A New Charm Quark Tagging Algorithm at the CMS Detector

K. Kovitanggoon[1*], B. Asavapibhop[1§], N. Suwonjandee[1], M. Verzetti[2] and S. Moortgat[3]

[1]Department of Physics, Faculty of Science, Chulalongkorn University, 254 Phyathai Road, Patumwan, Bangkok Thailand, 10330

[2]Department of Physics, University of Rochester, USA

[3]Department of Physics, Vrije Universiteit Brussel, Belgium

E-mail: kovitang.cern@gmail.com, burin.a@chula.ac.th[§]

Abstract

At the Compact Muon Solenoid (CMS), both Standard Model (SM) and Beyond Standard Model (BSM) physics processes can result the final states with charm quark jets. Charm quarks hadronize to D mesons which could travel some distance in the CMS silicon tracker before decaying into showers of detectable particles, called jets. Consequently, charm jets can be distinguished by particular properties such as secondary vertices from displaced tracks with respect to the primary interaction. The algorithm to identify charm jets, c-tagging algorithm, is invented based on Combined Secondary Vertex algorithm for b-tagging. C-tagging uses multivariate analysis (MVA) techniques to study a set of jet properties in order to identify jets originated from charm quarks. It is the first of its kind at the CMS collaboration. The c-tagging algorithm is integrated into the CMS software (CMSSW). It will be used in supersymmetry (SUSY) searches for new particles such as stop ($\tilde{t}$), the SUSY partner of standard model (SM) top, that may subsequently decay to a charm quark and the lightest supersymmetric particle (LSP), and for SM precision measurements in the data taking at the Large Hadron Collider (LHC) in 2015 and 2016.

Keywords: CMS: LHC: SUSY: C-tagging

Introduction

The Standard model (SM) is the underlying quantum field theory that deals with the fundamental building blocks of matter and their interactions. SM is an elegant framework and successfully describes three of the four known fundamental forces, the weak, the electromagnetic and the strong force. Although the SM has been established in 60s and 70s, it still holds up to many experiments since then. Particles predicted by the SM were discovered over the years at various experiments. However, it suffers from such shortcoming as the hierarchy problem and cannot include the fourth fundamental force, the gravity. Moreover some questions still cannot be explained with the SM, such as the unification of the four forces at high energy and dark matter. Supersymmetry (SUSY) as the extension of the SM is a popular candidate that postulates the existence of supersymmetric partners for every SM particle. SUSY provides the solution to the hierarchy problem and may provide a framework to integrate gravity into the forces unification. In some SUSY scenarios, it suggests the lightest supersymmetric particle (LSP or $\tilde{\chi}_1^0$) which is neutral and stable. The LSP is a candidate of dark matter. The decay mode of a stop particle is dominated by a charm quark and the LSP ($\tilde{t} \rightarrow c + \tilde{\chi}_1^0$).

The study involves the final state that is characterized by the jet from the hadronization process of the charm quark and the missing transverse energy (E_T^{miss}) from the undetected LSP. The aim of this paper is to present the status of the charm quark tagger in the CMS software framework (CMSSW).

Charm Quark Tagger in the CMS

In the SM, there are six flavours of quarks. They are grouped into three generations i.e. up and down quarks (*u* and *d*), charm and strange quarks (*c* and *s*), and also top and bottom quarks (*t* and *b*). All the quarks, except top quarks which decay in a time scale shorter than the hadronization scale, form hadrons and thus appear as jets after showering inside the detector. The charm quark is the third heaviest quark, $m_c \approx 1.3$ GeV. Considerably heavy and long lifetime, jets that decay from c quarks have the unique properties from those of jets that decay from lighter quarks and gluons. Following the high energy collisions, free quarks and luons are created. The color confinement, which only allows for colorless particles to exist, states that free quarks and gluons, containing color charge fragments, cannot exist in the free space, and thus they cannot be measured by particle detectors. They combine with quarks and anti-quarks, which are created in the vacuum to form the hadrons (the bound states of quarks and gluons). The decay of these hadronization processes can be detected as the tight cone of energy deposited in the calorimeter in term of jets.

The bound states of charm quarks exist in both forms of meson and baryons such as J/Ψ meson consisting of $c\bar{c}$, *D* meson consisting of *c* and another lighter quark or Λ_c^+consisting of *udc* quarks. Due to the high energy of the pp collisions at the LHC and low branching ratios of other bound states, *D* mesons are dominating at the CMS detector. The lifetime of *D* mesons is about 1.04 ps for $D^{\pm}$ and 0.41 ps for D^0. The corresponding relatively long flight paths lead to characteristic signatures of one displaced vertex, called secondary vertex and displaced tracks. The CMS silicon tracker allows the precise reconstruction of vertices and tracks which can be used to distinguish *c* jets from other jets.

The Compact Muon Solenoid (CMS)

The CMS detector is a large, high field superconducting magnet detector. The CMS main design priorities are a redundant muon tracking system, a good electromagnetic calorimeter and a high quality inner tracking system. The CMS structure consists of many cylindrical detecting layers, coaxial with the beam direction (barrel region), closed at both end with disks (endcap region), and large pseudorapidity calorimeter close to beam line (forward region). A schematic view of the CMS detector, which is 28.7 m long, 15 m in diameter, and 14,000 tons of the total weight is shown in Figure 1.

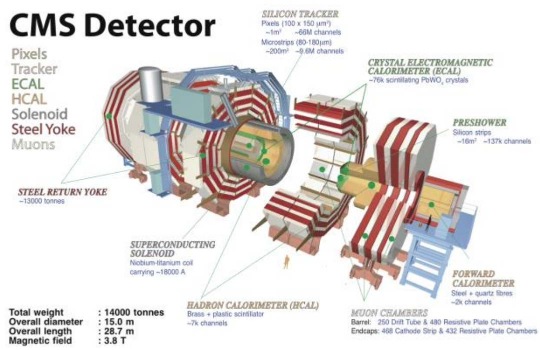

Figure 1 Illustration with details of the CMS detector.

Materials and Methods

Charm Quark Tagger Procedures

From the Monte Carlo (MC) simulation i.e. QCD, the jets with their relevant tagging variables, such as the signed impact parameter significance and the number of tracks associated with secondary vertex as shown in Figure 2 are extracted using the modified BTagAnalyzer. The output variables are subdivided in to different flavours (*c*, *b* and light or also referred to as *dusg*).

The training is performed by Multivariate method (MVA) using the Tool-kit for MVA (TMVA) package inside the ROOT data analysis framework. The well-known MVA technique, Boosted Decision Trees (BDT), is used. BDT uses successive decision nodes to categorize events as either being signal or background. Decision nodes are separated into CvsL and CvsB.

This cutting procedure is repeated for each of the sub-nodes and stops only when one of the nodes reaches a minimum number of events or certain signal purity. The combination of these individual decision trees (i.e. boosting) is done by penalizing after each individual decision tree misclassified events in the final leaves, giving them more weight in the next decision tree that is constructed. After the BDT is trained, jets can be successively subjected to different decision trees and are assigned an estimator (often called discriminator) values based on the number on of times they end up as signal or background.

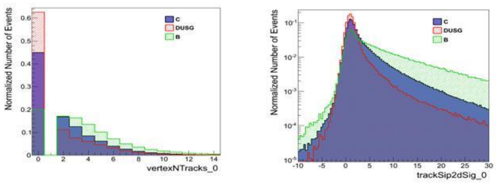

Figure 2 The number of track associated with SV and signed transverse impact parameter significance are shown. The clear distinctions between signal (*c*-jets) and background (*dusg* jets and *b* jets) can be seen.

Results and Discussion

The output discriminator distributions are shown in Figure 3 and the corresponding ROC curves showing the performances can be seen on Figure 4. It can be seen that the CvsL discrimination is actually more effective in the high purity region (large background rejection), while CvsB discrimination is more effective in the high efficiency region (large signal acceptance).

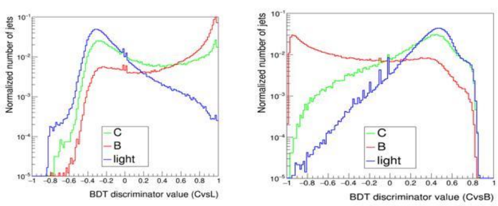

Figure 3 Overlay of the BDT discriminators for the different flavours for the CvsL (left) and for the CvsB (right) discrimination. Signal (charm quark) discriminator goes to 1 and background (light and bottom quark) discriminator goes to -1.

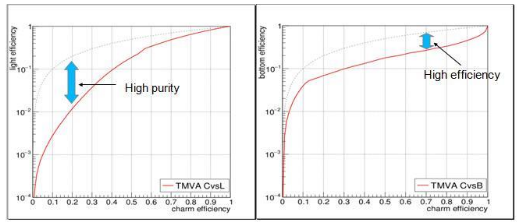

Figure 4 ROC curves showing the final performance of the CvsL (left) and CvsB (right).

Conclusions

For the first time c-tagger is available at the CMS. C-tagger will be one of the crucial elements for new physics search such as supersymmetric top. Also, it can be benefit for Standard Model precision measurement.

Acknowledgments

The author would like to thank to Department of Physics, Faculty of Science Chulalongkorn University and CUniverse research promotion project by Chulalongkorn University for financial support. I also would like to thank to CMS collaboration, in particular, our b-tagger and c-tagger colleagues and All SPC 2016 staffs for organizing this event. This research is supported by Rachadapisek Sompote Fund for Postdoctoral Fellowship, Chulalongkorn University.

References

[1] Halzen F. and Martin A. (1984). Quarks and leptons: an introductory course in modern Particle Physics, Wiley, New York USA.

[2] Dine M., Fischler W. and Srednicki M. (1981). Supersymmetric Technicolor. *Nucl. Phys. B, 189,* 575.

[3] Papucci M., Ruderman J.T. and Weiler A. (2012). Natural SUSY endures. *JHEP, 09*, 035.

[4] Sakai N. (1981). Naturalness in Supersymmetric Guts. *Z. Phys. C, 11,* 153.

[5] ATLAS Collaboration (2012). Search for a supersymmetric partner to the top quark in final states with jets and missing transverse momentum

at $\sqrt{s}$ =7 TeV with the ATLAS detector. *Rev. Lett., 109,* 211802.

[6] ATLAS Collaboration (2014). Search for pair-produced third-generation squarks decaying via charm quarks or in compressed supersymmetric scenarios in *pp* collisions at $s\sqrt{}$=8 TeV with the ATLAS detector. *Phys. Rev. D, 90,* 052008.

[7] Nakamura et al. (2010). Review of Particle Physics. *J. Phys. G, 37,* 075021.

[8] CMS Collaboration (2008). The CMS experiment at the CERN LHC. *JINST, 3,* S08004.