



Etude de l'identification de leptons tau pour la recherche de supersymétrie dans le cadre de l'expérience CMS au LHC

David Bodin

► To cite this version:

David Bodin. Etude de l'identification de leptons tau pour la recherche de supersymétrie dans le cadre de l'expérience CMS au LHC. Physique Atomique [physics.atom-ph]. Université de Strasbourg, 2012. Français. NNT : 2012STRAE050 . tel-00979482

HAL Id: tel-00979482

<https://theses.hal.science/tel-00979482>

Submitted on 16 Apr 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Strasbourg

Institut Pluridisciplinaire Hubert Curien

Thèse en vue de l'obtention du diplôme de
docteur de l'université de Strasbourg
en physique des particules

présentée par

David BODIN

Étude de l'identification de leptons τ pour la
recherche de supersymétrie dans le cadre de
l'expérience CMS au LHC

Cette thèse a été soutenue le jeudi 15 mars 2012.
Devant la commission d'examen constituée de

M.	Claude CHARLOT	Rapporteur externe
Mme	Isabelle RIPP-BAUDOT	Examineur
M.	Dirk ZERWAS	Rapporteur externe
M.	Ulrich GOERLACH	Directeur de thèse

Résumé

Ce mémoire de thèse présente une étude de l'identification des candidats τ par différents algorithmes dans les collisions pp au LHC et impliquant différents processus physiques : électrofaibles, production de quarks top, SUSY et QCD.

Les leptons tau sont identifiés géométriquement dans la voie de désintégration hadronique car les jets issus de ces désintégrations sont généralement plus étroits et mieux isolés que les jets QCD. Une discrimination supplémentaire est appliquée en exigeant la présence d'un des modes de désintégration hadronique dominants du τ contenant un ou trois pions chargés et jusqu'à deux pions neutres. Les jets QCD produits en très grande quantité dans les collisions proton-proton du LHC constituent une source de bruit de fond importante pour les analyses impliquant l'identification de leptons τ . L'analyse des jets QCD produits dans le LHC pour les études de mauvaise identification porte sur 36 pb^{-1} de données enregistrées par le détecteur CMS pendant l'année 2010.

Les efficacités d'identification et les taux de mauvaise identification nous révèlent une forte dépendance en les caractéristiques de l'évènement étudié : les efficacités sont globalement plus basses dans les évènements présentant une activité hadronique importante comme les évènements SUSY et $t\bar{t}$ que dans les évènements QCD mous et électrofaibles.

Les effets de l'environnement sur l'identification de τ sont paramétrés à l'aide de deux variables globales : ΣE_T et la multiplicité de hadrons chargés. Les efficacités chutent typiquement de 80% entre les multiplicités de hadrons les plus faibles et le plus fortes.

Mots clés

Lepton tau, jet, QCD, identification de lepton tau, supersymétrie (SUSY), LHC, CMS.

Abstract

The work presented in this thesis deals with tau identified by different algorithms in pp collision at LHC and in various kind of physics processes : Electroweak, top quark production, SUSY and QCD.

Tau leptons are geometrically identified in their hadronic decay mode because the jets coming from those decays are generally more collimated and isolated than QCD jets. An additional discrimination is applied in asking the presence of one of the dominant τ decay modes containing one or three charged pions and up to two neutral pions. The QCD jets produced in large quantities in LHC proton-proton collisions constitute a large background source for analyses involving tau-ID. The analysis of QCD jets produced in LHC for fake-rate studies processed 36 pb^{-1} of data registered during year 2010.

Identification efficiencies and fake-rate distributions reveals a strong dependency on event type : distributions are globally lower in events having a large hadronic activity like SUSY and $t\bar{t}$ events than in electroweak and QCD events.

The environment effects of the collision on the identification of tau leptons can be parameterised by two global variables : ΣE_T and charged hadron multiplicity. Efficiencies drop down by 80% between the lowest and the highest charged hadron multiplicities.

Keywords

Tau lepton, jet, QCD, tau-ID, supersymmetry (SUSY), LHC, CMS.

Remerciements

*“Mieux vaut allumer une bougie que maudire les ténèbres”
- Lao Tseu -*

Il y a tellement de monde qui m’ont aidé, que ça soit par les conseils, l’orientation et toutes sortes de coups de mains ou qui ont tout simplement rendu ces années agréables pour moi que j’espère n’oublier personne.

Pour commencer je souhaite remercier les chefs, tout d’abord mon chef direct M. Ulrich GOERLACH sans qui cette thèse n’aurait pas eu lieu et qui a su trouver le temps malgré ses responsabilités de me prodiguer de précieux conseils tout au long de mes recherches. Je tiens à remercier le grand chef de la tribu CMS de Strasbourg M. Daniel BLOCH pour m’avoir accueilli au sein de son laboratoire et pour m’avoir fourni mon outil de travail principal : un bon ordinateur. Ensuite je remercie tous les autres membres de CMS Strasbourg, tout d’abord les gens encore présents et à qui je posais plein de questions : Anne-Catherine LE BIHAN, Jeremy ANDREA, Eric CONTE, Jean-Laurent AGRAM, Pierre VAN HOVE, Eric CHABERT, Denis GELE. Courage vous allez en faire des belles découvertes sur ce LHC, je sais que certains d’entre vous l’attendent depuis très longtemps. Et les gens de passage : tout d’abord Ludovic HOUCHU qui m’a énormément aidé pendant mon stage de master puis pour le début de ma thèse, j’espère que tes enfants vont bien, un “arigatou” pour Yoshinari MIKAMI aka. Yoshi : I hope you enjoyed France, Anne-Fleur BARFUSS avec qui je pouvais boire une bière en suisse pendant mes shifts au CERN, Christina FERRO qui nous apporte le soleil de l’Italie : courage pour la fin de la thèse.

Et il y a ceux qui ne sont pas du groupe CMS, je remercie Jean-Pierre FROBERGER pour son aide aux mots croisés... aux problèmes de réseau interne, Sylvianne MOLINET qui m’a beaucoup aidé pour les problèmes d’émulateur Linux, Cedric MULLER et Sébastien GEIGER pour tous les problèmes de bug Windows (j’en vois déjà qui rigolent). Je remercie aussi Sébastien GREDER dont la maîtrise de ROOT est telle que je ne comprends tes codes (mais ils fonctionnent très bien).

Pour l’international je remercie les gens qui ont pris la peine de m’aider au CERN (ainsi que par de longs échanges de mails). So thank you Simone GENNAI for your precious advises and for coming to my half way presentation, thank you Christian VEELKEN and Evan K. FRIIS for all your help when I had a rough time running your tauanalysis package.

Enfin il y a tous les autres thésards qui sont venus, partis ou qui sont encore présents et qui allaient affronter avec moi la foule d’étudiants au RU le midi, m’infliger de sévères défaites au tarot et avec qui j’ai passé de très bon moments lors des pauses dans ma vie de thésard : Marie (que je remercie pour son template latex), Antonin, Florian, Momo, Didier, Estelle, Xitzel, Gautier, Alain le délégué syndical.

Enfin quelques remerciements spéciaux de non-scientifiques, mais qui ont fait partie de ma vie de thésard et qui l’ont rendue plus agréable : mes collocataires Etienne MALINGREY

le gars d’la Meuse, et le “gamin” Olivier RETEUNAUER, merci de m’avoir supporté tout ce temps. Et merci aussi à tous ceux avec qui j’ai passé de bons moments a Strasbourg : Le petit salé, le duke, le Francky et tous les “boite-à-bougiens” : Laura, Gala, Lilou, Nathes, Sarah, François. Et merci aussi les rockers : Louis le basseux, Henri le chanteux et Julien le gratteux (les autres rockers ayant déjà été remerciés).

Cela allant de soi, un grand merci à ma famille qui à su rester proche malgré la distance et qui m’a soutenu psychologiquement et matériellement dans les moments difficiles.

Table des matières

Introduction	1
1 La physique des collisionneurs	3
1.1 La physique des particules expérimentale	3
1.2 Le Modèle Standard	4
1.2.1 Les particules élémentaires	4
1.2.2 Les interactions fondamentales	5
1.2.2.1 Les interactions dans le Modèle Standard	6
1.2.2.2 L'interaction forte	8
1.2.2.3 L'interaction électrofaible	9
1.3 Au-delà du Modèle Standard	11
1.3.1 Aspects de la supersymétrie	12
1.3.2 Recherche expérimentale de la supersymétrie	13
1.4 Les collisions proton-proton	17
1.4.1 La Chromodynamique Quantique dans les collisions pp	19
1.4.1.1 Modélisation de la gerbe partonique	22
1.4.1.2 Fragmentation et hadronisation	24
1.4.2 Les interactions multiples	25
2 L'expérience CMS	28
2.1 Le Grand Collisionneur de Hadrons (LHC)	28
2.2 Le Solénoïde Compact à Muons (CMS)	30
2.2.1 Référentiel utilisé pour CMS	31
2.2.2 Conditions expérimentales requises	32
2.2.3 Conception du détecteur CMS	33
2.2.3.1 Système de trajectométrie interne	33
2.2.3.2 Calorimètre électromagnétique	34
2.2.3.3 Calorimètres hadroniques	36
2.2.3.4 Détecteurs à muons	37
2.2.3.5 Mesure d'un événement par le détecteur CMS	37
2.2.3.6 Systèmes de déclenchement	39
2.3 Reconstruction des données	40
2.3.1 Environnement des analyses	40
2.3.2 Reconstruction des trajectoires	41
2.3.3 Format des données	43
2.3.4 L'algorithme Particle Flow	43
2.3.5 Reconstruction des jets	44

3	L'identification du lepton τ	50
3.1	Le lepton τ	50
3.1.1	Caractéristiques	50
3.1.2	Détection des leptons tau	52
3.2	Les sources de leptons τ	55
3.3	L'identification des jets de τ	57
3.3.1	Identification géométrique basée sur des coupures	57
3.3.2	Identification analysant les modes de désintégration du τ	58
3.3.3	Identification via des réseaux de neurones	61
3.3.4	La réjection des électrons et des muons	64
3.4	Estimation du bruit de fond QCD	65
3.5	Architecture PAT et modules associés pour l'identification des τ	66
3.5.1	Création des objets <i>PAT</i>	67
3.5.2	L'environnement modulaire TauAnalysis	68
4	Efficacité et taux de mauvaise identification du lepton τ	70
4.1	Conditions d'analyse	70
4.1.1	Description des échantillons	70
4.1.1.1	Échantillons simulés Monte Carlo	70
4.1.1.2	Échantillons de données du LHC	73
4.1.2	Définition de l'efficacité d'identification	75
4.1.3	Définition du taux de mauvaise identification de τ	76
4.2	Efficacité d'identification du lepton tau	78
4.2.1	Spectres des variables associées aux jets de tau	78
4.2.2	Efficacités calculées avec les quantités générées	79
4.2.2.1	Dépendance en P_T	79
4.2.2.2	Dépendance en η	81
4.2.3	Efficacités calculées avec les quantités reconstruites	83
4.2.3.1	Dépendance en P_T	83
4.2.4	Dépendance en η	85
4.2.5	Dépendance en R_{Jet}	87
4.3	Taux de mauvaise identification du τ	87
4.3.1	Spectres des variables associées aux jets QCD	89
4.3.2	Dépendance en P_T	90
4.3.3	Dépendance en η	95
4.3.4	Dépendance en R_{Jet}	99
4.3.5	Influence du seuil de déclenchement	102
4.4	Conclusions sur les efficacités et les taux de mauvaise identification	106
5	Etudes de l'impact de la topologie sur l'identification du lepton τ	107
5.1	Conditions d'analyse	107
5.1.1	Caractérisation de l'activité d'un évènement	107
5.1.2	Critères de présélection des candidats τ	108
5.1.3	Critères d'identification des candidats τ	109
5.2	Influence de l'environnement sur les efficacités d'identification	109
5.2.1	Spectres en variables globales	110
5.2.2	Corrélations entre variables caractérisant l'énergie globale	115
5.2.3	Corrélations entre les multiplicités d'objets	118
5.2.4	Influence de l'environnement sur l'identification basée sur les coupures	120

5.2.4.1	Influence des variables caractérisant l'énergie globale . . .	120
5.2.4.2	Influence des multiplicités d'objets	125
5.2.5	Influence de l'environnement sur l'identification Hadron Plus Strips	130
5.2.5.1	Influence des variables caractérisant l'énergie globale . . .	131
5.2.5.2	Influence des multiplicités d'objets	132
5.2.6	Influence de l'environnement sur l'identification Tau Neural Classifier	137
5.2.6.1	Influence des variables caractérisant l'énergie globale . . .	137
5.2.6.2	Influence des multiplicités d'objets	139
5.3	Influence de la topologie des événements sur le taux de mauvaise identification	142
5.3.1	Identification basée sur les coupures	142
5.3.2	Identification Hadron Plus Strips	145
5.3.3	Identification Tau Neural Classifier	148
Conclusion		150
A Estimation des bruits de fond par la méthode de “fake-rate”		153
A.1	Conditions d'analyse	153
A.2	Resultats sur les échantillons $Z\tau\tau$	153
A.3	Resultats sur les échantillons QCD	154
B Taux de mauvaise identification des jets QCD triés selon leur impulsion transverse		157
B.1	Spectres des taux de mauvaise identification pour le jet de plus haut P_T . .	157
B.2	Spectres des taux de mauvaise identification pour le second jet de plus haut P_T	158
B.3	Spectres des taux de mauvaise identification pour le troisième jet de plus haut P_T	159
C Taux de mauvaise identification de jets QCD avec et sans coupure en énergie transverse manquante		162
Bibliographie		166

Table des figures

1.1	(source [8]) Distributions de la densité de partons $x.f(x)$ (où $f = u_v, d_v, \bar{u}, \bar{d}, s, c, b, g$).	6
1.2	Interactions fondamentales (source [9]).	7
1.3	Représentation tridimensionnelle du potentiel de Higgs.	10
1.4	Couplages de jauge ($1/\alpha$) pour le Modèle Standard (à gauche) et le Modèle SuperSymétrique Minimal (à droite). Le point noir représente les corrections dépendantes du Modèle en fonction de Q (Source [16]).	14
1.5	Cascade supersymétrique dans le processus $pp \rightarrow \tilde{q}\tilde{g}$.	15
1.6	Section efficace de différents processus du Modèle Standard en fonction de l'énergie dans le centre de masse (extrait de [20]) et de la nature de la collision ($\bar{p}p$ pour le Tevatron et pp pour le LHC).	18
1.7	Schématisation simplifiée d'une collision $pp \rightarrow F$.	19
1.8	Schématisation d'une collision $pp \rightarrow F$ avec radiation dans l'état initial et final. F est un état final avec fragmentation des quarks.	20
1.9	a) Diffusion élastique entre deux quarks (voie t), b) Fluctuations partoniques dans un proton.	20
1.10	Sections efficaces de production inclusives relatives des bosons W/Z + n jets mesurés par CMS[23] en 2010 (désintégration en muon).	21
1.11	a) Boucle quantique, le gluon est émis et absorbé par le même quark. b) Une "jambe", un gluon est émis par le quark mais non réabsorbé.	23
1.12	Fragmentation puis hadronisation des quarks par le modèle en cordes.	24
1.13	Schéma de la sélection itérative des saveurs et de l'impulsion des quarks produits dans une fragmentation par le modèle en cordes de Lund.	26
1.14	Section efficace totale et section efficace par parton[28] en fonction de $p_{\perp\min}$ pour le Tevatron (collisions proton-antiproton à 2,36 TeV) et le LHC (collisions proton-proton à 14 TeV).	27
2.1	Schéma en coupe d'un cryodipole du LHC[29].	29
2.2	Quantité de données délivrées par le LHC et enregistrées par CMS (provenant de [31]) en a) 2010, b) 2011.	30
2.3	Description du système de coordonnées utilisées dans CMS.	32
2.4	Représentation en coupe du détecteur CMS.	34
2.5	Schéma du trajectomètre.	35
2.6	Paramétrage de la résolution en énergie du calorimètre électromagnétique de CMS à l'aide d'un faisceau test d'électrons. La courbe du dessus est obtenue pour tous les évènements dans une surface de $20 \times 20 \text{ mm}$, la courbe du dessous est obtenue avec une sélection d'évènements sur une surface de $4 \times 4 \text{ mm}$.	36
2.7	Coupe longitudinale du détecteur CMS.	38

2.8	Exemple d'un évènement reconstruit par le détecteur CMS au cours de l'année 2010 (provenant de [39]).	42
2.9	Schéma des paramètres d'impact transverse et longitudinal.	42
2.10	En haut : Cas de mauvaise reconstruction de deux jets en cas de sensibilité à l'infrarouge, un rayonnement doux entraîne la fusion des deux jets reconstruits en un seul. En bas : cas de reconstruction d'un jet non existant en cas de sensibilité à la colinéarité, deux particules de faible énergie qui ne forment pas un jet (par exemple deux hadrons de l'évènement sous jacent) sont reconstruits comme un seul flux d'énergie et définissent un jet.	45
2.11	Principe général de l'algorithme sans "graine" (provenant de [43]).	47
2.12	Simulation de la reconstruction de jets par différents algorithmes (provenant de [45]).	49
3.1	Diagramme de Feynman au premier ordre de la désintégration du lepton τ	51
3.2	Exemple de désintégration hadronique du τ avec radiations correctives dans la voie S.	52
3.3	Schéma simplifié de la désintégration en deux corps du lepton τ	54
3.4	Masse invariante reconstruite des τ dans les modes de désintégration $\tau^\pm \rightarrow \pi^\pm \pi^0$ et $\tau^\pm \rightarrow \pi^\pm \pi^\mp \pi^\pm$ pour des évènements $Z \rightarrow \tau\tau \rightarrow \mu\tau_{had}$, extrait de [59].	55
3.5	Sources de τ dans le Modèle Standard. (a) Processus Drell-Yan, (b) Désintégration d'un boson W, (c) Désintégration d'un quark top.	56
3.6	Rapports d'embranchement des modes de désintégration du boson de Higgs (dans le Modèle Standard).	56
3.7	Isolation d'un candidat tau par la méthode basée sur les coupures.	59
3.8	Reconstruction d'un π^0 pour la méthode HPS.	59
3.9	Distribution des masses des candidats τ (échantillon $Z\tau\tau$ avec rejection de muon et d'électron) pour la méthode HPS avec les différentes catégories de la méthode HPS.	61
3.10	Distributions de la variable retournée par les réseaux de neurones pour les cinq modes de désintégration utilisée dans l'algorithme TaNC (provenant de [54]), pour des évènements $Z\tau\tau$ (en rouge) et QCD (en bleu).	64
3.11	Algorithme de création d'objets <i>PAT</i>	67
3.12	Architecture logique utilisée pour l'identification du lepton τ	69
4.1	Superpositions des sous-échantillons QCD constituant le Monte Carlo final : les spectres sont rééchelonnés en P_T en fonction de leur section efficace et du nombre d'évènements contenu dans le sous-échantillon.	72
4.2	Spectre en impulsion transverse (normalisée à partir de l'intégrale des données) des jets dans les données et les simulations QCD pour a) Tous les jets. b) Les jets "sonde".	74
4.3	Distribution du nombre d'évènements selon le seuil de déclenchement (en GeV) ayant été activé dans a) Les données 2010 b) Les simulations QCD Monte Carlo.	75
4.4	Distribution des distances entre les τ (information générée) et leur candidat reconstruit le plus proche dans divers échantillons.	77
4.5	Corrélations pour les τ dans un échantillon $Z\tau\tau$ (à gauche) et SUSY LM1 (à droite) entre les variables (information générée) a) P_T vs η , b) P_T vs R_{Jet} , c) η vs R_{Jet}	79

4.6	Efficacité d'identification idéale des τ générés dans un échantillon $Z \rightarrow \tau\tau$.	80
4.7	Distance angulaire entre les différents constituants du jet et la trace principale dans des événements $Z \rightarrow \tau\tau$ pour a) Les jets de τ b) Les jets QCD (cône signal de taille maximale $\Delta R = 0, 15$). Une coupure en $P_T > 1 \text{ GeV}$ est imposée sur les constituants.	80
4.8	Distance angulaire entre les différents constituants du jet et la trace principale dans des événements $Z \rightarrow \tau\tau$ pour a) Les jets de τ b) Les jets QCD (cône signal de taille maximale $\Delta R = 0, 3$). Une coupure en $P_T > 1 \text{ GeV}$ est imposée sur les constituants.	81
4.9	Distance angulaire entre les différents constituants du jet et la trace principale dans a) Les jets de τ b) Les jets de quarks. Dans des événements SUSY LM1 (cône signal de taille maximale $\Delta R = 0, 15$). Une coupure en $P_T > 1 \text{ GeV}$ est imposée sur les constituants.	82
4.10	Efficacité idéale par la méthode basée sur les coupures dans un échantillon $Z \rightarrow \tau\tau$	82
4.11	Spectres en impulsion transverse des jets de τ dans divers échantillons contenant des τ : a) Quantités générées, b) Quantités reconstruites.	83
4.12	Efficacités d'identification des candidats τ par la méthode basée sur les coupures, dans divers échantillons ($\Delta R_{\text{association}} < 0, 5$).	83
4.13	Efficacités d'identification des candidats τ , par la méthode basée sur les coupures avec association restreinte ($\Delta R_{\text{assoc}} < 0, 15$), dans divers échantillons.	84
4.14	Efficacités d'identification des candidats τ par la méthode HPS (Loose isolation), dans divers échantillons ($\Delta R_{\text{association}} < 0, 5$).	85
4.15	Efficacités d'identification des candidats τ par la méthode TaNC (Scenario à 1% de faux), dans divers échantillons ($\Delta R_{\text{association}} < 0, 5$).	85
4.16	Spectres en pseudo-rapacité des jets de τ dans divers échantillons contenant des τ : a) Information générée, b) Information reconstruite.	86
4.17	Efficacités expérimentales par la méthode basée sur les coupures dans divers échantillons ($\Delta R_{\text{association}} < 0, 5$).	86
4.18	Efficacités expérimentales par la méthode basée sur les coupures dans divers échantillons ($\Delta R_{\text{association}} < 0, 15$).	87
4.19	Efficacité expérimentale VS R_{Jet} ($P_T > 20 \text{ GeV}, \eta < 2, 4$) par la méthode basée sur les coupures dans divers échantillons.	88
4.20	Efficacité expérimentale VS R_{Jet} ($P_T > 20 \text{ GeV}, \eta < 2, 4$) par la méthode Hadron Plus Strips (isolation relâchée) dans divers échantillons.	88
4.21	Efficacité expérimentale VS R_{Jet} ($P_T > 20 \text{ GeV}, \eta < 2, 4$) par la méthode Tau Neural Classifier (scenario à 1% de mauvaise identification) dans divers échantillons.	89
4.22	Corrélations pour les jets QCD dans les collisions 2010 et dans un échantillon QCD (Seuil de déclenchement 30 GeV) entre les variables a) P_T vs $ \eta $, b) P_T vs R_{Jet} , c) η vs R_{Jet} . Une coupure en $P_T > 20 \text{ GeV}$ est imposée sur les candidats.	90
4.23	Corrélations pour les jets QCD dans les collisions 2010 et dans un échantillon QCD (Seuil de déclenchement 100 GeV) entre les variables a) P_T vs $ \eta $, b) P_T vs R_{Jet} , c) η vs R_{Jet} . Une coupure en $P_T > 20 \text{ GeV}$ est imposée sur les candidats.	91

4.24	Corrélations pour les candidats τ non-associés à un tau hadronique dans des échantillons $t\bar{t}$ (à gauche) et SUSY LM1 (à droite) entre les variables a) P_T vs $ \eta $, b) P_T vs R_{Jet} , c) η vs R_{Jet} . Une coupure en $P_T > 20 \text{ GeV}$ est imposée sur les candidats.	92
4.25	$\hat{P}_T$ des évènements sélectionnés en fonction des différents seuils de déclenchement pour jet.	93
4.26	Spectres en P_T des jets QCD dans divers échantillons : Sélection par déclenchement des jets avec un seuil de 100 GeV (QCD et Données), et rejections de muons et d'électrons appliquées à ces échantillons.	93
4.27	Taux de mauvaise identification VS P_T avec information reconstruite pour un déclencheur de haut niveau jet de 100 GeV avec la méthode basée sur les coupures.	94
4.28	Taux de mauvaise identification VS P_T avec information reconstruite pour un déclencheur de haut niveau jet de 100 GeV avec la méthode Hadron Plus Strips.	94
4.29	Taux de mauvaise identification VS P_T avec information reconstruite pour un déclencheur de haut niveau jet de 100 GeV avec la méthode Tau Neural Classifier.	95
4.30	Taux de mauvaise identification obtenus par la collaboration VS P_T pour les algorithmes HPS (isolation relâchée) et TaNC (isolation relâchée). Provenant de [59].	96
4.31	Spectres en η des jets QCD dans divers échantillons : Sélection par déclenchement des jets avec un seuil de 100 GeV (QCD et Données), et réjections de muons et d'électrons appliquées à ces échantillons.	96
4.32	Taux de mauvaise identification VS η avec information reconstruite pour un déclencheur de haut niveau jet de 100 GeV avec la méthode basée sur les coupures.	97
4.33	Taux de mauvaise identification VS η avec information reconstruite pour un déclencheur de haut niveau jet de 15 GeV avec la méthode basée sur les coupures.	97
4.34	Taux de mauvaise identification VS η avec information reconstruite pour un déclencheur de haut niveau jet de 100 GeV avec la méthode Hadron Plus Strips.	98
4.35	Taux de mauvaise identification VS η avec information reconstruite pour un déclencheur de haut niveau jet de 100 GeV avec la méthode Tau Neural Classifier.	98
4.36	Spectres en R_{Jet} des jets QCD dans divers échantillons : Sélection par déclenchement des jets avec un seuil de 100 GeV (QCD et Données), et réjections de muons et d'électrons appliquées aux échantillons contenant des τ (électrofaibles, $t\bar{t}$ et SUSY).	99
4.37	Taux de mauvaise identification VS R_{Jet} avec information reconstruite pour un déclencheur de haut niveau jet de 100 GeV avec la méthode Basée sur les coupures.	100
4.38	Taux de mauvaise identification VS R_{Jet} avec information reconstruite pour un déclencheur de haut niveau jet de 100 GeV avec la méthode Hadron Plus Strips.	100

4.39	Taux de mauvaise identification VS R_{Jet} avec information reconstruite pour un déclencheur de haut niveau jet de 100 GeV avec la méthode Tau Neural Classifier.	101
4.40	Taux de mauvaise identification pour la méthode basée sur les coupures avec différents seuils de déclenchement.	102
4.41	Contenu de l’anneau d’isolation avant l’isolation par l’algorithme cut-based pour les jets “étiquette” dans des échantillons de données et QCD Monte-Carlo.	103
4.42	Taux de mauvaise identification pour la méthode Hadron Plus Strips (isolation relâchée) avec différents seuils de déclenchement ($P_T > 20$ GeV, $ \eta < 2, 4$).	104
4.43	Taux de mauvaise identification pour la méthode Tau Neural Classifier (isolation relâchée) avec différents seuils de déclenchement ($P_T > 20$ GeV, $ \eta < 2, 4$).	104
4.44	Taux de mauvaise identification VS R_{Jet} pour le jet “sonde” de plus haute impulsion avec différents seuils de déclenchement, Identification basée sur les coupures.	105
5.1	Spectres en ΣE_T , les évènements dans les données et les Monte Carlo QCD sont sélectionnés par un seuil de déclenchement non corrigé (pour jet) de 30 GeV et 100 GeV.	110
5.2	Contenu en sparticules d’évènements SUSY LM1, à $\Sigma E_T \leq 350$ GeV et $\Sigma E_T \geq 350$ GeV.	111
5.3	Spectres en énergie transverse manquante (sans coupure pour les données et les Monte Carlo QCD), les évènements dans les données et les Monte Carlo QCD sont sélectionnés par un seuil de déclenchement non corrigé (pour jet) de 30 GeV et 100 GeV.	112
5.4	Spectres en énergie transverse manquante (avec une coupure $MET < 100$ GeV pour les données et les Monte Carlo QCD), les évènements dans les données et les Monte Carlo QCD sont sélectionnés par un seuil de déclenchement non corrigé (pour jet) de 30 GeV et 100 GeV.	112
5.5	Spectres en multiplicité de hadrons chargés de $P_T > 1$ GeV, les évènements dans les données et les Monte Carlo QCD sont sélectionnés par un seuil de déclenchement non corrigé (pour jet) de 30 GeV et 100 GeV.	113
5.6	Spectres en multiplicité de photons de $P_T > 1,5$ GeV, les évènements dans les données et les Monte Carlo QCD sont sélectionnés par un seuil de déclenchement non corrigé (pour jet) de 30 GeV et 100 GeV.	114
5.7	a) Spectres en multiplicité de jets de $P_T > 10$ GeV, b) Spectres en multiplicité de jets de $P_T > 30$ GeV. Les évènements dans les données et les Monte Carlo QCD sont sélectionnés par un seuil de déclenchement non corrigé (pour jet) de 30 GeV et 100 GeV.	115
5.8	Distributions dans les évènements $Z\tau\tau$ de ΣE_T VS a) Multiplicité de hadrons Chargés, b) Energie transverse manquante.	116
5.9	Distributions dans les évènements SUSY LM1 de ΣE_T VS a) Multiplicité de hadrons Chargés, b) Energie transverse manquante.	116
5.10	Distributions dans les données du LHC (seuil de déclenchement pour Jet de 30 GeV) de ΣE_T VS a) Multiplicité de hadrons Chargés, b) Energie transverse manquante.	116

5.11	Distributions dans les évènements QCD MC (seuil de déclenchement pour Jet de 30 GeV) de ΣE_T VS a) Multiplicité de hadrons Chargés, b) Energie transverse manquante.	117
5.12	Distributions dans les évènements $Z \rightarrow \tau\tau$ de la multiplicité de hadrons chargés VS a) Multiplicité de photons, b) Multiplicité de jets de $P_T > 10 \text{ GeV}$	118
5.13	Distributions dans les évènements SUSY LM1 de la multiplicité de hadrons chargés VS a) Multiplicité de photons, b) Multiplicité de jets de $P_T > 10 \text{ GeV}$	118
5.14	Distributions dans les données du LHC de la multiplicité de hadrons chargés VS a) Multiplicité de photons, b) Multiplicité de jets de $P_T > 10 \text{ GeV}$. . .	119
5.15	Distributions dans les évènements QCD MC de la multiplicité de hadrons chargés VS a) Multiplicité de photons, b) Multiplicité de jets de $P_T > 10 \text{ GeV}$	119
5.16	Efficacités d'identification basée sur les coupures VS ΣE_T pour des candidats de $P_T > 50 \text{ GeV}$, $ \eta < 1$	121
5.17	Efficacités d'identification basée sur les coupures VS ΣE_T pour des candidats de $20 \text{ GeV} < P_T < 50 \text{ GeV}$, $ \eta < 1$	121
5.18	Multiplicités en hadrons chargés ($P_T > 1 \text{ GeV}$) de l'anneau d'isolation pour les τ reconstruits (avant isolation) de $30 < P_T < 50 \text{ GeV}$ (à gauche) et $50 < P_T < 80 \text{ GeV}$ (à droite) dans des évènements de ΣE_T a)Inférieur a 150 GeV, b)Compris entre 150 et 300 GeV, c)Supérieur à 300 GeV.	123
5.19	Efficacités d'identification basée sur les coupures VS $E_{T,manquante}$ pour des candidats de $P_T > 50 \text{ GeV}$, $ \eta < 1$	124
5.20	Efficacités d'identification basée sur les coupures VS $E_{T,manquante}$ pour des candidats de $20 \text{ GeV} < P_T < 50 \text{ GeV}$, $ \eta < 1$	124
5.21	Efficacités d'identification basée sur les coupures VS $N_{\text{Hadrons Chargés}}$ pour des candidats de $P_T > 50 \text{ GeV}$, $ \eta < 1$	125
5.22	Efficacités d'identification basée sur les coupures VS $N_{\text{Hadrons Chargés}}$ pour des candidats de $20 \text{ GeV} < P_T < 50 \text{ GeV}$, $ \eta < 1$	126
5.23	Efficacités d'identification en fonction de P_T (à gauche) et η (à droite) dans différents intervalles en multiplicité de hadrons chargés.	127
5.24	Multiplicités en hadrons chargés ($P_T > 1 \text{ GeV}$) de l'anneau d'isolation pour les τ reconstruits (avant isolation) de $30 < P_T < 50 \text{ GeV}$ (à droite) et $50 < P_T < 80 \text{ GeV}$ (à gauche) dans des évènements avec des multiplicité de hadrons chargés a)Inférieure a 20, b)Comprise entre 20 et 40, c)Supérieure à 40.	128
5.25	Efficacités d'identification basée sur les coupures VS N_{Photons} pour des candidats de $P_T > 50 \text{ GeV}$, $ \eta < 1$	129
5.26	Efficacités d'identification basée sur les coupures VS N_{Jets} pour des candidats de $P_T > 50 \text{ GeV}$, $ \eta < 1$ VS a) La multiplicité de jets de plus de $P_T > 10 \text{ GeV}$, b) La multiplicité de jets de $P_T > 30 \text{ GeV}$	130
5.27	Efficacités d'identification Hadron Plus Strips VS ΣE_T pour des candidats de $P_T > 50 \text{ GeV}$, $ \eta < 1$	131
5.28	Efficacités d'identification Hadron Plus Strips VS ΣE_T pour des candidats de $20 \text{ GeV} < P_T < 50 \text{ GeV}$, $ \eta < 1$	132
5.29	Efficacités d'identification Hadron Plus Strips VS MET pour des candidats de $P_T > 50 \text{ GeV}$, $ \eta < 1$	133

5.30	Efficacités d'identification Hadron Plus Strips VS $N_{Hadrons\ Chargés}$ pour des candidats de $P_T > 50\ GeV$, $ \eta < 1$	134
5.31	Efficacités d'identification Hadron Plus Strips VS $N_{Hadrons\ Chargés}$ pour des candidats de $20\ GeV < P_T < 50\ GeV$, $ \eta < 1$	134
5.32	Efficacités d'identification Hadron Plus Strips VS $N_{Photons}$ pour des candidats de $P_T > 50\ GeV$, $ \eta < 1$	135
5.33	Efficacités d'identification Hadron Plus Strips VS N_{Jets} pour des candidats de $P_T > 50\ GeV$, $ \eta < 1$ VS a) La multiplicité de jets de plus de $P_T > 10\ GeV$, b) La multiplicité de jets de $P_T > 30\ GeV$	136
5.34	Efficacités d'identification Tau Neural Classifier VS ΣE_T pour des candidats de $P_T > 50\ GeV$, $ \eta < 1$	137
5.35	Efficacités d'identification Tau Neural Classifier VS ΣE_T pour des candidats de $20\ GeV < P_T < 50\ GeV$, $ \eta < 1$	139
5.36	Efficacités d'identification Tau Neural Classifier VS $N_{Hadrons\ Chargés}$ pour des candidats de $P_T > 50\ GeV$, $ \eta < 1$	140
5.37	Efficacités d'identification Tau Neural Classifier VS $N_{Photons}$ pour des candidats de $P_T > 50\ GeV$, $ \eta < 1$	140
5.38	Efficacités d'identification Tau Neural Classifier pour des candidats de $P_T > 50\ GeV$, $ \eta < 1$ VS a) La multiplicité de jets de plus de $P_T > 10\ GeV$, b) La multiplicité de jets de $P_T > 30\ GeV$	141
5.39	Taux de mauvaise identification basée sur les coupures VS ΣE_T pour des candidats de $P_T > 50\ GeV$, $ \eta < 1$, les données et QCD MC sont sélectionnés à l'aide des systèmes de déclenchement pour jet avec un seuil de $50\ GeV$	143
5.40	Multiplicités en hadrons chargés ($P_T > 1\ GeV$) de l'anneau d'isolation pour les jet QCD reconstruits (avant isolation) de $30 < P_T < 50\ GeV$ (à gauche) et $50 < P_T < 80\ GeV$ (à droite) dans des événements de ΣE_T a)Inférieur a $150\ GeV$, b)Compris entre 150 et $300\ GeV$, c)Supérieur à $300\ GeV$	144
5.41	Taux de mauvaise identification basée sur les coupures VS $N_{Hadrons\ Chargés}$ pour des candidats de $P_T > 50\ GeV$, $ \eta < 1$	145
5.42	Multiplicités en hadrons chargés ($P_T > 1\ GeV$) de l'anneau d'isolation pour les jets QCD reconstruits (avant isolation) de $30 < P_T < 50\ GeV$ (à droite) et $50 < P_T < 80\ GeV$ (à gauche) dans des événements avec des multiplicité de hadrons chargés a)Inférieure à 20 , b)Comprise entre 20 et 40 , c)Supérieure à 40	146
5.43	Taux de mauvaise identification Hadron Plus Strips VS ΣE_T pour des candidats de $P_T > 50\ GeV$, $ \eta < 1$	147
5.44	Taux de mauvaise identification Hadron Plus Strips VS $N_{Hadrons\ Chargés}$ pour des candidats de $P_T > 50\ GeV$, $ \eta < 1$	147
5.45	Taux de mauvaise identification Tau Neural Classifier VS ΣE_T pour des candidats de $P_T > 50\ GeV$, $ \eta < 1$	148
5.46	Taux de mauvaise identification Tau Neural Classifier VS $N_{Hadrons\ Chargés}$ pour des candidats de $P_T > 50\ GeV$, $ \eta < 1$	149
5.47	Limites d'exclusion dans l'espace $m_{1/2}$, m_0 pour la recherche de supersymétrie dans la voie hadronique (source [62]).	151
A.1	Spectre en impulsion transverse des candidats τ identifiés par la charge et le nombre de traces (en noir) puis estimés par la méthode de "fake-rate" (en rouge).	154

A.2	Spectre en impulsion transverse des candidats τ identifiés par isolation (en noir) puis estimés par la méthode de “fake-rate” (en rouge).	154
A.3	Spectre en impulsion transverse des candidats τ identifié par isolation + charge + nombre de traces (en noir) puis estimés par la méthode de “fake-rate” (en rouge).	155
A.4	Spectre en impulsion transverse des candidats τ identifiés par la charge et le nombre de traces (en noir) puis estimés par la méthode de “fake-rate” (en rouge).	155
A.5	Spectre en impulsion transverse des candidats τ identifiés par isolation (en noir) puis estimés par la méthode de “fake-rate” (en rouge).	156
A.6	Spectre en impulsion transverse des candidats τ identifié par isolation + charge + nombre de traces (en noir) puis estimés par la méthode de “fake-rate” (en rouge).	156
B.1	Taux de mauvaise identification par la méthode cut-based du jet de plus haute impulsion dans les évènements QCD MC et données en fonction de P_T , dans les évènements sélectionnés par un seuil de déclenchement de 15 GeV (à gauche) et par un seuil de 100 GeV (à droite).	157
B.2	Taux de mauvaise identification par la méthode HPS du jet de plus haute impulsion dans les évènements QCD MC et données en fonction de P_T , dans les évènements sélectionnés par un seuil de déclenchement de 15 GeV (à gauche) et par un seuil de 100 GeV (à droite).	158
B.3	Taux de mauvaise identification par la méthode TaNC du jet de plus haute impulsion dans les évènements QCD MC et données en fonction de P_T , dans les évènements sélectionnés par un seuil de déclenchement de 15 GeV (à gauche) et par un seuil de 100 GeV (à droite).	158
B.4	Taux de mauvaise identification par la méthode cut-based du jet de seconde plus haute impulsion dans les évènements QCD MC et données en fonction de P_T , dans les évènements sélectionnés par un seuil de déclenchement de 15 GeV (à gauche) et par un seuil de 100 GeV (à droite).	159
B.5	Taux de mauvaise identification par la méthode HPS du jet de seconde plus haute impulsion dans les évènements QCD MC et données en fonction de P_T , dans les évènements sélectionnés par un seuil de déclenchement de 15 GeV (à gauche) et par un seuil de 100 GeV (à droite).	159
B.6	Taux de mauvaise identification par la méthode TaNC du jet de seconde plus haute impulsion dans les évènements QCD MC et données en fonction de P_T , dans les évènements sélectionnés par un seuil de déclenchement de 15 GeV (à gauche) et par un seuil de 100 GeV (à droite).	160
B.7	Taux de mauvaise identification par la méthode cut-based du jet de troisième plus haute impulsion dans les évènements QCD MC et données en fonction de P_T , dans les évènements sélectionnés par un seuil de déclenchement de 15 GeV (à gauche) et par un seuil de 100 GeV (à droite).	160
B.8	Taux de mauvaise identification par la méthode HPS du jet de troisième plus haute impulsion dans les évènements QCD MC et données en fonction de P_T , dans les évènements sélectionnés par un seuil de déclenchement de 15 GeV (à gauche) et par un seuil de 100 GeV (à droite).	161

B.9	Taux de mauvaise identification par la méthode TaNC du jet de troisième plus haute impulsion dans les événements QCD MC et données en fonction de P_T , dans les événements sélectionnés par un seuil de déclenchement de 15 GeV (à gauche) et par un seuil de 100 GeV (à droite).	161
C.1	Taux de mauvaise identification par la méthode cut-based des jets sonde avec un seuil de déclenchement de 100 GeV. a) Aucune coupure en MET n'est appliquée, b) Une coupure $MET < 100 \text{ GeV}$ est appliquée sur les événements.	163
C.2	Taux de mauvaise identification par la méthode Hadron Plus Strips des jets sonde avec un seuil de déclenchement de 100 GeV. a) Aucune coupure en MET n'est appliquée, b) Une coupure $MET < 100 \text{ GeV}$ est appliquée sur les événements.	164
C.3	Taux de mauvaise identification par la méthode Tau Neural Classifier des jets sonde avec un seuil de déclenchement de 100 GeV. a) Aucune coupure en MET n'est appliquée, b) Une coupure $MET < 100 \text{ GeV}$ est appliquée sur les événements.	165

Liste des tableaux

1.1	Particules du Modèle Minimal Supersymétrique Standard[15].	12
1.2	Points de travail mSUGRA	14
1.3	Désintégrations supersymétriques impliquant des sleptons gauches (de [18]).	16
2.1	Caractéristiques du LHC [30].	29
2.2	Les différentes parties du trajectomètre	35
2.3	Interactions des particules avec les différentes parties du détecteur CMS. .	39
3.1	Caractéristiques du lepton τ	50
3.2	Modes de désintégration du W^- réel (poids relatif du mode concerné) . . .	51
3.3	Caractéristiques des coupures d'isolation.	58
3.4	Variables d'entrée pour les réseaux de neurones de la méthode Tau Neural Classifier et modes de désintégration concernés	63
4.1	Échantillons simulés Monte Carlo	72
4.2	Description des systèmes de déclenchement pour les jets	73
4.3	Échantillons de données	75
5.1	Détails de la régression exponentielle $\epsilon(\Sigma E_T) = e^{a\Sigma E_T+b}$ de la Fig. 5.16 et de la Fig. 5.17.	122
5.2	Détails de la régression exponentielle $\epsilon(MET) = e^{a.MET+b}$ de la Fig 5.19 et de la Fig 5.20.	122
5.3	Détails de la régression exponentielle $\epsilon(N_{Chargés}) = e^{aN_{Chargés}+b}$ de la Fig 5.21 et de la Fig 5.22	125
5.4	Détails de la régression exponentielle $\epsilon(N_\gamma) = e^{aN_\gamma+b}$ de la Fig 5.25	129
5.5	Détails de la régression exponentielle $\epsilon(N_{Jets,10GeV}) = e^{aN_{Jets,10GeV}+b}$ de la Fig 5.26	130
5.6	Détails de la régression exponentielle $\epsilon(\Sigma E_T) = e^{a\Sigma E_T+b}$ de la Fig 5.27 et de la Fig 5.28	131
5.7	Détails de la régression exponentielle $\epsilon(N_{Chargés}) = e^{aN_{Chargés}+b}$ de la Fig 5.30 et de la Fig 5.31	133
5.8	Comparaison des valeurs de a du paramétrage $\epsilon(N_{Chargés}) = e^{aN_{Chargés}+b}$ pour l'algorithme basé sur les coupures et l'algorithme HPS	135
5.9	Détails de la régression exponentielle $\epsilon(N_\gamma) = e^{aN_\gamma+b}$ de la Fig 5.32	135
5.10	Détails de la régression exponentielle $\epsilon(N_{Jets10}) = e^{aN_{Jets10}+b}$ de la Fig 5.33	136
5.11	Détails de la régression exponentielle $\epsilon(\Sigma E_T) = e^{a\Sigma E_T+b}$ de la Fig 5.34 . .	138
5.12	Détails de la régression exponentielle $\epsilon(N_{Chargés}) = e^{aN_{Chargés}+b}$ de la Fig 5.36	139

Introduction

Le travail présenté dans ce mémoire de thèse s'intéresse à l'identification des leptons τ produits dans les collisions proton-proton du collisionneur LHC (Large Hadron Collider). Le lepton τ est le seul lepton à pouvoir se désintégrer hadroniquement et nous l'identifions dans ce mode de désintégration. La motivation pour s'intéresser à ce lepton est qu'il est fréquemment présent dans l'état final lors de collisions impliquant des processus supersymétriques. La supersymétrie est une théorie s'étendant au-delà du Modèle Standard en physique des particules dont on recherche actuellement les traces expérimentales en observant les collisions proton-proton au collisionneur LHC à l'aide du détecteur CMS.

Les événements supersymétriques sont caractérisés par une énergie transverse et une multiplicité de jets élevées. Ils sont susceptibles de créer un environnement hadronique particulièrement actif dans le détecteur, ayant pour résultat un effet négatif notable sur l'identification de leptons τ . L'objectif de cette thèse est d'observer puis de caractériser l'impact que peut avoir un environnement hadronique important sur l'identification du lepton τ et le taux de mauvaise identification de jets hadroniques QCD en tant que τ . Cette étude s'est basée sur plusieurs échantillons simulés et sur les données issues du LHC en 2009 et 2010 ($36pb^{-1}$).

Le premier chapitre nous résume brièvement les concepts du Modèle Standard en physique des particules et nous présente les aspects phénoménologiques de la supersymétrie débouchant sur la production de leptons τ . La deuxième partie du chapitre 1 traite des collisions proton-proton, plus particulièrement des aspects des collisions proton-proton produisant des hadrons : les processus QCD induisant un événement sous-jacent.

Le deuxième chapitre est consacré à l'instrumentation et plus particulièrement aux caractéristiques du LHC et du détecteur CMS. Ce chapitre décrit dans un second temps les logiciels permettant de reconstruire un événement se produisant dans le détecteur : les méthodes de reconstruction des traces et des jets sont détaillées dans cette section, ces objets physique étant liés à l'identification du lepton τ .

Le troisième chapitre traite exclusivement du lepton τ . Nous y verrons ses caractéristiques, ses modes de désintégration et les aspects géométriques des jets dans le cas des désintégrations hadroniques. La deuxième partie de ce chapitre traite des méthodes d'identification du lepton τ employés au cours de ces études et présente les points importants de l'environnement d'analyse utilisé.

Le quatrième chapitre présente les résultats d'analyse obtenus par les études de l'identification du lepton τ et la mauvaise identification de jets QCD en fonction des caractéristiques des jets (impulsion, angle d'émission et largeur du jet). Les études concernant l'identi-

cation de τ emploient diverses sources de τ dans le Modèle Standard et des évènements supersymétriques. L'étude sur la mauvaise identification de jets QCD dans ce chapitre est basée sur l'analyse d'échantillons QCD simulés et les données issues du LHC.

Le cinquième et dernier chapitre nous présente les résultats obtenus par l'étude de l'identification de τ et de la mauvaise identification de τ en fonction de l'environnement créé par les évènements. L'activité d'un évènement est évalué par différentes variables globales susceptibles d'affecter l'identification de lepton τ (ou la mauvaise identification de jets QCD). L'hypothèse d'un paramétrage de l'impact d'une variable globale sur l'identification de τ est testée.

Chapitre 1

La physique des collisionneurs

1.1 La physique des particules expérimentale

L'étude des particules élémentaires est un domaine de la physique qui débuta il y a plus d'un siècle, commençant avec la découverte des constituants atomiques : l'électron en 1896 (Expérience de J.J. Thomson) et le noyau en 1909 (Expérience de Rutherford). Puis se poursuivant avec la découverte d'une sous-structure au noyau (proton et neutron) et enfin l'étude des rayons cosmiques qui permirent la découverte des muons et du positron : ce dernier démontrant l'existence de l'antimatière.

Cette science se développa avec la conception des premiers accélérateurs de particules, motivée par l'idée de reproduire les rayons cosmiques en laboratoire grâce à des cavités accélératrices (champ électrique alterné). Les premiers accélérateurs étaient linéaires et les particules n'étaient accélérées qu'une seule fois dans chaque cavité. Puis est venue l'idée d'accélérer une même particule plusieurs fois à l'aide de la même cavité accélératrice (cyclotrons et synchrotrons). Ce concept, qui aboutit en imposant aux particules accélérées une trajectoire circulaire à l'aide d'un champ magnétique, repoussa grandement les limites en énergie par particule. Enfin la conception d'accélérateurs encore plus puissants reposa sur le principe d'effectuer des collisions frontales (deux faisceaux de particules circulent en sens inverse et se croisent en des points définis) plutôt que sur cible fixe.

De nouvelles technologies apparurent au cours du temps, aussi bien au niveau des accélérateurs (des supraconducteurs pour produire des champs magnétiques de plusieurs Teslas et des cavités accélératrices plus puissantes) qu'au niveau des détecteurs (apparition de la micro-électronique permettant la conception de détecteurs avec une meilleure granularité, temporelle et énergétique), repoussant de plus en plus loin la limite en énergie et en intensité des collisions.

Les découvertes autour des accélérateurs furent nombreuses : les mésons et baryons de différentes familles débouchant sur le modèle des quarks, les bosons W et Z découverts d'abord au *SP $\bar{P}$ S* (CERN) puis mesurés avec précision au LEP (CERN), le quark top (découvert au Tevatron (FermiLab)), le lepton τ . Ces découvertes ont confirmé la validité expérimentale du Modèle Standard, mais ce modèle arrive à sa limite à l'échelle multi-TeV : on recherche alors les extensions de ce modèle. Le collisionneur LHC (CERN), qui est le collisionneur le plus puissant construit à ce jour, explore différentes théories au-delà du Modèle Standard.

1.2 Le Modèle Standard

Le Modèle Standard est la théorie communément admise et employée en physique des particules. Les bases de ce modèle ont été posées dans les années 1960 par A. Salam[1], S.L. Glashow[2] et S. Weinberg[3]. Les interactions fondamentales entre particules élémentaires sont traitées dans ce modèle par une invariance de jauge locale par les groupes de symétrie $SU(3)_C \times SU(2)_L \times U(1)_Y$, ces termes seront explicités dans la section traitant les interactions fondamentales. On a deux points importants dans ce Modèle :

- Les constituants élémentaires de la matière, aussi appelés particules élémentaires, sont les composants les plus simples dans le sens où ils ne possèdent pas de sous-structure.
- Les interactions fondamentales décrivent les processus subatomiques ayant lieu entre les particules élémentaires du Modèle Standard.

1.2.1 Les particules élémentaires

Les particules élémentaires sont les constituants fondamentaux du Modèle Standard (aucune sous-structure décrite pour ces particules), elles sont divisées en deux grandes catégories :

- Les **fermions** suivant la statistique de Fermi-Dirac, de spin demi-entier et soumis au principe d'exclusion de Pauli.
- Les **bosons** suivant la statistique de Bose-Einstein de spin entier et non soumis au principe d'exclusion de Pauli.

Les bosons véhiculent les interactions fondamentales et seront traités dans la section suivante. Les fermions, constituants de la matière, sont à leur tour divisés en deux grandes catégories : les quarks et les leptons. Les quarks sont soumis aux quatre interactions fondamentales (le modèle des quarks a été proposé dans les années 1960 par M. Gell-Mann[4], Y. Ne'eman[5] et G. Zweig[6]). Pour les leptons on distingue le cas des leptons massifs et chargés (électron, muon et tau) soumis aux interactions faible et électromagnétique et le cas des neutrinos sans masse ni charge électrique et soumis uniquement à l'interaction faible. Une distinction entre chiralité droite (Right : R) et gauche (Left : L) a dû être introduite suite à l'observation de la violation de parité par interaction faible[7]. L'électron gauche et le neutrino gauche possèdent les mêmes propriétés du point de vue de l'interaction faible, ils sont associés en un doublet d'isospin faible. Par extension le quark gauche up et le quark gauche down forment aussi un doublet d'isospin faible. Les particules de chiralité droite forment un singulet d'isospin faible. Les neutrinos sont considérés comme non massifs et n'existent donc pas dans l'état de chiralité droit.

Les fermions sont organisés en trois familles de deux doublets et de trois singulets d'isospin faible, ces familles sont identiques en tout point à l'exception faite de leur masse :

- Première famille : $\begin{pmatrix} \nu_{e,L} \\ e_L \end{pmatrix}, e_R, \begin{pmatrix} u_L \\ d_L \end{pmatrix}, u_R, d_R$
- Deuxième famille : $\begin{pmatrix} \nu_{\mu,L} \\ \mu_L \end{pmatrix}, \mu_R, \begin{pmatrix} c_L \\ s_L \end{pmatrix}, c_R, s_R$
- Troisième famille : $\begin{pmatrix} \nu_{\tau,L} \\ \tau_L \end{pmatrix}, \tau_R, \begin{pmatrix} t_L \\ b_L \end{pmatrix}, t_R, b_R$

• Les assemblages de particules élémentaires

Les particules composites sont des assemblages de particules élémentaires, par exemple les nucléons (proton et neutrons) composant l'atome sont constitués de quarks de la première famille. Les leptons ne donnent pas lieu à des particules composites (excepté dans le cas d'un atome qui est aussi constitué d'électrons), tandis que pour les quarks le contraire est observé : un quark seul n'existe pas en tant que tel, il est associé à d'autres partons et confiné dans un hadron.

Un hadron est un assemblage d'un quark et d'un antiquark dans le cas d'un méson et de trois quarks (ou trois antiquarks) dans le cas d'un (anti)baryon, ces quarks sont liés entre eux par des gluons. Il s'agit ici des quarks de valence responsables de la nature et des propriétés d'un hadron (charge, spin, étrangeté...). Mais les hadrons contiennent aussi une grande quantité de gluons et de paires quark-antiquark constituant une "mer" de particules virtuelles. Les quarks et les gluons confinés dans un hadron sont appelés "partons", le contenu en quark d'un hadron est décrit par les fonctions de densité de partons et dépend de l'échelle d'énergie à laquelle le hadron est étudié (au moment de la collision). La Fig. 1.1 représente ces fonction de densité de parton, x représente la fraction d'impulsion longitudinale du hadron emportée par le parton et $xf(x)$ sa densité aux échelles d'énergie $Q^2 = 20 \text{ GeV}^2$ et $Q^2 = 10 \text{ TeV}^2$ sur lesquelles le parton est étudié. La probabilité de trouver un parton emportant une fraction comprise entre x et $x + dx$ est donnée par $dx f(x)$. On voit dans la figure 1.1 que la densité des quarks de valence (notés u_v et d_v) varie très peu en fonction de l'échelle d'énergie tandis que les quarks lourds de la mer (charmés et beaux) ont une densité plus importante dans la mer à grand Q^2 qu'à bas Q^2 . La probabilité de retrouver des quarks lourds impliqués dans les collisions augmente à très haute énergie si leur fraction d'impulsion emportée x est faible.

1.2.2 Les interactions fondamentales

Les interactions fondamentales sont à l'origine des phénomènes physiques microscopiques et macroscopiques observés dans la nature. Ces interactions ont lieu à l'échelle microscopique grâce à l'échange de particules : les bosons (c.f. section suivante). Il existe 4 interactions fondamentales (résumées dans le tableau 1.2), :

- **Electromagnétisme** : L'interaction électromagnétique est présente dès qu'un système physique possède une charge électrique ou un moment magnétique. Cette interaction est bien connue et est largement décrite à l'échelle macroscopique par les lois de Maxwell. L'interaction électromagnétique est décrite à l'échelle microscopique par l'électrodynamique quantique (QED). La particule médiatrice de cette interaction est le photon.
- **Interaction forte** : Le noyau atomique est confiné dans un petit volume (une sphère d'un rayon de l'ordre du femtomètre) et est composé de nucléons possédant une charge

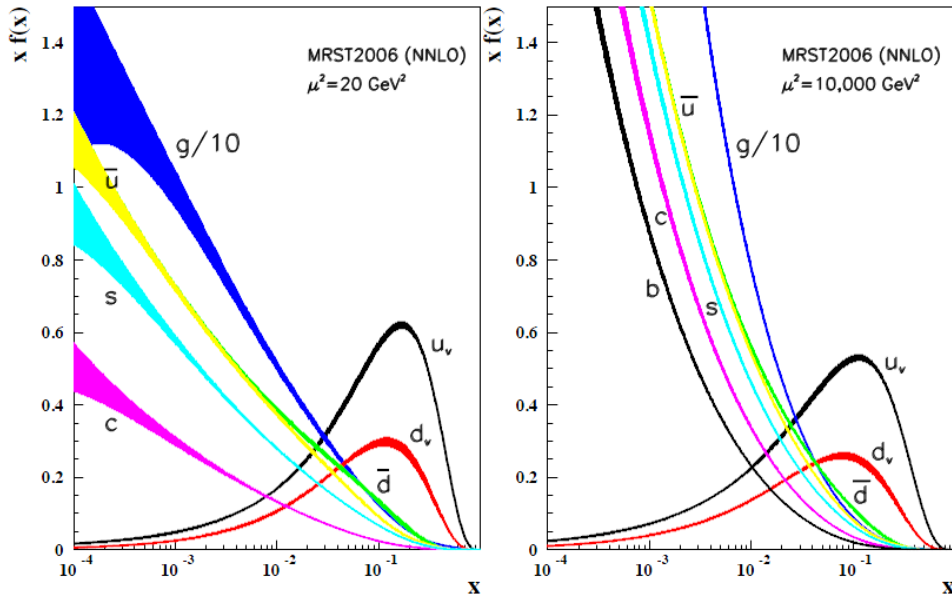


FIGURE 1.1 – (source [8]) Distributions de la densité de partons $x.f(x)$ (où $f = u_v, d_v, \bar{u}, \bar{d}, s, c, b, g$).

électrique positive ou neutre. Le noyau atomique ne pourrait donc pas exister si il n'y avait pas une interaction de force supérieure à la répulsion coulombienne (électromagnétique). L'interaction forte, décrite par la chromodynamique quantique (QCD), explique non seulement la cohésion du noyau atomique mais aussi le confinement des partons dans les hadrons (c.f. section 1.2.2). Les gluons sont les particules véhiculant l'interaction forte.

- **Interaction faible** : L'interaction faible intervient principalement dans les phénomènes de désintégration β , ce phénomène étant lui même dû à la désintégration d'un neutron en proton (ou d'un proton en neutron si le phénomène a lieu dans un noyau atomique). D'une manière plus générale l'interaction faible intervient dans les courants chargés et neutres et est véhiculée par les bosons $W^\pm$ (courant chargé) et Z^0 (courant neutre).
- **Gravitation** : La gravitation est responsable de l'attraction qui existe entre deux masses, elle est observée uniquement à l'échelle macroscopique car elle est beaucoup moins forte que les autres interactions. La particule qui véhicule l'interaction gravitationnelle est le graviton et n'a pas encore été découverte.

Le Modèle Standard n'inclue pas la gravitation, celle-ci est négligeable par rapport aux autres interactions dans les processus de physique subatomique.

1.2.2.1 Les interactions dans le Modèle Standard

Les interactions sont décrites dans le Modèle Standard via des transformations de jauge. Une particule libre de spin $\frac{1}{2}$ de masse m et de charge électrique e est décrite par un spineur de Dirac $\psi(x)$ et son lagrangien est le suivant :

Property \ Interaction	Gravitational	Weak (Electroweak)	Electromagnetic	Strong	
Acts on:	Mass – Energy	Flavor	Electric Charge	Color Charge	See Residual Strong Interaction Note
Particles experiencing:	All	Quarks, Leptons	Electrically charged	Quarks, Gluons	Hadrons
Particles mediating:	Graviton (not yet observed)	$W^+ W^- Z^0$	γ	Gluons	Mesons
Strength relative to electromag for two u quarks at: $\left\{ \begin{array}{l} 10^{-18} \text{ m} \\ 3 \times 10^{-17} \text{ m} \end{array} \right.$ for two protons in nucleus	10^{-41}	0.8	1	25	Not applicable to quarks
	10^{-41}	10^{-4}	1	60	
	10^{-36}	10^{-7}	1	Not applicable to hadrons	20

FIGURE 1.2 – Interactions fondamentales (source [9]).

$$\mathcal{L} = \bar{\psi}(x)(i\gamma^\mu \partial_\mu - m)\psi(x) \quad (1.1)$$

On applique une transformation globale de jauge à cette particule :

$$\psi(x) \rightarrow e^{i\alpha}\psi(x) \quad (1.2)$$

On constate alors que le Lagrangien reste invariant sous ces conditions car les phases s'annulent :

$$\bar{\psi}(x) \rightarrow e^{-i\alpha}\bar{\psi}(x) \quad (1.3)$$

$$\partial_\mu \psi(x) \rightarrow e^{i\alpha} \partial_\mu \psi(x) \quad (1.4)$$

Il en découle la conservation globale de la charge électrique. Cependant lors d'une transformation locale de jauge où la phase $\alpha(x)$ dépend des coordonnées d'espace temps x^μ le lagrangien n'est plus invariant car un terme supplémentaire apparaît avec la dérivée ∂_μ :

$$\partial_\mu \psi(x) \rightarrow e^{i\alpha(x)} \partial_\mu \psi(x) + i(\partial_\mu \alpha(x)) e^{i\alpha(x)} \psi(x) \quad (1.5)$$

Afin de conserver l'invariance du lagrangien sous une transformation locale il faut remplacer la dérivée ∂_μ par une dérivée covariante D_μ de manière que :

$$D_\mu \psi(x) \rightarrow e^{i\alpha(x)} D_\mu \psi(x) \quad (1.6)$$

La transformation considérée ici est décrite par le groupe de symétrie $U(1)$ des transformations locales unitaires à une seule dimension, comportant un seul générateur α . La condition d'invariance par cette transformation est satisfaite en introduisant un champ de jauge A_μ dans la dérivée covariante :

$$D_\mu = \partial_\mu - ieA_\mu \quad (1.7)$$

Dans le cas d'une transformation locale de jauge $\bar{\psi}(x) \rightarrow e^{-i\alpha(x)}\bar{\psi}(x)$, le champ A_μ se transforme aussi :

$$A_\mu \rightarrow A'_\mu = A_\mu - \frac{1}{e} \partial_\mu \alpha(x) \quad (1.8)$$

Le champ A_μ (A'_μ) décrit un boson vecteur (spin 1) de masse nulle, on y retrouve alors les caractéristiques du photon. Sous une transformation locale on voit apparaître un couplage entre un fermion et le champ électromagnétique décrit par le terme :

$$e \bar{\psi}(x) \gamma^\mu A_\mu \psi(x) \quad (1.9)$$

Les transformations locales décrites par le groupe de symétrie $U(1)$ peuvent être associées à l'interaction électromagnétique. Les interactions faible et forte sont aussi introduites dans le Modèle Standard par ces champs de jauge, nous allons les détailler dans les sections suivantes.

1.2.2.2 L'interaction forte

La théorie de jauge traitant l'interaction forte tire son origine de deux grands travaux : la découverte du phénomène de liberté asymptotique entre quark par D.J. Gross, F. Wilczek[10] et H.D. Politzer[11] dans les années 1970, et l'introduction pour les quarks d'un nombre quantique supplémentaire avec la symétrie $SU(3)$ par M. Gell-Mann, H. Fritzsch et H. Leutwyler[12] : la charge de couleur C dont les valeurs qu'on peut désigner par analogie comme Rouge, Bleu et Vert (un hadron est globalement sans couleur, ou de couleur "blanche"). L'introduction de ce nombre quantique a été motivé par la découverte des baryons Δ^{++} (uuu), Δ^- (ddd) et Ω^- (sss). Ces baryons contenant trois quarks de valence de même saveur et de même spin (3/2), le principe d'exclusion de Pauli ne serait pas respecté sans l'introduction d'un nouveau nombre quantique.

L'interaction forte entre charges de couleur est décrite par une invariance de jauge du lagrangien sous les transformations locales de jauge du groupe de symétrie $SU(3)_C$ (C étant la couleur) possédant 8 générateurs. Il y a donc l'introduction de 8 bosons de jauge G_μ^a dans la dérivée covariante pour l'interaction forte correspondant aux 8 gluons. La dérivée covariante devient pour l'interaction forte :

$$D_\mu = \partial_\mu - i g_s \lambda_a G_\mu^a \quad (1.10)$$

Où les λ_a sont les générateurs du groupe $SU(3)$ et g_s la constante de couplage de l'interaction forte.

Le couplage fort $\alpha_s = g_s/4\pi$ (ainsi que les autres couplages) évolue avec l'énergie :

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + \frac{\alpha_s(\mu^2)}{4\pi} b_0 \ln(\frac{Q^2}{\mu^2})} \quad \text{Avec } b_0 = -\frac{2}{3} N_f + \frac{11}{3} N_C \quad (1.11)$$

Où Q^2 est le moment transféré, μ une échelle d'énergie de référence, N_f le nombre de saveurs existant à l'échelle d'énergie considérée (de 3 à 6) et N_C le nombre de couleurs (=3).

Ceci implique que si $Q^2 \rightarrow \infty$ alors $\alpha_s \rightarrow 0$, nous sommes dans le cas de deux quarks infiniment proches et deviennent asymptotiquement libres. Dans le cas contraire si $Q^2 \rightarrow 0$ alors α_s augmente, le couplage fort devient plus important à mesure que la distance entre ces deux quarks augmente. C'est la raison pour laquelle un quark seul n'est jamais observé : il donnera lieu au phénomène de confinement et de fragmentation que nous verrons plus loin dans ce chapitre.

1.2.2.3 L'interaction électrofaible

Les observations des expériences antérieures en physique des particules ont démontré que le boson W de l'interaction faible ne se couple qu'aux fermions de chiralité gauche. Les fermions gauches sont regroupés en doublets d'isospin faible dans le Modèle Standard au vu des symétries observées expérimentalement entre les fermions d'un même doublet. L'invariance de jauge est obtenue pour l'interaction faible sous les transformations locale de symétrie $SU(2)$ avec trois générateurs $\vec{T}$ mélangeant les états gauche et droit. Cette invariance est obtenue pour l'interaction électrofaible à la condition que $m = 0$ dans le Lagrangien de Dirac 1.1. Tous les fermions ont alors une masse nulle dans le Lagrangien du Modèle Standard. Les champs de jauge de l'interaction faible apparaissent dans la dérivée covariante de jauge en faisant apparaître trois bosons de jauge $\vec{W}_\mu$.

Il y a de grandes similitudes entre le formalisme de l'interaction faible et celui de l'interaction électromagnétique, aussi ces deux interactions sont unifiées dans le Modèle Standard en interaction électrofaible. Le lagrangien doit alors être invariant sous une transformation locale de jauge $SU(2)_L \times U(1)_Y$ (L est la chiralité gauche et Y l'hypercharge faible). La dérivée covariante pour l'interaction électrofaible devient :

$$D_\mu = \partial_\mu + ig\vec{T} \cdot \vec{W}_\mu + ig'\frac{Y}{2}B_\mu \quad (1.12)$$

Où g et g' sont les deux constantes de couplage électrofaible. Cette dérivée covariante introduit 4 bosons de jauge vectoriels sans masse : les trois $\vec{W}_\mu$ et B_μ . Les bosons de l'interaction faible W/Z et le boson de l'interaction électromagnétique (photon) sont des mélanges de ces champs de jauge :

$$A_\mu = \cos\theta_W B_\mu + \sin\theta_W W_\mu^3 \quad (1.13)$$

$$Z_\mu = -\sin\theta_W B_\mu + \cos\theta_W W_\mu^3 \quad (1.14)$$

$$W_\mu^\pm = \frac{W_\mu^1 \mp iW_\mu^2}{\sqrt{2}} \quad (1.15)$$

Où θ_W est l'angle de mélange électrofaible (angle de Weinberg).

Les bosons comme les fermions apparaissent sans masse dans le Modèle Standard car les termes en $m\psi\bar{\psi}$ ne sont pas invariants sur transformations $SU(2)$ et les termes en mW^+W^- sont en conflit avec la renormalisabilité de la théorie. Expérimentalement on observe que les bosons W et Z ainsi que la plupart des fermions sont massifs, on a donc une symétrie brisée. Le mécanisme de Higgs[13] fournit un scénario à la brisure de symétrie électrofaible en introduisant un boson scalaire : le boson de Higgs qui se couple aux particules pour

leur donner leur masse.

• Le Mécanisme de Higgs

L'introduction d'un terme de masse pour les fermions dans l'équation 1.1 pose le problème que ce terme n'est pas invariant sous une transformation locale de jauge de $SU(2)_L \times U(1)_Y$. Le mécanisme de Higgs introduit un champ scalaire $\phi(x)$ (de spin 0) en auto-interaction via un potentiel $V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2$. L'état fondamental ($V(\phi) = 0$) admet un ensemble de solutions :

$$\phi^\dagger \phi = \frac{-\mu^2}{2\lambda} = \frac{v^2}{2} \quad (1.16)$$

Où $\mu^2 < 0$ et $\lambda > 0$ sont des paramètres inconnus et v est la valeur attendue du vide (vacuum expectation value : v.e.v). Cette symétrie est spontanément brisée par le choix d'un minimum (Fig. 1.3).

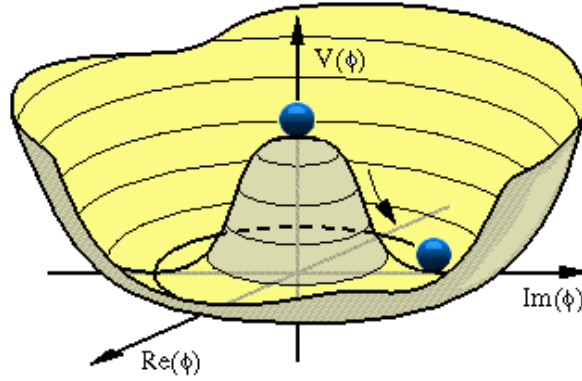


FIGURE 1.3 – Représentation tridimensionnelle du potentiel de Higgs.

Le champ de Higgs génère la masse des bosons $W^\pm$ et Z^0 autour de l'état fondamental ($\phi_0 = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix}$) décrites par la relation :

$$\phi(x) = \exp(-i\vec{\sigma} \cdot \vec{\xi}(x)) \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \quad (1.17)$$

Cette relation fait intervenir 4 champs : les trois champs réels $\vec{\xi}(x)$ sans signification physique que l'on appelle bosons de Goldstone, et un quatrième champ $H(x)$ correspondant au boson de Higgs. Les champs $\vec{\xi}(x)$ peuvent être éliminés de la relation 1.17 avec un changement de jauge approprié : le choix de cette jauge donne leur masse aux bosons faibles $W^\pm$ et Z^0 car il existe un couplage entre le champ de Higgs et les bosons de l'interaction électrofaible faisant apparaître des termes quadratiques correspondant à une masse via la dérivée covariante :

$$\mathcal{L}_{Higgs} = [D_\mu \phi]^\dagger [D^\mu \phi] - V(\phi) \quad (1.18)$$

Où D_μ est la dérivée covariante électrofaible vue dans la relation 1.12. Les termes de masse apparaissant par cette dérivée expriment la masse des bosons faibles en fonction des constantes de couplage électrofaibles g et g' et de la valeur attendue du vide v , on peut alors mesurer ces paramètres en mesurant la masse des bosons faibles. La masse du boson de Higgs dépend de μ^2 ($M_H = \sqrt{-2\mu^2}$), elle reste donc inconnue tant que ce boson n'est pas observé expérimentalement.

Des termes de masse apparaissent pour les fermions car on postule un couplage de Yukawa entre le champ de Higgs et les champs fermioniques dont la formule simplifiée (en réalité les couplages diffèrent selon que l'on étudie un quark de type up ou de type down) est la suivante :

$$\mathcal{L}_{Yukawa} = -g_Y(\bar{\psi}_L\phi\psi_R + \bar{\psi}_R\phi^\dagger\psi_L) \quad (1.19)$$

Les ψ_L sont les doublets de fermions gauches, les ψ_R sont les singulets droits et les g_Y sont les couplages de Yukawa des 3 leptons chargés et des 6 quarks (Yukawa généralisé pour les quarks). Les constantes de couplage sont données par : $g_Y = \frac{\sqrt{2}m_f}{v}$ où m_f est la masse du fermion. Il y a donc un couplage plus fort entre le champ de Higgs et les fermions les plus massifs (par exemple le quark top).

1.3 Au-delà du Modèle Standard

Le Modèle Standard a prouvé sa validité aux énergies atteintes dans les collisionneurs jusqu'à maintenant, cependant il possède plusieurs limitations :

- Certains points de la relativité générale tels que la gravitation et l'énergie noire ne sont pas inclus dans ce modèle.
- Il ne fournit pas de candidat à la matière noire.
- Il ne tient pas compte de la masse des neutrinos établie par les oscillations[14].
- Il donne lieu à un problème de hiérarchie (instabilité de la masse du Higgs par rapport aux correction radiatives dans le Modèle Standard).

De plus l'existence du champ de Higgs assurant la brisure de symétrie électrofaible dans ce modèle n'a pas encore été prouvée, ni infirmée.

Plusieurs théories viennent étendre le Modèle Standard vers les énergies plus hautes (et le corriger en cas de non-existence du boson de Higgs). Ces modèles sont divers et nombreux comme par exemple la théorie des cordes, les dimensions supplémentaires, la supersymétrie, la technicouleur, etc. Nous nous intéresserons ici à la supersymétrie (SUSY) qui est un modèle plébiscité par les théoriciens.

1.3.1 Aspects de la supersymétrie

La supersymétrie introduit une symétrie supplémentaire entre les fermions et les bosons, le nombre quantique associé est la R-parité[15] qui vaut 1 pour les particules du Modèle Standard et -1 pour les particules supersymétriques.

Cette théorie introduit pour chaque particule un partenaire supersymétrique appelé sparticule. Le superpartenaire d'un fermion est un boson, et vice-versa. Certains superpartenaires tels que le photino ou le Zino n'ont pas d'existence physique et ne peuvent pas être observés expérimentalement, ces superpartenaires de jauge se mélangent et donnent de nouvelles sparticules qui sont, elles, observables (Table 1.1).

TABLE 1.1 – Particules du Modèle Minimal Supersymétrique Standard[15].

Nom	Spin	R-parité	Valeurs propres de masse	Valeurs propres de jauge
Bosons de Higgs	0	+1	$h^0, H^0, A^0, H^\pm$	$H_u^0, H_d^0, H_u^\pm, H_d^\pm$
squarks	0	-1	$\tilde{u}_L, \tilde{u}_R, \tilde{d}_L, \tilde{d}_R$ $\tilde{s}_L, \tilde{s}_R, \tilde{c}_L, \tilde{c}_R$ $\tilde{t}_1, \tilde{t}_2, \tilde{b}_1, \tilde{b}_2$	$\tilde{u}_L, \tilde{u}_R, \tilde{d}_L, \tilde{d}_R$ $\tilde{s}_L, \tilde{s}_R, \tilde{c}_L, \tilde{c}_R$ $\tilde{t}_L, \tilde{t}_R, \tilde{b}_L, \tilde{b}_R$
sleptons	0	-1	$\tilde{e}_R^\pm, \tilde{e}_L^\pm, \tilde{\nu}_e$ $\tilde{\mu}_R^\pm, \tilde{\mu}_L^\pm, \tilde{\nu}_\mu$ $\tilde{\tau}_1, \tilde{\tau}_2, \nu_\tau$	$\tilde{e}_R^\pm, \tilde{e}_L^\pm, \tilde{\nu}_e$ $\tilde{\mu}_R^\pm, \tilde{\mu}_L^\pm, \tilde{\nu}_\mu$ $\tilde{\tau}_L, \tilde{\tau}_R, \nu_\tau$
neutralinos	1/2	-1	$\tilde{\chi}_1^0, \tilde{\chi}_2^0, \tilde{\chi}_3^0, \tilde{\chi}_4^0$	$\tilde{B}^0, \tilde{W}^0, \tilde{H}_u^0, \tilde{H}_d^0$
charginos	1/2	-1	$\tilde{\chi}_1^\pm, \tilde{\chi}_2^\pm$	$\tilde{W}^\pm, \tilde{H}_u^\pm, \tilde{H}_d^\pm$
gluino	1/2	-1	$\tilde{g}$	$\tilde{g}$
gravitino	3/2	-1	$\tilde{G}$	$\tilde{G}$

On peut faire sur la table 1.1 les observations suivantes :

- Le secteur de Higgs du Modèle Standard doit être modifié afin de s'accorder au modèle minimal supersymétrique, il y a 5 états propres de masse de higgs : h^0, H^0, H^+, H^- et A^0 .
- Il y a 4 neutralinos χ_i^0 : les états propres de jauge *Bino* $\tilde{B}^0$, *Wino* $\tilde{W}^0$ et *Higgsinos neutres* $\tilde{H}_u^0/\tilde{H}_d^0$ se mélangent afin de former les états propres de masse des neutralinos. Les *Bino* et *Wino* sont les partenaires SUSY des B^0 et W_0 qui sont les états propres de jauge des bosons électrofaibles du Modèle Standard.
- Il y a 2 charginos $\chi_i^\pm$, les états propres de masse des charginos sont dûs à des mélanges des états propres de jauge des *Wino* $\tilde{W}^\pm$ et des *Higgsinos chargés* $\tilde{H}_u^\pm/\tilde{H}_d^\pm$.
- Dans le Modèle Standard un fermion de chiralité gauche a la même masse que celui de chiralité droite et on peut passer de l'un à l'autre par renversement du spin. Dans la supersymétrie les superpartenaires de ces fermions sont de spin-0 ce qui conduit le superpartenaire d'un fermion droit à être totalement différent du superpartenaire d'un fermion gauche. Les sfermions apparaissent alors dans des états différents et avec des masses différentes selon la chiralité de leur partenaire du Modèle Standard.

- Le point ci-dessus admet des exceptions pour le neutrino ne pouvant pas être de chiralité droite. De plus pour les quarks beaux et top ainsi que pour le τ les états de chiralité droit et gauche se mélangent en état 1 et 2 dû à la masse importante de ces fermions.

Les intérêts de cette théorie sont les suivants :

- Elle unifie les constantes de couplages à l'échelle de grande unification. (Fig. 1.3.1).
- Elle résoud le problème de hiérarchie du Modèle Standard.
- Si la R-parité est conservée, elle fournit un candidat à la matière noire, la particule supersymétrique la plus légère : le neutralino 1 $\tilde{\chi}_1^0$ (c.f. table 1.1).
- Il est possible dans certains cas d'y inclure la gravitation.

Les expérimentateurs s'intéressent tout particulièrement au modèle minimal supersymétrique (MSSM). La supersymétrie doit être une symétrie brisée car les sparticules et leurs partenaires du Modèle Standard n'ont pas la même masse : la preuve expérimentale est le fait que les sparticules n'ont pas été observées dans les expériences précédentes explorant les basses masses. Le mécanisme de brisure étant inconnu, il y a 127 paramètres libres dans la supersymétrie générale. Le mécanisme mSUGRA (minimal SuperGRAvity) restreint le modèle supersymétrique à cinq paramètres (en plus de ceux du MS) et imposant un mécanisme pour la brisure de supersymétrie :

- **m_0 et $m_{1/2}$** : ce sont respectivement la masse commune des scalaires (spartenaires des fermions du MS) et des Gauginos (spartenaires des bosons) à l'échelle de grande unification.
- **A_0** : couplage scalaire universel trilinéaire à l'échelle de grande unification.
- **$\tan\beta$** : Rapport des valeurs attendues du vide des champs de Higgs (il y en a deux dans la théorie SUSY).
- **Signe de μ** : Signe du paramètre de masse de Higgs-Higgsino.

1.3.2 Recherche expérimentale de la supersymétrie

Un aspect particulièrement intéressant de cette théorie pour les expérimentateurs est sa vérifiabilité. Cette théorie prédit des particules dont la masse est de l'ordre du TeV/ c^2 appelée sparticules. Leur masse ouvre la possibilité de leur observation dans le LHC (Large Hadron Collider, c.f. chapitre 2).

Le choix des 5 paramètres cités dans la section 1.3.1 est libre tant qu'aucune sparticule n'a été observée (certaines contraintes sur ces paramètres ont pu être établies par les expériences menées au LEP[17]). Pour les recherches on définit un certain nombre de points de travail dits de basse masse (Low Mass points : LM) et de haute masse (HM), afin d'explorer des phénoménologies différentes. Ces points sont resumés dans le tableau

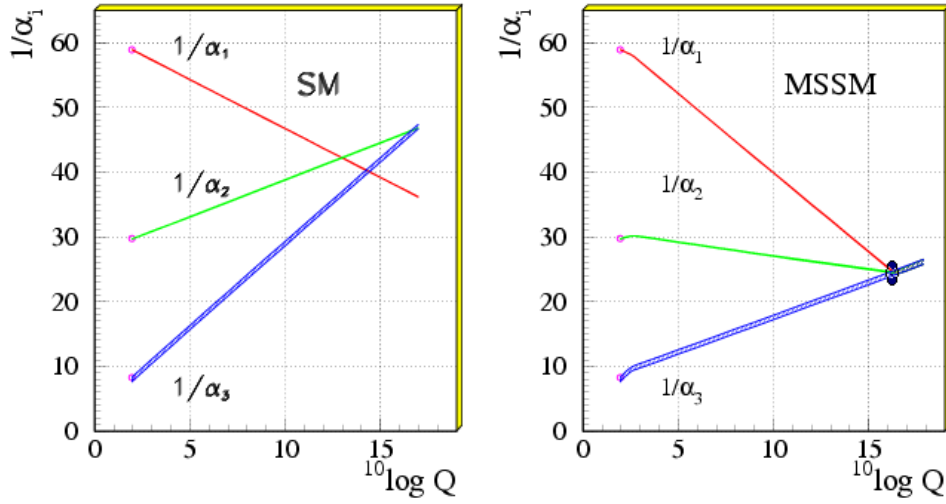


FIGURE 1.4 – Couplages de jauge ($1/\alpha$) pour le Modèle Standard (à gauche) et le Modèle SuperSymétrique Minimal (à droite). Le point noir représente les corrections dépendantes du Modèle en fonction de Q (Source [16]).

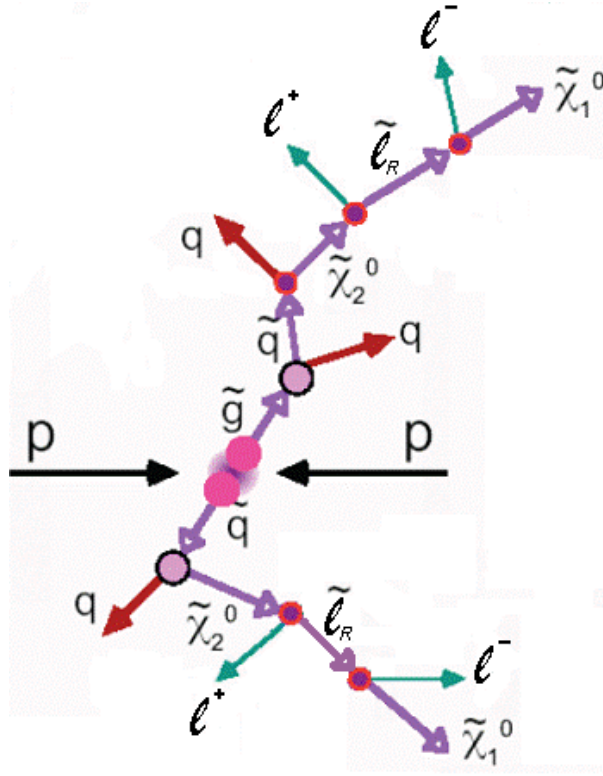
1.2.

TABLE 1.2 – Points de travail mSUGRA

Point	m_0 (GeV)	$m_{1/2}$ (GeV)	$\tan\beta$	A_0	$\text{signe}(\mu)$
LM1	60	250	10	0	+
LM2	185	350	35	0	+
LM3	330	240	20	0	+
LM4	210	285	10	0	+
LM5	230	360	10	0	+
LM6	85	400	10	0	+
LM7	3000	230	10	0	+
LM8	500	300	10	-300	+
LM9	1450	175	50	0	+
LM10	3000	500	10	0	+
HM1	180	850	10	0	+
HM2	350	800	10	0	+
HM3	700	800	10	0	+
HM4	1350	600	10	0	+

La production de particules supersymétriques (typiquement des squarks et/ou de gluinos) donne lieu à des cascades de désintégration très riches en particules. On a alors des signatures très claires pour la recherche de SUSY : une énergie transverse manquante élevée due à la non-détection des neutralinos stables en bout de chaîne (Fig 1.5), une multiplicité de jets élevée (au moins deux dus aux quarks) et de nombreux leptons.

La nature de ces cascades est dépendant des paramètres mSUGRA choisis, il y a cependant des processus communs à tous les points de travail :

FIGURE 1.5 – Cascade supersymétrique dans le processus $pp \rightarrow q\tilde{q}\tilde{g}$.

- La R-Parité est conservée.
- La particule supersymétrique la plus légère (donc stable) est le neutralino $\tilde{\chi}_1^0$.

Le choix des paramètres mSUGRA fixe la hiérarchie des masses et permet ou non certaines désintégrations :

- Les gluinos se désintègrent en squark ou vice-versa (dépendant de la hiérarchie des masses).
- La masse d'un squark dépend de sa chiralité, les squarks “droits” ont une désintégration quasiment exclusive en $\tilde{\chi}_1^0$: $\tilde{q}_R \rightarrow q\tilde{\chi}_1^0$. Les squarks “gauches” peuvent se désintégrer de la même façon que les squarks “droits” vers le neutralino 1 mais ont une fraction non-négligeable de désintégration en $\tilde{\chi}_2^0$: $\tilde{q}_L \rightarrow q\tilde{\chi}_2^0$ ou $\tilde{\chi}_1^\pm$: $\tilde{q}_L \rightarrow q\tilde{\chi}_1^\pm$.
- Les sleptons peuvent aussi être produits directement. Leurs largeurs de désintégration dépendent aussi de leur chiralité (c.f. table 1.1). Les sleptons “droits” se désintègrent quasiment exclusivement suivant $\tilde{l}_R \rightarrow l\tilde{\chi}_1^0$. Pour les sleptons “gauches”, tout dépend de la hiérarchie des masses. Les désintégrations possibles sont résumées dans la table 1.3.
- La hiérarchie des masses pour un type de superpartenaire (squark, sleptons) est souvent inversée par rapport aux particules du Modèle Standard, le sup (partenaire supersymétrique du up) est plus lourd que le stop (partenaire supersymétrique du top), les

sélectrons sont plus lourds que les staus...

TABLE 1.3 – Désintégrations supersymétriques impliquant des sleptons gauches (de [18]).

sparticule	modes de désintégration si $m_{\tilde{l}_L^\pm} > m_{\tilde{\chi}}$	modes de désintégration si $m_{\tilde{l}_L^\pm} < m_{\tilde{\chi}}$
$\tilde{l}_L^\pm$	$\tilde{l}_L^\pm \rightarrow l^\pm \tilde{\chi}_1^0$ $\tilde{l}_L^\pm \rightarrow \nu_l \tilde{\chi}_1^{\pm 0}$	$\tilde{\chi}_2^0 \rightarrow \tilde{l}_{R,L}^\pm l^\mp$ $\tilde{\chi}_1^\pm \rightarrow \tilde{l}_{R,L}^\pm \nu$
$\tilde{\nu}$	$\tilde{\nu} \rightarrow \nu_l \tilde{\chi}_2^0$ $\tilde{\nu} \rightarrow l^\pm \tilde{\chi}_1^\mp$	$\tilde{\chi}_2^0 \rightarrow \tilde{\nu} \nu$ $\tilde{\chi}_1^\pm \rightarrow \tilde{\nu} l^\pm$

Les désintégrations de sleptons vers les leptons impliquent très souvent leurs leptons associés. Les staus qui sont les sleptons les plus légers se désintègreront vers les taus.

De plus le neutralino $\tilde{\chi}_2^0$ et le chargino $\tilde{\chi}_1^\pm$ admettent des désintégrations en particules du Modèle Standard et en un neutralino $\tilde{\chi}_1^0$ [19] :

- Autres modes de désintégration du neutralino $\tilde{\chi}_2^0$ pouvant contribuer à la production de lepton τ :
 $\tilde{\chi}_2^0 \rightarrow Z^0 \tilde{\chi}_1^0 \rightarrow \tau^+ \tau^- \tilde{\chi}_1^0$.
 $\tilde{\chi}_2^0 \rightarrow H^0 \tilde{\chi}_1^0 \rightarrow \tau^+ \tau^- \tilde{\chi}_1^0$.
 $\tilde{\chi}_2^0 \rightarrow l^+ l^- \tilde{\chi}_1^0$, désintégration directe en leptons avec la possibilité d'obtention de τ .
- Autres modes de désintégration du chargino $\tilde{\chi}_1^\pm$ pouvant contribuer à la production de lepton τ :
 $\tilde{\chi}_1^\pm \rightarrow W^\pm \tilde{\chi}_1^0 \rightarrow \tau^\pm \nu_\tau \tilde{\chi}_1^0$
 $\tilde{\chi}_1^\pm \rightarrow H^\pm \tilde{\chi}_1^0 \rightarrow \tau^\pm \nu_\tau \tilde{\chi}_1^0$
 $\tilde{\chi}_1^\pm \rightarrow l^\pm \nu \tilde{\chi}_1^0$

Les points de travail de l'espace mSUGRA favorisant l'apparition de leptons τ sont les points LM1 et LM2 et leurs phénoménologies sont les suivantes :

• **LM1**

$m(\tilde{g}) \geq m(\tilde{q})$, le processus $\tilde{g} \rightarrow \tilde{q}q$ est dominant.

Rapports d'embranchements : $BR(\tilde{\chi}_0^2 \rightarrow \tilde{l}_R l) = 11, 2\%$, $BR(\tilde{\chi}_0^2 \rightarrow \tilde{\tau}_1 \tau) = 46\%$, $BR(\tilde{\chi}_1^\pm \rightarrow \tilde{\nu}_l) = 36\%$.

• **LM2**

$m(\tilde{g}) \geq m(\tilde{q})$, le processus $\tilde{g} \rightarrow \tilde{q}q$ est dominant ($\tilde{b}_1 b$ représente 25% des désintégrations de gluinos).

Rapports d'embranchements : $BR(\tilde{\chi}_0^2 \rightarrow \tilde{\tau}_1 \tau) = 96\%$ (dû à la valeur élevée de $\tan\beta$ qui favorise les désintégrations vers les stau), $BR(\tilde{\chi}_1^\pm \rightarrow \tilde{\tau}_1 \nu) = 95\%$.

Les autres points impliquent surtout des désintégrations vers les quarks top, les bosons faibles W et Z, et vers les bosons de Higgs. Le lepton τ apparaît alors à une fréquence moins importante dans ces points.

Une des méthodes expérimentales de recherche consiste donc à rechercher une signature d'évènement SUSY en observant plusieurs leptons (jusqu'à 4) associé à des jets QCD et

une énergie transverse manquante élevée. Il faut alors être capable d'identifier efficacement chacun de ces objets physiques, en particulier les leptons τ .

1.4 Les collisions proton-proton

Une collision entre deux particules est caractérisée par une section efficace analogue à la section efficace géométrique définie pour une collision en mécanique classique. La section efficace de collision est exprimée en barns ($1\text{ b} = 10^{-24}\text{ cm}^2$) : un système d'unités mieux adapté à la physique des collisionneurs que le système international. Le taux de collision est relié à la section efficace par la relation :

$$\frac{dN}{dt} = \mathcal{L}\sigma \quad (1.20)$$

Où σ est la section efficace de collision des deux particules (ici proton contre proton). La section efficace d'une collision proton-proton dépend de l'énergie dans le centre de masse comme le montre la figure 1.6 (σ_{tot}). Le paramètre $\mathcal{L}$ est la luminosité instantanée : c'est un paramètre relié au collisionneur (exprimée en $\text{cm}^{-2}\text{s}^{-1}$) calculé de façon générale par la relation 1.21. Des termes correctifs dûs à l'instrumentation doivent être apportés par la suite (c.f. section 2.1).

$$\mathcal{L} = fn \frac{N_1 N_2}{A} \quad (1.21)$$

Avec

f : fréquence de révolution

n : nombre de paquets dans l'anneau

$N_i, i = 1, 2$: nombre de particule par paquet

A : aire de la section efficace de collision du faisceau

Les sections efficaces et la fréquence d'occurrence (pour une luminosité instantanée $\mathcal{L} = 10^{34}\text{ cm}^{-2}\text{s}^{-1}$) de divers processus du Modèle Standard sont présentés dans la figure 1.6.

La production de jets QCD est très largement dominante avec une section efficace ($\approx \sigma_{tot}$) très supérieure à celle des autres processus (électrofaible, $t\bar{t}$). Les processus physiques recherchés (par exemple le Higgs) sont très rares : leur section efficace est d'au moins de 10 ordres de magnitude en dessous de la section efficace totale proton-proton. Pour les observer nous sommes contraints d'effectuer un nombre suffisant de collisions et donc de fonctionner à des luminosités très élevées. En conséquence il est probable à haute luminosité que plusieurs collisions aient lieu simultanément (pile-up), mais il est rare d'obtenir plus d'une collision "dure" i.e. de haute énergie (collisions principale). Les autres collisions sont "molles" et séparables de l'évènement à condition que les vertices d'interaction soient très bien reconstruits.

• Collisions entre partons

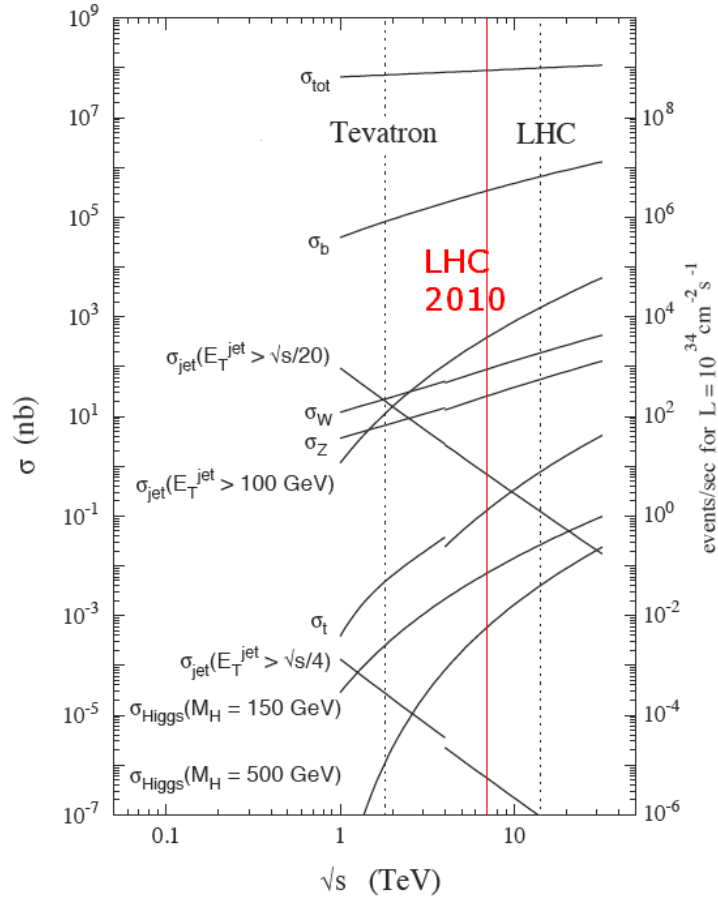


FIGURE 1.6 – Section efficace de différents processus du Modèle Standard en fonction de l'énergie dans le centre de masse (extrait de [20]) et de la nature de la collision ($\bar{p}p$ pour le Tevatron et pp pour le LHC).

Nous avons vu dans la section 1.2.1 que les collisions proton-proton à haute énergie impliquaient en réalité des collisions entre partons emportant une fraction de l'impulsion totale du proton donnée par les fonctions de densité de parton x . Ces densités ont pu être mesurées à l'aide d'expériences antérieures effectuées en collisionneur (HERA, EMC). L'évolution de ces densités en fonction de Q^2 ainsi que les interactions entre partons impliqués dans une collision sont décrites par la Chromodynamique Quantique [21] qui est la partie du Modèle Standard traitant l'interaction forte (On utilisera le terme anglophone QCD par la suite).

Le fait que les particules impliquées dans une collision soient des partons couplés à l'interaction forte a une conséquence expérimentale : la multiplicité de hadrons issue des collisions devient très importante pour la raison que les processus QCD deviennent dominants (Fig. 1.6). De plus d'autres phénomènes que nous verrons dans la section suivante (rayonnement de gluons, rupture de lien colorés) sont susceptibles d'augmenter la quantité de hadrons produits dans ces collisions.

1.4.1 La Chromodynamique Quantique dans les collisions pp

Les processus ayant lieu dans les collisions proton-proton sont calculés dans le cadre du Modèle Standard particulièrement par la QCD.

On décrira un processus $pp \rightarrow \text{interaction} \rightarrow F$ comme une interaction entre deux partons qui peuvent être un quark, un antiquark ou un gluon (Fig. 1.7). F est l'état final de la collision, c'est l'ensemble des particules obtenues après collision et que l'on pourra mesurer à l'aide d'un détecteur.

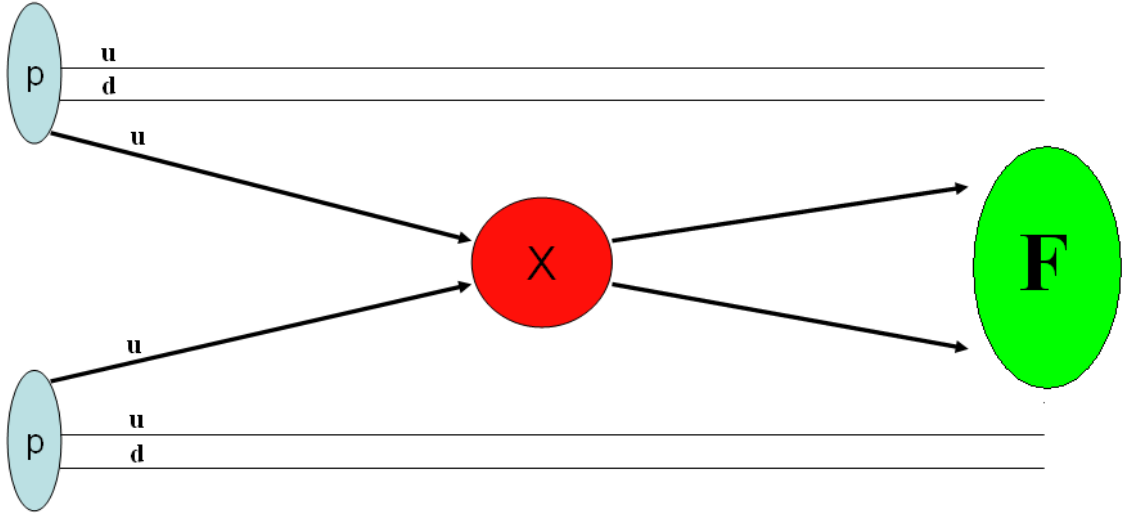


FIGURE 1.7 – Schématisation simplifiée d'une collision $pp \rightarrow F$.

Le processus donnant un état final F peut être soit un processus purement QCD par exemple la diffusion élastique de deux quarks (Fig. 1.9a), soit un processus électrofaible (par exemple $q\bar{q} \rightarrow Z^0$). D'autres processus sont envisageables au-delà du Modèle Standard (par exemple les événements SUSY décrits dans la section 1.4), mais nous resterons pour l'instant dans le cadre du Modèle Standard.

À ce processus peuvent s'ajouter des phénomènes radiatifs calculables perturbativement : des gluons sont irradiés par les partons (un gluon peut irradier un autre gluon) avant et après collision (Fig. 1.8) on parle alors de radiation d'état initial (Initial State Radiation : ISR) avant collision et de radiation d'état final (Final State Radiation : FSR) après collision.

Les gluons rayonnés par un parton peuvent être réabsorbés par ce même parton (boucle quantique, Fig. 1.9b). Ce phénomène aura une influence sur la section efficace de collision de ce parton, ces fluctuations quantiques sont aussi responsables de paires $q\bar{q}$ virtuelle (la "mer") dans le proton. Un gluon réel peut aussi être irradié à partir d'un parton et créer un jet.

Les quarks et les gluons ne peuvent pas exister sans être confinés dans un hadron, ils donneront lieu au phénomène de fragmentation puis d'hadronisation qui sont non pertur-

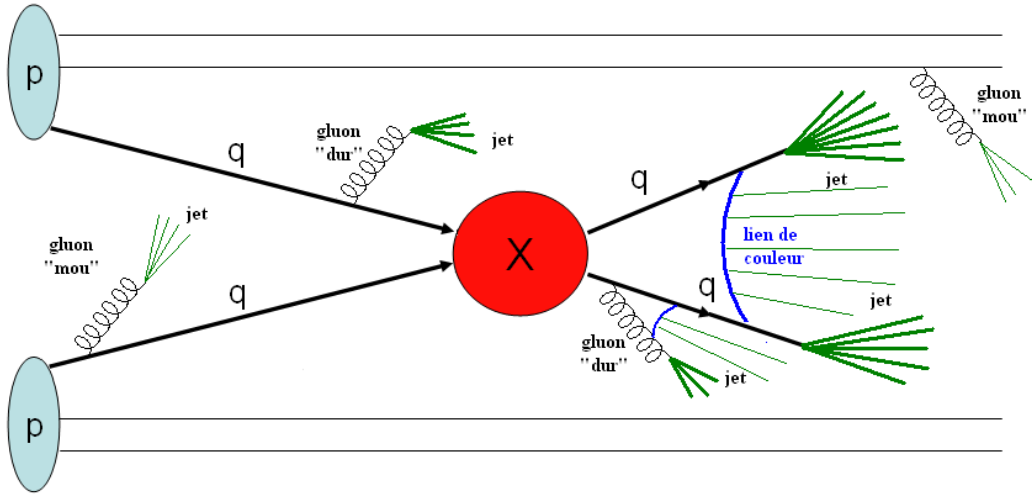


FIGURE 1.8 – Schématisation d’une collision $pp \rightarrow F$ avec radiation dans l’état initial et final. F est un état final avec fragmentation des quarks.

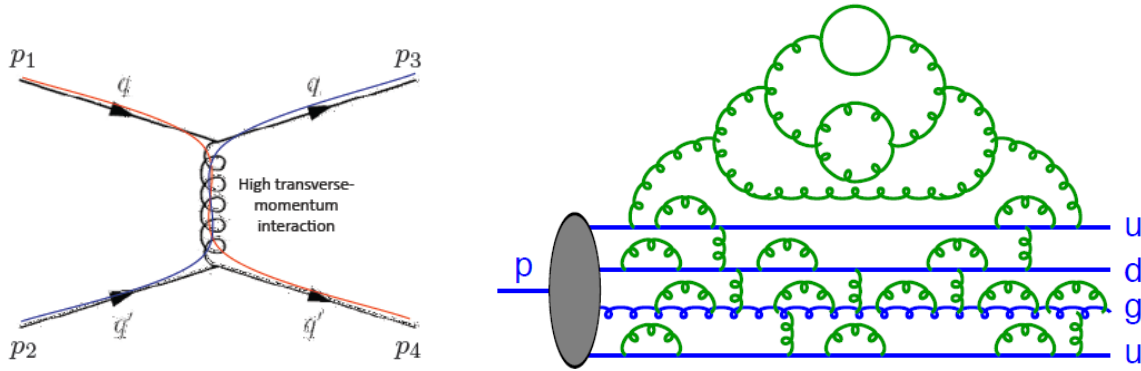


FIGURE 1.9 – **a)** Diffusion élastique entre deux quarks (voie t), **b)** Fluctuations partoniques dans un proton.

batifs. Ces phénomènes débouchent sur la création de gerbes de particules appelés “jets”, les quarks issus de la collision ainsi que les gluons rayonnés (ISR et FSR) créeront chacun un jet. Le processus partonique de haute énergie (QCD ou électrofaible) peut donc être accompagné d’un certain nombre de jets[22]. Par exemple dans un processus électrofaible (production de bosons W ou Z) l’irradiation d’un gluon “dur” arrive dans 10% des cas et ce chiffre est divisé par 10 à chaque nouveau gluon (1% pour deux gluons, 0,1% pour trois gluons, etc. c.f. Fig. 1.10).

La fragmentation et l’hadronisation créent aussi des particules supplémentaires hors jet dues à l’élongation du lien coloré qui existe entre deux quarks, ce point sera détaillé dans la section 1.3.1.2.

• Factorisation

La factorisation permet de diviser le calcul de la section efficace d’un processus QCD en

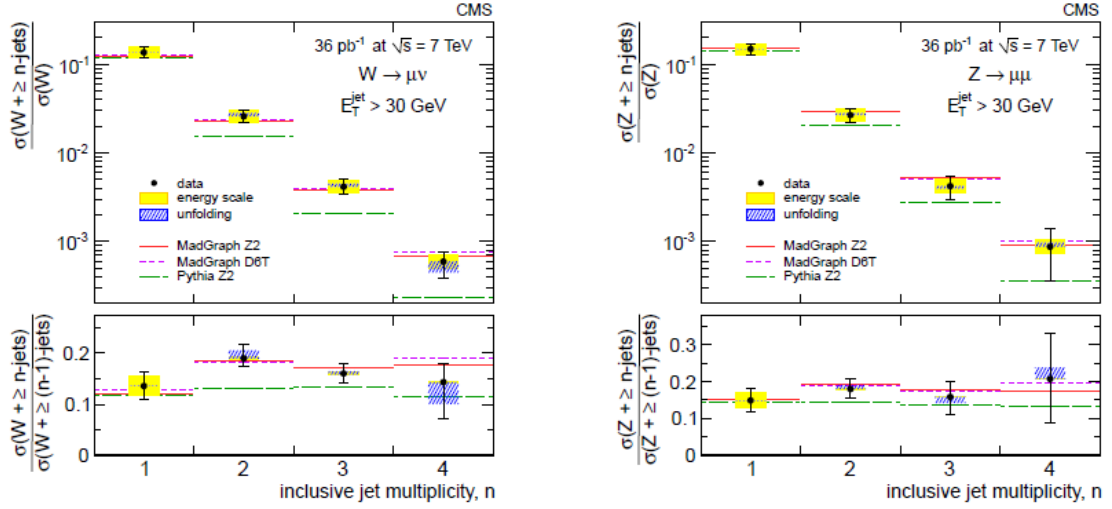


FIGURE 1.10 – Sections efficaces de production inclusives relatives des bosons W/Z + n jets mesurés par CMS[23] en 2010 (désintégration en muon).

deux parties : l'une concerne la partie partonique à haute énergie et est calculable perturbativement. L'autre concerne la sélection du parton impliqué dans le collision dans la phase initiale, et la fragmentation ainsi que l'hadronisation dans la phase finale. Cette partie du calcul est approximative et est paramétrée expérimentalement (tuning). La factorisation peut être appliquée successivement afin de subdiviser le calcul complet en termes plus simples et relativement indépendants.

En utilisant la factorisation colinéaire[25], la section efficace différentielle pour une observable $\mathcal{O}$ (par exemple la production d'un jet d'impulsion $\mathcal{O}$) dans les collisions hadron-hadron est la suivante :

$$\frac{d\sigma}{d\mathcal{O}} = \sum_{a,b} \int_0^1 dx_a dx_b \sum_F \int d\Phi_F f_a^{h_1}(x_a, \mu_F) f_b^{h_2}(x_b, \mu_F) \frac{d\hat{\sigma}_{ab \rightarrow F}}{d\mathcal{O}} D_F(\hat{\mathcal{O}} \rightarrow \mathcal{O}, \mu_F) \quad (1.22)$$

La partie partonique est le terme $d\hat{\sigma}$ qui est la section efficace partonique traitée perturbativement, le terme à longue distance est le produit $f_a f_b D_F$. Les différents termes de l'équation (1.22) sont explicités ainsi :

- La somme externe est celle de toutes les combinaisons de constituants partoniques a et b contenus dans les hadrons h_1 et h_2 respectivement, se collisionnant. La somme interne est celle de tous les processus donnant des états finals $ab \rightarrow F$ intégrés sur l'espace des phases $d\Phi_F$ des états finals.
- Les $f_j^{h_i}(x_j, \mu_F)$ sont les fonctions de densité de partons de type j emportant une fraction d'impulsion totale x_j du hadron h_i , ces fonctions sont non-perturbatives et leurs paramètres ont été déduits des mesures expérimentales des fonctions de structures à différentes échelles μ_F (Fig. 1.1).
- Le terme $D_F(\hat{\mathcal{O}} \rightarrow \mathcal{O}, \mu_F)$ est une fonction de fragmentation dont les paramètres sont a priori inconnus, mais qui peuvent être contraints avec des modèles. La fonction de frag-

mentation paramètre la transition à une échelle μ_F de l'état partonique à l'observable hadronique (état hadronique final après fragmentation et hadronisation).

La QCD perturbative nous donne la section efficace de diffusion partonique $d\hat{\sigma}_{ab \rightarrow F}$ à ordre fixé[26] :

$$d\hat{\sigma}_{ab \rightarrow F} = \frac{1}{2\hat{s}_{ab}} |\mathcal{M}_{ab \rightarrow F}(\Phi_F, \mu_F, \mu_R)|^2 \quad (1.23)$$

Où $|\mathcal{M}|^2$ est l'élément de matrice d'interaction au carré pour le processus $ab \rightarrow F$ (décrit par la phase ϕ_F) évalué aux échelles de factorisation et de renormalisation μ_F et μ_R . $\hat{s}_{ab} = (p_a + p_b)^2$ est l'énergie de la collision au carré dans le centre de masse des partons. Les $p_{a,b}$ sont les 4-impulsions des partons : $p_{a,b} = x_{a,b} P_{proton}$ où P_{proton} est la 4-impulsion du proton et $x_{a,b}$ la fraction emportée respectivement par les partons a et b impliqués dans la collision.

1.4.1.1 Modélisation de la gerbe partonique

La QCD perturbative nous permet de modéliser la création et l'évolution d'une gerbe partonique en apportant des corrections aux sections efficaces pour un processus donnant $pp \rightarrow F + N_{jets}$ (c.f. section 1.4.1). C'est la partie partonique décrite dans la section précédente (équation 1.22).

Schématiquement, on exprime la section efficace différentielle pour un observable $\mathcal{O}$ dans la production de $F + N_{jets}$ (suivant [21]) :

$$\left. \frac{d\sigma_F}{d\mathcal{O}} \right|_{ME} = \underbrace{\sum_{k=0}^{\infty} \int d\phi_{F+k}}_{\Sigma \text{ jambes}} \left| \underbrace{\sum_{l=0}^{\infty} \mathcal{M}_{F+k}^{(l)}}_{\Sigma \text{ boucles}} \right|^2 \delta(\mathcal{O} - \mathcal{O}(\phi_{F+k})) \quad (1.24)$$

k est le nombre de “jambes” i.e. le nombre de gluons rayonnés (Fig. 1.11) par le parton originel (tree-level).

l est le nombre de boucles, i.e. le nombre de gluons rayonnés et réabsorbés par le parton.

Les correction perturbatives QCD diverses peuvent être retrouvées tout en limitant les sommes de l'équation (1.24) en imposant certaines valeurs de $k + l$ qui sont :

- $k = 0, l = 0$: LO¹, pour la production inclusive de F.
- $k = n, l = 0$: LO, pour la production de F + n jets.
- $k + l \leq n$: N ^{n} pour F (incluant N ^{$n-1$} LO pour F + 1 jet, N ^{$n-2$} LO pour F + 2 jets, et ainsi de suite jusqu'à LO pour F + n jets).

• Les chaînes de Markov

Les générateurs d'évènements Monte Carlo utilisent une méthode itérative pour modéliser les gerbes partoniques : les chaînes de Markov.

1. Les physiciens des particules emploient les sigles anglophones LO (Leading Order) pour le cas sans perturbation, NLO (Next to Leading Order) pour le cas avec perturbations d'ordre 1, NNLO pour le cas avec perturbations d'ordre 2 etc...

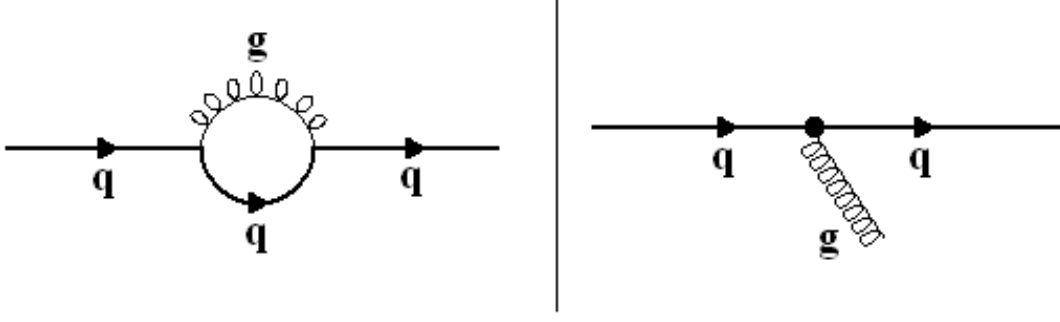


FIGURE 1.11 – **a)** Boucle quantique, le gluon est émis et absorbé par le même quark. **b)** Une “jambe”, un gluon est émis par le quark mais non réabsorbé.

Prenons la partie partonique (création de la gerbe partonique) de l’équation 1.22 sans perturbation (LO ou Born-level) pour un processus F (toujours suivant [21]), elle devient :

$$\left. \frac{d\sigma_F}{d\mathcal{O}} \right|_{\text{Born}} = \int d\phi_F |\mathcal{M}_F^{(0)}|^2 \delta(\mathcal{O} - \mathcal{O}(\phi_F)) \quad (1.25)$$

On remplace alors la fonction δ par un opérateur d’évolution $\mathcal{S}$ contenant implicitement la fonction δ , l’équation (1.25) devient alors :

$$\left. \frac{d\sigma_F}{d\mathcal{O}} \right|_{\mathcal{S}} = \int d\phi_F |\mathcal{M}_F^{(0)}|^2 \mathcal{S}(\phi_F, \mathcal{O}) \quad (1.26)$$

Cet opérateur est employé en chaîne itératives suivant un paramètre d’évolution Q_E . On discrétise l’évolution de la gerbe partonique en recalculant celle de chacun de ses éléments entre deux valeurs $Q_{E1} > Q_{E2}$:

$$\begin{aligned} \mathcal{S}(\phi_F, Q_{E1}, Q_{E2}, \mathcal{O}) &= \underbrace{\Delta(\phi_F, Q_{E1}, Q_{E2}) \delta(\mathcal{O} - \mathcal{O}(\phi_F))}_{F+0 \text{ exclusif à à-delà de } Q_{E2}} \\ &+ \underbrace{\sum_r \int_{Q_{E1}}^{Q_{E2}} \frac{d\phi_{F+1}^r}{d\phi_F} S_r(\phi_{F+1}) \Delta(\phi_F, Q_{E1}, Q_{E2}) \mathcal{S}(\phi_{F+1}, Q_{E1}, Q_{E2}, \mathcal{O})}_{F+1 \text{ inclusif à à-delà de } Q_{E2}} \end{aligned} \quad (1.27)$$

Où $\Delta(\phi_F, Q_{E1}, Q_{E2})$ est le facteur de *Sudakov*[24] définissant la probabilité qu’il n’y ait pas d’évolution (émission de parton) entre Q_{E1} et Q_{E2} :

$$\Delta(\phi_F, Q_{E1}, Q_{E2}) = \exp \left[- \sum_r \int_{Q_{E1}}^{Q_{E2}} \frac{d\phi_{F+1}^r}{d\phi_F} S_r(\phi_{F+1}) \right] \quad (1.28)$$

Le terme S_r est la fonction de radiation (splitting kernel), traditionnellement donnée par les équations de DGLAP¹.

L’équation (1.28) introduit le terme $S_r(\phi_{F+1})$ représentant la future possible évolution du système, et donc, l’itération suivante. La gerbe partonique n’évolue plus lorsque Q_E est de l’ordre du GeV, elle donne ensuite lieu aux phénomènes de fragmentation et d’hadronisation qui sont décrits dans la section suivante.

1. Dokshitzer-Gribov-Lipatov-Atarelli-Parisi

1.4.1.2 Fragmentation et hadronisation

Nous avons vu dans la section 1.2.2.2 que le couplage fort entre deux quarks croît avec la distance qui les sépare. Il apparaît un potentiel d'un champ coloré dipolaire (champ de l'interaction forte entre deux charges "de couleur") qui croît linéairement avec la distance entre ces charges si celle-ci dépasse 1 fm. Il survient alors les phénomènes de fragmentation (création de paires quark-antiquark) et de confinement (hadronisation) contrebalançant l'augmentation du potentiel fort croissant entre deux quarks qui s'éloignent. Ces phénomènes sont décrits par le modèle en cordes de l'hadronisation.

Plusieurs modèles basés sur des cordes ont été proposés dans les années 70-80. Le plus utilisé de nos jours est le modèle de Lund[27].

• Le modèle de Lund

On assimile le champ coloré à un élastique de tension κ . Le potentiel de ce champ augmente linéairement avec la distance séparant les chargés colorés, similairement à un élastique en élongation :

$$V(r) = \kappa r \quad (1.29)$$

Avec κ de l'ordre de $1 \text{ GeV}/\text{fm} \sim 0.2 \text{ GeV}^2$.

Toujours par analogie avec un élastique qui romps lorsque son énergie potentielle devient trop élevée, le lien coloré peut aussi rompre en convertissant son énergie potentielle en masse ($E_{Pot} > m_{q\bar{q}}$), créant ainsi une paire quark-antiquark. On aura deux nouveaux dipôles colorés qui pourront à leur tour se rompre si leur longueur devient trop importante, etc. (c.f Fig 1.4.1.2).

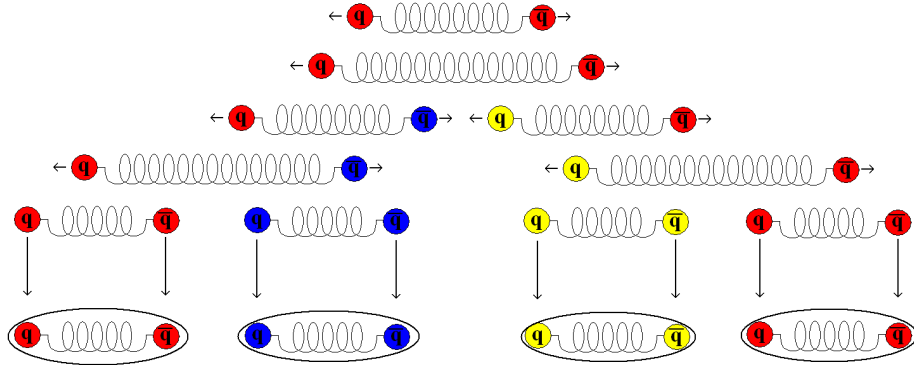


FIGURE 1.12 – Fragmentation puis hadronisation des quarks par le modèle en cordes.

Les valeurs caractéristiques des quarks (masse, impulsion transverse) sont données par une fonction de probabilité (suivant [27]) :

$$\text{Prob}(m_q^2, p_{\perp q}^2) \propto \exp\left(\frac{-\pi m_q^2}{\kappa}\right) \exp\left(\frac{-\pi p_{\perp q}^2}{\kappa}\right) \quad (1.30)$$

La masse des quarks est quantifiée, le spectre en $p_{\perp}$ (impulsion transverse à l'axe défini par le lien coloré) est indépendant de la saveur du quark avec une valeur moyenne universelle :

$$\langle p_{\perp q}^2 \rangle = \sigma^2 = \kappa/\pi \sim (250 \text{ MeV})^2 \quad (1.31)$$

Les mésons reçoivent une contribution en $p_{\perp}$ des deux côtés de la rupture de la liaison, d'où :

$$\langle p_{\perp h}^2 \rangle = 2\sigma^2 \quad (1.32)$$

La production de baryons peut également être incorporée dans ce modèle, on passe dans ce cas par un état “diquark”, un état intermédiaire de deux quarks faiblement liés, qui se liera à un autre quark pour donner un baryon. A noter que le rapport du nombre de diquark/quark n'est à priori pas connu et doit être déterminé expérimentalement.

Le dernier paramètre à connaître est la fraction z d'impulsion longitudinale du quark se fragmentant emporté par le hadron crée. Ce paramètre est déterminé par la fonction de fragmentation dite “fonction de fragmentation symétrique de Lund” indépendante de l'ordre dans lequel on traite les brisures de liens (toujours suivant [27]) :

$$f(z) \propto \frac{1}{z} (a - z)^a \exp\left(-\frac{b(m_h^2 + p_{\perp h}^2)}{z}\right) \quad (1.33)$$

Les paramètres a et b sont des paramètres libres. L'équation (1.33) dépend aussi de la masse du hadron, impliquant une fonction de fragmentation plus dure pour les hadrons plus lourds.

Pour la modélisation de la formation de hadrons on procède itérativement comme suit (Fig. 1.4.1.2) :

- Un quark q_0 produit dans un processus émerge de la gerbe partonique avec une énergie en dessous d'une valeur critique Q_{seuil} à laquelle peut se produire l'hadronisation. Ce quark à une 3-impulsion $\vec{p} = (p_{\perp 0}, p_+)$ où la troisième composante est donnée par $p_+ = E + p_z$.
- Une paire $q_1 \bar{q}_1$ est créée avec une énergie transverse relative $p_{\perp 1}$.
- Le quark q_0 se lie avec l'antiquark $\bar{q}_1$ pour former un méson d'impulsion $\vec{p} = (p_{\perp 0} - p_{\perp 1}, z_1 p_+)$. Le hadron aura donc emporté une fraction z_1 du quark issu de la gerbe.
- Le quark restant q_1 peut se lier à son tour à un antiquark $\bar{q}_2$ d'une autre paire quark-antiquark créée à partir du vide. En emportant une fraction z_2 de l'impulsion restante soit $z_2(1 - z_1)p_+$.

1.4.2 Les interactions multiples

Il est possible au sein d'une seule collision proton-proton d'avoir des interactions multiples de partons c'est-à-dire que les collisions se font deux partons contre deux partons, trois contre trois, etc. avec une probabilité de plus en plus faible au fur et à mesure que les partons impliqués sont plus nombreux[28].

Une approche intuitive pour ce modèle consiste à ne plus considérer que ce sont des paquets de protons qui circulent dans les anneaux des collisionneurs, mais directement des

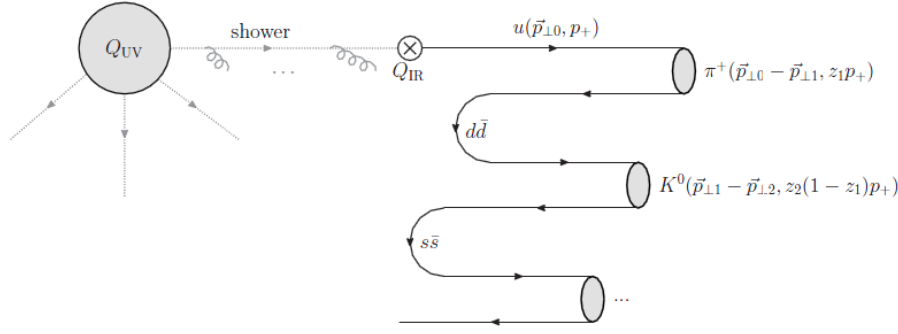


FIGURE 1.13 – Schéma de la sélection itérative des saveurs et de l’impulsion des quarks produits dans une fragmentation par le modèle en cordes de Lund.

paquets de partons : chaque collision parton-parton a une section efficace $\hat{\sigma}_{2 \rightarrow 2}$. Cette section efficace est celle de diffusion de 2 partons (d’où le terme $2 \rightarrow 2$) i.e. $qq' \rightarrow qq'$, $q\bar{q}' \rightarrow q\bar{q}'$, $q\bar{q} \rightarrow gg$, $qg \rightarrow qg$, $gg \rightarrow gg$ ou $gg \rightarrow q\bar{q}$ en tenant compte des corrections radiatives comme nous l’avons vu dans la section 1.4.1. Cette section efficace est intégrée entre un seuil $p_{\perp \min}$ et $p_{\perp \max}$ donnée par l’énergie dans le centre de masse $\sqrt{s}/2$ du collisionneur :

$$\hat{\sigma}_{2 \rightarrow 2}(p_{\perp \min}) = \int_{p_{\perp \min}}^{\sqrt{s}/2} \frac{d\sigma}{dp_{\perp}} dp_{\perp} = \langle n(p_{\perp \min}) \rangle \sigma_{\text{tot}} \quad (1.34)$$

A noter que la section efficace diverge lorsque $p_{\perp \min} \rightarrow 0$, entraînant aussi une section efficace $\sigma_{2 \rightarrow 2}$ supérieure à la section efficace totale (Fig. 1.4.2). Ceci s’interprète par le fait qu’une collision proton-proton contenant 2 interactions parton-parton est comptée deux fois pour $\sigma_{2 \rightarrow 2}$ mais une seule fois pour σ_{tot} . Le terme $\langle n \rangle(p_{\perp \min})$ nous donne la moyenne d’une distribution de Poisson du nombre d’interactions parton-parton au dessus de $p_{\perp \min}$ par collision hadron-hadron :

$$\mathcal{P}_n = [\langle n(p_{\perp \min}) \rangle]^n \frac{\exp[-\langle n(p_{\perp \min}) \rangle]}{n!} \quad (1.35)$$

Deux corrections importantes doivent encore être imposées à ce modèle :

1. La conservation de l’impulsion : la somme des impulsions de partons impliqués dans les interactions multiples ne peuvent pas dépasser celle du hadron mère. Ce qui a aussi pour conséquence de supprimer la large queue à haute valeurs de n (restriction du nombre de partons impliqués).
2. La possibilité de faire tendre $p_{\perp \min}$ vers 0 : la section efficace divergeant dans ce cas, on impose un “écranage de couleur” i.e. si la longueur d’onde $\sim 1/p_{\perp}$ d’un parton coloré échangé devient plus grande que la distance de séparation typique couleur-anticouleur. En conséquence on verra une chargée colorée “moyenne” (mélange de couleur) disparaissant lorsque $p_{\perp} \rightarrow 0$.

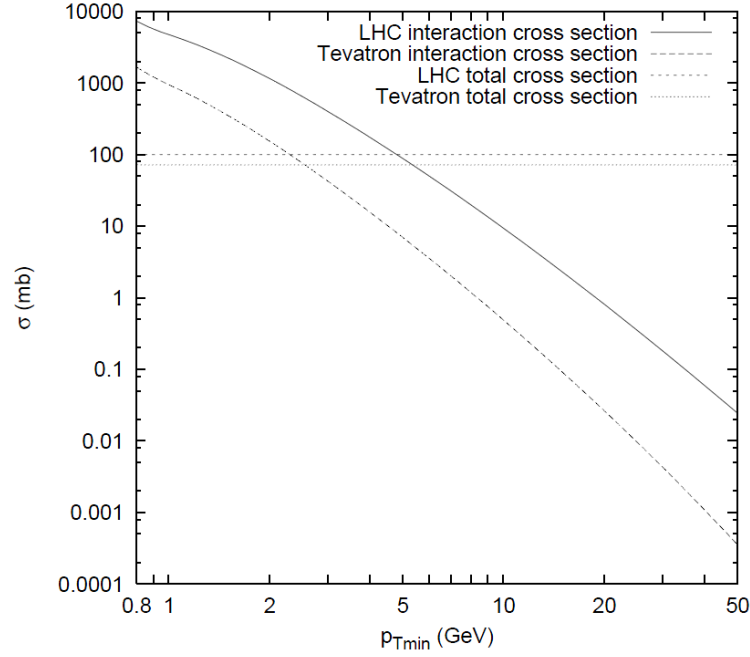


FIGURE 1.14 – Section efficace totale et section efficace par parton[28] en fonction de $p_{\perp\min}$ pour le Tevatron (collisions proton-antiproton à 2,36 TeV) et le LHC (collisions proton-proton à 14 TeV).

Phénoménologiquement, s'il se produit une interaction à partons multiples et que l'un des partons emporte une fraction x élevée de l'impulsion du proton, il en résultera que les autres partons emporteront une fraction x faible et produiront des collisions "molles". Pour la même raison, les hadrons issus de ces collisions "molles" entre partons auront une impulsion transverse faible et leur activité sera importante surtout dans les parties avant et arrière des détecteurs, mais ils participent cependant à l'évènement sous-jacent du fait du lien coloré existant entre ces deux partons.

• Activité hadronique dans les collisions proton-proton

Nous avons passé en revue les différents phénomènes conduisant à la production de hadrons dans les collisions proton-proton (rayonnement de gluons, fragmentation, évènement sous-jacent, collisions simultanées entre protons et collisions multiples entre partons). Tous ces phénomènes conduisent à la production de nombreux hadrons dans des collisions qui sont de faible énergie pour la plupart d'entre elles et à haute luminosité afin d'observer les processus rares.

L'observation de ces processus rares nécessite un collisionneur pouvant produire des collisions à très haute énergie et très haute luminosité, et des détecteurs très performants capables de distinguer les évènements dans un environnement hadronique à haute luminosité. Ce défi expérimental fut relevé et remporté par le collisionneur LHC et le détecteur CMS qui seront le sujet du chapitre suivant.

Chapitre 2

L'expérience CMS

2.1 Le Grand Collisionneur de Hadrons (LHC)

Le Grand Collisionneur de Hadrons ou Large Hadron Collider (LHC) consiste en l'intégration de deux synchrotrons dans une structure compacte commune (anneau constitué d'un assemblage de cryodipoles Fig. 2.1) effectuant des collisions proton-proton à très haute énergie. Situé au Centre Européen de la Recherche Nucléaire (CERN) à Genève, en Suisse, le LHC est le collisionneur le plus puissant jamais construit à ce jour. Le LHC et ses expériences associées sont le résultat d'une collaboration internationale de plus de 10000 scientifiques et ingénieurs venant de plus de 100 pays.

Les caractéristiques du LHC sont résumées dans la table 2.1. On y trouve ses dimensions, l'énergie portée par chaque proton, et les caractéristiques des faisceaux incluant les caractéristiques du point d'intersection de ces faisceaux. La luminosité instantanée $\mathcal{L}$ du LHC est donnée par la relation 2.1.

$$\mathcal{L} = \frac{\gamma f k_B N_p^2}{4\pi \epsilon_n \beta^*} F \quad (2.1)$$

Avec γ le facteur de Lorentz, f la fréquence de révolution, k_B le nombre de paquets, N_p le nombre de protons par paquet, ϵ_n l'emittance transverse normalisée du faisceau, β^* la fonction bêtatron au point d'interaction et F le facteur de réduction dû à l'angle de croisement des faisceaux.

La durée de vie de faisceau est le temps que l'on peut faire circuler ceux-ci sans observer une baisse notable de la luminosité : les faisceaux finissent par se dégrader à cause des interactions faisceau-faisceau (élargissement des faisceaux) causant une baisse d'intensité et donc de luminosité. Il faut alors stopper les faisceaux puis vider et remplir à nouveau le collisionneur pour une nouvelle prise de données.

Le nombre de collisions observé s'exprime en unité de luminosité intégrée (notée L) : une unité adaptée à la physique des hautes énergies car le nombre d'événements attendus pour un processus F avec une section efficace σ_F s'obtient par la relation :

$$N_F = \sigma_F \int \mathcal{L}.dt = \sigma_F \times L \quad (2.2)$$

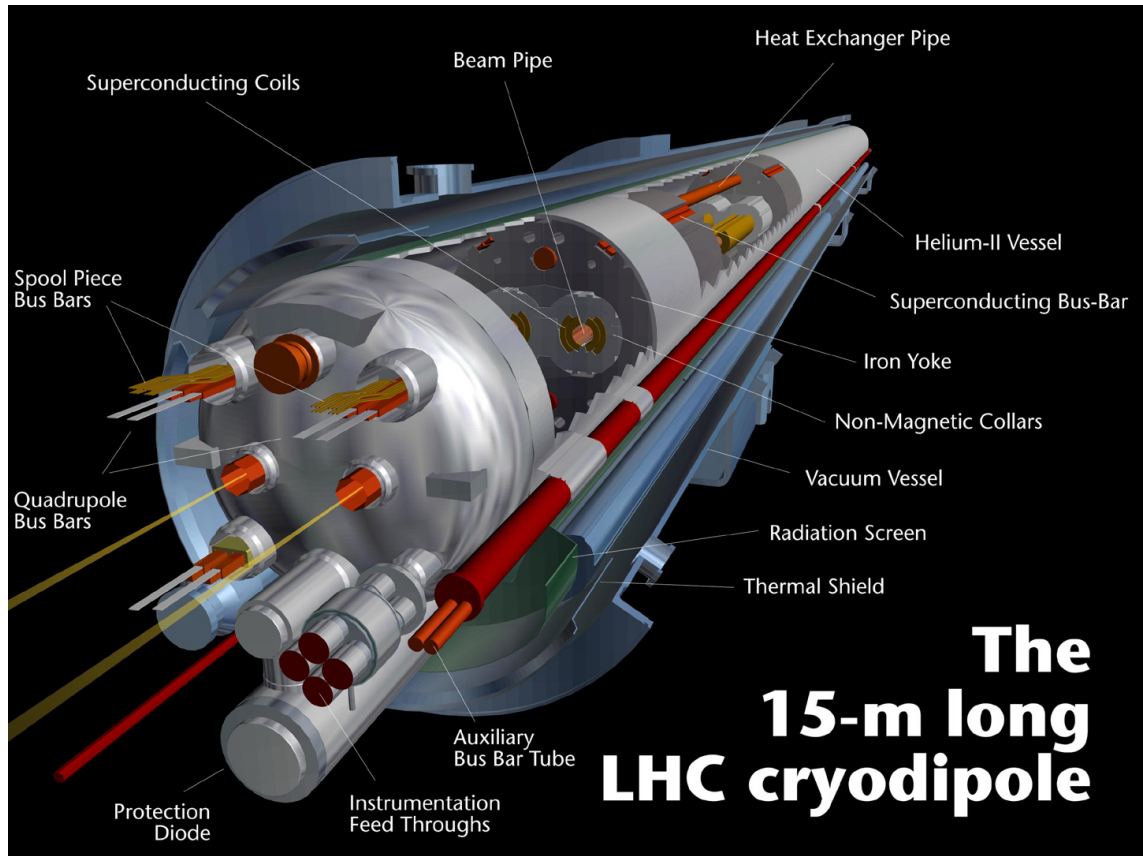


FIGURE 2.1 – Schéma en coupe d'un cryodipole du LHC[29].

TABLE 2.1 – Caractéristiques du LHC [30].

Circonférence		27 km
Profondeur		175 m
Énergie par proton	E	14 TeV
Énergie par proton en 2010-2011	E	7 TeV
Luminosité instantanée nominale	$\mathcal{L}_{nom}$	$10^{34} \text{ cm}^{-2} \cdot \text{s}^{-1}$
Temps de séparation des paquets		25 ns
Nombre de paquets	k_B	2808
Nombre de particules par paquet	N_P	1.15×10^{11}
Record de luminosité pour 2011	$\mathcal{L}$	$3,5 \cdot 10^{33} \text{ cm}^{-2} \cdot \text{s}^{-1}$
Collisions		
Valeur de β au PI	β^*	0.55 m
Emittance transverse normalisée du faisceau au PI	ϵ_n	$3,75 \text{ } \mu\text{m}$
Ecart-type sur la section du faisceau au PI	σ	$16.7 \text{ } \mu\text{m}$
Durée de vie moyenne du faisceau	τ_L	6-8 heures
Nombre de collisions/croisement (régime nominal)	n_c	≈ 20
Date de mise en service		20 Novembre 2009

La luminosité intégrée enregistré par CMS en 2010 et 2011 est montrée dans la figure 2.2, elle a dépassé les 5 fb^{-1} à la fin de l'année 2011.

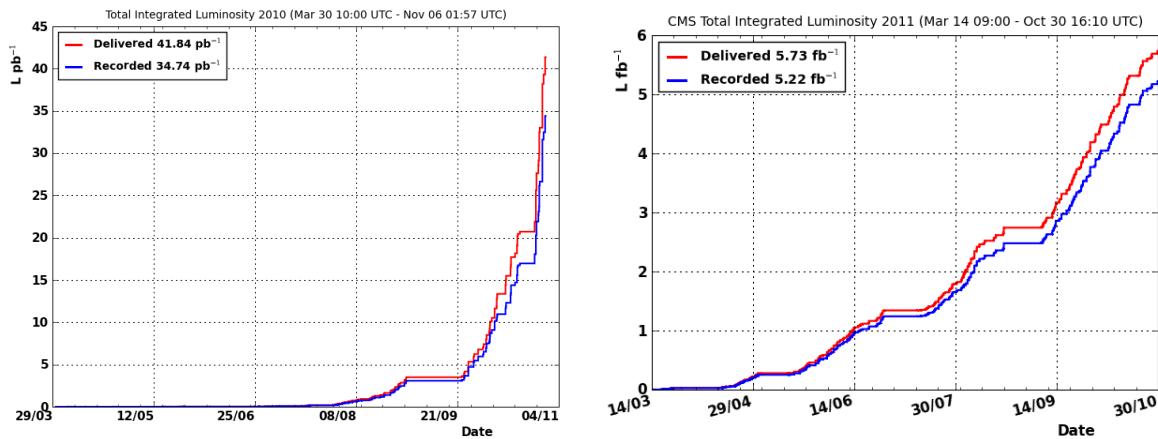


FIGURE 2.2 – Quantité de données délivrées par le LHC et enregistrées par CMS (provenant de [31]) en a) 2010, b) 2011.

Le LHC servira à vérifier les extensions au Modèle Standard qui se manifestent comme des événements rares superposés aux processus du Modèle Standard. Les sections efficaces des processus de MS ainsi que les prédictions concernant le boson de Higgs ont été présentées dans le chapitre 1 (Figure 1.6).

La section efficace totale σ_{tot} est équivalente à celle dite de “biais minimum”, ce sont pour la plupart des processus QCD mous, on fera alors souvent l’approximation $\sigma_{tot} \approx \sigma_{min\ bias}$. Les processus de biais minimum ont une section efficace 10^7 fois supérieure à celle de la production de bosons faibles W et Z, eux-mêmes ayant une section efficace 1000 fois supérieure à celle du top et 10000 fois supérieure à celle d’un boson de Higgs de 150 GeV. Les événements prédits par la supersymétrie ont une section efficace de production de l’ordre du femtobarn. Au pic du luminosité du LHC atteint en 2011 avec une luminosité de $\approx 3,5 \cdot 10^{33} \text{ cm}^{-2} \cdot \text{s}^{-1}$ (table 2.1) les événements de biais minimum ont une fréquence de $f_{min\ bias} \approx \mathcal{L} \times \sigma_{tot} = 350 \text{ MHz}$, les détecteurs présents sur le LHC doivent donc satisfaire aux deux prérequis suivants :

1. Ils doivent pouvoir faire la distinction entre deux croisements de paquets successifs (ce prérequis ne concerne pas tous les détecteurs situés sur le LHC).
2. Ils doivent être capables de distinguer les différents vertex d’interactions lors de collisions multiples entre protons, et de distinguer les collisions dures dans la grande multiplicité de collisions molles.

Quatre grandes expériences de physique des particules sont en cours sur le LHC : CMS, ATLAS, ALICE et LHCb. L’expérience CMS est présentée dans la section suivante.

2.2 Le Solénoïde Compact à Muons (CMS)

Le Solénoïde Compact à Muons ou Compact Muon Solenoid (CMS) est l’une des deux expériences (avec ATLAS) situées sur le LHC dédiées à la recherche du boson de Higgs

et de la physique au-delà du Modèle Standard. Les recherches conduites à l'aide de CMS sont les suivantes :

- Mesures de précision du Modèle Standard (MS) : Étude de la ChromoDynamique Quantique (QCD), du modèle électrofaible, de la physique des saveurs. Ces mesures servent de référence pour les recherches sur une physique au-delà du MS.
- Recherche du boson de Higgs : L'existence de ce boson fournit une explication à la brisure de la symétrie électrofaible dans le Modèle Standard ainsi qu'à la masse des bosons et des fermions.
- Recherche de la SUperSYmétrie (SUSY) : Une extension du Modèle Standard, les implications de ce modèle ont été décrites brièvement dans la section 1.3.
- Recherche de nouveaux bosons vecteurs massifs éventuels.
- Dimensions supplémentaires : théorie induisant l'existence de dimensions spatiales au-delà de 4.
- Physique des ions lourds : pour les recherches sur le Plasma de Quark et de Gluons.

2.2.1 Référentiel utilisé pour CMS

CMS est un détecteur de géométrie cylindrique typique pour les expériences menées sur les collisionneurs. Le détecteur CMS offre une couverture hermétique autour du point de collision de faisceaux car les particules issues des collisions vont dans toutes les directions de l'espace. Pour situer une particule dans l'espace on utilise le système de coordonnées sphérique (r, θ, ϕ) plutôt que le système de coordonnées cartésiennes (x, y, z) . Le point d'origine O du référentiel est le point nominal de croisement de faisceaux, l'axe z est celui des faisceaux incidents (c.f. Fig 2.3).

On définit alors le plan transverse xOy contenant l'angle ϕ , et le plan longitudinal xOz à l'axe des faisceaux contenant l'angle θ . La cinématique d'une particule peut être caractérisée de façon globale par trois paramètres : le module de son impulsion $\|\vec{p}\|$ et les deux angles polaire (θ) et azimutal (ϕ).

L'impulsion $\vec{p}$ d'une particule est décomposée en deux composantes : l'impulsion transverse P_T qui est la composante dans le plan xOy et la rapidité y définie par l'impulsion longitudinale P_L dans le plan xOz par la relation :

$$y = \frac{1}{2} \ln \left(\frac{E + P_L}{E - P_L} \right) \quad P_L : \text{Impulsion longitudinale} \quad (2.3)$$

Les particules étudiées sont ultra-relativistes, on se situe alors dans le cas limite où on peut faire les approximations $v \rightarrow c$ et $\beta = \frac{v}{c} \rightarrow 1$. On se situe alors dans le cas limite où $E \approx p$ et donc on remplace la rapidité y par la pseudo-rapacité η :

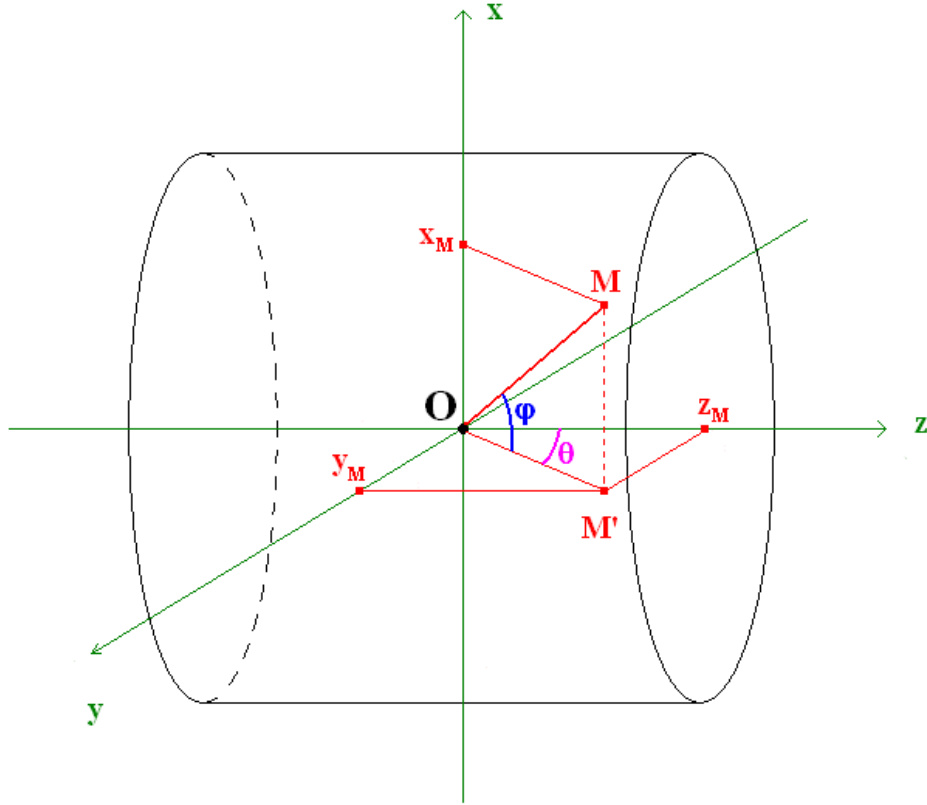


FIGURE 2.3 – Description du système de coordonnées utilisées dans CMS.

$$\eta = \frac{1}{2} \ln \left(\frac{|p| + P_L}{|p| - P_L} \right) = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right] \quad (2.4)$$

La plupart du temps il est impossible de mesurer la masse d'une particule car sa mesure nécessite la connaissance simultanée de l'impulsion et de l'énergie. On mesure alors la pseudo-rapacité d'une particule qui est une mesure valide pour les particules ultra-relativistes car l'énergie est remplacée par l'impulsion (scalaire) (équation 2.4). Le paramètre de pseudo-rapacité est égal à 0 si $\theta = \pm \frac{\pi}{2}$ (plan xOy) et infini si $\theta \rightarrow 0$ (plan xOz).

On utilise aussi l'impulsion transverse d'une particule et son énergie transverse : $E_T = E \sin(\theta)$ pour la raison que la composante totale du système $E_T = \sum_i \vec{E}_{T,i} = 0$ avant collision.

2.2.2 Conditions expérimentales requises

Les exigences nécessaires pour CMS afin d'atteindre les objectifs fixés par le programme de physique du LHC peuvent être résumées ainsi [32] :

- Une bonne identification des muons avec une couverture en rapidité s'étendant sur une région en $|\eta| < 2.5$, une bonne résolution sur la masse dimuons ($\approx 1\%$ à 100 GeV/c²),

et la possibilité de déterminer sans ambiguïté la charge des muons d'impulsion $p < 1 \text{ TeV}/c^2$.

- Une bonne résolution sur l'impulsion des particules chargées dans le trajectomètre interne et une bonne efficacité de reconstruction. Identification efficace des jets de τ et des quarks beaux. Excellente résolution requérant des détecteurs à pixels proches du point d'interaction pour reconstruire efficacement les vertex d'interaction et les vertex déplacés (secondaires) provenant de la désintégration d'une particule secondaire par exemple un hadron beau.
- Une bonne résolution en énergie électromagnétique, une bonne résolution sur la reconstruction de la masse diphoton et diélectron ($\approx 1 \%$ à $100 \text{ GeV}/c^2$), couverture en pseudo-rapacité large ($|\eta| < 2.5$), mesure de la direction des photons grâce à une localisation correcte des vertices d'interaction, isolation efficace des photons (et donc des $\pi^0 \rightarrow \gamma\gamma$) et des leptons à haute luminosité.
- Une bonne résolution sur l'énergie transverse manquante et la masse dijet, requérant des calorimètres hadroniques avec une large couverture hermétique ($|\eta| < 5$) et une segmentation latérale fine ($\Delta\eta \times \Delta\phi < 0.1 \times 0.1$)

2.2.3 Conception du détecteur CMS

Le détecteur CMS[33] est un détecteur cylindrique divisé en 4 grands systèmes de détection successifs entourant le point de collision (Fig. 2.4) qui sont en partant du point d'interaction : un système de trajectométrie d'une granularité très fine et composé de nombreuses couches, un calorimètre électromagnétique possédant une excellente résolution en énergie, un calorimètre hadronique hermétique (une couverture en angle solide de presque 4π stéradians) et un système de détection de muons performant.

Ce détecteur est plongé dans un champ magnétique intense de 4 Teslas produit par un aimant solénoïde supraconducteur afin de permettre la mesure des impulsions de particules chargées en mesurant le rayon de courbure de leur trajectoire.

2.2.3.1 Système de trajectométrie interne

C'est le système de détection le plus proche du point d'interaction. Il peut lui-même être divisé en trois sous-parties toutes dédiées à la trajectométrie des particules chargées :

- Trajectomètre pixel

Composé de trois couches concentriques en tonneau complétées par deux bouchons circulaires composés de disques de chaque côté. Les barils ont une longueur de 53 cm et des rayons respectifs de 4.4 cm, 7.3 cm et 10.2 cm. Les deux disques aux extrémités couvrent une surface délimitée par des rayons de 6 à 15 cm et sont placées à $|z| = 34.5 \text{ cm}$ et 46.5 cm .

Chaque pixel fait une taille de $100 \times 150 \mu\text{m}^2$. Un tonneau comprend 768 modules pixel arrangés en demi-échelles de 4 modules identiques. Les bouchons sont assemblés en forme

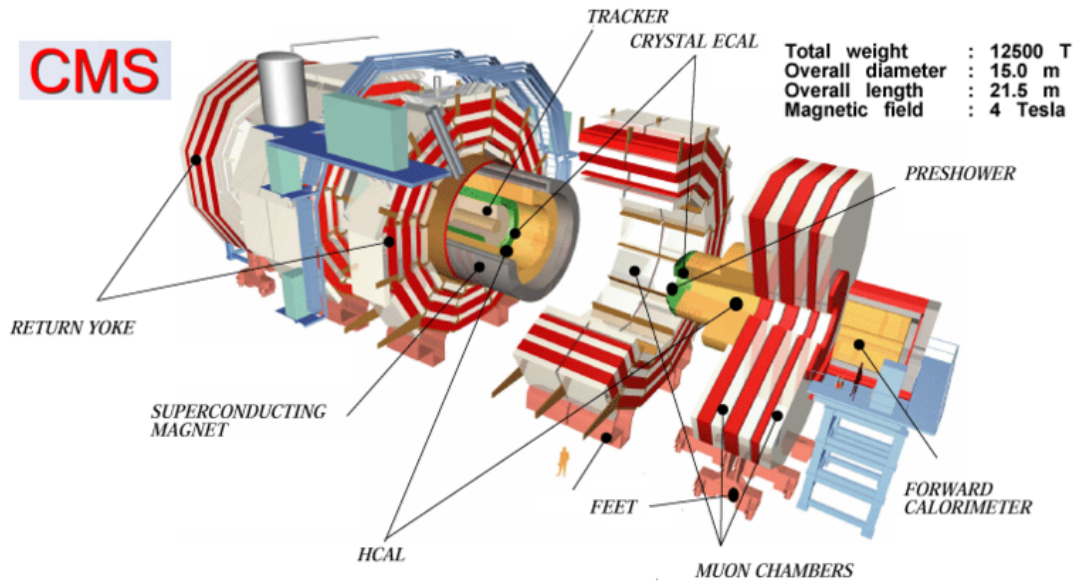


FIGURE 2.4 – Représentation en coupe du détecteur CMS.

de turbine et comprend 672 modules pixel avec 7 modules par lame. Ce détecteur utilise au total 66 millions de cellules de détection (pixel et silicium).

- Trajectomètre silicium à micropistes

Le trajectomètre est un assemblage de senseurs en silicium (micropistes). Il est divisé en deux parties (Fig. 2.5) : une partie interne à haute résolution et une partie externe de résolution très légèrement moindre.

- La partie interne est composée d'un tonneau de 4 couches (Tracker Inner Barrel : TIB) complétées par des bouchons de 3 disques (Tracker Inner Disks : TID) et occupe une longueur de $|z| < 65$ cm pour les tonneaux et $|z| < 110$ pour les disques. Chaque module à une surface active de 61×117 mm.
- La partie externe est composée de 6 couches complétées par 9 disques aux extrémités (Tracker End Cap : TEC) et occupent une distance par rapport aux vertex primaire de $|z| < 110$ cm pour les tonneaux et 120 cm $< |z| < 180$ pour les disques. Chaque module à une surface active de 94×186 mm.

Les caractéristiques des différentes parties sont données dans le tableau ci-dessous :

2.2.3.2 Calorimètre électromagnétique

Le calorimètre électromagnétique (ECAL) est un calorimètre hermétique et homogène dédié à la mesure des électrons et photons. Il consiste en un tonneau constitué de 61200 cristaux de tungstate de plomb (PbWO_4) fermé par des bouchons de 7324 cristaux chacun. La bonne performance du calorimètre électromagnétique est motivée par la recherche du boson de Higgs se désintégrant en deux photons, un état final important pour le Higgs de

TABLE 2.2 – Les différentes parties du trajectomètre

partie	nombre de détecteurs	épaisseur (μm)	pas d'une micropiste (μm)
TIB	2724	320	81/118
TOB	5208	500	81/183
TID	816	320	97/128/143
TEC	2512	320	96/126/128/143
TEC(2)	3888	500	143/158/183
Total	15148		

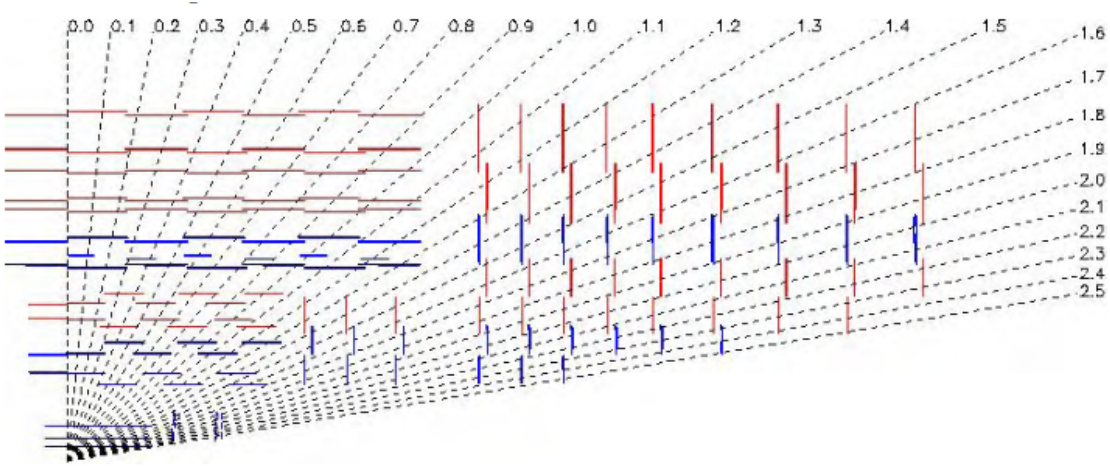


FIGURE 2.5 – Schéma du trajectomètre.

faible masse et reconnaissable dans l'environnement hautement hadronique du LHC.

Le choix s'est porté sur des scintillateurs en tungstate de plomb pour leur excellente résolution en énergie, leur courte longueur de radiation ($X_0 = 0.89 \text{ cm}$), leur réponse rapide (80% de la lumière est émise en moins de 25 ns) et leur tenue aux radiations (jusqu'à 10 MRad). Leur rendement est de 30 γ/MeV (30 photons sont émis pour 1 MeV déposé dans le cristal), ce signal est amplifié par des photoamplificateurs capables d'agir dans un champ magnétique intense : des photodiodes à avalanche en silicium sont utilisées dans le tonneau, et des phototriodes sous vide pour le bouchon. L'utilisation de cristaux de tungstate de plomb a permis la conception d'un calorimètre électromagnétique compact avec une résolution en énergie[34] :

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \quad (2.5)$$

Où S est le terme stochastique, N le bruit et C le terme constant. Ils ont été paramétrés grâce à un faisceau test d'électrons (Fig. 2.6).

-Le tonneau (EB) est divisé en 36 "supermodules" identiques de rayon interne de 129 cm et couvrant un intervalle en pseudo-rapacité $0 < |\eta| < 1.479$. Les cristaux utilisés pour le tonneaux ont une section de $22 \times 22 \text{ mm}^2$ et une longueur de 230 mm correspondant à 25.8 longueurs de radiation.

-Les bouchons (EE) situés à $|z| = 314 \text{ cm}$ couvrent une zone en pseudo-rapacité 1.479

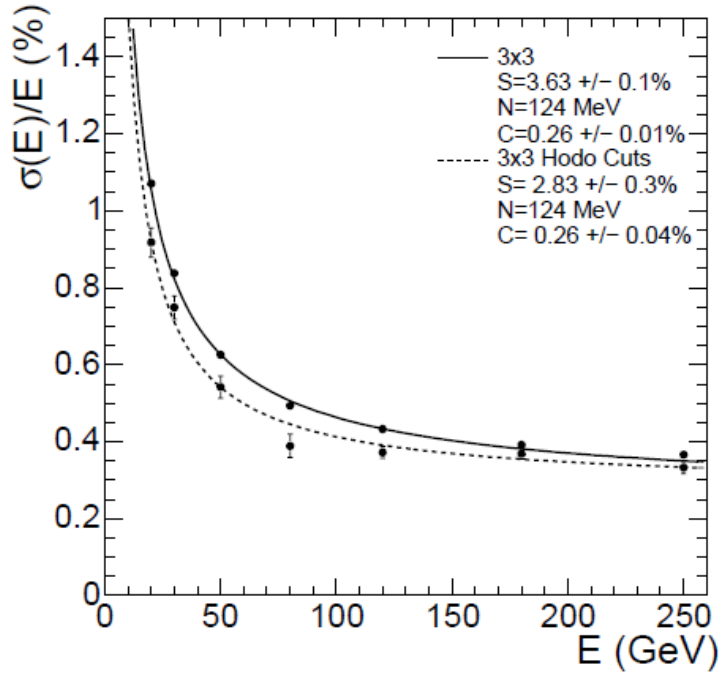


FIGURE 2.6 – Paramétrage de la résolution en énergie du calorimètre électromagnétique de CMS à l’aide d’un faisceau test d’électrons. La courbe du dessus est obtenue pour tous les évènements dans une surface de $20 \times 20 \text{ mm}$, la courbe du dessous est obtenue avec une sélection d’évènements sur une surface de $4 \times 4 \text{ mm}$.

$< |\eta| < 3.0$ et les cristaux utilisés pour les bouchons ont une section de $28.6 \times 28.6 \text{ mm}^2$ et une longueur de 220 mm correspondant à 24.7 longueurs de radiation.

Un mécanisme de pré-cascade (preshower) est placé devant la majeure partie des bouchons pour une meilleure résolution spatiale des photons afin de séparer les deux photons issus de la désintégration d’un π^0 ayant une impulsion élevée. Ce détecteur est constitué de détecteurs à pistes en silicium situé derrière des disques absorbeurs en plomb.

2.2.3.3 Calorimètres hadroniques

Le calorimètre hadronique (HCAL) est dédié à la mesure des énergies transverses des jets de hadrons et des hadrons isolés produits dans les collisions. Il est situé derrière le calorimètre électromagnétique et est constitué comme celui-ci d’un tonneau fermé par des bouchons à ses deux extrémités. Le calorimètre est un assemblage en alternance de plaques de laiton et de tuiles plastiques scintillantes lues par des fibres à déplacement de longueur d’onde (WaveLength-Shifting fibres : WLS).

Le tonneau (Hadron Barrel : HB) est constitué de 32 tours calorimétriques couvrant ensemble une région en pseudo-rapacité $|\eta| < 1.4$, chaque tour ayant une segmentation $\Delta\eta \times \Delta\phi < 0.087 \times 0.087$. Il y a une succession de 15 plaques de laiton de 5 cm d’épaisseur, plus 2 plaques d’acier inoxydable pour renforcer mécaniquement l’assemblage. La première plaque de scintillateur situé à la sortie du calorimètre électromagnétique est plus épaisse

(9 mm au lieu de 3.7 mm) et la lumière émise par celle-ci est multipliée par un facteur 1.5 afin de mesurer les hadrons de basse énergie issus des cascades se produisant dans l'armature entre les deux calorimètres (interaction nucléaire).

Les bouchons (Hadron Forward : HF) sont constitués de 15 tours calorimétriques couvrant la région en pseudo-rapacité $1.3 < |\eta| < 3.0$, pour les 5 tours externes la segmentation est 5° en ϕ et 0.087 en η tandis que pour les 8 tours internes la segmentation est de 10° en ϕ et varie de 0.09 à 0.35 en η .

Une autre partie du calorimètre hadronique (VF) se situe très en avant du détecteur (aux deux extrémités) et proche du faisceau pour couvrir la région en pseudorapacité comprise entre 3.0 et 5.0.

2.2.3.4 Détecteurs à muons

Le système de détection de muons est la partie la plus externe du détecteur. Parmi les particules produites dans les collisions, les muons sont les seules particules (à l'exception des neutrinos et neutralinos) pouvant traverser les calorimètres et le trajectomètre. Ceux-ci ne perdent que l'énergie d'ionisation (quelques GeV) dans les calorimètres contrairement aux hadrons. Le système de détection des muons est situé derrière l'aimant solénoïde où le champ magnétique est inversé et il est constitué de détecteurs gazeux de 3 types :

- Les tubes à dérive couvrant la région centrale du tonneau (MS1-4, $|\eta| < 1,2$) et prompts à fonctionner dans un environnement où l'activité est faible et le champ magnétique résiduel peu important.
- Les chambres cathodiques à piste couvrant les extrémités du tonneau (MF1-4 $1,6 < |\eta| < 2,4$) et pouvant fonctionner dans un environnement riche en neutrons et où le champ magnétique est plus fort.
- Les chambres à plateaux résistifs fonctionnant en parallèle des deux systèmes précédents et servant de système de déclenchement. Il couvre efficacement les pseudo-rapidités jusqu'à $|\eta| < 1,6$ (prévu initialement jusqu'à $|\eta| < 2,1$).

Les calorimètres et le système de détection de muons est présenté dans la figure 2.7.

2.2.3.5 Mesure d'un évènement par le détecteur CMS

Le détecteur CMS est capable de détecter la signature d'évènements faisant intervenir les particules suivantes (c.f. table 2.3) :

- **Électrons** Les électrons, en tant que particules chargées, laissent une trace dans le trajectomètre. Ils interagissent fortement avec la matière, une quantité d'énergie importante est dispersée par rayonnement de freinage dans le trajectomètre et le reste sera déposé dans le calorimètre électromagnétique.
- **Muons** Les muons laissent aussi une trace dans le trajectomètre en tant que particule chargée, mais interagissent peu avec les calorimètres. L'identification d'un muon fait principalement intervenir le système de détection de muons, mais la trajectoire est

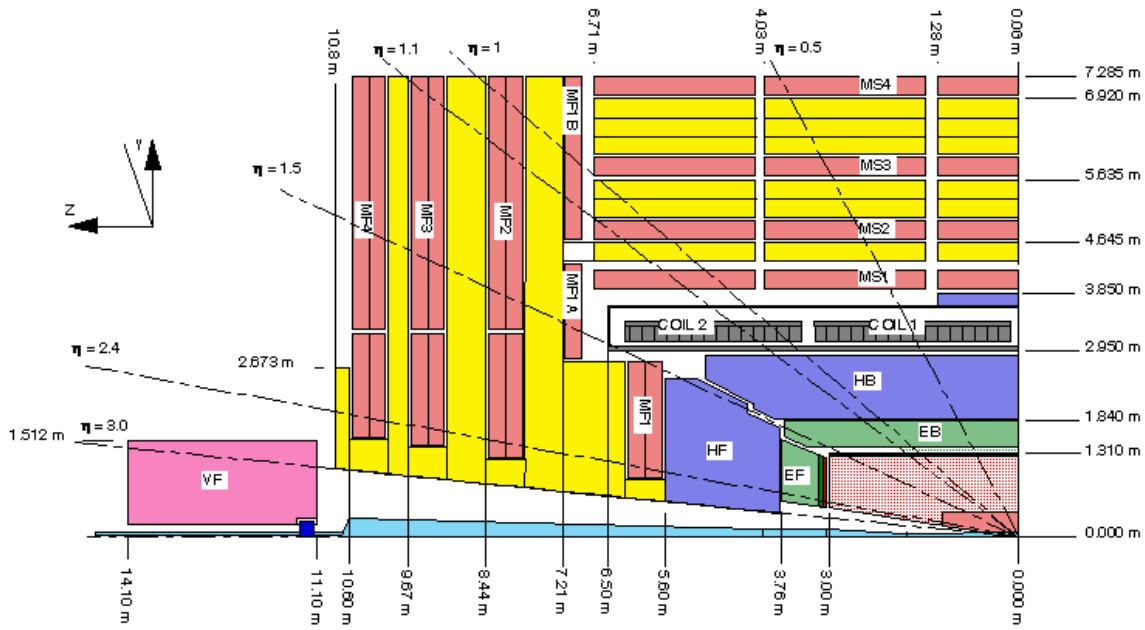


FIGURE 2.7 – Coupe longitudinale du détecteur CMS.

mesurée deux fois : une fois par le système de détection de muons et une fois par le trajectomètre.

- **Hadrons chargés** Les hadrons chargés laissent une trace dans le trajectomètre, une petite partie de leur énergie est dissipée dans le calorimètre électromagnétique mais le hadron parviendra jusqu'au calorimètre hadronique qui absorbera le reste, soit la majeure partie de son énergie.
- **Hadrons neutres** Les hadrons neutres sont invisibles dans le détecteur jusqu'aux calorimètres qui absorberont et mesureront l'énergie du hadron neutre de la même manière que les hadrons chargés. Aucune mesure de trajectoire n'est possible car les hadrons neutres sont invisibles dans le trajectomètre).
- **Photons** Les photons sont invisibles dans le trajectomètre mais sont stoppés par le calorimètre électromagnétique qui mesure leur énergie. Les photons peuvent se convertir en paire e^+e^- dans le trajectomètre, mais une reconstruction des vertex e^+/e^- est possible pour récupérer les mesures d'une bonne fraction de ces photons convertis.

À partir de ces objets simples, on peut reconstituer des objets physiques plus complexes :

- **Jets** Les jets sont des ensembles de particules contenant au moins un hadron (le jet peut être constitué de ce seul hadron) issues de la désintégration d'une particule initiale (par exemple un τ), ou d'un processus QCD (c.f. chapitre 1). Ils impliquent l'intégralité du détecteur CMS car ils peuvent contenir tous les objets simples cités précédemment. Les jets sont initialement uniquement constitués de hadrons, mais les mésons et baryons instables se désintègrent en particules élémentaires : des photons (provenant des π^0 très présents dans les jets hadroniques), et des leptons provenant de la désintégration de mésons charmés ou beaux.

- **Énergie transverse et énergie transverse manquante** En se plaçant dans le référentiel décrit dans la section 2.2.1. On a initialement une impulsion transverse $\Sigma \vec{P}_T = \vec{0}$ pour le système des deux protons se collisionnant. La conservation de l’impulsion veut donc que $\sum_i \vec{P}_{T,i} = \vec{0}$ pour toutes les particules i issues de la collision. Expérimentalement, les particules “invisibles” (neutrinos, neutralinos : si ils existent) ne sont pas détectés et donc seront soustraits à l’évènement, entraînant une impulsion transverse, et donc une énergie transverse manquante $\vec{E}_{T,miss} = -\Sigma \vec{E}_T$.

TABLE 2.3 – Interactions des particules avec les différentes parties du détecteur CMS.

Objet	Trajectomètre	Calo. EM	Calo. Hadr.	Syst. de muons
Électron	+++	+++		
Muon	+++	+	+	+++
Hadron chargé	+++	++	+++	
Hadron neutre		+	+++	
Photon	+	+++		

2.2.3.6 Systèmes de déclenchement

Le LHC a été conçu pour fonctionner à une luminosité de $10^{34} \text{ cm}^{-2} \cdot \text{s}^{-1}$ correspondant à une fréquence de collision de 1 GHz. La plupart des collisions ne sont d’aucun intérêt pour les recherches menées dans le LHC. Les systèmes de déclenchement de CMS ont le rôle de sélectionner et d’enregistrer uniquement les collisions intéressantes. Ces systèmes sont répartis sur deux niveaux :

• Les systèmes de déclenchement de niveau 1

Le système de déclenchement de niveau 1 est un système électronique utilisant les signaux d’une partie du détecteur CMS pour effectuer une sélection rapide des collisions[35]. La sélection des évènements au niveau 1 implique la calorimétrie et les détecteurs de muons ainsi que quelques corrélations entre ces systèmes. La granularité complète des calorimètres n’est pas disponible dans cette sélection, et on utilise alors seulement l’information grossière dans des larges éléments d’espace $\eta - \phi$: les tours calorimétriques.

La taille du détecteur et des cavernes souterraines imposent pour l’information un temps de transit aller-retour maximal de $3,2 \mu\text{s}$ depuis l’électronique du détecteur jusqu’aux cavernes contenant les systèmes logiques qui décideront si une collision doit être conservée ou non. La décision de retenir un évènement au niveau 1 est basée sur la présence d’objets “primaires de déclenchement” (trigger primitives) tels que des muons (une ou plusieurs trace(s) dans le détecteur de muons) ou des jets (tours calorimétriques) mais aussi des objets plus complexes : ΣE_T ou $E_{T,Manquante}$. Tous ces objets sont discriminés avec plusieurs seuils en P_T ou E_T .

Pendant ce temps toute l’information à haute résolution est conservée dans une mémoire tampon (pipeline) située sur le détecteur. Le déclenchement de niveau 1 fonctionne avec une fréquence moyenne de 50 kHz, mais est conçu pour atteindre une fréquence maximale

de 100 kHz.

Les événements retenus au niveau 1 doivent ensuite passer la sélection par le niveau supérieur du système de déclenchement.

- **Les systèmes de déclenchement de haut niveau (HLT)**

Les systèmes de déclenchement de haut niveau sont des systèmes utilisant des ordinateurs (et non plus des déclencheurs électroniques) pour analyser les événements sélectionnés par le niveau 1. Les objets primaires sont analysés pour appliquer des sélections supplémentaires sur l'événement en employant toute l'information du détecteur CMS : la trajectométrie complète et la granularité intégrale des calorimètres. Un déclenchement de haut niveau utilise comme point de départ les événements sélectionnés par le niveau 1 sur lesquels des critères de sélection supplémentaires sont appliqués pour sélectionner un type d'événement contenant un ou plusieurs objets physique particuliers. Les divers événements subissent alors une première classification en fonction de leur catégorie physique (échantillons muons, jets, énergie transverse manquante, etc.).

Les données ainsi classées sont divisées selon les résultats des systèmes de déclenchement à haut niveau en lots d'événements choisis pour des analyses particulières. Ainsi pour une analyse impliquant par exemple des jets, une année de collisions peut être analysée plus rapidement et avec un stockage moins important en ne regardant que les lots de données ayant été déclenchés par des jets.

Les études présentées dans ce mémoire impliquent des taus se désintégrant hadroniquement, nous avons donc utilisé surtout les événements sélectionnés par les système de déclenchement pour jet.

- **Restrictions sur l'acceptation d'événements**

Les systèmes de déclenchement de niveau 1 fonctionnent à une fréquence de 50 à 100 kHz tandis que la fréquence des collisions est de l'ordre du GHz. Les conditions de déclenchement choisies pour les différentes optiques de recherche doivent être optimisées de façon à ce que la fréquence d'enregistrement des événements ne dépasse pas les capacités techniques du détecteur[36]. On aura donc un facteur de diminution (downscale factor) sur le nombre d'événements enregistrés par certains systèmes de déclenchement. Ceci implique certaines précautions à prendre lorsque l'on analyse les échantillons de données que nous verrons dans le chapitre 4.

2.3 Reconstruction des données

2.3.1 Environnement des analyses

Au sein de la collaboration CMS, l'environnement standard d'analyse s'appelle CMSSW (pour CMS SoftWare), son architecture fonctionnant sur une base UNIX codée en C++ est modulaire et offre à l'utilisateur une grande liberté en lui permettant d'ajouter ses propres modules. Les modules sont accessibles via un service appelé CVS (Concurrent

Version System). Les modules d'analyse sont mis à jour en permanence pour inclure les améliorations et les nouvelles méthodes d'analyses développées par la collaboration.

CMSSW est compatible avec ROOT, un environnement d'analyse de données développé pour la physique des hautes énergies. Il permet de créer diverses collections de données : des histogrammes (de 1 à 3 dimensions), des graphiques, des fonctions, et aussi des “ntuples” contenant les informations événement par événement et préservant les corrélations entre les différentes variables et objets physiques. ROOT fournit aussi une interface graphique pour la présentation des données.

Pour les analyses à grande échelle, la grille de calcul (LHCGrid) relie les différents instituts de la collaboration CMS aux centres de calcul dans le monde entier. Pour contrôler ces analyses dans l'environnement CMSSW la collaboration utilise un outil appelé CRAB (Cms Remote Analysis Builder). Les calculs à grande échelle et la répartition des données sont indispensables au vu de la quantité de données accumulées par le LHC.

2.3.2 Reconstruction des trajectoires

Une trace est un ensemble d'impacts (ou coups) causé par le passage de particules chargées dans le trajectomètre de CMS (Fig 2.8). On les reconstruit par la méthode de “traçage itératif” [38] en associant les coups dans le trajectomètre en 7 étapes : on prend comme origine une “graine” qui est un triplet de coups alignés dans le détecteur à pixels pour extrapoler la trajectoire complète dans les détecteurs en silicium (étapes 0 et 1). On peut utiliser deux coups dans le détecteur à pixels pour augmenter l'efficacité de reconstruction (étape 2) dans le cas où la graine est mal définie. Les étapes 0 et 1 sont répétées (étape 3) en recherchant des traces déplacées par rapport au vertex d'interaction primaire (provenance potentielle d'un vertex d'interaction secondaire ou d'une désintégration). Dans l'étape 4 on utilise des triplets de coups impliquant les détecteurs à pixels et silicium pour récupérer les traces ayant manqué un coup dans le pixel ou celles se désintégrant près (quelques centimètres) du vertex d'interaction associé. Enfin les traces largement déplacées par rapport au vertex d'interaction primaire sont reconstruites (étapes 5 et 6) en n'utilisant que les détecteurs en silicium.

Les traces reconstruites sont finalement sélectionnées en appliquant des coupures de qualité sur celles-ci. La qualité d'une trace est évaluée à l'aide des paramètres suivants :

- **Nombre de coups** : Plus une particule chargée laisse de coups dans le trajectomètre, plus il y aura de points pour reconstruire la trajectoire (meilleure précision). Un des premiers requis pour obtenir une trace de qualité est le nombre minimal de coups ou le nombre maximal de coups manquants (points manquants sur la trajectoire).
- χ^2/ddl : Le χ^2 sur le nombre de degrés de liberté. Ce paramètre estime la déviation des coups par rapport à la trace reconstruite. Tous les coups d'une trace “parfaite” sont présents sur la trajectoire reconstruite ($\chi^2/ddl = 1$) mais dans le cas réel il peut y avoir des déviations des coups par rapport à la trace reconstruite. On requiert alors une valeur basse pour le paramètre χ^2/ddl . Le nombre de degrés de liberté est le nombre de coups moins le nombre de paramètres utilisés pour définir mathématiquement la trajectoire.

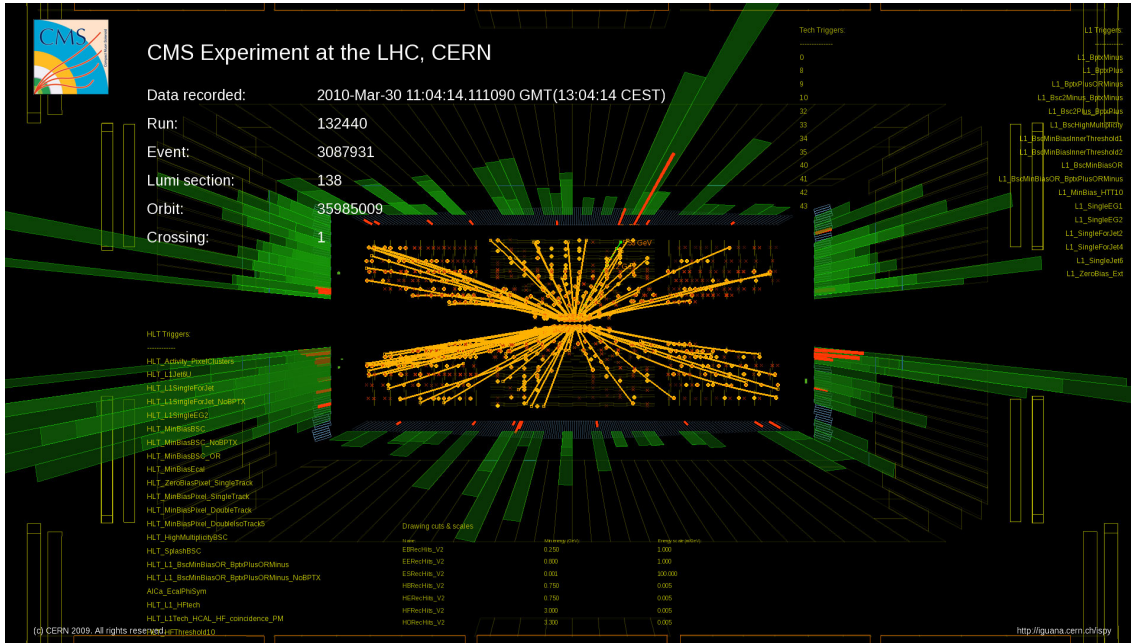


FIGURE 2.8 – Exemple d'un évènement reconstruit par le détecteur CMS au cours de l'année 2010 (provenant de [39]).

- **Paramètres d'impact (transverse et longitudinal)** : Les particules chargées suivent une trajectoire hélicoïdale dans le détecteur CMS : le rayon de courbure permet de mesurer l'impulsion de la particule. Le paramètre d'impact est la distance de plus courte approche entre la trace et le point défini par le vertex d'interaction primaire (c.f. Figure 2.9). On peut projeter la trace dans le plan xOy afin de mesurer le paramètre d'impact transverse noté d_{xy} , ou dans le plan xOz (plan longitudinal) dans lequel la trace est assimilable à une droite (si l'impulsion transverse de la particule est suffisante) : le paramètre d'impact longitudinal noté d_z est le déplacement de cette droite par rapport au vertex d'interaction primaire (selon l'axe z).

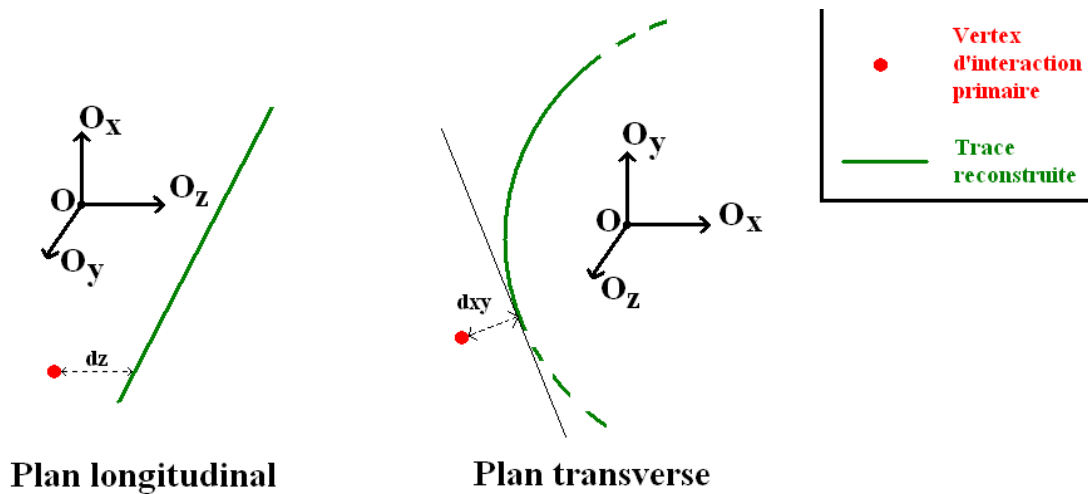


FIGURE 2.9 – Schéma des paramètres d'impact transverse et longitudinal.

2.3.3 Format des données

Pour la reconstruction des données, la collaboration CMS emploie différents algorithmes adaptés à la reconstruction d'un type de particule ou d'un objet physique complexe (comme décrit dans la section 2.2.3.5). Les données (Monte Carlo ou réelles) ont alors différents niveaux de reconstruction :

- **Raw (Brut)** : Aucune reconstruction n'a été effectuée, l'échantillon consiste donc en un ensemble de signaux provenant des modules de digitalisation (constructeur d'évènements), sans détermination de corrélation entre ceux-ci. C'est le format que l'on obtient juste après le passage par les déclencheurs de haut niveau. Un évènement au format RAW peut être "zero suppression", c'est-à-dire qu'aucun seuil (en énergie par exemple) n'est appliqué sur les détecteurs.
- **RECO** : Le format RECO contient la reconstruction complète de l'évènement, les signaux obtenus dans le format Raw sont calibrés et alignés afin d'obtenir des collections de traces et de grappes calorimétriques. Ces premières collections permettent la reconstruction des particules (électrons, muons, photons, etc.) et aussi d'objets physiques complexes (jets et énergie transverse).
→ Le format RECO contient les informations Particle Flow[40] (c.f. section suivante : 2.3.4).
- **AOD (Analysis Object Data)** : Similaire aux RECO, mais prévu uniquement pour l'analyse. Une partie de l'information inadaptée aux analyses physiques (objets du format RAW) est supprimée afin d'économiser de place.
→ Le format AOD contient les informations Particle Flow.
- **PAT (Physics Analysis Toolkit)** : Les échantillons traités par PAT sont l'aboutissement de la classification d'objets au sein d'un évènement, c'est une sélection d'objets essentiels pour une analyse physique spécifique. Un nettoyage supplémentaire du contenu des évènements est appliqué sur les objets physiques dans le but d'éviter un double comptage d'objets ambigus i.e. un objet appartenant à deux classes distinctes. L'utilisation de PAT est décrite plus largement dans la section 3.5.1.

2.3.4 L'algorithme Particle Flow

L'algorithme de reconstruction "Particle Flow" a une précision supérieure à l'algorithme "Calo" (reconstruction des objets physiques à l'aide des calorimètres uniquement). L'algorithme Particle Flow utilise les collections de traces, d'amas calorimétriques, et de signaux muoniques pour les relier entre eux, et remonter jusqu'au passage des particules (d'où le nom Particle Flow : Flux de particules). Pour chaque évènement l'algorithme procède ainsi :

Premièrement les muons globaux (i.e. une trace dans trajectomètre interne corrélée à une trace dans le détecteur de muons) sont identifiés et donnent lieu à des Particle Flow Muons (PFMuons), les traces correspondantes sont soustraites de l'évènement.

La reconstruction et l'identification des électrons vient ensuite, chaque trace de l'évènement

est soumise à une pré-identification en exploitant le trajectomètre comme un système de pré-cascade, les électrons tendent à donner des traces courtes en perdant de l'énergie par Bremsstrahlung (rayonnement de freinage). Une correction est appliquée sur ces traces pré-identifiées afin de suivre leur trajectoire jusqu'au calorimètre électromagnétique. Une identification finale est effectuée à l'aide de variables calorimétriques, les électrons ainsi identifiés deviennent des Particle Flow Electrons (PFElectrons). Les traces et les amas calorimétriques sont soustraits de l'évènement pour la suite.

Ensuite des critères restrictifs sont appliqués sur les traces restantes, on exige que l'incertitude sur la mesure de l'impulsion transverse soit inférieure à la résolution en énergie du calorimètre hadronique. Les traces respectant ces critères peuvent donner lieu à des hadrons chargés (Particle Flow Charged Hadrons : PFChargedHadrons), des photons (PFPhotons), des hadrons neutres (PFNeutralHadrons) ou plus rarement d'autres muons. Dans le cas où une trace est associée à plusieurs amas calorimétriques (hadronique comme électromagnétique), seul l'amas le plus proche est retenu et associé à cette trace.

Si aucune trace n'est associée à un amas calorimétrique, on considère que c'est un photon pour le ECAL et un hadron neutre pour le HCAL. Si un excédent d'énergie est observé dans un amas calorimétrique électromagnétique par rapport aux traces associées à cet amas, un objet Particle Flow Photon est créé pour correspondre à cette différence d'énergie. De façon similaire si un excédent est observé dans le calorimètre hadronique, un objet Particle Flow Neutral Hadron est créé.

Enfin, dans le cas où la somme des énergies des traces est largement supérieure à celle des calorimètres, une recherche avec des critères plus souples est effectuée pour trouver d'autres muons et des fausses traces. Les traces restantes non associées aux différents calorimètres sont considérées comme des objets Particle Flow Charged Hadrons. Ces traces peuvent dans certains cas être dus à des photons convertis dans le trajectomètre, l'algorithme Particle Flow récupère certains de ces photons convertis en analysant les traces "molles".

Un algorithme de reconstruction de jets est ensuite appliqué sur ces différents objets Particle Flow pour donner lieu à des objets PFJets. Plusieurs méthodes sont possibles pour reconstruire ces jets et seront décrites dans la section suivante.

2.3.5 Reconstruction des jets

La reconstruction d'un jet est une tâche plus délicate que l'identification d'un muon ou d'un photon : un jet n'est pas une particule simple, mais un regroupement de particules suivant un algorithme. Plusieurs conditions sont imposées pour valider la reconstruction d'un jet. Un algorithme de reconstruction de jet doit satisfaire idéalement les conditions générales suivantes[41] :

1. *Entièrement spécifié* : Le processus de construction et de sélection du jet, les variables cinématiques et les diverses corrections à apporter (c.à.d. le rôle de l'évènement sous-jacent) doivent être clairement et complètement définis. Si c'est nécessaire, l'algorithme doit prendre en compte les possibles scissions ou fusions de jet (split-

ting/merging).

2. *Sûreté infrarouge (infrared safety) et colinéaire (colinear safety)* : Le jet reconstruit par l'algorithme ne doit pas être modifié par la présence de particules de faible énergie (Fig. 2.10a) ou lors de l'émission de deux particule dans la même direction (Fig. 2.10b).
3. *Indépendance par rapport au détecteur* : Aucune dépendance de l'algorithme au nombre, au type et à la taille des cellules de détection.
4. *Indépendance par rapport à l'échelle* : L'algorithme doit être applicable au niveau du parton, de la particule ou du détecteur.

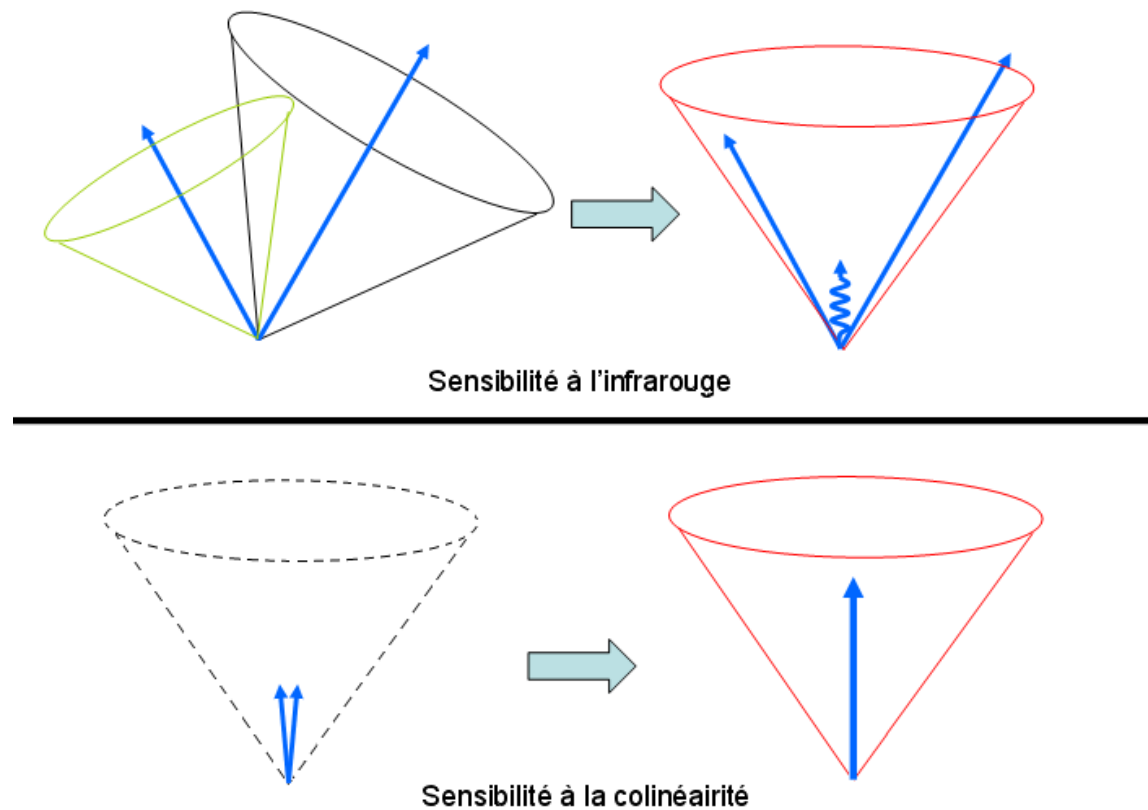


FIGURE 2.10 – En haut : Cas de mauvaise reconstruction de deux jets en cas de sensibilité à l'infrarouge, un rayonnement doux entraîne la fusion des deux jets reconstruits en un seul. En bas : cas de reconstruction d'un jet non existant en cas de sensibilité à la colinéarité, deux particules de faible énergie qui ne forment pas un jet (par exemple deux hadrons de l'évènement sous jacent) sont reconstruits comme un seul flux d'énergie et définissent un jet.

Tous les algorithmes doivent satisfaire aux deux premiers critères, les deux autres peuvent être approximativement satisfaits. Les algorithmes respectant ces conditions peuvent être

divisés en deux grandes catégories :

- **Les algorithmes de type Cône**

Les algorithmes de type cône sont inspirés par l'idée[42] de définir un jet comme un cône angulaire autour de la direction d'un flux d'énergie dominante. Pour trouver ces flux, les algorithmes de type cône prennent une partie ou la totalité des particules de l'évènement comme flux potentiel ou "graine" (i.e. les directions des cônes d'essai). Ensuite, ils établissent les listes de particules dans le cône d'essai, évaluent la somme de leur 4-impulsions et utilise la résultante de la somme comme nouvelle direction pour le cône d'essai. Cette procédure est réitérée jusqu'à ce que le cône soit "stable", c'est-à-dire que sa direction ne doit plus changer.

Un cône est dit "stable" si la condition suivante est respectée :

$$D(\sum_i p_i, a) = 0 \quad (2.6)$$

Où a est l'axe du cône d'essai et p_i la 4-Impulsion de chaque particule dans le cône d'essai. $D(\sum_i p_i, a)$ est une mesure de la distance angulaire entre les deux directions. Typiquement on définit la distance angulaire entre deux quadrivecteurs p et a par la relation :

$$D^2(\sum_i p_i, a) = (\phi_p - \phi_a)^2 + (y_p - y_a)^2 \quad (2.7)$$

ϕ est l'angle azimutal et y la rapidité de l'axe concerné.

Les deux relations 2.6 et 2.7 définissent le principe de l'algorithme de cône itératif (Iterative Cone Algorithm), cet algorithme est rapide par son temps de calcul, mais pose le problème de ne pas reconstruire correctement un jet dans lequel deux particules sont susceptibles d'être des "graines" (direction de deux particules symétriques entre elles par rapport à l'axe du jet). Ce problème peut être résolu en introduisant un point intermédiaire entre chaque paire de particules qui pourra constituer un nouvel axe pour un cône d'essai (Mid Point Iterative Cone). Mais ce problème observé pour des paires de particules peut se présenter des triplets, des quadruplets, etc.

Une méthode pour remédier à ces problèmes consiste à ne plus utiliser de "graine" du tout, ceci est réalisé par l'algorithme SIScone[43] (Seedless Infrared Safe Cone algorithm) dont le principe général de fonctionnement est le suivant :

L'algorithme prend toutes les particules de l'évènement, en fusionnant les particules colinéaires. On recherche les paires de particules distantes de moins de $2R$ (Fig. 2.11a)) et qui seront donc contenues dans un cône de rayon R . S'il n'y a pas de particules distantes de $2R$ dans le voisinage d'une particule i , la direction de cette particule i formera l'axe d'un cône d'essai.

Un cône d'essai contenant 2 particules sera soumis à plusieurs itérations pour inclure les particules dans le proche voisinage (une nouvelle particule sera sur le bord du cône), en conservant les deux particules initiales contenues dans le jet (Fig. 2.11b) et 2.11c)). Une

fois tous les cônes d'essai évalués (Fig. 2.11d)), le cône retenu (et définissant le jet reconstruit par la suite) sera celui qui est le plus stable.

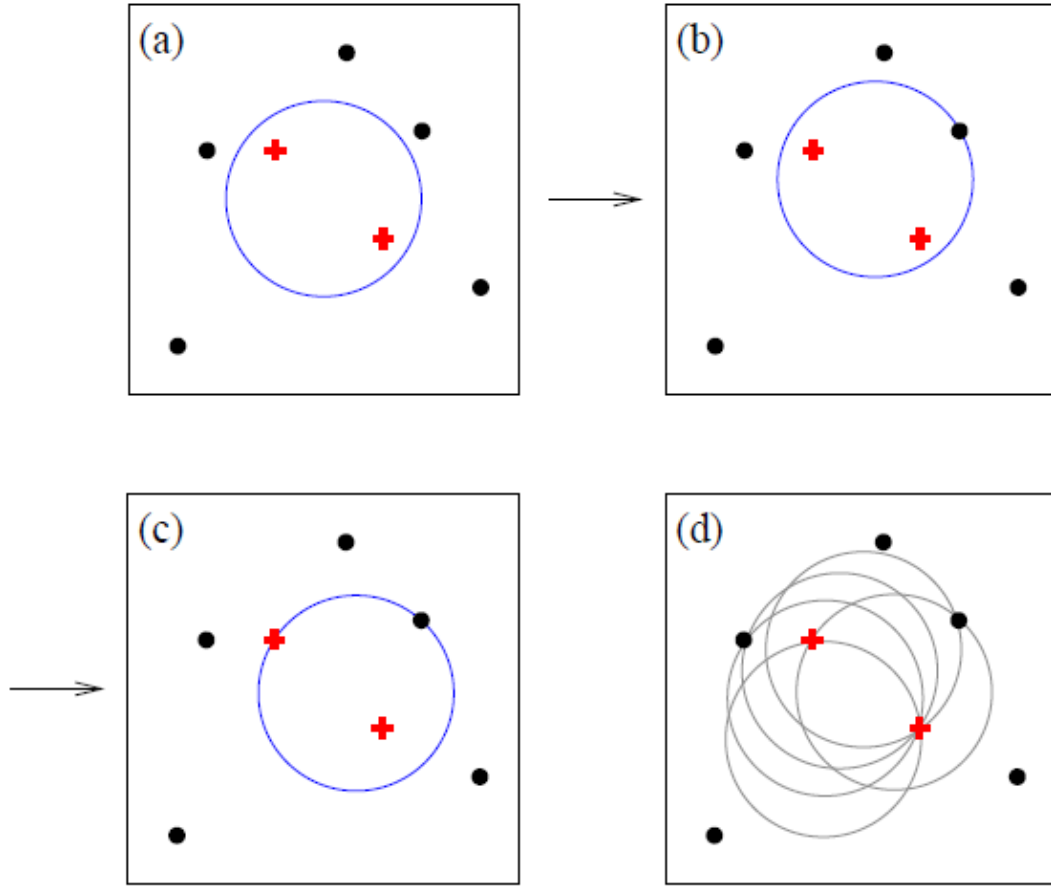


FIGURE 2.11 – Principe général de l'algorithme sans “graine” (provenant de [43]).

Quel que soit l'algorithme considéré, il est probable que des jets se recouvrent et partagent les mêmes particules entraînant ainsi un double comptage de celles-ci. Pour éviter ceci il faut pouvoir diviser ou fusionner les jets potentiels en associant les particules au jet approprié (en cas de division) ou déterminer si les deux jets sont en réalité un seul. Pour ce faire la stratégie consiste à appliquer des coupures en P_T (pour supprimer les jets de trop bas P_T même s'ils sont stables), ou associer les particules contenues dans deux cônes au jet le plus proche, tout en vérifiant la stabilité des jets après association.

• Les algorithmes de type k_T

Les algorithmes de type k_T se basent sur une collection d'entités, typiquement les particules contenues dans l'évènement (PFlow), et introduisent entre une entité i et une entité j la distance d_{ij} ainsi que la distance d_{iB} entre une entité i et le faisceau [44]. Ces grandeurs sont calculées par les relations suivantes :

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2} \quad (2.8)$$

$$d_{iB} = k_{ti}^{2p} \quad (2.9)$$

k_{ti} est l'énergie transverse de l'entité i et R un paramètre permettant de limiter la largeur des jets. Δ_{ij}^2 est calculée par la même relation que le D^2 (équation 2.7 en remplaçant les quadrivecteurs a, p par i, j) .

Le paramètre p est un paramètre supplémentaire introduit pour l'algorithme anti- k_t afin de gouverner la puissance relative de l'énergie par rapport aux échelles géométriques Δ_{ij}^2 : il est fixé à $p = 1$ pour l'algorithme k_t et -1 pour l'algorithme anti- k_t d'où son nom "anti". L'algorithme anti- k_t offre l'avantage de coupler les particules de faible impulsion (douces) avec les particules de grande impulsion (dures), offrant une reconstruction de jets similaire à l'algorithme *SISCone* avec un temps de calcul plus court.

L'algorithme calcule pour une entité i tous les d_{ij} avec les particules proches ainsi que d_{iB} , et il peut se présenter alors deux cas de figure :

1. $\min(d_{iB}, d_{ij}) = d_{ij}$: l'entité i et l'entité j sont fusionnées en une seule entité (pseudojet) pour l'itération suivante.
2. $\min(d_{iB}, d_{ij}) = d_{iB}$: le jet est stable, il est ajouté à la collection de jets reconstruits et soustrait de l'évènement pour l'itération suivante.

L'algorithme effectuera des itérations jusqu'à ce que la distance d_{iB} finisse par être minimale pour toutes les entités i , chacune de ces entités définit alors un jet.

La figure 2.12 nous présente la reconstruction des jets d'un même évènement par les différents algorithmes. L'algorithme *SISCone* est l'algorithme le plus juste, mais le choix de la collaboration CMS s'est porté sur l'algorithme anti- k_T afin d'accélérer la reconstruction tout en conservant une justesse très satisfaisante.

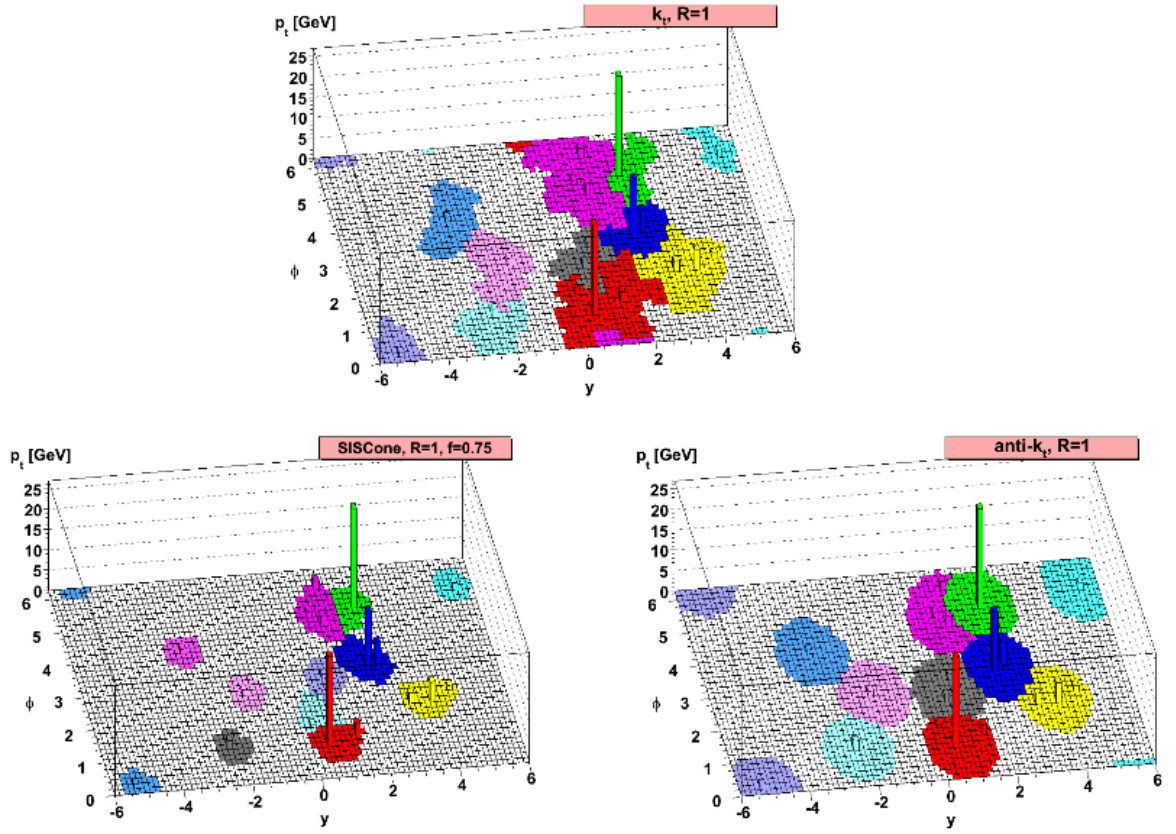


FIGURE 2.12 – Simulation de la reconstruction de jets par différents algorithmes (provenant de [45]).

Chapitre 3

L'identification du lepton τ

3.1 Le lepton τ

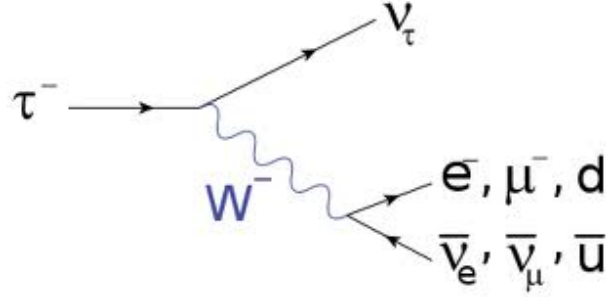
3.1.1 Caractéristiques

Le lepton tau est le plus lourd des leptons du modèle standard et le dernier lepton massif à avoir été découvert en 1975 à l'Accélérateur Linéaire de Standford (SLAC)[46]. Son neutrino associé : le neutrino tauique, fut découvert en 2000[47] par l'expérience DONUT (Direct Observation of NU Tau). C'est aussi le seul lepton pouvant se désintégrer en donnant des hadrons. Ses caractéristiques sont résumées dans le tableau 3.1.

TABLE 3.1 – Caractéristiques du lepton τ

Caractéristique	Grandeur physique	Valeur
Masse	m	1,777 GeV
Soumis aux interactions		Faible,Électromagnétique
Charge	q	± 1
Spin	s	1/2
Temps de vie	τ	$(290,6 \pm 1,0) \cdot 10^{-15}$ s
Distance de vol à 20 GeV	$c\tau\gamma$	871,1 μ m
Mode de désintégration	Résonance	Rapport d'embranchement
Leptoniques		
$\tau^- \rightarrow \mu^- \nu_\tau \nu_\mu$		17,36 %
$\tau^- \rightarrow e^- \nu_\tau \nu_e$		17,85 %
Hadroniques		
$\tau^- \rightarrow h^- \nu_\tau$		11,6 %
$\tau^- \rightarrow h^- \pi^0 \nu_\tau$	$\rho(770 \text{ MeV}/c^2)$	26,0 %
$\tau^- \rightarrow h^- \pi^0 \pi^0 \nu_\tau$	$a_1(1200 \text{ MeV}/c^2)$	10,8 %
$\tau^- \rightarrow h^- h^+ h^- \nu_\tau$	$a_1(1200 \text{ MeV}/c^2)$	9,8 %
$\tau^- \rightarrow h^- h^+ h^- \pi^0 \nu_\tau$	$a_1(1200 \text{ MeV}/c^2)$	4,8 %
Autres modes hadroniques		1,7 %
Total hadronique		64,8 %

Le lepton tau se désintègre par interaction faible (Fig 3.1.1) via un boson W virtuel (noté W^*), on parle d'interaction faible par courant chargé.

FIGURE 3.1 – Diagramme de Feynman au premier ordre de la désintégration du lepton τ .

Le W virtuel présenté dans la Fig. 3.1.1 a les mêmes propriétés qu'un W réel avec certaines contraintes imposées par la masse du τ : la masse invariante des particules issues de la désintégration de ce W^* ne peut dépasser celle du τ .

Les modes de désintégration d'un W réel sont résumés dans la table 3.2.

TABLE 3.2 – Modes de désintégration du W^- réel (poids relatif du mode concerné)

Leptons	Quark up	Quark charmé
$e^- \bar{\nu}_e$ (1)	$\bar{u}d$ ($3 V_{ud} ^2$)	$\bar{c}d$ ($3 V_{cd} ^2$)
$\mu^- \bar{\nu}_\mu$ (1)	$\bar{u}s$ ($3 V_{us} ^2$)	$\bar{c}s$ ($3 V_{cs} ^2$)
$\tau^- \bar{\nu}_\tau$ (1)	$\bar{u}b$ ($3 V_{ub} ^2$)	$\bar{c}b$ ($3 V_{cb} ^2$)

Les $V_{qq'}$ sont les éléments de matrice CKM[49] (Cabbibo-Kobayashi-Masakawa) donnant les couplages du courant chargé à un quark de type up (q) et un quark de type down q' . L'amplitude de désintégration d'un W^* en qq' est donc proportionnelle à $|V_{qq'}|^2$. Le 3 devant chaque $|V_{qq'}|^2$ provient des trois couleurs que peuvent prendre les quarks.

Les modes avec un quark charmé sont impossibles lors de la désintégration d'un lepton τ par courant chargé faible car le plus léger des mésons charmé, le $D^\pm$ a une masse de 1870 MeV : plus lourd que le tau (de masse 1,777 GeV). Les modes avec un quark beau sont interdits pour les mêmes raisons.

On peut alors remonter aux rapports d'embranchement du τ , les valeurs $|V_{ud}|$ et $|V_{us}|$ ont été mesurées dans des expériences précédentes de physique des particules : $|V_{ud}| = 0,97383$ et $|V_{us}| = 0,2272$ [49].

Pour le lepton τ on a alors les rapports d'embranchement suivants (calculées au premier ordre (LO)) :

$$\begin{aligned}
 BR_e = BR_\mu &= \frac{\Gamma_{e,\mu}}{\Gamma_e + \Gamma_\mu + \Gamma_{u\bar{d}} + \Gamma_{u\bar{s}}} = \frac{1}{2 + 3|V_{ud}|^2 + 3|V_{us}|^2} \approx 20,4\% \quad (3.1) \\
 BR_{u\bar{d}} &= \frac{\Gamma_{u\bar{d}}}{\Gamma_e + \Gamma_\mu + \Gamma_{u\bar{d}} + \Gamma_{u\bar{s}}} = \frac{3|V_{ud}|^2}{2 + 3|V_{ud}|^2 + 3|V_{us}|^2} \approx 57,0\% \\
 BR_{u\bar{s}} &= \frac{\Gamma_{u\bar{s}}}{\Gamma_e + \Gamma_\mu + \Gamma_{u\bar{d}} + \Gamma_{u\bar{s}}} = \frac{3|V_{us}|^2}{2 + 3|V_{ud}|^2 + 3|V_{us}|^2} \approx 2,9\%
 \end{aligned}$$

Les valeurs calculées intuitivement ici sont relativement similaires à celles affichées dans la table 3.1, les résonances ρ et a_1 sont des états excités du méson π^+ i.e. de la paire $u\bar{d}$. Les valeurs ne sont pas rigoureusement exactes ici car le calcul 3.1 n'est qu'au premier ordre et correspond purement au diagramme de Feynmann 3.1.1. Pour obtenir les valeurs exactes il faut tenir compte des ordres suivants i.e. des corrections radiatives[50] dont un exemple est montré dans la figure 3.1.1.

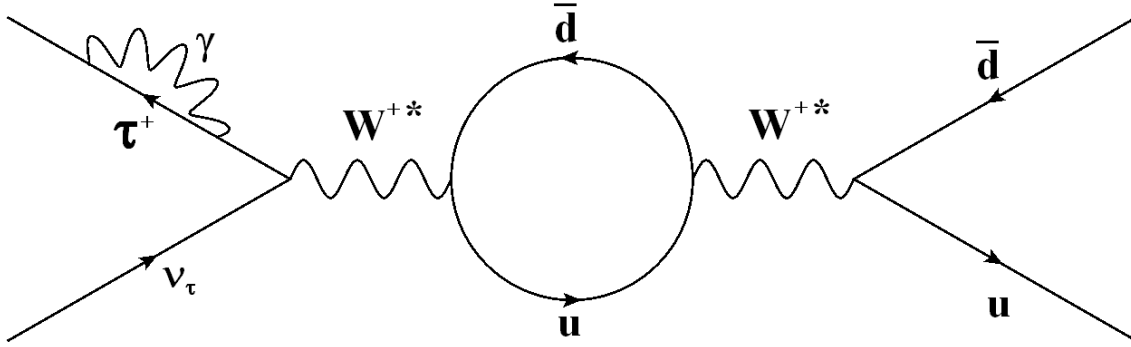


FIGURE 3.2 – Exemple de désintégration hadronique du τ avec radiations correctives dans la voie S.

La différence expérimentale entre les rapports d'embranchement trouvés pour les modes $\tau \rightarrow e\nu$ et $\tau \rightarrow \mu\nu$ est due au fait que en réalité $BR_e \neq BR_\mu$, il y a une faible différence de facteur d'espace des phases due à la différence de masse entre ces deux leptons.

3.1.2 Détection des leptons tau

En se basant sur les caractéristiques du τ nous pouvons discuter différentes méthodes d'identification de celui-ci :

- **Détection directe** : La détection directe du lepton τ est très difficile au vu de la distance de vol cette particule, cette méthode nécessite une excellente résolution spatiale de la trajectoire des particules chargées (trajectographe). La détection directe n'est pas réalisée dans CMS, mais elle reste possible et a déjà été employée (expérience OPERA utilisant des émulsions [48]).

• **Détection dans le canal leptonique** : Nous ne pouvons pas différencier les leptons issus de la désintégration d'un τ de ceux obtenus par production directe ou via la désintégration d'une autre particule à moins de mesurer la trajectoire du tau directement (c.f. détection directe). On peut cependant rechercher des paires de τ (issues de la désintégration d'un Z ou d'un Higgs) en recherchant un tau se désintégrant dans la voie hadronique et l'autre tau se désintégrant en un muon ou un électron (recherche d'un état final lepton + jet).

• **Détection dans le canal hadronique** : Les hadrons résultant de la désintégration du τ peuvent donner soit un meson π chargé, soit une résonnance (ρ ou a_1) se désintégrant à son tour en donnant peu de hadrons. On recherche alors des jets hadroniques (c.f. section 2.2.2.6) qui sont des objets physiques facilement détectables par CMS. La difficulté réside dans le fait que les sources de jets hadroniques sont nombreuses, notamment dans les collisions proton-proton où la QCD prédomine (c.f. section 2.1). Mais les jets formés par les τ ont de caractéristiques différentes de ceux formés par la QCD.

Le τ se désintègre vers les hadrons par interaction faible en deux quarks, ces deux quarks sont reliés entre eux par la couleur mais ne sont pas reliés au reste de l'évènement tandis que dans un évènement purement QCD tous les jets seront reliés entre eux par la couleur, se résultant par l'apparition de nombreux hadrons entre les jets. Nous avons donc dans le cas des τ des jets bien isolés contrairement aux jets QCD. De plus l'impulsion des quarks dans le centre de masse du τ est contraint par la faible masse de ce lepton, l'ouverture angulaire entre les constituants du jet sera donc très basse. On peut estimer l'angle maximal entre les produits de désintégration en effectuant les approximations suivantes (pour estimer l'ouverture maximale d'un jet hadronique de τ) :

- L'angle sera plus large dans le cas d'une désintégration en peu de hadrons légers : l'énergie totale du τ est conservée dans ses produits de désintégration, l'impulsion transverse de ces hadrons (dans le référentiel du τ au repos) sera plus importante si leur masse est faible.
- Afin de poser une limite, on considère un cas simple ou on néglige l'énergie emportée par le neutrino.

On considère alors un τ^- se désintégrant en $\pi^-\pi^0$. En négligeant le neutrino tauique on se retrouve dans le cas d'une désintégration à deux corps :

$$\begin{aligned}
 P_\tau^2 &= m_\tau^2 = (P_- + P_0)^2 \\
 &= \underbrace{P_-^2}_{m_{\pi^-}^2} + \underbrace{P_0^2}_{m_{\pi^0}^2} + 2P_- \cdot P_0 \\
 &= m_{\pi^-}^2 + m_{\pi^0}^2 + 2(E_-E_0 - \vec{p}_- \cdot \vec{p}_0) \\
 &= m_{\pi^-}^2 + m_{\pi^0}^2 + 2(E_-E_0 - \|\vec{p}_-\| \times \|\vec{p}_0\| \cos(\alpha))
 \end{aligned} \tag{3.2}$$

$$\tag{3.3}$$

Ici α est l'angle entre le π^- et le π^0 (Fig. 3.3).

Si E_+ et $E_0 \gg m_{\pi^-}, m_{\pi^0}$ alors $E_- \sim p_-$ et $E_0 \sim p_0$:

$$\begin{aligned}
 m_\tau^2 &= 2E_-E_0(1 - \underbrace{\cos(\alpha)}_{\sim 1-\alpha^2/2}) \\
 &= 2E_-E_0(1 - (1 - \frac{\alpha^2}{2})) \\
 &= 2E_-E_0\frac{\alpha^2}{2} = E_-E_0\alpha^2 \\
 \text{D'où :} \\
 \alpha^2 &= \frac{m_\tau^2}{E_-E_0} \Rightarrow \alpha \sim \frac{m_\tau}{\sqrt{E_-E_0}}
 \end{aligned}
 \tag{3.4}$$

$$\tag{3.5}$$

En dernière approximation on considère $E_+ \sim E_0 \sim E_\tau/2$ d'où :

$$\alpha_{max} \sim \frac{2m_\tau}{E_\tau} = \frac{3,6 \text{ GeV}}{E_\tau}
 \tag{3.6}$$

Ce calcul est basé sur beaucoup d'approximations, cependant cette formule correspond approximativement aux résultats expérimentaux (c.f. section 3.3.1). Le π^0 se désintègre en deux photons, ce qui aura pour effet de d'élargir le jet. On fixe alors la limite à :

$$\alpha_{max} = \frac{5,0 \text{ GeV}}{E_\tau}
 \tag{3.7}$$

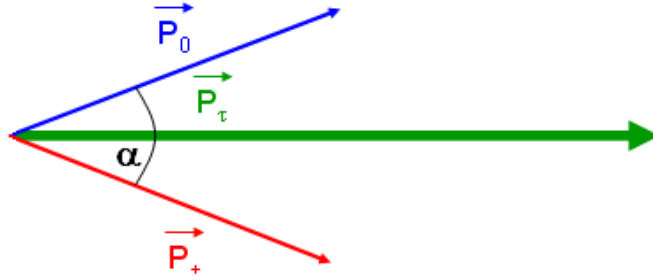


FIGURE 3.3 – Schéma simplifié de la désintégration en deux corps du lepton τ .

Dans la relation 3.6, nous avons négligé la présence du neutrino dans les produits de désintégration du τ car la largeur limite du jet est atteinte si ce neutrino n'emporte pas d'énergie. Mais la non-détection du neutrino tauique pose un problème lorsque l'on veut reconstruire la masse du τ comme le montre la figure 3.4. Nous remarquons dans cette figure que : premièrement le spectre en masse reconstruite du τ ne présente pas de pic sur sa masse réelle à 1,777 GeV et deuxièmement que la masse reconstruite dépend du mode de désintégration du τ car on reconstruit en réalité la masse des résonances ρ et a_1 .

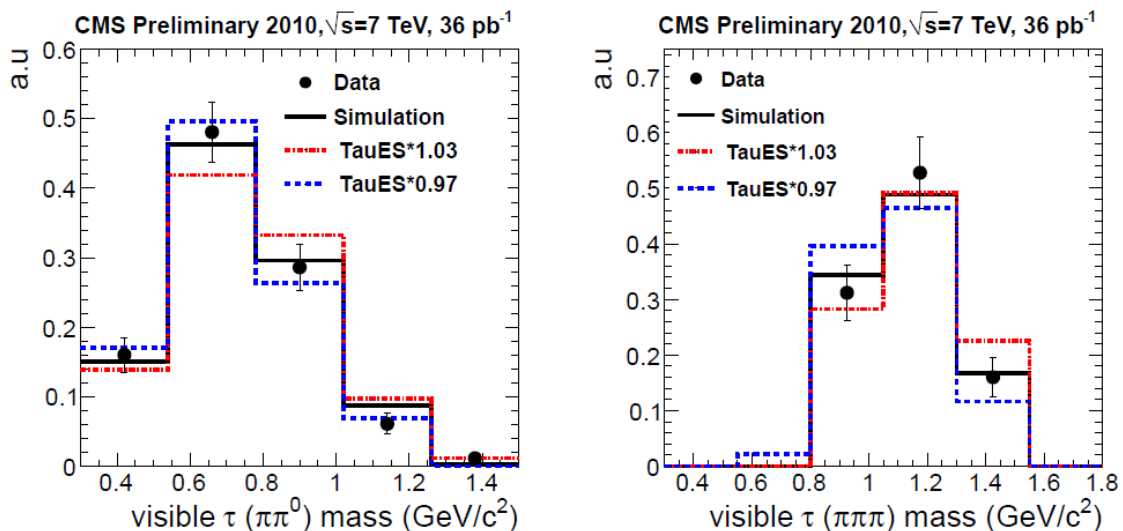


FIGURE 3.4 – Masse invariante reconstruite des τ dans les modes de désintégration $\tau^\pm \rightarrow \pi^\pm \pi^0$ et $\tau^\pm \rightarrow \pi^\pm \pi^\mp \pi^\pm$ pour des événements $Z \rightarrow \tau\tau \rightarrow \mu\tau_{had}$, extrait de [59].

3.2 Les sources de leptons τ

Plusieurs processus inélastiques sont susceptibles de donner des τ , certains de ces processus sont décrits par le Modèle Standard, d'autres sont décrits par des modèles théoriques à vérifier ou en cours de vérification dans le LHC.

Les sources de τ dans le Modèle Standard sont d'un nombre assez limité, ils sont issus de processus électrofaibles et sont donc régis par la théorie électrofaible. La production d'un ou deux leptons tau fera donc intervenir soit un photon, soit un des bosons vecteurs massifs (W ou Z).

Les différentes sources peuvent être résumées de la façon suivante (Figures 3.5 et 3.6) :

- Processus Drell-Yan, production de leptons via un photon ou un Z virtuel, l'apparition d'un boson Z réel est un cas particulier de processus Drell-Yan (Résonance).
- Désintégration d'un boson W : $W^- \rightarrow \tau^- \bar{\nu}_\tau$ et $W^+ \rightarrow \tau^+ \nu_\tau$.
- Désintégration d'un quark top : $t \rightarrow W^+ \bar{b} \rightarrow \tau^+ \nu_\tau \bar{b}$.
- Désintégration d'un boson de Higgs, ce boson n'est pas encore découvert mais la théorie prédit un couplage fort du Higgs aux particules massives et une désintégration favorisée vers les W,Z,top... qui sont des sources indirectes de τ . Mais il peut aussi se désintégrer directement en une paire de τ (toujours dû au couplage entre le Higgs et les particules massives). Les modes de désintégration du Higgs avec leurs rapports d'embranchement sont résumés dans la figure 3.6.

Chacun de ces processus du Modèle Standard a une signature bien particulière, on peut donc différencier les événements $W^\pm$ de ceux du top et des processus Drell-Yan. L'appari-

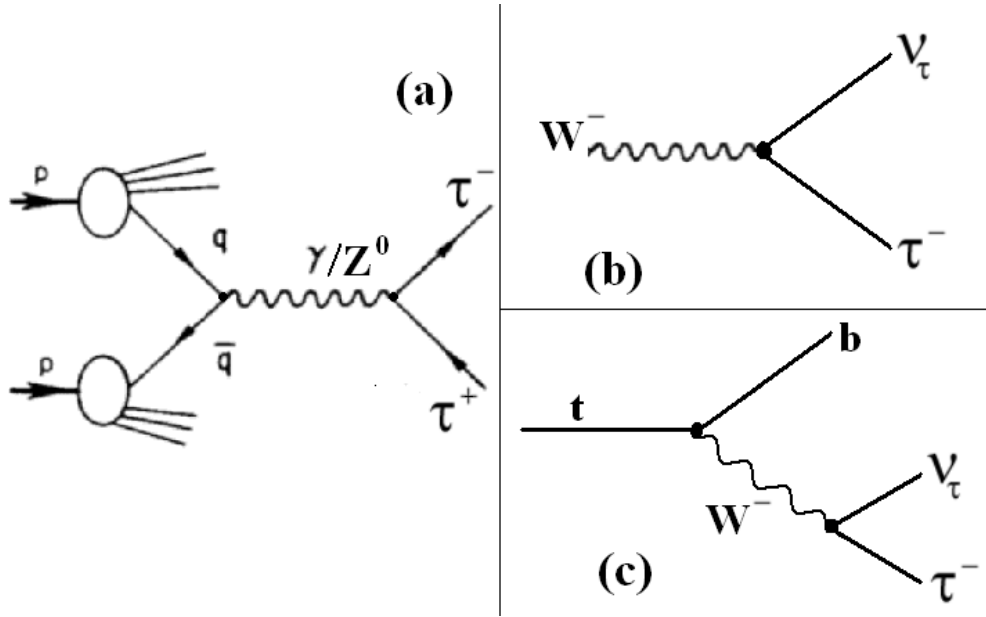


FIGURE 3.5 – Sources de τ dans le Modèle Standard. (a) Processus Drell-Yan, (b) Désintégration d'un boson W , (c) Désintégration d'un quark top.

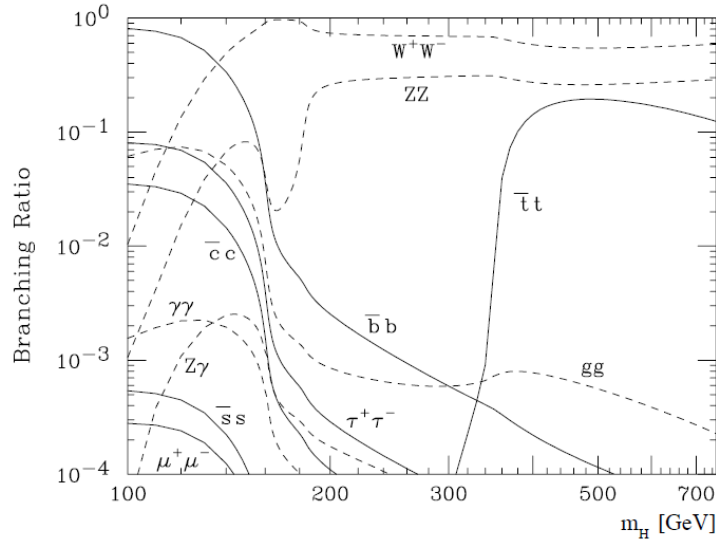


FIGURE 3.6 – Rapports d'embranchement des modes de désintégration du boson de Higgs (dans le Modèle Standard).

tion d'un boson Z réel est un cas particulier du processus Drell-Yan (résonnance), le boson Z peut être observé en reconstruisant la masse invariante du système formé par la paire de τ , la résonnance reconstruite de cette manière est large est de masse inférieure à celle du boson Z car une partie de la masse du système est emportée par les neutrinos tauiques obtenus lors de la désintégration des τ .

3.3 L'identification des jets de τ

Les jets hadroniques sont analysés à l'aide d'algorithmes d'identification afin de distinguer les jets provenant d'un τ de ceux issus d'autres sources comme les jets QCD qui représentent une source de bruit de fond. Une rejection des électrons et des muons permet de séparer les désintégrations hadroniques de τ .

Nous utilisons pour cette étude trois algorithmes :

- **Identification géométrique basée sur les coupures (cut-based)** effectue une différentiation entre les jets de τ et les jets de quarks en se reposant uniquement sur le fait que les jets de τ sont plus étroits et mieux isolés que ceux des quarks[52].
- **Hadron Plus Strips (HPS)** est une amélioration de l'algorithme basé sur les coupures, il analyse plus en profondeur la sous-structure du jet en reconstruisant l'un des mode de désintégration hadronique du τ .
- **Tau Neural Classifier (TaNC)** recherche aussi les jets étroits, isolés et compatibles avec l'un des modes de désintégration du τ . Cependant l'analyse n'est plus basée sur des coupures mais sur une analyse multivariable utilisant des réseaux de neurones.

Les performances de ces différents algorithmes seront décrits et discutés dans les chapitres 4 et 5.

3.3.1 Identification géométrique basée sur des coupures

Pour appliquer la méthode on procède comme suit :

- Etape 0 : On constitue une collection de candidats τ en recherchant tous les jets dans l'évènement, ceux-ci sont reconstruits par l'une des méthodes citées dans le chapitre 2.3.2. On utilisera préférentiellement l'algorithme anti-kt pour les avantages décrits dans la section 2.3.5.
- Etape 1 : Pour chaque jet de cette collection de jets, on rejette tous les jets n'ayant aucun hadron chargé.
- Etape 2 : On recherche dans chaque jet le hadron chargé de plus haute impulsion transverse (trace principale) avec un P_T au dessus d'un seuil de 5 GeV et situé à une distance $\Delta R < 0,1$ de l'axe du jet. ΔR étant la distance angulaire définie par la relation 3.8.

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \quad (3.8)$$

- Etape 3 : Un cône signal étroit (cône rouge dans la figure 3.7) est défini autour de la trace de plus haute impulsion (trace principale) dans lequel toutes les particules du jet doivent se situer. Les caractéristiques reconstruites du candidat τ sont obtenues à partir du contenu en particules du cône signal. Les caractéristiques de ces cônes sont résumées dans le tableau 3.3, le cône évolutif de taille variable est motivé par la relation 3.6. Les travaux présentés dans cette thèse utilisent un cône signal évolutif pour les hadrons chargés et un cône signal fixe pour les photons (le cône signal est plus large pour les photons afin de récupérer les π^0 dans le cas où les deux photons sont très séparés) . Le critère d'isolation

pour une particule P (photon ou hadron) est la distance angulaire ΔR entre la trace principale et la particule P.

- Etape 4 : Un cône d'isolation large est défini autour de la trace principale (cône bleu dans la figure 3.7). Aucune particule ne doit se trouver dans l'anneau défini par le cône d'isolation auquel on soustrait le cône signal. L'angle ΔR de ce cône d'isolation est fixe pour les hadrons chargés comme pour les photons et a une valeur $\Delta R = 0,5$.

TABLE 3.3 – Caractéristiques des coupures d'isolation.

Méthode avec cône fixe		
Cône signal Trajectographe		$\Delta R = 0,07$
Cône signal Calo. Hadronique		$\Delta R = 0,10$
Cône signal Calo. EM		$\Delta R = 0,15$
Cônes isolation		$\Delta R = 0,5$
Méthode avec cône évolutif		
Cône signal Trajectographe	formule	$\Delta R = \frac{5 \text{ GeV}}{E_T}$
	maximum	$\Delta R = 0,15$
	minimum	$\Delta R = 0,07$
Cône signal Calo. Hadronique		$\Delta R = 0,10$
Cône signal Calo. EM		$\Delta R = 0,15$
Cônes d'isolation		$\Delta R = 0,5$

- Etape 5 : Une sélection demandant une charge $|q| = 1$ et un nombre de trace égal à 1 ou 3 dans le cône signal (pour reproduire les modes de désintégration dominants du τ) est appliquée.
- Etape 6 : Une rejection des électrons et des muons est appliquée en utilisant plusieurs variables calorimétriques et associées au détecteur de muons. Cette rejection peut être appliquée pour présélectionner les candidats τ (dans les études d'échantillons Monte Carlo impliquant beaucoup de leptons légers dans leur état final) et intervient alors avant l'étape 1 (c.f. section 3.4).

3.3.2 Identification analysant les modes de désintégration du τ

La méthode Hadron Plus Strips[53] (HPS) à été conçue pour reconnaître les modes de désintégrations hadroniques dominants du τ (c.f. Table 3.3). C'est un algorithme d'identification basé sur des cônes, mais celui-ci n'utilise qu'un seul cône dans lequel on recherche les particules correspondant à l'un des 5 modes de désintégration dominants du τ dans la voie hadronique. La décision de conserver ce candidat se fait sur les objets physiques restant dans le jet (hors objets correspondant à l'un des 5 modes) : ce sont les particules non attribuées au candidat τ . Ce cône unique centré sur l'axe du jet a une largeur de :

$$\text{cone}(\Delta R) \leq \frac{2,8 \text{ GeV}}{p_T^{\text{jet}}} \quad (3.9)$$

Où p_T^{jet} est l'impulsion transverse du jet reconstruit.

L'identification des modes de désintégration se fait par la recherche d'une association de hadrons, i.e. des traces laissant leur énergie dans le calorimètre hadronique, et de "bandes" dans le calorimètre électromagnétique. Ces bandes sont des amas calorimétriques linéiques

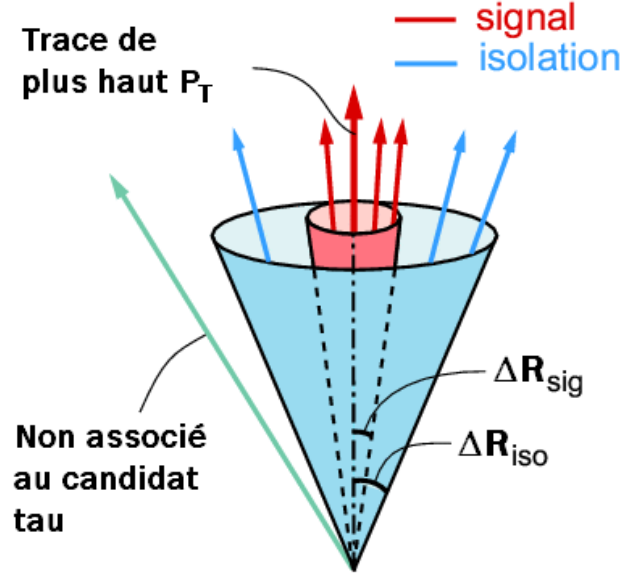
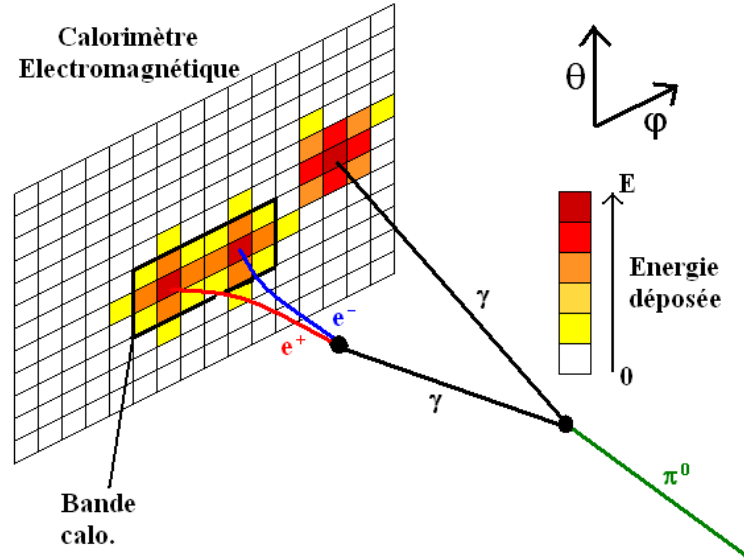


FIGURE 3.7 – Isolation d'un candidat tau par la méthode basée sur les coupures.

dont l'ouverture angulaire est en ϕ (c.f. Fig 3.8) afin d'avoir une reconstruction efficace des π^0 se désintégrant en deux photons, et en tenant compte des éventuelles conversions de photons en paires e^+e^- dont les traces s'ouvrent en ϕ . Une bande couvre un angle solide de $\Delta\eta \times \Delta\phi = 0,05 \times 0,2$ rad.

FIGURE 3.8 – Reconstruction d'un π^0 pour la méthode HPS.

Les modes de désintégration, dont les éléments doivent être recherchés dans le cône signal, sont les suivants :

- Hadron simple : Ce type reconstruit les désintégrations en un seul “prong” (i.e. une particule chargée) et les désintégrations $\tau^- \rightarrow h^- \pi^0$ où le π^0 a très peu d'énergie.
- Hadron plus une bande : Ce type a pour but de reconstruire la désintégration en un prong associé avec un π^0 dont les deux photons sont mal séparés. Cette bande prend en compte la possibilité que l'un ou les deux photons soient convertis en paires.
- Hadron plus deux bandes : Ce type reconstruit la désintégration $\tau^- \rightarrow h^- \pi^0$ dans le cas où les photons sont bien séparés.
- Trois hadrons : Ce type reconstruit les désintégrations en trois prongs, en requérant des charges compatibles $|\Sigma q| = 1$ et dont les traces proviennent d'un même vertex.

Les modes $\tau^- \rightarrow h^- h^+ h^- \pi^0$ et $\tau^- \rightarrow h^- \pi^0 \pi^0$ ne sont pas encore pris en compte par l'algorithme pour l'instant, mais peuvent être pris en compte pour améliorer l'efficacité.

Un candidat tau par la méthode HPS subit d'office une présélection, c'est un jet contenant l'un des 5 modes de désintégration pris en charge par l'algorithme, on doit donc au moins avoir un hadron chargé (similaire à la méthode “cut-based”) et on impose pour chacun de ces modes une fenêtre en masse invariante. De plus dans le cas où il y a la présence de deux bandes, une fenêtre contraint aussi la masse du π^0 : $0,05 \text{ GeV} < m_{\pi^0} < 0,2 \text{ GeV}$ pour retenir la masse diphoton la plus cohérente avec un π^0 . Les fenêtres en masse imposées pour les candidats τ sont les suivantes :

- Un hadron chargé : aucune restriction en masse, l'impulsion transverse du hadron en question doit être supérieure au seuil imposé pour le τ .
- Un hadron et une bande : $0,3 \text{ GeV} < m_{\pi^\pm \pi^0} < 1,3 \text{ GeV}$ (reconstruction de la résonance ρ).
- Un hadron et deux bandes : $0,4 \text{ GeV} < m_{\pi^\pm \pi^0 \pi^0} < 1,2 \text{ GeV}$ (reconstruction des résonances ρ et a_1).
- Trois hadrons : $0,8 \text{ GeV} < m_{\pi^\pm \pi^\mp \pi^\pm} < 1,5 \text{ GeV}$ (reconstruction de la résonance a_1).

Pour l'algorithme HPS, il existe trois niveaux de restriction pour l'isolation d'un candidat τ . L'isolation est appliquée sur les particules restantes une fois tous les constituants du jet susceptibles de provenir d'un τ identifiés et soustraits au jet, les seuils en P_T des particules restantes dans le cône sont :

- **Isolation relâchée** : Requiert l'absence de particules chargées avec $P_T > 1,0 \text{ GeV}/c$ et de photons avec $P_T > 1,5 \text{ GeV}/c$ dans le cône d'isolation.
- **Isolation moyenne** : Requiert l'absence de particules chargées avec $P_T > 0,8 \text{ GeV}/c$ et de photons avec $P_T > 0,8 \text{ GeV}/c$ dans le cône d'isolation.
- **Isolation étroite** : Requiert l'absence de particules chargées avec $P_T > 0,5 \text{ GeV}/c$ et de photons avec $P_T > 0,5 \text{ GeV}/c$ dans le cône d'isolation.

Les seuils en P_T utilisés pour l'isolation relâchée de l'identification HPS sont compatibles avec l'algorithme cut-based.

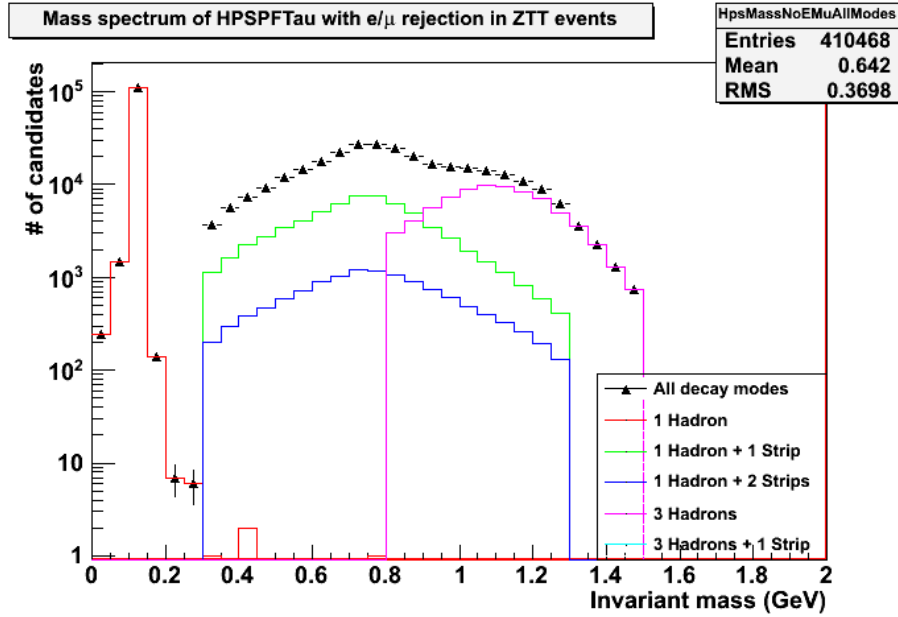


FIGURE 3.9 – Distribution des masses des candidats τ (échantillon $Z\tau\tau$ avec rejection de muon et d'électron) pour la méthode HPS avec les différentes catégories de la méthode HPS.

La figure 3.9 nous montre la reconstruction par l'algorithme HPS de la masse des candidats τ dans un échantillon $Z \rightarrow \tau\tau$. On y observe les résonnances intermédiaires issues de la désintégration du lepton τ : le meson π (pic à 140 MeV) le ρ (bosse à 770 MeV) et le a_1 (résonnance plus large : “épaule” à 1200 MeV).

3.3.3 Identification via des réseaux de neurones

Un algorithme d'identification de jet hadronique de τ appelé Classificateur Neural de Tau (Tau Neural Classifier : TaNC)[54] a été développé par la collaboration CMS. Cet algorithme, qui utilise comme entrée les candidats τ définis par la méthode cut-based, reconstruit les différents modes de désintégration hadronique dominants du τ en utilisant les candidats hadrons chargés et photons reconstruits par l'algorithme Particle Flow pour entraîner un ensemble de réseaux de neurones. Ceux-ci utilisent une collection de variables sensibles à certains modes de désintégration particuliers.

L'algorithme TaNC diffère de l'algorithme HPS par la méthode de reconstruction des π^0 . Le TaNC analyse les constituants photon d'un candidat τ et rejette d'abord les photons de $P_T < 0,5 \text{ GeV}/c$ pour supprimer le bruit dans le calorimètre dû aux photons de l'évènement sous-jacent. Les constituants photons obtenus après filtration sont associés par paires avec une masse invariante inférieure à $0,2 \text{ GeV}/c$ pour être cohérents avec l'hypothèse que cette paire de photons provienne de la désintégration d'un π^0 . Les constituants photons restants doivent satisfaire un prérequis de qualité : un photon célibataire ayant

une impulsion transverse dépassant 10% de l'impulsion totale du candidat τ est considéré comme π^0 , sinon il est filtré et ne sera plus comptabilisé comme constituant du jet.

La “trace principale” ainsi que les régions de signal et d'isolation sont celles définies par la méthode cut-based (Avec cône signal évolutif). Les réseaux de neurones ont été entraînés sur des échantillons $Z\tau\tau$ et QCD ($\hat{P}_T = 50 \text{ GeV}$) en utilisant comme entrée la collection de variables suivante :

- **ChargedOutlierAngleN** : ΔR entre le $N^{i\text{ème}}$ objet chargé dans la région de signal (trié par ordre croissant de P_T) et l'axe d'impulsion du candidat τ .
- **ChargedOutlierPtN** : Impulsion transverse du $N^{i\text{ème}}$ objet chargé dans la région d'isolation.
- **DalitzN** : Masse invariante de la “trace principale” (définie par la méthode cut-based) et du $N^{i\text{ème}}$ objet dans la région de signal.
- **Eta** : Pseudo-rapacité des particules dans la région de signal (région définie par la méthode cut-based).
- **InvariantMassOfSignal** : Masse invariante de toutes les particules reconstruites de la région de signal.
- **MainTrackAngle** : ΔR entre la “trace principale” et le 4-vecteur composite des constituants de la région de signal.
- **MainTrackPt** : Impulsion transverse de la “trace principale”.
- **OutlierNCharged** : Nombre d'objets chargés dans la région d'isolation.
- **OutlierSumPt** : Somme des impulsions transverses des objets de la région d'isolation.
- **PiZeroAngleN** : ΔR entre le $N^{i\text{ème}}$ objet π^0 dans la région de signal (trié par ordre croissant de P_T) et l'axe d'impulsion du candidat τ .
- **PiZeroPtN** : Impulsion transverse du $N^{i\text{ème}}$ objet π^0 dans la région de signal.
- **TrackAngleN** : ΔR entre le $N^{i\text{ème}}$ objet chargé dans la région signal (trié par ordre croissant de P_T) et l'axe d'impulsion du candidat τ en excluant la “trace principale”.
- **TrackPtN** : Impulsion transverse du $N^{i\text{ème}}$ objet chargé dans la région signal en excluant la “trace principale”.

Les réseaux de neurones regroupent tous ces paramètres en un seul dont la valeur est comprise entre 0 et 1 pour chacun des 5 modes de désintégrations hadroniques dominants du lepton τ : une valeur de 0 correspond à un jet totalement ressemblant à un jet QCD (quark-like) et une valeur de 1 correspond à un jet totalement ressemblant à un τ (tau-like). Cette variable est continue comme le représente la figure 3.10.

A partir de cette variable de sortie, on définit quatre points de travail tels que pour des candidats tau d'impulsion transverse $20 \text{ GeV} < P_T < 50 \text{ GeV}$, le taux de faux des di-jet QCD sont respectivement de 0,1%, 0,25%, 0,5% et 1%.

• L'entraînement des réseaux de neurones

Les échantillons utilisés pour entraîner les réseaux de neurones de la méthode TaNC sont typiques des signaux et bruits de fonds rencontrés pour les analyses physique impliquant des τ [54]. Ils sont constitués d'une part par des événements simulés $Z \rightarrow \tau\tau$ et d'autre part par des événements simulés de diffusions QCD $2 \rightarrow 2$ dures. Ces deux types d'échantillons sont divisés en 5 sous-ensembles correspondant aux 5 modes reconstruits de désintégration du lepton tau, chaque réseau de neurones de la méthode TaNC sera entraîné sur le sous-

TABLE 3.4 – Variables d'entrée pour les réseaux de neurones de la méthode Tau Neural Classifier et modes de désintégration concernés

Variable	$\pi^\pm \nu_\tau$	$\pi^\pm \pi^0 \nu_\tau$	$\pi^\pm \pi^0 \pi^0 \nu_\tau$	$\pi^\pm \pi^\mp \pi^\pm \nu_\tau$	$\pi^\pm \pi^\mp \pi^\pm \pi^0 \nu_\tau$
ChargedOutlierAngle1	•	•	•	•	•
ChargedOutlierAngle2	•	•	•	•	•
ChargedOutlierPt1	•	•	•	•	•
ChargedOutlierPt2	•	•	•	•	•
ChargedOutlierPt3	•	•	•	•	•
ChargedOutlierPt4	•	•	•	•	•
Dalitz1			•	•	•
Dalitz2			•	•	•
Eta	•	•	•	•	•
InvariantMassOfSignal		•	•	•	•
MaintTrackAngle		•	•	•	•
MainTrackPt	•	•	•	•	•
OutlierNCharged	•	•	•	•	•
OutlierSumPt	•	•	•	•	•
PiZeroAngle1		•	•		•
PiZeroAngle2			•		
PiZeroPt1		•	•		•
PiZeroPt2			•		
TrackAngle1				•	•
TrackAngle2				•	•
TrackPt1				•	•
TrackPt2				•	•

ensemble approprié.

Il existe une dépendance sur l'ouverture du jet en fonction de P_T et η . Cette dépendance est supprimée en appliquant un poids dépendant de ces deux paramètres aux candidats τ :

$$W(P_T, \eta) = \min(p_{sig}(P_T, \eta), p_{bf}(P_T, \eta)) \quad (3.10)$$

$$w_{sig}(P_T, \eta) = W(P_T, \eta)/p_{sig}(P_T, \eta) \quad (3.11)$$

$$w_{bf}(P_T, \eta) = W(P_T, \eta)/p_{bf}(P_T, \eta) \quad (3.12)$$

Où $p_{sig}(P_T, \eta)$ et $p_{bf}(P_T, \eta)$ sont les densité de probabilité des échantillons signal et bruit de fond après la selection du “pion principal” (candidat possédant un pion de $P_T > 5 \text{ GeV}$) et du mode de désintégration dominant.

→ Ces densités de probabilité ont été calculées pour des candidats de $P_T < 100 \text{ GeV}$, ceci aura des conséquences sur les analyses dont nous discuterons dans les chapitres 4 et 5.

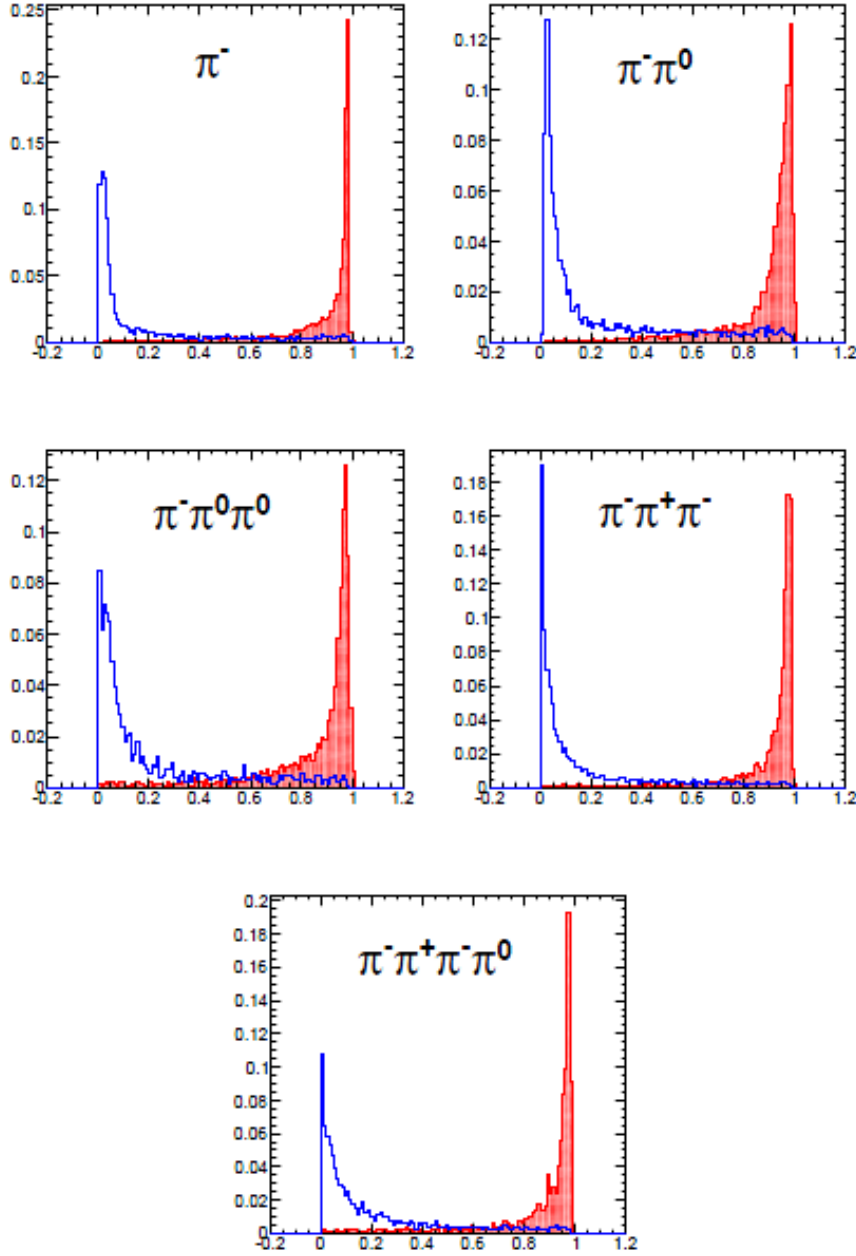


FIGURE 3.10 – Distributions de la variable retournée par les réseaux de neurones pour les cinq modes de désintégration utilisée dans l’algorithme TaNC (provenant de [54]), pour des évènements $Z\tau\tau$ (en rouge) et QCD (en bleu).

3.3.4 La réjection des électrons et des muons

• La réjection d’électrons

Les algorithmes d’identification de τ permettent de rejeter un maximum de jets QCD pour ne conserver que ceux issus de la désintégration d’un τ . Une fois le bruit de fond QCD supprimé, les électrons isolés produits dans les processus électrofaibles deviennent une source importante de faux τ [55].

L'algorithme Particle Flow fournit une pré-identification des électrons (c.f. section 2.3.4) basée sur une analyse multivariable rapide regroupant les informations du trajectographe et des calorimètres, incluant les photons émis par rayonnement de freinage des électrons. Cette préidentification a une efficacité de 90 à 95 % et supprime donc cette fraction du bruit de fond dû aux électrons. Pour rejeter les 5 à 10 % restants deux variables discriminantes supplémentaires sont introduites :

E/P : Somme des énergies des amas dans le calorimètre électromagnétique dans une bande d'ouverture $|\Delta\eta| \times \Delta\phi = 0.04 \times 0.5$ divisée par l'impulsion de la trace principale. Ce rapport doit être proche de 1 pour un électron isolé et dispersé autour d'une valeur plus petite pour les pions issus de la désintégration d'un τ .

$H_{3\times 3}/P$: Sommes des énergies des amas dans une tour calorimétrique hadronique (correspond à un rayon $\Delta R < 0.184$) centré sur le point d'impact extrapolé de la trace principale dans le calorimètre électromagnétique, divisé par l'impulsion de cette trace. Ce rapport doit être proche de 0 pour les électrons et distribué aléatoirement pour les pions.

Les coupures optimales pour la réjection des électrons sont :

- $E/P < 0,8$ et $H_{3\times 3}/P > 0,15$ pour les candidats τ non pré-identifiés comme électrons.
- $E/P < 0,9$ et $H_{3\times 3}/P > 0,05$ pour les candidats τ pré-identifiés comme électrons.

Dans le cas général, on ne considère pas les candidats pré-identifiés comme électron dans les analyses τ .

• La réjection de muons

Les muons peuvent aussi contaminer les analyses impliquant l'identification de τ . L'identification de muons par le détecteur CMS est très efficace (efficacité $> 99\%$). Aussi la réjection de muons se fait si la trace principale du candidat τ correspond à l'un des muons identifiés.

3.4 Estimation du bruit de fond QCD

La technique de "taux de faux" (fake-rate) fournit une méthode d'estimation par les données de la contribution du bruit de fond provenant d'une erreur d'identification d'un jet de quark ou de gluon dans les analyses impliquant des leptons tau dans l'état final[56].

Cette technique estime la probabilité qu'un jet QCD soit reconnu comme un τ par un algorithme d'identification (basée sur les coupures, HPS ou TaNC). La finalité de cette technique est de pouvoir estimer le bruit de fond QCD en ne se basant que sur les mesures faites au LHC. Il faut dans ce cas s'assurer de la sélection d'un lot d'évènements étant pour la majorité de la QCD car une sélection trop importante d'évènements contenant des jets de τ fausserait l'estimation du nombre de faux. Le but de cette méthode est d'avoir une vérification des estimations par simulations dans des conditions d'analyse particulières : l'algorithme considéré, la luminosité instantanée, le type de déclencheur choisi, etc...

Présélection des évènements QCD : On sélectionne les évènements multi-jet ($N_{jet} \geq 2$) largement dominés par la QCD. (c.f. section 2.1). Les évènements avec des bosons W et Z peuvent passer cette présélection, mais contribueront à une part très faible de l'échantillon

total (de l'ordre de 1 pour 10000). Des coupures supplémentaires peuvent être appliquées dans la présélection afin de supprimer davantage d'événements électrofaible par exemple un seuil maximal en énergie transverse manquante pour supprimer les événements $W \rightarrow \tau\nu$.

Estimation du taux de faux : Le taux est évalué en fonction de trois paramètres du jet, son impulsion transverse P_T , sa pseudo-rapacité η et la largeur du jet R_{jet} qui est défini par :

$$R_{jet} = \sqrt{E(\eta^2) + E(\phi^2)} \quad (3.13)$$

Où $E(\eta^2)$ (respectivement $E(\phi^2)$) sont les écarts-types en η (ϕ) des constituants du jet pondérés par leur énergie transverse E_T . Pour chaque triplet de ces trois variables discrétisées, on définit la probabilité qu'un jet soit identifié comme issu d'un τ :

$$P_{fr}(P_T^{jet}, \eta_{jet}, R_{jet}) := \frac{N_{jets}(P_T^{jet}, \eta_{jet}, R_{jet} | \text{candidat tau sélectionné})}{N_{jets}(P_T^{jet}, \eta_{jet}, R_{jet} | \text{candidat présélectionné})} \quad (3.14)$$

Une fois la statistique collectée suffisante, l'approche la plus simple pour avoir une estimation du nombre de jets QCD contribuant au bruit de fond dans les analyses impliquant l'identification de leptons τ est d'imposer le facteur de pondération à chaque candidat τ présélectionné comme étant égal à la probabilité qu'il soit reconnu comme tel.

$$w_{jet}^{simple}(P_T^{jet}, \eta_{jet}, R_{jet}) := P_{fr}(P_T^{jet}, \eta_{jet}, R_{jet}) \quad (3.15)$$

On peut donc avoir une estimation du nombre de jets QCD reconnus comme τ dans un intervalle en $P_T^{jet}, \eta_{jet}, R_{jet}$:

$$N_{fakes}(P_T, \eta, R_{jet}) = w_{jet}^{simple}(P_T^{jet}, \eta_{jet}, R_{jet}) \times N_{jets, presel.}(p_T, \eta, R_{jet}) \quad (3.16)$$

3.5 Architecture PAT et modules associés pour l'identification des τ

Nous avons vu dans la section 2 que les signaux mesurés par le détecteur CMS et passant la sélection des systèmes de déclenchement sont d'abord stockés au CERN (Tier 0) pour être ensuite envoyés aux instituts et universités (Tier 1 et 2) où les événements sont reconstruits, calibrés et analysés. Ces informations sont soumises à un premier traitement au niveau du Tier 0 pour conserver l'essentiel de l'information et optimiser leur stockage en convertissant les données sous un des formats vu dans la section 2 (AOD, RECO...).

Pour effectuer les études présentées dans les sections 4 et 5, le format de données utilisé est celui fourni par l'architecture *PAT* (Physics Analysis Toolkit).

3.5.1 Création des objets *PAT*

Physics Analysis Toolkit est une architecture adaptée à tous types d'analyses physiques pour CMS. Elle utilise comme entrée les objets reconstruit dans les formats AOD ou RECO (c.f. section 2.3.3). Un schéma du fonctionnement de PAT est présenté dans la figure 3.11[57] :

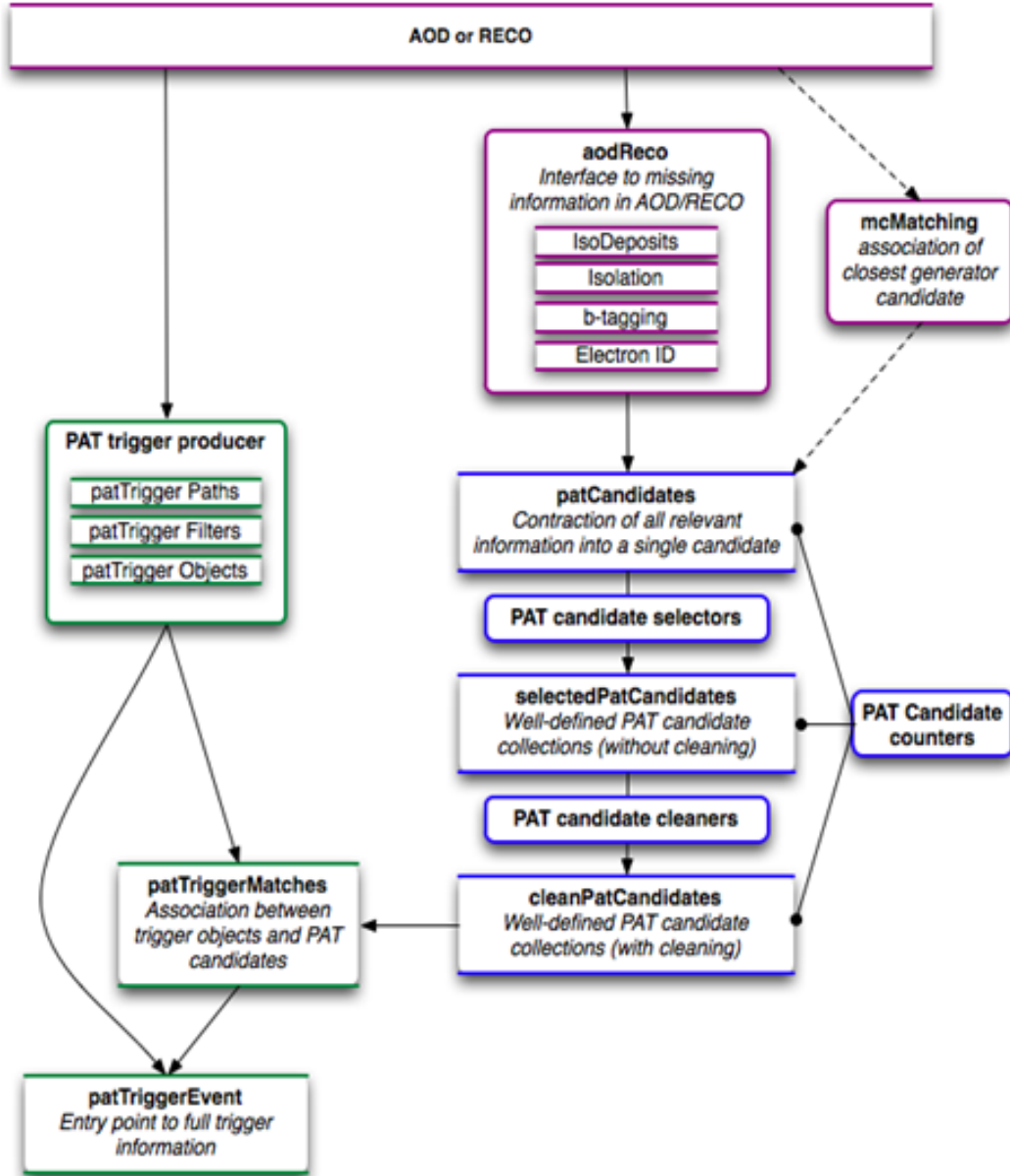


FIGURE 3.11 – Algorithme de création d'objets *PAT*.

Ce schéma nous présente les différentes étapes de la construction du format *PAT*, les objets reconstruits présents dans les formats AOD et RECO sont de deux types : Calo et ParticleFlow (PF). Les particules chargées (hadrons chargés et électrons) sont décrites dans le format de *Calo* par des grappes calorimétriques uniquement et leur mesure est moins précise à faible P_T . Les particules chargées *Particle Flow* sont décrites par leurs

trajectoires associées aux grappes calorimétriques, la mesure de leur énergie et de leur impulsion est beaucoup plus précise, c'est donc un meilleur choix pour les analyses physiques.

Pour les analyses impliquant des τ , on utilise une configuration de *PAT* qui crée une collection spécifique de candidats tau (objets *PAT* décrivant des τ) contenant toutes les informations nécessaires à son identification : des collections de variables discriminantes destinées à donner les décisions des différents algorithmes d'identification.

Pour les études présentées par la suite, un candidat tau *PAT* inclut la classe de candidats taus Particle Flow (tous les jets possédant au moins une trace i.e. une particule chargée) contenant les informations brutes Particle Flow, et sur lesquelles on peut ajouter les variables discriminantes *cut-based* et *TaNC*. Et pour l'identification *HPS* on crée une classe de candidats τ Particle Flow uniquement dédiés à l'algorithme *HPS*.

Une coupure cinématique ($P_T > 5 \text{ GeV}$ et $|\eta| < 2.5$) a été appliquée pour la création des candidats tau *PAT*, aucune coupure (Fig. 3.11 : PAT candidate cleaners) supplémentaire n'a été effectuée sur ces collections pour accepter le plus de candidats τ possible (en particulier dans le cas où ce candidat est compatible avec un autre type d'objet physique). Les présélections ainsi que les sélections de candidats sont appliquées ensuite manuellement par l'analyste (en utilisant les variables discriminantes).

3.5.2 L'environnement modulaire TauAnalysis

Le paquetage TauAnalysis est un module de CMSSW destiné aux analyses rapides pour la publication de données. Il allège à l'extrême le format *PAT* pour conserver uniquement un "ntuple" contenant l'essentiel, et permettre ainsi des représentations graphiques rapides des données relatives aux analyses impliquant des τ . Son principe de fonctionnement est le suivant :

- Collecte et recréation si nécessaire des objets τ Particle Flow et des objets associés relatifs à l'identification du lepton τ : les candidats spécifiques à l'identification *HPS*, les variables discriminantes des différents algorithmes (*Cut-based*, *HPS* et *TaNC*), les informations générées dans le cas où l'échantillon est une simulation Monte Carlo.
- Conversion au format *PAT* des objets *RECO* collectés lors de l'étape précédente : les objets *PAT* (c.f. section précédente) sont créés en conservant les informations désirées relatives à l'identification du τ . Les objets *PAT* sont clonés et ajustés pour chaque nouvel algorithme d'identification.
- Création des nTuples : les différentes variables associées aux objets *PAT* sont enregistrées dans un nTuple sous forme de "doubles" (variable numérique C++ avec double précision). Les objets *PAT* sont supprimés pour ne conserver que les variables relatives à l'identification de τ , tout en gardant les corrélations entre les différentes variables associées à un candidat.

→ A noter qu'une étape intermédiaire avant la création de figures graphiques a consisté en une copie brute des données des ntuples dans un "TTree" (objet *ROOT* similaire à un

ntuple) afin d'accélérer la lecture des données : le format TTree est un objet purement ROOT et fournit un certain confort d'utilisation car il est exploitable sans sortir de l'environnement ROOT.

L'avantage de la création d'un ntuple est le temps de calcul économisé lorsque l'on veut étudier l'identification du lepton τ sous plusieurs paramétrages différents. Cependant une partie de l'information est perdue dans un ntuple et celui-ci offre moins de flexibilité que la lecture directe d'objets contenus dans un échantillon au format AOD ou RECO (c.f. section 2.3.3).

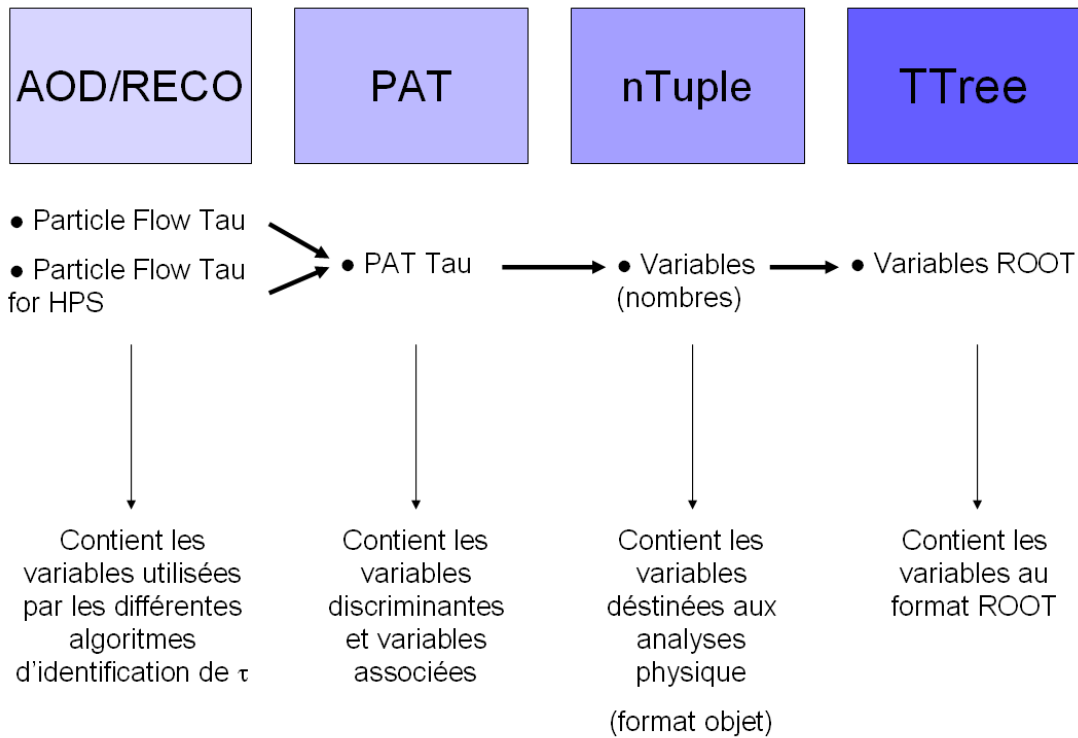


FIGURE 3.12 – Architecture logique utilisée pour l'identification du lepton τ .

Le schéma 3.12 nous résume les différentes étapes de l'analyse de données jusqu'à la création des figures présentées dans les sections 4 et 5. Une dernière étape consiste à analyser les entrées du ntuple afin de les regrouper sous forme d'histogrammes ROOT, pour les analyses physiques et finalement la présentation des résultats.

Chapitre 4

Efficacité et taux de mauvaise identification du lepton τ

4.1 Conditions d'analyse

4.1.1 Description des échantillons

Les échantillons étudiés dans ce chapitre sont de deux types :

- Des lots d'événements expérimentaux sélectionnés : ils comprennent l'ensemble des collisions enregistrées au LHC en 2010 (36 pb^{-1}) et triées par les systèmes d'acquisition de données (DAQ).
- Des simulations Monte Carlo fournissant un grand nombre d'événements contenant exclusivement des processus se rapportant à la physique du lepton τ (y compris des échantillons QCD pour les études de mauvaise identification). Les Monte Carlo simulent les processus du Modèle Standard et au-delà (SUSY), nous permettant de tester la réponse du détecteur CMS à ces événements particuliers.

4.1.1.1 Échantillons simulés Monte Carlo

• Les générateurs Monte Carlo

Les générateurs Monte Carlo simulent les interactions entre particules subatomiques en incluant les processus inclusifs comme les radiations d'état initial et final (c.f. section 1.3.1). Ils sont nécessaires pour comparer les modèles théoriques aux mesures expérimentales.

Les générateurs d'événements disponibles pour les simulations sont variés (PYTHIA, SHERPA, HERWIG, POMWIG, ExHuME, MADGRAPH, ALPGEN...). Ils diffèrent par les processus physiques simulés : création des gerbes partoniques et hadronisation, inclusion de modèles au-delà du Modèle Standard, prise en charge des interactions multiples, ordre des calculs perturbatifs, possibilité d'ajouter des processus définis par l'utilisateur.

La référence pour la génération d'événements reste PYTHIA qui inclut les gerbes partoniques, l'hadronisation par le modèle de Lund (Cordes, c.f. section 1.3.5.2), les interactions à partons multiples. La plupart des extensions au Modèle Standard y sont implémentées.

La désintégration de leptons τ est modélisée à l'aide d'une librairie supplémentaire (codée en C++) nommée TAUOLA décrivant la topologie complète de l'état final de ces désintégrations : le neutrino tauique, les résonnances intermédiaires et la structure complète de spin des produits de désintégrations.

Les interactions avec le détecteur CMS des particules générés sont simulées par d'autres logiciels comme GEANT4 et finalement la reconstruction complète d'un événement dans le détecteur CMS est effectuée sous l'environnement standard CMSSW (c.f. section 2.3.1).

• Les échantillons de simulation contenant des τ

Les processus du Modèle Standard et de SUSY donnant des τ sont décrits dans la section 3.2, ils seront utilisés pour les déterminations d'efficacité :

- Boson Z se désintégrant en deux τ : La probabilité qu'un τ se désintègre en hadrons est donné dans la section 3.1.2. Le boson Z peut être accompagné de un ou plusieurs jets QCD (pour les raisons citées dans la section 1.3.1).
- Boson W se désintégrant en un τ plus un neutrino tauique, il est souvent associé à la production d'un ou de plusieurs jets (pour les mêmes raisons que pour le boson Z).
- Production $t\bar{t}$: le quark top se désintègre en un boson W et un quark beau (Rapport d'embranchement proche de 100%) : $t\bar{t} \rightarrow W^+W^-b\bar{b} \rightarrow \ell^+\ell^-b\bar{b}$ ce qui implique une grande multiplicité de particules dans l'état final.
- Évènements SUSY LM1 et LM2 (c.f. section 1.4.2 et Fig 1.5) : ils peuvent donner 4 leptons et 2 à 4 quarks dans le cas d'une production squark-squark, gluino-gluino ou squark-gluino. Une production directe de stau, ou de neutralino est possible : leur masse est moins élevée que celle des squarks/gluinos et donc l'énergie totale sera moins importante dans ces évènements.

• Les échantillons de bruit de fond pour les analyses τ

Le bruit de fond principal à l'identification du lepton τ est dû aux jets QCD, les évènements QCD sont catégorisés en fonction de leur $\hat{P}_T$ ¹.

La section efficace des processus QCD ainsi que la topologie de l'évènement dépend de ce paramètre $\hat{P}_T$, les échantillons QCD sont simulés par intervalles de $\hat{P}_T$ jusqu'à l'échelle Multi-TeV. On peut alors reproduire les spectres des jets QCD dans les données du LHC en combinant les différents intervalles simulés et en les renormalisant avec leur section efficace (Relation 4.1 et Fig. 4.1) et leur nombre d'évènements.

$$N_{jets}(\hat{P}_T) = N_{jets}(\hat{P}_{Tref}) \times \frac{\sigma_{\hat{P}_T}}{\sigma_{\hat{P}_{Tref}}} \times \frac{N_{event,ref}}{N_{event}} \quad (4.1)$$

Les échantillons utilisés pour les différentes études sont résumés dans le tableau 4.1. Ce sont des simulations Monte Carlo incluant plusieurs sources de τ dans le cadre du Modèle Standard (Boson Z et W, $t\bar{t}$) et deux échantillons SUSY (LM1 et LM2). Les probabilités de mauvaise identification sont estimées à l'aide d'échantillons QCD allant d'un $\hat{P}_T$ de 15

1. Impulsion transverse des partons directement issus de la collision primaire.

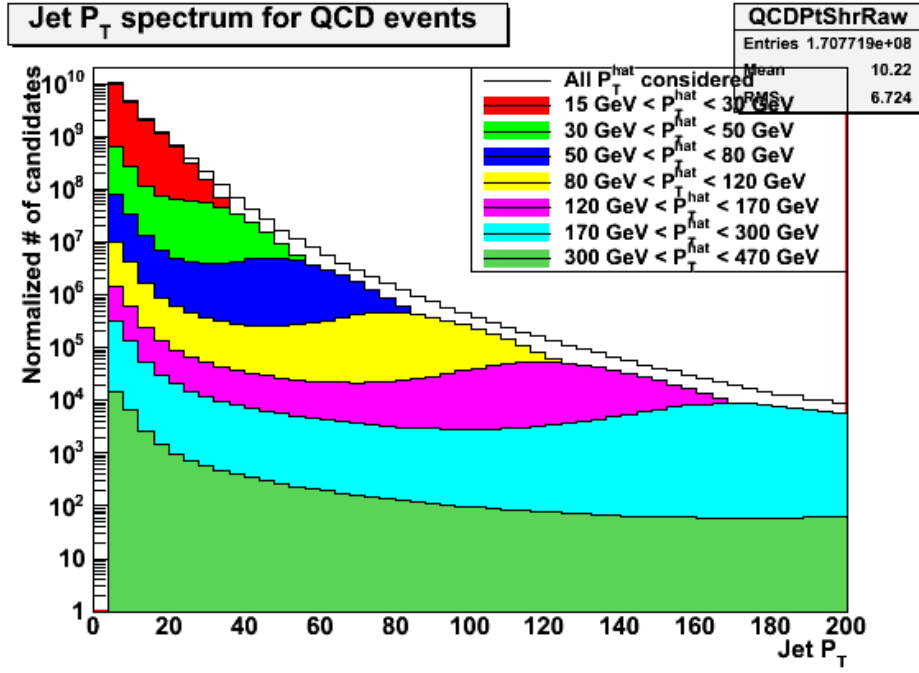


FIGURE 4.1 – Superpositions des sous-échantillons QCD constituant le Monte Carlo final : les spectres sont rééchelonnés en P_T en fonction de leur section efficace et du nombre d'évènements contenu dans le sous-échantillon.

GeV à 470 GeV.

TABLE 4.1 – Échantillons simulés Monte Carlo

Nom de l'échantillon	Processus	Générateur	$X_{section}$ (pb)	Nb d'évts
/QCD_Pt_15to30_Tune1.7TeV_pythia8/Fall10-START38_V12-v1/GEN-SIM-RECO	QCD $\hat{P}_T$ 15 à 30 GeV	Pythia 8	$7,883.10^8$	5420080
/QCD_Pt_30to50_Tune1.7TeV_pythia8/Fall10-START38_V12-v1/GEN-SIM-RECO	QCD $\hat{P}_T$ 30 à 50 GeV	Pythia 8	$5,033.10^7$	3264660
/QCD_Pt_50to80_Tune1.7TeV_pythia8/Fall10-START38_V12-v1/GEN-SIM-RECO	QCD $\hat{P}_T$ 50 à 80 GeV	Pythia 8	$6,027.10^6$	3191546
/QCD_Pt_80to120_Tune1.7TeV_pythia8/Fall10-START38_V12-v1/GEN-SIM-RECO	QCD $\hat{P}_T$ 80 à 120 GeV	Pythia 8	$7,510.10^5$	3190579
/QCD_Pt_120to170_Tune1.7TeV_pythia8/Fall10-START38_V12-v1/GEN-SIM-RECO	QCD $\hat{P}_T$ 120 à 170 GeV	Pythia 8	$1,119.10^5$	3030480
/QCD_Pt_170to300_Tune1.7TeV_pythia8/Fall10-START38_V12-v1/GEN-SIM-RECO	QCD $\hat{P}_T$ 170 à 300 GeV	Pythia 8	$2,409.10^4$	3191280
/QCD_Pt_300to470_Tune1.7TeV_pythia8/Fall10-START38_V12-v1/GEN-SIM-RECO	QCD $\hat{P}_T$ 300 à 470 GeV	Pythia 8	$1,201.10^3$	3121560
/DYToTauTau_M-20_TuneZ2.7TeV-pythia6-tauola/Fall10-START38_V12-v1/AODSIM	$Z \rightarrow \tau\tau + \text{jets}$	Pythia 6	1300	1980346
/WToTauNu_TuneZ2.7TeV-pythia6-tauola/Fall10-START38_V12-v1/AODSIM	$W \rightarrow \tau\nu + \text{jets}$	Pythia 6	7899	5221750
/TTJets_TuneD6T.7TeV-madgraph-tauola/Fall10-START38_V12-v2/AODSIM	$t\bar{t}$	Madgraph	94	1254843
/LM1.SUSY_sftsht.7TeV-pythia6/Fall10-E7TeV_ProbDist.2010Data.BX156.START38_V12-v1/GEN-SIM-RECO	SUSY LM1	Pythia 6	4,888	219190
/LM2.SUSY_sftsht.7TeV-pythia6/Fall10-E7TeV_ProbDist.2010Data.BX156.START38_V12-v1/GEN-SIM-RECO	SUSY LM2	Pythia 6	0,6027	210000

4.1.1.2 Échantillons de données du LHC

La probabilité d'occurrence d'un processus donné dans une collision proton-proton est relié à sa section efficace. Si on mesure toutes les collisions sans distinction il en résulte un mélange d'événements dominés par les processus communs (notamment la QCD à bas $\hat{P}_T$, c.f. Tab. 4.1). Il faut alors classifier les différents types d'événements pour diviser le total des données en échantillons adaptés à une ou plusieurs analyses physiques. Cet objectif est atteint par les systèmes de déclenchement à bas et haut niveau du LHC qui permettent non seulement de sélectionner uniquement les collisions de plus haute énergie, mais permettent aussi de classifier les événements selon leur contenu.

Pour sélectionner les événements QCD dans les données, on se base sur le système de déclenchement à haut niveau pour les jets, plusieurs seuils de déclenchement sont disponibles (Table 4.2).

• Le système de déclenchement à haut niveau pour les jets

Les systèmes de déclenchement de haut niveau (2 et 3) sont activés par un ou plusieurs objets physiques (particule, jet, énergie transverse manquante...) à la condition que ces objets aient aussi activé un déclenchement de niveau 1. En ce qui concerne les systèmes de déclenchement de niveau 1 pour les jets, ceux-ci sont activés si l'un des jets présent dans l'événement dépasse un seuil en énergie. Ce seuil s'applique à la mesure de jets non-corrigés en énergie (énergie du jet mesurée uniquement par les calorimètres). Ce jet doit ensuite activer les systèmes de déclenchement à haut niveau avec un seuil énergie plus élevé qui s'applique ici aussi à la mesure de jets non-corrigés en énergie.

TABLE 4.2 – Description des systèmes de déclenchement pour les jets

Déclencheur à haut niveau	Déclencheur niveau 1	Seuil en énergie non corrigé
HLTJet15U	L1SingleJet6U	$E > 15 \text{ GeV}$
HLTJet30U	L1SingleJet20U	$E > 30 \text{ GeV}$
HLTJet50U	L1SingleJet30U	$E > 50 \text{ GeV}$
HLTJet70U	L1SingleJet40U	$E > 70 \text{ GeV}$
HLTJet100U	L1SingleJet60U	$E > 100 \text{ GeV}$

• La sélection d'événements dans les données : la méthode de “tag and probe”

Pour mesurer une efficacité à partir des données, on utilise la méthode dite de “tag and probe” [58]. Cette méthode est utilisée pour mesurer les efficacités d'identification d'objets physiques détectés par CMS (électron, muons, jets...). Son principe général est le suivant :

On définit un objet “étiquette” (tag) comme un objet physique (particule ou jet) passant des sélections très strictes : le taux de faux doit être inférieur à 1%. Les systèmes de déclenchement ont le plus souvent une efficacité d'identification suffisante pour définir un objet “étiquette”. Cet objet est ensuite apparié à un objet “sonde” (probe) du même type qui sera sélectionné sur des critères beaucoup plus souples.

Dans le cas où les deux objets passent les critères de sélection “étiquette”, ils sont considérés aussi en tant que “sonde”.

• La méthode de “tag and probe” appliquée à la sélection d’évènements QCD

Pour sélectionner les évènements QCD on recherche comme signature les évènements contenant au moins 2 jets. Les processus hors QCD impliquant un état final multijet ont une section efficace négligeable devant celle des processus QCD (en particulier si l’énergie des jets est faible : biais minimum). Les jets “étiquette” sont sélectionnés par les systèmes de déclenchement à haut niveau (tableau 4.2). Tous les autres jets contenus dans l’évènement sont potentiellement appariés à ce (ou ces) jet(s) “étiquette” en y appliquant des coupures cinématiques ($P_T > 5 \text{ GeV}$ et $\eta < 2.5$) pour supprimer les jets QCD trop mous n’ayant aucune possibilité de passer les sélections du lepton τ (Gain de place sur les systèmes de stockage de données).

Seuls les jets “sonde” sont pris en compte dans les calculs de taux de faux afin de supprimer le biais causé par la sélection appliquée aux jets “étiquette” par les systèmes de déclenchement de haut niveau (sauf dans le cas où le jet appartient aux deux catégories).

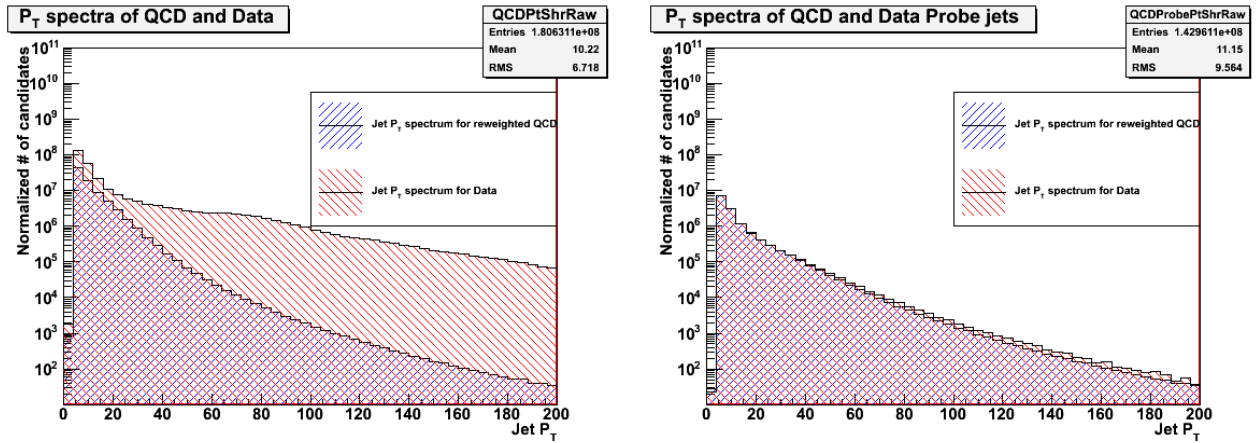


FIGURE 4.2 – Spectre en impulsion transverse (normalisée à partir de l’intégrale des données) des jets dans les données et les simulations QCD pour a) Tous les jets. b) Les jets “sonde”.

La figure 4.2a nous montre la superposition des spectres en P_T des tous les jets dans les données et dans les simulations MonteCarlo. La différence très nette entre données et Monte Carlo est due au contenu en jets qui n’est pas le même selon le seuil de déclenchement car chaque échantillon contient un mélange d’évènements sélectionnés par différents seuils de déclenchement (Fig. 4.3). En considérant les jets “sonde” associés aux jets “étiquette” n’utilisant qu’un unique seuil de déclenchement (parmi ceux cités dans la table 4.2) nous pouvons retrouver un contenu en jets similaire dans les données et les simulations Monte Carlo (Fig. 4.2b).

• Échantillons de données utilisés

Nous désirons sélectionner uniquement les échantillons contenant des jets QCD pour estimer les taux de faux τ dus à ces objets. Au fur et à mesure que le LHC produit des

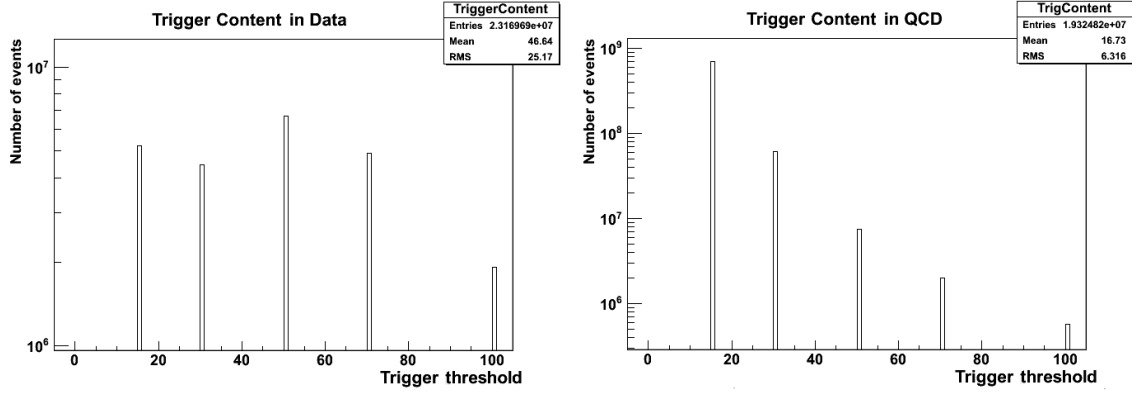


FIGURE 4.3 – Distribution du nombre d'événements selon le seuil de déclenchement (en GeV) ayant été activé dans a) Les données 2010 b) Les simulations QCD Monte Carlo.

données, les échantillons sont de plus en plus conséquents (la statistique augmente). Au début de l'année 2010 les échantillons Jet, Missing E_T et τ étaient regroupés dans le même lot de données, puis les τ ont été mis à part et enfin les échantillons contenant de l'énergie transverse manquante (Table 4.3). Les systèmes de déclenchement cités dans cette table sont ceux utilisés pour la première classification des données (sélection décidée par la collaboration pour la création des dataset), les seuils de déclenchement utilisés pour les calculs de taux de faux sont ceux du tableau 4.2.

TABLE 4.3 – Échantillons de données

Nom de l'échantillon	Sélection par déclenchement	Nombre d'événements
/JetMETTau/Run2010A-Nov4ReReco_v1/AOD	Jet et Energie transverse manquante	8470356
/JetMET/Run2010A-Nov4ReReco_v1/AOD	Jet et Energie transverse manquante	21800179
/Jet/Run2010B-Nov4ReReco_v1/AOD	Jet	16720183

→ Correspond à une luminosité intégrée de $36,1 \text{ pb}^{-1}$, la totalité des collisions du LHC en 2010

4.1.2 Définition de l'efficacité d'identification

L'efficacité d'un processus d'identification quelconque est la fraction des entités ayant passé les critères d'identification. Cette efficacité est souvent donnée en fonction d'un paramètre relatif à l'entité.

Pour le lepton τ on recherche les désintégrations hadroniques (c.f. section 3.1). Le nombre de τ réellement produits dans les collisions du LHC est à priori inconnu, mais le Modèle Standard nous fournit les différentes sections efficaces et rapports d'embranchement des processus produisant τ qui ont été vérifiées dans les expériences en collisionneur antérieures au LHC. On a alors une bonne estimation du nombre et des caractéristiques des τ issus des processus du Modèle Standard dans les simulations Monte-Carlo.

Les simulations contiennent toutes les informations qui nous seraient inconnues dans les véritables collisions, on peut alors calculer l'efficacité par rapport à des τ générés se

désintégrant hadroniquement en appliquant l'identification sur le candidat reconstruit associé i.e. celui qui serait mesuré expérimentalement : on utilise pour cela la distance ΔR donnée par la relation (3.8) entre le jet hadronique de τ généré que l'on veut étudier et le jet reconstruit le plus proche qui constituera notre candidat τ . On utilise les quantités générées (P_T , η ...) du jet hadronique de τ sans le neutrino tauique associé pour calculer les efficacités afin de limiter les effets expérimentaux.

L'efficacité est donnée en fonction d'un paramètre X par la relation :

$$\epsilon(X) = \frac{N_{candidats}(X) \mid \text{Passant le test d'identification}}{N_{\tau}(X) \mid \text{généré}} \quad (4.2)$$

Où " X " est une grandeur caractéristique du tau généré : impulsion transverse, pseudo-rapacité, angle azimutal, etc.

On peut diviser la relation 4.2 en deux termes, une efficacité d'association et une efficacité d'identification :

$$\epsilon(X) = \epsilon_{assoc}(X) \times \epsilon_{id}(X) \quad (4.3)$$

$\epsilon_{assoc}(X)$: **Efficacité d'association**, relatif à la fraction des candidats jet ayant une direction correspondante à celle du jet de τ généré. Les candidats sont sélectionnés par la condition $\Delta R_{Candidat, \tau \text{ généré}} \leq 0,5$ comme critère d'association réussie. Le choix de cette coupure est motivé par distribution ΔR entre le τ et sont plus proche candidat reconstruit (Fig. 4.4). On voit dans cette figure un léger excès de candidats situés entre $\Delta R = 0,15$ et $\Delta R = 0,5$ par rapport au τ généré pour les échantillons $t\bar{t}$ et SUSY que l'on n'observe pas dans les événements électrofaibles.

$$\epsilon_{assoc}(X) = \frac{N_{\tau \text{ généré}}(X) \mid \text{Associé à un candidat reconstruit par } \Delta R}{N_{\tau}(X) \mid \text{généré}} \quad (4.4)$$

$\epsilon_{id}(X)$: **Efficacité d'identification**, relatif à la fraction des candidats passant le test d'identification des jets hadroniques de τ . La sélection des candidats se fait par l'un des algorithmes d'identification du τ .

$$\epsilon_{id}(X) = \frac{N_{candidats}(X) \mid \text{Passant le test d'identification}}{N_{\tau}(X) \mid \text{associés}} \quad (4.5)$$

4.1.3 Définition du taux de mauvaise identification de τ

Le taux de mauvaise identification de τ dans une analyse est la fraction des entités indésirables à l'analyse passant tout de même les critères d'identification, constituant ainsi un bruit de fond.

Dans les analyses impliquant les leptons τ se désintégrant en hadrons, les jets QCD constituent le principal bruit de fond pour l'identification d'un lepton τ , les autres sources

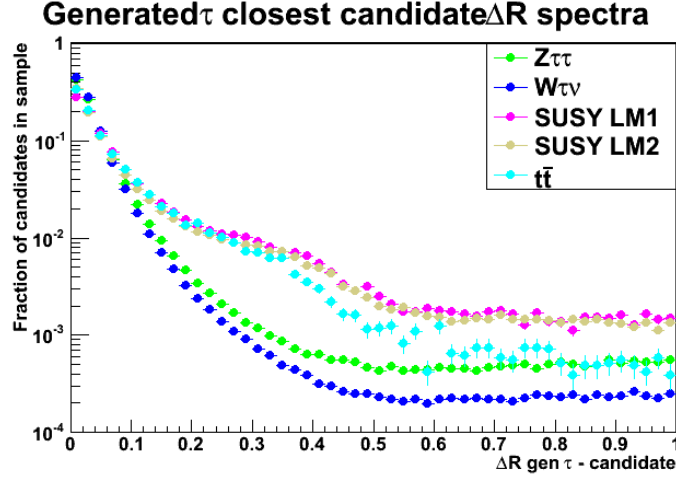


FIGURE 4.4 – Distribution des distances entre les τ (information générée) et leur candidat reconstruit le plus proche dans divers échantillons.

étant les électrons et les muons. La fraction de mauvaise identification est calculée de la façon suivante :

$$P_{faux}(X) = \frac{N_{candidats}(X) \mid \text{Passant le test d'identification}}{N_{candidats}(X) \mid \text{présélectionnés}} \quad (4.6)$$

Où X est ici aussi une caractéristique du jet. On peut alors étudier l'efficacité et le taux de mauvaise identification d'une méthode en fonction du même paramètre, pour estimer les performances de cette méthode. La présélection pour les τ dépend de l'échantillon considéré, les plus courantes sont des coupures en P_T , η car les jets d'impulsion transverse basse et de rapidité plus haute sont moins bien reconstruits que les jets centraux d'impulsion élevée.

• Présélection des candidats pour les calculs de mauvaise identification

La présélection des candidats est nécessaire lorsque l'on veut limiter la dépendance de l'identification en certains paramètres. Elle est donc spécifique à une étude donnée. Lorsque l'on veut étudier l'identification du lepton τ certaines présélections “générales” doivent être appliquées sur les candidats τ .

La présélection commune à toutes les études et tous les types d'échantillons est une sélection de candidats au-dessus d'un seuil en P_T et en dessous d'un seuil en η . Nous développerons les motivations de ces coupures dans les sections suivantes.

Pour les échantillons Monte-Carlo la présélection commune aux candidats jet QCD est l'exigence d'une non-association entre les candidats et les τ générés se désintégrant hadroniquement afin de ne pas accepter de τ dans la collection de candidats jet QCD. On utilise alors le critère de non-association $\Delta R > 0,5$ entre le candidat considéré et le τ généré le plus proche (se désintégrant hadroniquement).

Pour les échantillons QCD (données et MC), la présélection commune est l'exigence de candidats jets “sonde” comme nous l'avons vu dans la section 4.1.1.2. Le seuil de

déclenchement doit être le même pour les données et les Monte-Carlo pour effectuer des comparaisons entre la théorie (Monte-Carlo) et l'expérience (Données).

Enfin des présélections d'événements peuvent être appliquées dans le cas où on recherche des environnements particuliers pour ces événements, ceci se fait à l'aide de variables décrivant l'environnement. Ce point sera développé dans le chapitre 5.

4.2 Efficacité d'identification du lepton tau

4.2.1 Spectres des variables associées aux jets de tau

Dans la section 3.4 nous avons mentionné la reconnaissance de jets en fonction des paramètres P_T , η et R_{jet} (Définition de R_{jet} dans l'équation 3.13). C'est donc la réponse des algorithmes d'identification en fonction de ces différents paramètres qui sera le sujet d'études dans cette section. Les efficacités sont décrites en fonction de l'un de ces trois paramètres, nous recherchons les corrélations qui peuvent exister ceux-ci.

Le tracé de chacune de ces trois variables en fonction des autres (Fig. 4.5) nous révèle qualitativement les corrélations qui existent entre elles :

- **Corrélations entre P_T et η pour les τ** : Dans les événements électrofaibles les taus issus de la désintégration d'un W/Z sont relativement uniformes en rapidité en comparaison des événements supersymétriques impliquant des particules beaucoup plus lourdes et donc les τ sont caractérisés par une pseudo-rapidité $|\eta|$ plus centrale comme le montre la figure 4.5a). Les variables P_T et $|\eta|$ sont relativement corrélées dans le cas où on a initialement des particules lourdes produites dans la collision proton-proton (SUSY, $t\bar{t}$).

- **Corrélations entre P_T et R_{jet} pour les τ** : Ici nous mesurons les caractéristiques des jets de τ , on a donc une corrélation entre la largeur du jet et son impulsion transverse comme expliqué dans la section 3.1.2. : la variable R_{jet} (section 3.4 eq. 3.13) une mesure statistique de la largeur du jet prenant en compte l'éloignement ΔR entre l'axe d'un jet et ses constituants pondérée par l'impulsion transverse de ces mêmes constituants. En comparaison, la grandeur ΔR décrite dans la section 3.3.1 est une mesure purement angulaire de la largeur du jet car il s'agit de la distance entre l'axe d'un et son constituant le plus éloigné. On constate dans la figure 4.5b) la tendance globale à avoir des jets plus larges dans les événements SUSY que dans les événements électrofaibles. L'hypothèse est que cette tendance est un effet d'environnement contaminant le jet, ce sujet sera développé plus en détail dans le chapitre 5.

- **Corrélations entre η et R_{jet} pour les τ** : La figure 4.5c) nous montre que la largeur du jet hadronique de τ est indépendant de sa direction (en pseudo-rapidité), et ceci dans les événements électrofaibles comme dans les événements SUSY. L'élargissement de R_{jet} aux pseudo-rapidités $|\eta| \approx 1,5$ observé dans les événements $Z\tau\tau$ est un effet expérimental dû à la transition entre les tonneaux et les bouchons des différentes parties du détecteur CMS (c.f. Fig. 2.4) qui n'est pas visible pour les événements SUSY : ceux-ci occupant préférentiellement la région centrale.

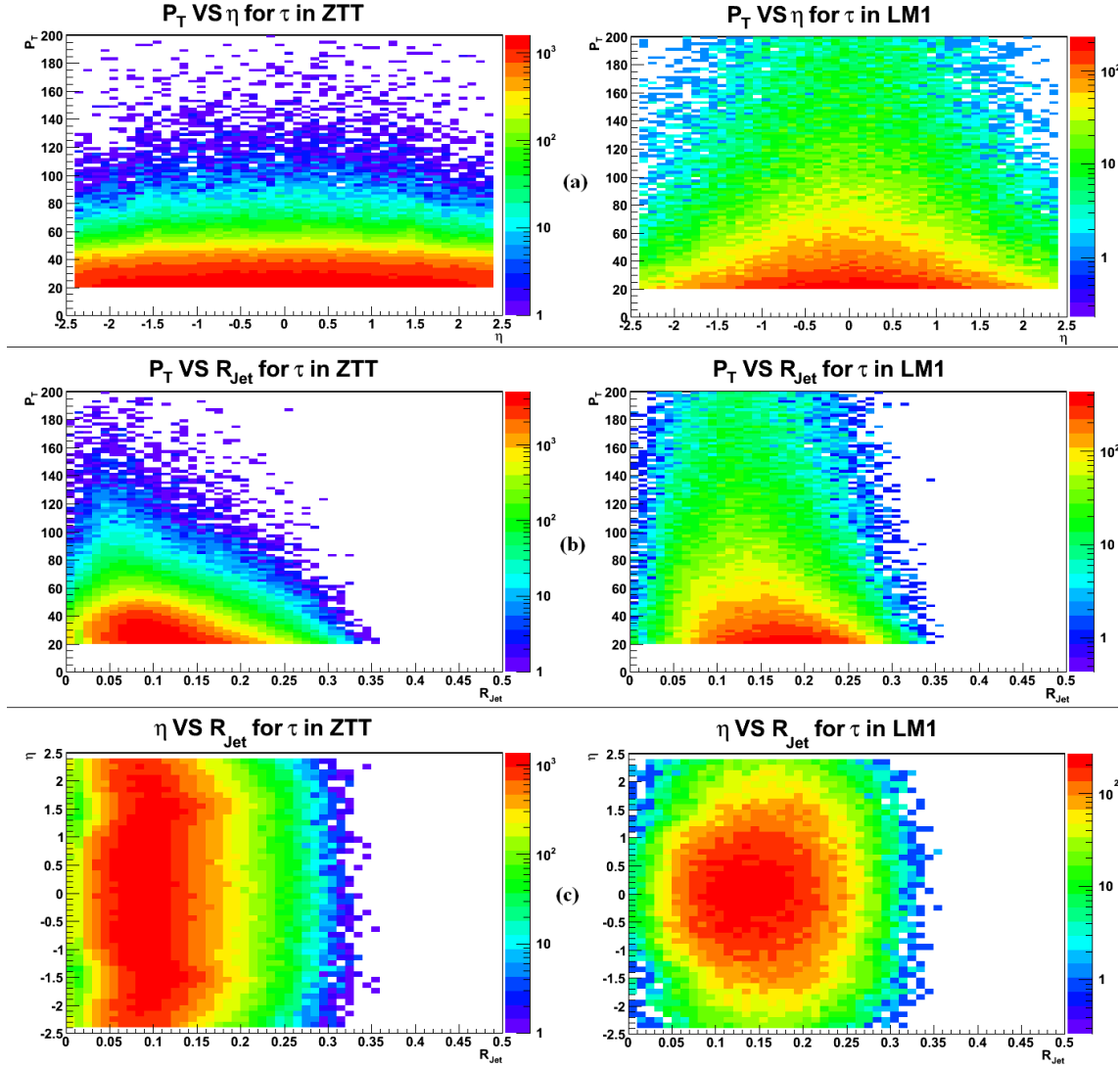


FIGURE 4.5 – Corrélations pour les τ dans un échantillon $Z\tau\tau$ (à gauche) et SUSY LM1 (à droite) entre les variables (information générée) a) P_T vs η , b) P_T vs R_{Jet} , c) η vs R_{Jet} .

4.2.2 Efficacités calculées avec les quantités générées

4.2.2.1 Dépendance en P_T

On étudie tout d'abord l'influence de P_T sur les efficacités d'identification du τ , en commençant par les efficacités "idéales" calculées uniquement en considérant les produits de désintégration des τ générés. La formule (4.2) admet alors comme numérateur le nombre de ces jets idéaux passant les critères d'identification, on a alors $\epsilon_{assoc} = 1$. Les résultats sont présentés dans la Fig 4.6 :

Les jets générés de τ sont identifiés dans ce cas idéal par la méthode basée sur les coupures avec une efficacité de 80% à partir de 30 GeV et 90% à partir de 50 GeV. Nous avons donc une limite physique qui est confirmée en observant l'écartement entre les constituants du jet et la trace principale dans les jets de τ et les jets de quarks (Fig 4.7).

La courbe en noir représente la limite du cône d'isolation et celle en rouge représente celle

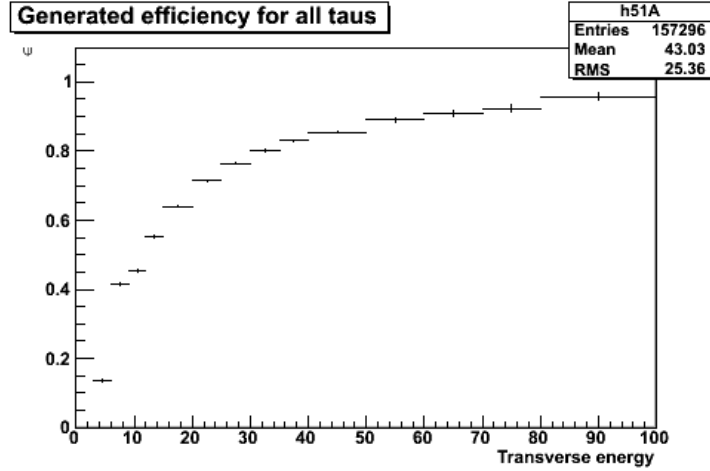


FIGURE 4.6 – Efficacité d'identification idéale des τ générés dans un échantillon $Z \rightarrow \tau\tau$.

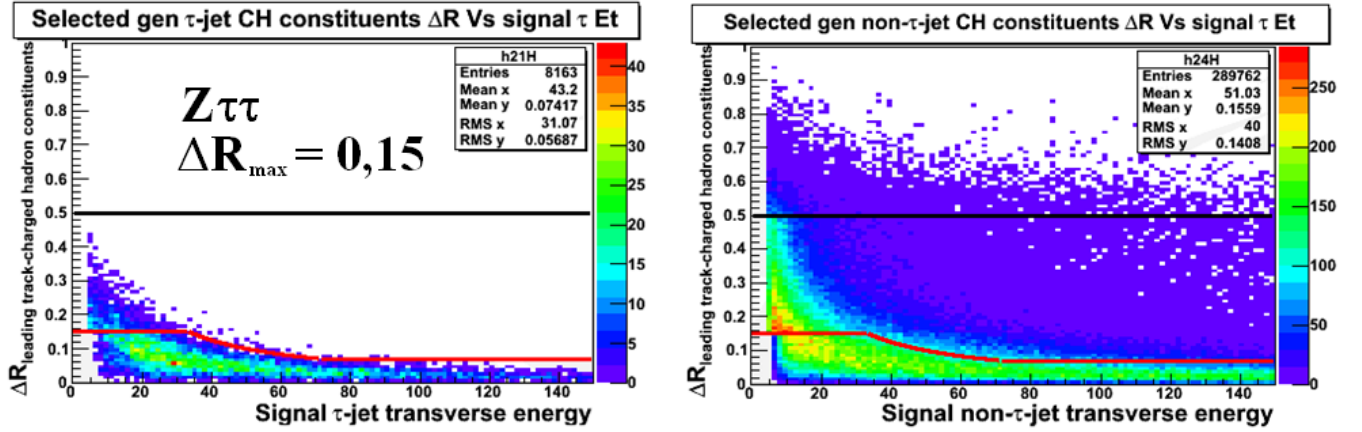


FIGURE 4.7 – Distance angulaire entre les différents constituants du jet et la trace principale dans des événements $Z \rightarrow \tau\tau$ pour a) Les jets de τ b) Les jets QCD (cône signal de taille maximale $\Delta R = 0,15$). Une coupure en $P_T > 1 \text{ GeV}$ est imposée sur les constituants.

du cône signal, ces cônes suivent les conditions définies dans la section 3.3.1. Le choix de la coupure imposant la taille maximale du cône signal à $\Delta R_{max} = 0,15$ est motivé par les tendances observées dans les figures 4.7 ($\Delta R_{max} = 0,15$) et 4.8 ($\Delta R_{max} = 0,30$). Ici le paramètre étudié est la somme des impulsions transverses dans le cône signal de taille maximale :

$$E_{T,signal} = \sum_i^N E_{T,i} \quad (4.7)$$

Où $E_{T,i}$ est l'énergie transverse du $i^{ème}$ constituant du jet situé à une distance angulaire $\Delta R < \Delta R_{max}$ de la trace de plus haute impulsion (trace principale). N est le nombre de constituants du jet respectant la condition de distance angulaire $\Delta R < \Delta R_{max}$.

Nous observons alors deux phénomènes lorsque la valeur de ΔR_{max} change :

- La valeur de l'impulsion transverse dans le cône signal change elle aussi : un cône signal plus large acceptera plus de particules et donc on aura un déplacement vers les plus haut $E_{T,signal}$ (somme des énergies transverses dans le cône signal) des candidats τ . Ce qui explique aussi les discontinuités observées dans les figures 4.7 et 4.8 au niveau de la limite du cône signal.
- Le terme E_T dans la formule calculant la taille évolutif changeant pour la raison citée précédemment, des candidats qui passaient la première sélection avec ($\Delta R_{max} = 0,15$) ne passent pas nécessairement la deuxième avec ($\Delta R_{max} = 0,30$), causant l'échec de l'identification observé dans la fig. 4.8 pour certains candidats de $P_{T,sig}$ compris entre 20 et 60 GeV.

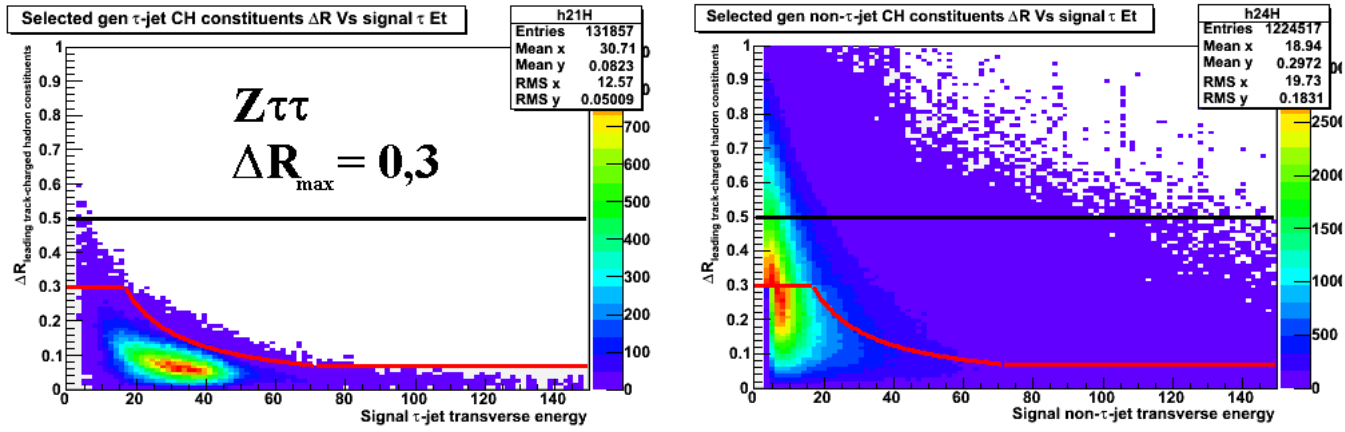


FIGURE 4.8 – Distance angulaire entre les différents constituants du jet et la trace principale dans des événements $Z \rightarrow \tau\tau$ pour a) Les jets de τ b) Les jets QCD (cône signal de taille maximale $\Delta R = 0,3$). Une coupure en $P_T > 1 \text{ GeV}$ est imposée sur les constituants.

La provenance du τ en question n'a pas d'influence sur la distribution angulaire des produit de désintégration, dans la Fig 4.9 on peut voir une étude similaire menée (avec un cône signal de taille maximale $\Delta R = 0,15$) sur des événements SUSY, les populations sont identiques à celle observées dans les événements $Z\tau\tau$.

Les méthodes HPS et TaNC sont elles aussi basées sur des cônes, il convient alors de ne pas définir celui-ci trop large.

4.2.2.2 Dépendance en η

Nous étudions ici si la pseudo-rapacité a un impact sur l'identification du lepton τ . La taille d'un jet ne dépend pas de sa direction cependant on constate une chute de l'efficacité "idéale" aux hautes pseudo-rapidités ($|\eta| > 2,5$) comme le montre la figure 4.10.

Cette baisse d'efficacité est due à la sensibilité de ΔR aux hautes valeurs de η , les $\Delta\eta$ deviennent plus importants dans ces régions et donc les ΔR aussi, les jets sont donc plus

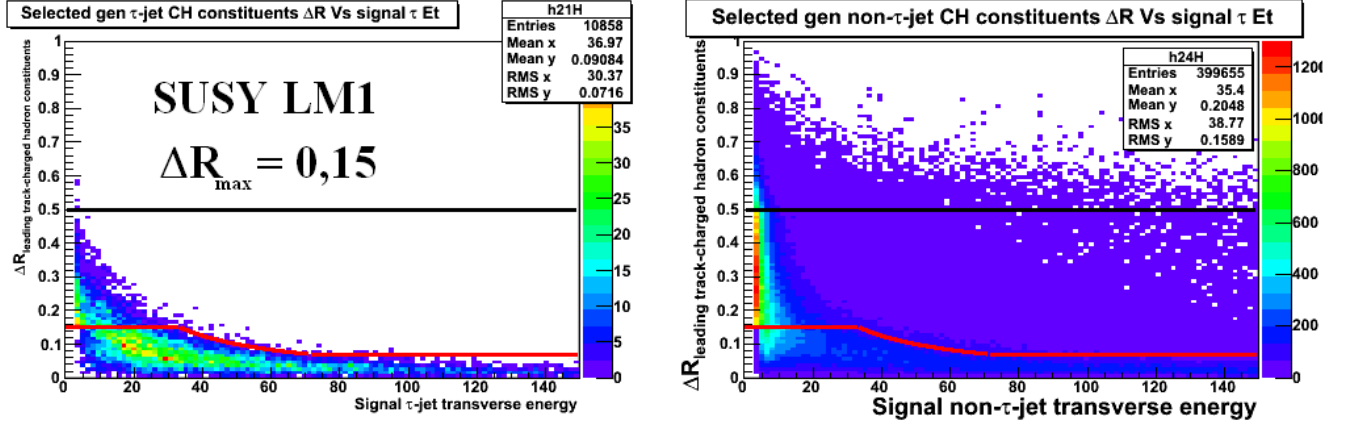


FIGURE 4.9 – Distance angulaire entre les différents constituants du jet et la trace principale dans a) Les jets de τ b) Les jets de quarks. Dans des événements SUSY LM1 (cône signal de taille maximale $\Delta R = 0,15$). Une coupure en $P_T > 1 \text{ GeV}$ est imposée sur les constituants.

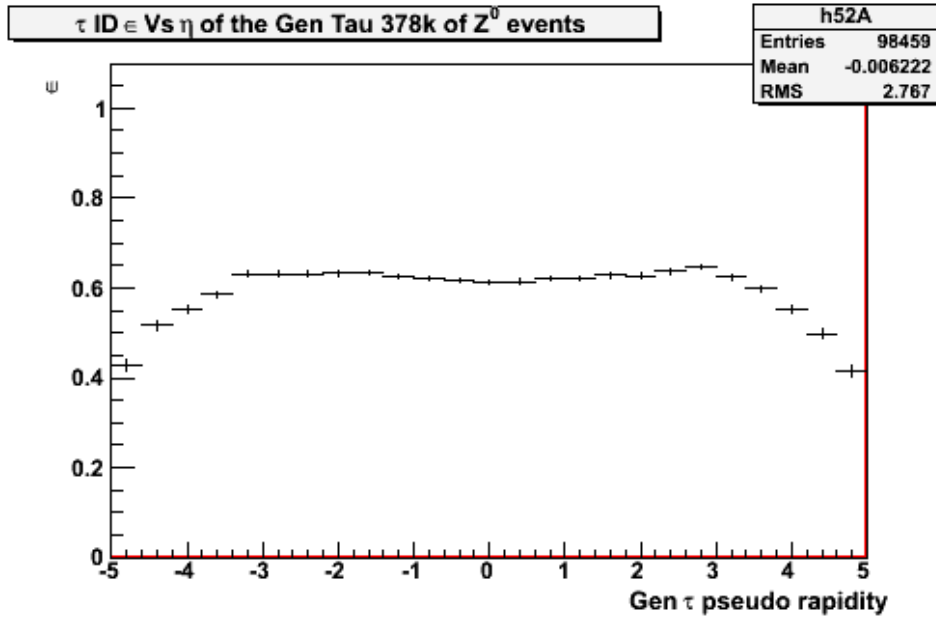


FIGURE 4.10 – Efficacité idéale par la méthode basée sur les coupures dans un échantillon $Z \rightarrow \tau\tau$.

larges et ne passent plus les tests d'identification.

Le détecteur CMS couvre efficacement la région en rapidité $|\eta| < 2,5$, il n'y a donc pas de raison physique que la pseudo-rapidité ait un effet important sur l'identification du lepton τ si on se limite à cette région en pseudo-rapidité, seuls des effets expérimentaux pourront influencer les niveaux d'efficacité d'identification en fonction de η .

4.2.3 Efficacités calculées avec les quantités reconstruites

4.2.3.1 Dépendance en P_T

Nous recherchons quel impact a l'impulsion transverse sur le succès de l'identification du lepton τ . Les spectres en P_T des jets de τ dépendent du type d'échantillon comme le montre la figure 4.11. On peut alors attendre des résultats différents si on intègre sur P_T et que ce paramètre a un impact sur l'identification du lepton τ .

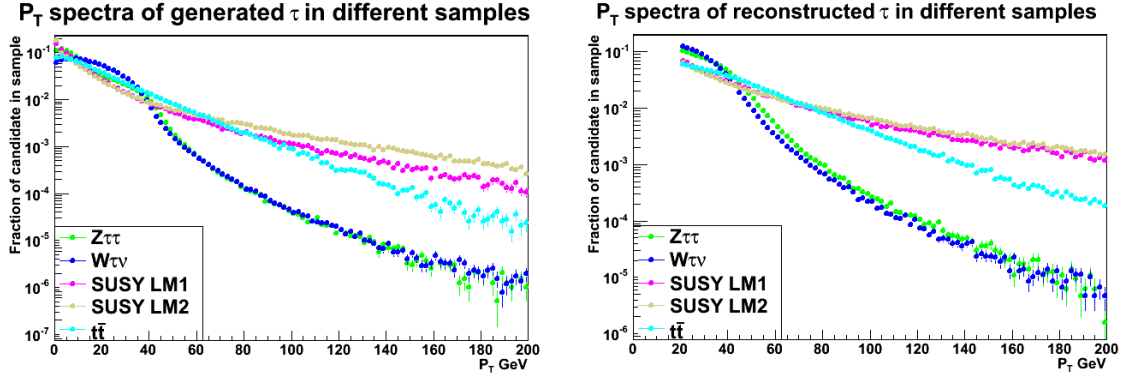


FIGURE 4.11 – Spectres en impulsion transverse des jets de τ dans divers échantillons contenant des τ : a) Quantités générées, b) Quantités reconstruites.

On détermine les efficacités avec la formule (4.2) sur les échantillons $Z\tau\tau$ et SUSY, les résultats présentés dans la Figure 4.12 nous montrent les résultats obtenus pour la méthode basée sur les coupures.

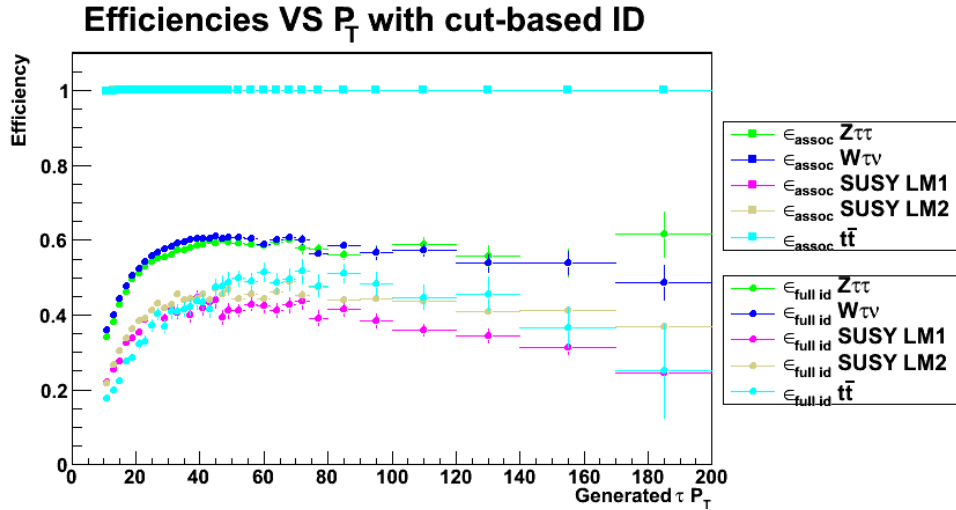


FIGURE 4.12 – Efficacités d'identification des candidats τ par la méthode basée sur les coupures, dans divers échantillons ($\Delta R_{association} < 0,5$).

Les efficacités “idéales” présentées dans la section précédente ne prennent en considération que les particules issues de la désintégration d'un τ , dans la réalité il faut aussi prendre

en compte la présence de hadrons et de photons supplémentaires du à la QCD et aux interactions multiples qui peuvent contaminer le cône signal comme le cône isolation, ce qui explique une efficacité maximale de 0,5 observée avec les quantités reconstruites alors qu'elle dépassait 0,9 dans le cas "idéal".

Les efficacités d'association sont excellentes (≈ 1) pour des candidats de plus de 20 GeV, y compris dans les événements SUSY et $t\bar{t}$ où la multiplicité globale de particules est beaucoup plus importante. La même étude menée avec une condition plus restrictive sur l'association ($\Delta R < 0.15$) entre candidat reconstruit et τ généré nous montre une baisse de cette efficacité d'association dans les événements SUSY. Ce comportement est attendu car nous avons vu dans la figure 4.4 que quasiment tous les τ reconstruits se situent à une distance $\Delta R < 0,15$ du τ généré alors qu'une fraction non négligeable de τ reconstruit sont au-delà de cette limite dans les événements SUSY et $t\bar{t}$. Cependant les efficacités après identification restent similaires, l'algorithme d'identification n'est pas sensible au choix de $\Delta R_{association}$. Afin d'accepter un maximum de candidats τ on fixe la valeur maximale du paramètre $\Delta R_{association} = 0,5$ pour la suite de l'étude.

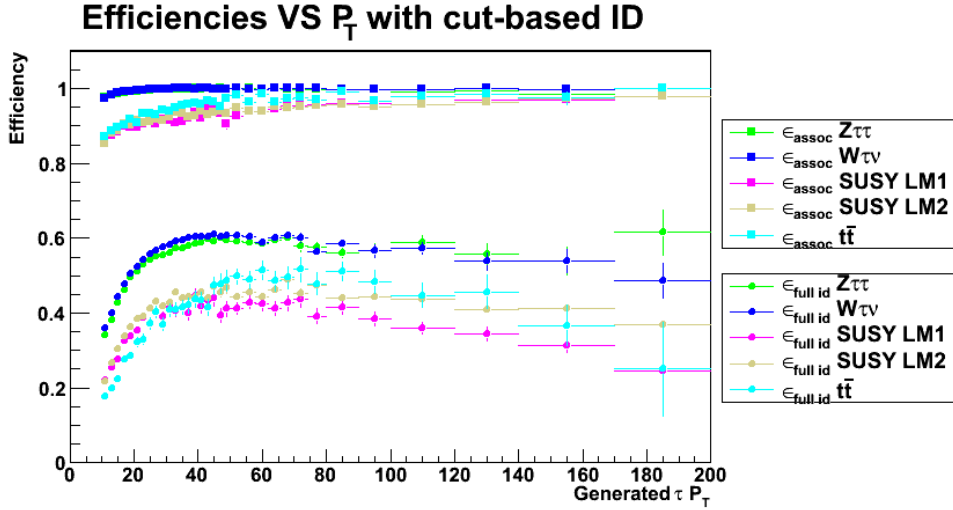


FIGURE 4.13 – Efficacités d'identification des candidats τ , par la méthode basée sur les coupures avec association restreinte ($\Delta R_{assoc} < 0,15$), dans divers échantillons.

Nous constatons donc un effet d'environnement sur l'identification du lepton τ dans les figures 4.12 et 4.13 : léger sur les efficacités d'association (visible seulement avec ΔR restreint), et plus important sur les efficacités d'identification, cet effet est indépendant de l'algorithme d'identification utilisé comme le montrent les résultats obtenus avec les méthodes HPS et TaNC (figures 4.14 et 4.15).

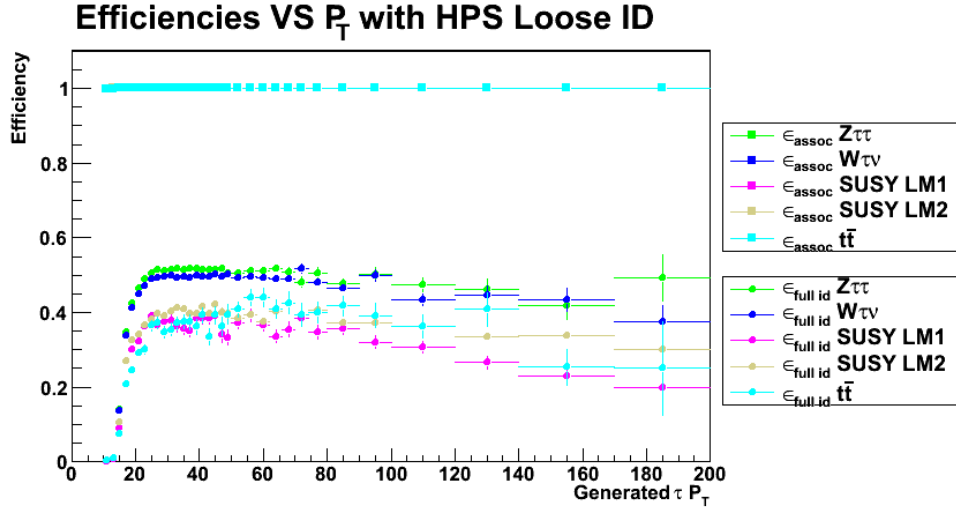


FIGURE 4.14 – Efficacités d'identification des candidats τ par la méthode HPS (Loose isolation), dans divers échantillons ($\Delta R_{association} < 0,5$).

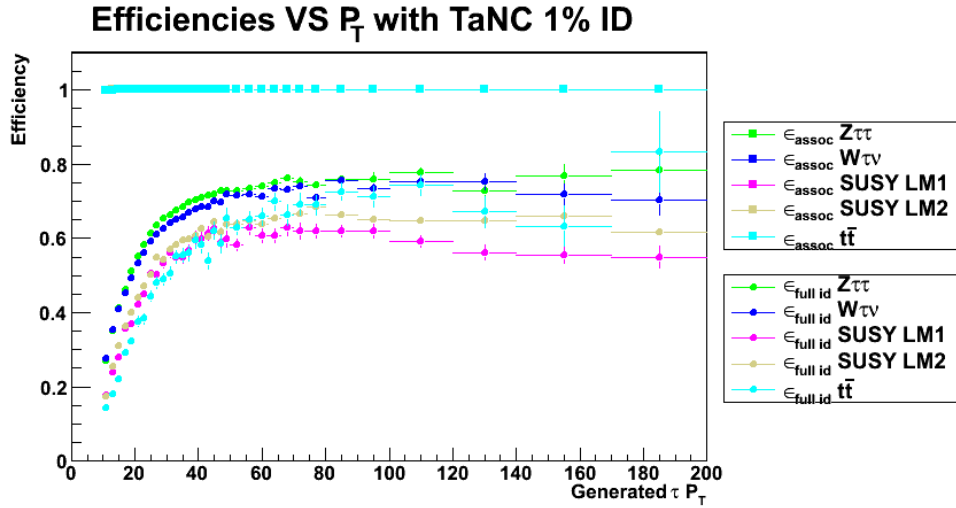


FIGURE 4.15 – Efficacités d'identification des candidats τ par la méthode TaNC (Scenario à 1% de faux), dans divers échantillons ($\Delta R_{association} < 0,5$).

4.2.4 Dépendance en η

Nous avons vu qu'idéalement l'efficacité d'identification est constante dans une zone en rapidité $|\eta| < 2,5$, mais on étudie ici la validité de cette hypothèse avec l'information reconstruite pour vérifier les effets expérimentaux. Les spectres en pseudo-rapidité sont présentés dans la figure 4.16 :

Pour les efficacités d'identification des candidats τ la couverture en η du détecteur CMS n'est pas totalement homogène à cause des transitions entre les bouchons et le tonneau comme nous l'avons vu dans la section 4.2 (Fig. 4.5c). La largeur du jet reconstruit est surévaluée dans cette région en pseudo-rapidité et ce phénomène a un impact sur les efficacités d'identification du lepton τ comme le montre la figure 4.17.

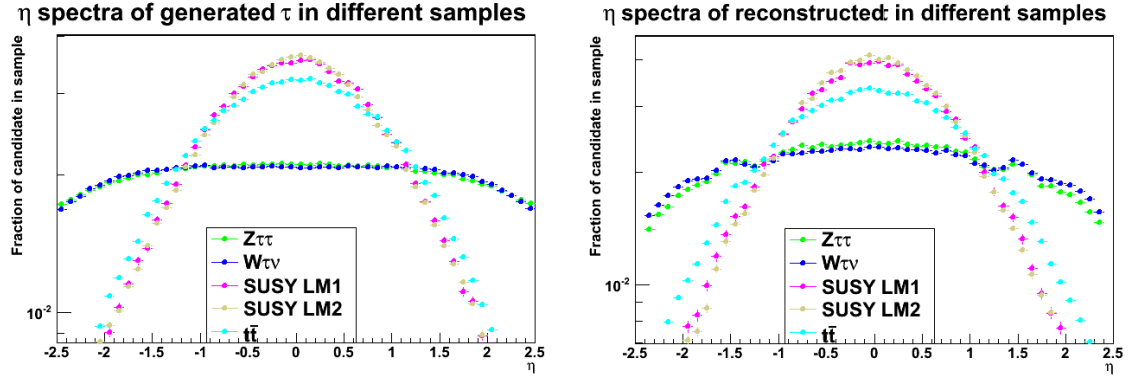


FIGURE 4.16 – Spectres en pseudo-rapidité des jets de τ dans divers échantillons contenant des τ : a) Information générée, b) Information reconstruite.

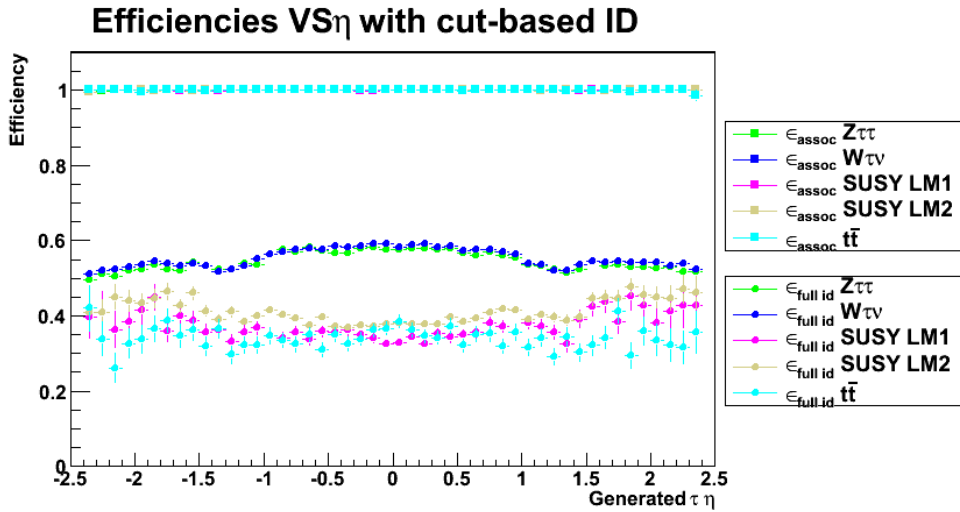


FIGURE 4.17 – Efficacités expérimentales par la méthode basée sur les coupures dans divers échantillons ($\Delta R_{association} < 0,5$).

On remarque aussi un effet expérimental moins important pour les événements SUSY, ceci pour la raison invoquée dans la section 4.2 : les particules desquelles sont issus les τ sont produites dans une région plus centrale pour les événements SUSY et la statistique concernée est moindre aux rapidités $|\eta| \approx 1,5$. La figure 4.18 nous montre que l'efficacité d'association chute dans la région centrale lorsqu'il s'agit d'événements actifs dans cette région (SUSY, $t\bar{t}$).

La tendance voulant que les τ présents dans les événements SUSY et $t\bar{t}$ soient moins bien identifiés que ceux des événements électrofaibles est elle aussi confirmée dans cet étude, particulièrement dans la région centrale ($|\eta| < 1$).

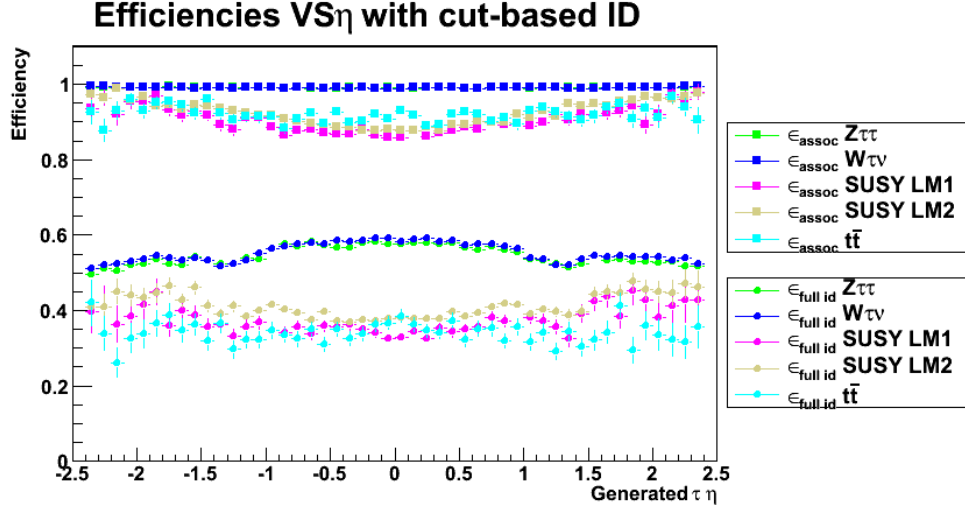


FIGURE 4.18 – Efficacités expérimentales par la méthode basée sur les coupures dans divers échantillons ($\Delta R_{association} < 0,15$).

4.2.5 Dépendance en R_{Jet}

Le paramètre R_{Jet} défini dans la section 3.4 est une définition de la largeur d'un jet hadronique. Les algorithmes d'identification du τ recherchent des jets étroits et/ou isolés et doivent donc être sensibles à ce paramètre. On attend alors une chute importante de l'efficacité d'identification et des taux de faux à mesure que R_{Jet} augmente.

Les efficacités d'identification du lepton τ en fonction de R_{jet} sont présentés dans les figures 4.19, 4.20 et 4.21 et nous confirment la tendance attendue : La figure 4.19 nous montre l'étendue de l'effet de R_{Jet} sur l'identification basée sur les coupures. L'efficacité d'identification chute en dessous de 10% pour $R_{Jet} \geq 0,2$ alors qu'elle est supérieure à 50% pour $R_{Jet} \leq 0,06$. Une dispersion entre les différents types d'échantillons est observé pour les jets larges ($R_{Jet} \geq 0,3$).

Les efficacités d'identification VS R_{Jet} par la méthode Tau Neural Classifier (Fig 4.21) nous montrent des tendances similaires à celles déterminées par la méthode basée sur les coupures. La disparité entre les différents échantillons pour les jets larges est aussi présente.

Les jets larges sont donc plus sensibles à l'environnement de l'évènement, ceci est le sujet d'études du chapitre 5.

4.3 Taux de mauvaise identification du τ

Dans cette section seuls les jets QCD sont considérés comme source de mauvaise identification. La réjection d'électrons et de muons se fait alors dans la présélection. L'échantillon de données du LHC étudié est l'ensemble des collisions de l'année 2010 avec sélection à l'aide des systèmes de déclenchement d'évènements contenant des jets. Les simulations Monte-Carlo utilisées pour cette même étude reproduisent les conditions de luminosité de cette même année. Pour les études portant sur les véritables données et les Monte-Carlo

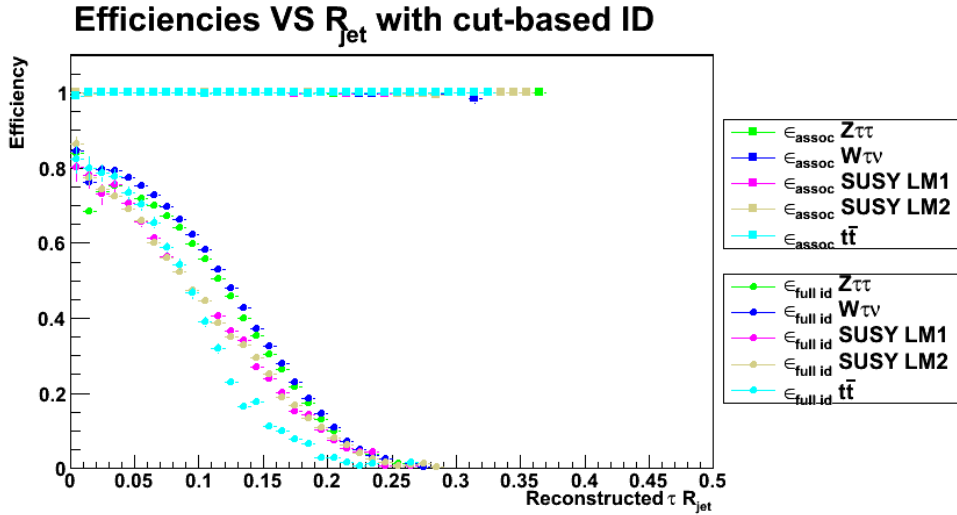


FIGURE 4.19 – Efficacité expérimentale VS R_{Jet} ($P_T > 20 \text{ GeV}$, $|\eta| < 2, 4$) par la méthode basée sur les coupures dans divers échantillons.

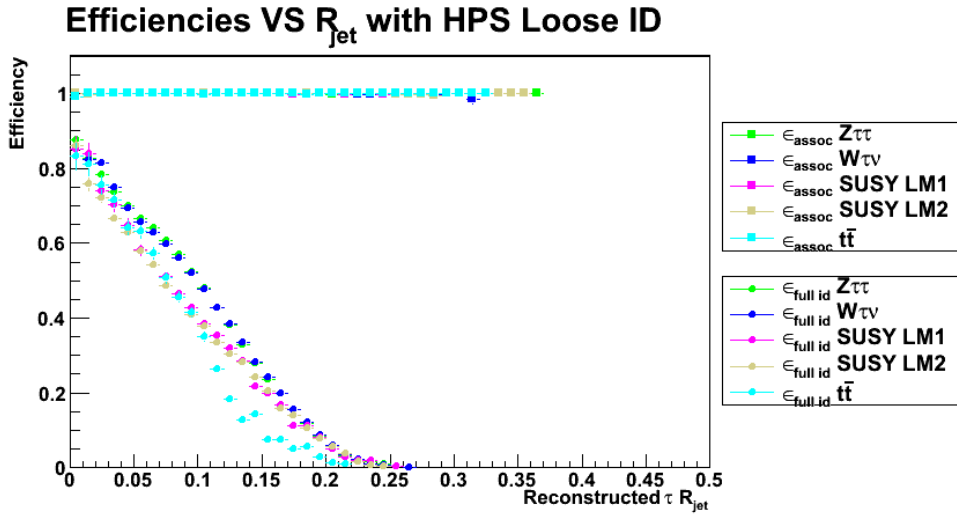


FIGURE 4.20 – Efficacité expérimentale VS R_{Jet} ($P_T > 20 \text{ GeV}$, $|\eta| < 2, 4$) par la méthode Hadron Plus Strips (isolation relâchée) dans divers échantillons.

QCD, seul les jets “sonde” sont pris en compte dans les calculs de taux de mauvaise identification (voir section 4.1.1.2 : méthode de “tag and probe”). Dans le cas d’un échantillon Monte-Carlo contenant des τ , les faux candidats τ i.e. les jets QCD soumis à l’identification d’un lepton τ sont présélectionnés par non-association de ce jet avec le tau le plus proche ($\Delta R_{\tau, cand} > 0.5$).

Pour l’étude des taux de mauvaise identification, les trois algorithmes : basé sur les coupures, Hadron Plus Strips et Tau Neural Classifier sont employés. Les taux de faux sont calculés pour les étapes 2,3,4 et 5 présentées dans la section 3.3.1 pour l’algorithme cut-based, pour les trois niveaux d’isolation dans la méthode HPS et avec les différents scénarii de taux de faux (0,1%, 0,25%, 0,5% et 1%) pour la méthode TaNC.

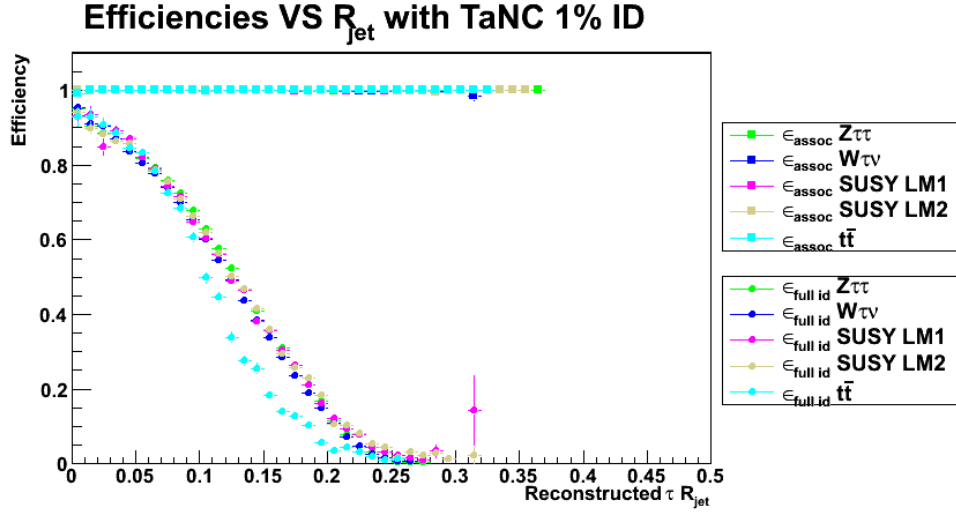


FIGURE 4.21 – Efficacité expérimentale VS R_{Jet} ($P_T > 20 \text{ GeV}$, $|\eta| < 2, 4$) par la méthode Tau Neural Classifier (scenario à 1% de mauvaise identification) dans divers échantillons.

4.3.1 Spectres des variables associées aux jets QCD

Les variables étudiées pour les jets QCD sont les mêmes que celle étudiées pour les jets de τ i.e. P_T , η et R_{Jet} . Ceci dans le but d’estimer le nombre de jets QCD passant les critères d’identification des algorithmes d’analyse de τ pour les mêmes paramètres.

Les corrélations (qualitativement) entre P_T , η et R_{Jet} pour les jets QCD sont présentées pour un seuil de déclenchement sur les jets “tag” de 30 GeV dans la figure 4.22, et de 100 GeV dans la figure 4.23.

- **Corrélations entre P_T et η pour les jets QCD** : Nous constatons pour les jets QCD des corrélations similaires entre les variables cinématiques à celles observées pour les jets de τ (Fig. 4.5). Pour un seuil de déclenchement de 30 GeV (Fig. 4.22a) nous avons une distribution relativement uniforme en rapidité de ces jets tandis que pour un seuil de déclenchement de 100 GeV (Fig. 4.23) nous observons des tendances approchant à celles observées pour les jets de τ issues de la désintégration de particules lourdes (Fig. 4.24a) : un seuil de déclenchement plus élevé les jets “tag” sélectionne des événements QCD avec des $\hat{P}_T$ différents (Fig. 4.25) et donc des spectres de jets “probe” différents.

- **Corrélations entre P_T et R_{jet} pour les jets QCD** : La largeur d’un jet QCD diminue à mesure que son impulsion transverse augmente comme nous le constatons dans les figures 4.22b) et 4.23b), l’explication de cet effet physique réside dans le fait que la largeur d’un jet diminue à mesure que P_T augmente (c.f. section 3.1.2, la relation 3.6 est étendue à plus de deux particules). La largeur moyenne, comparable à celle observée pour les jets QCD des événements $t\bar{t}$ et SUSY (Fig. 4.24b), est cependant plus importante pour les jets QCD que pour les jets hadroniques issus de la désintégration d’un lepton τ .

- **Corrélations entre $|\eta|$ et R_{jet} pour les jets QCD** : Nous constatons dans les figures 4.22c), 4.23c) que les variables $|\eta|$ et R_{jet} sont indépendantes, un phénomène déjà constaté pour les jets hadroniques de τ . Le même effet expérimental constaté dans la figure 4.5c) est

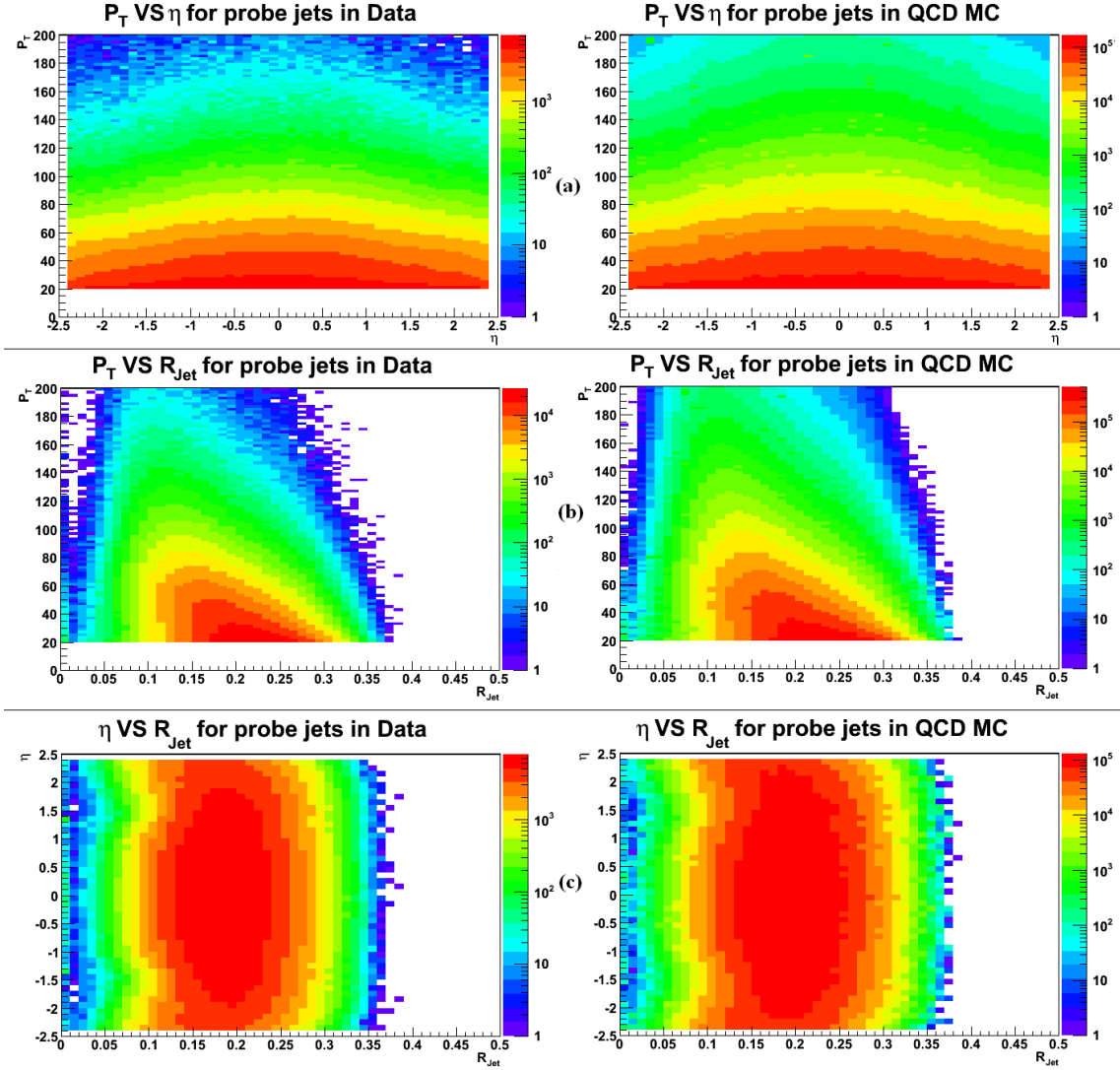


FIGURE 4.22 – Corrélations pour les jets QCD dans les collisions 2010 et dans un échantillon QCD (Seuil de déclenchement 30 GeV) entre les variables a) P_T vs $|\eta|$, b) P_T vs R_{Jet} , c) η vs R_{Jet} . Une coupure en $P_T > 20$ GeV est imposée sur les candidats.

présent ici aussi pour tous les échantillons (mais moins visible dans les événements SUSY), les raisons en sont les mêmes : la transition entre le tonneau et les bouchons entraîne des pertes de qualité dans la reconstruction des jets.

4.3.2 Dépendance en P_T

Les jets QCD ont une largeur qui décroît à mesure que leur P_T augmente, mais leur largeur reste cependant plus importante que celle des jets hadroniques de τ . On peut donc attendre un impact de ce paramètre sur le taux de mauvaise identification.

Les distributions en P_T des jets QCD pour différents échantillons sont présentées dans la figure 4.26). La fraction de jets QCD passant l'identification du lepton τ est calculée par la relation 4.6 pour la méthode basée sur les coupures

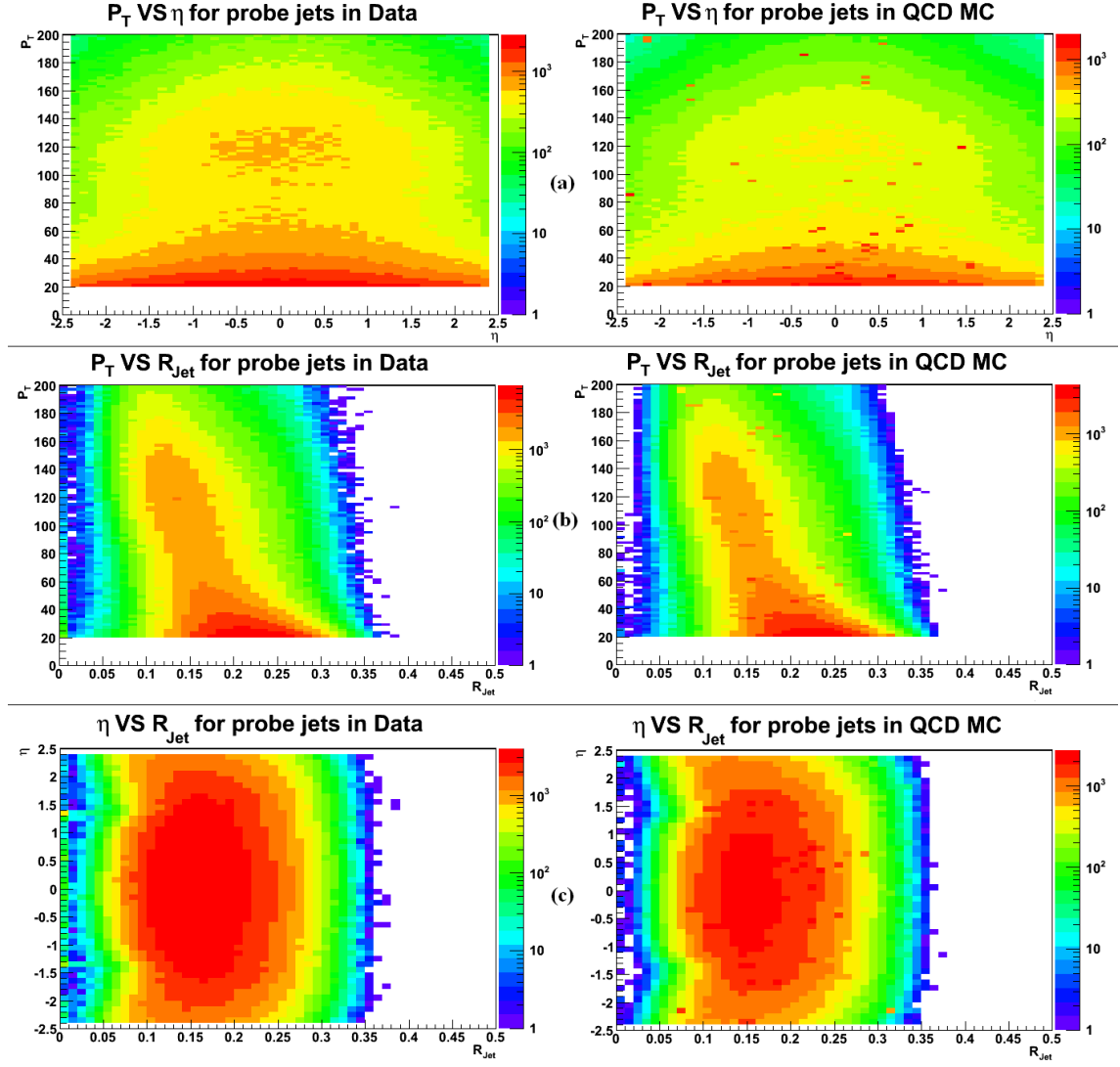


FIGURE 4.23 – Corrélations pour les jets QCD dans les collisions 2010 et dans un échantillon QCD (Seuil de déclenchement 100 GeV) entre les variables a) P_T vs $|\eta|$, b) P_T vs R_{Jet} , c) η vs R_{Jet} . Une coupure en $P_T > 20 \text{ GeV}$ est imposée sur les candidats.

dans la figure 4.27. Les différentes étapes de l'algorithme sont visibles dans cette figure.

La figure 4.27 nous montre un bon accord données Monte-Carlo pour des candidats de plus de 30 GeV. On constate une divergence entre MC et données aux alentours de 100 GeV. Cette divergence est due à la renormalisation d'un candidat passant les sélections dans l'échantillon QCD MC $\hat{P}_T = 15 \text{ GeV}$ (pondéré par un facteur ≈ 600). Le taux de mauvaise identification chute (d'un facteur ≈ 3) entre les jets de 20-30 GeV et ceux de plus de 80 GeV (méthode cut-based). Les distributions présentées dans la figure 4.23b nous montrent des jets plus fins à haut P_T qu'à bas P_T , or des jets globalement plus fins devraient passer plus facilement le test d'identification du lepton τ . On remarque que cette chute concerne les jets QCD après l'étape finale : l'isolation des jets.

L'explication de ce phénomène réside dans le fait que le cône signal évolutif a une largeur maximale de $\Delta R = 0,07$: ce seuil est atteint pour les jets de plus de 70 GeV. Nous avons

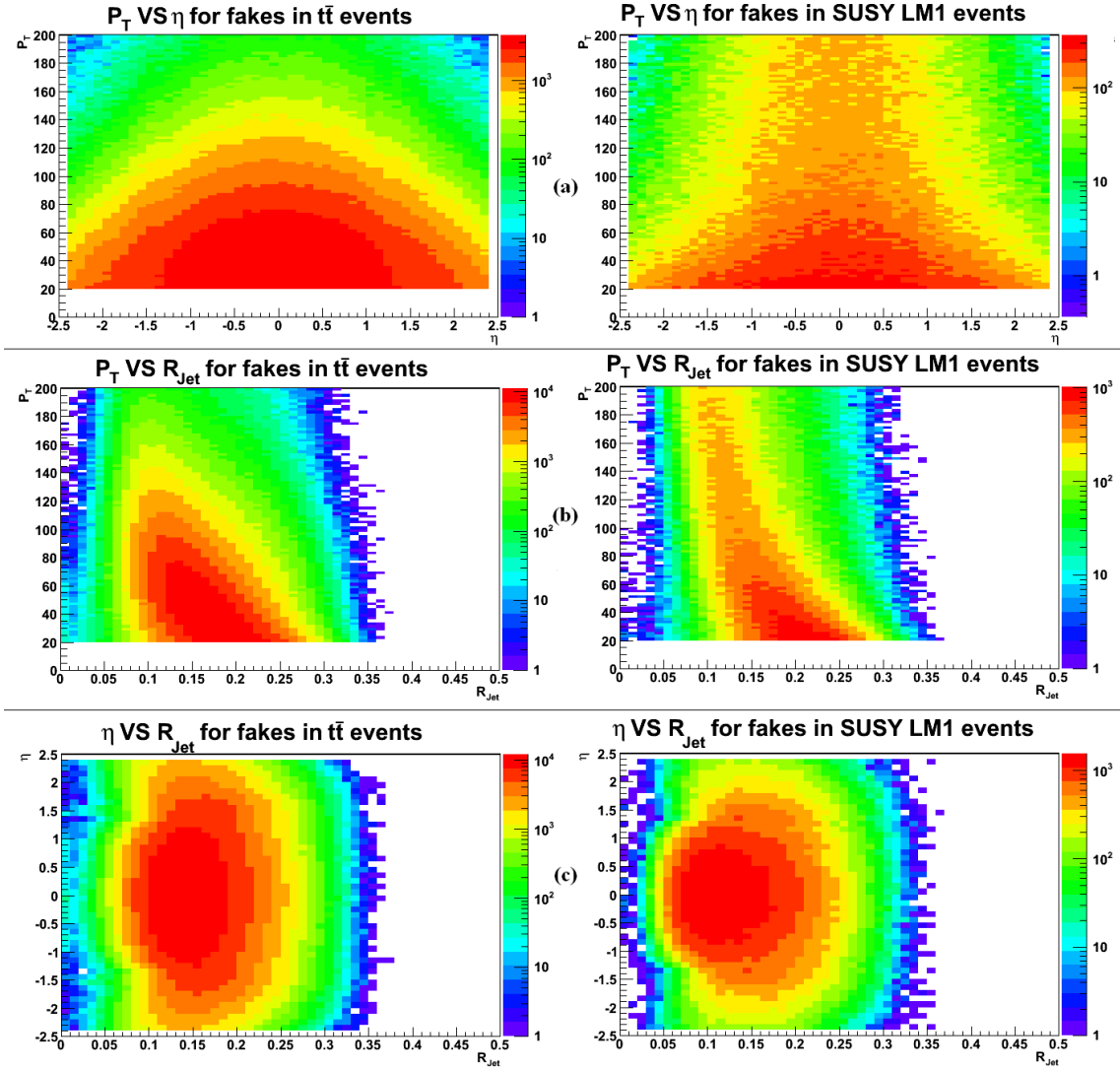


FIGURE 4.24 – Corrélations pour les candidats τ non-associés à un tau hadronique dans des échantillons $t\bar{t}$ (à gauche) et SUSY LM1 (à droite) entre les variables a) P_T vs $|\eta|$, b) P_T vs R_{Jet} , c) η vs R_{Jet} . Une coupure en $P_T > 20 \text{ GeV}$ est imposée sur les candidats.

vu dans la figure 4.23b que la majorité des jets QCD ont une largeur R_{jet} dépassant 0,1 pour des candidats de plus de 70 GeV. Les jets QCD sont généralement mal isolés (c.f. section 1.3.1.2) ce qui signifie que des particules nombreuses et/ou d'impulsion transverse importante se situent hors du cône signal, causant le rejet de ces candidats.

Pour les candidats de basse impulsion transverse ($P_T < 35 \text{ GeV}$), le cône signal à une taille de $\Delta R = 0,15$. Nous avons vu dans la figure 4.23b qu'une fraction importante de jets QCD ont une largeur $R_{Jet} < 0,15$. Cette fraction de jets est susceptible de satisfaire aux conditions d'identification de l'algorithme basé sur les coupures, ce qui entraîne un surcroît du nombre de faux dans cette région en P_T .

On observe pour le taux de faux avec la méthode HPS une tendance similaire : le taux de faux chute à mesure que le P_T des jets augmente (Figure 4.28).

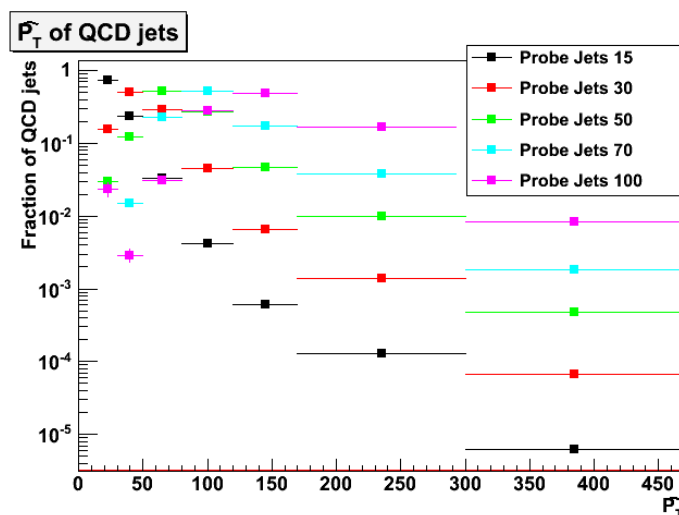


FIGURE 4.25 – $\hat{P}_T$ des évènements sélectionnés en fonction des différents seuils de déclenchement pour jet.

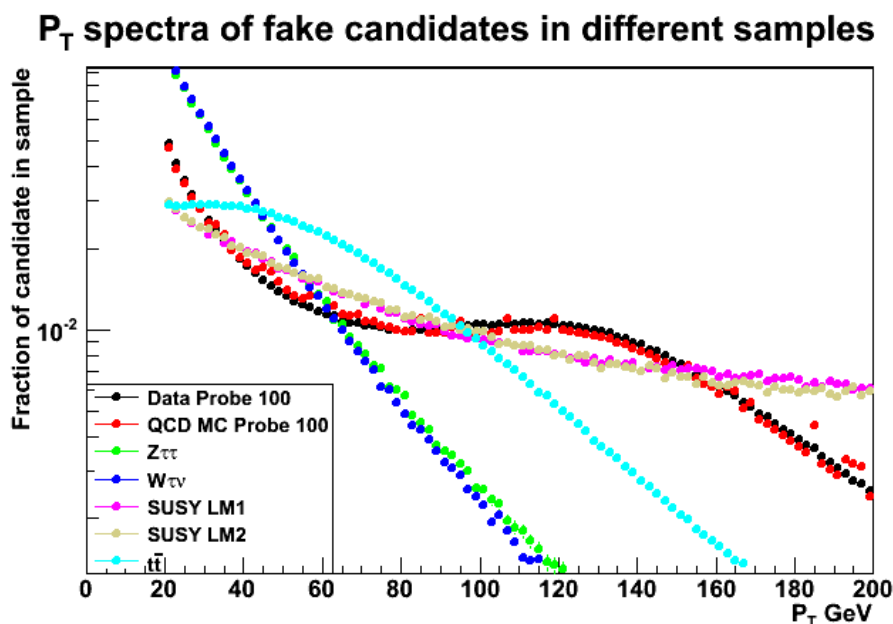


FIGURE 4.26 – Spectres en P_T des jets QCD dans divers échantillons : Sélection par déclenchement des jets avec un seuil de 100 GeV (QCD et Données), et rejections de muons et d'électrons appliquées à ces échantillons.

La méthode HPS est basée sur l'identification d'un des 5 modes de désintégration hadronique dominants du τ . Il est sensible à l'impulsion transverse des jets QCD car les hadrons seront plus collimés et leur multiplicité sera plus importante dans une région étroite. La présence de retrouver l'un des 5 modes identifiés par la méthode HPS devient alors moins probable. A bas P_T les hadrons présents dans le cône défini par la méthode HPS seront de plus basse impulsion transverse, ils sont alors moins susceptibles de causer la réjection du candidat et on observe ici aussi un surcroît de la fraction des jets QCD identifiés comme τ .

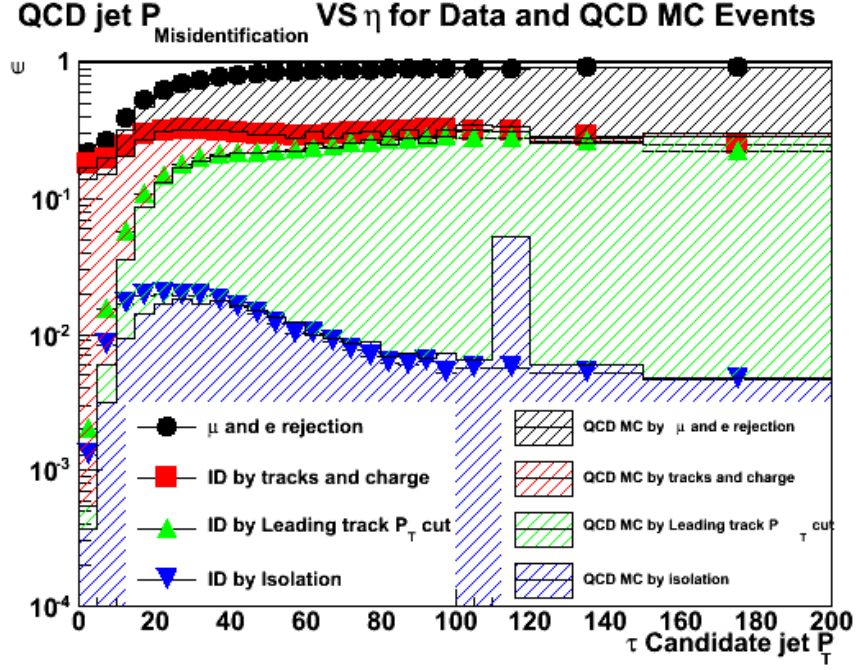


FIGURE 4.27 – Taux de mauvaise identification VS P_T avec information reconstruite pour un déclencheur de haut niveau jet de 100 GeV avec la méthode basée sur les coupures.

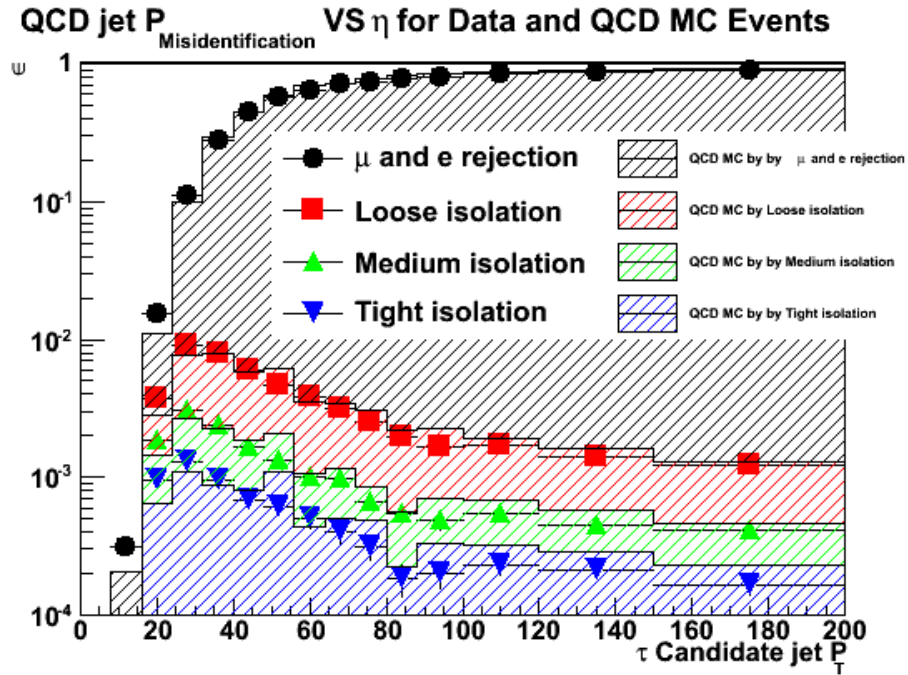


FIGURE 4.28 – Taux de mauvaise identification VS P_T avec information reconstruite pour un déclencheur de haut niveau jet de 100 GeV avec la méthode Hadron Plus Strips.

L'algorithme TaNC présente les mêmes tendances que les deux autres algorithmes à bas P_T pour la probabilité de mauvaise identification. On voit dans la figure 4.29 une augmen-

tation anormale de cette probabilité à partir de 100 GeV. Les résultats de la collaboration sont présentés pour les algorithmes HPS et TaNC (isolation relâchée) dans la figure 4.30. Les résultats obtenus pour HPS sont comparables à ceux présentés dans la figure 4.28 : un niveau de mauvaise identification de $\approx 10^{-2}$ pour les jets de moins de 50 GeV, et un niveau de mauvaise identification de $\approx 10^{-3}$ pour les jets de plus de 100 GeV. Les résultats obtenus pour l'algorithme TaNC sont légèrement différents : on constate dans la figure 4.30 une augmentation anormale de mauvaise identification des jets de plus de 100 GeV uniquement pour les événements Monte-Carlo (*QCDj simulation*). Le scénario d'identification choisi pour l'algorithme TaNC est "TaNC Loose" : les coupures appliquées sur les variables de sortie des réseaux de neurones sont ajustées de façon à obtenir un taux de faux de 1% qui est divisé par deux à chaque niveau successif d'isolation[60] (0,5% pour l'isolation moyenne et 0,25% pour l'isolation resserrée).

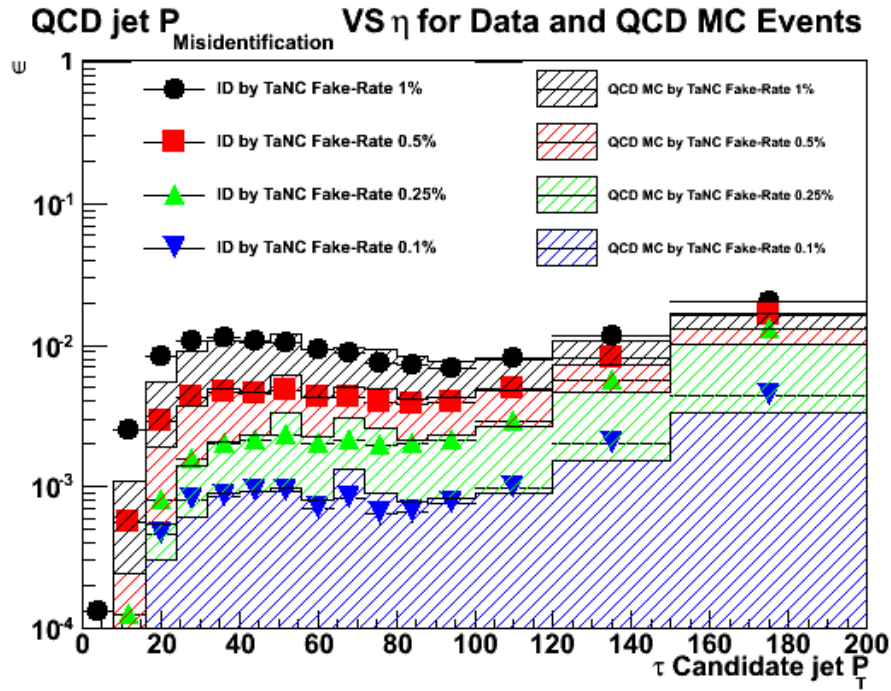


FIGURE 4.29 – Taux de mauvaise identification VS P_T avec information reconstruite pour un déclencheur de haut niveau jet de 100 GeV avec la méthode Tau Neural Classifier.

4.3.3 Dépendance en η

La distribution en η des jets QCD est montrée pour les différents échantillons dans la figure 4.31, on y observe les deux régimes différents observés de la Fig. 4.16 : les événements QCD et électrofaibles sont plus étendus en pseudo-rapidité que les événements $t\bar{t}$ et SUSY.

Les taux de faux en fonction de la pseudo-rapidité sont présentés pour l'algorithme basé sur les coupures dans la figure 4.32, pour la méthode HPS dans la figure 4.34 et pour la méthode TaNC dans la figure 4.35.

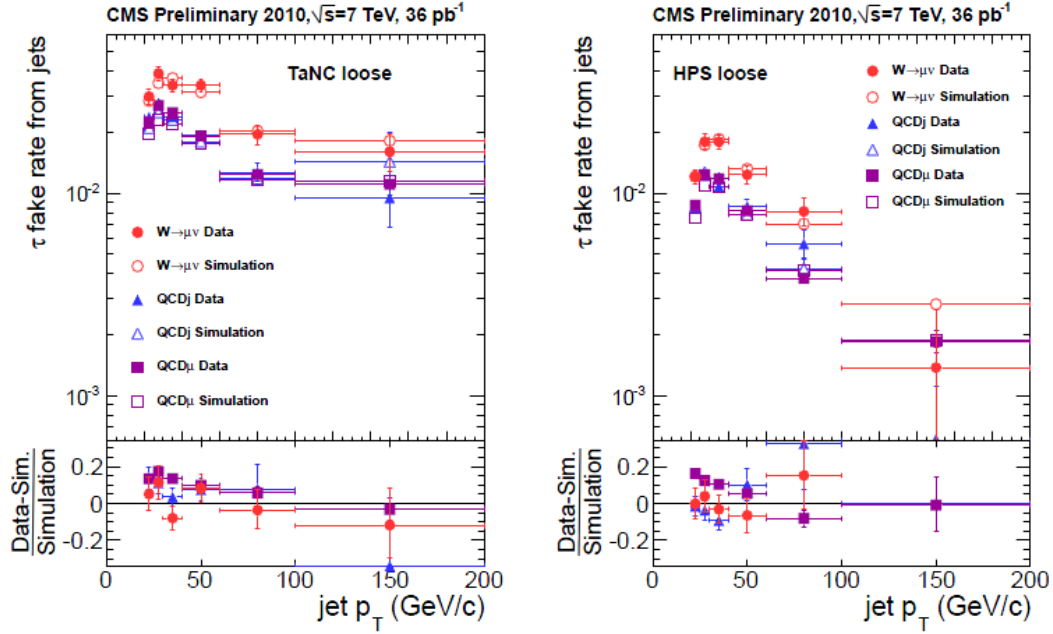


FIGURE 4.30 – Taux de mauvaise identification obtenus par la collaboration VS P_T pour les algorithmes HPS (isolation relâchée) et TaNC (isolation relâchée). Provenant de [59].

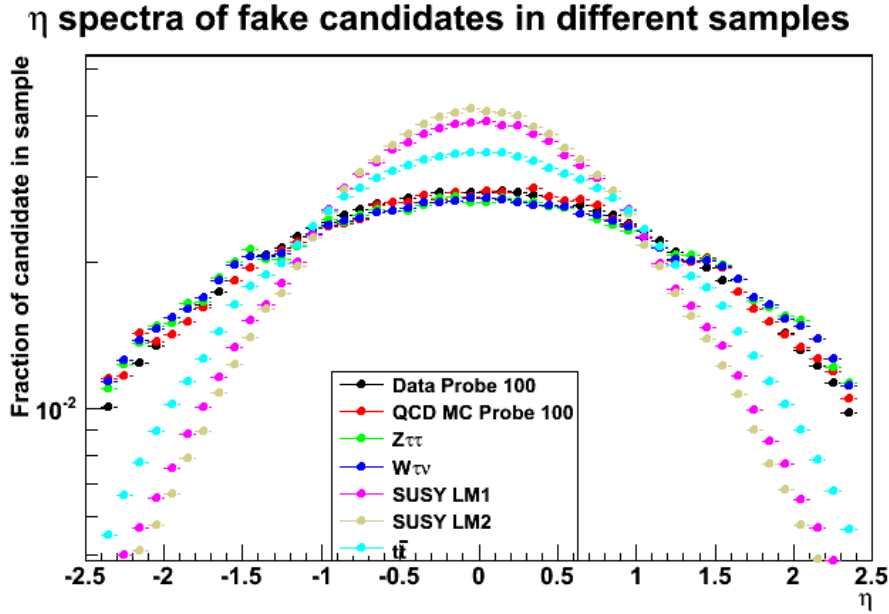


FIGURE 4.31 – Spectres en η des jets QCD dans divers échantillons : Sélection par déclenchement des jets avec un seuil de 100 GeV (QCD et Données), et réjections de muons et d'électrons appliquées à ces échantillons.

Les deux algorithmes cut-based et HPS montrent une tendance similaire : le taux de mauvaise identification augmente à mesure que $|\eta|$ augmente. Il s'agit ici d'un effet d'environnement : nous avons vu dans la figure 4.23a) que les jets sont de P_T plus élevés à mesure que η diminue pour les événements de haut $\hat{P}_T$ (dont la sélection est favorisée

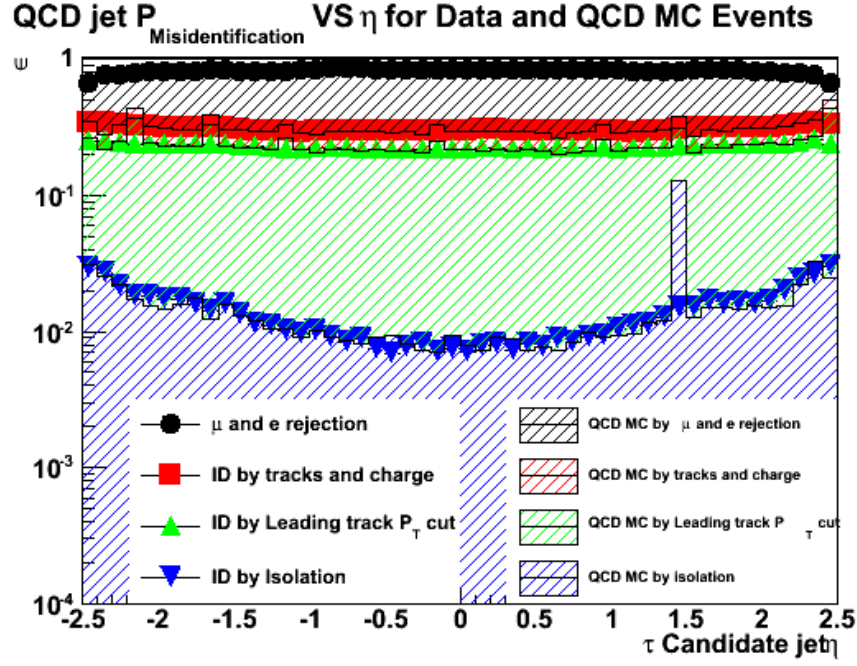


FIGURE 4.32 – Taux de mauvaise identification VS η avec information reconstruite pour un déclencheur de haut niveau jet de 100 GeV avec la méthode basée sur les coupures.

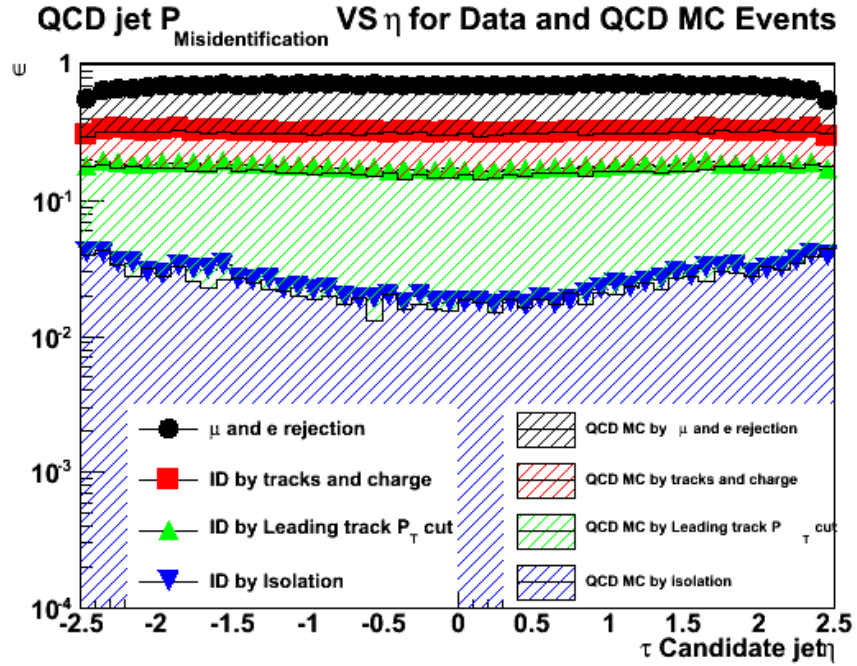


FIGURE 4.33 – Taux de mauvaise identification VS η avec information reconstruite pour un déclencheur de haut niveau jet de 15 GeV avec la méthode basée sur les coupures.

par des seuils de déclenchement élevés), on retrouve alors la tendance observée dans la sous-section précédente : les jets de plus haut P_T ont un taux de mauvaise identification plus bas. Un seuil de déclenchement plus bas nous montre une distribution plus plate de

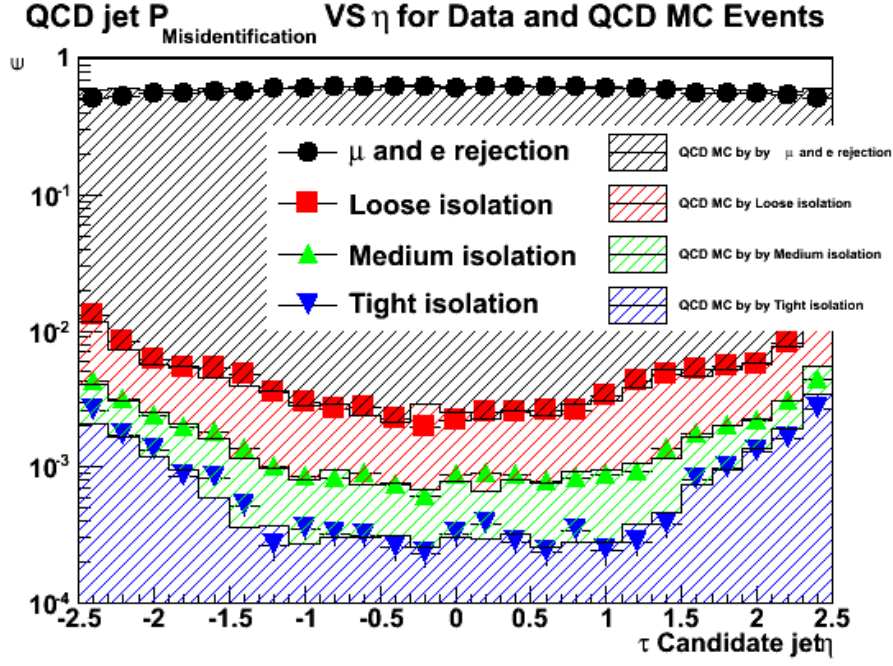


FIGURE 4.34 – Taux de mauvaise identification VS η avec information reconstruite pour un déclencheur de haut niveau jet de 100 GeV avec la méthode Hadron Plus Strips.

ce taux de mauvaise identification (Fig. 4.33).

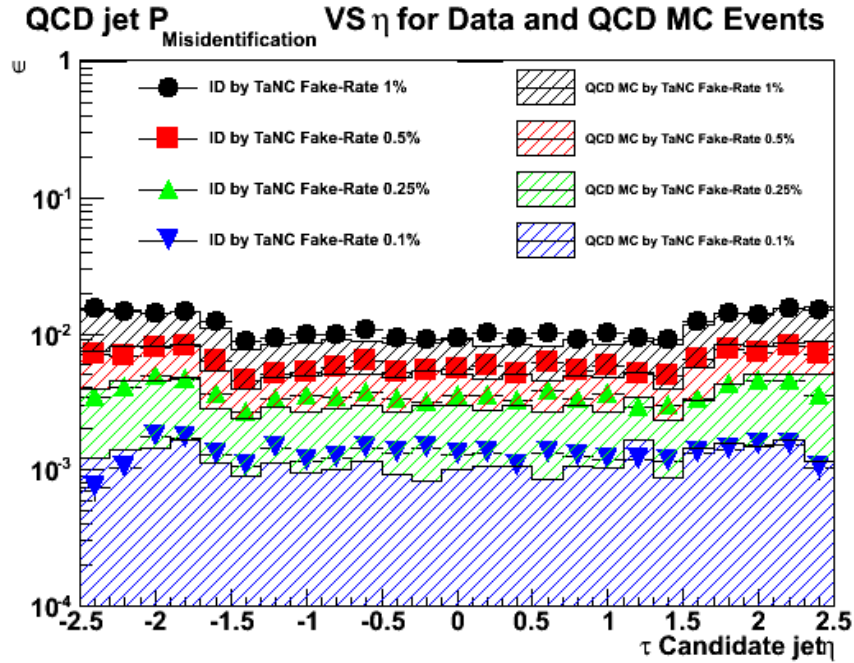


FIGURE 4.35 – Taux de mauvaise identification VS η avec information reconstruite pour un déclencheur de haut niveau jet de 100 GeV avec la méthode Tau Neural Classifier.

4.3.4 Dépendance en R_{Jet}

Le paramètre R_{Jet} (relation 3.13 dans le chapitre 3) décrit la largeur du jet et a un impact direct sur le succès de l'identification du τ assumant que les algorithmes d'identification recherchent des jets étroits et isolés. Les taux de faux doivent donc présenter les mêmes tendances que les efficacités d'identification en fonction de R_{Jet} : une chute drastique du taux de faux avec l'augmentation de R_{Jet} .

On veut d'abord déterminer les différences qui peuvent exister entre les divers échantillons concernant la largeur des jets QCD, ces distributions sont présentées dans la figure 4.36.

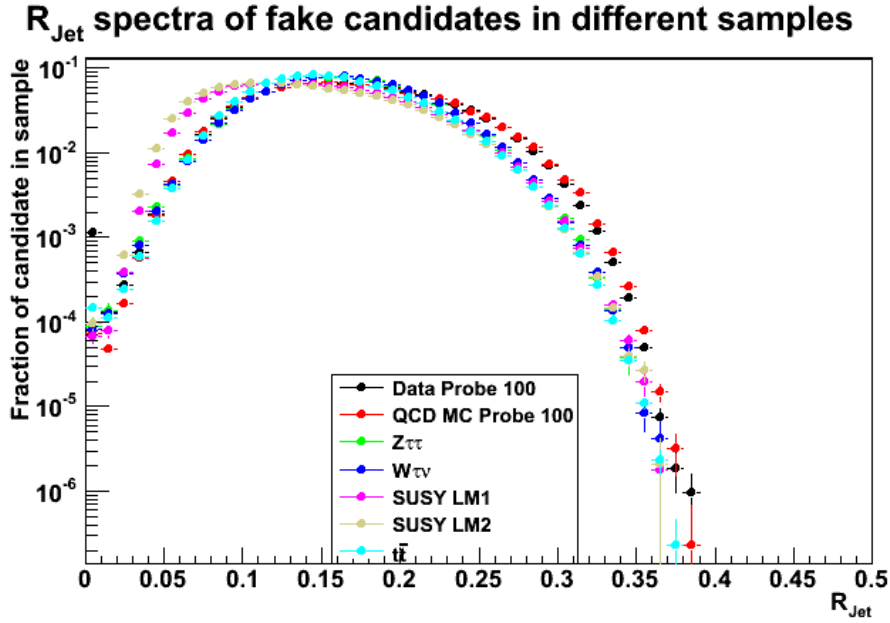


FIGURE 4.36 – Spectres en R_{Jet} des jets QCD dans divers échantillons : Sélection par déclenchement des jets avec un seuil de 100 GeV (QCD et Données), et réjections de muons et d'électrons appliquées aux échantillons contenant des τ (électrofaibles, $t\bar{t}$ et SUSY).

On constate dans cette figure 4.36 que les jets QCD présents dans les données (et dans les Monte-Carlo QCD) sont plus larges que ceux des échantillons électrofaibles, qui sont à leur tour plus larges que ceux des échantillons $t\bar{t}$ et SUSY (Fig. 4.24b). Les déterminations des probabilités de mauvaise identification sont présentées pour l'algorithme cut-based dans la figure 4.37, pour l'algorithme HPS dans la figure 4.38 et pour l'algorithme TaNC dans la figure 4.39.

La figure 4.37 nous montre une tendance similaire après identification complète (en bleu) à celle observée pour les efficacités d'identification, les niveaux de taux de faux dépassent même les 40% pour $R_{Jet} < 0,05$, ce qui est comparable à des niveaux d'efficacité d'identification, sans pour autant l'atteindre (60% pour des jets de τ aussi étroits). L'algorithme basé sur les coupures est très sensible à l'étroitesse des jets, mais sélectionne plus facilement les τ que les jets QCD à R_{Jet} égal.

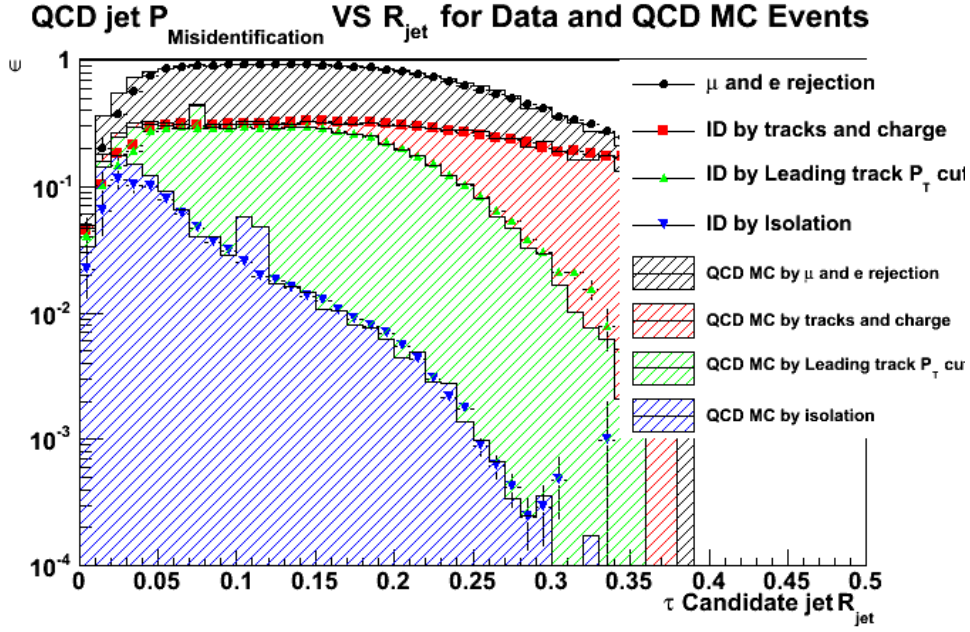


FIGURE 4.37 – Taux de mauvaise identification VS R_{Jet} avec information reconstruite pour un déclencheur de haut niveau jet de 100 GeV avec la méthode Basée sur les coupures.

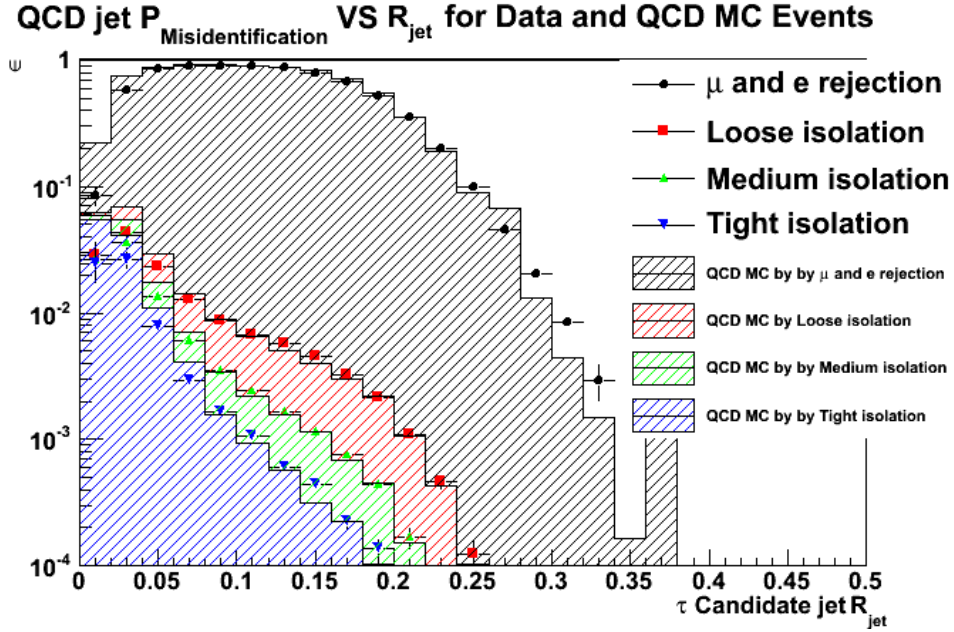


FIGURE 4.38 – Taux de mauvaise identification VS R_{Jet} avec information reconstruite pour un déclencheur de haut niveau jet de 100 GeV avec la méthode Hadron Plus Strips.

L'algorithme HPS atteint des probabilités de mauvaise identification de l'ordre de 5% pour des jets très étroits ($R_{Jet} < 0,05$), bien meilleures que celle obtenues par l'algorithme cut-based. On constate aussi que pour ces jets très étroits, le niveau d'isolation n'influence que très peu cette probabilité.

L'algorithme TaNC est aussi sensible au paramètre R_{Jet} et présente alors des probabilités

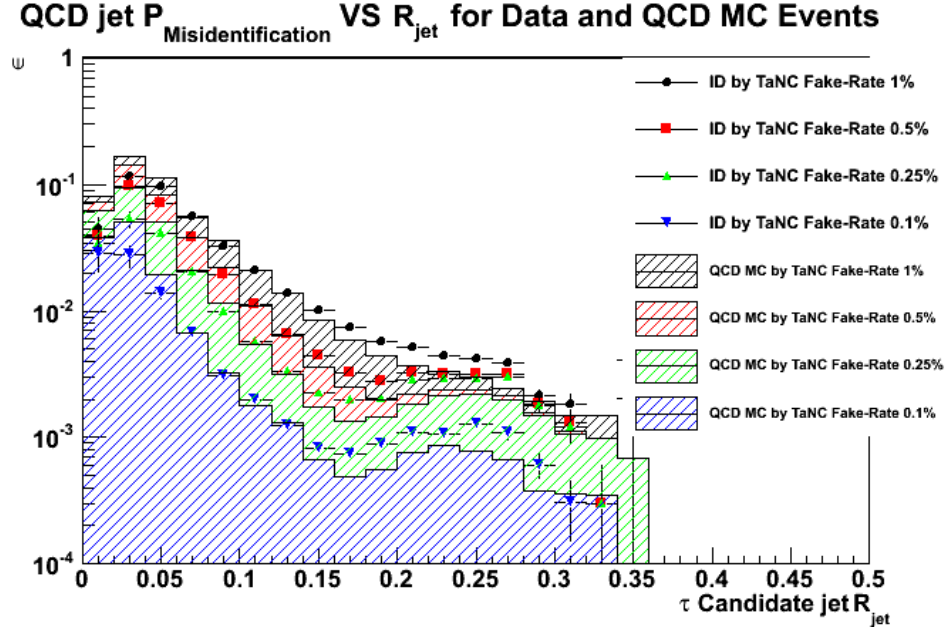


FIGURE 4.39 – Taux de mauvaise identification VS R_{Jet} avec information reconstruite pour un déclencheur de haut niveau jet de 100 GeV avec la méthode Tau Neural Classifier.

de mauvaise identification de l'ordre de 5 à 10 % pour des jets très étroits ($R_{Jet} < 0,05$). On constate une remontée anormale de cette probabilité pour des jets larges ($R_{Jet} > 0,15$). Ce problème est récurrent pour l'algorithme TaNC et l'hypothèse la plus probable pour expliquer ce phénomène est celle d'un surentraînement des réseaux de neurones.

4.3.5 Influence du seuil de déclenchement

Les évènements QCD sont sélectionnés dans les données à l'aide des systèmes de déclenchement. Nous avons vu que le choix du seuil de déclenchement favorisera la sélection d'évènements QCD dans certaines gammes de $\hat{P}_T$ comme présenté dans la figure 4.25. Un seuil élevé favorisera la sélection d'évènements contenant des jets "sonde" plus durs et plus centraux : un comportement retrouvé dans des évènements impliquant la désintégration de particules lourdes.

La conséquence sur la topologie de l'évènement d'un $\hat{P}_T$ plus élevé est la production de jets de plus haut P_T et de pseudo-rapidité $|\eta|$ plus basses avec pour résultat les influences sur les taux de mauvaise identification que nous avons vues dans les sections précédentes. Ici nous comparons les taux de mauvaise identifications en fonction de P_T , η et R_{Jet} sous les différentes sélections par déclenchement décrites dans la table 4.2. Nous vérifions aussi si le seuil de déclenchement n'a pas d'autre effet en superposant les taux de mauvaise identification VS P_T .

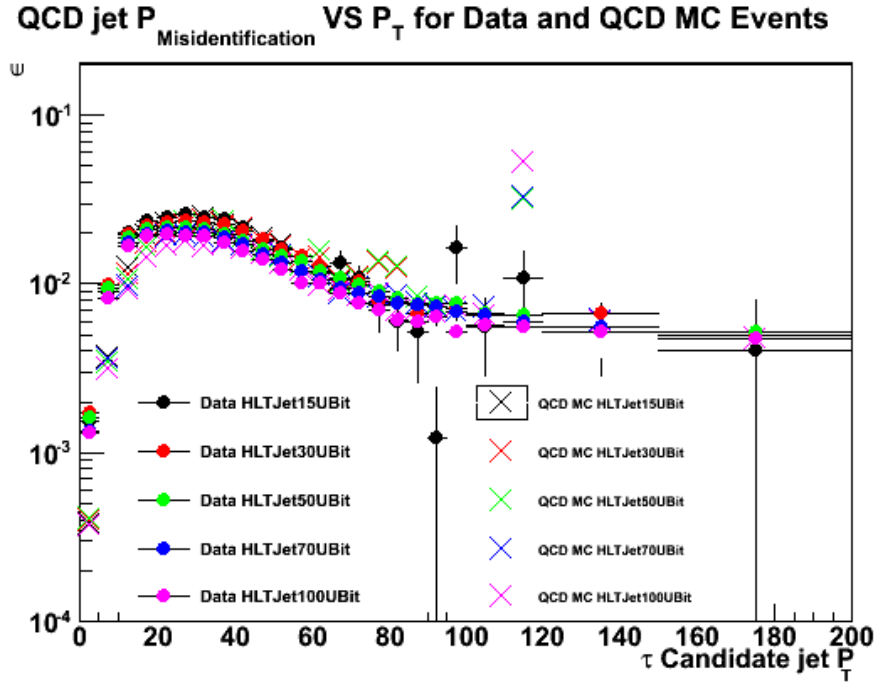


FIGURE 4.40 – Taux de mauvaise identification pour la méthode basée sur les coupures avec différents seuils de déclenchement.

La figure 4.40 nous montre que le seuil de déclenchement a bien un impact sur les taux de mauvaise identification mais il est relativement constant. Un seuil de déclenchement élevé induit aussi un effet d'environnement responsable d'une baisse globale du taux de mauvaise identification. Cet effet se traduit par l'apparition de particules supplémentaires dans les jets (Fig. 4.41).

Cette tendance est aussi observée pour les autres algorithmes (Fig 4.42 et 4.43).

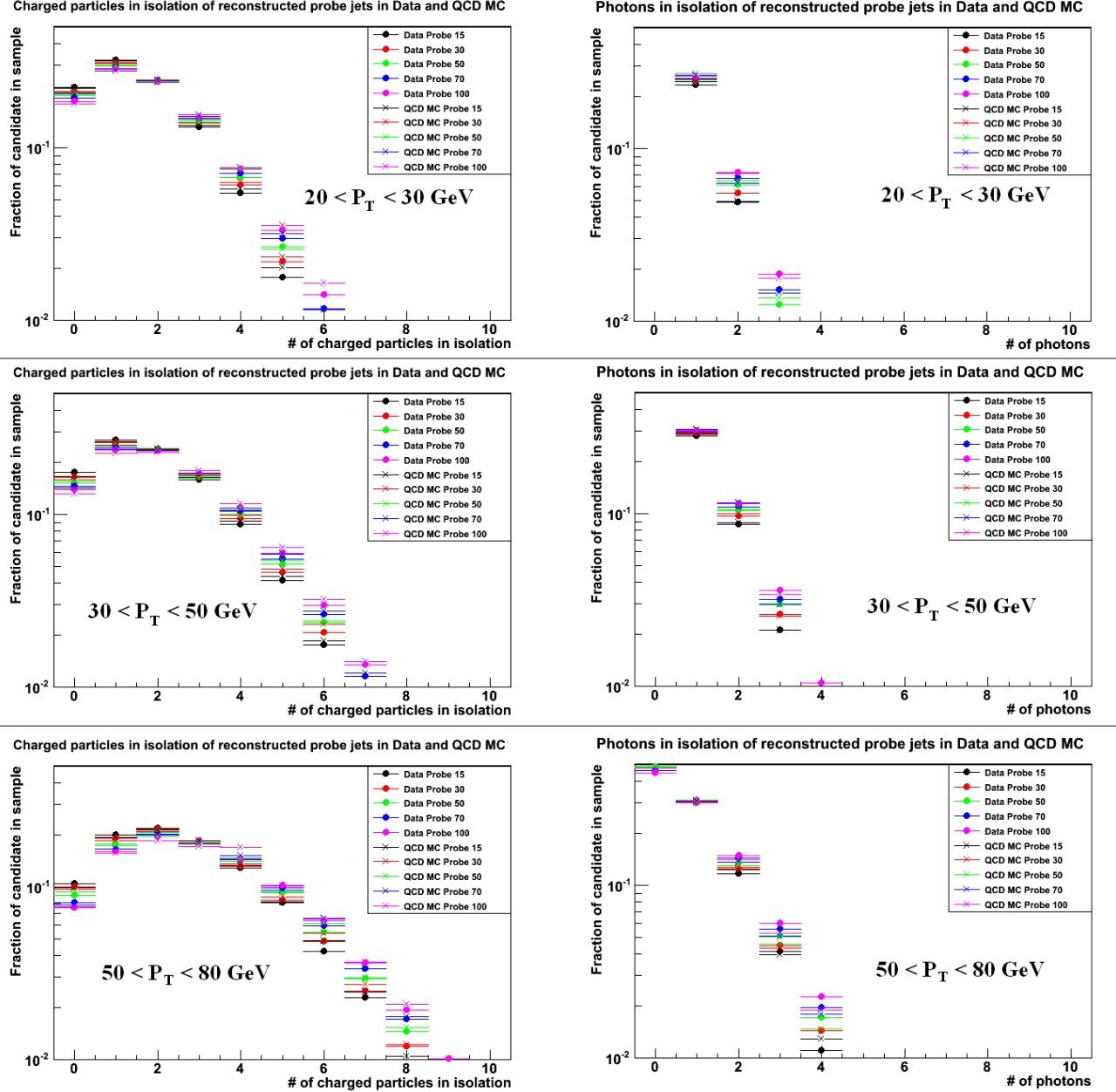


FIGURE 4.41 – Contenu de l’anneau d’isolation avant l’isolation par l’algorithme cut-based pour les jets “étiquette” dans des échantillons de données et QCD Monte-Carlo.

• Seuil de déclenchement et R_{Jet}

R_{Jet} est un paramètre particulièrement sensible au contenu de l’évènement. La largeur du jet de τ ou QCD étant liée à son impulsion transverse, le seuil de déclenchement aura donc un impact direct sur la sélection de jets plus ou moins larges : plus ce seuil est élevé, plus les jets sélectionnés seront étroits, les différences entre les différents seuils de déclenchement sont présentés dans la figure 4.44.

On constate en effet une plus grande multiplicité de jets étroits pour les seuils de déclenchement plus élevés, cependant la taux de mauvaise identification est beaucoup plus bas

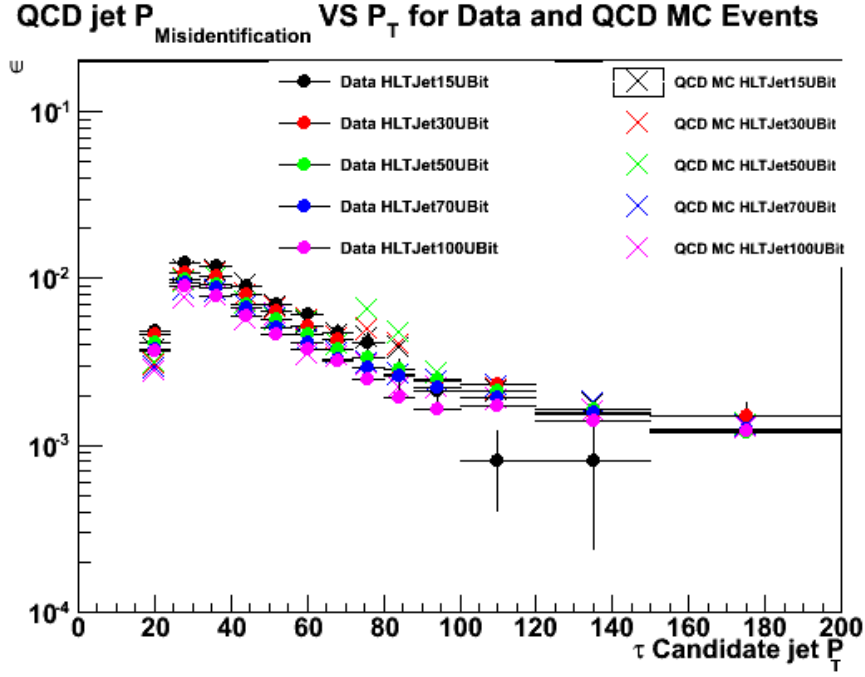


FIGURE 4.42 – Taux de mauvaise identification pour la méthode Hadron Plus Strips (isolation relâchée) avec différents seuils de déclenchement ($P_T > 20$ GeV, $|\eta| < 2, 4$).

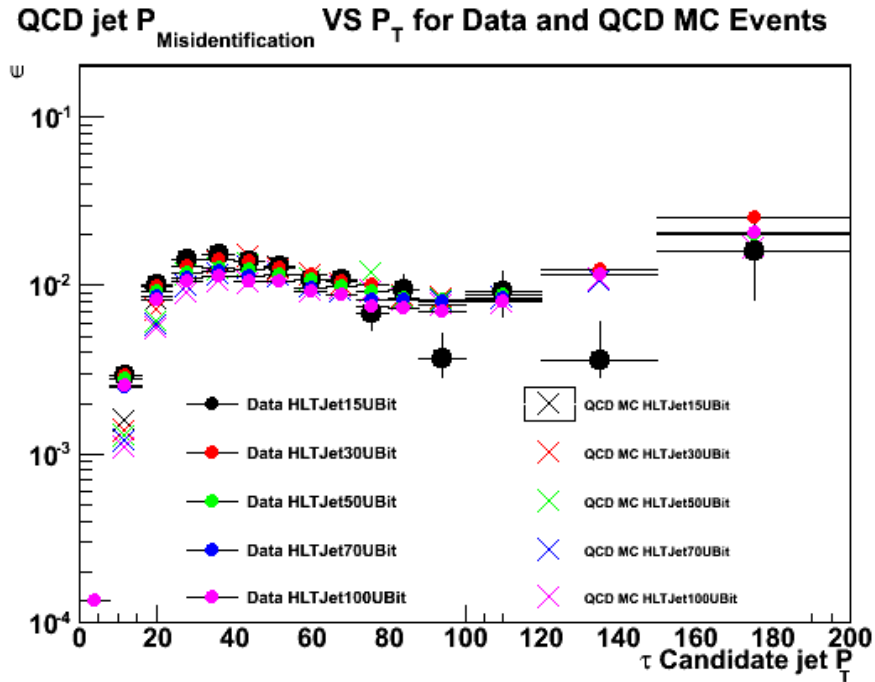


FIGURE 4.43 – Taux de mauvaise identification pour la méthode Tau Neural Classifier (isolation relâchée) avec différents seuils de déclenchement ($P_T > 20$ GeV, $|\eta| < 2, 4$).

pour les hauts seuils de déclenchement à R_{Jet} donné.

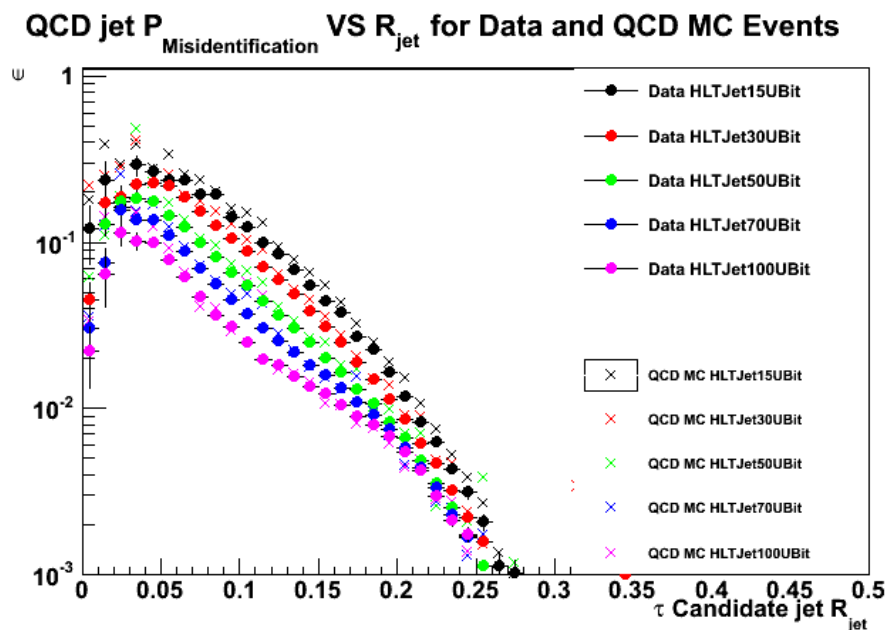


FIGURE 4.44 – Taux de mauvaise identification VS R_{Jet} pour le jet “sonde” de plus haute impulsion avec différents seuils de déclenchement, Identification basée sur les coupures.

R_{Jet} est en effet un paramètre basé sur les distributions en η, ϕ des constituants du jet, pondérées par l’énergie transverse de ces mêmes constituants. Un hadron de P_T petit devant le P_T du jet présent de l’anneau d’isolation contribuera peu au paramètre R_{Jet} , mais suffira à rejeter le candidat, d’où la baisse du taux de mauvaise identification.

• Conclusions sur les seuils de déclenchement

Les seuils de déclenchement ont une influence notable sur l’identification du lepton τ . Nous avons vu dans la section 2.2.3.6 que les systèmes de déclenchement fonctionnant à fréquence réduite induisent la sélection d’une fraction seulement de certains types d’événements. Nous avons observé dans le cas présent que les événements sélectionnés par les systèmes de déclenchement pour jet de bas P_T étaient en nombre sous-évalué par rapport au modèle des simulations. C’est un effet typique du facteur de diminution appliqué aux systèmes de déclenchements. Le choix de ce facteur étant motivé par le nombre de collisions molles beaucoup trop important aux luminosités effectives du LHC.

Pour conserver un bon accord entre données et simulations Monte-Carlo il ne faut utiliser qu’un unique seuil de déclenchement. On peut aussi pondérer les événements par un facteur inverse au facteur de diminution appliqué au seuil de déclenchement de manière à retrouver le contenu réel des collisions.

4.4 Conclusions sur les efficacités et les taux de mauvaise identification

Un bon accord des données avec les simulations Monte-Carlo est observé sous quasiment toutes les conditions : algorithme utilisé, variable considérée, seuil de déclenchement. Ceci nous révèle deux informations :

- L'échantillon de jets sélectionnés dans les données est bien un ensemble de jet QCD, les taux de mauvaise identification estimés dans les données sont donc ceux des jets QCD, ceci permettra de soustraire le bruit de fond concerné pour d'autres analyses de physique.
- Les simulations Monte-Carlo reproduisent de façon correcte les véritables données. Les études préliminaires menées à l'aide de ces simulations sont valides.

Nous avons observé des effets d'environnement dans les événements SUSY et QCD à haut $\hat{P}_T$ qui semblent indépendants des caractéristiques du jet. La méthode décrite dans la section 3.4 n'est applicable que dans une certaine limite car l'efficacité comme le taux de mauvaise identification ne dépend pas uniquement des paramètres P_T , η et R_{Jet} . Les efficacités et taux de mauvaise identification sont aussi sensibles au type d'évènement, il faut donc étudier plus en détail des paramètres caractérisant la topologie de l'évènement pour trouver le paramètre le plus significatif de l'effet qu'a l'environnement sur l'identification du lepton τ .

Chapitre 5

Etudes de l'impact de la topologie sur l'identification du lepton τ

5.1 Conditions d'analyse

5.1.1 Caractérisation de l'activité d'un évènement

Nous avons vu dans le chapitre précédent que le succès de l'identification d'un jet hadronique issu de la désintégration du τ n'est pas uniquement dépendant des paramètres caractérisant ce jet (P_T , η , R_{Jet}), mais est relié aussi au type d'environnement dans lequel est présent ce jet : Les jets de τ présentent des efficacités d'identification plus hautes dans les évènements électrofaibles (W et Z) que dans les évènements $t\bar{t}$ et SUSY ayant une multiplicité d'objets physiques plus importantes et une plus haute énergie globale (c.f chapitre 4 Fig. 4.12). Ce chapitre est consacré à la recherche et la caractérisation des sources de cette différence.

On caractérise l'activité de l'évènement par son contenu : le nombre d'objets physiques présent dans l'évènement et son énergie transverse totale (impliquant aussi l'énergie transverse manquante s'il y a lieu). Deux classes de variables globales nous servent à définir l'activité d'un évènement :

Multiplicité des objets : Ce sont les multiplicités d'un type d'objet physique : le nombre de hadrons chargés et neutres, de photons (provenant majoritairement des π^0), et de jets présents dans l'évènement. Des coupures peuvent être appliquées sur les caractéristiques de ces objets (en P_T et η notamment).

Énergie transverse après collision : C'est l'énergie de toutes les particules issues de la collision, elle est mesurée dans le plan transverse cité dans la section 2.2.1 : l'axe z dans le référentiel de CMS est l'axe des faisceaux. L'impulsion transverse des deux partons avant collision est alors $\Sigma \vec{P}_T = \vec{0}$, et donc la conservation de l'impulsion veut que $\Sigma \vec{P}_T = \vec{0}$ pour les particules issues de la collision. On caractérise alors la collision avec la somme scalaire des énergies transverses mesurant la fraction de l'impulsion longitudinale passée dans le plan transverse. On peut aussi déterminer l'énergie transverse manquante utilisée notamment dans les analyses SUSY[61].

On utilise pour caractériser l'évènement les variables suivantes :

- **Energie transverse totale ΣE_T :** Somme scalaire des énergies transverse de tous les

objets contenus dans l'évènement, cette variable est calculée en additionnant les énergies transverses scalaires des objets Particle Flow. $\Sigma E_T = \Sigma E \sin(\theta)$

- **Energie transverse manquante (MET)** : Cette grandeur est calculée par $MET = -\|\sum_i \vec{E}_{T,i}\|$, où $\vec{E}_{T,i}$ étant le 3-vecteur (projeté sur xOy) d'énergie transverse de la particule i contenue dans l'évènement. Cette variable nous donne l'énergie transverse emportée par l'ensemble de particules invisibles (neutrinos et hypothétiques neutralinos) contenues dans l'évènement ou l'imprécision sur les mesures.
- **Multiplicité de hadrons chargés ($N_{Charged}$)** : Nombre total de hadrons chargés présents dans l'évènement. Afin de rester en accord avec les critères d'identification du lepton τ une coupure en $P_T > 1 \text{ GeV}$ et $|\eta| < 2.4$ est appliquée sur ces hadrons, les autres ne sont pas comptés pour déterminer le paramètre $N_{charged}$.
- **Multiplicité de photons ($N_{photons}$)** : Nombre total de photons présents dans l'évènement. Une coupure en $P_T > 1,5 \text{ GeV}$ et $|\eta| < 2.4$ est appliquée sur ces photons.
- **Multiplicité de jets (N_{jets})** : Nombre total de jets présents dans l'évènement. Une coupure sur ces jets en $|\eta| < 2.4$ et en P_T sont appliquées : les seuils en P_T sont 10, 15, 20, 25 et 30 GeV.

5.1.2 Critères de présélection des candidats τ

Nous avons vu dans les sections précédentes la forte dépendance des performance du Tau-ID en impulsion transverse et en pseudo-rapidité. Le spectre en P_T , η des τ et des jets QCD dans un évènement dépend lui-même du type d'évènement considéré. Nous pouvons limiter l'influence de ces deux paramètres dans les études d'activité en se plaçant dans une zone où les performances varient peu et où la statistique est suffisante pour déterminer une tendance.

• Présélection des candidats relatifs aux calculs d'efficacité d'identification

Deux coupures sont imposées sur la présélection des taus générés pour le calcul des efficacités : $P_T > 50 \text{ GeV}$ et $|\eta| < 1$. La même étude à été également menée dans les régions à faible P_T ($20 \text{ GeV} < P_T < 50 \text{ GeV}$ et $|\eta| < 1$). Le choix des intervalles étudiés est motivé par les efficacités présentées dans le chapitre 4 (Fig. 4.12 et 4.17). Les efficacités varient moins (plateau) pour des candidats de $P_T > 50 \text{ GeV}$ et de pseudo-rapidité $|\eta| < 1$ que pour les candidats des autres intervalles.

Les échantillons utilisés pour les études d'efficacité sont des simulations Monte Carlo, nous utilisons pour les calculs d'efficacité la formule 4.2 vue dans le chapitre 4. Les objets présélectionnés (dénominateur de la formule 4.2) sont des τ générés se désintégrant dans la voie hadronique. Ce tau généré est sélectionné (numérateur de la formule 4.2) si son candidat associé passe les tests d'identification du lepton τ . Nous n'appliquons pas la réjection de muon et d'électrons dans la présélection des candidats car le critère d'association entre un tau généré et son candidat reconstruit suffit à rejeter ces leptons légers.

• Présélection des candidats relatifs aux calculs des taux de mauvaise identifi-

cation

Des coupures identiques à celle imposées pour les calculs d'efficacités sont appliquées aux candidats présélectionnés dans un intervalle en P_T et $|\eta|$.

Dans le cas où on étudie les faux (jet QCD) dans des échantillons qui ne sont pas de la QCD, la contribution des électrons et des muons au taux de mauvaise identification est significative. Aussi des échantillons contenant une importante part de ces leptons légers auront tendance à retourner un taux de mauvaise identification plus élevé que les échantillons purement QCD.

Tous les échantillons Monte Carlo utilisés pour cette étude sont susceptibles de donner des leptons légers dans l'état final : les τ se désintègrent en muon ou en électron dans 35% des cas. De plus les sparticules dans les événements supersymétriques peuvent se désintégrer directement en lepton léger (dans les cascades SUSY, c.f. section 1.4). Les réjections d'électrons et de muons sont alors imposées dans la présélection pour supprimer les différences pouvant exister dans les échantillons à cause de ces leptons légers.

5.1.3 Critères d'identification des candidats τ

Les candidats τ soumis aux différents algorithmes d'identification sont les candidats τ présélectionnés. On y inclut une condition supplémentaire pour s'assurer que le candidat est bien associé à un τ généré dans les calculs d'efficacité, et est suffisamment éloigné d'un τ généré pour le calcul des taux de mauvaise identification. Cette condition n'est demandée que dans les études utilisant les simulations Monte Carlo contenant des τ et est la suivante :

- Pour les calculs d'efficacité un candidat τ doit être situé à une distance $\Delta R < 0,5$ du τ généré auquel il est associé.
- Pour les calculs de taux de mauvaise identification, un jet est considéré comme n'étant pas associé à un tau si la distance qui les sépare $\Delta R > 0,5$.

Pour les données issues du LHC on considère tout objet jet comme un candidat potentiel (c.f. section 4.1.1.2). Nous avons vu dans le chapitre précédent que ce lot de données sélectionné par la méthode du “tag and probe” est largement dominé par la QCD et donc sera traité en tant qu'échantillon QCD.

5.2 Influence de l'environnement sur les efficacités d'identification

Nous avons vu dans la section 4.2 que les τ présents dans les échantillons SUSY et $t\bar{t}$ étaient moins efficacement identifiés que ceux issus d'événements électrofaibles. L'objectif de cette étude est de comprendre les raisons de cette baisse, et de quantifier l'impact de l'environnement des événements sur les efficacités d'identification du lepton τ .

5.2.1 Spectres en variables globales

L'étude des spectres de variables globales nous montre les différences qui existent entre les diverses topologies d'évènement. Les variables d'environnement étudiées sont celles présentées dans la section 5.1.1. Les échantillons utilisés sont ceux présentés dans la section 4.1.1 (table 4.1), tous les spectres sont normalisés à 1 afin d'obtenir la fraction des évènements ayant les mêmes caractéristiques.

- ΣE_T

Les spectres des candidats τ sont présentés dans la figure 5.1 où les jets sonde sont sélectionnés dans les données et QCD Monte Carlo par des jets étiquette avec des seuils de déclenchement de 30 et de 100 GeV.

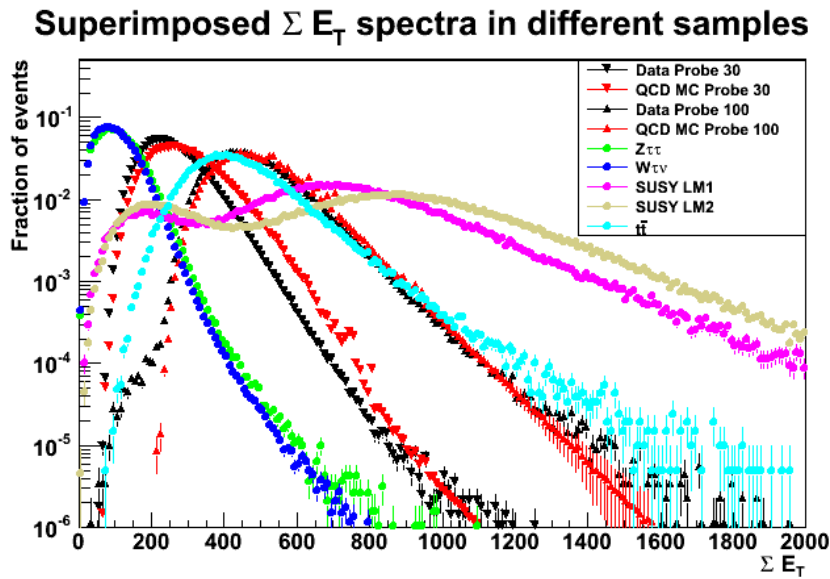


FIGURE 5.1 – Spectres en ΣE_T , les évènements dans les données et les Monte Carlo QCD sont sélectionnés par un seuil de déclenchement non corrigé (pour jet) de 30 GeV et 100 GeV.

On constate dans cette figure quatre tendances différentes selon les échantillons considérés : Les évènements électrofaibles contiennent une énergie transverse totale plus faible que les évènements QCD (données), ayant à leur tour une énergie totale plus faible que les évènements $t\bar{t}$. Les évènements SUSY présentent l'énergie totale la plus élevée, on constate deux régimes dans ces évènements, ils correspondent soit à la production directe de particules plus légères (neutralinos, charginos et sleptons) soit à la présence de cascade SUSY issues de la désintégration d'un squark ou d'un gluino, la figure 5.2 illustre ce phénomène.

Un excès d'évènements de basse énergie transverse est visible dans les données lorsqu'un seuil de déclenchement de 100 GeV est requis pour les jets étiquette.

Nous observons dans la figure 5.2 que les particules les plus présentes dans les évènements SUSY sont les neutralinos et les charginos légers ($\tilde{\chi}_1^0$, $\tilde{\chi}_2^0$ et $\tilde{\chi}_1^\pm$) mais le contenu en gluinos et en squarks de toutes les saveurs (u,d,s...), chiralités (Droite ou gauche), ou type (1 ou 2)

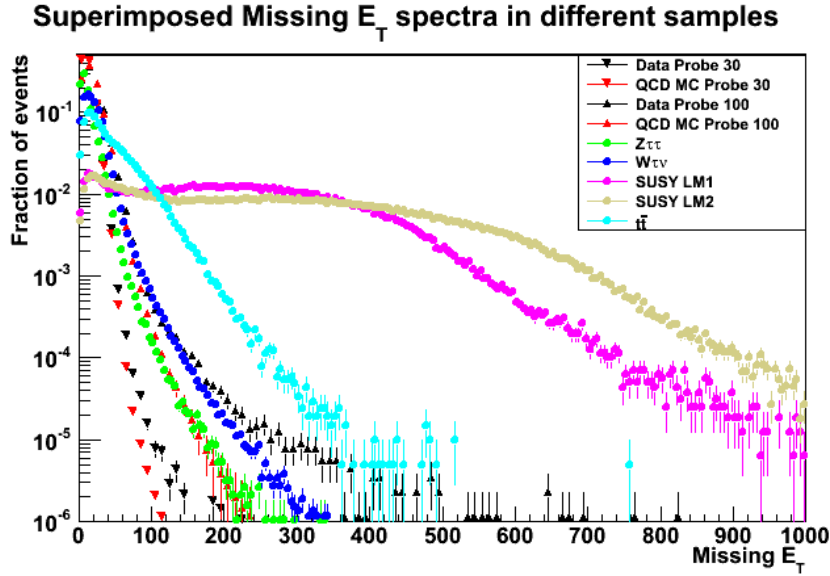


FIGURE 5.3 – Spectres en énergie transverse manquante (sans coupure pour les données et les Monte Carlo QCD), les événements dans les données et les Monte Carlo QCD sont sélectionnés par un seuil de déclenchement non corrigé (pour jet) de 30 GeV et 100 GeV.

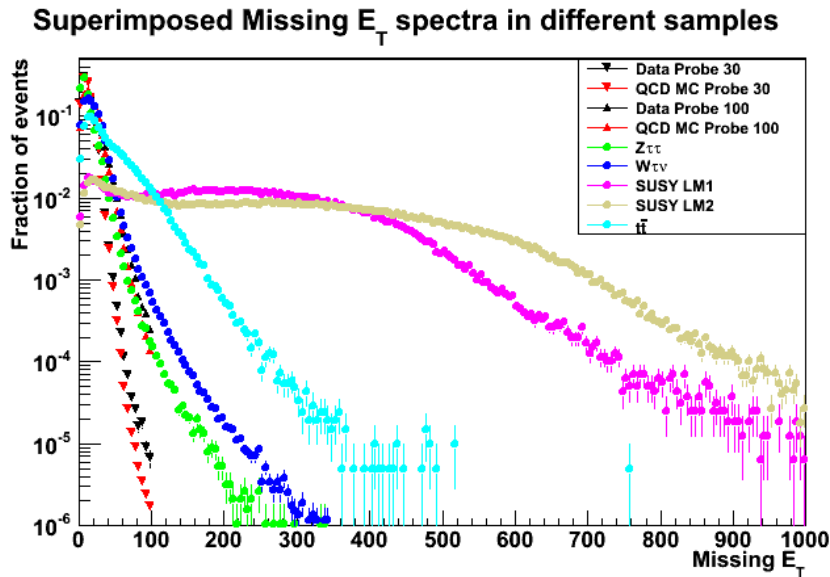


FIGURE 5.4 – Spectres en énergie transverse manquante (avec une coupure $MET < 100 \text{ GeV}$ pour les données et les Monte Carlo QCD), les événements dans les données et les Monte Carlo QCD sont sélectionnés par un seuil de déclenchement non corrigé (pour jet) de 30 GeV et 100 GeV.

étant plus lourd que les bosons W et Z l'énergie transverse sera plus importante. Enfin les événements SUSY présentent une queue de distribution d'énergie transverse manquante vers les valeurs élevées car les particules invisibles sont les neutralinos massifs (deux neutralinos par événement en plus des ν_τ) et emportent une énergie plus importante que les neutrinos.

Afin de se rapprocher de l'hypothèse que l'échantillon de données est un échantillon purement QCD, nous imposons la coupure $MET < 100 \text{ GeV}$ sur les données pour la mesure des distributions des autres variables globales. Ceci pour supprimer le biais dû à une sélection supplémentaire d'événements avec une énergie transverse manquante dans les données.

- Multiplicités en hadrons chargés et photons

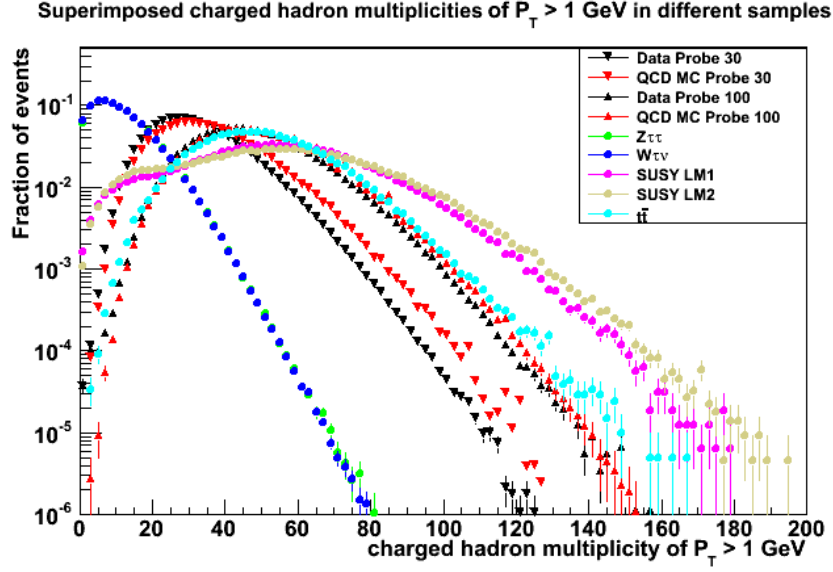


FIGURE 5.5 – Spectres en multiplicité de hadrons chargés de $P_T > 1 \text{ GeV}$, les événements dans les données et les Monte Carlo QCD sont sélectionnés par un seuil de déclenchement non corrigé (pour jet) de 30 GeV et 100 GeV.

La variable ΣE_T estime la dureté d'un événement mais ne donne pas d'indication sur l'occupation des particules de cet événement dans le détecteur. On cherche à caractériser l'importance de l'environnement en étudiant les multiplicités de particules impliquées dans l'identification de leptons τ : les hadrons chargés et les photons (c.f. section 3.1.2). Les multiplicités de hadrons chargés (on impose une coupure $P_T > 1 \text{ GeV}$ sur ces hadrons pour concorder avec les seuils imposés par les algorithmes d'identification de τ) sont présentées dans la figure (Fig. 5.5), et les multiplicités de photons dans la figure 5.6.

Les événements électrofaibles nous donnent la multiplicité de hadrons chargés la moins importante parmi les différents environnements ceci pour la raison que les bosons W et Z sont forcés dans leur désintégration en τ et donc les hadrons ne peuvent provenir que des τ ou de jets supplémentaires (production de W/Z + N jets du aux collisions multiples, aux radiations d'état initial et aux perturbations d'ordre supérieur). Les événements QCD présentent une activité plus importante car il s'agit d'événements impliquant les processus décrits dans le chapitre 1 : multi jets de faible P_T (minijets) et événement sous jacent (élongation du lien coloré) donnant dans l'état final de nombreux hadrons. De plus la multiplicité de particules dans les événements QCD est fortement liée au $\hat{P}_T$ de ces événements : les événements sélectionnés avec un seuil de déclenchement pour jet de 100

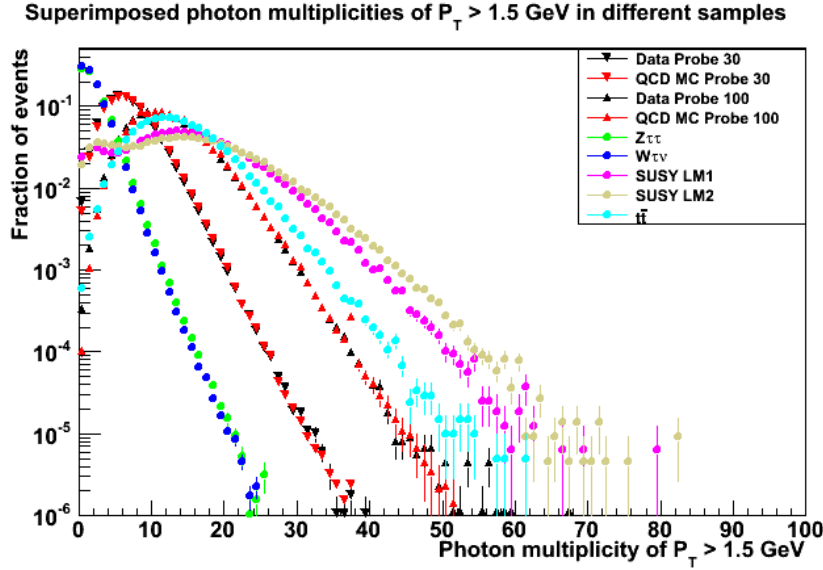


FIGURE 5.6 – Spectres en multiplicité de photons de $P_T > 1,5 \text{ GeV}$, les évènements dans les données et les Monte Carlo QCD sont sélectionnés par un seuil de déclenchement non corrigé (pour jet) de 30 GeV et 100 GeV.

GeV (sélectionnant les évènements QCD de haut $\hat{P}_T$, c.f. section 4.3.2 Fig. 4.25) donnent des multiplicités de particules plus importantes que dans les évènements de bas $\hat{P}_T$: les multiplicités de particules sont proches de celles observées dans les évènements $t\bar{t}$. Enfin les échantillons ayant la multiplicité de hadrons chargés la plus élevée sont les évènements $t\bar{t}$ (production de $t\bar{t} + N$ jets et fragmentation de quarks b lourds issus de la désintégration du top) et SUSY (dû aux cascades et à la masse des sparticules) : les évènements SUSY produisent de nombreux jets (jusqu'à 6 dans une production $\tilde{s} + \tilde{g}$) et peuvent être produits en association à d'autres jets (SUSY + N jets).

La multiplicité de photons, elle, provient majoritairement de la désintégration de π^0 . Le nombre de photons de haute énergie sera alors lié à l'activité hadronique de l'évènement (les détails sont dans la section suivante). La figure 5.6 nous montre pour les multiplicités de photons des résultats similaires à ceux des hadrons chargés : les évènements électrofaibles ont la multiplicité de photons la plus faible, les évènements QCD ont une multiplicité moyenne de photon et les évènements $t\bar{t}$ possèdent la multiplicité de photons la plus élevée.

La multiplicité de particules (hadrons plus photons) est une définition de l'activité globale d'un évènement. Mais les algorithmes d'identification de τ ne prennent pas en compte les hadrons et les photons de façon équivalente pour rendre leur décision, nous traiterons alors ces deux paramètres séparément.

• Multiplicité de jets

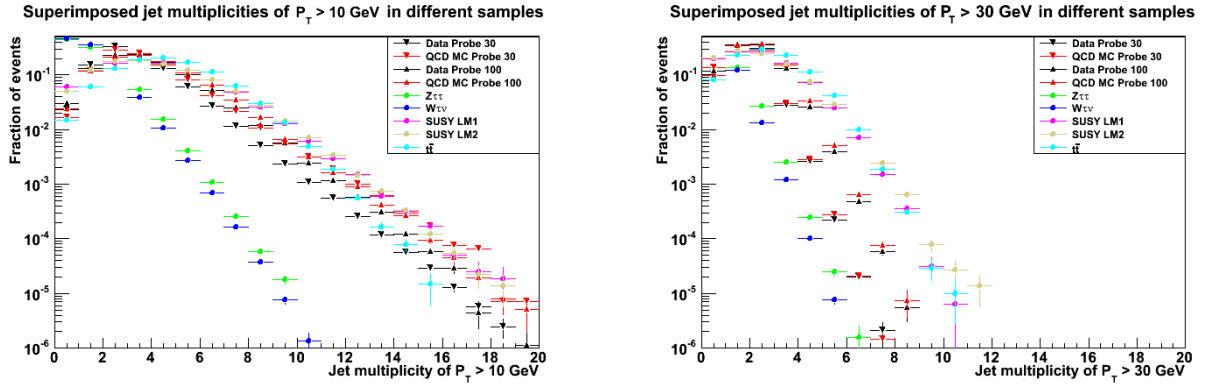


FIGURE 5.7 – a) Spectres en multiplicité de jets de $P_T > 10 \text{ GeV}$, b) Spectres en multiplicité de jets de $P_T > 30 \text{ GeV}$. Les événements dans les données et les Monte Carlo QCD sont sélectionnés par un seuil de déclenchement non corrigé (pour jet) de 30 GeV et 100 GeV.

Les événements électrofaibles présentent la multiplicité de jets la plus basse (Fig. 5.7), les jets sont dus soit à la désintégration du/des τ , soit aux quelques jets associé à la production du W/Z. Les événements QCD ont une multiplicité haute de jets de basse impulsion transverse, mais une multiplicité assez faible de jets de haute impulsion transverse (la fragmentation d'un quark peut donner plusieurs jets). Les événements top et SUSY ont encore la multiplicité de jet la plus importante, du à la physique qui gouverne ces processus.

5.2.2 Corrélations entre variables caractérisant l'énergie globale

Les deux variables caractérisant l'énergie globale sont ΣE_T et MET (section 5.1). L'énergie transverse manquante nous donne des indications sur les particules invisibles (neutrinos et/ou neutralinos). Nous étudions qualitativement les corrélations entre les deux variables caractérisant l'énergie transverse ΣE_T et MET pour deux échantillons contenant des τ ($Z\tau\tau$ Fig. 5.8 et SUSY LM1 Fig. 5.9), et les échantillons QCD (Monte Carlo Fig. 5.11 et données Fig. 5.10). Nous étudions ici aussi les corrélations entre la somme des énergies transverses ΣE_T et la multiplicité de hadrons chargés pour ces échantillons.

• **Corrélations entre ΣE_T et l'énergie transverse manquante :** La distribution de ΣE_T VS MET est large pour les 4 échantillons concernés : $Z\tau\tau$, SUSY LM1, QCD et Données. Il n'y a qualitativement pas de corrélation entre la somme des énergies transverses et l'énergie transverse manquante. La limite physique $MET = \Sigma E_T$ est observée, en particulier dans les événements SUSY pouvant avoir des MET très élevées. Les événements SUSY présentent deux régions correspondant aux deux régimes décrits dans la section 5.2.1.

• **Corrélations entre ΣE_T et la multiplicité de particules (hadrons chargés + photons) :** Il y a intuitivement une corrélation entre la multiplicité de particules dans l'événement et la somme des énergies transverses :

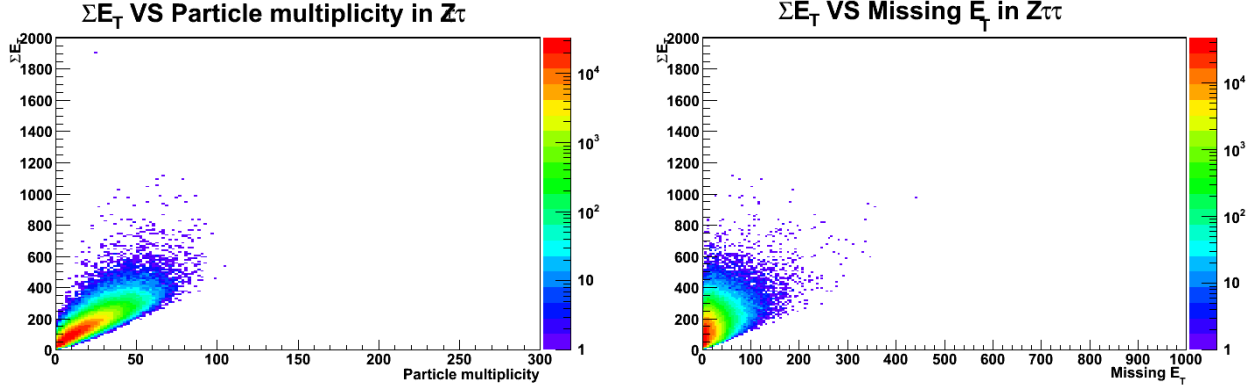


FIGURE 5.8 – Distributions dans les évènements $Z\tau\tau$ de ΣE_T VS a) Multiplicité de hadrons Chargés, b) Energie transverse manquante.

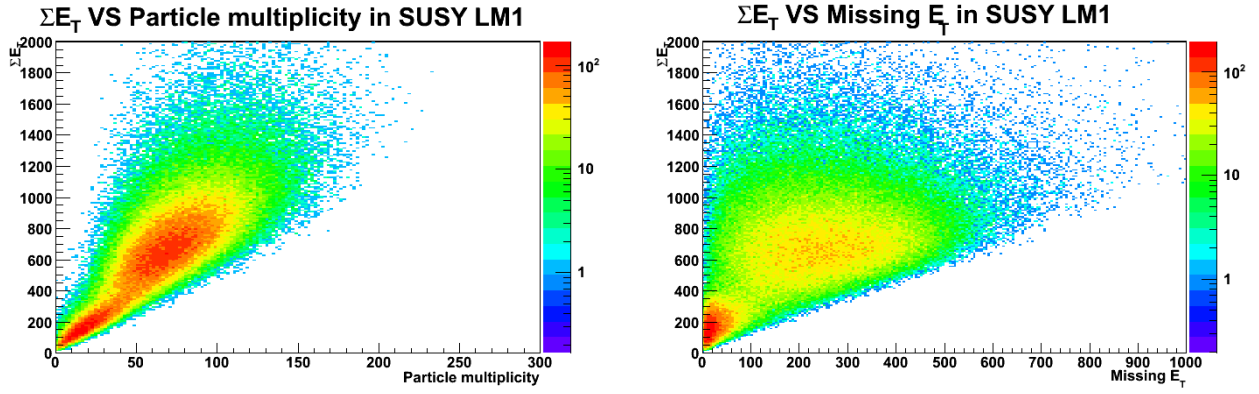


FIGURE 5.9 – Distributions dans les évènements SUSY LM1 de ΣE_T VS a) Multiplicité de hadrons Chargés, b) Energie transverse manquante.

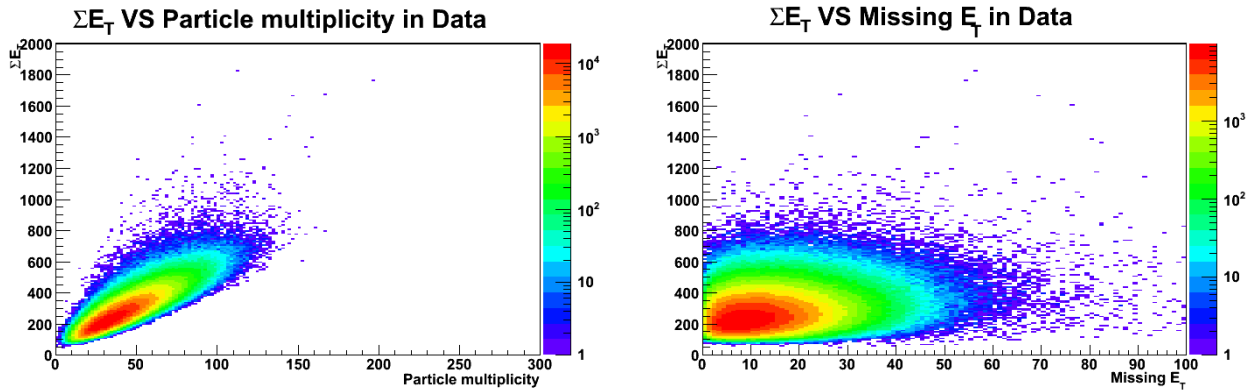


FIGURE 5.10 – Distributions dans les données du LHC (seuil de déclenchement pour Jet de 30 GeV) de ΣE_T VS a) Multiplicité de hadrons Chargés, b) Energie transverse manquante.

$$\Sigma E_T = N_{particules} \times \langle E_T \rangle \quad (5.1)$$

Où $\langle E_T \rangle$ est l'énergie transverse moyenne des particules dans l'évènement. On constate qualitativement que la distribution de ΣE_T est en effet fortement corrélée à la multiplicité de particules, la distribution de ΣE_T VS la multiplicité de particules (hadrons chargés +

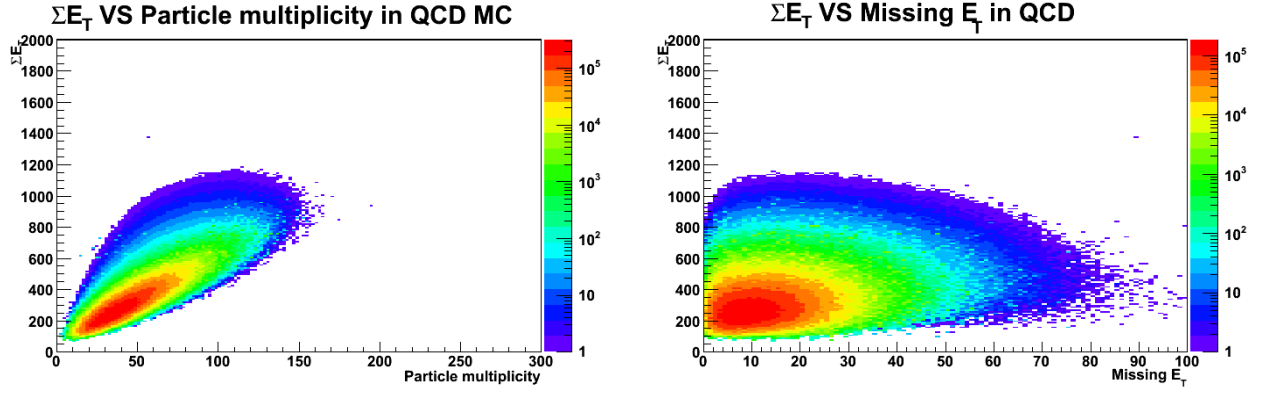


FIGURE 5.11 – Distributions dans les évènements QCD MC (seuil de déclenchement pour Jet de 30 GeV) de ΣE_T VS **a)** Multiplicité de hadrons Chargés, **b)** Energie transverse manquante.

photons) possède une certaine largeur dépendant de l'évènement car l'énergie transverse moyenne $\langle E_T \rangle$ est une distribution élargie.

5.2.3 Corrélations entre les multiplicités d'objets

Les trois multiplicités d'objets concernant cette étude sont la multiplicité de hadrons chargés ($P_T > 1 \text{ GeV}$, $|\eta| < 2,5$), de photons ($P_T > 1,5 \text{ GeV}$, $|\eta| < 2,5$) et de jets ($P_T > 10 \text{ à } 30 \text{ GeV}$, $|\eta| < 2,5$). Les corrélations entre ces variables sont montrées qualitativement pour les 4 échantillons $Z\tau\tau$, SUSY LM1, QCD et données.

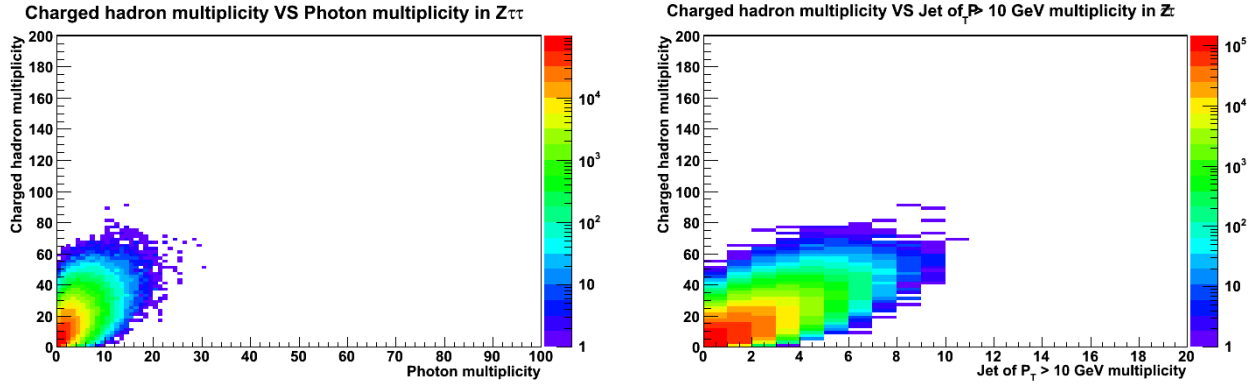


FIGURE 5.12 – Distributions dans les événements $Z \rightarrow \tau\tau$ de la multiplicité de hadrons chargés VS a) Multiplicité de photons, b) Multiplicité de jets de $P_T > 10 \text{ GeV}$.

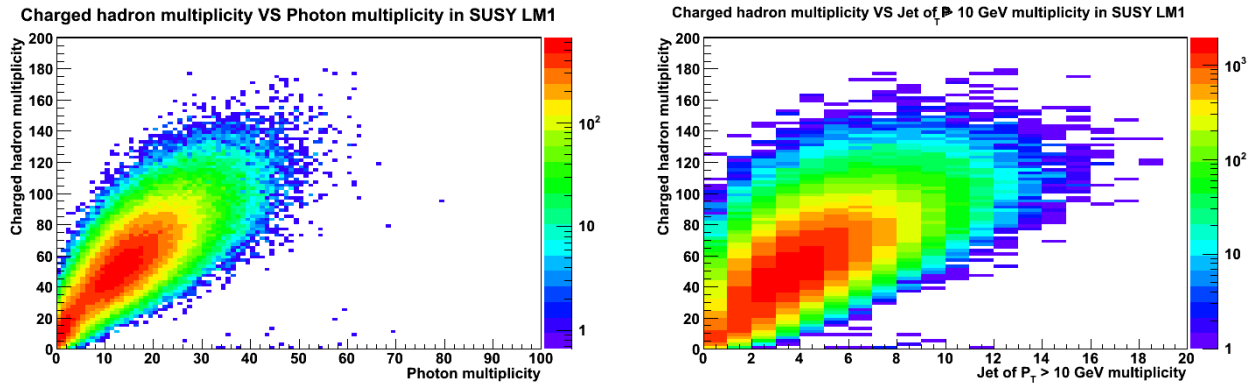


FIGURE 5.13 – Distributions dans les événements SUSY LM1 de la multiplicité de hadrons chargés VS a) Multiplicité de photons, b) Multiplicité de jets de $P_T > 10 \text{ GeV}$.

- **Corrélations entre la multiplicité de hadrons chargés et la multiplicité de photons :** Les jets de quarks et de gluons se fragmentent puis s'hadronisent par paires ou triplets de quarks. Lors de l'hadronisation un certain nombre de π^0 sont créés, leur nombre augmentant à mesure que le nombre de hadrons total dans ce jet augmente entraînant donc une augmentation de la multiplicité de photons. On observe alors une corrélation faible (très étendue) entre le nombre de hadrons chargés et le nombre de photons dans les événements $Z\tau\tau$ (Fig. 5.12), QCD (Fig. 5.15) et données (Fig. 5.14). Cette distribution est plus fine dans les événements SUSY (Fig. 5.13) car les multiplicités y sont plus importantes.

- **Corrélations entre la multiplicité de hadrons chargés et la multiplicité de jets :** La sources de jets diffèrent selon l'événement, les événements purement QCD ont tendance à donner une grande multiplicité de jets pour une multiplicités de hadrons

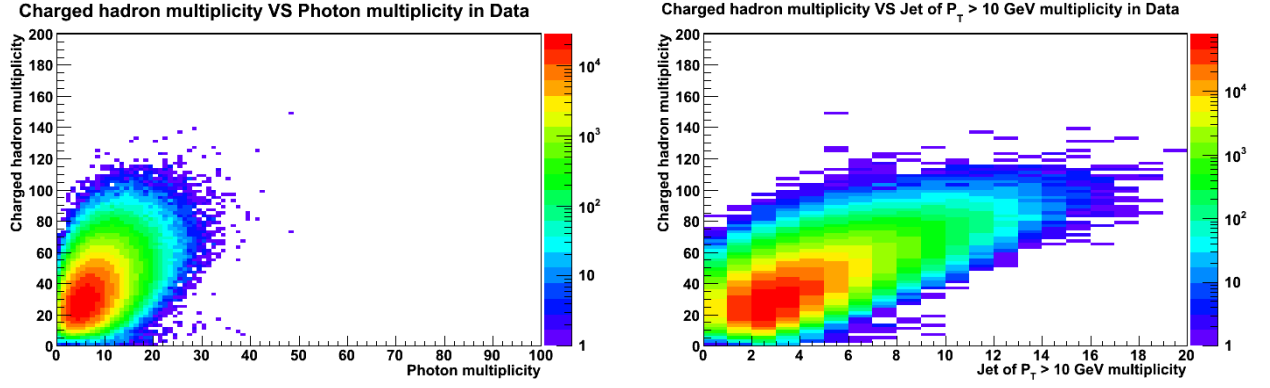


FIGURE 5.14 – Distributions dans les données du LHC de la multiplicité de hadrons chargés VS **a)** Multiplicité de photons, **b)** Multiplicité de jets de $P_T > 10$ GeV.

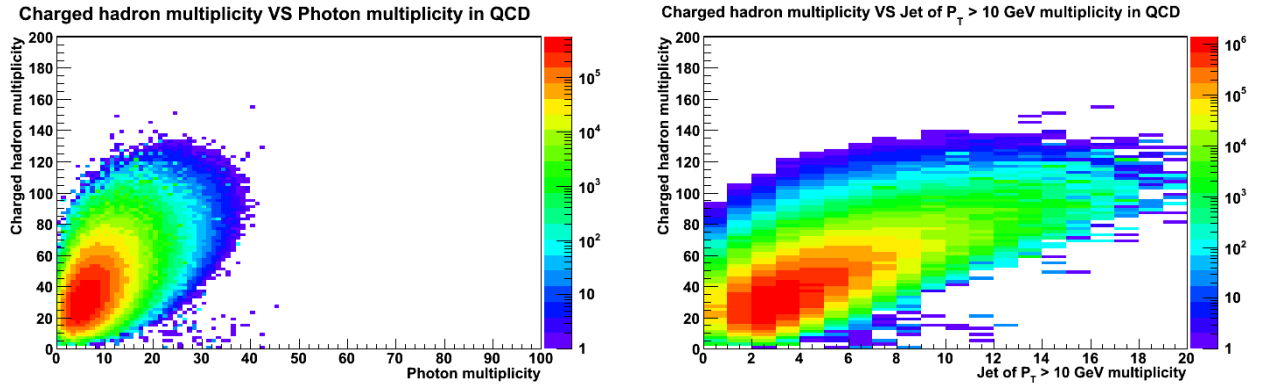


FIGURE 5.15 – Distributions dans les événements QCD MC de la multiplicité de hadrons chargés VS **a)** Multiplicité de photons, **b)** Multiplicité de jets de $P_T > 10$ GeV.

relativement basse comme le montrent les figures 5.15 et 5.14. Les événements SUSY contiennent de nombreux jets, mais ceux-ci sont issus de particules plus lourdes et donnent une multiplicité plus grande (Fig. 5.13). La multiplicité de jets est donc liée au processus physique et on observe alors une corrélation moins prononcée entre la multiplicité de hadrons chargés et la multiplicité de jets.

5.2.4 Influence de l'environnement sur l'identification basée sur les coupures

5.2.4.1 Influence des variables caractérisant l'énergie globale

Les coupures appliquées aux candidats τ (jets) choisies pour cette étude sont celles de la sélection complète :

- $P_T > 20 \text{ GeV}$ et $|\eta| < 2,4$, cette sélection est implicite pour les intervalles en P_T et η étudiés (c.f. section 5.1.2).
- Trace de plus haute impulsion proche de l'axe du jet ($\Delta R < 0,15$) et d'impulsion transverse supérieure à 5 GeV.
- Isolation par les traces avec cône signal évolutif (c.f. section 3.3.1).
- Isolation par le calorimètre électromagnétique avec cône signal fixe.
- Jet contenant 1 ou 3 particules chargées (traces) et dont la charge totale est de ± 1 .
- Réjection d'électrons et de muons.

• Influence de ΣE_T :

Nous avons vu dans la section précédente que la somme des énergies transverse est corrélée à l'activité hadronique de l'évènement. Nous observons alors une baisse des efficacités à mesure que l'énergie transverse augmente.

On teste l'hypothèse d'une régression exponentielle ($\epsilon_{ID}(X_G) = e^{aX_G+b}$ où X_G est une variable caractérisant l'environnement) pour paramétrer cette baisse d'efficacité. Les résultats de la régression sont présentés dans la Fig. 5.16 et le tableau 5.1 : les erreurs sont calculées avec le test du χ^2 par la routine de paramétrage de ROOT.

La validité d'un paramétrage commune aux cinq échantillons est estimée grâce à la moyenne sur les paramètres a et b (respectivement $\bar{a}$ et $\bar{b}$) de la régression, l'écart type sur a : σ_a (relation 5.2), et l'écart type relatif sur a $\sigma_{r,a} = \frac{\sigma_a}{\bar{a}}$ (les mêmes grandeurs statistiques sont déterminées pour le paramètre b).

$$\sigma_a = \sqrt{\frac{1}{n-1} \sum_i^n (a_i - \bar{a})^2} \quad (5.2)$$

La figure 5.16 nous montre l'influence de ΣE_T dans une zone en P_T et en $|\eta|$ où les efficacités varient peu. Une étude similaire menée dans un intervalle de P_T plus bas (Fig. 5.17) nous montre une déviation plus importante entre les différents échantillons : un écart type relatif sur la pente a de l'exponentielle de 45,7% pour l'intervalle $P_T > 50 \text{ GeV}$ contre 61,1% dans l'intervalle $20 \text{ GeV} < P_T < 50 \text{ GeV}$.

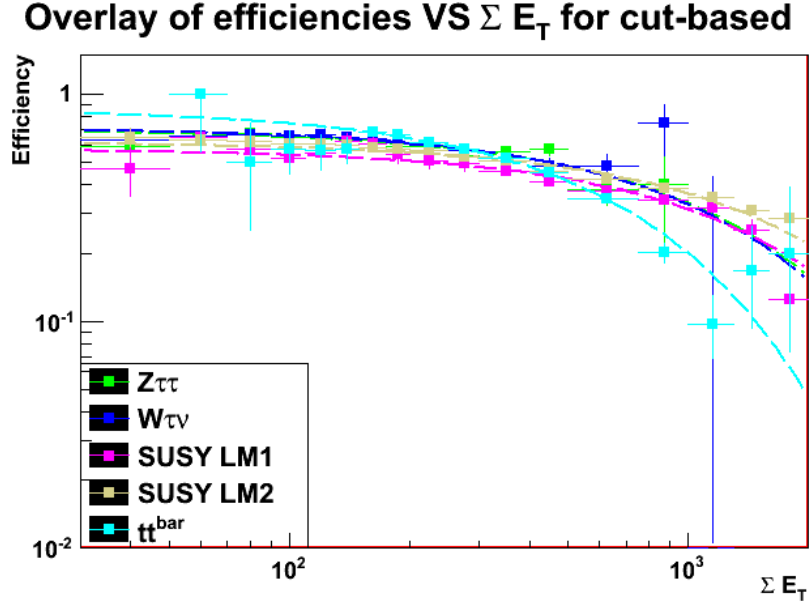


FIGURE 5.16 – Efficacités d'identification basée sur les coupures VS ΣE_T pour des candidats de $P_T > 50 \text{ GeV}$, $|\eta| < 1$.

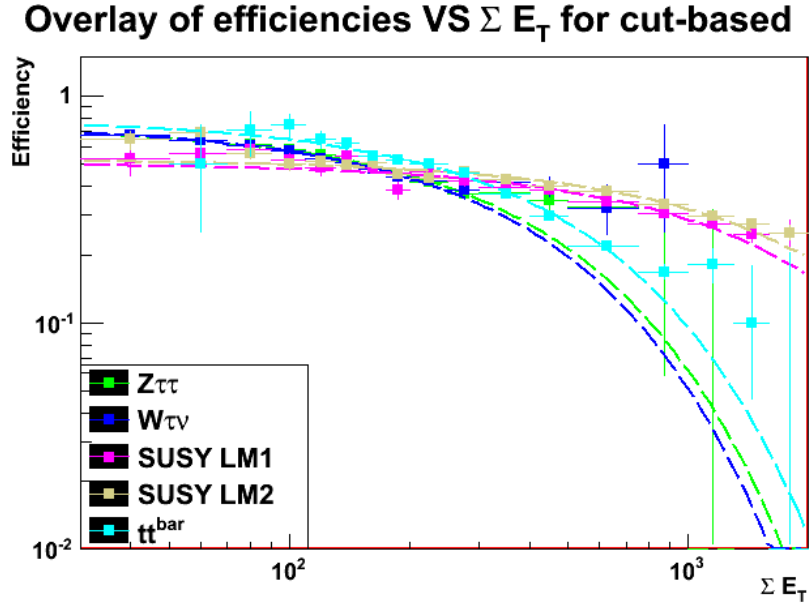


FIGURE 5.17 – Efficacités d'identification basée sur les coupures VS ΣE_T pour des candidats de $20 \text{ GeV} < P_T < 50 \text{ GeV}$, $|\eta| < 1$.

Les allures générales des courbes présentent quelques différences entre elles, on peut dégager de cette étude trois tendances vis-à-vis du paramètre ΣE_T :

- Les échantillons électrofaibles $Z\tau\tau$ et $W\tau\nu$ sont très similaires. Le paramètre a est très sensible au P_T des τ
- Les échantillons SUSY sont très similaires entre eux. Le paramètre a est peu sensible

TABLE 5.1 – Détails de la régression exponentielle $\epsilon(\Sigma E_T) = e^{a\Sigma E_T+b}$ de la Fig. 5.16 et de la Fig. 5.17.

Échantillon	a ($P_T \geq 50 \text{ GeV}$)	b ($P_T \geq 50 \text{ GeV}$)	a ($20 \leq P_T < 50 \text{ GeV}$)	b ($20 \leq P_T < 50 \text{ GeV}$)
$Z\tau\tau$	$-(7,4 \pm 0,9) \times 10^{-4}$	$-0,36 \pm 0,02$	$-(23,3 \pm 0,3) \times 10^{-4}$	$-0,326 \pm 0,004$
$W\tau\nu$	$-(7,6 \pm 0,9) \times 10^{-4}$	$-0,34 \pm 0,02$	$-(24,2 \pm 0,3) \times 10^{-4}$	$-0,324 \pm 0,003$
SUSY LM1	$-(6,1 \pm 0,4) \times 10^{-4}$	$-0,55 \pm 0,03$	$-(5,7 \pm 0,4) \times 10^{-4}$	$-0,68 \pm 0,02$
SUSY LM2	$-(5,2 \pm 0,2) \times 10^{-4}$	$-0,49 \pm 0,02$	$-(5,0 \pm 0,2) \times 10^{-4}$	$-0,65 \pm 0,02$
$t\bar{t}$	$-(14,6 \pm 0,6) \times 10^{-4}$	$-0,15 \pm 0,02$	$-(21,2 \pm 0,5) \times 10^{-4}$	$-0,23 \pm 0,02$
Moyenne	$-8,2 \times 10^{-4}$	$-0,38$	$-15,9 \times 10^{-4}$	$-0,64$
Écart type	$3,7 \times 10^{-4}$	$0,16$	$9,69 \times 10^{-4}$	$0,21$
Écart type relatif	45,7%	41,1%	61,1%	47,1%

au P_T des τ et a une valeur plus faible que dans les autres échantillons.

- L'échantillon $t\bar{t}$ a un comportement particulier. Le paramètre a est moyennement sensible au P_T des τ et a la valeur la plus élevée.

Les échantillons électrofaibles sont similaires aux échantillons SUSY pour les candidats de haut P_T et similaires à l'échantillon $t\bar{t}$ à bas P_T . Cependant la tendance observée est la même, l'efficacité décroît en fonction de ΣE_T pour les deux régions en P_T étudiées.

Le contenu de l'anneau d'isolation pour des candidats de bas P_T ($30 < P_T < 50 \text{ GeV}$) et haut P_T ($50 < P_T < 80 \text{ GeV}$) est présenté dans la figure 5.18 dans différentes gammes en ΣE_T , on constate que le contenu en hadrons chargés augmente lorsque ΣE_T augmente aussi, cependant les différents types d'événements n'ont pas la même sensibilité à ce paramètre.

• Influence de l'énergie transverse manquante :

On paramètre de façon similaire la variation de l'efficacité en fonction de l'énergie transverse manquante par une exponentielle de la forme $\epsilon(MET) = e^{a.MET+b}$ dans les figures 5.19 pour les τ de haute impulsion transverse ($P_T \geq 50 \text{ GeV}$) et la figure 5.20 pour les τ de basse impulsion transverse ($20 \leq P_T < 50 \text{ GeV}$). Les résultats du paramétrage sont présentés dans la table 5.2.

TABLE 5.2 – Détails de la régression exponentielle $\epsilon(MET) = e^{a.MET+b}$ de la Fig 5.19 et de la Fig 5.20.

Échantillon	a ($P_T \geq 50 \text{ GeV}$)	b ($P_T \geq 50 \text{ GeV}$)	a ($20 \leq P_T < 50 \text{ GeV}$)	b ($20 \leq P_T < 50 \text{ GeV}$)
$Z\tau\tau$	$+(0,2 \pm 3,2) \times 10^{-4}$	$-0,49 \pm 0,01$	$-(35,3 \pm 2,0) \times 10^{-4}$	$-0,533 \pm 0,004$
$W\tau\nu$	$-(0,2 \pm 2,2) \times 10^{-4}$	$-0,48 \pm 0,01$	$-(6,2 \pm 0,9) \times 10^{-4}$	$-0,530 \pm 0,003$
SUSY LM1	$-(4,9 \pm 1,1) \times 10^{-4}$	$-0,86 \pm 0,03$	$-(2,1 \pm 0,8) \times 10^{-4}$	$-1,02 \pm 0,02$
SUSY LM2	$-(6,4 \pm 0,5) \times 10^{-4}$	$-0,69 \pm 0,01$	$-(5,1 \pm 0,5) \times 10^{-4}$	$-0,88 \pm 0,01$
$t\bar{t}$	$-(6,0 \pm 1,7) \times 10^{-4}$	$-0,69 \pm 0,01$	$(11,7 \pm 1,3) \times 10^{-4}$	$-1,10 \pm 0,01$
Moyenne	$-3,5 \times 10^{-4}$	$-0,64$	$-12,1 \times 10^{-4}$	$-0,81$
Écart type	$3,2 \times 10^{-4}$	$0,16$	$13,4 \times 10^{-4}$	$0,27$
Écart type relatif	92,9%	25,1%	111%	33%

L'énergie transverse manquante influence peu ou pas l'identification du lepton tau. On remarque dans les figures 5.19 et 5.20 la décorrélation de l'énergie transverse manquante avec les autres variables globales avec les différences de niveau d'efficacités entre les échantillons électrofaibles W et Z et des échantillons présentant une activité hadronique plus impor-

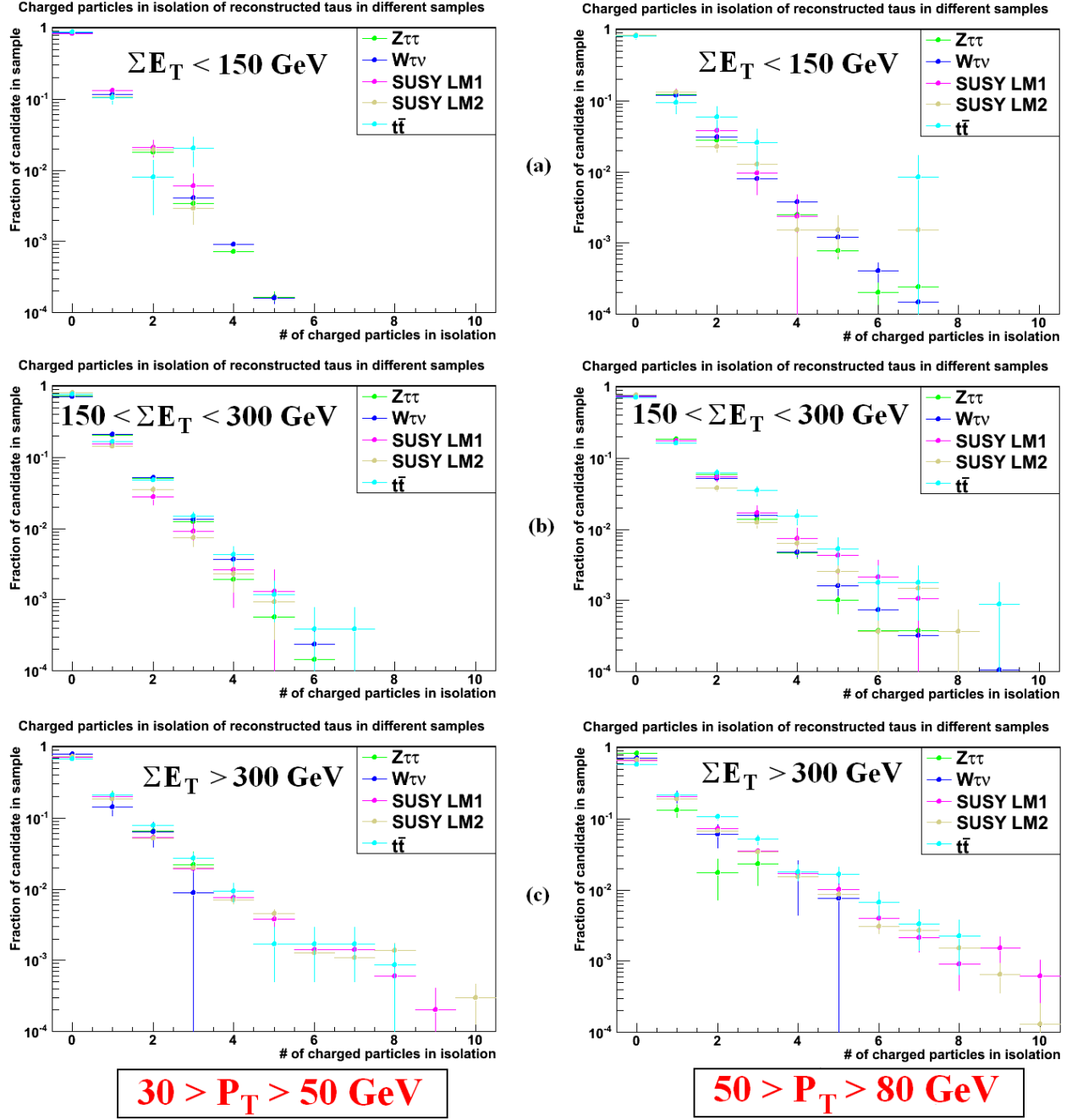


FIGURE 5.18 – Multiplicités en hadrons chargés ($P_T > 1 \text{ GeV}$) de l’anneau d’isolation pour les τ reconstruits (avant isolation) de $30 < P_T < 50 \text{ GeV}$ (à gauche) et $50 < P_T < 80 \text{ GeV}$ (à droite) dans des évènements de ΣE_T a) Inférieur à 150 GeV , b) Compris entre 150 et 300 GeV , c) Supérieur à 300 GeV .

tante : $t\bar{t}$ et SUSY.

On remarque aussi que le signe de la pente est mal défini et qu’il y a une grande disparité entre les échantillons (écart type relatif de 92,9% sur les pentes dans la Fig 5.19 et de 111% dans la Fig 5.20). En se plaçant dans l’intervalle $20 \text{ GeV} < P_T < 50 \text{ GeV}$ on constate aussi de plus grands écarts entre les divers échantillons. L’erreur sur le paramètre a de la

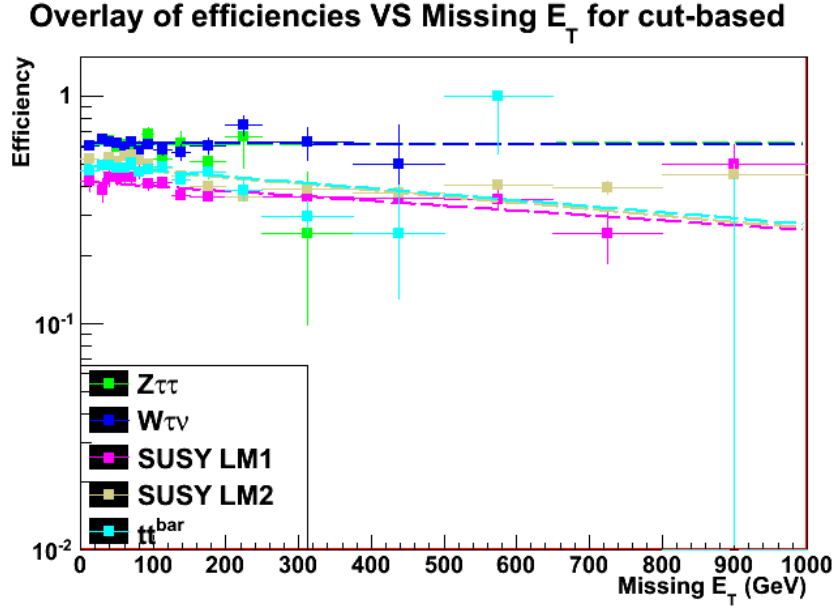


FIGURE 5.19 – Efficacités d'identification basée sur les coupures VS $E_{T,manquante}$ pour des candidats de $P_T > 50 \text{ GeV}$, $|\eta| < 1$.

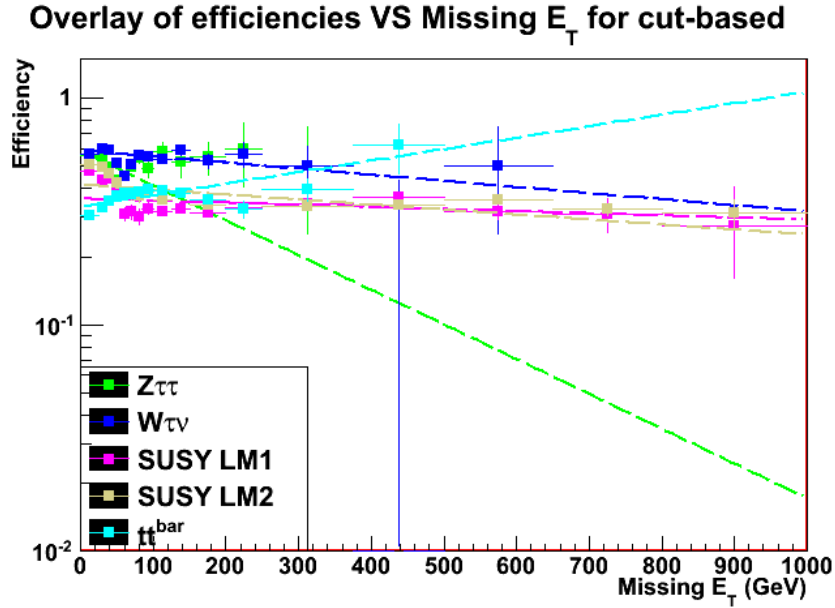


FIGURE 5.20 – Efficacités d'identification basée sur les coupures VS $E_{T,manquante}$ pour des candidats de $20 \text{ GeV} < P_T < 50 \text{ GeV}$, $|\eta| < 1$.

régression exponentielle et le fait que le signe ne soit pas défini correspond à un scénario où la pente est nulle, on en conclue alors que l'énergie transverse manquante est un paramètre qui n'influence pas l'identification basée sur les coupures.

5.2.4.2 Influence des multiplicités d'objets

L'identification basée sur les coupures est directement affectée par la multiplicité de hadrons chargés et de photons car ces particules supplémentaires ont une probabilité non négligeable d'échouer dans l'anneau d'isolation défini par cette méthode (dans le cône signal aussi, mais l'effet sera de fausser la mesure du jet).

Les figures 5.21, 5.25 et 5.26 nous montrent les niveaux d'efficacité dans les divers environnements en fonction de la multiplicité de divers objets : hadrons chargés, photons et jets (ayant une impulsion transverse de 10 GeV et 30 GeV). L'hypothèse testée ici est aussi celle d'un paramétrage par une régression exponentielle $\epsilon(N_{Objets}) = e^{aN_{Objets}+b}$, les résultats sont présentés dans la table 5.3 pour les hadrons chargés.

- Influence de la multiplicité de hadrons chargés :

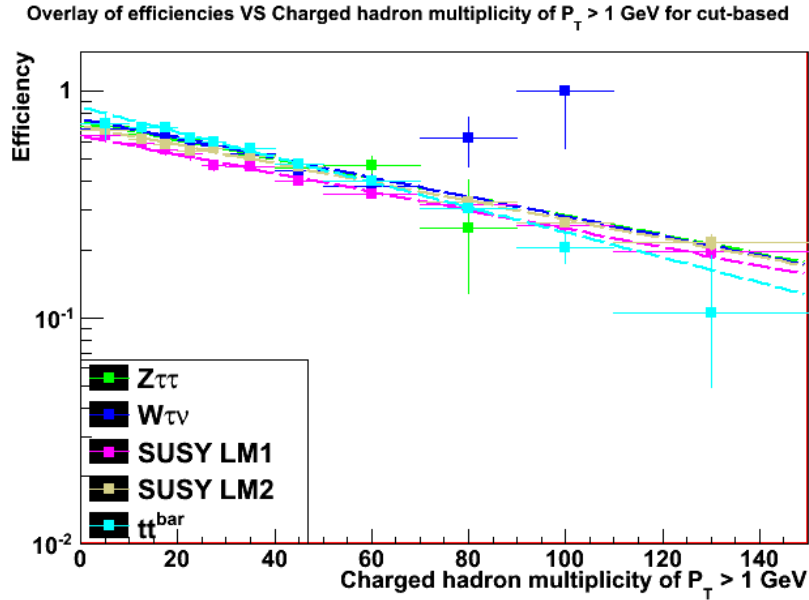


FIGURE 5.21 – Efficacités d'identification basée sur les coupures VS $N_{Hadrons\ Chargés}$ pour des candidats de $P_T > 50\text{ GeV}$, $|\eta| < 1$.

TABLE 5.3 – Détails de la régression exponentielle $\epsilon(N_{Chargés}) = e^{aN_{Chargés}+b}$ de la Fig 5.21 et de la Fig 5.22

Échantillon	a ($P_T \geq 50\text{ GeV}$)	b ($P_T \geq 50\text{ GeV}$)	a ($20 \leq P_T < 50\text{ GeV}$)	b ($20 \leq P_T < 50\text{ GeV}$)
$Z\tau\tau$	$-(9,6 \pm 0,73) \times 10^{-3}$	$-0,30 \pm 0,02$	$-(18,4 \pm 0,3) \times 10^{-3}$	$-0,359 \pm 0,004$
$W\tau\nu$	$-(9,9 \pm 0,68) \times 10^{-3}$	$-0,28 \pm 0,01$	$-(18,7 \pm 0,2) \times 10^{-3}$	$-0,328 \pm 0,002$
SUSY LM1	$-(9,4 \pm 0,56) \times 10^{-3}$	$-0,45 \pm 0,03$	$-(9,5 \pm 0,42) \times 10^{-3}$	$-0,55 \pm 0,02$
SUSY LM2	$-(9,5 \pm 0,29) \times 10^{-3}$	$-0,35 \pm 0,01$	$-(8,8 \pm 0,30) \times 10^{-3}$	$-0,54 \pm 0,02$
$t\bar{t}$	$-(12,7 \pm 0,50) \times 10^{-3}$	$-0,16 \pm 0,02$	$-(16,5 \pm 0,4) \times 10^{-3}$	$-0,31 \pm 0,02$
Moyenne	$-10,2 \times 10^{-3}$	-0,31	$-14,4 \times 10^{-3}$	-0,42
Écart type	$1,4 \times 10^{-3}$	0,11	$4,9 \times 10^{-3}$	0,12
Écart type relatif	13,6%	34,5%	33,8%	28,9%

L'efficacité en fonction de la multiplicité de hadrons chargés présente un écart type relatif léger entre les divers échantillons (13,6%), la multiplicité de hadrons chargés est un

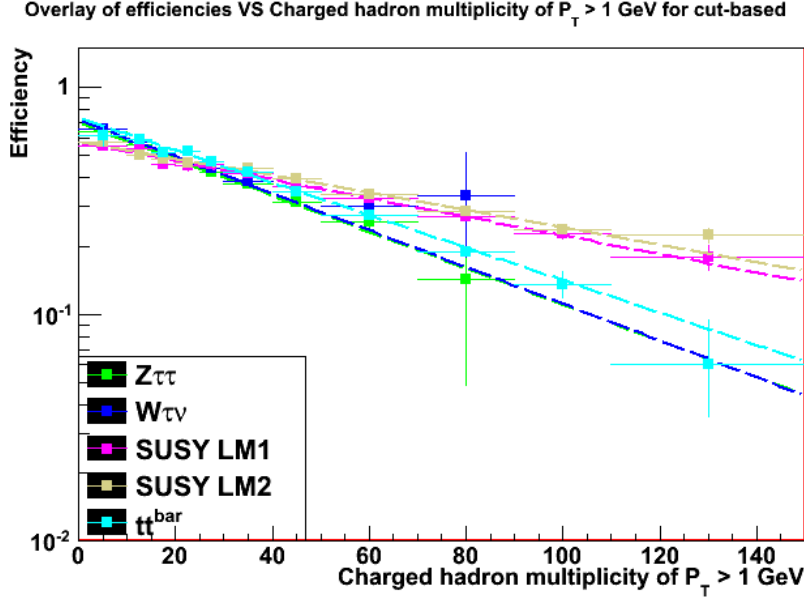


FIGURE 5.22 – Efficacités d'identification basée sur les coupures VS $N_{\text{Hadrons Chargés}}$ pour des candidats de $20 \text{ GeV} < P_T < 50 \text{ GeV}$, $|\eta| < 1$.

paramètre adapté à tous types d'échantillons sous l'identification basée sur les coupures. Nous vérifions dans la figure 5.22 que ce paramétrage est toujours valide dans l'intervalle $20 \text{ GeV} < P_T < 50 \text{ GeV}$.

On constate un écart type relatif plus important pour les jets de $20 \text{ GeV} < P_T < 50 \text{ GeV}$ que pour $P_T > 50 \text{ GeV}$ pour les pentes entre les différents échantillons. Cependant, le paramétrage en fonction de la multiplicité de hadrons chargés donne de meilleurs résultats que celle en fonction de ΣE_T pour les deux régions en P_T étudiées.

La figure 5.23 nous présente les efficacités en fonction de P_T et η dans différents intervalles en multiplicité de hadrons chargés, on voit dans cette figure que la baisse d'efficacité est généralisée pour les candidats de tous P_T et η lorsque la multiplicité de hadrons chargés augmente. Cette baisse est indépendante du processus physique étudié, il s'agit donc bien d'un effet d'environnement caractérisé par la multiplicité de hadrons chargés.

Le contenu de l'anneau d'isolation des candidats sous différentes multiplicités globales de hadrons chargés est présenté dans la figure 5.24. La multiplicité de hadrons chargés dans l'anneau d'isolation augmente avec la multiplicité globale. Les divergences entre les différents types d'événements sont moins prononcées que celles observées pour ΣE_T (c.f. 5.18), on constate effet que la fraction de candidats n'ayant aucun hadron chargé dans l'anneau d'isolation est similaire pour tous les échantillons dans un intervalle en $N_{\text{chargé}}$ donné. Cette fraction diminue lorsque $N_{\text{chargé}}$ augmente, ce qui confirme l'hypothèse de particules de l'environnement échouant dans l'anneau d'isolation.

La multiplicité globale de hadrons chargés est un très bon paramétrage de l'environnement pour l'identification du lepton τ et pour l'algorithme basé sur les coupures. Nous verrons plus en détail la validité de ce paramétrage pour les autres algorithmes et pour les taux

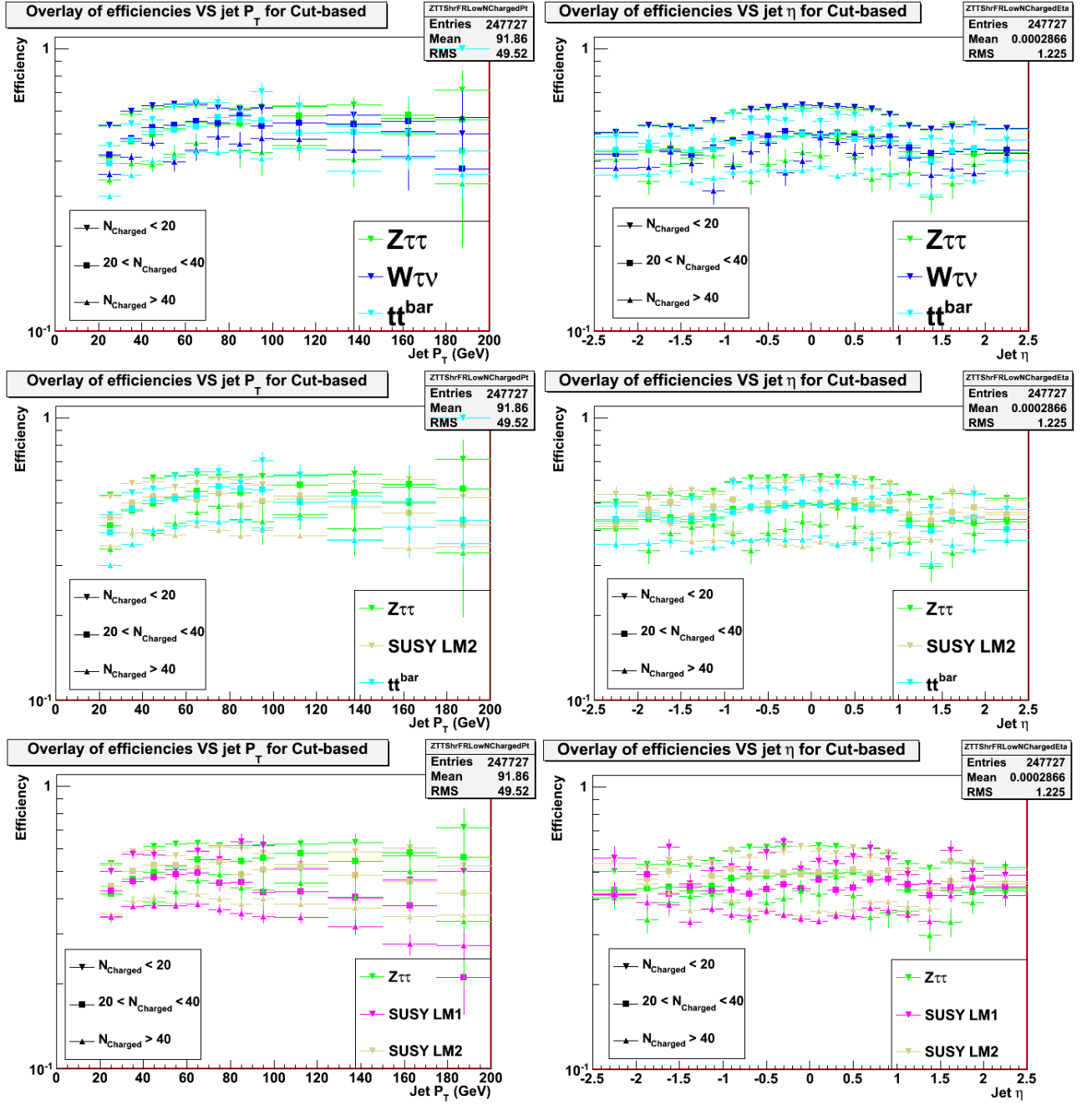


FIGURE 5.23 – Efficacités d'identification en fonction de P_T (à gauche) et η (à droite) dans différents intervalles en multiplicité de hadrons chargés.

de mauvaise identification dans les sections suivantes.

• Influence de la multiplicité de photons :

Nous vérifions au vu de la corrélation entre la multiplicité de hadrons chargés et celle de photons (c.f. 5.15) si le paramétrage de l'efficacité en fonction de la multiplicité de photons donne des résultats aussi satisfaisants. Les résultats sont présentés dans la Fig 5.25.

L'écart type relatif sur la pente est de 30,7% quand on se situe dans l'intervalle de

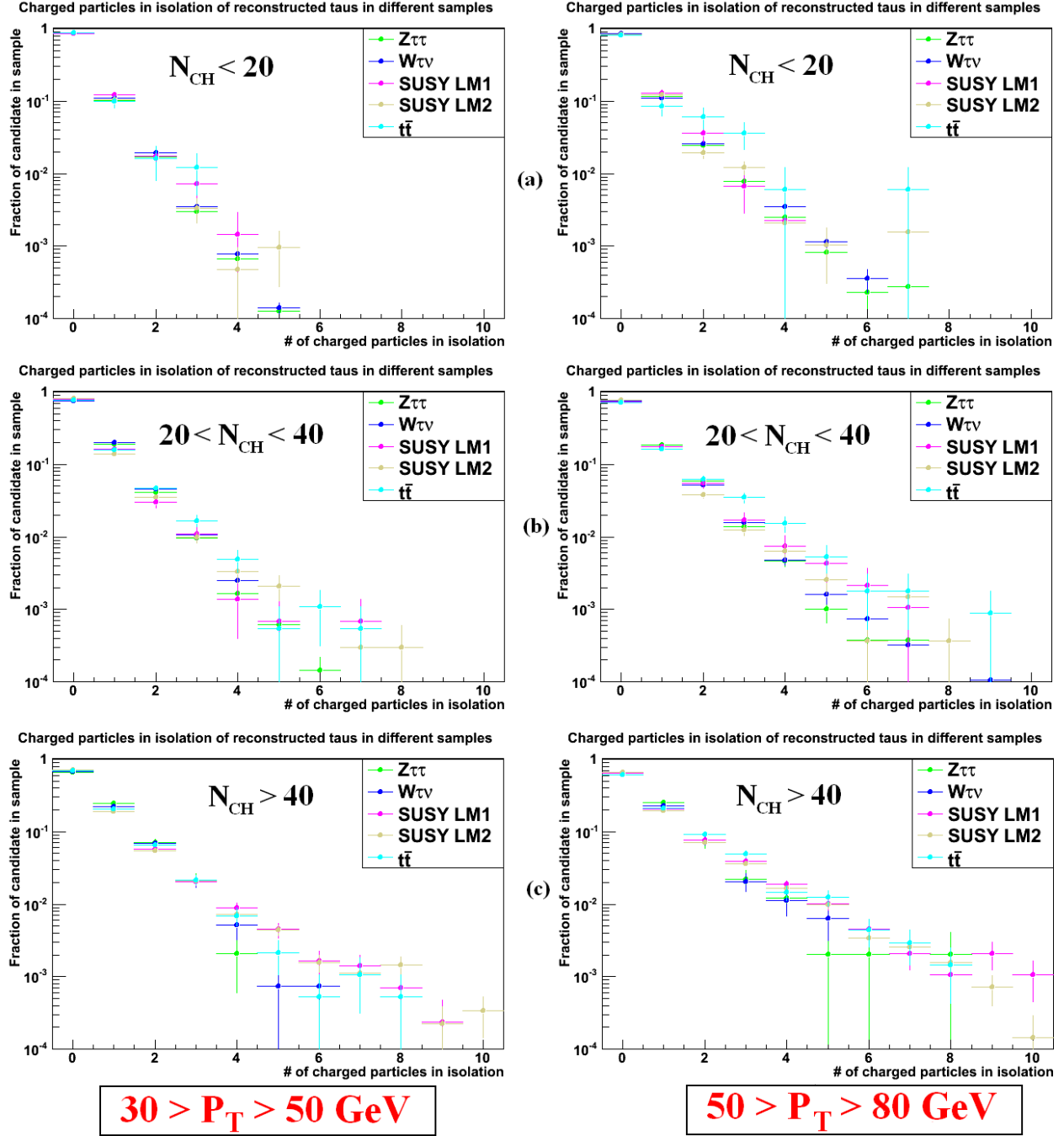


FIGURE 5.24 – Multiplicités en hadrons chargés ($P_T > 1 \text{ GeV}$) de l’anneau d’isolation pour les τ reconstruits (avant isolation) de $30 < P_T < 50 \text{ GeV}$ (à droite) et $50 < P_T < 80 \text{ GeV}$ (à gauche) dans des évènements avec des multiplicité de hadrons chargés a) Inférieure a 20, b) Comprise entre 20 et 40, c) Supérieure à 40.

$P_T > 50 \text{ GeV}$, c’est un écart plus important que celui mesuré pour la multiplicité de hadrons chargés, mais les tendances sur a sont plus proches entre les différents échantillons que celles données en fonction de ΣE_T (qui présentent un écart type relatif de 45,7%).

- Influence de la multiplicité de jets :

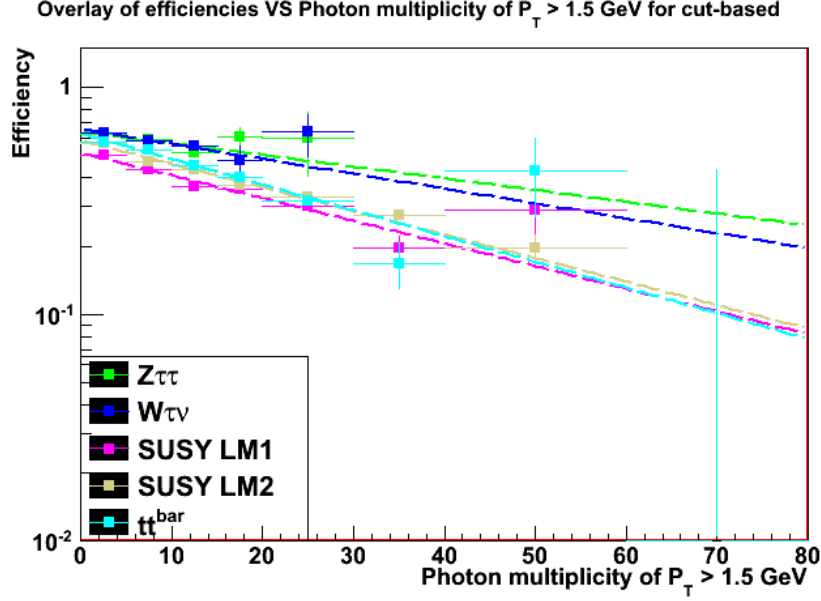


FIGURE 5.25 – Efficacités d'identification basée sur les coupures VS $N_{Photons}$ pour des candidats de $P_T > 50 \text{ GeV}$, $|\eta| < 1$.

TABLE 5.4 – Détails de la régression exponentielle $\epsilon(N_\gamma) = e^{aN_\gamma+b}$ de la Fig 5.25

Échantillon	a ($P_T \geq 50 \text{ GeV}$)	b ($P_T \geq 50 \text{ GeV}$)
$Z\tau\tau$	$-(11,9 \pm 2,9) \times 10^{-3}$	$-0,45 \pm 0,01$
$W\tau\nu$	$-(15,1 \pm 2,9) \times 10^{-3}$	$-0,42 \pm 0,01$
SUSY LM1	$-(23,0 \pm 1,8) \times 10^{-3}$	$-0,66 \pm 0,03$
SUSY LM2	$-(23,9 \pm 1,0) \times 10^{-3}$	$-0,54 \pm 0,01$
$t\bar{t}$	$-(26,1 \pm 1,5) \times 10^{-3}$	$-0,46 \pm 0,02$
Moyenne	$-20,0 \times 10^{-3}$	$-0,51$
Écart type	$6,14 \times 10^{-3}$	$0,097$
Écart type relatif	30,7%	19,2%

On vérifie ici l'influence de la dernière variable caractérisant l'environnement de l'évènement : la multiplicité de jets qui est présentée dans les Fig. 5.26a) et b).

L'erreur relative sur la valeur de chaque pente (dépassant les 100% pour $W\tau\nu$) ainsi que l'écart type relatif sont trop importants pour pouvoir conclure à un paramétrage en fonction de la multiplicité de jets. Nous observons des tendances similaires à celles déterminées pour les efficacités en fonction de l'énergie transverse manquante : on a le scenario d'une pente nulle signifiant que les efficacités d'identification (pour la méthode cut-based) sont peu affectées par la multiplicité de jets.

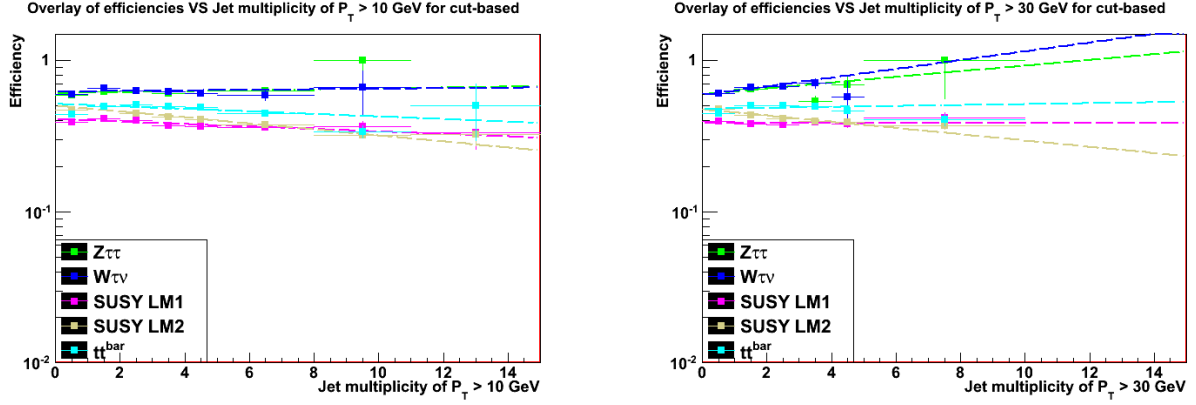


FIGURE 5.26 – Efficacités d’identification basée sur les coupures VS N_{Jets} pour des candidats de $P_T > 50 \text{ GeV}$, $|\eta| < 1$ VS a) La multiplicité de jets de plus de $P_T > 10 \text{ GeV}$, b) La multiplicité de jets de $P_T > 30 \text{ GeV}$.

TABLE 5.5 – Détails de la régression exponentielle $\epsilon(N_{Jets,10\text{GeV}}) = e^{aN_{Jets,10\text{GeV}}+b}$ de la Fig 5.26

Échantillon	a ($P_T \geq 50 \text{ GeV}$)	b ($P_T \geq 50 \text{ GeV}$)
$Z\tau\tau$	$-(7,0 \pm 6,0) \times 10^{-3}$	$-0,50 \pm 0,01$
$W\tau\nu$	$-(4,0 \pm 6,0) \times 10^{-3}$	$-0,47 \pm 0,01$
SUSY LM1	$-(18,8 \pm 6,2) \times 10^{-3}$	$-0,89 \pm 0,03$
SUSY LM2	$-(45,8 \pm 3,7) \times 10^{-3}$	$-0,68 \pm 0,02$
$t\bar{t}$	$-(19,4 \pm 4,3) \times 10^{-3}$	$-0,66 \pm 0,02$
Moyenne	$-19,0 \times 10^{-3}$	$-0,64$
Écart type	$16,5 \times 10^{-3}$	$0,17$
Écart type relatif	86,7%	26,4%

5.2.5 Influence de l’environnement sur l’identification Hadron Plus Strips

Les variables caractérisant l’évènement décrites dans la section 5.1.1. sont calculées à l’aide de l’information Particle Flow et les jets sont reconstruits par l’algorithme anti-kt (c.f. section 2.3.5). L’algorithme d’identification du lepton τ intervient après ces reconstructions, un changement d’algorithme d’identification ne modifie donc pas les mesures de ces variables d’environnement. Les spectres de ces variables présentés dans la section 5.2.1 sont aussi valables pour l’algorithme HPS.

Pour cette étude c’est l’identification “relâchée” de l’algorithme HPS qui est utilisée, les caractéristiques sont citées dans la section 3.3.2. Les seuils en impulsion transverse choisis pour rejeter un candidat τ avec ce niveau d’isolation sont identiques à ceux de l’identification basée sur les coupures. La présélection appliquée sur les hadrons chargés ($P_T > 1 \text{ GeV}$) et les photons ($P_T > 1,5 \text{ GeV}$) pour le calcul de la multiplicité reste valide.

5.2.5.1 Influence des variables caractérisant l'énergie globale

Nous tenterons ici aussi de modéliser les variations d'efficacité par une exponentielle $\epsilon_{HPS} = e^{aX_G+b}$ (X_G est l'une des variables globales vue dans la sous section précédente).

- Influence de ΣE_T :

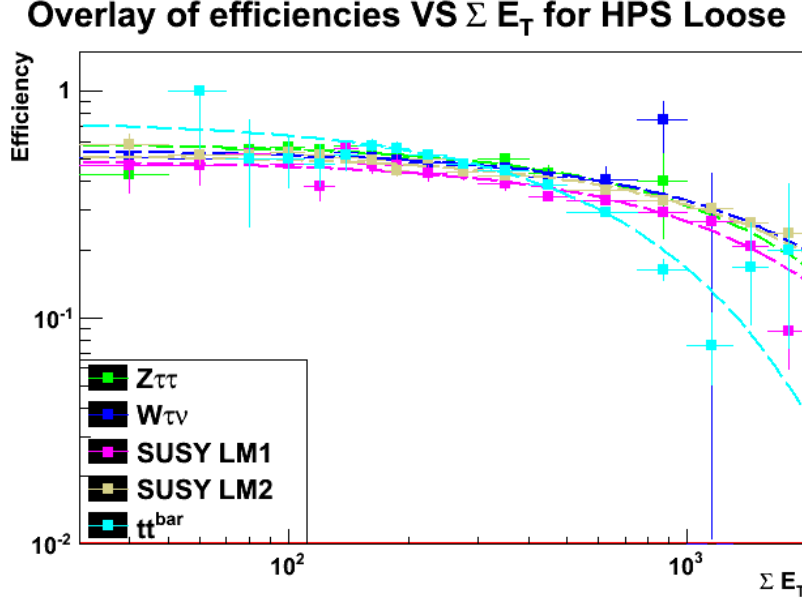


FIGURE 5.27 – Efficacités d'identification Hadron Plus Strips VS ΣE_T pour des candidats de $P_T > 50 \text{ GeV}$, $|\eta| < 1$.

TABLE 5.6 – Détails de la régression exponentielle $\epsilon(\Sigma E_T) = e^{a\Sigma E_T+b}$ de la Fig 5.27 et de la Fig 5.28

Échantillon	a ($P_T \geq 50 \text{ GeV}$)	b ($P_T \geq 50 \text{ GeV}$)	a ($20 \leq P_T < 50 \text{ GeV}$)	b ($20 \leq P_T < 50 \text{ GeV}$)
$Z\tau\tau$	$-(6,2 \pm 1,1) \times 10^{-4}$	$-0,53 \pm 0,02$	$-(21,5 \pm 0,4) \times 10^{-4}$	$-0,419 \pm 0,004$
$W\tau\nu$	$-(5,0 \pm 1,0) \times 10^{-4}$	$-0,60 \pm 0,02$	$-(21,7 \pm 0,3) \times 10^{-4}$	$-0,486 \pm 0,003$
SUSY LM1	$-(6,1 \pm 0,5) \times 10^{-4}$	$-0,71 \pm 0,03$	$-(5,7 \pm 0,4) \times 10^{-4}$	$-0,75 \pm 0,02$
SUSY LM2	$-(5,0 \pm 0,2) \times 10^{-4}$	$-0,65 \pm 0,02$	$-(4,6 \pm 0,2) \times 10^{-4}$	$-0,74 \pm 0,02$
$t\bar{t}$	$-(15,0 \pm 0,7) \times 10^{-4}$	$-0,30 \pm 0,03$	$-(21,1 \pm 0,6) \times 10^{-4}$	$-0,31 \pm 0,02$
Moyenne	$-7,5 \times 10^{-4}$	$-0,6$	$-14,9 \times 10^{-4}$	$-0,54$
Écart type	$4,2 \times 10^{-4}$	$0,2$	$8,9 \times 10^{-4}$	$0,20$
Écart type relatif	56,8%	28,2%	59,7%	36,5%

La variable ΣE_T (Fig. 5.27) a une influence sur l'identification HPS similaire à celle observée pour l'identification basée sur les coupures. L'écart type relatif évalué est plus important pour l'identification HPS (56,8% contre 45,7% pour l'identification "cut-based"). La tendance globale est cependant la même : une baisse d'efficacité modélisable par une exponentielle en fonction de ΣE_T et les trois régimes observés dans la section 5.2.4.1 (Fig. 5.16 et 5.17) pour les échantillons électrofaibles, les échantillons SUSY, et une divergence plus forte présente pour l'échantillon $t\bar{t}$.

La même étude pour l'intervalle en P_T compris entre 20 et 50 GeV est présentée par la figure 5.28 et la table 5.6. On observe pour cet intervalle le même comportement de

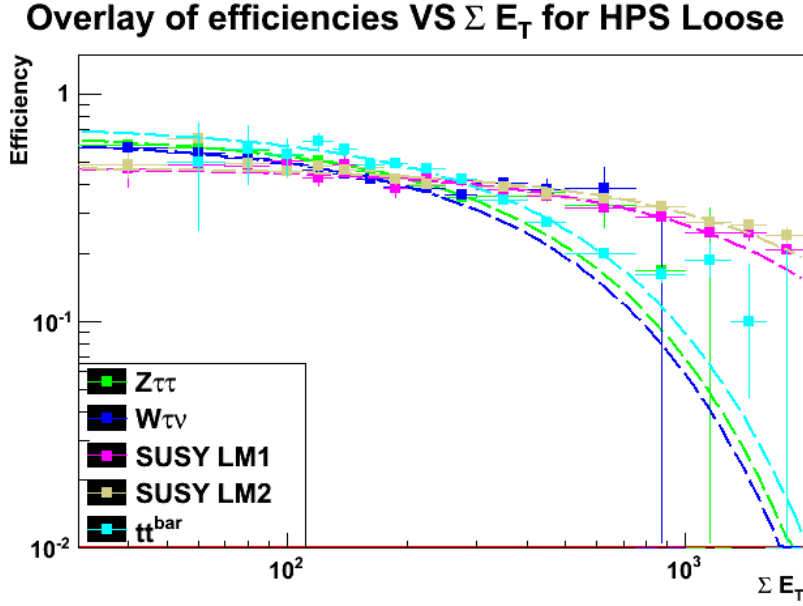


FIGURE 5.28 – Efficacités d'identification Hadron Plus Strips VS ΣE_T pour des candidats de $20 \text{ GeV} < P_T < 50 \text{ GeV}$, $|\eta| < 1$.

l'algorithme HPS face à la somme des énergies transverses ΣE_T que celui observé pour l'algorithme basé sur les coupures : la pente décroissante devient plus importante pour les échantillons électrofaibles pour atteindre une sensibilité dépassant celle de l'échantillon $t\bar{t}$. L'écart type relatif est aussi plus faible entre les différents échantillons que celui évalué pour la méthode basée sur les coupures : l'algorithme HPS est moins sensible à l'impulsion transverse (pour $P_T > 20 \text{ GeV}$, c.f. Fig. 4.12 et 4.14) que l'est celui basé sur les coupures.

L'identification basée sur les coupures n'était pas sensible à l'énergie transverse manquante, nous revérifions qualitativement que cette tendance est similaire pour l'algorithme HPS (Fig. 5.29).

Nous observons dans cette figure 5.29 des pentes très faibles et dont le signe n'est pas défini, nous en arrivons donc à la conclusion que nous avons tiré pour l'algorithme basé sur les coupures : l'énergie transverse manquante n'influence pas l'identification du lepton τ .

5.2.5.2 Influence des multiplicités d'objets

L'algorithme Hadron Plus Strips est affecté par des particules de l'environnement qui aboutissent dans le jet constituant notre candidat tau. Ces particules supplémentaires faussent la reconstruction du mode de désintégration et de plus l'isolation d'un candidat se fait sur son contenu en particule auquel on a ôté celles correspondant au mode de désintégration identifié entraînant un rejet supplémentaire dudit candidat. Nous effectuons l'étude des efficacités d'identification de τ en fonction de diverses multiplicités d'objets (hadrons chargés, photons et jets). Nous nous intéresserons particulièrement à la multiplicité de hadrons chargés afin de nous rapprocher de l'étude faite en fonction de cette multiplicité pour l'algorithme basé sur les coupures qui nous a donné le meilleur

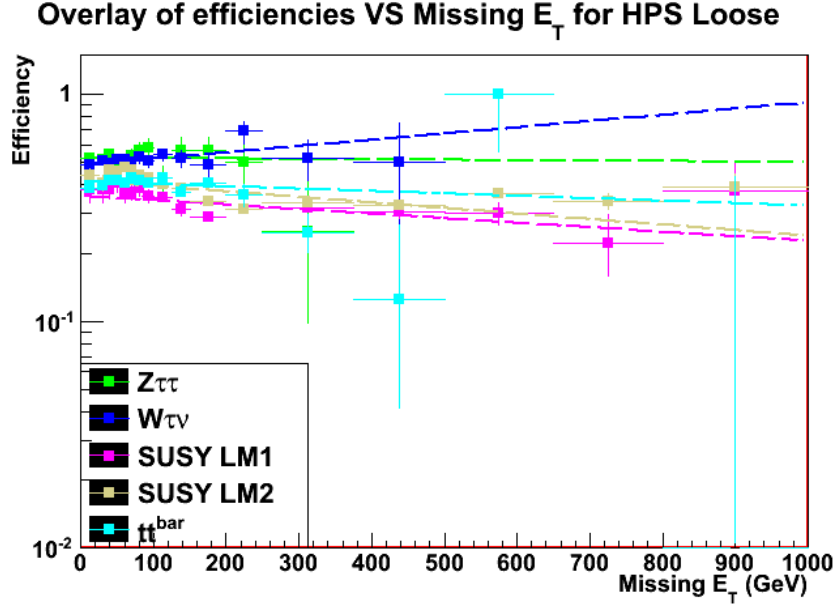


FIGURE 5.29 – Efficacités d'identification Hadron Plus Strips VS MET pour des candidats de $P_T > 50 \text{ GeV}$, $|\eta| < 1$.

paramétrage.

• Influence de la multiplicité de hadrons chargés :

Nous avons vu dans la section 5.2.4.2 que la multiplicité de hadrons chargés était le meilleur paramétrage de l'environnement des événements. Nous vérifions la validité de ce paramétrage pour l'algorithme HPS (Figures 5.30 et 5.31).

TABLE 5.7 – Détails de la régression exponentielle $\epsilon(N_{\text{Chargés}}) = e^{aN_{\text{Chargés}}+b}$ de la Fig 5.30 et de la Fig 5.31

Échantillon	a ($P_T \geq 50 \text{ GeV}$)	b ($P_T \geq 50 \text{ GeV}$)	a ($20 \leq P_T < 50 \text{ GeV}$)	b ($20 \leq P_T < 50 \text{ GeV}$)
$Z\tau\tau$	$-(8,9 \pm 0,9) \times 10^{-3}$	$-0,47 \pm 0,02$	$-(16,7 \pm 0,3) \times 10^{-3}$	$-0,453 \pm 0,004$
$W\tau\nu$	$-(8,9 \pm 0,9) \times 10^{-3}$	$-0,51 \pm 0,02$	$-(16,5 \pm 0,2) \times 10^{-3}$	$-0,491 \pm 0,003$
SUSY LM1	$-(9,3 \pm 0,6) \times 10^{-3}$	$-0,61 \pm 0,03$	$-(9,2 \pm 0,4) \times 10^{-3}$	$-0,63 \pm 0,02$
SUSY LM2	$-(9,4 \pm 0,3) \times 10^{-3}$	$-0,51 \pm 0,02$	$-(8,2 \pm 0,3) \times 10^{-3}$	$-0,64 \pm 0,02$
$t\bar{t}$	$-(12,9 \pm 0,6) \times 10^{-3}$	$-0,32 \pm 0,03$	$-(16,5 \pm 0,4) \times 10^{-3}$	$-0,38 \pm 0,02$
Moyenne	$-9,9 \times 10^{-3}$	$-0,48$	$-13,4 \times 10^{-3}$	$-0,52$
Écart type	$1,7 \times 10^{-3}$	$0,11$	$4,30 \times 10^{-3}$	$0,11$
Écart type relatif	17,2%	22,1%	32%	21,8%

Le paramétrage en fonction de la multiplicité de hadrons chargés est excellente pour l'algorithme HPS, de la même manière qu'elle l'était pour l'algorithme basé sur les coupures en présentant un écart-type relatif de 17,2% sur les pentes a des différents types d'événements. On observe aussi la déviation entre les trois types d'événements : électrofaibles, SUSY et $t\bar{t}$, un comportement observé pour l'algorithme basé sur les coupures. Les résultats sur le paramètre a sont très similaires entre les deux algorithmes comme le montre la table 5.8.

• Influence de la multiplicité de photons :

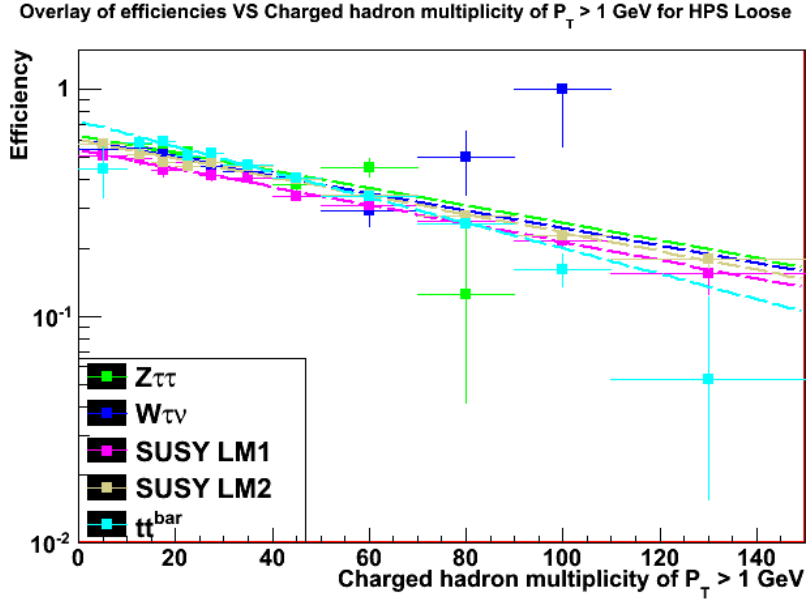


FIGURE 5.30 – Efficacités d'identification Hadron Plus Strips VS $N_{Hadrons\ Chargés}$ pour des candidats de $P_T > 50\ GeV$, $|\eta| < 1$.

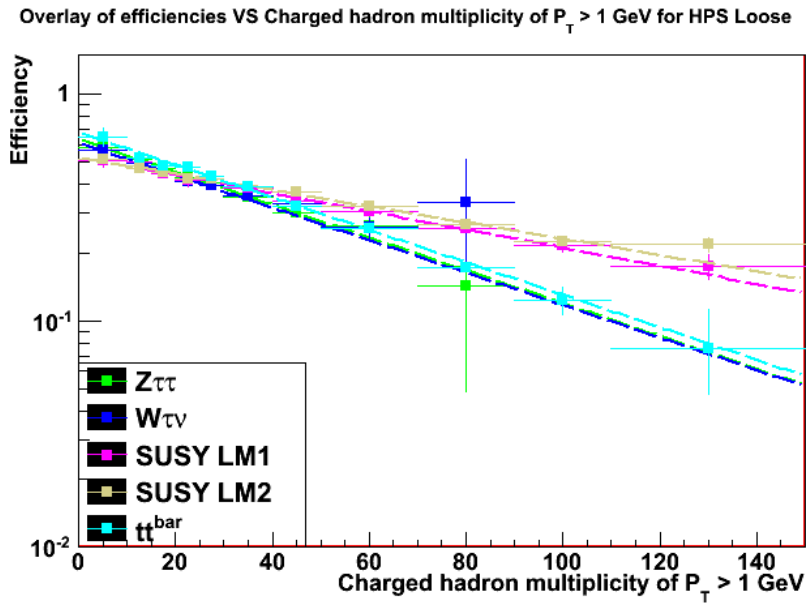


FIGURE 5.31 – Efficacités d'identification Hadron Plus Strips VS $N_{Hadrons\ Chargés}$ pour des candidats de $20\ GeV < P_T < 50\ GeV$, $|\eta| < 1$.

La multiplicité de photons est corrélée à celle des hadrons chargés, nous paramétrons dans la figure 5.32 la variation d'efficacité en fonction de la multiplicité de photons, les résultats sont présentés dans la table 5.9.

On retrouve dans cette étude les tendances observées dans la section 5.2.4.2 : une différence de pente entre les échantillons électrofaibles et les autres échantillons (SUSY et $t\bar{t}$), et un écart type plus important entre les pentes que celui observé dans les paramétrages avec la

TABLE 5.8 – Comparaison des valeurs de a du paramétrage $\epsilon(N_{Chargés}) = e^{aN_{Chargés}+b}$ pour l'algorithme basé sur les coupures et l'algorithme HPS

Échantillon	a ($P_T \geq 50 \text{ GeV}$) cut-based	a ($P_T \geq 50 \text{ GeV}$) HPS	a ($20 \leq P_T < 50 \text{ GeV}$) cut-based	a ($20 \leq P_T < 50 \text{ GeV}$) HPS
$Z\tau\tau$	$-(9,6 \pm 0,7) \times 10^{-3}$	$-(8,9 \pm 0,9) \times 10^{-3}$	$-(18,4 \pm 0,3) \times 10^{-3}$	$-(16,7 \pm 0,3) \times 10^{-3}$
$W\tau\nu$	$-(9,9 \pm 0,7) \times 10^{-3}$	$-(8,9 \pm 0,9) \times 10^{-3}$	$-(18,7 \pm 0,2) \times 10^{-3}$	$-(16,5 \pm 0,2) \times 10^{-3}$
SUSY LM1	$-(9,4 \pm 0,6) \times 10^{-3}$	$-(9,3 \pm 0,6) \times 10^{-3}$	$-(9,5 \pm 0,4) \times 10^{-3}$	$-(9,2 \pm 0,4) \times 10^{-3}$
SUSY LM2	$-(9,5 \pm 0,3) \times 10^{-3}$	$-(9,4 \pm 0,3) \times 10^{-3}$	$-(8,8 \pm 0,3) \times 10^{-3}$	$-(8,2 \pm 0,3) \times 10^{-3}$
$t\bar{t}$	$-(12,7 \pm 0,5) \times 10^{-3}$	$-(12,9 \pm 0,6) \times 10^{-3}$	$-(16,5 \pm 0,4) \times 10^{-3}$	$-(16,5 \pm 0,4) \times 10^{-3}$

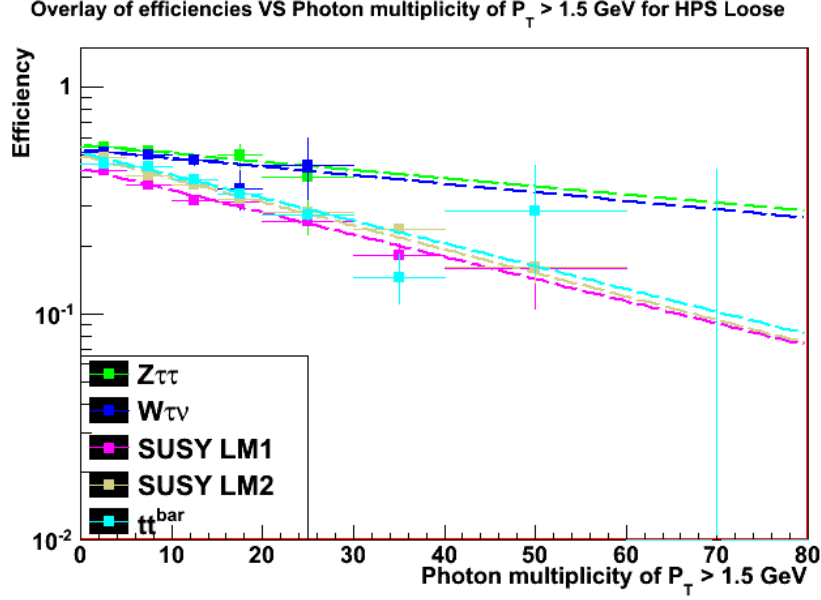


FIGURE 5.32 – Efficacités d'identification Hadron Plus Strips VS $N_{photons}$ pour des candidats de $P_T > 50 \text{ GeV}$, $|\eta| < 1$.

TABLE 5.9 – Détails de la régression exponentielle $\epsilon(N_\gamma) = e^{aN_\gamma+b}$ de la Fig 5.32

Échantillon	a ($P_T \geq 50 \text{ GeV}$)	b ($P_T \geq 50 \text{ GeV}$)
$Z\tau\tau$	$-(8,3 \pm 3,2) \times 10^{-3}$	$-0,59 \pm 0,02$
$W\tau\nu$	$-(8,6 \pm 3,4) \times 10^{-3}$	$-0,63 \pm 0,02$
SUSY LM1	$-(22,6 \pm 2,0) \times 10^{-3}$	$-0,82 \pm 0,03$
SUSY LM2	$-(23,9 \pm 1,1) \times 10^{-3}$	$-0,69 \pm 0,02$
$t\bar{t}$	$-(23,1 \pm 1,7) \times 10^{-3}$	$-0,66 \pm 0,02$
Moyenne	$-17,3 \times 10^{-3}$	$-0,68$
Écart type	$8,1 \times 10^{-3}$	$0,09$
Écart type relatif	46,5%	12,8%

multiplicité de hadrons chargés.

• Influence de la multiplicité de jets :

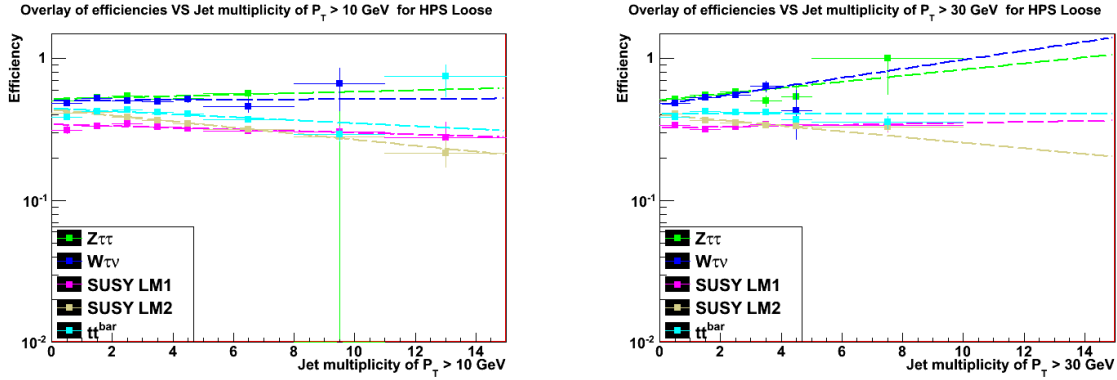


FIGURE 5.33 – Efficacités d'identification Hadron Plus Strips VS N_{Jets} pour des candidats de $P_T > 50 \text{ GeV}$, $|\eta| < 1$ VS a) La multiplicité de jets de plus de $P_T > 10 \text{ GeV}$, b) La multiplicité de jets de $P_T > 30 \text{ GeV}$.

TABLE 5.10 – Détails de la régression exponentielle $\epsilon(N_{Jets10}) = e^{aN_{Jets10}+b}$ de la Fig 5.33

Échantillon	a ($P_T \geq 50 \text{ GeV}$)	b ($P_T \geq 50 \text{ GeV}$)
$Z\tau\tau$	$+(12,0 \pm 7,1) \times 10^{-3}$	$-0,66 \pm 0,02$
$W\tau\nu$	$+(2,2 \pm 7,7) \times 10^{-3}$	$-0,68 \pm 0,02$
SUSY LM1	$-(13,5 \pm 6,9) \times 10^{-3}$	$-1,07 \pm 0,03$
SUSY LM2	$-(48,9 \pm 4,2) \times 10^{-3}$	$-0,83 \pm 0,02$
$t\bar{t}$	$-(24,1 \pm 5,0) \times 10^{-3}$	$-0,81 \pm 0,02$
Moyenne	$-14,5 \times 10^{-3}$	$-0,812$
Écart type	$23,8 \times 10^{-3}$	$0,162$
Écart type relatif	164%	20%

De la même manière qu'avec la méthode basée sur les coupures, la multiplicité de jets n'a pas d'influence sur les efficacités d'identification par la méthode HPS : le signe de la pente n'est pas défini, et les erreurs sur le paramètre a sont cohérents avec le scénario où la pente est nulle.

5.2.6 Influence de l'environnement sur l'identification Tau Neural Classifier

Les études menées à l'aide de l'algorithme basé sur les coupures et l'algorithme HPS ont montré que les variables les plus aptes à caractériser l'impact de l'environnement sur l'identification du lepton τ sont :

- La somme des énergies transverses ΣE_T dans le cas où l'on veut décrire l'évènement en terme d'énergie.
- La multiplicité de hadrons chargés dans le cas où l'on veut décrire l'évènement en terme d'occupation du détecteur.

Nous voulons vérifier que les paramétrages déterminés précédemment sont toujours valides pour l'algorithme Tau Neural Classifier.

5.2.6.1 Influence des variables caractérisant l'énergie globale

Les réseaux de neurones ont été entraînés de façon à suivre des scénarii identifiant les leptons τ avec des taux de faux de 0,1% ; 0,25% ; 0,5% et 1,0%. Les niveaux de taux de faux de l'identification basée sur les coupures est d'environ 1%, aussi c'est cette sélection qui est appliquée pour les études de l'identification TaNC dans les figures 5.34 et 5.35.

- Influence de ΣE_T :

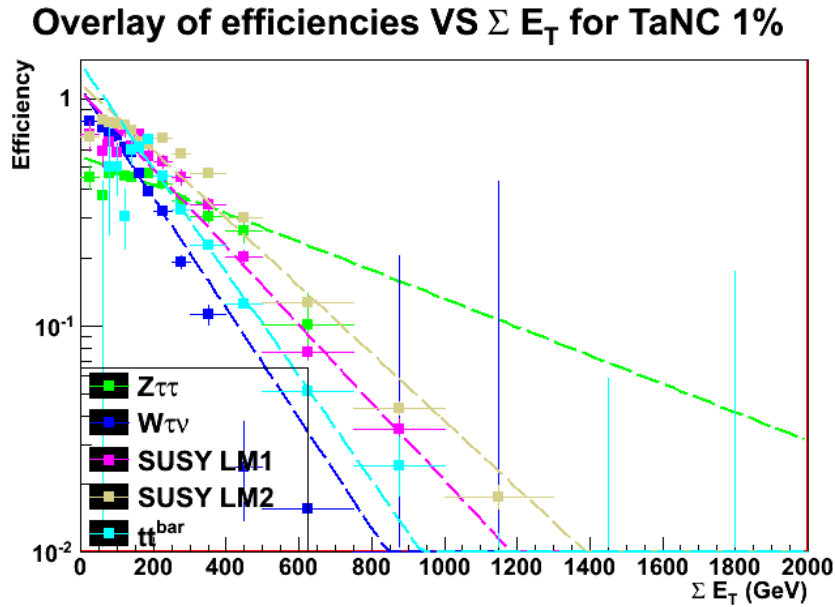


FIGURE 5.34 – Efficacités d'identification Tau Neural Classifier VS ΣE_T pour des candidats de $P_T > 50 \text{ GeV}$, $|\eta| < 1$.

Nous observons ici une chute dramatique de l'efficacité d'identification avec l'augmentation de ΣE_T ce qui signifie que l'algorithme TaNC est très sensible à l'environnement de

TABLE 5.11 – Détails de la régression exponentielle $\epsilon(\Sigma E_T) = e^{a\Sigma E_T + b}$ de la Fig 5.34

Échantillon	a ($P_T \geq 50 \text{ GeV}$)	b ($P_T \geq 50 \text{ GeV}$)
$Z\tau\tau$	$-(14,4 \pm 1,2) \times 10^{-4}$	$-0,58 \pm 0,02$
$W\tau\nu$	$-(5,6 \pm 1,1) \times 10^{-4}$	$+0,11 \pm 0,02$
SUSY LM1	$-(39,6 \pm 0,9) \times 10^{-4}$	$+0,83 \pm 0,03$
SUSY LM2	$-(34,3 \pm 0,4) \times 10^{-4}$	$+0,16 \pm 0,01$
$t\bar{t}$	$-(52,9 \pm 0,1) \times 10^{-4}$	$-0,36 \pm 0,04$
Moyenne	$-29,4 \times 10^{-4}$	$+0,03$
Écart type	$19,2 \times 10^{-4}$	$0,55$
Écart type relatif	65,3%	1589%

l'évènement. Ce comportement est très différent des deux autres algorithmes étudiés.

Les variations sur la pente a de l'exponentielle sont importantes : l'écart type relatif est de 58,4%, qui est la variation la plus forte observée. De plus la moyenne sur la constante b ainsi que son écart type relatif extrêmement important est le scénario d'une constante nulle, ce qui est physiquement impossible car l'efficacité d'identification ne peut pas être égale à 1 (c.f. section 4.2).

On peut donc poser l'hypothèse suivante concernant le Tau Neural Classifier : le modèle exponentiel n'est pas le meilleur paramétrage de l'impact qu'a l'environnement sur l'identification du τ par cet algorithme, ceci nous montre aussi que le Tau Neural Classifier ne se comporte pas comme les deux autres algorithmes étudiés (une pente moyenne VS ΣE_T de $-29,4 \times 10^{-4}$ contre $\approx 8 \times 10^{-4}$ pour les autres algorithmes).

Les réseaux de neurones de la méthode TaNC ont été entraînés sur des échantillons où les candidats τ ont une impulsion transverse comprise entre 20 et 50 GeV et dans des échantillons avec peu d'objets physique (QCD de bas $\hat{P}_T$ et $Z\tau\tau$). Les échantillons électrofaibles étant présents dans cette étude, on vérifie le comportement de la méthode TaNC pour des candidats τ de même impulsion transverse (Fig. 5.35).

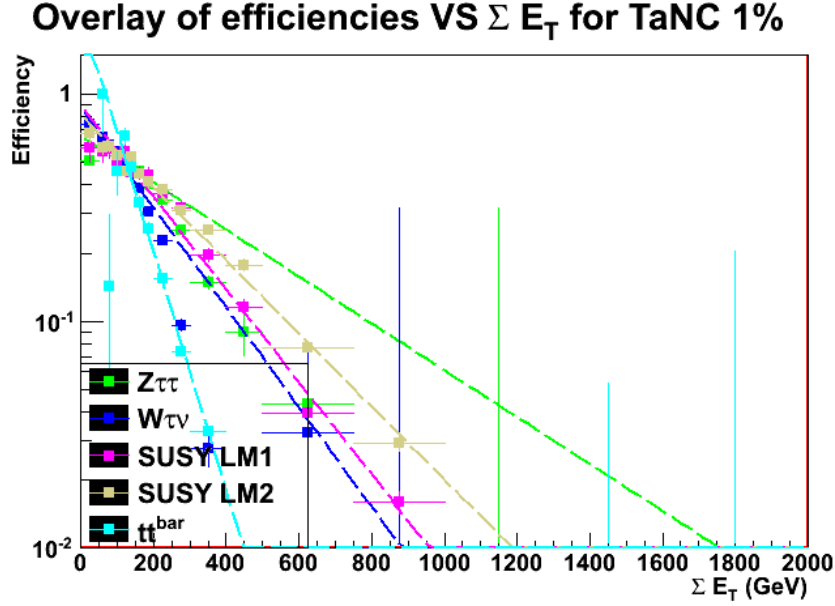


FIGURE 5.35 – Efficacités d'identification Tau Neural Classifier VS ΣE_T pour des candidats de $20 \text{ GeV} < P_T < 50 \text{ GeV}$, $|\eta| < 1$.

5.2.6.2 Influence des multiplicités d'objets

L'algorithme Tau Neural Classifier est un ensemble de réseaux de neurones entraînés à reconnaître l'un des 5 modes de désintégration dominants du τ . Il sera donc affecté négativement par la présence de particules supplémentaires dans ce jet car ceux-ci faussent le mode de désintégration reconstruit de la même façon que cette présence affectait l'algorithme HPS.

• Influence de la multiplicité de hadrons chargés :

Les études menées à l'aide des algorithmes basés sur les coupures et HPS ont démontré que le meilleur paramétrage de l'effet de l'environnement sur l'identification de lepton τ est celle de la multiplicité de hadrons chargés. Nous effectuons le même paramétrage avec l'algorithme Tau Neural Classifier dans la figure 5.36.

TABLE 5.12 – Détails de la régression exponentielle $\epsilon(N_{\text{Chargés}}) = e^{aN_{\text{Chargés}}+b}$ de la Fig 5.36

Échantillon	a ($P_T \geq 50 \text{ GeV}$)	b ($P_T \geq 50 \text{ GeV}$)
$Z\tau\tau$	$-(17,7 \pm 1,1) \times 10^{-3}$	$-0,49 \pm 0,02$
$W\tau\nu$	$-(25,0 \pm 1,0) \times 10^{-3}$	$-0,30 \pm 0,02$
SUSY LM1	$-(43,4 \pm 1,2) \times 10^{-3}$	$-0,05 \pm 0,04$
SUSY LM2	$-(40,5 \pm 0,5) \times 10^{-3}$	$-0,13 \pm 0,02$
$t\bar{t}$	$-(30,2 \pm 1,1) \times 10^{-3}$	$-0,28 \pm 0,04$
Moyenne	$-31,4 \times 10^{-3}$	$-0,25$
Écart type	$10,7 \times 10^{-3}$	$0,17$
Écart type relatif	34,1%	68%

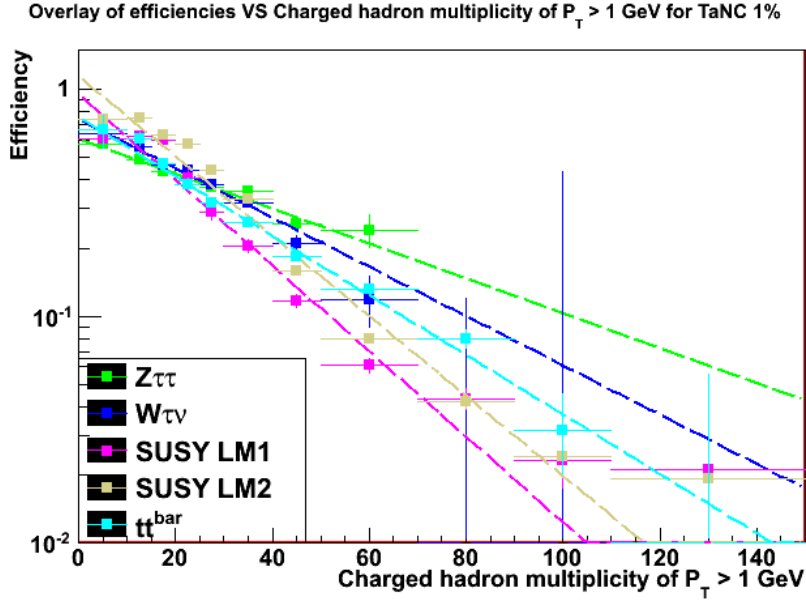


FIGURE 5.36 – Efficacités d'identification Tau Neural Classifier VS $N_{Hadrons\ Chargés}$ pour des candidats de $P_T > 50\text{ GeV}$, $|\eta| < 1$.

- Influence de la multiplicité de photons :

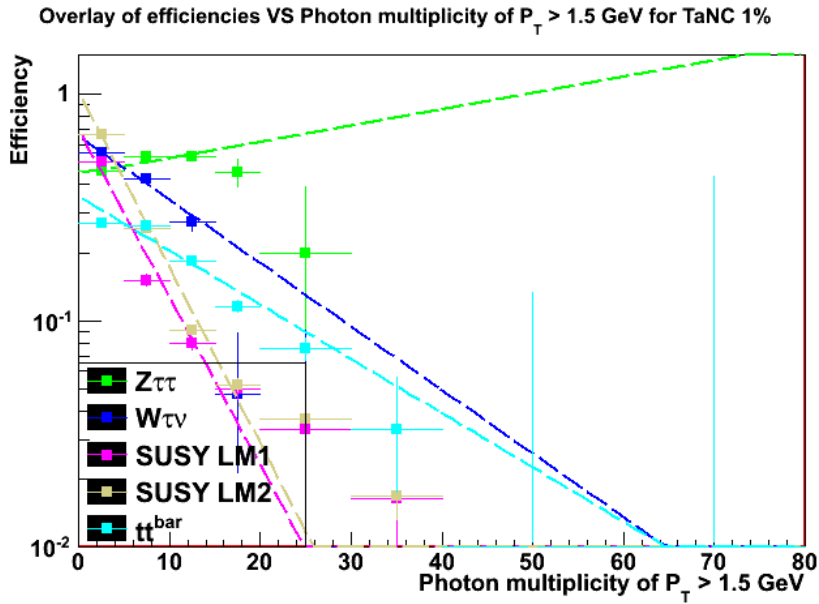


FIGURE 5.37 – Efficacités d'identification Tau Neural Classifier VS $N_{Photons}$ pour des candidats de $P_T > 50\text{ GeV}$, $|\eta| < 1$.

- Influence de la multiplicité de jets :

L'algorithme TaNC est très sensible aux variables d'environnement, y compris à la multiplicité de jets. Il est aussi très sensible au type d'évènement considéré au vu des divergences

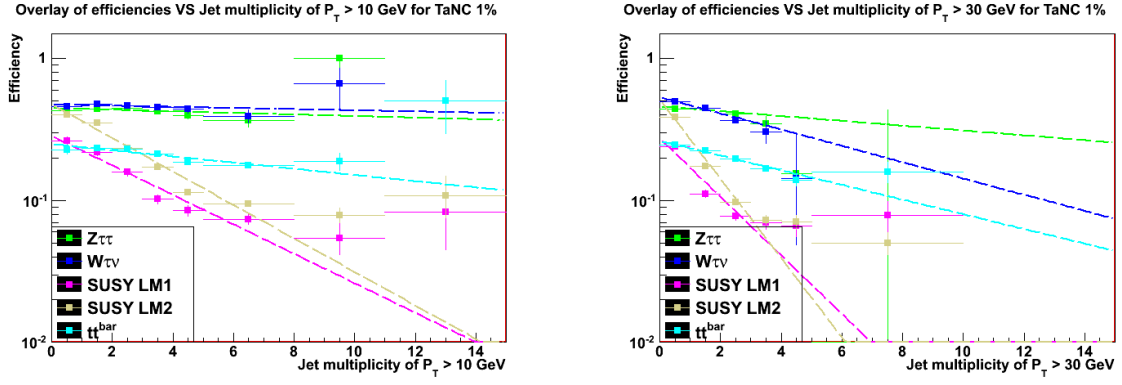


FIGURE 5.38 – Efficacités d'identification Tau Neural Classifier pour des candidats de $P_T > 50$ GeV, $|\eta| < 1$ VS **a)** La multiplicité de jets de plus de $P_T > 10$ GeV, **b)** La multiplicité de jets de $P_T > 30$ GeV.

entre les échantillons. un paramétrage est impossible à effectuer avec cet algorithme.

5.3 Influence de la topologie des évènements sur le taux de mauvaise identification

L'objectif de cette étude est de déterminer les similitudes entre l'impact de l'environnement sur les efficacités d'identification et celles sur les taux de mauvaise identification de jets QCD. Pour cette étude nous avons conservé uniquement les distributions de taux de faux VS ΣE_T et $N_{Hadrons\ Chargés}$ qui nous ont donné les meilleures paramétrages pour les algorithmes cut-based et HPS.

Dans le but de rechercher un paramétrage absolu de l'effet de l'environnement sur l'identification de τ , nous étudions les taux de mauvaise identification dans les mêmes intervalles en P_T et η que ceux étudiés dans la section 5.2. L'étude des taux de mauvaise identification nous permet d'utiliser les échantillons de données du LHC vus dans le chapitre 4 en tant qu'échantillon QCD : on prend comme candidats les jets sonde pour un unique seuil de déclenchement correspondant à l'intervalle en P_T étudié.

Dans les échantillons Monte Carlo, les candidats considérés comme jets QCD sont les candidats situés à une distance $\Delta R > 0.5$ du τ généré le plus proche. Ces candidats sont soumis à la réjection d'électrons et de muons vue dans le chapitre 3 afin de ne pas inclure dans le calcul les leptons légers fortement présents dans tous les échantillons contenant des τ . Ces réjections n'utilisent que l'information reconstruite pour rester cohérent avec les véritables données.

5.3.1 Identification basée sur les coupures

L'identification basée sur les coupures repose sur le contenu du jet. Aussi un jet QCD identifié comme τ est un jet collimé qui présente les mêmes caractéristique que celui d'un τ . L'environnement extérieur au jet doit avoir des effets similaires à ceux sur les efficacités d'identification.

• Influence de ΣE_T :

La figure 5.39 nous présente une décroissance rapide du taux de faux à basse ΣE_T (entre 0 et 400 GeV) puis une stabilisation vers les plus hautes ΣE_T . Nous observons qualitativement une tendance similaire pour tous les échantillons étudiés, mais des divergences subsistent dans les niveaux de mauvaise identification : les évènements purement QCD (données et MC) présentent les fractions de mauvaise identification les plus faibles que les échantillons électrofaibles, $t\bar{t}$ et SUSY. Le fait qu'un jet QCD passe ou non l'identification de τ dépend donc en partie de l'environnement de l'évènement mais il existe un facteur supplémentaire influençant la mauvaise identification des jets QCD.

L'étude du contenu de l'anneau d'isolation avant identification (Fig. 5.40) pour ces jets QCD dans différentes gammes en ΣE_T nous montre les effets d'environnement : l'occupation de l'anneau d'isolation devient plus importante pour tous les échantillons à mesure que ΣE_T augmente. On remarque quatre régimes dans ces graphiques :

- Les jets QCD des échantillons électrofaibles ont des anneaux d'isolation peu peuplés

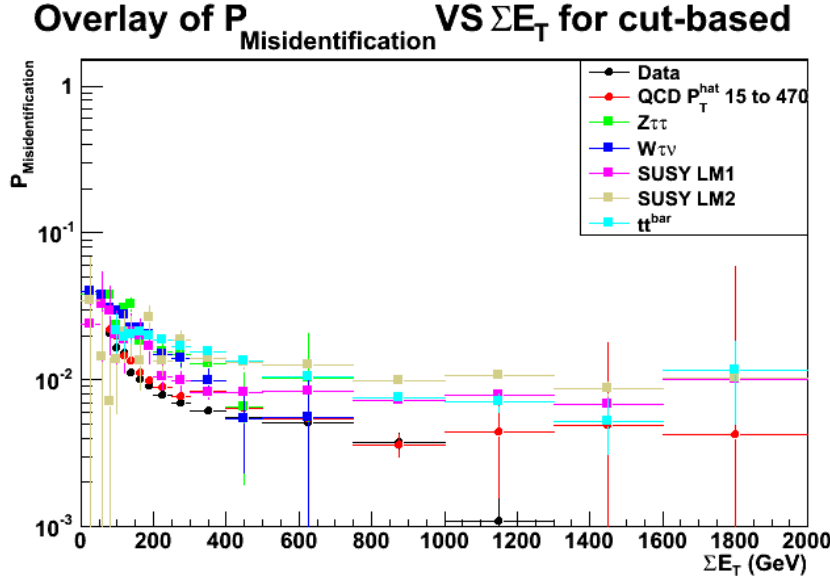


FIGURE 5.39 – Taux de mauvaise identification basée sur les coupures VS ΣE_T pour des candidats de $P_T > 50 \text{ GeV}$, $|\eta| < 1$, les données et QCD MC sont sélectionnés à l’aide des systèmes de déclenchement pour jet avec un seuil de 50 GeV.

(particulièrement visible à bas ΣE_T), ce qui explique les taux de faux observés élevés par rapport aux autres échantillons.

- Les échantillons SUSY dont les jets QCD ont des anneaux d’isolation très peuplés à haut ΣE_T , et peu peuplés à bas ΣE_T . On retrouve les deux grandes classes de processus SUSY : la production de squark/gluino d’une part, et la production directe de sleptons d’autre part.
- Les événements QCD (données et MC) avec des jets dont les anneaux d’isolation sont très peuplés et dont le taux de faux est très faible dans tous les cas de figure. Nous noterons cependant que les anneaux d’isolation sont moins peuplés que dans les événements SUSY à haut ΣE_T , mais on retrouve plus de candidats ayant au moins un hadron chargé dans l’anneau d’isolation dans les événements QCD que dans les événements SUSY.
- Les événements $t\bar{t}$ présentant des jets QCD peu peuplés dans l’anneau d’isolation, et donc un taux de faux important. On remarque que les jets QCD contenus dans ces événements sont peu sensibles à l’environnement.

• Influence de la multiplicité de hadrons chargés :

L’étude de l’influence de ΣE_T a révélé des disparités selon que l’on considère les efficacités d’identification de τ ou la mauvaise identification de jets QCD. Il faut donc étudier un paramètre influent sur l’identification de τ : la multiplicité d’objets, notamment les hadrons chargés.

Les résultats de la figure 5.41 nous présentent une baisse plus régulière du taux de faux

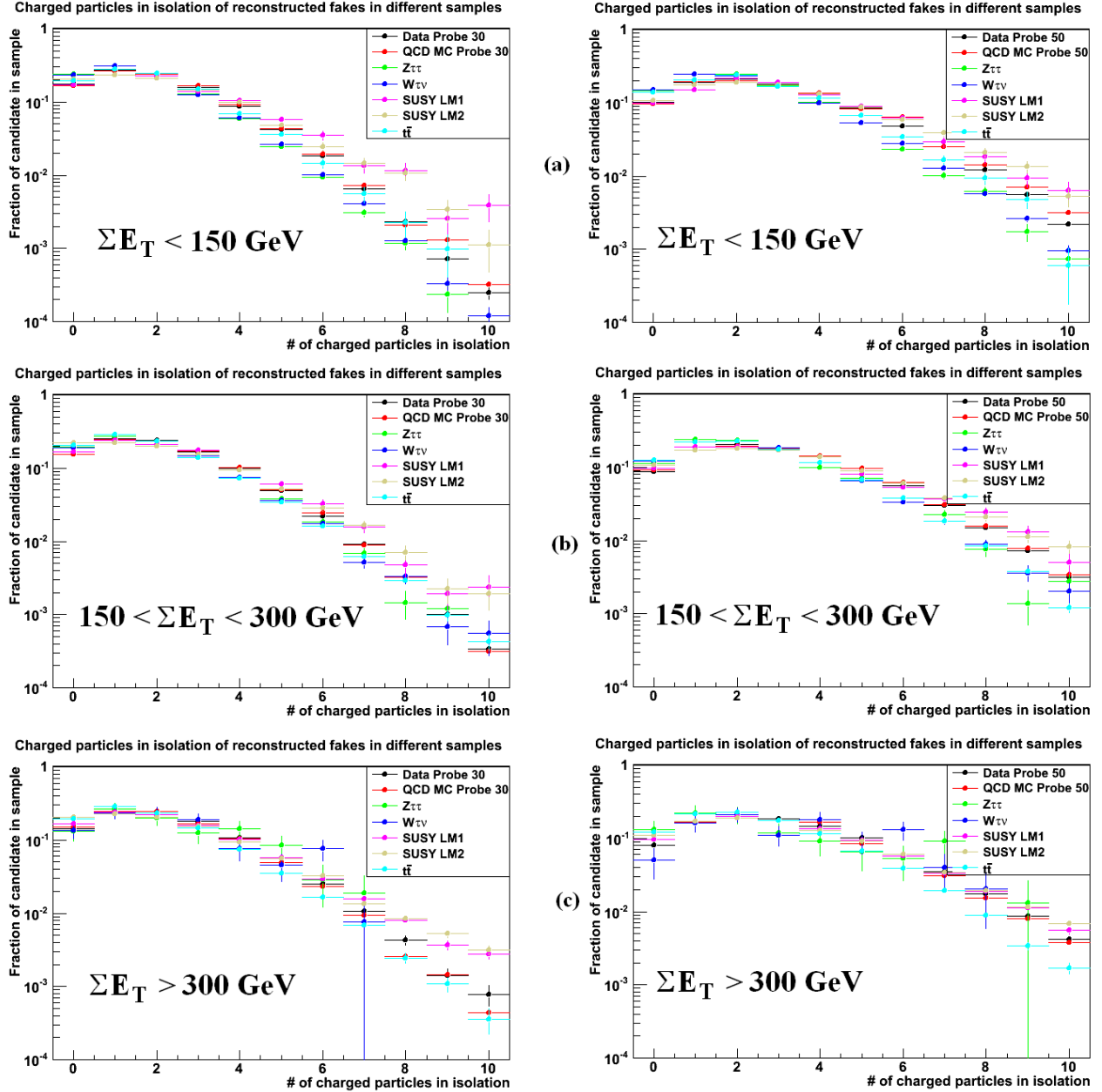


FIGURE 5.40 – Multiplicités en hadrons chargés ($P_T > 1 \text{ GeV}$) de l’anneau d’isolation pour les jet QCD reconstruits (avant isolation) de $30 < P_T < 50 \text{ GeV}$ (à gauche) et $50 < P_T < 80 \text{ GeV}$ (à droite) dans des évènements de ΣE_T a) Inférieur a 150 GeV, b) Compris entre 150 et 300 GeV, c) Supérieur à 300 GeV.

à mesure que $N_{\text{Hadrons Chargés}}$ augmente, ce qui est plus proche du comportement observé lors des déterminations d’efficacités d’identification dans des conditions similaires. Ici aussi nous constatons une tendance similaire pour la plupart des échantillons (seulement $t\bar{t}$ a un comportement très différent) mais avec des niveaux différents, ces divergences entre les différents type d’échantillons nous confirment l’hypothèse d’un troisième facteur influençant la mauvaise identification des jets QCD.

L’analyse du contenu des anneaux d’isolation (Fig. 5.42), nous montre les même tendances

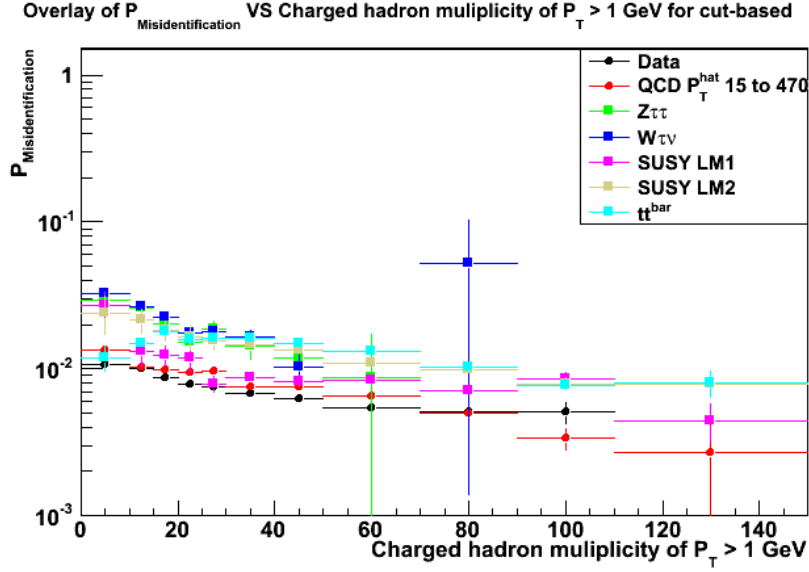


FIGURE 5.41 – Taux de mauvaise identification basée sur les coupures VS $N_{Hadrons\ Chargés}$ pour des candidats de $P_T > 50\text{ GeV}$, $|\eta| < 1$.

que celles observées pour l'étude de la probabilité de mauvaise identification en fonction de ΣE_T . Nous remarquons le cas à part de l'échantillon $t\bar{t}$ dont les populations de l'isolation sont les plus faibles lorsque la multiplicité de hadrons est haute (plus de 20 hadrons dans l'évènement) : on retrouve le comportement observé lors de l'étude en fonction de ΣE_T , i.e. la sensibilité faible à l'environnement des jets QCD de l'échantillon $t\bar{t}$ dans la figure 5.41.

5.3.2 Identification Hadron Plus Strips

L'étude vue dans la section précédente a aussi été menée pour l'algorithme Hadron Plus Strips pour deux raisons : premièrement dans le but de vérifier si l'effet de ΣE_T sur le taux de faux est dépendant de l'algorithme. Et deuxièmement pour tester l'hypothèse d'une dépendance en la multiplicité de hadrons qui ne dépend ni de l'algorithme, ni de la nature des candidats étudiés (jets de τ ou jets QCD).

• Influence de ΣE_T :

La figure 5.43 nous présente des résultats similaires à ceux observés pour la méthode basée sur les coupures :

- Baisse rapide de la probabilité de mauvaise identification aux bas ΣE_T de hadrons chargés puis stabilisation aux plus hautes ΣE_T .
- Séparation des évènements QCD mous, électrofaibles, SUSY et $t\bar{t}$.
- Légère insensibilité de l'échantillon $t\bar{t}$ à l'environnement.

• Influence de la multiplicité de hadrons chargés :

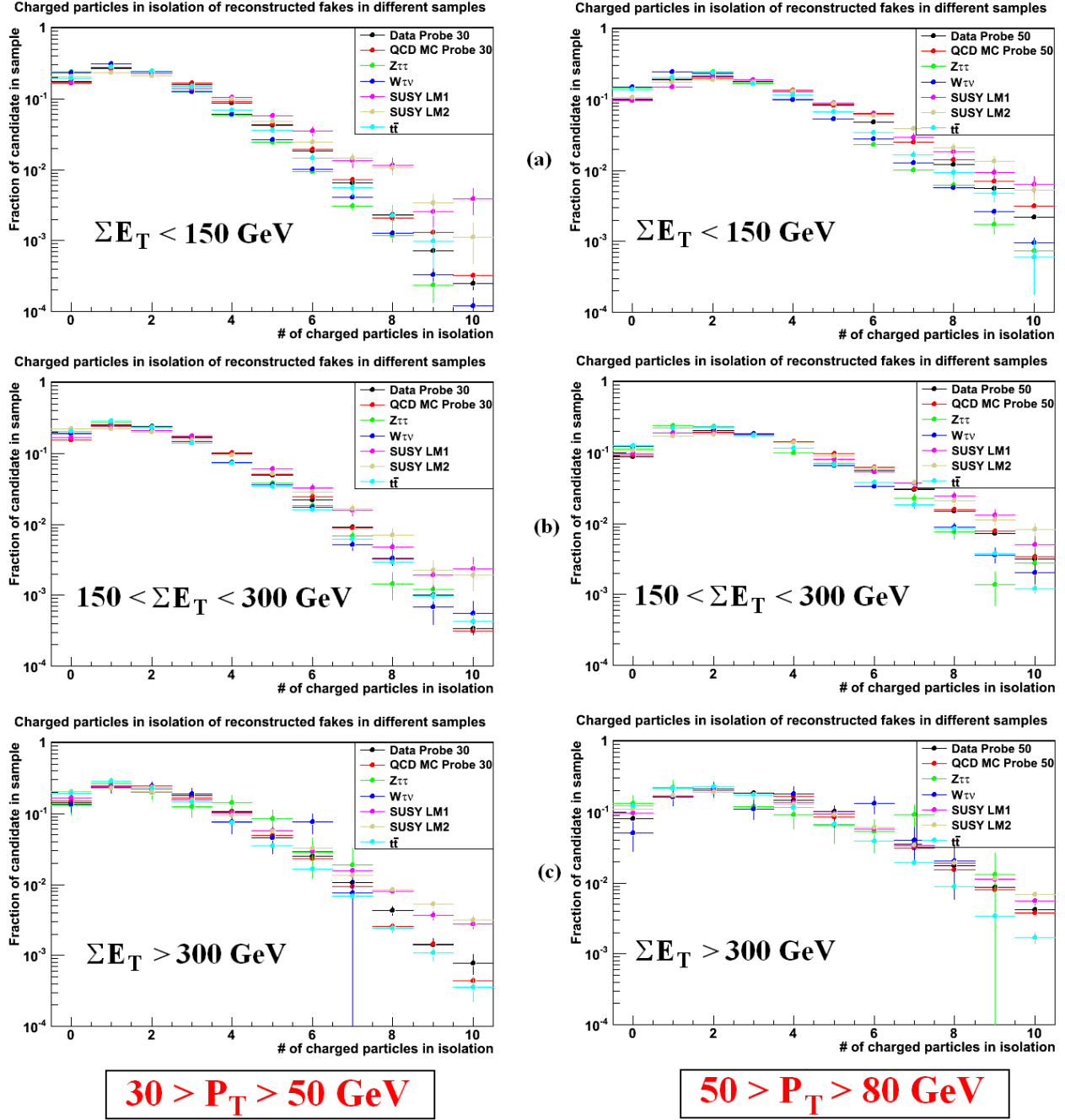


FIGURE 5.42 – Multiplicités en hadrons chargés ($P_T > 1 \text{ GeV}$) de l’anneau d’isolation pour les jets QCD reconstruits (avant isolation) de $30 < P_T < 50 \text{ GeV}$ (à droite) et $50 < P_T < 80 \text{ GeV}$ (à gauche) dans des évènements avec des multiplicité de hadrons chargés a) Inférieure à 20, b) Comprise entre 20 et 40, c) Supérieure à 40.

Afin de comparer les différents algorithmes, nous effectuons ici l’étude de la probabilité de mauvaise identification en fonction de la multiplicité de hadrons chargés, les résultats sont présentés dans la figure 5.44.

Toutes les tendances observées pour l’algorithme basé sur les coupures restent valides pour l’algorithme HPS (mais renvoient des probabilités de mauvaise identification plus faibles).

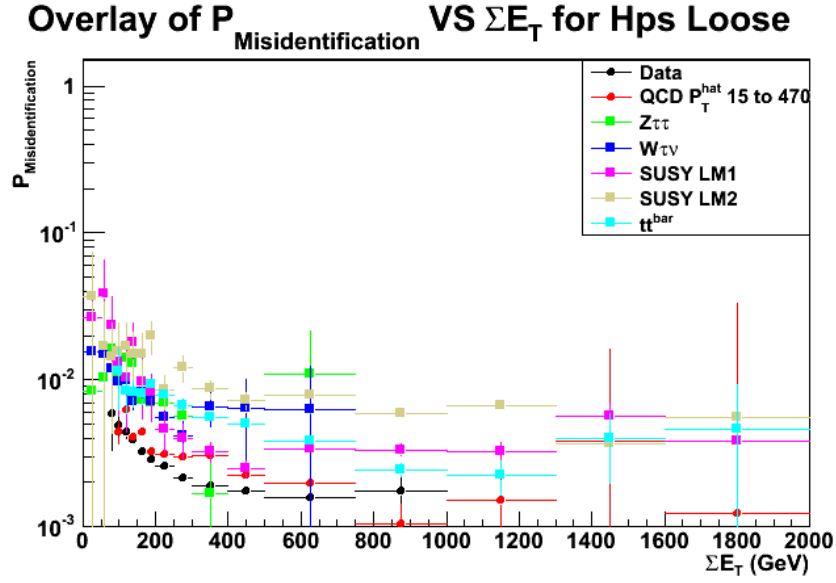


FIGURE 5.43 – Taux de mauvaise identification Hadron Plus Strips VS ΣE_T pour des candidats de $P_T > 50 \text{ GeV}$, $|\eta| < 1$.

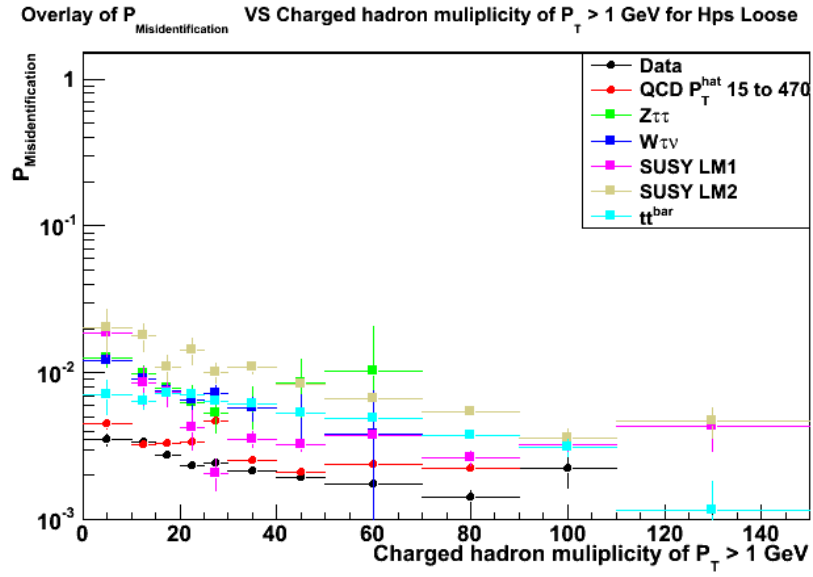


FIGURE 5.44 – Taux de mauvaise identification Hadron Plus Strips VS $N_{\text{Hadrons Chargés}}$ pour des candidats de $P_T > 50 \text{ GeV}$, $|\eta| < 1$.

Les algorithmes HPS et basé sur les coupures présentent la même sensibilité aux variables caractérisant la topologie de l'évènement en ce qui concerne la mauvaise identification de jets QCD en tant que τ .

5.3.3 Identification Tau Neural Classifier

Nous avons vu dans la section 5.2 que la sensibilité de l'identification du lepton τ à l'environnement de l'évènement était très dépendante de l'échantillon étudié. Il était impossible de caractériser d'une manière globale l'impact de cet environnement sur les efficacités d'identification, mais nous constatons une très forte sensibilité de l'algorithme TaNC aux variables d'environnement.

L'algorithme TaNC a été entraîné pour suivre des scenarii de bruit de fond ayant un certain niveau (nous étudions ici le scénario à 1% de faux), nous testons alors la sensibilité de l'algorithme TaNC à l'environnement pour la mauvaise identification de jets QCD.

- Influence de ΣE_T :

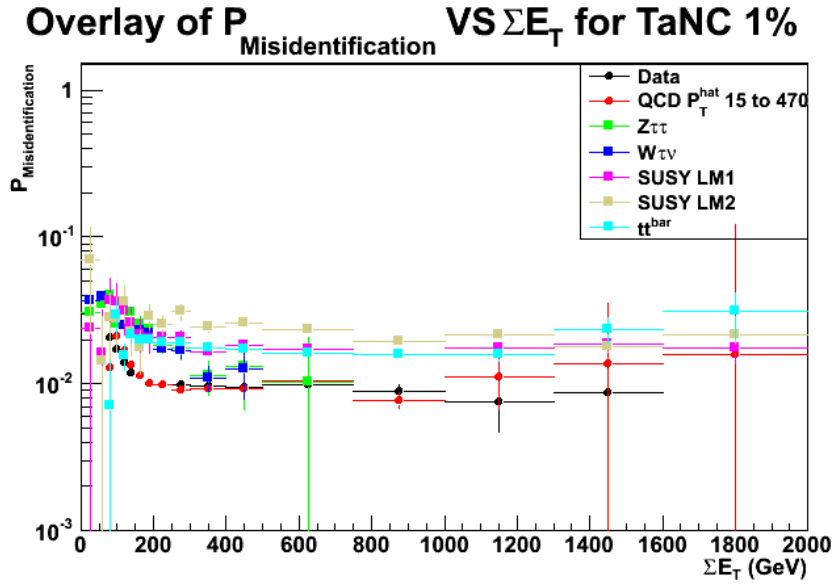


FIGURE 5.45 – Taux de mauvaise identification Tau Neural Classifier VS ΣE_T pour des candidats de $P_T > 50 \text{ GeV}$, $|\eta| < 1$.

La figure 5.45 présente une tendance un peu similaire aux deux autres algorithmes vus dans les sections précédentes, à l'exception du fait que les niveaux de taux de faux se stabilisent beaucoup plus vite (300 GeV contre quasiment 500 GeV pour cut-based et HPS). De plus le taux de mauvaise identification décroît d'un facteur 2 entre les basses et haute valeurs de ΣE_T alors que nous constatons une décroissance d'un facteur 5 à 10 pour les autres algorithmes. L'algorithme TaNC est donc moins sensible à la somme des énergies transverses que les autres algorithmes en ce qui concerne la mauvaise identification de jets QCD.

- Influence de la multiplicité de hadrons chargés :

Les probabilités de mauvaise identification présentées dans la figure 5.46 varient très peu avec la multiplicité de hadrons chargés. Ceci confirme le fait que l'algorithme TaNC est très peu sensible à l'environnement en ce qui concerne la réjection des jets QCD. L'algorithme TaNC étudié ici n'est pas adapté pour les analyses contenant une activité importante (mul-

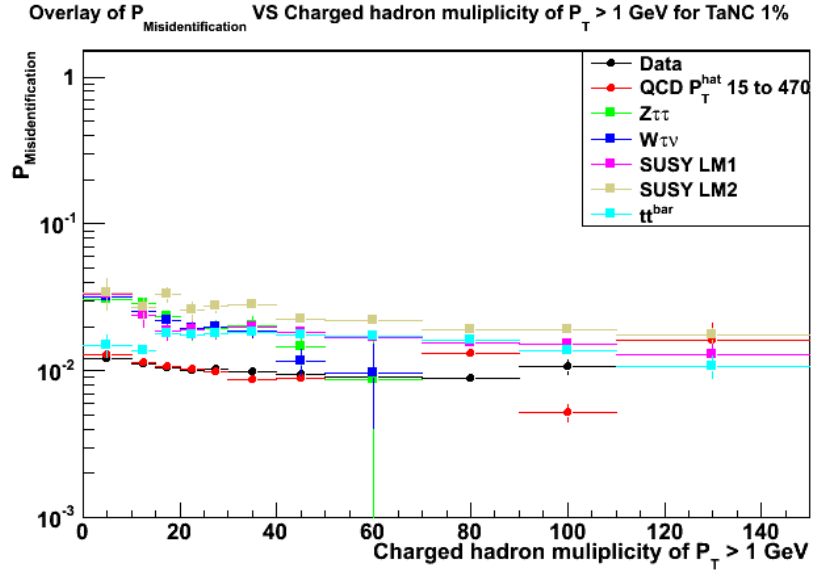


FIGURE 5.46 – Taux de mauvaise identification Tau Neural Classifier VS $N_{Hadrons Chargés}$ pour des candidats de $P_T > 50 GeV$, $|\eta| < 1$.

tiplicité d'objets importante avec une impulsion transverse élevée), les efficacités chutent drastiquement à haute activité tandis que les taux de faux restent quasiment constants.

Conclusion

Ce mémoire de thèse a présenté l'étude de l'identification du lepton τ dans des environnements divers afin de mieux comprendre ses applications pour la recherche de supersymétrie à l'aide du détecteur CMS au LHC. Le LHC est le collisionneur le plus puissant à ce jour et peut découvrir la supersymétrie ou dans le cas contraire rejeter certains scénarii. Les recherches de supersymétrie dans le canal $\tau^+\tau^- + \text{Missing } E_T$ menées par la collaboration en 2010-2011[61] (Fig. 5.47) ont déjà posé des limites sur les points de l'espace mSUGRA LM1 et LM2 étudiés dans cette thèse, mais d'autres scénarii de SUSY restent à prospecter et l'identification du lepton τ conserve toute son importance.

Les méthodes d'identification du lepton τ mises en place par la collaboration CMS sont performantes. Les efficacités d'identification se situent entre 40 et 70% dépendant de l'échantillon et de la méthode d'identification et les taux de mauvaise identification des jets QCD produits dans le LHC en très bon accord avec les simulations ont une moyenne de 1%. Cependant toutes ces méthodes sont basées sur les caractéristiques intrinsèques des jets hadroniques de τ : la largeur du jet et les particules contenues dans ce jet, ces deux paramètres étant sensibles à l'environnement hautement hadronique du LHC.

Cet environnement hautement hadronique a été caractérisé par la multiplicité de hadrons chargés présents dans les événements, les méthodes cut-based et HPS ont une sensibilité proche à cet environnement mais la méthode TaNC a une sensibilité différente. La baisse d'efficacité qu'entraîne une multiplicité de hadrons chargés importante a pu être paramétrée pour les deux premières méthodes. Les futures recherches de Supersymétrie basées sur l'identification des taus pourront alors tenir compte de l'activité hadronique très importante de ces événements. Les recherches actuellement poursuivies par la collaboration CMS utilisent la validation de l'identification de tau obtenue par les désintégrations du $Z\tau\tau + X$, où il y a suffisamment de données afin de comparer simulation et expérience. Néanmoins ces études ne sont pas effectuées à haute énergie transverse ou énergie transverse manquante, comme exigé dans les recherches de la Supersymétrie. Nous avons vu que l'énergie transverse manquante n'a quasi aucun effet sur l'identification des taus contrairement à l'énergie transverse totale, fortement corrélée avec la multiplicité des particules dans l'événement, qui a un fort impact sur l'identification en la diminuant fortement à haute multiplicité. Les corrélations paramétrées trouvées dans le chapitre 5 peuvent être utilisées afin d'extrapoler les efficacités validées (par la collaboration CMS) avec les événements $Z\tau\tau$ vers les environnements denses à haute ΣE_T que l'on retrouve dans les collisions où on recherche la Supersymétrie.

En ce qui concerne les perspectives de recherche, il faut apporter des corrections aux méthodes d'identification de τ afin d'y inclure les effets d'environnement. Des événements SUSY peuvent potentiellement ne pas être identifiés car les τ présents dans ces événements

ne sont plus reconnus. La probabilité de mauvaise identification d'un faux tau venant d'évènements QCD peut être bien contrôlée par des coupures cinématiques comme il a été démontré par des analyses existantes, contrairement aux faux taus venant des jets des évènements SUSY eux-mêmes ou de tout autre processus donnant lieu à des énergies transverses très importantes. Le bruit de fond électrofaible et $t\bar{t}$ contenant les vrais taus doit soigneusement être évalué pour tenir compte des modifications de l'identification des taus dans les évènements à haute multiplicité.

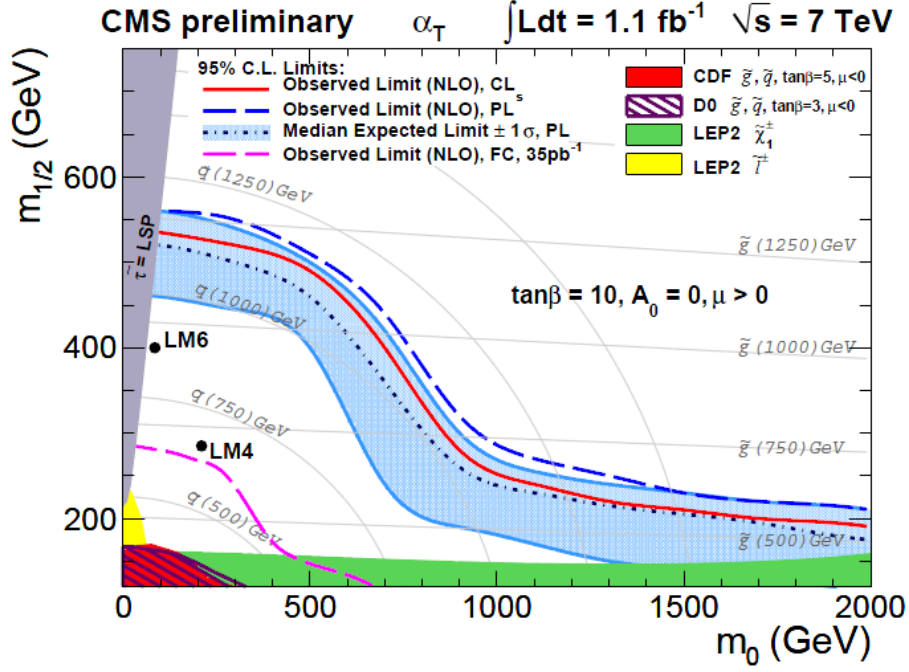


FIGURE 5.47 – Limites d'exclusion dans l'espace $m_{1/2}, m_0$ pour la recherche de supersymétrie dans la voie hadronique (source [62]).

Annexes

Annexe A

Estimation des bruits de fond par la méthode de “fake-rate”

A.1 Conditions d’analyse

La méthode pour estimer la contribution des jets QCD au bruit de fond pour l’identification de lepton τ à été décrite dans le chapitre 3 (section 3.4). Des histogrammes à trois dimensions (value maps) ont été créés, chaque intervalle contenant une valeur entre 0 et 1 définissant la probabilité que ce jet soit reconnu comme τ en fonction de P_T , η et R_{Jet} . Ces histogrammes ont été créés pour les jets QCD et pour les jets de τ et incluent plusieurs niveaux d’identification : identification par la charge et le nombre de trace, l’identification par isolation et la combinaison des deux.

Deux échantillons Monte Carlo ont été étudiés :

- Un échantillon $Z\tau\tau$: /RelValZTT/CMSSW_3_5_1-START3X_V21-v1/GEN-SIM-RECO
- Un échantillon QCD : /RelValQCD_FlatPt_15_3000/CMSSW_3_5_1-MC_3XY_V21-v1/GEN-SIM-RECO

A.2 Resultats sur les échantillons $Z\tau\tau$

Les comparaisons entre spectres des candidats τ en P_T évalués par la méthode vue dans le chapitre 3, et spectres calculés par la méthode d’identification basée sur les coupures sont présentés dans les figures A.1, A.2 et A.3 pour 3 niveaux d’identification : identification par la charge et les traces, identification par isolation puis la combinaison de ces deux critères.

Les distributions se révèlent entre très bon accord l’une envers l’autre pour tous les niveaux d’isolation. Cependant cette méthode est conçue pour évaluer la contribution des jets QCD, l’estimation des jets de τ est nécessaire dans le cas d’analyse de type CDF[56].

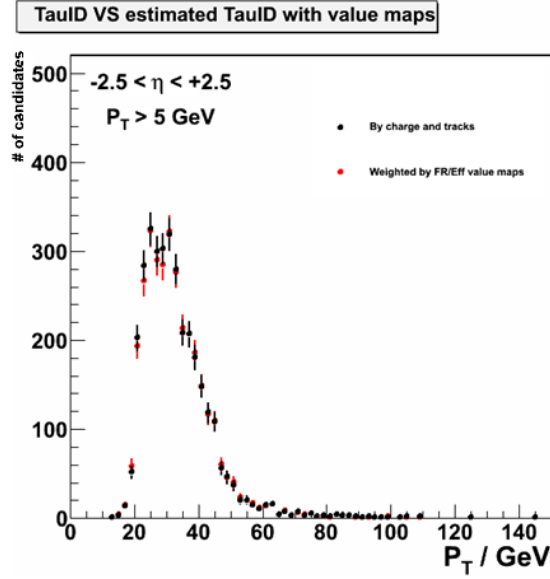


FIGURE A.1 – Spectre en impulsion transverse des candidats τ identifiés par la charge et le nombre de traces (en noir) puis estimés par la méthode de “fake-rate” (en rouge).

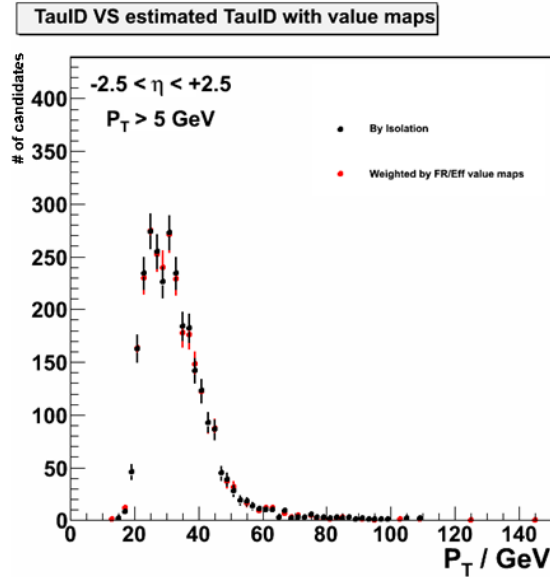


FIGURE A.2 – Spectre en impulsion transverse des candidats τ identifiés par isolation (en noir) puis estimés par la méthode de “fake-rate” (en rouge).

A.3 Resultats sur les échantillons QCD

Une étude similaire à été menée pour les jets QCD. Ces résultats sont présentés dans les figures A.4, A.5 et A.6.

Les distributions obtenues par identification et par évaluation sont en accord l’une envers l’autre dans le cas de l’identification par la charge et les traces. Cet accord est perdu lorsqu’on applique l’isolation.

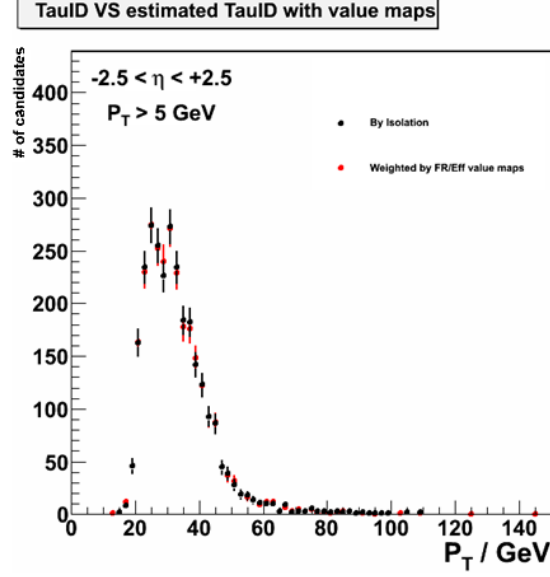


FIGURE A.3 – Spectre en impulsion transverse des candidats τ identifié par isolation + charge + nombre de traces (en noir) puis estimés par la méthode de “fake-rate” (en rouge).

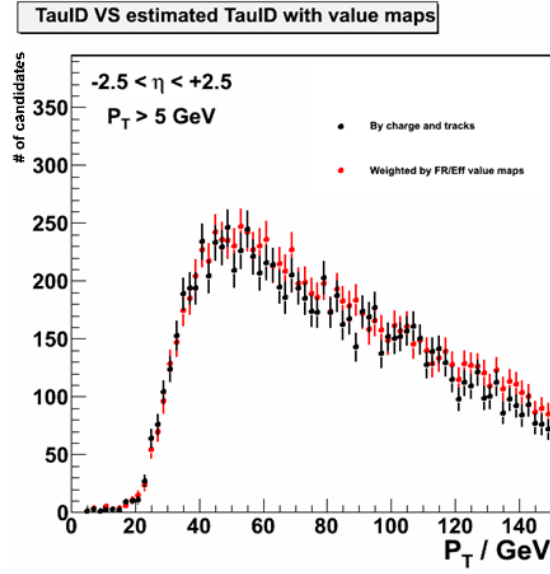


FIGURE A.4 – Spectre en impulsion transverse des candidats τ identifiés par la charge et le nombre de traces (en noir) puis estimés par la méthode de “fake-rate” (en rouge).

Nous pouvons juste poser l’hypothèse que comme la méthode de “fake-rate” ne prend pas en compte l’environnement, des effets notable sur l’isolation mais pas sur la mesure de la charge causent ce désaccord entre le spectre calculé et le spectre estimé.

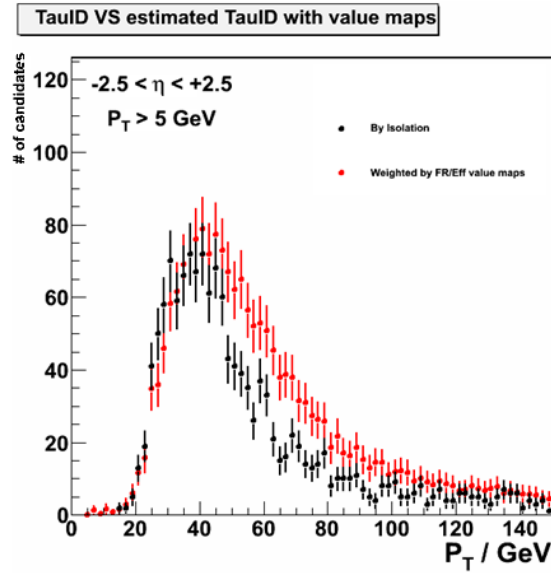


FIGURE A.5 – Spectre en impulsion transverse des candidats τ identifiés par isolation (en noir) puis estimés par la méthode de “fake-rate” (en rouge).

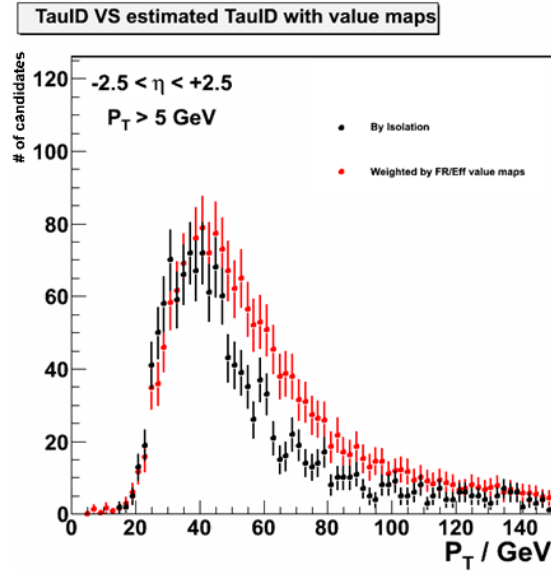


FIGURE A.6 – Spectre en impulsion transverse des candidats τ identifié par isolation + charge + nombre de traces (en noir) puis estimés par la méthode de “fake-rate” (en rouge).

Annexe B

Taux de mauvaise identification des jets QCD triés selon leur impulsion transverse

B.1 Spectres des taux de mauvaise identification pour le jet de plus haut P_T

Les jets de plus haute impulsion sont ceux qui activent les systèmes de déclenchement. Ils sont donc de basse impulsion transverse pour des seuil de 15 GeV, et de haute impulsion transverse pour un seuil de 100 GeV comme le montrent les figures B.1, B.2 et B.3.

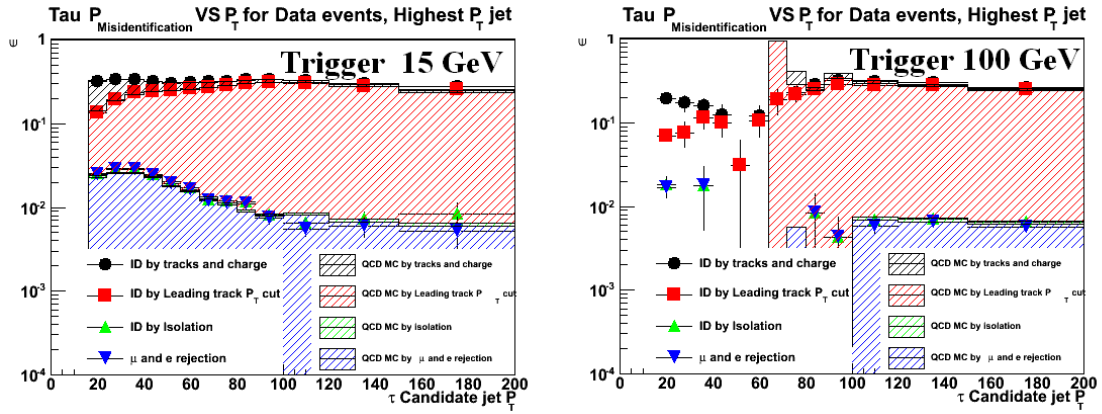


FIGURE B.1 – Taux de mauvaise identification par la méthode cut-based du jet de plus haute impulsion dans les événements QCD MC et données en fonction de P_T , dans les événements sélectionnés par un seuil de déclenchement de 15 GeV (à gauche) et par un seuil de 100 GeV (à droite).

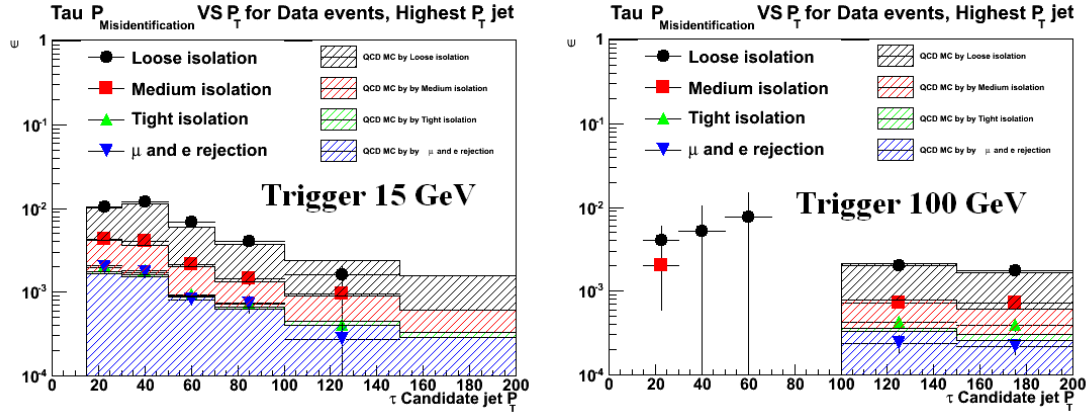


FIGURE B.2 – Taux de mauvaise identification par la méthode HPS du jet de plus haute impulsion dans les événements QCD MC et données en fonction de P_T , dans les événements sélectionnés par un seuil de déclenchement de 15 GeV (à gauche) et par un seuil de 100 GeV (à droite).

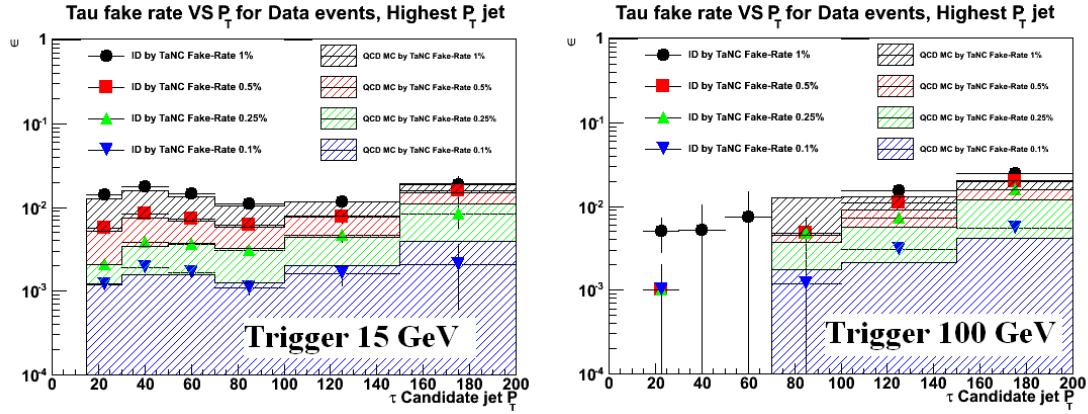


FIGURE B.3 – Taux de mauvaise identification par la méthode TaNC du jet de plus haute impulsion dans les événements QCD MC et données en fonction de P_T , dans les événements sélectionnés par un seuil de déclenchement de 15 GeV (à gauche) et par un seuil de 100 GeV (à droite).

B.2 Spectres des taux de mauvaise identification pour le second jet de plus haut P_T

Les jets de seconde plus haute impulsion transverse peuvent activer les systèmes de déclenchement, mais ne le feront pas la plupart du temps. Nous avons une distribution large en P_T pour ces jets.

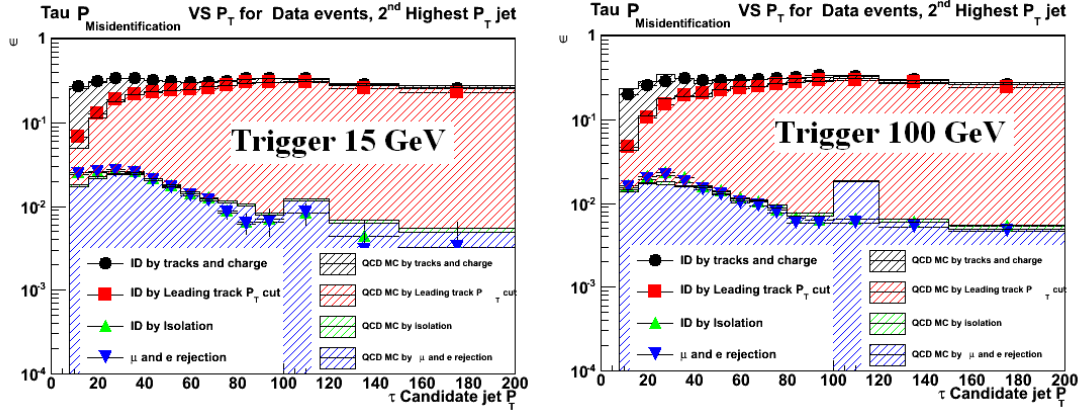


FIGURE B.4 – Taux de mauvaise identification par la méthode cut-based du jet de seconde plus haute impulsion dans les événements QCD MC et données en fonction de P_T , dans les événements sélectionnés par un seuil de déclenchement de 15 GeV (à gauche) et par un seuil de 100 GeV (à droite).

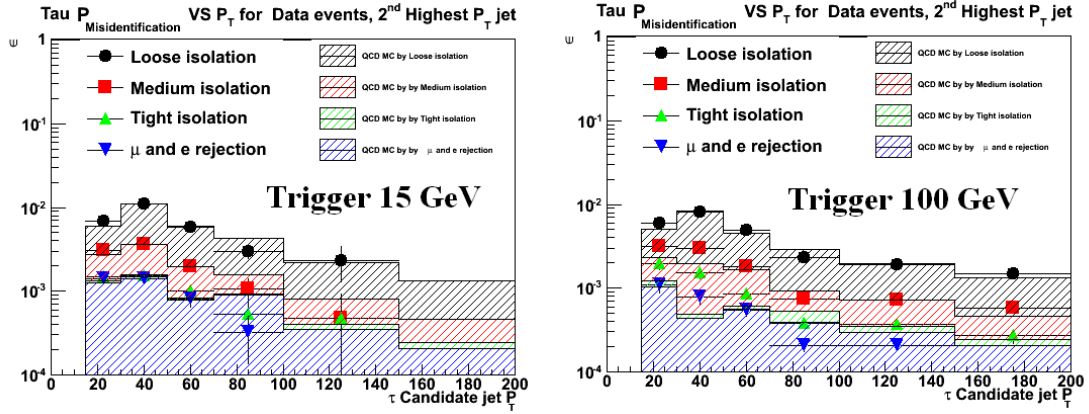


FIGURE B.5 – Taux de mauvaise identification par la méthode HPS du jet de seconde plus haute impulsion dans les événements QCD MC et données en fonction de P_T , dans les événements sélectionnés par un seuil de déclenchement de 15 GeV (à gauche) et par un seuil de 100 GeV (à droite).

B.3 Spectres des taux de mauvaise identification pour le troisième jet de plus haut P_T

Les jets de troisième plus haute impulsion transverse activent très rarement les systèmes de déclenchement, et ne s'obtiennent que dans le cas d'événements multi-jets. Ces jets sont alors de très basse impulsion transverse (Fig. B.7, B.8 et B.9)

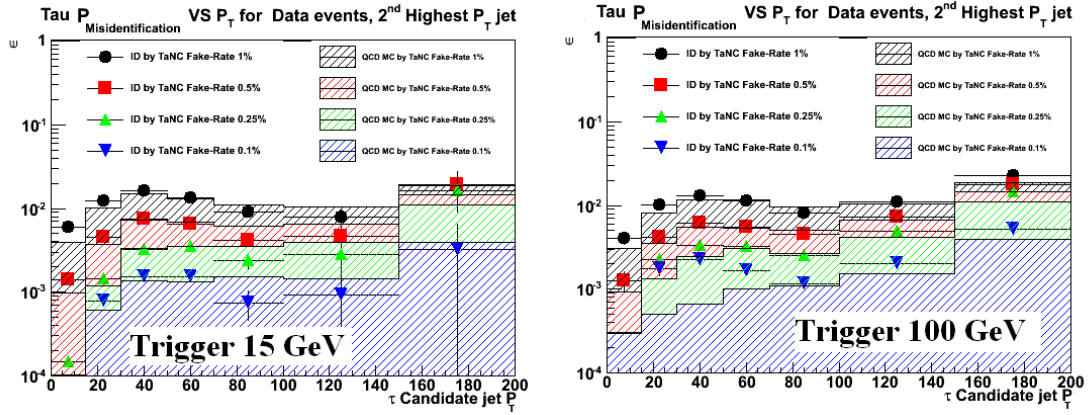


FIGURE B.6 – Taux de mauvaise identification par la méthode TaNC du jet de seconde plus haute impulsion dans les événements QCD MC et données en fonction de P_T , dans les événements sélectionnés par un seuil de déclenchement de 15 GeV (à gauche) et par un seuil de 100 GeV (à droite).

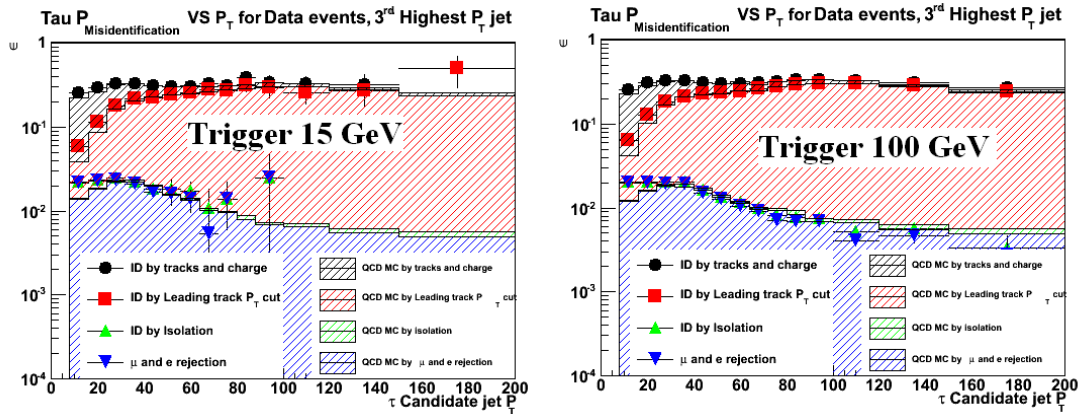


FIGURE B.7 – Taux de mauvaise identification par la méthode cut-based du jet de troisième plus haute impulsion dans les événements QCD MC et données en fonction de P_T , dans les événements sélectionnés par un seuil de déclenchement de 15 GeV (à gauche) et par un seuil de 100 GeV (à droite).

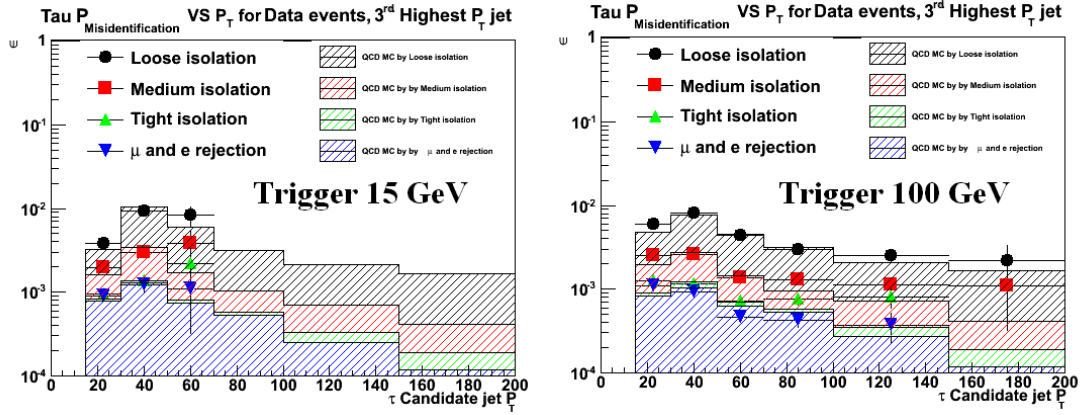


FIGURE B.8 – Taux de mauvaise identification par la méthode HPS du jet de troisième plus haute impulsion dans les événements QCD MC et données en fonction de P_T , dans les événements sélectionnés par un seuil de déclenchement de 15 GeV (à gauche) et par un seuil de 100 GeV (à droite).

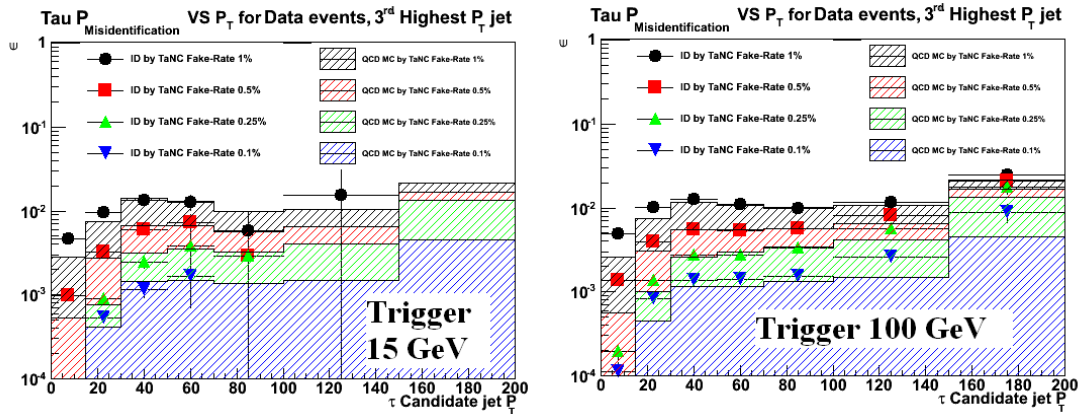


FIGURE B.9 – Taux de mauvaise identification par la méthode TaNC du jet de troisième plus haute impulsion dans les événements QCD MC et données en fonction de P_T , dans les événements sélectionnés par un seuil de déclenchement de 15 GeV (à gauche) et par un seuil de 100 GeV (à droite).

Annexe C

Taux de mauvaise identification de jets QCD avec et sans coupure en énergie transverse manquante

Nous avons vu dans le chapitre 5 (section 5.2.2) que les événements des données contiennent une queue de distribution en énergie transverse manquante que l'on ne retrouve pas dans les données Monte Carlo. Les taux de mauvaise identification vus dans le chapitre 4 sont présentés ici avec et sans une coupure en énergie transverse $MET < 100 \text{ GeV}$.

Cette coupure n'a pas montré son utilité comme on le constate dans les figures C.1, C.2 et C.3. Il y a en effet très peu de déviation avant et après application de la coupure

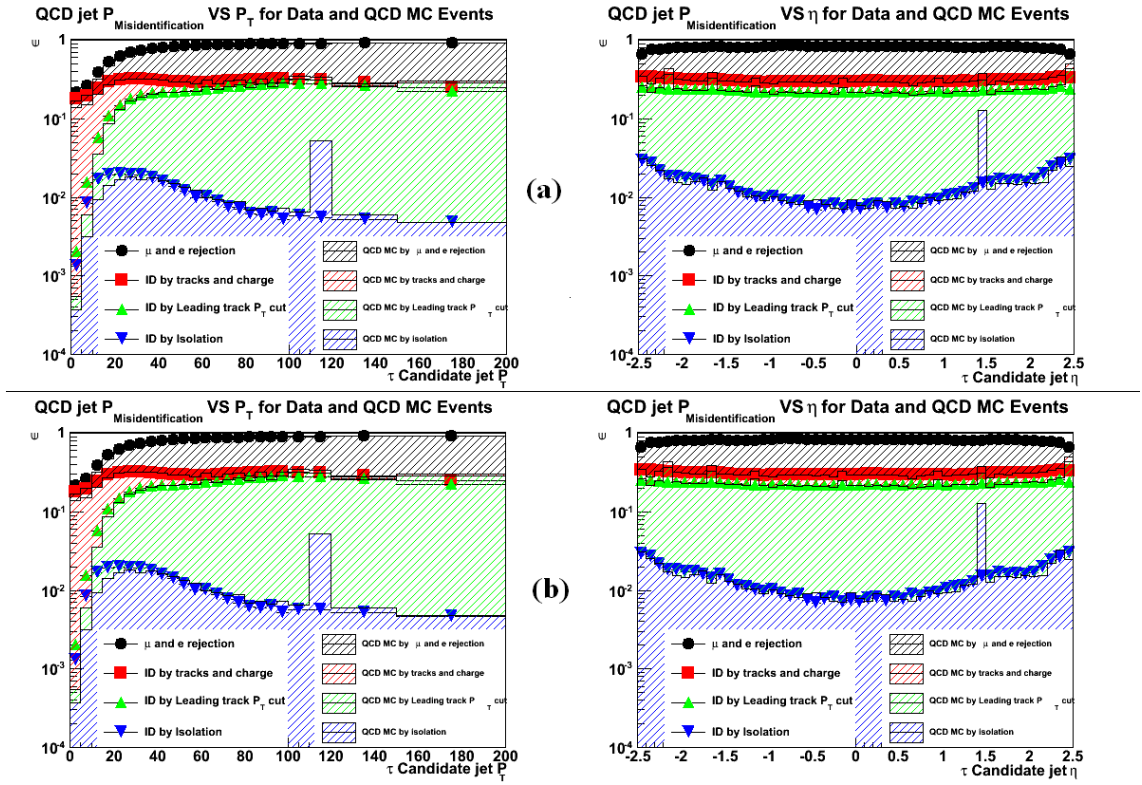


FIGURE C.1 – Taux de mauvaise identification par la méthode cut-based des jets sonde avec un seuil de déclenchement de 100 GeV. a) Aucune coupure en MET n'est appliquée, b) Une coupure $MET < 100 \text{ GeV}$ est appliquée sur les évènements.

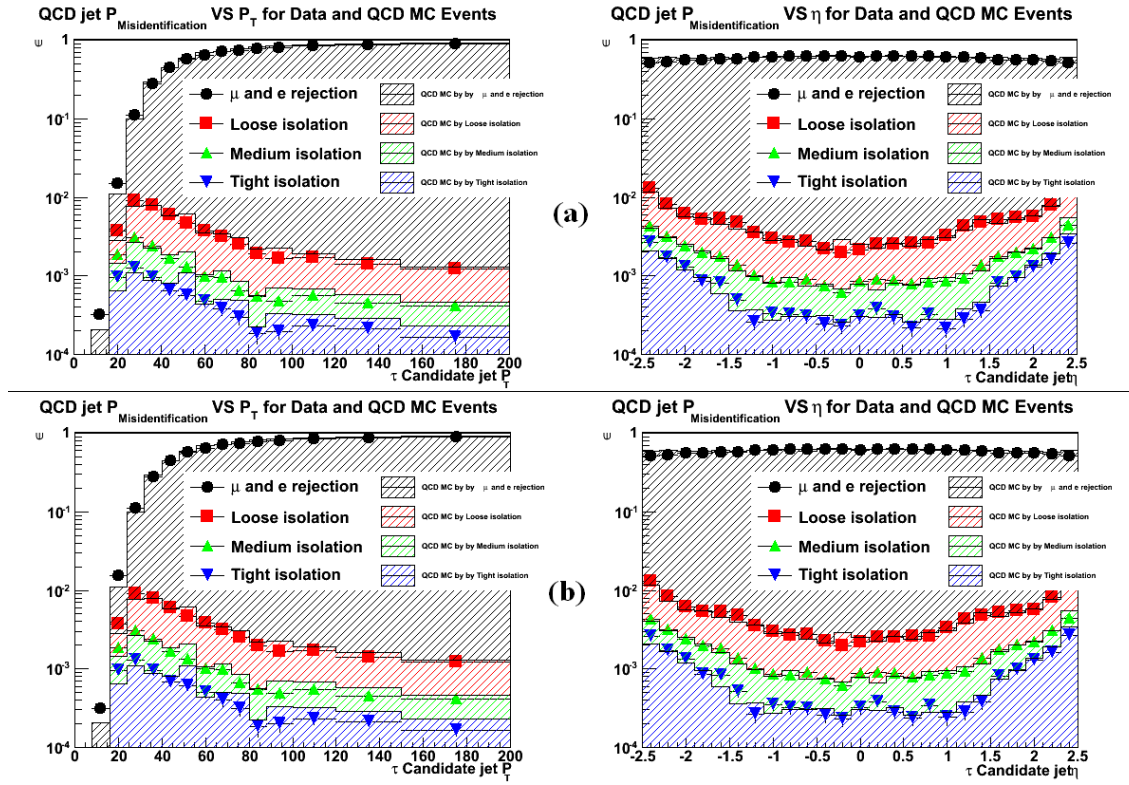


FIGURE C.2 – Taux de mauvaise identification par la méthode Hadron Plus Strips des jets sonde avec un seuil de déclenchement de 100 GeV. a) Aucune coupure en MET n'est appliquée, b) Une coupure $MET < 100$ GeV est appliquée sur les évènements.

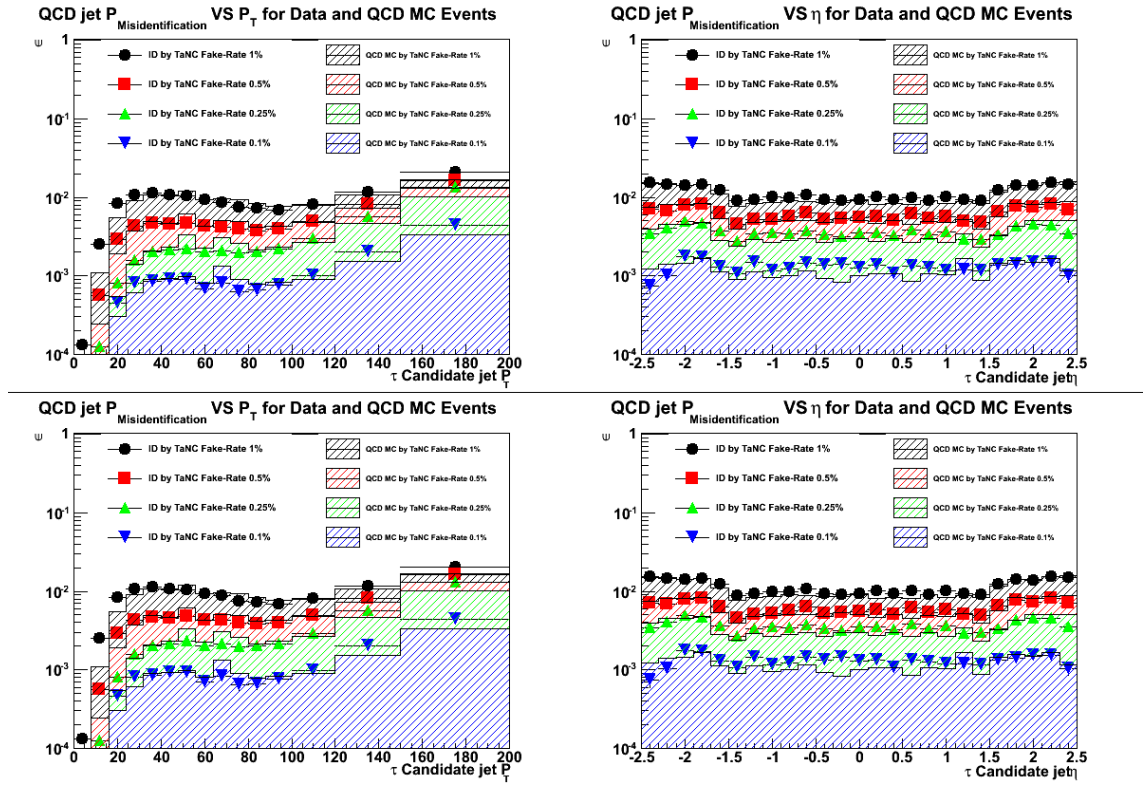


FIGURE C.3 – Taux de mauvaise identification par la méthode Tau Neural Classifier des jets sonde avec un seuil de déclenchement de 100 GeV. a) Aucune coupure en MET n'est appliquée, b) Une coupure $MET < 100 \text{ GeV}$ est appliquée sur les évènements.

Bibliographie

- [1] *Phys. Rev. Letter* 13, 168-171, 1964
A. Salam et al.
- [2] *Nucl. Phys.* 22, 579-588, 1961
S. Glashow et al.
- [3] *Phys. Rev.* 19, 1264-1266, 1967
S. Weinberg
- [4] *Phys. Rev.* 125, 1067-1084, 1962
M. Gell-Mann
- [5] *Nucl. Phys.* 26, 222-229, 1961
Y. Ne'eman
- [6] *CERN-TH-412*, 74pp, 1964
G. Zweig
- [7] *Phys. Rev.* 105, 1413-1414, 1957
C.S. Wu et al.
- [8] *Review of particle physics 2008*
Chapter 16. Structure functions
- [9] *Fundamental interactions and particle chart*
Contemporary Physics Education Project
- [10] *Phys. Rev. D.* 8, 3633-3652, 1973
D.J. Gross et al.
- [11] *Phys. Rev. Letter.* 30, 1346-1349, 1973
H.D. Politzer
- [12] *Phys. Letter B, Volume 47, Issue 4, p365-368 : Advantages of the color octet gluon picture* H.Fritzch, M. Gell-Mann et H. Leutwyler

-
- [13] *Phys. Rev. Letter 13(16) :508-509 Octobre 1964 : Broken Symmetries and the masses of gauge bosons*
Peter W.Higgs
- [14] *Phys. Rev. Letter 81, 1562-1567*
Super Kamiokande Collaboration (Y. Fukuda et al.)
- [15] *hep-ph/9709356 : A Supersymmetry Primer*
Stephen P. Martin
- [16] *Review of particle physics 2008*
Chapter 15. Grand Unified Theories
- [17] *arXiv :hep-ph/9603350 Combined Fit of Low Energy Constraints to Minimal Supersymmetry and Discovery Potential at LEP II*
W. de Boer, G. Burkart, R. Ehret, J. Lautenbacher, W. Oberschulte-Beckmann, U. Schwickerath
- [18] *arXiv :hep-ph/0402229v2 Sleptons at post-WMAP benchmark points at LHC (CMS)*
Yu.M. Andreev, S.I. Bitukov, N.V. Krasnikov
- [19] *CERN/LHCC 2006-21 CMS Physics Technical Design Report*
The CMS collaboration
- [20] *CERN/LHCC 2003-022 ATLAS TDR 016*
Chapter 4. Physics selection strategy
- [21] *QCD for Collider Physics*
P.Z. Skands
- [22] *QCD & Collider Physics*
R.K. Ellis, W.J. Stirling, B.R. Webber
- [23] *Jet production rates in association with W and Z bosons in pp Collisions at $\sqrt{7}$ TeV*
The CMS Collaboration
- [24] *Sudakov form factors*
J.C. Collins
- [25] *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol., 1996 : QCD and collider physics*
R.K. Ellis, W.J. Stirling, B.R. Webber
- [26] *Phys. Letters B 136, 99-104, 1984*
S.D. Ellis, D.G. Richards, W.J. Stirling

- [27] *The Lund model*
B. Andersson
- [28] *Multiple interactions and the structure of beam remnants* T. Sjöstrand, P.Z. Skands
- [29] http://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/components/magnets/type_of_magnet.htm
- [30] *CERN/LHCC 2006-001 CMS TDR 8.1 2 February 2006*
Chapter 1. Introduction, 1.1 The LHC machine
- [31] <http://cms-service-lumi.web.cern.ch/cms-service-lumi/overview.php>
Page web de la collaboration CMS pour communiquer le suivi de l'expérience.
- [32] *CERN/LHCC 2006-001 CMS TDR 8.1 2 February 2006*
Chapter 1. Introduction, 1.3.1 Summary of detector requirements
- [33] *CERN/LHCC 2006-001 CMS TDR 8.1 2 February 2006*
Chapter 1. Introduction, 1.5 CMS : The overall concept
- [34] *CERN/LHCC 2006-001 CMS TDR 8.1 2 February 2006*
Chapter 4. Electromagnétique calorimeter p168
- [35] *CERN/LHCC 2000-38, CMS TDR 6.1 : CMS, The TriDAS Project, Technical Design Report, Volume 1 : The Trigger Systems*
The TriDAS Project
- [36] *Physics at the terascale, 18.2.6 : Trigger rates and downscale*
Ian Brock, Thomas Schörner-Sadenius
- [37] *Review of particle physics : Particle physics summary tables, mesons p38-76*
Particle Data Group
- [38] *CMS Twiki iterativeTracking*
- [39] <http://lcg.web.cern.ch>
- [40] *CMS PAS PFT-09-001 : Particle Flow Event Reconstruction in CMS and Performance for Jets, Taus and E_T^{Miss}*
The CMS collaboration
- [41] *(Tevatron) Run II Jet physics*
G.C. Blazey, J.R. Dittmann, S.D. Ellis, V.D. Elvira, K. Frame, S. Grinstein, R. Hirosky, R. Piegaia, R. Snihur, V. Sorin, D. Zeppenfeld

-
- [42] *Phys. Rev. Lett.* 39 (1977) 1436.
G. Sterman et S. Weinberg
- [43] *arXiv :hep-ph/0704.0292 : A practical Seedless Infrared-Safe Cone jet algorithm*
Gavin P. Salam and Grégory Soyez
5
- [44] *CERN TH.6775/93 : Longitudinally-Invariant $k_{\perp}$ -Clusterings Algorithms for Hadron-Hadron Collisions*
S. Catani, Yu.L. Dokshitzer, M.H. Seymour, B.R. Webber
- [45] *JHEP04 (2008) 063 : The Anti- k_t jet clustering algorithm*
M. Cacciari, G.P. Salam and G. Soyez
- [46] *Phys.Rev.Lett.*35, 1489-1492 (1975) *Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation*
M.L. Pearl et al.
- [47] *arXiv :hep-ex/0012035 : Observation of Tau Neutrino Interactions*
DONUT Collaboration
- [48] *[http ://operaweb.lngs.infn.it](http://operaweb.lngs.infn.it) : Site web de l'expérience Opera*
- [49] *Review of particle physics, Chapter 11 : The CKM Quark mixing matrix*
Particle Data Group
- [50] *arXiv :hep-ph/0104232 Electroweak radiative corrections in high energy processes*
M. Melles
- [51] *[http ://www.physics.rutgers.edu/~zrwan/gg_h_tautau/plot/Result/DYTauTau.html](http://www.physics.rutgers.edu/~zrwan/gg_h_tautau/plot/Result/DYTauTau.html)*
Z. Wan
- [52] *CMS PAS PFT-10-004 : Study of reconstruction algorithms using pp collisions data collected at $\sqrt{s} = 7\text{TeV}$*
The CMS collaboration
- [53] *CMS AN-2010/082 : Prospects for measurement of $\sigma(pp \rightarrow Z).B(Z \rightarrow \tau^+\tau^-)$ with CMS in pp Collisions at $\sqrt{s} = 7\text{ TeV}$*
M. Bachtis,S. Dasu, A. Savin
- [54] *CMS AN-2009/000 : The Tau Neural Classifier algorithm : tau identification and decay mode reconstruction using neural networks*
J. Conway, E. K. Friis, M. Squires, and C. Veelken
UC Davis

- [55] *CMS PAS PFT-08-001 : CMS Strategies for tau reconstruction and identification using Particle Flow techniques*
The CMS collaboration
- [56] *CMS AN-2010/074 : Estimation of background contributions to Tau analyses via Fake-Rate technique*
J. Conway, E. K. Friis, and C. Veelken
- [57] *<https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkbookPATWorkflow> : Physics Analysis Toolkit (PAT) Workflow*
- [58] *CMS AN-2009/111 : Generic Tag And Probe tools for measuring efficiencies at CMS with early data*
N. Adam, J. Berryhill, V. Halyo, A. Kunt, K. Mishra
- [59] *CMS PAS-TAU-2011-001 : Performance of tau reconstruction algorithms in 2010 data collected with CMS*
The CMS Collaboration
- [60] *CMS TAU-11-001 : Performance of τ -lepton reconstruction and identification in CMS*
The CMS Collaboration
- [61] *CMS PAS-SUS-11-007 : Search for physics beyond the Standard Model in events with opposite-sign tau pairs and missing energy*
The CMS Collaboration
- [62] *CMS CR-2011/153 : Search for Supersymmetry at CMS in all-hadronic final state*
Mattias Schröder for the CMS Collaboration

