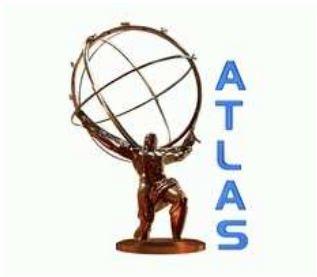
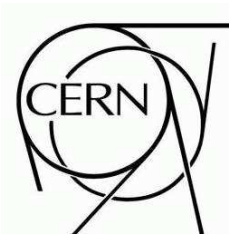


ATLAS NOTE



January 29, 2008

Jet-Vertex Association Algorithm

D.W. Miller[1], A. Schwartzman[1], D. Su[1]

[1]Stanford University/SLAC, Menlo Park, California 94025, USA

Abstract

This note describes methods developed for jet-vertex association and hard-scatter jet selection in events with multiple interactions. By combining tracks and their primary vertices with calorimeter jets we define a discriminant which measures the probability that a jet originated from a particular vertex. Jet selection based on this discriminant is shown to be insensitive to the contributions from simultaneous uncorrelated soft collisions that occur during pile-up. Performance of this algorithm is demonstrated in events at low luminosity (10^{33} cm^{-2}s^{-1}) for a variety of physics processes. Furthermore, the prospects for a jet-by-jet energy offset correction are discussed which has the potential to improve jet energy resolution in events with multiple interactions. Finally, improvements to the primary vertex identification from jet-vertex association are described. These methods are compatible with both calorimeter jets matched to tracks as well as track-jets, part of a new program of physics-jets intended to further integrate tracking information with calorimeter jets.

1 Introduction

Events containing jets from minimum bias interactions at the LHC will introduce challenges for hard-scatter jet identification, jet energy and missing energy ($\not\!\!E_T$) resolution. It is essential that physics analyses dependent on event jet multiplicity and jet energies be able to disentangle this background. In addition, removing jet energy contributions to the event from parasitic interactions will improve both the jet energy resolution and $\not\!\!E_T$ calculation.

The initial physics program of the LHC will not reach average luminosities at or much above the current Tevatron values, with target luminosities at approximately 10^{31} cm^{-2} s^{-1} [1]. However, the number of additional interactions is a function of individual bunch charge as well as the bunch spacing and intensity, which is expected to have unique parameters during the early years of machine running [1]. The possibility of additional diffractive (single or double) and inelastic proton-proton interactions at 14 TeV center-of-mass energy resulting in jet production is therefore possible even at these low luminosities. Any jet production in these events constitutes a (generally low-p_T) reducible background to jets originating in a hard-scattering interaction during the same bunch crossing.

The ATLAS tracking system (the Inner Detector, or ID) provides a very precise tool for understanding the composition of calorimeter jets and for reducing this background. Charged particle jet constituents that leave tracks in the ID provide 3-dimensional information (as opposed to only detector η and ϕ) on the jet origin and direction as a result of the vertexing provided by the tracks. The combination of tracking and calorimetry can therefore greatly enhance the identification and selection of jets.

The algorithm described in this note allows for the reduction of multiple interaction jet backgrounds thereby reducing the combinatorial background for multi-jet events without raising the nominal jet p_T threshold. Instead, by utilizing tracking information for jets (when available) it is possible to define a discriminant for each jet with respect to each identified primary vertex (PV) in the event. Once the hard-scattering PV is selected, the jet-vertex fraction, or JVF, maybe be used to select jets having a high likelihood of originating in that vertex.

2 Algorithm Description

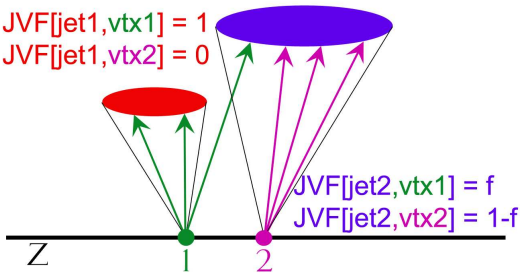
The jet-vertex association algorithm requires the collections of reconstructed tracks, jets and primary vertices in an event and is comprised of three distinct steps, (i) track and jet selection (ii) calorimeter jet-to-track matching and (iii) jet-vertex association. All tracks in the event must pass a set of quality criteria and are then matched to cal-jets which lie within the fiducial tracking volume ($|\eta| \leq 2.5$, see Section 3). Any track falling within a ΔR cone of a jet axis is associated to that jet. Each track matched to a jet is then required to have originated in a reconstructed PV in the event. Although no explicit secondary vertex (SV) finding is performed, all tracks which pass the selection criteria, including SV tracks, are considered by the algorithm. The combination of these steps results in the jet-vertex association.

The resulting discriminant is termed the jet-vertex fraction (JVF) and is defined for each jet with respect to each PV. JVF is the fraction of each jet's constituent transverse track-momentum contributed by each vertex. More specifically, it is the sum p_T of all matched-tracks from a given vertex divided by the total jet-matched track p_T. Formally, for a single jet jet_i the JVF with

respect to the vertex vtx_j in the event is written

$$JVF(\text{jet}_i, \text{vtx}_j) = \frac{\sum_k p_T(\text{trk}_k^{\text{jet}_i}, \text{vtx}_j)}{\sum_n \sum_l p_T(\text{trk}_l^{\text{jet}_i}, \text{vtx}_n)}, \quad (1)$$

where, jet i has a fraction $JVF(i, j)$ of it's total matched-track momentum originating in vertex j. Calorimeter jets which fall outside of the fiducial tracking region (see Section 3) or which have not been matched to tracks are assigned a $JVF = -1$. The event topology of interest is shown schematically in Figure 2 and described by Eq.(1).



3 Event Selection

The three primary steps involved in the jet-vertex association involve event and object selection criteria. These criteria have been optimized and studied first in single lepton $t\bar{t}$ events without additional interactions and then compared to the same events where additional minimum bias (MB) interactions were overlaid with a Poisson distribution per event (see Section 4.1 for a further discussion of the pile-up production mechanisms).

Tracks and jets are selected based on detector efficiency considerations, kinematics and quality criteria. Calorimeter jets seeded from towers with a cone size of 0.4 are used throughout the analysis. Tracking information is only valid in the fiducial tracking volume, where the track reconstruction efficiency drops significantly at large η and so tracks and jets are required to have $|\eta_{trk}| \leq 2.5$ and $|\eta_{jet}| \leq 2.0$. Transverse momentum (p_T) requirements are kept low for both jets ($p_T \geq 10.0$ GeV) and tracks ($p_T \geq 0.400$ GeV) in order to study algorithm efficiency at these values, although tracks with $p_T \geq 0.800$ GeV have been shown to improve jet-selection. The track selection criteria are listed in Table 1 below.

$\lvert z_0 \rvert$	≤ 200 mm
$\lvert \eta \rvert$	≤ 2.5
χ^2/ndf	≤ 3.0
Si hits	≥ 7 (Pix + SCT)
Pixel hits	$1 \leq N_{Pixel} \leq 5$
SCT hits	$5 \leq N_{SCT} \leq 20$
TRT hits	$1 \leq N_{TRT} \leq 40$

Table 1: Track selection criteria for all tracks considered for jet-association.

A simple lepton-jet overlap removal is also implemented such that jets matched to true electrons with $p_T^{true} \geq 10$ GeV and within $\Delta R_{jete} \leq 0.4$ are not considered valid jets for the algorithm.

4 Event Reconstruction and Topology

4.1 Pile-up event simulation

In order to simulate the effect of higher luminosities, ATLAS has taken the approach of overlaying additional interactions onto a given signal event during the digitization stage. Events of each type to be included are simulated independently, but necessarily with the same detector geometry. The overlaid events may include either or both MB and cavern background events (see Section 4.1.1 below). The actual electronics integration times of subdetectors in ATLAS can be significantly longer than a single bunch-crossing, and so a time-window is set in the digitization such that the expected signal and noise for each event are integrated over several bunch crossings. The time window is dependent on each subdetector. For example, the ATLAS Monitored Drift Tubes (MDT's) require an overall integration time window from -36 to +32 bunch-crossings (because of the ∼400-1200 ns drift time) [2], where the reference is the simulated signal event. Thus, "pile-up" correctly refers to the both the in-time MB and cavern background interactions (same bunch crossing as the signal interaction) as well as those out-of time (all others), although the sensitivity to these events is not identical for all subdetectors. All events simulated here assume an LHC bunch spacing of 25 ns, which need not be the case for the initial beam conditions.

4.1.1 Minimum bias events

For each bunch crossing in the integration window, including the signal event at $t = 0$, a Poisson distribution of MB events with a specified mean is digitized and reconstructed. The new Pythia model in PYTHIA 6.323 [3] for the underlying event with multiple parton-parton interactions is used where the allowed processes correspond to the total hadron-hadron cross-section for elastic, single diffractive, double diffractive, and non-diffractive events,

$$\sigma_{tot} = \sigma_{elas} + \sigma_{sd} + \sigma_{dd} + \sigma_{nd}.$$

Initial and final state radiation (ISR and FSR) are also included. Further details on the structure of the MB interactions can be found in [4]. By using the full hadron-hadron cross-section to generate the MB interactions, the pile-up simulation can therefore pick MB events from a pool of events at random without needing to fold-in cross-sections (since this is effectively done in the generation by Pythia). This also implies that the full spectrum of possible proton-proton events generated by Pythia will occur, weighted accordingly, and this will be an important consideration in assessing the results.

4.1.2 Cavern background events

Cavern background events consist mainly of thermalised slow neutrons, long-lived K^0's and low-energy photons escaping the calorimeters and the forward beam and shielding elements [5].

High cavern-background rates degrade the performance of the Muon System, yet do not significantly affect the ID or calorimeters. These events are inserted into each bunch-crossing at a constant rate depending on the luminosity, and at a random position within the experimental cavern. A nominal cavern background rate is then scaled by a "safety factor" (SF) which is recorded in the dataset name. Such a scaling factor is *not* applied to the MB events. This is done because cavern the background rates are not precisely known and to be able to test trigger and reconstruction software in more challenging environments.

4.2 Monte Carlo Samples

Several Monte Carlo datasets have been used to study events with pile-up (Sec. 4.3) and performance of the jet-vertex association and selection algorithm for various physics processes (Sec. 5). In particular, three independent datasets corresponding to $t\bar{t}$ events with and without MB pile-up and the individual MB events themselves are used to understand pile-up event topology and reconstruction. $t\bar{t}$ events with semileptonic decays are chosen as the signal process due to the presence of jets, its significance as an interesting physics channel, and for the expected improvements to be gained in jet selection and primary vertex identification (see Sec. 6.2). Comparisons are made between individual (i.e. no pile-up) $t\bar{t}$ events, $t\bar{t}$ overlaid with MB interactions and the individual MB events themselves in Figs. 1- 4. In addition to these $t\bar{t}$ events, 3 other physics processes of interest are investigated for the possible improvements in jet identification and selection as well as primary vertex selection by using the JVF discriminant.

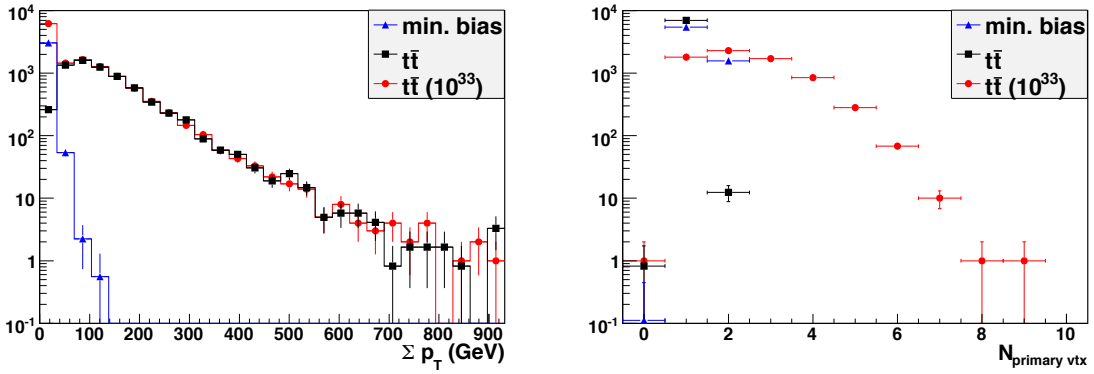
These datasets have all been centrally produced in the ATLAS production system and reconstructed using Release 12.0.5.2 or 12.0.6.X with details listed in Table 4 below. Note that the truth information for jets is inaccurate prior to Release 13.0.30. The di-jet sample used (J6) requires that the two leading jets have $560 \leq p_T \leq 1120$ GeV at the generator level, making this sample stand-out clearly above the MB background. Further studies are underway to investigate di-jet samples at lower jet p_T with pile-up.

Physics process	Luminosity ($\text{cm}^{-2}\ \text{sec}^{-1}$)	Dataset #	# events	Reconstruction release
MB (Sec. 4.1.1)	zero	5001	63K	12.0.6.1
$t\bar{t}$	zero	5200	8.5K	12.0.5.2
$t\bar{t}$	10^{33} or $\langle N_{MB}\rangle$=2.3	5200	7K	12.0.6.5
$H \rightarrow \gamma\gamma$	10^{33} or $\langle N_{MB}\rangle$=2.3	5316	3.7K	12.0.6.5
Di-jet (J6)	10^{33} or $\langle N_{MB}\rangle$=2.3	5015	9.2K	12.0.6.5
$t\bar{t}$ + jets	10^{33} or $\langle N_{MB}\rangle$=2.3	6371	8.5K	12.0.6.5

Table 2: Datasets used in investigating the event topology of pile-up events at ATLAS and in evaluating performance of the jet-vertex association algorithm. For full dataset names and the list of MB and cavern background files used for generating the pile-up events see Appendix A.

4.3 Event topology

Events with pile-up introduce higher charged particle multiplicities and therefore correspondingly higher track multiplicities and primary vertices. It is crucial for the performance of the



(a) Primary vertex Σp_T

(b) Primary vertex multiplicity

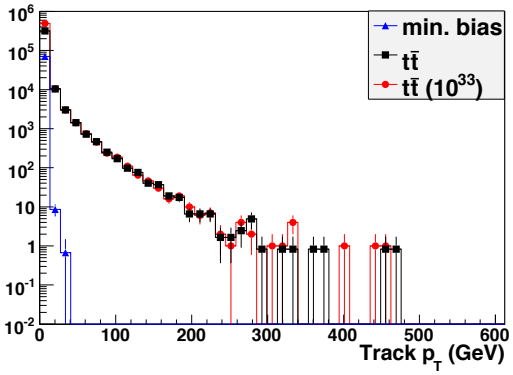
Figure 1: Primary vertex kinematics and multiplicity for single minimum bias, $t\bar{t}$ and $t\bar{t}$ with pile-up events, normalized to number of events.

jet-vertex association algorithm that the reconstruction and event topology be understood. In particular, the kinematic and multiplicity distributions for vertices, tracks, and jets are discussed.

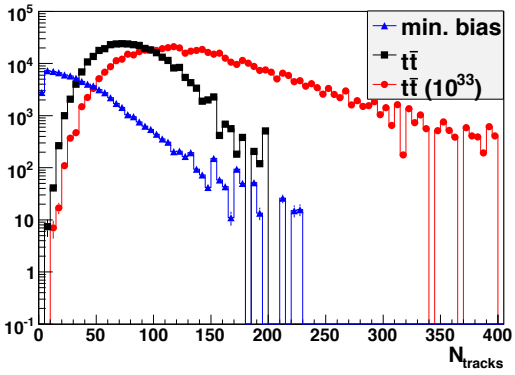
MB interactions contribute primarily low-Σp_T vertices as shown in Figure 1. These additional interactions increase the reconstructed PV multiplicity (Fig. 1(b)). A small fraction of the $t\bar{t}$ events also have multiple reconstructed vertices for two reasons. First, there is a non-negligible fraction of true displaced SV's that are still within the envelope considered by the primary vertexing algorithm to be good vertices and are thus considered as PV candidates. Secondly, there is an issue of "split vertices", in which a single vertex is reconstructed as multiple due to outlying tracks. In large part, these have been removed from the event by considering the χ^2/ndf of the vertex fit, attached track multiplicity and distance to higher Σp_T vertices. These result in a small source of inefficiency when a jet is matched to a track which has been fit to one of these split vertices (see Section 5).

All tracks must pass the selection criteria described in Section 3 and Table 1. Following these cuts, the reconstructed track p_T and multiplicity are shown in Figure 2 for $t\bar{t}$ and MB events. A minimum track p_T cut of 0.400 GeV is applied for these distributions in order to show the high track multiplicities at low-p_T, even in MB events. For the algorithm performance studies that follow, a higher p_T cut is used in order to improve jet-vertex association efficiency and reduce fake track contributions. The very long tail in track multiplicity for events with pile-up is due to the many low-p_T tracks in multiple MB interactions.

In Figure 3 are the reconstructed jet (cone, R=0.4) p_T and multiplicity for a low minimum p_T cut of 10 GeV. Individual MB interactions contribute comparatively fewer jets than a high-p_T event such as $t\bar{t}$, as expected. However, not only are multiple MB interactions present in each bunch crossing (here, $\langle N_{MB} \rangle$=2.3) but MB can also give you very soft jets which overlay with noise, or a soft gluon in the event. These effects can then also introduce additional jets into the event that will appear to be "min-bias jets." In Fig. 3(c) we indeed see that as the number of additional vertices in the event increase, the reconstructed jet multiplicity increases correspondingly. Even after increasing the minimum jet p_T cut from 10 to 40 GeV a noticeable increase remains. Correcting for this effect is a primary goal of this algorithm and is described in Sections 5.2, 5.3.

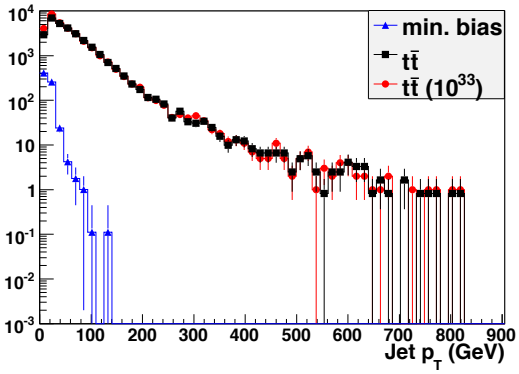


(a) Reconstructed track p_T

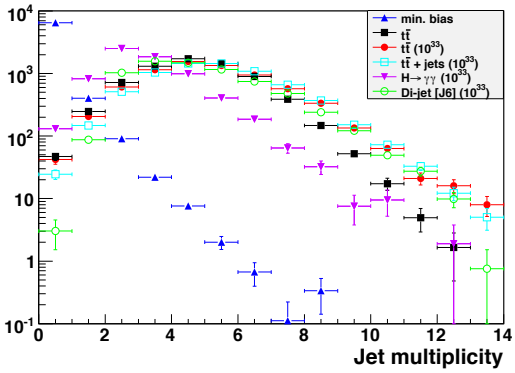


(b) Reconstructed track multiplicity

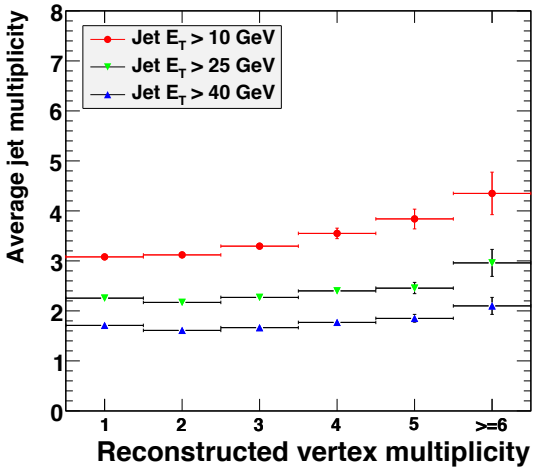
Figure 2: Track kinematics and multiplicity for single minimum bias, $t\bar{t}$ and $t\bar{t}$ with pile-up events, normalized to number of events.



(a) Reconstructed jet p_T

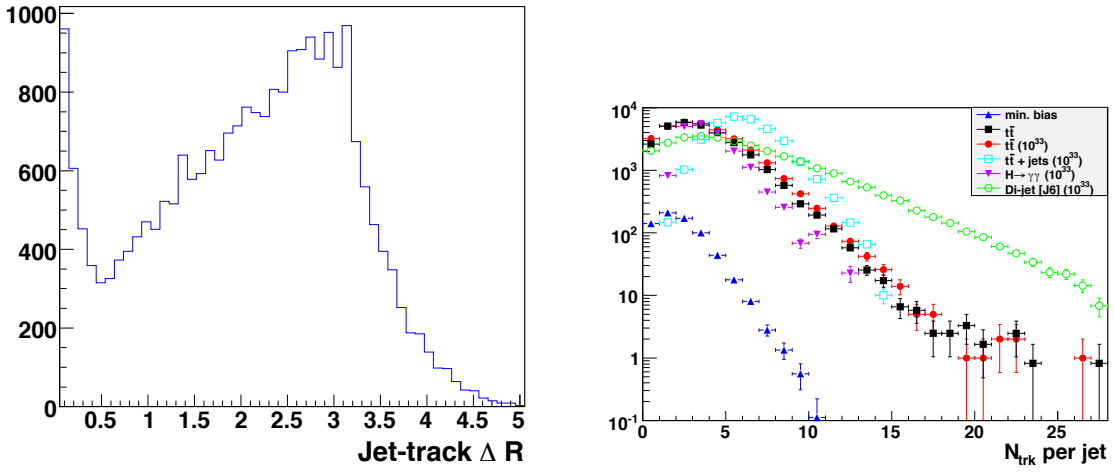


(b) Jet multiplicity



(c) Jet multiplicity vs. vertex multiplicity for $p_T^{\text{jet}} \geq 10, 25, 40$ GeV)

Figure 3: Jet kinematics and multiplicity for single minimum bias, $t\bar{t}$ and $t\bar{t}$ with pile-up events, normalized to number of events.



(a) ΔR between jets and tracks

(b) Matched track (jet-track) multiplicity per jet

Figure 4: The criterion for jet-track matching in the algorithm uses simple ΔR association. Shown in Fig. 4(a) is the separation between jets and all tracks in $t\bar{t}$ events without pile-up. A maximum $\Delta R = 0.4$ is used. Fig. 4(b) depicts the raw track multiplicity (criteria from Table 1 are not applied) for single minimum bias, $t\bar{t}$ and $t\bar{t}$ with pile-up events, normalized to number of events. The shape of the jet-track multiplicity is very similar for $t\bar{t}$ and MB jets at low N_{trk}.

5 Jet-Vertex Association Performance

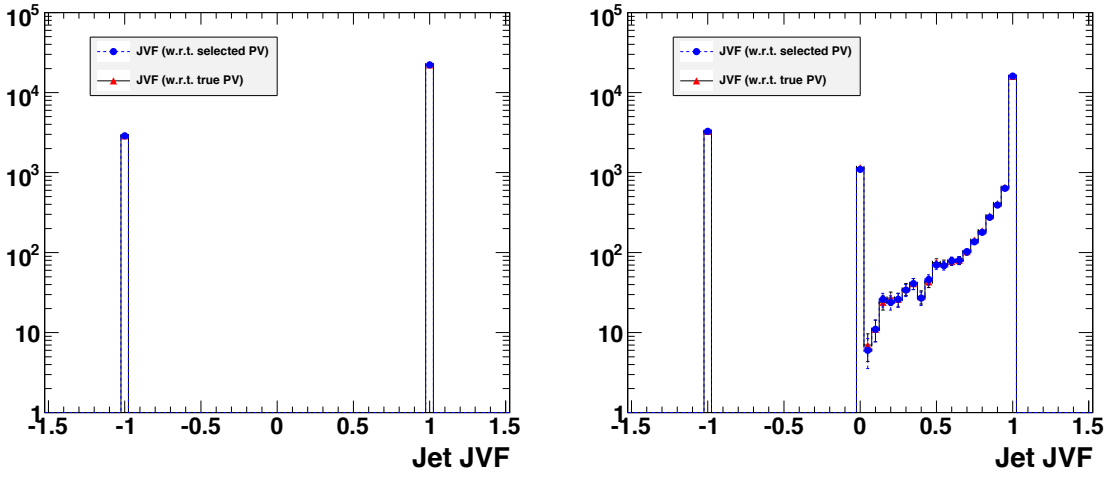
5.1 Jet-track association

Jets are first associated to tracks via a simple ΔR matching in the $\eta - \phi$ plane. The track parameters used for the matching are calculated at the distance of closest approach (DCA) to $(x, y, z) = (0, 0, 0)$ which is the default in ATHENA Release 12.0.6 [5]. Performing the matching using track parameters calculated at the origin is justified because the goal is to associate tracks to jets that were produced in the same PV, regardless of how the magnetic field affects their trajectories.

Using DCA parameters, tracks are matched to each jet with $\Delta R(\text{track,jet}) \leq 4$ (the so-called "jet-tracks"), as in Fig. 4, and used to calculate the fraction of track momemtum from each PV. The mean of the jet-track distribution is low compared to [6], for example, because each track is required to be used in a PV fit. Following the jet-track selection for each jet, no further tracks are considered by the algorithm.

5.2 Jet selection

Given the list of tracks matched to a jet and their associated vertices, the JVF can be calculated using Eq. 1. For all events that follow, the primary vertex selection was determined using criterion 4 of Section 6.2 due to it's lower vertex mis-ID rate for di-jet and $t\bar{t}$ events. The distributions of jet JVF for events with and without additional interactions are indicative of the underlying event topologies. Shown in Figure 5 are the JVF spectra for $t\bar{t}$ simulated with (Fig. 5(b), at $\langle N_{MB} \rangle$=2.3) and without (Fig. 5(a)) pile-up. Jets without associated tracking information are assigned a value of $JVF = -1$. Jets with tracking information obtain $0 \leq JVF \leq 1$,



(a) No pile-up

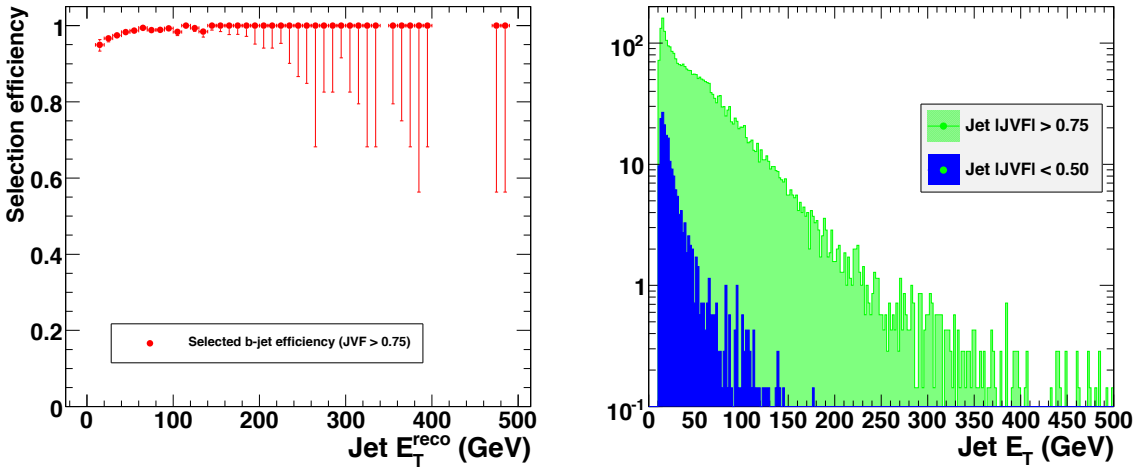
(b) With pile-up ($\langle N_{MB} \rangle$=2.3)

Figure 5: Distribution of the JVF discriminant variable for $t\bar{t}$ events without [5(a)] and with pile-up [5(b)]. The solid curve (▲'s) corresponds to JVF calculated using truth track association to the true PV of the event while the dashed line (■'s) represents JVF calculated using only reconstructed quantities; these two distributions overlay very closely, indicating well reconstructed jet JVF.

where a jet with all of its matched track-momentum originating in the selected PV will acquire $JVF = 1$. For the pile-up sample, jets with $JVF = 0$ originate in vertices not selected as the PV for that event. These jets, as well as those with large contributions from MB interactions in the event ($0 \leq JVF \leq 0.5$) constitute a non-negligible fraction of all jets and can be rejected using JVF. Evidence that these MB jets are not mostly spurious fakes or due to mis-associations in the algorithm is demonstrated by incrementally raising the jet p_T threshold and seeing a corresponding decrease in number of $JVF = 0$ jets. Additionally, track truth information indicates unambiguously whether a track matched to a given jet originates in a MB or signal vertex.

Jet backgrounds due to additional interactions, although small, can still be reduced by selecting jets based on their JVF. The selection efficiency of hard-scattering jets is important in assessing the performance of the jet-vertex association algorithm. By considering only jets matched to Monte Carlo truth B-hadrons from the primary interaction in $t\bar{t}$ events with well separated PV's ($min(\Delta Z) \geq 2$ cm), we can compare the number of jets selected using JVF to the total number of jets matched to B-hadrons. This B-jet selection efficiency is shown in Figure 6 for jet $JVF \geq 0.75$ and a low jet-$p_T \geq 10$ GeV cut. Also shown are the p_T distributions of the selected and rejected jets, demonstrating the expected low p_T spectrum for MB QCD jets. For this cut we obtain a total selection efficiency of 98.5% over the full jet p_T range above 10 GeV. Note that this is the efficiency to select *all* jets matched to B-hadrons in each event, not simply the n leading p_T jets in the event.

There are 31 jets in Figure 6(b) with $JVF \leq 0.5$ and $p_T \geq 100$ GeV whose origins must be explained. 17 of these jets are matched to a single track originating in a true MB interaction, 11 of which have track $p_T \geq 1$ GeV. 11 jets have 2-7 constituent tracks each with at least 40% of originating in a true MB interaction. These two groups constitute correctly identified jets with large MB contributions. The remaining 3 jets have either (i) been matched to a track which is



(a) B-jet selection efficiency

(b) Jet E_T of all JVF-selected jets

Figure 6: Efficiency is shown in 6(a) for selecting hard-scatter b-jets (matched to B-hadrons) using JVF. The total selection efficiency is 98.5% over the full b-jet p_T range above 10 GeV. The jet p_T distribution for the selected jets is shown in 6(b).

used in a fit to a split-vertex or (ii) been overlaid with fake tracks (the case for 1 jet with 4 tracks including 1 fake track). These three cases account for all of the 31 jets with $|JVF| \leq 0.5$ and $p_T \geq 100$ GeV.

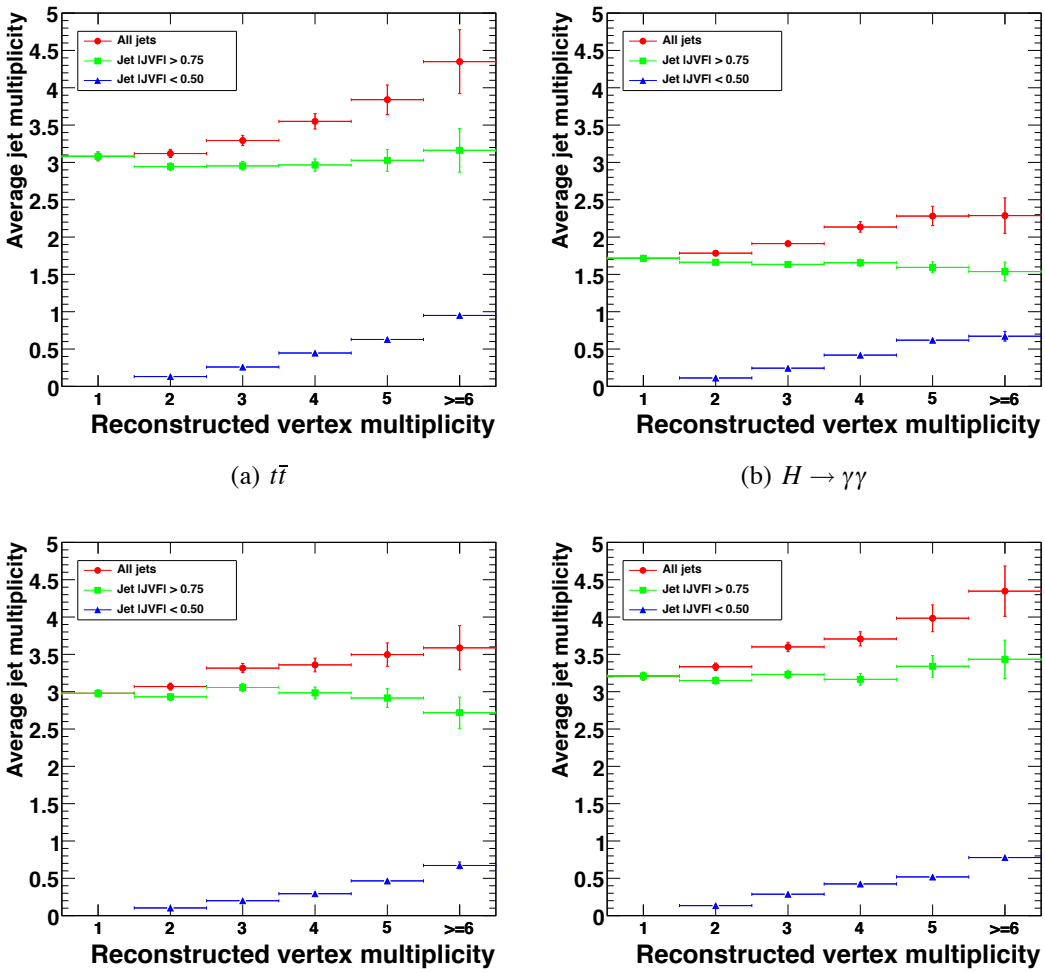
5.3 Jet multiplicity with additional interactions

Of primary importance for physics analyses is that jet selection be insensitive to the contributions from the simultaneous uncorrelated soft collisions that occur with pile-up interactions. Using the JVF discriminant with a suitable cut, the required insensitivity is recovered. Figure 7 shows the jet multiplicity distributions for $t\bar{t}$, $H \rightarrow \gamma\gamma$ ($m_H = 120$ GeV), $t\bar{t}$, di-jet (J6) and $W \rightarrow$jets events with pile-up ($\langle N_{MB} \rangle$=2.3). A flat distribution is recovered when jets with tracking information are required to have have $|JVF| \geq 0.75$ (green squares) compared to that with no selection applied (red circles). This result clearly highlights the importance of incorporating a jet-vertex selection algorithm in analyses which utilize jet counting to measure physical quantities. The MB interactions which contribute soft but non-negligible jets to the event must be identified and removed in order to maintain the high precision goal of measurements in ATLAS.

6 Further applications of jet-vertex association

6.1 Jet-by-jet energy offset correction using JVF

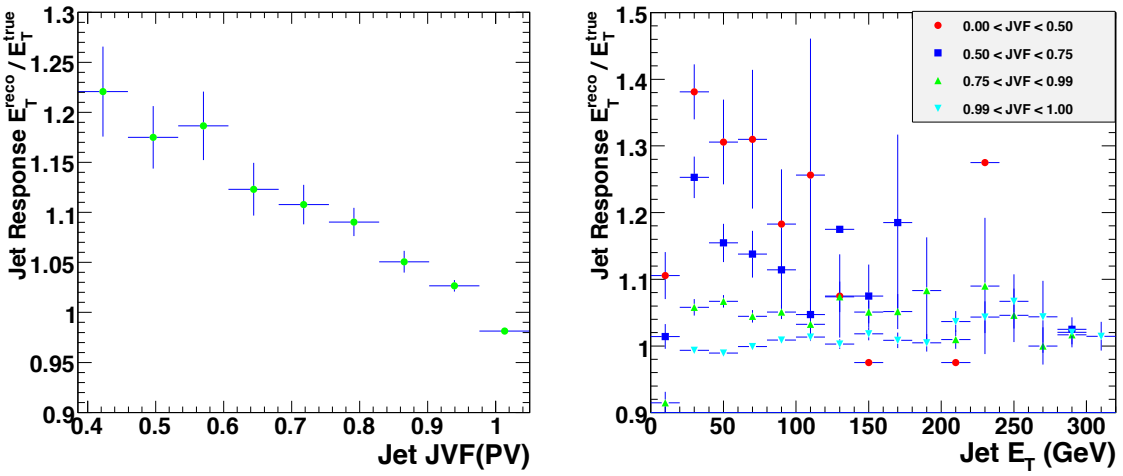
Standard methods have been developed both at CDF [7] and DØ [8] for correcting jet energies due to multiple interactions at various instantaneous luminosities. However, these methods take a simplified statistical approach and apply a constant offset correction to all jets based on studies of average energy depositions in the calorimeter during minimum or zero bias triggers.



(a) $t\bar{t}$ (b) $H \rightarrow \gamma\gamma$

(c) Di-jet (J6) (d) $t\bar{t}, W \rightarrow$jets

Figure 7: Jet multiplicity as a function of reconstructed vertex multiplicity with and without JVF selection. Even with a low $p_T \geq 10$ GeV cut applied, jet selection using JVF recovers the expected constant hard-scatter jet multiplicity for $t\bar{t}$, $H \rightarrow \gamma\gamma$ (m_H=120 GeV), di-jet (J6), and $t\bar{t}$, $W \rightarrow$jets events. "All jets" refers to all jets passing the selection criteria described in Sec. 3 and having $p_T \geq 10$, meaning that no track-association criteria are required. For JVF-selected jets, matched tracks are required to contribute $\geq$75% of the total track p_T from the selected primary vertex. Jets without tracking information ($JVF = -1$) are also included as selected jets.



(a) Jet response as a function of jet JVF

(b) Jet response as a function of jet E_T for JVF bins

Figure 8: Jet response $E_T^{\text{reco}}/E_T^{\text{true}}$ as a function of jet JVF and E_T.

This method can be improved by considering the track momentum contributed to each jet in the event from additional interactions, which is precisely what the JVF measures.

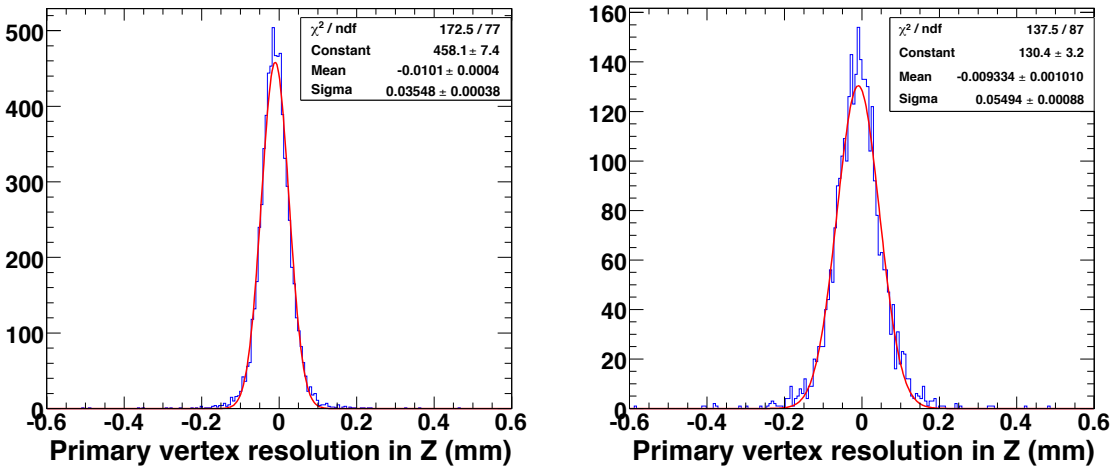
Using $t\bar{t}$ events reconstructed in Release 13.0.30.3 with $\langle N_{MB}\rangle$=2.3 (or a nominal luminosity of 10^{33}) we can make use of the truth jet collections which do not include the contributions from pile-up interactions. By matching reconstructed jets to truth jets (built without true pile-up particles) and requiring that $\Delta R(\text{reco}, \text{true}) \leq 0.4$ we can measure the jet response, defined as $R = E_T^{\text{reco}}/E_T^{\text{true}}$. Shown in Figure 8, is the jet response as a function of jet JVF. There is a clear and strong dependence on the jet JVF, indicating that a correction as a function of both jet E_T and JVF is possible. Such a correction will also clearly improve the $\not{E}_T$ resolution as well.

6.2 Jet-Based Primary Vertex Selection

Jet-vertex association, as defined by the JVF, can also be used to select vertices which have large contributions to jets in the event, and therefore aid in the selection of the primary hard-scattering vertex. In effect, this is taking into account what is seen in Figure 3(b), where hard-scattering interactions are more likely to give rise to jets in the event. Conventional vertex selection using $\sum p_T^2$ can benefit from the addition of jet topology information for certain physics processes.

As a very simple example of the type of improvement that can be gained by including jet production information in the selection of the PV we consider 4 criteria:

1. $\sum p_T^2$
2. $\sum p_T^2 \times \sum JVF$
3. $\sum p_T^2 \times \sum JVF/N_{trk}$
4. $\sum p_T^2 \times \sum \overline{JVF}/N_{trk}$



(a) $t\bar{t}$ (10^{33}) PV resolution in Z (b) $H \rightarrow \gamma\gamma$ (m_H=120 GeV) (10^{33}) PV resolution in Z

Figure 9: The PV resolutions for $t\bar{t}$ and $H \rightarrow \gamma\gamma$ (m_H=120 GeV) events with pileup. Shown here are the Z resolutions, whereas the X and Y resolutions are found to be 0.010 mm and 0.013 mm for $t\bar{t}$ and $H \rightarrow \gamma\gamma$ (m_H=120 GeV), respectively.

In each case, the candidate PV with the maximum value of that variable is selected as the PV in the event. Vertex $\sum p_T^2$ is the sum of all track p_T^2 for each track used in the fit to the vertex. The other two derived quantities, $\sum JVF$ and $\sum \overline{JVF}$, are a sum and weighted sum (by jet p_T) of the jet JVF for all jets associated to each vertex. Formally, these can be written as

$$\sum JVF(\mathrm{vtx}_j) = \sum_{\mathrm{jet}_i} JVF(\mathrm{jet}_i, \mathrm{vtx}_j) = \sum_{\mathrm{jet}_i} \frac{\sum_k p_T(\mathrm{trk}_k^{\mathrm{jet}_i}, \mathrm{vtx}_j)}{\sum_n \sum_l p_T(\mathrm{trk}_l^{\mathrm{jet}_i}, \mathrm{vtx}_n)}. \quad (2)$$

$$\sum \overline{JVF}(\mathrm{vtx}_j) = \sum_{\mathrm{jet}_i} JVF(\mathrm{jet}_i, \mathrm{vtx}_j) \times p_T(\mathrm{jet}_i) = \sum_{\mathrm{jet}_i} \frac{\sum_k p_T(\mathrm{trk}_k^{\mathrm{jet}_i}, \mathrm{vtx}_j)}{\sum_n \sum_l p_T(\mathrm{trk}_l^{\mathrm{jet}_i}, \mathrm{vtx}_n)} \times p_T(\mathrm{jet}_i). \quad (3)$$

In order to see the effect of scaling out the track multiplicity, we also divide by the number of attached tracks for each candidate PV. The selection criterion of Eq. 2 effectively measures the jet multiplicity contribution of vtx_j to the event, whereas Eq. 3 measures the jet energy contribution of vtx_j. Using these criteria, the selection efficiency and mis-identification (mis-ID) rates can be determined. We define the selection efficiency as the fraction of selected primary vertices which have been reconstructed within $\pm 300\mu$m of the true PV in Z (following the criterion used in [9]). For reference, the resolutions obtained for events at 10^{33} are shown in Figure 9, where the default vertexing in Release 12 is used (including a beam constraint). A vertex is considered mis-identified when the selected reconstructed vertex is closer to a pile-up MB truth vertex instead of the true PV.

Listed in Table 3 are the efficiencies and mis-ID rates (as %) for the Monte Carlo samples generated and for each of the PV selection criteria listed above. Keep in mind that this is a very simple example of including jet-vertex association information in the PV selection criteria, but already a slight improvement (at the per-mil level) can be seen for $t\bar{t}$ events with pileup. A more refined selection criteria can make use of a likelihood that combines these quantities in

Physics process		$\sum p_T^2$	$\sum p_T^2 \sum JVF$	$\sum p_T^2 \sum JVF/N_{trk}$	$\sum p_T^2 \sum \overline{JVF}/N_{trk}$
$t\bar{t}$	Sel. Eff %	99.89	99.89	99.89	99.89
	Mis ID %	0	0	0	0
$t\bar{t}$ (10^{33})	Sel. Eff %	99.46	98.24	99.54	99.57
	Mis ID %	0.36	1.57	0.26	0.21
$H \rightarrow \gamma\gamma$ (10^{33})	Sel. Eff %	92.42	81.45	85.20	84.66
	Mis ID %	5.52	16.63	12.72	13.35
Di-jet (J6) (10^{33})	Sel. Eff %	99.62	98.99	99.62	99.63
	Mis ID %	0.03	0.65	0.02	0.01
$t\bar{t}, W \rightarrow$jets (10^{33})	Sel. Eff %	99.75	98.82	99.68	99.7
	Mis ID %	0.11	1.04	0.18	0.19

Table 3: Primary vertex selection and mis-ID efficiencies for various selection criteria. For quantities utilizing jet JVF a $p_T^{\text{jet}} \geq 30$ GeV cut has been applied.

a more effective way. Note also that PV selection in $H \rightarrow \gamma\gamma$ (m_H=120 GeV) events does not benefit from the jet-vertex association, showing that this method will need to be optimized for each physics process. We also note that the split vertex removal mentioned in Section 4.3 only takes into account the $\sum p_T^2$ and ΔZ of vertices, so that reconstructed MB vertices which have a higher $\sum p_T^2$ than the hard-scatter PV will dominate. An improvement is expected if either the vertexing algorithm prevents such split vertices (and no merging is necessary) or the split-vertex removal takes into account the $\sum JVF$.

7 Summary

A method for jet-vertex association and jet-identification in events with multiple interactions has been developed for use in ATLAS. The vertex association is demonstrated and appears to be robust, resulting in a high hard-scattering jet selection efficiency for low luminosities. Use of the full jet-vertex information improves the selection efficiency for true hard-scattering vertices in events where multiple primary vertex candidates are reconstructed. Finally, we have shown a clear indication that a jet-by-jet offset energy correction for pile-up contributions is possible. The algorithm is currently implemented in Athena and progress towards providing an AlgTool within an upcoming Athena release is underway.

8 Acknowledgements

I am indebted to both my advisor Su Dong and Ariel Schwartzman for initiating this study and suggesting that ATLAS investigate the use of such an algorithm. I also would like to thank the conveners of the Jet/EtMiss working group for the comments and suggestions made during the course of this work. I would also like to thank very much every member of our group at SLAC for being thoroughly supportive and providing much needed feedback and assistance during this study.

A Datasets used

Physics process	Dataset	# events
MB (Sec. 4.1.1)	trig1_misal1_csc11.005001.pythia_minbias.recon.AOD.v12000601_tid005901	63K
$t\bar{t}$	trig1_misal1_csc11.005200.T1_McAtNlo_Jimmy.recon.AOD.v12000502	8.5K
$t\bar{t}(10^{33})^{\dagger}$	trig1_pile1sf05_misal1_csc11.005200.T1_McAtNlo_Jimmy.recon.AOD.v12000605_tid009270	7K
$H \rightarrow \gamma\gamma(10^{33})^{\ddagger}$	trig1_pile1sf05_misal1_mc12.005316.PythiaWH120gamgam.recon.AOD.v12000605_tid009265	3.7K
Di-jet (J6)$(10^{33})^{\ddagger}$	trig1_pile1sf05_misal1_csc11.005015.J6_pythia_jetjet.recon.AOD.v12000605_tid014764	9.2K
$t\bar{t}$ + jets$(10^{33})^{\ddagger}$	trig1_pile1sf05_misal1_mc12.006371.AlpgenJimmyttW1jets.recon.AOD.v12000605_tid009849	8.5K
$t\bar{t}(10^{33})^{\dagger}$	trig1_pile1sf02_misal1_mc12.005200.T1_McAtNlo_Jimmy.recon.AOD.v13003003_tid017023	9.5K
†MB overlay	misal1_csc11.005001.pythia_minbias.simul.HITS.v12000601	
†Cavern background overlay	misal1_csc11.007903.cavernbg_sf05.simul.HITS.v12000601	
‡MB overlay	misal1_csc11.005001.pythia_minbias.simul.HITS.v12003104	
‡Cavern background overlay	misal1_csc11.007903.cavernbg_sf05.simul.HITS.v12003105	

Table 4: Datasets used in investigating the event topology of pile-up events at ATLAS and in evaluating performance of the jet-vertex association algorithm. The simulated "HITS" files that were used in the pile-up overlay mechanism are also specified. Note that for all pile-up datasets except $t\bar{t}$ teh exact same sets of events were overlaid.

References

[1] M. Lamont, LHC Initial Commissioning, AB-Note-2006-XXXX.

[2] D. Sampsonidis,*et al*., Study of the response of the ATLAS Monitored Drift Tubes to heavily ionizing particles and of their performance with cosmic rays, NIM A **535** 2004 p260-264.

[3] T. Sjostrand, L. Lonnblad, S. Mrenna and P. Skands. PYTHIA 6.3 Physics and Manual hep-ph/0308153, 2003.

[4] Mininum bias event ATLAS CSC Note, 2008 (to be released)

[5] Estimation of Radiation Background, Impact on Detectors, Activation and Shielding Optimization in ATLAS, ATL-GEN-2005-001

[6] E. Hughes, Z. Marshall, and A. Schwartzman. Track-based improvement of jet resolution in ATLAS, ATL-COM-PHYS-2007-XXX

[7] A. Bhatti, *et al*., Determination of the Jet Energy Scale at the Collider Detector at Fermilab, arXiv:hep-ex/0510047v1

[8] DØ Collaboration, Measurement of the ppbar to ttbar production cross section at sqrt(s)=1.96 TeV in the fully hadronic decay channel, arXiv:hep-ex/0612040v1

[9] Section 10.2.4, Vertexing Performance, ATLAS Detector Paper, 2007