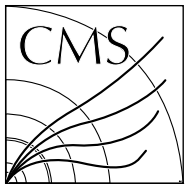
Available on CMS information server

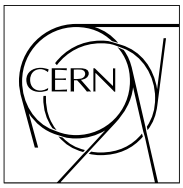
**CMS AN -2007/029**



**The Compact Muon Solenoid Experiment**

**Analysis Note**

The content of this note is intended for CMS internal use and distribution only



**19 October 2007**

# Measurement of jet energy scale corrections using top quark events

J. D'Hondt (editor), P. Van Mulders

*Vrije Universiteit Brussel, Belgium*

**Abstract**

A least-square kinematic fit using Lagrange multipliers is applied to enforce the mass constraints in the reconstructed $pp \rightarrow t\bar{t} \rightarrow q\bar{q}b\mu\nu_{mu}\bar{b}$ events. Both the W boson mass and the top quark mass are constrained to agree with their measured values. Residual corrections are estimated on the energy scale of the jet arising from both the light quarks in the W boson decay and the heavier bottom quark in the top quark decay. Utilizing the first integrated luminosity of 100pb$^{-1}$ of proton collisions at 14 TeV detected by the CMS detector, an uncertainty of xx % can be obtained on the jet energy scale from light quark jets and xx % on the jet energy scale from bottom quark jets. This analysis is based on detailed simulations of both the physics processes and the detector responds.

**DISCLAIMER: This note should be considered as a working document not to be subjected to any approval scheme, but only as an illustration of a new method to extract jet energy scale corrections. The plots are only as illustration and are in most of the cases to be considered as placeholders rather than results. The plots do not necessary reflect what is in the text or the caption but only show the reader want kind of plot is to be expected. The plots and some numbers are obtained from a signal $t\bar{t}$+0jet AlpGen sample using only part of the event selection described and a jet combination choosen via the generator information.**

# 1 Introduction

Compared to experiments at other colliders, the Compact Muon Solenoid (CMS) detector at the Large Hadron Collider (LHC) will collect a very large sample of events in which top quark are present. In the proton collisions at 14 TeV the Next-to-Leading Order (NLO) cross section of the process $pp \to t\bar{t}$ is about 830 pb [1]. Neglecting QCD corrections the top quark decays with a branching ratio of 100% in a W boson and a bottom quark, $t \to bW$[1)]. The masses of the both the W boson and the top quark are however measured precisely by the LEP and Tevatron experiments, and summarized in [2]. Therefore in the reconstruction of these processes using the CMS detector, the correct empirical masses equal to their world average values have to be obtained. Deviations from this could indicate that residual corrections on for example the energy scale of the hadronic jets in the top quark decay can be estimated. Applying on an event-by-event basis a kinematic fit forcing these mass constraint to the reconstructed event, these residual jet energy scale corrections can be estimated. The potential of this new method with 100pb$^{-1}$ of integrated luminosity is studied with detailed simulated events.

In Section 2 the different samples of simulated events are described. Section 3 elaborates on the reconstruction tools applied throughout the analysis including the resolutions obtained on the different reconstructed objects. The event selection criterium to reject W processes associated with hard jet and QCD multi-jet processes is detailed in Section 4. A thorough description of the kinematic fit is found in Section 6, followed with a study of the method used to extract the residual jet energy scale corrections in Section 7. The systematic uncertainties relevant for this measurement are mentioned in Section 8. The paper ends with Section 9 formulating conclusions on the potential of this new method.

# 2 Simulated samples

The AlpGen generator [3] has been utilized to simulate the physics processes of interest in this paper. In Table 1 a summary is given of the size of the different samples. The matching between the Matrix Element and Parton Shower initiated jets in AlpGen is performed with a threshold of 30 GeV/c.

In order to study the systematic influence of pile-up collisions, event samples are generated with and without the addition of low-luminosity pile-up collisions. This was done only for all $\mathrm{t\bar{t}}$ events using the `FastSim` framework[4]. All other event samples, for the study of systematic uncertainties, are also constructed using this `FastSim` framework. Due to the construction of the event selection criterium, all other background processes are assumed negligible.

# 3 Reconstruction of the event and object resolutions

The top quark has a branching ratio of about 100% to decay as $\mathrm{t} \to \mathrm{Wb}$, while the W boson decays has a leptonic branching ratio of BR($\mathrm{W} \to \mathrm{l}\nu_\mathrm{l}$)=$\frac{1}{3}$ and a hadronic branching ratio of BR($\mathrm{W} \to \mathrm{q\bar{q}}$)=$\frac{2}{3}$ (in the absence of QCD corrections). The generated final state topology of the semi-leptonic decay channel $\mathrm{pp} \to \mathrm{t\bar{t}} \to \mathrm{bq\bar{q}b}\mu\nu_\mu$ consists of four coloured partons of which two are heavy, a muon and a neutrino. The detector final state therefore can be characterized by four hadronic jets of which two originate from a heavy quark, an isolated muon and missing transverse energy. In this parer we consider the measurement of the jet energy scale corrections using the semi-leptonic $\mathrm{t\bar{t}}$ processes where the lepton is a muon, but the method is generally viable for the electron final state.

The events passing the single-muon High Level Trigger [5] are reconstructed using the general CMS Software framework CMSSW version 1-3-2 [6]. The jets are reconstructed from the combined electromagnetic and hadronic

[1)] Throughout the paper the charge conjugation will be implicit, hence indices indicating the charge will be omitted.

| | Number of events | Int.Luminosity $pb^{-1}$ | Cross-section pb |
|---|---|---|---|
| $t\bar{t} \rightarrow bq\bar{q}b\mu\nu_\mu$ | | | |
| $t\bar{t}$ + 0 jets | k | | |
| $t\bar{t}$ + 1 jets | k | | |
| $t\bar{t}$ + 2 jets | k | | |
| $t\bar{t}$ + 3 jets | k | | |
| $t\bar{t}$ + 4 jets | k | | |
| other $t\bar{t}$ decay channels | | | |
| $t\bar{t}$ + 0 jets | k | | |
| $t\bar{t}$ + 1 jets | k | | |
| $t\bar{t}$ + 2 jets | k | | |
| $t\bar{t}$ + 3 jets | k | | |
| $t\bar{t}$ + 4 jets | k | | |
| W + 4jets | k | | |
| W + 5jets | k | | |
| W + 6jets | k | | |
| Wbb + 2jets | k | | |
| Wbb + 3jets | k | | |
| QCD ?? | k | | |

Table 1: Overview of the number of analyzed simulated single-muon $t\bar{t}$ and background events with their corresponding integrated luminosity calculated with the indicated Leading-Order cross-sections. The samples of a certain process indicated with the highest jet multiplicity assumes the inclusion of all events with a higher multiplicity.

calorimeter deposites and clustered with the Iterative Cone algorithm using an opening angle of $\Delta R = 0.5$ in an $(\eta, \phi)$-metric. The energy scale of the jets are calibrated based on simulated events with a procedure which compensates the responds on $E_T$ in QCD events as a function of the transverse energy of the reconstructed jet. For the remainder of the paper the four jets with the highest transverse energy in the pseudo-rapidity range $|\eta| \leq 2.5$ are considered to be the hadronic decay products of the semi-leptonic $t\bar{t} \rightarrow bq\bar{q}b\mu\nu_\mu$ process. A muon candidate is formed when a muon track is reconstructed in the muon chambers and a matching track is found in the main tracker. These reflect the so-called global reconstructed muons in the CMS detector. In the analysis the muon with the largest transverse momentum in the pseudo-rapidity range $|\eta| \leq 2.5$ is used to reconstruct the topology of the top quark pair. The missing transverse energy is reconstructed as the complement to zero of the vectorial sum of the reconstructed jets in the event. In this process the muons are also included in the vectorial sum. To differentiate between light and heavy quark jets, the nominal track counting b-tagger is applied.

Important for the Least-Square kinematic fit to be described in Section 6 the resolutions on the kinematic parameters of the jet and muons need to be know. There are several options to parametrize the four-vector of the objects. Obviously one can parametrize this object as $(E_T, \theta, \phi)$ or $(E_T, \eta, \phi)$ vectors. For the jets however a four-vector parametrization can be applied describing the deviation from the measured momentum in a local coordinate system:

$$\vec{u}_1 = \frac{\vec{p}_m}{|\vec{p}_m|}, \quad \vec{u}_2 = \frac{\vec{u}_3 \times \vec{u}_1}{|\vec{u}_3 \times \vec{u}_1|}, \quad \vec{u}_3 = \frac{\vec{u}_z \times \vec{u}_1}{|\vec{u}_z \times \vec{u}_1|} \ . \tag{1}$$

where $u_z$ is the unit vector along the z-direction. The fitted momentum is parametrized with three parameters $a$, $b$, and $c$,

$$\vec{p}_f = a|\vec{p}_m|\vec{u}_1 + b\vec{u}_2 + c\vec{u}_3 \tag{2}$$

and the particle has a free floating fitted energy:

$$E_f = d \cdot E_m \ . \tag{3}$$

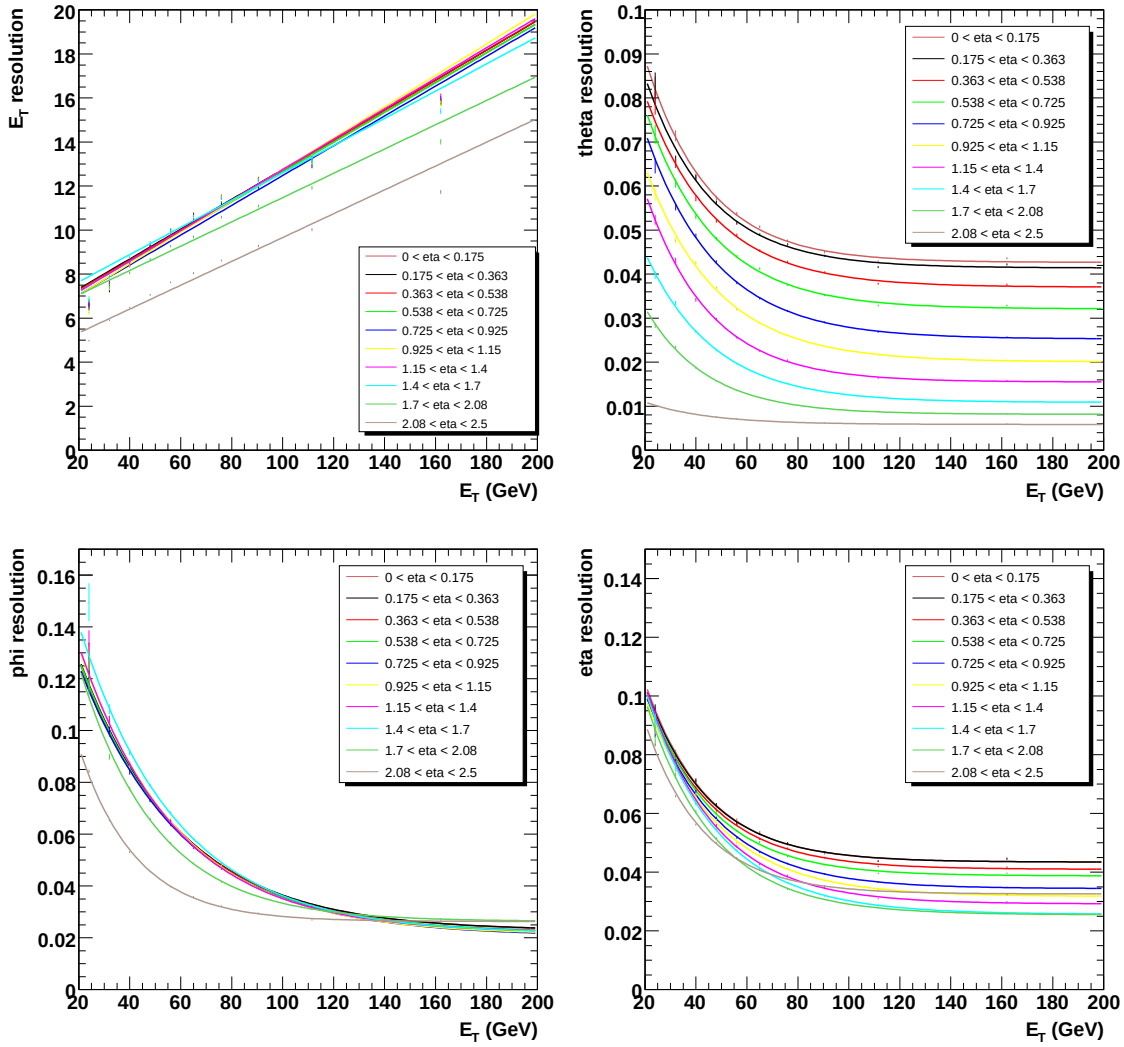



Figure 1: Resolutions on the $E_T$, $\theta$, $\phi$ and $\eta$ of the reconstructed jets arising from a quark in the W boson decay.

For this parametrization the resolutions on the parameters $(a, b, c, d)$ are determined. The resolution on the above mention parameters are determined for reconstructed jets, muons, electrons and where relevant the missing transverse energy. The reconstructed jet matching in an $(\eta, \phi)$ metric the primary parton better than 0.3 is linked to the parton. From the distribution of the residuals $(x_{parton} - x_{jet})$ the resolution is obtained by fitted the relevant range of the distribution with a Gaussian function. The square root of the estimated variance of this Gaussian is taken as the resolution. These resolutions are estimated in several bins of pseudo-rapidity $\eta$ and transverse momentum $p_T$. For the leptons a similar approach is taken using the couples of generated leptons from the W decay and reconstructed leptons matching to better than 0.1 in $(\eta, \phi)$ space. It was shown that the influence of the choice of matching criterium is neglegible.

Both the Layer-1 and Layer-2 of the Top Quark Analysis Framework (TQAF) is being applied as a platform for the analysis [7].

# 4 Event selection criterium

The event selection is based on the objects reconstructed as described in the previous section. The four jets with the highest transverse energy in $|\eta| \leq 2.5$ are required to have a transverse energy above 30 GeV/c. The muon with the highest transverse momentum in $|\eta| \leq 2.5$ is requested to have a transverse momentum above 20 GeV/c.

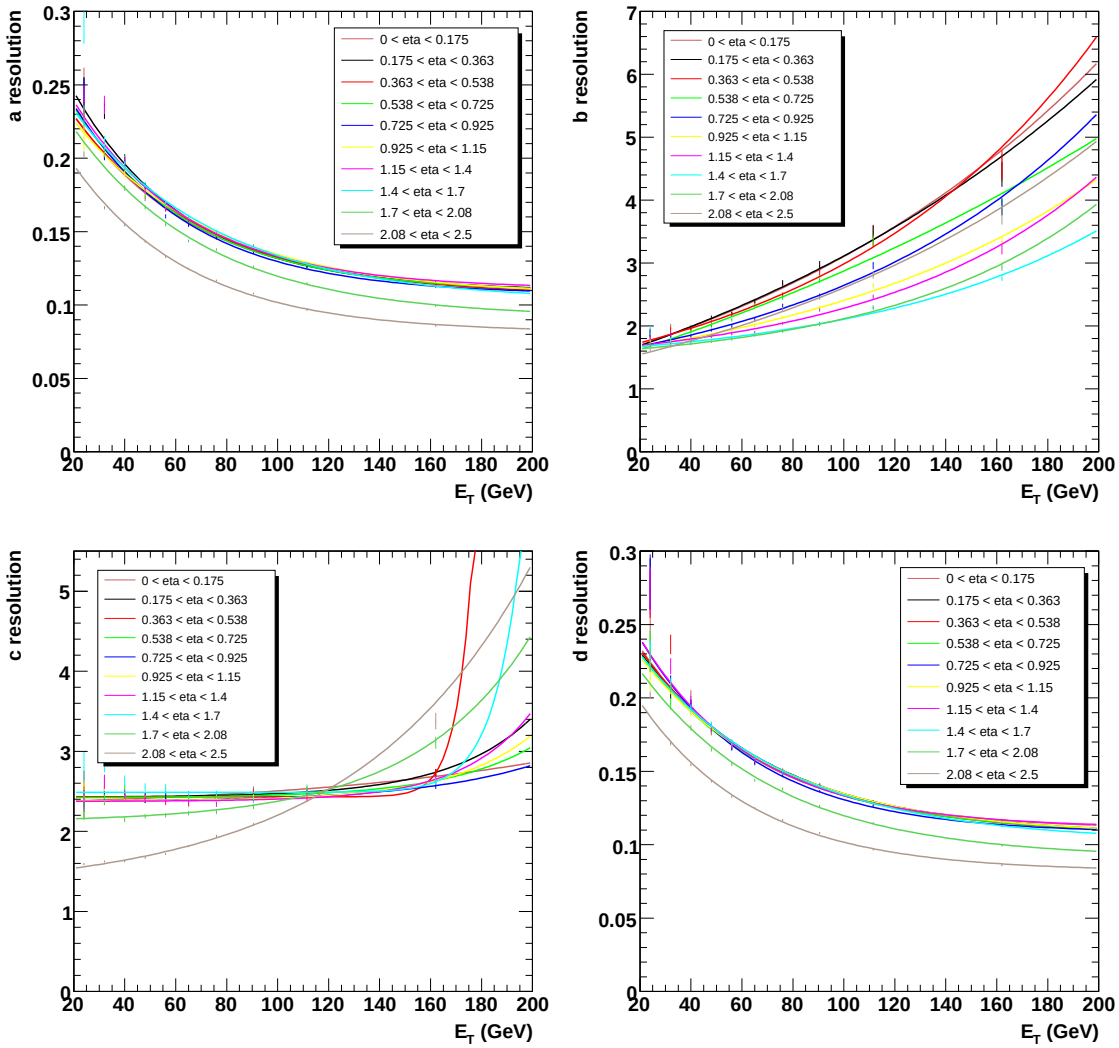



Figure 2: Resolutions on the $a$, $b$, $c$ and $d$ of the reconstructed jets arising from a quark in the W boson decay.

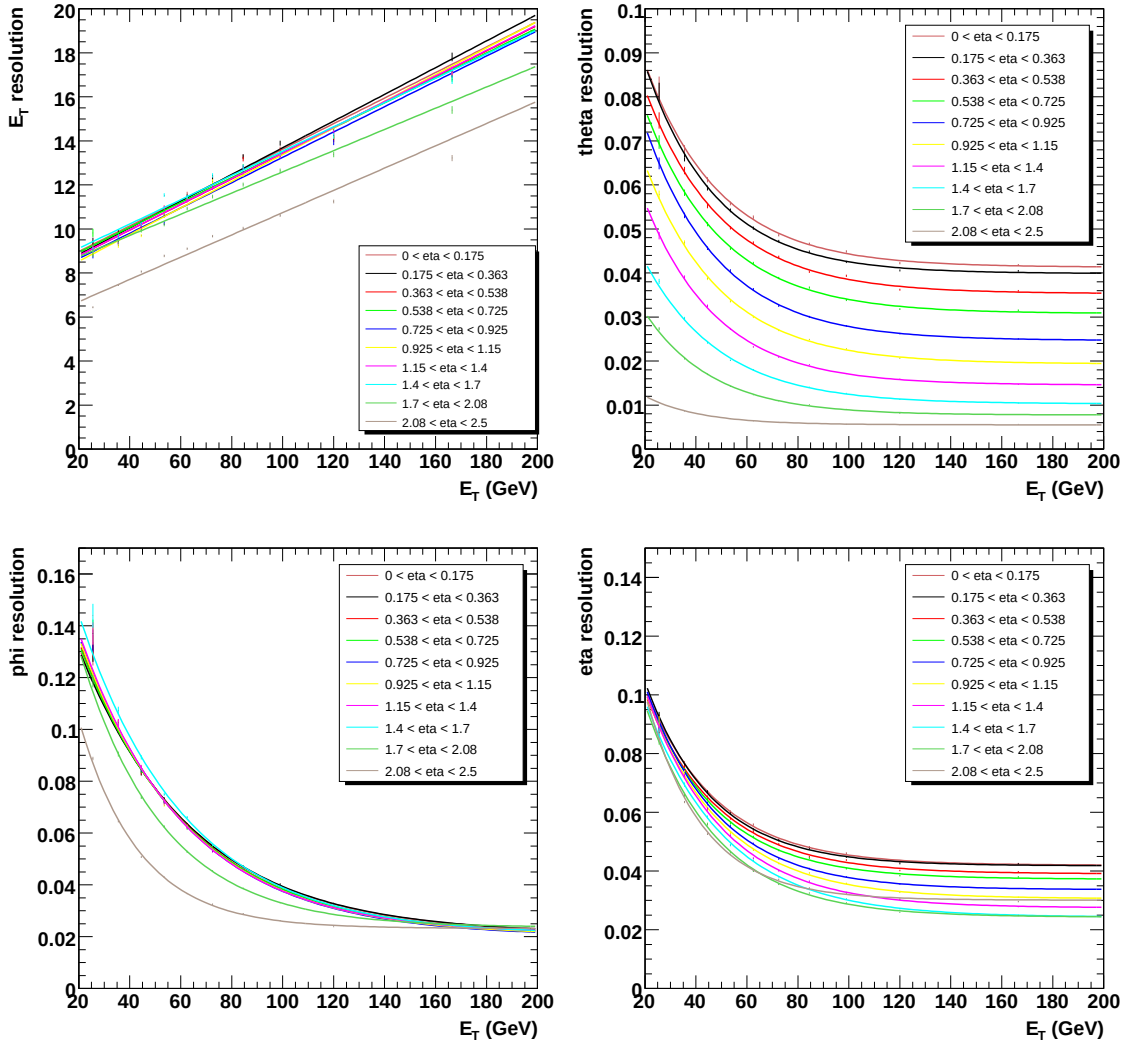



Figure 3: Resolutions on the $E_T$, $\theta$, $\phi$ and $\eta$ of the reconstructed jets arising from a bottom quark in the top quark decay.

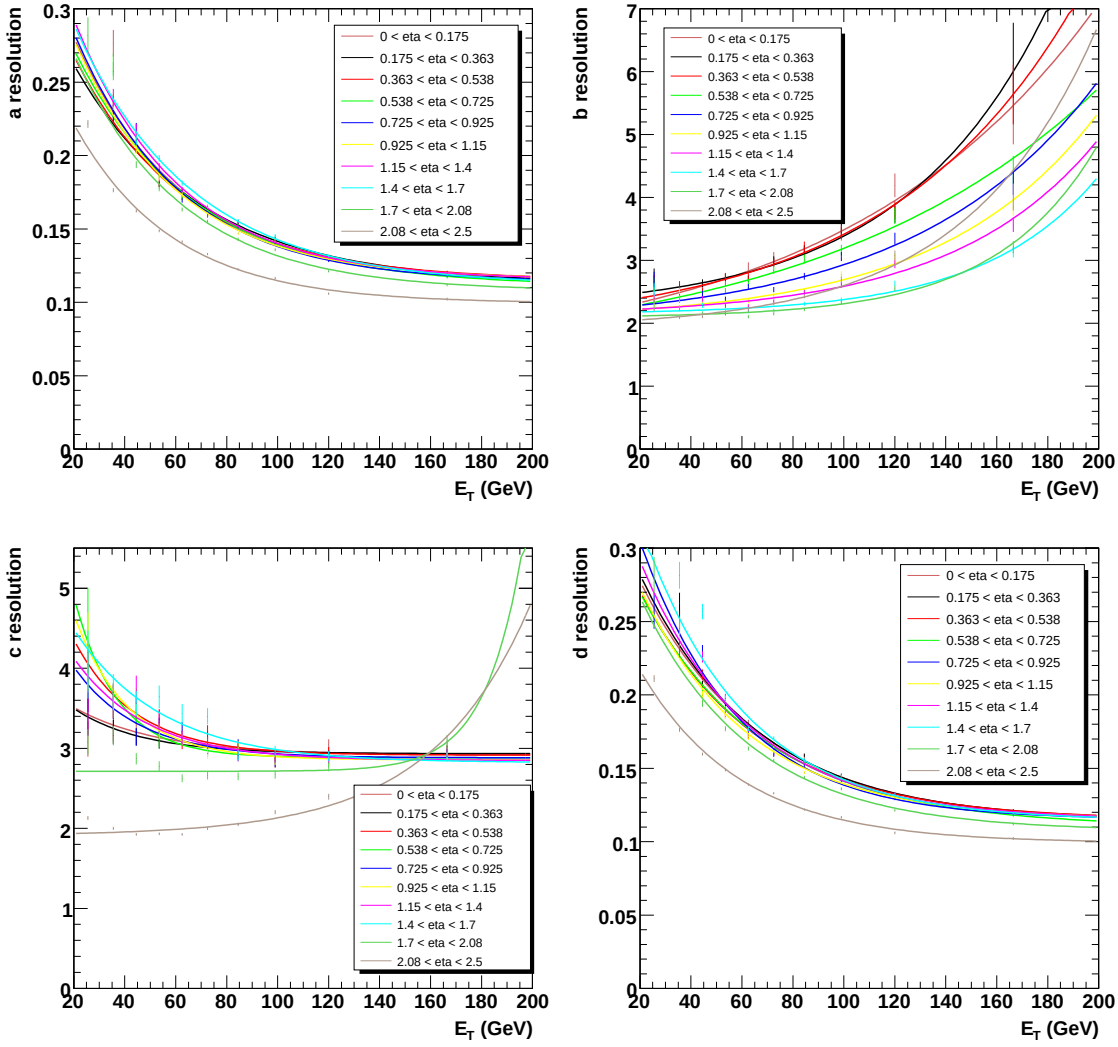



Figure 4: Resolutions on the $a$, $b$, $c$ and $d$ of the reconstructed jets arising from a bottom quark in the top quark decay.

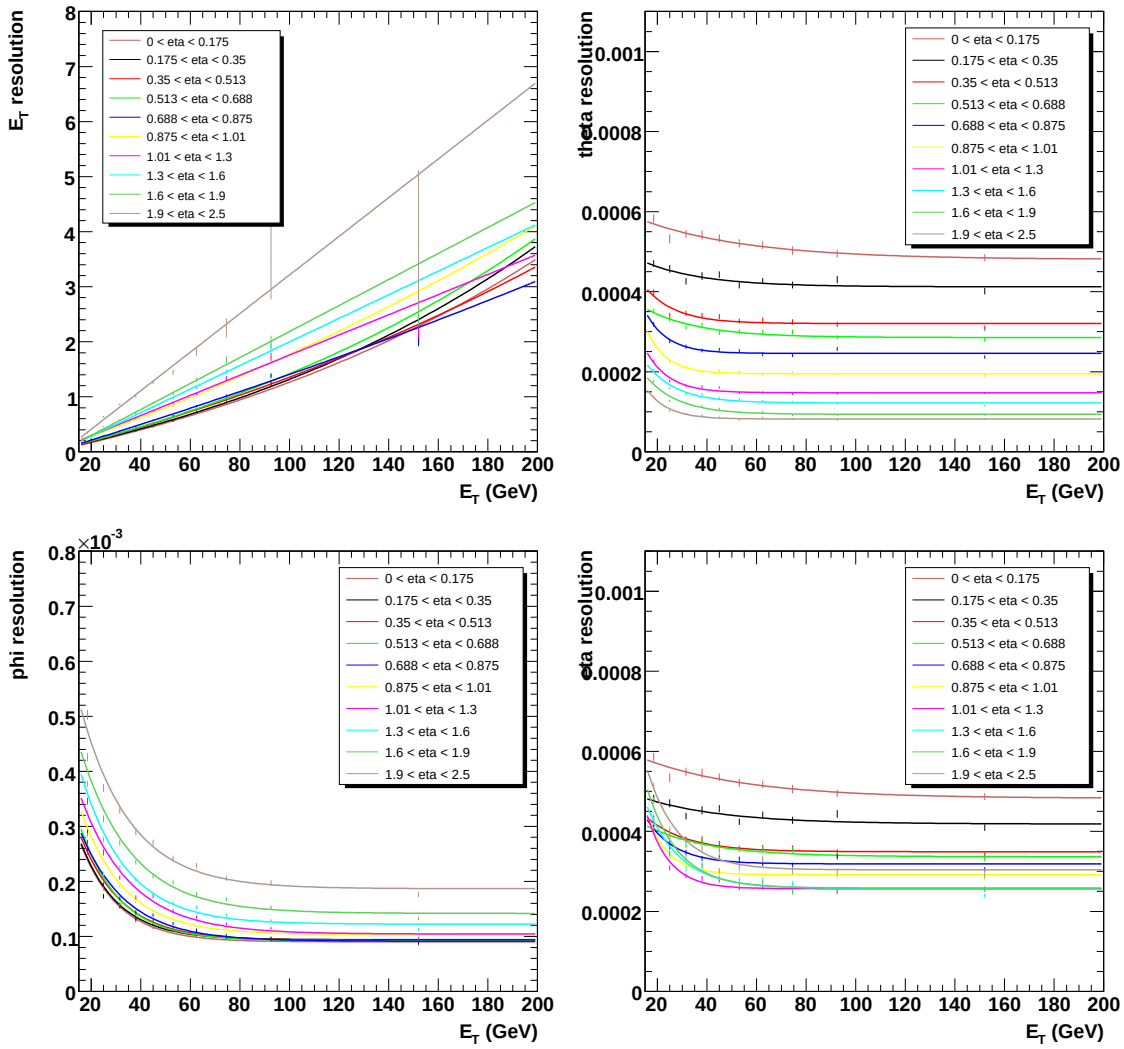



Figure 5: Resolutions on the $E_T$, $\theta$, $\phi$ and $\eta$ of the reconstructed muon arising from the W boson decay.

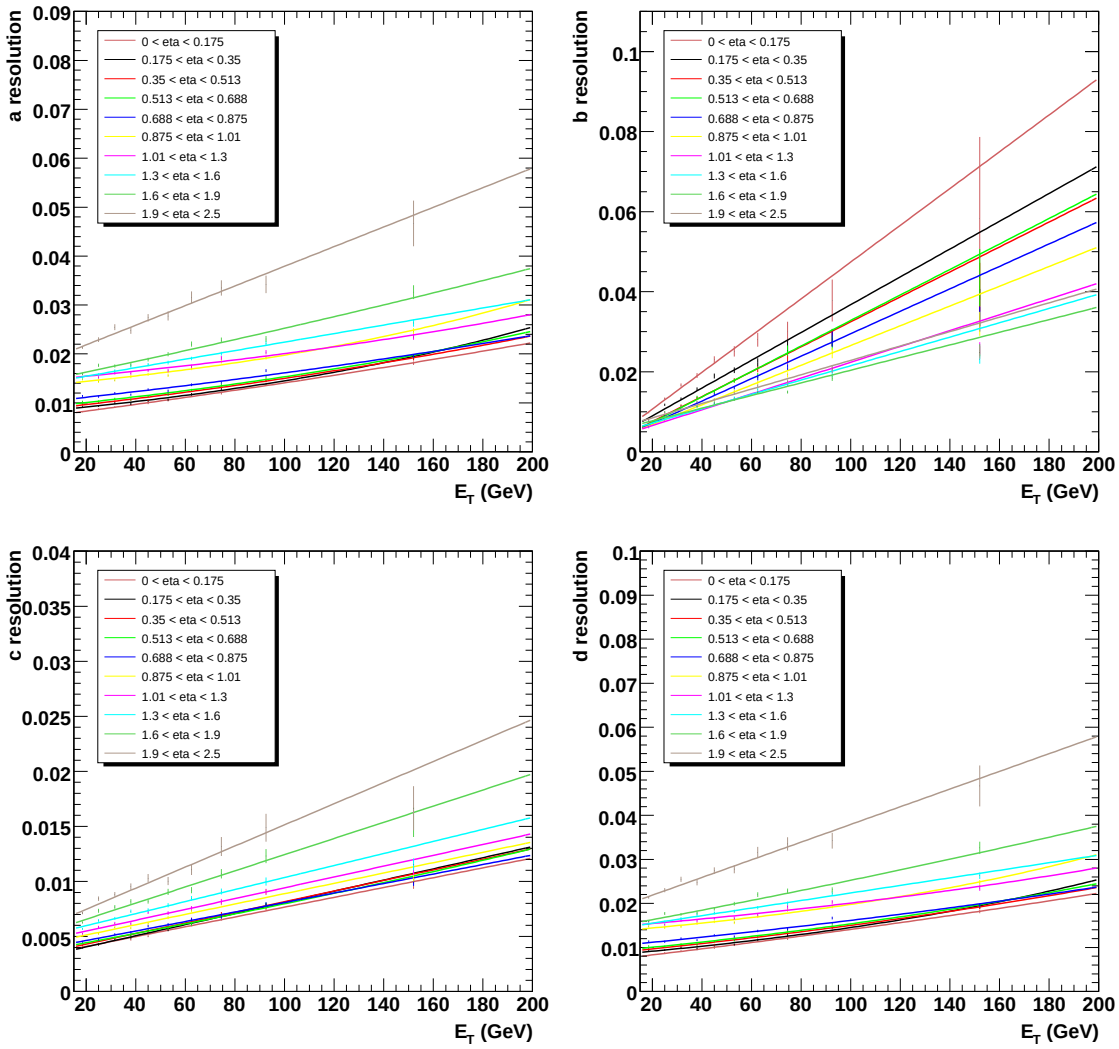



Figure 6: Resolutions on the $a$, $b$, $c$ and $d$ of the reconstructed muon arising from the W boson decay.

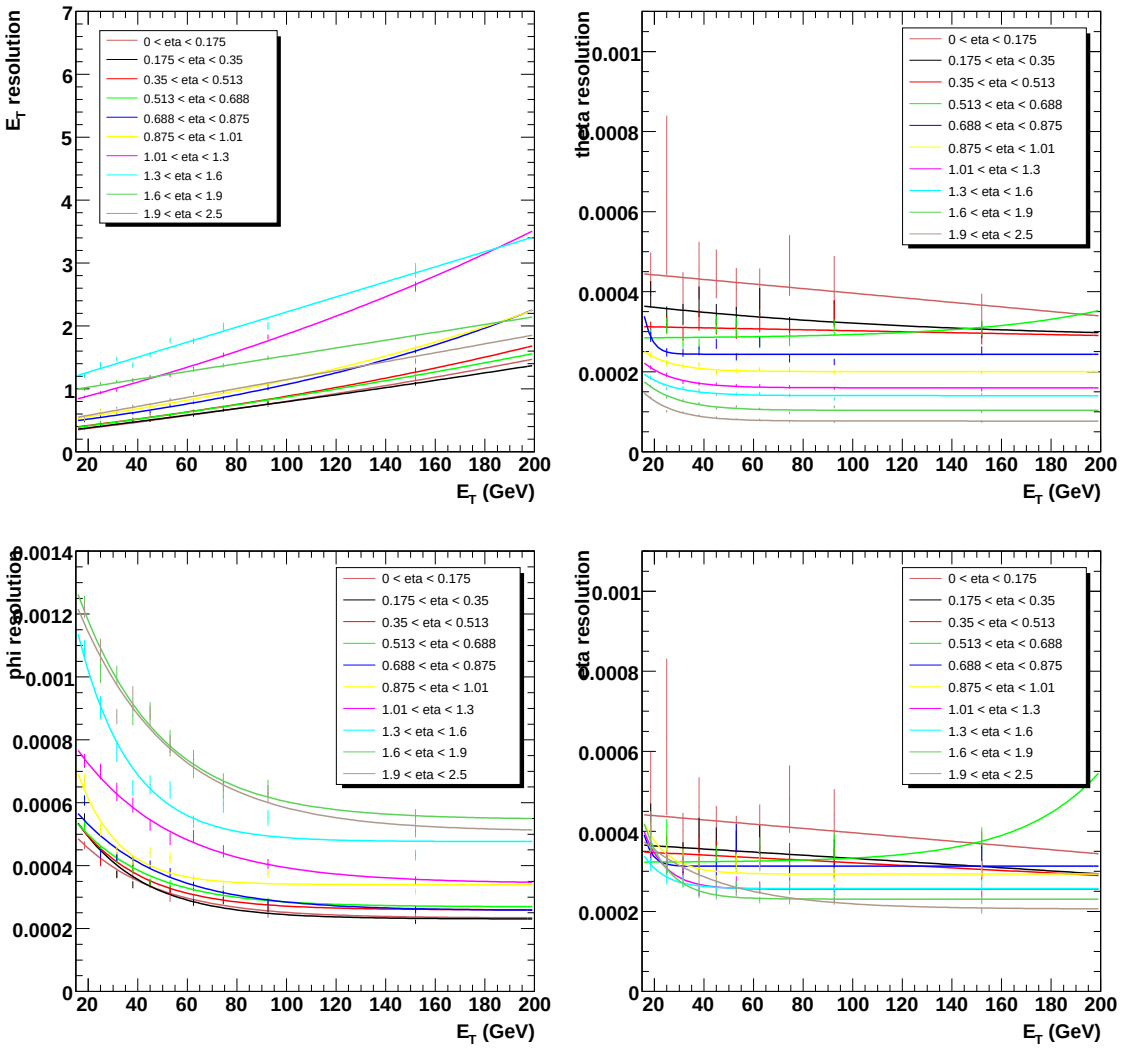



Figure 7: Resolutions on the $E_T$, $\theta$, $\phi$ and $\eta$ of the reconstructed electron arising from the W boson decay.

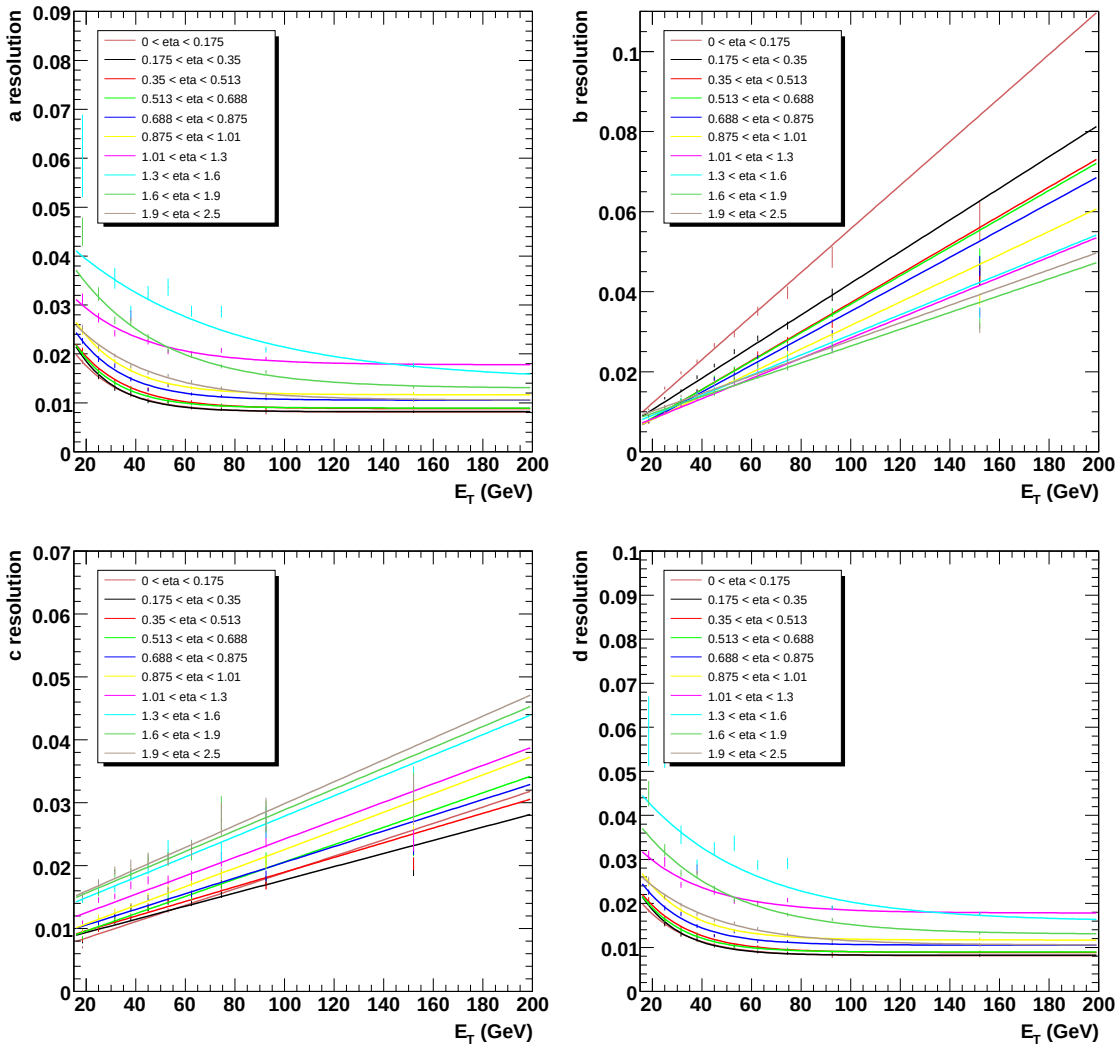



Figure 8: Resolutions on the $a$, $b$, $c$ and $d$ of the reconstructed electron arising from the W boson decay.

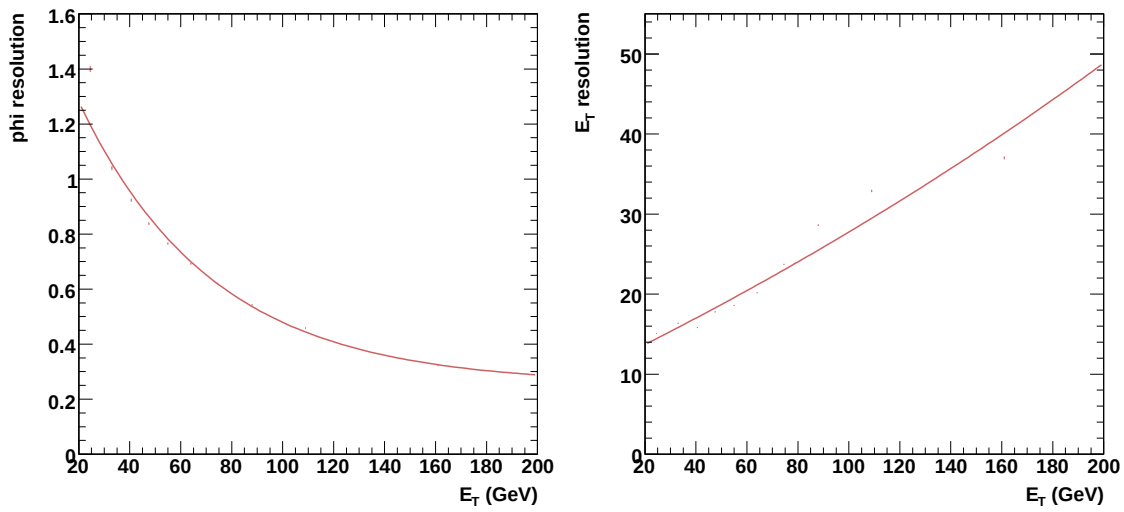



Figure 9: Resolutions on the $E_T$ and $\phi$ of the reconstructed missing transverse energy arising from the neutrino in the W decay.

The distributions of the $E_T$ of the jet with the fourth highest transverse energy in the event within $|\eta| \leq 2.5$ and the $p_T$ of the muon are shown in Figure 10 after the cut being applied. Similar the distributions of the other selection variables are shown after applying these initial selection cuts. The muon tracker isolation variable

$$p_t^{iso} = \sum_{\Delta R(track,muon)<0.3} p_{t,tracks} \tag{4}$$

and a calorimeter isolation variable

$$E_T^{iso} = \sum_{\Delta R(calo,muon)<0.3} E_{T,calo} \tag{5}$$

are required to be smaller than respectively xx and xx. The direction of the muon in the above equations are taken at the vertex. Distributions of these isolation variables can be found in Figure 11. The track counting b-tagger is used, defining a b-tagged jet with a discriminator value above xx. Two out of the four jets with the highest $E_T$ are required to be b-tagged. The distributions of the two highest b-tag discriminator values in the event are shown in Figure 12. To avoid difficulties in the interpretation of the result, the four jets are required not to overlap in the $(\eta, \phi)$ space where they are constructed. For this the distribution of the smallest angle between the four relevant jets is shown in Figure 13 together with the smallest angle between the direction of the choosen muon at the vertex and these four jets.

The cumulative efficiencies of the above sequential cuts are shown in Table 2. **some explanation...**

# 5 Jet combination procedure

For the selected events, four variables are being calculated to select the correct jet combination into a $\mathrm{t\bar{t}} \rightarrow \mathrm{bq\bar{q}b}\mu\nu_\mu$ topology. For each of the possible jet combinations the following variables are determined

- Sum of the transverse momenta of the two jets assumed b jets in the combination divided by the sum of the transverse momenta of the two other jets (notation $x_1$);
- Transverse momentum of the hadronic top quark in the combination scaled to the average transverse momentum of the hadronic top quark over all possible jet combinations (notation $x_2$);

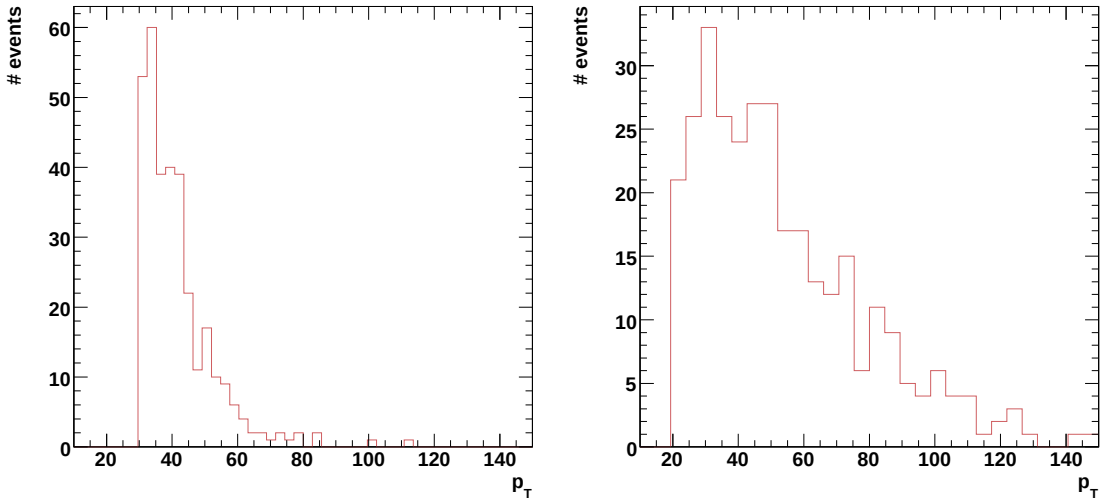



Figure 10: **PLACEHOLDER** Distribution of the transverse energy $E_T$ of the jet with the fourth highest $E_T$ in the event (left) and the transverse momentum $p_T$ of the muon (right).

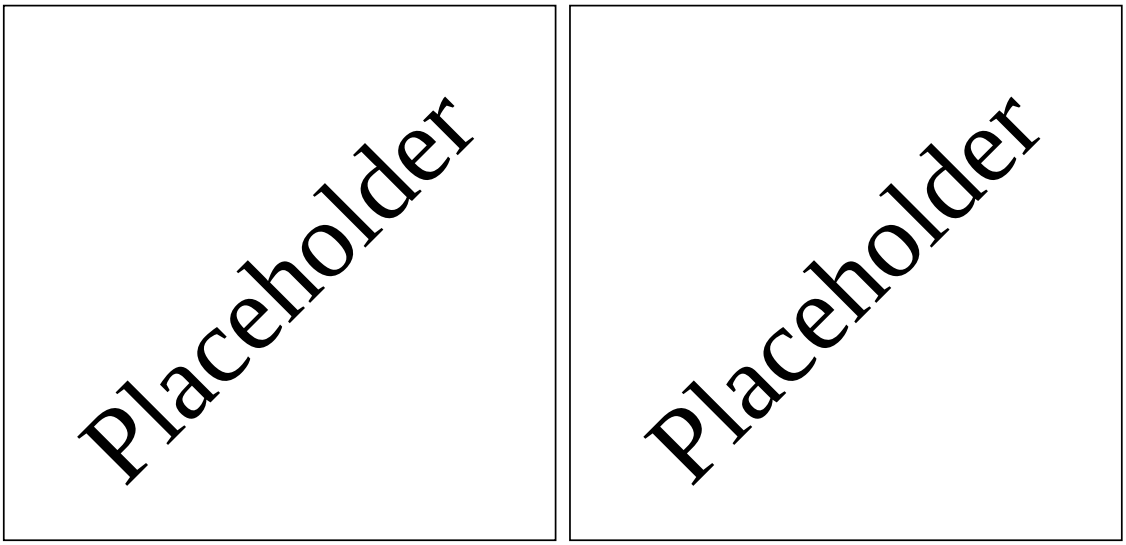



Figure 11: **PLACEHOLDER** Distribution of the tracker $p_T^{iso}$ (left) and calorimeter $E_T^{iso}$ (right) isolation variables as defined in the text.

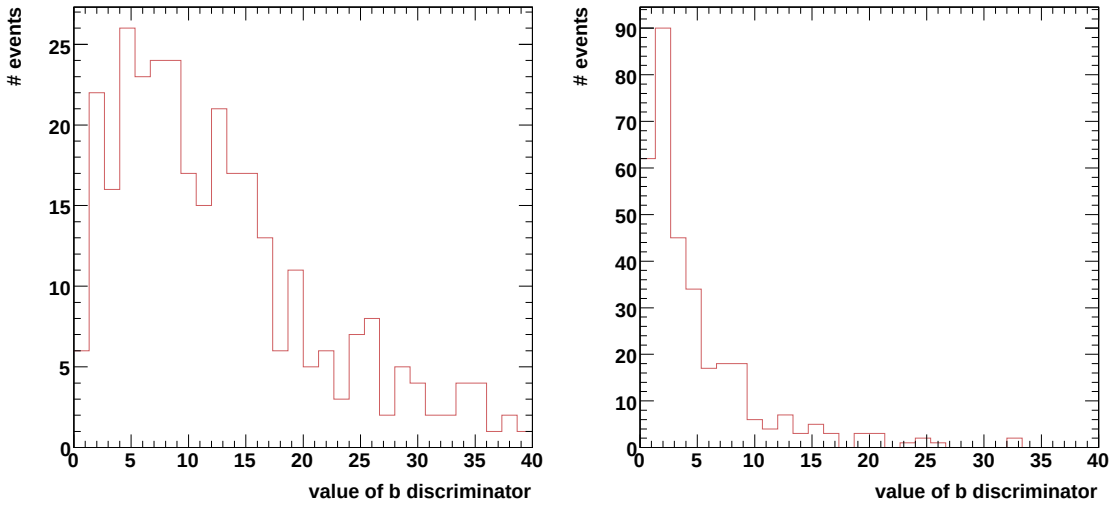



Figure 12: **PLACEHOLDER** Distribution of highest (left) and second highest (right) b-tag discriminator values on the event.

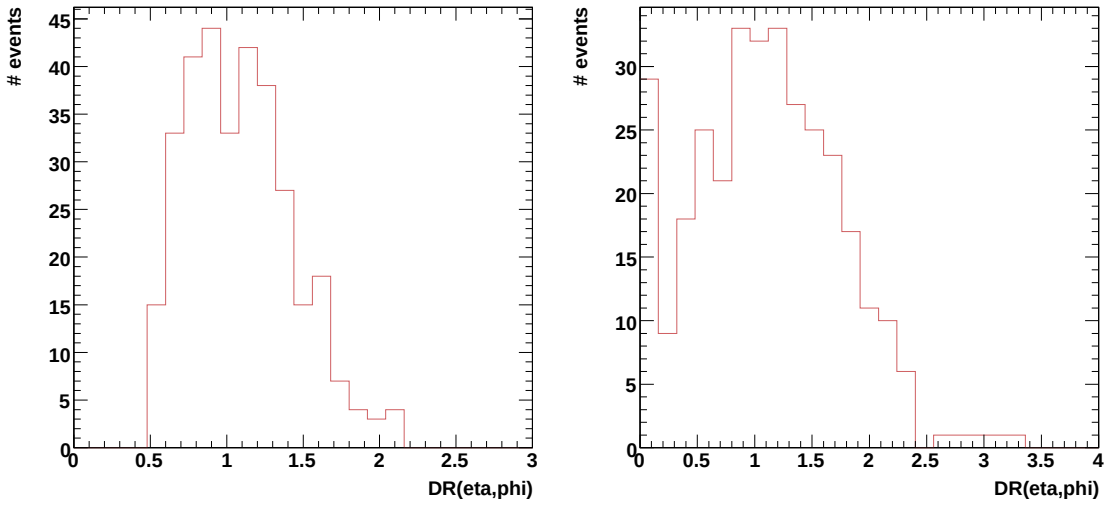



Figure 13: **PLACEHOLDER** Distribution of smallest angle between the four relevant jets in an $(\eta, \phi)$ metric (left) and the smallest angle in the same metric between the selected muon and the four jets (right).

| | $E_T^{jet}$ $p_T^{lepton}$ | lepton isolation | b-tag | jet overlap | CLR |
|---|---|---|---|---|---|
| $t\bar{t} \rightarrow bq\bar{q}b\mu\nu_\mu$ | | | | | |
| $t\bar{t}$ + 0 jets | | | | | |
| $t\bar{t}$ + 1 jets | | | | | |
| $t\bar{t}$ + 2 jets | | | | | |
| $t\bar{t}$ + 3 jets | | | | | |
| $t\bar{t}$ + 4 jets | | | | | |
| other $t\bar{t}$ decay channels | | | | | |
| $t\bar{t}$ + 0 jets | | | | | |
| $t\bar{t}$ + 1 jets | | | | | |
| $t\bar{t}$ + 2 jets | | | | | |
| $t\bar{t}$ + 3 jets | | | | | |
| $t\bar{t}$ + 4 jets | | | | | |
| W + 4jets | | | | | |
| W + 5jets | | | | | |
| W + 6jets | | | | | |
| Wbb + 2jets | | | | | |
| Wbb + 3jets | | | | | |
| QCD ?? | | | | | |

Table 2: Overview of the cumulative efficiencies of the sequential selection cuts applied on the event sample.

- The angle in $(\eta, \phi)$ space between the lepton and the bottom quark from the leptonic top quark decay (notation $x_3$);
- Sum of the b-tag discriminants of the two jets assumed b jets (notation $x_4$).

In reconstructed $t\bar{t} \rightarrow bq\bar{q}b\mu\nu_\mu$ events for which a correct jet combination exist (reconstructed jets are present which match the primary partons), the distributions of the variables are made for the correct and wrong jet combinations. These distributions can be found in Figure 14. After normalization these distributions reflect the probability density functions for signal combinations $P_i^S$ and background combinations $P_i^B$, hence a likelihood ratio $P_i(x_i)$ for each of the observables $x_i$ is calculated as

$$P_i(x_i) = \frac{f_i(x_i)}{1 - f_i(x_i)} \tag{6}$$

with $f_i(x_i)$ the fitted function over all bins of $x_i$ of $P_i^S(x_i)/(P_i^S(x_i) + P_i^B(x_i))$ which are determined in bins of the observable $x_i$. The product of the likelihood ratio values over all four observables reflect the value of the combinated likelihood ratio

$$P_C(\vec{x}) = \prod_{i=1}^{4} P_i(x_i) \tag{7}$$

and is the optimal use of the statistical power to separate signal and background jet combinations for these variables if they are uncorrelated. Assuming linear correlations, the correlations between the variables are found to be small as shown in Table 3. The distribution of this combined likelihood ratio variable $P_C(\vec{x})$ is shown in Figure 15. The best jet combination out of the 12 possibilities per event is choosen as the combination with the largest combined likelihood ratio value $P_C$, denoted as the unique value $P_C^{max}$ for each event.

In Figure 16 the distribution of the variable $P_C^{max}$ is shown for signal and background processes. This variable helps in rejecting background processes, and an extra sequential cut is made demanding the event to have a value of $P_C^{max}$ larger than xx. The cumulative efficiency of this cut is shown in Table 2.

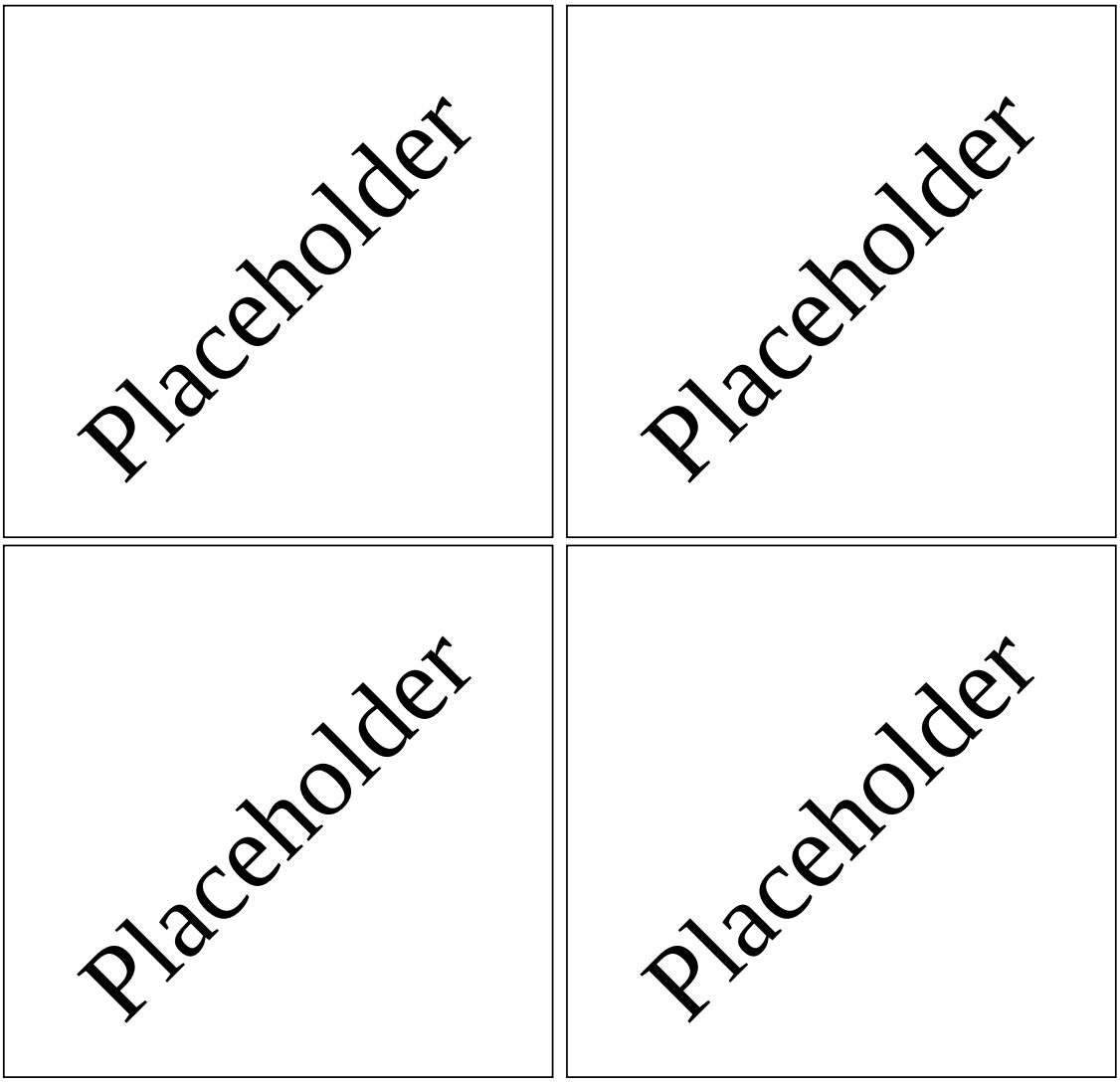



Figure 14: **PLACEHOLDERS** Distributions of the four relevant variables to distinguish between a correct and wrong jet combinations in $\mathrm{t\bar{t}} \to \mathrm{bq\bar{q}b}\mu\nu_\mu$ signal events.

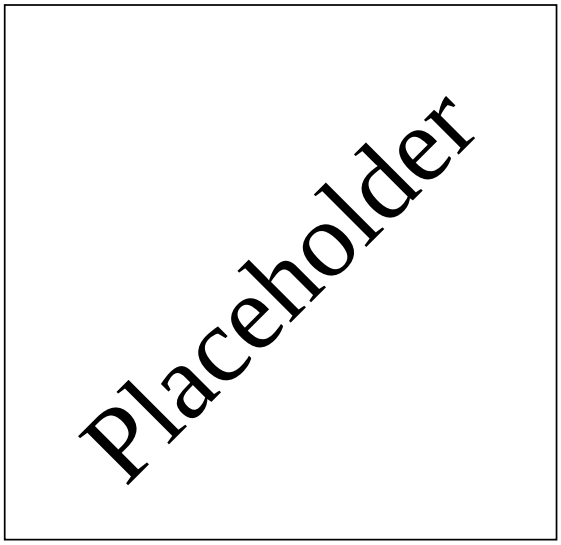



Figure 15: **PLACEHOLDERS** Distribution of the combined likelihood ratio value for correct and wrong jet combinations in $\mathrm{t\bar{t}} \to \mathrm{bq\bar{q}b}\mu\nu_\mu$ signal events.

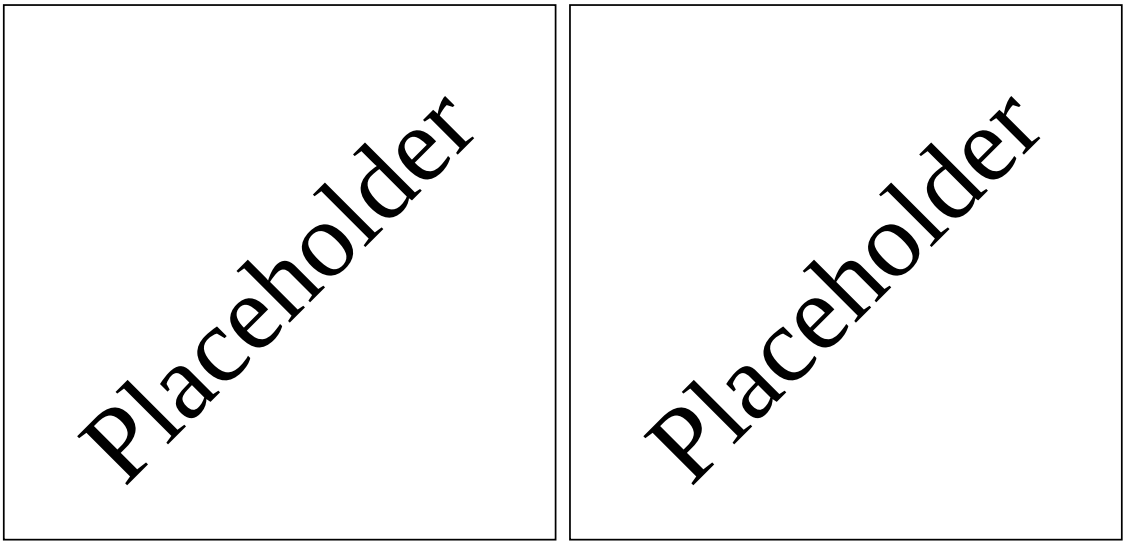



Figure 16: **PLACEHOLDERS** Distribution of the maximal combined likelihood ratio value $P_C^{max}$ over all jet combinations in the event (left) and for $\mathrm{t\bar{t}} \to \mathrm{bq\bar{q}b}\mu\nu_\mu$ signal events the efficiency and purity of choosing the correct jet combination when making a cut on this variable (right).

| | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|---|---|---|---|---|
| $x_1$ | 1 | - | - | - |
| $x_2$ | | 1 | - | - |
| $x_3$ | | | 1 | - |
| $x_4$ | | | | 1 |

Table 3: Linear correlations between the variables used in the likelihood ratio method to select the correct jet combination. The variables $x_i$ are defined in the text.

# 6 Kinematic fit

The general concept of non-linear least square fitting using Lagrange Multipliers applied in this paper is described in reference [8]. For convenience the main mathematical framework is repeated in this paper as it is the most essential tool applied in this paper for extracting the jet energy scale corrections from top quark events.

A physical problem in many cases consists of measured quantities as particles' four-vectors and unmeasured values. While the measured quantities typically represent 4-vector estimators of certain reconstructed objects (e.g jets, tracks, etc), the unmeasured quantities are supposed to mark the undetected particles of the underlying primary event structure (e.g. neutrinos). Also one often has a certain hypothesis for an event. Then constraints like energy and momentum conservation or mass constraints can be used to slightly change the measured values within the uncertainties to fulfil these kinematic requirements. This procedure one usually calls a *kinematic fit*, which is applied by a least square method. Below the basic framework for kinematic fitting is sketched.

Let us assume we have to solve a problem with $n$ measured parameters $\vec{y}$, $p$ unmeasured parameters $\vec{a}$ and $m$ constraints $\vec{f}$ as defined in eq. 8. These requirements will be fulfilled for the true parameters $\bar{a}$ and $\bar{y}$.

$$\begin{aligned} f_1(\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_p, \bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n) &= 0 \\ f_2(\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_p, \bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n) &= 0 \\ &\vdots \\ f_m(\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_p, \bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n) &= 0 \end{aligned} \tag{8}$$

In general the measured values $\vec{y}$ will not fulfil the constraints, so that one has to calculate corrections $\Delta\vec{y}$. Then the sum $\vec{y}' = \vec{y} + \Delta\vec{y}$ will fulfill the constraints. In the same step, the weighted sum

$$S(\vec{y}) = \Delta\vec{y}^T \mathcal{V}^{-1} \Delta\vec{y} \tag{9}$$

should be minimal with $\mathcal{V}$ being the covariance matrix of the measured parameters.

A general method to determine local extrema of non-linear functions of many variables is by using the definition of *Langrange Multipliers*. The likelihood of this new problem is defined as follows:

$$L(\vec{y}, \vec{a}, \vec{\lambda}) = S(\vec{y}) + 2\sum_{k=1}^{m} \lambda_k f_k(\vec{y}, \vec{a}) \tag{10}$$

with $\vec{\lambda}$ the Lagrange Multipliers. The necessary condition for a local minimum of this function is then equivalent to the condition for a minimum of $S(\vec{y})$ under the constraint $f_k(\vec{y}, \vec{a}) = 0$.

With linear constraints the solution can be found in one step, otherwise it has to be found iteratively. This means that in every iteration the problem is linearized. The linearization is repeated until the defined convergence criteria are fulfilled. These should guarantee first that the $\chi^2$-expression $S$ only changes by a given value $\epsilon_S$ from one iteration to the next and second that the constraints are fulfilled to better than a given value $\epsilon_F$:

$$\frac{S(n-1) - S(n)}{ndf} < \epsilon_S \qquad \text{and} \qquad F_{(n)} = \sum_{k=1}^{m} \left| f_k^{(n)}(\vec{y}, \vec{a}) \right| < \epsilon_F \tag{11}$$

In eq. 11 $n$ denotes the number of the iteration and $ndf$ is just the difference between the constraints and the amount of unmeasured quantities:

$$ndf = m - p \ . \tag{12}$$

To protect the fit from overshooting and jumping from a minimum, in every iteration $n$ the value of the constraints $F_{(n)}$ is compared with the same value from the previous step $F_{(n-1)}$. If $F_{(n)} > F_{(n-1)}$, the corrections to the previous step are divided by two and different values are calculated. If this new correction cannot decrease the constraints with respect to $F_{(n-1)}$, the corrections are reduced further until $F_{(n)}$ becomes smaller than $F_{(n-1)}$. A fixed limit is set on the total amount of times this procedure is applied in every fit.

The analytical nature of this procedure ensures high speed performance, as well as reliable convergences and, therefore, is probably the most common mathematical framework utilized for constraint kinematic reconstructions in particle physics.

In this paper the applied kinematic fit introduces the constraints of the top quark mass and the W boson mass in the hadronic decay chain the top quark. Therefore only the resolutions on the jets are relevant. The four vectors of the jet are parametrized is the dimensions of $(E_T, \theta, \phi)$. In this parametrization the jets are considered as massless objects as the momentum is obtained as $\vec{p} = (E_T cos\phi, E_T sin\phi, E_T cot\theta)$ and the energy as $E = E_T/sin\theta$. The convergence rate of the applied kinematic fit for the selected signal events and the choosen jet combination is about 99%. Alternative parametrizations as $(E_T, \eta, \phi)$ and the one mentioned in Section **??** can be considered. The resolution for all these parametrizations and of all objects relevant for top quark physics are shown in Section **??** as a function of the transverse energy or momentum and the pseudo-rapidity.

# 7 Method to measure the Jet Energy Scale

For the choosen jet combination in the selected events a kinematic fit was applied forcing the top quark mass and the W boson mass in the hadronic decay $t \to bW \to bq\bar{q}$. The kinematic fit provides a probability obtained from the $\chi^2$ value of the fit. This probability reflects how likely the reconstructed event in the choosen jet combination agrees with the hypothesis of the world average emperical determinations of the top quark and W boson masses. The jet energy scale of the jets can be adapted to maximize this probability over all selected events. This was obtained by determining this $\chi^2$ probability $P^{fit}$ is the dimensions of the three residual jet energy corrections, resulting in $P^{fit} = P^{fit}(\Delta E_b, \Delta E_q, \Delta_{\bar{q}})$. The energies and accordingly the momenta of the jet are shifted in steps of 0.5% between -25% and +25%, keeping the $E/|\vec{p}|$ ratio of the jet invariant. At each point in this three dimensional space of $(\Delta E_b, \Delta E_q, \Delta_{\bar{q}})$ the kinematic fit was applied. Hence for each selected event the kinematic fit is repeated about 1 million times, creating a rather CPU intensive analysis.

To obtain the a symmetry between the two dimensions of the jet arising from the W boson decay or the so-called light quark jets, the choice which of the two jets belongs to $\Delta E_q$ or $\Delta E_{\bar{q}}$ was randomized. In the limit of a large number of selected events the probability or $\chi^2$ distribution is symmetric between the $\Delta E_q$ or $\Delta E_{\bar{q}}$ dimensions.

Numerically it is more stable to combine the $\chi^2$ distributions of a selected events rather than to combined the corresponding probability distributions, because the sum is taken in the former rather than the product of small numbers in the latter option. Events are however only taken into account if they have in the range $-25\% \leq \Delta E_i \leq +25\%$ a maximal fit probability exceeding 0.1. The distribution of the maximum fit probability over all events in shown in Figure 17 together with the fit probability distribution when nu residual jet energy scale corrections are introduced. The uniform distribution expected for the latter distribution is nicely observed except the peak around 0 which is for wrongly reconstructed events or events in which a gluon was radiated from the hard partons in the top quark decay which obey the mass constraints. For the former distribution a peak around 1 is visible reflecting that when the jet energy scale corrections are adapted on an event-by-event basis the mass constraints in the events will indeed be fulfilled.

The best values of the residual corrections $(\Delta E_b, \Delta E_q, \Delta_{\bar{q}})$ are obtained when the combined $\chi^2(\Delta E_b, \Delta E_q, \Delta_{\bar{q}})$ distribution is minimized. These values minimizing the distribution are denoted as $((\Delta E_b)_{min}, (\Delta E_q)_{min}, (\Delta_{\bar{q}})_{min})$. They are obtained on the three dimensional grid of points where the kinematic fit has been performed, rather than by fitting a 3D ellipsiodal function through the points. The residual uncertainty from this approximation is negleg-ible given the smalness of the space between the grid points, being 0.5%, and the rather soft correlation between the $\Delta E_b$ and $\Delta E_{q,\bar{q}}$ residual corrections. In Figure 18 slices of the three dimensional $\chi^2$ distributions are shown. The distribution $\chi^2((\Delta E_b)_{min}, \Delta E_q, \Delta E_{\bar{q}})$ at the fixed value of $(\Delta E_b)_{min}$ and $\chi^2(\Delta E_b, \Delta E_q, (\Delta E_{\bar{q}})_{min})$ at the

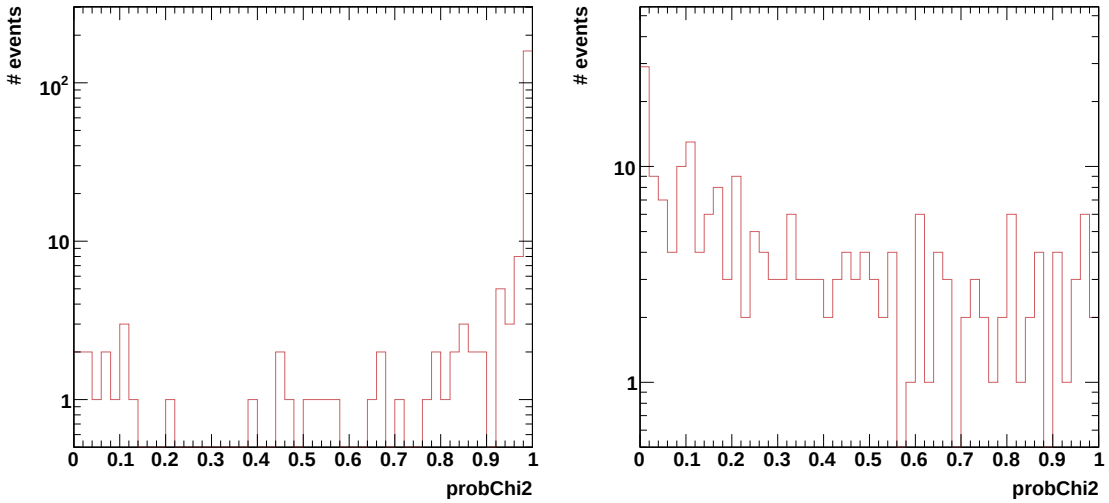



Figure 17: **PLACEHOLDERS** Distribution of the maximum fit probability in the range $-25\% \leq \Delta E_i \leq +25\%$ as described in the text (left) and the distribution of the fit probability when no residual jet corrections are introduced.

fixed value of $(\Delta E_{\bar{q}})_{min}$ (this would be equivalent with $\chi^2(\Delta E_b, (\Delta E_q)_{min}, \Delta E_{\bar{q}})$ given the constructed symmetry). From the first distribution the symmetry of the $\Delta E_q$ and the $\Delta E_{\bar{q}}$ dimensions is demonstrated. The second distribution illustrates the correlation between the residual correction for the bottom quark jet and the light quark jets from the W boson decay.

In order to estimate the final residual corrections the correlation between both corrections is neglected and the two dimensional distribution is further projected into one dimension as $\chi^2(\Delta E_b) = \chi^2(\Delta E_b, (\Delta E_q)_{min}, (\Delta E_{\bar{q}})_{min})$ and $\chi^2(\Delta E_q) = \chi^2(\Delta E_{\bar{q}}) = \chi^2((\Delta E_b)_{min}, \Delta E_q, (\Delta E_{\bar{q}})_{min})$. The parabolic tendency of the resulting $\chi^2$ distribution is shown in Figure 19. Given the very good agreement a parabol van be fitted analytically through three points around the minimum. From this the best estimated value and the uncertainty on the residual correction is obtained. For the reconstructed jets calibrated with the simulation based technique described, the residual corrections are

$$\Delta E_b = -0.3 \pm 1.2\%$$
$$\Delta E_q = -10.4 \pm 1.2\%$$

**THESE NUMBERS ARE HERE BY TAKING ONLY THE BEST JET COMBINATION USING MC INFORMATION AND APPLYING ONLY THE SELECTION CUTS ON THE TRANSVERSE ENERGY OR MOMENTUM OF JET AND MUON, HENCE JUST AN ILLUSTRATION OF THE VIABILITY OF THE METHOD**.

where the uncertainties reflects the expectation with a data set of 100pb$^{-1}$ of integrated luminosity. For the light quark jet energy scale correction the average is taken between the estimated $\Delta E_q$ and $\Delta E_{\bar{q}}$ corrections. The above numbers have to be compared with the expectation using the generation level information of the events. The applied mass constraints hold at the level of the primary partons, hence the energy of the reconstructed jet $E_{reco}$ (calibrated with methods based on simulation as explained in Section **??**) is compared with the energy of the matching primary parton $E_{parton}$. A matching criterium of $\Delta R \leq 0.3$ is applied. Using these matched parton-jet couples in the selected events, the true residual correction can be obtained from the distribution of $(E_{parton} - E_{reco})/E_{reco}$. This distribution is shown in Figure 20 for both the quarks in the W boson decay and the bottom quarks in the top quark decay. From a Gaussian fit of the peak of these distributions central expected values of $\Delta_{exp}E_b$ = -1.6 $\pm$ 1.1 % and $\Delta_{exp}E_q$ = -12.0 $\pm$ 0.9 % are obtained. Therefore it is concluded that the bias of these inclusive estimators are $E[\Delta_{exp}E_b - \Delta E_b]$ = -1 $\pm$ 2 % and $E[\Delta_{exp}E_q - \Delta E_q]$ = -2 $\pm$ 2 % respectively, therefore compatible with no bias.

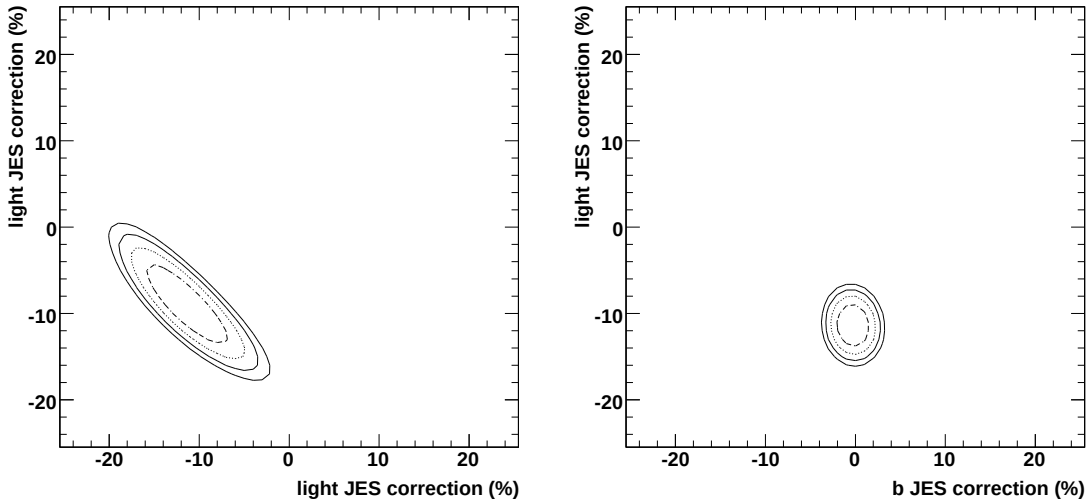



Figure 18: **PLACEHOLDERS** Distributions of $\chi^2((\Delta E_b)_{min}, \Delta E_q, \Delta E_{\bar{q}})$ at the fixed value of $(\Delta E_b)_{min}$ (left) and $\chi^2(\Delta E_b, \Delta E_q, (\Delta E_{\bar{q}})_{min})$ at the fixed value of $(\Delta E_{\bar{q}})_{min}$ (right). Both are represented as confidence intervals.

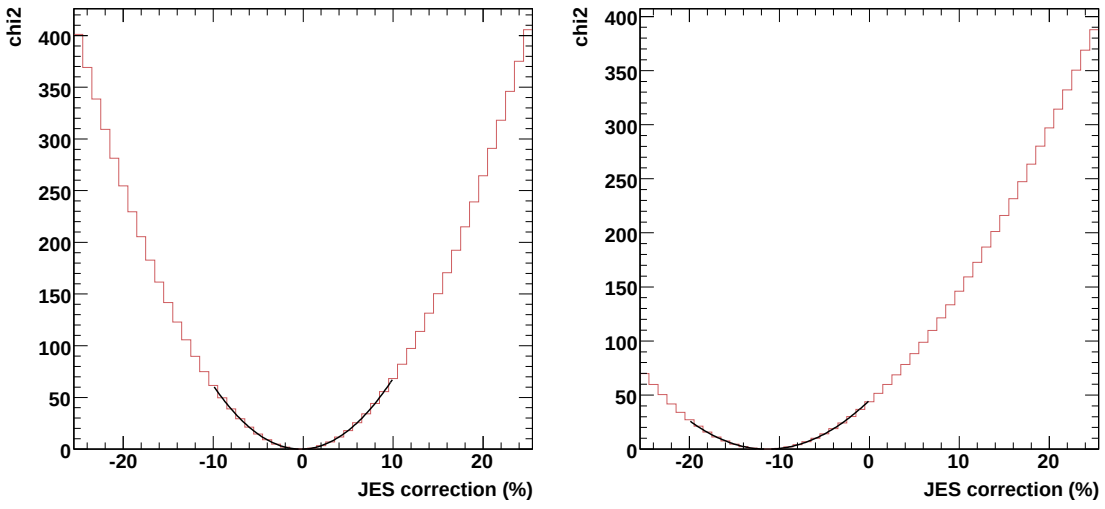



Figure 19: **PLACEHOLDERS** Distributions of $\chi^2(\Delta E_b)$ (left) and $\chi^2(\Delta E_q)$ (right) as described in the text.

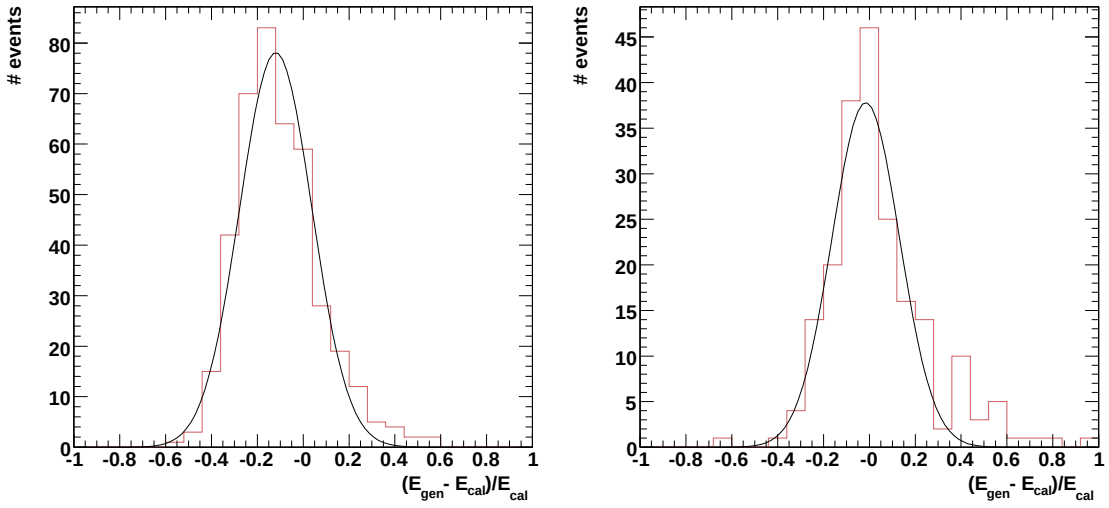



Figure 20: **PLACEHOLDERS** Distributions of $(E_{parton} - E_{reco})/E_{reco}$ for the jets arising from the light quarks in the W boson decay (left) and the jets arising from the bottom quarks in the top quark decay (right).

Resampling techniques are applied to check the consistency of the estimation of the quoted uncertainty. The pull distribution $(\Delta E^i_{b,q} - E[\Delta E_{b,q}])/\delta E^i_{b,q}$ from sampling over data sets $i \in \{1, ..., N\}$ each reflecting 100pb$^{-1}$ of integrated luminosity, is shown in Figure 21. **MORE**

# 8 Systematic uncertainties

Several systematic effects can influence the method described. They will be estimated applying FastSim tools. The main effect is expected to be the pile-up contribution although this could be considered as part of the residual jet energy scale correction to be estimated and differentiated versus the amount of reconstructed and identified primary vertices. **more to come**

# 9 Conclusion

Due to the relative large cross section of top quark pair processes at the LHC and the rather pure sample which is retrieved after the event selection, these top quark events serve excellent for the physics commissioning of the CMS experiment. With the large set of hadronic top quark decays reconstructed within the CMS detector in top quark pair events, is it demonstrated that residual jet energy scale corrections can be precisely estimated or a closure test of the jet energy scale procedure applied in CMS can be obtained. On an event-by-event a kinematic fit reveals a probability function in the dimensions of the residual jet energy scale corrections to agree with the W boson and top quark mass constraints present at quark level in the event. In this paper the inclusive residual jet energy scale corrections are estimated for both light quark and bottom quark jets on top of a jet calibration procedure based on simulated QCD events. The estimated absolute values of the residual corrections are is good agreement with the corrections expected when matching the reconstructed jet to the primary quark using the Monte Carlo information. Applying the method on a data set with an integrated luminosity of 100pb$^{-1}$ would result in uncertainties of xx% for light quark jets arising from the W boson decay and xx% for the bottom quark jets in the top quark decay. The method is expected to become dominated by systematic uncertainties from an integrated luminosity of xx pb$^{-1}$.

# References

[1] Theory prediction of cross section NLO.

[2] Recent Electro-Weak combination paper.

[3] AlpGen reference.

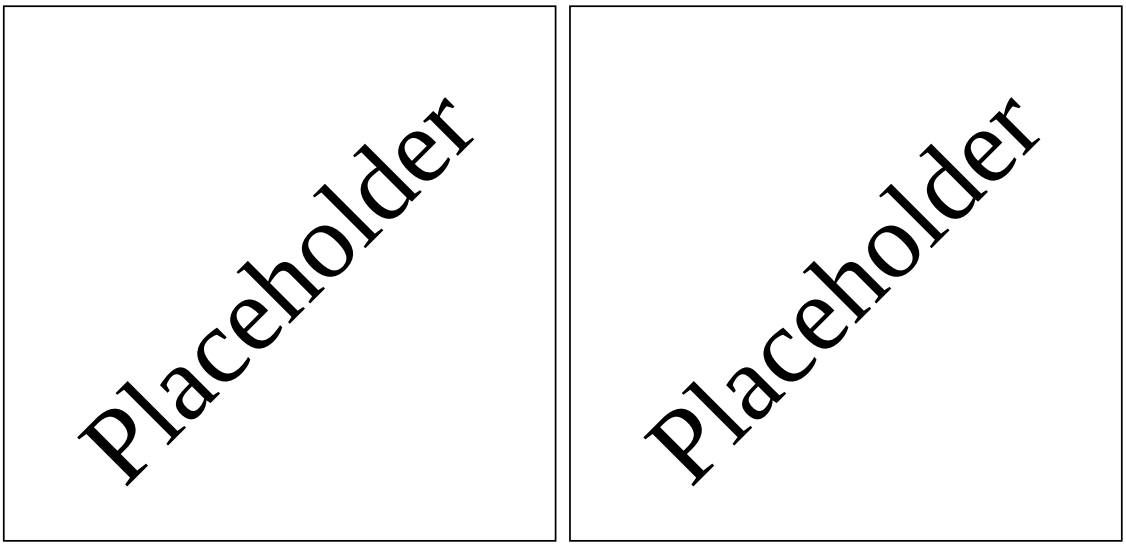



Figure 21: **PLACEHOLDERS** Distributions of the pull of the $\Delta E_b$ estimator (left) and the $\Delta E_q$ estimator (right). The distributions are fitted with a Gaussian function.

[4] FastSim reference (website ?).

[5] Trigger reference.

[6] CMSSW reference (website ?).

[7] TQAF reference (website ?).

[8] J. D'Hondt *et al.*, "Fitting of Event Topologies with External Kinematic Constraints in CMS", CMS Note 2006/023.