Boosted Decision Tree Aproach to Track Finding in the COMET Experiment

Ewen Lawson Gillies

Imperial College London
High Energy Particle Physics

May 30th, 2016

Overview

Developing an algorithm to distinguish between signal and background hits using Gradient Boosted Decision Trees and Reweighted Inverse Hough Transforms.

1. **Charged** Lepton Flavour Violation
2. The **Co**herent **M**uon to **E**lectron **T**ransition (COMET) experiment
3. **G**radient **B**oosted **D**ecision **T**rees (GBDT) and Hough Transforms in Track Finding

This algorithm was developed in collaboration with data scientists at Yandex, Russia's leading search engine.

Lepton Flavour Violation

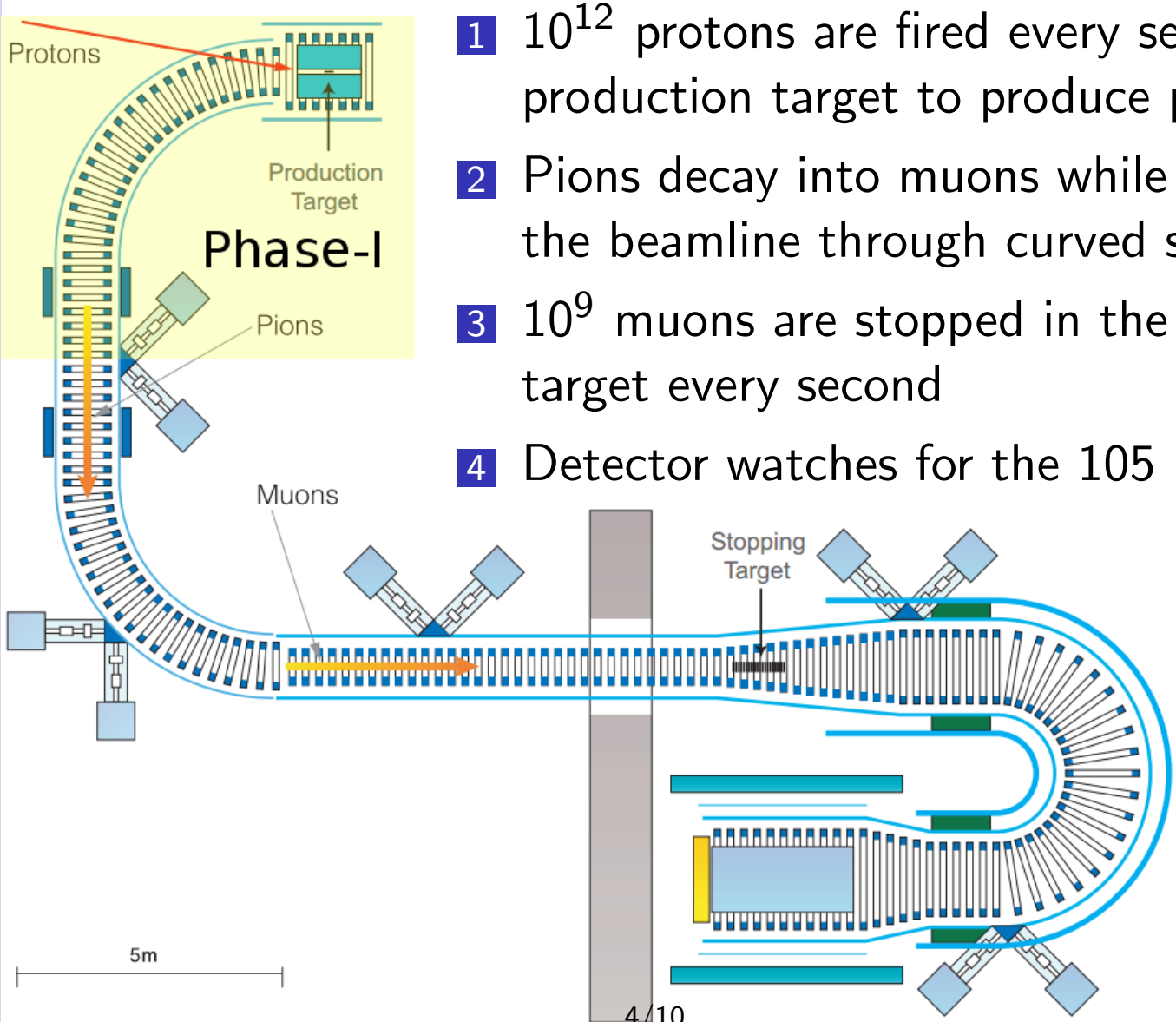
In the standard model, lepton numbers are conserved.
Muon Decay : $\mu^- \rightarrow \nu_\mu + e^- + \bar{\nu}_e$
Muon Capture : $\mu^- + N \rightarrow \nu_\mu + N'$

Do the charged leptons violate this conservation law of the Standard Model?

The three main places this is tested for are:

"μ to three e" : $\mu^+ \rightarrow e^+ + e^+ + e^-$
"μ to e–γ" : $\mu^+ \rightarrow e^+ + \gamma$
Muon to Electron Conversion : $\mu^- + N \rightarrow e^- + N$

COMET Design [1]



1. 10^{12} protons are fired every second at the production target to produce pions
2. Pions decay into muons while flying down the beamline through curved solenoids
3. 10^{9} muons are stopped in the aluminium target every second
4. Detector watches for the 105 MeV electrons

Classification Problem

"Is this wire a signal hit or a background hit?" This is *not* track fitting. This is hit point filtering.
Hit wires have three main features:

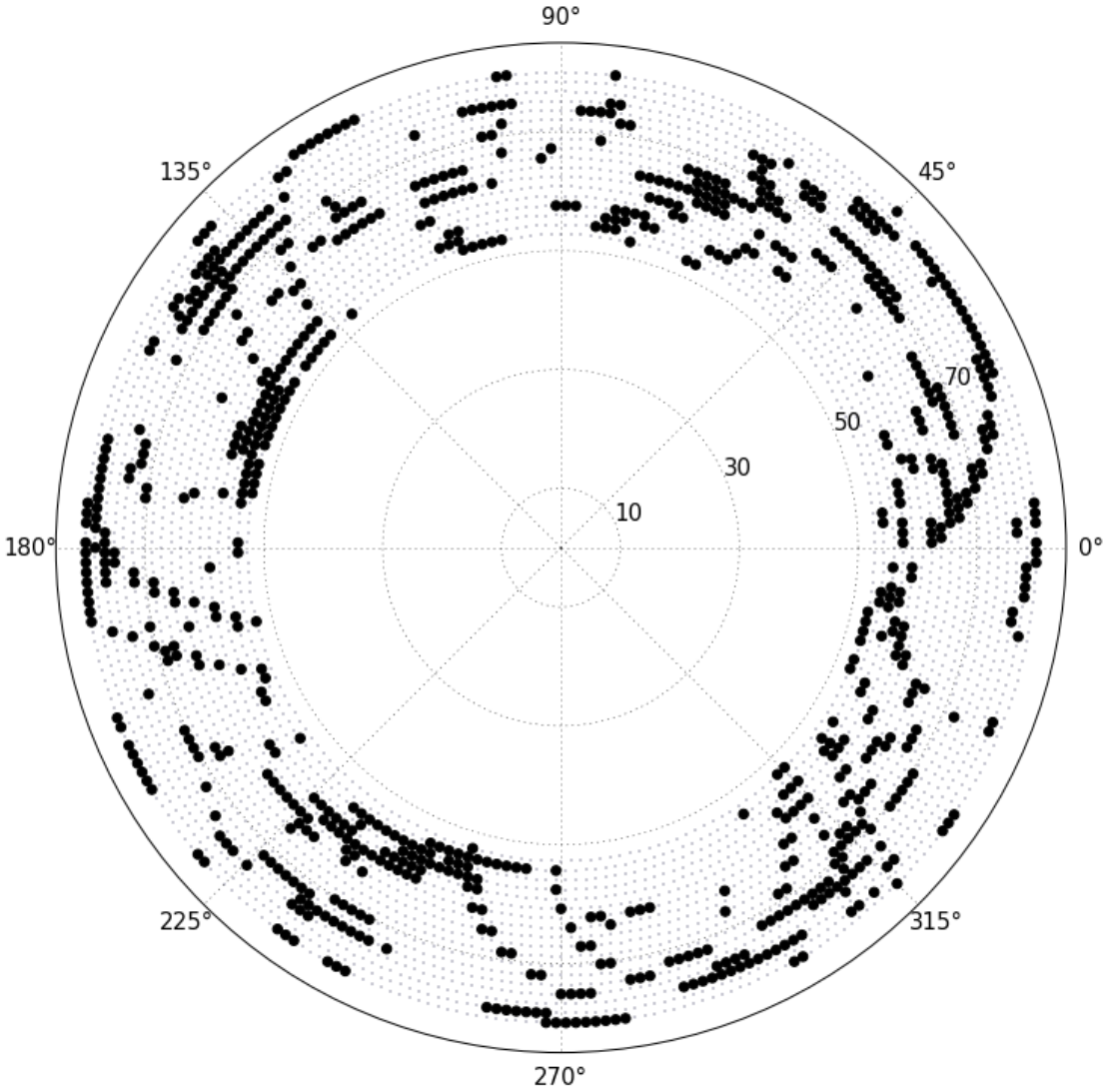
- Radial distance from centre.
- Energy deposited by charged particle.
- Timing of energy deposition.

Define categories of features:

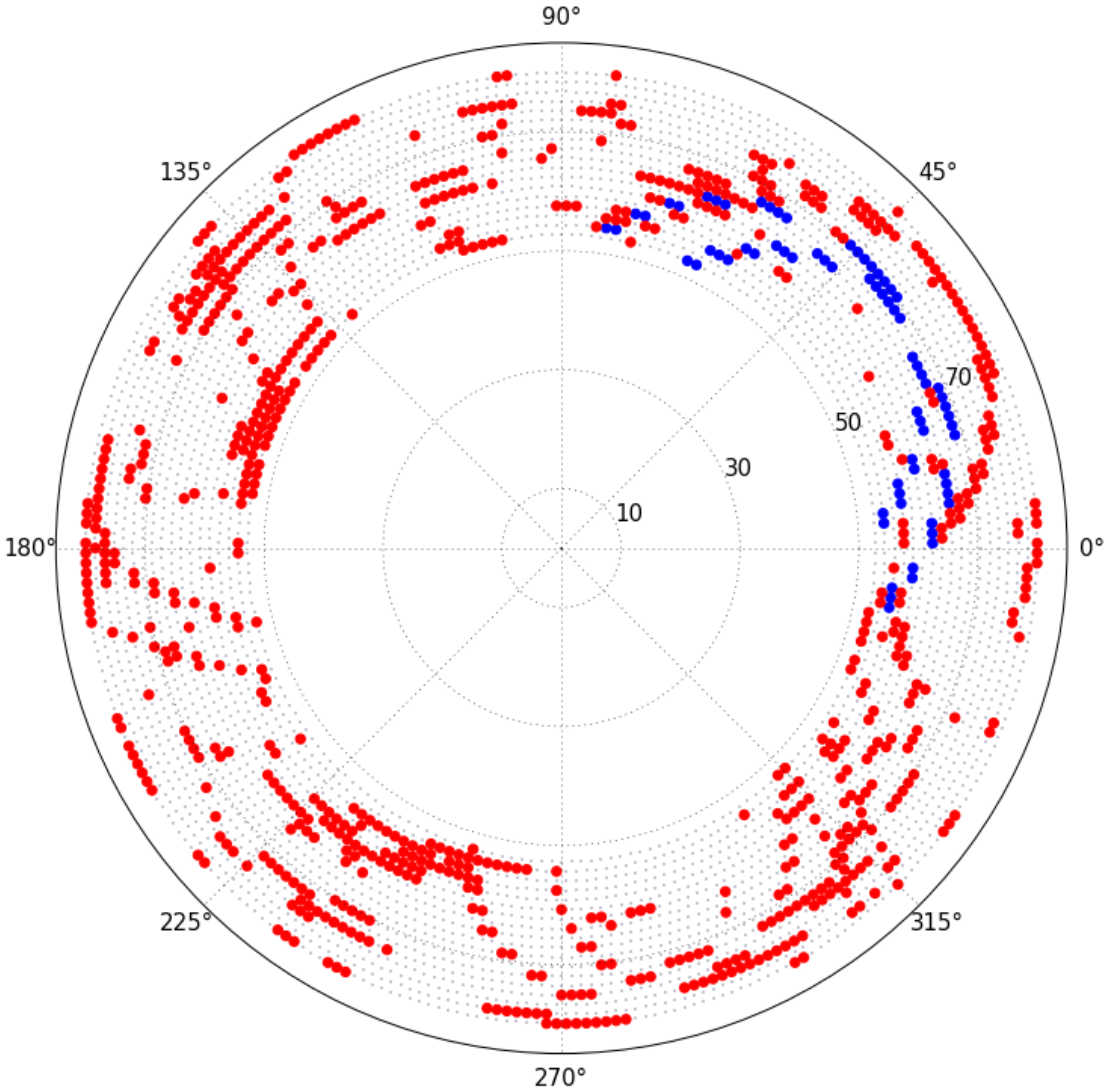
1. "Local" Features : Features on the wire itself
2. "Neighbour" Features : Features of adjacent wires
3. "Shape" Features : Check if the wire forms a circle with other hit wires

Combine the results into a classifier, remove background hits, and define signal tracks. Test and tune this against simulated data.

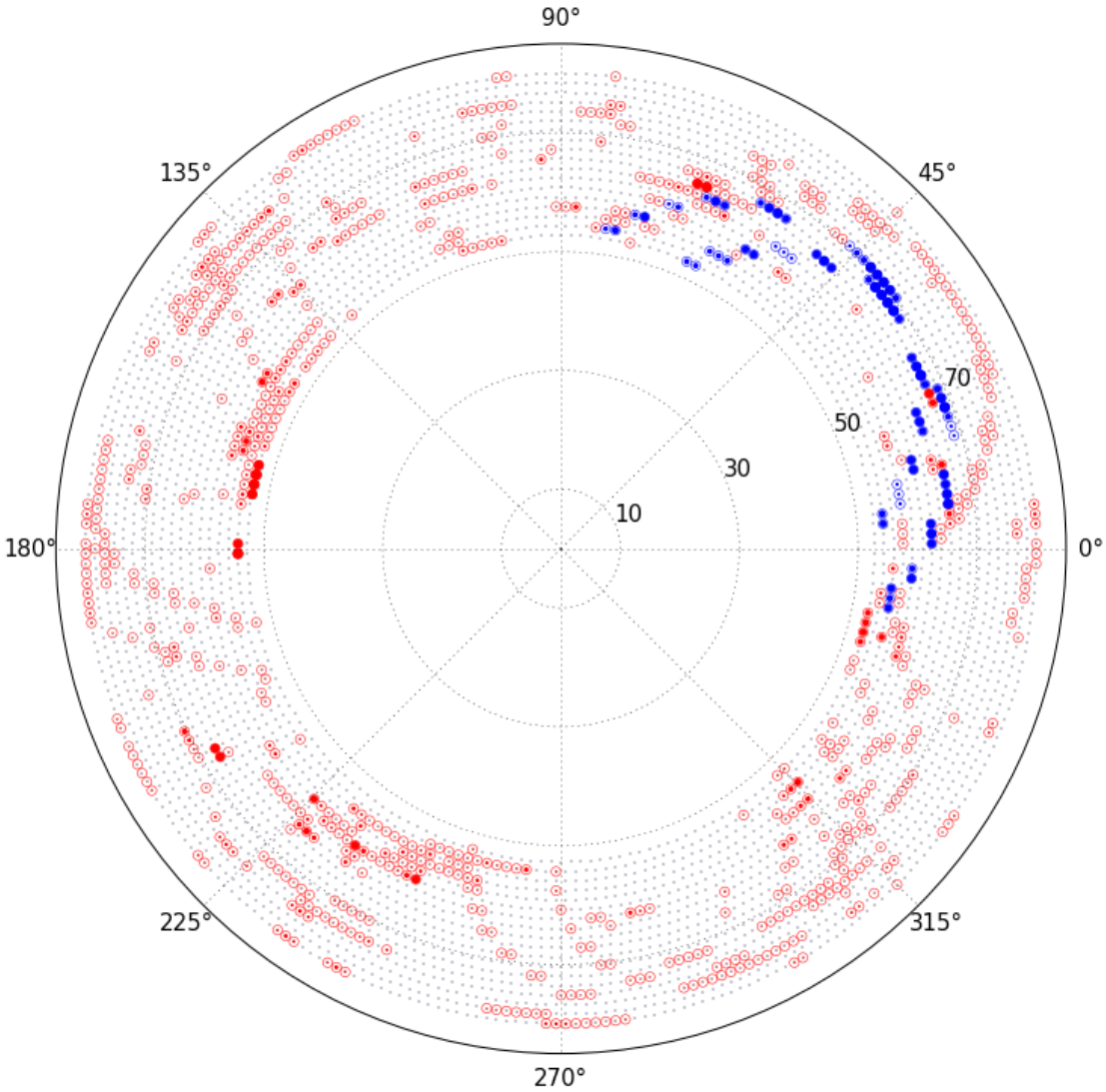
Typical Event [1]



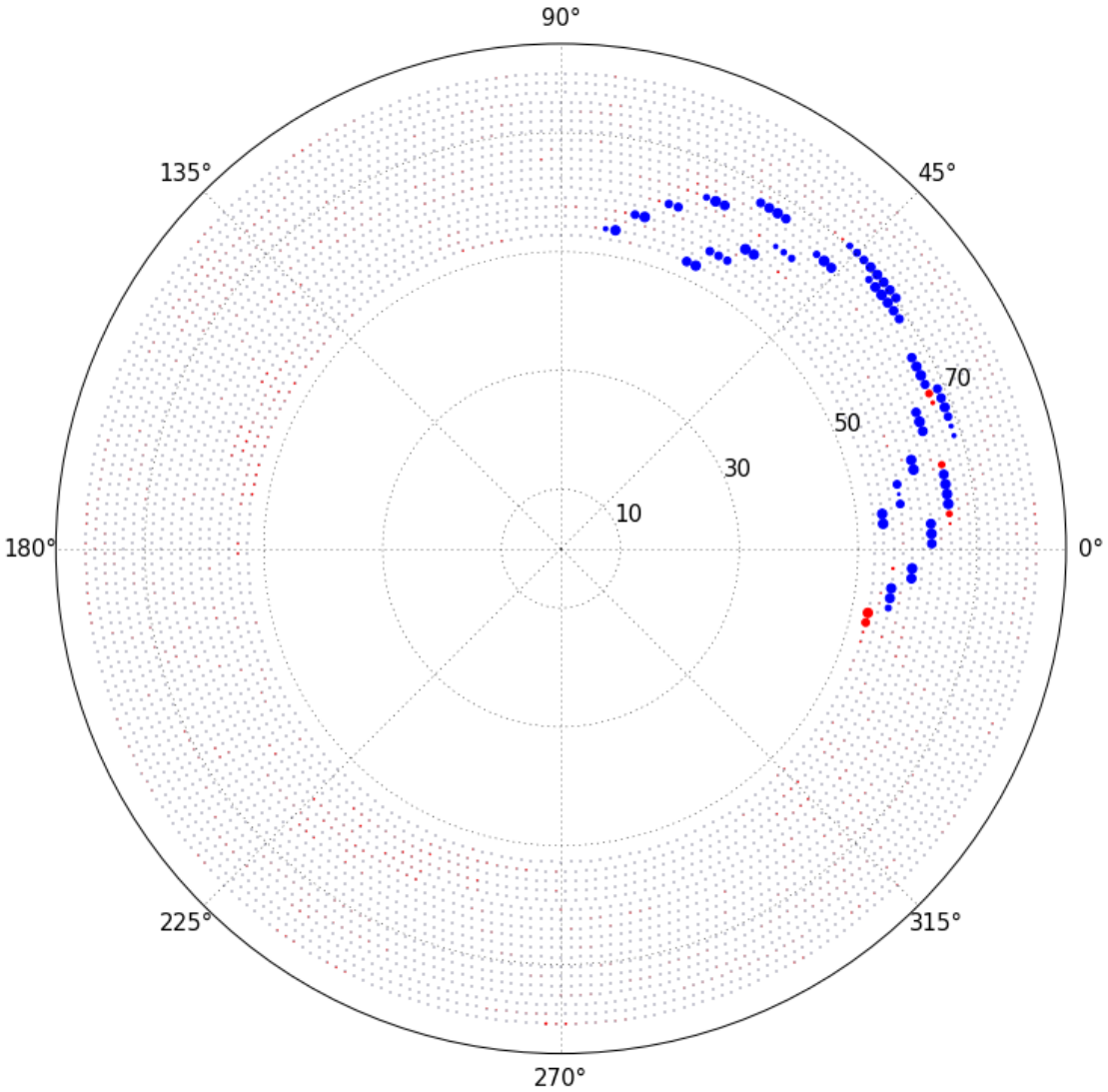
Typical Event [2]



Neighbour-Level Classification



Track-Level Classification



ROC Curves

Compare the performance of using only the energy deposit, local-level features, neighbour-level features, and track-level features.