Search for vector-like B quarks in CMS Run 2

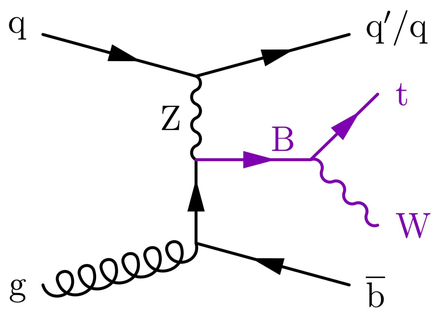
Dr. Julie Hogan, Kalin Johnson (Bethel University), Xiaohe Shen (Brown University)
on behalf of the CMS collaboration

Vector-Like Quarks

The **Compact Muon Solenoid experiment** (CMS) at the CERN Large Hadron Collider records proton-proton collision data in order to study the particles and forces that exist in very high energy conditions. The discovery of the Higgs boson in 2012 was a triumph for the field of particle physics, but pointed toward the **probable existence of unknown high-mass particles.**

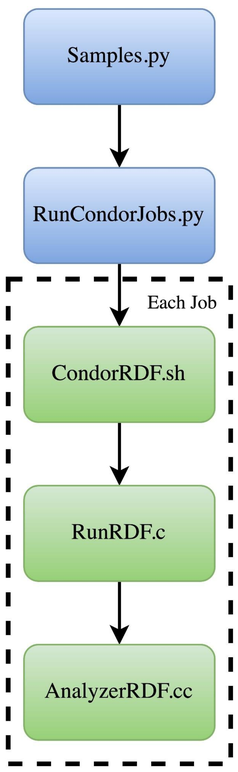
Vector-like quarks (VLQs) are heavy fermions that are predicted in several "new physics" models. Their decays to lighter particles like bottom or top quarks and W, Z, or Higgs bosons create exciting detector signatures that we can search for in CMS.

The VLQs "T" & "B" would be similar to standard model t & b quarks → but much much heavier!



Job Control Development

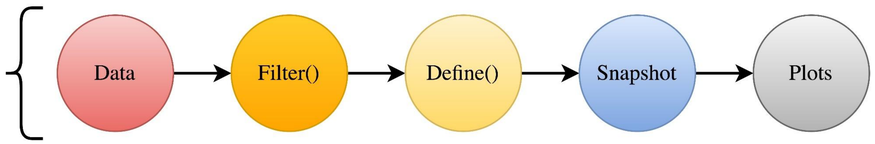
CMS physics analyses run best in parallel! Over the summer, I took the previous system for submitting HTCondor jobs and improved several steps in the process by adding and modifying code.



- The first file in the chain is a Python file, which contains a list of over 200 samples and dictionaries with information about those samples.
- Another Python program is then used to submit the condor jobs.
- Each condor job first runs a bash script, which initializes the analyzer.
- Finally, the analyzer in a C++ file is run on data from a specific sample.

- Inside the Analyzer: **RDataFrame**
  - RDataFrame provides common operations, low-level details, and looping over entries for ROOT analyses
  - The analyzer use RDataFrame to process the data, implement selection criteria, and build snapshots
  - The 2 key methods in RDataFrame are **.filter()** and **.define()**
  - **filter()** takes in an expression and returns whether events passed
  - **define()** creates a new column using known variables and functions

The analyzer results are stored via a "Snapshot" and can be plotted.



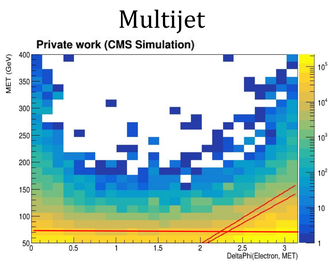
Optimizing the Selection

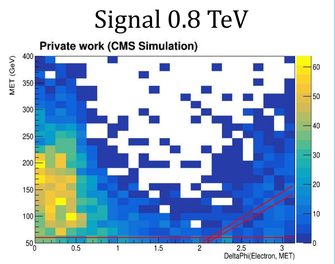
Multijet production is a difficult type of background for events with electrons. One way to reduce this is to study the relationship between MET & the electron.

We can drop background & keep signal with these requirements:

$$\text{MET} > 60$$

$$\text{MET} > \frac{130}{1.5}x - 130$$

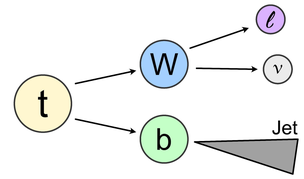


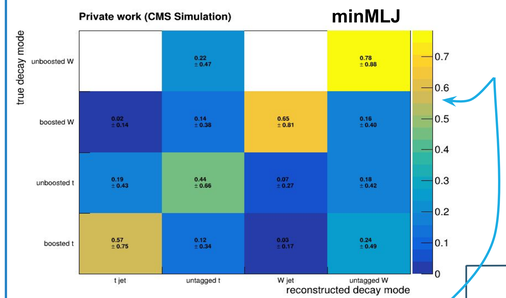


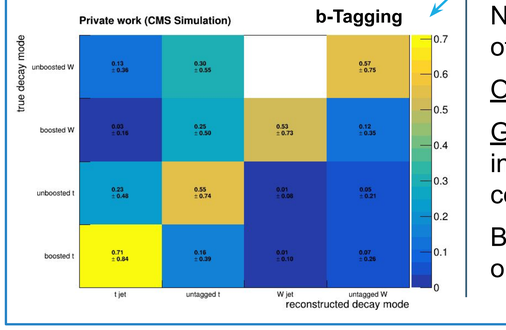
Top Quark Reconstruction

To choose our method for top quark reconstruction, we first compared two options:

1. **"minMLJ":** the smallest jet + lepton mass in the event is < 170 GeV
2. **"b-Tagging"**: a jet on the "lepton side" of the event is tagged as a b quark







Goal: sort events into the correct bucket on the diagonal line in confusion matrices.

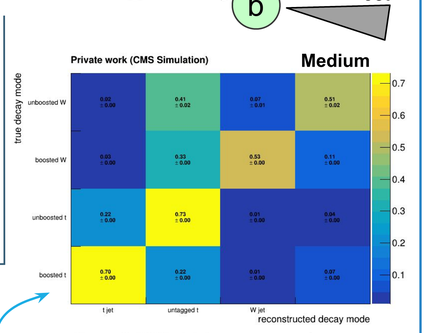
The results showed **b-Tagging** was the better choice.

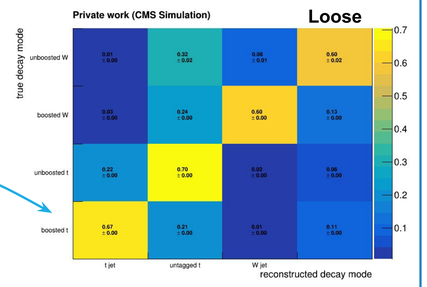
Next, we had to set the strictness of the b-tagged jet selection.

Options: **loose** vs **medium** cuts.

Goal: have most of the events fall in the boxes in a diagonal line in confusion matrices.

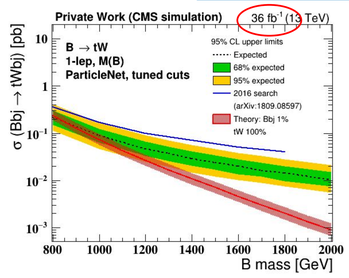
Based on the results, we decided on the **loose** criterion.

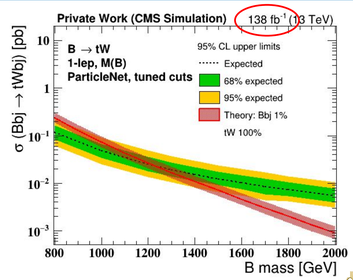




Further Development

- This search is actively in progress!
- We are currently developing a thorough background modeling technique
- We see signs of improvement compared to the previous analysis, which is represented by the blue line in the left plot
- **Using all of the Run 2 data from CMS will also improve the result!** (right plot)





Funding for this project was provided by Bethel University and National Science Foundation Award #2110972.