UNIVERSITY OF AMSTERDAM

Single Top Quark Production at the LHC. Date Processing and Cross Section Measurement

H. Lee

Summary

In the hadron collider, single top quark production through electroweak interaction servers as a tool for testing the Standard Model of particle physics and for probing new physics. Using the data collected by the ATLAS experiment in 2011, the study described in this thesis focuses on the search of single top quark production in the Wt-channel, a process that is not observed before the LHC.

Given that the signal strength is several orders of magnitude smaller than the background, the measurement of single top quark Wt-channel production is confronted with two major challenges. The first challenge in computing is to process as many events as possible in a given period of time so that the analysis is capable to handle a large amount of events for sufficient statistics, and to manage necessary iterations for improving the quality of measurement. The other challenge is to develop an efficient approach to disentangle the small amount of signal events from the background.

To tackle the challenge in computing, the offline data processing for enabling physics analysis is managed on a world-wide computing grid system. In 2011, about 47 thousand processor-years of computation have been accomplished by 78 million computing jobs distributed around the globe, producing petabytes of distributed data for physics analysis. The ATLAS computing reaches this scale by paying close attention to the operation of the grid services, thereby minimizing the effect of inevitable failures associated with distributed computing systems. The lessons learned from the operation provide valuable input to improve current systems for less manual intervention and better resource utilization.

The throughput of data analysis, defined as the number of events one manages to analyze per second, has been studied. Given the amount of computing resources, the throughput can be high on the grid, but degrades significantly when several physicists are competing for the available computing power. As the competition is unpredictable, the time it takes for an individual user to complete the data analysis is observed to be largely uncertain on the grid.

The "cloud system", a system based on the virtualization of computing resources, is adopted as an alternative approach for running the data analysis. The system guarantees constant availability of resources, which is beneficial for analyses aiming for a particular deadline. Concerning a small scale analysis, it is demonstrated that the throughput achieved on the cloud can be higher than on the grid. For a full scale analysis, it is observed that the cloud is, however, missing the scalability provided by the grid. Using the BigGrid HPC cloud, the missing scalability is quantified to be a factor of 5, a target to be reached by resolving the bottleneck of the centralized filesystem used for realizing non-interrupted resource availability.

The second part of this thesis describes the search of the single top Wt-channel signal with one charged lepton, missing transverse energy and multiple jets in the final state. For isolating the signal events in data, contributions of QCD multi-jet and W+jets backgrounds need to be determined from data as the simulation of these two processes has large uncertainty. Both the matrix method and the jet-lepton model are applied to estimate the size and the distribution of the QCD multi-jet events in data. After the determination of the quark flavor composition among the W+jets events, the actual size of the W+jets background in data is obtained by exploiting the asymmetry between W^+ and W^- production at the LHC.

For enhancing the single-to-background ratio, a likelihood discriminant is constructed from five physics observables that are carefully selected to achieve better signal discrimination as well as lower systematic uncertainty in the final measurement. By fitting the event distribution on the likelihood discriminant, the single top Wt-channel production cross section in the $p-p$ collisions at $\sqrt{s} = 7$ TeV is measured as

$$\sigma_{pp\to Wt} = 16.2 \pm 5.8 \text{ (stat.)} \pm 12.1 \text{ (syst.) pb.}$$

The central value of the measurement is close to the theoretical prediction of 15.7 ± 1.1 pb, however, the uncertainty is as large as 83%. With the inclusion of the full 2011 datasets, the statistical uncertainty appears not to be the limiting factor. The systematic uncertainty already dominates. The cross section uncertainties of the $t\bar{t}$, W+jets processes, and the modeling of initial-/final-state radiation are the top 3 sources of the systematic uncertainty. The large uncertainty leads to a signal significance of 1.35 σ, which is too small to claim an observation. Thus, the upper limit of the cross section is evaluated to be

$$\sigma_{pp\to Wt} < 41.8 \text{ pb}$$

at 95% CL_s.

Although improvements are needed for the future to reduce the systematic uncertainty, this study addresses the challenges in both computing and physics, performing one of the first measurements of the single top Wt-channel production cross section. It also establishes a basis for further research in this single top quark channel at the LHC.