A new technique for the reconstruction of the interaction vertices inside the CMS detector at HL-LHC

Giovanna Salvi[1,2], Stefano Moretti[1], Claire Shepherd-Themistocleous[2], Emmanuel Olaiya[2]

[1] University of Southampton, [2] Rutherford Appleton Laboratory

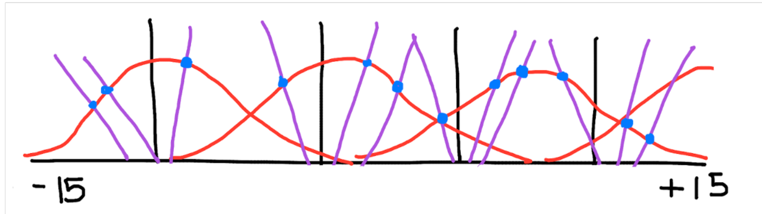
1 Introduction

- For the first time, High Luminosity (HL-LHC) will be able to use track information coming from the tracker at the L1 trigger
- The aim of this project is to develop a technique to identify the hard interaction vertices produced from the p-p collision. The idea is to implement it later on Field Programmable Gate Arrays (FPGAs) to be included in the L1 trigger
- We are currently exploring the potential of the approach. We report here the performances obtained so far and the comparison with much more sophisticated Machine Learning techniques

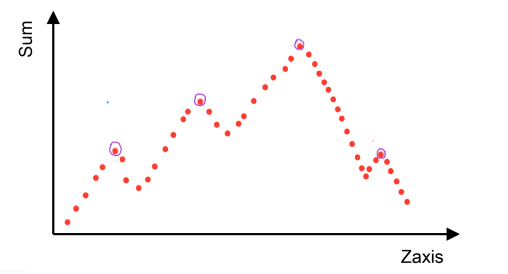
2 Peak Finding Algorithm (PFA)

- The algorithm uses a Gaussian shape distribution that moves along the Z axes of the CMS detector with a certain step-size. The Gaussian has a width that corresponds to $\pm 3\,\sigma$
- It scans over all the tracks it encounters and sums all the values of the tracks under the Gaussian shape distribution as: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

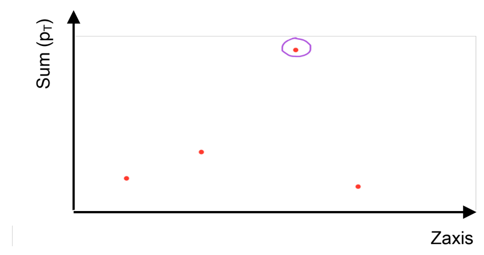
(where σ is the scan width, x is the Z value of the track and μ is the centre of the moving Gaussian)



- This creates a sum output where each value corresponds to a different Gaussian central value



- At this point, the algorithm repeats the scanning process, summing all the tracks only around the Z corresponding to the selected maxima, with the difference that this time the tracks are weighted by their p_T
- The greatest output after this process is recognized as the position of the hard vertex in the event



3 Monte Carlo Samples

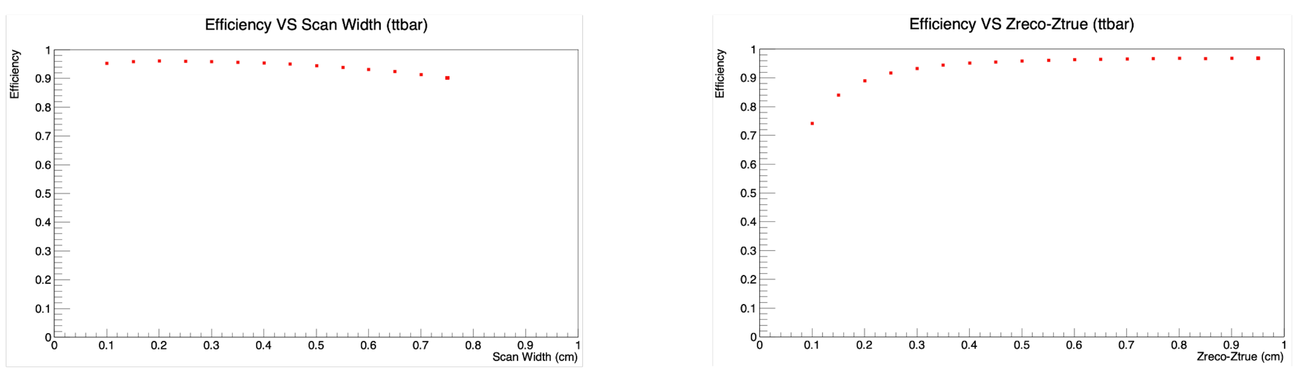
This algorithm has been applied to three different Monte Carlo (MC) samples containing HL-LHC events:

- **ttbar** sample (high multiplicity events)
- **gg-> H->$\mu^+\mu^-$** sample (low multiplicity events)
- **QCD** sample (high multiplicity events)

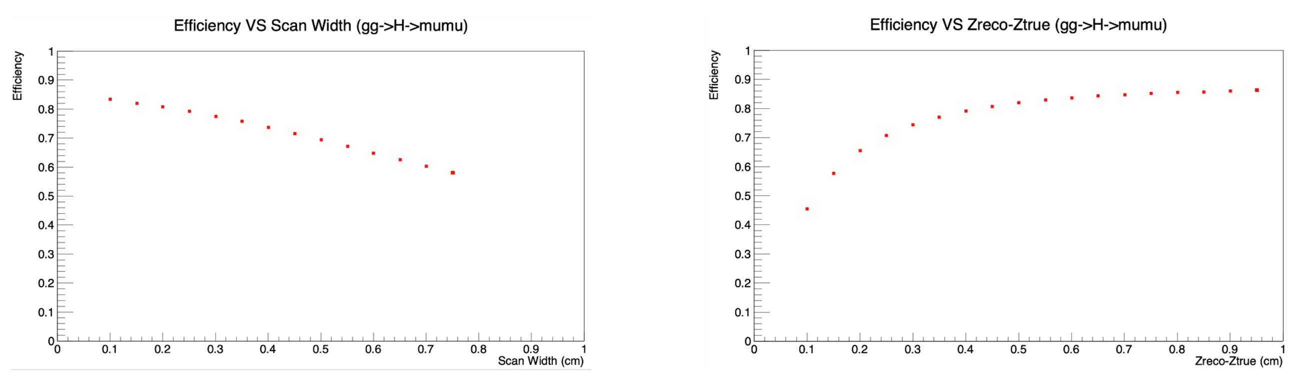
The same samples were also studied with Recurrent Neural Network (RNN), Convolutional Neural Network (CNN) and DBscan (which is another non-machine algorithm that looks at the density of the tracks to identify a vertex).

4 Performances of PFA

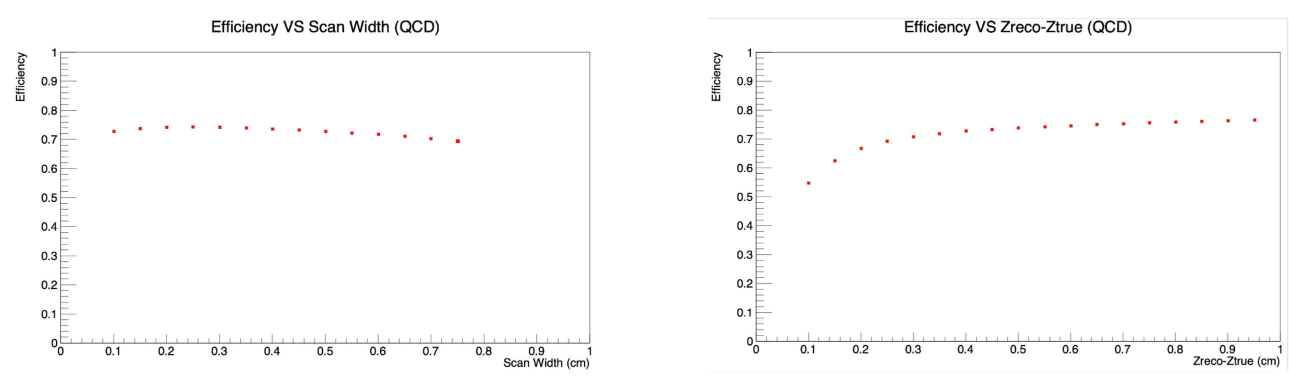
We report performances plots produced using the PFA in the three samples.
On the left-hand side, we have efficiency vs. scan width and on the right-hand side we have efficiency vs. Zreco-Ztrue where Zreco-Ztrue is the maximum distance between the reconstructed vertex and the simulated one.
We have set as reference values: scan width = 0.15 cm, Zreco-Ztrue = 0.5 cm, step size = 0.005 cm.



With the reference values the efficiency achieved is 0.959



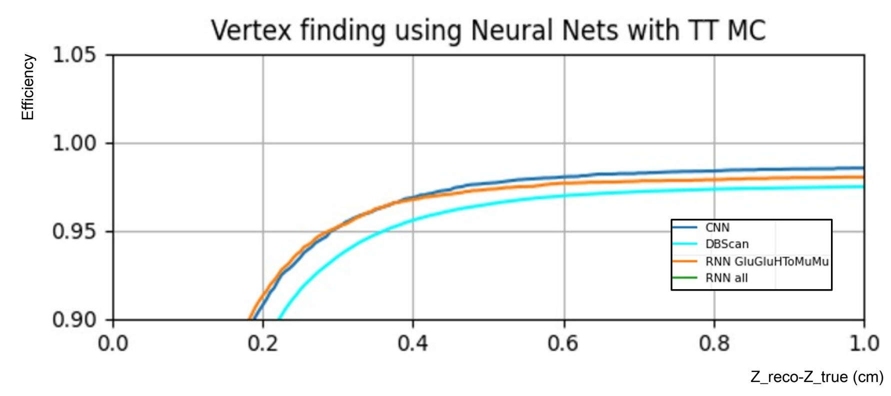
With the reference values the efficiency achieved is 0.819



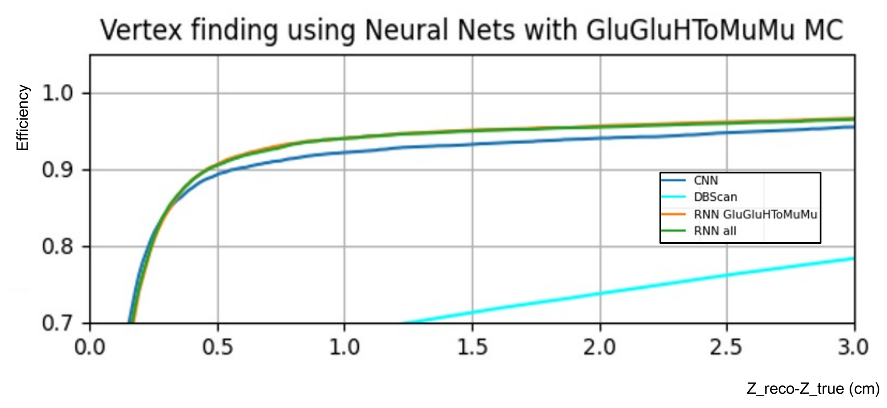
With the reference values the efficiency achieved is 0.742

5 Comparison with Machine Learning (ML) Techniques

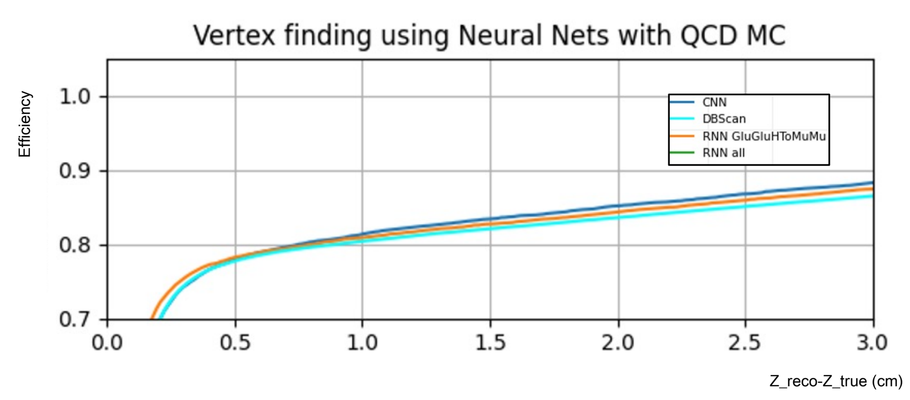
Plots showing Efficiencies vs. Zreco-Ztrue with ML Techniques are presented for the three samples. On the right-hand side of the plots, there is the comparison between CNN, RNN, the PFA and DBscan performances.



ttbar events (reference values):
- CNN-> efficiency 0.973
- RNN-> efficiency 0.977
- DBScan-> efficiency 0.965
- PFA-> efficiency **0.959**



gg->H->mumu events (reference values):
- CNN-> efficiency 0.931
- RNN-> efficiency 0.949
- DBScan-> efficiency 0.65
- PFA-> efficiency **0.819**



QCD events (reference values):
- CNN-> efficiency 0.782
- RNN-> efficiency 0.779
- DBScan-> efficiency 0.775
- PFA-> efficiency **0.742**

6 Conclusion and Future Work

- These first results are promising because even before adding the η information of the tracks (which is taken into account by RNN and CNN), the efficiency of the vertex reconstruction is very close to the one achieved by MC techniques. The latter, being more elaborated are not the most suitable to be implemented on FPGAs which instead require a simpler algorithm like the PFA
- Next step will be adding the eta information and see if we can further improve the performances of this approach, for example, making the width of this Gaussian dependent on the resolution of the tracks it encounters (and so dependent on η)
- Another procedure to reduce the amount of computation has been studied. This procedure parallelize some of the PFA steps asking for less computational resources and being even more appropriate for the implementation of FPGAs

UNIVERSITY OF Southampton

Science & Technology Facilities Council
Rutherford Appleton Laboratory

HEPP & APP Annual Conference 2022
3-6 April 2022, Rutherford Appleton Laboratory STFC