Eur. Phys. J. C manuscript No.
(will be inserted by the editor)

Boost 2013 Report Title

Report of BOOST2013, hosted by the University of Arizona, 12^{th}-16^{th} of August 2013.

D. Adams[1], A. Arce[2], L. Asquith[3], J. Backus Mayes[4], E. Bergeaas Kuutmann[5], J. Berger[6], D. Bjergaard[2], L. Bryngemark[7], A. Buckley[8], J. Butterworth[9], M. Cacciari[10], M. Campanelli[9], T. Carli[11], M. Chala[12], B. Chapleau[13], C. Chen[14], J.P. Chou[15], Th. Cornelissen[16], D. Curtin[17], M. Dasgupta[18], A. Davison[9], F. de Almeida Dias[19], A. de Cosa[20], A. de Roeck[11], C. Debenedetti[8], C. Doglioni[21], S. D. Ellis[22], F. Fassi[23], J. Ferrando[24], S. Fleischmann[16], M. Freytsis[25], M.L. Gonzalez Silva[26], S. Gonzalez de la Hoz[23], F. Guescini[21], Z. Han[27], A. Hook[4], A. Hornig[22], E. Izaguirre[4], M. Jankowiak[4], J. Juknevich[28], M. Kaci[23], D. Kar[24], G. Kasieczka[29], R. Kogler[30], A. Larkoski[4], P. Loch[31], D. Lopez Mateos[27], S. Marzani[32], L. Masetti[33], V. Mateu[34], D. W. Miller[35], K. Mishra[36], P. Nef[4], K. Nordstrom[24], E. Oliver Garcia[23], J. Penwell[37], J. Pilot[38], T. Plehn[29], S. Rappoccio[39], A. Rizzi[40], G. Rodrigo[23], A. Safonov[41], G. P. Salam[10,11], J. Salt[23], S. Schaetzel[29], M. Schioppa[42], A. Schmidt[29], J. Scholtz[22], A. Schwartzman[4], M. D. Schwartz[27], M. Segala[43], M. Son[44], G. Soyez[45], M. Spannowsky[32], I. Stewart[34], D. Strom[46], M. Swiatlowski[4], V. Sanchez Martinez[23], M. Takeuchi[29], J. Thaler[34], E. Thompson[1], N. V. Tran[36], C. Vermilion[25], M. Villaplana[23], M. Vos[23], J. Wacker[4], and J. Walsh[47]

[1] *Columbia University, Nevis Laboratory, Irvington, NY 10533, USA*
[2] *Duke University, Durham, NC 27708, USA*
[3] *Argonne National Laboratory, Lemont, IL 60439, USA*
[4] *SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*
[5] *Deutsches Elektronen-Synchrotron, DESY, D-15738 Zeuthen, Germany*
[6] *Cornell University, Ithaca, NY 14853, USA*
[7] *Lund University, Lund, SE 22100, Sweden*
[8] *University of Edinburgh, EH9 3JZ, UK*
[9] *University College London, WC1E 6BT, UK*
[10] *LPTHE, UPMC Univ. Paris 6 and CNRS UMR 7589, Paris, France*
[11] *CERN, CH-1211 Geneva 23, Switzerland*
[12] *CAFPE and U. of Granada, Granada, E-18071, Spain*
[13] *McGill University, Montreal, Quebec H3A 2T8, Canada*
[14] *Iowa State University, Ames, Iowa 50011, USA*
[15] *Rutgers University, Piscataway, NJ 08854, USA*
[16] *Bergische Universitaet Wuppertal, Wuppertal, D-42097, Germany*
[17] *YITP, Stony Brook University, Stony Brook, NY 11794-3840, USA*
[18] *University of Manchester, Manchester, M13 9PL, UK*
[19] *UNESP - Universidade Estadual Paulista, Sao Paulo, 01140-070, Brazil*
[20] *INFN and University of Naples, IT80216, Italy*
[21] *University of Geneva, CH-1211 Geneva 4, Switzerland*
[22] *University of Washington, Seattle, WA 98195, USA*
[23] *Instituto de Física Corpuscular, IFIC/CSIC-UVEG, E-46071 Valencia, Spain*
[24] *University of Glasgow, Glasgow, G12 8QQ, UK*
[25] *Berkeley National Laboratory, University of California, Berkeley, CA 94720, USA*
[26] *Universidad de Buenos Aires, AR-1428, Argentina*
[27] *Harvard University, Cambridge, MA 02138, USA*
[28] *Weizmann Institute, 76100 Rehovot, Israel*
[29] *Universitaet Hamburg, DE-22761, Germany*
[30] *Universitaet Heidelberg, DE-69117, Germany*
[31] *University of Arizona, Tucson, AZ 85719, USA*
[32] *IPPP, University of Durham, Durham, DH1 3LE, UK*
[33] *Universitaet Mainz, DE 55099, Germany*
[34] *MIT, Cambridge, MA 02139, USA*
[35] *University of Chicago, IL 60637, USA*
[36] *Fermi National Accelerator Laboratory, Batavia, IL 60510, USA*
[37] *Indiana University, Bloomington, IN 47405, USA*
[38] *University of California, Davis, CA 95616, USA*
[39] *Johns Hopkins University, Baltimore, MD 21218, USA*
[40] *INFN and University of Pisa, Pisa, IT-56127, Italy*
[41] *Texas A & M University, College Station, TX 77843, USA*
[42] *INFN and University of Calabria, Rende, IT-87036, Italy*
[43] *Brown University, Richmond, RI 02912, USA*
[44] *Yale University, New Haven, CT 06511, USA*
[45] *CEA Saclay, Gif-sur-Yvette, FR-91191, France*
[46] *University of Illinois, Chicago, IL 60607, USA*
[47] *University of California, Berkeley, CA 94720, USA*

Abstract Abstract for BOOST2013 report

Keywords boosted objects · jet substructure · beyond-the-Standard-Model physics searches · Large Hadron Collider

1 Introduction

Jet substructure has been around a while now, and it's time to study the correlations between the plethora of observables that have been developed and used. Previous BOOST reports [1,2,3] studied some of these things.

2 Monte Carlo Samples

Give details about how the samples we use have been generated.

3 Jet Algorithms and Grooming Approaches

Describe the jet algorithms and grooming approaches that we will use in the report. Give the nomenclature that we will use to refer to e.g. the groomed mass in the rest of the report.

4 Substructure Variables/Taggers

Describe the specific substructure variables and tagging approaches that we will be using in this report e.g. n-subjettiness, Q-jets, HTT, JH tagger. Give the nomenclature that we will use to refer to these variables/taggers in the rest of the report.

5 Boosted W-Tagging

In this section we study the performance of various jet algorithms in combination with jet substructure variables/taggers in terms of the identification of a boosted hadronically decaying W signal. For each jet algorithm we produce Receiver Operating Characteristic (ROC) curves that elucidate the performance of various variables that are capable of providing discrimination between a hadronic W signal and a QCD jet. These variables are then combined in a Boosted Decision Tree (BDT) and the performance of the resulting BDT discriminant explored through ROC curves to understand the degree to which variables are correlated and exploiting the same information. These studies are repeated in different kinematic regimes, to explore both the performance and correlations as a function of the jet boost, and where substructure approaches may break down.

5.1 Methodology

These studies use the $X \to WW$ samples as signal and the XXX samples to model the QCD background.

Jets are reconstructed using the XXX jet algorithms described in the previous section. The following event selection is then applied to these samples....(presumably this will vary depending on which kinematic bin is used, as will the actual samples used - maybe summarize in a table).

Figure 1 shows background versus signal in some basic kinematic distributions. *Do we want to reweight signal kinematics to background or vice versa? Do we want to study quarks/gluons separately?*

Go on to explain how we produce the ROC curves.

5.2 Performance at Moderate Boosts

(this section is to cover the W-tagging performance for jet p_{T} 200-300 GeV and 500-600 GeV using $\sqrt{s} = 8$ TeV samples)

5.2.1 Single Variable Performance

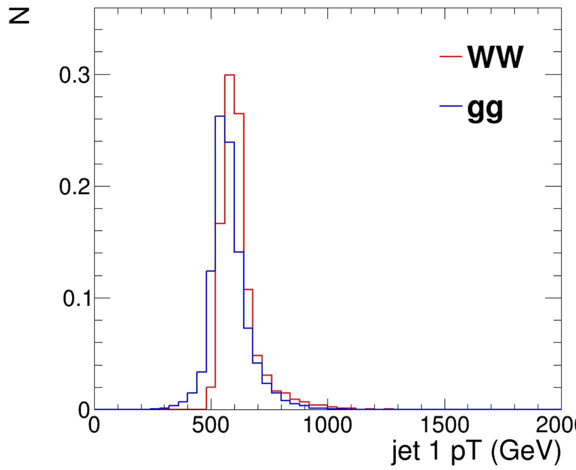
Show plots of signal versus background for all single variables investigated.

Figure 2 the compares signal and background in the mass distributions for the different groomers, and Figure 3 in the different substructure variables.

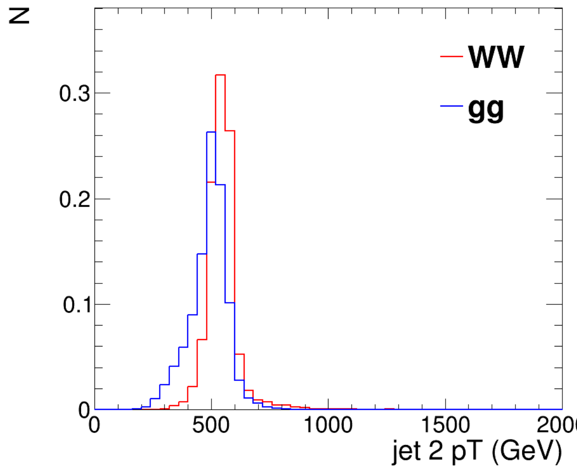
Figure 4 shows the single variable ROC curves in the p_{T} 500 GeV bin for the anti-k_{T} R=0.8 algorithm, compared to the ROC curve for a BDT combination of all the variables. One can see that the best performant single variables for a reasonable signal efficiency are the groomed/filtered masses, which all have a similar level of performance with the exception of the soft drop mass with $\beta = -1$. *Would be good to split this into two plots, one using the masses and one for other variables, or somehow make the mass and other variable curves more distinct from one another by using same colour for all the mass curves.*

We want to look also at:

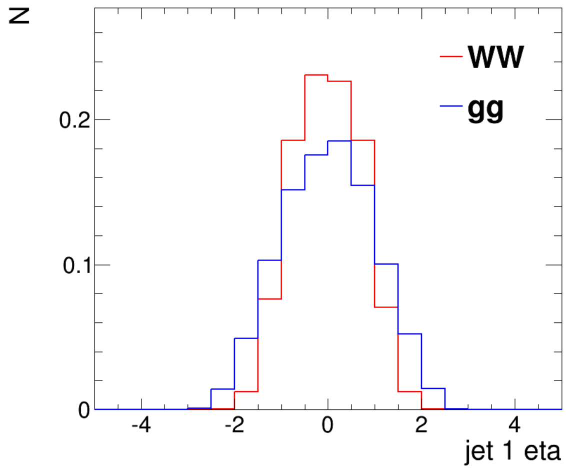
- *Dependence on R. So have the same single variable ROC for e.g. R=1.2, R=0.4. Then possibly have another plot which compares the best single variable (e.g. groomed mass) for different R.*



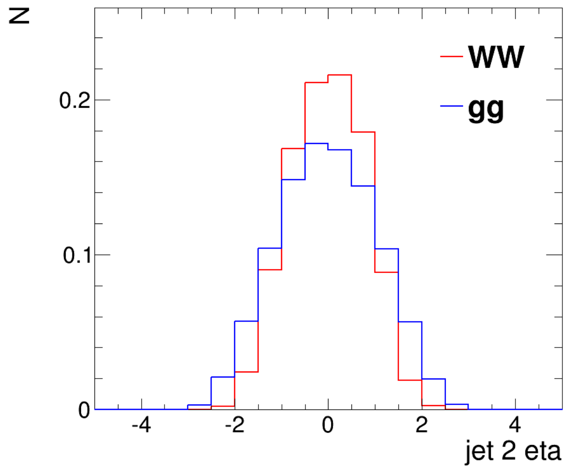
(a) Leading jet p_{T}



(b) Sub-leading jet p_{T}



(c) Leading jet η



(d) Sub-leading jet η

Fig. 1 Comparisons of the QCD background to the WW signal in the p_T 500 GeV bin using the anti-k_{T} R=0.8 algorithm: basic kinematic distributons.

– *Dependence on pT. Again want to repeat the plot for different kinematic bins, and then have a plot which compares the best performance in each kinematic bin to see the dependence of performance on kinematics.*

5.2.2 Combined Performance

Figure 5 shows the BDT combinations of each mass variable with every other variable considered in the p_T 500 GeV bin using the anti-k_{T} R=0.8 algorithm. *Can we drop the combinations of mass + mass from these plots to make them clearer? Also would be good to put the single variable mass curve on these plots, so you can see how much improvement the combination gives, and the "all variables" curve.*

No combination with other variables can recover the poor performance of the ungroomed mass and the soft drop mass with $\beta = -1$. The other groomed/filtered masses are all most improved by combination with the $C_2^{\beta=1}$ energy correlation function. *Show 2-D correlation plot of $C_2^{\beta=1}$ vs groomed mass - show that it is largely uncorrelated. Now show a plot which compares on one plot the best combined performance for each mass + X. e.g. mass + $C_2^{\beta=1}$, and compared also to the all variables curve. This plot is just for one R and one kinematic bin.*

Repeat these studies for different R and different kinematic bins. Finally make plots which compare best combined performance for different R and kinematics.

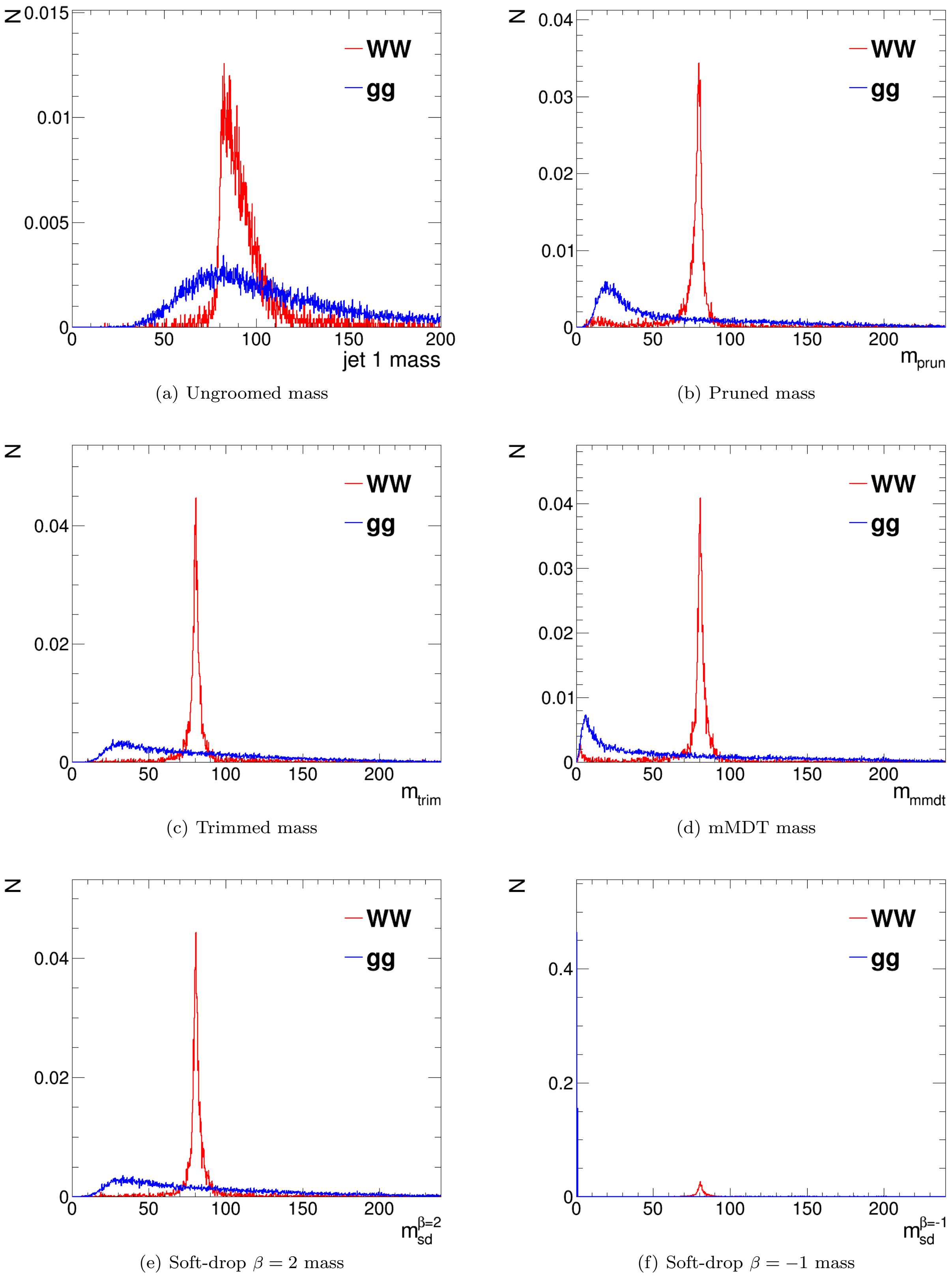
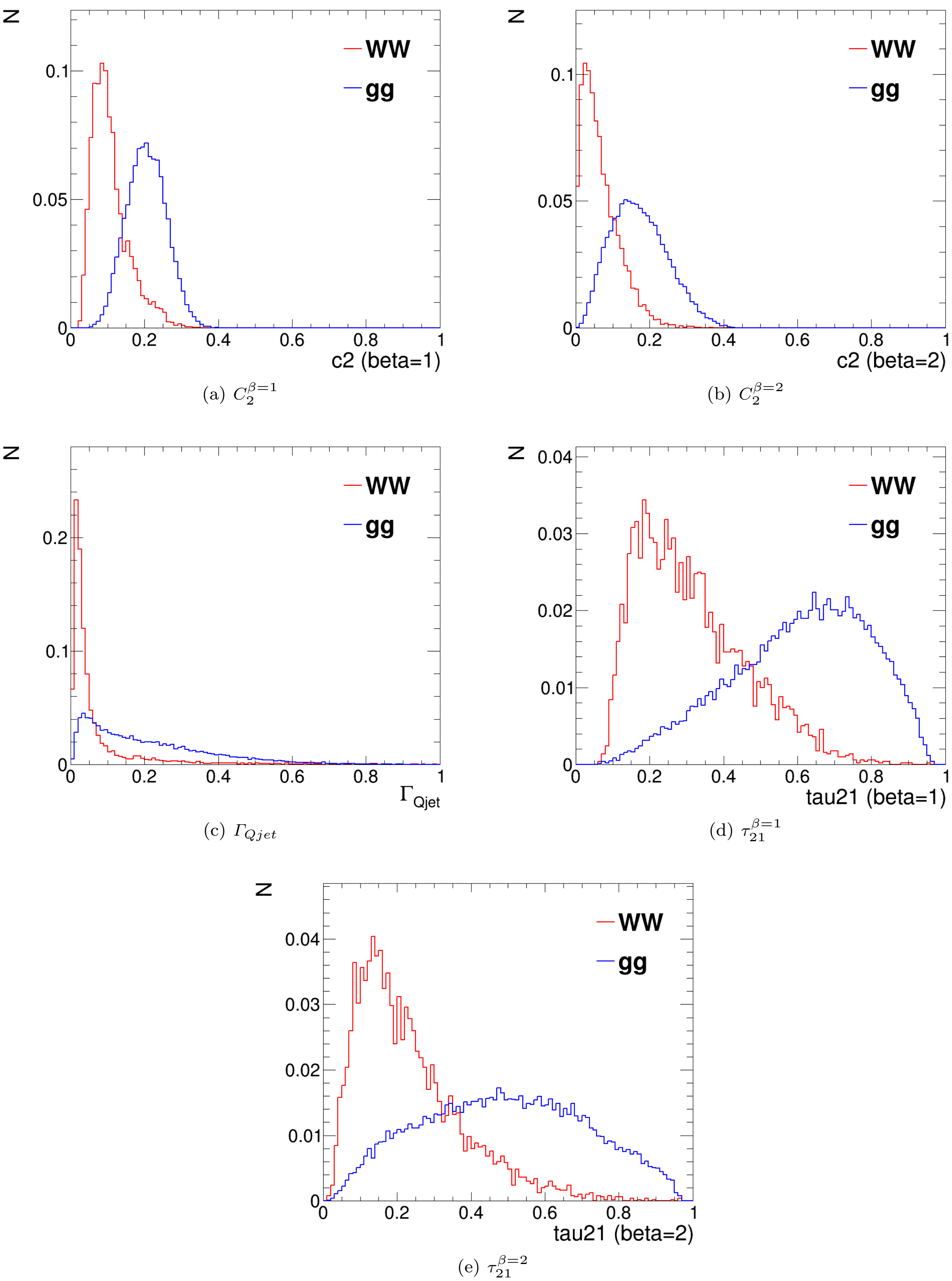


Fig. 2 Comparisons of the QCD background to the WW signal in the p_T 500 GeV bin using the anti-k_{T} R=0.8 algorithm: leading jet mass distributions.



(a) $C_2^{\beta=1}$

(b) $C_2^{\beta=2}$

(c) Γ_{Qjet}

(d) $\tau_{21}^{\beta=1}$

(e) $\tau_{21}^{\beta=2}$

Fig. 3 Comparisons of the QCD background to the WW signal in the p_T 500 GeV bin using the anti-k_{T} R=0.8 algorithm: substructure variables.

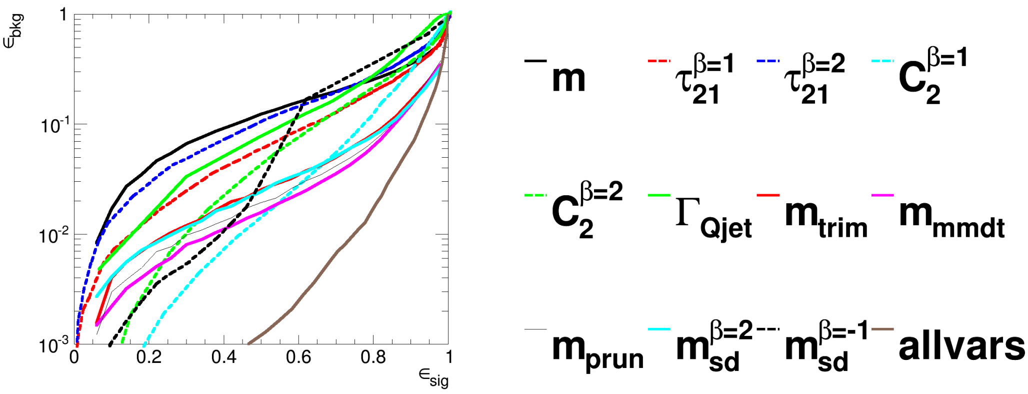
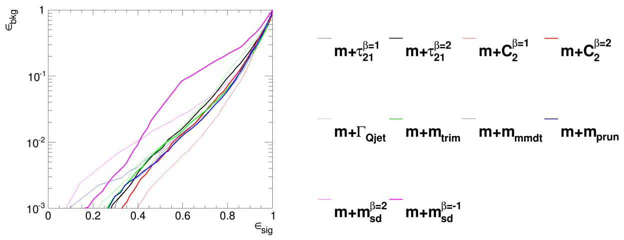
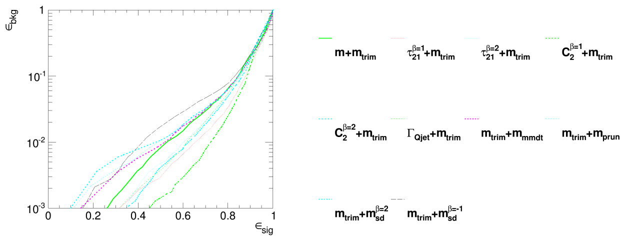


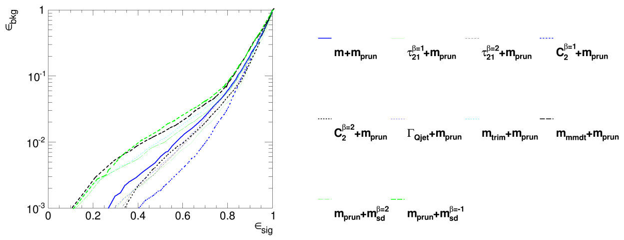
Fig. 4 The ROC curve for all single variables considered for W tagging in the p_T 500 GeV bin using the anti-k_{T} R=0.8 algorithm.



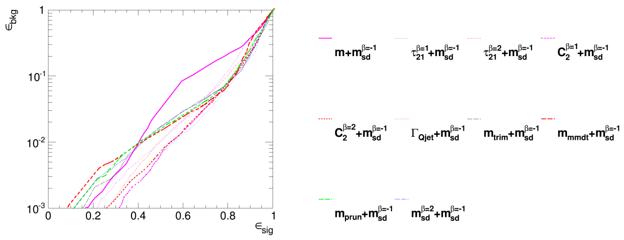
(a) Ungroomed mass + X



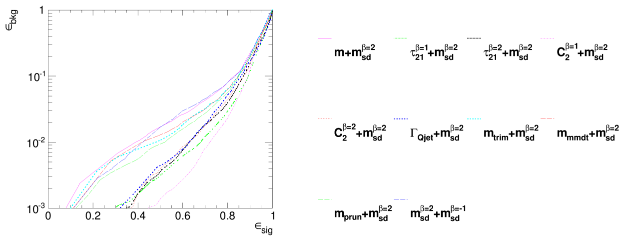
(b) Trimmed mass + X



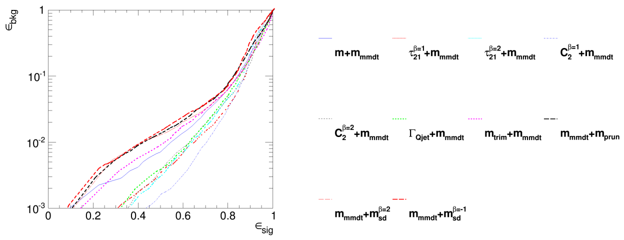
(c) Pruned mass + X



(d) Soft drop mass ($\beta = -1$) +X

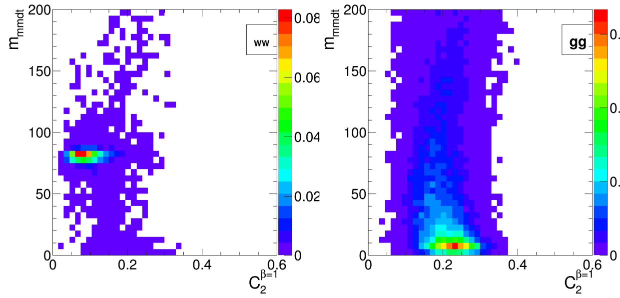


(e) Soft drop mass ($\beta = 2$) + X

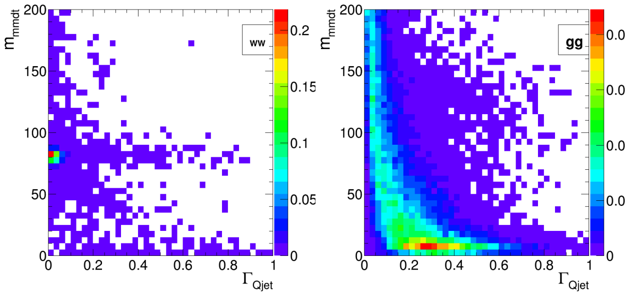


(f) mMDT mass + X

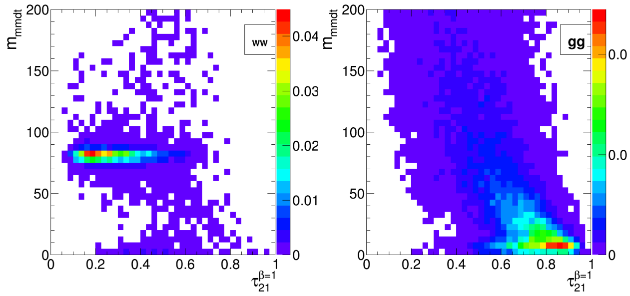
Fig. 5 The BDT combinations of each mass variable with every other variable considered in the p_T 500 GeV bin using the anti-k_{T} R=0.8 algorithm.



(a) mMDT mass vs $C_2^{\beta=1}$



(b) mMDT mass vs Γ_{Qjet}



(c) mMDT mass vs $\tau_{21}^{\beta=1}$

Fig. 6 2-D plots showing the correlation between mMDT mass and various substructure vairables in the p_T 500 GeV bin using the anti-k_T R=0.8 algorithm.

Do we want to look at other combinations of variables which don't involve mass? Practically I think we will always be making mass + X though.

5.3 Performance at High Boosts

(this section is to cover the W-tagging performance for jet p_T 1-1.1 TeV and > 1.5 TeV using $\sqrt{s} = 14$ TeV samples)

Maybe we don't need to divide into different medium/high boost sections.

6 Top Tagging

Top tagging studies go here.

7 Quark-Gluon Tagging

q/g tagging studies go here.

8 Summary & Conclusions

This report discussed the correlations between observables and looked forward to jet substructure at Run II of the LHC at 14 TeV center-of-mass collisions eneergies.

Acknowledgements

We thank the Department of Physics at the University of Arizona and for hosting the conference at the Little America Hotel. We also thank Harvard University for hosting the event samples used in this report. We also thank Hallie Bolonkin for the BOOST2013 poster design and Jackson Boelts' ART465 class (fall 2012) at the University of Arizona School of Arts VisCom program. (NEED TO ASK PETER LOCH FOR MORE ACKNOWLEDGEMENTS)

References

1. A. Abdesselam, E. B. Kuutmann, U. Bitenc, G. Brooijmans, J. Butterworth, et al., *Boosted objects: A Probe of beyond the Standard Model physics*, *Eur.Phys.J.* **C71** (2011) 1661, [arXiv:1012.5412].
2. A. Altheimer, S. Arora, L. Asquith, G. Brooijmans, J. Butterworth, et al., *Jet Substructure at the Tevatron and LHC: New results, new tools, new benchmarks*, *J.Phys.* **G39** (2012) 063001, [arXiv:1201.0008].
3. A. Altheimer, A. Arce, L. Asquith, J. Backus Mayes, E. Bergeaas Kuutmann, et al., *Boosted objects and jet substructure at the LHC*, arXiv:1311.2708.