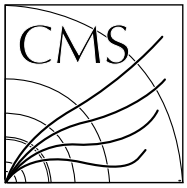
Available on CMS information server

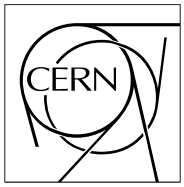
CMS NOTE 2006/059



The Compact Muon Solenoid Experiment

CMS Note

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland



1 May 2006

Optimization of Jet Reconstruction Settings and Parton-Level Corrections for the $t\bar{t}H$ Channel

A. Santocchia

Università di Perugia e Sezione dell'INFN, Perugia, Italy

Abstract

The typical final state for $t\bar{t}H$ associated production where the Higgs decays into a pair of b quark is an event with high jet multiplicity. The expected jet number for one (two) semi-leptonic top decay or fully hadronic top decay is 6(4) and 8 respectively. The choice of the jet algorithm is therefore crucial to improve the chance to detect a light Higgs boson in the top-associated H production. Previous studies both with fast and full simulated events show that different jet algorithms give different results in the selection procedure. To fully investigate the effect of different jet algorithm however is mandatory to use jets calibrated for detector effects and corrected for effects related to the parton-jet formation. Different jet finders and Monte-Carlo jet calibration parameters are studied relatively to the best performance for the $t\bar{t}H$ channel observability. The set of algorithm chosen is the iterative cone algorithm with cone size going from 0.30 up to 0.50 (with 0.05 step) and the inclusive $K_\perp$ algorithm with r=0.4. Different calibration parameters as a function of η, E_T and jet flavour are calculated. Finally, as an example, the jet-to-parton pairing efficiency for the fully hadronic $t\bar{t}H$ decay is computed.

1 Introduction

The aim of this note is to describe the procedure adopted for the jet calibration in the $t\bar{t}H$ analysis. The typical multi-jet final state for this channel suggests a careful choice of the jet reconstruction algorithm. The best approach is to try to optimize the algorithm parameters studying the effect of different algorithm values during the selection procedure of the signal events. A basic condition for implementing this optimization is the availability of a jet calibration procedure which allows for the correction of the raw measured jet energy.

Two different effect have to be taken into account: the effect due to the detector (particle-level correction) and to the fragmentation (parton-level correction). The first one is normally provided by the CMS calorimeters group while the latter is more channel-dependent and it's up to the analysis groups to use the common one or to develop ad-hoc corrections for their specific channels. Particle-level corrections available in CMS were limited to Iterative Cone Algorithm[1] with ΔR=0.5 and 0.7 and to the inclusive $K_\perp$ algorithm[2] with r=1; only recently also MidPoint[3] corrections have been added. These options however are not the optimal choice for a final state with high multiplicity jets. Past study for the $t\bar{t}H$ channel[4] have shown that a smaller cone size and the $K_\perp$ algorithm with r=0.4 could give better results. For the Physics Technical Design Report (PTDR)[5] both particle-level and parton-level correction have been computed for the requested jet algorithms.

The major objective for this study is to parametrize the calibration parameter as a function of the direction, energy and flavour of the jets for different jet algorithms. The adopted procedure can be summarized in the following steps:

- Detector Effect - (Particle-Level Correction to MonteCarlo Jets)
  1. Raw jets are reconstructed using the chosen algorithm from full simulated and reconstructed events
  2. MonteCarlo jets are reconstructed using generator level particles as input to the same jet algorithm
  3. MonteCarlo jets are b-tagged looking at the flavour of the particles which formed the jet
  4. Raw and MonteCarlo Jets are paired minimizing the ΔR distance
  5. All the paired jet with $\Delta R < 0.30$ are used to build a set of histograms mapping the $\eta - E_T$ plane with the E_T^{raw}/E_T^{MC} distribution separately for b-tagged and non-b-tagged Monte Carlo jet
  6. The Gaussian fit of this set of histograms is used to obtain the first calibration functions.
- Fragmentation (Parton-Level Correction)
  1. MonteCarlo Jets and partons are paired minimizing the ΔR distance
  2. The same mapping in the $\eta - E_T$ plane is used to build a similar set of histograms with the E_T^{MC}/E_T^{parton} distribution using a stronger cut in the pairing ($\Delta R < 0.15$)
  3. The Gaussian fit is again used to obtain the second calibration functions
- The combination of the first and second effect gives the final set of calibration functions

In the Particle-Level correction, different detector effects on b-jet and non-b-jet have to be evaluated. The major motivation is due to the higher probability to have μ-leptons and neutrinos from semi-leptonic decays of b hadrons in the b-jets with respect to the light-quark and gluon jets. The μ energy contribution is not correctly measured using the calorimeters and only more sophisticated techniques (like Energy Flow methods where muon chambers and calorimeter measurements are combined) could correct for this energy loss while the neutrinos contribution is completely lost.

The achievable precision for the described procedure depends mainly on two factor: the goodness of the detector simulation, which will improve during the data taking in the first few years of LHC running, and the precise knowledge of the theoretical fragmentation model. For the latter, the very first months of data taking will be crucial because of the tuning of the generators on real data. The actual choice adopted is just a extrapolation of the available data at lower energy scale. This model uncertainty is also affected by the poor knowledge we have of the underlying event contribution. All of these aspects are expected to be better tuned after the start of the LHC run and all the calibration functions will be re-calculated using real data.

The set of jet algorithm chosen in this study is the following: Iterative Cone Algorithm with ΔR=0.30, 0.35, ... 0.50 and inclusive $K_\perp$ algorithm with $r = 0.4$.

2 Data Sample and Results

The jet response has been studied using the same data sample produced for the $t\bar{t}H$ channel analysis. These samples have been generated with PYTHIA[6] and COMPHEP[7], simulated in the CMS detector by CMSIM[8] (version 133) and digitized by ORCA[9] (version 8.7.4). A total of 1.6 million events have been used from the ttH, ttbb, ttjj samples (COMPHEP) and di-jet in the $\hat{p}_T$ ranges from 120 to 170 GeV/c and greater than 170 GeV/c samples (PHYTIA)[10]. Raw jets have been reconstructed using a threshold $E_T > 0.5$ GeV and $E > 0.8$ GeV in the calorimeters to suppress the instrumental noise contribution. A raw jet was accepted if $E_T^{raw} > 5$ GeV; the E recombination scheme has been adopted together with a 0 GeV seed for cone algorithm[11].

MonteCarlo Jets were reconstructed using all generator-level stable particles (including muons and neutrinos) with $E_T > 0.5$ GeV. A jet is kept if $E_T^{MC} > 10$ GeV. To b-tag the MC jet, the charged particles energy has been used following this scheme:

- The set of generator level, stable particle forming the jet are scanned for particles decaying from b-flavoured hadrons;
- The energy is summed for particles decaying from b-flavoured hadrons and the ratio to the jet energy is calculated;
- If the ratio is higher than 0.1 the jet is classified as a b-jet

The same algorithm has been used to pair raw to MC jet and MC jet to parton; the idea is to minimize the sum of the ΔR for each possible pair (raw-MC jet or MC jet-parton). All the pairs with $\Delta R < 0.30(0.15)$ have been used for the transverse energy ratio distribution. The $|\eta|$-plane has been divided in 25 bin of 0.1 size covering the range $|\eta| < 2.5$ while the transverse energy has been mapped up to 600 GeV with 200 bin, 3 GeV size. In fig 1 the E_T^{raw}/E_T^{MC} ratio distribution for a typical bin is shown (0.5$< |\eta| <$0.6 and 45 GeV$< E_T^{raw} <$48 GeV). Figure 2 shows the same distribution for the E_T^{MC}/E_T^{parton}. The fit has been done in 2 steps: first the whole histogram interval has been used and then the fit range was changed to $[mean_{fit} - 2.5 \times \sigma_{fit}, mean_{fit} + 2.5 \times \sigma_{fit}]$ where $mean_{fit}$ and σ_{fit} are the values obtained from the first fit.

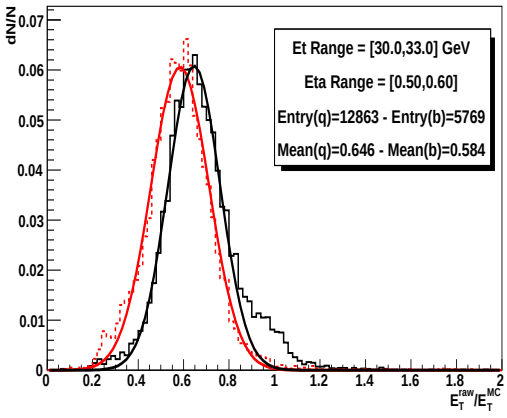


Figure 1: E_T^{raw}/E_T^{MC} Ratio Distribution for b-jets (dashed line) and non-b-jets (solid line)

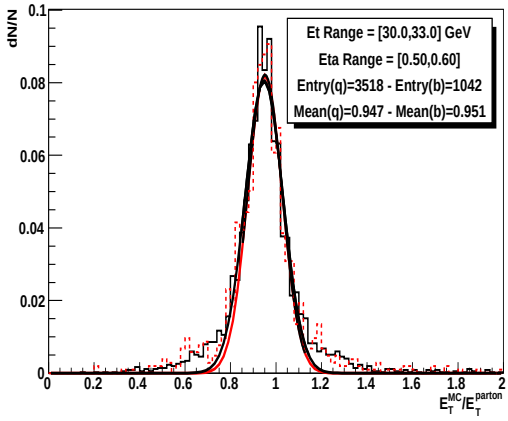


Figure 2: E_T^{MC}/E_T^{parton} Ratio Distribution for b-jets (dashed line) and non-b-jets (solid line)

Figures 3 and 4 show respectively the mean value (from the second fit) of the E_T^{raw}/E_T^{MC} and E_T^{MC}/E_T^{parton} for the same η-ring ($0.5 < |\eta| < 0.6$) as a function of the E_T^{raw} and E_T^{MC} respectively. Error bars on the mean values have been defined as $mean_{2nd\_fit}/\sqrt{N}$ where N is the number of entries in each histogram. Figure 3 show clearly the different behavior of b-jet and non-b-jet mainly at low E_T. In figure 4 this difference is negligible and the need to separate the two contributions is lost. The major effect is due to the different charged component of b-jet and non-b-jet: the magnetic field bending tends to sweep out the lower p_T tracks losing the energy contribution to the jets. The lost energy is higher for b-jets where the probability to have low p_T tracks is higher with respect to non-b-jet. Another effect is due to muon leptons within the jet which leave only a MIP signal in the calorimeter.

This effect is higher for b-jets because of the higher probability to have μ-leptons from b semileptonic decays. The distinction of b-jets from non-b-jets for the Parton-Level correction is not necessary and all jets are used together.

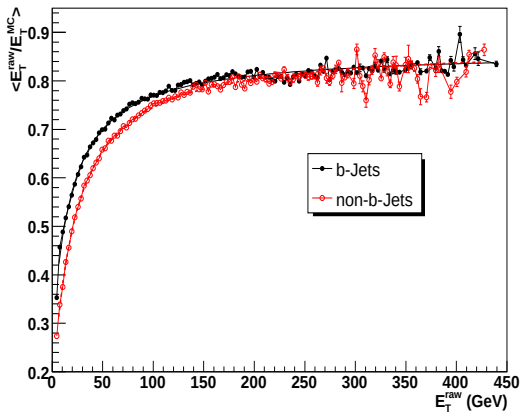


Figure 3: E_T^{raw}/E_T^{MC} Ratio for b-jets and non-b-jets as a function of E_T^{raw}

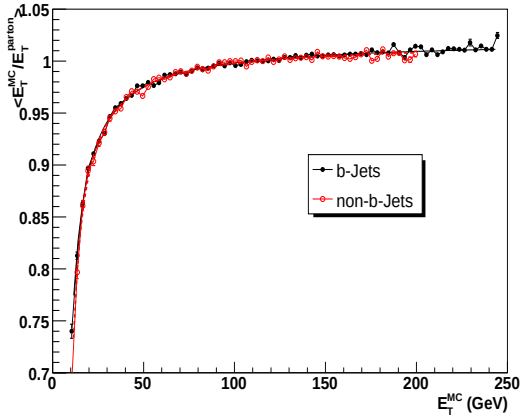


Figure 4: E_T^{MC}/E_T^{Parton} Ratio for b-jets and non-b-jets as a function of E_T^{MC}

3 Calibration Function

The same procedure has been followed for each η-bin and each plot is fitted with the following function:

$$\frac{E_T^{raw}}{E_T^{MC}} = \frac{1}{(a + bE_T^{raw})} + c \tag{1}$$

All the plots shown refers to $\Delta R = 0.5$ cone jets. Similar plots exist for the whole set of studied jet algorithm. The same fitting function has been used for the Parton-Level correction expressing E_T^{MC}/E_T^{parton} as a function of E_T^{MC}.

3.1 Particle-Level Correction

In figure 5 the three fitted parameters as a function of η are shown for the Particle-Level correction. The big change around $|\eta| = 1.5$ is due to the different material budget in this region where the tracker barrel-endcaps border is located.

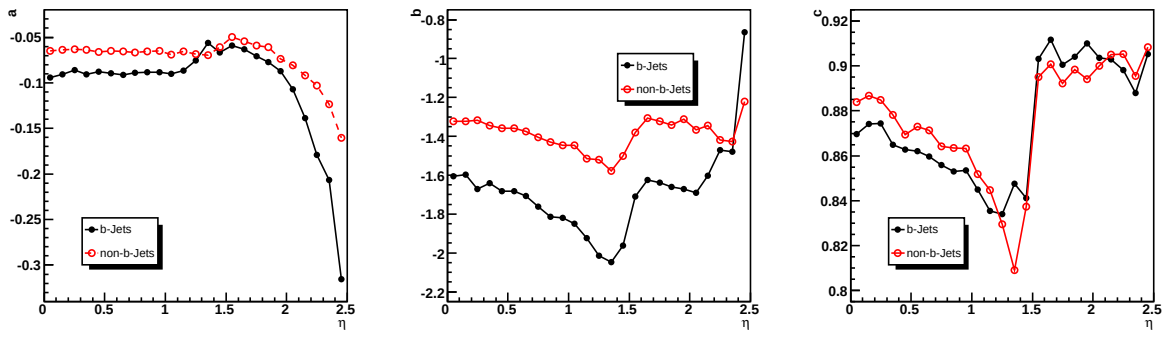


Figure 5: Raw-MC Fitted parameters (a, b and c from eq.1) for b-jets (dashed line) and non-b-jets (solid line) as a function of η

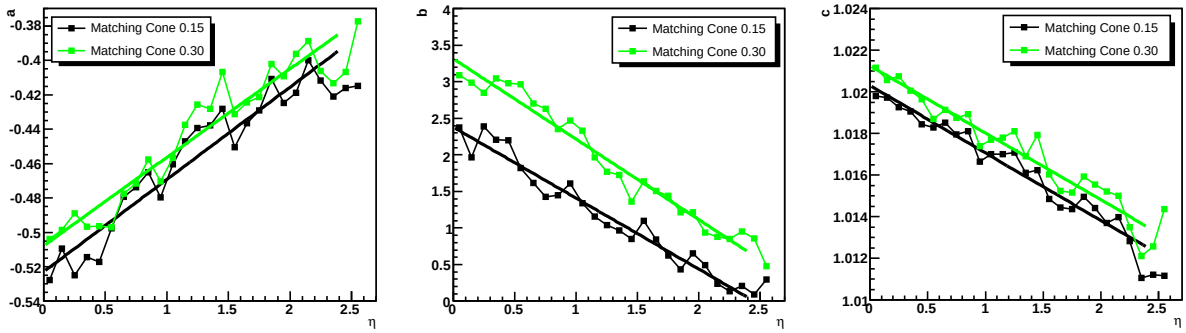


Figure 6: MC-Parton Fitted parameters (a, b and c from eq.1) for a matching cone 0.15 (solid line) and 0.30 (dashed line) as a function of η

3.2 MC jets-to-Parton Correction (Parton-Level)

The same three fitted parameters as a function of E_T^{MC} are shown in fig 6. These parameters reflect the physics model for parton showering and fragmentation function chosen and will have to be tuned with LHC real data. For this particular choice there is a linear relation in the central region and the three parameters have been fitted with a straight line.

To cross check the stability of the correction algorithm at parton level, two matching cone ($\Delta R = 0.15$ and 0.30) were used. A small shift is observed between the 2 fitted lines, the overall effect is below 0.5% in the central region for the whole E_T spectrum. For $|\eta| = 1.4$, only low E_T jets are affected (2% shift for $E_T = 20$ GeV) while for $|\eta| = 2.7$ the effect is higher (order of 40% for E_T =20 GeV decreasing to 15% and 7% respectively for 50 and 100 GeV jet E_T).

4 Comparison of ttH-based Calibration to the CMS di-jet-based Calibration

An independent data sample is used to compare the ttH method to a calibration[11] based on di-jet events, called in the following standard calibration. QCD di-jet events with $\hat{p}_T$ up to 600 GeV and $\Delta R = 0.5$ cone size have been used for this purpose. In figure 7 the mean E_T^{ttH} to E_T^{STD} ratio distributions for 3 η-bin ($0.0 < |\eta| < 0.1$, $1.2 < |\eta| < 1.3$ and $2.4 < |\eta| < 2.5$) as a function of E_T^{STD} are shown. In the tracker acceptance region the difference between standard and ttH calibration is within 5% for $E_T > 40$ GeV and non-b-jet while the ttH calibration gives always harder jets for lower transverse energy. The ttH calibration for b-jets produces as expected higher transverse energy. The different behavior at low E_T is due mainly to the different Parton-Level calibration applied: the standard calibration use a Parton-Level correction extracted by light-quark jets events.

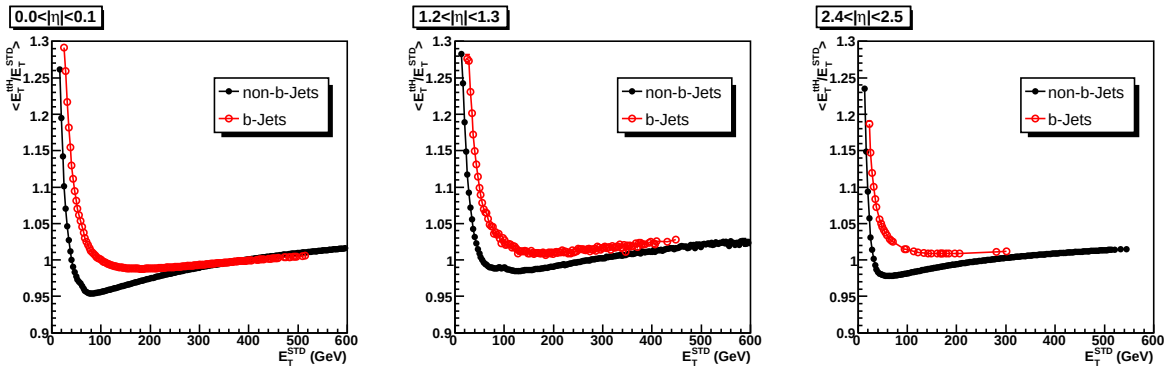


Figure 7: Mean ratio of ttH and standard calibration for different η bins as a function of standard calibration transverse energy (E_T^{STD})

5 Application to the $t\bar{t}H$ fully hadronic channel

The six different calibrations have been applied to the $t\bar{t}H$ fully hadronic decay channel, with a generated Higgs and Top masses of 120 and 175 GeV respectively, as a bench mark. The standard calibration to the iterative cone algorithm with $\Delta R = 0.5$ has been also considered as a reference using 10 GeV as minimum jet calibrated transverse energy. Signal events have been reconstructed and all the different jet algorithm have been applied together with the proper calibration functions. The eight most energetic jets in the tracker acceptance region have been paired to the eight partons in the $t\bar{t}H$ final state using generator information minimizing the ΔR(jet-parton). An event is then selected if all the jets have been paired to the parton with $\Delta R < 0.3$. Finally invariant masses for W bosons, t quarks and H boson have been build to compare the effect of different algorithms and calibrations. Figure 8 show the transverse energy distribution for the eight most energetic jets. No difference is visible up the 3th jet, then the standard calibration gives always a lower E_T as a consequence of the different treatment of muons and neutrinos within the jets and the separate calibration function for b and non-b jets. All the six chosen jet algorithm for $t\bar{t}H$ calibration give similar distribution and no evident differences are present. Figure 9 show the invariant mass for W, t-quark and H obtained with standard calibration and the ttH-calibration for a iterative cone algorithm and $\Delta R = 0.5$.

The results for all the jet algorithms are summarized in table 1. Invariant masses and widths of the fitted Gaussian peaks are compatible with the standard calibration. The ttH-calibration gives higher values for the masses peak, especially for the t and H particles. This is due to the different calibration functions used for b and non-b-jets. Mass resolutions together with jet-pairing efficiency could give a more clear idea of the different algorithm performance in a complex multi-jet events as the one used for this comparison.

	STD	ICA 0.30	ICA 0.35	ICA 0.40	ICA 0.45	ICA 0.50	K_T^{inc} $r = 0.4$
M_W	81.5	81.6	81.6	81.8	82.1	82.8	82.7
M_t	172.1	172.9	173.1	173.7	174.7	176.7	176.2
M_H	105.5	108.9	109.3	109.9	110.7	111.2	112.2
$\sigma(W)$	13.8	13.0	13.2	13.2	13.5	13.5	13.6
$\sigma(t)$	22.2	21.1	21.0	21.0	21.3	21.2	20.6
σ_H	19.3	18.8	19.1	19.0	19.1	19.3	18.6
Res(W)	0.170	0.159	0.162	0.162	0.164	0.164	0.152
Res(t)	0.129	0.122	0.121	0.121	0.122	0.120	0.117
Res(H)	0.183	0.173	0.174	0.173	0.173	0.174	0.166
Pair Eff.(%)	4.1	5.9	6.4	6.0	5.2	4.3	5.4

Table 1: Invariant masses for different jet algorithm: STD is standard calibration for 0.5 cone size; ICA is Iterative Cone Algorithm with ttH Calibration; K_T^{inc} is ttH calibration (Generated Higgs Mass is 120 GeV)

6 Conclusions

Detection of a light Higgs boson in the $t\bar{t}H$ associated production is a challenging analysis which require a carefull optimization of the signal selection procedure. This note describes the ad-hoc calibration procedure developed for the study of $t\bar{t}H$ channel with respect to different jet reconstruction algorithm. Particle-Level and Parton-Level corrections have been calculated as a function of η, E_T and jet flavour for Iterative Cone (ΔR = 0.30, 0.35 ... 0.50) and inclusive $K_\perp$ algorithm with r=0.4. Comparison to the standard calibration functions has been made and no major differences are present for E_T >40 GeV. The difference at lower transverse energy is due to the different parton-level correction used which has been tuned to the $t\bar{t}H$ channel. The results indicate that a cone size around 0.40 is more promising for events with high multi-jet topology in the final state. The particle-level correction developed for this study is however not final because of the missing end-caps region where the material-budget and the complicated geometry of the CMS calorimeters are more difficult to study.

7 Acknowledgments

I wish to thank Olga Kodolova and Chris Tully for fruitful discussions and valuable inputs to this note.

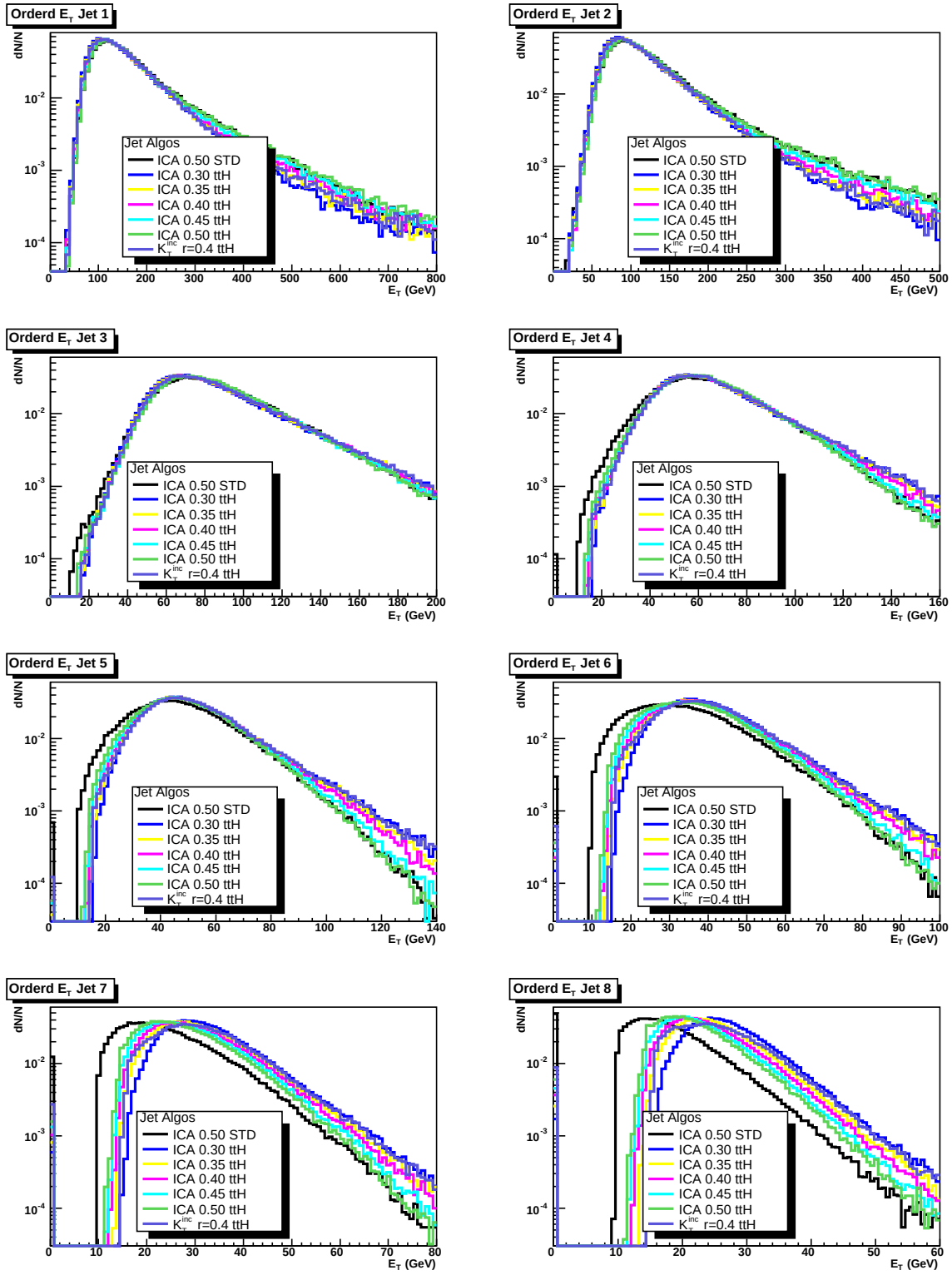


Figure 8: Ordered E_T Distribution for the $t\bar{t}H$ fully hadronic decay channel

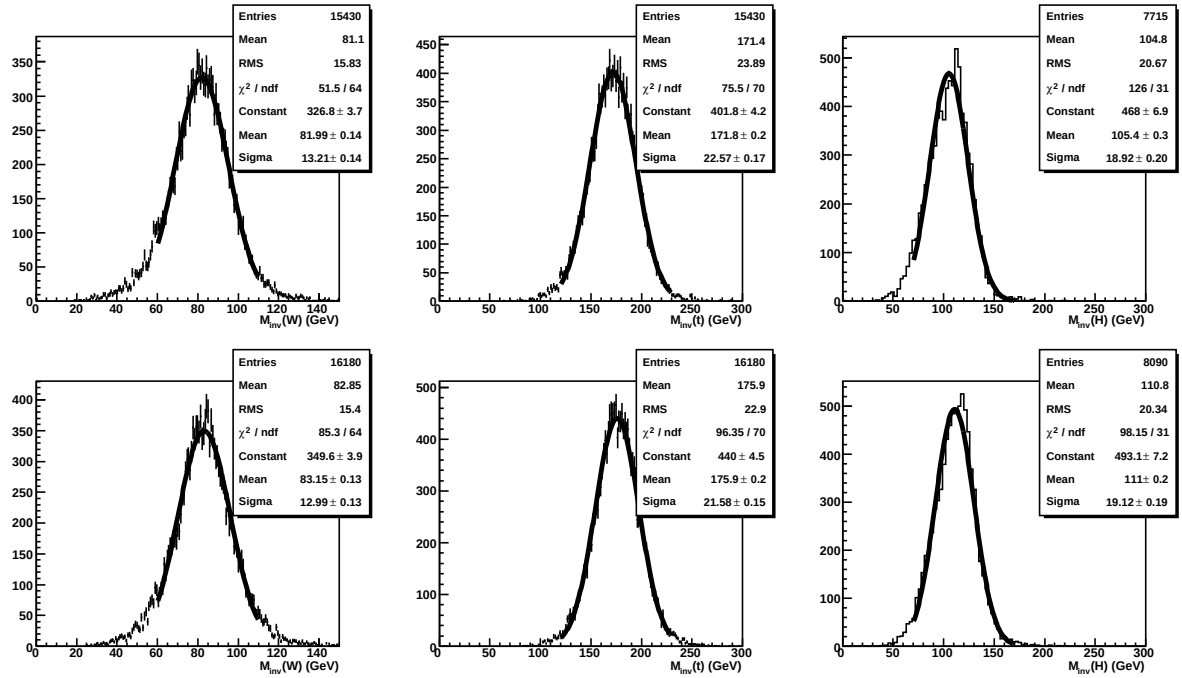


Figure 9: Invariant masses for W, t and H for the standard calibration (up) and tth-calibration (down) for cone algorithm and $\Delta R = 0.5$ (Generated Higgs Mass is 120 GeV)

References

[1] J.E. Huth et al., "Toward a standardization of jet definitions", Published in Snowmass Summer Study 1990:134-136

[2] J.M. Butterworth, J.P. Couchman, B.E. Cox and B.M. Waugh "KtJet : A C++ implementation of the $K_\perp$ clustering algorithm", Comp. Phys. Comm. vol 153/1 85-96 (2003)

[3] M.H. Seymour, "Jet shapes in hadron collisions: Higher orders, resummation and hadronization", Nucl. Phys. B513, 269 (1998)

[4] D. Benedetti, L. Fanò "Study of High Level b-trigger Selection of ttH Fully Hadronic Decays", CMS NOTE-2002/044

[5] CMS Coll., Physics TDR Vol.II, in preparation

[6] T. Sjostrand, L. Lonblad and S. Mrenna, "PYTHIA 6.2:Physics and manual", hep-ph/0108264

[7] A. Pukhov, E. Boos, M. Dubinin, V. Edneral, V. Ilyin, D. Kovalenko, A. Kryukov, V. Savrin, S. Shichanin, and A. Semenov, "CompHEP - a package for evaluation of Feynman diagrams and integration over multi-particle phase space. User's manual for version 33", hep-ph/9908288

[8] CMSIM home page: http://cmsdoc.cern.ch/cmsim/cmsim.html

[9] CMS Coll., "Object-oriented Reconstruction for CMS Analysis", Physics TDR Vol.I, Chapter 2, ORCA home page: http://cmsdoc.cern.ch/ORCA/

[10] D. Benedetti, S. Cucciarelli, C. Hill, J. Incandela, S.A. Koay, C. Riccardi, A. Santocchia, A. Schmidt, P. Torre, C. Weiser, "Search for $H \to b\bar{b}$ in association with a $t\bar{t}$ pair at CMS", CMS Note in preparation

[11] A. Heister, O. Kodolova, V. Konoplianikov, S. Petrushanko, J. Rohlf, C. Tully, and A. Ulyanov "Measurement of Jets with the CMS Detector at the LHC" CMS Note 2006/036