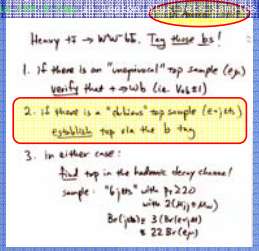
Development of B-Tagging Using Jet Probability at CDF and Some First Signs of Top

David Gerdes, University of Michigan

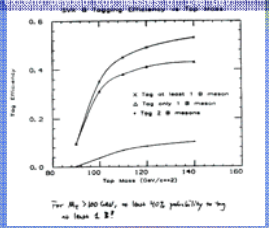
Abstract: During the top quark discovery period from 1992-95, I was a postdoc in the CDF group at the University of Michigan under the leadership of Dan Amidei and Myron Campbell. Other members included postdocs Paul Derwent, Soo Bong Kim, and Steve Vejcik. Shortly after I arrived in the fall of 1992, I began working on the top search using b-tagging with CDF's new silicon vertex detector. In March 1993, I came across an Aleph note by Dave Brown and Markus Frank that described a b-tagging algorithm based on combining track impact parameters into an overall "jet probability." I immediately started adapting this algorithm for CDF. Very soon we had the start of a promising top signal. This poster shows some of the early efforts to understand this algorithm, and how the top quark began to emerge from 13 pb^{-1} of data…

Early Inspiration

In 1988-90, Dan Amidei, who later became my postdoctoral mentor at Michigan, was a postdoc at Fermilab and was coordinating the construction of CDF's first Silicon Vertex Detector (SVX). By then, there were strong indications that the top quark was "heavy", with a mass greater than about 90 GeV. In this case, the top quark will decay to a real W and a b. In a CDF talk in March 1989, Dan laid out a top physics agenda that included the idea to separate top events from W+jets samples…



About a year later, he had worked out some of the details. This plot shows an estimate of the top b-tagging efficiency, using a generator-level study of a simple b-tag algorithm based on counting the number of displaced tracks. The estimate of ~40% efficiency is very close to what we achieved in Run I--but note that the plot does not account for the loss of efficiency with top mass!



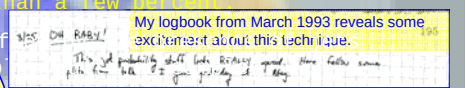
Jet Probability Basics

Top quark events contain multiple jets. Unfortunately so do the backgrounds.

Top quark events typically contain two b-quark jets. Background events are far less likely to contain any b's. By identifying b-jets in samples containing a mix of top and background, we enhance the top fraction significantly.

Jet probability provides one way of distinguishing b-jets from other jets, by exploiting the fact that the b-quark is long-lived. Tracks from b-jets tend to be displaced from the collision point. Using precision tracking information from the silicon vertex detector, we could characterize each jet in an event by a single number: the probability that it is consistent with originating from the collision point. "Normal" jets should have jet probabilities uniformly distributed between 0 and 1, while b-jets should have low jet probabilities, typically less than a few percent.

Top could therefore be identif… with low jet probability in W …

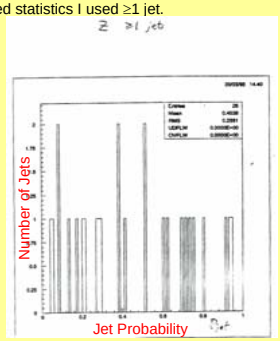


My logbook from March 1993 reveals some excitement about this technique.

Our First Look at W, Z+jets data: April 1993

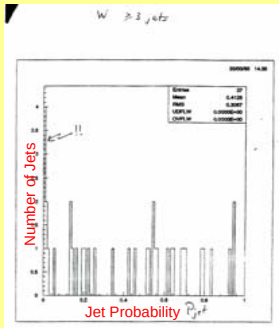
On April 1, 1993 I presented our first results with the new b-tag algorithm at a top group meeting. At that point we had 13 pb^{-1} of good data to analyze, and SVX tracking was pretty well understood. There were some tantalizing hints of top…

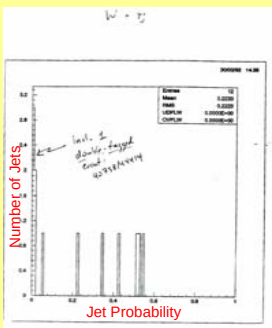
The Z+jets sample is a useful control sample because it contains many of the same backgrounds as the W+jets signal samples, but very little top. This plot shows the jet probability distribution in Z+jets data. It is a relatively flat distribution, with no events in the JP<1% bin. This told us that there were not a lot of b-quarks produced in association with Z (and presumably W) bosons. Ideally I would have used a ≥3 jet sample here, but because of limited statistics I used ≥1 jet.



Next I went for the gold: the W plus ≥3 jet sample. Here is where we expected the top signal to be strongest. Sure enough, there was an excess of events in the lowest-probability bin--which rated two exclamation points in my talk. This meant there were b-quarks in this sample!

In the W+4 jet subsample, where the top signal-to-background should be best, there was even a double-b-tagged event.





Conclusions

- By April 1993 we had hints of a top signal. But we needed more data and a better understanding of the backgrounds before we could be sure. It took another year before we published "Evidence for Top".
- Jet Probability was one of three b-taggers developed for this analysis. The interplay between the various taggers in the CDF b-tag Working Group fostered a much better understanding of b-tagging at hadron colliders, and led to a very stimulating (and competitive!) intellectual environment.
- The opportunity to participate in the top quark discovery was the thrill of my professional lifetime--so far!

Background Image: *TopMuon Event from the Fermilab Tevatron* (detail), from University of Michigan artist Jan-Henrik Andersen's Exhibition entitled "Sized Matter: Perceptions of the Extreme Unseen"