Stage de M2 physique théorique au Laboratoire de Physique Nucléaire et Hautes Energies: Mesure de la masse du quark Top dans l'expérience ATLAS.

Guillaume Beuf

5 avril 2006

Résumé :

Le but de ce stage était de contribuer à l'élaboration d'un programme d'analyse de données pour ATLAS pour l'étude du canal électron-jets des évènements top. L'étude a portée principalement sur les critères de sélection des évènements, et la reconstruction de la masse du top. Le programme a été testé sur données simulées.

Table des matières

1 Contexte du stage **3**
1.1 ATLAS et le LHC 3
1.1.1 Le LHC 3
1.1.2 ATLAS (A large Toroidal LHC ApparatuS) 3
1.2 La physique étudiée sur ATLAS 4
1.3 Le quark top 5
1.3.1 Interêt de l'étude du quark top 5
1.3.2 Les canaux de production et de désintégration du top . . 5
1.4 Insertion du stage dans l'expérience 6

2 Sélection des évènements **7**
2.1 Sélection des électrons 8
2.2 Sélection des jets 12
2.3 Critères globaux de sélection 15
2.3.1 Impulsion transverse manquante 15
2.3.2 Contenu en particules 16
2.3.3 *b*-tagging 17

3 Reconstruction **18**

1 Contexte du stage

1.1 ATLAS et le LHC

1.1.1 Le LHC

Le Large Hadron Collider est actuellement en chantier au CERN. Son démarrage est prévu pour l'été 2007. C'est un accélérateur circulaire réutilisant le tunnel du LEP. Du LEP au LHC, on passe de collisions électron - positon à des collisions proton - proton (ou noyau - noyau pendant certaines phases). Cela permet d'augmenter de 200 GeV à 14 TeV l'énergie dans le centre de masse, limitée dans ce type d'accélérateur par rayonnement synchrotron.

Quatre grandes expériences seront installées sur le LHC : LHCb, étudiant la physique des B, ALICE, étudiant le plasma de quarks et de gluons, et CMS et ATLAS, expériences généralistes.

1.1.2 ATLAS (A large Toroidal LHC ApparatuS)

Le détecteur : ATLAS sera le plus grand détecteur de particules jamais construit. Sa conception et son utilisation sont d'une complexité difficilement imaginable. Il mettra, par exemple, en jeu plusieurs centaines de millions de voies de mesures. Toutefois, le principe général est assez simple, et analogue à la grande majorité des détecteurs récents : c'est la succession de détecteurs de traces, de calorimètres et enfin de chambres à muons, tous cylindriques, emboités les uns dans les autres.

Le détecteur de traces est constitué de trois sous-détecteurs (Pixel, SCT[1], TRT[2]), de précision décroissante en s'éloignant du point d'interaction. Ils donnent des points par où sont passées les particules chargées. En reliant ces points, on reconstruit les trajectoires, et grâce au champ magnétique imposé dans cette partie du détecteur, on obtient la charge et l'impulsion des particules chargées.

Dans les calorimètres, formant les couches suivantes du détecteur, les particules issues de la collision interagissent avec le matériau pour former des gerbes de particules de plus basse énergie. On mesure alors l'énergie déposée par ces gerbes. Les gerbes correspondant aux photons, électrons et positons sont contenues dans le premier calorimètre, dit "électromagnétique". Les hadrons, par contre, forment des gerbes s'étalant dans les deux calorimètres successifs. Le deuxième est dit "hadronique".

Les seules particules (à part les neutrinos) traversant les calorimètres sans être arretées sont les muons (et anti muons). Ils arrivent dans la zone occuppée par les aimants toroïdaux donnant son nom au détecteur. Plusieurs couches de chambres à muons sont intercalées dans et autour de ces aimants. Cela permet de reconstruire les traces et de mesurer l'impulsion des muons.

[1] Semiconductor Tracker
[2] Transition Radiation Tracker

Grandeurs transverses : Dans le cas de collisions lepton-lepton, on peut exploiter les particules produites dans toutes les directions. Dans le cas de collisions hadron-hadron, comme ici, on doit se contenter de mesurer des grandeurs transverses par rapport aux faisceaux entrants. Les particules émises loin de cette direction proviendront de collisions parton-parton de haute énergie.

D'une part, l'évenement sous-jacent à ces collisions inélastiques, et aussi les collisions élastiques vont donner des hadrons proche de la direction axiale. Ainsi il y a un cône autour de l'axe qui va être noyé sous ce bruit de fond.

D'autre part, l'accélération des protons amplifie de presque quatre ordres de grandeur l'impulsion relative axiale des partons par rapport aux hadrons. Cela empèche de faire des bilans d'impulsion dans cette direction.

Par contre l'impulsion transverse relative des partons est négligeable. On peut alors mettre en évidence la production de neutrinos très énergétiques par une impulsion transverse manquante dans le bilan en impulsion d'un évènement.

Le système de coordonnées : Le repère universellement utilisé pour ATLAS est tel que l'axe des z est celui des deux faisceaux entrants, et donc les axes x et y définissent le plan transverse. On utilise aussi les coordonnées sphériques : ϕ correspond aux rotations autour de l'axe z, et θ l'angle par rapport à l'axe z. Plutôt que θ, on préfère utiliser la pseudo-rapidité η :

$$\eta = -\ln \tan\left(\frac{\theta}{2}\right) \tag{1}$$

1.2 La physique étudiée sur ATLAS

Le LHC va permettre d'explorer le domaine d'énergie correspondant à la brisure de la symétrie électrofaible. En effet, dans le modèle standard, la valeur moyenne dans le vide des champs de Higgs vaut 246 GeV. De nombreuses extensions ou alternatives au modèle standard font intervenir de la nouvelle physique à cette échelle d'énergie privilégiée. Ainsi, ATLAS et CMS devraient détecter le boson de Higgs s'il existe, ou bien nous permettre de trancher entre les mécanismes alternatifs de brisure de symétrie proposés. Ces expériences vont permettre de réfuter certains modèles théoriques, et d'imposer des contraintes très fortes sur les autres.

Par ailleurs, les expériences du LHC permettront d'améliorer notablement la précision des mesures pour de nombreux autres phénomènes liés au Modèle Standard déja observés, notamment dans la physique des bosons intermédiaires, des saveurs lourdes de quarks, et dans la chromodynamique quantique (QCD).

1.3 Le quark top

1.3.1 Interêt de l'étude du quark top

Le quark top est observé depuis plus de dix ans au Tevatron, mais il est encore relativement mal connu. ATLAS et CMS devraient être en mesure d'étudier systématiquement toutes ses propriétés. Comme il est le seul fermion dont la masse est de l'ordre de l'énergie de brisure de symétrie électrofaible, certains modèles théoriques alternatifs lui font jouer un rôle privilégié dans cette brisure.

Par ailleurs, la mesure de la masse du quark top à 1 GeV près, contre 5, 1 GeV aujourd'hui[3], est une des premières études qui pourront être réalisées sur ATLAS. Elle fait partie des paramètres libres du Modèle Standard. Pour plusieurs processus importants, les corrections radiatives sont dominées par la contribution d'une boucle top, ou top-bottom. Ainsi, à l'ordre sous dominant, la renormalisation de la masse du W dans le Modèle Standard donne une relation entre la masse du W, très bien connue, celle du top, et celle du boson de Higgs. Or, comme la masse du top intervient quadratiquement, et celle du Higgs logarithmiquement. Ainsi, en améliorant un peu la précision de la mesure de la masse du top, on améliore de beaucoup la précision de la prédiction sur la masse du Higgs dans le cadre du Modèle Standard.

1.3.2 Les canaux de production et de désintégration du top

Il existe deux modes de production pour ce quark :

Par interaction faible : Cela correspond à la désintégration d'un W virtuel en $t\bar{b}$ (ou $\bar{t}b$), ou bien à la collision d'un $\bar{b}$ avec un parton de haute énergie. Ce canal est peu probable, et n'a donc pas été détecté au Tevatron. De même, au LHC, il faudra attendre plusieurs années pour pouvoir l'étudier. Mais ce mode de production est le seul permettant certaines études particulières, comme celle de la polarisation des quarks top.

Par interaction forte : Cela correspond à la création d'une paire $t\bar{t}$ par fusion de quarks ou de gluons. C'est par ce mode de production que le top a été découvert, et ce mode que l'on va envisager pour le reste de cette étude.

La désintégration d'un quark top se fait quasi-exclusivement en b et $W+$. En effet, la matrice CKM ne mélange que très peu la troisième génération de quarks avec les deux autres. Le $W+$ se désintègre alors en une paire de quarks plus légers, ou en une paire de leptons. Dans le cadre de la production $t\bar{t}$, on distingue alors trois canaux de désintégrations : dilepton si les deux W se désintègrent en leptons, lepton-jets si l'un se désintègre en leptons et l'autre en quarks, et totalement hadronique si les deux W se désintègrent en quarks.

[3] Résultats combinés de DØ et de CDF pour le Run I du Tevatron.

Canal totalement hadronique : Sa signature (six jets) est particulièrement banale dans un collisionneur de hadrons. Il ne sera donc pas facile à distinguer de son bruit de fond. C'est pourquoi il n'est pratiquement pas étudié pour ATLAS, même si son rapport de branchement est de presque un demi.

Canal lepton-jet : C'est le canal le plus étudié. En effet, on a un lepton chargé et un neutrino, ce qui permet de sélectionner efficacement ces évènements.

Canal dilepton : Grâce aux deux leptons chargés créés, on peut distinguer de manière extrèmement efficace ce canal de son bruit de fond. Par contre, à cause de la présence de deux neutrinos, on ne peut pas reconstruire l'évènement de manière exacte, et on est obligé de recourir à des méthodes statistiques de vraisemblance pour obtenir la masse du top dans ce canal. Ce sera un des principaux domaines d'étude du groupe ATLAS du LPNHE.

En fait, il faut distinguer le type de leptons créés. En effet, l'électron, le muon et le tau ont des signatures très différentes. On détecte un muon de haute énergie dans ATLAS grâce aux chambres à muons principalement. Un électron, par contre sera signalé par un impact dans le calorimètre électromagnétique aligné avec une trajectoire donnée par les détecteurs de trace. Ces deux leptons donnent donc des signaux très distincts, et tous les deux aisément identifiables. Un tau, par contre, va se désintégrer en neutrino ν_τ et en un W, qui va lui même donner une paire de quarks ou de leptons légers. Cela donne dans les deux cas des processus nettement plus difficiles à sélectionner où à reconstruire.

Dans le cadre de ce stage, le canal électron-jet a été étudié, de manière à faire le lien entre des travaux passés (conception du calorimètre EM, et reconnaissance des électrons) et futurs (canal dilepton) de l'équipe au LPNHE.

(!!!!!!!!!!!!!!!!!!!!!!!!figure électron-jets)

1.4 Insertion du stage dans l'expérience

Pour étudier un processus à partir de données expérimentales en physique des particules, la première étape est de trier les évènements, pour ne retenir que ceux qui semblent correspondre au processus en question. Cette sélection doit augmenter le rapport signal sur bruit de plusieurs ordres de grandeur, pour rendre l'évènement étudié prépondérant dans les données choisies, pour réduire l'incertitude systématique des mesures. Elle doit aussi être efficace, c'est à dire ne pas trop rejeter d'évènements contenant le processus étudié, afin de ne pas augmenter l'incertitude statistique. Il faut donc trouver un compromis, minimisant l'incertitude totale. D'ailleurs, pour une expérience comme ATLAS, durant des années, les critères de sélection choisis évoluent a cours de l'expérience : plus les données arrivent, et plus on peut se permettre d'utiliser des critères restrictifs.

Il faut donc réaliser d'abord ce type d'étude sur des évènements simulés. Cela permet d'une part de débugger les programmes d'analyses réalisés et de tester leur pertinence, et d'autre part d'optimiser les critères de sélection utilisés. Mon stage correspondait à ce stade-là. Plusieurs séries d'évènements avaient été simulées grâce au générateur Monte Carlo Phytia. Ensuite le passage dans le détecteur et les prétraitements de données avaient été simulés, dans l'environnement Athena, pour constituer des données simulées au niveau CBNT[4]. Ces données contiennent les listes, par évènement, des impacts dans les calorimètres et des traces, avec toutes leurs propriétés intéressantes. On disposait de tels NTuples pour une série d'évènements electron-jets et pour une série d'évènements de bruit de fond de ce processus.

La structure générale du programme d'analyse du LPNHE pour les canaux lepton-jets et dilepton était déja écrite sous ROOT. Le but de mon stage était d'implémenter, pour le cas électron-jets, les sélections, les calculs de recontruction, et le tracé de certains histogrammes.

2 Sélection des évènements

La première étape du programme d'analyse physique réalise, pour chaque évènement, la liste des particules et jets détectés pouvant provenir du canal étudié. Ce tri est l'objet de cette section.

Comme on l'a vu au paragraphe 1.1.2, la direction axiale pose problème, alors on ne recherche ces signaux que suffisamment loin de l'axe des faisceaux, ce qui permet d'être sûr de n'avoir que des particules issues de processus à grand transfert d'impulsion. Cette restriction est sous la forme d'une coupure sur la pseudo-rapidité η. Les évènements top ne vont produire que des particules ou jets de haute énergie. Or, comme on ne les cherche que loin de l'axe, on impose une impulsion transverse p_T minimale, plutôt qu'une énergie minimale. Mais en plus de ces deux coupures simples, il y a des problèmes non triviaux : la distinction entre les leptons et les jets, et la définition même des jets.

Ensuite, une fois ces listes de "bonnes particules" formées, le programme impose des conditions plus globales. Pour l'étude du quark top dans le canal électron-jets, on ne garde que les évènements pour lesquels on a sélectionné au moins quatre jets et un électron, et où l'on devine l'emission d'un neutrino. Alors, on reconstruit l'ensemble du processus, en imposant d'autres conditions quand c'est possible, jusqu'à obtenir une valeur pour la masse du top dans cet évènement. Ceci sera l'objet de la section suivante.

[4] Combined N-Tuples : C'est un des formats de données utilisées pour les analyses physiques sur ATLAS. Il existe aussi AOD et ESD. Ils diffèrent par leur précision et leur niveau de prétraitement et de reconstruction.

2.1 Sélection des électrons

Dans les données CBNT, il y a, par évènement, la liste des objets dits $e\gamma$, qui sont les gerbes repérées dans le calorimètre électromagnétique. Cela peut correspondre à des photons, des électrons, ou bien à des hadrons. Tout d'abord, on applique les coupures déja mentionnées. on ne garde que les $e\gamma$ tels que :

$$|\eta| < 2,5 \quad (2)$$
$$p_T > 20 \text{ GeV} \quad (3)$$

Il faut alors étudier comment rejeter les photons et les hadrons parmi les $e\gamma$ et ne garder que les électrons. Une précédente stagiaire de l'équipe avait réalisé un programme similaire, mais dans un environnement informatique différent. Elle avait utilisé comme coupures :

$$\frac{EMCal}{EMCal + HADCal} > 0,9 \quad (4)$$
$$Isol < 0,15 \quad (5)$$
$$0,9 < \frac{p_{T\,Cal}}{p_{T\,trace}} < 1,3 \quad (6)$$

La première est une condition sur la fraction de l'énergie de la gerbe qui va être déposée dans le calorimètre électromagnétique. Pour un électron ou un photon, celle-ci est proche de 1, alors que pour un hadron, elle peut être plus faible : il y aura plus d'énergie déposée dans le calorimètre hadronique.

La deuxième condition porte sur la variable $Isol$, définie pour chaque objet $e\gamma$ au niveau CBNT. Plus elle est petite, plus la gerbe est isolée des autres gerbes dans le calorimètre, et donc moins elle a de chance de faire partie d'un jet. Ces deux conditions servent donc à rejeter les hadrons.

Dans les données CBNT, à chaque objet $e\gamma$ est associé la trace reconstruite la plus proche. La troisème condition compare le p_T mesuré grâce au calorimètre, et le p_T mesuré grâce au détecteur de trace. Pour une particule neutre comme le photon, qui ne laisse pas de trace dans le détecteur interne, la trace associée n'aura rien à voir avec l'objet $e\gamma$. Les p_T auront donc toutes les chances d'être assez différents, contrairement au cas des particules chargées comme l'électron.

Ces trois conditions permettent donc en principe de ne garder que les électrons. Mais c'est une sélection assez grossière, comme on le verra plus loin. On utilise plutôt la variable CBNT $IsEM$. Sa valeur est un nombre hexadécimal à trois chiffres, c'est à dire douze chiffres en binaire. Ceux-ci représentent les résultats de tests réalisés sur un objet $e\gamma$, permettant de déterminer si c'est un électron ou non. Ces tests généralisent les trois conditions précédentes.

En fait, seuls six bits sont vraiment utilisés. Trois servent a rejetter les hadrons, et ne font intervenir que les calorimètres :

– Le premier est analogue à la condition (4).

– Les hadrons créent des gerbes plus larges que les électrons et photons. Alors le deuxième est calculé en pondérant deux tests. L'un porte sur le rapport entre l'énergie mesurée dans un cluster de cellules, et la somme des énergies mesurées dans ce cluster et ses voisins. L'autre porte sur une variance spatiale de la gerbe, obtenue en pondérant chaque cellule par l'énergie qu'elle mesure.
– Le troisième bit rassemble cinq tests sur la forme fine de la gerbe, notamment deux qui rejettent les gerbes doubles, comme celle de deux photons issus d'un π^0.

Trois autre bits servent à distinguer les électrons des photons, en étudiant la trace associée à la gerbe au niveau des variables CBNT :

– Le premier teste la correspondance entre la trace et la gerbe, en pondérant les écarts entre les deux en ϕ, en η et en p_T.
– Le second teste la qualité de la trace dans les détecteurs internes. Pour tre validée selon ce critère, la trace doit être construite à partir d'au moins neuf points dans les sous-détecteurs Pixel et SCT, dont deux dans Pixel, dont un dans sa partie centrale. La trace doit aussi avoir un paramètre d'impact inférieur à un millimètre.
– Le troisième concerne la qualité de la trace dans le troisième détecteur de trace (TRT). Or ce sous détecteur pose des problèmes dans la version disponible des données simulées. On n'utilise donc pas ce critère dans cette étude.

Parmi les vrais électrons, beaucoup vont être sélectionnés par $IsEM$, d'autres seront rejetés. La fraction sélectionnée est par définition l'efficacité de la sélection. Puisque l'étude s'effectue sur des données simulées, on a accès à la vérité Monte Carlo, c'est à dire la nature et les propriétés des particules générées, avant simulation du détecteur. Alors on peut étudier l'efficacité de $IsEM$.

La figure 1 représente l'efficacité en fonction de η pour différents tests contenus dans $IsEM$, en prenant au départ les électrons déja sélectionnés selon les critères (2) et (3). On a enlevé tout une zone vers $1,5$, qui correspond à la jonction entre deux électrodes du calmorimètre électromagnétique. C'est une zone relativement aveugle du détecteur. Les premièrs, deuxièmes et quatrièmes graphiques représentent respectivement les trois sélections purement calorimétriques. Le cinquième les reprend toutes les trois. Toutes ces sélections ont une efficacité de l'ordre de $0,9$. Le sixième graphique représente l'efficacité des trois autres bits décrits. Il y a une baisse très importante due à des bugs dans la simulation du TRT. On retrouve cette baisse dans le dernier graphique, qui reprend les six tests. On utilisée dans l'analyse la sélection grâce à tous les bits décrits, sauf celui à propos du TRT. Elle correspond au huitième graphique.

Pour comparer la pertinence de la sélection par $IsEM$ et de celle par les conditions (4), (5), et (6), on construit les histogrammes des grandeurs sur lesquelles portent ces conditions, après sélection suivant (2) et (3), avant et après sélection par $IsEM$. Ils sont donnés sur les figures 2 à 4. Comme on dispose de signal simulé et de bruit simulé pour le canal électron-jet, on conduit en fait

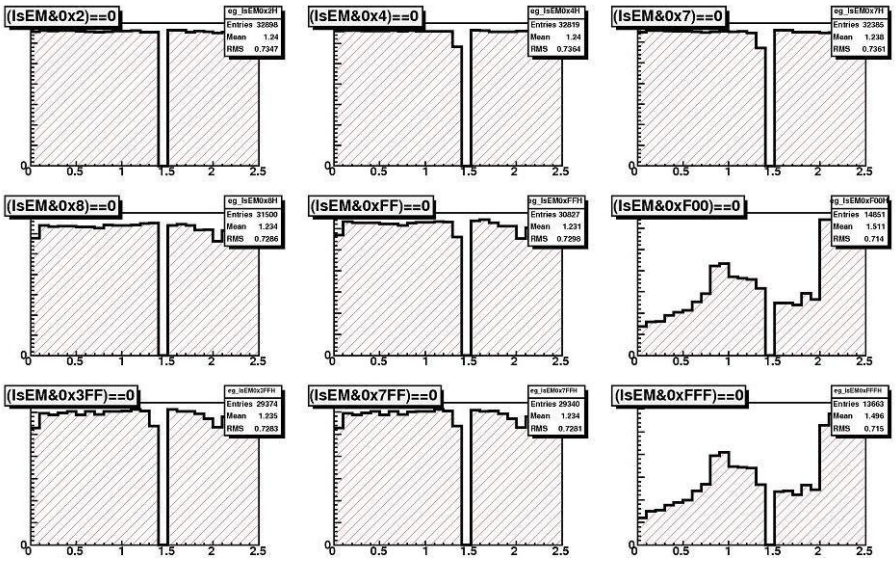

FIG. 1 – Efficacité de différentes sélections par $IsEM$.

l'analyse des deux indépendamment. Pour cette analyse des sélections des objets $e\gamma$, on utilise la vérité Monte Carlo de manière à ce que le signal ne contienne que des électrons, et le bruit que des hadrons légers, notamment des pions. Pour tous les graphiques à partir d'ici, l'histogramme du signal est en rouge, et celui du bruit en noir. Par ailleurs, jusquà la figure 8, sur chaque graphique, les histogrammes du bruit et du signal sont normalisés indépendamment.

Sur le premier graphique de la figure 2, on retrouve bien ce qui est attendu : pour les électrons du signal, la fraction électromagnétique de la gerbe est concentrée en 1, c'est à dire que la gerbe d'un électron est arrétée par le calorimètre électromagnétique, et ne parvient pas jusqu'au calorimètre hadronique. Et sur le deuxième graphique, on voit que le objets du bruit laissant une part non négligeable de leur énergie dans le calorimètre hadronique ont été supprimés. On remarque que dans le critère (4) la valeur du seuil aurait pu être choisie supérieure à $0,9$, par exemple $0,96$.

Par ailleurs, on voit sur l'ensemble de ces trois figures que le signal est très peu modifié par la sélection par $IsEM$, ce qui est en accord avec l'étude précédente sur l'efficacité. Une sélection suivant les conditions (4), (5) et (6) aurait été beaucoup moins efficace. En effet, on voit qu'elles auraient coupé les queues de distributions.

Grâce à ces trois figures, on remarque aussi que le bruit est notablement rejeté. Finalement, il ne contient bien plus que les hadrons qui ont des signatures analogues aux électrons.

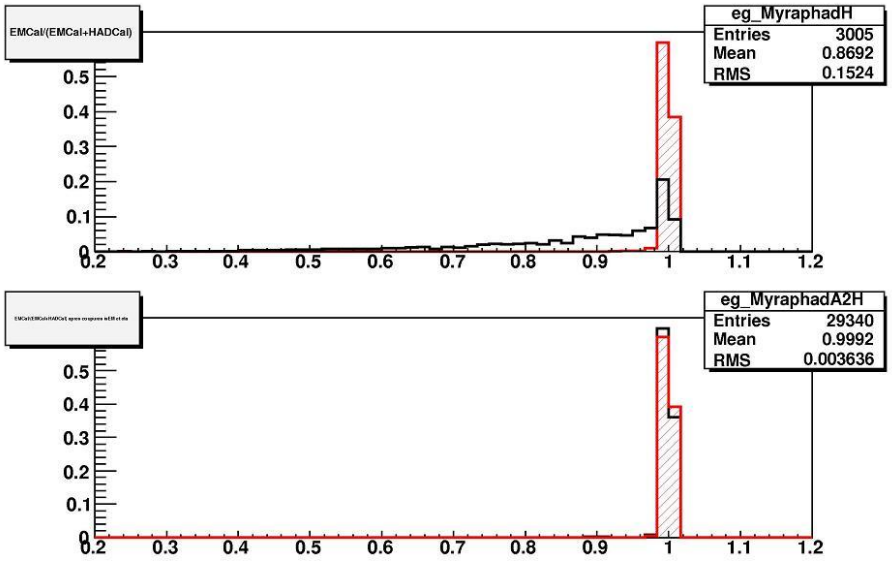

FIG. 2 – Fraction électromagnétique (cf condition (4)) avant et après sélection par $IsEM$.

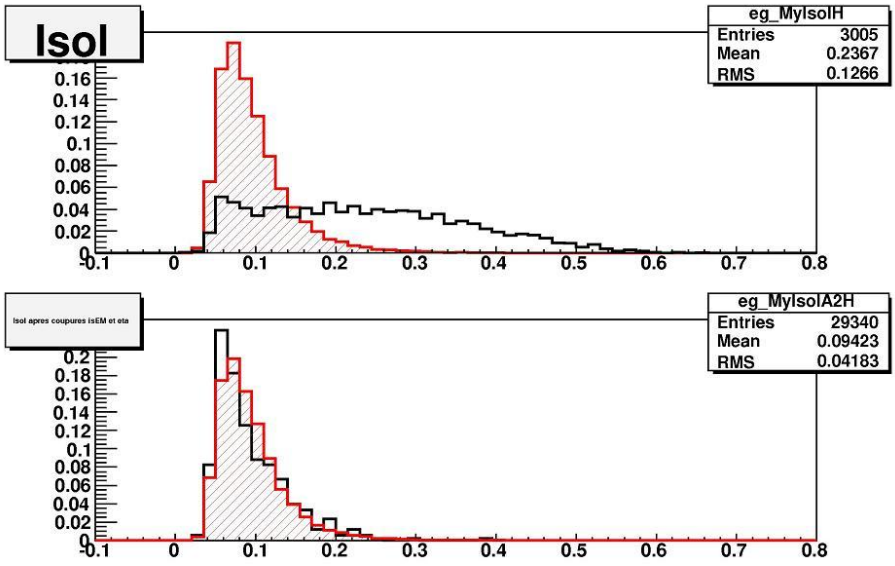

FIG. 3 – $Isol$ (cf condition (5)) avant et après sélection par $IsEM$.

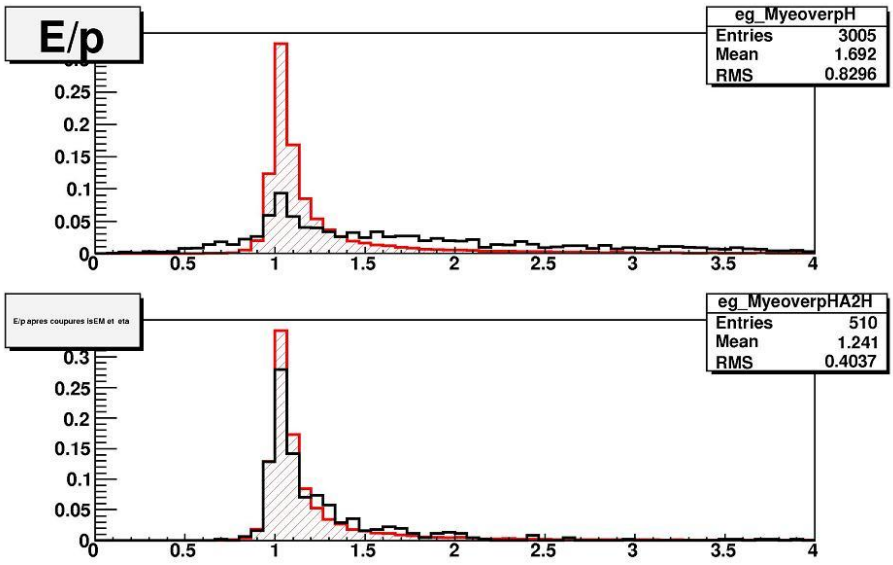

FIG. 4 – Rapport des p_T mesurés (cf condition (6)) avant et après sélection par $IsEM$.

Pour l'analyse du canal électron-jet, on construit alors, les listes de candidats électrons, pour chaque évènement du signal ou du bruit.

2.2 Sélection des jets

La notion de jet de hadron n'est pas bien définie, aussi bien sur le plan théorique, puisque l'hadronisation est encore mal comprise en chromodynamique quantique, que sur le plan phénoménologique. En effet, ce que l'on repère dans le détecteur, ce sont des gerbes dans les calorimètres, et des traces dans les détecteurs internes. Celles-ci sont souvent groupées. On forme des jets à partir de ces groupes, mais il n'y a pas de critère pratique pour décider de manière universelle comment distinguer plusieurs petits jets qui se chevauchent d'un jet plus large. Ce problème est relié notamment à celui, théorique, de la régularisation des divergences infrarouges et colinéaires de l'émission de gluons. A cause de cela, il existe plusieurs algorithmes différents pouvant reconstruire la liste des jets d'un évènement. Et tous donneront des résultats différents. Au moins trois algorithmes sont implémentés dans les programmes d'analyse d'ATLAS : Cone04, Cone07, et Kt. Au niveau CBNT, on dispose des jets reconstruits par ces trois algorithmes, et de leurs propriétés.

Les trois premiers graphiques des figures 5, 6 et 7 représentent les distributions en p_T, en η et en ϕ de ces trois types de jets. On voit clairement que

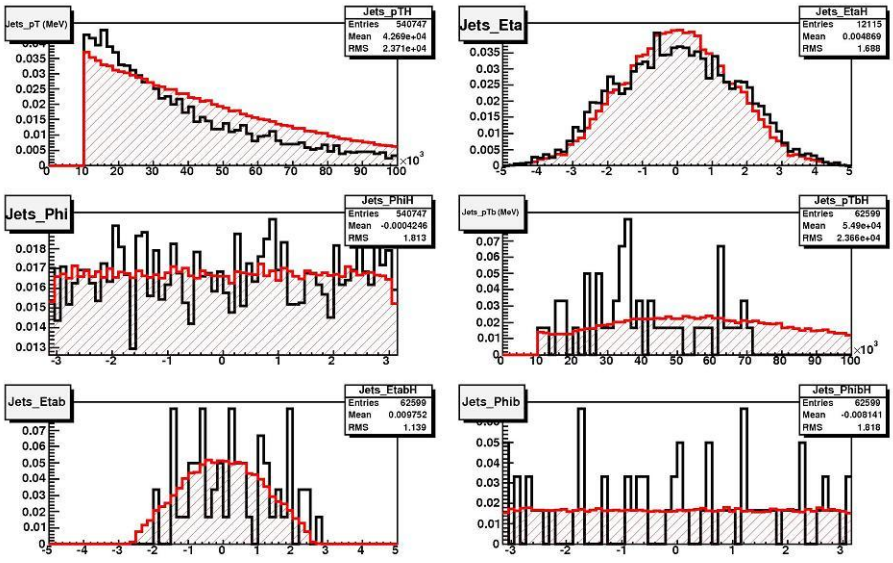

FIG. 5 – Caractéristiques des jets construits par l'algorithme Cone 04.

la distribution en p_T n'est pas la même. Les algorithmes Cone04 et Cone07 construisent les jets en prenant toutes les gerbes présentes dans un cône d'une certaine ouverture, différente selon l'algorithme. Ainsi, Cone07 construit moins de jets, et ceux-ci contiennent plus de particules, donc le p_T des jets sera plus grand. Et on voit sur les graphiques que Kt construit les jets les plus petits et nombreux, et les moins énergétiques. Le type de jet choisi dépend du processus étudié, et de l'énergie à laquelle on l'étudie. En effet, plus un processus produit des quarks et des gluons de haute énergie, plus ceux-ci vont former des jets fins. Donc Kt est plus adapté à l'étude de processus de très haute énergie, et Cone07 de plus basse énergie. Pour l'étude du canal électron-jet des évènements top, on utilise l'algorithme Cone04.

Il est possible, par une procédure dite *b-tagging*, de reconnaitre, parmi les jets, lesquels sont des jets issus de quarks b (ou $\bar{b}$). Cette procédure sera discutée plus en détail au 2.3.3. Les trois derniers graphiques des figures 5 à 7 représentent les distributions en p_T, η et ϕ des jets reconnus comme issus de b. Comme les évènements $t\bar{t}$ produisent deux jets de b, le signal contient beaucoup plus de jets reconnus comme jets de b que le bruit, ce qui sera utile plus tard.

Avant toute sélection, on dispose de nombreux jets construits, par évènement, par Cone04. La figure 8 représente ce nombre. On restreint alors l'étude aux jets transverses et très énergétiques en imposant les conditions :

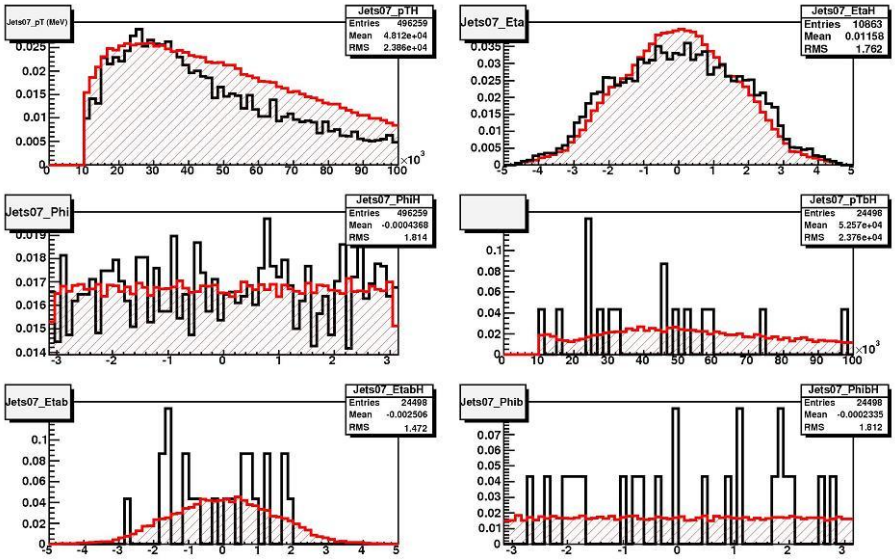

FIG. 6 – Caractéristiques des jets construits par l'algorithme Cone 07.

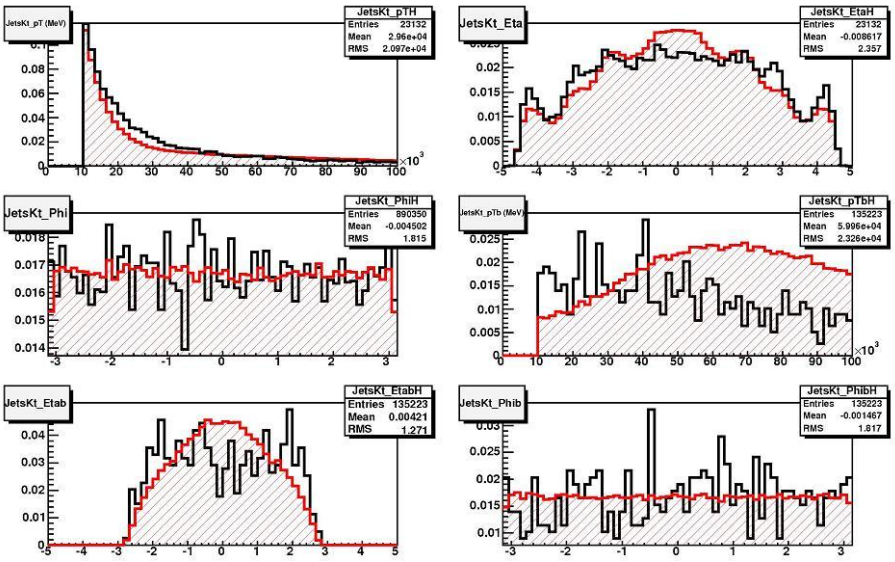

FIG. 7 – Caractéristiques des jets construits par l'algorithme Cone Kt.

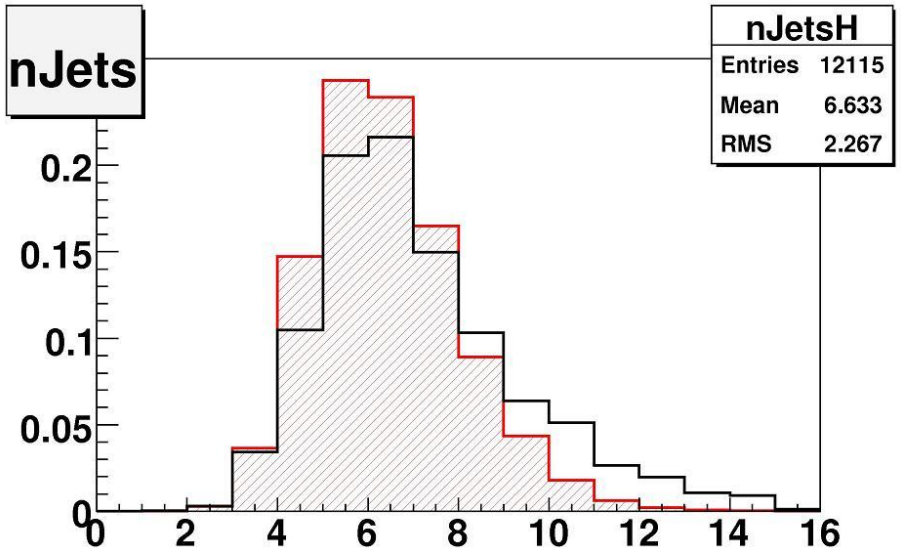

FIG. 8 – Nombre de jets construits par évènement par l'algorithme Cone 04.

$$|\eta| < 2,5 \quad (7)$$

$$p_T > 30 \text{ GeV} \quad (8)$$

On a alors une liste de candidats électrons, et une liste de jets pour chaque évènement. Il reste à sélectionner les évènements avant de tenter de reconstruire la masse du quark top.

2.3 Critères globaux de sélection

2.3.1 Impulsion transverse manquante

Fondamentalement, chaque cellule des calorimètres mesure une énergie. Comme les gerbes couvrent de nombreuses cellules, il faut sommer ces énergies. Or, on connait la position des cellules, et on sait d'où vient la particule donnant la gerbe. Alors plutôt que de faire une somme scalaire, on peut faire une somme vectorielle, et donc assimiler ces énergies mesurées à des impulsions, à une calibration du détecteur près. Alors, on peut faire un bilan d'impulsion sur chaque évènement, en excluant comme toujours un voisinage de l'axe des faisceaux. L'implusion transverse totale est évidemment nulle. Si cela ne semble pas être le cas, et qu'il y a une impulsion transverse manquante, c'est qu'un ou plusieurs neutrino de haute énergie se sont échappés. Cette impulsion transverse manquante fait partie des variables CBNT (sous le nom d'énergie transverse man-

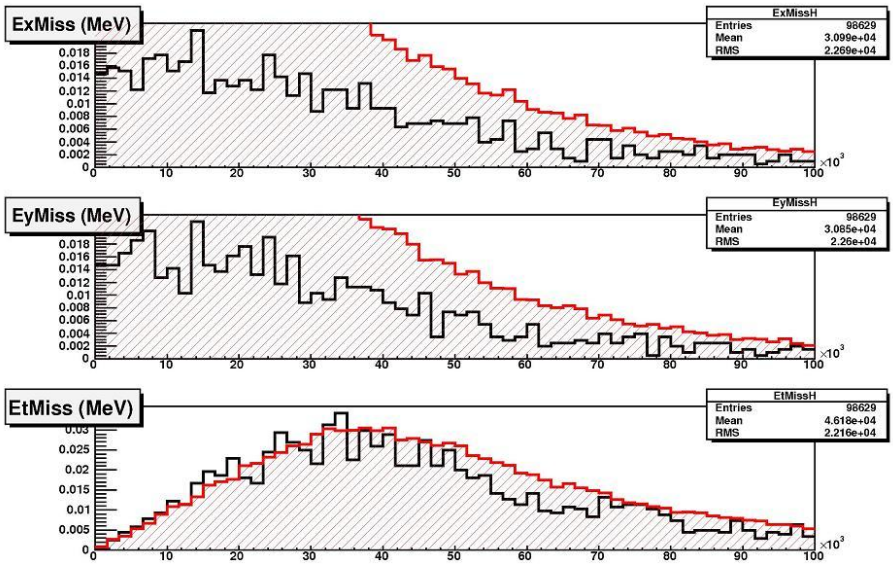

FIG. 9 – Impulsion transverse manquante : composantes suivant x et y, et norme.

quante). Un neutrino de grande énergie est émis lors des évènements électron-jets. Ainsi, la première sélection globale sur les évènements consiste à exiger la présence d'un neutrino dur (et suffisamment transverse) grâce au critère :

$$\sqrt{E_{Xmiss}^2 + E_{Ymiss}^2} > 20 \text{ GeV} \tag{9}$$

E_{Xmiss} et E_{Ymiss} sont les composantes de l'impulsion transverse manquante. Les distributions de ces deux grandeurs, et de la norme de l'impulsion transverse manquante sont données par la figure 9.

2.3.2 Contenu en particules

Ensuite, il faut vérifier qu'aucun des jets sélectionnés n'est un électron, identifié par erreur comme jet. On définit pour cela une distance angulaire :

$$\Delta R = \sqrt{(\eta_2 - \eta_1)^2 + (\phi_2 - \phi_1)^2} \tag{10}$$

Pour chaque jet sélectionné, on cherche alors le candidat électron le plus proche suivant cette distance. On rejette alors les jets pour lesquels cette distance minimale à un électron ne vérifie pas la condition :

$$\Delta R_{min} > 0,1 \tag{11}$$

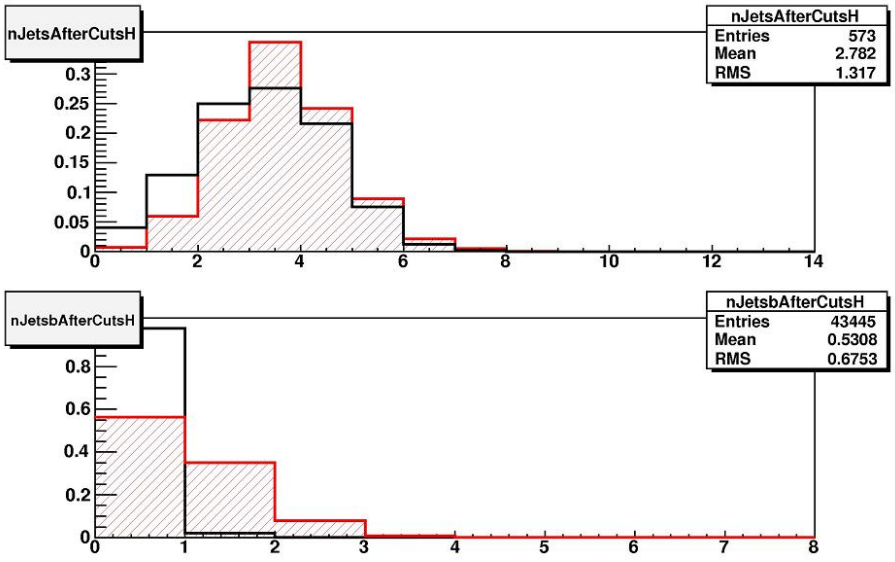

FIG. 10 – Nombre de jets et de jets de b sélectionnés par évènement.

Il reste à rejeter les évènements qui n'ont pas le contenu minimal requis pour le canal électron-jets, c'est à dire quatre jets, un électron et un neutrino. La présence d'un neutrino est en principe assurée par la condition 9. On ne garde alors que les évènements pour lesquels les nombres de candidats électrons et de jets sélectionnés vérifient :

$$N_{elec} \geq 1 \quad (12)$$

$$N_{jet} \geq 4 \quad (13)$$

La distribution de N_{jet} est donnée par la figure 10. On remarque que la majorité des évènements du signal ont moins de 4 jets sélectionnés. Ils doivent donc contenir des jets trop près de l'axe pour pouvoir être exploités.

2.3.3 b-tagging

Les évènements $t\bar{t}$ contiennent toujours deux jets de b (b et $\bar{b}$ en fait), ce qui est assez rare dans le bruit de fond. Or, il existe des méthodes pour reconnaitre les jets de b. On peut profiter de cette possibilité pour réduire encore le bruit de fond. Des algorithmes de b-tagging sont déja effectués lors de la création des variables CBNT. Pour chaque jet, on dispose alors d'un nombre qui quantifie sa vraisemblance d'être un jet de b. Ces algorithmes sont loin d'être totalement fiables, d'où le résultat sous la forme de vraisemblance plutôt que sous la forme

d'un booléen. J'ai choisi un seuil très élevé sur la vraisemblance, afin d'être presque sûr de n'avoir que des vrais jets de b parmi les jets désignés par le b-tagging. Sur la figure 10, on voit le nombre de jets de b ainsi reconnu parmi les jets précédemment sélectionnés. J'ai choisi le seuil en vraisemblance de manière à ce que le nombre d'évènements du signal contenant au moins 3 jets de b soit négligeable. On voit que si on rejette les évènements sans jets de b, on rejette environ la moitié de signal, et presque tout le bruit. Comme le b-tagging nécessite une bonne compréhension du détecteur, et plus de statistiques, il ne sera utilisé que dans un deuxième temps lors de l'exploitation des données d'ATLAS. C'est pourquoi j'ai mis seulement en option dans le programme la sélection des évènements par b-tagging :

$$N_{bjet} \geq 1 \tag{14}$$

3 Reconstruction

Il reste à faire la reconstruction proprement dite des évènements, c'est à dire identifier de quelle particule sont issus les jets, et calculer grâce à cela les masses des W et des t.

On traite d'abord le côte hadronique de l'évènement. Il s'agit dans un premier temps de trouver les trois jets provenant de la désintégration du même quark t (ou $\bar{t}$). Comme les quarks t et $\bar{t}$ sont émis avec une grande impulsion dans le référentiel de leur centre de masse, les particules issues de chacun doivent être proche en distance angulaire, dans ce référentiel. On doit donc avoir deux groupes de particules, formés dans des cônes opposés. Ce référentiel est aussi celui des partons interagissant, mais pas celui des protons, et il est inconnu a priori. Dans le référentiel du centre de masse des protons, qui diffère essentiellement par un boost axial, ces configurations angulaires ne se remarquent pas vraiment. On utilise un critère analogue, mais qui marche un peu mieux : on cherche le triplet de jets dont la somme vectorielle des p_T est maximale. Ce critère ne donne pas toujours trois jets issus d'un même quark t ou $\bar{t}$, mais seulement dans la plupart des cas, cela dépend notamment de la direction de production de ces quarks par rapport à la direction des faisceaux entrants.

On dispose des impulsions de ces trois jets grâce aux calorimètres. On néglige la masse des quarks donnant ces jets, ce qui permet d'obtenir les énergies. On fait la somme des quadriimpulsions des trois jets. Cela donne la quadriimpulsion et donc la masse du quark t du côté hadronique de l'évènement.

La distribution de cette masse est donnée par le premier graphique des figures 11 ou 12, selon que l'on utilise ou non le b-tagging. On remarque sur ces deux graphiques que le signal présente déja un pic bien prononcé, qui doit correspondre à la masse physique du quark top. Par ailleurs, pour tout les graphiques à partir de ceux-ci, les histogrammes du bruit sont normalisés de manière à ce que le maximum des distributions de signal et de bruit coïncident. On remarque que l'on nombre d'évènements de bruit est faible dans le cas sans b-tagging, ce qui prouve l'efficacité des sélections précédentes, et même égal à 1 dans le cas

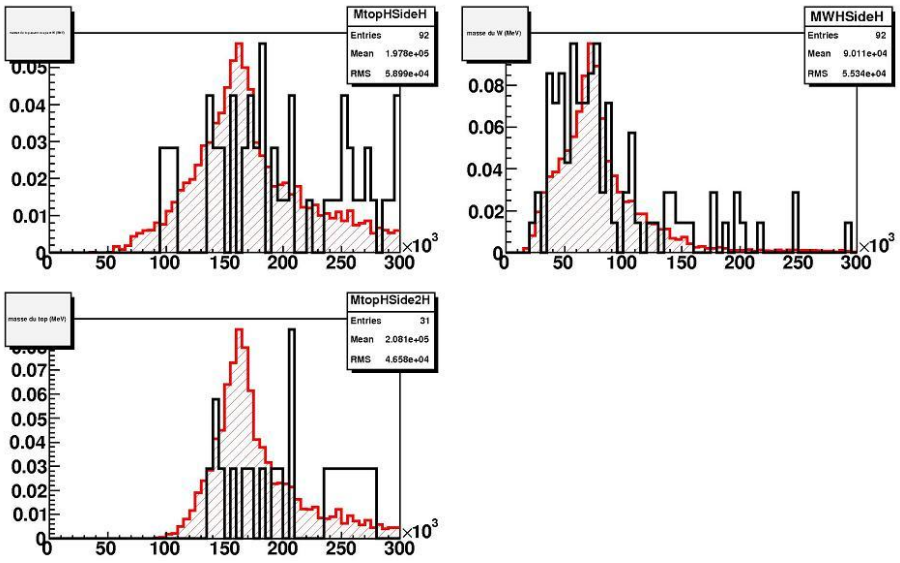

FIG. 11 – Masses reconstruites du côté hadronique sans b-tagging : top et W.

avec b-tagging.

Une fois que l'on a sélectionné ces trois jets dans la liste des bons jets, il faut chercher lequel est le jet de b, et lesquels proviennent du W. La meilleure méthode, pour effectuer cette identification, est de se placer dans le référentiel propre du t, c'est à dire le référentiel barycentrique des trois jets. On calcule les quadriimpulsions des trois jets dans ce référentiel par transformation de Lorentz. Dans ce référentiel, le b et le W sont émis dos à dos, et le W va se désintégrer en deux jets dans des directions presque toujours proches de la sienne. Il y aura donc ces deux jets proches l'un de l'autre, et le jet de b à l'opposé. Alors on applique un critère analogue à celui de l'étape précédente de la reconstruction : les jets issus du W sont les deux (parmi les trois du côté hadronique) dont la somme des impulsions transverses est maximale. On a alors la quadriimpulsion du W en sommant celles des deux jets qu'il donne, et on en déduit la masse du W.

Le deuxième graphique des figures 12 et 11 représente la distribution de la masse du W ainsi reconstruite, avec où sans b-tagging. On a bien, pour le signal, un pic pour une valeur de l'ordre de la masse du W[5].

Afin de réduire l'erreur systématique, on veut éliminer les évènements mal reconstruits, dont les jets ont mal été identifié. Pour cela, on rejette les évènements pour lesquels la masse du W diffère de plus 20 GeV de la valeur connue. Par

[5] $M_W = 80,425 \pm 0,038$ GeV, (Particle Data Group, 2004)

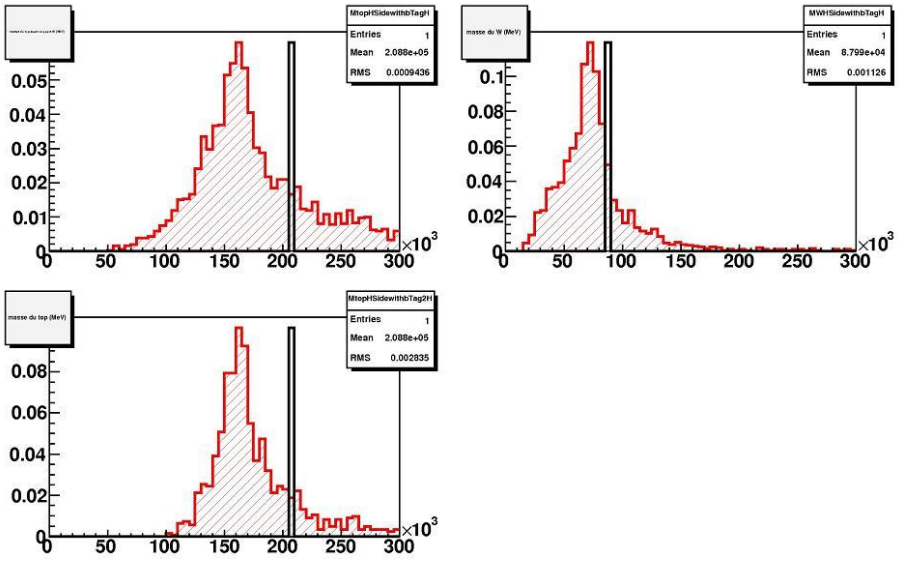

FIG. 12 – Masses reconstruites du côté hadronique avec b-tagging : top et W.

ailleurs, dans le cas avec b-tagging, on rejette les évènements pour lesquels un des jets issus du W est un jet de b.

Alors, on trace de nouveau la distribution de la masse du *top*, pour les évènements restants. Cela donne le troisième graphique des figures 11 et 12. On peut voir que ces dernières sélections ont nettement affiné le pic de masse du *top* trouvé. Il faudrait ensuite approximer ces histogrammes par des courbes de Breit-Wigner pour trouver la masse et la lageur totale du *top*. Mais on voit tout de même que masse trouvée est de l'ordre de 160 GeV, soit nettement en dessous de la valeur attendue, puisque le générateur Monte Carlo a créé ces évènements en prenant 175 GeV comme masse du *top*. Mais on peut voir que la masse du W a l'air d'être un peu sous estimée elle-aussi. Tout cela est donc probablement dû à l'absence de calibration des jets dans cette étude.