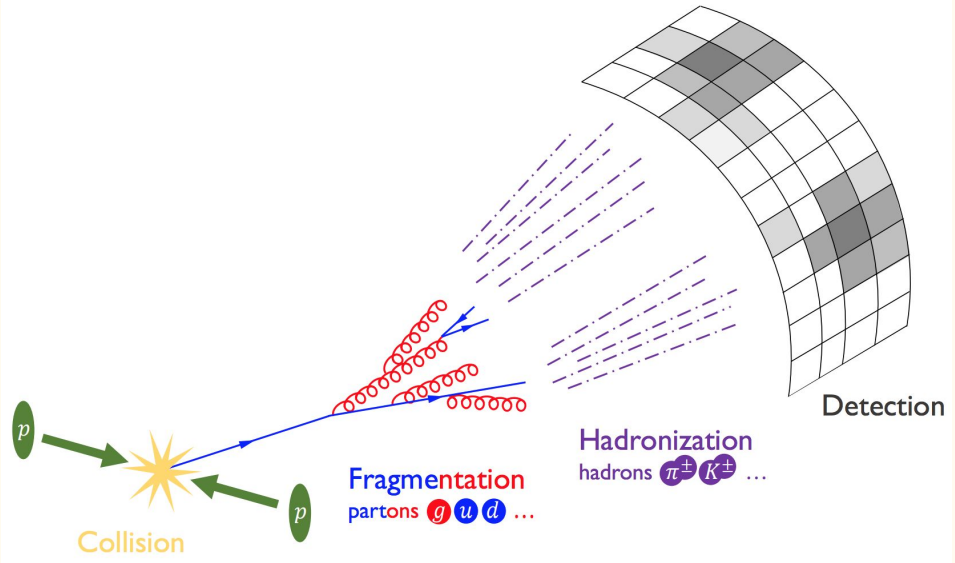
ML4HEP Workshop 2026 @ TIFR

Workshop exercise: Jet-Tagging

Saranya Ghosh
IIT Hyderabad

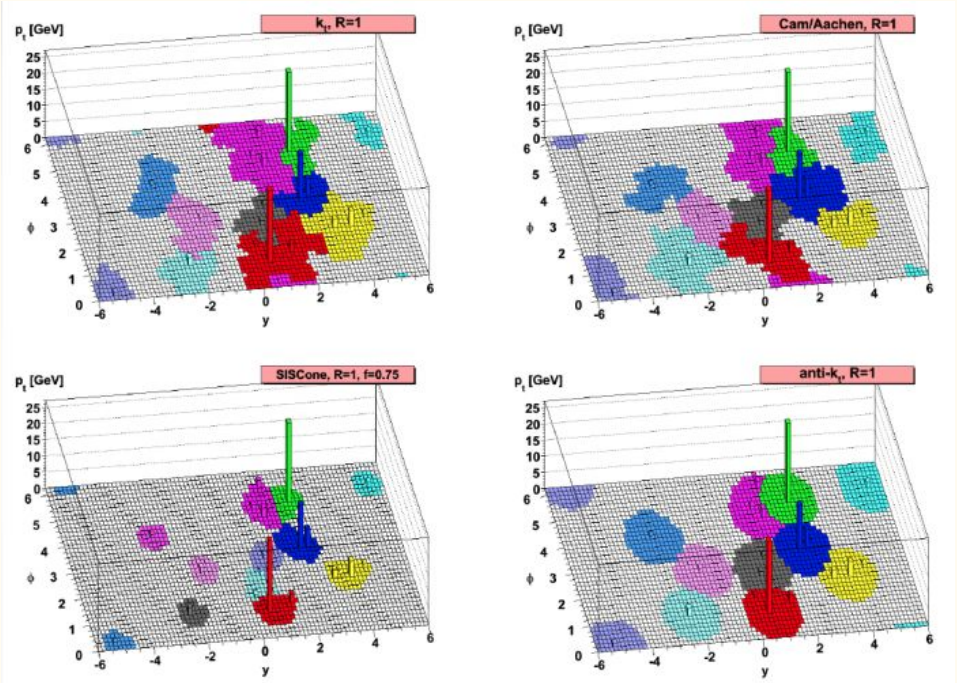
Jets

- A jet is a narrow cone of hadrons and other particles produced by the hadronization of quarks and gluons.
- Quarks and gluons produced in particle collisions cannot exist freely (QCD confinement) - they fragment and hadronize into detectable hadrons that travel together in the same direction.
- Narrow cone of multiple particles hits the detector from where jets can be reconstructed.



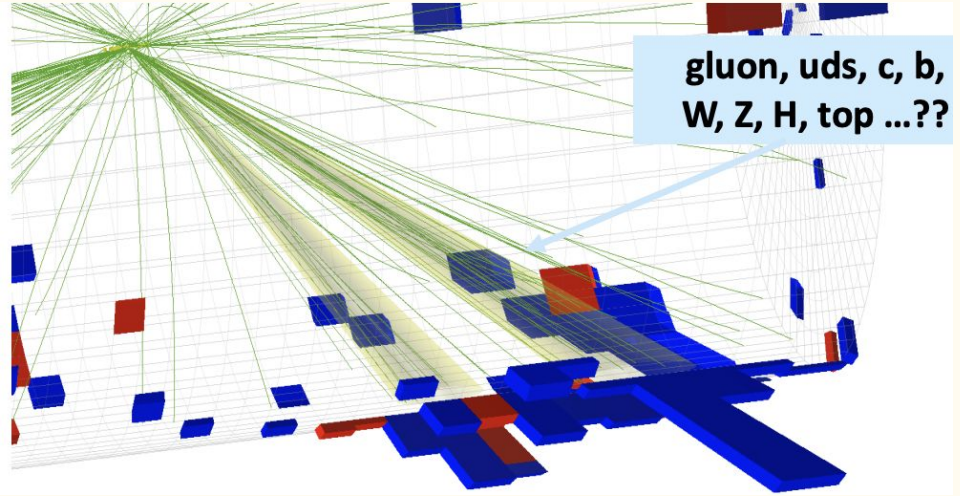
Jet reconstruction

- Jets are reconstructed based on hits / energy deposits in the detector clustered in a region.
- Different jet algorithms are used to define the kinematics and shape of the jets.
- More details on jet clustering algorithms: here
- The radius of jet cone R is a key parameter for such algorithms.



Jet tagging

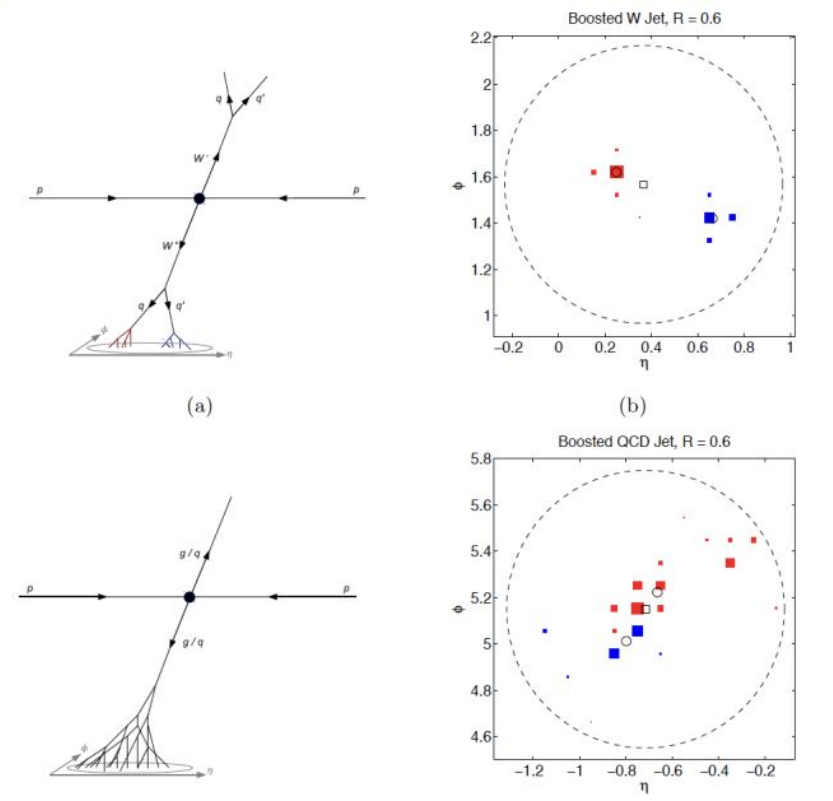
- Jets can originate from different initial particles: gluons; u,d,s quarks; b-quarks, c-quarks, W/Z, H etc
- Jets originating from different particles have different characteristics
- People have been tagging jets for more than 30 years at colliders starting with b jets at LEP and Tevatron.
- It is only recently at the LHC that we have begun to develop powerful tools with tagging capabilities for even more particles.



- This has the potential to open access to many new physics topics that had been written off before.

Exercise : separating jets - W vs QCD

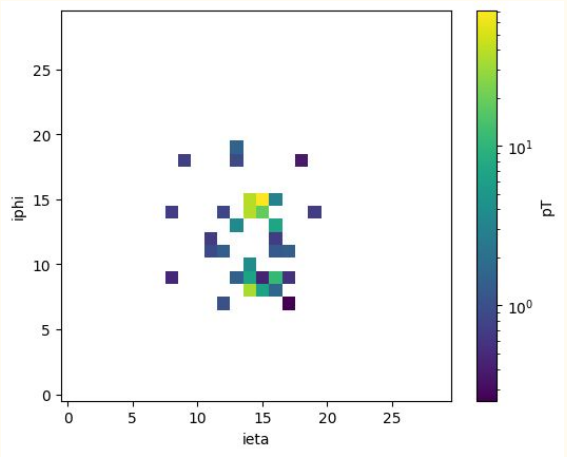
- Jets associated with W boson originate from the decay of W boson into quarks and subsequent hadronisation of quarks.
  If W is boosted, then it may appear as one jet to experimental observer
- QCD jets are those from gluons or light quarks which are the most dominant sources of jet production at the LHC
- An algorithm differentiating the two can have multiple use cases for SM and BSM studies with W bosons



Exercise : separating jets - W vs QCD

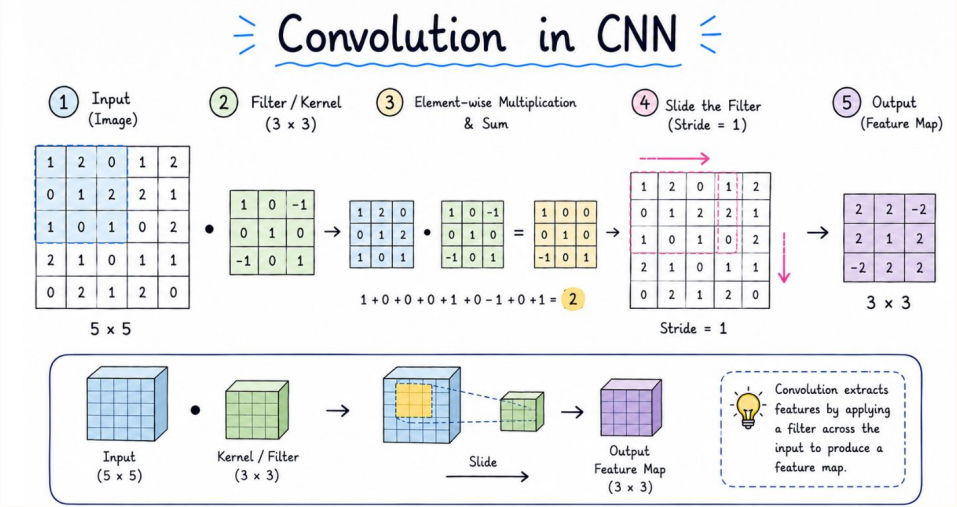
- Goal: to use “images” from the detector energy deposits to differentiate W and QCD jets
- Initial idea: to develop a CNN (convolutional neural network) to tag W jets against QCD jets
- Since it is an image differentiation problem - CNNs are apt. Can explore other models if successful implementation of CNN is achieved.

https://arxiv.org/pdf/1511.05190



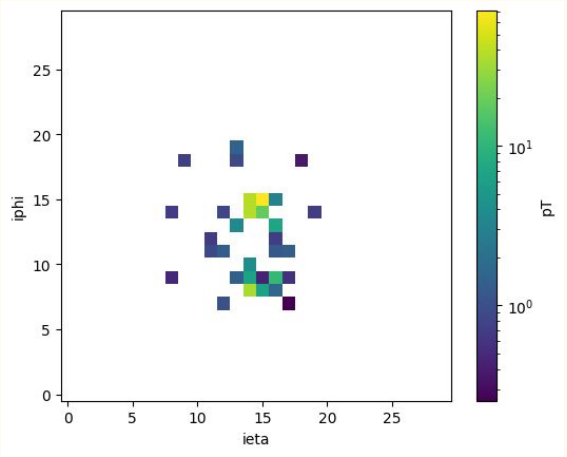
Convolutional Neural Networks

- Convolutional Neural Networks have been found to work well for Computer Vision problems
- Convolution is using a ‘kernel’ to extract certain ‘features’ from an input image.
- Covered during pre-school lectures: https://scitalks.tifr.res.in/event/9299/timetable/#20260522
- Another short reference: here



Implementation

- Information to form the “Images” available in file in folder: link
  - Labeled “QCD” and “TT”
  - # 1st dim is jet index
  - # 2nd dim is eta bin
  - # 3rd dim is phi bin
  - # 4th dim is pt value
- For some ideas on architecture, can base it on paper: https://arxiv.org/pdf/1511.05190



Summary & possible extensions

- Goal: to use “images” from the detector energy deposits to differentiate W and QCD jets
- Initial idea: to develop a CNN (convolutional neural network) to tag W jets against QCD jets
- If you can implement in time then the work can be extended to more complicated problems of jet classification using more advanced networks based on transformers
- Can be discussed further on Friday

Link:
https://drive.google.com/drive/folders/1Os805lDwRJGSMPsib2fkvAECi914NvKL?usp=sharing