>
</tr>
<tr>
<td>$</td>
<td>W_c - W_o</td>
</tr>
<tr>
<td>Num. shared clusters on each SCT layer</td>
<td>4</td>
</tr>
</tbody>
</table>

### 43 track variables fed into the BDT

#### Fixing the merged tracks

**ATLAS Simulation Preliminary**

Run 2, $\tau \rightarrow 3\nu \nu$, $|\eta|=0$

- $p_{T,\tau} = 800$ GeV
- Track's first cluster on SCT layer 1

**Wc-Wo SCT1**

- Merged Tracks (3.9% of sample)
- Single Particle Tracks

**ATLAS Simulation Preliminary**

Run 2, $\tau \rightarrow 3\nu \nu$, $|\eta|=0$

- $p_{T,\tau} = 800$ GeV
- Track's first cluster on pixel Layer 2

**Wc-Wo Pix2**

- Merged Tracks (3.9% of sample)
- Single Particle Tracks

**dE/dx**

- Fraction of hits / 0.5 MeV g cm$^{-2}$
Training the BDT

Simulation Preliminary
Run 2, $\tau \rightarrow 3\pi^+\nu_\tau, |\eta|=0$

- pink = ideal
- green = ATLAS reco
- blue = if BDT flags track as merged count as 2 tracks

Recover $\tau$s $\rightarrow$ larger with increasing $p_T$:

- $1 - \text{Fake Rate}$

ROC curve for BDT

Probability of correctly identifying merged track

Percent of $\tau$s with three found $\pi$s with three found $\tau$

- 3 truth-$\pi$s with $\geq 7$ hits
- Perfect merged track ID
- No pixel or SCT penalty
- Track merging BDT (cut at 0.1)
- Width-based SCT sharing
- Default Track Reco
Pion charge BDT

- Using same input but different training
  - on tracks identified as merged at truth-level and flagged by original track-merging BDT
- Only runs on merged tracks
- With charge BDT cut of 0, charge-tagging accuracy is over 85%

Can distinguish same vs opp. sign \( \pi \)

ROC curve for charge BDT

\( \text{ATLAS} \) Simulation Preliminary

Run 2, \( \tau \rightarrow 3\pi^\pm \nu_\tau, |\eta|=0 \)
Conclusions

• Increasing efficiency for tracks coming from B:
  • By loosening the cuts we manage to improve the efficiency by ~5/10% per track coming from high pT Bs → should result in even larger efficiency for B hadrons
  • The refit procedures allows us to remove some of the wrongly assigned tracks and thus improve the quality of tracks coming from B
  • The ROI selects tracks coming from B → high efficiency and purity
  • Further improvement possible with better optimisation

• Identifying merged tracks from $\tau$:
  • Promising results by using BDT to identify merged tracks → recover efficiency for 3 prong $\tau$s
  • Low impact on number of duplicate tracks
  • Can also be used to distinguish same vs opp. sign pions
BACKUP
Light-flavour jet rejection for $\varepsilon_b = 77\%$

- MV2 - 2016 configuration
- MV2 - 2017 configuration
- MV2Mu - 2017 configuration
- MV2MuRnn - 2017 configuration

$\sqrt{s} = 13\text{ TeV}, Z'$

Jet $p_T$ [GeV]

2017/2016

ATLAS Simulation Preliminary

ATLAS Experiment

Université de Genève
ATLAS Simulation Preliminary

Run 2, $\tau \rightarrow 3\pi^+\nu_\tau$, $|\eta|=0$

$p_{T,\tau} =$ 50, 400, 800, 1000 GeV

Percent of $\tau$'s with three found $\pi$'s [%]

- 3 truth-$\pi$'s with $\geq$ 7 hits
- BDT cut -0.05
- BDT cut 0.1
- Width-based SCT sharing
- Default Track Reco

Average $\Delta R$ between pions in event
Results for tracks from B hadrons

Increased number of tracks with wrong SCT hits $\rightarrow$ loosening the cuts allows for more errors but we have increased efficiency.
Little impact on fake rate