

Université Denis Diderot, Paris 7

Mémoire

présenté
pour obtenir

l'habilitation à diriger des recherches

par

Frédéric DERUE

*LPNHE - Laboratoire de Physique Nucléaire et de Hautes Énergies
IN2P3 - CNRS ; Université Pierre et Marie Curie-Paris 6 ; Université Denis
Diderot-Paris 7*

**Reconstruction et identification des électrons dans
l'expérience Atlas. Participation à la mise en place
d'un Tier 2 de la grille de calcul.**

CERN-THESIS-2008-030
11/03/2008



Composition du jury

MM.	Pascal	Debu	Président
	Daniel	Froidevaux	Rapporteur
	François	Le Diberder	Rapporteur
	Yves	Sirois	Rapporteur
	Philippe	Schwemling	Examineur

Soutenu le 11 mars 2008

AVANT PROPOS

Ce document présente une synthèse des travaux auxquels j'ai participé depuis six ans dans la collaboration Atlas pour la reconstruction et l'identification des électrons.

Je travaille dans le groupe Atlas du Laboratoire de Physique Nucléaire et de Hautes Energies de Paris, depuis mon entrée au CNRS en octobre 2002. Une partie de mes activités a porté sur l'acquisition de compétences en matière de logiciel et de calcul. Le chapitre consacré à l'expérience Atlas fait ainsi la part belle à la description du développement d'un nœud de la grille de calcul au laboratoire. L'essentiel du travail de développement logiciel a lui été consacré à la reconstruction et l'identification des électrons ; quatre chapitres de ce manuscrit y sont consacrés. Outre ma participation à l'étude des performances de l'algorithme d'identification des électrons isolés, j'ai contribué au développement de l'algorithme de reconstruction des électrons mous. Les applications les plus importantes de ce travail concernent l'utilisation des événements J/ψ pour étudier l'étalonnage du calorimètre électromagnétique à basse énergie et l'étiquetage des jets issus de l'hadronisation de quarks b . Enfin, depuis maintenant près de deux ans, une analyse visant à mesurer la masse du quark top avec les premières données du LHC est développée au laboratoire ; un chapitre est consacré à ma participation à cet effort.

L'essentiel du contenu des chapitres portant sur les performances de la reconstruction combinée des détecteurs provient de notes internes de la collaboration Atlas, publiées vers 2003 pour le Data Challenge 1¹). Ce manuscrit a été écrit en pleine analyse du Data Challenge "Computing System Commissioning (CSC)" qui verra l'écriture de plusieurs dizaines de notes, en particulier une mise à jour de toutes celles concernant les électrons. En plus d'une présentation synthétique des résultats déjà publiés, j'indique à la fin des chapitres les principaux développements effectués depuis pour les notes "CSC". Ce compromis, même s'il alourdit et brise un peu l'unité de lecture, permettra, je l'espère, à ce manuscrit de ne pas être totalement obsolète sitôt écrit.

Ce document n'aurait pas été complet sans une étude de données réelles. La reconstruction et l'identification des électrons de basse énergie a été étudiée avec les données du test en faisceau combiné d'Atlas effectué en 2004. Bien que les études soient encore en cours, un travail original et non encore publié est présenté ici, se limitant toutefois aux performances de l'identification des électrons. Comme nous le verrons ces résultats sont encourageant et sont un avant-goût des études à venir avec les premières données du collisionneur LHC.

¹) Les Data Challenge sont des séries de productions massives de données Monte Carlo, chacune étant dénotées par un numéro ou par un nom plus spécifique, qui seront décrites au paragraphe 1.6.

Enfin, dans les annexes, on trouvera un curriculum vitæ, la liste de mes activités d'enseignement et d'encadrement ainsi qu'une liste de publications, de notes internes et de présentations en conférences et séminaires. On y trouvera aussi un récapitulatif de mes activités de recherche ayant précédées mon entrée dans la collaboration Atlas : une thèse de doctorat, qui s'est déroulée de 1996 à 1999 au Laboratoire de l'Accélérateur Linéaire d'Orsay dans la collaboration Eros II sur la recherche de matière noire dans la Galaxie et trois années de post-doc au Service de Physique des Particules de Saclay et au Centre de Physique des Particules de Marseille sur l'étude de la violation de CP dans le système des kaons neutres (NA48) et des mésons B (LHCb). Le document se finit sur quelques perspectives sur mes activités de recherche à venir et sur les inévitables remerciements à ceux sans qui tout cela n'aurait pas été possible.

TABLE DES MATIÈRES

1. L'expérience Atlas	1
1.1 Le Large Hadron Collider	1
1.2 La collaboration Atlas	3
1.3 Le détecteur Atlas	4
1.3.1 Le trajectographe interne	5
1.3.2 Les calorimètres	9
1.3.3 Le spectromètre à muons	15
1.3.4 Le système de déclenchement	15
1.4 Le modèle de calcul d'Atlas	17
1.4.1 Environnement logiciel	18
1.4.2 Le formatage des données	19
1.4.3 La structure en Tiers et en Nuages	21
1.5 Le projet Grif	23
1.5.1 Contexte technique	23
1.5.2 Le projet Grif au LPNHE	27
1.5.3 Matériel et infrastructure	29
1.5.4 Plan de financement	35
1.5.5 Etat actuel de Grif	38
1.6 Utilisation des ressources de calcul	39
1.6.1 Participation aux "Data Challenge"	40
1.6.2 Publication des ressources des sites de production français	40
2. Canaux à électrons dans Atlas	43
2.1 Physique des bosons de jauge électrofaibles	43
2.2 Recherche du boson de Higgs	45
2.2.1 Physique au-delà du Modèle Standard	46
2.3 Physique des saveurs lourdes	46
2.3.1 La production de quarks b et c	46
2.3.2 Production de quarks top	48

3. Séparation électrons/jets	49
3.1 Les données utilisées	49
3.2 Reconstruction des électrons	50
3.2.1 Algorithme de reconstruction	51
3.2.2 Energie et position de la particule incidente	51
3.2.3 Reconstruction des électrons de conversion	52
3.3 Sélection des électrons au niveau du déclenchement	53
3.4 Identification des électrons	54
3.4.1 Variables caractérisant les électrons	54
3.4.2 Performances	61
3.5 Quelques développements et résultats récents	65
3.5.1 Développements algorithmiques	65
3.5.2 Identification des électrons	67
3.5.3 Isolation des électrons	68
4. Etudes sur le J/ψ	73
4.1 Les données utilisées	73
4.1.1 Cinématique des événements	74
4.1.2 Sélection des événements	74
4.2 Reconstruction et identification des électrons	75
4.2.1 Algorithme de reconstruction	75
4.2.2 Variables caractérisant les électrons	76
4.2.3 Identification des électrons	81
4.3 Reconstruction du J/ψ	83
4.3.1 Sélection des événements	83
4.3.2 Simulation du bruit de fond	84
4.3.3 Reconstruction de la masse invariante	86
4.4 Inter-étalonnage du calorimètre électromagnétique	89
4.5 Travaux plus récents	92
4.5.1 Système de déclenchement	92
4.5.2 Reconstruction et identification des électrons	93
4.6 Conclusion	94
5. Reconstruction des électrons dans les jets et étiquetage des jets par un électron mou	95
5.1 Les données utilisées	95
5.2 Reconstruction des électrons dans les jets	98
5.2.1 Description de l'algorithme	98
5.2.2 Performances	99

5.2.3	Facteur de rejet des électrons de conversion	100
5.3	Etiquetage des jets de b par un électron	103
5.4	Travaux récents	106
5.4.1	Développements algorithmiques	106
5.4.2	Méthodes d'analyses discriminantes	106
5.4.3	Performances	107
5.5	Conclusion	108
6.	Etudes sur le quark top	111
6.1	Les données utilisées	111
6.2	Reconstruction du signal $t\bar{t}$ avec les premières données	112
6.2.1	Sélection des événements	112
6.2.2	Reconstruction du top hadronique	114
6.3	Etiquetage des jets de b dans les événements $t\bar{t}$	118
7.	Identification des électrons de basse énergie dans les données du test en faisceau combiné H8 de 2004	119
7.1	La ligne de faisceau H8 du SPS	120
7.1.1	Le faisceau	120
7.1.2	Instrumentation de la ligne de faisceau	120
7.2	Mise en place du test en faisceau combiné	122
7.2.1	Le système de coordonnées	122
7.2.2	Les sous-détecteurs	122
7.2.3	Le système de déclenchement et d'acquisition	123
7.2.4	Reconstruction des données	124
7.3	Les données utilisées	124
7.4	Purification des faisceaux	125
7.4.1	Utilisation des informations du déclenchement	127
7.4.2	Contamination en muons	127
7.4.3	Utilisation des chambres à faisceaux	131
7.4.4	Récapitulatif des critères de sélection	132
7.5	Sélection des lots d'électrons et de pions	133
7.5.1	Sélection d'événements simples	133
7.5.2	Utilisation des compteurs Cerenkov	134
7.6	Etudes sur le TRT	138
7.7	Etudes de la forme des gerbes électromagnétiques	139
7.7.1	Profils longitudinaux	140
7.7.2	Profils latéraux	142
7.8	Identification des électrons	143

7.8.1	Variables caractérisant les électrons	144
7.8.2	Performances	147
7.9	Conclusion	149
8.	Conclusion	151
	<i>Annexe A : Curriculum Vitae</i>	<i>165</i>
	<i>Annexe B : Activités d’enseignement, d’encadrement et de diffusion de la connaissance</i>	<i>167</i>
	<i>Annexe C : Liste des publications</i>	<i>171</i>
	<i>Annexe D : La vie avant Atlas, l’Univers, le reste</i>	<i>177</i>
	<i>Annexe E : Résumé sur l’originalité des recherches</i>	<i>183</i>
	<i>Annexe F : Perspectives des activités de recherche</i>	<i>185</i>
	Remerciements	189

1. L'EXPÉRIENCE ATLAS

Ce chapitre est dédié à la présentation de l'expérience Atlas ainsi qu'à ma participation à la mise en place de celle-ci. Le texte n'a ni la prétention ni le but d'être exhaustif. Le fil conducteur est de rappeler un certain nombre d'informations qu'il sera nécessaire d'avoir en tête lorsqu'on abordera la reconstruction et l'identification des électrons. Un effort particulier est porté à la présentation des moyens de calcul et de stockage de l'expérience.

Le paragraphe 1.1 résume quelques informations relatives au LHC. Ensuite une présentation de la collaboration et du détecteur Atlas est donnée dans les paragraphes 1.2 et 1.3. Le paragraphe 1.4 décrit le modèle de calcul utilisé par l'expérience. Le paragraphe 1.5 décrit la participation de notre laboratoire au projet Grif de développement d'un nœud de la grille de calcul en Ile de France, c'est à dire d'un réseau permettant aux utilisateurs de disposer de manière transparente de puissance de calcul et de stockage. Enfin la section 1.6 décrit mes activités effectuées pour le calcul et la simulation de données Monte Carlo.

La documentation sur le LHC et l'expérience Atlas est vaste et intimidante. En plus des saines lectures que sont l'étude des nombreux Technical Design Report [1], du papier "détecteur" [2] en cours d'écriture ou de diverses notes techniques, rappelons qu'il n'est pas inutile de jeter un coup d'œil sur la littérature de vulgarisation, ce qui apporte un peu de fraîcheur et permet souvent d'être moins désarmé face aux questions innocentes, comme celles parfois posées par les étudiants en stage. Les références suivantes [3] [4] et [5] ont donc été largement pillées pour rendre ce chapitre plus agréable à lire.

1.1 *Le Large Hadron Collider*

Le Large Hadron Collider (LHC) [6] est l'accélérateur de nouvelle génération qui entrera en fonctionnement au Cern [7], le Laboratoire Européen de Physique des Particules, au printemps 2008. Il est situé dans le tunnel de 27 km de circonférence occupé jusqu'en 2002 par le collisionneur LEP (Large Electron Positron collider). Deux faisceaux de protons, ainsi que d'ions lourds dans certains cas, seront accélérés jusqu'à une énergie de 7 TeV, soit une énergie disponible dans le centre de masse de 14 TeV. Comme les protons sont des objets composites constitués de quarks et de gluons et non pas des objets élémentaires comme les électrons et positons, chaque collision proton-proton sera en fait une collision entre deux constituants appartenant à l'un ou l'autre proton. On observera donc des collisions quark-quark, quark-gluon ou gluon-gluon. Chaque quark ou gluon ne portant qu'une fraction de l'impulsion totale des

1. L'EXPÉRIENCE ATLAS

7 TeV des protons du faisceau, en pratique le LHC permettra d'explorer un domaine d'énergie allant jusqu'à 2-3 TeV. Par ailleurs, on ne sait pas, collision par collision, quelle est la fraction d'impulsion mise en jeu, ni quelle est la nature précise (quark ou gluon) des particules ayant interagi.

Une représentation du complexe d'accélérateurs permettant l'injection et l'accélération dans le LHC est montrée sur la figure 1.1. Les protons seront produits à une énergie de 50 MeV dans l'accélérateur linéaire, puis injectés dans le Proton Synchrotron Booster à 1,4 GeV. Le Proton Synchrotron (PS) accélérera alors les protons jusque 25 GeV. Le Super Proton Synchrotron (SPS) accélérera enfin les protons jusque 450 GeV, où ils seront fin prêts à être injectés dans le LHC. La construction du LHC s'est accompagnée de développements technologiques importants. Ainsi, pour maintenir les protons sur leur trajectoire, il faut produire un champ magnétique de 8,4 T sur une fraction appréciable de la circonférence de l'accélérateur. Ceci sera effectué par 1232 aimants dipolaires supraconducteurs d'une quinzaine de mètres de long chacun, refroidis à une température de 1,8 K par de l'hélium superfluide. S'y ajoutent près de 500 aimants quadripolaires supraconducteurs dont le but est de focaliser les faisceaux.

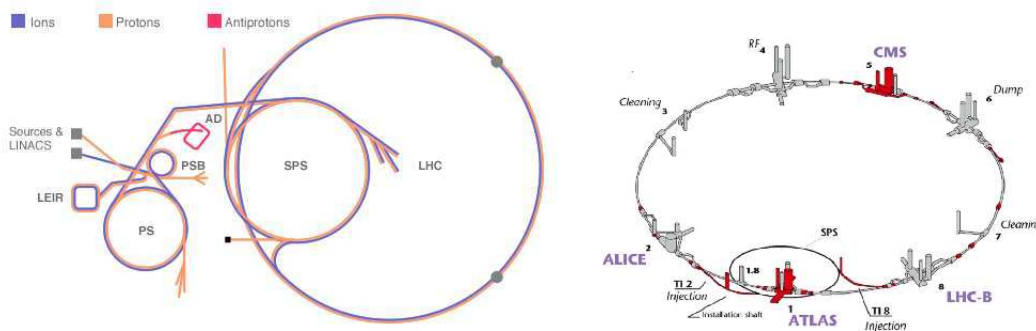


Fig. 1.1: Le complexe d'accélérateurs du Cern (à gauche) et une vue schématique du LHC et les emplacements destinés à chaque expérience (à droite).

Une caractéristique essentielle du LHC sera son importante luminosité, de cent à mille fois plus importante que celle de toutes les machines du même type qui l'ont précédé. La section efficace est fixée par l'interaction alors que la luminosité dépend de l'accélérateur et le taux observé d'événements est le produit des deux. Comme la section efficace des processus physiques que l'on veut étudier auprès du LHC est extrêmement faible, il faut une luminosité élevée, et donc beaucoup de particules par faisceau ainsi qu'une taille transverse faible.

Différentes phases sont attendues pour la mise en route du LHC : un démarrage à une très basse luminosité de $10^{31} \text{cm}^{-2} \text{s}^{-1}$ est attendue pour 2008, suivi dans les premières années d'une phase opératoire à une luminosité (dite basse) de $2 \times 10^{33} \text{cm}^{-2} \text{s}^{-1}$, pour finalement atteindre la luminosité nominale (dite aussi haute) de $10^{34} \text{cm}^{-2} \text{s}^{-1}$. A basse (haute) luminosité environ 10fb^{-1} (100fb^{-1}) de données seront collectées par année. Outre les processus intéressants, potentiellement porteurs de découvertes, se produiront inévitablement des interactions beaucoup

plus probables, avec des sections efficaces jusqu'à 10^9 fois plus importantes, bien connues et donc moins intéressantes. Pour voir quelques événements intéressants il faudra donc étudier un très grand nombre d'interactions. La section efficace d'interaction inélastique et non diffractive proton-proton au LHC est attendue à 70 mb. A haute luminosité le taux d'événements attendu est de 10^9 Hz. Les collisions peuvent être classées de la manière suivante :

1. des collisions molles, à longues distances entre deux protons. Les particules dans l'état final ont une impulsion longitudinale importante et une impulsion transverse faible. Ces événements, majoritaires, sont aussi appelés de biais minimum.
2. des collisions entre partons dures, frontales, dues à des interactions à courtes portées. Dans ces interactions le transfert d'impulsion peut être important, permettant la production dans l'état final de particules de grandes impulsions transverses et la création de nouvelles particules massives. Au LHC ces événements sont dominés par la production de jets QCD provenant de la fragmentation de quarks et de gluons dans l'état final. Ces événements rares, impliquant la production de nouvelles particules, ont toutefois une section efficace le plus souvent plusieurs ordres de grandeur plus basse que la production de jets. Ainsi les états finals hadroniques sont difficiles à utiliser pour détecter des événements rares car ils ont un bruit de fond important. On leur préférera souvent des états finals avec des désintégrations en leptons ou photons qui, si ils ont des rapports d'embranchement plus petits, ont aussi moins de bruit de fond.

Pour atteindre la luminosité requise il est nécessaire de délivrer de nombreux paquets de protons, un total de 2835, d'une densité importante, de l'ordre de 10^{11} protons par paquet, ainsi que de réduire l'espacement entre eux à 7,5 m, correspondant à une séparation temporelle de 25 ns. A chaque croisement des faisceaux de protons, le nombre d'événement moyen produits par des collisions molles à basse (haute) luminosité est d'environ quatre (vingt trois), avec un total de près de deux cents particules chargées dans la région de pseudorapidité $|\eta| < 2,5$. Ces événements sont appelés d'empilement. Les paramètres du détecteur ont été spécialement optimisés pour réduire leur impact sur les études de physique.

Les détecteurs installés auprès du LHC sont au nombre de cinq (cf. figure 1.1) : Alice [8], Atlas [9], CMS [10], LHCb [11] et Totem [12]. Les détecteurs Atlas et CMS sont des expériences généralistes, optimisées pour l'étude des interactions proton-proton et la recherche de nouvelle physique. Alice est optimisé pour l'étude des interactions entre ions lourds, alors que LHCb est quant à lui conçu pour l'étude des désintégrations des quarks b . Enfin, l'expérience Totem a pour but la mesure de la section efficace totale d'interaction proton-proton.

1.2 La collaboration Atlas

Atlas [13] est une collaboration internationale regroupant environ 2100 physiciens et ingénieurs provenant de 167 laboratoires dans 37 pays. En France [14], un département de

la Direction des Sciences de la Matière du Commissariat à l'Énergie Atomique ainsi que six Unités Mixtes de Recherche du CNRS/IN2P3 et des universités y participent.

1.3 Le détecteur Atlas

Les caractéristiques des collisions proton-proton auprès du LHC, et en particulier la luminosité importante ont imposé de très fortes contraintes sur la conception du détecteur. La luminosité s'est traduite par la nécessité de détecteurs à l'électronique extrêmement rapide, capable de traiter en temps réel le milliard d'interactions se produisant chaque seconde et d'en extraire les quelques interactions les plus intéressantes. La forte luminosité implique aussi le choix de techniques de détection et la conception de systèmes électroniques résistant aux radiations. En

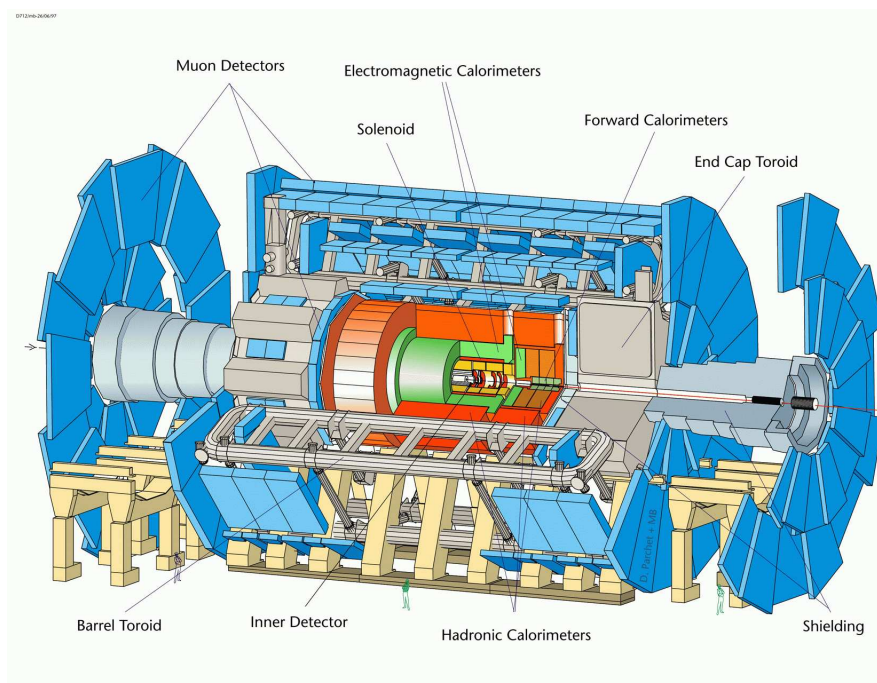


Fig. 1.2: Vue schématique du détecteur Atlas.

effet on s'attend à des doses allant de quelques Gray à des centaines de milliers de Gray selon la position par rapport au point d'interaction. Le flux annuel de neutrons pourra aller jusqu'à 10^{14} par cm^2 , ce qui est bien au-delà des doses pouvant être supportées sans problèmes par les technologies courantes. La quantité de rayonnement émise aux alentours du détecteur a par ailleurs comme conséquence pratique que l'accès ne sera pas possible en cours de fonctionnement de l'accélérateur, par exemple pour réparer ou changer une partie défectueuse. Il faut donc construire et préparer les détecteurs en mettant en œuvre des procédures de qualité et de suivi de production extrêmement sévères.

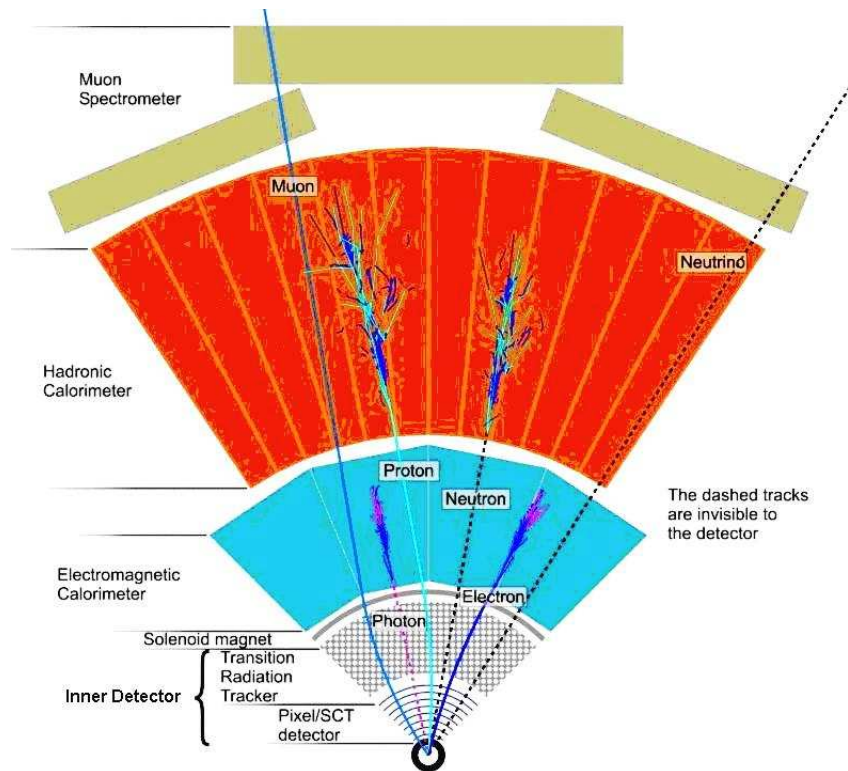


Fig. 1.3: Vue schématique du passage de particules à travers les couches du détecteur Atlas.

Le détecteur Atlas (cf. figure 1.2) est constitué d'une partie centrale cylindrique et de deux disques servant à fermer les extrémités du cylindre central. Une telle géométrie permet de réduire le plus possible la portion d'angle solide non instrumentée autour du point d'interaction. Il possède une structure en oignon caractéristique des expériences généralistes sur collisionneurs. Une vue schématique du passage de différents types de particules à travers les couches du détecteur est donnée sur la figure 1.3. En allant du point d'interaction vers l'extérieur on trouve ainsi successivement le trajectographe interne, baignant dans un champ magnétique solénoïdal, le système de calorimétrie et le spectromètre à muons, baignant dans un champ toroïdal. Le détecteur possède des mensurations impressionnantes : il pèse 7000 tonnes pour 44 mètres de long et 20 mètres de diamètre. Une description détaillée du détecteur est accessible dans le Technical Proposal [15], les Technical Design Reports [1] et le "papier détecteur" [2].

1.3.1 Le trajectographe interne

Le détecteur interne a pour but la reconstruction des traces des particules chargées, ayant une impulsion transverse minimum de 0,5 GeV/c et situées dans la zone $|\eta| < 2,5$, avec une mesure précise de leur impulsion et de leurs vertex primaires et secondaires. Il doit de plus aider à leur identification. Aussi bien ses matériaux que son électronique doivent pouvoir supporter

les hautes radiations pendant plus de dix ans et être refroidis pour évacuer la chaleur produite.

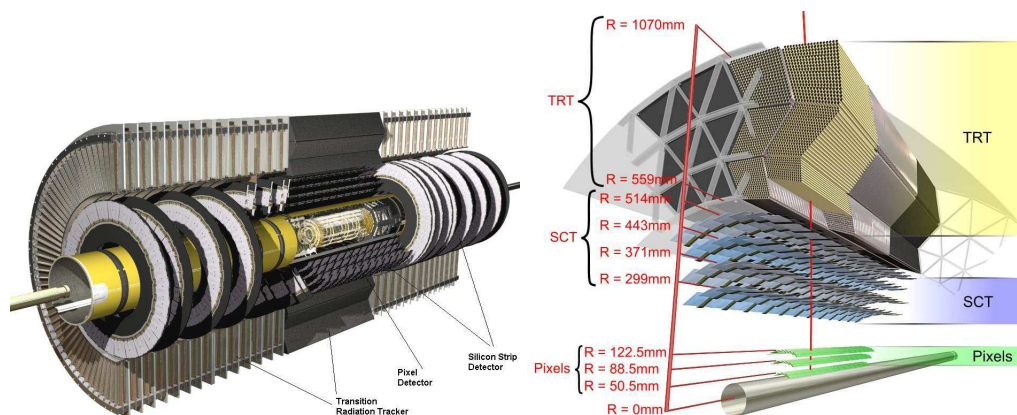


Fig. 1.4: Schéma tridimensionnel (à gauche) et coupe radiale (à droite) du trajectographe interne.

La figure 1.4 montre un dessin du détecteur interne. Il est contenu dans un cylindre d'une longueur de $\pm 3,4$ mètres, ayant un rayon de 1,15 mètres. Il est contenu dans un solénoïde supraconducteur délivrant un champ magnétique allant de 2 T en son centre, à 0,5 T en son extrémité. Ce champ assure la déviation des particules pour mesurer leur impulsion transverse. Il est composé de trois sous-détecteurs indépendants. Les principales contraintes sur ces détecteurs sont :

- une efficacité de reconstruction des traces supérieure à 95% sur toute la couverture pour des traces isolées ayant une impulsion transverse supérieure à 5 GeV/c. Le taux de fausses traces doit être plus bas que 1% du taux de signal ;
- une identification individuelle des particules dans des jets denses ;
- une mesure de l'impulsion sur une grande gamme. En dessous de $p_T = 0,5$ GeV/c les particules bouclent dans le champ magnétique et leur reconstruction n'est plus possible.
- une distinction entre les électrons et les photons qui créent des amas similaires dans le calorimètre électromagnétique ;
- une séparation entre les électrons et les jets supplémentaire à celle produite par les calorimètres. Typiquement pour une efficacité d'identification des électrons de $p_T > 20$ GeV/c on attend un facteur de rejet des jets QCD de l'ordre de 50 à 100 ;
- une identification du vertex primaire en présence de plusieurs vertex provenant d'événements de biais minimum empilés ;
- la reconstruction de vertex secondaires, déplacés d'environ 1 mm du vertex primaire, sera essentielle pour la physique du B , notamment pour la mesure du temps de vie des mésons B . Il a été montré dans les études passées [16] qu'une résolution sur le temps propre de 0,073 ps pouvait être obtenue ;
- l'étiquetage des jets de b par la reconstruction du vertex secondaire et l'identification des leptons provenant de la désintégration semileptonique d'un méson B .

L'environnement difficile et la présence d'événements d'empilement impliquent d'avoir un détecteur avec une grande granularité, de l'électronique, des services de lecture et de refroidissement situés à l'intérieur du volume du détecteur. La quantité de matériau dans le trajectographe interne est ainsi importante. La longueur de radiation X_0 intégrée a de fortes implications sur les performances de la trajectographie [2] :

- l'effet du rayonnement de freinage créé une asymétrie des amas dans le calorimètre pour les électrons, ainsi qu'une association avec une trace ayant une impulsion plus basse ;
- les photons de conversion augmentent le bruit de fond d'électrons secondaires au signal d'électrons et une augmentation de la diffusion multiple va dégrader les performances de la trajectographie et de la résolution du calorimètre ;
- l'analyse des photons directs sera plus difficile.

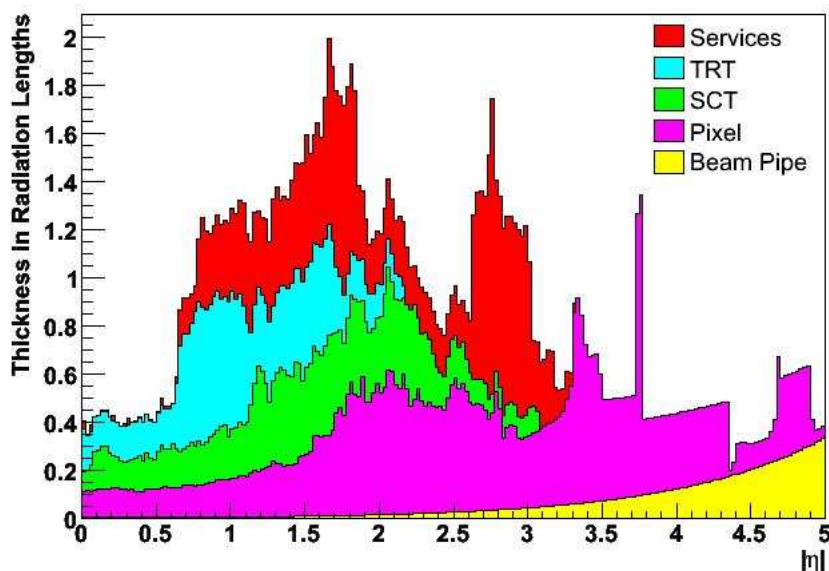


Fig. 1.5: Distribution de matière (en X_0) à la sortie du détecteur interne dans la géométrie dite "CSC". La distribution est montrée en fonction de η , intégrée en azimuth.

La figure 1.5 montre le X_0 intégré traversé par une trace rectiligne en fonction de $|\eta|$ à la sortie du détecteur interne, obtenu avec la version la plus à jour de la simulation, dite "CSC" (cf. paragraphe 1.6.1). La chose la plus frappante est la quantité de matière provenant des matériaux non actifs, tels les services et la structure à l'interface entre les parties tonneau et bouchons. De plus la quantité de matière a fortement augmenté dans les simulations au cours des années avec des simulations de plus en plus réalistes. Ceci aura des conséquences directes sur les performances de la reconstruction des électrons.

Les détecteurs de précision : pixel et silicium

Deux technologies sont utilisées pour le trajectographe de précision à silicium : un détecteur

à pixels situé au plus près du point d'interaction et un détecteur à bandes de silicium situé un peu plus loin. Le détecteur à pixels doit fournir des mesures d'une très grande granularité et d'une grande précision. Le système est composé de trois couches cylindriques de détection placées à des rayons de 5, 9 et 12 cm du faisceau et complétées par trois disques disposés à des distances de 50, 58 et 65 cm du point central. Il contient 80 millions de pixels, chacun ayant une taille de $50\ \mu\text{m}$ dans la direction $R\phi$ et $400\ \mu\text{m}$ en z . Les requêtes les plus fortes sur ce détecteur proviennent de la reconstruction des vertex primaires et secondaires. En particulier, la couche de pixels la plus proche du point d'interaction est la plus importante pour la précision de la reconstruction du paramètre d'impact des traces. Aussi cette couche est appelée couche b ou couche "vertex". Le système se trouvant à un plus grand rayon, entre 25 et 50 cm, est moins segmenté. Appelé SCT (Semi-Conductor tracker), ce détecteur est composé de près de six millions de bandes de silicium d'une longueur de 12 cm et d'une largeur de $80\ \mu\text{m}$. Il contient quatre couches cylindriques de deux plans de détection décalées d'un angle stéréo de 40 mrad, complétées de chaque côté par neuf disques. La résolution sur chaque point de mesure est d'environ $22\ \mu\text{m}$ dans le plan transverse et $580\ \mu\text{m}$ dans la direction longitudinale.

Au total c'est près de onze mesures de précision par trace qui sont ainsi obtenues, trois dans les pixels et huit dans le SCT.

Le trajectographe à rayonnement de transition

Le trajectographe à rayonnement de transition (TRT) est basé sur l'utilisation de tubes à pailles. Il permet d'allier une reconstruction des traces chargées à une capacité d'identification des électrons jusque $|\eta| < 2,0$. Les pailles sont en effet remplies d'un mélange gazeux contenant du xénon pour l'absorption des photons de rayonnement de transition créés par les électrons dans le radiateur situé entre les pailles. Comparativement aux détecteurs à silicium, ce type de détecteur permet un suivi des particules avec moins de matière et à un moindre coût. De plus cette technique est intrinsèquement résistante aux radiations.

Au total il y a environ 370000 pailles d'un diamètre de 4 mm, placées de manière radiale dans les bouchons et le long de l'axe du faisceau dans le tonneau. Ces directions ont été choisies pour maximiser le nombre de pailles traversées dans toutes les directions pointant à partir du point d'interaction. La partie centrale du TRT contient 50000 pailles de 144 cm de longueur couvrant l'intervalle $|\eta| < 0,7$. Celles-ci sont divisées en deux en leur milieu et lues à chaque bout afin de réduire le taux d'occupation. Chaque bouchon contient 125000 pailles réparties sur vingt roues. Au centre de chaque paille il y a un fil en or de $0,5\text{-}0,7\ \mu\text{m}$ couvert de $31\ \mu\text{m}$ de tungstène. Le grand nombre de points de mesure le long de la trace, en moyenne 36, avec une précision individuelle de $130\ \mu\text{m}$, permet d'améliorer les performances de l'algorithme de reconstruction des traces par recherche de formes.

Les performances de ce détecteur ont été étudiées à l'aide de plusieurs démonstrateurs dans différents tests en faisceau sur la ligne H8 provenant du SPS au Cern. Une description détaillée peut être trouvée par exemple dans [17], notamment pour l'identification des électrons. La fi-

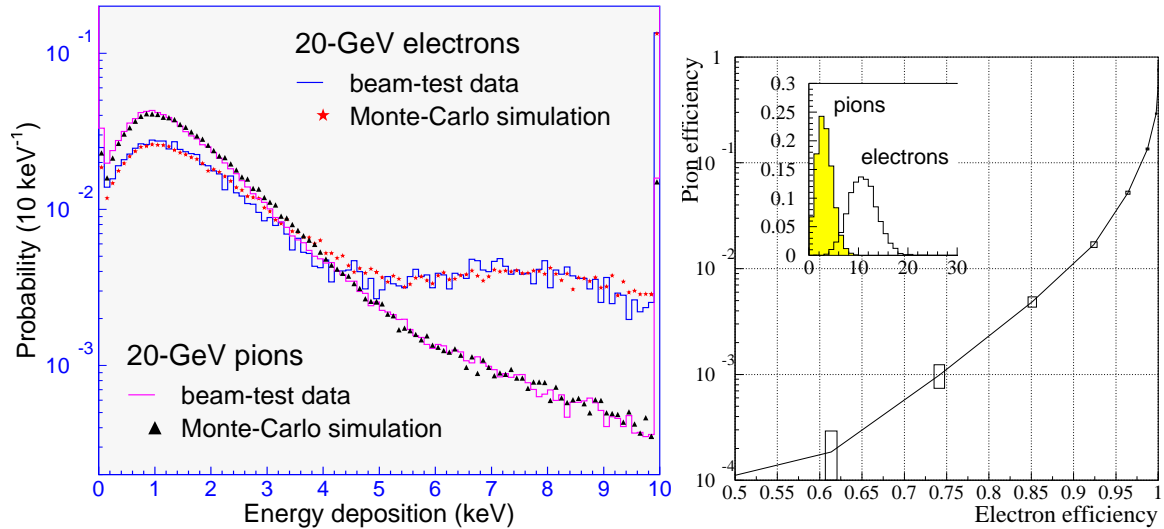


Fig. 1.6: Spectre en énergie obtenu sur une seule paille et son radiateur pour des données et de la simulation, pour des pions et des électrons de 20 GeV (à gauche) ; l'efficacité d'identification des pions en fonction de l'efficacité d'identification des électrons est montrée à droite. Figure extraite de [18].

Figure 1.6 à gauche montre, pour un électron de 20 GeV, le dépôt d'énergie d'un photon de rayonnement de transition, qui est typiquement de 8-10 keV [18]. Au contraire, pour une particule au minimum d'ionisation, comme pour un pion de même énergie, le dépôt moyen est alors seulement de 1-2 keV. On peut alors compter, pour une particule donnée, le nombre de points de mesure au-dessus d'un seuil haut, typiquement de 5 keV. La distribution de ce nombre pour des électrons et des pions est montrée dans le cartouche de la figure de droite et est clairement discriminante. En coupant sur cette variable, on peut construire une courbe donnant l'efficacité d'identification des pions en fonction de l'efficacité d'identification des électrons. Pour une efficacité d'identification des électrons de 90%, l'efficacité d'identification des pions est d'environ 1.2%. On obtient donc un facteur de rejet d'environ 75 contre des pions de 20 GeV dans un champ magnétique de 0,8 T, pour une géométrie correspondant à celle du détecteur interne d'Atlas autour de $|\eta| \sim 1,2$.

1.3.2 Les calorimètres

Une vue des calorimètres d'Atlas est présentée sur la figure 1.7. Ces détecteurs consistent en un calorimètre électromagnétique couvrant la région en pseudorapidité $|\eta| < 3,2$, un calorimètre hadronique couvrant la même région et un calorimètre avant couvrant la région $3,1 < |\eta| < 4,9$.

Le calorimètre électromagnétique

Le rôle du calorimètre électromagnétique est de mesurer avec précision l'énergie et la direc-

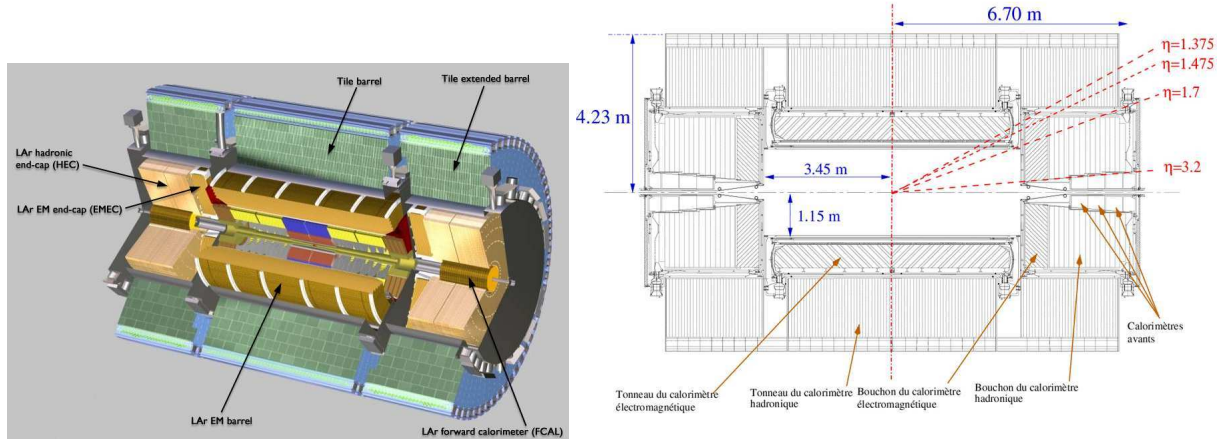


Fig. 1.7: Représentation en coupe des différents calorimètres.

tion des électrons et des photons et, dans une moindre mesure, contribuer aux mêmes mesures pour les jets. Il doit être capable de fournir des informations rapides au système de déclenchement, comme l'énergie, la position dans un amas de cellules données. La calorimétrie électromagnétique est basée sur la mesure de l'ionisation provoquée par les particules dans un empilement de plomb et d'argon liquide. Ce type de détecteur offre les avantages d'une réponse linéaire très uniforme et d'une grande résistance aux radiations.

Ce détecteur a un diamètre de 2,5 mètres et pèse environ 2000 tonnes. Il est composé d'un cylindre (tonneau) couvrant la partie $|\eta| < 1,475$ fermé par deux roues (bouchons) qui permettent une couverture angulaire jusqu'à la région $|\eta| < 3,2$. Ils sont répartis dans trois cryostats distincts, laissant une région pauvrement instrumentée, entre $1,37 < |\eta| < 1,52$, où passent les câbles et les services du détecteur interne. Chaque cellule du calorimètre comporte un absorbeur, une électrode et un espace rempli d'argon liquide. Chaque absorbeur est constitué de deux feuilles d'inox de 0,1 mm d'épaisseur, d'un tissage de fibre de verre imprégné de colle (prepreg) et de plaques de plomb prises entre des feuilles d'inox et de prepreg. Une électrode de lecture est située entre deux absorbeurs. Chaque électrode est constituée de trois couches de cuivre séparées les unes des autres par de la colle epoxy et une couche de polyimide isolant, du kapton. Les deux couches de cuivre externes sont connectées à la haute tension, tandis que les absorbeurs sont reliés à la masse. L'ionisation engendrée par les particules chargées traversant l'argon liquide est collectée par les électrodes grâce au champ électrique ainsi créé. Lorsque les charges électriques dérivent dans l'argon elles induisent par un couplage capacitif un courant sur la couche interne. Ce courant est relu par l'intermédiaire de cartes sommatriques directement enfichées sur les électrodes, puis transmis aux préamplificateurs et à l'électronique de lecture dans les châssis frontaux. La spécificité du calorimètre électromagnétique est sa structure en accordéon. Cette géométrie des électrodes et des absorbeurs permet une couverture azimuthale

homogène et élimine les espaces non instrumentés. L'alimentation haute tension et la lecture du signal se faisant à l'avant et à l'arrière du calorimètre, le câblage est réduit au minimum.

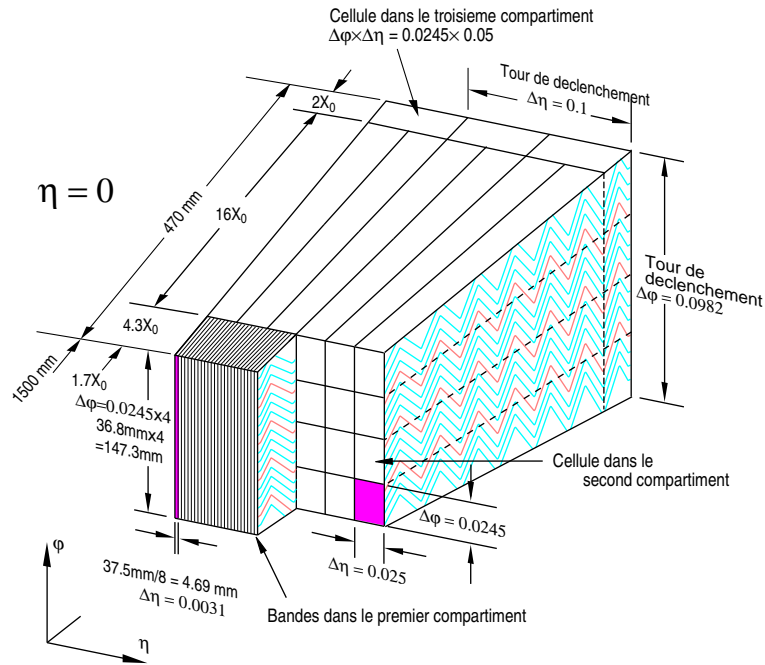


Fig. 1.8: Structure en compartiments du calorimètre électromagnétique et granularité des cellules.

Le groupe du LPNHE a été impliqué dans la construction (production, tests, assemblage et intégration) de ce sous-détecteur. On pourra se reporter aux présentations plus détaillées de l'article [19] ou faites par mes collègues, à travers, par exemple, leurs propres manuscrits d'habilitation [20] [21] [22], voire même ceux faits dans d'autres laboratoires [23] [24] [25].

Si le premier objectif d'un calorimètre électromagnétique est de mesurer précisément l'énergie déposée par les photons et les électrons, cet aspect a été peu abordé dans mes études. L'accent a été mis sur son utilisation pour leur identification en étudiant le profil latéral et longitudinal de la gerbe électromagnétique. Pour cela il faut pouvoir faire une reconstruction en trois dimensions de la gerbe et segmenter le calorimètre en cellules de lecture, chacune de ces cellules ayant sa propre électronique. Dans la direction suivant l'accordéon cette segmentation est faite en regroupant le signal de plusieurs électrodes. Dans les deux autres directions elle est obtenue par un découpage en pavé sur l'électrode. Enfin, dans la région dédiée à la physique de précision ($|\eta| < 2,5$) le calorimètre électromagnétique est divisé en quatre sous-compartiments :

- le pré-échantillonneur : il est composé d'argon liquide et d'une électrode de lecture, sans absorbeur. L'information délivrée par cette couche permet de corriger les effets de pertes d'énergie dans la matière morte située avant le calorimètre ;
- le premier compartiment : il est équipé de cellules très fines en η définies par la gravure des électrodes. Elles représentent une profondeur longitudinale de $4,3 X_0$. Seize électrodes

consécutives sont sommées suivant ϕ pour le tonneau et douze pour les bouchons. La fonction des cellules du premier compartiment est tout d'abord de mesurer la direction des gerbes électromagnétiques en combinaison avec le second compartiment. La coordonnée en ϕ n'est cependant pas utilisable car les photons convertis s'ouvrent du fait du champ magnétique et produisent des amas ayant des largeurs similaires à celles des amas de π^0 . Ce compartiment permet en particulier l'identification des photons avec une efficacité d'environ 90% tout en rejetant les π^0 d'un facteur trois. Ceci a été étudié avec des simulations [26] [27] mais aussi lors des tests en faisceau [28]. La granularité du premier compartiment est la suivante : $\Delta\eta \times \Delta\phi = 0,003 \times 0,1$. La profondeur des cellules ainsi que leur nombre par cellule de second compartiment ont été choisis de manière à optimiser le rejet des π^0 . La profondeur des cellules a ainsi été fixée à $6 X_0$ et une granularité de huit cellules.

- Le second compartiment : il est équipé de cellules ayant une granularité de $\Delta\eta \times \Delta\phi = 0,025 \times 0,025$, pour un rayon de Molière de 2,5 cm. Ce choix permet d'optimiser la taille des amas avec un compromis entre, d'une part les queues de basse énergie et la dimension longitudinale des gerbes et d'autre part le bruit électronique et l'empilement. Par ailleurs, cette granularité permet une bonne séparation entre les gerbes électromagnétiques et celles produites par les pions. Comme pour le premier compartiment, la granularité en η est définie par la gravure des électrodes. En ϕ quatre électrodes du tonneau et trois pour les bouchons sont connectées entre elles pour former une cellule de lecture. La profondeur totale du calorimètre jusqu'à la fin du second compartiment est de $22 X_0$ à $\eta = 0$. Elle augmente à grande rapidité jusqu'à atteindre $25 X_0$.
- Le troisième compartiment : La segmentation en ϕ est identique à celle du second compartiment. Etant donné que l'essentiel de l'énergie des gerbes est contenu dans le second compartiment et que le troisième n'est pas nécessaire à la mesure de la direction des gerbes, une granularité plus grossière en η est suffisante : $\Delta\eta \times \Delta\phi = 0,05 \times 0,025$.

La table 1.1 résume les caractéristiques de ces différents compartiments. Le dessin du calorimètre électromagnétique est dirigé par les requêtes de résolution en énergie et en direction dans les désintégrations du boson de Higgs en électrons et en photons, $H \rightarrow ZZ^* \rightarrow 4e$ et $H \rightarrow \gamma\gamma$. La gamme dynamique du calorimètre va d'une énergie transverse d'environ 10 MeV, correspondant au bruit électronique, à quelques TeV pour les produits de désintégration d'un boson vecteur lourd. La principale contrainte concerne la résolution en énergie :

$$\frac{\sigma}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c, \quad (1.1)$$

avec $a \sim 10\% \text{ GeV}^{1/2}$ et $c = 0,7\%$. Le paramètre a est le terme d'échantillonnage et dépend du rapport entre l'épaisseur de matière active et l'épaisseur totale en X_0 , b est le terme de bruit venant de l'électronique et de l'empilement et c est le terme constant global. Le premier terme est un reflet de la géométrie du calorimètre et peut donc être estimé précisément à l'aide de la

compartiment	tonneau		bouchons	
Nombre de compartiments et couverture en $ \eta $				
pré-échantillonneur	1	$ \eta < 1,52$	1	$1,35 < \eta < 1,8$
calorimètre	3	$ \eta < 1,35$	2	$1,375 < \eta < 1,5$
	2	$1,35 < \eta < 1,475$	3	$1,5 < \eta < 2,5$
			2	$2,5 < \eta < 3,2$
granularité $\Delta\eta \times \Delta\phi$ en fonction de $ \eta $				
pré-échantillonneur	$0,025 \times 0,1$		$0,025 \times 0,1$	
1 ^{er}	$0,025/8 \times 0,1$	$ \eta < 1,4$	$0,05 \times 0,1$	$1,375 < \eta < 1,425$
	$0,025 \times 0,1$	$1,4 < \eta < 1,475$	$0,025 \times 0,1$	$1,425 < \eta < 1,5$
			$0,025/8 \times 0,1$	$1,5 < \eta < 1,8$
			$0,025/6 \times 0,1$	$1,8 < \eta < 2,0$
			$0,025/4 \times 0,1$	$2,0 < \eta < 2,4$
			$0,025 \times 0,1$	$2,4 < \eta < 2,5$
			$0,1 \times 0,1$	$2,5 < \eta < 3,2$
2 ^{ème}	$0,025 \times 0,025$	$ \eta < 1,4$	$0,05 \times 0,1$	$1,375 < \eta < 1,425$
	$0,075 \times 0,025$	$1,4 < \eta < 1,475$	$0,025 \times 0,1$	$1,425 < \eta < 2,5$
			$0,1 \times 0,1$	$2,5 < \eta < 3,2$
3 ^{ème}	$0,05 \times 0,05$	$ \eta < 1,35$	$0,05 \times 0,025$	$1,5 < \eta < 2,5$
			$0,1 \times 0,1$	$2,5 < \eta < 3,2$

Tab. 1.1: Couverture en pseudorapidité, segmentation longitudinale et granularité en unités de $\Delta\eta \times \Delta\phi$ des cellules, pour les parties tonneau et bouchons du calorimètre électromagnétique.

simulation. Le terme b prend en compte les performances de l'électronique et de la reconstruction du signal. Le terme constant c domine à haute énergie. Il doit être meilleur que 0,7% pour remplir le cahier des charges du canal de désintégration du Higgs en deux photons. L'expérience des tests en faisceau a démontré qu'un terme constant meilleur que 0,5% peut être obtenu dans des régions de taille $\Delta\eta \times \Delta\phi = 0,2 \times 0,4$. Dans la zone d'acceptance du calorimètre utilisé pour les études de précision ($|\eta| < 2,5$), il existe 448 régions ainsi définies. Cependant de possibles non-uniformités de région à région ne peuvent être exclues. Celles liées à la mécanique du détecteur peuvent être mesurées et corrigées. Cependant, l'étalonnage final sera effectué uniquement *in situ*, puisque par exemple la quantité de matériau présent devant le calorimètre électromagnétique sera différent dans l'expérience finale par rapport aux tests en faisceau. La précision nécessaire sur cet inter-étalonnage de région à région doit être de l'ordre de 0,5%.

L'échelle absolue en énergie, ainsi que l'étalonnage *in situ*, seront déterminés par la reconstruction du boson Z^0 et à plus basse énergie par celle du J/ψ ou de Υ . La précision de l'échelle en énergie doit être de l'ordre de 0,1%, encore une fois pour respecter le cahier des charges du canal $H \rightarrow \gamma\gamma$. La linéarité de la réponse doit être meilleure que 0,5%, dans la gamme d'énergie allant jusqu'à 300 GeV, pour permettre une résolution optimale de la masse pour les modes $H \rightarrow \gamma\gamma$ et $H \rightarrow 4e$. Enfin, une précision de 0,02% sera nécessaire, à basse luminosité, pour la

mesure de la masse du boson W .

Au niveau angulaire et direction, on attend les résolutions suivantes :

$$\begin{aligned} \frac{\sigma_\eta}{E} &= \frac{3 \times 10^{-4}}{\sqrt{E(\text{GeV})}}, & \frac{\sigma_\theta(\text{mrad})}{E} &= \frac{50}{\sqrt{E(\text{GeV})}}, \\ \frac{\sigma_\phi(\text{mrad})}{E} &= \frac{15}{\sqrt{E(\text{GeV})}}, & \frac{\sigma_z(\text{cm})}{E} &= \frac{1,4}{\sqrt{E(\text{GeV})}}. \end{aligned}$$

Une autre caractéristique très importante pour la physique dans Atlas est relative à la précision avec laquelle on peut associer un dépôt calorimétrique avec un temps absolu. Cette mesure permet potentiellement de différencier des dépôts d'énergie ayant pour origine des interactions primordiales différentes. La résolution souhaitée est de l'ordre de la centaine de picosecondes (ps). Cette caractéristique a été étudiée en détail lors des tests en faisceaux. La résolution obtenue est [29] :

$$\frac{\sigma_t(\text{ps})}{E} = \frac{2500}{E(\text{GeV})} \oplus 70.$$

Le calorimètre hadronique

L'objectif principal de ce détecteur est la reconstruction des jets. Il doit être suffisamment épais pour contenir l'ensemble du dépôt en énergie des hadrons.

Le calorimètre hadronique couvre la région $|\eta| < 1,7$ pour la partie tonneau, $1,5 < |\eta| < 3,2$ pour les bouchons et $3,1 < |\eta| < 4,9$ pour le calorimètre avant. Dans sa partie centrale, 2900 tonnes d'un sandwich composé de fer et de tuiles scintillantes convertit l'énergie déposée en son sein en lumière, qui est alors détectée par des tubes photomultiplicateurs. Dans sa partie plus à l'avant, une calorimétrie faite de cuivre et d'argon liquide (HEC) a été choisie pour sa meilleure tenue aux radiations. La taille de ces 15000 cellules est $\Delta\eta \times \Delta\phi = 0,1 \times 0,1$ jusqu'à $|\eta| < 2,5$ et plus grossière ensuite, avec une taille de $0,2 \times 0,2$. Un total de dix à quinze longueurs d'absorption a été choisi pour être placé en amont du spectromètre à muons.

Sa conception est optimisée pour permettre une reconstruction de l'échelle en énergie des jets à 1% près. La résolution nominale en énergie de la calorimétrie hadronique doit atteindre $\frac{\sigma}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%$, dans la région centrale correspondant à $|\eta| < 3$. La non-linéarité de la réponse en énergie doit elle être contrôlée à quelques pourcents. La reconstruction du boson W dans le mode de désintégration hadronique est aussi importante et pourra servir à étalonner la réponse en énergie de ce détecteur. La découverte d'un Higgs de grande masse se désintégrant en $W \rightarrow \text{jetjet}$ de haute impulsion transverse implique la reconstruction de la masse jet-jet et de l'étiquetage des jets en avant.

Enfin, en ce qui concerne l'identification des électrons, le calorimètre hadronique joue un rôle important puisqu'il sert dans l'algorithme d'identification des électrons et de rejet des jets.

Le calorimètre à l'avant

Une excellente herméticité des calorimètres est requise pour la reconstruction de l'énergie manquante. La présence d'un détecteur à très petit angle, couvrant la région $3,1 < |\eta| < 4,9$, doit

permettre d'atteindre ce but. Le détecteur est un calorimètre à échantillonnage très dense au cuivre dans le premier compartiment et au tungstène dans les deux suivants. Le milieu actif est l'argon liquide et il est inséré dans les mêmes cryostats que les calorimètres électromagnétique et hadronique bouchons.

1.3.3 Le spectromètre à muons

Le rôle du spectromètre à muons est bien sur de reconstruire la trajectoire des muons et de mesurer leur impulsion, à l'aide de la déflexion produite par un champ magnétique toroïdal. Ce détecteur est équipé de chambres de haute précision pour la reconstruction de la trajectoire des muons ainsi que d'autres chambres utilisées dans le système de déclenchement. Le champ magnétique toroïdal est créé pour la partie centrale par un système de huit bobines supraconductrices montées autour du calorimètre hadronique, de diamètres internes et externes de 9,4 m et 20,1 m. Le champ magnétique maximal est de 3,9 T. Dans la partie bouchon, de chaque côté du détecteur se trouvent également huit bobines, constituant un aimant de 1,65 m de diamètre interne et de 10,7 m de diamètre externe. Trois stations cylindriques de tubes à dérive (MDT, Monitored Drift Tube) permettent de reconstruire les traces dans la partie centrale du détecteur, $|\eta| < 1,0$. Ces stations sont également utilisées dans la partie bouchon, sauf pour la station la plus proche du point d'interaction. Les CSC (Cathode Strip Chambers) sont des chambres multi-fils fonctionnant en mode proportionnel. Elles sont utilisées à grand η et à proximité du point d'interaction sur la couche la plus interne. L'information de déclenchement est donnée par des chambres à plaques résistives (RPC, Resistive Plate Chamber) dans la partie centrale et par des chambres à intervalle étroit (Thin Gap chamber) dans les parties bouchons. Ces chambres fournissent également une information sur la position.

Comme nous le verrons plus loin l'information de déclenchement sur les muons sera utilisée dans la sélection des lots d'événements pour les études sur les électrons dans les événements J/ψ ou les jets.

1.3.4 Le système de déclenchement

Le système de déclenchement est basé sur trois niveaux de sélection qui permettent de passer du taux de collisions de 40 MHz à environ 100 Hz d'événements enregistrés [30]. Le déclenchement de niveau 1 (LVL1, Level 1) est une sélection basée sur les informations d'une partie des détecteurs. Il doit effectuer une décision rapide, en $2,5 \mu\text{s}$ et n'utilise que les informations provenant des calorimètres et du spectromètre à muons. Pour les calorimètres, la granularité est réduite par la sommation de signaux de plusieurs canaux. Dans le spectromètre à muons on utilise les chambres de déclenchement. Ces régions sont appelées Régions d'Intérêt. On recherche dans celles-ci des muons, des électrons ou des photons de hautes impulsions transverses, des leptons τ se désintégrant en hadrons, de l'énergie transverse manquante et

de l'énergie totale importante. Le déclenchement possède différents menus pour un type de signal ou bien combine différents objets en coïncidence. Les seuils appliqués dans ces menus sont optimisés suivant le scénario de luminosité attendue. Une chaîne de capacités et de

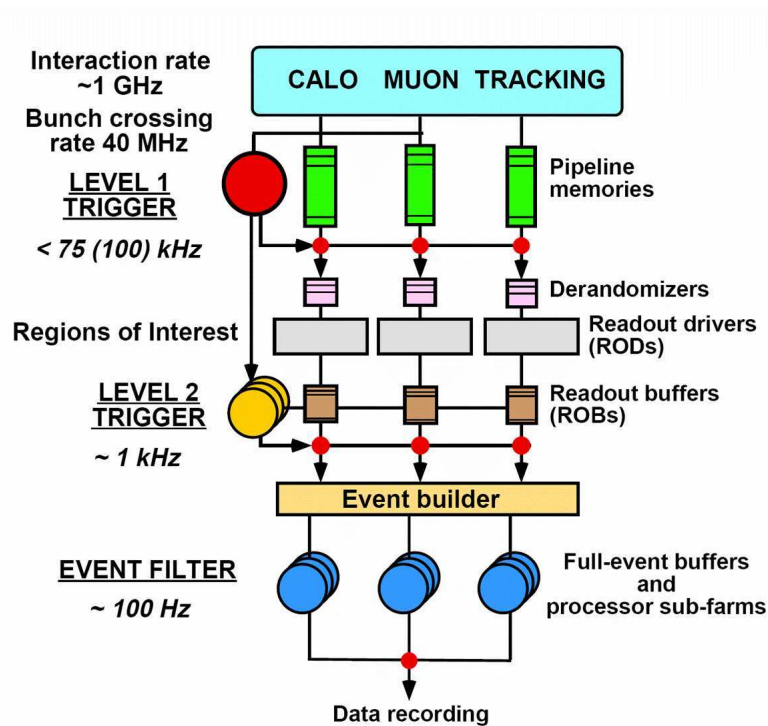


Fig. 1.9: Représentation du système de déclenchement et des taux d'événements acceptés.

résistances stocke l'information de chaque canal de lecture jusqu'à réception du signal de déclenchement. Si aucun signal n'est reçu l'information est perdue. S'il y a déclenchement, l'information est soumise aux critères du déclenchement du niveau supérieur. Le taux maximal de déclenchements de niveau un au niveau de l'électronique de lecture est de 75 kHz, une augmentation jusqu'à 100 kHz étant envisageable.

Le déclenchement de haut niveau (HLT, High Level Trigger) est constitué de deux niveaux avec une logique implémentée à un niveau purement logiciel. Les régions d'intérêt du niveau un sont utilisées pour définir les zones dans lesquelles une reconstruction plus détaillée va être entreprise. Au niveau deux (LVL2, Level 2), la granularité totale du détecteur est utilisée et les informations des détecteurs peuvent être combinées. Le taux d'événements est réduit à 1-2 kHz, avec un temps moyen d'exécution de 10 ms. Enfin, au filtre d'événements (EF, Event Filter), l'ensemble de l'événement est étudié avec des algorithmes et des méthodes d'analyse similaires à ceux utilisés hors-ligne. Il dispose des plus récentes données d'étalonnage et d'alignement ainsi que de la carte de champ magnétique. Ceci permet de réduire le taux d'événements à environ 100 Hz et permet le stockage sur bande pour une utilisation ultérieure.

Menu de déclenchement	Type d'événements	Taux (Hz)
e25i	$W \rightarrow e\nu, Z^0 \rightarrow e^+e^-,$ quark top, $H \rightarrow WW^*/ZZ^*, W', Z'$	40
2e15i	$Z^0 \rightarrow e^+e^-, H \rightarrow WW^*/ZZ^*$	<1
μ 20i	$W \rightarrow \mu\nu, Z^0 \rightarrow \mu^+\mu^-,$ quark top, $H \rightarrow WW_{x^*}/ZZ^*, W', Z'$	40
2 μ 10	$Z^0 \rightarrow \mu^+\mu^-, H \rightarrow WW^*/ZZ^*$	<1
γ 60i	prod. directe de photons, $H \rightarrow \gamma\gamma$	25
2 γ 20	$H \rightarrow \gamma\gamma$	2
j400	QCD, SUSY, nouvelles résonances	10
3j165	QCD, SUSY, nouvelles résonances	10
4j110	QCD, SUSY, nouvelles résonances	10
j70+xE70	SUSY	20
τ 35i+xE45	qqH($\tau\tau$), $W \rightarrow \tau\nu, Z^0 \rightarrow \tau^+\tau^-$, SUSY à grand $\tan(\beta)$	5
2 μ 6+ $\mu^+\mu^-$	physique du B	10
+coupures en masse		
autres		20

Tab. 1.2: Principaux menus de déclenchement pour la phase de basse luminosité, type d'événements ainsi sélectionnés et taux d'événements sélectionné. Pour chaque menu la première lettre indique le type d'objet physique sur lequel le déclenchement est opéré (e pour un électron, j pour un jet, xE pour de l'énergie manquante, le second l'énergie transverse pour laquelle le déclenchement est efficace, et enfin la lettre i indique un critère d'isolation.

Les principaux menus du système de déclenchement, pour la phase de basse luminosité, sont décrits dans la table 1.2 [31]. Pour chaque menu la première lettre indique le type d'objet physique sur lequel le déclenchement est opéré (e pour un électron, j pour un jet, xE pour de l'énergie manquante, le second l'énergie transverse pour laquelle le déclenchement est efficace, et enfin la lettre i indique un critère d'isolation. Nous verrons dans le paragraphe 3.3 le détail du déclenchement utilisé pour les canaux avec des électrons isolés.

1.4 Le modèle de calcul d'Atlas

Le modèle de calcul d'Atlas décrit les définitions et les choix concernant l'organisation, les technologies et les caractéristiques du projet logiciel. Il a été établi pour le "Computing Technical Proposal" et le Computing Technical Design Report" [32]. Depuis, il est régulièrement mis à jour [33]. La présentation de cette partie est issue de ces lectures ainsi que des références suivantes [34], [35] et [36]. Les grands principes du modèle de calcul sont brièvement donnés ici :

- un logiciel doté d'une architecture dynamique et non pas monolithique ;
- la technologie orientée objet pour le langage de programmation ;
- une organisation des moyens de calcul basé sur le schéma des grilles de calcul ;
- un stockage de données selon deux technologies complémentaires : les fichiers

séquentiels structurés, essentiellement pour les données de physique, bien adaptés aux contraintes de taille et de performance d'accès et les bases de données relationnelles pour les données qui demandent des mécanismes de recherche très sophistiqués.

1.4.1 Environnement logiciel

Athena [37] est l'environnement logiciel de la collaboration Atlas. Ses fondations reposent sur Gaudi [38], développé initialement par la collaboration LHCb et maintenant amélioré conjointement par les deux collaborations. Il doit être capable de traiter les données provenant des systèmes de déclenchement et d'acquisition et de fournir les données traitées aux physiciens ainsi que les outils d'analyse nécessaires à la production de résultats de physique. Athena a été utilisé depuis plusieurs années pour les productions et analyses de données simulées mais aussi lors de la prise de données du test en faisceau de 2004 et de rayons cosmiques actuellement en cours. Il est écrit en langage C++ et s'est inspiré du logiciel ATRECON [39], écrit lui en langage Fortran et utilisé pour produire les résultats du "Technical Design Report Detector and Physics Performance". Le code d'Athena est composé de modules spécialisés pour des tâches précises. Certains composants externes sont toutefois en Fortran, comme des générateurs Monte Carlo, ou en Java, comme Atlantis le visualisateur graphique d'événements. Athena est produit dans une version stable environ tous les six mois, mais des versions intermédiaires existent, utilisées principalement par les développeurs du code : des versions de développement à un rythme approximatif d'une tous les mois et des versions quotidiennes, construites chaque nuit.

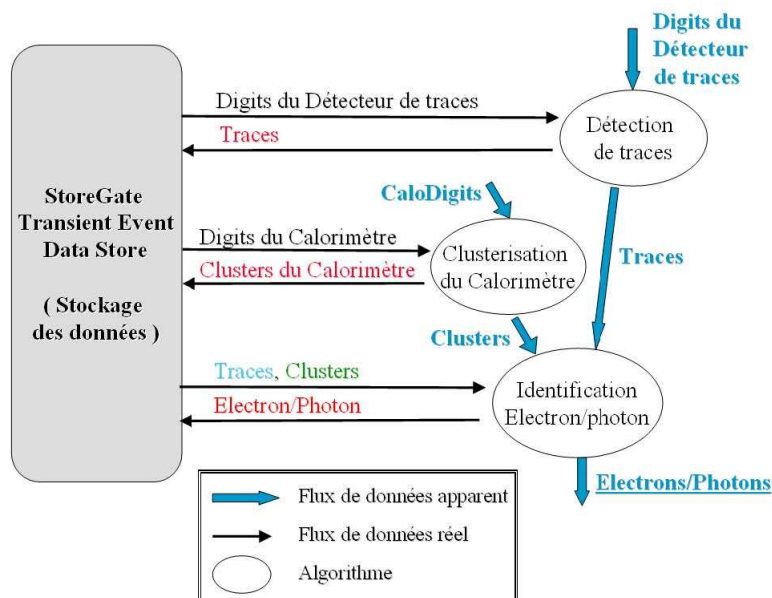


Fig. 1.10: Flux de données pour une partie de la séquence d'exécution d'Athena du calorimètre électromagnétique (figure extraite de [35]).

1.4.2 Le formatage des données

Après le déclenchement et le filtrage en ligne, les données d'Atlas sont écrites à une fréquence d'environ 200 Hz, incluant un lot de données sélectionnées pour l'étalonnage. Sachant que la taille de chaque événement brut est de l'ordre de 1,6 mega-octets, cela se traduit par environ un peta-octet par année. Les données brutes sont écrites sous forme d'un flux de données suivant un format dédié. Elles sont transférées à la ferme de calcul du Cern dans des fichiers d'une taille maximale de deux giga-octets. Chacun contiendra des événements d'une même période de prise de données, utilisant le même déclenchement dans un même remplissage de l'accélérateur.

Le stockage est alors effectué en une série d'étapes successives, chacune ayant un format spécifique, commençant par les données brutes ou les données de simulation et se poursuivant par la reconstruction jusqu'à un format plus adapté aux analyses. La table 1.3 indique la taille attendue pour les différents formats d'événements, le temps de calcul et les paramètres opérationnels du LHC qui permettent d'estimer les besoins en puissance de calcul et en espace de stockage.

Item	unité	valeur
taille des données brutes	Mo	1,6
taille des ESD (2008)	Mo	1,0
taille des ESD (2009)	Mo	0,8
taille des AOD	Mo	0,1
taille des TAG	Mo	0,01
taille de la vérité Monte Carlo	Mo	0,40
taille des ESD simulés	Mo	0,5
temps pour la reconstruction (1 ev)	kSI2k-sec	15
temps pour la simulation (1 ev)	kSI2k-sec	400
temps pour l'analyse (1 ev)	kSI2k-sec	0,5
taux d'événements après EF	Hz	200
temps d'opération	secondes/jour	50000
temps d'opération	jours/an	200
temps d'opération (2008)	jours/an	50
statistique	événements/jour	10^7
statistique	événements/an	2×10^9

Tab. 1.3: Taille attendue pour les différents formats d'événements, le temps de calcul et les paramètres opérationnels du LHC correspondants. Le temps de calcul est exprimé en SpectInt2000 (SI2K) ; un kSI2k, soit 1000 SpecInt 2000, correspond à peu près à la puissance d'un processeur Pentium Xeon 2,8 GHz.

La simulation

La simulation est une activité très consommatrice en puissance de calcul, mais produisant peu d'entrées/sorties. Typiquement, le temps de calcul attendu pour simuler un événement est de 400 kSI2k-seconde (cf. Table 1.3). La plupart des simulations sont effectuées dans une activité bien programmée par les groupes de physique. Les productions effectuées jusqu'à ce jour, ainsi que ma participation à celles-ci, seront détaillées dans le paragraphe 1.6.1. On utilise une description complète du détecteur Atlas permettant de simuler la réponse de l'appareillage au passage des particules. Depuis 2003 la collaboration utilise Geant 4 [40] pour la description de la géométrie ainsi que la simulation du détecteur.

La reconstruction

La phase de reconstruction analyse les données brutes et écrit les ESD (Event Summary Data). Cette étape permet de traduire le signal électronique d'un élément de détecteur en un point de mesure dans les trajectographes ou en un dépôt d'énergie dans une cellule des calorimètres. Ensuite, des algorithmes de recherche de forme vont construire des traces et des amas de cellules dans les calorimètres. On peut alors obtenir des informations permettant d'estimer le type de la particule associée à l'objet reconstruit : photons, électrons, muons, tau, hadrons ou neutrinos, indirectement par l'énergie transverse manquante. On doit estimer leurs quadri-vecteurs impulsion-énergie et si possible leurs vertex d'origine. Les ESD doivent contenir suffisamment de détails sur les objets pour que des raffinements de la reconstruction combinée (identification des particules, réajustement des traces, étalonnage des jets...) puissent être effectués à nouveau, au fur et à mesure des améliorations des algorithmes ou des étalonnages et alignements. Néanmoins, on ne saurait espérer refaire depuis zéro la recherche de formes dans les traces ou bien un nouvel étalonnage de toutes les cellules, car cela demanderait de sauvegarder trop de données. La reconstruction complète devra être effectuée depuis les données brutes. Le budget de taille initial pour les ESD était de 500 kilo-octets par événements. Il a été dépassé à l'heure actuelle, et les révisions des modèles utilisent 800 kilo-octets.

Une deuxième étape, issue des ESD, aboutit aux AOD (Analysis Object Data) avec des objets prêts à être analysés. Ceux-ci doivent donner un accès uniforme aux particules reconstruites (cinématique, variables d'identification...), avec des pointeurs de navigation vers les objets ESD pour une information plus détaillée. Le budget de taille pour les AOD est de 100 kilo-octets par événement. Les AOD risquent d'évoluer assez rapidement en raison d'optimisation des algorithmes d'identification, l'ajout ou le retrait de variables considérées comme indispensables ou inutiles etc. Il doit donc être possible de re-générer plusieurs fois les AOD, avec une version différente du logiciel, à partir des mêmes ESD. Lors de la création des AOD, des méta-données (TAG) seront aussi écrites. Elles permettront un accès rapide à tout événement. Le budget de taille attendu pour les TAG est de 1 kilo-octet par événement. Finalement un format appelé DPD

(Derived Physics Data), permettra de contenir les informations finales de l'analyse pour les utilisateurs sous la forme de n-tuples et d'histogrammes. Ce(s) format(s) sera à définir suivant la pratique des physiciens et pourra différer d'une analyse à une autre.

Avant la définition de ces formats, la collaboration Atlas avait développé un Ntuple Combiné (CBNT en format HBOOK [41] puis en format Root [42]) où figurent à peu près les mêmes informations qu'au niveau des ESD. Ce format de données était celui accessible au moment de mon entrée dans la collaboration ; ainsi une grande partie des analyses présentées dans ce manuscrit a été effectuée à partir de ces CBNT. En parallèle la quasi-totalité de ces études a été depuis développée à partir des AOD.

L'analyse

Les activités d'analyse sont divisées en deux classes. La première regroupe une activité routinière et pré-établie à travers les groupes de travail. Les tâches seront effectuées en utilisant de petits lots de manière chaotique, c'est à dire sans agenda établi. Les tâches s'appliquant à de grandes quantités de données seront approuvées par les groupes de travail. Le modèle de calcul suppose qu'il y a environ vingt groupes travaillant au même moment et que chacun tournera sur l'ensemble des données quatre fois dans l'année. En ce qui concerne le stockage, le modèle suppose que seuls les résultats de deux tâches seront gardés au même moment. La seconde classe d'analyse est faite par des utilisateurs individuels et est, par nature, chaotique. Elle inclut une analyse directe des AOD et d'une partie des ESD. Les analyses "dites distribuées" pourront s'effectuer avec les outils développés par la collaboration comme Ganga [43] et Panda [44].

1.4.3 La structure en Tiers et en Nuages

Les données produites par le LHC seront distribuées selon un modèle hiérarchique à quatre niveaux notés Tier-0 à Tier-3. Localisé au CERN, le Tier-0 est responsable du stockage sur bande magnétique et de la distribution des données brutes. Il produit la première reconstruction de toutes les données en plus de celles dédiées aux études d'étalonnage et de celles en vue d'une analyse rapide permettant de valider la reconstruction. Les ressources nécessaires en 2008, estimées d'après la révision du modèle de calcul [33] sont présentées dans la table 1.4. Les ressources sont exprimées pour la puissance de calcul en SpectInt2000¹⁾) et les espaces de stockage en tera-octets. Les lots de ESD, AOD et TAG sont distribués vers les Tier-1.

Les centres Tier-1, au nombre d'une dizaine, sont répartis à travers le monde. Chacun possède une capacité de stockage suffisante pour héberger environ 20% des données brutes. Afin de garantir la pérennité des données stockées dans un centre Tier-1, elles sont recopiées

¹⁾ La puissance d'un processeur de calcul est définie en SpectInt200 ; un kSI2k, soit 1000 SpecInt 2000, correspond à peu près à la puissance d'un processeur Pentium Xeon 2,8 GHz.

	calcul (MSI2K)	disque (To)	bande (To)
Tier-0	3,7	0,15	2,4
CAF	2,1	1,0	0,4
Tier-1	18,1	10,0	7,7
Tier-2	17,5	7,7	-
Total	41,4	18,9	10,5

Tab. 1.4: Puissance de calcul (en mega-SpectInt2000) et l'espace de stockage (en tera-octets) attendus en 2008 pour Atlas dans les différents Tiers.

vers un autre centre Tier-1 qui sera son partenaire privilégié. Les centres Tier-1 ont en charge la reconstruction des données et la mise à disposition des ESD pour les centres Tier-2.

Les Tier-2, au nombre d'une trentaine, sont localisés au niveau des régions et sont utilisés par une communauté allant de vingt à cents physiciens. Ils doivent fournir les moyens de traitement et de stockage permettant d'effectuer les tâches d'analyse et de simulation décidées au niveau de la collaboration. Leurs activités sont décrites dans un Memorandum Of Understanding (MoU) [45]. Ces centres ne possèdent pas de stockage de masse sur bande magnétique, coûteux et très lourd à mettre en œuvre. Les données à analyser seront donc stockées uniquement sur disque. Les données produites lors de la simulation sont recopiées vers les centres Tier-1. La puissance disponible dans les centres Tier-2 sera du même ordre que les centres Tier-1.

La facilité d'analyse du Cern (CAF), comme son nom le suggère, a pour but principal de permettre les activités d'analyse pour une grande communauté d'utilisateurs. Elle servira aussi de plateforme pour le développement de code et les études d'étalonnage. En effet elle aura un accès particulièrement simple aux données brutes de par sa proximité avec le Tier-0. Elle devrait être accessible à l'ensemble de la communauté Atlas, mais avec une grande priorité aux personnes ayant un rôle bien défini dans les développements algorithmiques, l'étalonnage et l'alignement.

Les centres Tier-3 représentent l'ensemble des autres ressources fournies à des utilisateurs locaux mais ne font pas partie d'un MoU. Les ressources ainsi disponibles devraient être du même ordre que l'ensemble des Tier-2. Leur taille va dépendre de la communauté d'utilisateur et des moyens qu'elle peut y attribuer.

En plus de cette structure en Tiers, une structure en "Nuages" implique des liens d'interdépendance entre les centres. Le Nuage Français est constitué tout d'abord d'un Tier-1, situé au CCIN2P3 à Villeurbanne [46]. Ce Tier-1 est associé aux Tier-1 de BNL (Brookhaven) [47] et de FZK (Kalsruhe) [48]. Le Nuage est ensuite constitué de plusieurs Tier-2 et Tier-3. Les activités des sites français sont coordonnées au sein de la collaboration LCG France [49]. Les ressources disponibles pour Atlas dans les différents centres peuvent être obtenues dans [50] et [51]. Trois centres Tier-2 sont situés en France, à Clermont-Ferrand (LPC) [52], à Annecy (LAPP) [53] et en région Ile de France (cf. la description de Grif au paragraphe 1.5). Trois

autres sont situés en Chine (Pékin) [54], au Japon (Tokyo) [55] et en Roumanie [56]. Enfin des Tier-3 français sont situés à Grenoble (LPSC) [57] et Marseille (CPPM) [58], ainsi que dans les différents Tier-2. Le Centre de Calcul comprend une Facilité d'Analyse (AF), de la taille d'un Tier-2, non compris dans le MoU, dédiée pour l'essentiel aux analyses effectuées par les groupes français.

1.5 *Le projet Grif*

Le projet de Grille au service de la Recherche en Ile de France (Grif) [59] a pour objectif la mise en place d'un nœud de la grille de calcul fédérant les grands centres de recherche de la région parisienne autour d'un outil commun dans le cadre de la grille européenne EGEE et de la grille mondiale WLCG. Le projet au LPNHE [60] est décrit en détails dans les notes suivantes [61] [62] [63] [64].

1.5.1 *Contexte technique*

Les projets de recherche et développement liés aux grilles sont nombreux. En France, les actions de recherche ont été fédérées au travers des "Action Concertée Incitative" Grid [65], Masses de Données [66], Grid5000 [67] et maintenant aussi des projets labélisés par l'Agence Nationale pour la Recherche. Depuis 2004, l'INRIA pilote le réseau d'excellence CoreGRID [68]. L'effort est mis sur les aspects ingénierie et organisation des architectures matérielles pour aller vers des infrastructures de production. Différents projets de grille de production ont vu le jour depuis quelques années, et sont maintenant détaillés.

La grille européenne

Depuis plusieurs années l'Europe soutient l'initiative de création d'une grille de calcul à travers différents projets se succédant sur des périodes de deux à quatre ans. Les projets MONARC [69] et DataGRID [70] ont permis de développer un démonstrateur de l'intergiciel. Les projets EGEE [71] (Enabling Grids for European E-science) (2004-2006), EGEE II (2006-2008) qui vient de s'achever, et leur successeur EGEE III doivent permettre l'arrivée à maturité d'une grille de calcul au service de toutes les communautés scientifiques européennes avec une forte ouverture vers les pays d'Amérique du nord et d'Asie. Ils ont en charge le développement de l'intergiciel, le déploiement et le suivi opérationnel de la grille de calcul. Ils contribuent au financement d'une part importante du personnel nécessaire à ces opérations. EGEE est la plus grande infrastructure de calcul distribué du monde actuellement déployée et en cours de production. La grille de calcul EGEE regroupe plus de cinquante mille processeurs et neuf peta-octets de stockage répartis sur près de cent quatre vingt nœuds en Europe, en Amérique et en Asie. EGEE compte plusieurs milliers d'utilisateurs provenant d'une dizaine de communautés scien-

tifiques internationales les plus diverses, allant de la physique des hautes énergies à la finance en passant par la biologie, la médecine, la pharmacologie, l'astrophysique, la cosmologie, la chimie, sciences de la Terre, les bibliothèques numériques, etc. Ces utilisateurs soumettent plusieurs milliers de travaux par jour à l'infrastructure qui fonctionne vingt quatre heures sur vingt quatre, sept jours sur sept.

La grille mondiale WLCG

Les physiciens des particules ont choisi de mettre en commun des outils de calculs et de stockage répartis sur l'ensemble de la planète (Europe, Asie et continent américain) en se fédérant autour du projet LHC Computing Grid (LCG) [72] dont le modèle de calcul pour l'expérience Atlas a été présenté plus haut. Ce projet, dont l'objectif est de mettre en place l'infrastructure nécessaire pour la simulation, le traitement et l'analyse des données des quatre expériences LHC, est intimement lié au projet EGEE. Il utilise et participe au développement des logiciels de grille du projet d'EGEE.

National et European Grid Initiative

Des réflexions sont en cours autour de la création d'une "National Grid Initiative" (NGI) française, point d'appui d'une "European Grid Initiative" (EGI). La fédération EGEE-France et la communauté de la physique des particules sont particulièrement actifs dans ce domaine. Cela fédère les attentes de nombreuses communautés impliquées dans des projets à dimension grille. Les besoins les plus fréquents exprimés concernent les plate-formes (opérationnelles) d'expérimentation, la formation, la pérennisation des infrastructures, l'interopérabilité et la standardisation.

Le CNRS vient de créer, au mois de septembre 2007, l'Institut des Grilles [73], unité propre du CNRS (UPS3107) rattachée au département MPPU. Ce nouvel institut fédère l'ensemble de ses actions concernant les grilles de calcul. Plus de quinze laboratoires du CNRS issus de quatre départements scientifiques et de deux instituts nationaux participent à cet institut. L'objectif est de consolider les infrastructures existantes dans le domaine des grilles de calcul, de renforcer la recherche sur ce thème et d'augmenter la synergie entre les différents acteurs. Cette action s'intègre dans une démarche plus générale de mise à disposition de moyens de calcul extrêmement puissants au service des communautés scientifiques. Cette nouvelle structure du CNRS servira de point de contact vis-à-vis des partenaires, notamment européens, pour tous les projets de grilles dans ces domaines. Il a vocation à fédérer, en accord avec les autres organismes, participants et les départements concernés.

Renater

Le groupement d'Intérêt Public Renater [74] (Réseau National de Télécommunications pour la Technologie l'Enseignement et la Recherche) a été déployé au début des années 90 pour fédérer les infrastructures de télécommunication pour la recherche et l'éducation. Les organismes membres du GIP Renater sont de grands organismes de recherche dont le CEA et le CNRS ainsi que le Ministère de l'Education Nationale de l'enseignement supérieur et de la Recherche. Plus de 800 sites sont raccordés via les réseaux de collectes régionaux au réseau national Renater, partie française du réseau européen Geant [75]. Ce réseau suit l'évolution des technologies et des capacités des infrastructures disponibles. Il est partenaire direct ou indirect de plusieurs des projets mentionnés ci-dessus. Tout cela a conduit à la mise en place de l'architecture projet de Renater où les liaisons "fibre noire" permettent d'obtenir 10 Gbit/s, en particulier en région Île de France.

La fédération Grif

Le projet de Grille au service de la Recherche en Ile de France (Grif) a pour objectif la mise en place d'un nœud de la grille de calcul fédérant les grands centres de recherche de la région parisienne autour d'un outil commun dans le cadre de la grille européenne EGEE et de la grille mondiale WLCG. La réalisation d'un démonstrateur a débuté au printemps 2005 grâce à l'association de trois laboratoires de physique des hautes énergies de la région, le CEA/IRFU²⁾ à Saclay, le Laboratoire de l'Accélérateur Linéaire d'Orsay et le Laboratoire de Physique Nucléaire et de Hautes Energies de Paris. En septembre 2005, l'Institut de Physique Nucléaire d'Orsay et le Laboratoire Leprince Ringuet de l'école Polytechnique ont rejoint le projet. A l'été 2007, le laboratoire AstroParticule et Cosmologie (APC) de Paris a contacté Grif afin de s'y intégrer. Le projet Grif apparaît donc comme une fédération de ressources informatiques de laboratoires partenaires. La nature modulaire de la grille permet en principe de constituer un agrégat de nœuds de tailles très variables. Cette grande flexibilité permet aux différents partenaires de contribuer à hauteur de leurs moyens propres. Bien qu'initié par la physique des hautes énergies, le projet Grif est ouvert à d'autres communautés de chercheurs. Par exemple, le Groupement d'intérêt Scientifique "Institut des systèmes Complexes" [76] apporte, depuis 2007, des ressources très importantes dans le cadre de Grif.

Grif a l'ambition de mettre à la disposition des physiciens des hautes énergies, dès le démarrage du LHC, l'outil indispensable à leurs travaux d'analyse et de simulation (centre Tier-2) tout en préservant leurs besoins locaux (centre Tier-3) ainsi que la participation des laboratoires au projet EGEE.

Dans le cadre du projet WLCG, la composante Tier-2 de Grif est accessible à l'ensemble des membres des expériences Atlas (2100 collaborateurs), CMS (1300 collaborateurs), LHCb

²⁾ Institut de Recherche sur les lois Fondamentales de l'Univers, anciennement Dapnia

(650 collaborateurs) et Alice (1000 collaborateurs). La communauté de physiciens membres des expériences auprès du LHC en région Île de France est d'environ 150 personnes dont près de 70 sont dans l'expérience Atlas. Une partie du travail collectif d'analyse de données et de simulation Monte Carlo de ces expériences doit être prise en charge dans la partie Tier-2 tout en préservant des ressources pour les groupes locaux. Le dimensionnement de Grif est actuellement dominé par les besoins des expériences auprès du LHC. L'objectif est d'atteindre une ressource correspondant à un Tier-2 moyen pour les quatre expériences tel que résumé dans la table 1.5, soit typiquement pour 2008 une puissance de calcul de près de 2100 kilo-SpectInt200 et un espace de stockage de plus de 550 tera-octets. L'évolution de ces besoins au cours du temps est aussi déterminée par les modèles de calcul des expériences auprès du LHC.

année	Atlas		LHC	
	calcul (kSI2K)	stockage (To)	calcul (kSI2K)	stockage (To)
2007	78	42	710	137
2008	584	259	2067	565
2009	899	443	3566	838
2010	1719	744	6653	1263
2011	2305	1043	8653	1718
2012	2891	1341	10760	2222

Tab. 1.5: Evolution attendue d'ici 2012 d'un Tier-2 moyen pour Atlas et pour les quatre expériences LHC, suivant les modèles de calcul de ces expériences. Les ressources sont la puissance de calcul (en kilo-SpectInt2000) et l'espace disque (en tera-octets).

Le développement du projet Grif peut être divisé en trois grandes phases. La première phase s'est déroulée de 2005 à la fin 2006 et a permis la réalisation d'un démonstrateur de taille réduite. La phase deux se déroule depuis et devrait finir en 2009. Elle correspond à la mise en production du site, sa mise à niveau pour les besoins du démarrage du LHC et l'exploitation des ressources pour la phase de basse luminosité. Cette ressource multi-site doit apparaître comme unique vue de la grille. Si le problème de la gestion de la charge entre les sites est relativement facile dans le contexte de la grille en déployant un élément de calcul par site, il n'en est pas de même en ce qui concerne le positionnement des données. Les modèles de calcul devront aussi être validés durant la phase deux pour préciser les besoins de la phase suivante. Les objectifs dépendront donc fortement du fonctionnement du collisionneur et des premiers résultats de physique. Une troisième phase commencera vers 2010-2011 avec l'exploitation des ressources durant la phase de haute luminosité du LHC.

Les équipes de chaque laboratoire forment une seule équipe structurée par un comité de pilotage technique. Elles se rencontrent périodiquement (tous les mois) pour faire le bilan des actions en cours et planifier celles à venir. Les équipes ont chacune leur responsable technique qui coordonne le travail entrepris par son laboratoire. Les outils modernes de déploiement, de monitoring et de publication des ressources sont mis en œuvre de façon centralisée sur des

serveurs gérés en commun afin d'éviter la duplication inutile des efforts, tout en veillant à ne pas introduire de point faible unique (single point of failure) dans le système. Les équipes techniques ne sont donc pas juxtaposées mais constituent une seule équipe solidaire et structurée par le comité de pilotage technique.

Les représentants scientifiques de chaque laboratoire se réunissent au moins une fois par trimestre pour évoquer les aspects financiers et de fonctionnement du projet. Ils désignent un responsable de projet pour une durée d'un an renouvelable, dont le rôle est la représentation du projet au niveau WLCG et EGEE. Les représentants scientifiques ont à charge de veiller à l'esprit d'ouverture du projet. Ils doivent entre autre s'assurer que les ressources sont allouées conformément à l'origine des financements et aux choix faits par les laboratoires concernant les différents groupes d'utilisateurs.

Les laboratoires ont en charge de déployer les ressources acquises. Ils fournissent les infrastructures et les installations indispensables au fonctionnement du matériel tel que les réseaux, l'alimentation électrique, la climatisation... Les frais de fonctionnement et le financement de l'augmentation de la capacité d'accueil des salles machines sont à la charge des laboratoires. Le partage des ressources financières provenant d'une aide dont la demande est faite en commun se fera au prorata des ressources déployées dans chaque laboratoire, selon les modalités définies dans un Groupement d'Intérêt Scientifique, en cours d'élaboration, afin de stimuler les efforts de financement propre.

1.5.2 *Le projet Grif au LPNHE*

Le projet Grif au LPNHE vise à répondre aux besoins informatiques, de calcul et de stockage pour l'essentiel, de thématiques variées concernant diverses communautés scientifiques.

Besoins des différentes communautés

Le groupe Atlas est constitué de 18 physiciens (permanents, thésards et post-docs). Le projet Grif doit répondre aussi bien aux besoins du groupe en termes de Tier-2 que de Tier-3 : ce sera en effet une ressource essentielle pour effectuer des analyses et simulations d'ordre privées ou de manière plus rapide. De plus le Tier-3 pourra bénéficier de l'élément de stockage du Tier-2. Le groupe LHCb du laboratoire est constitué de cinq physiciens. Le modèle de calcul de LHCb prévoit l'utilisation des Tier-2 uniquement pour la simulation Monte Carlo. De ce fait il n'y a pas de demandes d'espace disque. De plus la collaboration LHCb ne prévoit pas l'utilisation de Tier-3. Le LPNHE n'a pas de groupes Alice et CMS. Cependant, des groupes étant présents dans d'autres laboratoires membres de Grif, il est jugé souhaitable de répartir les ressources à disposition pour les quatre expériences sur l'ensemble des laboratoires plutôt que les concentrer sur un seul site. Ceci permettra plus de souplesse dans l'utilisation des ressources et favorisera l'esprit de collaboration et d'entraide. L'expérience CMS ayant des besoins simi-

laire à ceux d'Atlas, la collaboration CMS aura accès aux ressources du LPNHE. Il va de soi que les ressources mises ainsi à disposition sont compensées par un autre site.

Année	calcul (kSI2K)	stockage (To)
2008	516	141
2009	891	209
2010	1663	316
2011	2163	429
2012	2690	555

Tab. 1.6: Evolution attendue du démarrage du LHC à 2012 du projet Grif au LPNHE, suivant les modèles de calcul des expériences LHC. Les ressources sont la puissance de calcul (en kilo-SpectInt2000) et l'espace disque (en tera-octets).

En tant que membre de EGEE le LPNHE a ouvert ses ressources à des communautés présentes sur le site de Jussieu comme les sciences de la Terre avec l' "organisation virtuelle" Earth Science Research [77] à l'Institut Pierre Simon Laplace (IPSL). Le projet Grif au LPNHE a aussi pour but de favoriser l'accès aux ressources de la grille pour d'autres expériences du laboratoire. C'est le cas de la collaboration Auger qui a accès aux ressources de Grif au LPNHE depuis le début 2008. D'autres, comme Hess, sont en train de porter leurs logiciels sur la grille. Le groupe de simulations physiques des systèmes biologiques (SIMBIOPHYS) a lui-aussi des besoins de calculs importants et pourra aussi utiliser cette ressource. Enfin, le LPNHE ouvre une partie de ses ressources à d'autres laboratoires, en particulier ceux situés sur le campus de Jussieu ; citons en particulier le LPTHE Paris membre de la Fédération de Recherche sur les Interactions Fondamentales [78] avec le LPNHE et le LPT-ENS.

Répartition des ressources

Le projet prévoit que 70% des ressources sont affectées à la création d'un Tier-2 pour le projet WLCG. Les 30% restants sont dédiés à la création d'un Tier-3 pour l'expérience Atlas, aux besoins propres du LPNHE et à la participation du laboratoire dans la grille EGEE. L'évolution des ressources attendues au LPNHE est montrée dans la table 1.6. L'objectif du dimensionnement de la partie Tier-2 de Grif au LPNHE est de 25% d'un Tier-2 moyen. Pour la partie Tier-2, 65% des ressources sont dédiées à la collaboration Atlas, 20% à LHCb et 15% à CMS. Pour la partie Tier-3, 50% des ressources sont dédiées à la collaboration Atlas, les 50% restants aux besoins propres du LPNHE et à la participation du laboratoire dans la grille EGEE.

Constitution de l'équipe

L'équipe Grif du LPNHE est constituée d'ingénieurs informaticiens et d'un physicien. Elle bénéficie en outre de l'aide des services généraux du laboratoire pour tous les problèmes d'infrastructure. L'équipe technique a en charge le déploiement et la mise en œuvre du matériel

informatique. L'équipe est constituée, en 2006-2007, de deux ingénieurs informaticiens permanents, représentant une personne-temps-équivalent. Des ingénieurs sous contrat temporaire ont aussi travaillé sur le projet, représentant un peu moins d'une personne-temps-équivalent par an. L'équipe devrait être renforcée en 2008 par l'arrivée d'une personne en contrat à durée déterminée de deux ans dans le cadre du projet EGEE III.

Le responsable technique au LPNHE a pour rôle de mettre en place les moyens de fonctionnement du projet, de veiller à ce que ce dernier se déroule conformément aux prévisions et d'alerter la cellule de suivi des projets du LPNHE dans le cas contraire. Il coordonne l'action des différents intervenants. Le responsable scientifique du projet au LPNHE est a priori un chercheur impliqué dans une des expériences LHC et donc un utilisateur de la grille. Ce dernier est chargé d'informer et de coordonner les demandes des utilisateurs du laboratoire, ainsi que d'établir et de suivre les demandes budgétaires. Je suis le responsable scientifique pour le LPNHE depuis le début du projet.

1.5.3 Matériel et infrastructure

Dans un nœud de grille, il y a besoin de trois types d'ingrédients : des serveurs de calcul, des serveurs de stockage et de la connectique réseau (cf. figure 1.11). Pour chacun il existe différentes solutions technologiques qui ont à la fois des performances, des implémentations et des qualités de service différentes. Il faut tenir compte du matériel existant afin de garantir la compatibilité et des contraintes liées aux infrastructures telles que l'espace disponible, les alimentations électriques et la capacité de refroidissement de la salle machine.

Serveurs de calcul

L'élément de calcul (CE, Computing Element) est l'ensemble des services donnant accès à un ordonnanceur de tâches (système de "batch"), optimisé selon les besoins, tournant sur la ferme de calcul. Outre le mécanisme d'ordonnancement, il doit permettre une publication des ressources ainsi qu'un mécanisme d'authentification et d'autorisation, basé sur des certificats délivrés par une autorité compétente. Les serveurs de calcul doivent :

- posséder au moins 2 Go de mémoire vive par cœur pour être capable de simuler les événements du LHC ;
- avoir deux disques durs de 80 Go (système et swap1) et 160 Go (swap2 et scratch) ;
- avoir une carte réseau au Gbit/s.

Différentes technologies sont à envisager :

- l'utilisation de serveurs en châssis 1U ;
- l'utilisation de serveurs "lame" qui partagent entre eux les blocs d'alimentation, les ventilateurs, les unités disquettes, les commutateurs, les ports. Les lames sont regroupées dans un serveur 7U pouvant contenir 14 lames.

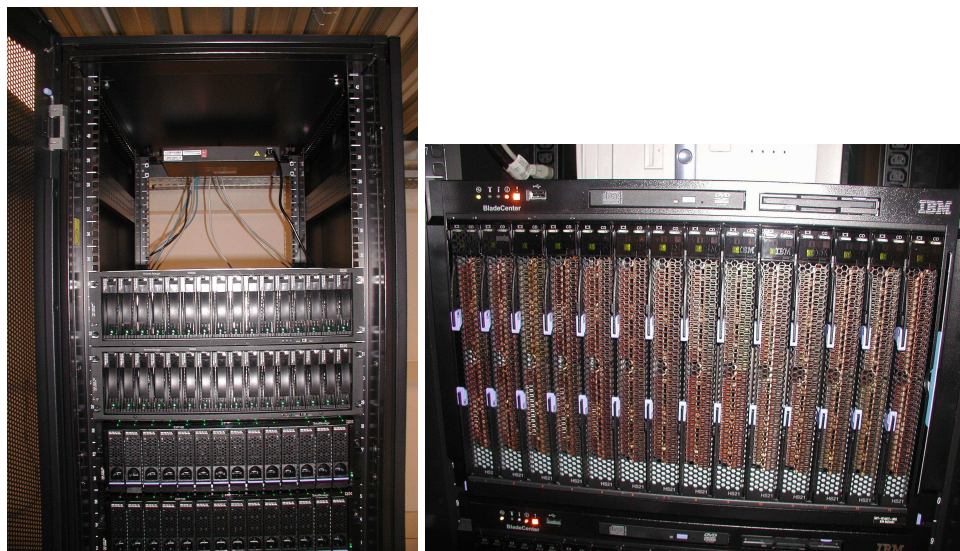


Fig. 1.11: Serveurs de stockage (à gauche) et de calcul (à droite) au LPNHE.

Un processeur 1U consomme typiquement 500 W. Dans le cas de l'utilisation de serveurs lames il faut compter environ 250 W par lame auquel il faut ajouter la consommation du serveur lui-même, typiquement 0,8 kW. Le tableau 1.7 résume les consommations électriques attendues pour différents types de technologie. Les lames quadri-cœurs bi-processeurs sont disponibles à partir de 2007-2008. Les lames 8-cœurs bi-processeurs devraient être disponibles à partir de 2010.

Matériel	Calcul (kSI2K)	Consommation électrique			
		kW/lame	kW/serveur	kW (total)	kW / kSI2K
1 bi-cœur 1U	2,8	-	-	1	0,36
14 bi-cœurs 1U	39	-	-	14	0,36
serveur bi-cœur IBM LS20	39	0,243	0,7	4,1	0,105
14 bi-cœurs bi-pro 1U	78	-	-	28	0,36
seveur bi-cœur bi-pro IBM LS21	78	0,195	0,7	3,4	0,044
serveur bi-cœur bi-pro IBM HS21	101	0,250	0,8	4,3	0,043
14 bi-cœurs bi-pro 1U	202	-	-	56	0,36
serveur quadri-cœur bi-pro	202	0,250	0,8	4,3	0,022
14 8-cœurs bi-pro 1U	404	-	-	112	0,36
serveur 8-cœur bi-pro	404	0,250	0,8	4,3	0,011

Tab. 1.7: Pour différentes technologies de serveurs de calcul on présente la puissance de calcul (kSI2K), la puissance électrique consommée par serveur, totale et moyenne. Les différentes technologies sont disponibles à partir d'une certaine date uniquement.

Le choix pour les serveurs de calcul a porté sur l'utilisation de serveurs de lames. Ils présentent un coût à l'achat élevé mais ils permettent par ailleurs de réduire les coûts d'infrastructure.

Serveurs de stockage

L'élément de stockage (SE, Storage Element) permet un stockage de masse avec uniquement des disques, en ce qui concerne les Tier-2, et un accès standardisé aux données grâce à une interface de type SRM (Storage Resource Manager), associé à système de gestion des données, DPM (Data Protection Manager) sur les Tier-2, dCache [79] ou castor [80] sur les Tier-1. Il inclut un service GridFTP permettant le transfert en entrée et en sortie vers la grille et un mécanisme d'authentification, d'autorisation et de publication des comptes. Différents types de serveurs de stockage sont disponibles :

- Direct Attached Storage (DAS) : Un DAS est un dispositif dans lequel les disques sont directement attachés aux serveurs qui les pilotent. L'accessibilité des disques est en conséquence limitée aux serveurs auxquels ils sont attachés. Des logiciels permettent d'agréger un certain nombre de serveurs dans un espace de stockage en exposant une vue plus ou moins uniforme de l'espace.
- Network Attached Storage (NAS) : un NAS est un dispositif de stockage en réseau. Il s'agit d'un serveur à part entière disposant de son propre système d'exploitation. Les caractéristiques de ce type de configuration sont assez proches de celles d'un DAS.
- Storage Area Network (SAN) : c'est un dispositif de stockage (baies de disques, dérouleurs de bandes) et de serveurs (fournissant la vue du système de fichiers) qui communiquent via un réseau de stockage dédié. Il permet une simplification de la gestion du stockage puisqu'il est inutile de modifier la configuration physique, comme le câblage par exemple, pour effectuer l'allocation/désallocation de l'espace pour un serveur. Il permet une meilleure évolutivité et flexibilité. Il présente de hautes performances aussi bien en accès séquentiel qu'en aléatoire. Cependant ce matériel est plus onéreux avec en contrepartie un meilleur service de la part des constructeurs comme la surveillance et une intervention directe faite par le constructeur sur le site.

Le choix pour les serveurs de stockage a porté dès le début du projet sur le SAN, principalement en raison de la faible taille de l'équipe technique. La technologie Raid-5 (Redundant Array of Inexpensive Disks) permet de constituer une unité de stockage à partir de plusieurs disques durs. On obtient ainsi une plus grande sécurité des données car si l'un des disques tombe en panne les données sont sauvegardées sur les autres. Le tableau 1.8 résume les consommations électriques attendues pour différents types de technologie. Les disques d'une capacité de 1 To devraient être disponibles à partir de 2008, ceux d'une capacité de 2 To à partir de 2010. Ces technologies ne sont disponibles qu'à partir d'une certaine année qui est indiquée dans la table. Il est à noter qu'il sera impératif de vérifier au préalable si les vitesses d'accès à ces disques ne seront pas

1. L'EXPÉRIENCE ATLAS

un facteur limitant et s'il n'est pas préférable d'avoir des données réparties sur plus de disques mais plus rapidement accessibles.

Matériel	Stockage	Consommation électrique	
	(To)	kW/tiroir	kW/To
Tiroir - disques 250 Go	4	0,4	0,1
Tiroir - disques 500 Go	8	0,45	0,056
Tiroir - disques 750 Go	12	0,45	0,038
Tiroir - disques 1000 Go	16	0,50	0,031
Tiroir - disques 2000 Go	32	0,50	0,016

Tab. 1.8: Pour différentes technologies de serveurs de stockage on présente l'espace de stockage (To), la puissance électrique consommée par tiroir et moyenne. Les différentes technologies sont disponibles à partir d'une certaine date uniquement.

Les aspects réseaux

Le modèle de calcul implique une bande passante de 1 Gbit/s pour les Tier-2. Avec un centre multi-site dont chacun est connecté à 1 Gbit/s, nous devrions donc dépasser la connexion minimale requise. Le projet Grif s'appuie fortement sur le réseau national Renater [74] dont la présence en région Ile de France est particulièrement dense (cf. figure 1.12). Jusqu'à l'été 2006,

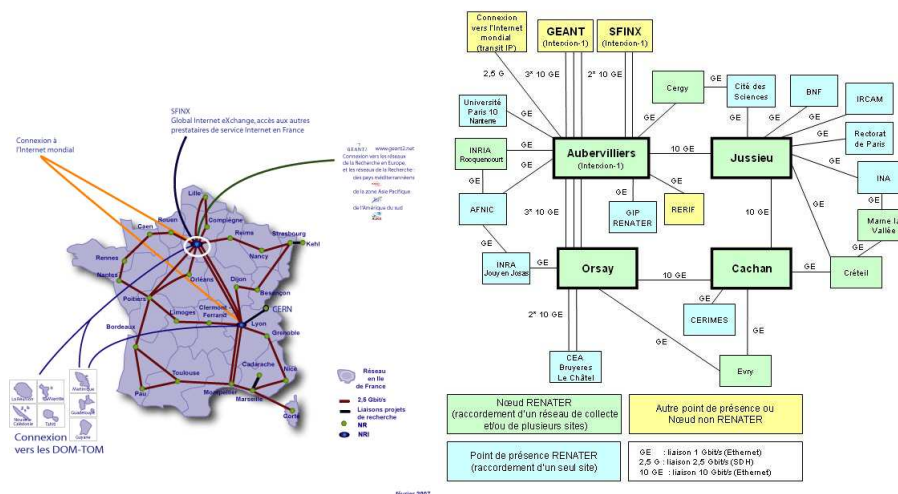


Fig. 1.12: Etat du réseau Renater à l'été 2007.

le LPNHE possédait un routeur CISCO connecté d'un côté directement à la plaque régionale Renater à Jussieu par une liaison en fibre optique à 1 Gbit/s et de l'autre côté à un commutateur du réseau du LPNHE, avec des ports entrée/sortie à 1 Gbit/s, cœur du réseau. Or, le routeur en question était géré par l'équipe réseau du CCIN2P3. Tous les types de machines et

de protocoles étaient au sein d'une même plage d'adresse et aucun filtrage n'était effectué. Les serveurs de Grif, qui sont équipés d'une carte ethernet 1 Gbit/s ont été directement connectés à un commutateur à 1 Gbit/s et ont bénéficié du transfert des données à 1 Gbits/s de bout en bout. Afin d'améliorer la sécurité, le contrôle du réseau et pour donner une certaine autonomie au laboratoire, un deuxième routeur, en cascade avec le premier géré par Lyon, a été introduit et connecté au LAN (Local Area Network) du laboratoire par un port à 1 Gbit/s. Ce second routeur est lui entièrement géré par l'équipe réseau du LPNHE. Une plage d'adresse est réservée à Grif. L'équipe Grif du LPNHE souhaite sortir les équipements informatiques Grif du LAN du LPNHE pour être connectés directement au routeur géré par le CCIN2P3 en raison du trafic nécessitant une bande passante importante. Or ce routeur n'a plus de patte de libre à 1 Gbit /s. Nous sommes en pourparlers avec le CCIN2P3 pour y remédier. Entre temps le projet d'interconnexion entre les différents sites de Grif à 10 Gbit/s est apparu. Les connexions se feront par fibre optique 10 Gbit/s (par " fibre noire "), le LAL étant le point de concentration pour Grif. Les équipements réseau, c'est-à-dire le routeur et le cœur de réseau, doivent être mis à niveau en 2008.

Infrastructures

Une salle machine doit accueillir le projet Grif au LPNHE. Pour la première année la salle machine 103 du groupe informatique du laboratoire a pu abriter les premiers éléments de calcul et de stockage. Entre temps, une salle dédiée, la salle 115 du laboratoire (cf. figure 1.13), a été choisie pour accueillir le projet jusqu'au déménagement du laboratoire, vers 2010. Cette salle a une surface d'environ 50 m². Elle est séparée en deux parties, deux tiers pour les machines et le reste, séparé par une cloison vitrée, permettant d'effectuer des opérations de suivi et de contrôle des équipements dans un environnement moins bruyant.

Le tableau 1.9 résume les besoins d'infrastructure liés à l'évolution du projet Grif sur la période 2007-2012. A la consommation électrique des serveurs de calcul il faut ajouter la consommation des serveurs de stockage. La consommation attendue dépend assez des choix technologiques effectués. Dans le tableau 1.9 il a été supposé que seule la technologie de serveurs de calcul de type lame a été achetée, avec des lames quadri-cœurs bi-processeurs jusqu'en 2009 et 8-cœurs bi-processeurs au-delà. De même pour les serveurs de stockage, des tiroirs contenant des disques d'une capacité de 1 To sont pris en compte à partir de 2008 et de 2 To à partir de 2010. Le nombre de prises électriques installées doit être dimensionnée en conséquence des besoins. Typiquement un châssis contiendra un serveur de calcul de type "lame" et est complété par un nombre suffisant de disques, sans toutefois le remplir. Chaque châssis est relié à deux prises électriques type 32 A. La salle 115 devrait pouvoir supporter l'alimentation d'une quinzaine de telles prises électriques.

Du fait de la présence de chantiers sur le campus de Jussieu, il est indispensable de fournir un réseau ondulé à la pièce Grif. Si pour l'année 2005, le matériel acheté comportait un onduleur

Année	Consommation électrique (kW)	réseau ondulé kVA	Evacuation thermique (kW)	Nombre de châssis	Dimensions (m ²)
2005	4,5	8	4	1	4
2006	15,6	25	12	2	8
2007	23,5	37	19	3	12
2008	30,1	50	24	4	16
2009	40	64	32	6	24
2010	45	72	36	7	28
2011	55	88	44	8	32
2012	65	104	52	9	36

Tab. 1.9: Dimensionnement des infrastructures nécessaires au projet Grif au LPNHE en terme de consommation électrique, de réseau ondulé, d'évacuation thermique et de dimension de la salle machine.

intégré de 8 kVA cette solution ne pouvait être continuée, aussi bien pour des raisons de coût que de sécurité. Aussi en 2006 un onduleur central d'une capacité de 50 kW (80 kVA) a été installé dans une pièce à part du laboratoire. Il permet une autonomie d'une dizaine de minutes permettant d'assurer un arrêt conforme des serveurs.

Une armoire de climatisation de 25 kW frigorifique a été installée ainsi qu'un ventilateur convecteur de 3 kW frigorifique pour la salle de suivi. Elle est raccordée au réseau d'eau glacée du campus de Jussieu. L'installation de trois ventilateurs convecteurs pourrait être effectuée pour porter la puissance frigorifique jusque 35 kW. Ceci devrait permettre de subvenir aux besoins jusqu'en 2009-2010. Pour avoir une puissance supérieure, l'installation d'une nouvelle conduite d'eau glacée sera certainement nécessaire. Une telle installation pourrait cependant être nécessaire uniquement pour l'année 2010 et dépend fortement de la date exacte de déménagement du laboratoire. Un bilan du fonctionnement de la climatisation permettra d'estimer si un tel investissement est réellement indispensable.

Un faux plancher a été installé dans la salle des serveurs avec des dalles de 600 × 600 mm de classe de charge 1A (1300 kg/m²), et environ 170 mm de hauteur du plancher. Un châssis pèse environ 500 kg et prend une surface au sol de moins de 1 m². Toutefois, les baies doivent être espacées pour permettre la circulation aussi bien des personnes que de l'air dans la salle. Aussi il faut compter environ 4 m² par baie.

La salle 115 est située dans les sous-sols de Jussieu. S'il en est de même pour la salle machine à partir de 2010 il faudra prendre en compte les risques d'inondation : non seulement ceux issus d'une rupture de canalisation, mais surtout ceux d'une crue particulièrement importante de la Seine. Dans le cas d'une crue "record" les sous-sols de Jussieu pourraient être entièrement sous les eaux, ainsi qu'une partie du rez de chaussée. Une telle inondation devrait être annoncée au moins deux jours à l'avance. Enfin, la salle doit être protégée contre les incendies.

La salle informatique qui accueillera le projet Grif dans les locaux du futur laboratoire

consiste en deux pièces de 100 m² chacune. Une puissance électrique de 200 kW sera disponible. La pièce accueillera l'ensemble des services informatiques du laboratoire.



Fig. 1.13: La salle machine accueillant le projet Grif au LPNHE.

1.5.4 Plan de financement

Estimation des coûts des matériels

Le plan de financement doit prendre en compte l'achat des différents types de matériel mais aussi les coûts d'infrastructure. Les équipements lourds concernant les aspects réseaux sont fournis par Renater et le CCIN2P3. Les serveurs ont un prix qui change au cours du temps en fonction des évolutions technologiques, avec des variations fortes certaines années. Le coût dépend aussi, bien que dans des proportions moindres, de la quantité de matériel acheté simultanément et du partenaire s'adressant au constructeur. Les membres de Grif ont effectué une étude pour les années 2005 et 2006 et ont ainsi obtenu un prix "moyen" pour l'année 2006 de 0,5 k€ par kSI2K pour les serveurs de calcul et de 2 k€ par tera-octet pour les serveurs de disque. A partir de ces prix il est possible d'estimer le coût des ressources de calcul et de stockage nécessaires pour suivre l'évolution des besoins et ainsi d'effectuer un plan de financement qui sera présenté dans le paragraphe 1.5.4. Nous avons choisi de prendre des décroissances de prix plus faibles que celles observées jusqu'ici, soit 35% par an sur le prix des serveurs de calcul et 50% pour les serveurs de stockage. Nous devons également tenir compte de la durée de vie

de ce genre de matériel qui est estimée à trois ans pour les serveurs de calcul et à quatre ans pour les serveurs de stockage. Ces durées de vie plus faible d'un an que celles considérées dans les études menées au Cern constituent une marge de sécurité supplémentaire.

Les serveurs de calcul choisis pour le projet au LPNHE sont en moyenne un peu plus chers que ceux envisagés dans le projet Grif car nous prévoyons d'utiliser préférentiellement des lames. Ceci est dicté par la volonté de minimiser la consommation électrique et l'espace physique occupé, permettant ainsi de limiter les coûts d'infrastructure. Cette politique a pour conséquence d'abaisser la taille de la ressource de Grif, aussi bien en terme de serveur de stockage que de serveur de calcul, non seulement au LPNHE mais aussi pour l'ensemble du projet. La politique de répartition des ressources financières obtenues en commun (comme le financement Sesame) s'effectue au prorata du matériel acheté et non pas du financement injecté. Il est important d'essayer de baisser les prix si nous voulons atteindre les objectifs. L'essentiel du financement est obtenu par de l'argent provenant de l'Université Paris 6. Aussi avons nous été soumis au marché public de cette université qui était remporté par une seule entreprise. Depuis 2007, le fait d'avoir un marché ouvert à plusieurs entreprises a permis de faire "jouer la concurrence". Une autre voie passe par des commandes communes des différents membres de Grif, voire avec le Tier-1 de Lyon. Ceci est en pratique difficile à réaliser car les financements arrivent à différentes époques de l'année. De plus, pour minimiser les coûts d'infrastructure, le matériel que nous désirons acheter n'est pas toujours celui que désirent les autres laboratoires. Ainsi par exemple en 2007 un prix intéressant a été obtenu au LAL pour des serveurs de calcul 1 U bénéficiant d'un prix avantageux du fait d'une commande très importante provenant des achats du LAL et du GIS ISC. Du fait de choix différents nous n'avons pas pu bénéficier de cette offre.

Le plan de financement est effectué en deux étapes.

Investissements pour la période 2005-2008

La principale source de financement provient de fonds du Ministère de l'Education Nationale de l'enseignement supérieur et de la Recherche, transmis à l'Université Pierre et Marie Curie dans le cadre de la réorganisation de son pôle de physique. Une demande de 490 k€ répartie sur quatre ans (2005-2008) a été effectuée. Les fonds demandés ont été obtenus en 2005 mais pas en 2006³⁾. Le LPNHE a contribué avec 217 k€ de crédits propres en 2006. En 2007 les groupes Atlas et SIMBIOPHYS ont contribué pour près de 30 k€. Des demandes de Bonus Qualité Recherche (BQR) de l'ordre de 15 k€ ont été effectuées aux deux universités de tutelle du LPNHE, l'Université Pierre et Marie Curie et l'Université Denis Diderot mais n'ont pas abouties. Un total de près de 30 k€ a été obtenu par le projet EGEE-France pour les années 2006 et 2007. Il est à noter que ce financement peut/doit être utilisé pour des missions,

³⁾ Au moment de la rédaction, nous sommes toujours en attente du versement des sommes pour 2007.

Sources	2005 Reçu (k€)	2006 Reçu (k€)	2007 Prévu (k€)	2008 Prévu (k€)
Ministère/UPMC	96		130	130
LPNHE	10	217		
Atlas LPNHE	2		27	
SIMBIOPHYS LPNHE			3	
Sesame			18	
GIS P2I			16	12
EGEE France		18	15	
Total	108	235	209	142
Intégrale	108	343	552	694

Tab. 1.10: Sources de financement reçues et demandées (en k€) pour le projet Grif au LPNHE.

des moyens humains (stagiaires), pour l'achat de serveurs spécifiques (comme les interfaces utilisateurs) pour les projets EGEE ou pour l'aménagement de la salle machine. En 2006 et 2007 notre fournisseur de matériel a reversé une taxe professionnelle (~ 15 k€ à chaque fois) au LPNHE, ce qui a permis le renouvellement du matériel informatique pour l'une des salles d'enseignement du laboratoire.

D'autres ressources sont issues d'une demande commune de financement auprès de la région Île de France, dite de Soutien aux Equipes Scientifiques pour l'Acquisition de Moyens Expérimentaux (Sesame [81]). Celles-ci ont pour but de permettre au projet Grif d'atteindre sa pleine taille au moment du démarrage du LHC. Un total de 138 k€ a été obtenu et versé en 2007. La répartition de ce financement suit la proportion du matériel apporté par chaque partenaire dans le projet, intégré sur la période d'utilisation. Ceci devrait représenter environ 23 k€ TTC (18 k€ HT). Une nouvelle demande a été effectuée en 2007 mais n'a pas aboutie. De même, un financement de 100 k€ (70 k€) a été obtenu en 2007 (2008) par le projet Grif auprès du GIS Physique des Deux Infinis, soit 16 k€ (12 k€) pour le LPNHE. Le tableau 1.10 indique les différentes sources de financement obtenues et attendues pour les années 2005 à 2008. Le tableau 1.11 présente le coût estimé du projet Grif au LPNHE pour la période 2005-2012.

Estimation du coût pour la période 2009-2012

Pour la période 2009-2012, le tableau 1.11 prend en compte le renouvellement du matériel. Les sommes à trouver sont du même ordre que celles obtenues pour la phase 2005-2008. Certaines ressources financières sont déjà attendues. Le projet LCG-France a jusqu'ici permis l'émergence du Tier-1 au Centre de Calcul de l'IN2P3. A partir de 2009, LCG France devrait pouvoir contribuer financièrement à la pérennisation des différents centres Tier-2 développés en France. L'enveloppe globale ainsi que la répartition sont en cours de discussion mais un total de 50 k€ par an pour le LPNHE semble raisonnable.

Pour développer la ressource, il est nécessaire d'effectuer une demande de financement

1. L'EXPÉRIENCE ATLAS

Année	calcul (k€)	stockage (k€)	Infrastructure (k€)	Total/an (k€)	Total (k€)
2005	45	53	10	108	108
2006	160	32	43	235	343
2007	56	128	6	190	533
2008	60	60	10	130	663
2009	138	43		182	845
2010	141	49		190	1035
2011	73	36		109	1144
2012	90	36		126	1270

Tab. 1.11: Coût estimé du projet Grif au LPNHE pour la période 2005-2008. Le premier chiffre donne la quantité de matériel, le chiffre entre parenthèse est le prix (TTC).

propre au niveau de l'Université Pierre et Marie Curie. Pour ce faire un Plan Pluri-Formations est en cours d'élaboration et sera déposé, en partenariat avec d'autres laboratoires, au printemps 2008, et ce pour un total d'environ 550 k€.

1.5.5 Etat actuel de Grif

Fin 2004, les deux nœuds de la grille pré-existants (IRFU et LAL) proposaient une ressource correspondant à 1 To de stockage et 20 kSI2K de puissance de calcul. En décembre 2006, à la fin de la phase démonstrateur, le projet Grif comportait 137 To de stockage et 1176 kSI2K de puissance de calcul. Le LPNHE contribue largement au succès de Grif avec son équipe technique et le déploiement d'une ressource correspondant à 25 To de stockage et 250 kSI2K de puissance de calcul. La contribution du LPNHE était particulièrement importante pour Atlas avec 160 kSI2K et 16 To, soit près de 44% des ressources de calcul et 25% de l'espace disque.

Laboratoire	Atlas Tier-2		Atlas Tier-3		Total	
	calcul (kSI2K)	stockage (To)	calcul (kSI2K)	stockage (To)	calcul (kSI2K)	stockage (To)
IPNO	-	-	-	-	300	10
IRFU	125	22	11	2	284	20
LAL	60	26	-	-	170	65
LLR	-	-	-	-	160	20
LPNHE	120	13	39	3	262	22
Total	315	61	50	5	1176	137

Tab. 1.12: Ressources disponibles en 2007, en puissance de calcul (kSI2K) et en espace de stockage (en tera-octets), pour chaque membre de Grif et au total. Les contributions respectives d'Atlas Tier-2, Atlas Tier-3 et le total sont montrées.

Pour le printemps 2008, la puissance de calcul de Grif va considérablement augmenter

comme montré sur la table 1.13. Pour la puissance de calcul, cela est dû en grande partie à la mise en route de serveurs de calcul de l’Institut des Systèmes Complexes au LAL (puissance non LHC rappelons-le), mais aussi de la progression des ressources en vue du démarrage du LHC. Un effort important sera mené tout particulièrement sur l’espace disque. D’un point de vue strictement financier il aurait été idéal d’acheter le matériel au dernier moment. Cependant, il a été nécessaire d’acquérir des compétences et donc de construire dans les laboratoires des ressources de tailles conséquentes bien avant le démarrage du LHC. De plus les demandes de moyens financiers sont à effectuer parfois plusieurs années à l’avance dans le cadre de plan pluriannuels. Rappelons aussi que depuis le démarrage du projet LCG France le calendrier du LHC s’est décalé de près d’un an. Ceci explique en grande partie le fait que dès 2006 le projet a dépassé les besoins de calcul pour 2007 et presque atteint les besoins en espace disque.

Laboratoire	calcul (kSI2K)	stockage (To)
APC	224	7
IPNO	636	24
IRFU	636	96
LAL	3050	105
LLR	640	254
LPNHE	500	115
Total	5686	601

Tab. 1.13: Ressources attendues en 2008 pour le démarrage du LHC, en puissance de calcul (kSI2K) et en espace de stockage (en tera-octets), pour chaque membre de Grif et au total.

Même si ces ressources sont importantes, il sera certainement utile au tout début du LHC d’avoir des ressources supplémentaires pour faire face aux besoins de multiples re-reconstruction imprévues des données ainsi qu’à des formats de données moins compacts que prévu. Rappelons, enfin, que les modèles de calcul impliquent une croissance constante et très importante des ressources chaque année et qu’il n’est pas inutile de “prendre un peu d’avance”.

1.6 Utilisation des ressources de calcul

Ce dernier paragraphe résume mes activités de production de données Monte Carlo. Ce travail à la base orienté uniquement sur les données qui m’intéressaient directement a peu à peu évolué vers un travail de suivi des performances des différents sites de production du Nuage Français et en premier lieu du site de Grif au LPNHE.

1.6.1 Participation aux 'Data Challenge'

Le LHC Computing Review [82] a recommandé en 2001 aux différentes expériences d'effectuer des Data Challenge (DC) de tailles et de complexité de plus en plus importantes. Ces Data Challenge sont en pratique des productions de données Monte Carlo les plus réalistes possibles⁴. Ces données permettent à la fois d'optimiser les détecteurs, d'étudier les performances combinées de ceux-ci et de préparer les analyses de physique.

Je ne décrirai pas ici en détail tous ces Data Challenge. J'ai participé directement à la production de ces données de diverses manières :

- dans le cadre du Data Challenge 1 (2002-2003), par la production des données qui m'intéressaient directement et qui seront présentées ultérieurement. Ces données ont été produites en utilisant le Centre de Calcul de Lyon, en dehors de la grille ;
- dans le cadre des Data Challenge Rome (2004) et DC3 (2006) à l'aide de production effectuées avec la grille de calcul. Ces productions, sous la responsabilité de groupes de physique, étaient encore limitées aux données m'intéressant directement ;
- dans le cadre du Computing System Commissioning (2007) à l'aide de l'"exécuteur" LyonDQ2 [83], et maintenant Panda [84], installés au Centre de Calcul. Il s'agit de systèmes permettant de contrôler la production centrale d'Atlas, distribuée sur tous les centres de calcul, c'est à dire l'envoi et le suivi des tâches correspondant à différentes étapes de la production Monte Carlo : génération, simulation, empilement, reconstruction. Le travail est réparti entre différentes personnes en charge de l'exécuteur qui doivent surveiller la bonne exécution des tâches et l'efficacité des sites les recevant.

Le travail de production Monte Carlo est ainsi passé d'une tâche individuelle, sur un centre de calcul, à une tâche collective, distribuée sur de nombreux centres.

1.6.2 Publication des ressources des sites de production français

Une étape importante pour pouvoir contrôler le bon fonctionnement d'un site ainsi que les productions qui y sont effectuées est de publier les informations nécessaires sur le portail d'opérations de la grille EGEE [85]. Ceci a été effectué par les équipes techniques de Grif. La figure 1.14 donne la fraction des tâches envoyées en France par les différents sites. Près de 13% de ces tâches ont été envoyées depuis janvier 2006 sur les sites de Grif. On rappelle que la fédération étant vue de l'extérieur comme une ressource unique on ne peut détailler les contributions de chaque membre⁵. Parmi les différentes tâches, environ 28% provenaient de la collaboration Atlas.

⁴ La version de la géométrie utilisée dans la simulation est très dépendante des versions de Data Challenge utilisées et en général correspond à un détecteur de plus en plus réaliste. Dans l'usage courant on parle de géométrie "TDR", "DC1" etc, faisant référence à la version de la simulation utilisée effectivement durant le Data Challenge correspondant.

⁵ Un travail est en cours au laboratoire pour y remédier.

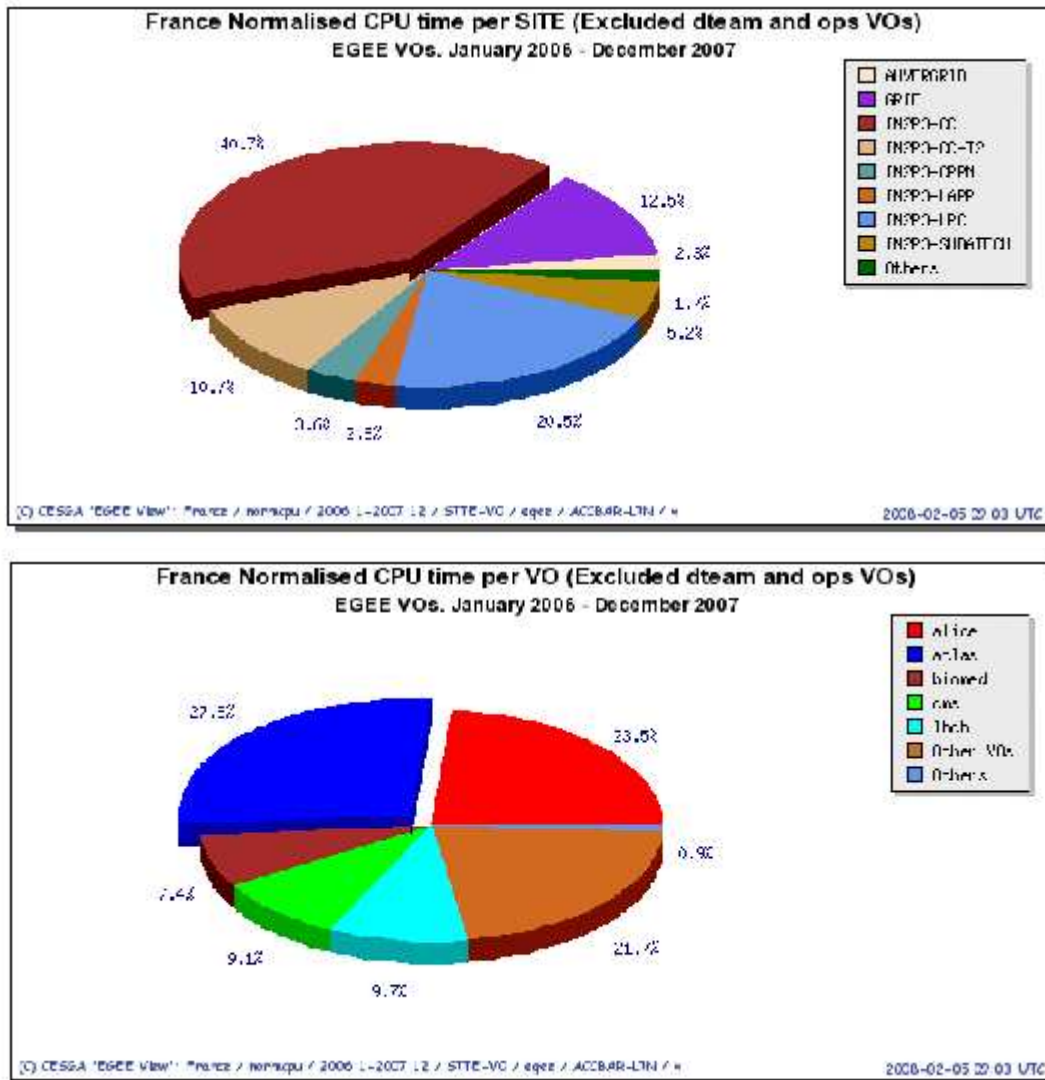


Fig. 1.14: Publication des tâches effectuées de janvier 2006 à juillet 2007 par les centres de calcul français en fonction des différentes organisations virtuelles (en haut) et fonction des sites de production (en bas).

La figure 1.15 montre la progression de la puissance de calcul consommée par les différentes organisations virtuelles utilisant les ressources de Grif pour chaque mois écoulé depuis le mois de janvier 2006. Depuis janvier 2006, l'expérience Atlas a utilisé environ 34% des ressources de Grif. La compréhension précise de cette courbe nécessite une compréhension détaillée à la fois de l'état de chaque site de Grif (quelles sont les ressources disponibles et quand ?) et les besoins des différentes expériences (sont-elles en production ou non ?).

La collaboration Atlas France s'est structurée de manière à obtenir une collaboration efficace en matière de calcul entre les membres des laboratoires français du CEA et de l'IN2P3. Le mandat du groupe Calcul Atlas France [86] est en particulier d'analyser les besoins des groupes

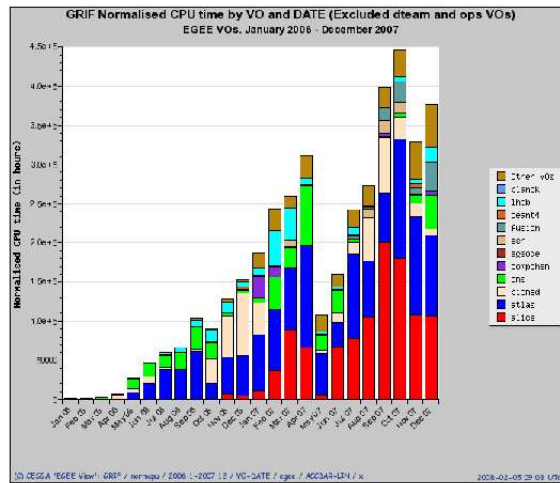


Fig. 1.15: Puissance de calcul (en heures) utilisée par chaque organisation virtuelle utilisant les ressources de Grif pour chaque mois écoulé depuis le mois de janvier 2006.

français sur tous les aspects du calcul hors-ligne dans Atlas, de communiquer les besoins au Centre de Calcul et de fournir des recommandations sur l'évolution et l'utilisation du Centre de Calcul comme Tier-1 d'Atlas, ainsi que des centres Tier-2 et Tier-3. Le groupe a aussi mis en place la production centralisée de données de simulation.

Le site Grif du LPNHE a largement contribué à l'effort de production de données Monte Carlo de l'expérience Atlas. Il a fonctionné avec une très bonne efficacité. Un certain nombre de dysfonctionnements ont bien sûr eu lieu au début et des périodes d'arrêt pour maintenance ont lieu de temps en temps. Le fonctionnement permanent du site, aussi bien pour la production d'Atlas que pour celle d'autres collaborations, est d'ores et déjà riche d'enseignements et permettra au site et à l'équipe en charge d'améliorer les performances en vue du démarrage du LHC. Il reste cependant encore beaucoup de travail pour améliorer les performances de la production d'Atlas, qui est la convolution des efficacités des différents sites, des entrées et sorties vers les Tier-1 et des systèmes de suivi des tâches. La grille de calcul a plus que prouvé son intérêt pour la production et la reconstruction des données de simulation. Un défi tout aussi important, et demandant une implication croissante des groupes d'analyse, concerne la migration de nos analyses de physique des centres habituels (Cern et Lyon) vers une utilisation massive de la grille. C'est sur ce chantier que doivent désormais tout particulièrement porter nos efforts.

2. CANAUX À ÉLECTRONS DANS ATLAS

Les événements contenant des électrons dans l'état final représentent une signature caractéristique de nombreuses analyses de physique du programme de l'expérience Atlas. Les électrons isolés de haute impulsion transverse permettent un déclenchement relativement aisé. Une excellente connaissance de l'énergie des électrons est nécessaire pour effectuer des mesures de précision, comme par exemple avec les bosons de jauge électrofaibles (cf. paragraphe 2.1). Une bonne résolution en énergie et une bonne linéarité doivent être atteintes dans une gamme en énergie allant de quelques GeV à plusieurs TeV. Un autre exemple, développé au paragraphe 2.2, concerne la recherche du boson de Higgs. Son observation requiert des performances excellentes sur les électrons de basses et hautes énergies. Le paragraphe 2.2.1 évoquera l'importance de la reconstruction des électrons pour la recherche de nouvelle physique, comme par exemple pour résoudre de possibles résonances étroites de nouvelles particules au sein d'un bruit de fond important. Une source importante d'électrons provient des désintégrations semi-leptoniques des quarks lourds, sujet développé au paragraphe 2.3, donnant accès à un vaste programme d'études. A l'opposé un grand nombre de ces canaux de physique seront utilisés pour parfaire notre connaissance des performances des détecteurs mis en œuvre comme le trajectographe interne et le calorimètre électromagnétique. L'étude des résonances se désintégrant en paires e^+e^- devrait permettre d'extraire un lot pur d'électrons de haute impulsion transverse avec le boson Z mais aussi à plus basses énergies avec le J/ψ ou le Υ . Une mesure absolue de l'énergie pourra être étalonnée avec ces événements, en exploitant la connaissance précise de la masse de ces particules. Enfin, la reconstruction d'électrons dans les jets signera la présence d'un quark b .

Toutefois, tous ces électrons sont difficiles à identifier au LHC du fait d'un très grand bruit de fond QCD provenant de jets de hautes impulsions transverses. Nous étudierons donc au travers de différents chapitres les algorithmes développés pour reconstruire et identifier ces électrons, tout en rejetant efficacement ce bruit de fond, à la fois avec des données simulées mais aussi avec de vraies données issues du test en faisceau combiné qui a eu lieu en 2004 au Cern.

2.1 *Physique des bosons de jauge électrofaibles*

La production de leptons isolés et à grandes impulsions transverses est caractéristique de processus du Modèle Standard faisant intervenir la production ou l'échange des bosons de jauge W^\pm et Z^0 . Ces processus constituent les principaux bruits de fonds de la recherche du boson

2. CANAUX À ÉLECTRONS DANS ATLAS

de Higgs et d'une nouvelle physique dans les voies leptoniques. Les taux de production pour les principaux processus au LHC sont présentés dans la figure 2.1, pour la phase de haute luminosité. Les désintégrations leptoniques des bosons de jauge constituent ainsi la principale source d'électrons primaires et isolés.

La masse du W est l'un des paramètres fondamentaux du Modèle Standard et sa mesure précise, $\delta m_W/m_W \sim 2 \times 10^{-4}$, est à l'origine des contraintes les plus fortes sur les performances du détecteur Atlas. La section efficace de production $pp \rightarrow W + X$ avec $W \rightarrow lv$ où $l = e, \mu$ est 30 nb. A la luminosité instantanée de $2 \times 10^{33} \text{cm}^{-2}\text{s}^{-1}$, le taux attendu pour la production d'électrons issus de désintégrations de W^\pm est d'environ 20 Hz. Avec une statistique accumulée de 10fb^{-1} la précision statistique attendue sur la masse du W est de 2 MeV [16]. L'une des principales sources d'incertitudes systématiques est l'échelle en énergie des électrons.

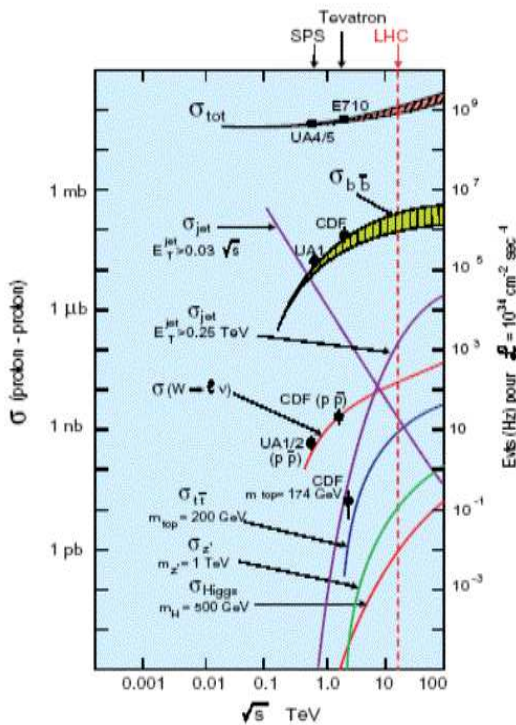


Fig. 2.1: Sections efficaces inclusives proton-(anti)proton pour les principaux processus physiques au LHC. Les taux en événements par seconde et par an pour une luminosité instantanée de $10^{34} \text{cm}^{-2}\text{s}^{-1}$ sont indiqués sur l'échelle de droite.

De par sa signature relativement simple et propre, l'étude des événements Z , dans la première phase de prise de données du LHC, a plusieurs motivations importantes. Le taux de production de paires d'électrons issues de désintégrations de Z^0 est de 3,5 Hz. Ce processus fournit donc une source abondante d'électrons utilisés pour l'étude des performances de l'algorithme de reconstruction; en particulier, il permet de mesurer avec les données l'efficacité d'identification des électrons au niveau du déclenchement, de la reconstruction et de l'analyse. Il permet d'étudier l'alignement entre le trajectographe interne et le calorimètre électromagnétique. La détermination d'une échelle d'énergie commune pour le détecteur interne et le calorimètre électromagnétique est d'une importance cruciale pour toutes les analyses utilisant les informations provenant de ces deux détecteurs. Ce sera le cas lors de l'utilisation des électrons pour le calcul d'une masse invariante ou pour la mesure de l'impulsion transverse. Une possibilité d'établir une telle échelle d'énergie commune réside justement dans les désintégrations de résonances en paires électron-positron.

Il sera par ailleurs important de vérifier que les différentes régions des détecteurs répondent de la même manière, de telle sorte que des événements provenant d'un processus donné peuvent être traités de manière cohérente, indépendamment d'où ils ont été mesurés. La connaissance précise de la masse du Z permet ainsi d'inter-étalonner le calorimètre électromagnétique. On contrôle

alors l'uniformité de la réponse du détecteur. Enfin ces processus fournissent des sources abondantes d'électrons qui sont utilisés pour l'étalonnage et la détermination des efficacités de reconstruction. Le chapitre 3 décrit l'algorithme de reconstruction des électrons isolés, nécessaires à ces études, ainsi que ses performances.

2.2 Recherche du boson de Higgs

La recherche du boson de Higgs est l'un des enjeux majeurs du LHC. Il peut être produit par fusion de gluons, par fusion de bosons vecteurs ou en association avec un boson W ou Z. D'autres processus moins importants sont la fusion $t\bar{t}$ et la radiation de bosons de Higgs par les bosons vecteurs. Selon sa masse, différents processus pourraient permettre son observation.

Dans la gamme de masse $80 \text{ GeV}/c^2 < m_H < 120 \text{ GeV}/c^2$, le mode de désintégration $H \rightarrow b\bar{b}$ est dominant. Ce canal souffre cependant d'un bruit de fond important mais qui peut être réduit considérablement si le boson de Higgs est produit en association avec un boson vecteur ou une paire $t\bar{t}$. La reconstruction des électrons intervient alors dans la reconstruction du W, du quark top ou même dans l'étiquetage des quarks b . Un autre canal dans cette fenêtre en masse est le canal $H \rightarrow \gamma\gamma$. Il présente un faible rapport d'embranchement et un bruit de fond important. Il reste tout de même un canal très prometteur comme montré par exemple dans la thèse [87]. Pour ce canal, un bon véto des traces d'électrons est nécessaire et plus encore une bonne reconstruction des électrons de conversion.

Dans la gamme de masse $120 \text{ GeV}/c^2 < m_H < 2m_Z \simeq$, le canal $H \rightarrow ZZ^* \rightarrow 4l^\pm$ offre une signature propre due à l'impulsion transverse des leptons. Une bonne résolution en énergie est néanmoins nécessaire. Le bruit de fond principal vient de la production directe de paires ZZ^* . De plus amples détails sur ce canal peuvent être obtenus par exemple dans la thèse [88]. On notera que le spectre en impulsion transverse de l'électron de plus bas p_T peut descendre jusque 7 GeV. Il est donc important d'être capable d'identifier les électrons jusqu'à des impulsions transverses faibles pour augmenter la statistique dans ce canal.

Dans la fenêtre $2m_Z < m_H < 800 \text{ GeV}/c^2$ s'ouvre le canal $H \rightarrow ZZ \rightarrow 4l^\pm$, où le boson de Higgs se désintègre en deux Z réels, ce qui permet d'avoir un rapport d'embranchement élevé et un bruit de fond très faible. Pour des masses encore plus importantes le canal $H \rightarrow WW \rightarrow l^+l^- \nu\nu$ a un rapport d'embranchement six fois plus important que le canal $H \rightarrow ZZ \rightarrow 4l^\pm$. Il est caractérisé par deux leptons ayant une impulsion transverse importante ainsi que de l'énergie transverse manquante.

On retrouvera les performances de l'algorithme de reconstruction des électrons isolés dans le chapitre 3, ainsi que l'étiquetage des jets de b par la présence d'un électron comme dans le canal $WH \rightarrow \mu\nu b\bar{b}$ dans le chapitre 5.

2.2.1 Physique au-delà du Modèle Standard

Dans le domaine de la physique au-delà du Modèle Standard, les leptons produits dans les désintégrations des charginos et des neutralinos fournissent des signatures claires pour la recherche de la supersymétrie. De plus, dans la désintégration $\tilde{\chi}^{\pm}\tilde{\chi}^0 \rightarrow 3l + E_T^{miss}$, le spectre en p_T du lepton le plus mou descend jusque 5 GeV/c. Les théories de grande unification sont une extension possible au Modèle Standard. Elles impliquent l'existence d'un boson de jauge neutre supplémentaire, le Z' . Les canaux les plus prometteurs sont ceux où la particule se désintègre en paires de leptons. On trouvera une étude de ce canal dans [89]. Ce canal possède plusieurs bruits de fond, notamment le Drell-Yan, $q\bar{q} \rightarrow (Z, \gamma) \rightarrow l^+l^-$. L'étude de ce même continuum Drell-Yan à des masses invariantes de l'ordre du TeV permettra la recherche de sous-structures et de résonances.

2.3 Physique des saveurs lourdes

L'étude la production de quarks lourds (charme, beauté et top), mettant en jeu leurs masses, permet d'aborder les prédictions du Modèle Standard à un ordre perturbatif élevé et de donner de nouvelles informations sur les effets de QCD à longue distance. La section efficace totale de production du charme est d'environ 7,8 mb, celle de la production de b est de 0,5 mb. La section efficace de production d'une paire de top n'est quant à elle que d'environ 0,8 mb.

2.3.1 La production de quarks b et c

La production de charme et de beauté sera l'une des plus copieuses sources de hadrons au LHC. Le nombre d'états de quarkonium comme le J/ψ ou l' Υ sera de même très important. Les modes de désintégrations en deux leptons permettra de séparer de manière relativement aisée ces événements d'un bruit de fond hadronique très important. En complément du $Z^0 \rightarrow e^+e^-$, ces désintégrations permettront d'établir une échelle d'énergie commune au détecteur interne et au calorimètre électromagnétique. Elles permettront de vérifier l'étalonnage et la linéarité du calorimètre, et ce pour des électrons dans une gamme d'énergie plus basse. En plus de ces motivations expérimentales, la production d'états de quarkonium lourds est d'une grande importance pour les études de phénoménologie pour tester la chromodynamique quantique. En particulier des mesures de CDF [90] ont montré il y a quelques années un excès dans la production directe de J/ψ qui ont amené une révision des modèles de production. De plus amples détails sur les études sur la production de quarkonia dans Atlas peuvent être trouvés dans [91]. Dans les canaux de désintégration en diélectrons, la section efficace intégrée pour la production directe de J/ψ (Υ) après le déclenchement au niveau un sur les deux muons, avec un $p_T > 3$ GeV/c, est de 116 nb (48 nb). La détection de ces événements nécessite la mise au point d'un

déclenchement spécifique à partir de deux amas électromagnétiques de bas E_T au niveau un¹⁾.

Programme de physique du B

La fraction d'événements d'Atlas contenant une paire $b\bar{b}$ est attendue à environ 1%. Durant la période de basse luminosité, les paires $b\bar{b}$ seront produites à un taux d'environ 10^6 Hz, mais seules environ 10 Hz pourront être écrites sur bandes pour le programme de physique du B. Le système de déclenchement pour la physique du B est essentiellement développé autour du déclenchement sur deux muons dans l'état final. Ces paires $b\bar{b}$ sélectionnées sont produites typiquement à des impulsions transverses moyennes dans la région centrale en rapidité. En comparaison des usines à B, le LHC peut étudier les désintégrations rares du B_s^0 ($B_s^0 \rightarrow \phi\gamma$, $B_s^0 \rightarrow \phi\mu^+\mu^- \dots$) et du Λ_b ($\Lambda_b \rightarrow \Lambda J/\psi$, $\Lambda_b \rightarrow \Lambda\mu^+\mu^- \dots$), les désintégrations rares des mésons $B_{d,s}^0$ ayant des rapports d'embranchement particulièrement faibles ($\leq 10^{-9}$) comme le $B_{d,s}^0 \rightarrow \gamma\mu^+\mu^-$ et le $B_{d,s}^0 \rightarrow \mu^+\mu^-$ ainsi que la mesure, avec une grande précision, des distributions différentielles pour les désintégrations semi-leptoniques des $B_{d,s}^0$ qui sont très sensibles aux extensions du Modèle Standard. Si la plupart des études ont été jusqu'à lors effectuées avec des muons dans l'état final, de nombreuses études similaires pourront l'être en utilisant les désintégrations $J/\psi \rightarrow e^+e^-$. En effet, comme les mésons B sont produits par paires, il est alors possible d'utiliser le muon provenant de la désintégration semi-leptonique d'un des B et de regarder la désintégration $J/\psi \rightarrow e^+e^-$ dans l'événement. C'est le cas des analyses sur les études de violation de CP, qui pourront être étalonnées au tout début grâce à la mesure précise de $\sin(2\beta)$ avec $B_d^0 \rightarrow J/\psi K_s$. Ensuite, des canaux plus intéressants, sensibles à la nouvelle physique, comme le $B_s^0 \rightarrow J/\psi\phi$, permettront d'accéder à l'angle γ , à la différence de largeur $\Delta\Gamma_s$ et au paramètre de mélange $x_s = \Delta m_s/\Gamma_s$. Enfin, et en particulier pour ce qui nous concerne, ces J/ψ seront très importants pour les études d'étalonnage du calorimètre électromagnétique car ils sont produits en grand nombre et le déclenchement est plus aisé que pour la production directe. Un bref résumé du programme de physique du B peut être obtenu dans la référence [92] et dans le chapitre correspondant du TDR [93].

Etiquetage des quarks b

Les désintégrations du quark top, du boson de Higgs et de particules supersymétriques impliquent souvent la présence d'un quark b dans l'état final. Une technique efficace pour sélectionner des événements avec des hadrons beaux consiste à étiqueter la présence d'un lepton issu de la désintégration semileptonique du quark b. Les leptons produits sont alors de petite impulsion transverse, localisés dans des jets. Comme nous le verrons au chapitre 5 cette méthode permet d'obtenir de bonnes performances. Toutefois elle est fortement pénalisée par le faible

¹⁾ Les premiers résultats sont attendus dans la référence [108].

rapport d'embranchement $b \rightarrow e$ ou $b \rightarrow c \rightarrow e^2$). Cette méthode servira ainsi à compléter les performances obtenues avec les méthodes utilisant la reconstruction des vertex secondaires, ou tout au moins à les vérifier.

La reconstruction des électrons dans les jets ne se limite cependant pas à l'étiquetage des jets de b . En effet, les désintégrations semileptoniques des quarks b et c seront l'une des principales sources déléctrons, permettant de nombreuses études d'étalonnage.

2.3.2 Production de quarks top

Dans un collisionneur proton-proton comme le LHC, les quarks top peuvent être produits soit par production célibataire électrofaible, comme récemment mis en évidence au Tevatron, soit par production par paires, à travers l'interaction forte. Avec une section efficace de production de paires $t\bar{t}$ prédite à 833 pb, le LHC va produire environ huit millions de paires $t\bar{t}$ par an, dans sa phase de basse luminosité. Une mesure précise de la masse de cette particule permettrait d'accroître la précision des incertitudes dans les calculs théoriques, ainsi que de contraindre la masse du boson de Higgs. Dans le Modèle Standard le quark top se désintègre quasi exclusivement en $t \rightarrow Wb$. Dans environ les deux tiers des cas, les deux bosons W vont se désintégrer hadroniquement en paires de jets ou au moins l'un des W se désintégrant en $W \rightarrow \tau\nu$. Ces événements, s'ils sont nombreux, sont toutefois difficiles à extraire parmi un bruit de fond de jets QCD important. Le canal "en or" est celui où un seul des W se désintègre semileptoniquement, en un électron ou un muon. Un lepton énergétique associé à une grande énergie transverse manquante est une signature très claire permettant un rejet important du bruit de fond. Le canal où les deux W se désintègrent semileptoniquement a un rapport d'embranchement faible (environ 5%), mais aussi un bruit de fond peu important, essentiellement dû au Z . Ici la présence des deux neutrinos dans l'état final rend la mesure de la masse invariante plus difficile.

Le groupe du LPNHE a entamé voici près de deux ans un programme de recherche visant à mesurer, avec les premières données du LHC, la masse du quark top dans le canal en dileptons. Le chapitre 6 résume quelques travaux préparatoires auxquels j'ai participé.

²⁾ Les mesures des rapports d'embranchement [94] pour ces désintégrations sont $Br(b \rightarrow \bar{l}) = (10,71 \pm 0,22)\%$, $Br(b \rightarrow c \rightarrow l^+) = (8,01 \pm 0,18)\%$ et $Br(b \rightarrow \bar{c} \rightarrow t) = (1,62^{+0,44}_{-0,36})\%$.

3. SÉPARATION ÉLECTRONS/JETS

Comme nous l'avons vu dans le chapitre précédent une reconstruction efficace et précise des électrons isolés, ayant une impulsion transverse importante, typiquement plus grande que 20 GeV/c, est nécessaire pour des analyses cruciales pour le programme de physique du LHC, comme le top, W, le Z ou le $H \rightarrow 4e$, pour ne citer qu'eux. L'environnement difficile dans lequel les analyses seront effectuées implique des contraintes sévères sur la reconstruction des électrons et des photons. Un facteur de rejet des jets de l'ordre de 10^5 , tout en maintenant une bonne identification des électrons isolés, est nécessaire pour de nombreuses recherches. Ce chapitre est consacré à la description de l'algorithme de reconstruction des électrons isolés et à la présentation de ses performances en terme d'identification des électrons et de pouvoir de rejet des jets. Le travail original à la base de ce chapitre a été effectué en 2003-2004 dans le cadre du Data Challenge 1 [95] [96]. L'analyse des données produites lors du Data Challenge "Computing System Commissioning" était en cours au moment de la rédaction de ce manuscrit. J'ai donc décidé de résumer les résultats obtenus avec les données DC1 et à la fin du chapitre d'évoquer les travaux et améliorations attendues avec les données CSC.

3.1 Les données utilisées

Pour les études sur la séparation électrons/jets (cf. paragraphe 3.4) on utilise des lots de données produits dans le cadre du Data Challenge 1 (cf. paragraphe 1.6), avec Pythia 6.203 [97]. Les événements ont été simulés avec Geant 3 [98], basé sur la version 3.2.1 du programme Fortran d'Atlas ATLASIM/Dice [99]. Des événements d'empilement ont été ajoutés pour les scénarii de basse et haute luminosité. Ces données ont été reconstruites avec la version d'Athena 6.0.4. On utilise des électrons célibataires ayant une pseudo-rapacité $|\eta| < 2,5$ et une énergie transverse $E_T=25$ (30) GeV, simulés avec une luminosité basse (haute) (cf. table 3.1). Pour le bruit de fond on utilise des événements de di-jets QCD incluant des radiations dans les états initiaux et finals. Au niveau partonique, chaque jet doit avoir une impulsion transverse minimale $p_T \geq 17$ (25) GeV/c et une rapidité comprise dans $\pm 2,7$. En supplément des événements di-jets, d'autres processus ont été inclus tels que les événements avec des photons directs ou la production de bosons Z et W. Les événements qui ne passeraient pas le déclenchement de niveau un sont immédiatement rejetés : au moins une région de taille $\Delta\eta \times \Delta\phi = 0,12 \times 0,12$ doit être trouvée dans laquelle la somme des énergies transverses de toutes les particules stables, à l'exception des neutrinos et des muons, doit être supérieure à 17 GeV. Seuls 8,33% (10,20%)

3. SÉPARATION ÉLECTRONS/JETS

données DC1				
données	luminosité	E_T [GeV]	lot de données	nombre d'événements
électrons célibataires	basse	= 25	002026	5000
dijets	basse	≥ 17	002000	13750000
électrons célibataires	haute	= 30	002021	10000
dijets	haute	≥ 25	002001	4416150
données CSC				
électrons célibataires	-	= 25	007003	7000
électrons célibataires	-	= 20	007002	50000
électrons célibataires	-	= 15	007001	35000
électrons célibataires	-	= 10	007000	50000
électrons célibataires	-	= 7 – 80	007020	50000
dijets	-	≥ 17	005802	
$Z^0 \rightarrow e^+e^-$	-		005152	25000

Tab. 3.1: Lots de données utilisés, luminosité correspondante, seuil en E_T , numéro associé pour la production et nombre d'événements simulés.

de tous les événements produits à basse (haute) luminosité sont acceptés et simulés.

En ce qui concerne les évolutions récentes de l'algorithme de reconstruction des électrons présentées au paragraphe 3.5.1, on utilise des données produites dans le cadre du "Computing System Commissioning". Les données ont été produites avec la version d'Athena 12.0.6. Elles ont été simulées avec la version 12.0.6 avec Geant 4 [40] sans événements d'empilement superposés. Enfin elles ont été reconstruites avec la version d'Athena 12.0.6. On utilise des lots d'électrons célibataires produits à différentes énergies transverses (cf. tableau 3.1) ainsi qu'un lot d'événements dijets. Enfin on utilise un lot d'événements contenant la production de bosons Z^0 produits avec Mc@NLO [100]. Un filtre implique que les événements produits contiennent au moins un électron (ou positron) ayant une pseudo-rapacité $|\eta| < 2,7$ et une impulsion transverse $p_T > 10$ GeV/c. Un lot d'événements de bruit de fond de jets QCD est aussi utilisé.

3.2 Reconstruction des électrons

L'algorithme principal de reconstruction des électrons est basé sur la reconstruction d'amas de cellules dans le calorimètre électromagnétique et de leur association avec un trace chargée reconstruite dans le trajectographe interne. Cet algorithme a été développé au niveau hors-ligne pour permettre une reconstruction optimale de la quadri-impulsion des électrons (et des photons) pour toute la gamme en impulsion et en rapidité, quelque soit la luminosité. Pour ce faire l'information des deux détecteurs est combinée. Ainsi les électrons seront identifiés avec le moins de bruit de fond possible, sachant que l'optimum à atteindre entre l'efficacité d'identification et le rejet du bruit de fond est en fin de compte dépendant de l'analyse effectuée.

3.2.1 Algorithme de reconstruction

Les amas électromagnétiques sont recherchés avec un algorithme de fenêtre glissante qui balaye le calorimètre avec une fenêtre de taille fixe dans laquelle on somme, à chaque position, les énergies des cellules sur toute la profondeur du calorimètre. La granularité de la deuxième couche du calorimètre électromagnétique, où les électrons et les photons déposent l'essentiel de leur énergie est de $\Delta\eta \times \Delta\phi = 0,025 \times 0,025$. Des amas rectangulaires sont définis en nombre de cellules. Leur dimension a été optimisée pour la meilleure reconstruction des électrons et des photons isolés. La fenêtre glissante choisie a une taille de $\Delta\eta \times \Delta\phi = 0.125 \times 0.125$, soit 5×5 en unités de cellules du second compartiment. Lorsque l'énergie comprise dans la fenêtre est maximale dans les deux directions η et ϕ , le barycentre de l'amas 3×3 au centre de la fenêtre 5×5 définit la graine de départ. Un seuil en énergie transverse est appliqué sur cette tour 3×3 à $E_T > 3$ GeV. Les amas finaux sont des amas rectangulaires définis autour de cette graine : pour les électrons les dimensions optimales [16] sont de 3×7 , pour les photons de 3×5 dans le tonneau et 5×5 dans les bouchons. Ces dimensions ont été choisies comme un compromis pour capturer la plus grande partie de l'énergie de la particule initiale tout en minimisant les effets de bruit électronique, d'empilement et d'événement sous-jacent. L'extension est généralement plus grande en ϕ pour pouvoir capturer les photons de rayonnement de freinage ainsi que les électrons provenant de conversion de photon.

La reconstruction des traces dans le détecteur interne est réalisée jusqu'à une pseudorapidité de 2,5 avec trois couches de pixels, dont la plus interne, la couche b contribue le plus à la précision du paramètre d'impact, puis quatre couches de détecteur à silicium à bandes stéréo, suivi par un détecteur à pailles donnant typiquement 36 mesures par temps de dérive, en plus d'une identification des électrons par rayonnement de transition des électrons.

Les électrons et les photons sont modélisés par le même objet à cette étape de la reconstruction complète qui est appelé communément "egamma". La combinaison des différentes informations reconstruites par chaque sous-détecteur permet d'établir un certain nombre d'hypothèses de travail quant à l'identité des particules ayant interagies.

3.2.2 Energie et position de la particule incidente

Connaissant l'énergie reconstruite dans les cellules, les amas permettent de reconstruire l'énergie et la position de la particule incidente. L'énergie dans les différentes couches du calorimètre est la somme des énergies reconstruites dans les cellules. La partie essentielle étant déposée dans le plomb, élément passif, un poids doit être appliqué à l'énergie mesurée dans l'argon liquide. Ce poids est l'inverse de la fraction d'échantillonnage et l'énergie totale de la particule incidente devient donc $E_{tot} = E_{act} / f_{samp}$, avec $f_{samp} = E_{act} / (E_{act} + E_{pas})$, où E_{tot} est l'énergie totale, f_{samp} la fraction d'échantillonnage, E_{act} et E_{pas} les énergies déposées dans les éléments actifs et passifs respectivement. La mesure de l'énergie reconstruite est une somme

pondérée des énergies reconstruites dans chaque compartiment :

$$E_{rec} = \lambda (b + w_0 E_0 + E_1 + E_2 + w_3 E_3), \quad (3.1)$$

où E_0 , E_1 , E_2 et E_3 sont les énergies reconstruites dans chaque compartiment et λ , b , w_0 et w_3 sont des paramètres dépendants de η appelés poids longitudinaux. La fraction d'échantillonnage variant selon la couche considérée, l'application de ces poids permet une optimisation à la fois de la linéarité et de la résolution en énergie. Ces poids sont obtenus par une comparaison avec l'énergie vraie dans les données Monte Carlo.

La position de l'électron (ou du photon) est reconstruite en ϕ et en η en déterminant le barycentre des positions des cellules de l'amas pondérées par l'énergie contenue dans les cellules : $\eta = \sum_i (\eta_i E_i) / E_{tot}$ et $\phi = \sum_i (\phi_i E_i) / E_{tot}$, où ϕ_i et η_i sont les positions de la cellule i de l'amas considéré, E_i l'énergie reconstruite dans la cellule. Elle est déterminée dans chaque compartiment du calorimètre. Pour mesurer avec précision la position et l'énergie d'une particule dans le calorimètre électromagnétique, certains effets doivent être corrigés. Ces effets peuvent être dus à la structure du calorimètre, à sa composition ou encore aux algorithmes utilisés pour former les amas de cellules. Ils peuvent en plus dépendre de l'énergie ou du type de particules (électrons ou photons). La mesure de la direction en η de la particule se fait en ajoutant la mesure de la profondeur de la gerbe.

3.2.3 Reconstruction des électrons de conversion

La reconstruction des électrons de conversion ne fait pas à vraiment parler partie des études d'identification auxquelles je me suis intéressées jusqu'ici. Elle est par contre particulièrement importante pour les études sur le canal $H \rightarrow \gamma\gamma$, auxquelles j'ai par ailleurs participé dans le cadre du Data Challenge 1 avec les études de séparation photon/jet [26] [27]. Toutefois, de manière à être un peu complet, il est utile à ce moment de l'exposé d'évoquer cette classe particulière d'électrons.

Dans Atlas, environ 30% des photons se convertissent dans la cavité interne ($R < 115$ cm). Près de 75% de ces conversions s'effectuent dans le détecteur interne ($R < 80$ cm et $|z| < 280$ cm) dans lequel elles peuvent être identifiées. Les conversions s'effectuant en dehors de cette région sont peu dangereux car les électrons ainsi créés sont alors peu déviés dans la direction ϕ avant d'entrer dans le calorimètre, et ainsi ressemblent à des photons non convertis. On peut alors reconstruire des particules de faibles masses invariantes. L'algorithme recherche des paires de traces chargées de charges opposées et effectue un ajustement à un vertex commun. Si le χ^2 de cet ajustement est inférieur à un seuil (typiquement 200), alors la trace est considérée comme étant une conversion et donc a priori plutôt un objet de type "Photon". Suivant la valeur du rayon de conversion, les traces sont reconstruites soit en utilisant l'ensemble du détecteur interne ($R_C \leq 40$ cm pour les conversions rapides) ou en utilisant seulement les informations du TRT ($40 \leq R_C \leq 80$ cm pour les conversions tardives). Au final cet algorithme permet de reconstruire

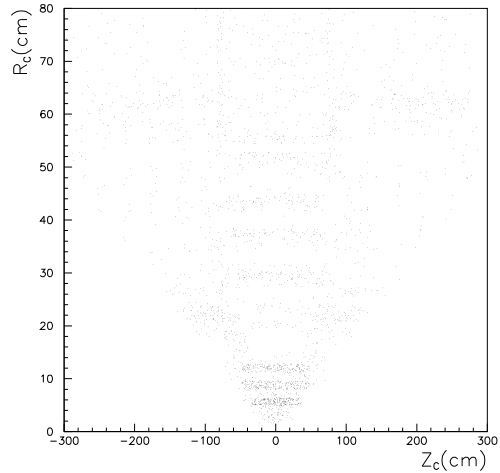


Fig. 3.1: Vertex de désintégration de photons convertis dans des données $H \rightarrow \gamma\gamma DC1$.

environ 80% des photons de conversion, ayant une impulsion transverse de l'ordre de 50 GeV/c et se convertissant dans un rayon plus petit que 40 cm.

3.3 Sélection des électrons au niveau du déclenchement

Le système de déclenchement pour les électrons inclut le menu à objet simple $em25i$ ($em30i$) à basse (haute) luminosité et le menu pour les objets doubles $2em15i$ ($2em20i$). Les critères suivants sont requis au niveau 1 [30] :

- l'énergie transverse reconstruite dans le calorimètre électromagnétique dans une tour de déclenchement 1×2 , correspondant à une taille $\Delta \eta \times \Delta \phi = 0,1 \times 0,1$, doit être plus grand que 19 (20) GeV pour un déclenchement sur un objet simple. Le seuil requis en E_T pour un objet double est de 9 (16) GeV.
- l'isolation électromagnétique, reconstruite comme la somme sur 12 tours autour de cœur d'une tour de déclenchement de taille 2×2 doit être plus petite que 3 (5 GeV). Le seuil appliqué pour les objets doubles est de 8 (4) GeV.
- l'isolation hadronique, reconstruite comme la somme sur 12 tours autour de cœur d'une tour de déclenchement de taille 2×2 doit être plus petite que 2 (3 GeV). Le seuil appliqué pour les objets doubles est de 4 (2) GeV.
- le cœur de la fuite hadronique, reconstruite dans une tour de déclenchement de taille 2×2 derrière l'amas électromagnétique est plus petit ou égal à 2 (2) GeV. Le seuil appliqué pour les objets doubles est de 4 (2) GeV.

Dans la version d'Athena utilisée à l'époque (6.0.4), le code de trajectographie du niveau 2 (LVL2) avait un problème impliquant une perte d'environ 4% des traces (cf. [101] pour plus

de détails) et ne pouvait donc être utilisé. Au niveau du filtre d'événements (EF) les électrons sont sélectionnés à partir d'un seuil en énergie transverse E_T et d'après la forme des gerbes électromagnétiques et des informations du détecteur interne. Les critères de sélection sont quasi-identiques à ceux utilisés dans l'analyse hors ligne mais avec des seuils moins élevés. Le seuil en E_T est de 22 (27) GeV pour le déclenchement sur des objets simples et de 12,5 (17) GeV pour des objets doubles. Les performances du déclenchement pour des électrons célibataires ont été étudiées pour les deux scénarii de luminosité. L'efficacité d'identification des électrons est de $\sim 78\%$. Ceci correspond à une efficacité de $\sim 81\%$ par rapport au niveau LVL1.

3.4 Identification des électrons

L'identification des électrons est basée sur l'utilisation de variables caractérisant le passage des électrons dans le détecteur interne et le développement des gerbes électromagnétiques dans les calorimètres. Des coupures sur ces variables, ou sur une fonction de vraisemblance construite à partir d'elles, sont optimisées pour permettre de garder une haute efficacité d'identification des électrons, y compris en présence d'événements d'empilement. De plus ces critères doivent permettre de rejeter avec la plus grande efficacité possible les pions et autres hadrons, en particulier formant un jet.

3.4.1 Variables caractérisant les électrons

L'identification des électrons isolés est basée essentiellement sur l'utilisation du calorimètre électromagnétique (ECAL) pour trouver des événements contenant un amas électromagnétique de grande énergie transverse. Une trace de bonne qualité est alors sélectionnée et doit correspondre à cet amas à la fois spatialement (η, ϕ) mais aussi en énergie. Finalement l'information du TRT est utilisée pour améliorer l'identification des électrons.

La figure 3.2 montre les distributions en E_T , après le déclenchement de niveau un, pour tous les amas reconstruits dans le calorimètre électromagnétique, dans des lots d'électrons célibataires et dans un lot d'événements di-jets. Les amas de bas E_T , correspondant aux particules de basses énergies provenant des événements d'empilement, sont rejetés. Pour le lots de jets le premier pic correspond au seuil de reconstruction des amas qui est d'environ 5 GeV et le second provient de l'application du déclenchement de niveau un qui enrichit le lot avec des événements ayant au moins un amas au dessus du seuil de déclenchement. Seuls les amas électromagnétiques ayant une énergie transverse $E_T > 17$ (25) GeV à basse (haute) luminosité sont considérés par la suite.

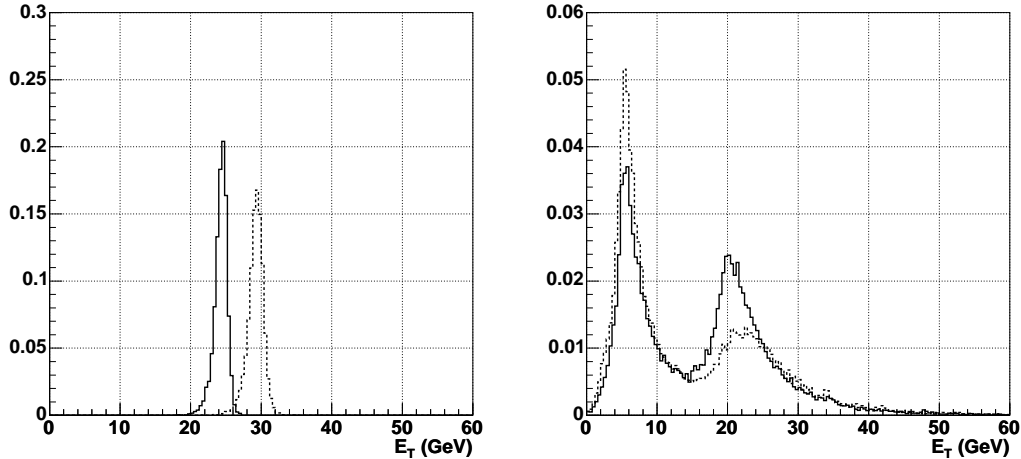


Fig. 3.2: Distribution de l'énergie transverse E_T de tous les amas après application du déclenchement LVL1 dans le calorimètre électromagnétique. La figure de gauche montre des électrons célibataires d'énergie $E_T = 25$ GeV. La figure de droite montre la distribution pour des jets.

Utilisation des informations calorimétriques

Les électrons sont vus comme des objets étroits dans le calorimètre électromagnétique alors que les jets tendent à avoir un profil plus large. Ainsi la forme des gerbes peut être utilisée pour rejeter les jets. Les informations pertinentes sont celles du premier et du second compartiment du calorimètre électromagnétique, ainsi que le premier compartiment du calorimètre hadronique.

Fuites hadroniques :

Les gerbes électromagnétiques ne déposent qu'une faible fraction de leur énergie dans le calorimètre hadronique, typiquement moins de 2% pour des électrons, soit par leur faible énergie soit étant arrêtés par les parois du cryostat pour les électrons de plus haute énergie. Les fuites hadroniques sont définies comme le rapport entre l'énergie transverse reconstruite dans le premier compartiment du calorimètre hadronique dans une fenêtre $\Delta\eta \times \Delta\phi = 0,2 \times 0,2$ et l'énergie transverse reconstruite dans le calorimètre électromagnétique. Ce rapport est montré dans la figure 3.3 pour des électrons isolés ayant une énergie transverse de 25 GeV et un lot de dijets.

Utilisation du second compartiment du ECAL :

Les variables suivantes sont utilisées :

- Les électrons déposent la majeure partie de leur énergie dans une fenêtre ayant une taille $\Delta\eta \times \Delta\phi = 3 \times 7$, exprimée en nombre de cellules. Le développement latéral $R_\eta(37)$ est donné par le rapport de l'énergie reconstruite dans une fenêtre 3×7 sur l'énergie reconstruite dans un fenêtre 7×7 . Cette variable (cf. figure 3.4 à gauche) montre un pic près de l'unité pour les électrons dû à leur faible développement latéral. De larges queues dans cette distribution sont attendues, à de faibles valeurs de $R_\eta(37)$ pour les jets.

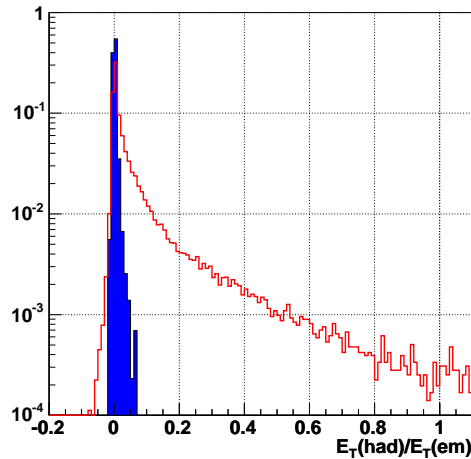


Fig. 3.3: Fuites hadroniques, définies comme le rapport entre l'énergie transverse reconstruite dans le premier compartiment du calorimètre hadronique dans une fenêtre $\Delta\eta \times \Delta\phi = 0,2 \times 0,2$ sur l'énergie transverse reconstruite dans le calorimètre électromagnétique. Les distributions sont montrées pour des électrons (histogramme hachuré) et des jets (histogramme en trait plein). Seul le déclenchement de niveau LVL1 a été appliqué. Les distributions sont normalisées à l'unité.

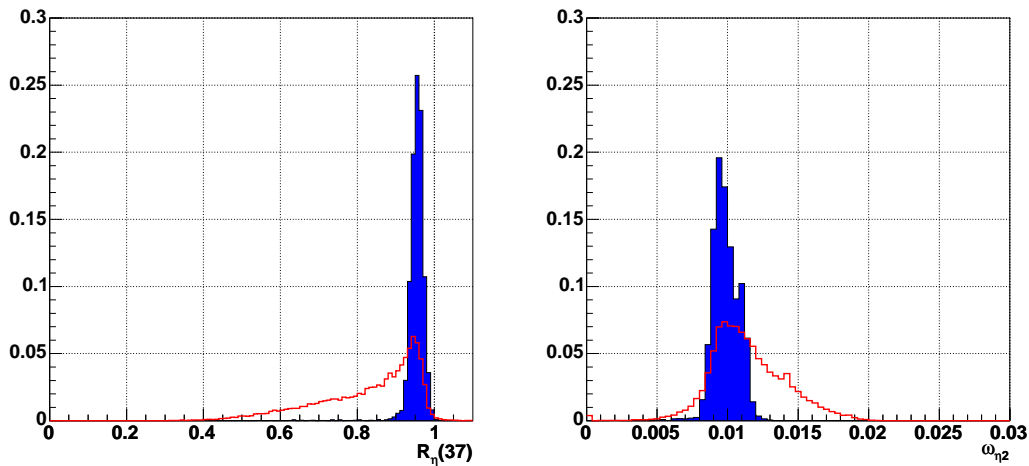


Fig. 3.4: Développement latéral de la gerbe $R_\eta(37)$ (gauche) et largeur latérale $\omega_{\eta 2}$ (droite). Les distributions sont montrées pour des électrons (histogrammes hachurés) et des jets (histogrammes en trait pleins). Seul le déclenchement de niveau LVL1 a été appliqué. Les distributions sont normalisées à l'unité.

- La largeur latérale (cf. figure .3.4 à droite) est calculée dans une fenêtre de 3×5 cellules, en utilisant une somme sur la position η de chaque cellule, pondérée par son énergie :

$$\omega_{\eta 2} = \sqrt{\frac{\sum E_c \times \eta^2}{\sum E_c} - \left(\frac{\sum E_c \times \eta}{\sum E_c}\right)^2}.$$

Utilisation du premier compartiment du ECAL :

Les critères précédents permettent le rejet des jets contenant des pions de haute énergie et formant une gerbe large. Les jets contenant un ou plusieurs η , π^0 etc., forment maintenant la contribution principale de ceux pouvant ressembler à des électrons (ou des photons). Le premier compartiment, avec sa granularité très fine en pseudorapidité, peut être utilisé pour détecter des sous-structures dans la gerbe. On peut isoler les π^0 's qui, se désintégrant en deux photons, vont présenter de telles sous-structures. La forme latérale de la gerbe est exploitée lorsqu'un minimum d'énergie (au moins 0,5% du total) est reconstruit dans ce compartiment et pour $|\eta| < 2,35$ où la granularité des cellules est suffisamment fine. Pour tous les critères suivants, l'énergie de deux cellules en ϕ est sommée.

- Les jets contenant des désintégrations de π^0 présentent souvent deux dépôts d'énergie importants. L'étude de la gerbe est effectuée dans une fenêtre $\Delta\eta \times \Delta\phi = 0,125 \times 0,2$ (cf. figure 3.5) autour de la cellule la plus chaude pour chercher un deuxième maximum. Si plus de deux maxima sont trouvés, le second maximum le plus important est utilisé.

Deux variables sont utilisées :

- la différence $\Delta E = E_{\max 2} - E_{\min}$ entre l'énergie associée au second maximum $E_{\max 2}$ et l'énergie reconstruite dans la cellule ayant la plus faible valeur, situé entre le premier et le second maximum E_{\min} (cf. figure 3.5 à gauche).

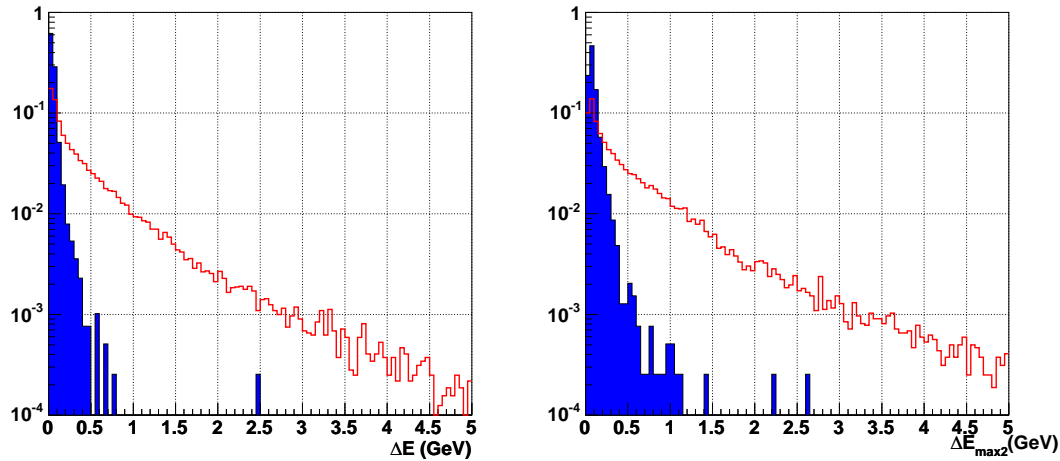


Fig. 3.5: à gauche : différence ΔE entre l'énergie associée au second maximum $E_{\max 2}$ et l'énergie reconstruite dans la cellule ayant la plus faible valeur, situé entre le premier et le second maximum E_{\min} . À droite : $\Delta E_{\max 2}$. Les distributions sont montrées pour des électrons (histogrammes hachurés) et des jets (histogrammes en trait plein). Seul le déclenchement de niveau LVL1 a été appliqué. Les distributions sont normalisées à l'unité.

- $\Delta E_{\max 2} = E_{\max 2} / (1 + 9(5) \times 10^{-3} E_T)$, avec E_T l'énergie transverse de l'amas utilisé du calorimètre électromagnétique. Pour s'assurer de ne pas être sensible à des fluctuations, la valeur du second maximum d'énergie doit être plus grande qu'un certain seuil qui

dépend linéairement de l'énergie transverse (cf. figure 3.5 à droite).

- La largeur totale de la gerbe est déterminée dans une fenêtre de taille $\Delta\eta \times \Delta\phi = 0,0625 \times 0,2$, correspondant typiquement à 40 cellules en η . Elle est calculée à partir de : $\omega_{\text{tot1}} = \sqrt{\sum E_i \times (i - i_{\text{max}})^2 / \sum E_i}$, où i est le numéro de cellule considérée et i_{max} celui du premier maximum local. Cette largeur est montrée pour des électrons et des jets dans la figure 3.6.

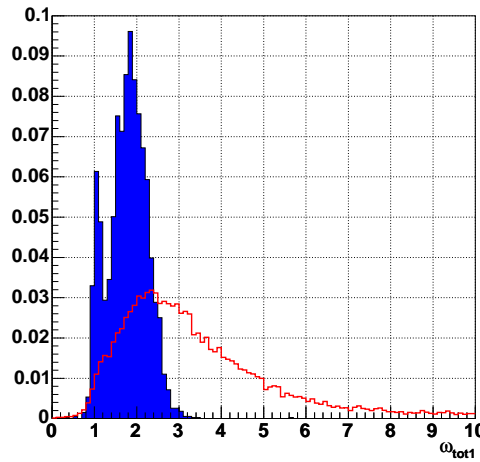


Fig. 3.6: Largeur totale de la gerbe ω_{tot1} dans le premier compartiment du calorimètre électromagnétique. La structure en double pic pour les électrons reflète la granularité variable dans les bouchons du calorimètre électromagnétique. Les distributions sont montrées pour des électrons (histogramme hachuré) et des jets (histogramme en trait plein). Seul le déclenchement de niveau LVL1 a été appliqué. Les distributions sont normalisées à l'unité.

Après l'application de ces critères, seules des gerbes très fines, ne contenant qu'un seul maximum sont conservées. Les deux variables suivantes sont utilisées pour une étude fine de la forme des gerbes :

- F_{side} donne la fraction d'énergie contenue en dehors du cœur de la gerbe. Elle est calculée à partir de la formule suivante : $F_{\text{side}} = [E(\pm 3) - E(\pm 1)] / E(\pm 1)$, où $E(\pm n)$ est l'énergie reconstruite dans $\pm n$ cellules autour de celle de plus haute énergie. Sa distribution est montrée sur la figure 3.7 à gauche.
- La largeur de la gerbe dans les trois cellules autour du maximum est montré sur la figure 3.7 à droite. Elle est donnée par la formule suivante :

$\omega_{\text{3strips}} = \sqrt{\sum E_i \times (i - i_{\text{max}})^2 / \sum E_i}$, où i est le numéro de cellule utilisé et i_{max} celui du plus énergétique. Cette variable est intéressante dans certains intervalles en η .

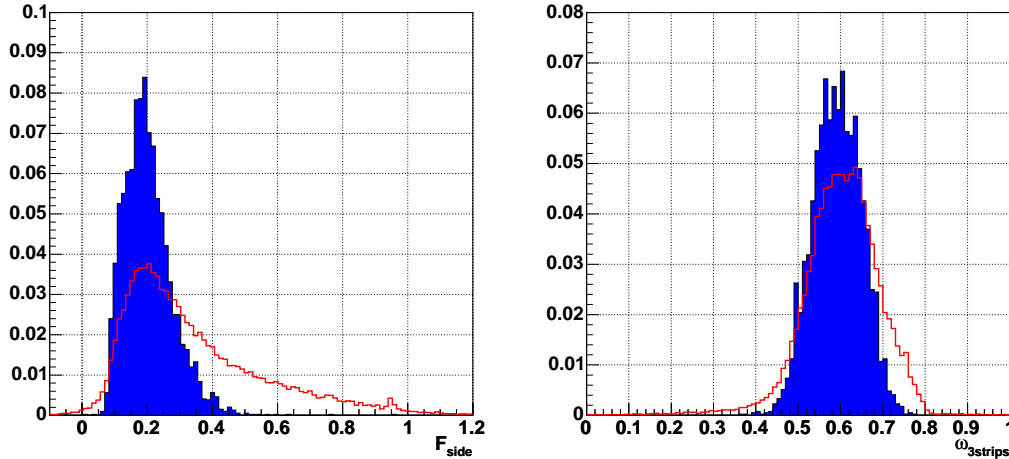


Fig. 3.7: Fraction d'énergie F_{side} reconstruite en dehors du cœur de la gerbe (à gauche) et largeur de la gerbe dans trois cellules $\omega_{3\text{strips}}$ (à droite), dans le premier compartiment du calorimètre électromagnétique. Les distributions sont montrées pour des électrons (histogrammes hachurés) et des jets (histogrammes en trait plein). Seul le déclenchement de niveau LV1 a été appliqué. Les distributions sont normalisées à l'unité.

Utilisation des informations du détecteur interne

Le bruit de fond restant est dominé par des photons convertis et des jets à faible multiplicité contenant des mésons π^0 de haut p_T . Celui-ci peut être encore plus réduit en requérant la présence de traces de bonne qualité pointant vers un amas électromagnétique, avec un bon accord entre les mesures d'énergie et d'impulsion. Seules les traces ayant un $p_T > 5$ GeV/c sont conservées. Une association angulaire est effectuée dans une fenêtre $\Delta\eta = \pm 0.05$, $\Delta\phi = \pm 0.1$ entre l'amas électromagnétique et la position de la trace extrapolée jusqu'au calorimètre¹⁾. Dans le cas où plus d'une trace est trouvée, celle de plus grand p_T est conservée. Si le rapport E/p est plus petit que 4, l'association avec la trace est considérée être un succès.

Un ensemble de coupures de bonne qualité des traces est de plus appliqué :

- au moins deux points de mesure sont requis dans le détecteur à pixel dont au moins un dans la couche "b" ;
- au moins neuf points de mesure dans les détecteurs de précision (pixels + SCT) ;
- le paramètre d'impact transverse $|d_0| \leq 1$ mm.

Utilisation des informations combinées

Le rejet des jets peut être grandement amélioré en s'assurant de la cohérence entre les informations du calorimètre électromagnétique et du détecteur interne. On vérifie tout d'abord l'association angulaire (cf. figure 3.8) :

¹⁾ Nous verrons au paragraphe 3.5.2 que la procédure d'association amas-trace a depuis un peu évolué.

3. SÉPARATION ÉLECTRONS/JETS

- $|\Delta\eta| = |\eta_{\text{cluster}}^{\text{strips}} - \eta_{\text{ID}}|$, où $\eta_{\text{cluster}}^{\text{strips}}$ est calculé dans le premier compartiment du calorimètre électromagnétique où la granularité est très fine et η_{ID} est la pseudorapidité de la trace extrapolée jusqu'au calorimètre ;
- $|\Delta\phi| = |\phi_{\text{cluster}}^{\text{middle}} - \phi_{\text{ID}}|$, où $\phi_{\text{cluster}}^{\text{middle}}$ est calculé dans le second compartiment du calorimètre électromagnétique et ϕ_{ID} est l'azimuth de la trace extrapolée jusqu'au calorimètre.

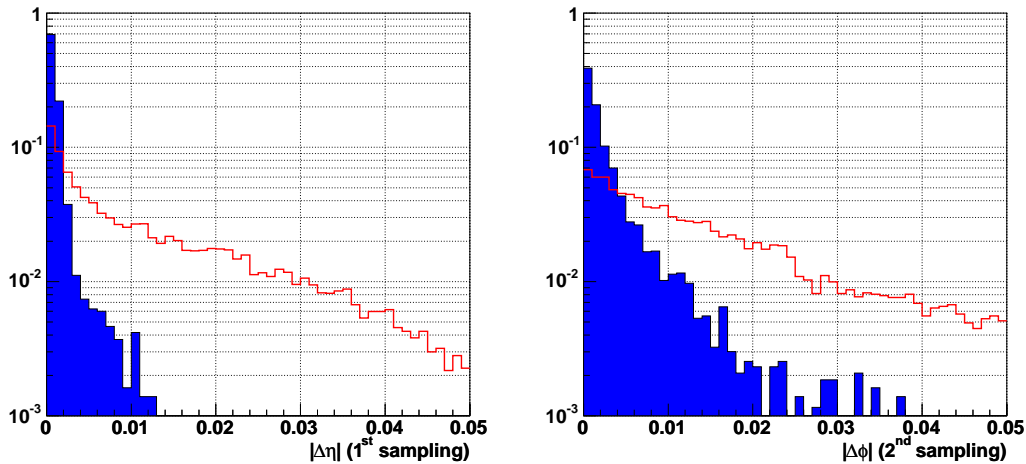


Fig. 3.8: Différence angulaire en pseudorapidité ($|\Delta\eta|$) et en azimuth ($|\Delta\phi|$) entre la position de la trace extrapolée au calorimètre électromagnétique et celle de l'amas électromagnétique. Les distributions sont montrées pour des électrons (histogrammes hachurés) et des jets (histogrammes en trait plein). Seul le déclenchement de niveau LVL1 a été appliqué. Les distributions sont normalisées à l'unité.

L'énergie E reconstruite dans le calorimètre électromagnétique est alors comparée à l'impulsion p reconstruite dans le détecteur interne (cf. figure 3.9 à gauche). Dans le cas d'un électron, l'impulsion doit être en accord avec l'énergie. Cependant, de large queues à de basses et grandes valeurs du rapport E/p peuvent être vues. Celles-ci sont dues à des électrons de conversion ainsi qu'à du rayonnement de freinage.

Utilisation du TRT

Le paragraphe 1.3.1 a détaillé la capacité du TRT à identifier les électrons. Un pouvoir de rejet supplémentaire de la contamination en hadrons chargés est en effet obtenu en rejetant les traces ayant une faible fraction de points de mesure de haut seuil dans le TRT. La figure 3.9 (à droite) montre le rapport $N_{\text{HTR}}/N_{\text{straw}}$ entre le nombre de points de mesure de haut seuil N_{HTR} et le nombre total dans le TRT N_{straw} .

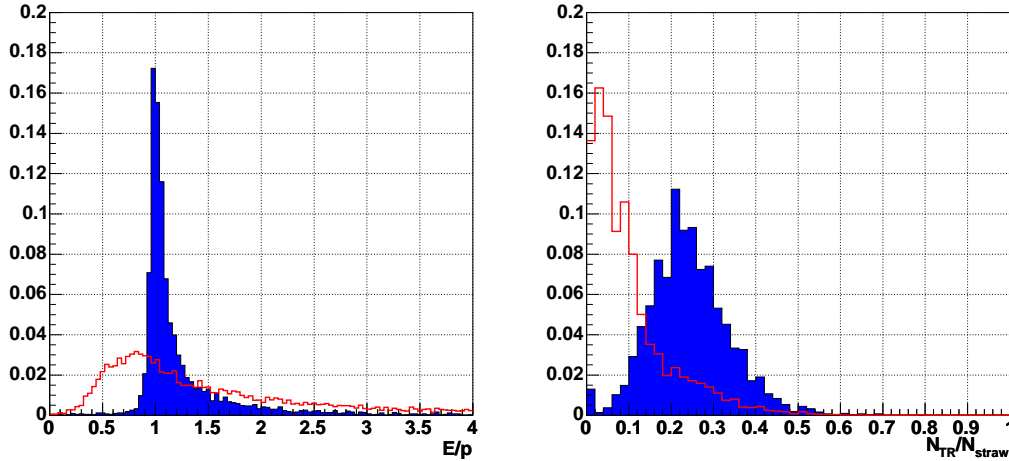


Fig. 3.9: Rapport E/p entre l'énergie reconstruite dans le calorimètre et l'impulsion reconstruite dans le trajectographe (à gauche) et rapport $N_{\text{HTR}}/N_{\text{straw}}$ (à droite) entre le nombre de points de mesure de haut seuil dans et le nombre de pailles touchées dans le TRT. Les distributions sont montrées pour des électrons (histogrammes hachurés) et des jets (histogrammes en trait plein). Seul le déclenchement de niveau LVL1 a été appliqué. Les distributions sont normalisées à l'unité.

3.4.2 Performances

La figure 3.10 montre un profil des valeurs moyennes des variables calorimétriques en fonction de la pseudo-rapidité de l'amas. Les variations suivent celles de la granularité, l'épaisseur de plomb et de matériau en amont du calorimètre électromagnétique. Aussi les seuils appliqués ont été optimisés dans différents intervalles²⁾. Les deux premiers intervalles, (0;0,8) et (0,8;1,37), couvrent la partie tonneau du calorimètre électromagnétique. La zone jonction et pauvrement instrumentée, située entre les parties tonneau et bouchons, entre $1,37 < |\eta| < 1,52$ est exclue de l'analyse. Dans les bouchons les intervalles (1,52;1,8), (1,8;2,0), (2,0;2,35) et (2,35;2,47) sont essentiellement déterminés par les variations de granularité du premier compartiment. Les quantités calculées dans le premier compartiment ne peuvent être utilisées que dans les régions $|\eta| < 1,37$ et $1,52 < |\eta| < 2,35$, puisque au-delà les régions du premier compartiment ont une granularité trop grande. Les seuils appliqués ont été optimisés pour obtenir une efficacité d'environ 95% par rapport au niveau LVL1, indépendamment de la pseudo-rapidité. Les seuils ont été définis pour être au plus près de la limite des distributions pour le signal. Ainsi il n'y aura pas de pertes importantes par rapport aux seuils définis au niveau du filtre à événements, mais le gain dans le pouvoir de rejet des jets sera augmenté (d'environ un facteur 10). A haute luminosité, les gerbes électromagnétiques étant plus larges, les seuils des coupures sont alors relâchés.

²⁾ Une optimisation en fonction de l'énergie est en cours de développement comme montré au paragraphe 3.5.3)

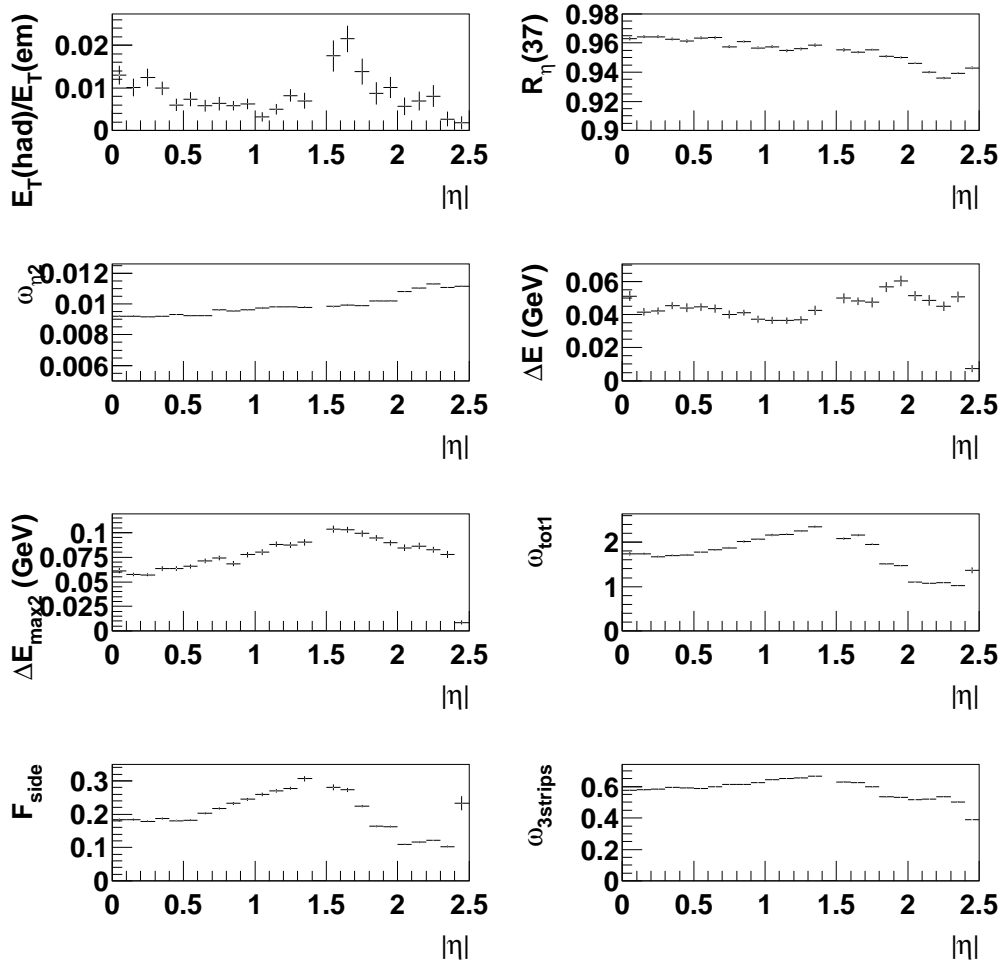


Fig. 3.10: Distributions pour des électrons célibataires de $E_T = 25$ GeV de la valeur moyenne de chaque variable discriminante “calorimétrique” en fonction de la pseudo-rapidité $|\eta|$.

Efficacité d'identification des électrons

L'efficacité d'identification est définie comme $\varepsilon_e = \frac{N_e^t}{N_e}$, où N_e est le nombre d'électrons, comptés à partir du bloc de la vérité Monte Carlo et une association angulaire dans une fenêtre $\Delta\eta = \pm 0.05$, $\Delta\phi = \pm 0.1$ entre l'amas et la position de la trace. N_e^t est le nombre d'électrons identifiés. La table 3.2 montre les efficacités obtenues pour des électrons de 25 GeV (30 GeV) à basse (haute) luminosité, sur la gamme en pseudorapidité $|\eta| < 2,47$. La zone jonction dans le calorimètre, située entre $1,37 < |\eta| < 1,52$, est exclue. L'efficacité totale ε_e est de $\sim 73\%$ ($\sim 71\%$) pour des électrons de 25 GeV (30 GeV). Elle est constante suivant η . L'efficacité

ε_e (in %)	low luminosity	high luminosity
LVL1	$95,8 \pm 0,3$	$95,9 \pm 0,2$
Calo	$91,2 \pm 0,4$	$90,0 \pm 0,3$
ID	$81,3 \pm 0,5$	$78,7 \pm 0,4$
ID-Calor	$76,4 \pm 0,6$	$73,0 \pm 0,4$
TRT	$73,5 \pm 0,6$	$71,3 \pm 0,4$

Tab. 3.2: Efficacité ε_e d'identification des électrons célibataires. La zone jonction dans le calorimètre, située entre $1,37 < |\eta| < 1,52$, est exclue.

d'identification des performances est vérifiée avec les électrons de signal dans le lot de jets inclusifs. Comme la fraction des événements avec des électrons de signal est très faible, les erreurs statistiques seront importantes. Pour les électrons provenant de désintégration de W/Z , l'efficacité est de $69 \pm 5\%$. Cela correspond à 43 W se désintégrant semileptoniquement. Le bruit de fond QCD à ce signal est initialement $\sim 10^5$ fois plus important. Comme nous allons le voir, dans le lot de jets QCD, après application des critères d'identification, seulement 33 amas sont mal identifiés comme des électrons. La pureté du lot d'électrons issus de W , définie comme le nombre de candidats électrons divisé par le nombre total de candidats sélectionnés, est de ~ 0.6 .

L'application des critères de sélection aux électrons provenant de la désintégration semileptonique de quarks b et c donne un total de 46 candidats. Cela correspond à une efficacité qui est seulement de $4.4 \pm 0.6\%$ car ces électrons ne sont pas isolés et de basse impulsion transverse, et donc ont peu de chance de passer le déclenchement de niveau 1. L'efficacité, normalisée au nombre de candidats passant ce déclenchement de niveau 1 est de $54 \pm 7\%$.

Facteur de rejet des jets QCD et rapport signal sur bruit

Comme décrit dans le paragraphe 3.1, le lot d'événements dijets comporte aussi des photons directs, du rayonnement de quark et de la production de W , Z et quark top. De plus une partie des événements contient des électrons provenant de la désintégration semileptonique de quarks lourds. Pour les étiqueter on utilise le fait que les événements provenant directement d'un quark doivent avoir un faible rayon au niveau du vertex. Ainsi un électron qui a un rayon, estimé à partir des informations du bloc vérité, plus petit que 10 mm est considéré comme étant de signal et est donc exclu de l'analyse. Une méthode alternative consiste à couper sur une impulsion transverse $p_T \geq 16$ GeV/c et de vérifier que cette particule ne provient pas d'une conversion. Les deux méthodes donnent le même résultat. Tous ces événements ne sont pas pris en considération dans la mesure du facteur de rejet des jets QCD. Par contre ils sont pris en compte dans la mesure du rapport signal sur bruit, c'est à dire sur le nombre d'électrons restants dans le lot de jets inclusif sur le nombre de hadrons chargés.

3. SÉPARATION ÉLECTRONS/JETS

Le calcul du facteur de rejet des jets QCD est fait par rapport au nombre de jets ayant une énergie transverse plus grande que 17 GeV. Les jets sont reconstruits avec le programme de simulation rapide ATLFast. Les fichiers utilisés pour cette analyse ne sont pas complets puisqu'ils passent à travers un filtre décrit précédemment, et ne contiennent qu'une fraction $\varepsilon_f = 8,33\%$ de tous les événements engendrés. Le nombre initial d'événements est égal à N_i/ε_f . Le nombre initial de jets N_{jets} est déterminé par ailleurs sur un sous-lot d'événements. Ces nombres sont basés sur l'étude de $N_{tot} = 2 \times 10^6$ événements dijet, avant l'application du filtre sur les particules. Le facteur de normalisation est alors N_{jets}/N_{tot} . Le nombre total de jets utilisé dans cette étude est alors donné par : $N_{jet}^i = \frac{N_i \times N_{jets}}{\varepsilon_f \times N_{tot}}$. Le facteur de rejet $R(\text{jet})$ sur les jets QCD est défini par le rapport du nombre initial de jets N_{jet}^i sur le nombre de jets mal identifiés N_{jet}^{res} :

$$R(\text{jet}) = \frac{N_{jet}^i}{N_{jet}^{res}}. \quad (3.2)$$

Le facteur de rejet après le niveau calorimétrique est de ~ 3000 . Après les critères d'association ID-Calo les hadrons chargés restent le principal bruit de fond. Le rapport signal sur bruit est alors 2 : 1 pour un facteur de rejet des jets QCD de l'ordre de 0.6×10^5 . Le signal provient alors pour moitié d'électrons provenant de la désintégration semileptonique de quarks lourds et d'électrons isolés provenant de la désintégration de bosons W et Z . Le facteur de rejet des jets QCD peut être grandement amélioré en utilisant le pouvoir de rejet du TRT. Pour une efficacité d'identification des électrons d'environ 70% un facteur de rejet des jets QCD de près de 2×10^5 est atteint. Il reste alors seulement 33 candidats dans le lot de jets QCD, dont 7 dans le tonneau et 26 dans les parties bouchons. Cette différence peut être comprise en regardant la distribution de E/p dans la figure 3.9. Cette variable montre de larges queues dans la distribution pour des électrons dues aux électrons de conversion et au rayonnement de freinage. Sept candidats ont une énergie transverse plus basse que 22 GeV, *i.e* sont sélectionnés uniquement par le menu de déclenchement à deux objets. Parmi les candidats restants 23 sont des électrons provenant de photons de conversion. La moitié de ces conversions sont très asymétriques, *i.e* le membre de la paire e^+e^- qui est sélectionné par l'algorithme transporte la plus grande partie de l'impulsion transverse. L'autre membre est une trace d'électron avec une petite impulsion transverse (qqg GeV/c). Ces conversions sont difficiles à reconstruire. Les résultats de l'algorithme d'identification des conversions n'étaient pas disponibles à l'époque de l'analyse et n'ont donc pas pu être utilisés. Cinq des événements sont des hadrons ($2 \pi^+$, $1 \pi^-$ et $2 K^+$) qui transportent plus d'un tiers de l'énergie de l'amas. L'événement restant correspond à une collection de traces hadroniques, aucune ne dominant en impulsion.

Ainsi, en enlevant les photons de conversion, on pourrait obtenir l'identification d'un lot pur d'électrons inclusif tout en obtenant un facteur de rejet des jets QCD de l'ordre de 10^6 .

3.5 Quelques développements et résultats récents

Depuis ces études de nombreuses évolutions ont été effectuées par le groupe électron/photon afin d'améliorer les performances de la reconstruction et de l'identification des électrons. Les résultats de ces développements seront présentés et incorporés dans les analyses "CSC" qui sont encore en cours au moment de l'écriture du manuscrit. En pratique j'ai surtout participé à la maintenance du programme, en particulier en ce qui concerne le calcul des formes des gerbes électromagnétiques et la partie reliée à l'identification des électrons. Je n'ai donc pas participé aux développements décrits ci-après mais ceux-ci sont importants pour la suite de l'exposé. De plus je présente certains résultats préliminaires que j'ai obtenu avec la version de production d'Athena (12.0.6), concernant encore une fois l'identification des électrons, permettant de voir l'évolution des performances depuis l'analyse DC1.

3.5.1 Développements algorithmiques

Construction de l'objet 'electron

Comme décrit dans le paragraphe 1.4 sur le modèle de calcul d'Atlas, les formats de données ESD et AOD permettent d'accéder aux informations des objets reconstruits egamma dans une forme complète au niveau ESD ou dans la forme "Electron" et "Photon" au niveau des AOD. A un objet sont associées des informations générales :

- un quadri-vecteur formé de la reconstruction de l'énergie dans l'amas du calorimètre et des angles provenant soit des informations du trajectographe soit des informations du calorimètre ;
- un identifiant de l'auteur de l'algorithme de reconstruction, soit commençant par un amas électromagnétique soit commençant par une trace ;
- l'amas associé à l'objet ;
- la trace associée, le cas échéant, à l'objet ;
- la conversion associée, le cas échéant, à l'objet ;
- des informations donnant les résultats des différents algorithmes d'identification.

De plus des informations détaillées sont disponibles sur :

- la forme des gerbes électromagnétiques ;
- l'association amas-trace ;
- la possible identification d'une conversion ;
- la reconstruction de l'impulsion en prenant en compte le rayonnement de freinage.

Si une trace est associée alors cet objet formera un "Electron", sinon il formera un "Photon". La classification des électrons de conversions comme "Photon" est en cours d'élaboration.

Formation des amas de cellules

La taille des amas rectangulaires créés par l'algorithme de fenêtre glissante est à optimiser selon le type de particule et selon la partie (tonneau/bouchon) où l'on se trouve. Ce résultat découle des études effectuées dans le cadre du TDR de physique avec la version fortran du programme de reconstruction. Dans la version d'Athena disponible durant de nombreuses années et en particulier pour l'essentiel des résultats présentés dans ce document, une seule taille 5×5 était gérée par l'algorithme de reconstruction des électrons. C'est uniquement pour les analyses CSC en cours que les différentes tailles d'amas est à nouveau disponible.

Par ailleurs l'algorithme de reconstruction des amas de cellules de taille variable, appelé topologique, qui a été développé pour la reconstruction des jets, a été adapté pour la reconstructions des amas électromagnétiques. La recherche des gerbes électromagnétiques s'effectue en recherchant des cellules graines dans le compartiment central au-dessus d'un certain seuil du bruit dans la cellule. Deux autres seuils permettent à l'amas d'accréter des cellules. Cet algorithme n'est pas encore utilisable par défaut pour la reconstruction des électrons. Toutefois des résultats relatifs à son utilisation peuvent être obtenus dans la référence suivante [88].

Le rayonnement de freinage et les électrons de conversion

La mesure précise des objets électromagnétiques est perturbée par la présence de matière devant le calorimètre, entre 0,3 et 1,5 longueurs de radiation à la sortie du détecteur interne comme montré sur la figure 1.5, avec un maximum autour de $\eta = 1,5$, correspondant au passage de câbles et de tuyaux de refroidissement du détecteur interne entre les bouchons et l'extrémité du tonneau. Une telle longueur de radiation signifie que la gerbe électromagnétique peut commencer dans le détecteur interne et laisser une partie de son énergie en amont du calorimètre. Cet effet est corrigé en pondérant en fonction de l'énergie du pré-échantillonneur et du premier compartiment du calorimètre. De surcroît la conversion d'un photon ou l'émission d'un photon de rayonnement de freinage dur à petit rayon peut, en raison du champ magnétique, causer une fuite d'énergie en dehors de l'amas, causant une traîne vers les basses énergies. De nombreuses études sont en cours avec le développement de plusieurs algorithmes. Ces développements ne seront disponibles que pour les notes CSC et pas pour les études de ce manuscrit.

Les nouveaux développements au niveau de l'algorithme de reconstruction des traces permettent d'améliorer la reconstruction des conversions, voire rendent possible la reconstruction des reconstructions tardives, c'est à dire dans les dernières couches de silicium, voire même dans le TRT seul. Ces informations ont manqué à la plupart des analyses auxquelles j'ai participé dans le cadre de DC1, en particulier celles concernant l'identification des photons, tout particulièrement pour la recherche du boson de Higgs en deux photons.

3.5.2 Identification des électrons

Je présente ici quelques résultats et développement récents auxquels j'ai participé. Une description exhaustive des résultats les plus récents (mais obtenus après l'écriture de ce document) de l'algorithme de reconstruction des électrons de haute impulsion transverse, avec les données CSC, peut être obtenu dans la référence [102].

Evolutions récentes de l'identification

Par rapport à l'étude présentée plus haut une différence majeure provient de l'évolution du modèle d'événements pour les objets "Electron". En effet ceux-ci supposent désormais implicitement l'association d'une trace à l'amas électromagnétique, alors que cette association était effectuée plus tardivement auparavant. Cette association est effectuée en deux étapes. En premier lieu une association avec des critères assez lâches : un rapport $E/p < 10$ et une association angulaire $\Delta\eta < 0,05$ et $\Delta\eta < 0,01$ entre la position de l'amas et la position de la trace au niveau du vertex. Ensuite des critères plus stricts sont appliqués, correspondants à ceux décrits au paragraphe 3.4.1 : $|\Delta\eta| = |\eta_{\text{cluster}}^{\text{strips}} - \eta_{\text{ID}}| < 0,025$ et $-0,05 < \Delta\phi = \phi_{\text{cluster}}^{\text{middle}} - \phi_{\text{ID}} < 0,025$. Ainsi, avec des électrons célibataires ayant une énergie transverse égale à 25 GeV, reconstruits avec la version d'Athena 12.0.6, seuls 87% d'entre eux ont une association trace-amas. Le tableau 3.3 donne l'efficacité d'identification des électrons alors obtenue. Cette efficacité n'est plus montrée avec une factorisation d'effets qui suivent l'utilisation successive des sous-détecteurs mais suivant des critères de qualité associés à l'identification [103] :

- "Loose" : application des critères hadroniques et du deuxième compartiment du calorimètre électromagnétique
- "Medium" : application de tous les critères calorimétriques et des critères de bonne qualité des traces, à l'exception de celui sur la couche b du détecteur à pixels.
- "Tight" : l'ensemble des critères d'identification.

ϵ_e (in %)	low luminosity
Loose	$98,4 \pm 0,2$
Medium	$79,9 \pm 0,5$
Tight	$65,2 \pm 0,6$

Tab. 3.3: Efficacité ϵ_e d'identification des électrons célibataires d'énergie transverse $E_T = 25$ GeV. La zone jonction dans le calorimètre, située entre $1,37 < |\eta| < 1,52$, est exclue.

Cette efficacité est plus basse d'environ 10% que celle obtenue dans le cadre de DC1. On notera toutefois les choses suivantes. On a vu sur la figure 1.5 que la quantité de matière présente dans la simulation Monte Carlo a augmenté, et ce de manière assez importante, entre les Data Challenge DC1 et CSC. Les seuils des critères, en particulier pour les variables calorimétriques, sont donc à réoptimiser. De plus les critères sur le TRT ont été considérablement modifiés. On

avait remarqué que dans l'étude DC1, l'efficacité d'identification du TRT seul était de 95% alors qu'elle aurait dû être optimisée à 90%. Ces seuils ont été réoptimisés et renforcés permettant d'atteindre les performances attendues de ce détecteur. Ainsi pour tous ces électrons célibataires, en prenant en compte l'efficacité d'association trace-amas et la zone jonction du calorimètre électromagnétique, l'efficacité d'identification des électrons est d'environ 55%.

Depuis les études DC1, de nombreuses autres méthodes d'identification ont été développées. Elles n'utilisent plus des coupures sur chaque variable discriminante, mais la forme des distributions : rapport de vraisemblance, réseau de neurones, "H-matrix". Notons que la méthode utilisant les coupures reste la référence, de par sa relative facilité d'utilisation mais aussi par la robustesse des résultats et le haut niveau de performance toujours obtenu.

3.5.3 Isolation des électrons

A toutes ces méthodes il peut être utile d'ajouter un critère supplémentaire demandant que les électrons soient isolés. Deux critères ont été développés :

- une isolation au niveau calorimétrique : le critère est basé sur la somme des énergies transverses déposées autour du centre de l'amas dans un cône $R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ dans les calorimètres. On lui retire la composante de l'électron, dans une fenêtre 5×7 dans le calorimètre électromagnétique.
- une isolation au niveau du trajectographe : le critère est basé sur le nombre de traces chargées ayant une impulsion totale transverse p_T supérieure à un seuil, estimé autour du centre de l'amas.

Ces deux critères sont particulièrement utiles pour combattre les événements d'empilement. Cela a été montré en particulier pour la séparation photon/jet et le canal $H \rightarrow \gamma\gamma$ dans les analyses ayant suivies le TDR [104] pour l'isolation calorimétrique et DC1 [26] pour l'isolation par le trajectographe.

Reconstruction du boson Z^0

Nous utilisons un lot de 25000 événements $Z \rightarrow e^+e^-$ produits pour le Data Challenge CSC, correspondant à 18 pb^{-1} de données prises à basse luminosité. La masse invariante des deux électrons de désintégration est bien connue et distribuée selon une Breit-Wigner. On va donc pouvoir sélectionner à l'intérieur de ce pic un lot pur d'électrons pour la mesure de l'efficacité de l'algorithme d'identification des électrons directement sur les données. La figure 3.11 montre la masse invariante des paires de candidats électrons de charges opposées sélectionnées. L'efficacité de sélection des Z est d'environ 25%, les pertes provenant pour l'essentiel des pertes d'acceptance géométrique et sur la sélection des électrons. Pour mesurer l'efficacité nous utilisons la méthode "tag and probe" : chaque paire est formée d'un électron bien identifié et d'une trace de charge opposée sur laquelle les critères d'identification n'ont pas été appliqués. La

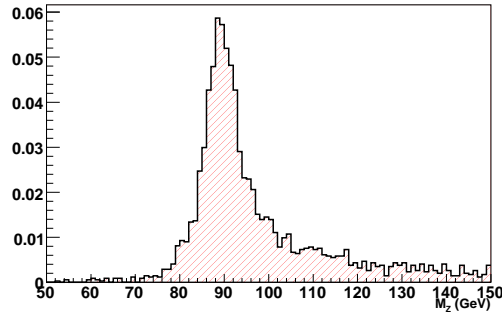


Fig. 3.11: Masse invariante des paires d'électrons de charges opposées dans le lot $Z \rightarrow e^+e^-$.

masse invariante est formée et si celle-ci est proche de la masse du Z , en pratique dans l'intervalle $80\text{-}100 \text{ GeV}/c^2$, alors l'événement est conservé et on ne s'intéresse qu'aux traces sur lesquelles les critères d'identification n'ont pas été appliqués. L'efficacité d'identification des électrons est mesurée comme le rapport du nombre de traces passant alors les critères sur le nombre total. Cette efficacité est alors mesurée à 71% en bon accord avec une mesure effectuée en utilisant les informations de la vérité Monte Carlo.

De plus on peut contraindre la masse invariante des électrons reconstruits à la masse du boson Z dans une fenêtre en masse autour du pic du Z reconstruit. Différentes méthodes permettent de ramener l'énergie des deux électrons à la masse du Z soit en appliquant une contrainte fixe, $M_Z = 91, 187 \text{ GeV}/c^2$ selon le Particle Data Group, soit une contrainte prenant en compte la largeur du Z . Comme nous l'avons vu dans le paragraphe 1.3.2, le calorimètre électromagnétique est constitué de 448 régions de taille $\Delta\eta \times \Delta\phi = 0,2 \times 0,4$ dans lesquelles le terme constant "local" de la résolution en énergie est meilleur que 0,5%. La contrainte sur la masse du Z peut aussi être utilisée pour inter-étalonner ces différentes régions pour obtenir un terme constant "global" meilleur que 0,7%, comme nécessaire dans le cahier des charges de l'expérience. Cette même méthode sera utilisée au chapitre suivant, à plus basse énergie, avec des événements J/ψ et sera décrite à ce moment là.

L'impact du bruit de fond n'a pas encore été pris en compte. Celui-ci provient d'une part du processus $Z \rightarrow \tau\tau$ qui peut donner deux électrons dans l'état final. Toutefois vu la valeur du rapport d'embranchement $\tau \rightarrow e\nu_e$ ce bruit de fond doit être de l'ordre de 1% du signal. Comme nous l'avons vu l'identification des électrons (en appliquant les critères "Tight") donne un facteur de rejet des jets QCD de l'ordre de 10^{-5} , ainsi le bruit de fond di-électron provenant des événements dijets doit être négligeable.

Plus de détails sur l'extraction du signal de Z^0 et son utilisation peuvent être obtenus par exemple dans les références [105] et [106].

Optimisation à basse énergie

Une grande partie des analyses utilisant des leptons requiert en tout premier lieu une impulsion transverse plus grande que 20 ou 25 GeV, correspondant aux seuils des menus de déclenchement sur les objets isolés. Toutefois, comme vu au chapitre 2, pour certains canaux le spectre en impulsion transverse peut aller bien plus bas et il est alors intéressant d'essayer d'améliorer les performances de l'identification à basse énergie. Comme on le voit sur la figure 3.12 ou sur la table 3.4 l'efficacité d'identification des électrons chute en dessous de 20 GeV. Cependant les seuils des critères d'identification, en particulier pour les variables calorimétriques, n'ont pas été optimisés dans ces gammes. Cette chute peut être en partie compensée en optimisant les seuils des critères d'identification, en particulier pour les variables calorimétriques. Il n'a pas été trouvé de variable expliquant l'essentiel de la perte, mais une optimisation globale de l'ensemble des seuils appliqués permet de gagner typiquement 5-8% de gain en efficacité. Par contre vers 10 GeV on constate une baisse supplémentaire due au plus faible nombre d'association existante entre une trace et un amas. Une étude plus approfondie sera nécessaire pour essayer d'optimiser les critères d'association.

ε_e	$E_T = 10$ GeV		$E_T = 15$ GeV		$E_T = 20$ GeV
	standard	optimisé	standard	optimisé	
Association	81,8 ± 0,2		85,5 ± 0,2		86,5 ± 0,2
Loose	77,3 ± 0,2	79,2 ± 0,2	82,4 ± 0,2	83,5 ± 0,2	84,5 ± 0,2
Medium	54,5 ± 0,2	64,3 ± 0,2	62,6 ± 0,2	69,3 ± 0,2	68,7 ± 0,2
Tight	40,6 ± 0,6	48,4 ± 0,2	48,0 ± 0,3	53,6 ± 0,3	54,2 ± 0,2

Tab. 3.4: Efficacité d'identification des électrons célibataires d'énergie transverse de 10 à 20 GeV. La zone jonction dans le calorimètre, située entre $1,37 < |\eta| < 1,52$, est exclue.

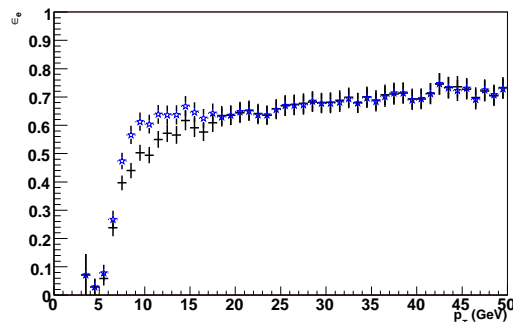


Fig. 3.12: Efficacité d'identification du lot d'électrons célibataires d'énergie transverse 7-80 GeV des données CSC en fonction de l'énergie transverse. L'efficacité est montrée pour les coupures (standard) (croix) et optimisées (étoiles).

Enfin à encore plus basse énergie transverse l'efficacité de l'association va encore baisser suivant l'efficacité de reconstruction des amas électromagnétiques eux-mêmes. Il faudra alors envisager un autre algorithme de reconstruction des électrons dont la description va faire l'objet du chapitre suivant.

4. ETUDES SUR LE J/ψ

Comme nous venons de le voir, les performances de l'algorithme d'identification des électrons isolés décroissent fortement à basse impulsion transverse. Or la détermination d'une échelle d'énergie commune au détecteur interne et au calorimètre électromagnétique implique des études non seulement aux énergies typiques du boson Z^0 mais aussi à plus basse énergie avec le J/ψ ou le Υ . L'étude de la reconstruction du J/ψ a été à l'origine de ma participation au développement et à la maintenance de l'algorithme de reconstruction et d'identification des électrons de basses impulsions transverses, ce qui représente une part extrêmement importante de mon activité. Elle fait suite à des études effectuées dans le groupe pour la préparation du TDR de physique [16].

L'essentiel du travail présenté dans ce chapitre a fait l'objet d'une note Atlas [107]. Les travaux actuels sur le J/ψ et l' Υ feront l'objet d'une note Atlas CSC [108], en préparation, dont je suis éditeur. Ils rassembleront les études sur le déclenchement, la reconstruction des traces en prenant en compte le rayonnement de freinage, la reconstruction et l'identification des électrons et enfin la reconstruction du J/ψ . Quelques résultats récents et préliminaires sont tout de même évoqués dans le paragraphe 4.5.

4.1 Les données utilisées

L'essentiel des études auxquelles j'ai participé sont basées sur des lots de données produits dans le cadre du Data Challenge 1, avec Pythia 6.203 [97]. Les événements ont été simulés avec Geant 3 [98], basé sur la version 5.0.0 d'Athena, sans événements d'empilement superposés et reconstruits avec la version 7.8.0. Le signal utilisé provient d'événements $bB_d \rightarrow \mu(6)J/\psi(e^+e^-)K_S(\pi^+\pi^-)$ (cf. Tab. 4.1). Un seuil sur l'impulsion transverse de 2 GeV/c est appliqué pour engendrer les électrons. Le bruit de fond provient d'événements $b\bar{b} \rightarrow \mu 6X$. Il contient des muons provenant de la désintégration semi-leptonique de hadrons B , avec un seuil sur l'impulsion transverse du muon de 6 GeV/c. Un même lot a été produit dans le cadre du Data Challenge dit "Rome", simulé avec la version d'Athena 10.0.1 et reconstruit avec la 11.0.41. Enfin dans le cadre d'études plus récentes, dans le cadre du Data Challenge 3 et du Computing System Commissioning, on utilise un lot de $bB_d \rightarrow \mu(6)J/\psi(e^+e^-)X$ produit et reconstruit avec la version 13.0.30.

Processus	Data Challenge	Lot de données	Nombre d'événements
$bB_d \rightarrow \mu(6)J/\psi(e^+e^-)K_S(\pi^+\pi^-)$	DC1	013901	250×10^3
$b\bar{b} \rightarrow \mu(6)X$	DC1	012500	90×10^3
$bB_d \rightarrow \mu(6)J/\psi(e^+e^-)K_S(\pi^+\pi^-)$	Rome	018102	20,000
$bB_d \rightarrow \mu(6)J/\psi(e^+e^-)X$	DC3	018102	250×10^3
$bB_d \rightarrow \mu(6)J/\psi(e^+e^-)X$	CSC	013901	50×10^3

Tab. 4.1: Processus, numéro dans la production Monte Carlo et nombre d'événements utilisés.

4.1.1 Cinématique des événements

La figure 4.1 montre les distributions d'impulsion transverse et de pseudorapidité pour les électrons de signal et les pions du lot de bruit de fond. Les électrons de signal sont produits principalement dans la région centrale en pseudorapidité et ont un spectre en impulsion transverse plus dur que les événements de bruit de fond.

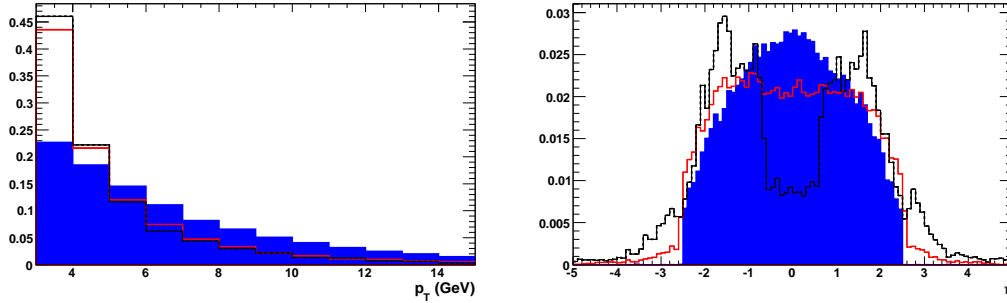


Fig. 4.1: Impulsion transverse reconstruite p_T (à gauche) et pseudorapidité $|\eta|$ (à droite) pour les électrons de signal (histogrammes hachurés), les électrons de conversion (histogrammes en trait tireté) et les pions du lot de bruit de fond (histogrammes en trait plein).

4.1.2 Sélection des événements

Le déclenchement pour la physique du B devrait couvrir 5 à 10% des ressources totales pour l'expérience Atlas. Le cœur du déclenchement est basé sur un muon d'une impulsion transverse suffisante au niveau un. Différents seuils sont implémentés. Un seuil de 6 GeV/c devrait induire un taux d'événements de 21 kHz à basse luminosité, parmi lesquels environ 15% contiennent effectivement un quark b . L'étude s'est limitée aux événements J/ψ produits à partir de la désintégration d'une paire $b\bar{b}$. Il est alors possible d'utiliser le muon provenant de la désintégration semi-leptonique d'un des B et de regarder la désintégration $J/\psi \rightarrow e^+e^-$ dans l'événement. Le cas de la production directe de J/ψ , où le déclenchement ne peut se faire qu'à partir des deux amas électromagnétiques, n'a pas été abordé. On trouvera tout de même

une étude effectuée au laboratoire dans le rapport de stage [109], et les résultats de l'analyse CSC [108].

La section efficace pour $pp \rightarrow b\bar{b}$ avec un b donnant une muon de $p_T > 6$ GeV/c et une désintégration du méson en $B \rightarrow J/\psi$ et $J/\psi \rightarrow e^+e^-$ est d'environ 2×10^{-4} μb . Un total de 4×10^6 événements est attendu en un an de prise de données à basse luminosité. Avec une efficacité de reconstruction typique de $\sim 20\%$, on attend environ $\sim 8 \times 10^5$ événements complètement reconstruits.

4.2 Reconstruction et identification des électrons

L'algorithme de reconstruction des électrons décrit au chapitre 3 est efficace pour reconstruire des électrons isolés, ayant une impulsion transverse d'au moins 10 GeV/c. Cependant il n'est pas adapté pour des électrons isolés de bas p_T , et pour des électrons localisés dans des jets comme ceux issus de la désintégration semileptonique de quarks de saveurs lourdes. En effet, les gerbes issues des hadrons et des électrons tendent alors à se chevaucher lorsque le jet est collimé, ce qui ne permet pas de distinguer les caractéristiques d'une gerbe électromagnétique.

Une méthode alternative pour la reconstruction des électrons consiste à utiliser comme graine de départ une trace dans le détecteur interne. Cette trace est extrapolée jusqu'au calorimètre et un amas électromagnétique est construit autour de sa position. La méthode permet de reconstruire des électrons ayant une impulsion transverse aussi basse que possible.

4.2.1 Algorithme de reconstruction

Etant donné le grand nombre de traces pour chaque événement, il est important de ne chercher à reconstruire que des traces ayant a priori des chances d'être ensuite identifiées comme des électrons utilisables. En pratique, seules les traces ayant une impulsion transverse plus grande que 2 GeV/c sont considérées. Les critères standards de *bonne qualité* sont appliqués :

- au moins deux points de mesure dans les pixels, dont au moins un dans la couche "b" ;
- au moins neuf points de mesure dans les détecteurs de précision (pixels + SCT) ;
- le paramètre d'impact transverse $|d_0| \leq 1$ mm.

Deux autres critères permettent de réduire le nombre de traces de hadrons :

- au moins un point de mesure à haut seuil dans le TRT le long de la trajectoire ;
- au moins vingt points de mesure dans le TRT le long de la trajectoire.

Toutes ces traces sont extrapolées jusqu'aux différents compartiments du calorimètre électromagnétique. Dans chacun, on recherche la cellule la plus chaude dans une petite fenêtre en η et ϕ autour du point d'impact : 3×3 dans le pré-échantillonneur, 3×1 dans le premier compartiment, 3×3 dans le second et 3×1 dans le troisième. Puis un amas est formé de taille : 3×5 dans le pré-échantillonneur, 19×5 dans le premier compartiment, 9×13 dans le second

et 5×13 dans le troisième¹⁾.

La figure 4.2 montre le rapport entre l'énergie reconstruite dans le calorimètre électromagnétique et l'énergie engendrée, en fonction de $|\eta|$, pour tous les électrons de signal et pour ceux ayant une énergie inférieure à 8 GeV. Sans les corrections, l'énergie reconstruite est environ 5% trop basse. Ceci peut atteindre 10% pour des électrons ayant une énergie inférieure à 8 GeV. Les corrections permettent d'atténuer cet effet, mais il y a toujours une perte à basse énergie. De plus on remarque des structures dans la distribution du rapport en fonction de la pseudo-rapidité qui sont proches de la distribution de matière en amont du calorimètre électromagnétique. Enfin ces corrections n'ont pas été étudiées ni optimisées pour un lot de si basse énergie, comme montré dans [111].

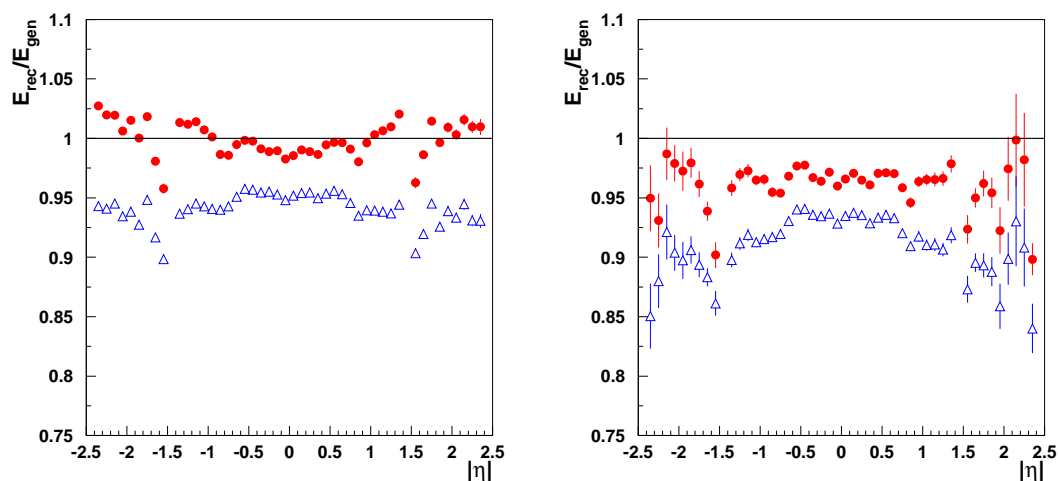


Fig. 4.2: Rapport entre l'énergie reconstruite dans le calorimètre électromagnétique et l'énergie engendrée, en fonction de $|\eta|$, pour tous les électrons de signal (à gauche) et pour ceux ayant une énergie inférieure à 8 GeV (à droite). Les rapports sont montrés sans les corrections (triangles) et après application des corrections (cercles) sur l'énergie.

4.2.2 Variables caractérisant les électrons

L'identification des électrons repose sur les mêmes caractéristiques que celles qui ont été décrites dans le chapitre précédent et une procédure similaire est utilisée pour l'identification²⁾.

¹⁾ Nous verrons au paragraphe 4.5.2 que la construction de cet amas, alors identique à celle des versions type "TDR" [110] a été depuis considérablement simplifiée en un amas rectangulaire de taille 5×5 .

²⁾ Le nombre et le choix des variables utilisées ont un peu évolué au cours du développement. Dans la mesure du possible ce choix a tendu vers une simplification et un lot de variable le plus proche possible de l'identification standard, et ce sans changement important dans les performances. La présentation effectuée ici correspond au choix le plus récent.

Utilisation des informations calorimétriques

Dans le calorimètre électromagnétique les électrons sont des objets étroits tandis que les jets tendent à avoir un profil plus large. Cependant, les électrons ne sont pas bien isolés pour le canal qui nous concerne, $b\bar{b} \rightarrow J/\psi X$. Il est donc important de s'intéresser au cœur de la gerbe dans les différents compartiments. Ceci expliquera le choix parfois légèrement différent dans les variables utilisées par rapport à l'algorithme standard. On notera alors que les distributions obtenues pour les électrons du J/ψ sont bien plus larges, avec des queues de distributions plus étalées, que pour les électrons isolés (cf. paragraphe 3.4.1). Ces mêmes variables sont alors moins discriminantes. Enfin on notera assez souvent des pics dans les distributions à des valeurs égales à zéro ou à un (en fait une valeur par défaut).

Utilisation du troisième compartiment du ECAL :

La figure 4.3 montre la variable de fuite hadronique définie précédemment (cf. Fig 4.3) pour les électrons du J/ψ et les pions du $b\bar{b}$. Celle-ci est calculée dans une fenêtre de taille $\Delta\eta \times \Delta\phi = 0,2 \times 0,2$ qui inclut donc la partie hadronique qui englobe l'électron. La variable perd donc de son sens et surtout de son intérêt pour des électrons non-isolés de par la granularité du calorimètre hadronique. En lieu et place nous utilisons la fraction d'énergie reconstruite dans le troisième compartiment du calorimètre électromagnétique $E_3(\text{core})/E$ qui a une granularité plus fine. L'énergie $E_3(\text{core})$ est l'énergie reconstruite dans une fenêtre 3×3 dans le troisième compartiment. Ceci peut paraître surprenant au premier abord car celui-ci est généralement présenté comme utile pour les électrons de haute énergie, mais pour mesurer leur énergie. Cette distribution est montrée sur la partie droite de la figure 4.3.

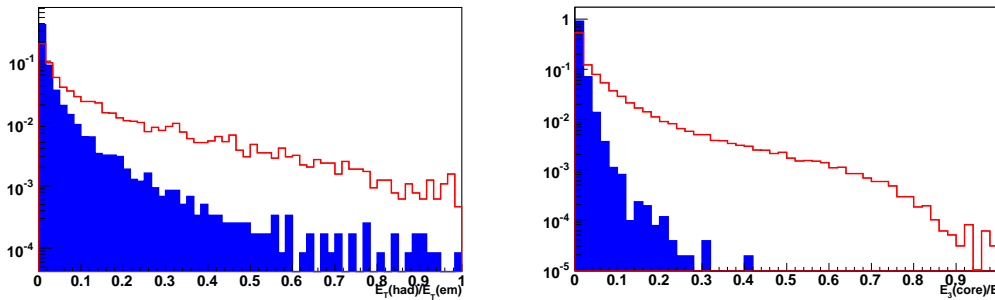


Fig. 4.3: A gauche : fuites hadroniques, définies comme le rapport entre l'énergie transverse reconstruite dans le premier compartiment du calorimètre hadronique, dans une fenêtre $\Delta\eta \times \Delta\phi = 0,2 \times 0,2$ sur l'énergie transverse reconstruite dans le calorimètre électromagnétique. A droite : fraction d'énergie $E_3(\text{core})/E$ reconstruite dans le cœur de la gerbe du troisième compartiment du calorimètre. Les distributions sont montrées pour des électrons (histogrammes hachurés) et des pions du lot de bruit de fond (histogrammes en trait plein). Elles sont normalisées à l'unité.

Utilisation du second compartiment du ECAL :

On utilise les deux mêmes variables que pour l'identification des électrons isolés : $R_\eta(37)$ pour

la description du développement latéral de la gerbe et $\omega_{\eta 2}$ pour sa largeur. Leurs distributions sont montrées sur la figure 4.4.

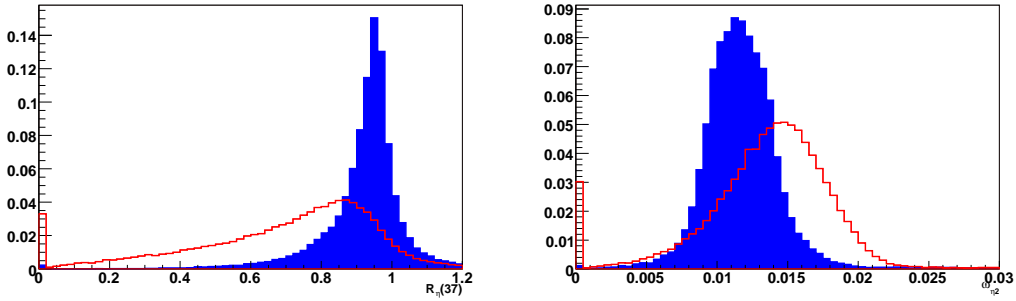


Fig. 4.4: Développement latéral de la gerbe $R_{\eta}(37)$ (gauche) et largeur latérale $\omega_{\eta 2}$ (droite). Les distributions sont montrées pour des électrons (histogrammes hachurés) et pour des pions du lot de bruit de fond (histogrammes en trait plein). Elles sont normalisées à l'unité.

Utilisation du premier compartiment du ECAL :

Les critères précédents permettent le rejet des pions de hautes énergies et formant des gerbes larges. Pour exclure les pions de basses énergies, nous utilisons la fraction d'énergie reconstruite dans le cœur de la gerbe $E_1(core)/E$. L'énergie $E_1(core)$ est l'énergie reconstruite dans une fenêtre 15×3 dans le premier compartiment. La distribution correspondante est montrée sur la figure 4.5 pour des pions et des électrons.

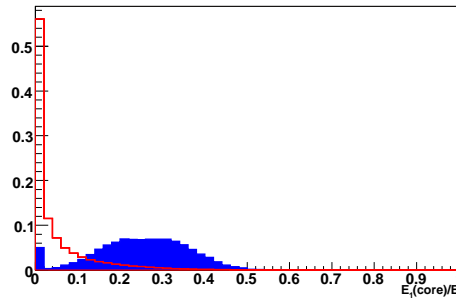


Fig. 4.5: Fraction d'énergie $E_1(core)/E$ reconstruite dans le cœur de la gerbe dans le premier compartiment du calorimètre électromagnétique. Les distributions sont montrées pour des électrons (histogramme hachuré) et des pions (histogramme en trait plein). Elles sont normalisées à l'unité.

La forme latérale de la gerbe est maintenant exploitée avec la largeur totale de la gerbe $\omega_{tot 1}$ et la largeur estimée dans trois cellules $\omega_{3strips}$. Leurs distributions sont montrées sur la figure 4.6. Les autres variables du premier compartiment utilisées dans l'identification des électrons isolés ne sont pas discriminantes pour des électrons non-isolés.

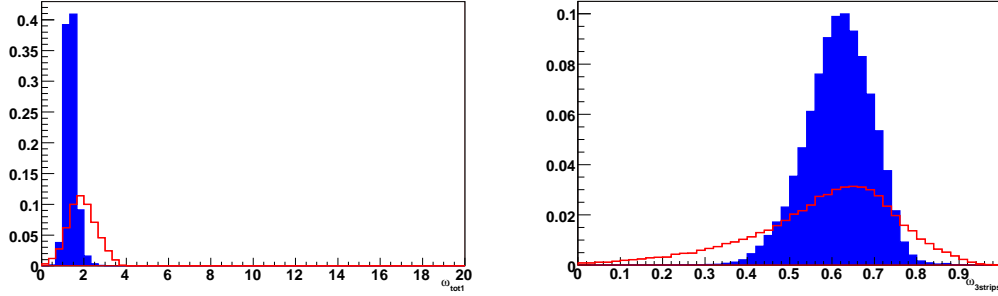


Fig. 4.6: Largeur totale de la gerbe ω_{tot1} (à gauche) et dans trois strips $\omega_{3strips}$ (à droite) dans le premier compartiment du calorimètre électromagnétique. La structure en double pic pour les électrons reflète la granularité variable dans les bouchons du calorimètre électromagnétique. Les distributions sont montrées pour des électrons (histogrammes hachurés) et des pions du lot de bruit de fond (histogrammes en trait plein). Elles sont normalisées à l'unité.

Utilisation des informations du détecteur interne

Contrairement à l'algorithme de reconstruction standard, celui utilisé ici a pour graine de départ une trace. L'association trace-amas électromagnétique n'est alors plus implicite mais explicite. Un ensemble de coupures de bonne qualité des traces a donc déjà été déjà appliqué. On vérifie l'association angulaire entre la trace et le calorimètre électromagnétique (cf. la figure 4.7). On utilise la variable donnant la différence de position en η calculée sur \pm sept cellules autour la trace extrapolée dans le calorimètre électromagnétique. La variable est calculée par $|\Delta\eta| = \frac{\sum_{i=i_m-7}^{i=i_m+7} E_i \times (i - i_m)}{\sum_{i=i_m-7}^{i=i_m+7} E_i}$, où i_m est le point d'impact de la trace extrapolée, E_i est l'énergie reconstruite dans la i -ème cellule dans la direction de η pour une valeur de ϕ constante.

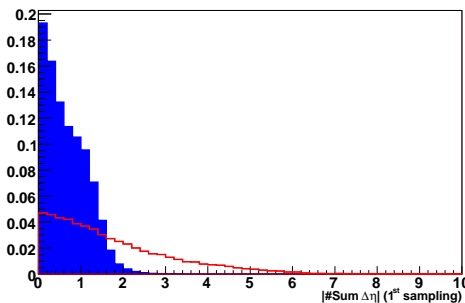


Fig. 4.7: Différence angulaire en pseudorapacité ($|\Delta\eta|$) entre la position de la trace extrapolée au calorimètre électromagnétique et celle de l'amas électromagnétique. La différence est calculée sur \pm sept cellules autour la trace extrapolée. Les distributions sont montrées pour des électrons (histogramme hachuré) et des pions (histogramme en trait plein). Elles sont normalisées à l'unité.

Suivant ces critères, l'énergie E reconstruite dans le calorimètre électromagnétique est comparée à l'impulsion p reconstruite dans le détecteur interne (cf. figure 4.8 à gauche). Dans le cas d'un électron, l'impulsion doit être en accord avec l'énergie. Cependant, de large queues à de basses et grandes valeurs du rapport E/p peuvent être vues. Celles-ci sont dues à des électrons de conversion ainsi qu'à du rayonnement de freinage.

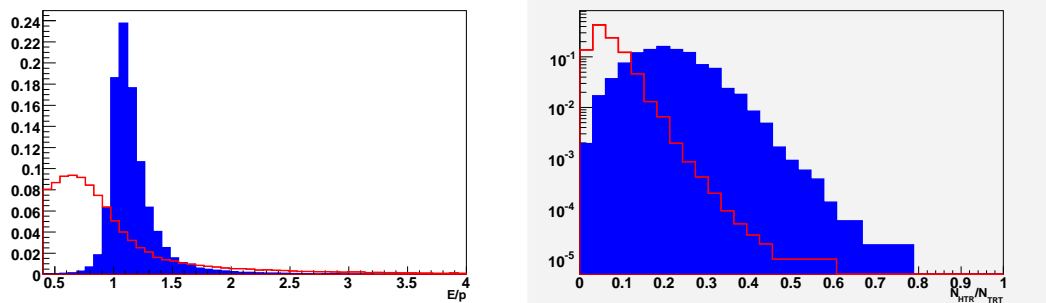


Fig. 4.8: Rapport E/p entre l'énergie reconstruite dans le calorimètre électromagnétique et l'impulsion reconstruite dans le trajectographe (à gauche) et rapport $N_{\text{HTR}}/N_{\text{straw}}$ (à droite) entre le nombre de points de mesure de haut seuil dans et le nombre de pailles touchées dans le TRT. Les distributions sont montrées pour des électrons (histogrammes hachurés) et des pions (histogrammes en trait plein). Elles sont normalisées à l'unité.

On utilise l'information provenant du TRT. La figure 4.8 (à droite) montre le rapport $N_{\text{HTR}}/N_{\text{straw}}$ entre le nombre de points de mesure à haut seuil N_{HTR} et le nombre total de points de mesure dans le TRT N_{straw} .

Finalement on utilise la distribution du paramètre d'impact transverse de la trace. L'algorithme a en effet été développé pour identifier des électrons dans les jets. Cette variable permet donc de différencier les électrons issus de la désintégration semi-leptonique de quarks b de pions. La distribution obtenue est montrée sur la figure 4.9.

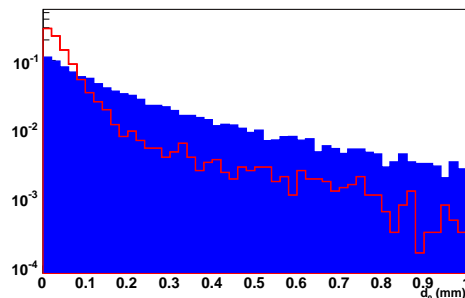


Fig. 4.9: Distributions du paramètre d'impact transverse (d_0) pour des électrons (histogramme hachuré) et des pions (histogramme en trait plein). Elles sont normalisées à l'unité.

4.2.3 Identification des électrons

L'efficacité d'identification des électrons ε_e est définie comme $\varepsilon_e = \frac{N_e^t}{N_e}$, où N_e est le nombre d'électrons, ayant une trace de bonne qualité, i.e pour laquelle la trace a été extrapolée. N_e^t est le nombre d'électrons identifiés après les coupures de la sélection. Cette efficacité permet d'estimer les performances de l'algorithme. Elle ne correspond pas à l'efficacité totale de reconstruction et d'identification des électrons. L'identification est basée sur la valeur d'une fonction discriminante associée à cette trace. Les traces qui ont une valeur plus petite qu'un certain seuil sont rejetées. La courbe de facteur de rejet des pions en fonction de l'efficacité d'identification des électrons est obtenue en faisant varier la valeur du seuil. Le facteur de rejet des pions est défini comme $R_\pi = N_\pi/N_\pi^t$, où N_π est le nombre de traces de pions de bonne qualité et N_π^t est le nombre de traces de pions de bonne qualité identifiées comme un électron.

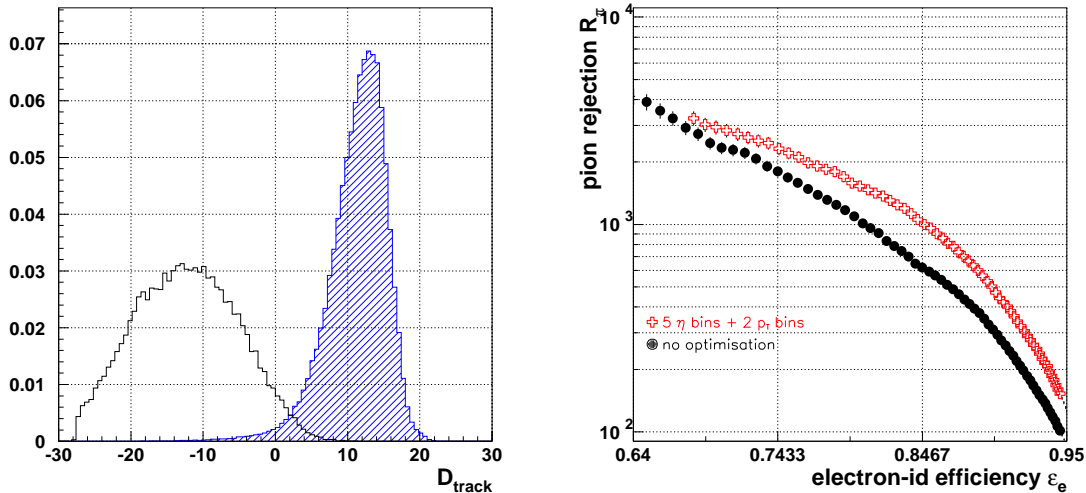


Fig. 4.10: A gauche : Fonction discriminante X_{RL} pour des électrons (histogramme hachuré) et des pions. A droite : facteur de rejet des pions R_π dans les événements $b\bar{b} \rightarrow \mu(6)X$ en fonction de l'efficacité d'identification des électrons de signal dans le lot $B_d \rightarrow J/\psi(ee)K_S$. Les courbes de réjection obtenues avec et sans optimisation des PDFs sont montrées.

Comme on a pu le constater, les variables utilisées sont bien moins discriminantes que pour des électrons de haut p_T . Une série de coupures ne pourra donner de bonnes performances. Il nous faut exploiter la totalité de l'information disponible et en particulier la forme des distributions. De nombreuses méthodes d'analyse statistique existent pour cela. On utilise ici la plus commune, à savoir un rapport de vraisemblances construit à partir des distributions des variables discriminantes pour le signal et le bruit de fond qui doivent être séparées :

$$X_{RL} = \frac{\prod_i g_i^s(x_i)}{\prod_i g_i^b(x_i)}, \quad (4.1)$$

où $g_i^s(x_i)$ (resp. $g_i^b(x_i)$) est la distribution de densité de probabilité à une dimension de la variable x_i pour les événements de signal (resp. de bruit de fond). Ces distributions, appelées fonctions de densité de probabilité (PDF - Probability Density Function) sont construites à partir d'une fraction (environ 20%) des données, qui ne sont alors plus utilisées pour estimer les performances. Comme les variables utilisées dépendent de la pseudo-rapidité, les PDFs ont été construites dans les mêmes intervalles que ceux définis pour l'identification des électrons isolés. La dépendance avec l'impulsion transverse est moins marquée. Deux intervalles en p_T sont tout de même considérés pour optimiser l'identification : $2 \leq p_T < 10$ GeV/c et $p_T \geq 10$ GeV/c. Les distributions de X_{RL} obtenues pour des électrons et des pions sont montrées sur la figure 4.10. Ce rapport ne prend pas en compte les corrélations entre les variables. Cela se reflète en partie dans les distributions qui présentent un recouvrement et des queues importantes. La table 4.2 et

optimisation	ϵ_e		
	70 %	80 %	90 %
aucune	2400 ± 170	1050 ± 50	293 ± 7
2 intervalles en p_T	2627 ± 190	1180 ± 58	345 ± 9
5 intervalles en η et 2 en p_T	2900 ± 305	1550 ± 150	450 ± 13

Tab. 4.2: Facteur de rejet des pions R_π pour différentes valeurs de ϵ_e , avec une construction des PDFs optimisée par intervalles de η et p_T .

la figure 4.10, à droite, détaillent les performances obtenues pour différentes valeurs d'efficacité d'identification. Pour une efficacité d'identification des électrons $\epsilon_e = 80\%$, le facteur de rejet des pions est d'environ 1500. L'optimisation des fonctions de densité de probabilité à la fois en η et p_T permet d'améliorer considérablement les performances.

Les performances peuvent être vérifiées avec le lot inclusif d'événements $b\bar{b}$. Pour un événement donné, la méthode décrite précédemment est appliquée. En particulier de nouvelles PDFs sont déterminées. La figure 4.11 montre le facteur de rejet individuel des pions en fonction de l'efficacité d'identification des électrons. La plupart des électrons sont originaires de désintégrations semi-leptoniques de quarks lourds. Cependant, certains peuvent aussi provenir de conversions ou de désintégrations Dalitz. Les résultats sont donc détaillés pour les électrons de signal (points) et pour tous les électrons (carrés). Pour une efficacité d'identification des électrons de $\sim 90\%$ le facteur de rejet des pions est de ~ 480 , en bon accord avec les résultats de la table 4.2. Le facteur de rejet est $\sim 40\%$ plus bas lorsque les PDFs sont construites avec tous les électrons. La figure 4.12 montre le facteur de rejet des événements de bruit de fond en fonction de l'efficacité d'identification des événements de signal (avec $b \rightarrow e$). Les performances sont plus basses que pour les particules individuelles car les événements de bruit de fond peuvent être sélectionnés par des traces mal-identifiées comme des électrons.

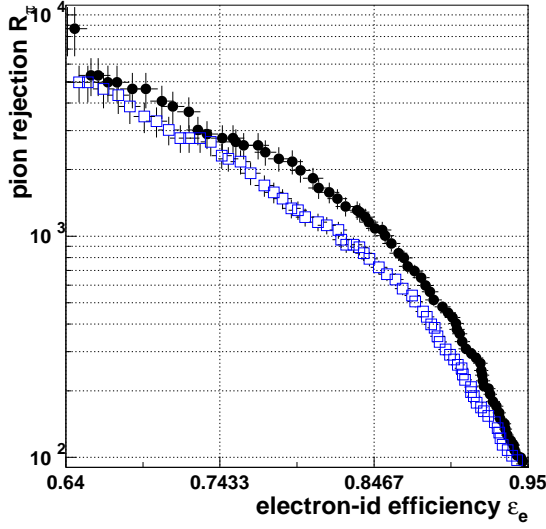


Fig. 4.11: Facteur de rejet des pions R_π en fonction de l'efficacité d'identification des électrons dans les événements $b\bar{b} \rightarrow \mu(6)X$. Les points correspondent aux électrons de signal et les carrés à tous les électrons.

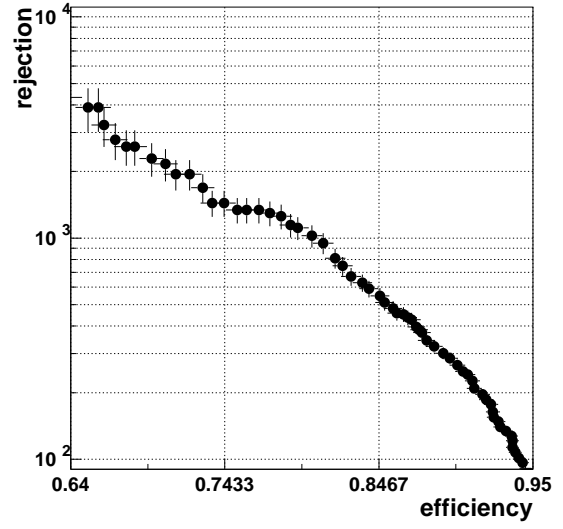


Fig. 4.12: Facteur de rejet des événements sans désintégration $b \rightarrow e$ en fonction de l'efficacité d'identification des événements avec $b \rightarrow e$.

4.3 Reconstruction du J/ψ

4.3.1 Sélection des événements

Pour un événement donné on apparie tous les électrons sélectionnés avec une efficacité fixée à 90%. Ces paires permettent de calculer l'acceptance $a_{J/\psi} \sim 32\%$ des événements J/ψ qui est une mesure des pertes dues aux coupures de qualité sur les traces, l'identification des électrons ainsi que le rejet des électrons se trouvant dans la zone pauvrement instrumentée du calorimètre électromagnétique. Les paires de traces reconstruites ayant des charges opposées sont ajustées à un même vertex commun. Les ajustements réussis, ayant un χ^2 par degré de liberté inférieur à 6, sont retenus. La longueur L_{xy} ajustée doit être plus grande que $250 \mu\text{m}$. Elle est calculée par : $L_{xy} = \frac{\vec{D} \cdot \vec{p}_T(J/\psi)}{\|\vec{p}_T(J/\psi)\|}$, où D est la distance entre le premier et le second vertex et $p_T(J/\psi)$ est l'impulsion transverse reconstruite du J/ψ .

La figure 4.13 montre la distribution de masse invariante reconstruite des paires de candidats électrons, sur les événements de signal, après l'application de ces différents critères. Un bruit de fond combinatoire important est visible lorsqu'on demande seulement des traces de bonnes qualités de charges opposées. Il est grandement réduit lorsqu'on applique les critères d'identification des électrons (histogramme en trait plein). Enfin l'histogramme hachuré montre la masse

invariante obtenue après l'application de toutes les coupures d'analyse.

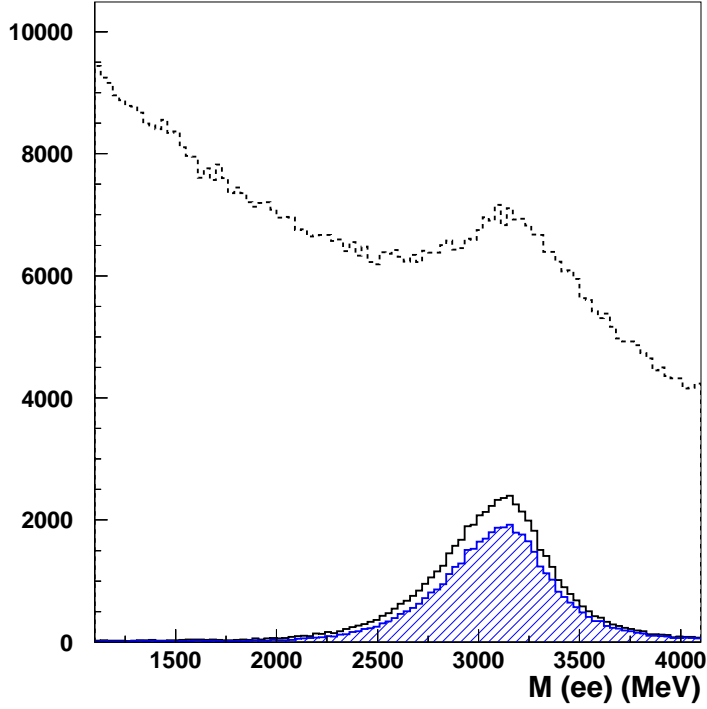


Fig. 4.13: Distributions de la masse invariante pour toutes les combinaisons de traces dans les événements de signal, pour toutes les paires de charges opposées après les coupures de qualité des traces (histogramme en trait tireté), après l'identification des électrons (histogramme avec ligne solide), et après application de toutes les coupures d'analyse (histogramme hachuré).

L'efficacité $\epsilon_{J/\psi}$ de reconstruction du J/ψ est définie comme le rapport entre le nombre de candidats passant tous les critères d'analyse, situés dans une fenêtre de masse $2,1 - 4,1 \text{ GeV}/c^2$, sur le nombre de candidats passant les critères d'acceptance. Elle est typiquement de 60%, correspondant à une efficacité totale $a_{J/\psi} \times \epsilon_{J/\psi} \sim 20 \%$.

Dans la figure 4.14 les distributions des distances entre les deux traces, $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, sont montrées pour toutes les combinaisons. Pour les événements de signal, presque aucune combinaison n'a de grande valeur de ΔR , au contraire du bruit de fond. Le critère supplémentaire, $\Delta R < 1,5$, n'a presque pas d'influence sur le nombre d'événements de signal et, de plus, n'affecte pas la forme de la distribution des événements de bruit de fond.

4.3.2 Simulation du bruit de fond

Le bruit de fond consiste en événements $b\bar{b}$, où l'un des hadron- B est forcé à se désintégrer en un muon et l'autre hadron peut se désintégrer librement. Lorsqu'on calcule le rapport signal

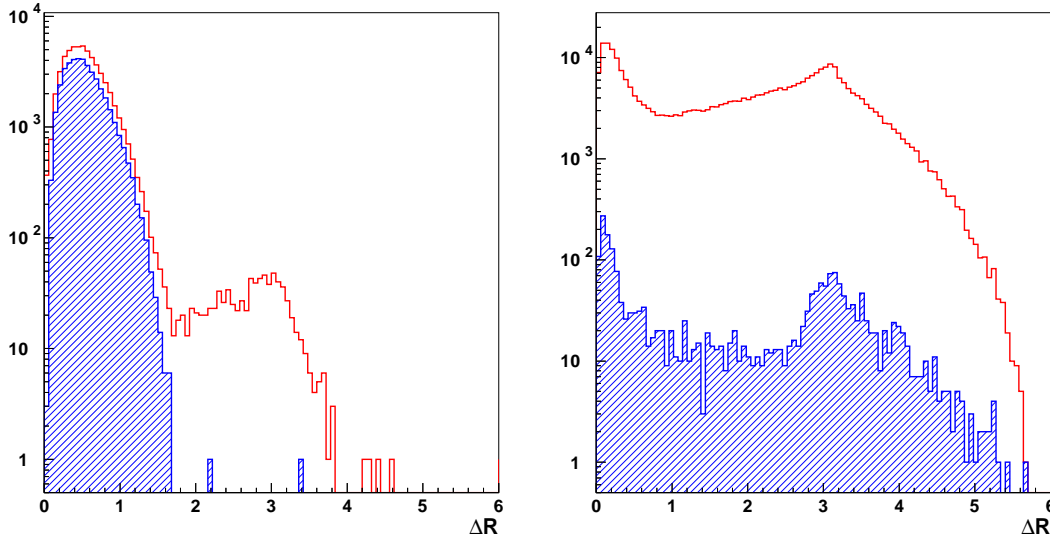


Fig. 4.14: Distributions ΔR pour les événements de signal (à gauche) et de bruit de fond (à droite). Les distributions sont montrées pour toutes les combinaisons (histogramme en trait plein) et après toutes les critères d'analyse (histogramme hachuré), sauf pour la coupure sur ΔR . Pour les événements de bruit de fond, l'identification des électrons n'est pas appliquée.

sur bruit, le facteur suivant doit être pris en compte :

$$R_{BR} = \frac{Br(B \rightarrow X)}{Br(B \rightarrow J/\psi X) \times Br(J/\psi \rightarrow e^+e^-)} = \frac{1}{0,011 \times 0,06} = 1500. \quad (4.2)$$

Etant donné le nombre d'événements de signal, environ $\sim 250 \times 10^3$, le nombre d'événements de bruit de fond nécessaire ne peut être simulé dans un temps raisonnable. Aussi nous devons basé notre étude sur une simulation rapide pour ceux-ci. Pour ce faire, l'analyse est appliquée sur les 90×10^3 événements de bruit de fond ayant subi la simulation complète. Dans la fenêtre de masse $2,1 - 4,1 \text{ GeV}/c^2$, après toutes les étapes de l'analyse, il ne reste que trois événements. Ceci correspond à une efficacité $\varepsilon_{b\bar{b}} \sim 3.3 \times 10^{-5}$. La rapport de branchement effectif à prendre en compte est donc :

$$R_{BR}^{eff} = R_{BR} \times \varepsilon_{b\bar{b}} \times \varepsilon_{J/\psi} \times a_{J/\psi}. \quad (4.3)$$

La forme de la distribution à produire est obtenue uniquement par des pions, en supposant que ceux-ci sont bien représentatifs du spectre. La distribution de masse invariante des paires de candidats est montré sur la figure 4.15 et est ajustée, dans la fenêtre de masse, par une fonction composée de deux exponentielles décroissantes. Une fois la forme de la distribution connue, ainsi que le nombre d'événements à simuler, ceux-ci sont engendrés par la technique Monte Carlo, avec la méthode d'acceptation-rejet. Le nombre d'événements complètement simulés ayant passé l'analyse étant petit, il y a une grande erreur statistique sur le nombre d'événements à produire. Pour estimer la sensibilité des performances à cette incertitude, celles-ci seront es-

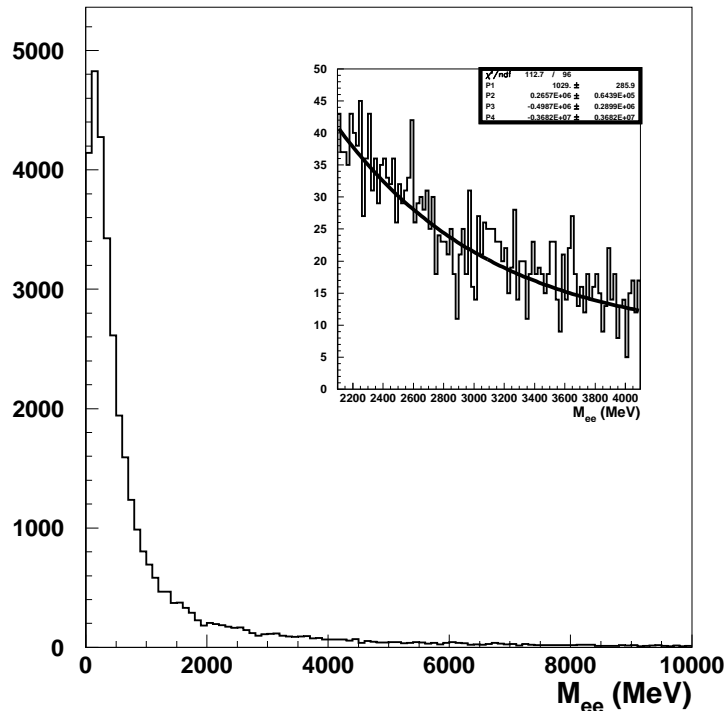


Fig. 4.15: Distribution de la masse invariante des paires de traces issues de $b\bar{b} \rightarrow \mu(6)X$. Un zoom dans la fenêtre d'intérêt est montré ainsi qu'un ajustement de deux exponentielles décroissantes.

timées avec des scénarii extrêmes, où des nombres d'événements minimaux et maximaux, à 95% de niveau de confiance, seront engendrés selon une statistique Poissonienne.

4.3.3 Reconstruction de la masse invariante

La masse du J/ψ peut être déterminée par l'énergie et la direction des électrons. Les mesures les plus précises seront effectuées à l'aide du détecteur interne uniquement, avec le risque de biaiser la direction de la trace par le rayonnement de freinage. Pour les études d'étalonnage du calorimètre électromagnétique nous utiliserons l'énergie estimée dans ce détecteur. Par la suite l'efficacité d'identification des électrons est optimisée à 90%.

La figure 4.16 à gauche montre la distribution de masse invariante des paires d'électrons reconstruite en n'utilisant que les informations du détecteur interne. La fonction ajustée pour le signal est une distribution de Breit-Wigner de forme $\sim \Gamma/(\Delta m_0^2 + (\Gamma/2)^2)$ à gauche du pic de masse m_0 et une gaussienne de largeur σ_{droite} à droite. Le paramètre σ_{droite} caractérise la résolution effective de la distribution de masse invariante, tandis que Γ est une mesure de la détérioration de la résolution due au rayonnement de freinage. Le bruit de fond est ajusté par la somme de deux exponentielles décroissantes. La forme de la distribution n'est pas très bien reproduite, en particulier pour le pic. Les valeurs ajustées des paramètres $\Delta m_0 = m_0 - M_{J/\psi}$,

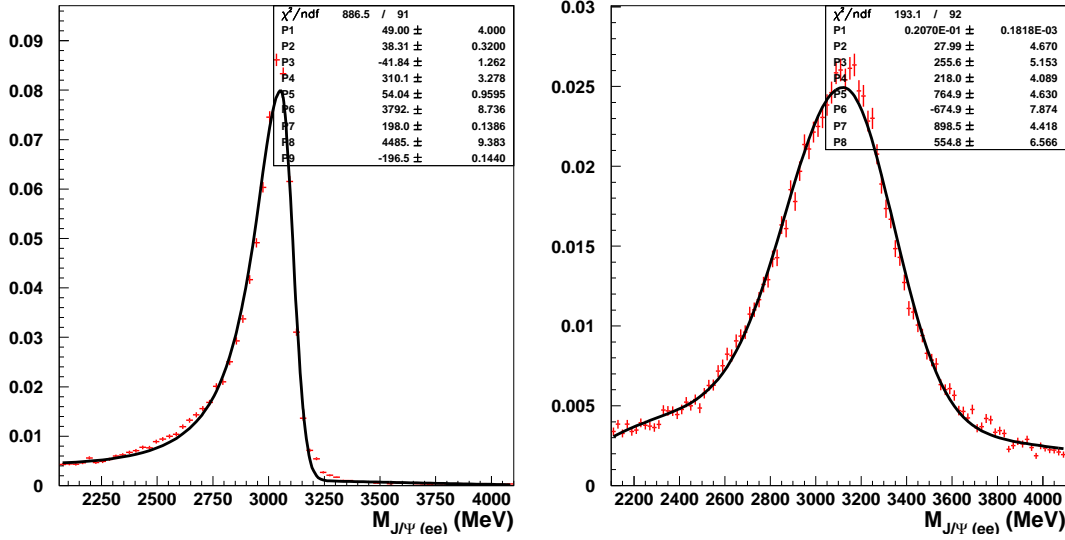


Fig. 4.16: Distribution de masse invariante des paires d'électrons en utilisant uniquement les informations du détecteur interne (à gauche) et avec l'énergie reconstruite par le calorimètre électromagnétique (à droite). L'efficacité d'identification des électrons est de 90%. Le signal est ajusté, dans le cas de l'utilisation uniquement des informations du détecteur interne, par une distribution de Breit-Wigner à gauche du pic et par une gaussienne à droite. Lors de l'utilisation des informations calorimétriques le signal est ajusté par deux gaussiennes à gauche et à droite du pic. Le bruit de fond est ajusté par la somme de deux exponentielles décroissantes.

où $M_{J/\psi} = 3096 \text{ MeV}/c^2$ est la masse nominale du J/ψ , Γ et σ_{droite} sont données dans la table 4.3. Ces valeurs sont données de manière séparées pour le tonneau du TRT, lorsque les deux électrons ont leur pseudorapidité η dans $\pm 0,7$ et pour les bouchons du TRT, lorsqu'au moins l'un des électrons a $|\eta| > 0,7$. Les paramètres de la fonction de Breit-Wigner ne sont pas trop

η intervalle	Δm_0 MeV/ c^2	Γ MeV/ c^2	σ_{droite} MeV/ c^2
tout	$-41,8 \pm 1,3$	$310,1 \pm 3,3$	$54,0 \pm 1,0$
tonneau	$-36,0 \pm 2,0$	$200,7 \pm 3,7$	$47,2 \pm 1,2$
bouchon	$-58,8 \pm 2,0$	$387,7 \pm 6,1$	$66,4 \pm 1,5$

Tab. 4.3: Résultats de l'ajustement de la distribution de masse invariante des paires d'électrons par une fonction se comportant comme une distribution de Breit-Wigner à gauche du pic et comme une gaussienne à droite, en utilisant uniquement les informations du détecteur interne. L'efficacité d'identification des électrons est de 90%.

affectés par la présence des événements de bruit de fond. Le décalage de la masse est d'environ $42 \text{ MeV}/c^2$, plus important dans les bouchons que dans la tonneau. Le paramètre Γ est d'environ $300 \text{ MeV}/c^2$, bien plus important dans la partie bouchon que dans la tonneau. La largeur de la gaussienne, estimée à partir de la partie droite de la distribution, est de $54 \text{ MeV}/c^2$.

La figure 4.16 à droite montre la distribution de la masse invariante des paires de candidats reconstruite avec les informations du calorimètre électromagnétique. L'efficacité $\epsilon_{J/\psi}$ de reconstruction du J/ψ , telle que définie ci-dessus, est d'environ 64%. La distribution est légèrement asymétrique. L'ajustement utilise alors la somme de deux gaussiennes, à gauche et à droite du pic, avec des largeurs différentes, σ_{gauche} et σ_{droite} . Le paramètre σ_{droite} caractérise la résolution effective de la masse invariante tandis que σ_{gauche} est une mesure de la détérioration de la résolution due au rayonnement de freinage. Les valeurs ajustées des paramètres sont montrées dans la table 4.4. Les performances sont données de manière séparées pour la partie tonneau du calorimètre électromagnétique, lorsque les deux électrons ont leur pseudorapidité η dans $\pm 1,4$, pour la région des bouchons, lorsqu'au moins l'un des électrons a $|\eta| > 1,4$. La table 4.4 donne les paramètres de la gaussienne ajustée. La masse du J/ψ est décalée de $\sim 37 \text{ MeV}/c^2$. Ce décalage est particulièrement important dans les bouchons où il atteint $\sim 80 \text{ MeV}/c^2$. La résolution, obtenue à partir de σ_{droite} , est de $\sim 220 \text{ MeV}/c^2$.

η intervalle	Δm_0 MeV/c^2	σ_{gauche} MeV/c^2	σ_{droit} MeV/c^2
tout	$28,0 \pm 4,7$	$255,6 \pm 5,1$	$218,0 \pm 4,1$
tonneau	$10,3 \pm 5,2$	$249,8 \pm 4,6$	$234,2 \pm 4,7$
bouchon	$72,0 \pm 10,5$	$235,2 \pm 9,9$	$165,8 \pm 9,5$

Tab. 4.4: Résultats de l'ajustement de deux gaussiennes, où la mesure de la direction est prise du détecteur interne et l'énergie reconstruite du calorimètre électromagnétique. L'efficacité d'identification des électrons est de 90%.

La sensibilité des paramètres est étudiée en minimisant et maximisant la quantité de bruit de fond qui est ajoutée. La table 4.5 donne les résultats de ces ajustements. Les paramètres de la partie gaussienne de l'ajustement ne sont pas sensibles au niveau de bruit de fond utilisé. Les

bruit de fond	Δm_0 (MeV/c^2)	σ_{gauche} (MeV/c^2)	σ_{droit} (MeV/c^2)	S/B
minimal	$26,6 \pm 4,1$	$257,9 \pm 3,6$	$222,3 \pm 3,6$	3,5
optimal	$23,1 \pm 4,1$	$257,1 \pm 3,4$	$226,4 \pm 3,5$	2,2
maximal	$33,2 \pm 4,8$	$253,8 \pm 4,5$	$209,0 \pm 4,4$	1,0

Tab. 4.5: Résultats de l'ajustement pour la reconstruction de la masse invariante du J/ψ avec les informations du calorimètre électromagnétique pour différents niveaux de bruit de fond.

nombre d'événements de signal et de bruit de fond reconstruits peuvent être obtenus de ces ajustements. Ainsi l'efficacité de reconstruction des événements de J/ψ est de 54%. Environ 10% des candidats J/ψ sont mal identifiés comme étant du bruit de fond, essentiellement sur les bords de l'intervalle de masse utilisé, $2,1-4,1 \text{ GeV}/c^2$. Le rapport signal sur bruit peut ainsi être estimé à $S/B = 2,2 \pm 1,2$, où l'erreur statistique provient des estimations effectuées à partir des nombres extrêmes de bruit de fond mis dans la simulation.

4.4 Inter-étalonnage du calorimètre électromagnétique

Nous avons vu aux paragraphes 1.3.2 et au chapitre 2 toute l'importance que revêt la détermination de l'échelle d'énergie du calorimètre électromagnétique, son inter-étalonnage et particulièrement le besoin d'une étude à basse énergie, en particulier en utilisant le J/ψ . Cette étude est le but initial de mon travail sur la reconstruction des électrons. Pour déterminer les coefficients d'inter-étalonnage à basse énergie du calorimètre électromagnétique, nous suivons la méthode utilisée pour les événements $Z \rightarrow e^+e^-$, décrite dans [112].

Pour chaque région i un biais α_i est engendré suivant une distribution gaussienne centrée à 0, avec une variance σ_{bias}^2 . Tous les électrons tombant dans la région i ont leur énergie E_i biaisée suivant : $E_i^{bias} = E_i \times (1 + \alpha_i)$. La masse invariante M_{ij} de deux électrons de régions i et j est donnée par $M_{ij} = \sqrt{2E_i E_j (1 - \cos \theta)}$, où θ est l'angle entre eux. La masse invariante biaisée M_{ij}^{bias} est reliée à la masse originale M_{ij} par :

$$M_{ij}^{bias} = M_{ij} \times \left(1 + \frac{\alpha_i + \alpha_j}{2}\right) = M_{ij} \times \left(1 + \frac{\beta_{ij}}{2}\right), \quad (4.4)$$

où $\beta_{ij} = \alpha_i + \alpha_j$ si les termes de second ordre sont négligés. La masse invariante biaisée pour un couple de régions (i, j) dépend linéairement du seul paramètre β_{ij} . Elle peut être ajustée à une distribution de référence, avant tout biais. Les paramètres β_{ij} et leurs erreurs σ_{ij}^β peuvent alors être déterminés. En testant tout le domaine des valeurs de β_{ij} , une minimisation de χ^2 entre les distributions de référence et biaisée est effectuée, permettant l'extraction des paramètres β_{ij} et σ_{ij}^β . Pour retrouver les paramètres α_i , la méthode des moindres carrés est appliquée. La précision de l'inter-étalonnage est donnée par la largeur de la fonction gaussienne ajustée à la distribution des "pull" des biais α_i^{rec} reconstruits :

$$\text{pull}_\alpha = \frac{\alpha_i^{rec} - \alpha_i}{\sigma_{\alpha_i^{rec}}} \quad (4.5)$$

Pour obtenir la distribution de référence, les événements de bruit de fond doivent être soustraits ; ceci est une complication supplémentaire par rapport au cas du $Z \rightarrow e^+e^-$ qui en est presque exempt. Des histogrammes d'événements de signal et de bruit de fond sont construits pour chaque couple de région (i, j) à partir des paramètres des ajustements effectués. Le bruit de fond est soustrait de la distribution de masse biaisée en utilisant l'histogramme de masse de référence du bruit de fond. La distribution de masse invariante ainsi obtenue est ajustée sur la distribution de référence. Seuls les couples de régions ayant un nombre d'événements suffisant, quinze en pratique, sont utilisés. Les valeurs de β_{ij} et leurs incertitudes σ_{ij}^β sont ainsi extraites de la procédure de minimisation de χ^2 décrite ci-dessus. Cette minimisation est effectuée dans un intervalle étroit en masse, typiquement entre 2, 1 et 4, 1 GeV/c². Les valeurs de β_{ij} sont alors trouvées facilement. Cette méthode est a priori sensible à la forme de la distribution de bruit de fond qui a été simulée, ainsi qu'au faible nombre d'événements de bruit de fond ayant passé la

simulation complète. Cependant, la minimisation est effectuée dans une fenêtre de masse étroite, où les paramètres de l'ajustement gaussien ne sont pas sensibles à la forme du bruit de fond. Ainsi la forme et la quantité de bruit de fond ne devraient pas être un point critique de l'analyse.

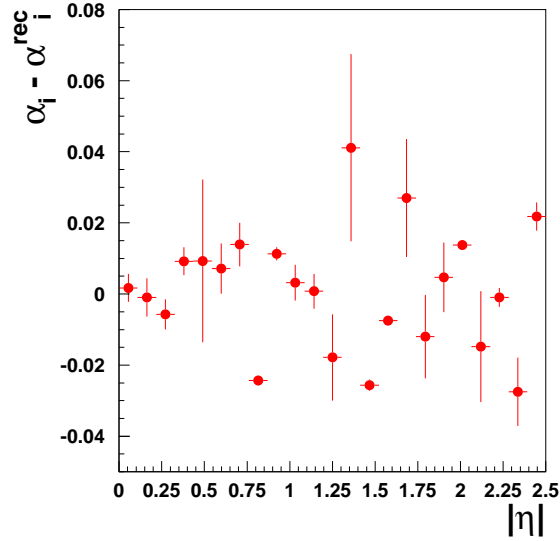


Fig. 4.17: Différences, pour les événements de signal, entre les biais α_i injectés et les α_i^{rec} reconstruits, dans un anneau azimuthal de taille $\Delta\eta = 0, 1$.

Une autre complication est que les électrons de basse énergie sont très sensibles à la quantité de matière présente en amont du calorimètre électromagnétique. Pour tester la sensibilité des performances à cet effet, un total de cinquante régions est défini dans $|\eta| < 2, 5$. A chaque région i correspond un anneau azimuthal de taille $\Delta\eta = 0, 1$. La méthode est appliquée sans injecter de biais, $\alpha_i = 0$, sur les événements de signal uniquement. Dans la figure 4.17 les différences entre les biais α_i injectés et les α_i^{rec} reconstruits sont présentés. Des structures similaires peuvent être observées dans la figure 4.2 du fait de la reconstruction non optimale de l'énergie des électrons. Comme l'effet du matériel en amont du calorimètre électromagnétique ne doit pas être confondu avec l'inter-étalonnage, la méthode ne sera appliquée qu'aux événements ayant les deux électrons dans $|\eta| < 0, 8$.

De plus, comme montré sur la figure 4.14, la forme de la distribution des distances ΔR entre les paires de traces est différente pour les événements de signal et de bruit de fond. Ceci implique une variation des rapports signal sur bruit (S/B) dans les différentes régions (i, j) . Si la distance entre les régions est trop grande alors le rapport S/B est trop petit et ce couple de régions n'est pas pris en compte. Du fait de la variation des rapports signal sur bruit, différents histogrammes de masse de référence sont construits pour huit régions en η . Dans la figure 4.18 quelques exemples de distributions de référence sont présentées pour différentes valeurs

de $\Delta\eta(i, j)$, pour les événements de signal uniquement et avec le bruit de fond inclus. Les ajustements aux distributions de masse invariante sont montrés.

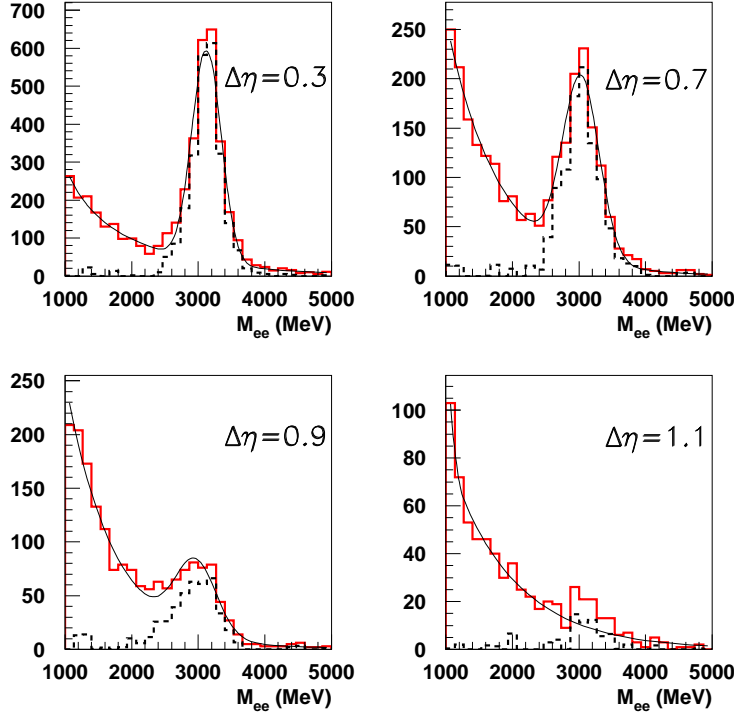


Fig. 4.18: Les distributions de référence pour différents $\Delta\eta(i, j)$, pour le signal uniquement (histogramme tireté) et avec le bruit de fond inclus (histogramme plein). Les ajustements à ces distributions sont montrés.

Du fait de la faible statistique disponible, seules 32 régions sont définies, 8 en η et 4 en ϕ , ce qui correspond à une taille de $\Delta\eta \times \Delta\phi = 0,2 \times 1,57$. Dans l'intervalle $|\eta| < 0,8$, un total de 22×10^3 événements de signal survivent aux coupures effectuées. Pour chacune des trente deux régions un biais $\sigma_{bias} = 2.5\%$ est injecté. Le nombre de couples de régions (i, j) non vides est de 143. La figure 4.19 montre les distributions des pulls pour les β_{ij}^{rec} et α_i^{rec} reconstruits. La largeur de la gaussienne ajustée à ces distributions est proche de l'unité. La distribution des résidus des α est présenté sur la figure 4.20. Le biais peut être retrouvé avec une précision donnée par la largeur de la gaussienne ajustée sur cette distribution, soit $0,6\% \pm 0,3\%$. La statistique utilisée, 22×10^3 événements dans $|\eta| < 0,8$, est équivalente à 308×10^3 événements dans 448 régions. Elle peut être extrapolée à 0,4 % avec 800×10^3 événements, en bon accord avec les résultats reportés dans le TDR [16]. Lorsqu'on effectue cette extrapolation on ne doit pas oublier que les événements reconstruits dans $|\eta| < 0,8$ sont les mieux reconstruits et que la précision sur l'inter-étalonnage ainsi obtenue est sans doute un peu surestimée. Ce résultat

a été vérifié en changeant le nombre de régions de 32 à 16, donc en augmentant la statistique à l'intérieur des régions. Une précision de $\sigma = 0,3\%$ est alors obtenue en bon accord avec celle obtenue avec les trente deux régions. En utilisant uniquement les événements de signal une précision de $\sigma = 0,3\%$ est aussi obtenue indiquant la faible sensibilité du résultat à la quantité de bruit de fond. Finalement, différentes valeurs du biais σ_{bias} ont été utilisées, de 1,5% et 5%, donnant des résultats similaires à ceux obtenus avec un biais de $\sigma_{bias} = 2,5\%$.

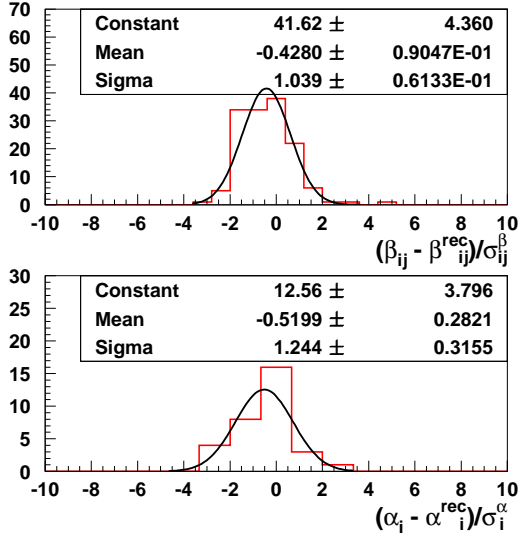


Fig. 4.19: Distributions des "pull" de β et α pour les 32 régions utilisées dans $|\eta| < 0,8$.

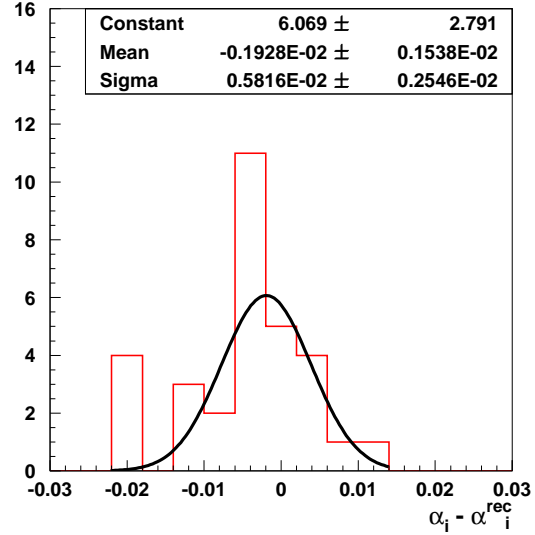


Fig. 4.20: Distribution des résidus sur α pour les 32 régions dans $|\eta| < 0,8$.

4.5 Travaux plus récents

Les travaux sur le J/ψ auxquels j'ai participé depuis ont été pour l'essentiel une vérification des performances avec des versions d'Athena et de géométrie du détecteur plus récentes [111].

4.5.1 Syst`eme de d`eclenchement

Au démarrage du LHC, la luminosité sera de $10^{31} \text{cm}^{-2} \text{s}^{-1}$. La stratégie, développée récemment [113] [114] [115], pour déclencher sur des événements J/ψ ou $\Upsilon \rightarrow e^+e^-$ est basée sur le menu 2e5, au niveau L1, sur l'utilisation de régions d'intérêt électromagnétiques à très bas E_T (en pratique deux amas au-dessus d'un seuil de 3 GeV), et sur l'utilisation aux autres niveaux d'informations d'identification des électrons dans le calorimètre électromagnétique et dans le détecteur interne (avec des seuils montant alors à 5 GeV). Ainsi, après le niveau 1 et la sélection des événements hors ligne, un total d'environ $1,7 \times 10^5 J/\psi$ et $3,4 \times 10^5 \Upsilon$ est attendu pour une luminosité intégrée de 100pb^{-1} et un taux de déclenchement de fond de l'ordre de

6 Hz [108].

Durant la phase de basse luminosité ($10^{33} \text{cm}^{-2} \text{s}^{-1}$), cette même stratégie ne pourra être appliquée car le taux de déclenchement sera alors bien trop important. Deux stratégies sont envisagées pour étendre les menus de physique du B à des états finals comportant des désintégrations électromagnétiques. Dans les deux cas un seul muon est utilisé au premier niveau et est confirmé au second. Il est alors possible, au déclenchement de haut niveau, de reconstruire les traces dans l'ensemble du détecteur interne (méthode dite "Full scan") ou bien de se restreindre dans un volume autour d'une région d'intérêt de bas E_T identifiée au niveau un (méthode "RoI-guided"). Cette dernière méthode devrait être plus rapide mais ce gain va dépendre du nombre de régions à prendre en considération et donc du seuil en E_T considéré. Les développements et études les plus récents [116] [117] sont prometteurs.

4.5.2 Reconstruction et identification des électrons

Comme nous venons de le voir, la forme de l'amas électromagnétique utilisé ressemble à un ovoïde qui suit la forme de la gerbe électromagnétique. Si cette forme peut être adaptée pour collecter les informations nécessaires à l'identification de la particule, elle ne l'est pas pour la reconstruction de l'énergie ; les méthodes standards utilisées dans Atlas supposant un amas rectangulaire. L'expérience du test en faisceau, qui sera décrit au chapitre 7 a aussi montré que la reconstruction était très sensible à l'alignement du détecteur interne et du calorimètre électromagnétique. Aussi, par souci de simplification, la reconstruction de l'amas a été changée et simplifiée. La trace de la particule étudiée est extrapolée jusqu'au deuxième compartiment du calorimètre électromagnétique. Autour de ce point, un amas rectangulaire est construit de taille $\Delta\eta \times \Delta\phi = 0,125 \times 0,125$, i.e 5×5 dans le second compartiment.

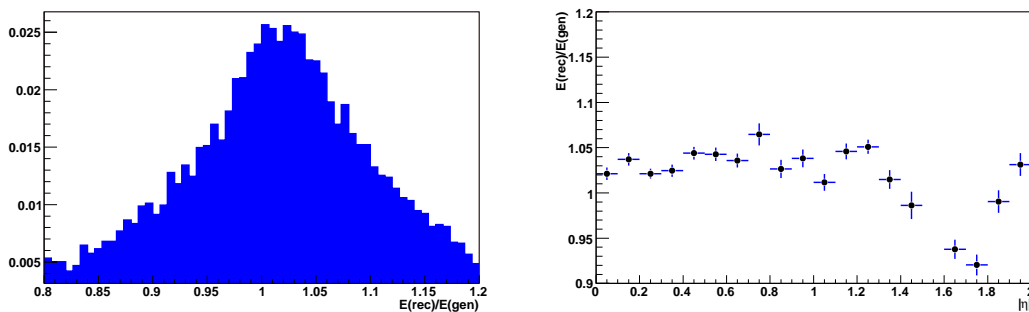


Fig. 4.21: Rapport entre l'énergie reconstruite dans le calorimètre électromagnétique et l'énergie engendrée (à gauche) et profil en fonction de $|\eta|$ (à droite) pour les électrons de signal dans les données CSC.

La figure 4.21 montre le rapport entre l'énergie reconstruite dans le calorimètre électromagnétique et l'énergie engendrée (à gauche) et profil en fonction de $|\eta|$ (à droite) pour

les électrons de signal dans les données CSC. Si la qualité de la reconstruction de l'énergie ne semble guère avoir progressée il n'en est pas de même de l'implémentation des corrections apportées. En effet si dans les études DC1 quelques corrections ont été péniblement appliquées au niveau de nos macros Paw, désormais la même liste de corrections que l'amas standard est appliquée. Il sera désormais possible d'étudier correctement chaque correction, et de vérifier leur pertinence à très basse énergie et d'estimer si des corrections spécifiques à cet amas doivent être appliquées. Enfin, comme nous l'évoquerons dans le chapitre 7 l'étude des données de basse énergie du test en faisceau combiné de 2004 permettra d'étudier la reconstruction de l'énergie ne conditions réelles.

En ce qui concerne la prise en compte du rayonnement de freinage dans la mesure de l'impulsion, de nombreux développements ont été effectués et ont montré [118] qu'il est possible d'améliorer la résolution de la masse reconstruite du J/ψ typiquement de 20%.

4.6 Conclusion

L'utilisation du J/ψ pour l'étude de l'étalonnage du calorimètre électromagnétique a été jusqu'ici à la base de mes activités de recherche. L'effort a été jusqu'ici porté sur la reconstruction des électrons à basse énergie et la mise au point d'une analyse complète montrant la capacité d'une reconstruction de la masse du J/ψ par Atlas pour atteindre ce but. Outre la continuité de cet effort, en particulier pour la reconstruction de l'énergie, il sera nécessaire dans les mois qui viennent d'intégrer les développements effectués par d'autres groupes (déclenchement, rayonnement de freinage) à ceux que j'ai présenté dans ce chapitre. Ce sera l'une des perspectives de mes activités de recherche que je présenterai dans l'annexe 8 en vue de l'analyse des premières données d'Atlas.

Le développement de cet algorithme de reconstruction des électrons mous a eu par ailleurs une autre application, à savoir la reconstruction des électrons dans les jets et l'étiquetage des jets de b qui vont être décrits dans le chapitre qui suit.

5. RECONSTRUCTION DES ÉLECTRONS DANS LES JETS ET ÉTIQUETAGE DES JETS PAR UN ÉLECTRON MOU

La désintégration du quark top, du boson de Higgs ou de particules supersymétriques implique souvent la présence d'un quark b dans l'état final. L'étiquetage des jets de b peut s'effectuer à partir d'un lepton (un électron pour ce qui nous intéresse ici) issu de la désintégration semileptonique du quark b . Ce besoin a été à l'origine du développement de l'algorithme de reconstruction des électrons dans les jets [110] [119] [120]. Il a été montré en particulier dans [121] que l'utilisation de cet algorithme pouvait augmenter le facteur de rejet des jets légers de $\sim 10\%$. Depuis, les algorithmes standards d'étiquetage ont évolué [122] et leurs performances sont meilleures. L'étiquetage des jets de b par des leptons mous est toutefois toujours important, en particulier en vue du démarrage du détecteur, de manière à vérifier les performances des algorithmes standard.

Les lots de données utilisés sont décrits au paragraphe 5.1. La reconstruction et l'identification des électrons dans les jets est présentée dans le paragraphe 5.2. Le paragraphe 5.3 présente l'algorithme d'étiquetage des jets de b et ses performances. Les performances de la reconstruction des électrons dans les jets et de l'étiquetage des jets de b sont celles de la note DC1 [123]. Cette analyse reste encore à l'heure actuelle la plus complète, en particulier car c'est la seule où l'effet des événements d'empilement a pu être étudié. A la fin du chapitre le paragraphe 5.4 évoque les principales évolutions des algorithmes et de leurs performances depuis cette étude.

5.1 Les données utilisées

On utilisera la production d'un boson de Higgs associé à la production d'un boson W , avec une désintégration semi-leptonique en muon de ce dernier $W \rightarrow \mu\nu$. Ces lots ont été produits dans le cadre du Data Challenge 1, avec Pythia 6.203 [97]. Les événements ont été simulés avec Geant 3 [98], basé sur la version 3.2.1 du programme fortran d'Atlas ATLASIM/Dice [99]. Des événements d'empilement ont été ajoutés pour les scénarii de basse et haute luminosité. Ces données ont été reconstruites avec la version d'Athena 7.8.0. La masse du boson de Higgs considéré est $m_H = 120 \text{ GeV}/c^2$. Le lot de signal est composé d'événements $H \rightarrow b\bar{b}$ et les lots de bruit de fond d'événements $H \rightarrow c\bar{c}$ et $H \rightarrow u\bar{u}$. On utilise aussi un lot de $H \rightarrow b\bar{b}$, filtré pour qu'au niveau du générateur les deux quarks b se désintègrent en électron. Il est utilisé pour obtenir un large échantillon d'électrons pour étudier les variables discriminantes. Les coupures de pré-sélection pour ce lot sont les suivantes : les deux quarks b (avant radiation dans l'état

5. RECONSTRUCTION DES ÉLECTRONS DANS LES JETS ET ÉTIQUETAGE DES JETS PAR UN ÉLECTRON MOU

final) doivent avoir une impulsion transverse $p_T^b > 15$ GeV/c et une pseudo-rapidité $|\eta_b| < 2,5$; au moins l'un des électrons doit être dans un cône $\Delta R = \sqrt{[\Delta\eta(e,b)]^2 + [\Delta\phi(e,b)]^2} < 0.4$ autour de l'axe du quark b (avant radiations dans l'état final). La table 5.1 détaille les informations relatives aux lots utilisés.

Processus	lot de données	Nb d'événements		
		pas d'empilement	basse luminosité	haute luminosité
$H \rightarrow b\bar{b}$	002055	20,000		
$H \rightarrow c\bar{c}$	002057	20,000		
$H \rightarrow u\bar{u}$	002056	50,000		25,000

Tab. 5.1: Lots de données utilisées : le processus, le numéro de lot pour la géométrie considérée et le nombre d'événements.

Les jets sont reconstruits avec un algorithme standard de cône, de taille $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.4$. Les jets utilisés pour les études d'étiquetage ont une impulsion transverse minimale de 15 GeV/c et une pseudo-rapidité $|\eta| < 2,5$. Les traces associées à un jet doivent être à une distance inférieure à 0.4 de l'axe du jet. Pour les études de performance, le type du parton (ici du quark) associé au jet est effectué à partir des informations de la vérité Monte Carlo. Un jet est labélisé comme provenant d'un quark b si un quark ayant une impulsion transverse $p_T > 5$ GeV/c (après les FSR) est trouvé dans un cône $\Delta R = \sqrt{[\Delta\eta(e,b)]^2 + [\Delta\phi(e,b)]^2} < 0,2$ autour de l'axe du jet¹⁾. La mise du label pour les jets c et τ est effectuée de la même façon. Les jets restants sont labélisés comme des jets légers, sans faire de distinction entre u, d, s, g .

La multiplicité moyenne des traces dans les jets, avec au moins une trace reconstruite ayant un $p_T > 2$ GeV/c, dans un cône $\Delta R < 0,4$ autour de l'axe du jet augmente du scénario de basse à haute luminosité : pour les jets de b de 3,7 à 4,3, pour les jets de u de 3,4 à 3,8 et pour les jets de c de 3,5 à 4,0. Ceci rendra l'étiquetage plus difficile à haute luminosité. La fraction des jets labélisés contenant des traces d'électrons avec $p_T > 2$ GeV/c est donnée dans la table 5.2 pour différents types de jets et pour différents scénarii de luminosité. Les électrons formant le signal à considérer pour les analyses sur le quark b proviennent alors de :

- désintégrations semileptoniques directes ($b \rightarrow e$) et en cascade ($b \rightarrow c \rightarrow e$) de hadrons- B ; Les mesures des rapports d'embranchement [94] pour ces désintégrations sont $Br(b \rightarrow l^-) = (10,71 \pm 0,22)\%$, $Br(b \rightarrow c \rightarrow l^+) = (8,01 \pm 0,18)\%$ et $Br(b \rightarrow \bar{c} \rightarrow l^-) = (1,62^{+0,44}_{-0,36})\%$.
- $b \rightarrow \tau \rightarrow e^-$ et $b \rightarrow (J/\psi, \Upsilon') \rightarrow e^+e^-$. Les mesures des rapports d'embranchement [94] pour ces désintégrations sont $Br(b \rightarrow \tau \rightarrow e^-) = (0,419 \pm 0,055)\%$ et $Br(b \rightarrow (J/\psi, \Upsilon') \rightarrow e^+e^-) = (0,072 \pm 0,006)\%$.

Les électrons de bruit de fond proviennent des processus suivants :

¹⁾ Depuis ces études, nous utilisons la taille "standard" de 0.3.

- désintégrations Dalitz de π^0 ($\pi^0 \rightarrow e^+e^-\gamma$);
- conversions de γ se produisant dans les matériaux du détecteur interne;
- désintégrations de hadrons légers.

La source dominante d'électrons dans les jets de b est donc la désintégration semileptonique des hadrons B et D . Les désintégrations semileptoniques de hadrons D apparaissent fréquemment dans les jets de c . Les jets légers sont presque exempts d'électrons de signal.

Type de jet	hadrons- B	hadrons- D	conversions- γ et π^0 Dalitz	Autres sources
jets- b	6, 6/6, 7/6, 5	4, 2/4, 3/4, 2	1, 5/1, 6/2, 0	0, 4/0, 4/0, 4
jets- u	-/-/-	-/-/-	1, 6/1, 7/2, 4	0, 1/0, 1/0, 1
jets- c	-/-/-	4, 8/4, 7/4, 6	1, 4/1, 4/1, 4	0, 1/0, 1/0, 1

Tab. 5.2: Fraction des jets avec des électrons (en %), pour différentes origines. Les nombres sont donnés pour des événements sans empilement, à basse et haute luminosité.

Dans les jets légers les sources principales d'électrons sont ceux provenant de conversion et les désintégrations Dalitz. La présence de tout électron dans les jets à rejeter rend l'étiquetage difficile. Comme on peut le conclure de la table 5.2 le meilleur facteur de rejet peut être obtenu pour les jets légers et le moins bon pour les jets de c . Pour les événements avec empilement la fraction des jets avec des électrons de signal ne change pas, mais la fraction des jets avec des électrons de bruit de fond augmente dans le cas des événements $H \rightarrow u\bar{u}$ et $b\bar{b}$ rendant l'étiquetage moins efficace.

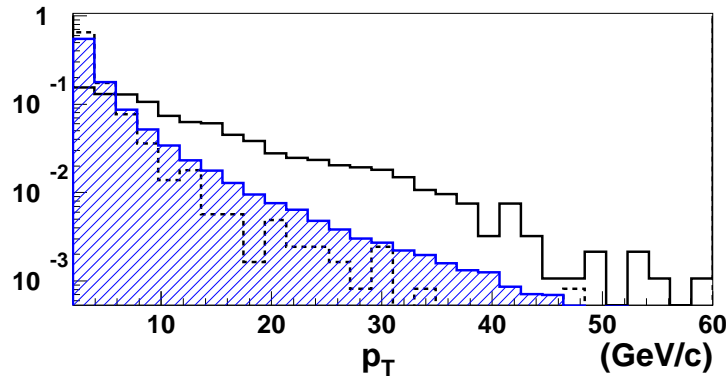


Fig. 5.1: Distributions normalisées de l'impulsion transverse p_T pour des électrons du lot $H \rightarrow \bar{b}b$ (histogramme en trait plein), de bruit de fond (histogramme en trait tireté) et pour des pions du lot de bruit de fond (histogramme hachuré), pour des données à basse luminosité.

Tout électron trouvé dans le lot de jets de bruit de fond va rendre très difficile son rejet. Dans la figure 5.1 on montre les distributions d'impulsion transverse p_T pour les électrons de signal et de bruit de fond. Les particules dans les jets reconstruits ont des impulsions transverse plutôt faibles. Ces quantités sont peu modifiées par les scénarii de luminosité, comme on peut

5. RECONSTRUCTION DES ÉLECTRONS DANS LES JETS ET ÉTIQUETAGE DES JETS PAR UN ÉLECTRON MOU

le voir d'après les informations de la table 5.3. L'empilement va cependant induire une plus grande multiplicité des traces dans les jets. Ainsi, l'utilisation de traces de mauvaises qualité ou de faux candidats va être plus importante. Les électrons seront par ailleurs encore moins isolés. Tout ceci va rendre le travail de l'algorithme d'identification des électrons plus difficile et rendre les performances moins bonnes.

	$\langle p_T \rangle$ (GeV/c)			$\langle n_{trace} \rangle$
	électrons de signal	électrons de bruit de fond	pions	
$H \rightarrow bb$	10, 8/11, 0/11, 1	4, 3/4, 4/4, 3	5, 3/4, 9/4, 8	9, 3/9, 4/10, 7
$H \rightarrow u\bar{u}$	5, 1/5, 6/5, 8	4, 8/4, 6/5, 1	6, 1/6, 2/6, 0	8, 9/9, 0/10, 5
$H \rightarrow c\bar{c}$	8, 2/8, 1/8, 0	4, 0/4, 2/3, 7	5, 2/5, 1/5, 0	9, 1/9, 2/10, 7

Tab. 5.3: Valeur moyenne de l'impulsion transverse p_T des électrons de signal et de bruit de fond, multiplicité moyenne en nombre de traces, $\langle n_{trace} \rangle$, pour différents types d'événements. Tous les nombres sont donnés sans empilement et pour les scénarii de luminosité basse et haute.

Le rapport d'embranchement inclusif $BR^{\text{all}}(b \rightarrow e)$ est défini comme la fraction de jets de b reconstruits avec un électron autour du jet. Le rapport d'embranchement exclusif de la désintégration $b \rightarrow e$, $BR^{\text{exclu}}(b \rightarrow e)$, inclue la désintégration semi-leptonique directe de hadrons B ($b \rightarrow e$) et les désintégrations en cascade de hadrons charmés ($b \rightarrow c \rightarrow e$) et est normalisé au nombre de jets reconstruits et identifiés comme des b .

5.2 Reconstruction des électrons dans les jets

5.2.1 Description de l'algorithme

L'algorithme de reconstruction des électrons dans les jets est le même que celui décrit pour la reconstruction des électrons du J/ψ . On se reportera donc au paragraphe 4.2 pour une description de celui-ci. Seules les traces ayant une impulsion transverse ayant $p_T > 2$ GeV/c et $|\eta| \leq 2,4$ sont gardées pour l'analyse. De manière à rejeter la contribution des traces mal reconstruites, les fausses traces ou les traces dues à l'empilement, les critères suivants sont appliqués :

- au moins deux points de mesure dans les pixels, dont au moins un dans la couche "b" ;
- au moins neuf points de mesure dans les détecteurs de précision (pixels + SCT) ;
- le paramètre d'impact transverse $|d_0| \leq 1$ mm ;
- le critère d'ajustement de la trace $\chi_{\text{fit}}^2 \leq 3$;
- un critère sur la position selon l'axe z du vertex primaire, $|z_0 - z_{\text{vertex}}| \sin \theta \leq 0.15$ cm, pour combattre les traces d'empilement ;

Deux autres critères permettent de réduire le nombre de traces de hadrons tout en gardant les traces dues à des électrons :

- au moins un point de mesure à haut seuil dans le TRT le long de la trajectoire ;
- au moins vingt points de mesure dans le TRT le long de la trajectoire.

De même les variables utilisées pour caractériser les électrons sont identiques à celles utilisées pour les électrons provenant du J/ψ . On se reportera à la figure 5.2 pour voir les distributions correspondantes.

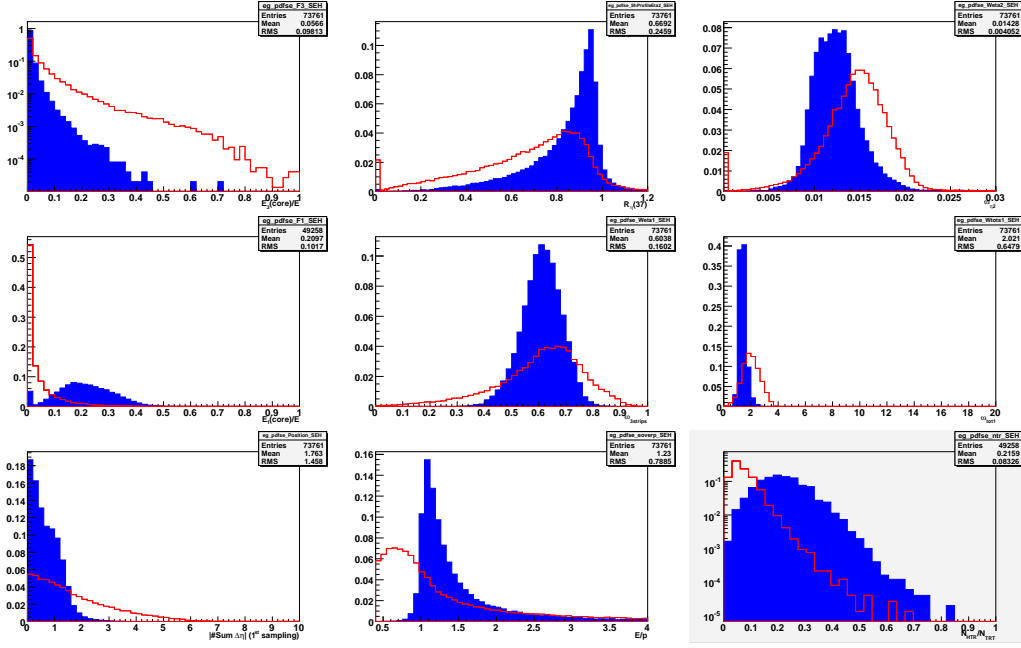


Fig. 5.2: Variables discriminantes utilisées pour l'identification des électrons dans les jets. On se reportera au paragraphe 4.2.2 pour la description de celles-ci. Les distributions sont montrées pour des électrons (histogrammes hachurés) et des pions du lot de bruit de fond (histogrammes en trait plein). Les distributions sont normalisées à l'unité.

5.2.2 Performances

L'efficacité d'identification des électrons ε_e est définie comme $\varepsilon_e = \frac{N_e^t}{N_e}$, où N_e est le nombre d'électrons, ayant une trace de bonne qualité, i.e pour laquelle la trace a été extrapolée. N_e^t est le nombre d'électrons identifiés après les coupures de la sélection. Le facteur de rejet des pions est défini comme $R_\pi = N_\pi/N_\pi^t$, où N_π est le nombre de traces de pions de bonnes qualités et N_π^t est le nombre de traces de pions de bonnes qualités identifiées comme électrons. On utilise à nouveau le rapport de vraisemblances X_{RL} déjà rencontré au paragraphe 4.2.3 et défini dans l'équation 4.1. Les facteurs de rejet des pions, dans différents lots de données, sont montrés dans la table 5.4 et dans la figure 5.3 pour différentes valeurs de l'efficacité ε_e d'identification des électrons. Par exemple, pour des événements sans empilement, pour une efficacité $\varepsilon_e = 80\%$, le facteur de rejet des pions est de 245 ± 17 dans le lot $H \rightarrow b\bar{b}$, 561 ± 35 dans le lot $H \rightarrow u\bar{u}$

5. RECONSTRUCTION DES ÉLECTRONS DANS LES JETS ET ÉTIQUETAGE DES JETS PAR UN ÉLECTRON MOU

luminosité	ε_e	R_π		
		$H \rightarrow bb$	$H \rightarrow u\bar{u}$	$H \rightarrow c\bar{c}$
	90 %	71 ± 3	129 ± 4	121 ± 6
pas d'empilement	80 %	245 ± 17	561 ± 35	508 ± 54
	70 %	514 ± 51	1439 ± 142	1040 ± 158
	90 %	39 ± 1	66 ± 1	60 ± 2
basse	80 %	180 ± 11	420 ± 21	345 ± 28
	70 %	455 ± 43	1312 ± 116	862 ± 109
	90 %	20 ± 1	30 ± 1	28 ± 1
haute	80 %	73 ± 3	133 ± 5	117 ± 5
	70 %	171 ± 10	405 ± 26	308 ± 22

Tab. 5.4: Facteurs de rejet des pions, dans différents lots de données, pour différentes valeurs de l'efficacité ε_e d'identification des électrons. Les résultats sont donnés pour différents scénarii de luminosité.

et 508 ± 54 dans le lot $H \rightarrow c\bar{c}$. Les différences de performances sont dues pour l'essentiel à l'utilisation de la distribution du paramètre d'impact transverse, qui est très différente entre les pions de jets légers et de jets de b . Les résultats sont aussi indiqués pour les scénarii de basse et haute luminosité. Le facteur de rejet des pions est environ 30% (70%) plus bas dans le scénario à basse (haute) luminosité que dans le cas où il n'y a pas d'événements d'empilement. La figure 5.4 montre la distribution du nombre de points de mesure de haut seuil dans le TRT pour des électrons de signal et pour des pions du lot de bruit de fond pour les deux scénarii de luminosité. On voit nettement que cette variable devient moins discriminante. L'effet d'empilement est moins dramatique pour les variables calorimétriques car les électrons étant situés dans des jets, ces variables sont déjà moins discriminantes que pour les électrons isolés.

5.2.3 Facteur de rejet des électrons de conversion

Une des sources principales de bruit de fond provient des électrons de conversion, de désintégration Dalitz et de pions neutres ($\pi^0 \rightarrow e^+ e^- \gamma$). Les critères de *bonne qualité* appliqués aux traces utilisées permettent le rejet des traces issues des événements d'empilement et de sélectionner en particulier celles qui ont traversé l'ensemble du trajectographe interne. La reconstruction des électrons de conversion a déjà été évoqué au paragraphe 3.2.3. Cet algorithme permet de reconstruire environ 80% des photons de conversion, ayant une impulsion transverse de l'ordre de 50 GeV et se convertissant dans un rayon plus petit que 10 cm [26]. Cependant les photons et les pions neutres dans les lots de jets ont des impulsions transverses bien plus faibles, en moyenne de 2,6 GeV pour les photons et 3,8 GeV pour les π^0 [110]. En général l'impulsion du photon n'est pas divisée de manière équitable entre les deux membres de la paire. Si l'un des partenaires est très énergétique, le second pourra ne pas avoir de trace reconstruite si son

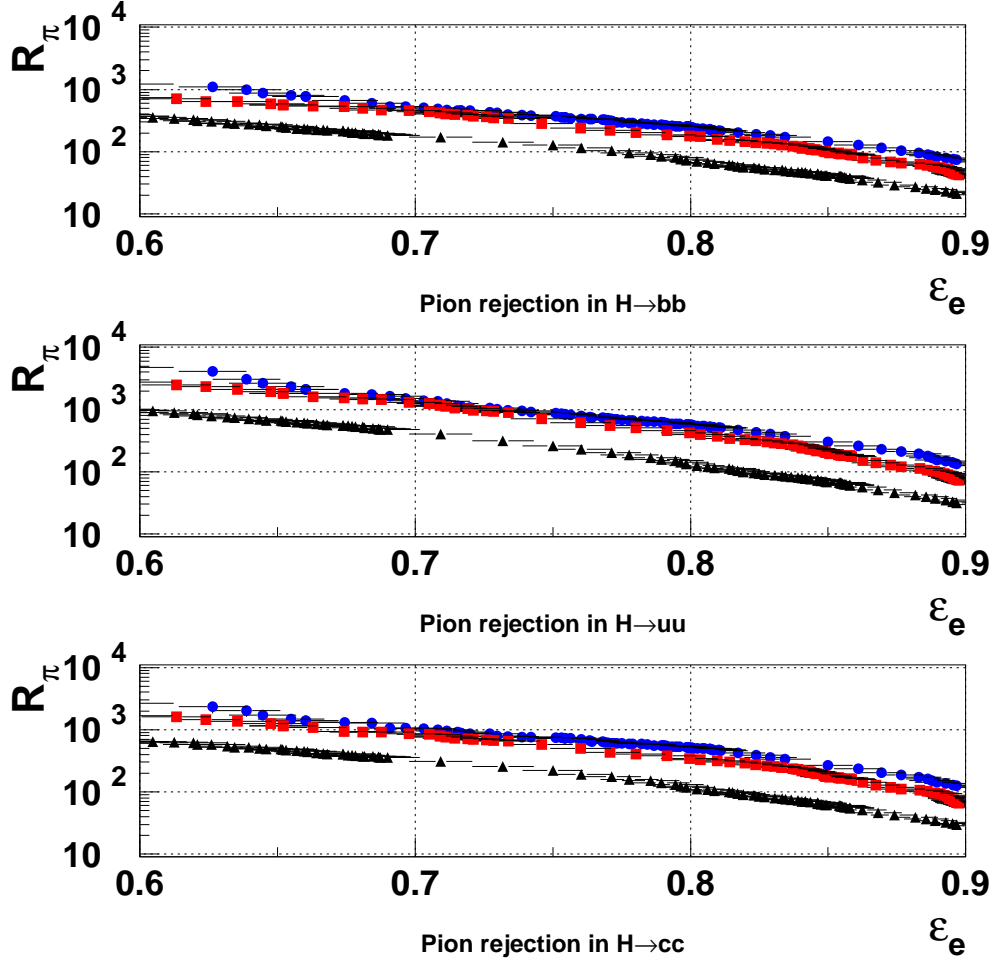


Fig. 5.3: Facteur de rejet de pions R_π en fonction de ϵ_e pour les scénarii sans événements d'empilement (cercles), à basse luminosité (carrés) et avec la haute luminosité (triangles).

impulsion transverse est inférieure à 0,5 GeV/c. De plus, la haute densité en traces dans les jets qui nous intéressent peut induire une mauvaise association dans les points de mesure formant les traces, ce qui va perturber l'algorithme de reconstruction des conversions et donner une efficacité de reconstruction plus faible que pour des photons isolés de haute énergie. Ainsi l'algorithme de reconstruction des conversions sera moins efficace que pour les photons du canal $H \rightarrow \gamma\gamma$.

Ainsi pour le lot utilisé, une efficacité de reconstruction des conversions de 50% est attendue en utilisant le seul critère sur χ_{conv}^2 pour des traces de *bonnes qualités*. On définit un facteur de rejet des conversions R_{conv} dans le lot $H \rightarrow u\bar{u}$, en utilisant l'algorithme de reconstruction des électrons dans les jets. Celui-ci peut être estimé avec ou sans l'ajout de l'algorithme de

5. RECONSTRUCTION DES ÉLECTRONS DANS LES JETS ET ÉTIQUETAGE DES JETS PAR UN ÉLECTRON MOU

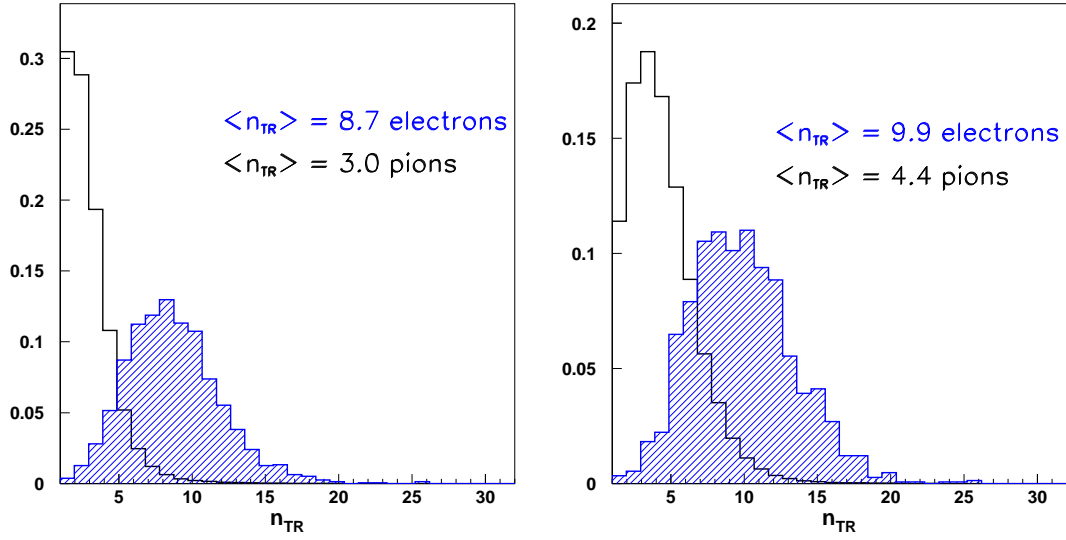


Fig. 5.4: Le nombre de points de mesure de haut seuil dans le TRT pour des électrons de signal (histogramme hachuré) et pour des pions. Les distributions sont montrées pour des données à basse luminosité (à gauche) et à haute luminosité (à droite).

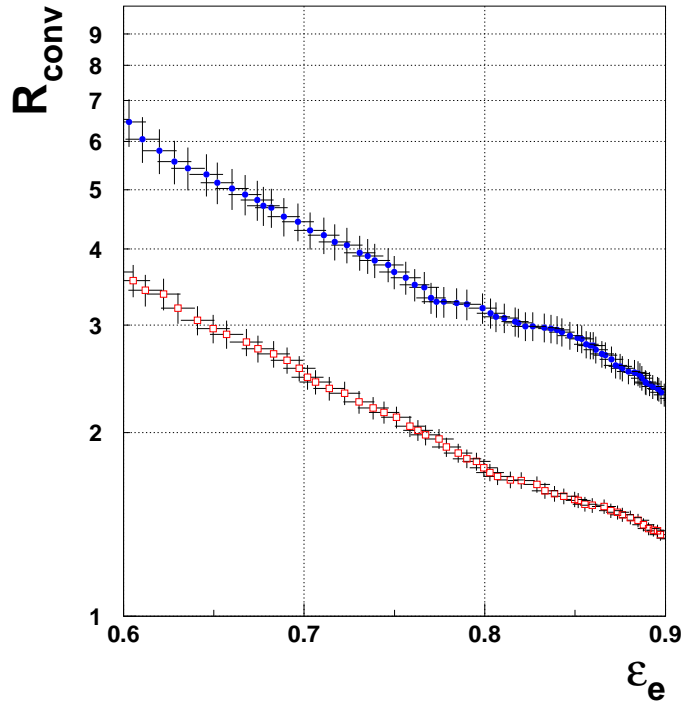


Fig. 5.5: Facteur de rejet des conversions R_{conv} dans le lot $H \rightarrow u\bar{u}$ en fonction de ϵ en utilisant (cercles) ou non (carrés) l'algorithme de reconstruction des conversions.

reconstruction des conversions. Ce facteur de rejet est montré sur la figure 5.5 en fonction

de l'efficacité d'identification des électrons de signal. Pour une efficacité $\varepsilon_e = 80\%$ le facteur de rejet des conversions est de $3,2 \pm 0,2$ en utilisant la reconstruction des conversions et de $1,8 \pm 0,1$ sans l'utiliser.

5.3 Etiquetage des jets de b par un électron

L'algorithme d'étiquetage de jets de b est basé sur la procédure de reconstruction et d'identification des électrons non isolés :

- pour chaque trace dans le jet, la valeur de la fonction X_{RL} d'identification des électrons est mesurée en utilisant le maximum de vraisemblance.
- pour chaque jet la trace ayant la plus grande valeur de X_{RL} est choisie et cette valeur donne la fonction de discrimination du jet, $D_{jet} = \max(X_{RL})$;
- pour un seuil D_{jet}^{thr} un jet ayant $D_{jet} \geq D_{jet}^{thr}$ est étiqueté comme étant un jet beau.

L'efficacité de l'algorithme d'étiquetage des jets de b est définie comme $\varepsilon_b^{alg} = \frac{N_b^t}{N_b^e}$, où N_b^t est le nombre de jets de b étiquetés et N_b^e est le nombre de jets de b labélisés, avec au moins un électron ayant $p_T > 2$ GeV/c, après les coupures de qualité des traces.. Le facteur de rejet des jets est défini par $R_{jet} = \frac{N_j}{N_j^t}$, où j est le type du jet (léger, c), N_j est le nombre de jets labélisés du type j dans le lot $H \rightarrow j\bar{j}$ avec au moins une trace de bonne qualité ayant $p_T > 2$ GeV/c à l'intérieur du cône du jet, et N_j^t est le nombre de jets labélisé du type j étiquetés comme des jets de b .

On considère la fraction des jets de b ayant un électron dans un cône $\Delta R \leq 0,4$ autour de l'axe du jet. Cette fraction est normalisée au nombre total de jets labélisés comme des b , ayant au moins une trace de *bonne qualité* d'impulsion transverse $p_T > 2$ GeV/c dans le cône du jet. Cette fraction est d'environ 13% pour le lot $H \rightarrow b\bar{b}$. L'efficacité totale de l'algorithme d'étiquetage ε_b peut alors être obtenu en multipliant l'efficacité ε_b^{alg} par cette fraction. Ainsi, pour une efficacité $\varepsilon_b^{alg} = 60\%$, l'efficacité totale de l'algorithme est de $\sim 7,8\%$. La figure 5.6 montre les facteurs de rejet obtenus pour les jets légers et c en fonction de l'efficacité ε_b^{alg} de l'algorithme d'étiquetage des b . Les performances sont données pour les différents scenarii et pour les deux méthodes d'analyse, le rapport de vraisemblances et réseau de neurones qui est décrit dans [123]. Les facteurs de rejet, pour une efficacité de 60%, sont donnés dans la table 5.5. Pour une efficacité d'étiquetage des jets de b de 60%, le facteur de rejet des jets légers, sans

	sans empilement	basse luminosité	haute luminosité
$R_{\text{légers}}$	151 ± 11	136 ± 9	104 ± 11
R_c	35 ± 2	36 ± 2	33 ± 3

Tab. 5.5: Facteurs de rejet des jets légers et c pour une efficacité de l'algorithme d'étiquetage des b de 60%.

événements d'empilement, est de 151 ± 11 . Ce facteur décroît de 10% pour le scenario à basse

5. RECONSTRUCTION DES ÉLECTRONS DANS LES JETS ET ÉTIQUETAGE DES JETS PAR UN ÉLECTRON MOU

luminosité et de $\sim 30\%$ pour la haute luminosité. Comme nous l'avons vu précédemment, ces pertes sont dues en grande partie au caractère moins discriminant du TRT. Le facteur de rejet des jets c est d'environ 35 et ne montre pas de dégradation avec les différents scénarii. Au niveau des points de fonctionnement attendus des algorithmes d'étiquetage des jets de b , à savoir vers 50-60%, le gain du réseau de neurones reste assez modeste, mais devient important à haute luminosité.

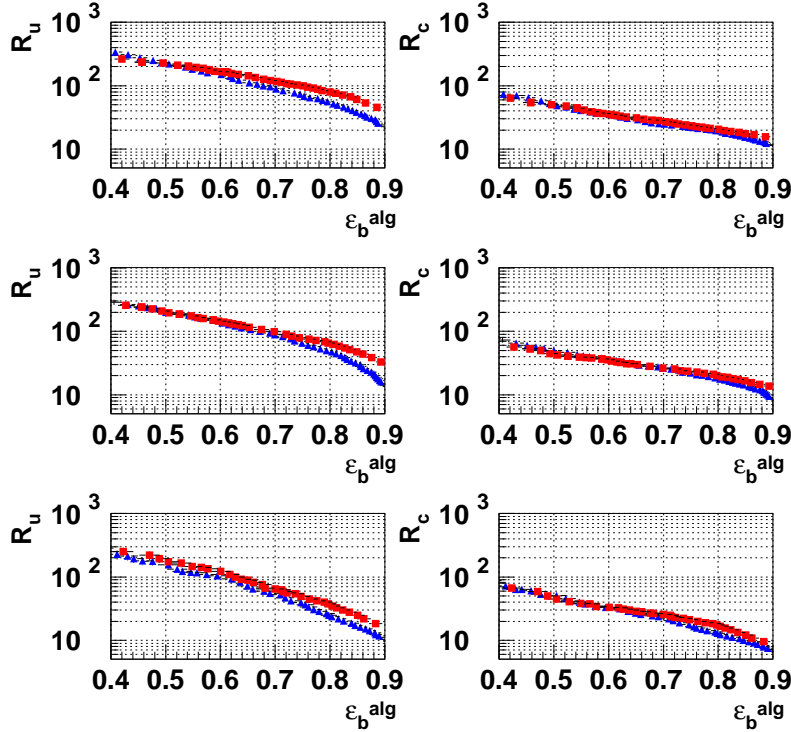


Fig. 5.6: Facteur de rejet, R_{jet} , pour les jets légers et c en fonction de l'efficacité ϵ_b^{alg} de l'algorithme d'étiquetage des b . Les résultats sont donnés pour la méthode du maximum de vraisemblance (triangles) et du réseau de neurones (carrés). Les résultats sont donnés pour les scénarii sans événements d'empilement (en haut), à basse luminosité (au milieu) et avec la haute luminosité (en bas).

La figure 5.7 montre le facteur de rejet R_{jet} pour les jets légers et c , pour une efficacité d'étiquetage des jets de b de 60%, en fonction de la pseudorapidité. La perte de performance observée dans la partie bouchon était aussi observée au niveau du facteur de rejet des pions.

La table 5.6 donne l'origine et fraction des jets étiquetés pour une efficacité de $\epsilon_b^{alg} = 60\%$. L'essentiel des jets est étiqueté par de vrais électrons, indépendamment du type de jet. Ceci indique que la procédure d'identification des électrons fonctionne avec une haute efficacité et une bonne pureté. Pour les jets de b et de c , la plupart des électrons ($\sim 90\%$ pour les b) qui permettent l'étiquetage sont des électrons de signal originaires de désintégrations de hadrons

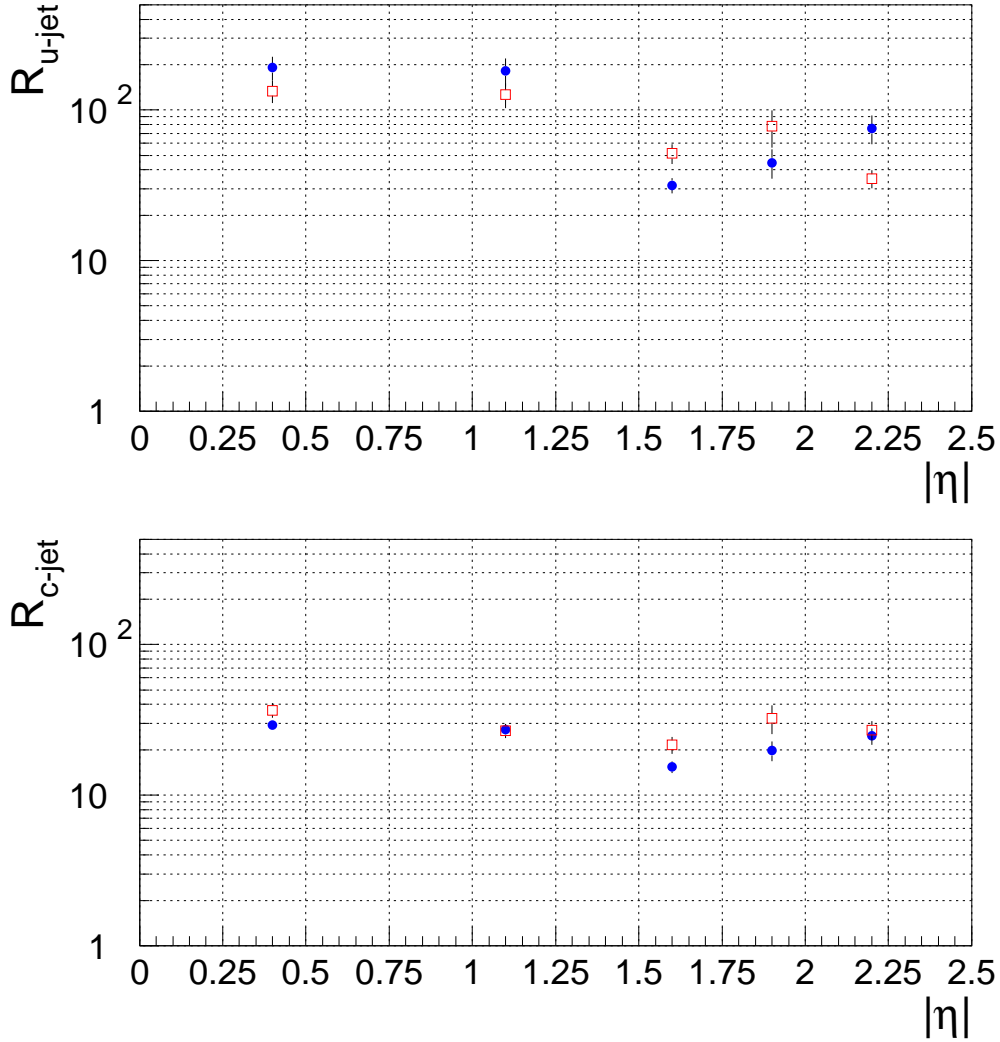


Fig. 5.7: Facteur de rejet R_{jet} , pour une efficacité d'étiquetage des jets de b de 60%, en fonction de la pseudorapidité. Les performances sont celles de l'analyse DC1 [124] (carrés) et du TDR [119] (cercles).

Type du jet	Origine et fraction des jets étiquetés [%]						
	tous e	e de b	e de d	e de *	autres e	π	autres
b	99,4	61,5	28,4	6,6	2,9	0,3	0,3
léger	69,1	0,0	0,0	65,7	3,4	16,9	14,0
c	82,4	0,0	58,2	13,5	0,7	7,7	9,9

Tab. 5.6: Origine et fraction des jets étiquetés pour $\epsilon_b^{\text{alg}} = 60\%$ à basse luminosité. Le symbole “ e de *” correspond à des électrons de conversion ou d'une désintégration Dalitz.

B ou D . Pour le lot de jets légers, ceux-ci sont étiquetés dans près de 66% des cas par des électrons de conversion. Comme montré dans [124] l'utilisation de l'algorithme de reconstruction des conversions (cf. paragraphe 5.2.3) permet d'augmenter le facteur de rejet des jets légers d'environ 30%.

Environ 30% des jets légers sont étiquetés par un pion ou un hadron. Des améliorations dans l'algorithme de reconstruction des électrons dans les jets pourraient permettre d'abaisser ce chiffre et donc d'obtenir de meilleures performances.

5.4 Travaux récents

5.4.1 Développements algorithmiques

L'essentiel des modifications apportées à l'algorithme a déjà été évoqué dans le paragraphe 4.5. L'évolution la plus spécifique à l'étude des jets est venue du trop grand nombre d'objets "Electron" créés dans les événements di-jets, l'essentiel d'entre eux étant en fait des pions. En effet, les critères d'identification sont accessibles mais ne sont pas appliqués lors de la création de l'objet "Electron". Un jet étant constitué d'un grand nombre de traces on peut alors se retrouver avec un très grand nombre d'objets "Electron". Ce nombre est d'ores et déjà réduit par les critères de *bonne qualité* des traces utilisées et par le seuil en impulsion transverse : le nombre moyen d'objet "Electron" reconstruit par jet est alors de près de 5 pour un seuil de 1 GeV/c et de 2 pour un seuil de 2 GeV/c. Ce nombre passe à 1,4 en demandant un nombre de points de mesure minimal dans le TRT. Toutefois ce nombre est encore trop élevé. Aussi a-t-il fallu appliquer des critères de pré-sélection sur les variables discriminantes elle-même : On demande que le rapport E/p soit plus grand que 0,7 : on passe alors à 0,9 "Electron" par jet. Des coupures sur $E_3(\text{core})/E < 0,5$ et $E_1(\text{core})/E > 0,03$ ont finalement été introduites : le nombre moyen d'objets "Electron" par jet est alors "seulement" de 0,4. Ces critères ont quelques conséquences sur les performances de l'algorithme : il n'est alors plus possible d'obtenir une efficacité de l'algorithme plus grande que 90%. Le point de fonctionnement pour les analyses sur le WH ou le J/ψ est à 80%. Celui-ci permet d'avoir un facteur de rejet des pions suffisant. Il n'est pas affecté par les coupures de pré-sélection.

La partie du code qui concerne l'étiquetage des jets de b proprement dit a été développé dans Athena et fait maintenant partie des algorithmes disponibles pour le groupe de performances combinés "Flavour Tagging" [125]. On utilise aussi une variable plus caractéristique des jets, à savoir la distribution d'impulsion transverse de la trace par rapport à l'axe du jet.

5.4.2 Méthodes d'analyses discriminantes

En plus des analyses utilisant des coupures sur chaque variable discriminante, de nombreuses méthodes d'analyses multidimensionnelles permettent d'utiliser la forme des distribu-

tions des variables, leurs corrélations etc. La méthode utilisée le plus dans nos études est le rapport de vraisemblance. Suivant les notes consacrées à l'identification des électrons de bas p_T et dans les jets [124] [123] [107] [126] [111], on trouvera différentes définitions de ce rapport, des études complémentaires sur le pouvoir discriminant de chaque variable, sur le (faible) niveau de corrélation entre les variables utilisées, ainsi que les améliorations possibles avec l'utilisation d'un réseau de neurones. L'une des difficultés rencontrées dans ces méthodes concerne la construction des fonctions de densité de probabilité et leur utilisation dans Athena. De manière générale il serait souhaitable que les fonctions de densité de probabilité soient construites par intervalle de pseudorapacité et d'impulsion transverse, voire pourquoi pas pour les électrons dans les jets par intervalle d'impulsion transverse par rapport à l'axe du jet. Or une telle optimisation n'a été possible jusqu'ici que pour les études DC1 J/ψ , car cela nécessite beaucoup de données pour reproduire les queues de distribution/ Une construction aisée, rapide, robuste et reproductible de ces fonctions est un problème important. Cependant, jusqu'ici les méthodes utilisées dans Atlas en général et Athena en particulier sont assez dépendantes de la ou des personnes qui effectuent l'analyse. Ainsi, dans la reconstruction des électrons il y a deux outils permettant d'effectuer un maximum de vraisemblance, l'un pour la reconstruction à haut p_T et un autre à bas p_T . Les deux outils ne sont pas codés de la même manière et les formats de données ne sont pas exactement les mêmes. De manière similaire l'outil permettant d'effectuer le rapport de vraisemblance au niveau de l'étiquetage des jets de b n'est pas exactement le même qu'au niveau de la reconstruction des électrons. Il y a donc un besoin d'un outil unifié pour construire et étudier les fonctions de densité de probabilité, étudier et comparer les différentes méthodes et enfin utiliser toutes ces méthodes dans Athena. Un tel outil est depuis près d'un an disponible.. Il s'agit de TMVA [127] qui permet d'effectuer ces analyses dans Root. Du travail est à effectuer pour comparer les résultats obtenus jusqu'ici avec ceux de TMVA, mais cet outil devrait à terme permettre, je l'espère, de résoudre une grande partie des problèmes techniques rencontrés et ainsi permettre de se concentrer sur des aspects plus difficiles.

5.4.3 Performances

Les performances de l'algorithme de reconstruction des électrons de bas p_T sont vérifiées pour chaque version d'Athena ou par chaque évolution de l'algorithme sur le lot WH . Les performances de ce lot sont donc sous surveillance étroite. La table 5.7 récapitule les performances mesurées sur le lot WH (avec une masse de Higgs de $120 \text{ GeV}/c^2$) sur les données DC1 [123], Rome [111] et avec des données plus récentes produites pour les Data Challenge 3 (CSC), reconstruites et analysées avec la version 12.0.6 (13.0.30).

On remarquera que les performances ont largement baissé pour l'analyse des données CSC [128]. A cela il y a deux raisons, toutes deux liées aux électrons de conversion, dont l'une est assez "technique". Comme indiqué dans la section 1.3.1, la description du détecteur a considérablement évolué avec les Data Challenge pour devenir de plus en plus réaliste. En

5. RECONSTRUCTION DES ÉLECTRONS DANS LES JETS ET ÉTIQUETAGE DES JETS PAR UN ÉLECTRON MOU

particulier la quantité de matière devant le calorimètre électromagnétique a augmenté, impliquant plus de conversions. Ceux-ci se retrouvant en plus grande abondance dans des jets légers vont rendre plus difficile leur rejet. Par ailleurs, dans les dernières versions d’Athena (13.0.30) la collection “Electron” ne contient plus de candidats qui sont associés à une conversion reconstruite par l’algorithme adéquat, qui sont alors mis dans la collection “Photon”. Pour éviter de retrouver ces conversions lorsqu’on ajoute les candidats reconstruits par l’algorithme “soft électrons” il a été décidé d’enlever, de toute collection, tout “soft électron” qui est associé à une conversion. Or cette association enlève près de 10% des électrons de signal dans les jets, un électron dans un jet se retrouvant suffisamment proche d’un pion pour qu’un candidat conversion soit malencontreusement reconstruit. Cette perte de signal induit elle-aussi une perte des performances comme montré dans [128]. Des travaux sont en cours à la fois pour avoir un nouvel algorithme de reconstruction des conversions et pour avoir des critères supplémentaires pour effectuer l’association électron-conversion.

Lot	Data Challenge	R_u
$H \rightarrow jj$ ($m_H=120$ GeV/ c^2)	DC1	151 ± 11
	Rome	133 ± 5
	DC3	150 ± 5
	CSC	110 ± 2
dijets	CSC	100 ± 2

Tab. 5.7: Facteurs de rejet des jets u pour une efficacité de l’algorithme d’étiquetage des b de 60% pour différents lots de données.

Les performances de l’algorithme dans le canal $t\bar{t}H$, avec une géométrie DC1, peuvent par ailleurs être trouvées dans la référence [129]. Dans le cas d’événements dijets, le facteur de rejet des jets de u est alors de près de 100. La multiplicité de traces dans les jets étant plus faible, nous aurions dû obtenir des performances meilleures. Toutefois on se rappellera que l’identification des électrons a été optimisée pour les électrons dans le lot WH et non pas de dijets, ce qui explique certainement cela.

5.5 Conclusion

La mise au point de l’algorithme d’étiquetage des jets de b par un électron mou a été effectuée en deux étapes : la reconstruction des électrons dans les jets qui s’est effectuée de paire avec celle des électrons issus du J/ψ et la mise au point de l’algorithme d’étiquetage des jets proprement dit. Cet algorithme a jusqu’ici été testé sur le lot de données simulées de production associée d’un boson de Higgs et d’un W qui, outre son intérêt pour la recherche du boson de Higgs, est le lot “standard” de suivi des performances de cet algorithme dans le groupe de performances “Flavour Tagging” [125]. Il est maintenant nécessaire d’effectuer un recadrage

de ces études vers une utilisation à la fois plus orientée vers l'analyse des premières données et vers des canaux de physique utilisés par le groupe du LPNHE. En ce sens les études les plus récentes ont permis l'étude de la reconstruction des électrons dans les événements dijets, qui sera l'une des sources principales d'électrons, bien que non isolés, et vers l'étude et l'utilisation de l'étiquetage des quarks b dans les événements issus de la production d'une paire $t\bar{t}$. En effet, le groupe du LPNHE a entamé il y a maintenant deux ans le développement d'une analyse de mesure de la masse du quark top. Le chapitre qui suit est consacré à ma participation à cet effort.

6. ETUDES SUR LE QUARK TOP

Le groupe du LPNHE a entamé voici près de deux ans un programme de recherche pour mesurer, avec les premières données du LHC, la masse du quark top dans le canal en dileptons. De nombreux stages ont été effectués sur le sujet dans le canal semileptonique [130] [131] [132], et dans le canal en dileptons [133]. Une thèse est désormais en cours sur la mesure de la masse du quark top en dileptons [134]. L'étude brièvement résumée dans ce chapitre a été réalisée dans le cadre d'un stage de niveau maîtrise d'une durée de six mois que j'ai encadré [131] et s'est largement inspirée de celle présentée dans [135].

Le paragraphe 6.1 décrit les données utilisées. Le paragraphe 6.2 montre une sélection des événements $t\bar{t}$ dans le canal semileptonique en utilisant les particules reconstruites (leptons et jets) et de l'énergie manquante. Elle montre qu'il sera possible de voir le signal $t\bar{t}$ avec les premières données d'Atlas, sans utiliser d'étiquetage des jets de b ou un étalonnage précis de l'échelle en énergie des jets. Dans le rapport de stage [131], une procédure simplifiée d'étiquetage des jets de b et d'étalonnage de l'énergie des jets avait ensuite été introduite. Elle permettait de montrer la capacité d'Atlas à reconstruire la masse du quark top avec une bonne précision, une fois la reconstruction des effets détecteurs bien compris. Les résultats ne seront toutefois pas détaillés ici. Enfin, le paragraphe 6.3 évoque brièvement des travaux en cours visant à mesurer l'efficacité de l'étiquetage des jets de b par un électron mou en utilisant ces mêmes événements $t\bar{t}$ avec les premières données.

6.1 Les données utilisées

Avec une section efficace de production de paires $t\bar{t}$ prédite à 833 pb, le LHC va produire dans sa phase de basse luminosité environ huit millions de paires $t\bar{t}$ par an. Dans le Modèle Standard le quark top se désintègre quasi exclusivement en un boson W et un quark b . Le canal "en or", dit semileptonique, est celui où un seul des W se désintègre en un électron ou un muon. Un lepton énergétique, associé à une grande énergie transverse manquante permet d'obtenir une signature très claire permettant un rejet important du bruit de fond. Le signal est constitué de près de deux cent milles événements $t\bar{t}$. La version d'Athena 11.0.42 a été utilisée pour la génération, la simulation et la reconstruction de ces événements dans le cadre du Data Challenge 3. Une masse de quark top de 175 GeV a été utilisée avec le générateur Monte Carlo MC@NLO [100] [136], version 3.1. Ce générateur calcule le processus dur de production $t\bar{t}$ à l'ordre suivant principal, c'est à dire en incluant au niveau des éléments de matrice les dia-

grammes qui produisent un gluon additionnel dans l'état final et qui utilise les densités de parton définies par CTEQ6M [137]. Les éléments de matrice calculés par MC@NLO sont processés par Jimmy, une adaptation de Herwig [138] qui inclut les événements sous-jacents, pour inclure la fragmentation et l'hadronisation. Les données correspondent au lot Atlas 005200 et contient les événements semileptoniques et dileptoniques. Le bruit de fond est constitué d'un lot de W+4 jets (lot Atlas 003017) qui représente le principal bruit de fond. Au moment de l'étude, ce bruit de fond n'avait pas encore été simulé. Un lot d'environ cent mille événements produit lors du Data Challenge "Rome" a donc été utilisé.

6.2 Reconstruction du signal $t\bar{t}$ avec les premières données

Comme les quarks top sont abondamment produits au LHC, on peut se demander s'il est possible d'observer ce signal avec un détecteur qui n'a pas encore atteint ses performances optimales. Pour ce faire nous effectuons une analyse sans utiliser l'information de l'étiquetage des jets de b et sans étalonnage de l'énergie des jets. Ce type d'analyse sera effectué dans la phase dite de compréhension du détecteur.

6.2.1 Sélection des événements

On peut définir un candidat semileptonique comme un événement avec un seul bon lepton (électron ou muon), quatre jets et de l'énergie transverse manquante.

Sélection des jets

La reconstruction des jets est faite avec un algorithme de cône de taille 0,4. Les électrons et les photons peuvent alors aussi être reconstruits comme des objets de type "Jet". La figure 6.1 à gauche montre la distribution de la distance ΔR entre un jet et l'amas électromagnétique le plus proche, en fonction du rapport $E(\text{cl})/E(\text{jet})$ de l'énergie reconstruite pour l'amas électromagnétique $E(\text{cl})$ et celle du jet $E(\text{jet})$. Si l'objet a une distance inférieure à 0,2, avec un rapport des énergies supérieur à 0,8, il est enlevé de la liste des "Jets". Il nous faut maintenant étiqueter leurs saveurs. Les algorithmes d'étiquetage des jets de b implémentés dans Athena n'étaient pas disponibles pour cette analyse¹⁾. Aussi l'étiquetage des jets de b est basé sur la distribution de la figure 6.1 à droite montrant la distance entre un "vrai" quark b , provenant de la désintégration du top et le jet le plus proche. Le jet est étiqueté comme b si la distance est inférieure à 0,2. L'efficacité de cet étiquetage est égale à 90%, donc bien supérieur à ce qui est attendu dans Atlas (de l'ordre de 50-60%). Les autres jets sont appelés jets légers. L'énergie des jets n'est pas modifiée ; il n'y a donc pas d'étalonnage de l'échelle en énergie des jets.

¹⁾ Celle-ci a en effet été effectuée au niveau CBNT/ESD alors que l'étiquetage des jets de b n'est effectuée qu'à l'étape suivante. Depuis les analyses du groupe ont été portées au niveau AOD

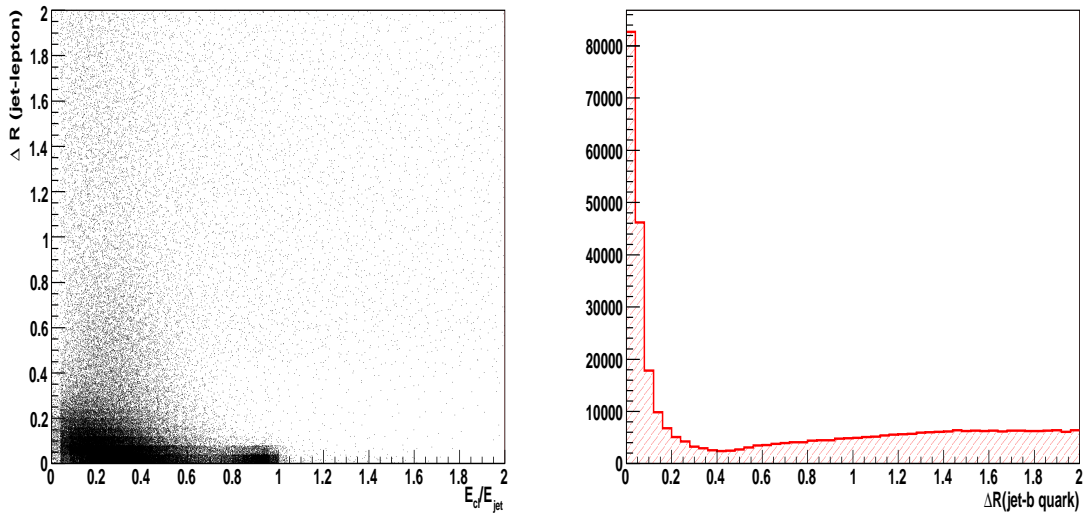


Fig. 6.1: A gauche : distance ΔR entre un jet et l’amas électromagnétique le plus proche en fonction du rapport E_{cl}/E_{jet} de l’énergie reconstruite pour l’amas électromagnétique E_{cl} et celle du jet E_{jet} . A droite : distribution de l’impulsion transverse p_T de tous les jets (à gauche) et de la distance ΔR entre un quark- b et le jet le plus proche (à droite) dans le lot de signal.

Sélection des leptons

Seuls des leptons ayant une impulsion transverse plus grande que 20 GeV/c sont considérés. D’une part à un tel seuil les algorithmes de reconstruction sont efficaces mais de plus cette valeur correspond aux seuils des menus de déclenchement sur des objets simples. Le déclenchement n’est cependant pas pris en compte explicitement dans cette analyse. On a utilisé l’algorithme de reconstruction et d’identification des électrons isolés décrit au chapitre 3. Les électrons situés dans la zone de transition du calorimètre électromagnétique, $1,37 < |\eta| < 1,52$, ne sont pas pris en compte. La figure 6.2 montre l’efficacité d’identification des électrons obtenue en fonction de leur impulsion transverse dans le lot $t\bar{t}$. Pour des électrons ayant une impulsion transverse plus grande que 20 GeV/c, celle-ci est d’environ 70%. D’autre part, on utilise l’algorithme MuonBoy [139] de reconstruction des muons dans le spectromètre. La figure 6.2 à droite montre l’efficacité de reconstruction des muons en fonction de leur impulsion transverse. Elle est de $\sim 90\%$ pour $p_T > 20$ GeV/c.

Energie transverse manquante

La figure 6.3 montre l’énergie transverse manquante E_T^{miss} reconstruite dans des événements $Z^0 \rightarrow e^+e^-$ et dans les événements “leptons plus jets” où un neutrino est présent. Une coupure au-dessus de 20 GeV permet de sélectionner les événements dans lesquels il y a des neutrinos et donc où on attend de l’énergie transverse manquante.

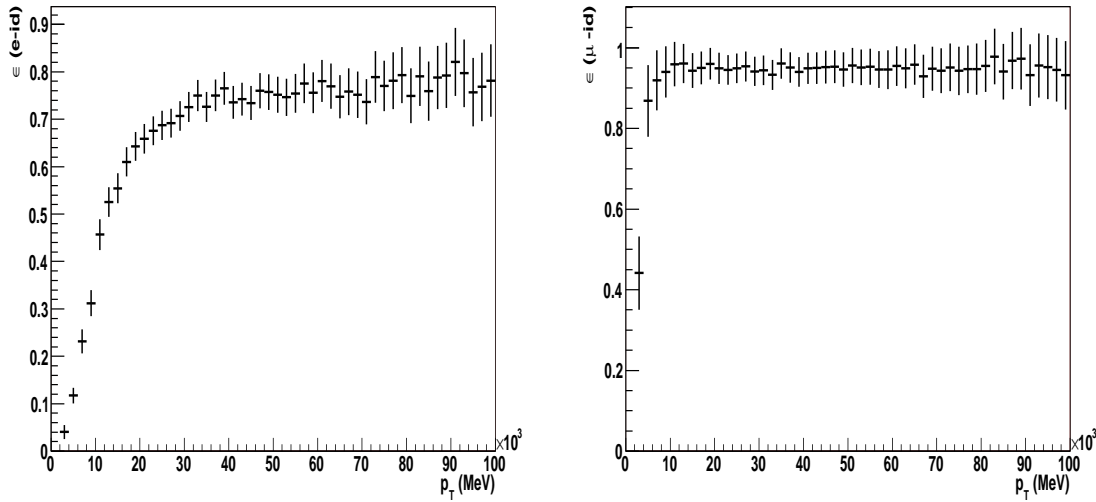


Fig. 6.2: Efficacité d'identification des électrons (à gauche) et des muons (à droite) en fonction de leur impulsion transverse pour des événements $t\bar{t}$.

Récapitulatif de la sélection

Les critères suivants sont appliqués

- l'énergie transverse manquante E_T^{miss} doit être plus grande que 20 GeV ;
- il doit y avoir au moins un lepton ayant $p_T > 20$ GeV/c et $|\eta| < 2,5$.
- il doit y avoir au moins quatre jets ayant $E_T > 20$ GeV et $|\eta| < 2,5$.

Comme expliqué dans [135], un bruit de fond important dû aux désintégrations semileptoniques des quarks b pourrait provenir du lot purement hadronique $t\bar{t}$. Pour le réduire on demande que le nombre de leptons isolés soit égal exactement à un. Le lepton est dit isolé si sa distance au jet le plus proche est plus grande que 0,2 (cf. figure 6.4). Environ 30% des événements de signal passent les critères de sélection contre 18% pour le bruit de fond W+4jets.

6.2.2 Reconstruction du top hadronique

Une fois la sélection des objets effectuée, il faut vérifier si l'événement sélectionné est compatible avec une paire $t\bar{t}$ avec une sélection des trois jets formant le top qui se désintègre de manière purement hadronique (cf. figure 6.5) et un test de compatibilité de deux de ces jets avec une origine provenant du boson W . En l'absence d'étiquetage des jets de b il existe une ambiguïté sur la combinaison de trois jets à utiliser. On considère toutes les permutations de trois jets parmi les quatre sélectionnés. Celle qui a la plus grande impulsion transverse est choisie comme étant le top "hadronique".

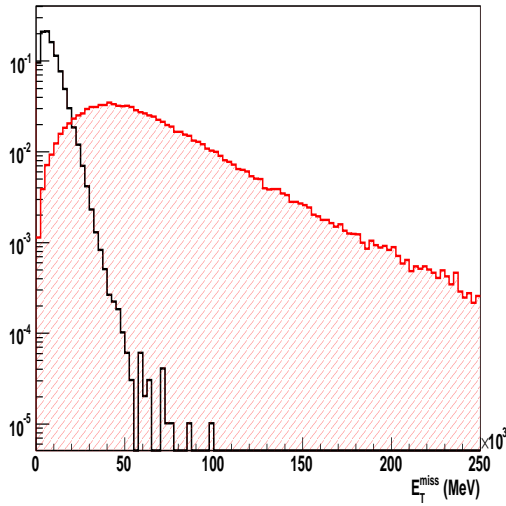


Fig. 6.3: Distribution de E_T^{miss} dans des événements $Z^0 \rightarrow e^+e^-$ et “leptons plus jets”.

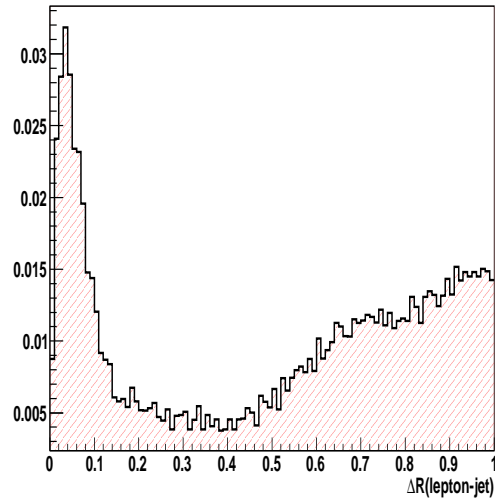


Fig. 6.4: Distance entre les leptons et le jet le plus proche.

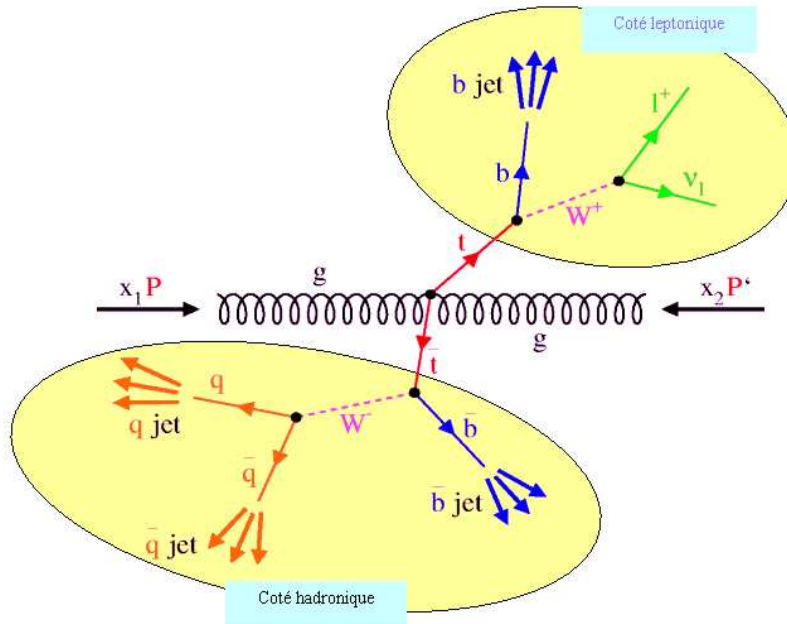


Fig. 6.5: Représentation schématique de la désintégration d'une paire $t\bar{t}$ dans le canal semileptonique.

On reconstruit aussi la masse du boson W “hadronique” originaire de la désintégration de ce quark top. Celui-ci est pris comme étant la permutation de deux jets, parmi trois, qui a la plus grande impulsion transverse. Pour estimer l’efficacité de cette sélection, nous avons comparé

l'impulsion transverse reconstruite à celle du vrai W . Si l'impulsion transverse est reconstruite à mieux que 20% nous considérons l'apariement comme réussi. L'efficacité ainsi trouvée est d'environ 40%. Des estimations basées sur la position ou la masse reconstruite donnent des résultats similaires. La figure 6.6 montre la distribution de masse invariante du système de deux jets ainsi obtenue, pour des coupures sur l'énergie transverse des jets sélectionnés de 20 et 30 GeV. La table 6.1 donne le nombre d'objets W ainsi reconstruits, le nombre d'événements dans le lot de bruit de fond et le rapport signal sur bruit. Pour le seuil à 20 GeV le pic de masse du boson W est presque invisible, masqué par les événements de bruit de fond. Le rapport signal sur bruit n'est que de 0,5. Pour un seuil à 30 GeV le pic apparaît et le rapport signal sur bruit est de 3,5. Un ajustement de la distribution est alors effectué avec une gaussienne pour le signal et un polynôme d'ordre trois pour le bruit de fond. La masse reconstruite du W est alors d'environ $73 \text{ GeV}/c^2$, en l'absence de tout étalonnage de l'énergie des jets. Avec un seuil encore plus élevé le rapport signal sur bruit augmente peu mais le nombre de W chute de plus de 40%.

coupure E_T jet	Tous les événements			$50 < M_W < 90 \text{ GeV}/c^2$		
	N_W	N_{bkg}	S/B	N_W	N_{bkg}	S/B
20	6473	13945	0,5	3365	5460	0,6
30	4112	1166	3,5	2137	475	4,5
40	2306	549	4,2	1122	216	5,2

Tab. 6.1: Nombre de W hadroniques, d'événements de bruit de fond et rapport signal sur bruit. Les nombres sont donnés pour trois seuils appliqués sur l'énergie transverse des jets.

On peut alors reconstruire la masse du quark top à partir des trois jets sélectionnés. De plus nous pouvons purifier le lot en utilisant la reconstruction du W hadronique. On définit un lot de bons W , bien reconstruits, en sélectionnant les événements dans la fenêtre de masse $50 < M_{jj} < 90 \text{ GeV}/c^2$. La reconstruction du top est alors restreinte à ce lot d'événements. La figure 6.7 montre la distribution du signal attendu, pour une luminosité intégrée de 100 pb^{-1} , c'est à dire après quelques jours de prise de données à basse luminosité. La contribution du bruit de fond le plus important, issu d'événements $W+4$ jets, reste faible. Un ajustement de la distribution est alors effectué avec une gaussienne pour le signal et un polynôme d'ordre trois pour le bruit de fond. Pour un seuil sur l'énergie transverse des jets de $30 \text{ GeV}/c$, la valeur du pic, correspondant à la masse reconstruite du quark top, est de $169,6 \pm 0,5 \text{ GeV}/c^2$, avec une largeur de $\sim 19 \text{ GeV}/c^2$. Comme pour la reconstruction de la masse du W on notera que la masse reconstruite du quark top est plus basse que celle utilisée à la génération ($175 \text{ GeV}/c^2$) mais tout de même assez proche malgré l'absence d'étalonnage de l'énergie des jets. Le nombre total d'événements top reconstruit est estimé dans la fenêtre de masse $100 < M_{top} < 220 \text{ GeV}/c^2$. Un nombre total de 1267 événements est ainsi reconstruit, avec une contribution de 144 événements de bruit de fond. Ceci correspond à une efficacité de reconstruction de l'ordre de 1% et un rapport signal sur bruit d'environ 9. La table 6.2 récapitule les résultats ainsi obtenus.

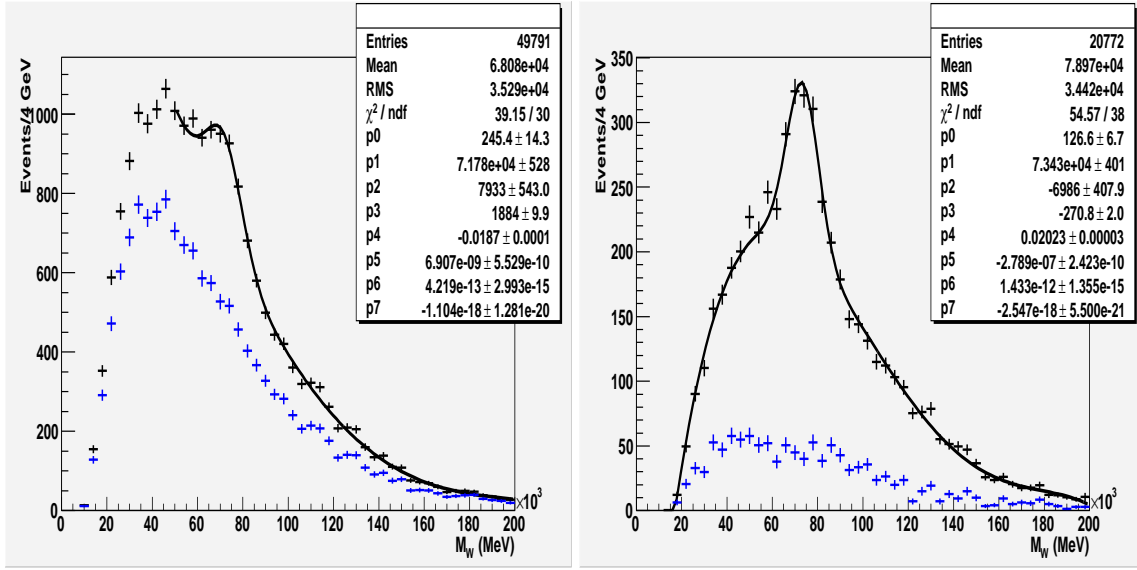


Fig. 6.6: Distributions de la masse invariante reconstruite pour la paire de jets légers candidate pour former le boson W. Les distributions sont montrées pour des coupures sur l'énergie transverse des jets de 20 GeV (à gauche) et 30 GeV (à droite). La courbe en trait tireté montre la distribution obtenue pour les événements de bruit de fond uniquement.

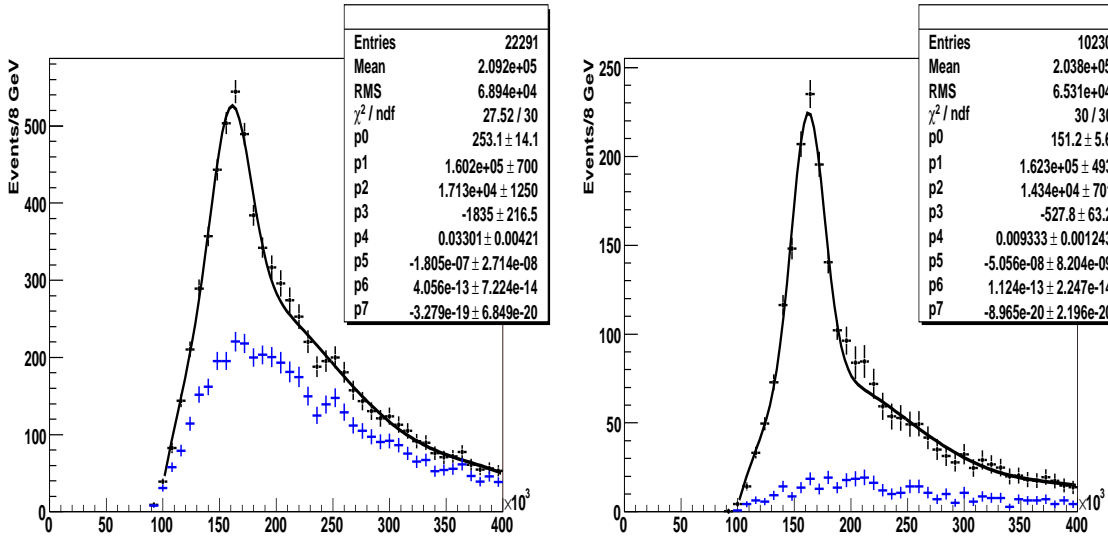


Fig. 6.7: Distributions de la masse invariante pour les trois jets formant le quark top hadronique, pour des coupures sur l'énergie transverse des jets de 20 GeV (à gauche) et 30 GeV (à droite), et des "bons" bosons W reconstruits avec une masse $50 < M_{jj} < 90 \text{ GeV}/c^2$. La courbe en trait tireté montre la distribution obtenue pour les événements de bruit de fond uniquement.

La sélection d'un lot d'événements de top quark dans le mode semileptonique peut donc être effectuée de manière relativement aisée, en quelques jours de prise de données, et ce même

coupure E_T jet	Tous les événements			$50 < M_W < 90 \text{ GeV}/c^2$		
	N_{top}	N_{bkg}	S/B	N_{top}	N_{bkg}	S/B
20	3721	5663	0,6	2108	1998	1,0
30	2097	386	5,4	1267	144	8,8
40	944	136	6,9	609	54	11,1

Tab. 6.2: Nombre total d'événements top reconstruits ($100 < M_{top} < 220 \text{ GeV}/c^2$), contribution du bruit de fond et rapport signal sur bruit pour une luminosité intégrée de 100 pb^{-1} . Les nombres sont aussi donnés en ayant sélectionné les bons W , dans la fenêtre $50 < M_W < 100 \text{ GeV}/c^2$.

sans une compréhension approfondie du détecteur. Une étude comme la mesure de la section efficace pourra alors être entreprise. La mesure de la masse du quark top nécessitera des études détaillées sur la sélection des leptons, l'étalonnage en énergie des jets et l'étiquetage des jets de b . A partir de la sélection présentée plus haut un grand nombre d'études permettra une meilleure compréhension du détecteur. Les références [135] [140] [141] explorent par exemple l'étalonnage de l'énergie des jets légers avec la reconstruction du W hadronique avec les premières données. Nous allons maintenant nous intéresser à des études plus récentes qui ont pour but d'étudier avec les premières données l'étiquetage des jets de b .

6.3 Etiquetage des jets de b dans les événements $t\bar{t}$

Le grand nombre d'événements $t\bar{t}$ qui pourra être sélectionné avec les premières données d'Atlas permettra de sélectionner un lot de jets enrichi en jets de b . Cela est vrai pour le canal semileptonique comme nous venons de le montrer mais aussi dans le canal en dileptons où les deux seuls jets sont alors des jets de b . Sur ce lot il est alors possible d'étudier les algorithmes d'étiquetage des jets de b , en particulier en ce qui nous concerne l'étiquetage par un électron mou présenté au chapitre 5.

Des études préliminaires sont en cours avec les données de la production Monte Carlo CSC. Elles montrent que pour une efficacité d'identification des jets de b de l'ordre de 60%²⁾ un pouvoir de rejet des jets de u de l'ordre de 100 est attendu.

Toutefois de nombreux travaux et développements sont encore nécessaires pour valider ce résultat. Surtout il sera nécessaire de développer une mesure de ces performances non plus à partir de l'information de la vérité Monte Carlo mais à partir de la sélection des événements top décrite dans ce chapitre. Ceci fera partie de l'évolution de mes activités de recherche telles que décrites dans le paragraphe 8.

²⁾ On rappelle que cette efficacité permet de mesurer l'efficacité de l'algorithme et correspond à une efficacité totale d'environ 8%

7. IDENTIFICATION DES ÉLECTRONS DE BASSE ÉNERGIE DANS LES DONNÉES DU TEST EN FAISCEAU COMBINÉ H8 DE 2004

Les périodes de test en faisceau sont cruciales puisqu'elles permettent de vérifier les performances des détecteurs. Elles sont aussi l'occasion de faire face à l'avance à des problèmes non attendus qui sinon apparaîtraient lors de la phase opérationnelle de l'expérience. Un test en faisceau dit combiné a permis d'étudier de mai à novembre 2004 une tranche complète du détecteur Atlas. Il s'est déroulé sur la ligne de faisceau H8 du Cern. Les sous-détecteurs comportaient une électronique de lecture finale ou quasi-finale, intégrée pour la première fois avec le système d'acquisition et de déclenchement. Ce test a permis d'étudier les performances individuelles de chaque sous-détecteur et leur reconstruction combinée. Des faisceaux de différentes particules (pions, électrons, protons, photons et muons) ont été utilisés à diverses énergies, polarités et conditions expérimentales. Il s'agissait donc d'une opportunité unique d'acquérir de l'expérience en conditions réelles et ce avant la prise de données de 2008.

De nombreuses études sont en cours sur les données de très basses énergies, allant de 1 GeV à 9 GeV. Dans ce chapitre nous allons étudier les performances de l'algorithme de reconstruction et d'identification des électrons de basses énergies qui a été présenté dans les chapitres précédents. D'autres études sont en cours dans la collaboration avec ces mêmes données portant par exemple sur les performances du TRT, la prise en compte du rayonnement de freinage dans la mesure de l'impulsion des électrons ou une reconstruction précise de l'énergie des électrons dans le calorimètre électromagnétique.

Le paragraphe 7.1 décrit le faisceau et son instrumentation et le paragraphe 7.2 la mise en place du test. Le paragraphe 7.3 décrit les données utilisées pour cette étude. Une méthode permettant d'obtenir un lot propre d'électrons et de pions a été développée en utilisant les informations des différents détecteurs présents sur la ligne de faisceau. Elle est décrite dans les paragraphes 7.4 et 7.5. Avec ces lots on peut étudier la reconstruction dans le TRT et dans le calorimètre électromagnétique. Enfin, fort de ces études, le chapitre se finira sur la mesure des performances de l'algorithme d'identification des électrons en mesurant, pour différents lots à des énergies allant de 1 à 20 GeV, l'efficacité d'identification des électrons et le facteur de rejet des pions. Ces performances seront étudiées en utilisant le TRT seul et le calorimètre électromagnétique seul. La rédaction de chapitre a largement bénéficié des références [142], [143], [144] et [31]. Ce travail sera décrit dans une note en préparation [145].

7.1 La ligne de faisceau H8 du SPS

7.1.1 Le faisceau

La ligne de faisceau H8 du Cern (cf. figure 7.1) produit des hadrons, des électrons ou des muons avec des énergies allant de 1 GeV jusqu'à 350 GeV. Le faisceau H8 est créé en extrayant vers la zone nord un faisceau de protons, d'une énergie allant jusqu'à 400 GeV, du Super Proton Synchrotron (SPS) où il est dirigé sur la cible primaire T4 faite de béryllium d'une épaisseur maximale de 300 mm. Les intensités typiques du faisceau primaire sont de quelques 10^{12} protons par bouffée. La longueur des trains de paquets est de 4,8 s et ils sont espacés de 16,8 s. Il y a quatre lignes de faisceau, deux issues de la cible T4 (H6 et H8) et deux de la cible T2 (H2 et H4). De la cible primaire T4 le faisceau secondaire a une impulsion allant de 10 GeV/c jusqu'à l'impulsion maximale du SPS de 450 GeV/c. Une deuxième cible, placée environ 130 m plus bas que la cible T4, produit un faisceau tertiaire 'pur' d'électrons (cible en plomb plus air) ou de pions (cible en polyethylene plus plomb).

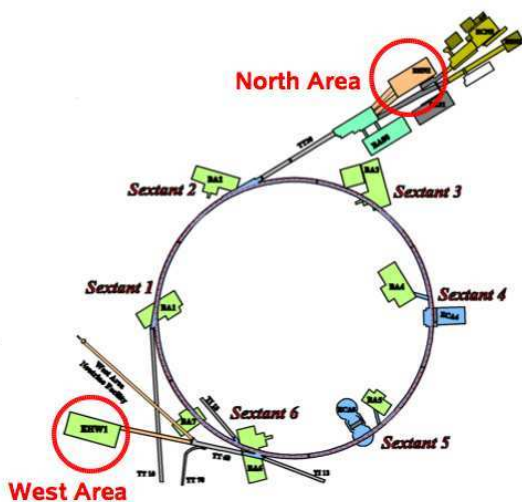


Fig. 7.1: La ligne de faisceau H8.

Cette ligne de faisceau, dite de haute énergie, n'est pas adaptée à la production de pions de basse énergie (<10 GeV). Très peu de pions pourraient en effet parcourir avant de se désintégrer les 500 mètres qui séparent la cible T4 des sous-détecteurs d'Atlas. Pour la prise de données à très basse énergie, qui va nous intéresser par la suite, le faisceau est dévié vers une autre cible (T48), à moins de 50 mètres des sous-détecteurs. Un dispositif de focalisation et de sélection en impulsion est placé après la cible secondaire. Les particules de haute énergie (<80 GeV) issues de la cible secondaire sont arrêtées par un absorbeur. après celui-ci, seuls les muons sont susceptibles de continuer leur parcours.

7.1.2 Instrumentation de la ligne de faisceau

L'instrumentation de la ligne de faisceau (cf. figure 7.2) consiste en scintillateurs, des chambres et des détecteurs Cerenkov. Ces détecteurs sont utilisés pour le déclenchement, mesurer la qualité des faisceaux et obtenir la position et le type des particules.

Il y a trois compteurs Cerenkov, d'une longueur d'un mètre. Le premier (CHRV1) est le plus éloigné. Les deux autres sont placés juste après la dernière courbe de la ligne de faisceau dédiée à la prise de données à très basse énergie. L'un est placé sur la partie droite de la ligne de faisceau de haute énergie, en cas d'impulsion plus grande que 10 GeV/c, l'autre sur le chemin dévié

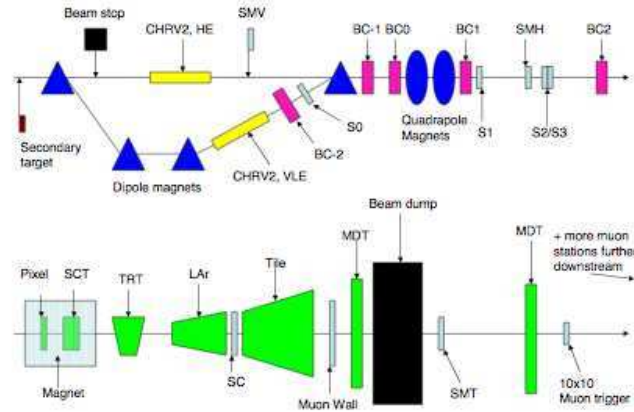


Fig. 7.2: Schéma de l'instrumentation de la ligne de faisceau H8 et des éléments de détecteur d'Atlas.

pour le mode à très basse énergie. Ces deux compteurs sont appelés CHR2. Les compteurs Cerenkov sont utilisés pour identifier le type des particules.

Cinq chambres à fils, appelées BC-2 à BC2, sont placées sur la ligne de faisceau. Ces chambres ont une surface active d'environ $11 \times 11 \text{ cm}^2$ et une résolution spatiale de l'ordre de $200 \mu\text{m}$. BC-2 est placée le long de la ligne de faisceau à basse énergie juste après le deuxième compteur Cerenkov. BC-1 est placée après le dernier aimant dipole, BC0 avant les deux aimants quadripoles, BC1 après le dernier quadripole et BC2 après les scintillateurs S2/S3.

Plusieurs scintillateurs sont placés le long de la ligne de faisceau. En suivant celle-ci, le premier est placé juste après la CHR2, sur la ligne de faisceau de haute énergie. Ce scintillateur, appelé Muon Veto (SMV), permet l'étiquetage de muons non désirés provenant en droite ligne à haute énergie. La dimension de ce scintillateur est de $81 \times 34 \times 2 \text{ cm}^3$, pour respectivement la hauteur, la largeur et l'épaisseur. Le scintillateur S0 est placé sur le chemin à très basse énergie, juste après BC-2 et CHR2. Il est utilisé pour vérifier l'intensité et la qualité du faisceau. Les dimensions de S0 sont de $10 \times 10 \times 0,6 \text{ cm}^3$. Le scintillateur S1 est placé juste après BC1 et le dernier aimant sur la ligne de faisceau. Il est utilisé pour vérifier l'intensité et la qualité du faisceau ainsi que pour le déclenchement. Les dimensions de S1 sont de $10 \times 10 \times 0,6 \text{ cm}^3$. Les scintillateurs S2 et S3 sont placés environ à mi-distance entre BC1 et BC2. Il sont aussi utilisés pour vérifier l'intensité et la qualité du faisceau ainsi que pour le déclenchement. Ces scintillateurs sont tous les deux équipés de deux photomultiplicateurs. Les dimensions de S2 et S3 sont de $5 \times 5 \times 1 \text{ cm}^3$. Le Muon Halo (SMH) est placé juste avant S2 et S3. Il est utilisé pour étiqueter les muons du halo et les autres particules. Les dimensions de SMH sont de $30 \times 32 \times 1,5 \text{ cm}^3$, avec un trou d'un diamètre de 3,4 cm centré sur le faisceau. Le Muon Tag (SMT) est placé plus loin après le premier mur d'arrêt du faisceau. Il peut être utilisé pour le déclenchement sur les muons ou pour les étiqueter. Les dimensions du SMT sont de $40 \times 40 \times 2 \text{ cm}^3$.

7.2 Mise en place du test en faisceau combiné

7.2.1 Le système de coordonnées

Pour s'assurer de la compatibilité des données du test en faisceau avec celles acceptées par l'environnement logiciel d'Atlas, la tranche de détecteurs assemblée en H8 a été mise par convention en correspondance avec le premier secteur en η et en ϕ , comme montré sur la partie hachurée (en jaune) de la figure 7.3. Dans le référentiel dit "de référence" pour tout le dispositif

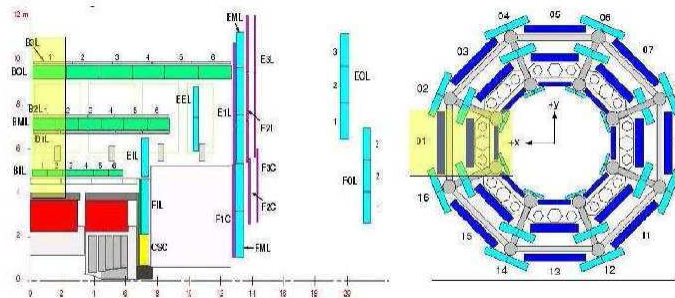


Fig. 7.3: Les référentiels en tranche et de référence du test en faisceau H8. La partie hachurée sert de correspondance entre le référentiel local du test en faisceau et celui utilisé par Atlas.

expérimental, le faisceau va à travers les différents sous-détecteurs (de gauche à droite) le long de l'axe x parallèle à l'axe du faisceau H8 ; l'axe y allant verticalement vers le ciel et l'axe z horizontal vers le sud-est. Toutes les distances sont comptées à partir du point d'interaction, très exactement à la surface de l'aimant du détecteur interne (MBPSID).

7.2.2 Les sous-détecteurs

La mise en place du test en faisceau peut être visualisée sur la figure 7.4. Les différents détecteurs sont le trajectographe interne, le calorimètre électromagnétique, le calorimètre hadronique et le spectromètre à muons.

Le trajectographe interne est composé de trois sous-systèmes. Au plus près du point d'interaction on trouve trois couches de six modules de pixels ayant une très grande granularité et quatre couches de détecteurs à bandes de silicium (SCT) faites de deux modules chacun. Ils sont placés dans l'aimant MBPSID. Celui-ci engendre un champ magnétique axial de 1,4 Tesla produit par un solénoïde. La distance entre la dernière couche de pixel et la première de SCT est la même que dans Atlas (175 mm). Finalement on trouve deux modules du détecteur à pailles avec des tubes à dérive espacés de radiateurs à effet de transition (TRT). Celui-ci est situé 55 mm après la dernière couche de SCT, soit plus loin que ce qui est prévu dans Atlas (40 mm). En particulier, dû à sa taille, le TRT est situé en dehors de l'aimant MBPSID.

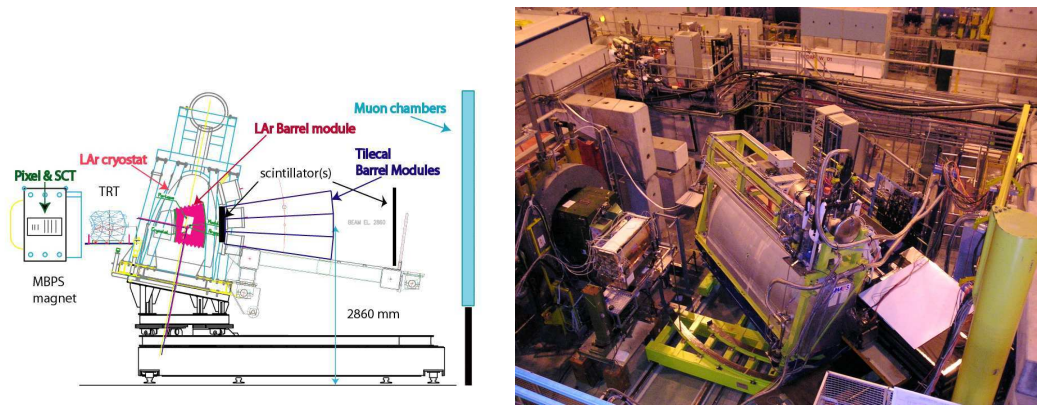


Fig. 7.4: Représentation schématique (à gauche) et photographie (à droite) de la mise en place du test en faisceau. Les faisceaux de particules rencontrent tout d'abord les modules de pixel et le SCT et continuent ensuite vers le TRT, le calorimètre électromagnétique (LAr), le calorimètre hadronique à tuiles scintillantes (Tilecal) et le spectromètre à muons. Dans le système de coordonnées utilisé, le faisceau se meut dans la direction des x positifs, l'axe y pointant vers le haut, l'axe z pointant hors du plan, qui est aussi la direction des η positifs. La table soutenant les calorimètres pouvait être translatée et tournée pour simuler une arrivée des particules du point d'interaction avec différentes valeurs de η .

Un module démonstrateur du calorimètre électromagnétique à argon liquide est placé ensuite à l'intérieur de son cryostat. Il est situé à 782 mm de la dernière couche du TRT. Pour le calorimètre hadronique à tuiles, trois modules du tonneau et trois modules du tonneau étendu ont été utilisés. Tous les modules du calorimètre hadronique, à l'exception du module tonneau le plus en dessous, sont des modules de production. Lors de la prise de données avec les modules du tonneau étendu, un petit calorimètre électromagnétique supplémentaire était placé devant, pour augmenter la couverture du détecteur. Quelques mètres en arrière de la table (environ 28 m) on trouve une partie du tonneau du spectromètre à muons, avec une chambre Monitored Drift Tube (MDT) BOS, et quatre modules de stations MDT (sept chambres) placés un peu plus loin. Enfin la partie boucon était placée encore plus loin avec une Cathode Strip Chamber (CSC), six chambres MDT placées à trois stations, un triplet de Thin Gap Chamber (TGC) et deux doublets de TGC. Il y avait aussi un système de déclenchement sur les muons avec deux scintillateurs de taille 10×10 cm, deux stations de déclenchement de Resistive Plate Counters (RPC). Enfin deux aimants étaient installés.

7.2.3 Le système de déclenchement et d'acquisition

Durant la prise de données, le plus récent démonstrateur du système d'acquisition et de déclenchement a été utilisé. Le déclenchement principal sur le faisceau est constitué par les scintillateurs S1, S2 et S3. Les autres scintillateurs SMH, SMV et SMT peuvent cependant aussi être utilisés. Les signaux des scintillateurs étaient envoyés vers un châssis dans l'une des

salles de contrôle où la logique du déclenchement était construite. L'acquisition était déclenchée lorsqu'une particule produisait un signal dans ces trois scintillateurs et que tous les détecteurs étaient en état de mesurer un nouveau signal. Une fenêtre en temps était alors ouverte et les signaux provenant de tous les sous-détecteurs étaient enregistrés sur un disque local. Finalement les fichiers de sortie étaient transférés sur le système de stockage de masse du Cern. Leur format de données est très similaire à celui attendu pour les données d'Atlas sur collisionneur.

7.2.4 Reconstruction des données

Les données sont reconstruites par le logiciel Athena, décrit au paragraphe 1.4.1. Le format des données reconstruites est le Ntuple combiné Root, évoqué au paragraphe 1.4.2. Celui-ci contient l'ensemble des informations de reconstruction des traces chargées, des cellules et des amas électromagnétiques, ainsi que les algorithmes de reconstruction et d'identification des électrons décrits dans les chapitres précédents. Nous utiliserons par la suite l'algorithme initié par une trace dans le détecteur interne et optimisé pour les électrons de basse énergie.

L'algorithme a en effet permis de reconstruire avec succès des électrons. Toutefois un changement a été nécessaire. En effet, comme vu au chapitre 4, dans sa forme utilisée pour les Data Challenge 1 ou pour les premières reconstruction des données du test en faisceau, l'amas électromagnétique créé ressemblait à un ovoïde qui suit la forme de la gerbe électromagnétique. La trace était extrapolée dans chaque compartiment et une cellule chaude était recherchée autour dans une petite fenêtre, typiquement de trois cellules. Une collection de cellules était alors prise en compte, dans chaque compartiment, autour de cette cellule chaude. Toutefois cette manière de faire rendait la reconstruction très sensible à l'alignement entre le trajectographe interne et le calorimètre électromagnétique, et ce particulièrement dans le premier compartiment où cette taille de trois cellules est très petite, problème inexistant dans les données Monte Carlo utilisées alors qui simulaient un détecteur parfait. Si le logiciel Athena a permis de compenser, après études, l'alignement et l'extrapolation des traces, j'ai aussi procédé à une modification importante de la formation de l'amas, telle que décrite dans le paragraphe 4.5.2, à savoir la formation de l'amas autour de la position extrapolée de la trace dans le deuxième compartiment. Cette modification, allant dans le sens de la simplification, était à la fois indispensable pour l'analyse du test en faisceau, pour celles des données Monte Carlo actuelles avec une géométrie réaliste et a permis aussi d'appliquer les corrections apportées à l'énergie reconstruite de la particule incidente estimées par la collaboration pour un amas rectangulaire.

7.3 Les données utilisées

Les données prises durant le test en faisceau combiné 2004 se répartissent en huit périodes qui se distinguent par la configuration des détecteurs dans la ligne de faisceau, ainsi que par la nature du faisceau lui-même. Les données utilisées pour cette étude ont été prises durant la

période de prise de données huit, pour l’essentiel à très basse énergie. Le faisceau rencontrait le calorimètre électromagnétique à un angle fixe et différentes positions en η , ici $\eta = 0,45$ et $\phi = 0$. Chaque prise de données contenait environ 10^5 particules et ont été sélectionnées suivant leur type de particule (électron ou pion) et suivant leur énergie, de 1 GeV à 9 GeV (cf. table 7.1). Un lot d’électrons et de pions pris à 20 GeV dans la prise de données à “haute énergie” est aussi utilisé. Toutes les données utilisées dans cette étude ont été prises sans champ magnétique.

E (GeV)	Faisceau nominal	Numéro	Nombre d’événements
1	e	2102101	100×10^3
2	e	2102115	80×10^3
		2102117	120×10^3
		2102118	80×10^3
3	e	2102098	200×10^3
5	e	2102097	200×10^3
9	e	2102096	100×10^3
20	e	2102397	100×10^3
20	π	2102393	14×10^3
20	π	2102394	12×10^3
20	π	2102396	51×10^3

Tab. 7.1: Données du test en faisceau utilisées pour les études de séparation électron/pion au niveau hors-ligne. L’énergie et la composition principale des faisceaux est donnée pour chaque cas.

Des données Monte Carlo ont été simulées avec un programme basé sur Geant 4 dédié à ce test en faisceau. Deux lots de données, électrons et pions, ont été simulés aux mêmes énergies et positions de collision avec des conditions de mise en place de la ligne de faisceau et du détecteur semblable à celles existant durant la prise de données. Cependant, la simulation et la reconstruction du lot final étant en cours au moment de la rédaction de ce document, ces données n’ont pu être utilisées.

7.4 Purification des faisceaux

Le faisceau de la période 8 se compose d’un mélange de pions et d’électrons venant de la ligne de basse énergie. Après l’absorbeur, seuls des muons sont susceptibles de poursuivre leur trajectoire sur la ligne principale. Le scintillateur Muon Veto aurait dû permettre d’identifier ces muons. Ce scintillateur, ainsi que S0, n’étant malheureusement pas fonctionnels durant la période 8, une contamination du faisceau par des muons dont l’énergie peut atteindre 80 GeV est possible. On attend en plus une fraction croissante de muons issus de la désintégration de pions à mesure que l’énergie incidente diminue.

7. IDENTIFICATION DES ÉLECTRONS DE BASSE ÉNERGIE DANS LES DONNÉES DU TEST EN FAISCEAU COMBINÉ H8 DE 2004

La figure 7.5 montre, pour les différents lots de données, la distribution de l'énergie reconstruite dans le calorimètre électromagnétique pour chaque amas issu de l'algorithme de reconstruction des électrons de basse énergie. Le pic vers l'énergie nominale du faisceau correspond

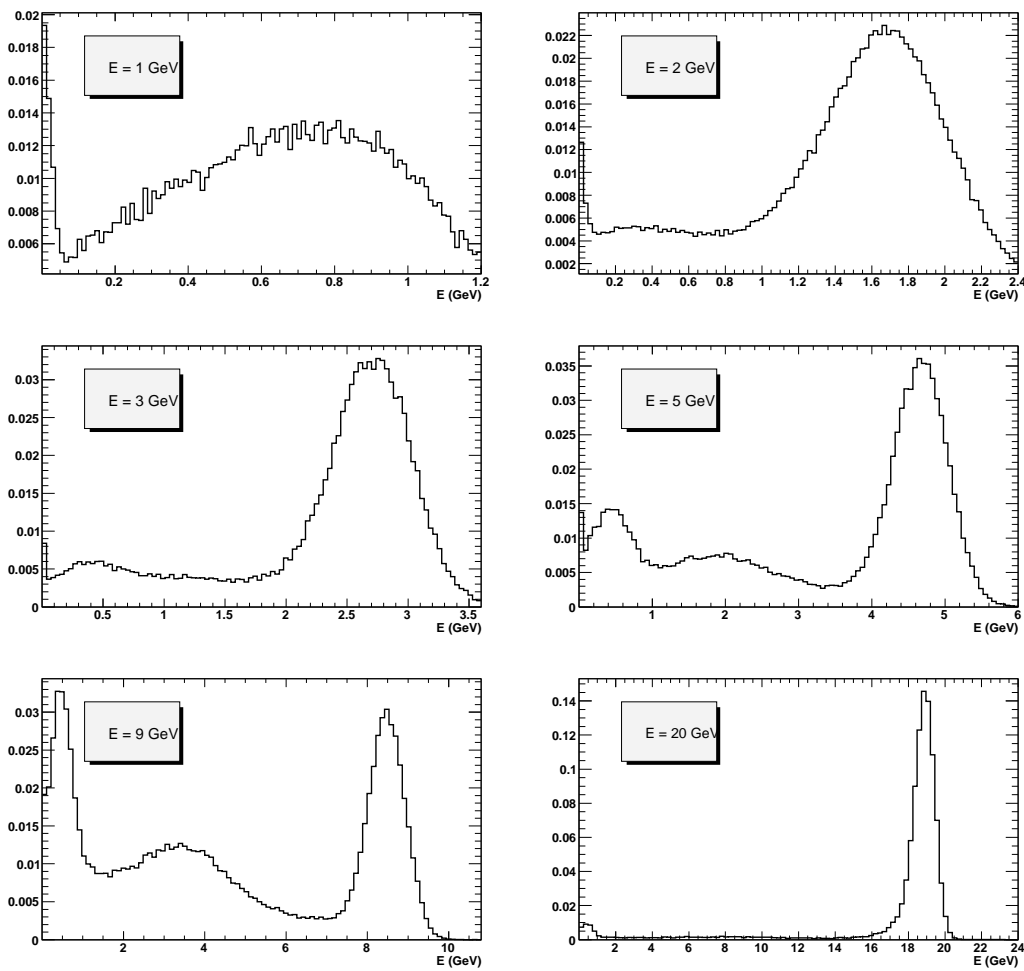


Fig. 7.5: Distribution de l'énergie totale reconstruite dans le calorimètre électromagnétique pour chaque amas issu de l'algorithme de reconstruction des électrons de basse énergie, dans les données de 1 GeV (en haut à gauche) à 20 GeV (en bas à droite), avant l'application des critères de purification du faisceau.

aux électrons. La résolution en énergie se dégradant à basse énergie, la largeur de ce pic augmente pour les petites énergies. La distribution située vers la moitié de l'énergie nominale est constituée de pions. Aux très basses énergies, de par la plus faible résolution en énergie, cette distribution est de plus en plus située sous le pic du signal. On notera en particulier qu'à 1 GeV on ne peut plus séparer les électrons des pions. Enfin, le pic à très basse énergie est dû à des muons. L'algorithme de reconstruction des électrons étant initié par une trace ce pic est de plus en plus faible à basse énergie alors que, comme nous le verrons, la contamination en muons du

faisceau augmente.

Parmi les différents détecteurs présents le long de la ligne de faisceau, les suivants ont été utilisés pour séparer ces trois populations : un compteur Cerenkov, les scintillateurs Muon Tag et Muon Halo et les chambres à faisceau.

7.4.1 Utilisation des informations du déclenchement

Le premier critère appliqué est de s'assurer que le déclenchement a bien fonctionné. Ceci permet d'éliminer une fraction de quelques pourcents des événements.

7.4.2 Contamination en muons

Comme nous l'avons vu dans les distributions de la figure 7.5 la contamination du faisceau en muons est importante et ce particulièrement à basse énergie. Ces muons peuvent provenir de la ligne de haute énergie ou former un halo autour des particules qui déclenchent le système d'acquisition. De plus une certaine fraction de muons vient de la désintégration des pions de basse énergie, principalement suivant le processus $\pi^\pm \rightarrow \mu^\pm \bar{\nu}$. Les muons de désintégration produits avant la sélection en impulsion passent par la ligne à basse énergie et ont une énergie égale à l'énergie nominale du faisceau. Des muons peuvent être également produits après la sélection en impulsion. La fraction de muons de désintégration présents dans le faisceau à basse énergie a été étudiée dans [143]. Le rapport du nombre de muons sur le nombre de pions serait d'environ 30% à 9 GeV, 40% à 5 GeV et atteindrait 70% à 3 GeV. D'autres idées pour étudier cette contamination proviennent de [144]. Quatre critères sont appliqués pour enlever la contamination due aux muons.

Utilisation des scintillateurs

La figure 7.6 montre le signal reconstruit dans les scintillateurs Muon Tag (SMT) et Muon Halo (SMH) pour un faisceau de électrons et de pions de 1 GeV et 9 GeV. Seuls les muons de haute énergie sont susceptibles de produire un signal. Dans les lots de 1 GeV et 2 GeV un halo de muons est présent et visible dans la distribution de coups dans le scintillateur Muon Halo, à de très faibles valeurs. Les objets ayant un signal dans le SMH inférieur à 900 coups ADC ou un signal dans le SMT supérieur à 450 coups ADC sont considérés comme des muons et rejetés. La contribution des muons aux lots de basse énergie reste cependant très faible, de l'ordre de 10%, sauf dans le lot de 1 GeV où elle atteint près de 40%.

Energie déposée dans le calorimètre électromagnétique

Dans un calorimètre, un muon est (presque) une particule au minimum d'ionisation, ce qui a pour conséquence que les muons ne produisent pas de gerbes électromagnétiques. Typi-

7. IDENTIFICATION DES ÉLECTRONS DE BASSE ÉNERGIE DANS LES DONNÉES DU TEST EN FAISCEAU COMBINÉ H8 DE 2004

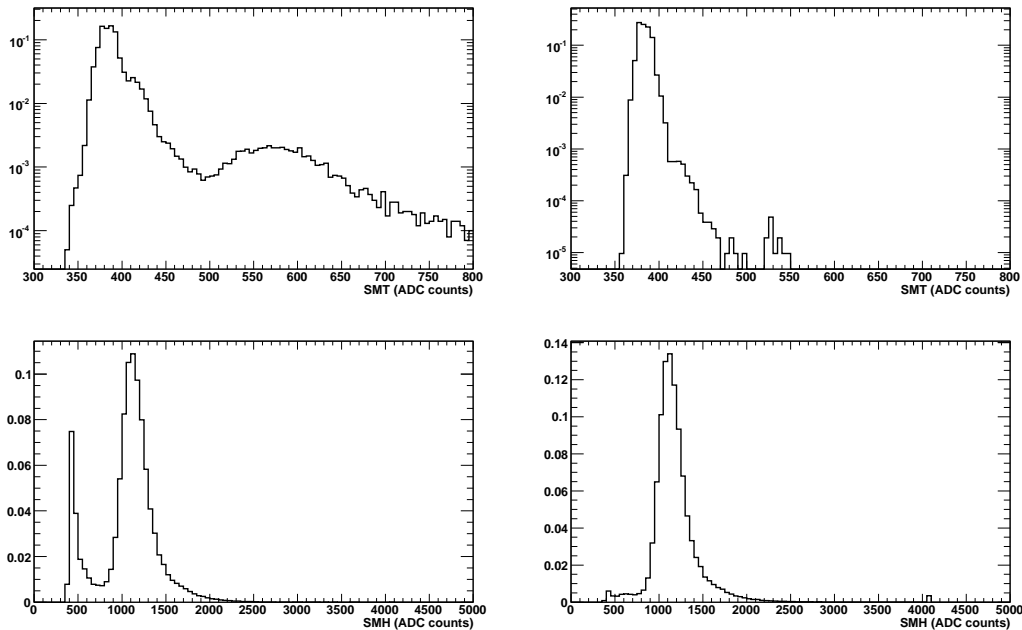


Fig. 7.6: Signal (en coups ADC) dans le scintillateur Muon Tag (en haut) et Muon Halo (en bas) pour un faisceau d'électrons de 1 GeV (à gauche) et 9 GeV (à droite).

quement, les muons traversent une ou deux cellules du second compartiment du calorimètre électromagnétique et déposent peu d'énergie. Ces objets peuvent alors être identifiés et se présentent comme des sondes très fines permettant d'explorer les propriétés du calorimètre comme par exemple son uniformité [146]. Ceci a déjà été utilisé dans les tests en faisceaux précédents ainsi qu'actuellement durant la prise de données de rayons cosmiques. Cette identification a été intégrée dans l'algorithme appelé "LArMuID" [147] qui reconstruit des amas électromagnétiques de très petites tailles. La figure 7.7 montre pour les objets ainsi reconstruits l'énergie et le nombre de cellules de l'amas dans le second compartiment du calorimètre électromagnétique. On y reconnaît les caractéristiques attendues, à savoir des objets de petites tailles déposant peu d'énergie. Ainsi, toute particule associée à un objet LArMuID ayant moins de quatre cellules dans le second compartiment du calorimètre électromagnétique et moins de deux cellules dans le premier, est considérée comme étant un muon.

Pour illustration, la figure 7.8 montre la distribution d'énergie reconstruite dans le calorimètre électromagnétique dans les données à 9 GeV après application des critères de purification avec et sans le critère "LArMuID". Son effet est clairement visible et permet de réduire la contamination en muons du lot de pions. Son effet est très dépendant de l'énergie du lot étudié. A 20 GeV, on utilise deux lots indépendants dominés l'un par des électrons, l'autre par des pions. Le critère élimine environ 10% des données "électrons" et près de 30% des données "pions". A plus basse énergie, c'est en moyenne 30% des événements qui sont ainsi rejetés,

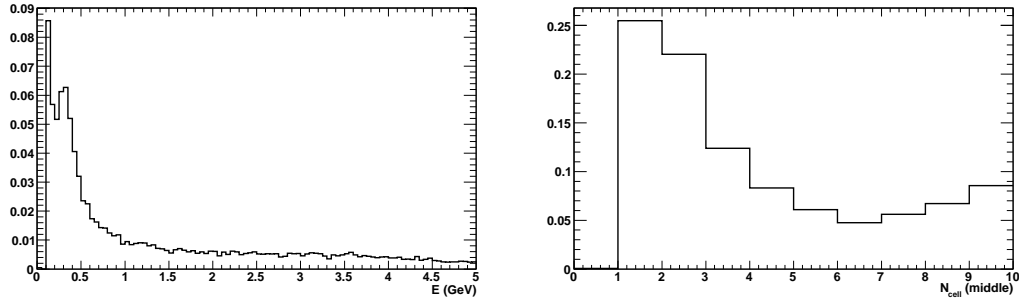


Fig. 7.7: Distribution, pour les objets reconstruits par LArMuId, de l'énergie reconstruite (à gauche) et du nombre de cellules (à droite) dans le second compartiment du calorimètre électromagnétique.

chiffre pouvant atteindre plus de 50% à 1 GeV.

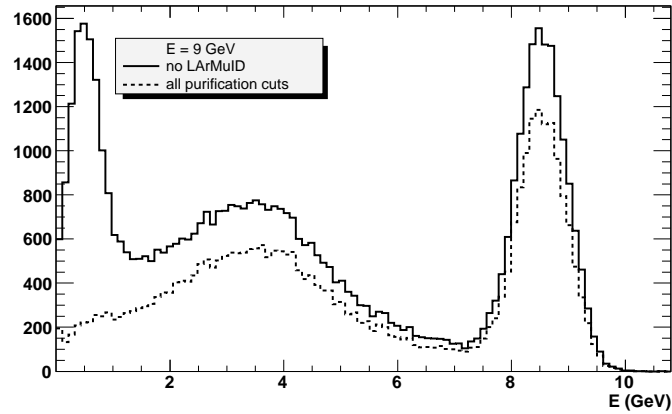


Fig. 7.8: Distribution de l'énergie totale reconstruite pour les électrons et les pions dans le calorimètre électromagnétique dans les données de 9 GeV, après l'application de tous les critères de purification du faisceau (histogrammes hachurés) et sans application du critère.

Energie d'éposée dans le dernier compartiment longitudinal du calorimètre à tuiles

La figure 7.9 montre la distribution en énergie dans la dernière couche (couche D) du calorimètre hadronique à tuiles scintillantes dans les données de 9 GeV. Etant donnée la faible énergie des pions, on s'attend à ce que ceux-ci perdent leur énergie avant d'atteindre cette couche. Pour les muons de contamination ($E \leq 80$ GeV), au contraire, la probabilité d'être arrêtés avant la couche D est beaucoup plus faible, même aux très basses énergies (jusqu'à 2,5 GeV typiquement). Les muons peuvent être rejetés en demandant que l'énergie reconstruite soit inférieure à 150 MeV [143]. Typiquement ce critère rejette environ 10% des événements.

7. IDENTIFICATION DES ÉLECTRONS DE BASSE ÉNERGIE DANS LES DONNÉES DU TEST EN FAISCEAU COMBINÉ H8 DE 2004

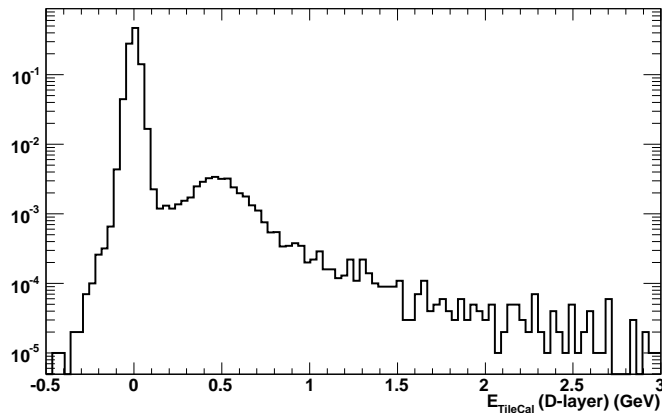


Fig. 7.9: Distribution en énergie dans la couche D du calorimètre à tuiles dans le lot à 9 GeV. Le signal entre 150 MeV et 1 GeV correspond aux muons contaminant le faisceau. Les électrons et les pions n'atteignant pas cette couche, ils contribuent au pic de bruit proche de 0.

Phase du signal dans les cellules du calorimètre à tuiles

Les muons de contamination provenant de la ligne à haute énergie n'ont a priori aucune raison d'être synchronisés avec les pions de basse énergie [143]. Lorsqu'un pion de basse énergie déclenche le système d'acquisition, il peut arriver qu'un muon venant de la ligne à haute énergie passe au même moment, i.e pendant l'ouverture de la fenêtre d'acquisition. La différence de temps entre le déclenchement de l'acquisition et la phase dans le calorimètre hadronique à tuiles¹⁾ est constante pour la particule ayant déclenché l'acquisition. Par contre, cette différence

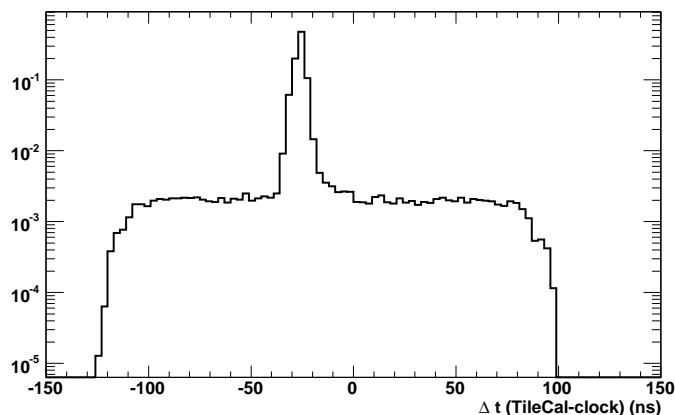


Fig. 7.10: Différence de temps entre l'arrivée du signal dans le calorimètre à tuiles et le déclenchement de l'acquisition dans un faisceau d'électrons de 9 GeV.

¹⁾ Dans une cellule du calorimètre à tuiles, la phase correspond à l'instant où le signal est maximum

est quelconque pour un muon de contamination, passant fortuitement pendant l'ouverture de la fenêtre d'acquisition. La figure 7.10 montre la distribution de la différence de temps entre le déclenchement de l'acquisition (horloge) et la phase du signal dans toutes les cellules du calorimètre à tuiles touchées dans un faisceau de 9 GeV. Elle indique la présence de particules hors temps dans ce faisceau. Par la suite, un événement sera dit hors temps si au moins une cellule du calorimètre a enregistré un signal pour lequel la différence de temps ($t_{\text{TileCal}} - t_{\text{déclenchement}}$) se situe hors de la fenêtre [-40 ns, -20 ns]. Typiquement ce critère rejette environ 10% des événements.

7.4.3 Utilisation des chambres `a faisceaux

Nous allons nous intéresser en particulier à la chambre BC-2 qui se trouve sur la ligne à basse énergie. Aucun signal n'y est donc attendu dans le cas des muons en temps. Dans la figure 7.11 le pic proche de 0 dans la distribution de la coordonnée X correspond au cas où la chambre n'a enregistré aucun signal, soit en raison de son inefficacité, soit parce qu'aucune particule ne l'a traversée.

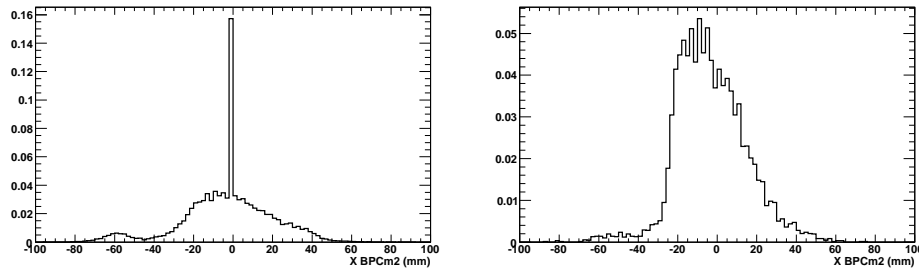


Fig. 7.11: Distribution de la position X dans la chambre BC-2 dans un faisceau d'électrons de 9 GeV. La distribution est montrée avant l'application des critères de sélection (à gauche) et après de tous les critères à l'exception du critère sur la dispersion spatiale du faisceau.

Environ 15% des données à 9 GeV qui passent dans BC-2 ne sont pas détectés par cette dernière. Ceci donne une estimation de l'efficacité de celle-ci. L'efficacité est supérieure pour les électrons que pour les pions de 10 à 20% suivant l'énergie du faisceau. Une raison possible est que les pions étant plus proches du minimum d'ionisation leur efficacité soit effectivement plus faible. Cela pourrait de plus expliquer la perte de plus de la moitié des événements dans le faisceau à 1 GeV.

La figure 7.11 représente la distribution en position dans le plan horizontal, pour la chambre BC-2, toutes les coupures précédemment décrites pour la sélection ayant été appliquées. Une dernière coupure est envisagée afin de limiter la dispersion du faisceau selon l'axe horizontal (selon l'axe vertical le faisceau est moins divergent et aucune coupure n'est nécessaire). L'extension spatiale du faisceau est limitée à la région [-40 mm, 20 mm] dans BC-2.

7. IDENTIFICATION DES ÉLECTRONS DE BASSE ÉNERGIE DANS LES DONNÉES DU TEST EN FAISCEAU COMBINÉ H8 DE 2004

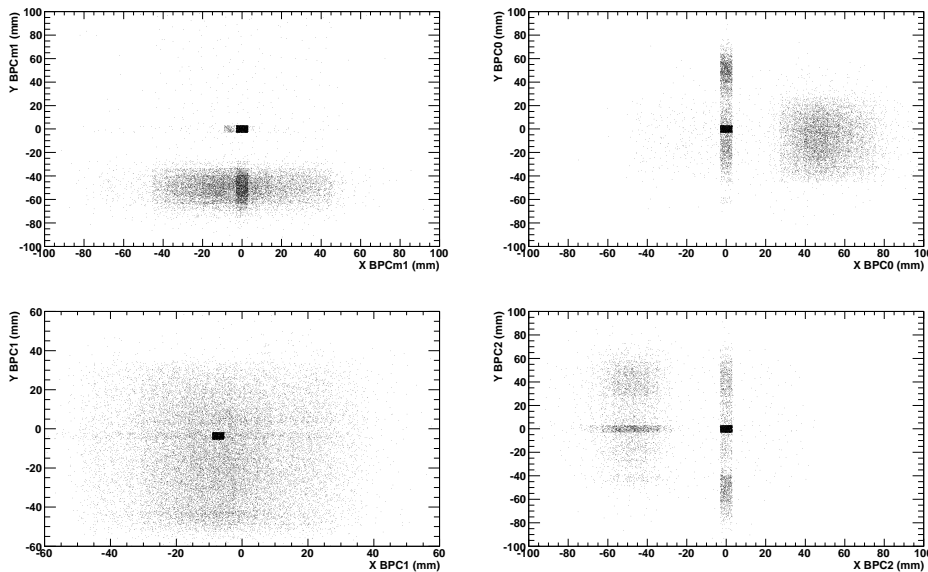


Fig. 7.12: Distributions des positions Y vs X dans les chambres à faisceau BC-1 à BC2 dans le faisceau de 9 GeV.

La figure 7.12 montre la distribution de position (Y vs X) des points de mesure dans les différentes chambres à faisceau. Deux populations sont visibles : l'une est située près de l'axe du faisceau, l'autre est répartie dans les chambres. Il pourrait sembler judicieux là aussi d'exiger la présence d'un signal dans chacune des chambres et une dispersion limitée. Cependant les pertes en nombre d'événements seraient alors énormes (plus des trois quarts) et il ne resterait plus suffisamment de particules pour effectuer les études sur la séparation électrons/pions.

Le fait de demander un signal dans BC-2 est donc la seule garantie qu'un électron ou un pion soient effectivement passés par la ligne à basse énergie. Pour le lot de données à 20 GeV, la chambre BC-2 n'étant pas dans la ligne de faisceau à haute énergie, ces critères ne sont pas appliqués.

7.4.4 Récapitulatif des critères de sélection

Le tableau 7.2 rappelle les critères appliqués pour la purification des faisceaux et pour chacun indique le nombre et la fraction d'objets ainsi sélectionnés. D'une manière générale, les mêmes critères sont appliqués quelque soit l'énergie. Dans la série des quatre critères pour éliminer les muons, le critère le plus important est celui recherchant un "petit amas électromagnétique", correspondant à l'utilisation de l'algorithme LArMuID qui permet d'éliminer une grande partie du pic observé à basse énergie dans la figure 7.5. Au total c'est près de 40% des objets initiaux qui sont identifiés à des muons, cette fraction étant plus grande à basse énergie là où la contamination attendue en muons est la plus importante. Les critères sur

les chambres à faisceau permettent d'éliminer chacun environ 15% d'objets mal reconstruits. Au final, l'ensemble de ces critères permet de sélectionner environ la moitié des objets initiaux. Cette proportion est plus grande à 20 GeV et beaucoup plus faible à 1 GeV.

critères	E_{faisceau} (GeV)						
	1	2	3	5	9	20-e	20- π
nombre d'événements	100000	430622	100000	198807	103995	60000	41300
Fraction (%) d'objets sélectionné							
déclenchement	87,8	94,4	95,3	96,0	96,4	97,2	97,2
SMT	71,2	82,9	95,3	87,9	94,2	96,6	93,5
LArMuID	43,6	77,4	76,5	69,5	59,9	89,3	71,2
$E_D(\text{TileCal})$	78,9	92,3	93,8	93,8	92,3	95,8	88,1
$t_{\text{TileCal}} - t_{\text{déclenchement}}$	78,9	90,4	90,3	90,8	85,0	95,3	83,4
Muons	34,2	65,6	72,5	60,2	52,4	87,8	61,2
BC-2 \neq 0	56,6	79,1	83,5	84,5	84,5	-	-
$-40 < BC - 2_x < 20$ mm	83,7	89,4	90,7	91,2	88,3	-	-
Tous	27,3	56,4	63,8	53,3	46,4	87,8	61,2

Tab. 7.2: Liste des critères utilisés pour purifier les faisceaux de 1 à 20 GeV, le nombre d'événements et la fraction d'objets sélectionnés par chaque critère de manière indépendante des autres.

7.5 Sélection des lots d'électrons et de pions

Pour estimer les performances de l'algorithme d'identification des électrons il reste à séparer les événements "purifiés" en un lot d'électrons et un lot de pions.

7.5.1 Sélection d'événements simples

L'analyse est effectuée sur des objets simples, c'est à dire un seul amas électromagnétique et une seule trace associée. En pratique cela revient à demander un seul objet de type "Electron de bas p_T ". De plus, des critères sur la qualité des traces sont requis :

- un seul segment dans le silicium (SCT+pixel) dans l'événement ;
- un total de 2 ou 3 points de mesure dans les pixels dans le segment de silicium ;
- au moins 4 points de mesure dans le SCT dans le segment de silicium ;
- la position y_0 dans le silicium doit être typiquement $-20 < y_0 < 5$ mm (les seuils dépendent de l'énergie).
- un seul segment de TRT dans l'événement ;
- au moins 30 points de mesure dans le TRT.

La figure 7.13 montre les distributions de certaines de ces variables. Ces critères éliminent environ un tiers des événements à 9 GeV et 20 GeV mais près de 70% à 2 GeV et 90% à

7. IDENTIFICATION DES ÉLECTRONS DE BASSE ÉNERGIE DANS LES DONNÉES DU TEST EN FAISCEAU COMBINÉ H8 DE 2004

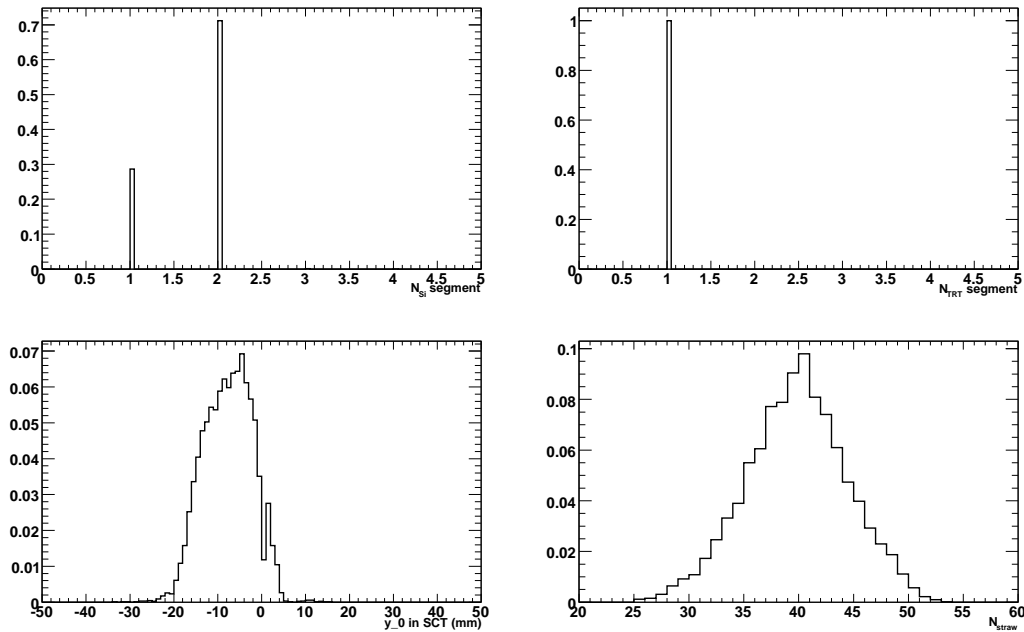


Fig. 7.13: Distributions des variables décrivant la qualité des traces dans le détecteur interne comme le nombre de segments dans le silicium (en haut à gauche), le nombre de segments dans le TRT (en haut à droite), la position y_0 dans le silicium (en bas à gauche) et le nombre de points de mesure dans le TRT (en bas à droite).

1 GeV, pour l'essentiel dû à la demande de l'existence d'une trace. Les critères de qualité peuvent sembler sévères puisque presque la moitié des traces n'y survivent pas. Toutefois elles représentent l'unique assurance que les études sur le TRT seront effectuées sur de bonnes traces.

7.5.2 Utilisation des compteurs Cerenkov

Les compteurs Cerenkov peuvent être utilisés pour sélectionner les lots d'électrons et de pions. La figure 7.14 montre le signal dans le compteur CHRV2 dans un faisceau d'électrons de 9 GeV. Les électrons doivent avoir au moins 700 coups ADC dans le CHRV2. Dans le lot de données à 20 GeV ce seuil est augmenté à 800 coups ADC. Les pions donnent dans le compteur CHRV2 un signal autour de 550 coups ADC correspondant au piédestal du détecteur. Le lot de pions est sélectionné en demandant moins de 600 coups ADC dans CHRV2.

La figure 7.15 montre la distribution de la fraction de points de mesure de haut seuil dans le TRT pour les lots d'électrons et de pions sélectionnés avec le Cerenkov dans les données à 3 et 9 GeV. La distribution montre une queue importante pour les pions sous la distribution des électrons. L'essentiel de ces "pions" ont de plus une énergie totale proche de l'énergie nominale du faisceau, ainsi que des répartitions des dépôts dans les différents compartiments du calorimètre électromagnétique qui sont typiques des électrons. On peut faire le même constat

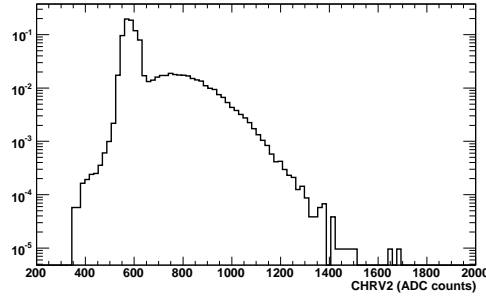


Fig. 7.14: Signal dans le compteur Cerenkov CHR2 dans le faisceau d'électrons de 9 GeV.

pour toutes les données à basse énergie, mais à 20 GeV les lots semblent déjà bien définis.

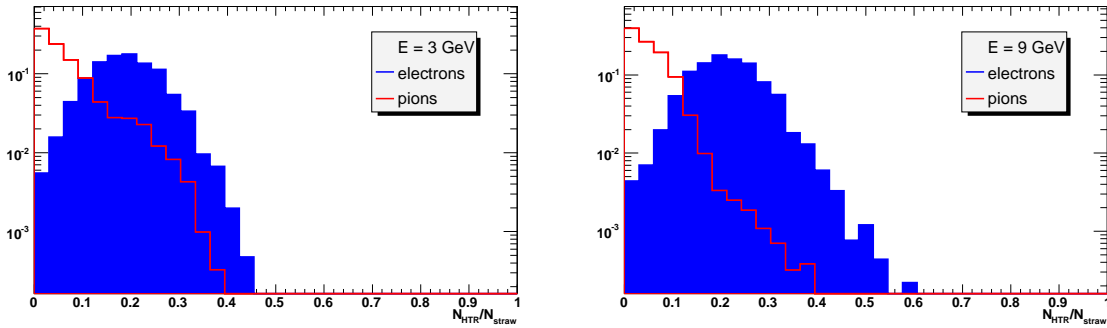


Fig. 7.15: Distributions de la fraction de points de mesure de haut seuil dans le TRT pour les lots d'électrons et de pions sélectionnés. Les distributions sont montrées avant (à gauche) et après (à droite) l'utilisation des critères sur le calorimètre électromagnétique. Les distributions sont montrées pour des pions et pour des électrons (histogramme hachuré), dans les données à 3 GeV (à gauche) et à 9 GeV (à droite).

La figure 7.16 montre les distributions du rapport E_1/E_{run} en fonction de $\sum_i E_i/E_{\text{run}}$, où E_i est l'énergie reconstruite dans le compartiment i du calorimètre électromagnétique et E_{run} l'énergie nominale du faisceau du lot étudié, dans les données de 1 GeV (en haut à gauche) à 20 GeV (en bas à droite). Les distributions sont montrées après application du critère sur le scintillateur Cerenkov pour sélectionner les électrons (étoiles) et les pions (cercles). On y reconnaît les électrons en haut à droite et les pions en bas à gauche. La contamination entre les deux populations est là aussi clairement visible, particulièrement à basse énergie.

Pour finaliser la sélection des lots il est donc le plus souvent nécessaire d'utiliser en plus les informations du TRT et des calorimètres. Ceci ne pose pas de problèmes lorsqu'on veut étudier de manière indépendante les performances de l'identification du TRT ou du calorimètre électromagnétique, mais est problématique pour étudier les performances combinées.

7. IDENTIFICATION DES ÉLECTRONS DE BASSE ÉNERGIE DANS LES DONNÉES DU TEST EN FAISCEAU COMBINÉ H8 DE 2004

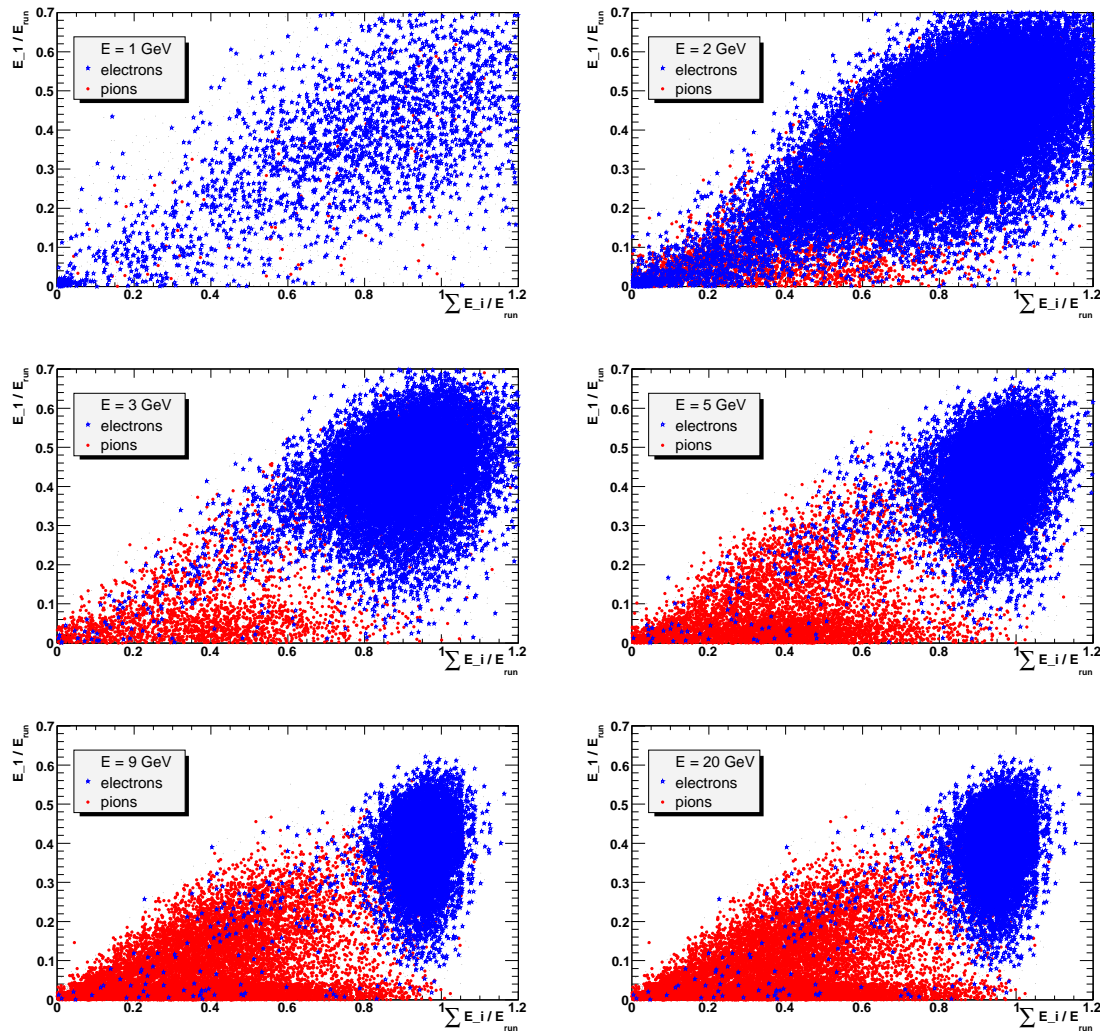


Fig. 7.16: Distribution du rapport E_1/E_{run} en fonction de $\sum_i E_i/E_{run}$, où E_i est l'énergie reconstruite dans le compartiment i du calorimètre électromagnétique et E_{run} l'énergie nominale du faisceau du lot étudié, dans les données de 1 GeV (en haut à gauche) à 20 GeV (en bas à droite). Les distributions sont montrées après application du critère sur les scintillateurs Cerenkov pour sélectionner les électrons (étoiles) et les pions (cercles).

Utilisation du TRT

Pour étudier les performances de l'algorithme d'identification des électrons en utilisant uniquement les informations du calorimètre électromagnétique on définit les lots d'électrons et de pions non seulement avec les informations des compteurs Cerenkov mais aussi avec le TRT, comme par exemple dans la référence [31] qui étudie les performances de l'identification des électrons au niveau du système de déclenchement. Un critère simple, noté cut_{TRT} , consiste à sélectionner les électrons comme le lot qui a une fraction de points de mesure de haut seuil dans

le TRT supérieure à 20% et les pions comme ceux qui ont une fraction inférieure à 10%.

Utilisation du calorimètre électromagnétique

Pour étudier les performances de l'algorithme d'identification des électrons en utilisant uniquement les informations du TRT nous pouvons définir les lots d'électrons et de pions avec les compteurs Cerenkov et le calorimètre électromagnétique. Ce type de sélection se trouve par exemple dans les références [148] et [149] qui étudient exclusivement les performances du détecteur interne. Le critère adopté, noté *cutLAR*, consiste à couper dans les distributions des rapports E_1/E_{run} et $\sum_i E_i/E_{\text{run}}$, où E_i est l'énergie reconstruite dans le compartiment i du calorimètre électromagnétique et E_{run} l'énergie nominale du faisceau du lot étudié. Les seuils appliqués dépendent de l'énergie et sont décrits dans la table 7.3. La figure 7.17 montre les

E_{faisceau} (GeV)	E_1/E_{run}		$\sum E_i/E_{\text{run}}$	
	e	π	e	π
1	$> 0,2$	$< 0,1$	$> 0,7$	$< 0,4$
2	$> 0,2$	$> 0,1$	$> 0,7$	$< 0,4$
3	$> 0,2$	$< 0,1$	$> 0,7$	$< 0,7$
5	$> 0,2$	$< 0,1$	$> 0,8$	$< 0,7$
9	$> 0,2$	$< 0,1$	$> 0,85$	$< 0,7$
20	$> 0,2$	$< 0,1$	$> 0,9$	$< 0,8$

Tab. 7.3: Liste des critères utilisés pour purifier les faisceaux et sélectionner les lots d'électrons et de pions en utilisant les informations du calorimètre électromagnétique (*cutLAR*).

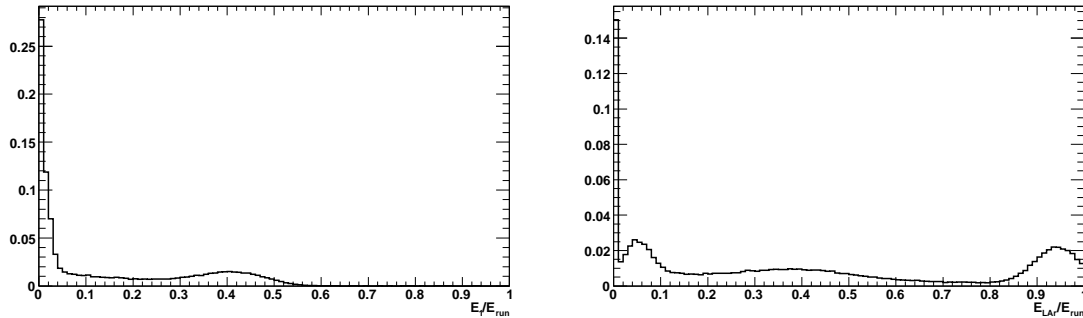


Fig. 7.17: Distributions des rapports E_1/E_{run} et $\sum E_i/E_{\text{run}}$, où E_i est l'énergie reconstruite dans le compartiment i du calorimètre électromagnétique et E_{run} l'énergie nominale du faisceau du lot étudié, ici 9 GeV, après application de tous les critères de purification.

distributions de ces variables avant toutes coupures et après l'application des critères de purification. En pratique on résoud ainsi le problème observé dans la distribution de fraction de points de mesure de hauts seuils dans le TRT de la figure 7.15.

7. IDENTIFICATION DES ÉLECTRONS DE BASSE ÉNERGIE DANS LES DONNÉES DU TEST EN FAISCEAU COMBINÉ H8 DE 2004

Différentes méthodes peuvent être trouvées pour estimer la pureté des lots [143] [149]. Elles n'ont cependant pas été appliquées, essentiellement par manque de temps. Elles permettraient cependant uniquement de quantifier ce dont on peut juger visuellement, à savoir que les lots d'électrons et de pions sont très purs à haute énergie, et le sont de moins en moins à basse énergie.

Le tableau 7.4 rappelle les critères appliqués pour la sélection des lots d'électrons et de pions. Pour chacun il indique le nombre et la fraction d'événements ainsi sélectionnés. On remarquera que le lot d'électrons est assez important, de quelques milliers à quelques dizaines de milliers, ce qui est plus que suffisant pour étudier les variables discriminantes à utiliser, la forme des gerbes électromagnétiques ou mesurer l'efficacité de l'algorithme d'identification. Par contre le nombre de pions est assez faible pour estimer un facteur de rejet. Plus l'énergie du faisceau est basse moins il y a de pions dans le lot. Ainsi à 9 GeV, seuls quelques dizaines de pions sont présents. Aussi ce lot ne sera-t-il pas utilisé par la suite pour les études de performance.

E_{faisceau} (GeV)		Tous les critères	1 amas/trace	qualité des traces	Cerenkov	$cutTRT$	$cutLAR$
1	e	27267	9671	3394	2697	375	1570
	π				121	69	29
2	e	192385	106086	61442	43414	7357	16111
	π				5711	1823	1224
3	e	63837	38030	22846	15494	6341	14378
	π				3237	2566	1518
5	e	62112	41590	26935	13987	11387	12943
	π				7811	11890	4831
9	e	48235	42355	32733	9804	5219	9096
	π				15797	27767	8418
20	e	52671	45576	32676	27767	10742	8112
	π	25276	16311	10065	8418	23942	5842

Tab. 7.4: Liste des critères utilisés, nombre et la fraction d'objets sélectionnés par chaque critère appliqué les uns après les autres. La signification des critères est indiquée dans le texte. Ces chiffres sont donnés pour les lots de données de 1 GeV à 20 GeV.

7.6 Etudes sur le TRT

La figure 7.18 montre, pour les différents lots étudiés, le rapport $N_{\text{HTR}}/N_{\text{straw}}$ entre le nombre de points de mesure de haut seuil dans le TRT N_{HTR} sur le nombre total de pailles touchées N_{straw} . Des études approfondies ont été effectuées par les membres de la collaboration TRT [150]. En particulier une procédure de nettoyage des pailles défectueuses a été introduite et permet d'améliorer un peu les performances de ce détecteur [148]. Par manque de temps cette

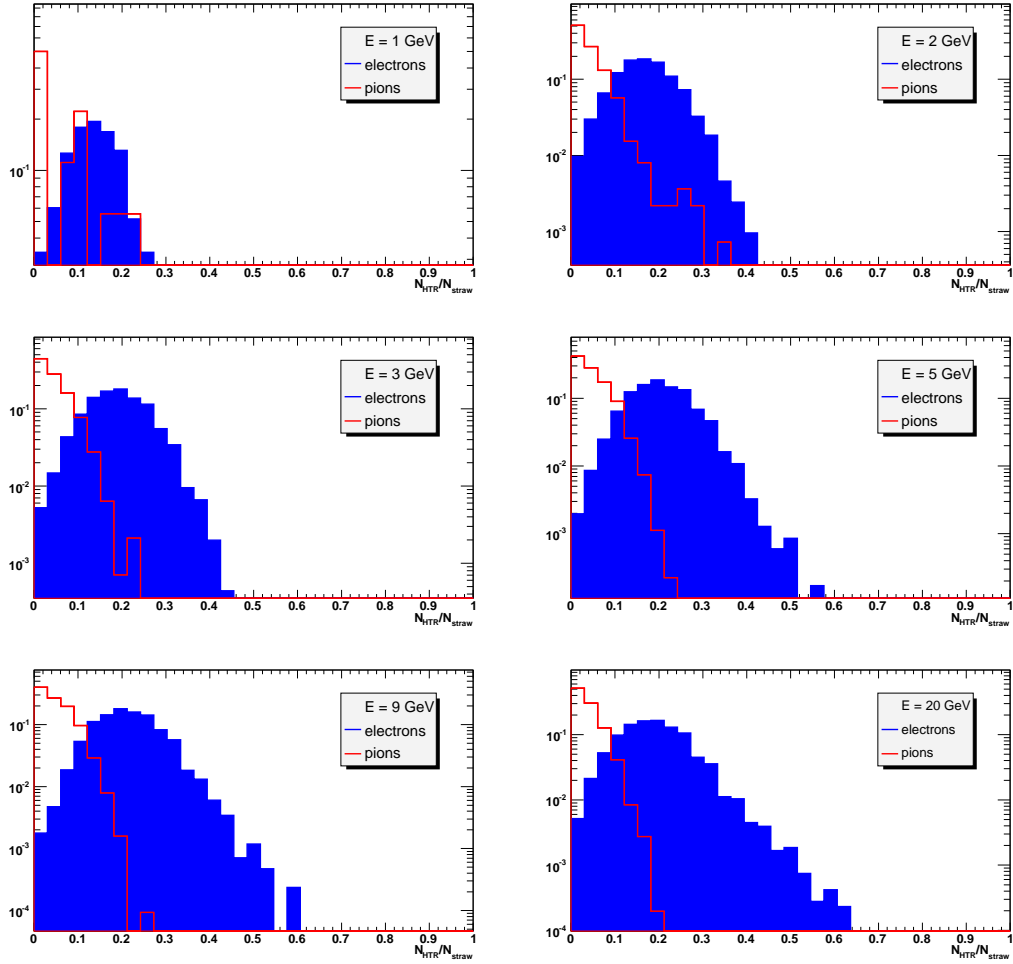


Fig. 7.18: Distribution du rapport $N_{\text{HTR}}/N_{\text{straw}}$ entre le nombre de points de mesure de haut seuil dans la TRT N_{HTR} sur le nombre total de pailles touchées N_{straw} pour des pions et des électrons (histogramme hachuré) pour des données de 1 GeV (en haut à gauche) à 20 GeV (en bas à droite).

procédure n'a pas été appliquée dans cette analyse.

7.7 Etudes de la forme des gerbes électromagnétiques

La forme des gerbes électromagnétiques a déjà été étudiée pour des électrons dans les tests en faisceau effectués en 2002 [151], dans la gamme d'énergie allant de 10 à 245 GeV. Cette étude peut être effectuée avec les données étudiées dans une gamme d'énergie plus basse.

La distribution d'énergie totale reconstruite dans les lots d'électrons et de pions sélectionnés est montrée dans la figure 7.19 dans les faisceaux de 1 GeV à 20 GeV .

7. IDENTIFICATION DES ÉLECTRONS DE BASSE ÉNERGIE DANS LES DONNÉES DU TEST EN FAISCEAU COMBINÉ H8 DE 2004

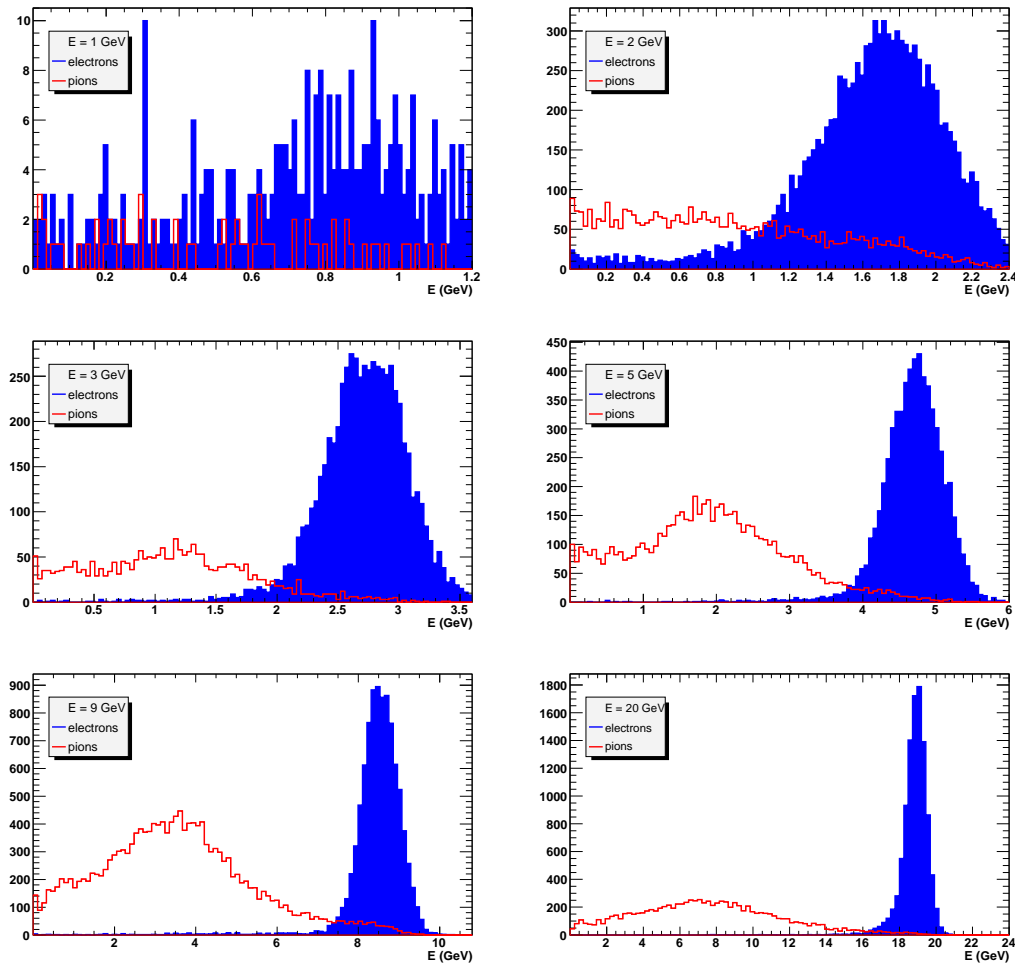


Fig. 7.19: Distribution de l'énergie totale reconstruite pour les électrons (histogramme hachuré) et les pions dans le calorimètre électromagnétique dans les faisceaux de 1 GeV (en haut à gauche) à 20 GeV (en bas à droite), après l'application des critères de purification du faisceau.

7.7.1 Profils longitudinaux

Pour une gerbe initiée par un électron l'énergie déposée atteint rapidement un maximum dans le second compartiment du calorimètre électromagnétique pour ensuite s'atténuer lentement. Plus l'énergie de la particule initiale est importante plus celle-ci va pénétrer profondément dans le calorimètre. Aux énergies relativement basses qui nous intéressent ici, toute la gerbe est contenue dans le calorimètre électromagnétique et presque aucune énergie n'est déposée dans le troisième compartiment.

Le profil longitudinal de la gerbe électromagnétique est montré sur la figure 7.20 pour des électrons et des pions, incidents à $\eta = 0,45$, provenant des faisceaux d'énergies allant de 3 GeV à 20 GeV. Le profil est obtenu en calculant, pour chaque compartiment, la fraction d'énergie qui y est reconstruite E_{comp}/E_{tot} , où $E_{tot} = \sum_{comp} E_{comp}$. Les électrons déposent l'essentiel de leur

énergie dans les deux premiers compartiments du calorimètre électromagnétique. Au delà de 9 GeV la gerbe électromagnétique démarre dans le premier compartiment et a son dépôt d'énergie maximum dans le second compartiment. Plus l'énergie baisse, plus la gerbe a eu un démarrage précoce. En-dessous de 5 GeV le maximum est situé dans le premier compartiment.

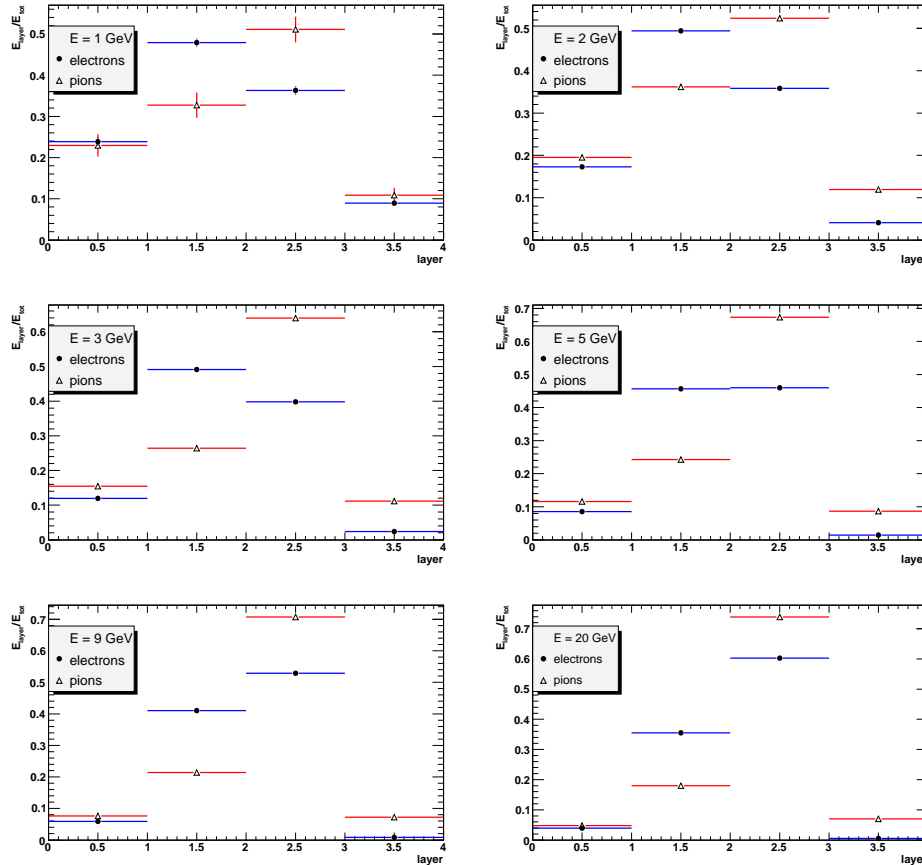


Fig. 7.20: Profils longitudinaux pour des électrons (cercles bleus) et des pions (triangles rouges) dans les faisceaux de 3 GeV (en haut à gauche) à 20 GeV (en bas à droite).

Les pions déposent moins de la moitié de leur énergie dans le calorimètre électromagnétique. Les gerbes hadroniques débutent plus tard que les gerbes électromagnétiques. Le dépôt d'énergie maximal reste dans le deuxième compartiment et il est presque identique dans le premier et le troisième compartiment. Avec l'augmentation de l'énergie du faisceau on peut constater une augmentation du rapport entre l'énergie reconstruite dans le deuxième compartiment et celle du premier.

Cette différence est utilisée pour la séparation entre les électrons et les pions à travers deux variables : la fraction d'énergie reconstruite dans le premier compartiment et celle reconstruite dans le troisième compartiment.

7. IDENTIFICATION DES ÉLECTRONS DE BASSE ÉNERGIE DANS LES DONNÉES DU TEST EN FAISCEAU COMBINÉ H8 DE 2004

7.7.2 Profils latéraux

La forme de la distribution d'énergie mesurée dans la direction η , pour des électrons et des pions d'énergie allant de 1 GeV à 20 GeV, est montrée sur la figure 7.21. La position η est obtenue par rapport au barycentre de la gerbe et est exprimée en unités de cellules du premier compartiment. Le profil latéral est alors obtenu en mesurant la fraction d'énergie reconstruite dans chaque cellule E_{cell}/E_{comp} . La gerbe étant étroite pour un électron, l'extension radiale

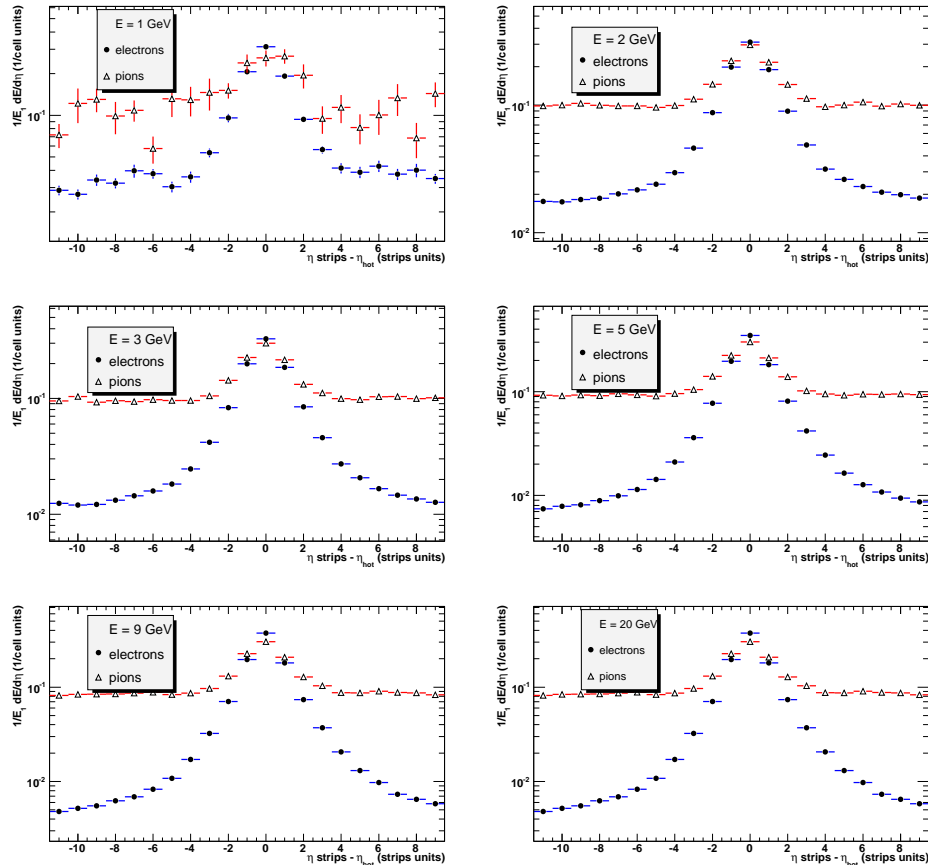


Fig. 7.21: Profils latéraux dans le premier compartiment du calorimètre électromagnétique pour des électrons (cercles bleus) et des pions (triangles rouges) dans les faisceaux de 1 GeV (en haut à gauche) à 20 GeV (en bas à droite).

de la gerbe dépend peu de l'énergie. La distribution est légèrement asymétrique. Cet effet est dû à la géométrie du calorimètre électromagnétique. Les cellules sont projectives vers le point d'interaction. Aussi les cellules du premier compartiment ont un petit angle dans la direction η . Ainsi si les cellules sont perpendiculaires au front de la gerbe à $\eta \sim 0$ elles ne le sont plus à de plus grandes valeurs de η .

Les profils pour les pions sont plus larges que ceux des électrons. En effet les interactions nucléaires responsables des gerbes hadroniques impliquent une plus grande dispersion angulaire

que les processus électromagnétiques (rayonnement de freinage et création de paires).

De la même manière, les profils latéraux dans le deuxième compartiment, sont montrés pour les électrons et les pions sur la figure 7.22. Comme pour le premier compartiment, le

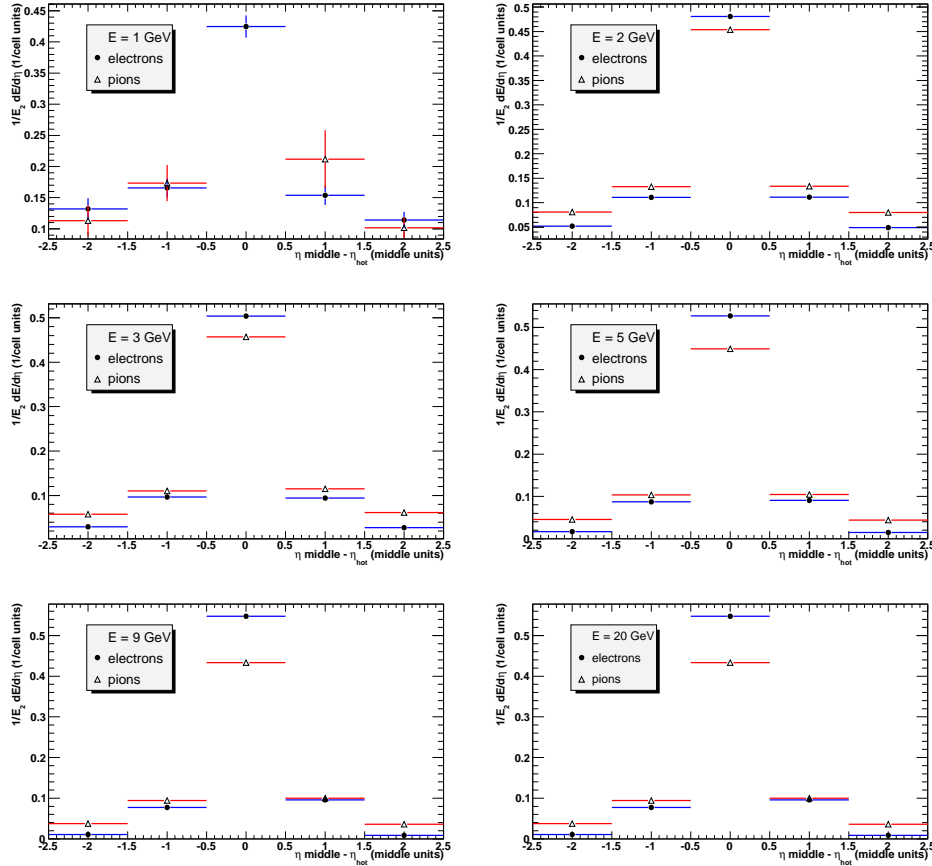


Fig. 7.22: Profils latéraux dans le deuxième compartiment du calorimètre électromagnétique pour des électrons (cercles bleus) et des pions (triangles rouges) dans les faisceaux de 1 GeV (en haut à gauche) à 20 GeV (en bas à droite).

profil est plus large pour les pions que pour les électrons. Cette particularité est utilisée pour l'identification des électrons dans quatre variables discriminantes : les largeurs de la gerbe $\omega_{\eta 1}$ et $\omega_{\text{tot}1}$ dans le premier compartiment ainsi que le développement latéral de la gerbe R_{η} (37) et sa largeur $\omega_{\eta 2}$ dans le deuxième compartiment.

7.8 Identification des électrons

L'identification des électrons repose sur les mêmes caractéristiques que celles qui ont été décrites dans les chapitres précédents et une procédure similaire est utilisée pour l'identification. Les électrons étant isolés nous utiliserons donc les variables décrites au paragraphe 3.4 pour

l'identification de tels électrons.

7.8.1 Variables caractérisant les électrons

Fuites hadroniques :

Les gerbes électromagnétiques ne déposent qu'une faible fraction de leur énergie dans le calorimètre hadronique, typiquement moins de 2% pour des électrons. Les fuites hadroniques sont définies comme le rapport entre l'énergie transverse reconstruite dans le premier compartiment du calorimètre hadronique dans une fenêtre $\Delta\eta \times \Delta\phi = 0,2 \times 0,2$ et l'énergie transverse reconstruite dans le calorimètre électromagnétique. Ce rapport est montré dans la figure 7.23 pour des électrons et des pions du lot de données à 9 GeV.

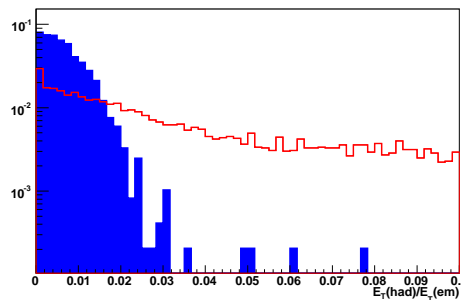


Fig. 7.23: Fuites hadroniques, définies comme le rapport entre l'énergie transverse reconstruite dans le premier compartiment du calorimètre hadronique dans une fenêtre $\Delta\eta \times \Delta\phi = 0,2 \times 0,2$ et sur l'énergie transverse reconstruite dans le calorimètre électromagnétique. Les distributions sont montrées pour des pions et pour des électrons (histogramme hachuré) dans les données à 9 GeV. Elles sont normalisées à l'unité.

Utilisation du second compartiment du ECAL :

Les variables suivantes sont utilisées (cf. figure 3.4) :

- Les électrons déposent la majeure partie de leur énergie dans une fenêtre ayant une taille $\Delta\eta \times \Delta\phi = 3 \times 7$, exprimée en nombre de cellules. Le développement latéral $R_\eta(37)$ est donné par le rapport de l'énergie reconstruite dans une fenêtre 3×7 sur l'énergie reconstruite dans une fenêtre 7×7 .
- La largeur latérale est calculée dans une fenêtre de 3×5 cellules, en utilisant une somme sur la position η de chaque cellule, pondérée par son énergie : $\omega_{\eta 2} =$

$$\sqrt{\frac{\sum E_c \times \eta^2}{\sum E_c} - \left(\frac{\sum E_c \times \eta}{\sum E_c}\right)^2}.$$

Utilisation du premier compartiment du ECAL :

Pour exclure les pions de basse énergie, nous utilisons la fraction d'énergie reconstruite dans le premier compartiment du calorimètre électromagnétique E_1/E . La distribution correspondante est montrée sur la figure 7.25 pour des pions et des électrons.

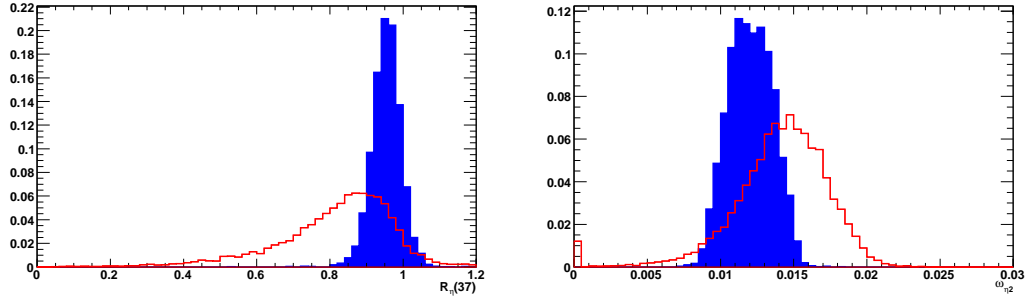


Fig. 7.24: Développement latéral de la gerbe $R_{\eta}(37)$ (gauche) et largeur latérale $\omega_{\eta 2}$ (droite). Les distributions sont montrées pour des pions et pour des électrons (histogramme hachuré), dans les données à 9 GeV. Elles sont normalisées à l'unité.

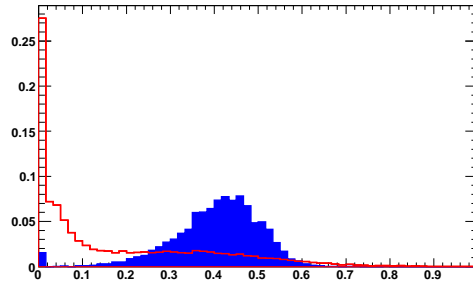


Fig. 7.25: Fraction d'énergie E_1/E reconstruite dans le premier compartiment du calorimètre électromagnétique. Les distributions sont montrées pour des pions et pour des électrons (histogramme plein), dans le lot de données à 9 GeV. Elles sont normalisées à l'unité.

Les variables décrites dans le paragraphe 3.4 permettant de rechercher des sous-structures dans les gerbes et optimisées pour la séparation entre les photons et les π^0 ne sont pas pertinentes dans notre analyse. Aussi le choix des variables est restreint à (cf. figure 7.26) :

- La largeur totale de la gerbe est déterminée dans une fenêtre de taille $\Delta\eta \times \Delta\phi = 0,0625 \times 0,2$, correspondant typiquement à 40 cellules en η . Elle est calculée à partir de : $\omega_{\text{tot1}} = \sqrt{\sum E_i \times (i - i_{\text{max}})^2 / \sum E_i}$, où i est le numéro de cellule considérée et i_{max} celui du premier maximum local.

- La largeur de la gerbe dans les trois cellules autour du maximum est donnée par la formule suivante :

$$\omega_{\text{3strips}} = \sqrt{\sum E_i \times (i - i_{\text{max}})^2 / \sum E_i}, \text{ où } i \text{ est le numéro de cellule utilisé et } i_{\text{max}} \text{ celui du plus énergétique.}$$

Utilisation des informations du détecteur interne :

On rappelle tout d'abord, encore une fois, que cet algorithme a pour graine de départ une trace dans le détecteur interne. L'association trace/amas est donc déjà effectuée, ainsi que l'applica-

7. IDENTIFICATION DES ÉLECTRONS DE BASSE ÉNERGIE DANS LES DONNÉES DU TEST EN FAISCEAU COMBINÉ H8 DE 2004

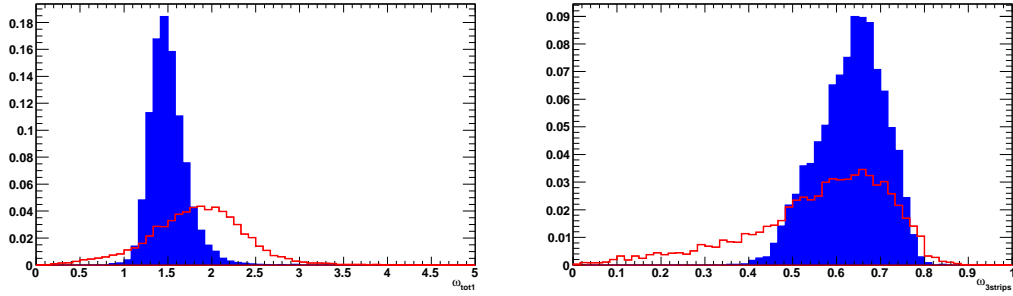


Fig. 7.26: Largeur totale de la gerbe ω_{tot1} (à gauche) et dans trois strips $\omega_{3strips}$ (à droite) dans le premier compartiment du calorimètre électromagnétique. Les distributions sont montrées pour des pions et pour des électrons (histogramme plein), dans le lot de données à 9 GeV. Elles sont normalisées à l'unité.

tion de critères sur la qualité des traces ainsi que décrit plus haut.

Le rejet des pions peut être grandement amélioré en s'assurant de la cohérence entre les informations du calorimètre électromagnétique et du détecteur interne. On vérifie donc l'association angulaire (cf. figure 7.27) :

- $\Delta\eta = \eta_{cluster}^{strips} - \eta_{ID}$, où $\eta_{cluster}^{strips}$ est calculé dans le premier compartiment du calorimètre électromagnétique où la granularité est très fine et η_{ID} est la pseudorapidité de la trace extrapolée jusqu'au calorimètre ;
- $\Delta\phi = \phi_{cluster}^{middle} - \phi_{ID}$, où $\phi_{cluster}^{middle}$ est calculé dans le second compartiment du calorimètre électromagnétique et ϕ_{ID} est l'azimuth de la trace extrapolée jusqu'au calorimètre.

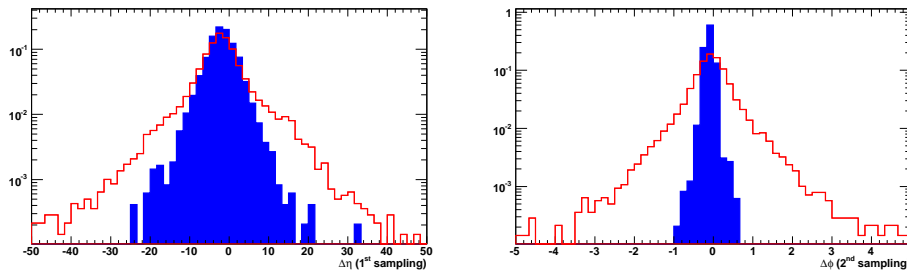


Fig. 7.27: Différence angulaire en pseudorapidité ($|\Delta\eta|$) et en azimuth ($|\Delta\phi|$) entre la position de la trace extrapolée au calorimètre électromagnétique et celle de l'amas électromagnétique. Les distributions sont montrées pour des pions et pour des électrons (histogramme hachuré), dans le lot de données à 9 GeV. Elles sont normalisées à l'unité.

Les données utilisées n'ont pas de champ magnétique. On ne peut donc reconstruire l'impulsion de la trace et par conséquent utiliser le rapport énergie sur impulsion.

Enfin, on utilise l'information provenant du TRT tel que montré dans la figure 7.18.

7.8.2 Performances

L'identification est basée sur l'utilisation d'une variable discriminante basée sur un rapport de vraisemblances construit à partir des distributions des variables discriminantes pour le signal et le bruit de fond :

$$X_{\text{RL}} = \frac{\prod_i g_i^s(x_i)}{\prod_i g_i^b(x_i)}, \quad (7.1)$$

où $g_i^s(x_i)$ (resp. $g_i^b(x_i)$) est la distribution de densité de probabilité à une dimension de la variable x_i pour les événements de signal (resp. de bruit de fond). Ces distributions, appelées fonctions de densité de probabilité (PDF - Probability Density Function) ont été présentées dans les figures 7.23 à 7.27. L'efficacité d'identification des électrons ε_e est définie comme $\varepsilon_e = \frac{N_e^t}{N_e}$, où N_e est le nombre total d'électrons sélectionnés. N_e^t est le nombre d'électrons identifiés après les critères de sélection. Les candidats qui ont une valeur de X_{RL} plus petite qu'un certain seuil sont rejetées. Le facteur de rejet des pions est défini comme $R_\pi = N_\pi/N_\pi^t$, où N_π est le nombre de pions sélectionnés et N_π^t est le nombre de pions identifiés comme un électron. La courbe de facteur de rejet des pions en fonction de l'efficacité d'identification des électrons est obtenue en faisant varier la valeur du seuil.

Analyse avec uniquement le TRT

Pour estimer les performances de l'identification des électrons uniquement avec les informations du TRT, on applique le critère de sélection *cutLAR* défini plus haut pour définir les lots d'électrons et de pions. Les performances obtenues pour les différents lots sont récapitulées dans la table 7.5. Ainsi, par exemple à 9 GeV, pour une efficacité d'identification des électrons de 90%, le facteur de rejet des pions est d'environ 35. La figure 7.28 à gauche montre la distribution du facteur de rejet des pions obtenu en fonction de l'efficacité d'identification des électrons, lorsqu'on utilise uniquement les informations du TRT. Ces résultats sont en accord avec ceux obtenus par d'autres groupes [148] et [149]. Le facteur de rejet obtenu à 9 GeV est cependant environ 30% plus bas, ce qui pourrait s'expliquer par l'absence de nettoyage des pailles défectueuses dans le TRT de l'analyse présentée ici comme expliqué plus haut. La figure 7.28 à droite montre la distribution du facteur de rejet des pions en fonction de l'énergie nominale du faisceau, pour une efficacité d'identification de 90%. Le facteur de rejet est assez dépendant de l'énergie, l'émission de photons de rayonnement de transition n'ayant pas encore atteint son maximum à 2-3 GeV puis vers 20 GeV étant sensible à la remontée relativiste dans les pertes d'énergie par dE/dx .

Analyse uniquement avec le calorimètre électromagnétique

Pour estimer les performances de l'identification des électrons uniquement avec les informations du calorimètre électromagnétique, on applique le critère de sélection *cutTRT* définie plus

7. IDENTIFICATION DES ÉLECTRONS DE BASSE ÉNERGIE DANS LES DONNÉES DU TEST EN FAISCEAU COMBINÉ H8 DE 2004

ε (%)	E_{faisceau} (GeV)					
	2	3	5	9	20	
80	37 ± 18	44 ± 10	90 ± 15	110 ± 15	100 ± 15	100 ± 15
90	15 ± 2	16 ± 2	28 ± 2	35 ± 2	28 ± 2	25 ± 2

Tab. 7.5: Facteur de rejet obtenu pour différentes valeurs de l'efficacité d'identification des électrons, en utilisant uniquement les informations du trajectographe à rayonnement de transition.

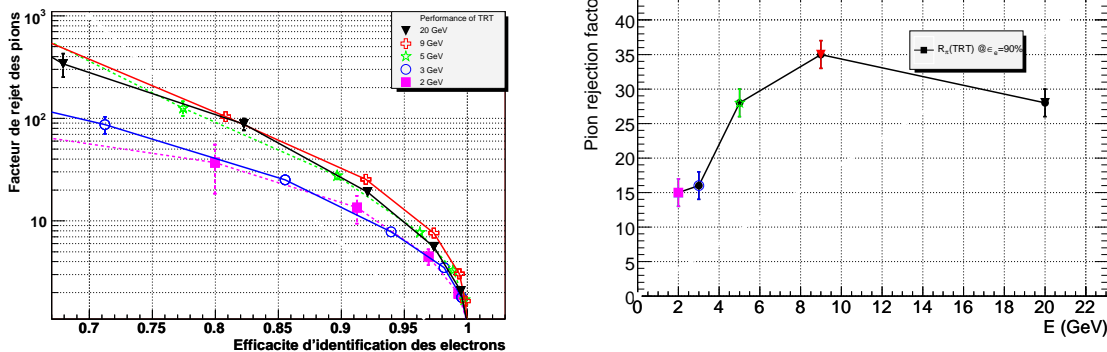


Fig. 7.28: Facteur de rejet des pions en fonction de l'efficacité d'identification des électrons (à gauche) pour différents lots de données et en fonction de l'énergie nominale du faisceau (à droite) pour une efficacité de 90%. Seule l'information du trajectographe à rayonnement de transition est utilisée.

haut pour définir les lots d'électrons et de pions. Les performances obtenues pour les différents lots sont récapitulées dans la table 7.6. Ainsi, par exemple à 9 GeV, pour une efficacité d'identification des électrons de 90%, le facteur de rejet des pions est, comme pour le TRT, d'environ 50. La figure 7.29 à gauche montre la distribution du facteur de rejet des pions en fonction de l'efficacité d'identification des électrons, pour différents lots de données, lorsqu'on utilise uniquement les informations du calorimètre électromagnétique. La figure 7.29 à droite montre le facteur de rejet des pions en fonction de l'énergie nominale du faisceau, pour une efficacité d'identification de 90%. Celle-ci croît régulièrement, augmentant d'un facteur 10 entre 2 GeV et 20 GeV.

ε (%)	E_{beam} (GeV)				
	2	3	5	9	20
80	27 ± 5	40 ± 5	70 ± 8	150 ± 15	
90	12 ± 1	24 ± 3	41 ± 3	70 ± 5	

Tab. 7.6: Facteur de rejet obtenu pour différentes valeurs de l'efficacité d'identification des électrons, en utilisant uniquement les informations du calorimètre électromagnétique.

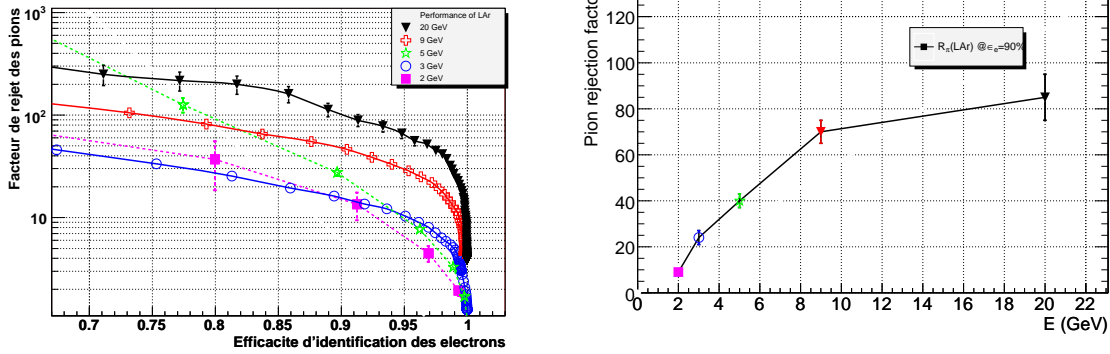


Fig. 7.29: Facteur de rejet des pions en fonction de l'efficacité d'identification des électrons (à gauche) pour différents lots de données et en fonction de l'énergie nominale du faisceau (à droite) pour une efficacité de 90%. Seule l'information du calorimètre électromagnétique est utilisée.

7.9 Conclusion

Le test en faisceau combiné 2004 a été l'occasion de tester pour la première fois l'algorithme de reconstruction et d'identification des électrons de bas p_T avec des données réelles. Les performances du TRT ont déjà été mesurées en test faisceau. De même la reconstruction et de l'identification des électrons a été étudiée avec précision, mais à des énergies plus élevées, avec le calorimètre électromagnétique. Ce test en faisceau a été la première occasion de reconstruire des électrons, de manière combinée, avec ces deux détecteurs, et ce tout particulièrement dans la gamme d'énergie allant de 1 GeV à 9 GeV. Des études spécifiques sur le TRT et la forme des gerbes électromagnétiques ont été effectuées mais sont restées limitées en l'absence de données de simulation correspondantes. L'identification des électrons isolés est basée sur les mêmes variables que celles définies dans les chapitres précédents. et du calorimètre électromagnétique, de manière séparé. Rappelons encore une fois, que l'estimation de ces performances est effectuée à partir d'un seul et même objet commun "Electron de bas p_T ". Les performances combinées n'ont pas pu être mesurées essentiellement car les faisceaux utilisés n'étaient pas purs et que l'instrumentation présente le long de la ligne de faisceau n'a pas permis d'obtenir de lots purs d'électrons et de pions. De plus les données utilisées contenaient trop peu d'événements pour mesurer des facteurs de rejet des pions de l'ordre de 1000.

L'analyse des données du test en faisceau est encore en cours. Les prochains mois devraient voir la finalisation des analyses effectuées par les différents groupes, en particulier pour les données à basse énergie. En particulier il serait nécessaire d'ajouter une comparaison entre les données et la simulation et d'avoir une meilleure estimation de l'efficacité des critères de sélection et de la pureté des lots ainsi obtenus.

8. CONCLUSION

Ce mémoire a présenté ma participation à la préparation de l'expérience Atlas. Les enjeux du LHC sont multiples et d'une importance cruciale pour la discipline : étude de la brisure de la symétrie électrofaible, recherche et étude du boson de Higgs, mesures de précision du Modèle Standard et si possible mise en évidence d'une nouvelle physique au-delà. Dans la plupart de ces études les électrons joueront un rôle fondamental, permettant le déclenchement du système d'acquisition, l'étude des performances des sous-détecteurs et la reconstruction des états finals.

La reconstruction des électrons commence par celle d'objets électromagnétiques de grandes énergies dans le calorimètre puis en leur associant une trace de bonne qualité. L'environnement difficile dans lequel les analyses seront effectuées implique des contraintes sévères sur l'identification des électrons. Il est possible de garder une bonne identification des électrons isolés tout en maintenant un facteur de rejet des jets de l'ordre de 10^5 . Outre ma participation à la mesure des performances de cet algorithme et à la maintenance du code de reconstruction Athena, j'ai par ailleurs contribué au développement d'un algorithme de reconstruction des électrons débutant lui par une trace dans le détecteur interne. Cet algorithme permet de reconstruire des électrons à des impulsions transverses bien plus basses que l'algorithme standard et possiblement imbriqués dans des jets. L'application principale de cet algorithme est la reconstruction d'événements J/ψ , permettant à terme d'effectuer un étalonnage in-situ du calorimètre électromagnétique, à basse énergie, complémentaire de celle utilisant les événements Z^0 . Une autre application est l'étiquetage des jets de b par la présence d'un électron mou. Si cet algorithme souffre du relativement faible rapport d'embranchement de la désintégration semileptonique du quark b , ses performances lui permettront d'être, avec les premières données, un outil appréciable de vérification des algorithmes d'étiquetage standards, et à terme d'améliorer les performances globales. L'algorithme a par ailleurs été utilisé avec succès sur les données du test en faisceau combiné 2004, permettant de mesurer ses performances avec des faisceaux d'énergie allant de 1 GeV à 20 GeV.

Ces études ont été effectuées, pour la plupart, avec des données de simulation Monte Carlo. Le développement d'un nœud de la grille de calcul au laboratoire a permis de participer à cet effort et devrait permettre de contribuer au succès de nos analyses avec les données du LHC. Tous ces développements pourront s'intégrer dans une analyse de mesure de la masse du quark top avec les premières données du LHC en développement dans le groupe du LPNHE.

BIBLIOGRAPHIE

- [1] Liste des TDR de la collaboration Atlas,
<http://atlas.web.cern.ch/Atlas/internal/tdr.html>.
- [2] ATLAS collaboration, *The ATLAS experiment at the CERN Large Hadron Collider*,
subm. to JINST
[https://twiki.cern.ch/twiki/bin/view/Atlas/AtlasTechnicalPaper\(2009\)](https://twiki.cern.ch/twiki/bin/view/Atlas/AtlasTechnicalPaper(2009)).
- [3] Elémentaire, revue d'information scientifique,
<http://elementaire.web.lal.in2p3.fr>.
- [4] L'actualité du LAL, numéro 42 : ATLAS au cœur de la matière,
http://www.lal.in2p3.fr/rubrique.php3?id_rubrique=27.
- [5] Ph. Schwemling, *Les enjeux du nouveau collisionneur LHC*, Images de la physique
(2005), <http://www.cnrs.fr/publications/imagesdelaphysique/2005.htm>.
- [6] Portail internet du Large Hadron Collider, <http://lhcb.web.cern.ch/lhcb/>.
- [7] Portail internet du Cern, <http://cern.ch/>.
- [8] Portail internet de l'expérience Alice, <http://aliceinfo.cern.ch>.
- [9] Portail internet de l'expérience Atlas, <http://atlas.web.cern.ch>.
- [10] Portail internet de l'expérience CMS, <http://cms.cern.ch/>.
- [11] Portail internet de l'expérience LHCb, <http://lhcb.web.cern.ch/lhcb/>.
- [12] Portail internet de l'expérience Totem, <http://totem.web.cern.ch/Totem/>.
- [13] Portail internet de la collaboration, <http://atlas.ch/>.
- [14] Portail internet de la collaboration Atlas France, <http://atlas-france.in2p3.fr>.
- [15] ATLAS collaboration, *Technical Proposal for a General-Purpose pp Experiment at the Large Hadron Collider at CERN*, CERN/LHCC 94-43, LHCC/P2 (1994).
- [16] ATLAS collaboration, *TDR ATLAS detector and physics performance, vol. 1*,
CERN/LHC 99-14, ATLAS-TDR-14 (1999).
- [17] T. Akesson et al. (ATLAS Coll.), *ATLAS Transition Radiation Tracker test-beam results*,
Nucl. Instrum. Meth. **A522** (2004) 50–55.
- [18] V. Mitsou et al. (ATLAS Coll.), *The ATLAS Transition Radiation Tracker*, ATL-CONF-
2003-012 (2003), proc. of ICATPP 2003, .

- [19] Aubert, B et al.- ATLAS Electromagnetic Liquid Argon Calorimeter Group., *Construction, assembly and tests of the ATLAS electromagnetic barrel calorimeter*, Nucl. Instrum. Methods Phys. Res. **A 558** (2006) 388–418, CERN-PH-EP-2005-034.
- [20] D. Lacour, *Les électrodes du calorimètre électromagnétique du détecteur Atlas*, HDR, Université Paris 7, 2002.
- [21] Ph. Schwemling, *Construction du calorimètre électromagnétique de l'expérience Atlas et minimisation du terme constant de la résolution*, HDR, Université Paris 7, 2002, http://lfnhe-atlas.in2p3.fr/pubdocs/present/These_Habilitation_PhS.ps.
- [22] B. Laforge, *Le calorimètre électromagnétique d'Atlas et son électronique. Préparation des analyses de physique*, HDR, Université Paris 6, 2004.
- [23] L. Serin, *Calorimétrie électromagnétique d'Atlas*, HDR, Université de Paris Sud, 2006.
- [24] F. Hubaut, *Préparation de l'expérience Atlas auprès du futur grand collisionneur de protons LHC : performances du calorimètre électromagnétique et potentiels pour la physique du quark top*, HDR, Université de la Méditerranée - Aix-Marseille II, CPPM-H-2007-01, 2007, <http://tel.archives-ouvertes.fr/tel-00144390>.
- [25] P. Pralavorio, *Calorimétrie électromagnétique et mesures de précision avec le détecteur Atlas auprès du collisionneur LHC*, HDR, Université de la Méditerranée - Aix-Marseille II, CPPM-H-2007-03, 2007, <http://tel.archives-ouvertes.fr/tel-00158826>.
- [26] Escalier M., Derue F., Fayard L., Kado M., Laforge B., Reifen C., Unal G., *Photon/jet separation with DC1 data*, 2005, ATL-PHYS-PUB-2005-018.
- [27] Reifen C., Derue F., Escalier M., Fayard L., Kado M., Laforge B., Unal G., *γ/π^0 separation in the 1st compartment of the ATLAS EM Calorimeter with DC2 data*, 2005, ATL-PHYS-INT-2005-005.
- [28] J. Colas et al. (ATLAS EMLarg Coll.), *Position resolution and particle identification with the ATLAS EM calorimeter*, Nucl. Instrum. Meth **A550** (2005) 96–115.
- [29] I. Nikolic, L. Serin, *Time resolution of the ATLAS Liquid Argon Calorimeter*, 2004, ATLAS-COM-LARG-2004-002.
- [30] ATLAS computing group, *ATLAS Computing Technical Design Report*, CERN-LHCC-2005-022, ATLAS-TDR-17 (2005).
- [31] Pérez-Réale, V, *Electron/photon Identification and Standard Model Higgs Process Studies at the High Level Trigger for the ATLAS experiment*, Ph.D. thesis, Université de Berne, 2006, CERN-THESIS-2006-055.
- [32] ATLAS computing group, *ATLAS Computing Technical Design Report*, CERN-LHCC-2005-022, ATLAS-TDR-17 (2005).
- [33] Révision des modèles de calcul, <http://lcg.web.cern.ch/LCG/MB/revise%5Fresources/>.

- [34] D. Rousseau, *La reconstruction dans Atlas*, HDR, Université de Paris Sud, 2007.
- [35] Binet S., *Environnement logiciel et étalonnage de l'échelle en énergie des jets dans l'expérience Atlas*, Ph.D. thesis, Université Blaise Pascal, 2006, PCCF T 0601.
- [36] M. Lechowski, *Test du modèle du Petit Higgs dans Atlas au LHC. Simulation de la numérisation du calorimètre électromagnétique.*, Ph.D. thesis, Université de Paris Sud, 2005, LAL 05-20.
- [37] ATHENA Developers guide,
<http://atlas.web.cern.ch/Atlas/GROUPS/SOFTWARE/OO/architecture/General/Documentation/AthenaDeveloperGuide-8.0.0-draft.pdf>.
- [38] G. Barrand et al., *Gaudi - A Software Architecture and Framework for building HEP Data Processing Applications*, International Conference on Computing in High Energy Physics (CHEP) (2000).
- [39] ATRECON,
<http://atlas.web.cern.ch/Atlas/GROUPS/SOFTWARE/DOCUMENTS/reconstruction.html>.
- [40] S. Agostinelli et al., *Geant4, a simulation toolkit*, Nuclear Instruments and Methods in Physics Research **A 506** (2003) 250–303, <http://geant4.web.cern.ch/geant4>.
- [41] Portail internet de Hbook, http://wwwinfo.cern.ch/asdoc/hbook_html3/hboomain.htm.
- [42] Portail internet de Root, <http://root.cern.ch>.
- [43] Analyse distribuée avec Ganga,
<https://twiki.cern.ch/twiki/bin/view/Atlas/DistributedAnalysisUsingGanga&lang=en>.
- [44] Analyse distribuée avec Panda,
<https://twiki.cern.ch/twiki/bin/view/Atlas/PanDA>.
- [45] Memorandum of Understanding for Collaboration in the Deployment and Exploitation of the Worldwide LHC Computing Grid, CERN-C-RRB-2005-01/ref, 2007,
<http://lcg.web.cern.ch/lcg/C-RRB/MoU/WLCGMoU.pdf>.
- [46] Centre Tier-1 du CCIN2P3, http://lcg.in2p3.fr/wiki/index.php/Tier_1 :CCIN2P3.
- [47] Centre Tier-1 BNL, <http://www.acf.bnl.gov/>.
- [48] Centre Tier-1 FZK (Karlsruhe), http://www.fzk.de/fzk/idcplg?IdcService=FZK&node=0730&lang=en#P32_3575.
- [49] LCG France project,
<http://lcg.in2p3.fr>, <https://edms.in2p3.fr/document/I-003682>.
- [50] LCG France Tier-2 Tier-3 resource planning 2006-2010,
<https://edms.in2p3.fr/document/I-008837>.
- [51] LCG France status report : Tier-2 et Tier-3, conseil scientifique de l'IN2P3 02/07/2007.

- [52] Centre Tier-2 de Clermont, http://lcg.in2p3.fr/wiki/index.php/Tier_2 :LPC.
- [53] Centre Tier-2 du LAPP, http://lcg.in2p3.fr/wiki/index.php/Tier_2 :LAPP.
- [54] Centre Tier-2 de Pékin, <http://www.ihep.ac.cn/english/index.htm>.
- [55] Centre Tier-2 de Tokyo, <http://www.icepp.jp/index.html>.
- [56] Centre Tier-2 de Roumanie.
- [57] Centre Tier-3 du LPSC, http://lcg.in2p3.fr/wiki/index.php/Tier_3 :LPSC.
- [58] Centre Tier-3 du CPPM, http://lcg.in2p3.fr/wiki/index.php/Tier_3 :CPPM.
- [59] F. Derue, M. Jouvin, J-P. Meyer, C. Charlot, Le projet GRIF, 2005, <https://edms.in2p3.fr/document/I-010718/1>.
- [60] Portail internet de Grif au LPNHE, <http://lpnhe-grif.in2p3.fr>.
- [61] F. Derue, L. Martin, D. Terront, Le projet GRIF au LPNHE. Document A v2 : description scientifique du projet, 2007, <https://edms.in2p3.fr/document/I-012261/1>.
- [62] F. Derue, L. Martin, D. Terront, Le projet GRIF au LPNHE. Document B v2 : mise en œuvre du projet, 2007, <https://edms.in2p3.fr/document/I-012262/1>.
- [63] F. Derue, L. Martin, D. Terront, Le projet GRIF au LPNHE. Document C v2 : déploiement du matériel, 2007, <https://edms.in2p3.fr/document/I-012263/1>.
- [64] F. Derue, L. Martin, D. Terront, Le projet GRIF au LPNHE. Document D v2 : étude de risques, 2007, <https://edms.in2p3.fr/document/I-012264/1>.
- [65] Action Concertée Incitative GRID, <http://www-sop.inria.fr/aci/grid/public/>.
- [66] Action Concertée Incitative Masse de Données, <http://acimd.labrid.fr>.
- [67] Grid'5000, a large scale nation wide infrastructure for Grid research, <http://www.grid5000.org>.
- [68] The European Research Network on Foundations, Software Infrastructures and Applications for large scale distributed, GRID and Peer-to-Peer Technologies, <http://www.coregrid.net>.
- [69] MONARC project, <http://monarc.web.cern.ch/MONARC/>.
- [70] DataGRID project, <http://eu-datagrid.web.cern.ch/eu-datagrid/>.
- [71] EGEE project, <http://www.eu-egee.org/>.
- [72] LCG project, <http://lcg.web.cern.ch/LCG/>.

- [73] L'Institut des Grilles,
<http://idgrilles.lal.in2p3.fr>.
- [74] Le Réseau National de Télécommunications pour la Technologie l'Enseignement et la Recherche, <http://www.renater.fr/>.
- [75] Le réseau Geant, <http://www.dante.net/server/show/nav.007>.
- [76] Institut des Systèmes Complexes, <http://iscpif.fr>.
- [77] Earth Science Research,
<http://www.ipgp.jussieu.fr/moguilny/GRILLE/>.
- [78] Fédération de Recherche sur les Interactions Fondamentales,
<http://www.lpthe.jussieu.fr/fed/>.
- [79] Le stockage dCache,
<http://cc.in2p3.fr/docenligne/424> ; <http://www.dcache.org/>.
- [80] Cern Advanced STORage manager,
<http://castor.web.cern.ch/castor/>.
- [81] procédure de Soutien aux Equipes Scientifiques pour l'Acquisition de Moyens Expérimentaux de la Région Ile de France :,
<http://www.iledefrance.fr/?IDINFO=6666>.
- [82] Report of the Steering Group of the LHC Computing Review,
http://cern.ch/lhc-computing-review-public/Public/Report_final.PDF.
- [83] Manuel utilisateur de l'exécuteur LyonDQ2,
<https://atlas-france.in2p3.fr/cgi-bin/twiki/bin/view/Atlas/ProdMcFr>.
- [84] Manuel utilisateur de l'exécuteur Panda à Lyon,
<https://atlas-france.in2p3.fr/cgi-bin/twiki/bin/view/Atlas/PandaLyon>.
- [85] Portail internet pour l'accounting EGEE,
http://www3.egee.cesga.es/gridsite/accounting/CESGA/egee_view.php.
- [86] Page web du groupe Calcul Atlas France,
<https://atlas-france.in2p3.fr/cgi-bin/twiki/bin/view/Atlas/Informatique>.
- [87] M. Escalier, *Recherche expérimentale de la brisure spontanée de symétrie électrofaible dans le canal $H \rightarrow \gamma\gamma$ et d'une solution au problème de hiérarchie dans Atlas. Participation à la préparation de l'électronique du calorimètre électromagnétique.*, Ph.D. thesis, Université de Paris Sud, 2005, CERN-THESIS-2005-023, CNRS/tel-00009292, LPNHE-7855.
- [88] N. Kerschen, *Etude d'un algorithme de reconstruction des électrons dans l'expérience Atlas auprès du LHC*, Ph.D. thesis, Université de Paris Sud, 2006, dapnia-06-21-T.

- [89] Gaumer O., *Etude de la production de Z' au LHC. Mesure de l'uniformité et de la résolution en énergie des modules du calorimètre en faisceau-test*, Ph.D. thesis, Université de Savoie, 2004.
- [90] F. Abe et al. (CDF Collaboration), *Phys. Rev. Lett* **69** (1992) 3704.
- [91] Etzion, E ; Ginzburg, J ; Kartvelishvili, V G ; Price, D, *Heavy Quarkonium Production Simulations at ATLAS*, 2007, ATL-PHYS-INT-2007-001.
- [92] Derue, F., B-physics overview in ATLAS, (QCD 06, Montpellier, France, 2006), ATL-PHYS-PUB-2007-002.
- [93] ATLAS collaboration, *TDR ATLAS detector and physics performance, vol. 2*, CERN/LHC 99-15, ATLAS-TDR-15 (1999).
- [94] The ALEPH Collaboration, the DELPHI Collaboration, the L3 Collaboration, the OPAL Collaboration, the LEP Electroweak Working Group, the SLD Electroweak and Heavy Flavour Groups, *Precision Electroweak Measurements on the Z Resonance*, Physics Report **427** (2006) 5–6, CERN-PH-EP/2005-041, SLAC-R-774.
- [95] Derue F., Serfon C., *Electron identification with DC1 data*, 2004, ATLAS-PHYS-2004-012.
- [96] Derue F., Electron and photon identification in ATLAS, (Les Diablerets 2005, Hadron collider physics, 2005), pp. 107–111, ATLAS-PHYS-2004-012.
- [97] Sjostrand T. et al., *Pythia*, *Comp. Phys. Comm.* **82** (1994) 74, Sjostrand T., *Pythia 6.206*, LU TP 01-02 [hep-ph/0108264] 2002.
- [98] GEANT 3, <http://www.eu-egee.org/>.
- [99] Portail internet sur ATLSIM, <http://atlas.web.cern.ch/Atlas/GROUPS/SOFTWARE/DOCUMENTS/simulation.html#atlsim>.
- [100] S. Frixione and B.R. Webber, *Matching NLO QCD computations and parton showers in heavy flavour production*, *JHEP* **0206** (2002) 029.
- [101] Diaz-Gomez M. et al., *The e/γ analysis framework*, 2004, ATL-COM-DAQ-2004-010.
- [102] Benslama et al., *Electron reconstruction and identification with the ATLAS detector*, 2008, Atlas CSC EG-1 note in preparation.
- [103] Turlay E., Zerwas D., *Electrons ID studies : loose/medium/tight cuts*, réunion egamma, 20 mars 2007 (2007).
- [104] Wielers M., *Isolation of photons*, 2002, ATL-PHYS-2002-004.
- [105] Besson N., Boonekamp M., *Determination of the absolute lepton scale using Z boson decays. Application to the measurement of m_W* , 2006, ATLAS-PHYS-PUB-2006-007.
- [106] Di Caccio L., Goy C., Jézéquel S., *Early extraction of the $Z \rightarrow e^+e^-$ signal using the Atlas detector*, 2007, ATLAS-PHYS-INT-2007-006.

- [107] Derue F., Kaczmarska A., Schwemling, Ph., *Reconstruction of DCI $J/\psi \rightarrow e^+e^-$ decays and use for the low energy calibration of the ATLAS electromagnetic calorimeter*, 2006, ATL-PHYS-PUB-2006-004.
- [108] Berger-Hryn'ova T., Bouhova-Thacker E., Cochran J., Derue F., Kaczmarska A., Kartvelishvili V., Kirk J., Lytken E., Nelson A., *Reconstruction of $J/\psi \rightarrow e^+e^-$ and $\Upsilon \rightarrow e^+e^-$ with CSC data*, 2007, Atlas CSC EG7 note in prepration.
- [109] Rocheman S., *Calibration du calorimètre électromagnétique d'Atlas avec des particules J/Ψ et Upsilon*, Stage, DEA, 2003,
<http://lfnhe-atlas.in2p3.fr/TheseStage/Stages/documents/rapport-Rocheman-2003.ps.gz>.
- [110] Kaczmarska A., *Identification of low energy electrons in ATLAS experiment at LHC*, Ph.D. thesis, Institute of Nuclear Physics, Cracovie, 2000.
- [111] Cochran, J.; Derue, F.; Kaczmarska, A.; Wolter, M., *Identification and use of low p_T electrons with Rome data*, 2006, ATL-PHYS-PUB-2006-026.
- [112] Djama, F., *Using $Z^0 \rightarrow e^+e^-$ for Electromagnetic Calorimeter calibration*, 2004, ATLAS-LARG-2004-008.
- [113] Pérez-Réale, V., *Electron and photon trigger menus at 10^{31} luminosity*, réunion egamma, 5 juin 2007 (2007), <http://indico.cern.ch/conferenceDisplay.py?confId=16155#4>.
- [114] Tetiana Hryn'Ova, *Electron trigger selection optimisation and expected rates at low luminosity*, réunion egamma, 5 juin 2007 (2007),
<http://indico.cern.ch/conferenceDisplay.py?confId=16155#4>.
- [115] Tetiana Hryn'Ova, *Status of Di-electron trigger studies*, réunion egamma, 29 aout 2007 (2007),
<http://indico.cern.ch/conferenceDisplay.py?confId=16155#4>.
- [116] Kirk, J; Baines, J T M; Watson, A T, *ATLAS B-physics Trigger Studies using EM and Jet RoIs*, 2006, ATL-DAQ-PUB-2006-004.
- [117] Scheirich, D; Pérez-Réale, V, *Level2 Trigger Study of $J/\Psi \rightarrow e+e-$ Using Level1 EM RoI's*, 2007, ATL-COM-DAQ-2007-010.
- [118] Kartvelishvili V., *ID brem recovery and kalman-DNA fit*, réunion egamma, 29 mars 2007 (2007), <http://indico.cern.ch/getFile.py/access?contribId=145&sessionId=17&resId=0&materialId=slides&confId=5060>.
- [119] Kaczmarska A., *Tagging b -jets using low- p_T electrons (full rapidity range)*, 2000, ATLAS-PHYS-2000-012.
- [120] Jagielski S., Kaczmarska A., Wolter M., *Tagging b -jets using low- p_T electrons*, Acta Phys. Polon. **b30** (1999) 931.

- [121] Wolter M., Kaczmarska A., *Combining b-tagging methods using a neural network approach*, 2000, ATLAS-INDET-2000-023.
- [122] S. Corréard et al., *b-tagging with DC1 data*, 2004, ATL-PHYS-2004-006.
- [123] Bold, T., Derue F., Kaczmarska A., Stanecka, E., Wolter, M., *Pile-up studies for soft electron identification and b-tagging with DC1 data*, 2006, ATL-PHYS-PUB-2006-001.
- [124] Derue F., Kaczmarska A., *Soft electron identification and b-tagging with DC1 data*, 2004, ATLAS-PHYS-2004-036.
- [125] Portail internet du groupe Flavour Tagging,
<https://twiki.cern.ch/twiki/bin/view/Atlas/FlavourTagging>.
- [126] O. Ali, *Identification des électrons non isolés dans Atlas*, Stage, Licence de Physique PhiTEM, ENS de Cachan-UPMC, 2003,
<http://lpnhe-atlas.in2p3.fr/TheseStage/Stages/documents/rapport-OALi-2003.ps.gz>.
- [127] A. Hocker et al., *TMVA, Toolkit for Multivariate Data Analysis with Root*, arXiv physics/0703039, , <http://tmva.sourceforge.net/docu/TMVAUsersGuide.pdf>.
- [128] Derue F., Kaczmarska A., *Soft electron b-tagging with CSC data*, 2008, Atlas CSC BT-4 note in preparation.
- [129] Facius K., *Discovery Potential in the ATLAS Experiment for a Light Higgs Boson in the Channel $t\bar{t}$, $h \rightarrow b\bar{b}$ and the Influence of Electron b-tagging*, Master's thesis, Niels Bohr Institute (Copenhagen), 2005,
http://hep.nbi.dk/theses_files/master/katrine_facius.ps.gz.
- [130] Bordoni S., *Selezione di eventi $t\bar{t}$ nell'esperienza Atlas presso l'LHC*, Stage, Tesi di Laurea Triennale, 2005,
http://lpnhe-atlas.in2p3.fr/TheseStage/Stages/documents/rapport_Stefania.pdf.
- [131] A. Cimmarusti, *Etudes du quark top dans Atlas*, Stage, Maitrise de l'universit des Andes, Merida, Venezuela (HELEN), 2006,
http://lpnhe-atlas.in2p3.fr/TheseStage/Stages/documents/stage_Cimmarusti-2006.pdf.
- [132] Beuf G., *Etudes du quark top dans Atlas*, Stage, Master 2 de physique théorique, 2006,
<http://lpnhe-atlas.in2p3.fr/TheseStage/Stages/documents/>.
- [133] Biteau J., *Sélection du signal $t\bar{t}\bar{b}$ en dileptons avec le détecteur ATLAS au Cern*, Stage, L3 -Phytem Cachan, 2007,
http://lpnhe-atlas.in2p3.fr/TheseStage/Stages/documents/rapport_de_stage_BiteauJonathan.pdf.
- [134] Cavalleri P., *Mesure de la masse du quark top dans le canal en dilepton dans Atlas*, Ph.D. thesis, Université Pierre et Marie Curie, 2008, en préparation.
- [135] Etievre A-I., Meyer J-P., Schwindling J., *Top quark measurement in the lepton plus jets channel using full simulation*, 2005, ATL-PHYS-INT-2005-022.

- [136] S. Frixione, P. Nason and B.R. Webber, *Matching NLO QCD computations and parton showers in heavy flavour production*, JHEP (2003) 007.
- [137] CTEQ Parton Distribution Functions : version 6,
<http://user.pa.msu.edu/wkt/cteq/cteq6/cteq6pdf.html>.
- [138] G. Corcella et al., *HERWIG 6 : An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **0101** (2001) 010.
- [139] ATLAS Muon Saclay group,
<http://atlas-samusog.web.cern.ch/atlas-samusog/muonboy/Muonboy.htm>.
- [140] Verkerke V., van Vulpen I., *Commissioning ATLAS using top quark production*, 2007, ATL-COM-PHYS-2007-023.
- [141] Cogneras E., *Production de paires de Top et effet de Nouvelle Physique. Calibration des jets légers avec le processus W en jet-jet. Mesure de la masse du Top*, Ph.D. thesis, Université Blaise Pascal, Clermont-Ferrand, 2007.
- [142] B. Di Girolamo et al., *Beamline instrumentation in the 2004 combined ATLAS testbeam*, 2005, ATL-TECH-PUB-2005-001.
- [143] Giangiobbe V., *Étude en faisceau-test de la réponse des calorimètres de l'expérience Atlas du LHC à des pions chargés, dénergie comprise entre 3 et 350 GeV.*, Ph.D. thesis, Université Blaise Pascal, 2006, PCCF T 0608.
- [144] T. Carli, Status of data/MC comparisons for low-energy muons, electrons and pions in barrel CTB and status of high energy pions, ATLAS Overview Week 2007, <http://indico.cern.ch/getFile.py/access?contribId=53&sessionId=16&resId=1&materialId=slides&confId=11266>.
- [145] Derue F., Kaczmarska A., *Reconstruction and identification of low energy electrons in CTB data*, 2007, Atlas note in prepration.
- [146] A. Camard, *Caractérisation et exploitation des performances du calorimètre électromagnétique d'Atlas : études des muons et mise à profit de la résolution en temps*, Ph.D. thesis, Université Pierre et Marie Curie, 2004.
- [147] M. Ridet, LArMuID update, LAr Commissioning Analysis of cosmic muons 25/10/2005, <http://indico.cern.ch/getFile.py/access?contribId=s1t1&resId=0&materialId=0&confId=a056793>.
- [148] T. Vu Ahn, Atlas physics workshop, June 2005, Electron identification performance in CTB at low energies, <http://indico.cern.ch/getFile.py/access?contribId=s8t5&sessionId=s8&resId=0&materialId=0&confId=a044738>.
- [149] Mehlhase, S ; Petersen, T C, *A probability based approach to PID in the TRT detector of ATLAS*, 2006, ATL-COM-INDET-2006-017.

- [150] Klinkby, E ; Kittelmann, T H, *A Study of TRT Noise in 2004 Test Beam Data*, 2006, ATL-COM-INDET-2006-015.
- [151] ATLAS Electromagnetic Barrel Calorimeter Collaboration (M. Aharrouche et al.), *Energy linearity and resolution of the ATLAS electromagnetic barrel calorimeter in an electron test-beam*, Nucl.Instrum.Meth **A568** (2006) 601–623.
- [152] C. Becavin, *Identification des électrons et réjection des jets dans Atlas*, Stage, Licence de Physique, Université Paris 7, 2004,
<http://lpnhe-atlas.in2p3.fr/TheseStage/Stages/documents/rapport-CBecavin-2004.ps.gz>.
- [153] E. Dahan, *Analyse des données de test en faisceau H8 de l'expérience Atlas*, Stage, DEUG MIME première année, Université Paris 6, 2005,
<http://lpnhe-atlas.in2p3.fr/TheseStage/Stages/documents/rapport-EvaDahan.ps>.
- [154] Amérique latine - Formation Académique,
http://ec.europa.eu/europeaid/projects/alfa/index_fr.htm.
- [155] High Energy Physics Latinamerican-European Network,
<http://www.roma1.infn.it/exp/helen/>.
- [156] A. Cournol, *Premières études avec les données d'Atlas : reconstruction du boson Z*, Stage, M1, Université Paris 6, 2006,
http://lpnhe-atlas.in2p3.fr/TheseStage/Stages/documents/stage_Cimmarusti-2006.pdf.
- [157] Les conférences NEPAL (Noyaux Et Particules Au Lycée),
http://www.in2p3.fr/physique_pour_tous/aulyce/nepal/index.htm.
- [158] Portail internet du groupe Atlas LPNHE, <http://lpnhe-atlas.in2p3.fr/>.
- [159] Derue F., *Recherche d'événements de microlentille gravitationnelle dans les bras spiraux de la galaxie avec EROS II*, Ph.D. thesis, Université Paris 11, 1999, LAL 99-14.
- [160] Derue F. et al. (EROS coll.), *Observation of microlensing towards the galactic spiral arms, EROS II 2 years survey*, Astronomy and Astrophysics **351** (1999) 87.
- [161] Derue F. et al. (EROS coll.), *Observation of microlensing towards the galactic spiral arms, EROS II 3 years survey*, Astronomy and Astrophysics **373** (2001) 126.
- [162] Derue F. et al. (EROS coll.), *Observations of periodic variable stars towards the galactic spiral arms by EROS II*, Astronomy and Astrophysics **389** (2002) 149.
- [163] Rahal Y., *Recherche de microlentilles gravitationnelles vers les bras spiraux de la Galaxie et spectroscopie de supernovae dans EROS II*, Ph.D. thesis, Université Paris 11, 2003, LAL 03-85.
- [164] Afonso C. et al. (EROS coll.), *Bulge microlensing optical depth from EROS 2 observations*, Astronomy and Astrophysics **404** (2003) 145.

- [165] Tisserand P. et al. (EROS coll.), *Limits on the Macho Content of the Galactic Halo from the EROS-2 Survey of the Magellanic Clouds*, *Astronomy and Astrophysics* **469** (2007) 387.
- [166] J.H. Christenson, J.W. Cronin, V.L. Fitch et R. Turlay, *Evidence for the 2π decay of the K_2^0 meson*, *Phys. Rev. Lett.* **13** (1964) 138–140.
- [167] J.R. Batley et al. et al. (NA48 Coll.), *A precision measurement of the direct CP violation in the decay of neutral kaons into two pions*, *Phys. Lett.* **B 544** (2002) 97–112.

Frédéric DERUE

Tél : 01.44.27.75.77 / Courriel : derue@lpnhe.in2p3.fr
Nationalité française
Né le 07/01/1974 (34 ans)



Expérience professionnelle (11 ans)

- 2002-** **Chargé de recherche de 1^{ère} classe** - collaboration Atlas (1900 physiciens)
Laboratoire de Physique Nucléaire et de Hautes Énergies de Paris
Etudes sur le quark top
Reconstruction des électrons
Développement d'un noyau de la grille de calcul
- 2001 - 2002** **Attaché Temporaire d'Enseignement et de Recherche** - collaboration LHCb (600 physiciens)
Centre de Physique des Particules de Marseille
Optimisation des paramètres du système de déclenchement à muons de niveau 0
- 1999 - 2001** **Service national et contrat post-doctoral** - collaboration NA48 (150 physiciens)
Service de Physique des Particules de Saclay (CEA/DSM/DAPNIA)
Responsable de la production de la simulation complète du détecteur
Comparaison des acceptances des détecteurs avec différentes simulations
- 1996 - 1999** **Thèse de doctorat** - collaboration EROS II (50 physiciens)
Laboratoire de l'Accélérateur Linéaire d'Orsay
Mise au point d'algorithmes de sélection et d'analyse des données
Interprétation des résultats dans le cadre d'un modèle numérique de la Galaxie

Cycle d'études (Bac+8)

- 1996 - 99 **Doctorat en physique expérimentale des particules** - Université Paris Sud
1993 - 96 Magistère de Physique Fondamentale - Université Paris Sud
1995 - 96 DEA de physique des particules (Champs, particules, matière) - Université Paris Sud

Informatique

Systèmes Linux, Windows
Langages Fortran, C/C++, HTML, shell
Logiciels LaTeX, Word, Excel, PAW
Simulation Monte Carlo pour des détecteurs —
et propagation de particules (Geant)
Calcul Grille de calcul EGEE/LCG

Encadrement

Enseignement 200 heures en DEUG et Master
Stages 9 stagiaires
Projet Responsable scientifique de
la grille de calcul au LPNHE Paris

Publications

Présentations 15 conférences et séminaires
Rédaction 1 thèse, 29 publications

Divers

Langues Anglais (bon niveau)
Internet Webmestre de trois sites



ACTIVITÉS D'ENSEIGNEMENT, D'ENCADREMENT ET DE DIFFUSION DE LA CONNAISSANCE

Enseignement

J'ai effectué des vacances à l'Université d'Évry val d'Essonne lors du second semestre 2001 en deuxième année de DEUG MIAS. J'ai assuré dix séances de Travaux Dirigés d'une heure trente chacune. Les Travaux Dirigés de "Vibration-Propagation" portaient essentiellement sur l'application pratique des oscillateurs harmoniques au cas de vibrations de molécules diatomiques. J'ai aussi participé à la surveillance de deux Devoirs Surveillés. Enfin, j'ai effectué dans la même section seize heures de Travaux Pratiques. Ceux-ci portaient à la fois sur les oscillateurs (couplés et de torsion) ainsi qu'en Mécanique des Fluides (tube Venturi).

J'ai travaillé durant un an à l'Université de la Méditerranée, au Centre de Physique des Particules de Marseille, en tant qu'Attaché Temporaire d'Enseignement et de Recherche. J'ai effectué une charge de 96 heures d'enseignement (équivalent TD) au niveau des premières et deuxièmes années de DEUG Sciences de la Vie. Au premier semestre j'ai effectué une charge de 32 heures dédiée à l'Optique ondulatoire pour les DEUG deuxième année avec des Travaux Pratiques et des Travaux Dirigés. Au second semestre j'ai effectué des TD et TP d'Optique Géométrique, de Mécanique des Fluides et de Bioélectricité pour les DEUG première année. Au total mon service d'enseignement a avoisiné les 120 heures (équivalent TD).

Depuis mon entrée au CNRS, j'ai effectué tout le mois de septembre 2005 des Travaux Pratiques de physique nucléaire au Master de deuxième année "Noyaux Particules Astroparticule et Cosmologie" pour l'Université Pierre et Marie Curie, pour un total de 86 heures.

Encadrement de stages

Depuis mon entrée au CNRS j'ai encadré plusieurs stages, pour l'essentiel des stages d'été :

- le stage de huit semaines de Olivier Ali à l'été 2003 [126] au niveau Licence de Physique PhiTEM de l'ENS Cachan et de l'Université Paris 6. Le stage a porté sur l'identification des électrons non isolés.
- le stage de huit semaines de Christophe Becavin à l'été 2004 [152], au niveau Licence de Physique de l'Université Paris 7. Le stage a porté sur l'identification des électrons isolés et sur le pouvoir de rejet des jets.

-
- le stage de six semaines d'Eva Dahan à l'été 2005 [153], au niveau première année de DEUG MIME, Université Paris 6. Le stage a porté sur l'identification des électrons dans les données du test en faisceau combiné 2004.
 - le stage de six mois d'Andrés Cimmarusti, de mai à novembre 2006 [131]. L'étudiant provenait de l'Université de Los Andes, Mérida au Venezuela. Le stage s'est inscrit dans le cadre du programme européen Alfa [154] Helen (High Energy Latino-American European Network) [155]. Je suis allé à la soutenance à l'Université de Los Andes. Le stage a porté sur l'identification des électrons isolés et sur la préparation des analyses visant à reconstruire la masse du quark top.
 - le stage de sept semaines d'Anne Cournol à l'été 2007 [156], au niveau M1 de l'Université Paris 6. Le stage a porté sur l'identification des muons et des électrons, ainsi que sur la reconstruction du boson Z^0 dans ses canaux de désintégrations leptoniques.

Dans le cadre de mes activités dans la mise en place de la ressource de la grille de calcul dans le projet Grif j'ai été le responsable de stage de deux boursiers Helen :

- stage de juin 2006 à mars 2007 (8 mois) de Victor Mendoza, ingénieur au Centre de Calcul National de Mérida (Venezuela) ;
- stage d'août 2007 à mars 2008 (6 mois) de Laudin Molina, étudiant de niveau maîtrise d'informatique, Université de Los Andes, Mérida (Venezuela).

Administration de projet

Je suis le représentant scientifique du LPNHE pour le projet de nœud de grille de calcul Grif qui a été présenté au paragraphe 1.5. A ce titre, je suis chargé d'informer et de coordonner les demandes des utilisateurs du laboratoire, ainsi que d'établir et de suivre les demandes budgétaires. Je travaille en collaboration étroite avec les ingénieurs informaticiens formant l'équipe technique du projet au laboratoire. A ce titre je suis aussi représentant du LPNHE pour la grille EGEE.

Rédaction de documents

J'ai participé, dans le groupe Atlas, à l'écriture de documents pour demander des postes ou des financements, tout particulièrement :

- demande auprès de l'Union Européenne pour une bourse Marie Curie, obtenue pour la période mars 2003-mars 2005, sur l'identification des électrons, l'étiquetage des jets de b , l'étalonnage du calorimètre électromagnétique et l'analyse des données du test en faisceau combiné 2004.
- deux demandes - non réussies - de postes en contrat à durée déterminée auprès de

l'Agence Nationale pour la Recherche pour la grille de calcul (2005) et pour les études sur le quark top (2006).

- la préparation en 2007 d'un Plan Pluri-Formation auprès de l'Université Pierre et Marie Curie pour le développement de la grille de calcul.

Diffusion de la connaissance

J'ai participé à différentes activités de diffusion de la connaissance scientifique :

- j'ai été responsable durant trois ans de l'organisation de la "réunion du vendredi" du LPNHE, forum hebdomadaire, d'une durée d'une heure, auquel l'ensemble du laboratoire est convié. Le début de la réunion est l'occasion pour la direction de faire part des nouvelles concernant le laboratoire, ainsi que de brèves interventions pour présenter les nouveaux arrivants ou l'état d'avancement d'un projet. Cette partie est suivie par un ou plusieurs exposés sur un thème de physique ou sur une réalisation technique, un compte-rendu de conférence ou le bilan d'activité d'un service.
- j'ai participé à des activités de communication, aussi bien au niveau d'Atlas, comme la participation à l'organisation d'ateliers (un atelier egamma de deux jours en 2004, et l'Overview Week d'Atlas qui s'est organisée à Paris en octobre 2005) ou à des manifestations plus orientées vers le grand public. J'ai participé en particulier à toutes les Fêtes de la Science, qui ont vu chaque année plusieurs centaines de personnes visiter le laboratoire. J'ai effectué quelques conférences Nepal [157] auprès de lycées d'Ile de France.
- je suis webmestre des sites "Atlas LPNHE" [158], "Atlas France" [14] et "Grif LPNHE" [60] contenant des informations à la fois de travail pour les groupes concernés mais aussi, dans la mesure du possible, à vocation pour un public plus étudiant.



LISTE DES PUBLICATIONS

Principales publications (avec comité de lecture)

- 1 *The ATLAS experiment at the CERN Large Hadron Collider*, ATLAS Coll. (1900 auteurs), en préparation
- 2 *A precise determination of the direct CP violation parameter $Re(\epsilon'/\epsilon)$* , Lai A., Marras D. et al (NA48 Coll.) (150 auteurs), Eur. Phys. J. C 22,231-254 (2001).
- 3 *Observations of periodic variable stars towards the galactic spiral arms by EROS II*, Derue F., Marquette J-B. et al (EROS Coll.) (49 auteurs), Astronomy & Astrophysics 389 :149,2002.
- 4 *Observations of microlensing towards the galactic spiral arms, EROS II, 3 years survey*, Derue F., Afonso C. et al (EROS Coll.) (45 auteurs), Astronomy & Astrophysics 373 :126,2001.
- 5 *Observations of microlensing towards the galactic spiral arms, EROS II, 2 years survey*, Derue F., Afonso C. et al (EROS Coll.) (45 auteurs), Astronomy & Astrophysics 351 :87,1999.

Liste complète des publications dans Eros II

- 1 *EROS 2 photometry of probable R Coronae Borealis stars in the Small Magellanic Cloud*, Tisserand P. et al (EROS Coll.), Astron. & Astrophys. 424 :245-252,2004.
- 2 *Type Ia supernova rate at redshift of 0.1*, Blanc G. et al. (EROS Coll.), Astron. & Astrophys. 423 :881,2004
- 3 *Bulge microlensing optical depth from EROS 2 observations*, C. Afonso et al. (EROS Coll.), Astron. & Astrophys.404 :145-156,2003
- 4 *EROS II proper motion survey : constraints on the halo white dwarfs*, Goldman B., Alard C. et al (EROS Coll.), Astronomy & Astrophysics 389 L :69,2002.
- 5 *Observations of periodic variable stars towards the galactic spiral arms by EROS II*, Derue F., Marquette J-B. et al (EROS Coll.), Astronomy & Astrophysics 389 :149,2002.
- 6 *Blue irregular variable stars in the Small Magellanic Cloud from EROS II*, Beaulieu, J.-P. et al. (EROS Coll.), Astronomy & Astrophysics 380 :168,2001.

-
- 7 *Photometric constraints on microlens spectroscopy of EROS-BLG-2000-05*, Afonso C. et al (EROS Coll), Astronomy & Astrophysics 378 :1014,2001
 - 8 *Observations of microlensing towards the galactic spiral arms, EROS II, 3 years survey*, Derue F., Afonso C. et al (EROS Coll.), Astronomy & Astrophysics 373 :126,2001.
 - 9 *Type Ia supernova rate at $z \sim 0.1$* , D.Hardin et al (EROS Coll), Astronomy & Astrophysics 362 :419,2000.
 - 10 *Not enough stellar Mass Machos in the Galactic Halo*, Lasserre T., Afonso C. et al. (EROS Coll.), Astronomy & Astrophysics 355 :L39, 2000.
 - 11 *Observations of microlensing towards the galactic spiral arms, EROS II, 2 years survey*, Derue F., Afonso C. et al (EROS Coll.), Astronomy & Astrophysics 351 :87,1999.
 - 12 *EROS 2 proper motion survey : a field brown dwarf and an L dwarf companion to LHS 102*, Goldman B., Delfosse X. et al. (EROS Coll.), Astronomy & Astrophysics 351 :L5,1999.
 - 13 *Combined analysis of the Binary-Lens Caustic-Crossing Event MACHO 98-SMC-1*, (The EROS, MACHO/GMAN, MPS, OGLE, and PLANET Collaborations), Astrophysical Journal 532 :340, 2000.
 - 14 *A slope variation in the PL relation for short period SMC Cepheids.*, Bauer F., Beaulieu J.-P. et al. (EROS Coll.), Astronomy & Astrophysics, 348 :175, 1999.
 - 15 *Microlensing towards the Small Magellanic Cloud, EROS II two year analysis.*, Afonso C., Alard C. et al. (EROS Coll ?), Astronomy & Astrophysics, 344 :L63, 1999.
 - 16 *EROS II intensive observation of the caustic crossing of microlensing event MACHO SMC-98-1.*, Afonso C. et al. (EROS Coll), Astronomy & Astrophysics, 337 :L17, 1998.
 - 17 *Microlensing towards the Small Magellanic Cloud, EROS II 1st year analysis.* Palanque-Delabrouille N. et al. (EROS Coll.), Astronomy & Astrophysics, 332 :1, 1997.

Liste complète des articles dans NA48

- 1 Measurement of the branching ratio and form factors for the decay $K(L) \rightarrow \pi^+ \pi^0 e^- + \nu(e)$ (anti- $\nu(e)$), J.R. Batley et al. (NA48 Coll.), Phys.Lett.B595 :75-85,2004
- 2 First observation of the $K(S) \rightarrow \pi^0 \gamma$ decay, A. Lai et al. (NA48 Coll.), Phys.Lett.B578 :276-284,2004
- 3 Search for CP violation in $K_0 \rightarrow 3\pi^0$ decay, A. Lai et al. (NA48 Coll.), Phys.Lett.B610 :165-176,2005
- 4 Precise measurements of the $K_S \rightarrow \gamma\gamma$ $K_L \rightarrow \gamma\gamma$ decay rates, A. Lai et al (NA48 Coll.), Phys.Lett.B551 :7-15,2003
- 5 A precision measurement of the direct CP violation in the decay of neutral kaons into two pions, J.R. Batley et al. et al. (NA48 Coll.), Phys. Lett., B 544 (2002) 97-112

-
- 6 A measurement of the KS lifetime, A. Lai et al. (NA48 Coll.), Physics Letters B537 (2002) 28-40.
 - 7 New measurement of the eta and K0 masses, A. Lai et al. (NA48 Coll.), Phys.Lett.B533 :196-206,200
 - 8 A precise measurement of the direct CP violation parameter $Re(\varepsilon'/\varepsilon)$ A. Lai, Marras D. et al. (NA48 Coll.), Eur. Phys. J. C 22,231-254 (2001).

Liste complète des articles dans LHCb

- 1 LHCb Inner Tracker Technical Design Report, LHCb TDR 008, CERN-LHCC-2002-029

Liste complète des articles dans Atlas

- 1 *The ATLAS experiment at the CERN Large Hadron Collider*, ATLAS Coll. (1900 auteurs), en préparation
- 2 ATLAS Computing Technical Design Report, ATLAS Coll., CERN-LHCC-2005-022
- 3 B Physics overview in ATLAS, F. Derue (ATLAS Coll.), ATL-PHYS-PUB-2007-022, inproc. QCD06 MontPellier
- 4 Electron and photon identification in ATLAS , F. Derue (ATLAS Coll.), ATL-PHYS-CONF-2005-015, inproc. HCP05, Les Diablerets (Suisse)

Présentations et compte-rendus de conférences

- 1 Décembre 2006, Cecalcula, Mérida (Venezuela) Computing grids for Atlas at the LHC
- 2 Décembre 2006, Reunion de Fisica de Altas Energias, FAE06, Caracas (Venezuela) Physics with first data in Atlas at the LHC
- 3 Juillet 2006, QCD06, Montpellier (France) Overview of B-physics in ATLAS
- 4 Juillet 2005, Hadron Collider Physics, Les Diablerets 2005 (Suisse) Electron and photon identification in ATLAS
- 5 Mars 2002, Les Rencontres de physique de la vall d'Aoste, La Thuile 2002 (Italie), New measurement of e'/e by the NA48 experiment at CERN
- 6 Mars 2000, Lens 2000 Moriond, Les Arcs (France), Results of the EROS II collaboration towards the Galactic spiral Arms
- 7 Janvier 1998, Fourth international workshop on gravitational microlensing, Paris (France), Preliminary results of EROS II towards the spiral arms

Présentations en séminaires et ateliers

- 1 Mai 2003, Atlas Physics Workshop Athènes (Grèce), e/γ identification (low / high

-
- / very high pT)
- 2 Février 2003, Institute of Nuclear Physics, Cracovie (Pologne), A status of the Atlas calorimetry
 - 3 Avril 2002, Laboratoire de Physique des Particules d'Annecy (France), Le trigger à muons de niveau 0 de LHCb
 - 4 Avril 2002, Laboratoire de Physique Nucléaire et de Hautes Energies de Paris (France), Le trigger à muons de niveau 0 de LHCb
 - 5 Juin 2001, Séminaire interne SPP, Saclay (France), Nouvelles mesures de la violation directe de CP dans le système des kaons neutres par les expériences NA48 et KTeV
 - 6 Mai 2001, Laboratoire de Physique Corpusculaire de Clermont-Ferrand (France), La violation de CP à travers les résultats des expériences NA48 et KTeV
 - 7 Mai 2001, Centre de Physique des Particules de Marseille (France), La violation de CP à travers les résultats des expériences NA48 et KTeV
 - 8 Avril 2001, Laboratoire de Physique des Particules d'Annecy (France), La violation de CP à travers les résultats des expériences NA48 et KTeV
 - 9 Avril 2001, Laboratoire de Physique Nucléaire et des Hautes Energies de Paris (France), La violation de CP à travers les résultats des expériences NA48 et KTeV
 - 10 Décembre 1998, Séminaire interne LAL, Orsay, Recherche de déviations aux effets de microlentille ordinaires dans les candidats trouvés par EROS II en direction des bras spiraux

Notes internes Eros II

- 1 Catalogue d'étoiles variables vers les bras spiraux., Derue F., Sep 99.
- 2 Le candidat de grande amplification vers γ Sct., Derue F., Sep 99.
- 3 Aide à l'utilisation des programmes d'analyse vers les Bras Spiraux, Derue F., Mai 99.
- 4 Aide à l'utilisation du programme de simulation de la Galaxie, Derue F., Mai 99.
- 5 Etat des travaux effectués pour l'estimation de la distance des étoiles vers les Bras Spiraux., Derue F., Oct 98.
- 6 Ajustements d'effets de lentilles gravitationnelles sur les candidats Bras Spiraux., Derue F. & Moniez M., Oct 98.
- 7 Analyse des 2 premières saisons vers les bras spiraux., Derue F., Moniez M., Sep 98.
- 8 Estimation la ligne de base d'une courbe de lumière., Derue F., Moniez M., Oct 97.

Notes internes NA48

- 1 Influence of drift chamber inefficiencies on the double ratio for 1998/99 data based on NASIM030., Derue F., NA48 note 01-11.

-
- 2 Acceptance correction for 1998 and 1999 data with NMC and correlated NASIM events., Derue F., NA48 note 01-08.
 - 3 NASIM User's Guide, De Beer M. & Derue F., NA48 note 00-23.

Notes internes LHCb

- 1 Performance of the muon trigger with a realistic simulation, Aslanides E., Derue F., Le Gac R. et al., 2002, LHCb-2002-041
- 2 Specification of the muon trigger processing board, Aslanides E., Cachemite J-P et al., 2002, LHCb-2002-003

Notes internes Atlas

- 1 Investigating the possibility of the top quark mass measurement with the matrix element method in the dilepton final state with the ATLAS detector at the LHC, Cavalleri, P; Derue, F; Lacour, D; Nikolic-Audit, I; Ridet, M; Trincaz-Duvoid, S, ATL-COM-PHYS-2007-080
- 2 Search for a SM Higgs decaying to two photons, Bettinelli M. et al., ATL-PHYS-PUB-2007-013
- 3 Writing B physics analysis code in Athena, Anastopoulos C. et al., ATL-COM-SOFT-2006-014
- 4 Identification and use of low pT electrons with Rome data Cochran J., Derue F., Kaczmarek A., Wolter M., ATL-PHYS-PUB-2006-026
- 5 Search for a Standard Model Higgs boson in the ATLAS experiment on the $H \rightarrow \gamma\gamma$ channel Escalier M., Derue F., Fayard L., Kado M., Laforge B., Reifen C., Unal G., ATL-PHYS-INT-2006-012
- 6 γ/π^0 separation in the 1st compartment of the ATLAS EM Calorimeter with DC2 data Reifen C., Derue F., Escalier M., Fayard L., Kado M., Laforge B., Unal G., ATL-PHYS-INT-2005-005
- 7 Photon/jet separation with DC1 data Escalier M., Derue F., Fayard L., Kado M., Laforge B., Reifen C., Unal G., ATL-PHYS-PUB-2005-018
- 8 Pile-up studies for soft electron identification and b-tagging with DC1 data, Bold T., Derue F., Kaczmarek A., Stanecka E., Wolter M., ATL-PHYS-PUB-2006-001
- 9 Reconstruction of DC1 J/Psi- γ ee decays and use for the intercalibration of the Atlas electromagnetic calorimeter, Derue F., Kaczmarek A., Schwemling Ph., ATL-PHYS-PUB-2006-004
- 10 Electron-jet separation with DC1 data, Derue F., Serfon C., ATL-PHYS-PUB-2005-016
- 11 Soft-electron identification and b-tagging with DC1 data, Derue F., Kaczmarek A., ATL-PHYS-2004-026



LA VIE AVANT ATLAS, L'UNIVERS, LE RESTE

Je propose ici un bref résumé de mes activités de recherche passées. En effet, avant Atlas, il y a eu les études sur la matière noire (l'Univers) et sur la violation de la symétrie CP (le reste).

Travail de thèse : la recherche de matière noire (LAL, 1996-1999)

Ma thèse de doctorat s'est déroulée de 1996 à 1999 au Laboratoire de l'Accélérateur Linéaire d'Orsay dans la collaboration Eros II [159]. Cette expérience a recherché des événements de microlentille gravitationnelle provoqués par le passage d'objets compacts et massifs devant des étoiles lumineuses. Ces objets étaient susceptibles d'expliquer la composante baryonique de la matière noire au niveau de notre galaxie. La probabilité d'observer un tel phénomène étant très faible, la collaboration Eros II a effectué chaque nuit un suivi photométrique de plusieurs millions d'étoiles dans différentes régions du ciel : le Petit et le Grand Nuage de Magellan, le centre galactique et quatre directions vers les bras spiraux de la galaxie. L'étude selon ces dernières directions devait permettre de donner une estimation de la quantité de matière dans le disque, dans une région éloignée de la structure en barre du centre galactique. Cette estimation est donnée en mesurant la probabilité pour qu'une source ponctuelle subisse, à un instant donné, un effet de microlentille avec une amplification supérieure à 34%¹⁾. Cette probabilité, appelée profondeur optique, peut alors être comparée à celle estimée grâce à une modélisation de la structure et de la distribution en masse dans la galaxie.

La thèse a porté sur l'analyse de trois ans de données, sur neuf millions d'étoiles vers les bras spiraux et a constitué la première recherche d'effets de microlentille vers ces lignes de visée. Les observations ont été effectuées depuis l'Observatoire Européen Austral de La Silla au Chili. J'ai eu la possibilité de participer durant trois mois aux observations faites sur le télescope Marly, ayant un miroir d'un mètre de diamètre et une caméra couvrant sur le ciel un champ de $0,7^\circ \times 1,4^\circ$. La réduction des données, représentant un demi téraoctet, pour créer les courbes de lumière, c'est à dire la distribution du flux de lumière reconstruit pour chaque étoile en fonction du temps, s'est échelonnée sur la moitié de la thèse.

Une analyse a été développée, basée sur la recherche des caractéristiques des événements de microlentille, à savoir une amplification du flux lumineux de l'étoile située en arrière plan

¹⁾ ou plus exactement à la position de la source projetée dans le plan de la lentille, c'est à dire son paramètre d'impact, inférieur à un rayon d'Einstein.

unique au cours de l'expérience, achromatique et compatible avec la forme théorique de l'effet. Elle a permis de sélectionner sept candidats d'une durée moyenne de cinquante jours. A cette occasion j'ai mis au point un filtre permettant de préselectionner les 10% d'étoiles les plus variables les plus intéressants, en implémentant différents algorithmes de sélection non spécifiques. Ceci a facilité par la suite les dernières étapes de l'analyse.

Une analyse détaillée des courbes de lumière des candidats a été effectuée comme la recherche d'un effet de parallaxe, c'est à dire la mise en évidence de la déformation de la courbe de lumière par la révolution de la Terre autour du Soleil. Ceci permet d'apporter des contraintes supplémentaires sur les caractéristiques physiques de la lentille. La non-observation d'une telle déviation sur un candidat ayant une amplification maximale ~ 27 et d'une durée caractéristique de 70 jours nous a permis d'obtenir un intervalle de confiance sur la masse de la lentille, exprimée en unité de masse solaire M_{\odot} : $0.84 > M/M_{\odot} > 2.7 \times 10^{-3}$ à 95% de niveau de confiance, typique d'une petite étoile du disque. La courbe d'amplification d'un autre candidat présentait une modulation par rapport à l'ajustement d'un effet de lentille ordinaire, qui peut s'interpréter de façon naturelle par un effet de lentille sur une source binaire, avec une période orbitale de 50 ± 3 jours, là aussi un effet attendu dans le disque.

Contrairement aux autres cibles étudiées par Eros II, nous n'avions que peu d'informations sur la distance de la population d'étoiles surveillée vers les bras spiraux. Pour affiner nos estimations par des méthodes indépendantes, j'ai mené différents travaux exploratoires dont un programme de recherche d'étoiles variables qui a donné un catalogue de plus de mille étoiles variables qui pourront permettre d'affiner l'estimation de la distance des populations d'étoiles suivies.

J'ai participé au développement d'une simulation de la galaxie avec les trois composantes habituelles : un bulbe central en forme de barre, un disque et un halo sombre. La composante qui contribue le plus ($\sim 90\%$) à la profondeur optique vers les bras spiraux est bien sur le disque. Le principal modèle envisagé ne présentait pas de matière noire dans le disque. La profondeur optique attendue, moyennée sur tous les champs était de 0.55×10^{-6} pour ce modèle. Les incertitudes systématiques étaient importantes. Pour une variation de la distance moyenne des étoiles sources de 30%, soit 2 kpc (kilo-parsecs), ce qui représentait une estimation assez conservatrice de notre incertitude, l'erreur sur la profondeur optique était de 50%. De même, une incertitude sur la densité de matière du disque de 50%, représentant les estimations les plus extrêmes trouvées alors dans la littérature, menait à une incertitude sur la profondeur optique de 50%. La profondeur optique mesurée, moyennée sur tous les champs observés, était $\tau_{GSA} = 0.45^{+0.24}_{-0.11} \times 10^{-6}$. Si nos mesures étaient compatibles avec le modèle d'une manière générale, la faible statistique à notre disposition à ce moment, ainsi que notre mauvaise connaissance de la distribution de distance des étoiles sources ne nous a pas permis d'aller plus en avant dans nos conclusions.

Cette thèse a donné lieu à l'écriture de sept notes internes et de trois articles, deux sur les résultats des effets de microlentilles [160]- [161] et un sur un catalogue d'étoiles variables [162].

Fin 2003 l'analyse finale des données dans la direction des bras spiraux a été effectuée dans le cadre d'une thèse [163] analysant seize millions de courbes de lumières prises sur une durée de sept années. Un total de vingt quatre candidats a été sélectionné. Sur les sept de ma thèse, seuls trois ont survécu. Deux candidats, qui étaient de faible amplification, ne satisfont plus - de peu - à certains critères. Deux autres ont présenté de nouvelles fluctuations quelques années après. Cela avait déjà été observé dans les directions des Nuages de Magellan et indique la présence d'un bruit de fond inconnu et non prévu au début des expériences, d'une classe d'étoiles variables dont les fluctuations ressemblent à celles d'un effet de microlentille. La profondeur optique mesurée est de $\tau_{GSA} = 0.38 \pm 0.08 \times 10^{-6}$, affinant considérablement la mesure effectuée en 1999. Ces résultats associés à ceux de la direction du centre galactique [164] sont compatibles avec les modèles de Galaxie sans matière noire. D'une manière générale les résultats de l'expérience Eros montrent que les Machos, dans la gamme $10^{-7} < M/M_0 < 5$, ne peuvent expliquer le problème de la matière noire dans la galaxie [165].

Études post-doctorales : la violation de CP

La différence de comportement entre la matière et l'antimatière joue un rôle clef dans l'évolution de l'univers. A l'instant de sa création, il était composé de matière et d'antimatière en égale quantité. Aujourd'hui l'antimatière a disparu. Pour comprendre cette évolution, un processus traitant différemment la matière et l'antimatière a dû apparaître dans la première microseconde.

Une différence de comportement entre la matière et l'antimatière a été observée dès 1964 [166]. Il s'agissait de la découverte de la violation de CP dans le système des mésons K neutres. Mais sa grandeur est trop faible pour expliquer la disparition de l'antimatière au cours de l'évolution de l'univers. Le Modèle Standard permet de rendre compte de la violation de CP à travers l'existence d'une phase qui apparaît dans la matrice de Cabibbo-Kobayashi-Maskawa, mais son origine est mystérieuse. Il prédit aussi des effets violant CP, importants dans les systèmes des mésons beaux. Par ailleurs, ce modèle théorique n'est encore pas assez contraint expérimentalement, en ce qui concerne la partie violation de CP, laissant la porte ouverte à d'autres sources de violation de CP qui pourraient trouver leurs origines dans les modèles supersymétriques.

L'expérience NA48 (Post-doc au CEA/Dapnia, 1999-2001)

J'ai travaillé deux années au CEA/Dapnia dans la collaboration NA48 de septembre 1999 à juillet 2001, la première année dans le cadre de mon service militaire, la deuxième année avec

un contrat post-doctoral.

L'expérience NA48 au Cern a cherché, entre autres choses, à déterminer l'amplitude de la violation directe de CP dans le système des kaons neutres. La recherche de la violation directe de CP se résume expérimentalement à mesurer une déviation relativement à l'unité du double rapport des taux de désintégrations :

$$R \equiv \frac{\Gamma(K_L \rightarrow \pi^0 \pi^0) / \Gamma(K_S \rightarrow \pi^0 \pi^0)}{\Gamma(K_L \rightarrow \pi^+ \pi^-) / \Gamma(K_S \rightarrow \pi^+ \pi^-)} = \left| \frac{\eta_{00}}{\eta_{+-}} \right|^2 \approx 1 - 6Re \left(\frac{\varepsilon'}{\varepsilon} \right). \quad (8.1)$$

La problématique expérimentale est donc le comptage précis de nombres d'événements dans quatre modes de désintégration. En effet, si $Re(\varepsilon'/\varepsilon)$ doit être connu à $2-3 \times 10^{-4}$ près, ce qui est le but de l'expérience, R et *a fortiori* chacun de ces nombres doivent l'être à mieux que le pour mille. L'expérience NA48 a utilisé une méthode permettant de minimiser la dépendance envers la simulation MonteCarlo. Pour cela, d'une part on n'utilisait que les désintégrations de K_L se produisant dans la région fiducielle peuplée par les désintégrations de K_S , d'autre part les événements K_L étaient pondérés par un poids fonction du temps propre τ . Ce poids était proportionnel au rapport K_S/K_L attendu au temps τ , les distributions en temps propre devenant alors similaires en K_L et en K_S . Les acceptances étant alors presque identiques, la précision systématique de la mesure n'était pas basée sur une simulation précise de l'expérience.

Pour corriger les différences résiduelles d'acceptance, l'expérience avait recours à une simulation Monte Carlo des quatre types de désintégrations. La réponse des détecteurs était simulée et les événements étaient soumis à la même chaîne de reconstruction et de sélection que les données réelles. Les valeurs de R mesurées dans chaque intervalle d'énergie devaient donc être corrigées par le double rapport des acceptances : $\Delta R = (29.0 \pm 4.1(stat) \pm 4.0(sys)) \times 10^{-4}$. En l'absence de pondération la correction globale serait de l'ordre de 2%, et pourrait atteindre plus de 10% dans certains intervalles. Cette correction était la plus importante appliquée à la valeur du double rapport R .

Deux simulations avaient été développées : la première, à Saclay, était basée sur Geant 3, la deuxième était une simulation rapide : certaines interactions dans la matière étaient négligées (interaction hadronique, rayons delta), les champs magnétiques étaient décrits de manière rudimentaire (application de kicks plutôt que la carte de champ), l'appareillage était décrit sommairement. Ce Monte Carlo était près de quatre fois plus rapide que la simulation détaillée. Seul ce Monte Carlo rapide était utilisé jusqu'alors pour estimer la correction d'acceptance. La simulation détaillée n'était elle utilisée principalement que pour l'étude des désintégrations rares des kaons (un autre volet de NA48) et des bruits de fond. Mon travail a consisté à comparer les corrections d'acceptance obtenues avec les deux simulations. Un important travail de production de données, ainsi que la partie débogage qui est allée de pair, m'ont accaparé pendant près de la moitié de mon temps. De manière à obtenir une très bonne précision statistique j'ai produit les mêmes événements que dans la simulation rapide. Les premiers résultats

préliminaires ont été issus de la production de 20 millions d'événements par mode sur une petite ferme de calcul à Saclay comprenant quelques processeurs, ce qui avait pris plus de trois mois. Les incertitudes statistiques étant bien trop importantes il a été nécessaire d'apprendre à produire massivement et rapidement. Pour 1998 j'ai produit 100 millions d'événements par mode en août puis en novembre 2000. Pour 1999 j'ai produit 200 millions d'événements par mode en décembre 2000. Ces productions ont été effectuées au Centre de Calcul de Lyon. La dernière production a été effectuée en une dizaine de jours, soit un gain en rapidité de près d'un facteur 100 par rapport aux premières études. Nous avons pu ainsi étudier leur devenir dans les deux simulations en utilisant la grande corrélation entre les deux lots. La différence des simples rapports d'acceptance des modes neutres pour 1999 était compatible avec zéro. En ce qui concerne les modes chargés le résultat obtenu était sensible à des effets fins de résolution sur l'énergie et le vertex de désintégration des kaons. L'effet sur R restant cependant comparable aux autres erreurs systématiques il a été décidé d'appliquer une correction à la correction d'acceptance égale à la moitié de la différence obtenue dans le mode chargé en 1998 et 1999 : $\Delta R = (-2.3 \pm 2.3) \times 10^{-4}$, le résultat dans le mode neutre a été considéré comme une vérification du résultat obtenu avec la simulation rapide. Ce travail a donc permis d'une part de confirmer la faible sensibilité à une simulation détaillée du détecteur et a amené à une correction faible de la correction d'acceptance.

J'ai aussi étudié l'effet des inefficacités des chambres à fils sur les valeurs de R obtenues en 1998 et 1999. Les inefficacités des fils ont été estimées à Saclay sur les données et mises dans la simulation détaillée fil par fil (en divisant la prise de données en 15 blocs). J'ai simulé cinq millions d'événements dans chacun des modes chargés en 1998 et 1999 pour des valeurs nominales et maximales des inefficacités. L'effet obtenu est compatible avec zéro.

Ces travaux ont été inclus dans la dernière publication de NA48 donnant une valeur du paramètre $Re(\epsilon'/\epsilon)$ [167]

$$Re(\epsilon'/\epsilon) = (15.3 \pm 2.6) \times 10^{-4}, \quad (8.2)$$

qui prouve à plus de 5σ l'existence de la violation directe de CP.

L'expérience LHCb (ATER au CPPM, 2001-2002)

J'ai travaillé une année au Centre de Physique des Particules de Marseille en tant qu'Attaché Temporaire d'Enseignement et de Recherche de septembre 2001 à juillet 2002.

L'étude de la violation de CP et des désintégrations rares dans le secteur de la beauté doit permettre de tester le Modèle Standard comme étant la formulation théorique de la violation de CP et, peut-être, de mettre en évidence des processus nouveaux. C'est à ce programme que s'attaque la collaboration LHCb. Au LHC, les B ont tendance à être produits à petit angle par rapport à l'axe du faisceau. L'expérience LHCb consiste en un spectromètre avant à un

bras dont l'acceptance angulaire va de 10 mrad à 300 mrad. Il est constitué d'un détecteur de vertex en silicium, d'un système de reconstruction des traces chargées partiellement placé dans un aimant, d'un RICH et d'aérogel pour l'identification des particules, de deux calorimètres, l'un électromagnétique avec un détecteur de pied de gerbe et l'autre hadronique et enfin d'un détecteur de μ .

Le groupe du Centre de Physique des Particules de Marseille est chargé du développement du système de déclenchement à muons de niveau zéro. Le groupe a aussi développé une application orientée objet qui simule le système de déclenchement à muons et des outils d'évaluation des performances de ce système dans les canaux de physique les plus prometteurs. C'est sur ce dernier aspect qu'a porté ma contribution soit l'évaluation des performances en fonction de différents paramètres : caractéristiques du détecteur à muons ; propriétés du halo de muons générés par la machine ; algorithmes et détails de l'implantation ; avec bruit de fond optimal et maximal. J'ai participé activement à la production et à l'analyse des miniDST nécessaires à nos études, représentant plus de 1 téra-octet de données.

Les événements de signal considérés sont des B se désintégrant semileptoniquement en muon, des b en $J/\Psi K_s$ ou des B en $\mu^+ \mu^-$. Le bruit de fond est constitué de collisions proton-proton inélastiques génériques appelées "minimum bias". Ces événements sont injectés dans la simulation du système de déclenchement. Celui-ci possède une bande passante d'environ 200 kHz qui correspond à un taux de rétention d'événements minimum bias de 2%. Ce taux étant fixe, on détermine la valeur de la coupure appliquée sur l'impulsion transverse et la taille des fenêtres de recherche des candidats. On mesure alors le nombre d'événements intéressants sélectionnés et donc l'efficacité du système. J'ai travaillé sur l'estimation de l'effet des limitations du nombre de candidat pouvant être traité à la fois. Environ 3% des événements sélectionnés présentent un tel surnombre. J'ai ainsi pu montrer que ces effets dans le système de déclenchement à muons sont largement dominés par des effets de diaphonie dans les chambres à muons. J'ai mis au point un algorithme permettant d'éliminer ces traces fantômes qui a ensuite été validé au niveau de l'implantation hardware.

Depuis, le groupe du CPPM a continué la mise au point du système de déclenchement qui a abouti aux différents modules et cartes nécessaires. Il est aussi impliqué dans des analyses sur la violation de CP et l'étiquetage des mésons B .

RÉSUMÉ SUR L'ORIGINALITÉ DES RECHERCHES

L'origine de la masse des particules élémentaires est liée au mécanisme de brisure de la symétrie électrofaible. Dans le Modèle Standard celui-ci prédit l'existence d'une nouvelle particule, le boson de Higgs. L'étude de cette brisure sera l'un des enjeux majeurs de l'expérience Atlas auprès du Large Hadron Collider, au Cern à partir de 2008. Dans la plupart des cas, les recherches seront limitées par notre connaissance des performances du détecteur, telles que la précision avec laquelle l'énergie des particules est reconstruite ou l'efficacité avec laquelle elles sont identifiées. Mes études ont porté sur la reconstruction et l'identification des électrons avec des données réelles et de simulation. La production d'électrons isolés est caractéristique de processus faisant intervenir la production ou l'échange des bosons de jauge W^\pm et Z^0 . Ces processus constituent les principaux bruits de fonds de la recherche du boson de Higgs et d'une nouvelle physique dans les voies leptoniques. Les électrons issus du Z^0 ou à plus basse énergie du J/ψ pourront de plus permettre d'étudier les performances du détecteur. Par ailleurs, les électrons dans les jets signent leurs saveurs comme provenant de l'hadronisation de quarks b . Le groupe Atlas du LPNHE développe une mesure de la masse du quark top avec les premières données du LHC qui pourra bénéficier de ces études. Il est le seul fermion élémentaire ayant une masse de l'ordre de l'échelle électrofaible. Il se trouve ainsi être un secteur privilégié pour l'étude de cette brisure afin de contraindre la masse du boson de Higgs, voire de mettre en évidence de possibles effets non prévus par le Modèle Standard. Enfin, l'analyse des données d'Atlas nécessitera l'utilisation de ressources de calcul et de stockage importantes nécessitant le développement d'une grille de calcul mondiale dont un des nœuds est développé au laboratoire.



PERSPECTIVES DES ACTIVITÉS DE RECHERCHE

La physique des particules s'attache à développer notre connaissance des quarks (les constituants des neutrons et des protons), des leptons (comme l'électron) et des interactions qui les gouvernent. Ces particules sont considérées, à ce jour, comme les constituants élémentaires de la matière. Les avancées, étroitement liées, de la théorie et des observations expérimentales, ont conduit à ce que l'on appelle le "Modèle Standard". Ce modèle décrit la vision que nous avons de l'univers et repose sur douze constituants élémentaires, répartis en trois familles. Le Modèle Standard explique trois des quatre interactions fondamentales, à l'exception de la gravitation. L'interaction électrofaible est le résultat de l'unification de l'interaction électromagnétique et de l'interaction faible, véhiculées respectivement par le photon et les bosons vecteurs W^\pm et Z^0 . Les gluons véhiculent l'interaction forte qui est responsable de la cohésion des noyaux. Ce modèle n'a pas été mis en défaut par les observations expérimentales. Cependant, nous savons qu'il ne constitue pas, à lui seul, la théorie ultime - de la physique des particules - puisque, dans sa forme actuelle, certaines questions importantes restent sans réponse :

- Comment pouvons-nous unifier l'interaction électrofaible avec l'interaction forte et l'interaction gravitationnelle ? Le nombre élevé de paramètres fondamentaux est également un indice qui tend à faire penser que le Modèle Standard n'est qu'une partie d'un modèle plus vaste, dont la Super-Symétrie est l'une des extensions les plus populaires. Elle prédit, entre autres choses, l'existence d'une cohorte de nouvelles particules dites supersymétriques, partenaires de celles formant la matière ordinaire. Elles sont des candidats intéressants pour expliquer la matière noire, sujet auquel je me suis confronté durant ma thèse de doctorat. Cette unification pourrait aussi conduire à la découverte de dimensions cachées, bouclées sur elles-mêmes.
- Comment expliquer la disparition de l'antimatière alors qu'à sa naissance l'univers contenait autant de matière que d'antimatière ? Ce mystère pourrait trouver sa solution grâce à l'étude de la violation des symétries discrètes CP de conjugaison de charge et de parité. J'ai abordé ces études durant mes années d'études post-doctorales. L'amplitude de la violation de CP observée jusqu'à ce jour dans les systèmes des kaons et des mésons beaux neutres ne permet cependant pas d'expliquer cette disparition.
- Pourquoi - et comment - les particules élémentaires ont-elles une masse ? Et pourquoi ces masses sont-elles si différentes ? allant de quelques eV, ou même moins pour les neutrinos,

à quelques centaines de GeV. Le mécanisme de Higgs pourrait répondre partiellement à ces questions, en permettant d'expliquer comment ces particules acquièrent une masse. Une conséquence remarquable est qu'il prédit l'existence d'une particule scalaire neutre : le boson de Higgs. Cette particule est à ce jour le dernier élément manquant du Modèle Standard.

En 2008, les études sur le Modèle Standard vont connaître un regain d'activité avec le démarrage du collisionneur de protons LHC (Large Hadron Collider), au Cern près de Genève. Il offrira par rapport aux machines hadroniques actuelles un gain d'un ordre de magnitude en énergie et de deux ordres de grandeur dans le taux de collisions proton-proton. Il y a ainsi de grands espoirs d'améliorer notre connaissance du Modèle Standard et de découvrir le boson de Higgs, dont la masse serait comprise entre $114 \text{ GeV}/c^2$ (limite expérimentale) et 1 TeV (limite théorique) c'est à dire dans la gamme d'énergie accessible au LHC. On peut, de plus, espérer voir apparaître de la nouvelle physique, dite au-delà du Modèle Standard. Ces sujets de physique seront l'enjeu de l'expérience Atlas.

Le détecteur Atlas est l'un des plus grands et plus complexes jamais élaboré. Il a été conçu en prenant en compte un large éventail d'études physiques qui ont amenées à la nécessité d'une trajectographie efficace, d'une excellente calorimétrie et d'un système de reconstruction performant des muons. L'observation de nouvelles particules s'effectuera par la reconstruction des produits de désintégration (électrons, photons, muons, jets...) ainsi que la conservation de l'énergie. L'observation indirecte s'effectuera à travers l'observation d'anomalies dans les distributions de quantités physiques mesurées.

Le groupe Atlas du LPNHE est constitué de onze physiciens permanents dont sept enseignants-chercheurs et quatre chercheurs au CNRS. Le groupe accueille aussi quatre étudiants en thèse. L'équipe technique est composé d'une dizaine d'ingénieurs et techniciens des services d'électronique, de mécanique et d'informatique. Le groupe est impliqué dans la construction du calorimètre électromagnétique d'Atlas. Celui-ci joue un rôle central en physique des hautes énergies : il permet le déclenchement du système d'acquisition, la reconstruction et l'identification de particules neutres ou chargées comme les photons, les π^0 ou les électrons, ainsi que de particules interagissant peu avec la matière à travers la mesure d'une énergie manquante. Le laboratoire a aussi participé aux tests en faisceau de ce détecteur. Depuis 2005 le laboratoire est aussi impliqué dans la mise au point d'un nœud de la grille de calcul en Ile de France, c'est à dire d'une ressource de calcul et de stockage, en collaboration avec d'autres instituts. Le laboratoire participe aussi au développement du logiciel orienté objet pour la reconstruction et l'analyse des données. Les analyses de physique en préparation comprennent l'étude du phénomène de brisure spontanée de la symétrie électrofaible avec la recherche du boson de Higgs dans son canal de désintégration en deux photons, mais aussi à travers l'étude les mécanismes de production et de propagation des bosons W . Le groupe participe aussi à la

recherche de neutrinos lourds. Enfin, un groupe a débuté une analyse pour mesurer la masse du quark top.

Le quark top est le seul fermion élémentaire ayant une masse de l'ordre de l'échelle électrofaible. Il se trouve ainsi être un secteur privilégié pour l'étude de la brisure de la symétrie électrofaible. Les motivations pour une détermination précise de la masse du quark top sont nombreuses. Il s'agit d'un paramètre fondamental du Modèle Standard et doit donc être mesurée avec la plus grande précision possible. Une mesure précise de la masse du top (de l'ordre de 2 GeV) devrait permettre d'effectuer des tests de cohérence interne rigoureux du modèle et de contraindre certains paramètres du Modèle comme la masse du boson de Higgs. De même, les modèles au-delà du Modèle Standard visant à expliquer de manière plus fondamentale l'origine et la hiérarchie en masse observée des fermions peuvent être bien étudiées si on connaît la masse du quark top avec une précision de l'ordre de 1 GeV. Le quark top est une particule instable qui se désintègre spontanément, et quasi-exclusivement, en quarks b et en bosons W qui forment alors des leptons et des jets de particules stables. La préparation des analyses de physique se basent sur l'expertise des laboratoires sur la reconstruction et l'identification de ces particules. Après la prise de données, l'interprétation des résultats s'effectuera dans le cadre de modèles théoriques.

Dans la plupart des cas, la précision sur la mesure des propriétés des particules étudiées sera limitée par notre connaissance des performances du détecteur Atlas, telles que la précision avec laquelle l'énergie des produits de désintégrations est reconstruite ou l'efficacité avec laquelle ces produits de désintégration sont identifiés. Cette connaissance passe à la fois par l'expertise déjà acquise lors de la construction des détecteurs, de leurs tests en faisceaux, mais aussi par le suivi précis de leurs performances avec les premières données du LHC. Les paires de quarks top-anti-top vont aussi fournir des outils importants pour la collaborations Atlas, au démarrage de la prise de données, pour mieux comprendre le détecteur. Ainsi, ce lot de données permettra d'étudier l'étalonnage de l'énergie des jets ou l'étiquetage des jets de b .

Au démarrage du LHC, le quark top sera observé pour la première fois en Europe, avec la possibilité d'étudier ses propriétés avec un luxe de détails, puisque le LHC produira en un jour dix fois plus de quarks top produits en un an au Tevatron, le collisionneur proton-anti-proton, situé près de Chicago, où cette particule a été découverte en 1995. Les perspectives de mes activités de recherche sont liées aux études sur le quark top, en particulier à la mesure de la masse du quark top dans le canal en dileptons. Ce canal, où les deux W se désintègrent semileptoniquement, a un rapport d'embranchement faible (environ 5%), mais aussi un bruit de fond peu important, essentiellement dû au Z . Ici la présence des deux neutrinos dans l'état final rend la mesure de la masse invariante plus difficile. Le groupe du LPNHE a entamé voici près de deux ans un programme de recherche visant à mesurer, avec les premières données du LHC, la masse du quark top dans le canal en dileptons. Pour une luminosité intégrée de 10 fb^{-1} , corres-

pendant à un an de prise de données à basse luminosité, environ 400000 événements dilepton sont attendus, donnant une erreur statistique sur la masse du top de 1 GeV. Les incertitudes systématiques sont elles de l'ordre de 2 GeV.

J'envisage de travailler sur le quark top, particulièrement dans le canal en di-électrons. Cette étude bénéficiera de la chaîne de sélection des événements, de l'analyse et du développement de la méthode de mesure de la masse du quark top qui est actuellement en cours de développement au LPNHE dans le canal en "muon plus électron". Au niveau des outils, un effort particulier sera porté sur la reconstruction des électrons, en particulier sur leur identification. Celle-ci pourra être étudiée avec des événements $Z^0 \rightarrow e^+e^-$. Le bruit de fond aux études sur le quark top proviendra en partie des canaux $Z^0 \rightarrow e^+e^-$ et $Z^0 \rightarrow \tau^+\tau^-$ et devra être étudié. Je compte aussi poursuivre mes études sur l'étalonnage du calorimètre électromagnétique, particulièrement à basse énergie avec le $J\psi \rightarrow e^+e^-$. Enfin, l'étiquetage des quarks b sera étudié, particulièrement avec des électrons de basses impulsions transverses. Ces études seront particulièrement intéressantes et importantes au démarrage de la prise de données.

J'envisage de diriger une thèse sur la mesure de la masse du quark top dans le canal en di-électrons avec les premières données de l'expérience Atlas. Le démarrage du LHC sera ainsi une opportunité unique de participer aux premières prises de données de l'une des plus grandes ressources mises en activité pour la physique des hautes énergies ces dix dernières années.

REMERCIEMENTS

Je remercie MM. Daniel Froidevaux, François Le Diberder et Yves Sirois d'avoir accepté d'être rapporteurs de cette habilitation. Je remercie M. Philippe Schwemling d'avoir accepté de participer à ce jury. Enfin, je remercie M. Pascal Debu, directeur du LPNHE, d'avoir accepté la présidence de ce jury.

Je tiens à remercier M. Jean-Eudes Augustin, directeur du LPNHE au moment de mon entrée au CNRS, pour m'avoir accueilli dans son laboratoire (au moment des "négociations" post-concours ce fut sans doute le coup de téléphone le plus court et le plus fructueux que j'ai pu avoir). J'aimerais saluer l'ensemble des membres du LPNHE Paris pour ces déjà quelques années passées au sein de ce beau campus de Jussieu, mais ne tiens pas à remercier monsieur l'architecte pour tous ces matins d'hiver pluvieux où l'on marche avec précautions sur le parvis très glissant, avec pour seul bruit les croassements des corbeaux qui tournent autour de la tour centrale (les volatiles ont en fait disparus depuis quelques mois, chassés par les travaux...). Blague à part, une fois qu'on a arrêté de se perdre dans les couloirs, il y a plein de personnes super qui font qu'on est content de venir au labo.

Je remercie bien évidemment les membres du groupe Atlas du LPNHE avec qui je travaille tous les jours. Le groupe s'est considérablement renforcé d'années en années, nous avons appris à travailler ensemble, et je gage que l'approche du démarrage du LHC est loin d'être la seule raison de cette croissance mais que nous y contribuons aussi. J'ai tout de même une pensée émue pour celles et ceux qui auront à leur tour la joie de préparer leur habilitation et leur confirme qu'effectivement ça prend beaucoup de temps.

Je remercie les membres du projet Grif (ceux du LPNHE, mais aussi les autres !) et du CAF (Calcul Atlas France) qui m'ont permis de survivre dans ce milieu hostile qu'est le calcul et la grille (promis j'arrête le batch à Lyon très bientôt).

J'ai profité de l'expérience et des conseils de très nombreuses personnes de la collaboration Atlas, et si je ne les nommerai pas toutes (rappelons que la liste des signataires d'Atlas est un article à elle seule) j'aimerais tout de même remercier tout particulièrement David Rousseau et Dirk Zerwas pour leur immense aide.

Bien évidemment, je remercie Anna Kaczmarek à la fois pour tout le travail effectué ensemble (à peu près tout ce qui concerne les "soft électrons"), basé sur son expérience passée (vous trouverez de grandes similitudes entre le travail présenté dans ce manuscrit et de nom-

breux passages de sa thèse) et pour les nombreuses journées thématiques à Paris et Cracovie alliant la culture (musées, châteaux-forts et vestiges gallo-romains quand on les trouve) et l'accessoire (la bière s'appécie-t-elle mieux en écoutant du punk ou de la musique médiévale ? et des études approfondies des mérites respectifs du vin et de la vodka...).

J'ai une pensée aussi pour mes anciens camarades d'université et de DEA qui se retrouvent dans Atlas (aussi des fois dans d'autres manips) et qui donnent un côté "réunion des anciens combattants" au travail...

Pour finir, j'aimerais remercier tous ceux qui m'ont aidé, par leur soutien direct ou leurs conseils, à obtenir ce poste au CNRS et en tout premier lieu les membres des groupes Eros (LAL, SPP), NA48 (SPP, LAL) et LHCb (CPPM).



Résumé : L'origine de la masse des particules élémentaires est liée au mécanisme de brisure de la symétrie électrofaible. Son étude sera l'un des enjeux majeurs de l'expérience Atlas auprès du Large Hadron Collider, au Cern à partir de 2008. Dans la plupart des cas, les recherches seront limitées par notre connaissance des performances du détecteur, telles que la précision avec laquelle l'énergie des particules est reconstruite ou l'efficacité avec laquelle elles sont identifiées. Ce mémoire d'habilitation présente un travail portant sur la reconstruction des électrons dans Atlas avec des données simulées et des données prises durant le test en faisceau combiné qui s'est déroulé en 2004. L'analyse des données d'Atlas nécessite l'utilisation de ressources de calcul et de stockage importantes qui a impliqué le développement d'une grille de calcul mondiale dont un des nœuds est développé au laboratoire. Le manuscrit présente aussi l'effort effectué au LPNHE Paris pour la mise en place d'un Tier 2 en région Ile de France.

Mots-clefs : LHC, ATLAS, électrons, grille de calcul, Tier-2

Abstract The origin of the mass of elementary particles is linked to the electroweak symmetry breaking mechanism. Its study will be one of the main efforts of the ATLAS experiment at the Large Hadron Collider of CERN, starting in 2008. In most cases, studies will be limited by our knowledge of the detector performances, as the precision of the energy reconstruction or the efficiency to identify particles. This manuscript of habilitation presents a work dedicated to the reconstruction of electrons in the ATLAS experiment with simulated data and data taken during the combined test beam of 2004. The analysis of the ATLAS data implies the use of a huge amount of computing and storage resources which brought to the development of a world computing grid. This manuscript presents also the effort done at the LPNHE Paris for the setup of a Tier 2 in the grand Paris area.

Key-words : LHC, ATLAS, electrons, computing grid, Tier-2