LHCb Topological Trigger Reoptimization

Philip Iten\textsuperscript{1}, Tatiana Likhomanenko\textsuperscript{2,3}, Egor Khairullin\textsuperscript{2}, Andrey Ustyuzhanin\textsuperscript{2,3}, Mike Williams\textsuperscript{1}

\textsuperscript{1} Massachusetts Institute of Technology, US
\textsuperscript{2} Yandex School of Data Analysis, RU
\textsuperscript{3} NRC ”Kurchatov Institute”, RU

April 14, 2015
What is Topological Trigger?

- Generic trigger for decays of beauty and charm hadrons
- It is designed to be an inclusive trigger line to efficiently select any B decay with at least 2 charged daughters
- Look for 2, 3, 4 track combinations in a wide mass range
- Designed to efficiently select decays with missing particles
- Use fast-track fit to improve signal efficiency and minbias rejection

**Goal:** improve topological trigger efficiency for Run-2
What tells us an event contains interesting physics?

- A combination of displacement from PV and high PT

Run-2 strategy:

- HLT-1 track is looking for either one super high PT or high displacement track
- HLT-1 2-body SVR classifier is looking for two tracks making a vertex
- HLT-2 improved topo classifier uses full reconstructed event to look for 2, 3, 4 and more tracks making a vertex

NOTE: tracking thresholds are quite different in Run-1 and Run-2
Two, three or four tracks are combined to form a SVR.

Each secondary vertex in Monte Carlo data is preselected in such a way, that all tracks must be matched to particles from the signal decay (true match preselection).
The trigger is designed to allow for the omission of one or more daughters when forming the trigger candidate.
Omission of Daughters

\[ m_{\text{corrected}} = \sqrt{m^2 + |p'_{T\text{missing}}|^2 + |p'_{T\text{missing}}|} \]
**Machine Learning Specific Problem: data structure**

- Signal samples are simulated 13-TeV B decays of various topologies
- Background sample is generic Pythia 13-TeV proton-proton collisions
- Most events have many secondary vertices SVRs (not all events have an SVR)

Event is represented as set of SVR's

Truth match to signal

other preselections

BDT

If at least one SVR in event passed all stages, whole event passes trigger
Machine Learning Specific Problem: FOM

- FOM is the overall efficiency, calculated for passed events, not SVRs.
- Output rate must be limited.
- Restriction is imposed on background events efficiency $FPR = 0.2\%$ (corresponds to 2 kHz).
HLT2: threshold instability

\[
\begin{align*}
B^0 & \rightarrow K^+ [K^- \pi^+] \mu^+ \\
B^+ & \rightarrow D^- [K^+ \pi^-] \\
B^0 & \rightarrow D^- [K^+ \pi^-] \mu^+ \nu^+ \\
B^+ & \rightarrow K^- K^+ \\
B^0 & \rightarrow D^+_s [K^- K^+ \pi^-] \\
B^+_s & \rightarrow (S)[K^- K^+ \pi^-] \\
\end{align*}
\]
HLT2: efficiency vs output rate

![Graph showing the efficiency versus output rate for various B decays.](image_url)
**Bonsai BDT (BBDT):**
- Used in Run-1 for online processing
- Features hashing before training by yourself
- Convert decision trees to n-dimensional table making it essentially infinitely fast
- Predict operation takes one reading from this table

**But:**
- We are limited in the table size (or count of bins for each feature)
- Discretization reduces efficiency

**MatrixNet (MN) post-prunning:**
- Another strategy for online processing
- Features also hashing with amount count of bins for each variable
- Post-pruning of the decision trees to speedup prediction operation (less count of trees)
- Online predict event by all trees
BBDT vs Post-pruning Efficiencies

- $B^0 \rightarrow K^- \pi^+ \mu^+ \mu^-$
- $B^0 \rightarrow D^+ [K^+ \pi^-] D^- [K^+ \pi^-]$
BBDT vs Post-prunning Efficiencies: ROCs

ROC for events

- base MN
- BBDT MN-5
- BBDT MN
- Prunned MN
- rate: 4 kHz
- rate: 2.5 kHz

TPR vs FPR
Ratio of Run-2 over Run-1 for HLT2/HLT1 efficiencies

<table>
<thead>
<tr>
<th>mode</th>
<th>2.5 kHz</th>
<th>4.0 kHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B^0 \rightarrow K^* [K^+ \pi^-] \mu^+ \mu^-$</td>
<td>1.64</td>
<td>1.72</td>
</tr>
<tr>
<td>$B^+ \rightarrow \pi^+ K^- K^+$</td>
<td>1.59</td>
<td>1.65</td>
</tr>
<tr>
<td>$B^0_s \rightarrow D^- [K^+ K^- \pi^-] \mu^+ \nu_\mu$</td>
<td>1.14</td>
<td>1.47</td>
</tr>
<tr>
<td>$B^0_s \rightarrow \psi(1S) [\mu^+ \mu^-] K^+ K^- \pi^+ \pi^-$</td>
<td>1.62</td>
<td>1.71</td>
</tr>
<tr>
<td>$B^0_s \rightarrow D^- [K^+ K^- \pi^-] \pi^+$</td>
<td>1.46</td>
<td>1.52</td>
</tr>
<tr>
<td>$B^0 \rightarrow D^+[K^- \pi^+ \pi^+] D^- [K^+ \pi^- \pi^-]$</td>
<td>1.40</td>
<td>1.86</td>
</tr>
</tbody>
</table>

Note that the denominator is reconstructible with $PT(B) > 2$ GeV, $\tau(B) > 0.2$ ps.
1. New HLT scheme in Run-2: sophisticated HLT1 (classifier) and HLT2-Topo
2. Overall (HLT2/HLT1) efficiency improvement: 15-60% for 2.5 kHz (50-80% for 4 kHz) vs Run-1
3. Timing comparison of MatrixNet BBDT vs post-pruning is in progress
4. Looking forward to data taking!
Thank you for attention!
Backup
1 track:
- Tracks preselections:
  - $PT > 500$ MeV;
  - $IP_{\chi^2} > 4$;
  - $track_{\chi^2/ndof} < 3$;
- BDT uses $PT$, $IP_{\chi^2}$
- Output rate 100 kHz

2-body SVR:
- Tracks preselections:
  - $PT > 500$ MeV;
  - $IP_{\chi^2} > 4$;
  - $track_{\chi^2/ndof} < 2.5$;
- SVR preselections:
  - $PT > 2$ GeV;
  - $vertex_{\chi^2} < 10$;
  - $1 < MCOR$ GeV;
  - $2 < \eta < 5$ (PV to SVR)
- Don’t use MCOR in BDT (from a systematics perspective)
- BDT variables: sum $PT$, $vertex_{\chi^2}$, $FD_{\chi^2}$, $N$(tracks with $IP_{\chi^2} < 16$)
- Output rate 50 kHz
The same preselections as for 2-body SVR

- Changed track $PT > 200$ MeV
- Added $MCOR < 10$ GeV
- Added $N(\text{tracks with } IP_{\chi^2} < 16) < 2$
- Used any min $PT$
- BDT variables: $n, MCOR, \text{sum } PT, vertex_{\chi^2}, \eta, FD_{\chi^2}, \text{min } PT, IP_{\chi^2}, N(\text{tracks with } IP_{\chi^2} < 16), N(\text{tracks})$
- Output rate 2-4 kHz
HLT2: n-bodies comparison for other modes

![Graph showing efficiency for different n-bodies modes](image-url)
HLT2: models comparison

The diagram shows the efficiency of different models for various decay channels:

- $B^0 \rightarrow K^+ K^-$
- $B^- \rightarrow D^- K^+ K^-$
- $B^+ \rightarrow K^+ \pi^-$
- $B^+ \rightarrow K^- K^+ \pi^-$
- $B^0_s \rightarrow \psi(1S) \mu^+ \bar{\nu}_\mu$
- $B^0_s \rightarrow D^- K^+ K^+ \pi^-$

Different models are represented by different colors:
- 'MN reweight', 2500.0
- 'blend', 2500.0
- 'TMVA', 2500.0
- 'AdaBoost', 2500.0
- 'MN', 2500.0
HLT2: models comparison for other modes

![Graph showing efficiency for various decay modes]
HLT2: efficiency vs output rate for other modes
BBDT vs Post-pruning efficiencies for other modes

T. Likhomanenko
tatiana.likhomanenko@cern.ch

LHCb Topological Trigger Reoptimization 28 / 16