

Université de Montréal

**Discrimination des quarks et des gluons dans les événements à quatre
jets**

par

Pierre-Hugues Beauchemin

Département de physique

Faculté des arts et des sciences

**Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en physique**

Septembre, 2000

©Pierre-Hugues Beauchemin, 2000



Université de Montréal

Département de physique et des sciences de la Terre

Date

QC
3
U54
2000
N. 022

Reçu

Faculté des arts et des sciences
Département de physique
Centre de recherche en physique

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de
Maîtrise en sciences (M.Sc.)
en physique

Septembre 2000



© Pierre-Hugues Benoit, 2000

Remerciements

Ce mémoire représente pour moi l'aboutissement de deux ans d'effort. Il a été rendu possible grâce à la complicité de nombreuses personnes à qui je dois une fière chandelle. En premier lieu je tiens à remercier cordialement mon directeur de recherche Monsieur Viktor Zacek, qui fut pour moi un maître à penser. En effet, par sa passion de la physique et son intérêt pour les diverses branches de ce domaine de connaissance, il a su m'ouvrir l'esprit à un univers beaucoup plus vaste que celui que j'envisageais à la fin de mon baccalauréat, contribuant ainsi à me former un esprit de synthèse. En m'apprenant également à développer une intuition des divers phénomènes physiques étudiés, j'ai pu ainsi saisir le dynamisme de la recherche scientifique. Monsieur Zacek m'a donc permis d'acquérir la maturité nécessaire aux études supérieures

Je tiens ensuite à remercier Monsieur Hannes Jeremie, responsable du groupe de recherche dans lequel je me suis intégré. Nos nombreuses discussions sur l'orientation de mon travail ont servi à donner un but à ce travail, une place dans l'immense effort international de recherche en physique des particules. C'est également à Monsieur Jeremie que je dois mes deux séjours au CERN qui furent pour moi autant de sources d'enrichissement intellectuel et personnel.

Je remercie également Monsieur Georges Azuelos pour tout l'aide tant académique que technique qu'il m'a offerte. Son sens critique exemplaire a su m'éviter de fâcheuses erreurs méthodologiques alors que ses connaissances informatiques m'ont permis de comprendre et d'utiliser des programmes tels que PYTHIA et JETNET. Je n'oublierai jamais sa patience et son affabilité (d'ailleurs il a accepté d'être mon codirecteur de recherche pour le doctorat).

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé:

**Discrimination des quarks et des gluons dans les événements à quatre
jets**

présenté par:

Pierre-Hugues Beauchemin

a été évalué par un jury composé des personnes suivantes:

Georges Azuelos, président-rapporteur
Viktor Zacek, directeur de recherche
Claude Leroy, membre du jury

Mémoire accepté le:

Sommaire

L'étude de la désintégration hadronique du boson Z^0 en quatre jets permet de tester la nature non-abélienne de la théorie de jauge de l'interaction forte, la chromodynamique quantique. En effet, la valeur de la constante de Casimir T_F est directement proportionnelle à la fraction d'événements à quatre quarks f_q contenus dans un échantillon d'événements à quatre jets. En prenant le rapport de f_q^{exp} sur f_q^{the} (ce qu'on définit comme R_{4q}) on peut ainsi obtenir une évaluation directe de la fiabilité du secteur CDQ du Modèle Standard.

Pour obtenir une telle mesure, il s'agit de séparer les événements $q\bar{q}q'\bar{q}'$ des événements $q\bar{q}gg$, puis de les compter. Or de récentes mesures de R_{4q} effectuées par la collaboration OPAL au CERN (OPAL-PN321) avec l'aide des corrélations angulaires Bengtsson-Zerwas, ont donné une valeur de $R_{4q} = 2.00 \pm 0.17(\text{stat})_{-0.36}^{+0.27}(\text{syst})$ ce qui est incompatible avec la prédiction de $R_{4q} = 1.00$ du Modèle Standard. Plusieurs causes peuvent expliquer ce surplus. Il est entre autre possible que cela vienne d'un effet indéterminé du détecteur ou de la méthode de mesure qui déformerait les résultats par l'analyse des corrélations angulaires entre les jets. Il est donc important de trouver une méthode d'évaluation de f_q^{exp} qui n'utilise pas de telles corrélations angulaires.

Pour ce faire, nous avons utilisé un ensemble de variables topologiques discriminant les événements de types $q\bar{q}q'\bar{q}'$ de ceux de types $q\bar{q}gg$. Parmi les variables retenues dans notre analyse, on compte la sphéricité, le thrust ainsi que la multiplicité de chaque jet. Un réseau de neurones est ensuite utilisé pour effectuer la discrimination voulue. Il s'agit en fait d'un perceptron multicouche "feed-forward" à apprentissage par rétro-propagation. La réponse que donnera le

réseau de neurones à un échantillon d'événements à quatre jets est caractérisée par une courbe de la pureté (p_q) en fonction de l'efficacité (ϵ_q). Cette courbe permettra de déduire le nombre d'événements à quatre quarks contenus dans un échantillon d'événements à quatre jets.

Mots clés: jets, discrimination, classification, variables, réseau de neurones

Table des matières

Sommaire	iii
Table des matières	v
Liste des tableaux	ix
Liste des figures	x
Remerciements	xiv
Introduction	1
Chapitre 1: Contexte théorique	4
1.1 Interaction electro-faible	4
1.1.1 Théorie des champs quantiques	4
1.1.2 Le Modèle Glashow-Salam-Weinberg	9
1.2 Interaction forte	14
1.2.1 Chromodynamique quantique	14
1.2.2 Modèle de partons	17
1.2.3 Hadronisation	18
Chapitre 2: Contexte expérimental	25
2.1 LEP	26

2.2	OPAL	29
2.2.1	Détecteur central	30
2.2.2	Calorimètre électromagnétique	32
2.2.3	Calorimètre hadronique	33
2.2.4	Détecteur à muons	33
2.2.5	Détecteurs avant	34
2.3	Traitement des données	34
2.4	Les succès du LEP	36
Chapitre 3: Processus CDQ à quatre partons		40
3.1	Configuration à quatre partons	40
3.1.1	Processus à quatre partons	40
3.1.2	Fractions apparentes des 4 quarks	42
3.1.3	Une porte au-delà du Modèle Standard	44
3.2	Phénoménologie des jets	45
3.3	Méthode de discrimination des $q\bar{q}q'\bar{q}'$	49
3.3.1	Angle Bengtsson-Zerwas	49
3.3.2	Variables topologiques	51
Chapitre 4: Variables topologiques		53
4.1	Critère de sélection des variables	53

4.2	Variables globales	57
4.2.1	Sphéricité	57
4.2.2	Sphérocité	59
4.2.3	Thrust	62
4.2.4	Acoplanarité	64
4.2.5	Moments de Fox-Wolfram	65
4.3	Variables de jets	71
4.3.1	Multiplicité	71
4.3.2	Impulsion transverse au carré	72
4.3.3	Énergie	74
4.3.4	Paramètre de résolution	74
Chapitre 5: Réseaux de neurones artificiels (ANN)		79
5.1	Genèse des réseaux de neurones	81
5.2	Fonctionnement d'un perceptron multicouche	84
Chapitre 6: Classification des événements à quatre jets à l'aide d'un réseau de neurones artificiel		90
6.1	Caractérisation du réseau de neurones utilisé	91
6.1.1	Architecture du réseau	91
6.1.2	JETNET	91
6.1.3	Caractérisation empirique	93

6.2 Analyse de la réponse du réseau	98
Conclusion	110
Bibliographie	112

Liste des tableaux

3-I	Définition du paramètre de résolution y_{ij} et du schème de recombinaison de différents algorithmes d'identification des jets.	46
4-I	Figure de mérite des sept variables globales sélectionnées pour discerner les événements à quatre quarks parmi ceux à 4-jets. . .	70
4-II	Figure de mérite des quatre variables de jet sélectionnées pour discerner les quarks des gluons.	78
6-I	Valeurs choisies des divers paramètres libres fixant la configuration du réseau de neurones que nous utiliserons pour notre analyse. . .	96

Liste des figures

1-1	Comparaison du schéma des lignes de champ de la force forte avec celui des lignes de champ de la force électromagnétique.	19
1-2	Étapes de la création d'événements multihadroniques issus de la désintégration du Z^0	20
2-1	Schéma du complexe d'accélérateurs conduisant au LEP.	27
2-2	Détecteur OPAL	30
2-3	Exemple de visualisation GOPAL d'un événement à quatre partons produit par simulation Monte Carlo.	36
2-4	Comparaison des mesures de la section efficace du pic de résonance du Z^0 en fonction de l'énergie avec les prédictions théoriques pour deux, trois et quatre familles de neutrino.	38
3-1	Diagrammes de Feynman des événements à quatre partons.	41
3-2	Nombre relatif de deux, trois et quatre jets reconstruits pour différentes valeurs de y_{cut}	48
3-3	Schéma de l'angle χ_{BZ} défini par les plans formés des jets 1,2 et 3,4.	51

4-1	Distribution de la fonction s de la sphéricité pour les événements à deux quarks et deux gluons (traits pleins) ainsi que pour les événements à quatre quarks (traits pointillés).	58
4-2	Distribution de la fonction D de la sphérocité pour les événements à deux quarks et deux gluons (traits pleins) ainsi que pour les événements à quatre quarks (traits pointillés).	61
4-3	Distribution du thrust principal pour les événements à deux quarks et deux gluons (traits pleins) ainsi que pour les événements à quatre quarks (traits pointillés).	63
4-4	Distribution de l'acoplanarité pour les événements à deux quarks et deux gluons (traits pleins) ainsi que pour les événements à quatre quarks (traits pointillés).	64
4-5	Distribution du deuxième moment de Fox-Wolfram pour les événements à deux quarks et deux gluons (traits pleins) ainsi que pour les événements à quatre quarks (traits pointillés).	66
4-6	Distribution du troisième moment de Fox-Wolfram pour les événements à deux quarks et deux gluons (traits pleins) ainsi que pour les événements à quatre quarks (traits pointillés).	67
4-7	Distribution du quatrième moment de Fox-Wolfram pour les événements à deux quarks et deux gluons (traits pointillés) ainsi que pour les événements à quatre quarks (traits pleins).	68
4-8	Distributions de la multiplicité normalisée ($\frac{m}{\ln E}$) du jet i ($i=1,2,3,4$) lorsque celui-ci est un gluon (traits pointillés) ainsi que lorsqu'il est un quark (traits pleins).	73

4-9	Distributions de l'impulsion transverse au carré normalisée du jet i ($i=1,2,3,4$) lorsque celui-ci est un gluon (traits pointillés) ainsi que lorsqu'il est un quark (traits pleins).	75
4-10	Distributions de l'énergie du jet i ($i=1,2,3,4$) lorsque celui-ci est un gluon (traits pointillés) ainsi que lorsqu'il est un quark (traits pleins).	76
4-11	Distributions de y^{34} pour les événements $q\bar{q}gg$ (traits pleins) ainsi que pour les événements $q\bar{q}q'\bar{q}'$ (traits pointillés).	77
5-1	Dépendance sur la température (pente) de la fonction de sortie sigmoïde.	87
5-2	Schéma de la propagation des données dans un réseau feed-forward à cinq entrées, une couche cachée et une sortie.	88
6-1	Courbe de la variation de la figure de mérite de la sortie du réseau en fonction du nombre de cycles d'entraînement (epoch).	97
6-2	Distribution de la multiplicité normalisée pour chacun des quatre jets lorsque l'événement est un $q\bar{q}gg$ (traits pleins) ou lorsque c'est un $q\bar{q}q'\bar{q}'$ (traits pointillés).	100
6-3	Distribution de l'impulsion transverse au carré pour chacun des quatre jets lorsque l'événement est un $q\bar{q}gg$ (traits pointillés) ou lorsque c'est un $q\bar{q}q'\bar{q}'$ (traits pleins).	101
6-4	Distribution de l'énergie pour chacun des quatre jets lorsque l'événement est un $q\bar{q}gg$ (traits pointillés) ou lorsque c'est un $q\bar{q}q'\bar{q}'$ (traits pleins).	102

- 6-5 Courbes de sortie du réseau de neurones. La courbe en traits pleins correspond à la sortie du réseau lorsqu'on lui présente des événements $q\bar{q}gg$, alors que la courbe en traits pointillés correspond à la réponse du réseau aux événements $q\bar{q}q'\bar{q}'$ 103
- 6-6 Courbe de l'efficacité en fonction de la pureté de la classification des événements à quatre jets par le réseau de neurones. 105
- 6-7 Courbes de sortie du réseau de neurones retraçant les événements à quatre partons. La courbe en traits pleins correspond à la sortie du réseau lorsqu'on lui présente des événements $q\bar{q} + q\bar{q}g$, alors que la courbe en traits pointillés correspond à la réponse du réseau aux événements $q\bar{q}gg + q\bar{q}q'\bar{q}'$ 108
- 6-8 Courbe de l'efficacité en fonction de la pureté de la classification des événements à quatre partons par le réseau de neurones. 109

Si Monsieur Azuelos m'a aidé dans mes problèmes techniques à Montréal, à Genève j'ai eu constamment recours à l'aide d'Anna Macchiolo pour la réalisation de nombreuses tâches. Je la remercie donc chaleureusement et lui souhaite la meilleure des chances pour son nouveau poste en Italie.

Je tiens à remercier mes amis Ervig, Jean-François, Jeff, Maxime et Thomas pour leur aide (correction, L^AT_EX, etc), mais surtout pour leur support et leur présence.

Je veux finalement adresser une pensée spéciale à ma famille que j'adore, c'est-à-dire à mon père Pierre, ma mère Marie, mes soeurs Geneviève et Marie-Élyse, ainsi qu'à Phanie, l'amour de ma vie, avec qui je compte un jour fonder une famille.

A tous, mille fois MERCI!

“Ainsi les idées, dont j’usais précédemment pour me figurer un cheval que je n’avais pas encore vu, étaient de purs signes, comme les empreintes sur la neige étaient des signes de l’idée de cheval: et on use des signes et des signes de signes dans le seul cas où les choses nous font défaut.”

Umberto Eco
(Le nom de la rose)

Introduction

Dans sa tentative de fournir une description aussi précise et complète que possible de la nature, la physique moderne a connu de retentissants succès. Son astuce: penser le monde comme une constitution de structures fondamentales indivisibles et décrire leurs interrelations. C'est là le paradigme de la physique des particules. La pierre de touche de cette description est la partition de tous les constituants fondamentaux de l'univers en deux classes distinctes: les bosons et les fermions. Bien que techniquement parlant ces termes renvoient à la nature symétrique ou anti-symétrique (respectivement) des fonctions d'onde décrivant l'état des particules formant ces deux classes, les fermions sont considérés comme les particules constituant la matière, alors que les bosons sont les particules qui véhiculent les interactions entre les divers constituants de la matière. Tout le reste de l'ensemble théorique décrivant le monde subatomique consiste à décrire ces interactions et les conséquences qu'elles engendrent.

Pour rendre compte de la panoplie des phénomènes observés, ces deux classes de particules ont elles-mêmes été subdivisées en sous-classes. C'est ainsi que l'on considère que les fermions sont formés de leptons (électron, muon, tau et leur neutrino associé) ainsi que de quarks (up, down, strange, charm, bottom et top), alors que les particules formant la classe des bosons se regroupent selon l'interaction qu'elles décrivent. On compte ainsi quatre de ces interactions fondamentales: l'interaction électromagnétique (véhiculée par le photon), l'interaction faible (assurée par les bosons massifs W^+ , W^- et Z^0), l'interaction forte (véhiculée par les gluons) et finalement l'interaction gravitationnelle (véhiculée par le graviton). Il faut cependant noter qu'en raison de sa faible portée sur les objets microscopiques, l'interaction gravitationnelle n'est pas jumellée au trois

autres dans les descriptions de la physique des particules. Le corpus théorique permettant de faire des prédictions physiques à partir de ces considérations, s'appelle le Modèle Standard. La portée de tout projet de recherche en physique des particules doit donc être située par rapport à cet ensemble théorique. C'est pourquoi le premier chapitre y constituera une introduction sommaire.

Or depuis sa genèse, le Modèle Standard a connu énormément de succès se reflétant dans la précision des vérifications de ses prédictions. Bon nombre de ces réussites sont imputables aux résultats obtenus par les expériences du LEP au CERN. Elles ont entre autre permis de:

1. mesurer avec précision la largeur du pic de résonance de la section efficace (σ) du boson de jauge Z^0 , permettant ainsi de prouver l'existence de seulement trois familles de neutrinos (ν_e , ν_μ et ν_τ);
2. mesurer avec précision l'angle de mélange $\sin(\theta_w)$, paramètre d'unification des interactions électromagnétique et faible;
3. mesurer α_s , la constante de couplage de l'interaction forte;
4. vérifier la consistance du Modèle Standard;
5. établir les contraintes sur une éventuelle mesure de la masse du quark top (qui a été confirmé récemment à CDF) ainsi que celle du boson de Higgs.

Comme le présent travail prendra sa signification à partir de ce contexte expérimental, le deuxième chapitre sera donc consacré à la présentation des expériences du LEP et à l'exposition de leurs principaux résultats.

Cependant si les prédictions concernant la physique de l'interaction électrofaible peuvent être effectivement déduites de la théorie, il est plus difficile d'étudier directement l'interaction forte à cause de son confinement à basse énergie (c'est-à-dire dans les hadrons comme le proton, par exemple). Une des astuces permettant

l'étude de cette interaction est d'analyser la formation d'amas de traces de particules qu'on appelle jets dans le détecteur et de considérer ces objets comme la représentation empirique de structures plus fondamentales: les quarks. L'étude de l'interaction forte entre les quarks passera donc par l'étude de la structure des jets dans un événement donné. Il est entre autre intéressant d'étudier la structure des événements à quatre jets car cela permet de spécifier certains paramètres fondamentaux de la théorie de l'interaction forte: les constantes de Casimir ou facteurs de couleur. Le troisième chapitre présentera donc une introduction à cette étude qui constituera en fait l'origine de la problématique qui nous intéresse: construire une méthode de discrimination des événements à quatre jets en deux classes distinctes: $q\bar{q}gg$ ou $q\bar{q}q'\bar{q}'$.

Or le fondement de notre méthode de classification de ces deux types d'événements sera le pouvoir de séparation d'un ensemble de variables topologiques. Le quatrième chapitre sera donc destiné à la présentation de ces variables. Cependant, l'originalité de la méthode présentée sera d'utiliser un réseau de neurones pour tenir compte simultanément du pouvoir de séparation de chacune de ces variables. C'est pourquoi on réservera le cinquième chapitre à une brève introduction de cet outil d'analyse. On montrera finalement au sixième chapitre comment on peut, à partir des variables topologiques sélectionnées et d'un réseau de neurones correctement configuré, obtenir une évaluation du nombre d'événements à quatre quarks contenus dans un échantillon d'événements à quatre jets. On verra, à travers tout ce travail, que cela constitue en fait une méthode permettant de mettre en évidence la possibilité pour de la physique au-delà du Modèle Standard; mais avant d'en arriver là, attardons nous tout d'abord à ce contexte théorique qu'est le Modèle Standard.

Chapitre 1

Contexte théorique

1.1 Interaction electro-faible

1.1.1 Théorie des champs quantiques

Le cadre conceptuel sur lequel le Modèle Standard se fonde pour décrire les particules fondamentales et leurs interactions est la théorie des champs quantiques. En effet, cette théorie permet de bien décrire le comportement des systèmes quantiques relativistes, ce qui est nécessaire compte tenu du fait que les phénomènes décrits par la physique des particules se produisent à très petite échelle ($< 1 \text{ fm}$) et à hautes énergies (ou mettent en jeu des vitesses proches de celle de la lumière comme dans le cas de la désintégration β). L'avantage d'une telle théorie sur une mécanique quantique relativiste est que sa formulation covariante la rend automatiquement invariante de Lorentz et qu'elle constitue une théorie à plusieurs particules. En effet, comme en théorie des champs chaque particule de même nature constitue une excitation d'un même champ défini en tout point de l'espace-temps [9], cette théorie permet donc d'interpréter de façon satisfaisante le fait que les relations d'Einstein $E = mc^2$ et d'Heisenberg $\Delta E \cdot \Delta t \geq \hbar$ permettent la création des paires particule-antiparticule spontanément en un même point de l'espace (la relation d'incertitude spécifiant que si l'énergie n'est pas suffisante, les particules créées seront virtuelles); ce qui n'est pas le cas de la mécanique quantique relativiste qui est forcée d'admettre dans son interprétation des états d'énergie négative ainsi que d'autres inconsistances comme le viol de la causalité

relativiste. Le formalisme de la théorie des champs fournit également un cadre naturel permettant de faire des calculs sur les transitions entre états de nombres de particules différents. Finalement, cette théorie, en considérant le produit de champs différents (au même point espace-temps, pour préserver la causalité) permet de rendre compte des diverses transitions entre des états de particules de nature différentes dans ses équations. Tout cela donne donc de bonnes indications montrant que la théorie des champs et son formalisme constituent le cadre théorique adéquat à la description des phénomènes en physique des hautes énergies.

Prenons l'exemple de l'électrodynamique quantique (EDQ) qui est la théorie formulée en terme de champs permettant de décrire l'interaction électromagnétique. Le lagrangien décrivant cette interaction est donné par [20]:

$$\mathcal{L}_{EDQ} = \bar{\psi}(i \not{\partial} - m)\psi - \frac{1}{4}(F_{\mu\nu}^2) - e\bar{\psi}\gamma^\mu\psi A_\mu \quad (1.1)$$

Ce lagrangien permettrait en fait de décrire toute la physique des particules si seule l'interaction électromagnétique existait. En effet, si on applique l'équation d'Euler-Lagrange de chaque champ à ce lagrangien, on obtiendra les équations de Maxwell, l'équation de Dirac pour des fermions libres et une équation d'interaction photons-fermions. Or considérer les termes d'interaction séparément des termes libres revient en fait à se placer dans une représentation interactive où les interactions entre les fermions sont traitées comme des termes de correction sur leur mouvement libre. Les prédictions physiques sont alors faites à partir de l'hamiltonien du système par calcul des perturbations, méthode qui est également utilisée en mécanique quantique non-relativiste [3]. Ce formalisme permet donc d'obtenir le degré de précision voulu sur les prédictions physiques (en choisissant l'ordre de nos calculs). Le formalisme de la théorie des champs consistera alors à développer une méthode de calcul permettant, à partir d'un lagrangien décrivant la dynamique du système étudié, d'obtenir des prédictions physiques par calcul des perturbations.

Comme les interactions entre particules élémentaires sont étudiées à partir d'expériences de diffusion, les quantités mesurables que l'on peut calculer à l'aide du formalisme de la théorie des champs sont les sections efficaces (σ), les taux de désintégration (Γ) et les temps de vie (τ). Pour définir une section efficace, considérons deux paquets de particules de longueur ℓ_A et ℓ_B , de densité ρ_A et ρ_B et de section transverse A , diffusant l'un sur l'autre. La section efficace est alors définie par [9]:

$$\sigma \equiv \frac{\text{Nombre de particules diffusées}}{\rho_A \rho_B \ell_A \ell_B A} \quad (1.2)$$

Comme le nombre de particules diffusées dépend de la probabilité que l'état initial à m particules diffuse pour donner l'état final à n particules, on a donc que la section efficace est proportionnelle au carré de l'élément de la matrice de transition de l'état initial à l'état final, c'est-à-dire que $\sigma \propto |\mathcal{M}|^2$. En fait, pour une diffusion de deux électrons initiaux (orientés, disons, selon l'axe z) en deux particules finales de masse m dans le centre de masse, on a que:

$$\left(\frac{d\sigma}{d\Omega}\right)_{CM} = \frac{|\mathcal{M}(p_A, p_B \rightarrow p_1, p_2)|^2}{64\pi^2 E_{cm}^2} \cdot \sqrt{1 - \frac{4m^2}{E_{cm}^2}} \quad (1.3)$$

On peut de même définir le taux de désintégration par [9]:

$$\Gamma \equiv \frac{\text{Nombre de désintégrations de particules } \mathcal{A} \text{ par unité de temps}}{\text{Nombre de particules } \mathcal{A} \text{ présentes}} \quad (1.4)$$

ce qui peut se calculer de la façon suivante:

$$\Gamma = \frac{(2\pi)}{2m_A} \int \left(\prod_f d^3 p_f \frac{1}{2E_f} \right) 01 \delta^{(4)}(p_A - \Sigma p_f) |\mathcal{M}(m_A \rightarrow p_f)|^2 \quad (1.5)$$

Notons que si on ne considère pas la somme sur tous les états finaux possibles, mais qu'on considère plutôt un état final donné, on obtiendra alors le taux de

branchement du canal de désintégration donné. Bien sûr la somme sur tous les canaux de branchement possibles redonne le taux de désintégration de la particule A . La dernière grandeur mesurable, le temps de vie, est simplement défini par l'inverse du taux de désintégration. Ainsi, section efficace, taux de désintégration et temps de vie constituent les grandeurs mesurables et calculables permettant à la théorie des champs de faire des prédictions physiques vérifiables expérimentalement.

Pour calculer ces grandeurs à partir d'un lagrangien comme celui de l'EDQ, il faut pouvoir calculer ces éléments de la matrice de transition entre l'état initial et l'état final. Pour ce faire, il s'agit d'utiliser les règles de Feynman pour dessiner les diagrammes possibles décrivant un processus à l'ordre voulu, de traduire ensuite ces diagrammes en produits de propagateurs et de vertex associés, de sommer la contribution de chacun des diagrammes considérés, d'élever au carré en sommant sur les degrés de liberté du système (spin, couleur) et finalement d'intégrer sur les impulsions des boucles internes [9].

Or ce calcul mène parfois à des infinis. En effet, il existe deux types de diagrammes de Feynman: les diagrammes d'arbre et les diagrammes à boucles. Ceux du premier type constituent les diagrammes dominant un processus donné, alors que ceux du second type fournissent les corrections d'ordres supérieurs sur les amplitudes des termes principaux de cette même réaction. C'est ainsi qu'on appelle "corrections radiatives" l'ensemble des diagrammes d'ordre supérieur, comportant une ou plusieurs boucles. On s'attend donc à ce que la valeur des amplitudes de ces diagrammes soit proportionnelle à une puissance donnée (dépendamment de l'ordre considéré) de la faible constante de couplage ($\frac{1}{137}$ à 1 eV pour l'interaction électromagnétique), n'ajoutant ainsi qu'une petite correction à la valeur de l'amplitude de la somme des diagrammes dominants (pour la réaction considérée). Cependant les calculs sur les diagrammes de ce genre conduisent à des infinités. Ainsi, pour éviter que les prédictions physiques découlant de ces

calculs perdent tout leur sens, il faut adjoindre à notre théorie un processus qu'on appelle *renormalisation* et qui permet de rendre toutes nos prédictions physiques finies et consistantes. Ce procédé d'élimination des infinis est justifié par le fait que le calcul des perturbations ne permet qu'une approximation de la prédiction physique et que si on veut être cohérent, il faut tenir compte de tous les ordres pour faire une prédiction exacte. Il constitue donc un critère de consistance interne nécessaire à toute théorie des interactions fondamentales.

Tout le formalisme développé pour la physique de l'interaction électromagnétique reste également valable pour la physique des autres interactions: chaque interaction peut correctement être décrite par un champ interagissant avec des fermions. Il s'agit en fait de trouver le lagrangien adéquat à la description d'une interaction donnée. Une fois ce lagrangien obtenu, on peut en déduire les prédictions physiques à l'aide du formalisme de la théorie des champs. De plus, il a été montré (t'Hooft, 1971) que chacune de ces théories est renormalisable. La théorie des champs constitue donc un outil de description efficace des diverses interactions fondamentales car elle permet de faire des calculs effectifs pour obtenir des prédictions physiques mesurables. Tout le travail consiste donc à trouver ces lagrangiens qui modélisent une interaction donnée.

1.1.2 Le Modèle Glashow-Salam-Weinberg

On a dit qu'une particule est décrite par un état d'excitation d'un champ (c'est-à-dire par un des opérateurs d'échelle de création (a_p^\dagger) constituant le champ) en un point donné de l'espace-temps. Ainsi, en appliquant un champ sur le vecteur d'état du vide, on obtiendra un vecteur d'état à n particules d'impulsion p_1, p_2, \dots, p_n , correspondant aux n excitations de ce champ [9]. Comme l'état $|p_1 p_2 \dots p_n\rangle$ ainsi produit est un vecteur dans un espace d'Hilbert généralisé (qu'on appelle espace de Fock), on a donc qu'un champ est un opérateur agissant sur cet espace. Or, tout comme en mécanique quantique [3], ces opérateurs sont définis à une phase près, c'est-à-dire que la transformation suivante

$$\psi \rightarrow \psi' = e^{i\alpha} \psi \quad (1.6)$$

effectuée sur un champ fermionique devra laisser le lagrangien (ainsi que les équations qui en découlent) invariant sous cette transformation [10]. On dira alors que le lagrangien est invariant sous une transformation de jauge globale. Ce type de transformation constitue en fait la clé de voûte du Modèle Standard. En effet, si on demande que la phase de la transformation de jauge dépende des coordonnées x^μ , c'est-à-dire si on impose une transformation de jauge locale au champ, on aura que le lagrangien d'un système libre ne sera pas invariant. En effet, prenons le lagrangien pour des particules de Dirac libres, c'est-à-dire:

$$\mathcal{L}_{Dirac} = \bar{\psi}(i \not{\partial} - m)\psi \quad (1.7)$$

et appliquons aux champs la transformation de jauge locale $\psi \rightarrow \psi' = e^{i\alpha(x)}\psi$ ainsi que son complexe conjugué. Le lagrangien deviendra alors [20]:

$$\begin{aligned} \mathcal{L}' &= i\bar{\psi}' \not{\partial} \psi' - m\bar{\psi}' \psi' \\ &= i e^{-i\alpha(x)} \bar{\psi} \not{\partial} e^{i\alpha(x)} \psi - m e^{-i\alpha(x)} \bar{\psi} e^{i\alpha(x)} \psi \\ &= i\bar{\psi} \not{\partial} \psi - m\bar{\psi} \psi - \bar{\psi} \gamma^\mu \partial_\mu \alpha(x) \psi \\ &\neq \mathcal{L} \end{aligned}$$

Par ailleurs, on sait par l'électromagnétisme classique, que la transformation de jauge du potentiel électromagnétique est donnée, dans sa formulation covariante, par [33]:

$$A_\mu \rightarrow A'_\mu = A_\mu + \partial_\mu \chi(\vec{x}, t) \quad (1.8)$$

où χ est une fonction quelconque de l'espace et du temps, c'est-à-dire que χ est une phase locale. Ainsi, si on remplace la dérivée partielle ∂_μ dans le lagrangien par la dérivée covariante $D_\mu \equiv \partial_\mu + ieA_\mu$ et qu'on impose simultanément une transformation de jauge sur les champs ψ , $\bar{\psi}$ et A_μ , le lagrangien restera invariant. Cependant, le nouveau lagrangien

$$\mathcal{L} = \bar{\psi}(\not{D} - m)\psi \quad (1.9)$$

n'est plus celui qui décrit des particules libres, mais bien celui qui, par l'ajout du terme $\bar{\psi}\gamma^\mu\psi A_\mu$, décrit l'interaction entre fermions et photons. On parle alors du champ A_μ comme d'un champ de jauge. Il représente, dans le cas de l'électromagnétisme, les photons.

Or lorsque les équations du mouvement sont laissées invariantes par une transformation on dit que cette transformation est une symétrie du système. Ainsi, demander par exemple que le lagrangien électromagnétique soit invariant sous une transformation de jauge locale (c'est-à-dire qu'il ne soit pas modifié par l'ajout d'une phase complexe) revient à lui imposer la symétrie du groupe de transformation $U(1)$. On conclut donc que, simplement en imposant une symétrie de jauge à notre système, on peut retrouver tout le formalisme de l'électrodynamique quantique, c'est-à-dire que pour déterminer le lagrangien décrivant l'interaction électromagnétique, il suffit d'imposer la symétrie de jauge $U(1)$ au système. C'est ce qu'on appelle la théorie de Yang-Mills [20]. Mais qu'en est-il des autres interactions fondamentales? Peut-on, en imposant une symétrie de jauge, déduire le lagrangien décrivant l'interaction faible? La réponse est positive et c'est ce qui fait que les transformations de jauge constituent la clé de voûte du Modèle Standard.

En fait, il existe une théorie de jauge qui permet de décrire simultanément l'interaction électromagnétique et l'interaction faible. La symétrie qui permet cette théorie de jauge est celle associée au groupe $SU(2) \otimes U(1)$. La transformation de jauge découlant de cette symétrie est donc définie ainsi [20]:

$$\psi \rightarrow \psi' = \exp \left(igI_k \alpha^k + ig' \frac{1}{2} Y \alpha^0 \right) \psi \quad (1.10)$$

où les I_k sont les composantes de l'isospin faible (partie $SU(2)$) et sont proportionnelles aux matrices de Pauli ($I_k = \tau_k/2$), où Y est l'hypercharge faible (partie $U(1)$) et où g et g' sont les constantes de couplage associées aux deux groupes de transformation. On parle alors d'une théorie de l'interaction électrofaible car les deux interactions sont unifiées en une seule. Cette théorie est cependant plus complexe que le simple cas de l'électromagnétisme pour deux raisons:

1. le groupe $SU(2)$ est non-abélien, ce qui autorisera des interactions entre les bosons médiateurs de l'interaction faible;
2. les bosons médiateurs de l'interaction faible doivent avoir une masse pour rendre compte de la faiblesse de cette interaction comparativement à la force électromagnétique.

Or dans une théorie de jauge, toutes les particules sont sans masse. Il faut donc trouver un mécanisme permettant de donner une masse aux particules. C'est ce qu'accomplit le mécanisme de Higgs. Il s'agit en fait d'introduire un nouveau champ scalaire à quatre états, le champ de Higgs donné par [8]:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \eta \end{pmatrix} \exp \left(ig \frac{1}{2} \tau^k \theta_k / v \right) \quad (1.11)$$

et d'ajouter au lagrangien les termes d'interaction de ce nouveau champ avec les fermions et avec les bosons de jauge. L'ajout de ces termes brisera la symétrie. On parlera alors de théories de jauge à symétrie spontanément brisée. On dit qu'il y a brisure spontanée de symétrie lorsque les équations régissant un phénomène

présentent une symétrie sans que celle-ci reste apparente au niveau du phénomène observé à proprement dit. Cela permet notamment d'expliquer pourquoi, à notre échelle, les forces électromagnétique et faible nous apparaissent comme deux interactions fondamentalement différentes, même si elles viennent préalablement d'une structure commune. Notons finalement que les trois états décrit par θ_k seront absorbés par trois bosons de jauge qui deviendront massifs. Il restera alors un boson massif η observable physiquement: le boson de Higgs.

Le lagrangien complet de l'interaction électrofaible avec brisure spontanée de symétrie est donc [20]:

$$\begin{aligned}
\mathcal{L} &= \mathcal{L}_G + \mathcal{L}_{F-G} + \mathcal{L}_{H-G} + \mathcal{L}_H & (1.12) \\
&= -\frac{1}{4}\text{Trace}[W^{\mu\nu}W_{\mu\nu}] - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} \\
&\quad - \sum_F \bar{\psi}_F \gamma^\mu D_\mu \psi_F \\
&\quad + (D^\mu \phi)^\dagger (D_\mu \phi) \\
&\quad + \mu^2 (\phi^\dagger \phi) - \lambda (\phi^\dagger \phi)^2
\end{aligned}$$

avec:

$$W_k^{\mu\nu} = \partial^\mu W_k^\nu - \partial^\nu W_k^\mu - g\epsilon^{lm} W_l^\mu W_m^\nu \quad (1.13)$$

$$B^{\mu\nu} = \partial^\mu B_\nu - \partial^\nu B_\mu \quad (1.14)$$

et

$$D_\mu = \partial_\mu - igI_k W_\mu^k - ig'\frac{1}{2}YB_\mu \quad (1.15)$$

Les W sont les champs de jauge de $SU(2)$ alors que B est celui de $U(1)$. Cependant, ils ne représentent pas encore les trois bosons de l'interaction faible et le photon. C'est en fait une combinaison de ces quatre champs qui servira à décrire les véritables bosons d'interaction. Ainsi [20],

$$W_\mu^+ = (-W_\mu^1 + iW_\mu^2)/\sqrt{2} \quad (1.16)$$

$$W_\mu^- = (-W_\mu^1 - iW_\mu^2)/\sqrt{2} \quad (1.17)$$

$$Z_\mu^0 = \frac{gYB_\mu + g'W_\mu^0}{\sqrt{g'^2 + g^2Y^2}} \quad (1.18)$$

$$A_\mu = \frac{g' B_\mu - g Y W_\mu^0}{\sqrt{g'^2 + g^2 Y^2}} \quad (1.19)$$

Pour simplifier on définit:

$$\sin \theta_w = \frac{g}{\sqrt{g^2 + g'^2}} \quad \text{et} \quad \cos \theta_w = \frac{g'}{\sqrt{g^2 + g'^2}} \quad (1.20)$$

Comme $g \sin \theta_w$ est le couplage électromagnétique que l'on connaît déjà, on a donc $g \sin \theta_w = e$. De plus, comme le photon n'a pas de masse, on a que

$$g \sin \theta_w = g' \cos \theta_w \quad (1.21)$$

ainsi que

$$M_W = M_Z \cos \theta_w \quad (1.22)$$

L'angle de Weinberg (θ_w) constitue donc un angle de mélange des symétries $SU(2)$ et $U(1)$ permettant d'exprimer les deux constantes de couplage en fonction d'un seul et même paramètre. C'est ainsi qu'on peut parler d'une unification des interactions faible et électromagnétique. La valeur de cet angle n'est cependant pas prédite par la théorie et doit être déterminée par l'expérience. Il constitue l'un des 24 paramètres libres du Modèle Standard.

On a donc vu que pour trouver le lagrangien adéquat à la description des phénomènes impliquant les interactions faible et forte il suffit d'imposer que le système soit invariant de jauge sous la symétrie du groupe $SU(2) \otimes U(1)$ et que, moyennant une nouvelle particule et un paramètre libre, on peut unifier ces deux interactions. Reste maintenant à voir ce qu'il en est de l'interaction forte.

1.2 Interaction forte

1.2.1 Chromodynamique quantique

La chromodynamique quantique est la théorie de l'interaction forte qui est responsable de la cohésion entre les quarks à l'intérieur des hadrons. Chaque quark est en fait porteur d'un nombre quantique particulier ayant trois valeurs possibles qu'on appelle "couleurs". Les champs de quarks sont donc représentés par un spineur $\psi_j^{(\alpha)}$ où $\alpha = u, d, s, c, b, t$ est la saveur et où $j = 1, 2, 3$ (correspondant à bleu, rouge et vert) est l'indice de la couleur. C'est sur ce degré de liberté qu'agit l'interaction forte via le boson médiateur de cette force: le gluon. La couleur est cependant un degré de liberté interne aux hadrons qui a d'abord été introduit pour assurer l'antisymétrisation complète de la fonction d'onde des quarks qui les constituent (puisque'ils sont des fermions). Ainsi les types de combinaisons possibles des spineurs de quarks assurant l'invariance de la fonction d'onde des hadrons sous une transformation de couleur sont: $\bar{q}^i q_i$ représentant les mésons, $\epsilon^{ijk} q_i q_j q_k$ représentant les baryons et finalement $\epsilon_{ijk} \bar{q}^i \bar{q}^j \bar{q}^k$ qui représente les anti-baryons [9]. On remarque que ces fonctions d'onde sont sommées sur tous les états de couleur, ce qui fait que les baryons sont des singulets de couleur. Il est donc impossible de voir directement ce degré de liberté sur les particules détectables, ce qui fait que la couleur reste un degré de liberté interne aux particules mesurables.

Bien que, contrairement à l'interaction électrofaible, l'interaction forte n'a pu encore être unifiée aux autres, on peut quand même la déduire de l'invariance d'un lagrangien sous une transformation de jauge appliquée aux champs de quarks. La symétrie permettant la transformation adéquate est celle du groupe $SU(3)$. Cette transformation de jauge est donc donnée par [20]:

$$\psi \rightarrow \psi' = \exp\left(\frac{i}{2} \sum_k \lambda_k \alpha_k\right) \psi \quad (1.23)$$

nécessitant la définition de la dérivée covariante suivante [20]:

$$\partial_\mu \rightarrow D_\mu \equiv \partial_\mu + \frac{i}{2}g \sum_{j=1}^8 \lambda_j G_\mu^j. \quad (1.24)$$

pour avoir invariance des équations du système. Les λ_k représentent les huit matrices hermétiques de traces nulles du groupe $SU(3)$ dans la représentation 3×3 , alors que les G_μ^j sont les champs des huit bosons de jauge de l'interaction forte. Bien que, tout comme dans le cas des photons, les gluons sont des bosons de jauge sans masse (la CDQ est une généralisation de l'EDQ au groupe $SU(3)$) le caractère non-abélien du groupe de symétrie $SU(3)$ implique une distinction fondamentale entre la nature du photon et celle des gluons: ces derniers sont porteurs d'une charge de couleur. En fait, chaque gluon porte une couleur et une anticouleur, ce qui leur permet d'interagir ensemble. En effet, le champ de jauge de l'interaction forte se transforme ainsi:

$$G_\mu^j \rightarrow G'^j_\mu = G_\mu^j - \frac{1}{g} \partial_\mu \alpha_j - \sum_{k,l} f_{jkl} \alpha_k G_\mu^l \quad (1.25)$$

où les f_{jkl} sont les constantes de structure de l'algèbre de Lie du groupe $SU(3)$ ($[\lambda_j, \lambda_k] = 2if_{jkl}\lambda_l$) [9, 20, 28, 8]. Or comme le terme d'interaction des gluons dans le lagrangien CDQ (l'équivalent du terme générant les équations de Maxwell en EDQ) est donné par $-\frac{1}{4}\text{Trace}[F^{\mu\nu}F_{\mu\nu}]$ et que le terme $F^{\mu\nu}$ est défini par:

$$[D_\mu, D_\nu]\psi = -igF_{\mu\nu}\psi \quad (1.26)$$

$$\Rightarrow F_{\mu\nu}^j = \partial_\mu G_\nu^j - \partial_\nu G_\mu^j - g \sum_{k,l} f_{jkl} G_\mu^k G_\nu^l \quad (1.27)$$

on a donc que le lagrangien contiendra non-seulement des termes en G^2 comme en EDQ, mais également des termes en G^3 et G^4 , correspondant, par les règles de Feynman, à des graphes à trois ou quatre gluons, c'est-à-dire à des interactions entre les gluons.

Cet aspect fondamental de la théorie a des conséquences observables. En effet, les calculs effectués avec les théories de jauge non-abéliennes font intervenir des constantes caractéristiques du groupe de symétrie qu'on appelle constantes

de Casimir. On peut les définir à partir des générateurs $(T^j)_{ab}$ de l'algèbre de Lie du groupe considéré de la façon suivante [9, 28, 23]:

$$\sum_{j=1}^{N_A} (T^j T^{j\dagger})_{ab} = \delta_{ab} C_F \quad (1.28)$$

$$\sum_{j,k=1}^{N_A} (f^{jkl} f^{jkm}) = \delta^{lm} C_A \quad (1.29)$$

$$\sum_{j,k=1}^{N_A} T_{ab}^j T_{ba}^k = \delta_{ab} T_F \quad (1.30)$$

Dans le cas de $SU(3)$ on a que

$$C_F = 4/3 \quad C_A = 3 \quad T_F = 1/2 \quad (1.31)$$

Or les amplitudes relatives des processus $q \rightarrow qg$, $g \rightarrow gg$ et $g \rightarrow q\bar{q}$ dépendent respectivement des facteurs C_F, C_A et T_F . Ainsi, une mesure empirique de la valeur de ces constantes pour l'interaction forte permet de tester le caractère non-abélien de cette interaction.

Pour écrire le lagrangien de l'interaction forte complètement invariant sous la transformation de jauge (1.23), il s'agit d'appliquer le même procédé utilisé pour écrire le lagrangien de l'EDQ. Le résultat est le suivant [20, 8]:

$$\mathcal{L}_{CDQ} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi - \frac{1}{2}g \sum_j (\bar{\psi}\gamma^\mu \lambda_j \psi) G_\mu^j - \frac{1}{4} \sum_j G_{\mu\nu}^j G_j^{\mu\nu} \quad (1.32)$$

Avec la construction des lagrangiens donnés aux équations (1.12) et (1.32), on est arrivé à une théorie complète des trois interactions fondamentales. Chaque élément de cette théorie a été décrit comme un champ quantique découlant de l'exigence de covariance et d'invariance de jauge de la théorie sous les transformations du groupe de symétrie $SU(3) \otimes SU(2) \otimes U(1)$. C'est là toute l'essence du Modèle Standard. On a également vu que la théorie des champs nous permettait de faire de façon non-équivoque toutes les prédictions physiques nécessaires pour rendre compte des mesures possibles en physique des hautes énergies. Cela n'est cependant pas complètement vrai pour le cas de l'interaction forte. Regardons pourquoi de plus près.

1.2.2 Modèle de partons

On a vu que les prédictions physiques sont faites à partir d'un lagrangien interactif par le processus de Feynman qui est une forme de calcul perturbatif. Ainsi comme chaque terme d'ordre supérieur est proportionnel à une puissance de la constante de couplage, pour que le calcul soit possible, c'est-à-dire pour que les calculs convergent vers une approximation de la solution exacte en considérant un nombre fini de termes, il faut que cette constante de couplage soit plus petite que 1. Sinon, on dit que le système est non-perturbatif, c'est-à-dire qu'il faut considérer un nombre infini de termes pour pouvoir effectuer une prédiction physique, enlevant au calcul par perturbation son pouvoir prédictif. Or justement, à faible énergie ($q^2 < 1 \text{ GeV}$), la constante de couplage de l'interaction forte (α_s) est plus grande que 1. La théorie de l'interaction forte n'est donc pas une théorie perturbative. Cependant la CDQ perturbative permet quand même de faire des prédictions importantes pour les hadrons finaux.

En effet, on a dit que pour qu'une théorie ait du sens il faut qu'elle soit renormalisable. Or cela consiste à inclure dans les calculs de prédictions physiques tous les diagrammes des processus qu'une détection ne distinguerait pas, éliminant ainsi tous les infinis possibles. Pour ce faire, il s'agit de redéfinir les paramètres physiques et de les évaluer expérimentalement à une certaine échelle d'énergie. Toutes les autres valeurs que ces paramètres peuvent prendre seront alors en fonction des valeurs établies expérimentalement à cette échelle d'énergie.

Les paramètres physiques qu'il faut renormaliser sont: la masse, l'énergie et les constantes de couplage de chacune des interactions fondamentales. On parle alors de couplage effectif. Pour connaître l'expression renormalisée de la constante de couplage forte (en fonction d'une variable d'échelle d'énergie μ), il faut résoudre l'équation du groupe de renormalisation suivante pour la constante

α_s [9, 2, 23]:

$$\mu^2 \frac{\partial \alpha_s}{\partial \mu^2} = -b_0 \alpha_s^2 \quad (1.33)$$

$$b_0 = \frac{33 - 2n_f}{12\pi} \quad (1.34)$$

dont la solution est:

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + \alpha_s(\mu^2) b_0 \ln(Q^2/\mu^2)} \quad (1.35)$$

et remplacer $\alpha_s(\mu^2)$ par la valeur de la mesure de la constante de couplage forte à l'échelle prévue. On voit donc que $\alpha_s = \alpha_s(Q^2)$ et que si Q^2 augmente, alors α_s diminuera. Il existe donc un seuil d'énergie (Λ) pour lequel α_s devient forte de sorte que pour toutes les échelles d'énergie supérieures, la constante de couplage de l'interaction forte sera faible, rendant possible les calculs par perturbation. Cela apparaît clairement si on exprime la constante de couplage forte en fonction du paramètre d'échelle fixe Λ au lieu du paramètre arbitraire μ :

$$\alpha_s(Q^2) = \frac{2\pi}{b_0 \log(Q/\Lambda)} \quad (1.36)$$

Cette propriété de la constante de couplage s'appelle liberté asymptotique car les quarks et les gluons deviennent alors libres. Or comme les grandes valeurs de Q^2 correspondent aux échelles de courtes distances, on considère donc que les quarks et les gluons sont libres dans les hadrons. On appelle ces particules asymptotiquement libres des partons. La nature partonique des quarks et des gluons à l'intérieur des hadrons a été corroborée par les expériences de diffusion profondément inélastique.

1.2.3 Hadronisation

Comme la couleur est un degré de liberté interne aux hadrons et que les partons portent une charge de couleur, les partons sont donc des particules "piégées" dans les hadrons. On dit qu'ils sont confinés. Ainsi, les hadrons

constituent des systèmes de quarks liés par des gluons. La propriété de confinement est en fait l'inverse de la liberté asymptotique et est liée à l'augmentation de la valeur de α_s avec la distance.

Pour bien comprendre le confinement, on peut se représenter l'état interne d'un méson comme un état de quark-antiquark (singulet de couleur) lié par un potentiel de la forme:

$$V(r) = -\frac{4}{3} \frac{\alpha_s}{r} + \lambda r \quad (1.37)$$

où λ ($\approx 1 \text{ GeV fm}^{-1}$) est relié au confinement. Ainsi, pour de petites distances ($r \leq 0.1 \text{ fm}$) α_s est petit et le potentiel est de forme coulombienne. On se retrouve alors dans une situation complètement analogue à l'EDQ, ce qui fait que la CDQ perturbative devient applicable, nous permettant de considérer que l'interaction entre deux quarks est médiée par l'échange d'un gluon. Par contre, pour de grandes distances ($r \geq 1 \text{ fm}$), α_s devient plus grand que 1 et le potentiel croît linéairement, comme si les quarks étaient reliés par un ressort. À cause des interactions gluon-gluon, les lignes de champ sont concentrées en une *corde*, comme le présente le schéma de la figure 1-1. Lorsque l'énergie emmagasinée dans

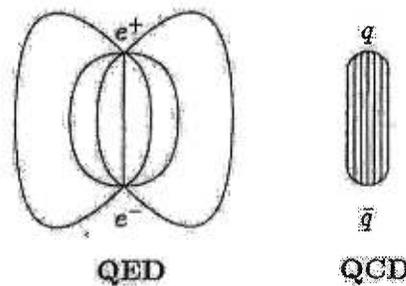


Figure 1-1. Comparaison du schéma des lignes de champ de la force forte avec celui des lignes de champ de la force électromagnétique.

le champ de force devient plus grande que deux fois la masse d'un quark, la corde se brise, libérant l'énergie emmagasinée sous forme de création d'une paire quark-antiquark. On dit qu'il y a eu *fragmentation* [2, 34]. Ce processus se produira

tant et aussi longtemps qu'il y aura de l'énergie disponible dans le système. La fragmentation aura donc pour conséquence de transformer un état lié de quark-antiquark en deux gerbes de particules se dirigeant en directions opposées. On appelle ces gerbes de particules des *jets*.

À partir de ce schéma représentatif, on peut séparer la description de la production d'événements multihadroniques issus de la désintégration du boson de jauge Z^0 en quatre étapes distinctes [29, 23]:

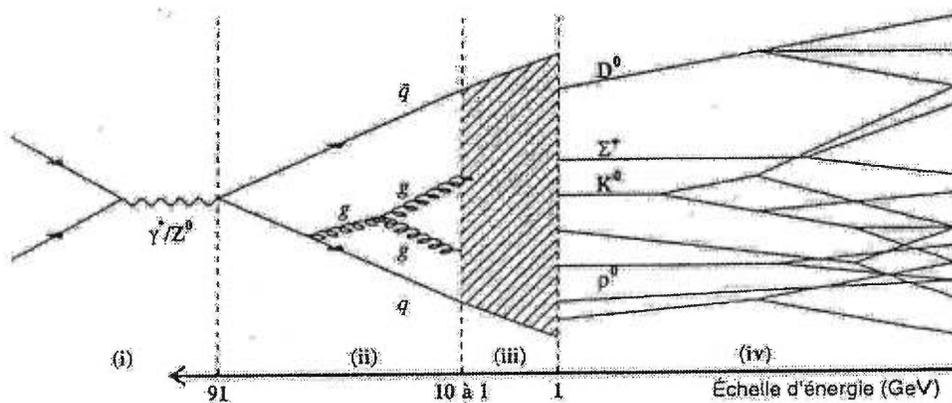


Figure 1-2. Étapes de la création d'événements multihadroniques issus de la désintégration du Z^0 .

1. **Production d'une paire $q\bar{q}$** Au cours de cette étape, le Z^0 produit par l'annihilation d'un électron avec un positron (à 90 GeV dans le centre de masse) se désintègre en une paire quark-antiquark. Ce processus est décrit par la théorie électrofaible.
2. **Niveau partonique perturbatif** Dans cette seconde phase, les quarks parents émettent des gluons durs qui constitueront alors autant de partons. Ces gluons peuvent également se matérialiser en une paire de $q\bar{q}$. À ce niveau α_s est faible de sorte que tous les calculs peuvent être faits à partir de la CDQ perturbative. On peut donc obtenir une évaluation partonique

de la section efficace de la désintégration multihadronique du Z^0 . Notons que le nombre de partons présents à la fin de cette phase dépend de l'ordre considéré dans les calculs de l'élément de matrice de transition (\mathcal{M}).

3. **Hadronisation et fragmentation** Cette phase constitue la phase non-perturbative du processus. Aucun calcul ne peut être fait pour prévoir l'état final. C'est au cours de cette étape que les jets sont formés à partir des partons produits à l'étape précédente. Cette fragmentation des liaisons partoniques en jets d'hadrons, également appelée hadronisation, peut être représentée par trois modèles empiriques différents permettant de reproduire les données à partir des partons prévus par la CDQ. Ces modèles sont [2]:
- L'hadronisation indépendante* Ce modèle d'hadronisation a été originellement présenté par Feynman et Field en 1978 pour fournir une description utile de plusieurs caractéristiques du comportement des jets. Le mécanisme en est simple: considérons qu'un quark q_1 initial s'éloigne rapidement du reste du système de couleur. Le confinement l'empêchera alors de s'échapper du système de sorte qu'une partie de l'énergie contenue dans le champ de couleur se matérialisera en une paire $q_2\bar{q}_2$ dont l'antiquark se recombinera avec le quark initial pour former un méson sans couleur. La 4-impulsion du quark q_1 sera alors partagée de façon probabiliste entre le méson créé et le quark q_2 restant, la fonction de probabilité étant donnée par:

$$f(z) = 1 - a - 3a(1 - z)^2 \quad (1.38)$$

où z est la fraction d'impulsion transmise au méson. La fragmentation procédera alors de façon récursive sur le quark q_2 restant et s'arrêtera lorsque le quark q_n restant ne sera plus assez énergétique pour développer un autre méson. Cette chaîne d'hadronisation s'applique simultanément et indépendamment à chaque parton initial produit lors de la phase perturbative. Les quarks restants à la fin de chaque chaîne de fragmentation seront appariés entre eux de sorte qu'il n'y ait plus de partons libres. Ce modèle

est problématique en ce qu'il ne permet pas de conserver simultanément l'énergie et l'impulsion longitudinale.

Hadronisation de gerbes Ce modèle a été proposé comme alternative au précédent par Odorico, Fox et Wolfram en 1980. Dans ce modèle, les partons initiaux sont considérés comme des particules virtuelles ayant la même masse que le jet qu'ils produiront. Les équations d'Altarelli-Parisi [9] prédisent alors que ces partons initiaux évoluent vers un état moins énergétique en émettant des gluons par radiation Bremsstrahlung suivant la fonction de fragmentation suivante:

$$P_{q \rightarrow qg}(z) = \frac{4}{3} \left(\frac{1+z^2}{1-z} \right) \quad (1.39)$$

où z est la fraction de 4-impulsion portée par le quark (pour le gluon sera donc $1-z$). Le processus se poursuivra jusqu'à ce que l'énergie des partons originaux soit suffisamment faible. On se retrouvera alors avec une série de gerbes contenant un amas de particules possédant une charge de couleur. Ces particules se combineront alors de façon imprévisible en hadrons mesurables, formant du coup autant de gerbes de particules que de partons initiaux.

Hadronisation de cordes Ce modèle, développé par Anderson en 1980, est basé sur l'idée de champ d'interaction de type *corde* comme présenté plus haut. En fait, il s'agit de considérer que le flux chromodynamique entre deux quarks partoniques devient non-négligeable lors de cette phase d'hadronisation non-perturbative. Ce champ de couleur constituant la corde de force, se scindera alors en deux pour diminuer son énergie interne selon le même processus que celui décrit par le processus de fragmentation. Évidemment, la fragmentation des diverses cordes continuera tant qu'il y aura de l'énergie disponible pour former de nouveaux hadrons. L'énergie non-matérialisée restante servira alors à alimenter l'oscillation harmonique des quarks dans les hadrons, manifestant ainsi les degrés de liberté internes. Anderson a

montré en 1983 que la fragmentation des cordes se produit selon la fonction suivante:

$$f(z) = \frac{(1-z)^a}{z} \exp\left(\frac{-b(m^2 + p_t^2)}{z}\right) \quad (1.40)$$

où z est encore une fois la fraction de 4-impulsion portée par le quark, m est la masse du hadron formé, p_t son impulsion transverse par rapport à l'axe du jet ainsi formé et finalement a et b constituent des paramètres ajustables permettant de reproduire les données. On remarque que cette fonction ne dépend pas de la nature du parton parent (quark ou anti-quark). Elle ne favorise donc pas un bout ou l'autre de la corde primaire lors de la fragmentation. On dit alors que cette fonction a une symétrie droite-gauche. La fonction de fragmentation ne variera donc que pour des cordes de masses différentes.

Comme chacun des modèles qui viennent d'être présentés comportent des paramètres libres, cela indique donc que ces modèles ne sont que des représentations construites pour ajuster les prédictions théoriques calculables aux états finaux effectivement mesurables. On les utilise surtout dans les générateurs d'événements Monte Carlo.

4. **Désintégration des particules instables** Les hadrons produits lors de l'étape précédente ont des temps de vie différents. Si ce temps de vie est trop court, la particule ne parcourra pas une distance suffisante dans le détecteur pour être mesurée. Il faut donc plutôt considérer ses diverses produits de désintégration dans les prédictions physiques. Bien que cela puisse effectivement être calculé par la théorie de l'interaction électrofaible ou par celle de l'interaction forte, il est plus pratique d'utiliser, pour les prédictions physiques, les différents rapports d'embranchement fournis par l'expérience.

Cette étude des différentes phases d'une réaction comportant des quarks ainsi que des différents modèles de fragmentation met en évidence l'importance

d'étudier les jets. En effet, on a vu que les jets étaient le produit de l'hadronisation des partons. Ainsi, bien que la CDQ ne permette pas de prévoir un état final déterminé, elle permet cependant de prévoir les jets. On dit alors que la prédiction de la CDQ sur la section efficace d'une réaction est inclusive car elle porte sur la formation de jets dont le contenu n'est pas connu, c'est-à-dire qui inclut ou admet dans sa prédiction tous les contenus de particules possibles issus de l'hadronisation. Cependant toutes les propriétés que l'on peut calculer sur les partons de l'étape perturbative sont également valables pour les jets (énergies, impulsions, etc). On a donc que:

$$\sigma(e^+e^- \rightarrow q\bar{q}) \approx \sigma(e^+e^- \rightarrow 2\text{jets}) \quad (1.41)$$

Or comme $\sigma(e^+e^- \rightarrow q\bar{q})$ est calculable pour l'étape perturbative, les quarks sont donc considérés comme libres et on a

$$\sigma(e^+e^- \rightarrow q\bar{q}) = \sigma(e^+e^- \rightarrow \mu^+\mu^-) \cdot 3 \cdot \sum_f Q_f^2 \quad (1.42)$$

Cela on sait le calculer! Bien sûr l'hadronisation ajoute une certaine correction ou incertitude (car on ne peut évaluer cette correction) sur les propriétés de jets prévus par la CDQ, mais cette théorie fournit quand même l'élément le plus important du calcul. Les jets constituent donc une manifestation observable des propriétés de la CDQ, permettant à la théorie d'être confrontée à l'expérience.

Avant de voir pourquoi cela constitue un intérêt de recherche pour nous, arrêtons-nous un peu sur le contexte expérimental de cette recherche.

Chapitre 2

Contexte expérimental

Pour comprendre le contexte dans lequel les différents tests ou simulations de tests du Modèle Standard peuvent être effectués, il faut connaître les expériences qui peuvent permettre de tels tests. En physique des hautes énergies, les complexes expérimentaux sont formés d'un accélérateur et d'un détecteur. L'accélérateur détermine les réactions qui sont produites et donc que l'on peut étudier, alors que le détecteur détermine ce qui est mesurable ainsi que la précision des différentes mesures possibles. Accélérateur et détecteur précisent donc le cadre dans lequel s'effectue une recherche en physique des hautes énergies.

Or un des complexes d'accélérateurs qui a le plus contribué aux succès du Modèle Standard est celui du LEP au CERN. Opérationnel depuis 1989, le complexe LEP (LEP II depuis 1996) constitue une véritable usine à bosons de jauge électrofaibles permettant d'étudier avec précision un grand nombre de phénomènes découlant des réactions $e^+e^- \rightarrow Z^0$ (LEP) et $e^+e^- \rightarrow W^+W^-$ (LEP II), y compris certains phénomènes d'ordres supérieurs en CDQ. Le LEP fournit donc un cadre permettant l'étude des événements à quatre jets et c'est pourquoi, avant d'aborder l'importance d'une telle étude, on va tout d'abord décrire brièvement en quoi consiste l'accélérateur du LEP ainsi qu'un de ces détecteurs: le détecteur OPAL. Nous soulignerons également au passage les principales réussites des expériences au LEP.

2.1 LEP

Le LEP est un anneau d'accélération de 27 km de circonférence, situé à environ 100 m sous terre et qui est à cheval sur la frontière Franco-Suisse en banlieue de la ville de Genève. Cependant, lorsqu'on parle du LEP, on réfère en fait à un complexe d'accélérateurs [20, 27, 34, 29]. En effet, tout le processus d'accélération des particules s'effectue le long d'une chaîne d'accélérateurs dont le LEP n'est que l'étape finale. Les électrons sont d'abord générés puis accélérés à une énergie de 200 MeV dans un accélérateur linéaire (le LIL). Une partie des électrons est alors projetée sur une cible de tungstène permettant la production des positrons. Électrons et positrons sont ensuite accélérés jusqu'à 600 MeV avant d'être stockés par paquets dans un anneau d'accumulation (le EPA). Ces paquets seront injectés dans un premier synchrotron (le PS) puis dans un second (le SPS), où ils atteindront 20 GeV. Les paquets d'électrons et de positrons seront finalement injectés quatre par quatre dans le LEP où ils seront amenés à leur énergie maximale [27, 29]. La figure 2-1 présente un schéma du complexe LEP en ses principales composantes.

L'anneau d'accélération du LEP consiste en fait en un anneau de stockage c'est-à-dire en un synchrotron servant à accélérer et stocker des particules [27, 34]. Il constitue également un collisionneur car il possède deux anneaux de faisceaux opposés qui se rencontrent en plusieurs zones d'interaction. Ces zones correspondent aux quatre détecteurs situés sur le parcours du LEP, soit: ALEPH, DELPHI, L3 et OPAL. Ainsi, pour assurer sa double fonction d'accélérateur et de collisionneur, le LEP est formé de deux sections avec des cavités de radiofréquence qui permettent l'accélération des particules à chacun de leur passage, de 3 368 dipôles permettant de maintenir les paquets de particules sur des orbites circulaires et de 808 quadrupôles assurant la focalisation des paquets sur toute la circonférence [29, 23].

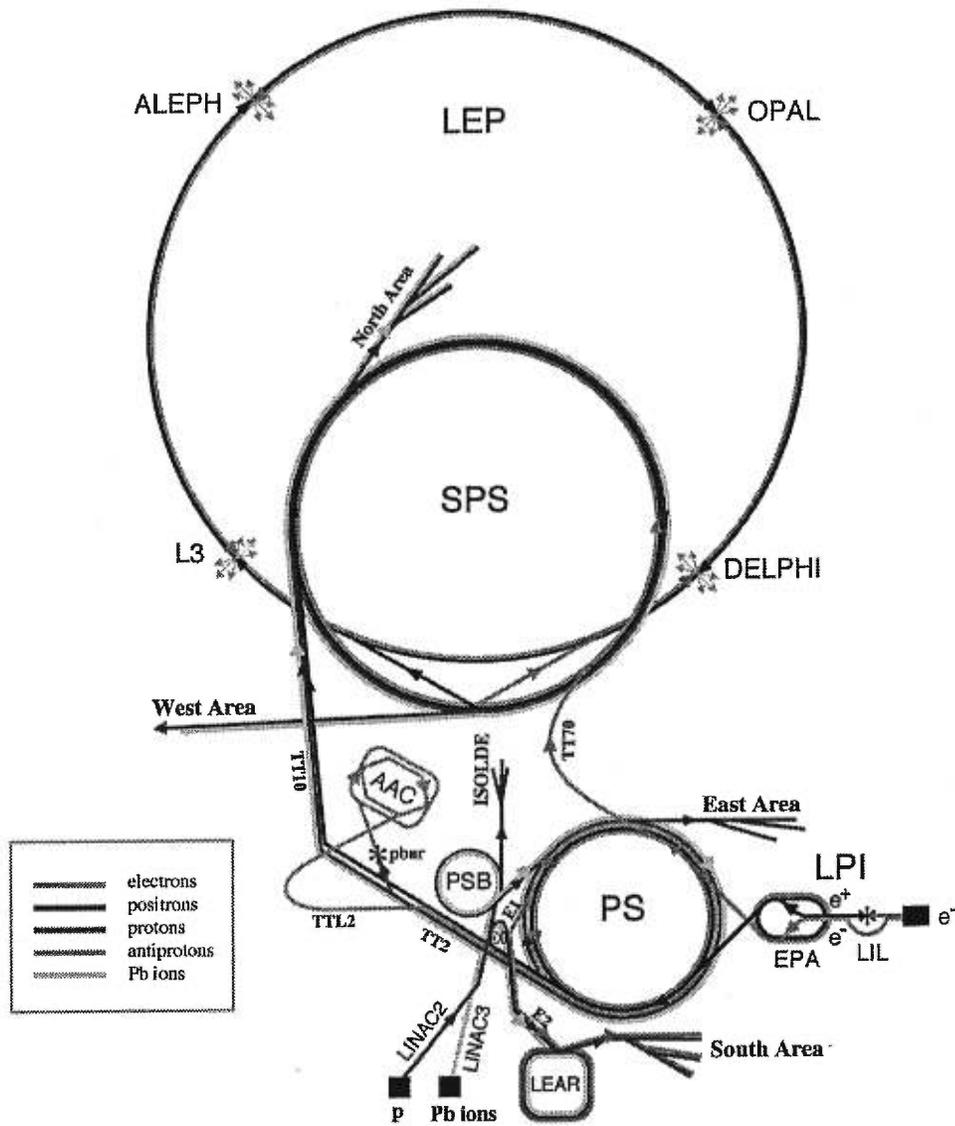


Figure 2-1. Schéma du complexe d'accélérateurs conduisant au LEP.

Bien que le LEP ait été originalement conçu pour atteindre des énergies d'un peu plus de 90 GeV nécessaires pour la production du Z^0 , des énergies de 206 GeV dans le centre de masse (103 GeV par faisceau) ont récemment été obtenues avec le complexe LEP II. En fait, la grande circonférence du LEP ainsi que la puissance maximale des aimants permettent au LEP d'atteindre des énergies de 240 GeV [20, 27, 14]. Cependant, comme l'accélérateur est circulaire, les électrons accélérés perdent beaucoup d'énergie sous forme de radiation synchrotrone — dont la perte est donnée par: $\Delta E = \frac{4\pi}{3} \alpha \frac{\hbar c}{r} \gamma^4 \beta^3$ — ce qui limite l'énergie maximale à environ 200 GeV.

La gamme d'énergie ainsi accessible par l'accélérateur aura une incidence sur le taux de production des différents bosons de jauge produits au LEP. En effet, ce taux de production dépend de la section efficace de la réaction considérée, section efficace qui elle-même dépend de l'énergie à laquelle les particules incidentes sont accélérées. Comme notre travail consistera à analyser la désintégration hadronique du Z^0 , on ne s'intéressera qu'aux événements qui se produisent à 92 GeV, ce qui correspond à l'énergie pour laquelle la section efficace de la réaction $e^+e^- \rightarrow Z^0$ est maximale.

Cependant, connaître la section efficace d'une réaction n'est pas suffisant pour déterminer son taux de production, il faut également connaître la luminosité de l'accélérateur. Cette luminosité est définie par [20, 27, 34, 14]:

$$\mathcal{L} = \frac{N_a N_b f}{a} \quad (2.1)$$

où $N_{a,b}$ est le nombre de particules contenues dans chaque paquet accéléré, f la fréquence des collisions et a la surface des faisceaux à la zone de collision. Pour le cas de la production de Z^0 on aura donc que le taux de réaction est déterminé par:

$$\text{taux} = \mathcal{L} \cdot \sigma(e^+e^- \rightarrow Z^0) \quad (2.2)$$

Ainsi, avec une fréquence de collision des paquets de $22\mu s$, le LEP (I) peut atteindre une luminosité de plus de $1.7 \times 10^{31} \text{ cm}^{-2} \text{ s}^{-1}$ [14], assurant un taux

de production de Z^0 de l'ordre de $0.4s^{-1}$. Énergie dans le centre de masse et luminosité sont donc les facteurs déterminants qui ont permis de produire jusqu'à maintenant près de 4.3 millions de Z^0 .

2.2 OPAL

Le détecteur OPAL¹ (Omni Purpose Apparatus for LEP) est, comme son nom l'indique, un appareil multifonctionnel construit dans le but de détecter avec grande précision tous les types d'interaction produite lors des collisions $e^+e^- \rightarrow Z^0$ au LEP. Il s'agit en fait d'un détecteur cylindrique composé de plusieurs sous-détecteurs répartis en couches successives assurant une couverture d'environ 4π d'angle solide autour du point d'impact. Parmi ces sous-détecteurs on compte: un détecteur central, un calorimètre électromagnétique, un calorimètre hadronique, un détecteur à muons et deux détecteurs avants. Ils permettent entre autre de:

- reconstruire les traces des particules chargées,
- déterminer leur impulsion,
- identifier la nature des particules chargées (en déterminant dE/dx),
- identifier les électrons et les photons et mesurer leur énergie,
- reconstruire les vertex primaires et secondaires dans la région d'interaction e^+e^- ,
- mesurer l'énergie des hadrons,
- identifier les muons,
- mesurer la luminosité absolue par la détection des réactions de diffusion Bhabha avec les détecteurs avants.

¹Toute la description du détecteur qui va suivre s'en réfère à: [25, 26, 34, 29]

Le schéma du détecteur est présenté à la figure 2-2.

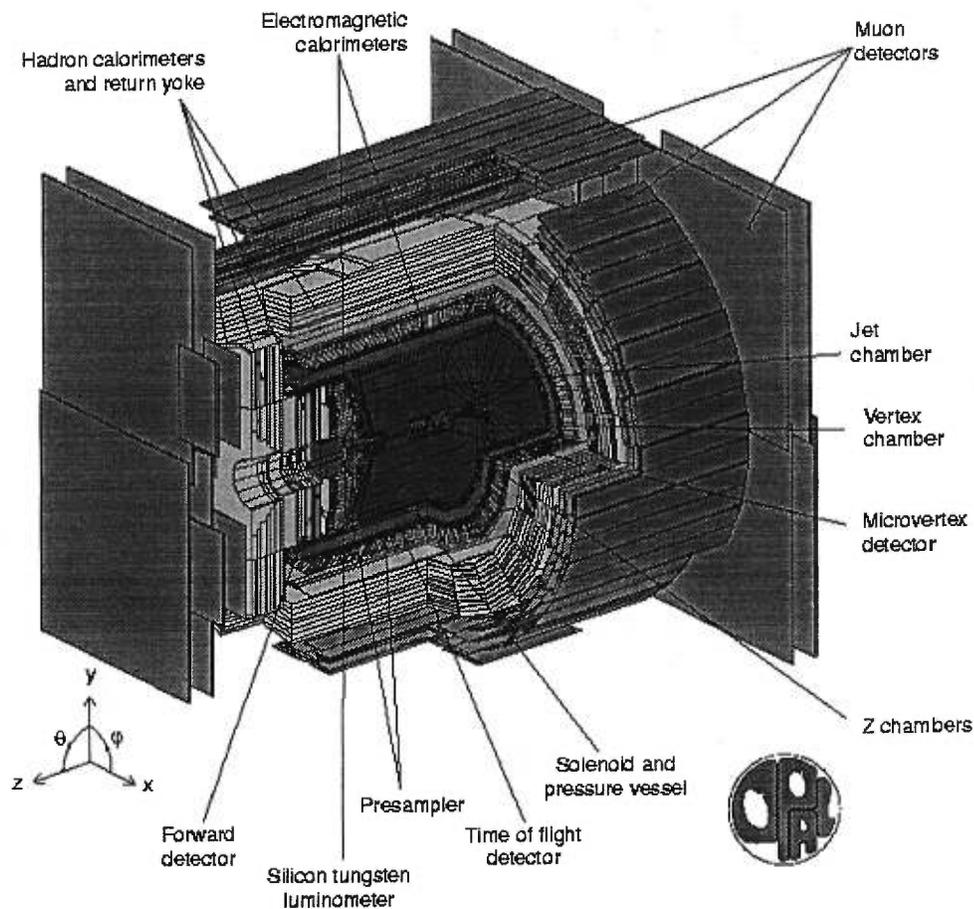


Figure 2-2. Schéma d'ensemble du détecteur OPAL.

2.2.1 Détecteur central

C'est la partie centrale du détecteur qui comprend un détecteur microvertex au silicium et trois chambres à dérive: le détecteur vertex, la chambre à jets et les chambres Z. Toute cette partie du détecteur est entourée d'un solénoïde produisant un champ magnétique axial de 0.435 T qui déviara les particules chargées, permettant, par le calcul de la courbure, de déterminer leur impulsion.

Détecteur microvertex

Ce détecteur est composé de deux barils concentriques constitués d'un ensemble de lattes. Chacune de ces lattes est composée de trois détecteurs à microbandes orientés de façon à permettre une lecture tridimensionnelle des coordonnées des traces chargées. La très bonne résolution spatiale (de l'ordre de $5 \mu\text{m}$ pour la coordonnée ϕ et de $12 \mu\text{m}$ dans la direction z) de cette géométrie permet donc de mesurer avec grande précision les vertex des mésons lourds à très courte durée de vie (de l'ordre de $1 \times 10^{-12} \text{s}$) ainsi que d'extrapoler celle des leptons tau ($3 \times 10^{-13} \text{s}$).

Détecteur vertex

Le détecteur à vertex est une chambre à dérive de 1 m de long sur 0,47m de diamètre contenant un mélange gazeux (88.2% d'argon, 9.8% de méthane et 2.0% d'isobutane) à une pression de quatre bars permettant, par ionisation du gaz, de tracer la trajectoire des particules chargées. Il est constitué de deux étages ayant chacune 36 sections. La combinaison de l'information provenant de ces cellules, dont celles du premier étage sont axiales tandis que celles du deuxième sont stéréoscopiques, permet une mesure précise de la coordonnée z près de la région d'interaction.

Chambre à jets

Ce détecteur consiste en une chambre à dérive cylindrique de 4 m de longueur sur 3.7 m de diamètre externe et dont le diamètre interne est fixé à 0.5 m par la chambre à vertex qui lui est concentrique. Constituée de 24 sections de 159 fils d'anode tendus parallèlement à la direction du faisceau, la chambre à jets permet de mesurer les coordonnées dans le plan $r\phi$ de l'interaction entre une particule chargée et le gaz (qui est le même que celui de la chambre à vertex) en calculant le temps de dérive des électrons produits par ionisation du gaz jusqu'aux fils d'anode. La coordonnée z est, quant à elle, obtenue en calculant le

rapport de la charge intégrée (déduite de $\Delta V = \frac{\int dQ}{C}$, où C est la capacité du fil d'anode), accumulée aux deux extrémités du fil d'anode touché. Finalement, en additionnant ces deux valeurs de la charge intégrée, on peut calculer la perte d'énergie dE/dx de la particule dans la chambre. Ainsi, cette connaissance des mesures des coordonnées $r\phi z$ nous permet de déterminer l'impulsion et la trajectoire des particules, alors que celle du dE/dx nous permet une bonne identification des particules.

Chambres Z

Les chambres Z sont des panneaux de chambres à dérive de 4 m de long formant un baril qui ceinture la chambre à jets. Chacun des panneaux est constitué de 8 cellules de 50 cm sur 50 cm et chaque cellule est formée de 6 fils d'anode tendus perpendiculairement à l'axe z , ce qui permet une mesure directe de cette coordonnée. Les chambres Z sont donc conçues pour donner une précision sur la mesure de la coordonnée z comparable à la précision de la mesure des coordonnées du plan $r\phi$ par les chambres à vertex et à jets. Cela permet également d'améliorer la résolution de la mesure de la masse invariante des particules.

2.2.2 Calorimètre électromagnétique

Ce détecteur a été conçu pour identifier les électrons et les photons ainsi que pour mesurer l'énergie déposée par les gerbes électromagnétiques créées par ces particules. Il est composé d'un baril et de bouchons de verre de plomb couvrant 98% de tout l'angle solide entourant l'aimant du détecteur central. Vu la présence de cet aimant et d'une enceinte pressurisée entre le détecteur central et le calorimètre électromagnétique, les gerbes commencent à se développer avant d'atteindre le calorimètre. Un prééchantillonneur composé de scintillateurs est donc installé avant les blocs de verre au plomb, permettant ainsi de mesurer la

position des gerbes et ainsi que l'énergie perdue.

2.2.3 Calorimètre hadronique

Le calorimètre hadronique mesure l'énergie des hadrons qui émergent du calorimètre électromagnétique et aide à l'identification des muons (parce que les muons vont laisser une trace beaucoup plus fine que la gerbe produite par les hadrons dans ce sous-détecteur). Il est construit en trois sections: le baril, les bouchons et les pôles. Le baril est constitué de 9 couches de chambre à fils opérant en mode étincelles, alternées avec 8 plaques de fer, jouant le rôle d'absorbeurs. Les bouchons sont pour leur part constitués de 36 chambres où circule un mélange gazeux (75% d'isobutane et 25% d'argon), alors que les pôles sont constitués de chambres minces à multifils, augmentant la couverture de l'angle solide de $|\cos \theta| = 0.91$ à 0.99 . En principe, seuls les muons et les neutrinos sont capables de traverser le calorimètre hadronique.

2.2.4 Détecteur à muons

Le détecteur à muons est le sous-détecteur le plus éloigné du point d'interaction des faisceaux. Il est formé d'un baril et de deux bouchons (comme la configuration cylindrique du détecteur l'impose) couvrant 93% de l'angle solide du détecteur. Le baril est composé de 110 chambres à dérive de 1.2 m de largeur sur 90 mm d'épaisseur. Les bouchons sont constitués de quatre couches de chambres à fils perpendiculaires à l'axe du faisceau. Le détecteur à muons permet ainsi une efficacité d'identification des muons isolés de plus de 3 GeV, sur 93% de l'angle solide, de 100% et une probabilité de moins de 1% qu'un pion isolé de 5 GeV soit identifié comme un muon.

2.2.5 Détecteurs avant

Les détecteurs avant sont situés près du tube à faisceaux aux deux extrémités du détecteur. Ils sont constitués d'un calorimètre avant, d'une chambre à tubes (chambre formée de trois plans de tubes de bronze espacés de 1 mm, recueillant les événements collinéaires aux faisceaux), d'une chambre à dérive et d'un préchantillonneur. Leur utilité est de permettre la mesure de la luminosité du LEP en détectant les électrons et les positrons produits lors de la diffusion Bhabha aux angles rasant l'axe des faisceaux ainsi que les électrons qui proviennent des interactions $\gamma\gamma$.

2.3 Traitement des données

Une fois que les données ont été recueillies par le détecteur, il faut les traiter. En effet, il faut éliminer l'information redondante, reconstruire la trajectoire des diverses particules à l'aide de l'information fournie par chaque sous-détecteur et déterminer les quadrivecteurs énergie-impulsion de chaque particule.

Une première sélection de données est réalisée dès leur acquisition. Cette présélection sert en fait à maximiser la probabilité que l'événement observé corresponde à un processus physique pertinent, issu de la désintégration d'un boson de jauge et non d'un signal de bruit de fond. Le système assurant cette présélection est le déclencheur ou "trigger". Ce "trigger" constitue en fait un ensemble de conditions déterminant les événements qui seront enregistrés par le système d'acquisition de données (DAQ). Les trois principales conditions imposées par ce système de déclenchement sont:

1. l'énergie enregistrée par le calorimètre électromagnétique doit être d'au moins 6 GeV,
2. au moins deux traces doivent être détectées,

3. la trace des particules doit pointer en direction du point d'interaction.

Si le "trigger" permet de sélectionner des événements avant leur enregistrement, le travail de reconstruction est, quant à lui, réalisé après l'enregistrement des données, essentiellement grâce à un ensemble de programmes informatiques. Le programme traitant les données pour reconstruire les événements détectés s'appelle ROPE [24]. ROPE constitue en fait un ensemble de modules offrant un cadre standard pour l'analyse d'événements OPAL. Parmi les routines qui composent l'ensemble du programme, il y a celles qui permettent de contrôler les processeurs, (ensemble de routines permettant de calculer les quantités physiques relatives à chacun des sous-détecteurs), celles qui relient l'information traitée par les processeurs des différents sous-détecteurs pour former les traces entières des particules détectées et finalement celles, plus générales, qui permettent une analyse physique de l'événement observé [24]. Toute l'information brute traitée par ROPE sera ensuite enregistrée sur un DST (Data Summary Tape) qui contiendra donc l'information nécessaire pour interpréter la physique d'un événement comme l'énergie visible, la quadri-impulsion et la courbure d'une trace chargée, l'énergie des gerbes électromagnétiques et hadroniques, la présence de muons, etc.

Comme les différents éléments du détecteur ainsi que les différents modules de reconstruction produisent la mesure d'une grandeur physique donnée, il faut donc tenir compte des effets que ces instruments et méthodes de mesure ont sur une prédiction physique pour que celle-ci puisse être pertinemment comparée à la mesure expérimentale obtenue. Cela peut se faire par simulation. Ainsi avant de comparer une prédiction théorique calculée à l'aide de la CDQ perturbative avec sa mesure expérimentale il faut non seulement simuler l'effet physique de l'hadronisation, mais aussi l'effet technique du processus de mesure. Ces simulations se font en deux étapes. La première utilise un générateur Monte Carlo comme PYTHIA 6.0 (qui est un assemblage du PYTHIA originale et de JETSET) pour modéliser l'hadronisation, tandis que les effets du détecteur sont tenus en

compte par le programme GOPAL qui est une version du programme GEANT3 appliqué au détecteur OPAL. La figure 2-3 présente une simulation GOPAL d'un événement à quatre jets. JETSET, GOPAL et ROPE nous permettent donc de faire une prédiction physique complète et de la comparer à une mesure expérimentale.

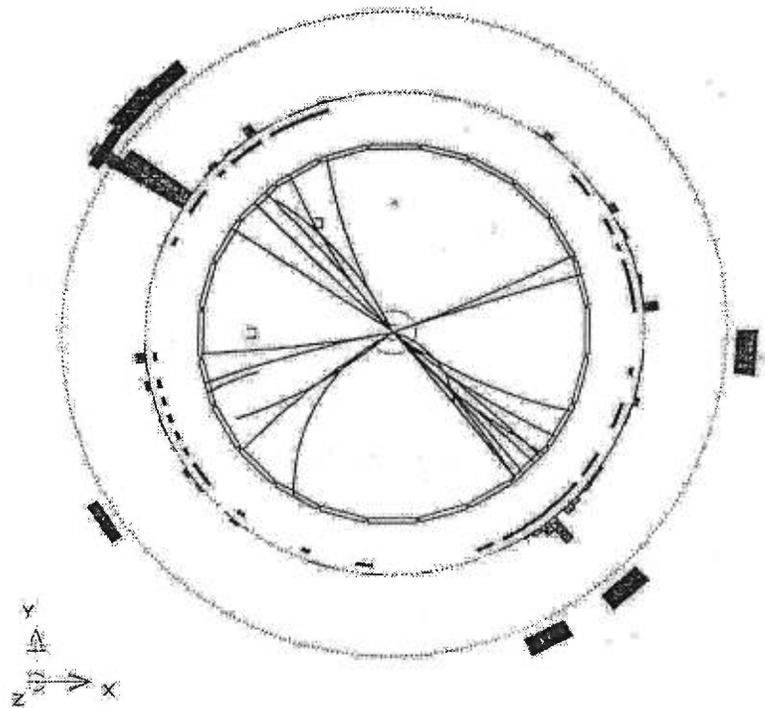


Figure 2-3. Exemple de visualisation GOPAL d'un événement à quatre partons produit par simulation Monte Carlo.

2.4 Les succès du LEP

Bien que les bosons de jauge W^+ , W^- et Z^0 aient été découverts en 1982 au collisionneur $p\bar{p}$ du CERN (expériences UA1 et UA2), les expériences du LEP ont été déterminantes dans la corroboration du Modèle Standard en ce qu'elles ont permis une étude systématique de la physique de ces bosons de jauge. En

effet la résolution des détecteurs a permis une mesure de la largeur (Γ) du pic de résonance du Z^0 (ce qui correspond à son taux de désintégration $R = \Gamma/\hbar$) dans la section efficace des réactions e^+e^- à une précision de l'ordre de 0.5%. Outre le fait que la précision de cette mesure a permis d'étudier les phénomènes découlant de la désintégration du Z^0 , la mesure précise des différents canaux de désintégration de cette particule a servi à prouver qu'il n'existe que trois familles de neutrinos. En effet, la conception des détecteurs du LEP comme OPAL ne leur permet pas de détecter les neutrinos. Par contre, on peut obtenir une mesure de Γ_Z^{TOT} , $\Gamma_Z^{q\bar{q}}$ et de $\Gamma_Z^{l^+l^-}$ ce qui nous permet d'obtenir une mesure du taux de désintégration du Z^0 en particules non détectables [20, 34]:

$$\Gamma_Z^{inv} = \Gamma_Z^{TOT} - \Gamma_Z^{q\bar{q}} - \Gamma_Z^{l^+l^-} \quad (2.3)$$

Ainsi, en prenant le rapport de cette valeur avec la valeur attendue du taux de désintégration du Z^0 en une paire neutrino et antineutrino ($\Gamma_Z^{\nu\bar{\nu}}$) on obtient une mesure du nombre de familles ayant un neutrino sans masse et un lepton plus lourd. Or la valeur de cette mesure telle que fournie par les expériences du LEP est de [14]:

$$\Gamma_Z^{inv}/\Gamma_Z^{\nu\bar{\nu}} = 3.04 \pm 0.04 \quad (2.4)$$

ce qui prouve donc qu'il y a seulement trois familles de leptons (qui sont connus par ailleurs). La figure 2-4 présente les résultats prouvant cette conclusion.

Les expériences du LEP ont également permis d'obtenir une mesure très précise de l'angle de Weinberg ($\sin \theta_w$). Comme la valeur de cette mesure n'est pas prédite par le Modèle Standard et que plusieurs prédictions dépendent de ce paramètre, il est donc important de mesurer $\sin \theta_w$ avec le plus de précision possible. Cela permet entre autre d'obtenir une valeur précise de la masse des bosons de jauge ainsi que des couplages électrofaibles. La valeur la plus précise obtenue par les expériences du LEP pour l'angle de Weinberg est de $\sin \theta_w = 0.2312 \pm 0.0001$. On peut obtenir cette valeur en mesurant [14] l'asymétrie avant-arrière dans les réactions $e^+e^- \rightarrow Z^0 \rightarrow \ell\bar{\ell}$ et en ajustant la section efficace

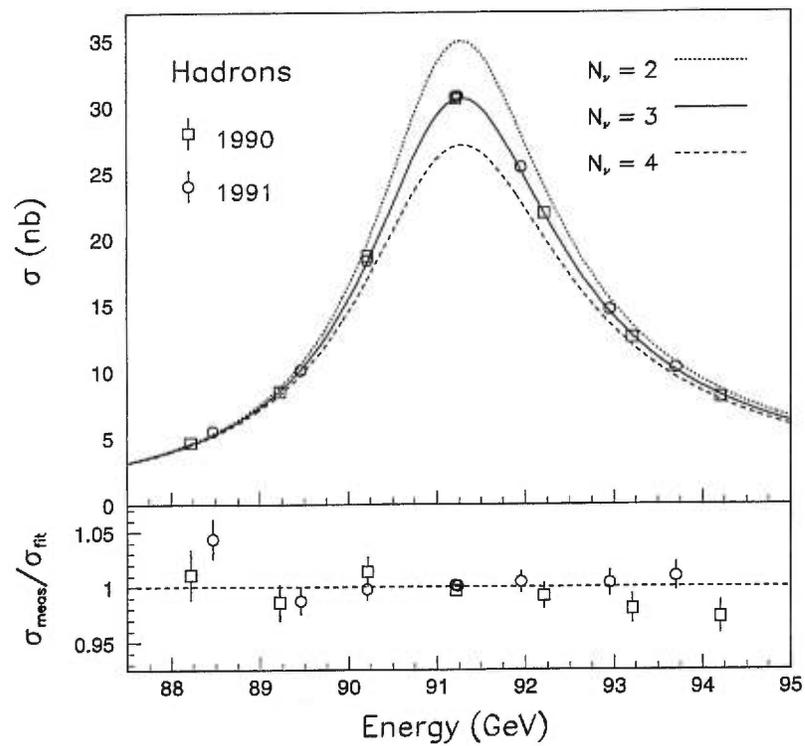


Figure 2-4. Comparaison des mesures de la section efficace du pic de résonance du Z^0 en fonction de l'énergie avec les prédictions théoriques pour deux, trois et quatre familles de neutrino.

mesurée du Z^0 avec la valeur prédite en fonction des différentes valeurs possibles de $\sin \theta_w$ compatibles avec l'asymétrie mesurée.

Parmi les autres mesures de précision effectuées au LEP citons la valeur de la constante de couplage forte (α_s). Cette valeur, qui est de $\alpha_s(M_Z) = 0.119 \pm 0.002$ à l'échelle de la masse du Z^0 [14], a été obtenue en mesurant le rapport du nombre d'événements à trois jets sur le nombre d'événements à deux jets, puisque le nombre de trois jets dépend de cette constante.

Maintenant que l'on a situé ce travail dans son cadre théorique ainsi qu'expérimental, regardons de plus près en quoi consiste le phénomène CDQ qui a motivé le développement de la méthode d'analyse que nous présenterons par la suite.

Chapitre 3

Processus CDQ à quatre partons

3.1 Configuration à quatre partons

La grande précision des expériences du LEP et le fait que cette usine à Z^0 a produit un très grand nombre d'événements $e^+e^- \rightarrow Z^0 \rightarrow q\bar{q}$ permet maintenant d'analyser les processus d'ordres supérieurs de la CDQ. L'étude des événements du deuxième ordre en α_s est particulièrement intéressante puisqu'elle permet l'étude des vertex fondamentaux $g \rightarrow gg$ et $g \rightarrow q\bar{q}$ liés, comme on l'a vu, aux constantes de Casimir. Elle constitue même une porte pour l'étude de la physique au-delà du Modèle Standard.

3.1.1 Processus à quatre partons

Outre les corrections sur l'amplitude des événements à deux ou trois partons, les processus de deuxième ordre en α_s admettent également des événements à quatre partons. Les diagrammes de Feynman contribuant à ces processus sont présentés à la figure 3-1 ([20, 30]).

Ces partons se fragmenteront par la suite en jets modélisés par les processus d'hadronisation présentés au premier chapitre. Comme la dynamique de ces partons est décrite par la CDQ perturbative, les quatre jets produits par la fragmentation des quatre partons parents pourront théoriquement être différenciés les uns des autres. Par contre, comme l'hadronisation est un processus non-

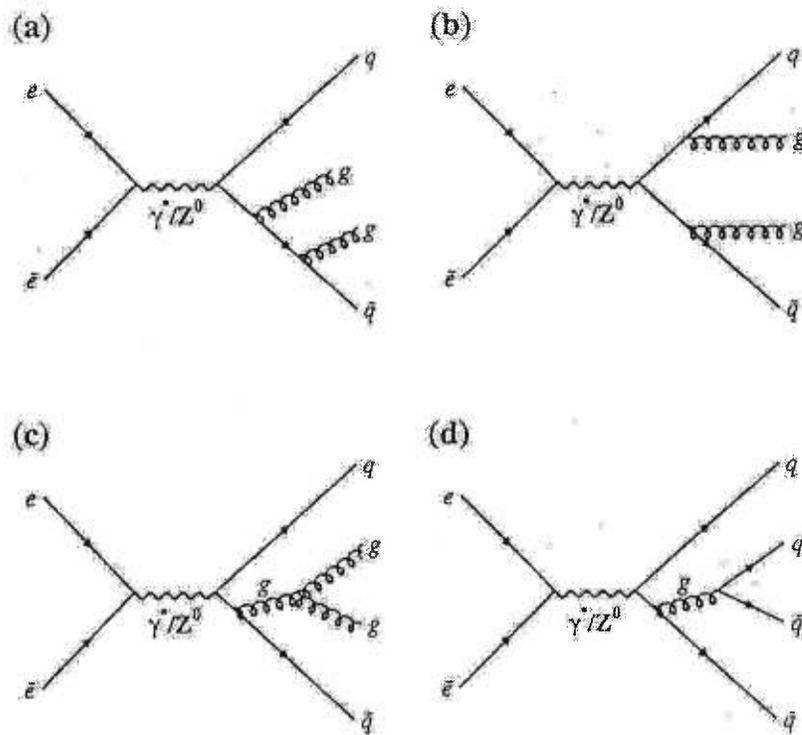


Figure 3-1. Diagrammes de Feynman des événements à quatre partons.

perturbatif, toutes les particules issues de la fragmentation d'un parton donné se confondront en un seul jet. Ainsi, la figure 3-1 présente les différents processus de deuxième ordre en α_s , conduisant à des événements à quatre jets.

3.1.2 Fractions apparentes des 4 quarks

On vient de voir que les événements de deuxième ordre à quatre jets admettent seulement deux états finaux possibles: $q\bar{q}q'\bar{q}'$ et $q\bar{q}gg$. Or, comme le vertex de $q \rightarrow qg$ est proportionnel à C_F , celui de $g \rightarrow gg$ est proportionnel à C_A et que celui de $g \rightarrow q\bar{q}$ est proportionnel à T_F , on a donc que les sections efficaces associées à la production de ces deux types d'événements dépendront différemment des facteurs de Casimir. En fait, on a que [30]:

$$\frac{1}{\sigma_0} d\sigma_{q\bar{q}gg} = \left(\frac{\alpha_s C_F}{\pi}\right)^2 \left[A + \left(1 - \frac{C_A}{2C_F}\right) B + \frac{C_A}{C_F} C \right] \quad (3.1)$$

et

$$\frac{1}{\sigma_0} d\sigma_{q\bar{q}q'\bar{q}'} = \left(\frac{\alpha_s C_F}{\pi}\right)^2 \left[n_f \frac{T_F}{C_F} D + \left(1 - \frac{C_A}{2C_F}\right) E \right] \quad (3.2)$$

où $\sigma_0 = \sigma(e^+e^- \rightarrow q\bar{q})$, les valeurs théoriques de C_A , C_F et T_F sont respectivement de $\frac{3}{4}$, 3 et $\frac{1}{2}$ et où A, B, C, D, E sont des fonctions cinétiques complexes dépendant de:

$$s_{ij} = (p_i + p_j)^2, \quad s_{ijk} = (p_i + p_j + p_k)^2 \quad \text{et} \quad y_{ij} = \frac{s_{ij}}{Q^2}.$$

Bien sûr, on a que:

$$\begin{aligned} d\sigma_{4p} &= d\sigma_{q\bar{q}gg} + d\sigma_{q\bar{q}q'\bar{q}'} \\ &= \sigma_0 \left(\frac{\alpha_s}{\pi}\right)^2 C_F \left[C_F A + \left(C_F - \frac{C_A}{2}\right) B + C_A C + n_f T_F D + \left(C_F - \frac{C_A}{2}\right) E \right] \end{aligned}$$

Or, comme $\left(C_F - \frac{C_A}{2}\right)$ est très petit (contribue que pour 0.2% de la section efficace totale) on peut donc négliger ce terme de sorte que [17, 23]:

$$d\sigma_{4p} \propto \sigma_0 C_F (C_F A + C_A C + n_f T_F D) \quad (3.3)$$

On peut réécrire la section efficace différentielle totale ainsi:

$$d\sigma_{4p} = K(F + n_f T_F D) \quad (3.4)$$

car A et C sont toutes les deux associées au même type d'événement: ceux du type $q\bar{q}gg$.

Comme la fraction d'événements appartenant à la distribution D est proportionnelle à la fraction d'événements à quatre quarks et que D est également proportionnelle à T_F , on pourra donc, par l'étude des événements à quatre quarks, analyser un des paramètres fondamentaux de la CDQ. En effet, si on définit f_q comme étant la fraction d'événements de type $q\bar{q}q'\bar{q}'$, on aura donc que [17]:

$$\frac{1}{\sigma_0} d\sigma_{4p} = (1 - f_q)F/F_0 + f_q D/D_0 \quad (3.5)$$

où $\sigma_0 = F_0 + n_f T_F D_0$. Si l'on divise (3.4) par σ_0 et qu'on réorganise les termes pour obtenir une forme similaire à (3.5), on obtiendra alors:

$$\frac{1}{\sigma_0} d\sigma_{4p} = \frac{1}{1 + n_f T_F \frac{D}{F_0}} F/F_0 + \frac{1}{\frac{F_0}{n_f T_F D_0} + 1} D/D_0 \quad (3.6)$$

On en déduit donc que [17, 23]:

$$f_q = \frac{1}{\frac{F_0}{n_f T_F D_0} + 1} \quad (3.7)$$

et

$$T_F = \frac{f_q}{n_f(1 - f_q)} \frac{F_0}{D_0} \quad (3.8)$$

nous offrant ainsi l'opportunité de tester la CDQ. Il s'agira en fait de calculer le rapport de T_F^{exp} sur T_F^{theo} et de vérifier que ce rapport est bien compatible avec

1. Or, comme

$$\frac{T_F^{exp}}{T_F^{theo}} = \frac{f_q^{exp}(1 - f_q^{theo})}{f_q^{theo}(1 - f_q^{exp})} \approx \frac{f_q^{exp}}{f_q^{theo}} \equiv R_{4q} \quad (3.9)$$

(car $f_q \ll 1$), il s'agira d'établir une méthode permettant d'évaluer f_q^{exp} et de le comparer à f_q^{theo} . Les calculs précédents ont cependant été obtenus par la CDQ perturbative. Il faudra donc, si on veut comparer de façon consistante f_q^{theo}

avec f_q^{exp} (qui provient de l'étude des événements à quatre jets mesurés dans le détecteur), obtenir la valeur de f_q^{theo} en simulant (par Monte Carlo) le processus de fragmentation, ainsi que les effets du détecteur, ce qui aura pour conséquence d'ajouter une incertitude à la valeur de R_{4q} obtenue.

3.1.3 Une porte au-delà du Modèle Standard

Bien que cette mesure offre l'intérêt d'obtenir une vérification supplémentaire du Modèle Standard, son intérêt primordial repose dans sa portée non-standard. En effet, les nombreuses expériences étudiant l'interaction forte sont toutes compatibles avec un modèle d'une théorie de jauge non-abélienne décrivant cette interaction. On ne s'attend donc pas à ce qu'une mesure de R_{4q} soit inconsistante avec toutes ces expériences. On considère en fait que $T_F^{exp} = T_F^{theo}$, c'est-à-dire qu'une mesure de T_F soit en accord avec les autres expériences testant la CDQ. Cependant, il est admissible que $f_q^{exp} \neq f_q^{theo}$ sans contredire les corroborations de la nature CDQ de l'interaction forte. En effet, si f_q^{exp} est contaminée par un canal de désintégration différent de $Z^0 \rightarrow q\bar{q}q'\bar{q}'$, on aura donc que $f_q^{exp} > f_q^{theo}$ et donc que $R_{4q} > 1$. Comme aucun canal standard ne permet cette ambiguïté, une mesure de $R_{4q} > 1$ sera donc un signal pour de la nouvelle physique. Le meilleur candidat prévoyant un excès dans la mesure de R_{4q} vient de la Supersymétrie.

Ce modèle théorique constitue une extension du Modèle Standard: en plus de comporter chacune des particules et interactions du Modèle Standard, le modèle Supersymétrique admet un partenaire pour chacune de ces particules ainsi que les termes d'interaction qui s'imposent entre ces particules nouvelles et les bosons de jauge standard. Ainsi, chaque boson est associé à un nouveau fermion et vice versa. Or, une des nouvelles particules considérées dans cette théorie est le partenaire supersymétrique du gluon: le *gluino* \tilde{g} (de spin $\frac{1}{2}$). Cette particule peut être produite par paire via la désintégration d'un gluon. Comme le gluino

constitue une particule d'interaction forte de spin $\frac{1}{2}$ issue de la désintégration d'un gluon, la signature du vertex $g \rightarrow \tilde{g}\tilde{g}$ est donc sensiblement la même que celle du vertex $g \rightarrow q\bar{q}$, empêchant à l'observation de les différencier. Une telle production de gluinos aurait donc pour effet d'augmenter la fraction apparente d'événements à quatre quarks, parmi les événements à quatre jets. Ainsi, une mesure de R_{4q} non-compatible avec 1 pourrait être une indication de l'existence du gluino et donc de physique au-delà du Modèle Standard. Il est alors impératif de pouvoir séparer les événements à quatre quarks des événements à deux quarks et deux gluons parmi l'ensemble des événements à quatre jets car cela nous permettra de mesurer f_q^{exp} et donc R_{4q} . Un tel problème nécessite cependant qu'on se penche tout d'abord un peu sur la phénoménologie des jets.

3.2 Phénoménologie des jets

On sait que la théorie de la CDQ perturbative permet de prévoir la topologie des jets ainsi que certaines de leurs caractéristiques cinématiques, mais qu'elle ne peut prévoir leur constitution ainsi que la distribution des particules dans ceux-ci. En effet, cela dépend de la fragmentation dont on a au plus un modèle plus ou moins réaliste du mécanisme. Les particules dans un jet peuvent ainsi avoir différentes valeurs d'impulsion et d'énergie qu'il ne nous est pas possible de prévoir. Il n'est donc pas évident de reconnaître les jets à partir des données d'observation ou encore de l'état final d'une simulation Monte Carlo (qui simule en fait l'état final des particules dans le détecteur). En fait, plus le nombre n de jets (issus de n partons) est grand, plus les particules seront dispersées dans l'état final, compliquant ainsi la reconnaissance des jets. De même, deux partons très rapprochés donneront deux jets collinéaires, rendant très difficile leur identification. Il faut donc se doter d'un algorithme qui, à partir des observables des particules mesurables, permettra de reconnaître les jets.

La méthode usuelle de reconstruction des jets est de recombinaison, grâce à un

paramètre de résolution parfois appelé *métrique*, les particules une par une en jets bien définis. Il s'agit en fait de calculer, pour chaque paire de particules (i, j) , la valeur du paramètre de résolution généralement appelé y_{ij} [16]. La paire pour laquelle la valeur de y_{ij} est minimale sera alors recombinaison en un seul amas et le processus sera ensuite réitéré pour les $n - 1$ particules restantes. Le processus s'arrêtera lorsque la valeur minimale de y_{ij} sera supérieure à un paramètre libre de coupure qu'on appellera y_{cut} . Le nombre de jets sera alors le nombre de particules ou d'amas qui n'auront pas été recombinaison. Ce nombre ne sera jamais inférieur à 2, car le processus à l'ordre le plus bas comporte deux partons.

Chaque algorithme d'identification des jets doit comporter deux prescriptions: une précisant la valeur du paramètre de résolution y_{ij} et l'autre spécifiant comment recombinaison les caractéristiques des particules mesurables en "particule-amas". Les principaux algorithmes d'identification des jets avec leurs critères de résolution et de combinaison sont résumés dans le tableau (3-I) [16, 31, 23].

Algorithmes	Résolution y_{ij}	Combinaison
JADE	$y_{ij} = \frac{2E_i E_j (1 - \cos \theta_{ij})}{E_{vis}^2}$	$P^\mu = P_i^\mu + P_j^\mu$
EO	$y_{ij} = \frac{2E_i E_j (1 - \cos \theta_{ij})}{E_{vis}^2}$	$E_k = E_i + E_j$ $\vec{p}_k = \frac{E_k}{ \vec{p}_i + \vec{p}_j } (\vec{p}_i + \vec{p}_j)$
PO	$y_{ij} = \frac{2E_i E_j (1 - \cos \theta_{ij})}{E_{vis}^2}$	$\vec{p}_k = \vec{p}_i + \vec{p}_j$ $E_k = \vec{p}_k $
Durham	$y_{ij} = \frac{2\min(E_i^2, E_j^2)(1 - \cos \theta_{ij})}{E_{vis}^2}$	$P^\mu = P_i^\mu + P_j^\mu$

Tableau 3-I. Définition du paramètre de résolution y_{ij} et du schème de recombinaison de différents algorithmes d'identification des jets.

L'intérêt de l'algorithme JADE est qu'il permet d'obtenir une équivalence entre le nombre de partons générés (par élément de matrice) et le nombre de jets reconstruits après l'hadronisation, de sorte que les effets de la CDQ peuvent être étudiés avec une dépendance minimale sur le modèle de fragmentation par corde [31]. Les algorithmes E0 et P0, qui ont la même définition que JADE pour le paramètre de résolution, tirent leur intérêt de leur schème de recombinaison: contrairement à JADE, ces algorithmes ne donnent pas de masse aux jets, lorsqu'ils recombinent les particules en jets, ce qui permet une comparaison directe avec les prédictions de la CDQ, qui sont faites pour des partons sans masse [31]. Finalement, l'algorithme DURHAM permet de réduire les problèmes dus à l'émission de gluons à faible énergie [31]. En effet, pour des raisons de renormalisation et de qualité de la définition des détecteurs, les événements à deux quarks avec émission d'un gluon de très faible énergie doivent être reconstruits comme des événements à deux jets (puisque les diagrammes de Feynman de ces processus contribuent à la section efficace des événements à deux partons) et non comme des événements à trois jets. DURHAM facilite cela. Ainsi, la reconstruction des jets à l'aide de DURHAM est moins affectée par les corrections radiatives à un ordre fixe en α_s . De plus, cet algorithme a une meilleure résolution en angle. Cela diminue l'effet des particules mal assignées à un jet donné. Pour ces raisons, DURHAM est l'algorithme le plus utilisé actuellement et c'est celui qui sera utilisé dans la présente analyse.

On a vu que la recombinaison des particules en jets cessait lorsque la valeur minimale du paramètre de résolution pour les paires restantes devenait plus grande que la valeur d'un paramètre de coupure libre: y_{cut} . Ainsi, le nombre de jets reconstruits pour un événement donné n'est pas absolu. Si la valeur pour lequel l'algorithme recombine $n + 1$ -jets en n -jets est $y^{n,n+1}$ et que celle qui recombine n -jets en $n - 1$ est $y^{n-1,n}$, alors pour n'importe quelle valeur de y_{cut} prise dans l'intervalle de valeurs $[y^{n,n+1}, y^{n-1,n}]$, l'événement sera un n -jets. Mais comme en général deux événements n'ont pas la même topologie et la même

composition en particules, les valeurs de $y^{n,n+1}$ et de $y^{n-1,n}$ différeront donc d'un événement à l'autre. Ainsi, une valeur de y_{cut} choisie ne se situera pas toujours dans un intervalle $[y^{n,n+1}, y^{n-1,n}]$ fixe, ce qui fait que certains événements seront vus comme des n -jets, d'autres comme des $n - 1$ -jets, etc [16]. La figure 3-2 montre le nombre relatif de deux, trois et quatre jets reconstruits pour différentes valeurs de y_{cut} lorsqu'on génère $q\bar{q} + q\bar{q}g + q\bar{q}gg + q\bar{q}q'\bar{q}'$.

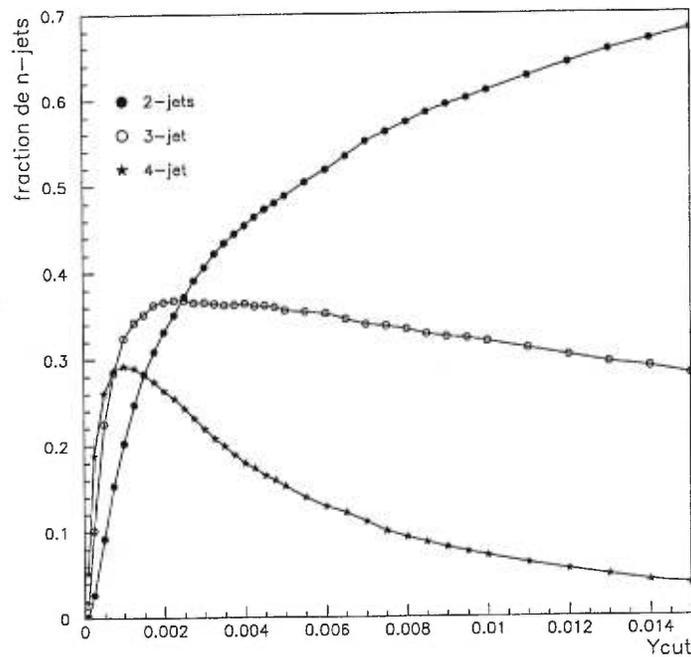


Figure 3-2. Nombre relatif de deux, trois et quatre jets reconstruits pour différentes valeurs de y_{cut} .

Il est donc important de faire un choix de coupure (en choisissant une valeur de y_{cut} ou en imposant un seuil aux $y^{n,n+1}$ considérés) adéquat à l'étude des différents phénomènes de la CDQ. Une fois que la sélection des jets est terminée, c'est-à-dire une fois que l'on a modélisé les événements détectés (ou simulés) en une configuration à n -jets représentant n partons, on est prêt pour l'analyse CDQ.

3.3 Méthode de discrimination des $q\bar{q}q'\bar{q}'$

Comme on est intéressé à la détermination de R_{4q} , l'analyse consistera à séparer les événements à quatre quarks de ceux à deux quarks et deux gluons parmi un ensemble d'événements reconstruits en quatre jets. Il s'agira en fait de choisir des observables qui réagissent différemment aux configurations de ces deux types d'événements, permettant ainsi de les séparer l'un de l'autre. Bien sûr, ces observables ne permettront pas de discerner la nature partonique d'un événement spécifique, mais ils permettront de séparer un échantillon d'événements à quatre jets en deux classes distinctes avec un certain niveau de confiance déterminé.

Les corrélations angulaires entre les jets constituent de bons candidats d'observables susceptibles de dresser un échantillon fiable d'événements à quatre quarks. En effet, puisque les gluons sont polarisés, les différences de spins entre les quarks et les gluons produiront des topologies différentes entre les événements $q\bar{q}q'\bar{q}'$ et $q\bar{q}gg$ [11, 17]. Cela conduit donc à des corrélations angulaires différentes entre les jets qu'il faut mettre en évidence grâce à la construction de variables angulaires à partir des observables de ces jets. La variable la plus utilisée pour étudier ces corrélations est l'angle de Bengtsson-Zerwas.

3.3.1 Angle Bengtsson-Zerwas

Proposée en 1988 par M. Bengtsson et P.M. Zerwas [11] cette variable constitue un bon observable révélant des corrélations angulaires entre les jets car elle s'appuie sur le fait que les gluons durs émis des quarks primaires, dans les événements $e^+e^- \rightarrow q\bar{q}g$, sont fortement linéairement polarisés dans le plan de l'état final $q\bar{q}g$. Cette polarisation peut être évaluée par la CDQ en fonction de la section efficace différentielle dans le plan de l'état final ainsi que de celle

dans le plan perpendiculaire par [11, 17]:

$$P(x_g) = \frac{(d\sigma_{\parallel} - d\sigma_{\perp})}{(d\sigma_{\parallel} + d\sigma_{\perp})} = 2 \frac{(1 - x_g)}{(x_q^2 + x_{\bar{q}}^2)} \quad (3.10)$$

où $x_i = 2E_i/\sqrt{s}$ est l'énergie d'échelle des quarks ou des gluons dans le référentiel du laboratoire.

Or la fragmentation d'un gluon polarisé linéairement en ces partons fils dépend de l'angle azimutal entre le plan de l'état final et le vecteur de polarisation. Cela introduira donc une asymétrie dans les fonctions de fragmentation des gluons, dont dépendent les équations d'Altarelli-Parisi, de sorte que ces fonctions deviendront [11, 17]:

$$P'_{g \rightarrow gg} = \frac{6}{2\pi} \left(\frac{(1 - z + z^2)^2}{z(1 - z)} + z(1 - z) \cos 2\chi \right) \quad (3.11)$$

$$P'_{g \rightarrow q\bar{q}} = \frac{n_F}{2\pi} \left(\frac{1}{2}(z^2 + (1 - z)^2) - z(1 - z) \cos 2\chi \right) \quad (3.12)$$

où z est la fraction de 4-impulsion portée par le gluon et où Bengtsson et Zerwas ont défini cet angle χ en fonction des variables d'impulsion des jets de la manière suivante [11, 17, 6]:

$$\cos \chi_{BZ} = \frac{|\vec{p}_1 \times \vec{p}_2 \cdot \vec{p}_3 \times \vec{p}_4|}{|\vec{p}_1 \times \vec{p}_2| |\vec{p}_3 \times \vec{p}_4|} \quad (3.13)$$

le premier jet étant le plus énergétique, tandis que le quatrième jet est le moins énergétique.

On remarque que le terme d'asymétrie dans l'équation de fragmentation $g \rightarrow q\bar{q}$ est opposé à celui de $g \rightarrow gg$. Ainsi, pour une asymétrie maximale (à $z = \frac{1}{2}$), les jets de quarks secondaires (c'est-à-dire issus de la désintégration d'un gluon) auront plus tendance à se trouver dans le plan perpendiculaire au vecteur de polarisation que les jets de gluons. Les jets de gluons issus de deux rayonnements successifs des quarks parents (diagrammes (a) et (b) de la figure 3-1) seront, pour leur part, situés dans un plan distribué isotropiquement par rapport au plan de polarisation. Ainsi une distribution d'angle χ_{BZ} présentant un plus grand nombre d'événements pour des angles tendant vers 90° sera une indication

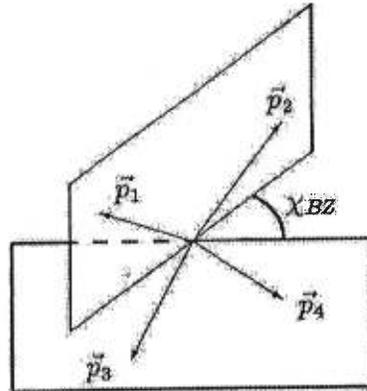


Figure 3-3. Schéma de l'angle χ_{BZ} défini par les plans formés des jets 1,2 et 3,4.

d'un fort taux de $q\bar{q}q'\bar{q}'$, alors qu'à l'inverse une distribution plus uniforme de χ_{BZ} indiquera une forte présence de $q\bar{q}gg$. Il suffira donc de comparer et d'ajuster la distribution de l'angle χ_{BZ} obtenue des données avec une distribution témoin construite à partir de la théorie (grâce à un simulateur Monte Carlo), pour avoir une évaluation de la proportion d'événements à quatre quarks dans un échantillon d'événements à quatre jets.

3.3.2 Variables topologiques

Les récentes études menées par la collaboration OPAL sur la mesure de R_{4q} à l'aide des distributions d'angles χ_{BZ} tendent à faire croire à un surplus apparent de R_{4q} par rapport à ce qui est attendu par la théorie de l'interaction forte au deuxième ordre en α_s [17]. Les raisons possibles d'un tel excès peuvent être:

1. L'existence d'une nouvelle particule ayant des propriétés similaires à celle du quark;
2. Un effet non-identifié du détecteur ou de la méthode de mesure produisant des déformations dans les distributions angulaires expérimentales utilisées

- pour déterminer le nombre apparent d'événements à quatre quarks;
3. L'insuffisance des calculs de la CDQ au deuxième ordre en α , pour rendre compte de la grande précision des données expérimentales;
 4. Les différents bruits de fond contaminant l'échantillon de quatre quarks produit.

Il est donc important d'étudier avec minutie les raisons 2), 3) et 4) avant de conclure à l'existence d'une nouvelle particule. Il est entre autre possible de construire une méthode de classification des événements à quatre jets qui soit indépendante des méthodes par corrélations angulaires. Ceci constitue en fait l'objectif visé par le présent travail. Il s'agira donc pour nous de faire un judicieux choix de variables topologiques possédant toutes un certain pouvoir de discrimination. Il suffira ensuite "d'additionner" ces séparateurs à l'aide d'un réseau de neurones (tenant compte des corrélations possibles entre ces variables) pour en déduire le pouvoir de séparation de l'ensemble des variables choisies. Nous devrons finalement établir une évaluation de ce pouvoir de séparation, évaluation qui sera nécessaire à la détermination du nombre d'événements à quatre quarks contenus dans un échantillon d'événements à quatre jets. Introduisons donc maintenant les variables topologiques choisies et évaluons leur pouvoir de discrimination.

Chapitre 4

Variables topologiques

4.1 Critère de sélection des variables

La première étape de notre méthode de discrimination des événements à quatre quarks consiste à choisir des variables permettant de séparer les événements $q\bar{q}q'\bar{q}'$ des événements $q\bar{q}gg$. Le choix de ces variables est guidé par les quatre critères suivant [2]:

1. ces variables doivent être des observables réels dont la section efficace différentielle ($\frac{d\sigma}{dA}$ où A est une des variables choisies) doit être calculable par la CDQ perturbative;
2. elles doivent être fonction des quadri-vecteurs des particules détectées;
3. elles doivent être le plus insensible possible aux effets non-perturbatifs qui ne dépendent pas de la nature des partons hadronisés
4. **ces variables doivent donner des distributions différentes selon la nature des événements.**

Le premier critère sert à assurer que les variables choisies sont mesurables et qu'on peut prévoir leur distribution par la théorie. En fait, comme les générateurs Monte Carlo servent à simuler des événements physiques à partir des calculs théoriques, ce critère nous assurera que les distributions des variables choisies pourront être obtenues par des simulateurs comme PYTHIA. On veut certes

pouvoir obtenir ces distributions par la théorie pour construire notre méthode de discrimination, mais on veut éventuellement les utiliser pour construire un échantillon d'événements à quatre quarks dans les données OPAL, ce qui fait que l'on doit pouvoir également construire ces variables à partir de l'information fournie par une mesure expérimentale. C'est ce qui justifie notre deuxième critère. Le troisième critère est, quant à lui, motivé par le fait que l'on veut séparer les quarks des gluons, c'est-à-dire que l'on veut étudier la nature partonique d'un événement (ce qui permet par ailleurs la comparaison avec les prédictions de la CDQ qui, on l'a déjà dit, ne sont possible qu'au niveau partonique). Il est donc nécessaire que les variables dépendent le moins possible de ce qu'on ne peut pas prévoir. Dans le cas contraire, les modèles de fragmentation auraient un effet déterminant sur ces distributions, diminuant le pouvoir réel de séparation de nos variables. Finalement le quatrième critère est évidemment nécessaire en ce qu'il représente la justification d'un tel choix de variables: pouvoir discerner les événements à quatre quarks, parmi ceux à quatre jets. Notons qu'on pourrait parler d'un cinquième critère qui est implicitement inclus dans la motivation de ce travail: les variables choisies ne doivent pas dépendre des corrélations angulaires entre les jets.

Ces variables peuvent être de deux types:

1. des variables globales qui caractérisent la structure générale d'un événement,
2. des variables de jets qui caractérisent la topologie générale des jets.

Les variables globales présentent un avantage sur les variables de jets parce qu'elles ne dépendent pas directement d'un paramètre arbitraire comme l'est le y_{cut} . Cependant, les variables de jets sont également pertinentes parce qu'elles reflètent la nature partonique d'un événement. Il existe un grand nombre de telles variables qui satisfont aux trois premiers critères. Notre travail a donc consisté à

construire leur distribution dans le cas précis des événements à quatre jets (fidèles aux prédictions par calculs des éléments de matrices au deuxième ordre en CDQ perturbative) et à choisir celles qui satisfont au quatrième critère.

Imposer que les distributions des variables choisies soient différentes est cependant un critère qualitatif. Il faut donc se doter d'un critère quantitatif pour évaluer le pouvoir de séparation de deux distributions différentes. Les critères quantitatifs permettant de comparer deux distributions de valeurs sont cependant de nature statistique et ne dépendent pas directement du contenu physique de ces distributions. Le choix du critère sera donc arbitraire et relatif, c'est-à-dire qu'il servira seulement à établir une relation de comparaison entre les distributions de variables. Nous avons choisi comme critère de comparaison le calcul des **figures de mérite**. Si on suppose qu'on a une variable A dont la distribution normalisée pour les événements $q\bar{q}gg$ est notée g et que celle des événements $q\bar{q}q'\bar{q}'$ est notée q , alors la mesure de la différence entre ces deux distributions telle que donnée par la figure de mérite est définie par [1, 5]:

$$F = \gamma\bar{\gamma} \sum_i \left(\frac{(g_i - q_i)^2}{\gamma g_i + \bar{\gamma} q_i} \right) \quad (4.1)$$

où γ est la fraction du nombre d'événements $q\bar{q}gg$ contenus dans les deux distributions comparées, $\bar{\gamma}$ est son complémentaire ($\bar{\gamma} = 1 - \gamma$), g_i la valeur de g dans l'intervalle i de valeurs de la variable A et q_i l'équivalent de g_i pour la distribution $q\bar{q}q'\bar{q}'$. Comme on générera autant d'événements $q\bar{q}gg$ que d'événements $q\bar{q}q'\bar{q}'$ pour construire les distributions des variables choisies, γ et $\bar{\gamma}$ seront tous les deux fixés à $\frac{1}{2}$. De plus comme les g_i et les q_i sont normalisées ($g_i = \frac{\#q\bar{q}gg \in \text{bin}(i)}{\#q\bar{q}gg \text{ total}}$), on a donc que F variera entre 0 et 1, 0 correspondant à deux distributions identiques et 1 à deux distributions qui ne se recoupent pas du tout. Finalement, comme F dépend de la définition des intervalles de valeurs (bin) des variables A , il faudra faire le même choix d'intervalles pour chacune des variables sélectionnées, ce qui sera possible car chacune de ces variables est normalisée. Cela renforce le caractère relatif et comparatif de F .

Ainsi le critère qui a dicté notre choix de variables en fonction de la figure de mérite est le suivant:

$$F \geq 0.015$$

Ce choix n'est pas complètement arbitraire (bien qu'il ne soit pas critique) car il est motivé par le fait que les distributions, dont les figures de mérite ont une valeur plus petite que 0.015, n'ajoutent pratiquement pas au pouvoir de séparation du réseau de neurones utilisé pour sommer les séparateurs, le poids étant surtout porté par les variables dont la figure de mérite est de valeur plus élevée (typiquement deux à quatre fois plus grande). Cela confirme ainsi la valeur relative de ce critère de résolution qu'est la figure de mérite. Présentons donc les variables retenues.

4.2 Variables globales

4.2.1 Sphéricité

La sphéricité est définie, de façon analogue au tenseur d'inertie en mécanique classique, c'est-à-dire par le tenseur de rang 2 suivant [12, 2]:

$$S_{ij} = \frac{\sum_{r=1}^N p_i^r p_j^r}{\sum_{r=1}^N |p^r|^2} \quad i, j = x, y, z \quad (4.2)$$

où N est le nombre de particules et p^r est l'impulsion de la r -ième particule finale.

On peut écrire ce tenseur sous forme d'une matrice 3×3 . Cette matrice étant réelle et symétrique, elle peut être diagonalisée avec des valeurs propres réelles sur la diagonale. Si deux de ces valeurs propres sont différentes, alors les vecteurs propres qui leurs sont associés seront orthogonaux. On pourra donc associer à chaque événement un ensemble de valeurs Q_1, Q_2 et Q_3 ($Q_1 \leq Q_2 \leq Q_3$) correspondant aux valeurs propres de S et attacher à cet événement un nouveau système de coordonnées dans l'espace de la sphéricité. Ce nouveau système de coordonnées minimise en fait le carré des moments perpendiculaires à chacun des trois axes. Si on examine les cas limites, un événement dont $(Q_1, Q_2, Q_3) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ dans l'espace des vecteurs propres de S correspondra à une distribution isotropique des particules dans un détecteur, alors qu'un événement pour lequel $Q_1 = Q_2 = 0$ et $Q_3 = 1$ correspondra à un événement complètement linéaire. La sphéricité permet donc d'associer à un événement un système de coordonnées décrivant les trois axes d'un ellipsoïde dont l'axe principal est de longueur Q_1 , l'axe majeur de longueur Q_2 et l'axe mineur de longueur Q_3 . On appelle traditionnellement ces trois axes la longueur, la largeur et l'épaisseur de l'espace d'impulsion de l'événement analysé. La figure 4-1 présente les distributions de la variable s , construite à partir des valeurs propres de la matrice de sphéricité selon la relation $s = \frac{3}{2}(Q_1 + Q_2)$ ($s = 0$ correspondant à un événement linéaire et $s = 1$ correspondant à un événement isotropique), pour les événements à quatre

quarks et ceux à deux quarks et deux gluons.

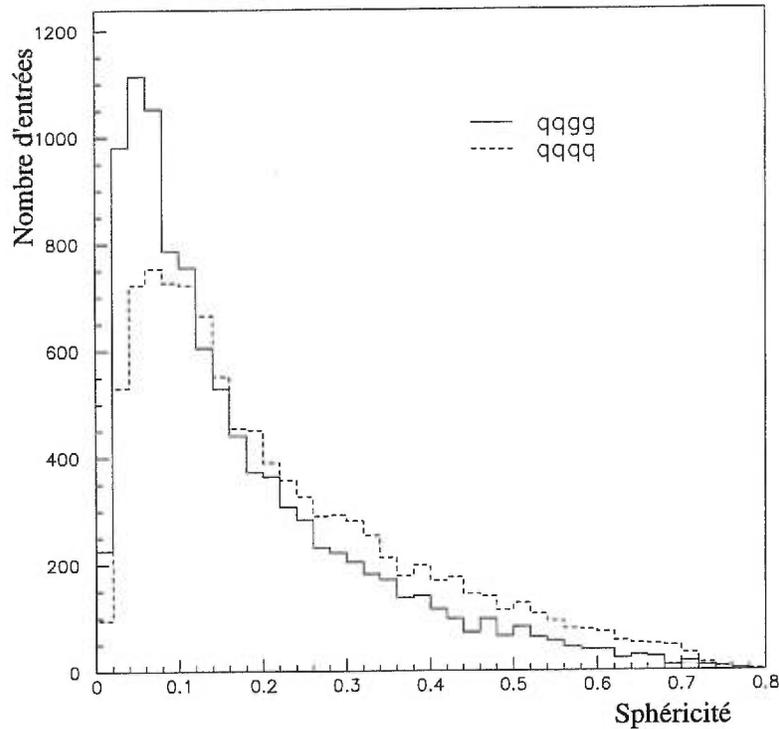


Figure 4-1. Distribution de la fonction s de la sphéricité pour les événements à deux quarks et deux gluons (traits pleins) ainsi que pour les événements à quatre quarks (traits pointillés).

On remarque sur cette figure qu'il y a bien une distinction dans les distributions de cette variable pour les deux types d'événements qu'on veut séparer, ce qui en fait une bonne variable de discrimination. Cette variable ne se conforme cependant pas complètement aux conditions qu'on a imposées au choix de variables à la section (4.1). En effet, elle ne satisfait pas au premier critère, qui demandait que la variable soit complètement calculable par la CDQ perturbative, car la sphéricité est sensible aux divergences infrarouges, c'est-à-dire à l'émission de gluons mous [19, 2]. Cependant comme PYTHIA (qui est le générateur utilisé pour reproduire les événements utilisés pour construire ces distributions) ne calcule qu'au deuxième ordre les éléments de la matrice de transition de l'état

initial à l'état final, sans inclure les corrections radiatives aux diagrammes d'arbre des événements à quatre partons ni les contributions dues à l'émission de gluons mous, le calcul de $\frac{d\sigma}{ds}$ ne fera pas intervenir de telles divergences et nos distributions pourront quand même être produites [31]. De plus, comme l'expérience est aveugle à ces gluons mous à cause du seuil de détection du détecteur (OPAL), on pourra éventuellement utiliser cette distribution de la sphéricité pour étudier les vraies données. Par contre, si on voulait faire des prévisions théoriques plus précises en incorporant les corrections radiatives des ordres supérieurs, il faudrait alors omettre la sphéricité dans notre choix de variables.

Or ce qui fait que la sphéricité est divergente infrarouge c'est qu'elle dépend du carré de l'impulsion total. En effet, lors d'un processus de désintégration en particules colinéaires (ce qui est le cas de la radiation des gluons mous) l'impulsion est conservée, mais le carré de l'impulsion ne l'est pas (la somme des carrés des impulsions des particules finales est plus grande que le carré de l'impulsion de la particule initiale; c'est ce qu'on appelle, mathématiquement, "l'inégalité du triangle"). Ainsi, comme la sphéricité dépend du carré de l'impulsion des particules finales, elle dépendra donc du détail des désintégrations colinéaires des particules instables, c'est-à-dire du détail de l'émission de gluons mous, ce qui en fera une mauvaise variable lorsqu'on considère les corrections radiatives [19, 2, 31]. On peut cependant corriger cette lacune en définissant une autre variable qui n'a pas ce problème et qui sera donc préservée des divergences infrarouges. Cette variable est la sphérocité.

4.2.2 Sphérocité

La sphérocité est définie, de façon analogue à la sphéricité, par le tenseur [15, 2]:

$$S_{ij} = \frac{\sum_{r=1}^N p_i^r p_j^r / |p^r|}{\sum_{r=1}^N |p^r|} \quad i, j = x, y, z \quad (4.3)$$

On peut, tout comme pour la sphéricité, diagonaliser la matrice associée à ce tenseur et y trouver trois valeurs propres Q_1 , Q_2 et Q_3 , ainsi qu'un système d'axes orthogonaux associés à l'événement. Encore une fois, le cas limite $Q_1 = Q_2 = Q_3 = \frac{1}{3}$ décrira une distribution complètement isotropique de l'événement dans l'espace des impulsions, alors que $Q_1 = Q_2 = 0$ et $Q_3 = 1$ correspondra à un événement linéaire. Or en calculant l'équation caractéristique du déterminant de cette matrice (permettant de calculer les valeurs propres), on voit apparaître deux combinaisons non-linéaires des valeurs propres de la sphéricité qui, bien que non-invariante sous rotation dans l'espace des impulsions, permettent toutefois une bonne discrimination des événements hadroniques (c'est-à-dire des événements à deux, trois et quatre jets). En fait, ces deux quantités sont:

$$C = 3(Q_1Q_2 + Q_2Q_3 + Q_1Q_3) \quad (4.4)$$

$$D = 27(Q_1Q_2Q_3) \quad (4.5)$$

où C est non-négligeable pour les événements à trois jets, alors que D est non-négligeable seulement pour les événements à quatre jets [2, 15, 31]. Nous avons donc choisi cette fonction D comme variable de discrimination. Notons que plus l'événement sera linéaire, plus D tendra vers 0 et au contraire, plus l'événement sera isotropique, plus D tendra vers 1. On s'attend également à ce que cette distribution soit de forme similaire à celle de la fonction s de la sphéricité (car ces deux variables sont sensibles aux mêmes caractéristiques de la topologie générale d'un événement), mais avec une distribution de valeur plus compressée vers $D = 0$, dû au fait que D est le produit des valeurs propres et non la somme. La figure 4-2 comparant la distribution de cette variable pour les événements $q\bar{q}gg$ et $q\bar{q}q'\bar{q}'$ confirme cela.

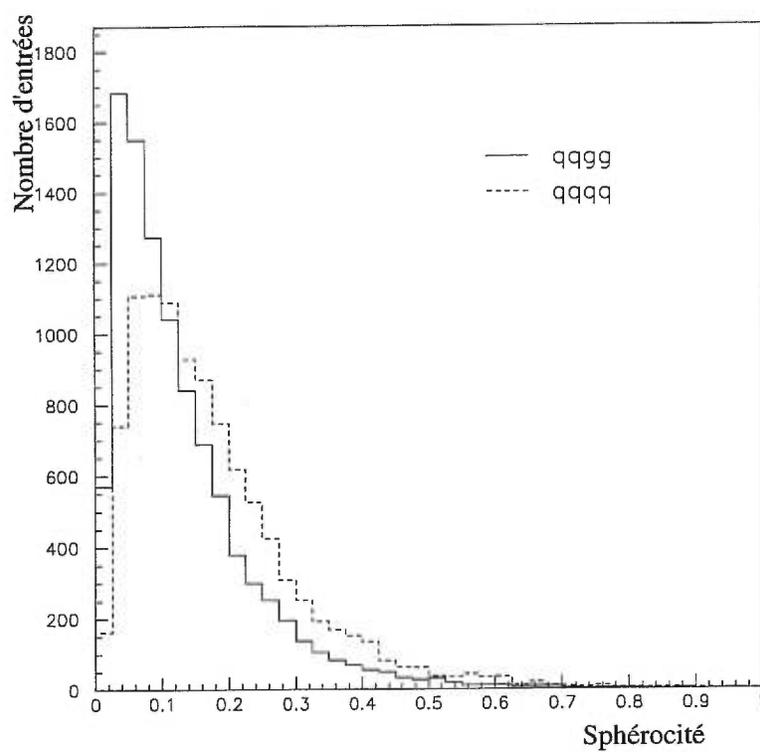


Figure 4-2. Distribution de la fonction D de la sphéricité pour les événements à deux quarks et deux gluons (traits pleins) ainsi que pour les événements à quatre quarks (traits pointillés).

4.2.3 Thrust

Le thrust est définie par [15, 2]:

$$T = \text{Max} \left(\frac{\sum_i^N |\vec{p}_i \cdot \hat{n}|}{\sum_i |\vec{p}_i|} \right) \quad (4.6)$$

Le vecteur unitaire \hat{n} , pour lequel la somme des produits scalaires avec l'impulsion de chacune des particules sera maximale, s'appellera l'axe principal du thrust et sera noté \hat{n}_T . Cet axe correspond à l'axe des partons dans les événements à deux jets. On peut de même définir le thrust majeur T_M en reprenant la même définition que celle du thrust principal, mais en imposant l'optimisation dans l'espace perpendiculaire à l'axe du thrust principal. On appellera finalement "thrust mineur" la variable T_m calculée le long de l'axe perpendiculaire au plan formé par le thrust principal et le thrust majeur. Ainsi, \hat{n}_T , \hat{n}_{T_M} et \hat{n}_{T_m} forment un système d'axes orthogonaux dont les valeurs de T , T_M et T_m qu'ils définissent servent encore une fois à classifier les différentes topologies d'un événement. Notons cependant que les variables T_M et T_m sont surtout sensibles aux événements à trois jets et ne seront donc pas de bons discriminants pour nous.

Bien que le thrust soit stable pour les émissions de gluons mous et pour la désintégration des particules instables, ce qui fait qu'il satisfait aux conditions imposées à notre choix de variables, le thrust a tout de même un petit défaut: le processus de maximisation définissant le thrust est itératif. La valeur de T ne peut donc être obtenue analytiquement ce qui fait qu'il est possible que le calcul de T s'arrête après un certain temps dans un maximum local, sans avoir trouvé le maximum absolu [2, 31]. Pour pallier à cela le plus possible, il s'agit de se donner plusieurs axes \hat{n} différents comme point de départ et arrêter les calculs lorsque tous les processus d'itération convergeront vers le même axe principal. La distribution du thrust pour les deux types d'événements à discriminer est donnée à la figure 4-3.

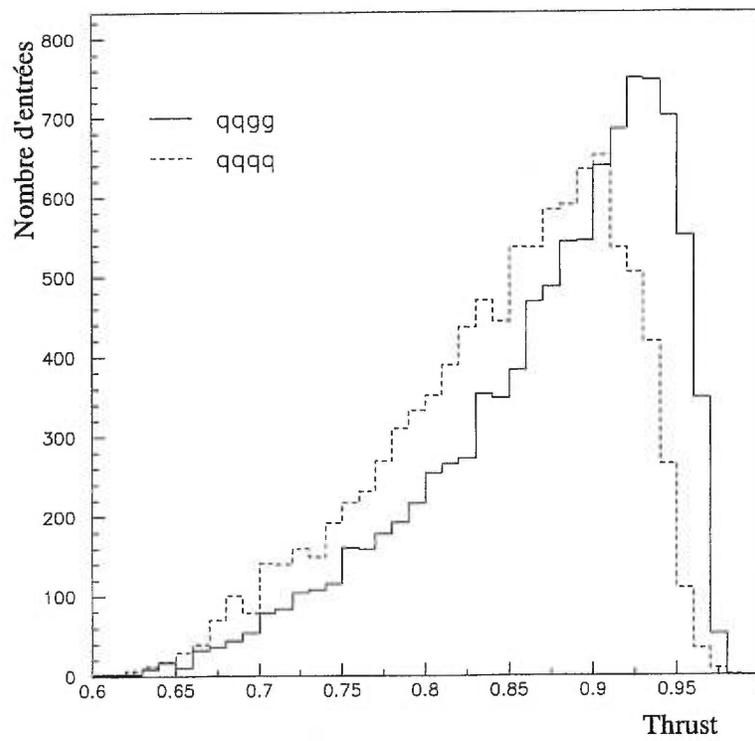


Figure 4-3. Distribution du thrust principal pour les événements à deux quarks et deux gluons (traits pleins) ainsi que pour les événements à quatre quarks (traits pointillés).

4.2.4 Acoplanarité

Cette variable est définie par [2]:

$$\mathcal{A} = 4 \min \left(\frac{\sum |p_{\perp i}|}{\sum |p^r|} \right)^2 \quad (4.7)$$

Au lieu de minimiser sur un axe comme le fait le thrust, l'acoplanarité minimise sur un plan. Ainsi, plus \mathcal{A} tendra vers 0, plus l'événement sera planaire. Comme cette variable est non-négligeable pour les événements qui ne sont pas planaires, elle prendra donc des valeurs différentes de 0 que pour les événements à quatre jets. Les distributions de cette variable sont reproduites à la figure 4-4.

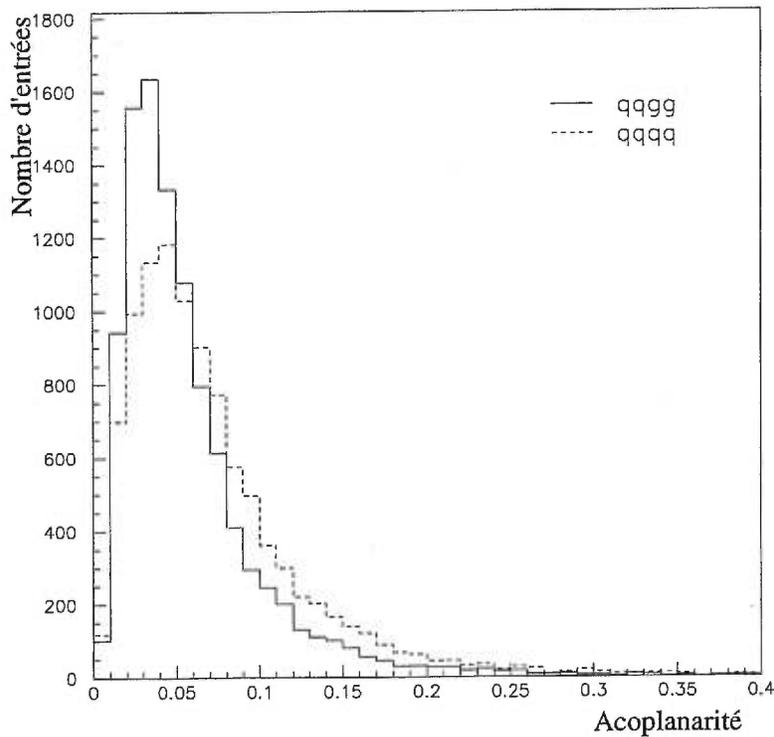


Figure 4-4. Distribution de l'acoplanarité pour les événements à deux quarks et deux gluons (traits pleins) ainsi que pour les événements à quatre quarks (traits pointillés).

4.2.5 Moments de Fox-Wolfram

Une autre façon de décrire la forme globale d'un événement est d'y calculer les moments suivants définis par Fox et Wolfram [13, 2]:

$$H_\ell = \sum_r \sum_s \frac{p^r p^s}{E_{cm}^2} P_\ell(\cos \theta_{rs}) \quad (4.8)$$

où θ_{rs} est l'angle entre la paire de particules (r,s) considérées et P_ℓ est le ℓ ème polynôme de Legendre. Comme H dépend de ℓ , on a donc un moment différent par polynôme de Legendre différent, ce qui fait autant de variables différentes, potentiellement discriminatrices. L'évaluation du pouvoir de discrimination de ces variables nous a cependant conduit qu'à retenir seulement le deuxième, le troisième et le quatrième moment ($\ell = 2, 3, 4$) comme variables de discrimination. Les figures 4-5, 4-6 et 4-7 présentent ces distributions. À cause de leur dépendance en $\cos \theta_{rs}$, ces variables renseignent sur la distribution angulaire des particules détectées (ou détectables). Par exemple, pour le deuxième moment, on a que:

$$P_2 = \frac{1}{2}(3 \cos(\theta_{rs})^2 - 1) \quad (4.9)$$

Ainsi, si l'événement est complètement linéaire, alors l'angle entre chaque paire de particule sera de 0° ou de 180° de sorte que $H_2 = 1$. Par contre, si l'événement est complètement isotropique, alors les valeurs de P_2 positives compenseront donc exactement les valeurs de P_2 négatives de sorte que la somme H_2 sera nulle. La forme des polynômes de Legendre fait que ces cas limites sont les mêmes pour tous les polynômes H_ℓ pour lesquels ℓ est paire [13]. Par contre, pour $\ell = 3$ on a que $H_3 = 0$ pour les événements linéaires et que $H_3 < \frac{5}{8}$ pour les événements sphériques.

Il faut noter que ces variables n'ont rien à voir avec les variables de corrélations angulaires entre les jets. En effet, ce type de variables (l'angle de Bengtsson-Zerwas par exemple) dépend du spin des partons initiaux, alors que les moments de Fox-Wolfram ne renseignent que sur la distribution des particules dans le

détecteur. Notons finalement que H_ℓ est analytique pour tous les ℓ , car les moments de Fox-Wolfram ne sont pas construits à partir d'un quelconque processus d'optimisation.

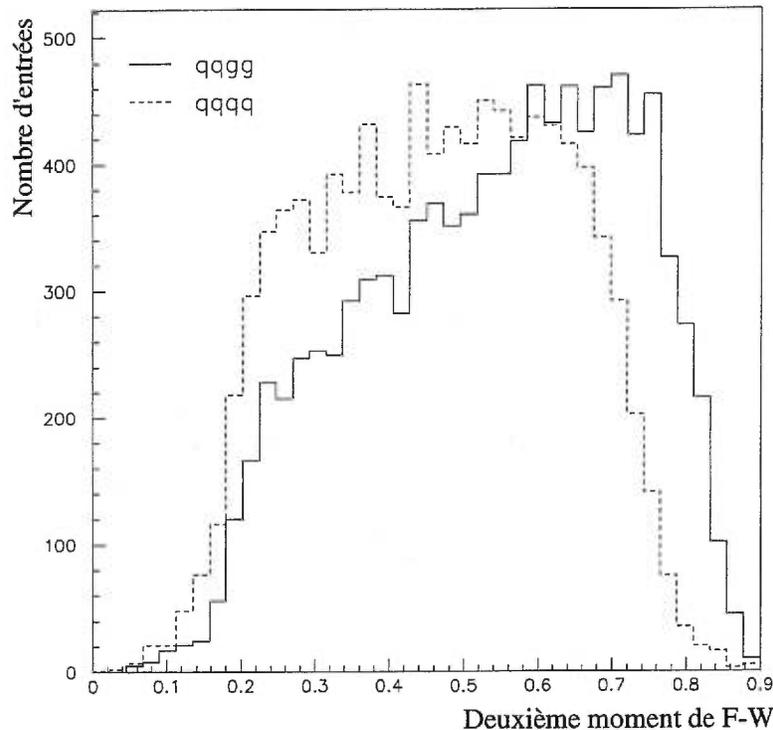


Figure 4-5. Distribution du deuxième moment de Fox-Wolfram pour les événements à deux quarks et deux gluons (traits pleins) ainsi que pour les événements à quatre quarks (traits pointillés).

En considérant les distributions des variables présentées jusqu'ici, il est possible d'établir une interprétation de la topologie générale des événements à quatre jets. En fait, il apparaît que ces événements sont en majorité presque planaires (faible acoplanarité) et possèdent un axe plus prononcé (thrust tendant vers 1). De plus, les particules sont distribuées de sorte que les événements semblent être très linéaires (faibles valeurs de s , D et du troisième moment de Fox-Wolfram et forte valeur du deuxième moment de Fox-Wolfram). Cela

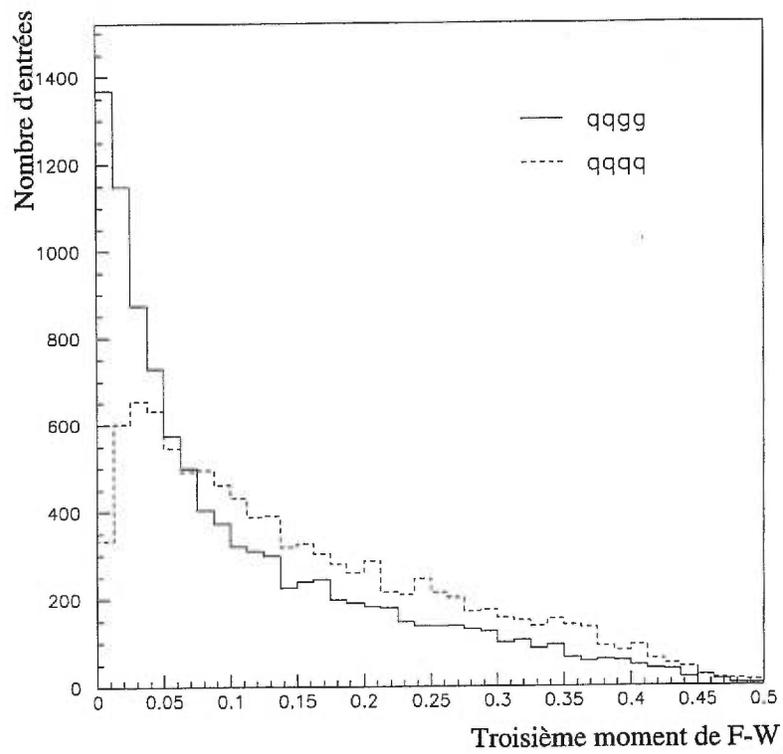


Figure 4-6. Distribution du troisième moment de Fox-Wolfram pour les événements à deux quarks et deux gluons (traits pleins) ainsi que pour les événements à quatre quarks (traits pointillés).

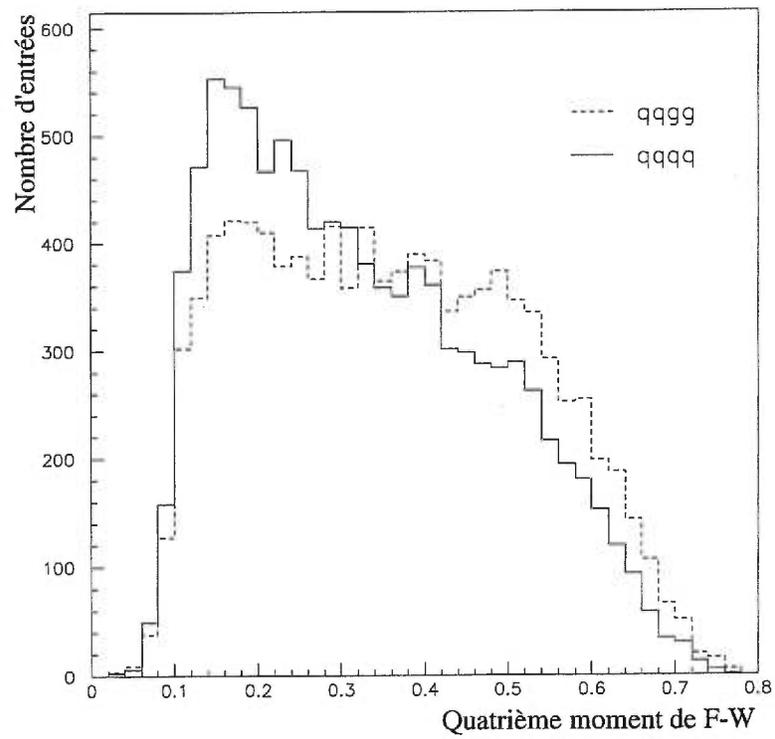


Figure 4-7. Distribution du quatrième moment de Fox-Wolfram pour les événements à deux quarks et deux gluons (traits pointillés) ainsi que pour les événements à quatre quarks (traits pleins).

s'explique par le fait que la probabilité qu'un quark émette un gluon de faible impulsion transverse est plus forte que la probabilité qu'il émette un gluon très dur. Ainsi, les jets issus de la fragmentation d'un des deux quarks produits par la désintégration du Z^0 et du gluon émis par ce quark auront tendance à être colinéaires et il en va de même avec l'autre quark produit par le Z^0 . Cela conduit donc à des événements à quatre jets qui sont mal définis. On verra que cette interprétation sera confirmée par le fait que la majorité de ces événements à quatre jets possèdent une faible valeur de Y^{34} . Cet effet de confusion des événements à quatre jets en événements à deux larges jets sera cependant un peu moins prononcé dans les événements à quatre quarks que dans ceux à deux quarks et deux gluons car les quarks secondaires proviennent du partage d'un gluon émis par un quark parent, diminuant ainsi l'effet de colinéarité de ce gluon. Cela explique donc pourquoi nos variables peuvent être utilisées pour discriminer les événements de type $q\bar{q}gg$ de ceux de type $q\bar{q}q'\bar{q}'$. Si on ne voulait étudier que les événements dont les quatre jets sont bien définis, il faudrait alors effectuer des coupures sur les valeurs prises par les variables topologiques présentées. On pourrait aussi faire une coupure sur la valeur de Y^{34} , ce qui aurait pour conséquence de déplacer les maxima des distributions de s , D et \mathcal{A} vers 1 et le maximum de T vers $\frac{1}{2}$. En fait, les coupures typiques permettant d'avoir des jets bien identifiables éliminent près de 75% des événements à quatre jets.

Maintenant que l'on a présenté et interprété les variables globales utilisées pour la classification des événements à quatre jets ainsi que leurs distributions pour chaque type d'événements, présentons l'évaluation de leur pouvoir relatif de séparation grâce à la comparaison des figures de mérite calculées pour chaque variable. Le tableau 4-I présente cette comparaison. Rappelons encore une fois que les valeurs de F telles qu'indiquées dans ce tableau n'ont de signification que lorsqu'on les compare entre elles.

Variables	Figures de mérite F
Thrust	0.064
2 ^e Moment F-W	0.061
D (Sphérocité)	0.058
3 ^e Moment F-W	0.058
Acoplanarité	0.035
Sphéricité	0.031
4 ^e Moment F-W	0.017

Tableau 4-I. Figure de mérite des sept variables globales sélectionnées pour discerner les événements à quatre quarks parmi ceux à 4-jets.

4.3 Variables de jets

Les variables de jets, contrairement aux variables globales, ne témoignent pas de la topologie d'un événement dans son ensemble mais renseignent sur la structure de chaque jet. Ainsi, on doit obtenir une distribution de chaque variable pour chaque jet. Chaque variable sélectionnée sera donc munie d'un indice (i) spécifiant le jet sur lequel la distribution de la variable est établi. Nous dirons que $i = 1$ correspond au jet le plus énergétique et que $i = 4$ correspond au jet le moins énergétique. Cependant les variables de jets ne serviront pas à discerner directement un événement à quatre quarks parmi un échantillon à quatre jets. Chaque variable servira plutôt à discerner si un jet donné semble plus être un jet issu de la fragmentation d'un gluon ou de la fragmentation d'un quark. Les distributions qui suivent compareront donc la variable A calculée sur le jet i lorsque ce dernier vient d'un quark avec cette même variable, sur le même jet, lorsque celui-ci vient d'un gluon.

4.3.1 Multiplicité

Le nombre de particules contenues dans un jet constitue une façon de discriminer les jets de gluon des jets de quark. En effet, en tenant compte des diagrammes de corrections et des schèmes de fragmentation, on évalue que la multiplicité dans les jets de gluon (m_g) est proportionnelle à $(\ln E)^\gamma$ (où E est l'énergie du jet et γ est un facteur de proportionnalité qui prend sa valeur dans l'intervalle $[0.5, 1]$), alors que dans les jets de quark, la multiplicité (m_q) est proportionnelle à $\frac{e^{\gamma \ln v}}{r}$, où $r = 1,84 \pm 0.02$ lorsque le nombre de saveurs considérées est cinq [18]. On voit donc que l'on s'attend à une plus grande multiplicité dans les jets de gluon que dans les jets de quark. On voit de plus que la multiplicité dépend de façon logarithmique sur l'énergie. Comme la dépendance sur l'énergie dans les jets de gluon est exactement la même que celle dans les

jets de quark, on utilisera plutôt $\frac{m_i}{\ln E_i}$ comme variable de discrimination. Cela nous permettra en outre de diminuer beaucoup la corrélation directe entre deux de nos variables, augmentant ainsi le pouvoir de séparation de notre ensemble de discriminateurs. La figure 4-8 compare les distributions de la multiplicité normalisée des jets de gluon avec celles des jets de quark.

4.3.2 Impulsion transverse au carré

Lorsque l'on considère le mouvement des partons dans la limite asymptotiquement libre, on considère que ces partons se propagent selon un axe donné, que l'on pose comme étant l'axe d'un jet. Cependant, on sait qu'au moment de la fragmentation, on ne peut plus les considérer comme libres. Ainsi, même s'ils se déplacent selon un certain axe, la relation d'incertitude d'Heisenberg autorise également une faible composante de déplacement dans le plan perpendiculaire à la direction de propagation. Ainsi, en se fragmentant, les partons produiront des jets certes orientés selon l'axe du mouvement des partons parents, mais les particules composant ces jets posséderont également une composante d'impulsion transverse aux axes des jets. Cette impulsion transverse est même inférieurement bornée par le principe de Pauli empêchant les particules du jets de se trouver au même point de l'espace-temps. Bien sûr, par conservation nette de l'impulsion pour un intervalle de temps significatif, la somme de chaque impulsion transverse doit être nulle. Par contre, la somme du carré de chaque impulsion (p_T^2) sera non-nulle. Les modèles de fragmentation prévoient par ailleurs un plus grand étalement pour les jets de gluon que pour les jets de quark. La somme sur chaque particule d'un jet des impulsions carrées perpendiculaires à l'axe de ce jet constitue donc une autre variable permettant de séparer les quarks des gluons. Notons que cette variable ne contredit pas le troisième critère imposé à notre choix de variables car bien que le p_T^2 dépende de la fragmentation, cette dépendance n'est pas insensible à la nature des partons, satisfaisant ainsi ce critère. Notons

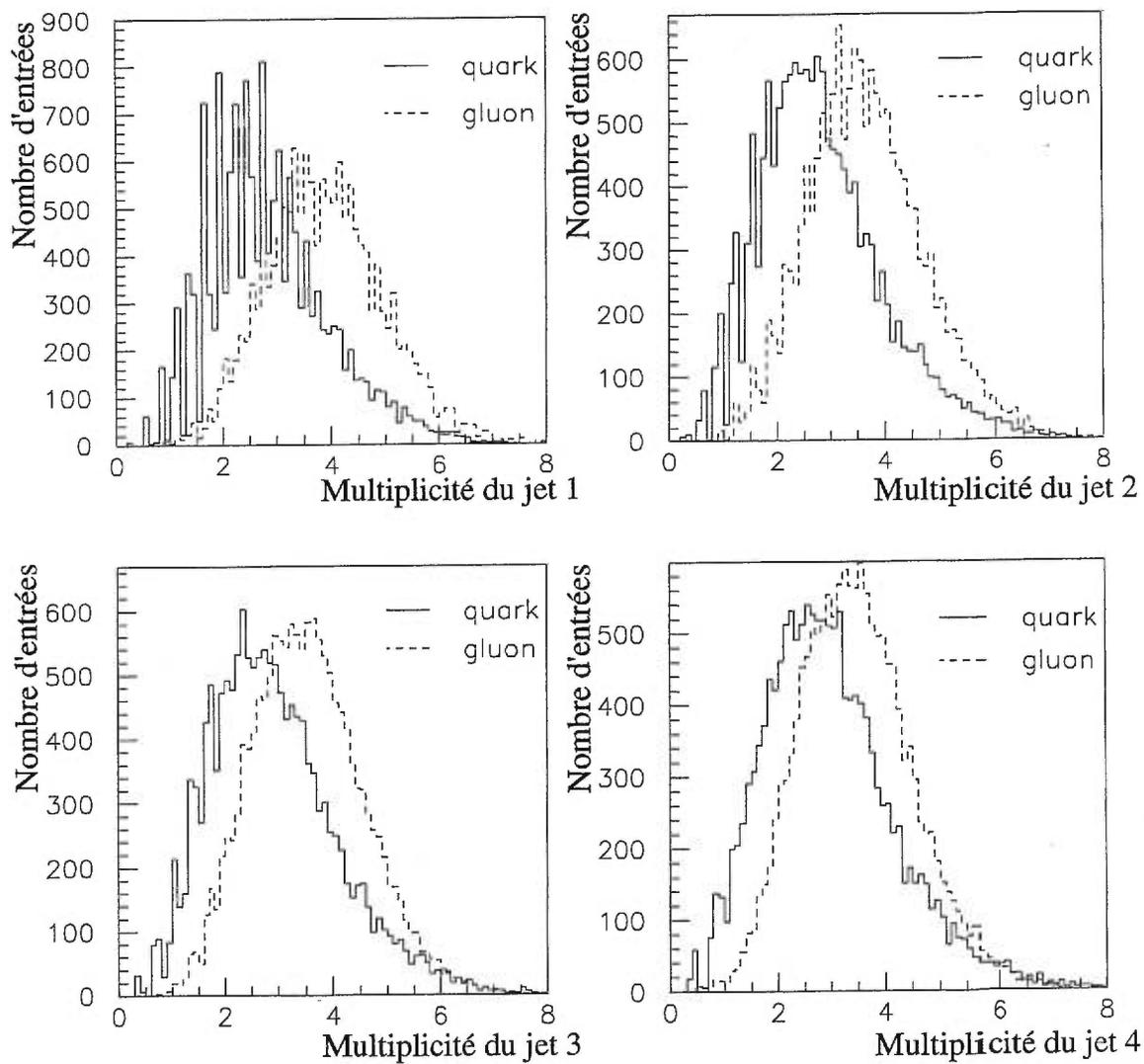


Figure 4-8. Distributions de la multiplicité normalisée ($\frac{m}{\ln E}$) du jet i ($i=1,2,3,4$) lorsque celui-ci est un gluon (traits pointillés) ainsi que lorsqu'il est un quark (traits pleins).

finalement que pour éviter des corrélations avec l'énergie, on a en fait utilisé $\frac{p_T^2}{p^3}$ comme variable. La distribution de cette variable pour chaque jet est donnée à la figure 4-9.

4.3.3 Énergie

Les normalisations effectuées sur la multiplicité et l'impulsion transverse au carré permettent d'inclure l'énergie des jets comme variable discriminante et cela sans que la corrélation entre cette variable et les deux autres (pour chaque jet) soit directe. La figure 4-10 présente les distributions des énergies de chaque jet de gluon et de quark. On remarquera une certaine corrélation entre l'énergie du jet i et celle du jet $i + 1$. Cela est dû au fait que la numérotation des jets est faite en fonction de leur énergie respective. Cette corrélation ne diminue cependant pas notre pouvoir de résolution car elle agit strictement sur les intervalles de valeurs de l'énergie et non sur la forme relative des distributions, de sorte que la distribution E_i et la distribution E_{i+1} ne contiennent pas la même information.

4.3.4 Paramètre de résolution

Avant de pouvoir construire chacune des variables de jets telles que décrites précédemment, il faut tout d'abord reconstruire les jets à partir de la topologie générale de l'événement. Or comme on s'intéresse aux processus d'arbre du deuxième ordre en α_s , tous les événements considérés seront alors reconstruits en quatre jets. La valeur minimale du paramètre y_{ij} lorsqu'un événement est reconstruit en quatre jets (y^{34}) dépendra donc de la topologie générale de l'événement, de sorte que y^{34} prendra des valeurs différentes pour des événements différents. En fait, la distribution de cette variable est différente selon que l'événement est un $q\bar{q}gg$ ou un $q\bar{q}q'\bar{q}'$. On l'ajoutera donc à notre ensemble de variables discriminantes, ce qui complétera la liste. La figure 4-11 présente les

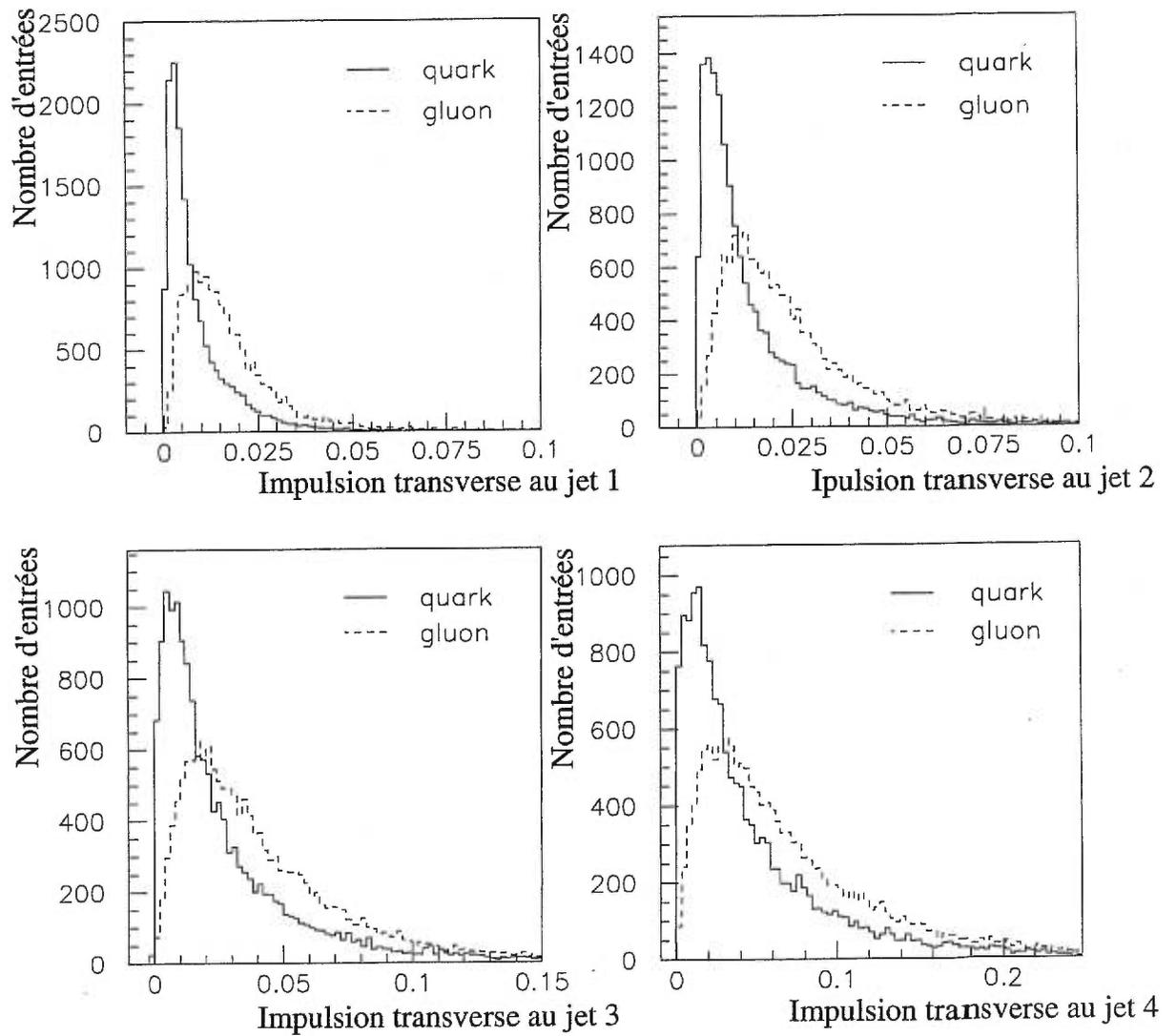


Figure 4-9. Distributions de l'impulsion transverse au carré normalisée du jet i ($i=1,2,3,4$) lorsque celui-ci est un gluon (traits pointillés) ainsi que lorsqu'il est un quark (traits pleins).

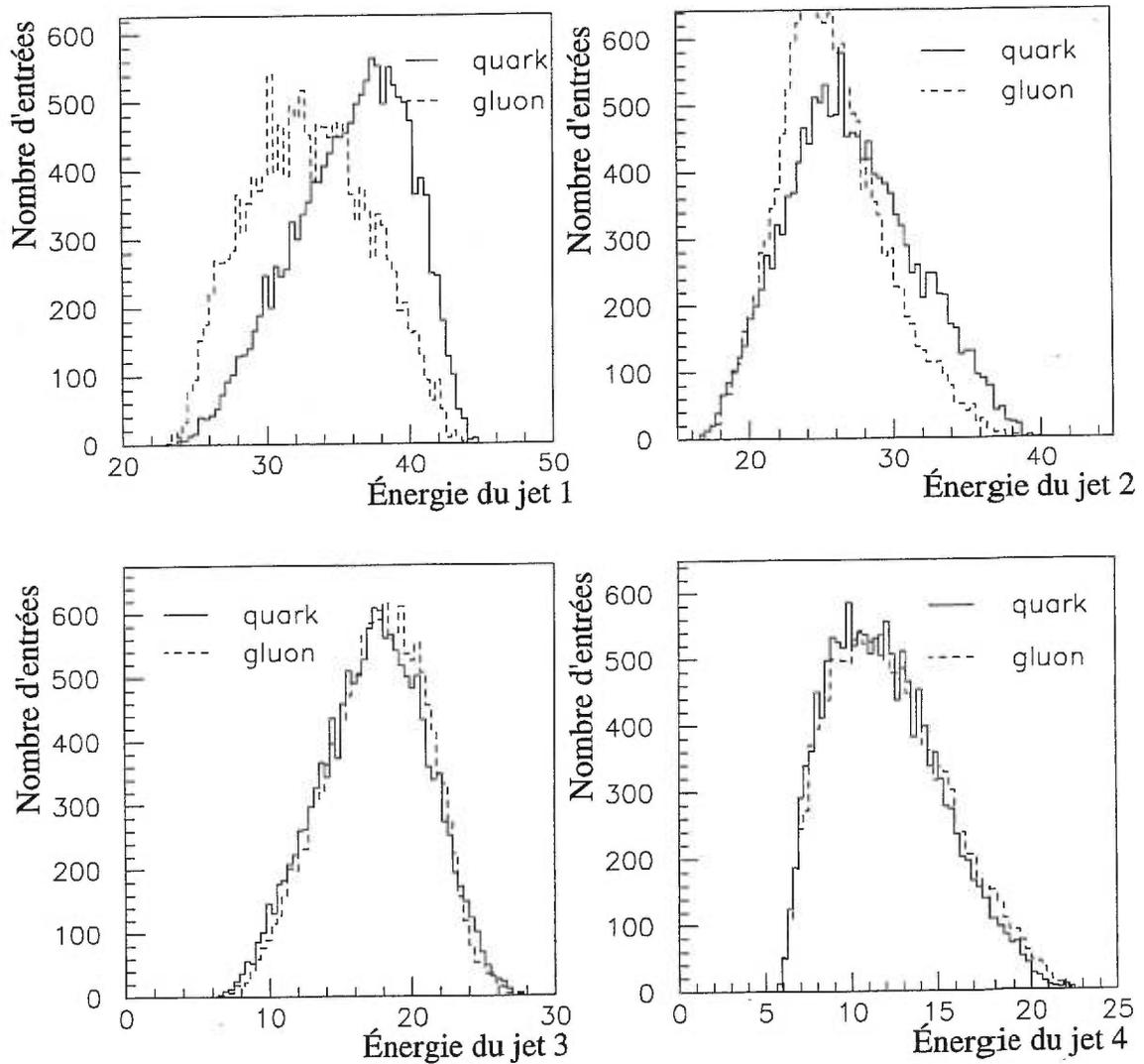


Figure 4-10. Distributions de l'énergie du jet i ($i=1,2,3,4$) lorsque celui-ci est un gluon (traits pointillés) ainsi que lorsqu'il est un quark (traits pleins).

distributions de cette variable pour les deux types d'événements à quatre jets. On remarque sur cette figure que les valeurs de Y^{34} sont faibles, ce qui confirme le fait que la majorité des événements à quatre partons ont une topologie semblable à celle des événements à deux jets. Notons de plus que cette variable n'est pas tout à fait une variable de jets car elle est définie pour un événement complet et non pour les jets qui y sont reconstruits; mais elle n'est pas non plus une variable globale car elle dépend du schème de reconstruction des jets.

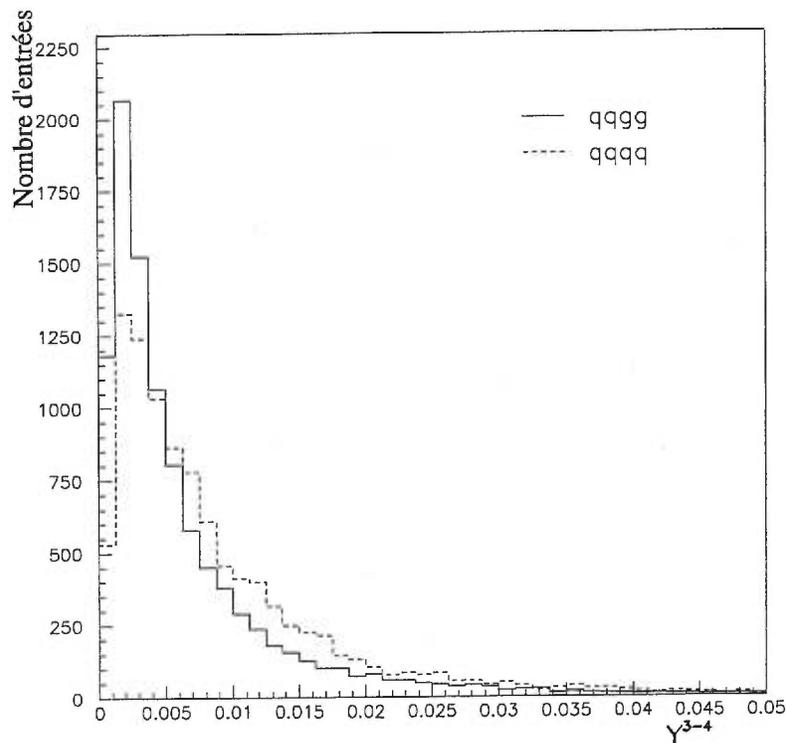


Figure 4-11. Distributions de Y^{34} pour les événements $q\bar{q}g\bar{g}$ (traits pleins) ainsi que pour les événements $q\bar{q}q'\bar{q}'$ (traits pointillés).

Le tableau suivant (4-II) présentera, comme on l'a fait pour les variables globales, les figures de mérite des différentes variables de jets, permettant ainsi de comparer leur pouvoir de séparation relatif. Ces différentes variables, ainsi que les variables globales, seront ensuite introduites dans un réseau de neurones qui

tiendra compte du pouvoir de séparation de chacune pour établir une classification des événements à quatre jets. Avant de voir comment s'effectue cette analyse, regardons donc tout d'abord ce qu'est un réseau de neurones.

Variables	Figures de mérite F			
	Jet 1	Jet 2	Jet 3	Jet 4
Multiplicité	0.327	0.232	0.156	0.118
Impulsion transverse	0.318	0.250	0.161	0.128
Énergie	0.174	0.057	0.009	0.007
y^{34}	0.0422			

Tableau 4-II. Figure de mérite des quatre variables de jet sélectionnées pour discerner les quarks des gluons.

Chapitre 5

Réseaux de neurones artificiels (ANN)

Le mot clé du travail présenté dans ce mémoire est **classification**. En effet, la méthode qui y est développée consiste à classifier un échantillon d'événements donnés en deux catégories différentes. C'est cette classification qui permettra d'obtenir une mesure de la valeur du nombre d'événements à quatre quarks contenus dans un échantillon d'événements à quatre jets et donc une évaluation de la possibilité de découvrir de la nouvelle physique au LEP. Or établir une classification consiste en fait à trouver une prescription effective permettant de déterminer dans quelle classe donnée chaque objet d'un ensemble de départ devra être assimilé. Une classification constitue donc un algorithme de séparation ou de tri d'objets en catégories connues, c'est-à-dire que c'est un algorithme de reconnaissance.

L'instrument de reconnaissance le plus puissant que l'on connaisse est le cerveau humain. En effet, le cerveau ne se limite pas, comme le ferait un logiciel conventionnel, à analyser un objet avec l'ensemble complet des critères permettant d'identifier cet objet, il peut également anticiper sur la reconnaissance d'objets dont il n'a pas une connaissance claire. Par exemple, un enfant n'a pas besoin de connaître et d'analyser le bagage génétique complet de ses parents pour les reconnaître à chaque fois qu'il les voit. Il a simplement à être en contact avec eux un certain nombre de fois, dans certaines situations données, pour les reconnaître dans n'importe quelle autre situation possible. L'algorithme de reconnaissance du cerveau n'a donc besoin que d'apprendre à partir de certains

exemples distinctifs pour accomplir une tâche de classification. C'est cette capacité du cerveau que l'on a cherché à mettre à profit dans l'élaboration de l'algorithme de classification qu'est le réseau de neurones artificiels (ANN). Ainsi un réseau de neurones est un ensemble d'algorithmes caractérisé par une méthode d'apprentissage par exemples connus et par une capacité de généralisation à des situations inconnues, permettant une classification optimale d'objets en classes distinctes. Le réseau de neurones constitue donc l'outil adéquat qui nous permettra de dresser un échantillon d'événements à quatre quarks à partir d'un échantillon d'événements à quatre jets.

5.1 Genèse des réseaux de neurones

L'idée d'emprunter les caractéristiques physiologiques du cerveau comme bases de nouveaux processus permettant d'améliorer la performance des ordinateurs dans les problèmes de reconnaissance de modèles complexes a été amenée par le neuropsychiatre W. S. MacCulloch et le mathématicien W. Pitts dès 1943 [21, 22, 5]. Ils développèrent alors un calculateur constitué d'un réseau de processeurs à deux états ($E=1$ correspondant à l'émission d'un signal, $E=0$ correspondant à un état d'inactivité) permettant de calculer des fonctions booléennes simples. Chacun de ces processeurs était caractérisé par un certain seuil déterminant, à partir d'entrées binaires, l'état dans lequel il devait se placer. Ainsi, par analogie au système nerveux, MacCulloch et Pitts appelèrent ces processeurs des *neurones*. Les réseaux de neurones actuels utilisent en fait une généralisation de cette définition de neurones. On considère donc qu'un neurone est une cellule opératoire caractérisée par [32]:

1. un ensemble de connexions d'entrées x_1, x_2, \dots, x_n indexées par un ensemble de nombres réels w_{ij} qu'on appelle *poids* (correspondant aux synapses d'une cellule nerveuse);
2. une fonction d'entrée v , assurant la combinaison des signaux d'entrée en un potentiel d'activation;
3. une fonction d'activation ϕ qui génèrera la sortie y du neurone grâce à la relation $\phi(v) = y$.

Six ans plus tard, le psychologue Donald Hebb publia un article insistant sur l'importance des connexions entre les synapses des cellules nerveuses pour tout processus d'apprentissage. Cette caractéristique neurologique fut alors mise à profit par Rosenblatt qui améliora l'idée de McCulloch et Pitts et forma ainsi, en 1958, le premier réseau de neurones opérationnel: le *perceptron*. Il s'agissait

simplement de relier tous les neurones d'entrée à chacun des neurones de sortie, établissant ainsi une structure de réseau. De façon générale, on peut donc définir un réseau de neurones comme un ensemble de neurones interconnectés, caractérisé par une certaine architecture et par un certain type de fonction v et ϕ spécifiques.

Le perceptron de Rosenblatt fut construit dans le but de développer un modèle des phénomènes de la vision. Il s'agissait en fait de l'utiliser pour classer de façon formelle chaque stimulus d'un ensemble W dans une des classes de stimuli $C(W)$ prédéterminées. Pour ce faire, Rosenblatt proposa un algorithme permettant de fixer la configuration des neurones à partir d'un ensemble restreint d'exemples de stimuli W' dont les classes sont connues d'avance. De plus, Rosenblatt proposa d'effectuer une réitération récursive de cet algorithme afin d'assurer au réseau une meilleure réaction aux mêmes données, c'est-à-dire d'assurer une plus juste évaluation des classes $C(W')$ par le réseau. Par tout ce processus de configuration, il a donc réussi à formaliser un processus d'*apprentissage* et à l'appliquer aux réseaux de neurones artificiels.

De façon plus technique [21] on peut dire que si $f(x) = y$ constitue une surjection classifiant l'ensemble X de toutes les entrées possibles $x \in X$ en classes $y \in Y$ et que $F(x) = y$ constitue la surjection partielle $(F; A \subseteq X \rightarrow Y)$ représentant le filtrage y d'une sélection d'entrées x par le réseau de neurones, on dira alors que l'apprentissage (supervisé) constitue le processus d'ajustement des paramètres libres de F (c'est-à-dire des poids w_{ij} des neurones) permettant de faire tendre F vers f sur le domaine A . L'apprentissage, qui sert donc à fixer la configuration optimale de la fonction de classification F d'un réseau de neurones, ne constitue cependant que le premier des deux modes de fonctionnement d'un tel réseau. Le deuxième mode de fonctionnement d'un réseau de neurones est la *généralisation*. C'est là que réside toute la puissance de cet outil. En effet, la généralisation est le processus par lequel le réseau étend F sur tout le domaine X , c'est-à-dire utilise sa "connaissance" de f sur A pour extrapoler F sur tout

X. Cette procédure permettra donc d'exécuter la tâche de classification voulue à partir d'entrées inconnues et cela avec un pourcentage de réussite équivalent à celui de la classification des éléments de A . Ainsi, structure neuronale, algorithme d'apprentissage et généralisation constituent les caractéristiques des réseaux de neurones qui permettent de mettre à profit certaines capacités du cerveau humain dans l'élaboration de tâches de classification.

Avec les développements de l'informatique et de la théorie de l'information des années 60-70, les réseaux de neurones sont cependant devenus très généraux [21]. Le perceptron, qui permet de réaliser, par apprentissage supervisé, des tâches de classification, ne constitue plus qu'un des types de réseaux de neurones possibles. Pour effectuer une tâche voulue, il faut donc utiliser le type de réseau qui soit le plus adéquat à la réalisation de cette tâche. Or comme on veut utiliser des distributions de variables construites à partir d'événements à quatre jets connus (entrées) pour entraîner un réseau de neurones à séparer ceux qui sont de type $q\bar{q}q'\bar{q}'$ de ceux qui sont de type $q\bar{q}gg$, le perceptron constituera donc le modèle de réseau de neurones le plus adéquat à la réalisation de cette discrimination d'événements. Cependant, le perceptron de Rosenblatt, dans lequel chaque noeud d'entrée est relié à chaque noeud de sortie, constitue un réseau trop simple pour réaliser des classifications complexes. En effet, il ne peut être entraîné que pour reconnaître des classes séparées linéairement [21, 22]. L'interaction entre les entrées du réseau empêche donc à l'apprentissage du perceptron de Rosenblatt de converger. Cela est malheureux car les variables de discrimination que l'on utilise dans notre analyse sont corrélées. Cependant ce problème peut être surmonté en utilisant un perceptron multicouche, c'est-à-dire un perceptron comportant, outre les couches de neurones d'entrée et de sortie, des couches cachées de neurones faisant le pont entre chacun des neurones d'entrée et chacun des neurones de sortie. Ce type de réseau est standard pour toutes tâches de classification et son utilisation est largement répandue dans les analyses en physique des hautes énergies [7, 32, 5]. Étudions le donc d'un peu plus près.

5.2 Fonctionnement d'un perceptron multicouche

Le perceptron multicouche possède une architecture qui permet un traitement de l'information qu'on qualifie de "feed-forward" parce qu'il ne garde pas en mémoire les valeurs de sortie et les états d'activation des neurones des étapes précédentes, c'est-à-dire qu'il ne fait que procéder les entrées vers la sortie, sans retour à une étape de calcul déjà réalisée. On qualifie par ailleurs son algorithme d'apprentissage d'algorithme de "rétro-propagation". En effet, la valeur de sortie obtenue par le réseau est comparée à la valeur visée pour calculer une fonction d'erreur dont la minimisation sera assurée par une modification adéquate des poids de chaque neurone. L'apprentissage part donc de la sortie et revient vers les poids pour optimiser la qualité de la classification, d'où "rétro-propagation". Ces deux processus assureront la structure générale d'opération d'un perceptron multicouche. Regardons donc plus en détail comment elles se mettent en oeuvre dans le fonctionnement d'un tel réseau. Pour simplifier la présentation de tout l'algorithme, considérons un réseau possédant une seule couche interne. Cela ne constitue cependant pas un choix arbitraire car il faut savoir que l'efficacité de l'apprentissage par rétro-propagation diminue lorsque le réseau comporte plus d'une couche cachée [5]. On fixera sur ce réseau n noeuds pour la couche d'entrée, m pour la couche cachée et un seul pour la couche de sortie (sortie binaire). Finalement nous considérerons p exemples d'entrées à partir desquels le réseau devra apprendre. Notons au passage que ce choix de structure présente l'intérêt inavoué de correspondre au réseau qui sera utilisé dans notre analyse.

On représente dans un premier temps nos p exemples connus par un ensemble de p vecteurs à n composantes. On associera à chacun de ces vecteurs une étiquette d_q représentant la valeur de sortie désirée, c'est-à-dire la classe dans laquelle le vecteur q est associé. Dans notre cas, les p vecteurs utilisés pour l'entraînement (ou apprentissage) correspondront à p événements dont la nature est connue d'avance. Les n composantes de chaque vecteur correspondront aux

valeurs des n variables calculées pour l'événement utilisé. Finalement d_q indiquera si l'événement q est un $q\bar{q}gg$ (dans ce cas $d_q = 0$) ou un $q\bar{q}q'\bar{q}'$ (dans ce cas $d_q = 1$). On générera ensuite un ensemble aléatoire de $n \times m$ valeurs réelles prises entre 0 et 1. Chacune de ces valeurs constituera un des poids reliant les neurones d'entrée aux neurones internes, w_{ij} étant le poids de la connexion entre le i^e neurone d'entrée et le j^e neurone interne. Une fois que la matrice des poids et les vecteurs d'entrée auront été définis, on sera maintenant prêt à entamer la procédure d'apprentissage. Celle-ci se développe en 8 étapes [21, 22, 5]:

1. On associe chaque coordonnée du vecteur d'entrée \vec{x}_1 à un noeud ou neurone de la couche d'entrée:

$$\vec{x} = (x_1, x_2, \dots, x_n)$$

2. On calcule la valeur d'entrée v_j du j^e neurone interne à partir de chacune des valeurs d'entrée grâce à la relation:

$$v_j = \sum_{i=1}^n x_i w_{ij} + \theta_j$$

où le seuil θ_j est traité comme un poids supplémentaire facilitant la convergence. Notons qu'en raison de la forme de cette fonction d'entrée interne, on dit que le perceptron multicouche est formé de neurones *produits scalaires*.

3. On génère la réponse de ce j^e neurone interne grâce à l'application de la fonction d'activation ϕ de ce neurone à son potentiel d'activation v_j . La sortie du neurone sera donc:

$$s_j = \phi(v_j)$$

Comme on veut que la réponse du réseau soit binaire, on utilisera une fonction d'activation sigmoïde (figure 5-1):

$$\phi(x) = \frac{1}{1 + \exp \frac{-2x}{T}}$$

où T s'appelle la température du réseau.

4. On répète le calcul des étapes 2) et 3) pour chaque noeud de la couche interne.
5. On calcule la valeur y de sortie du réseau à partir des réponses de chaque neurone interne:

$$y = g(\sum_j w_j s_j)$$

où les w_j représentent les poids des connexions des neurones internes au neurone de sortie et g est, comme ϕ , une fonction sigmoïde. Cette sortie sera une valeur réelle comprise dans l'intervalle $[0, 1]$.

6. On compare la sortie y du réseau avec la valeur cible d_1 et on définit le terme d'erreur:

$$\epsilon_1 = |d_1 - y|$$

7. On minimise la fonction d'erreur

$$E = \frac{1}{2} \epsilon_1^2 \iff E = E(\vec{w})$$

qui réajustera la valeur de tous les poids w .

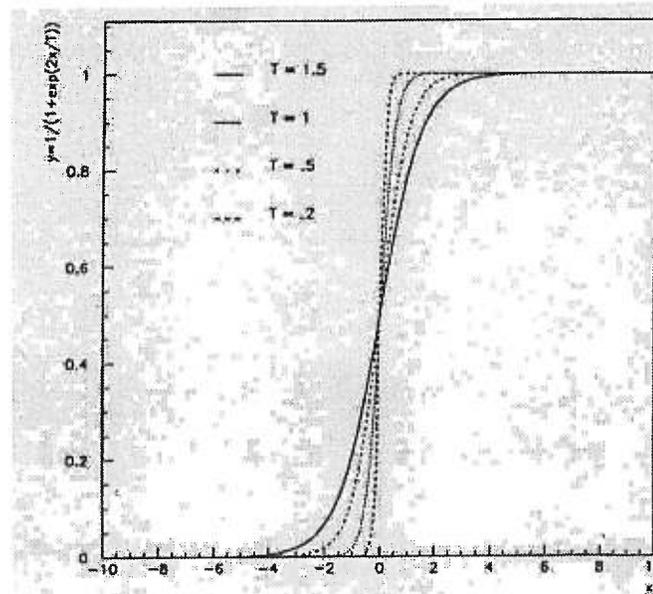


Figure 5-1. Dépendance sur la température (pente) de la fonction de sortie sigmoïde.

8. On recommence toute la procédure avec le vecteur \vec{x}_2 et on réitérera ainsi de suite jusqu'à \vec{x}_p .

La fonction d'erreur définie en 7) peut être vue comme une surface dans l'espace des poids. On peut donc minimiser cette fonction par la méthode de *descente du gradient*. Cette méthode consiste à partir des poids $\vec{w}(q)$ de l'étape q pour descendre le long de cette surface par petits sauts en direction de la pente. Lorsque le minimum sera atteint, on redéfinira les poids de l'étape $q + 1$ par la relation:

$$\vec{w}(q+1) = \vec{w}(q) - \eta \nabla E(w(q)) = \vec{w}(q) - \eta \frac{\partial E}{\partial w_{ij}} + \alpha \Delta \vec{w}(q-1) \quad (5.1)$$

où les paramètres η et α ($\alpha < 1$) déterminent la stabilité et la rapidité de convergence du vecteur poids et où $\Delta \vec{w}(q-1)$ est la modification des poids de l'étape $q-1$.

Lorsque la procédure aura été appliquée au dernier vecteur d'entrée \vec{x}_p , les poids recalculés seront alors fixés et le réseau sera maintenant prêt à être

utilisé sur des exemples d'événements inconnus, c'est-à-dire que le réseau sera prêt à fonctionner en mode généralisation. La procédure de fonctionnement du réseau dans ce mode d'opération est exactement la même que celle du mode d'apprentissage, sauf pour l'étape 7) qui est abandonnée. On dit donc que les six premières étapes ainsi que la huitième constituent les étapes de fonctionnement d'un réseau "feed-forward" représentées au schéma 5-2, alors que la septième étape caractérise l'apprentissage par rétro-propagation.

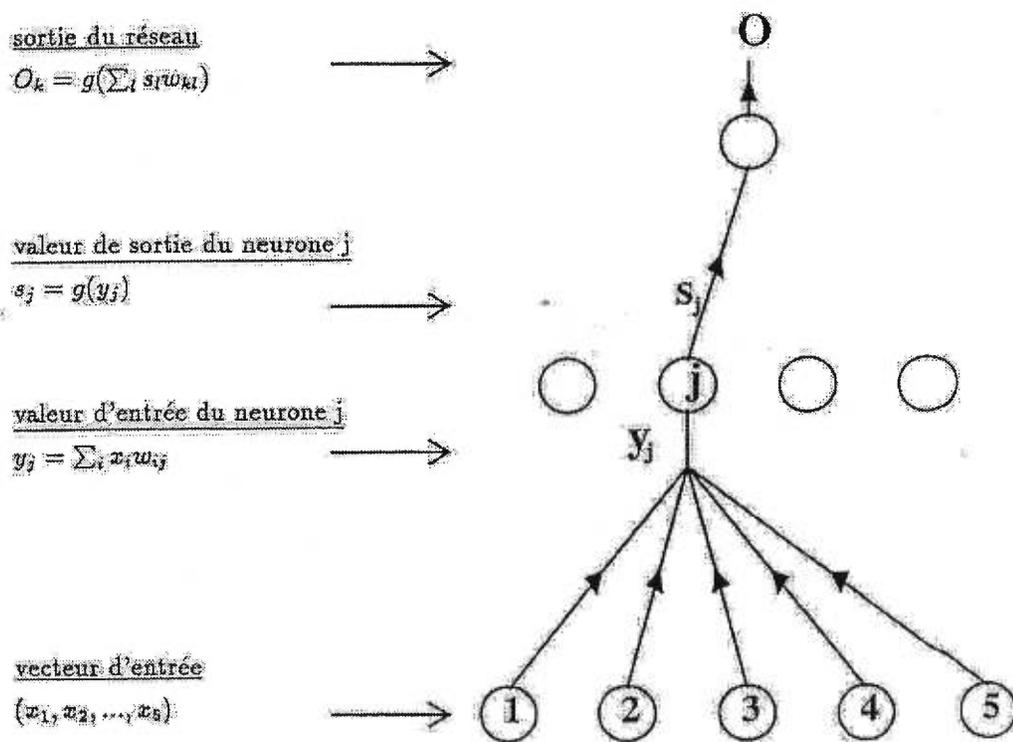


Figure 5-2. Schéma de la propagation des données dans un réseau feed-forward à cinq entrées, une couche cachée et une sortie.

Bien que le réseau que l'on vient de décrire constitue un cas particulier de réseaux destinés à effectuer une tâche quelconque de classification, il ne constitue pas encore la spécification complète nécessaire pour réaliser effectivement et optimalement notre classification précise. En effet, la procession des entrées dans le réseau et le calcul de la minimisation d'erreurs propre à l'apprentissage dépendent

de paramètres libres qui doivent être fixés avant le début de l'entraînement de façon à ce que la procédure de classification soit optimale. Deux de ces paramètres ont déjà été discutés: il s'agit du nombre de noeuds dans chaque couche ainsi que de la valeur initiale attribuée aux poids. Cependant, avant d'entamer le processus d'apprentissage il faut également fixer la température T indiquant la pente de la fonction sigmoïde, le taux d'apprentissage η déterminant le nombre d'itérations nécessaires avant un ajustement significatif des poids et finalement le paramètre cinétique α spécifiant la vitesse de convergence du processus d'apprentissage. Il n'y a aucune loi générale déterminant les valeurs que ces paramètres doivent prendre pour optimiser la performance du réseau. Il faudra donc les déterminer empiriquement. Une fois que différents essais sur le réseau auront été effectués pour déterminer la valeur de ces paramètres, le réseau sera alors complètement spécifique au problème considéré.

Chapitre 6

Classification des événements à quatre jets à l'aide d'un réseau de neurones artificiel

Ce qui nous motive depuis le début de ce travail c'est, on le répète, de développer une méthode de mesure de la variable R_{4q} qui soit indépendante des méthodes de mesure standard utilisant des corrélations angulaires entre les jets. Pour bien comprendre ce problème et en délimiter l'analyse, on a tout d'abord spécifié le contexte théorique dans lequel il apparaît, c'est-à-dire l'étude de la CDQ perturbative dans le Modèle Standard, ainsi que le contexte expérimental dans lequel l'analyse se situe, c'est-à-dire l'expérience OPAL du LEP au CERN. On a ensuite exposé les prémisses de cette méthode: utiliser un ensemble de variables topologiques discriminant entre les événements $q\bar{q}gg$ et les événements $q\bar{q}q'\bar{q}'$ et additionner le pouvoir de séparation de chacune de ces variables pour dresser une classification de ces deux types d'événements au sein d'un échantillon d'événements à quatre jets. On a finalement présenté un outil permettant d'exécuter cette classification: le réseau de neurones. Il nous reste donc à montrer comment cet outil nous permet d'obtenir une évaluation effective du nombre d'événements à quatre quarks contenus dans un échantillon à quatre jets. Cependant, on a vu que les réseaux de neurones sont très dépendants des types de problèmes de classification particuliers. Ainsi, avant de voir comment on obtient cette évaluation de la fraction apparente d'événements à quatre quarks, il faut tout d'abord établir la caractérisation complète du réseau de neurones effectivement utilisé pour résoudre ce problème.

6.1 Caractérisation du réseau de neurones utilisé

6.1.1 Architecture du réseau

L'architecture globale du réseau utilisé dans cette analyse consiste, comme on l'a déjà mentionné au chapitre précédent, en un perceptron comportant une seule couche cachée et un seul noeud à la sortie. Le nombre de noeuds dans la couche d'entrée étant fixé par le nombre de variables de discrimination utilisées, il ne reste plus qu'à fixer le nombre de noeuds dans la couche interne de sorte que la réponse du réseau soit optimale. Cependant, dans notre cas, on a vérifié que la sortie du réseau ne varie pratiquement pas lorsque le nombre m de noeuds de la couche interne est pris dans l'intervalle $[\frac{n}{2}, \frac{3n}{2}]$ où n est le nombre de variables. On a donc choisi $m = n$. Ainsi, comme on utilise 20 variables discriminantes, notre réseau est donc formé de 20 noeuds dans la couche d'entrée, 20 noeuds dans la couche interne et un seul dans la couche de sortie.

6.1.2 JETNET

Le réseau de neurones que nous utilisons dans cette analyse est construit à partir des bibliothèques du programme JETNET3.0 [4] qui permet toutes les libertés nécessaires à la résolution optimale de notre problème particulier. En effet, la procédure "feed-forward" du réseau est fixée par l'emploi des sous-routines JNINIT, JNTRAL, JNTEST, JNDUMP et JNREAD définies dans les bibliothèques de JETNET. Le reste des variables nécessaires à la construction et à la caractérisation complète du réseau sont contenues dans le bloc commun JNDAT1, également défini dans les bibliothèques de ce programme. Il s'agit donc pour l'utilisateur de configurer son réseau simplement en modifiant les différentes valeurs prises par les composantes des vecteurs MSTJN et PARJN qui sont contenues dans JNDAT1. Dans notre cas, pour caractériser complètement notre réseau, il suffit de fixer les valeurs

de seulement dix composantes de ces vecteurs dont les correspondances sont les suivantes:

- $MSTJN(1)$ = nombre de couches dans le réseau
- $MSTJN(2)$ = nombre de cycles par mise à jour
- $MSTJN(3)$ = fonction de réponse g et ϕ
- $MSTJN(10)$ = nombre de noeuds dans la couche d'entrée
- $MSTJN(11)$ = nombre de noeuds dans la couche interne
- $MSTJN(12)$ = nombre de noeuds dans la couche de sortie
- $PARJN(1)$ = paramètre d'apprentissage η
- $PARJN(2)$ = paramètre cinétique α
- $PARJN(3)$ = paramètre de température T

Toutes les autres composantes des vecteurs $MSTJN$ et $PARJN$ seront laissées à leur valeur de défaut, définies pour un perceptron multicouche standard (feed-forward, rétro-propagation), ce qui correspond au cas qui nous intéresse. Notons que l'on a déjà mentionné que la fonction de réponse du perceptron à classification binaire doit être une fonction sigmoïde. Cela fixe donc la valeur de $MSTJN(3)$ à 1, correspondant à ce type de fonction. Notons également que le manuel d'utilisation de JETNET [4] suggère d'être très prudent dans le choix des paramètres α et η lorsqu'on utilise la procédure normale d'apprentissage (c'est-à-dire l'algorithme de rétro-propagation, correspondant à $MSTJN(5)=0$) car un choix non-conservateur peut avoir une très grande influence sur la performance du réseau. On s'en est donc tenu aux valeurs de défaut, standard pour ce type de réseau et qui ont d'ailleurs déjà été utilisées dans des analyses similaires à la nôtre en physique de hautes énergies (cf. [5]).

6.1.3 Caractérisation empirique

Pour connaître les autres paramètres du réseau optimisant son efficacité de discrimination, il faudra introduire les valeurs des variables discriminatrices événement par événement dans la couche d'entrée et examiner cas par cas comment le réseau réagit aux différentes valeurs des paramètres qui y sont temporairement fixées. C'est pourquoi on parle de caractérisation empirique de la valeur des paramètres libres du réseau.

Les valeurs des variables de chaque événement présenté seront introduites dans le réseau via un n -tuple préalablement généré par PYTHIA6.0. Pour la période d'apprentissage, le réseau devra en fait lire aléatoirement les éléments de deux n -tuples différents: un présentant les variables calculées strictement pour les événements $q\bar{q}gg$ et l'autre présentant les variables calculées à partir des événements $q\bar{q}q'\bar{q}'$. Lorsque l'apprentissage sera terminé, le réseau devra opérer en mode test, c'est-à-dire qu'il devra lire deux autres n -tuples (un pour chaque type d'événement) et établir la classification à partir des poids obtenus lors de l'apprentissage. Comme la nature des événements présentés lors du test nous est déjà accessible avant même leur passage dans le réseau (parce que nous avons produit les événements d'où proviennent les n -tuples), il suffit de les étiqueter de sorte que l'on puisse évaluer le pouvoir de généralisation du réseau en évaluant la même fonction d'erreur que celle définie pour l'apprentissage, mais sans toutefois remodifier les poids. Toute cette procédure de test est exécutée par la sous-routine JNTEST de JETNET qui nous donne, à la fin de l'exécution du programme, le pourcentage de réussite de l'apprentissage ainsi que celui du test. L'optimisation de ce pourcentage de réussite constituera donc un critère permettant de fixer empiriquement la valeur des paramètres présentés plus haut.

La décision de ce qu'est une réussite est cependant arbitraire lorsque la classification n'est pas parfaite, c'est-à-dire lorsque les courbes de sortie (réponse

du réseau) des deux types d'événements étiquetés se chevauchent. En effet, l'association de la sortie à une classe ou à une autre dépend d'une coupure qui fixe la décision. On peut par exemple dire que tous les événements pour lesquels la sortie du réseau est supérieure à 0.4 seront considérés comme des quatre quarks (dont la valeur cible ou idéale est 1) alors que les autres seront considérés comme des deux quarks et deux gluons. On aurait cependant tout aussi bien pu choisir 0.515 comme valeur de coupure, changeant du coup la valeur du nombre d'événements assignés à chaque classe. Le pourcentage de réussite peut donc varier (d'environ 5% dans notre cas) sur la gamme de ces coupures. Ainsi, si la valeur du pourcentage de réussite du réseau varie de façon substantielle (c'est-à-dire de l'ordre de 5% ou davantage) en fonction des valeurs que peuvent prendre un paramètre donné, alors on peut utiliser ce critère d'optimisation du pourcentage de réussite du réseau pour fixer la valeur de ce paramètre. Par exemple, dans le cas du paramètre déterminant le nombre de cycles par mise à jour, on a obtenu (après 10 000 cycles d'entraînement) que:

- $MSTJN(2)=1$ \Rightarrow 63.4% de réussite maximale
- $MSTJN(2)=10$ \Rightarrow 58.2% de réussite maximale
- $MSTJN(2)=100$ \Rightarrow 57.4% de réussite maximale

On a donc choisi 1 comme valeur de ce paramètre.

Par contre, lorsque la variation de ce pourcentage de réussite n'est pas suffisamment grande, on ne peut plus l'utiliser comme critère permettant de fixer un paramètre libre du réseau. Il faut donc trouver une autre détermination de la qualité d'un réseau et voir comment ce nouveau facteur de qualité réagit aux variations du paramètre à fixer. Or la capacité du réseau à fournir des distributions de sortie distinctes les unes des autres (c'est-à-dire qui se chevauchent peu) est une autre qualité d'un réseau de neurones. Elle indique la capacité du

réseau d'émettre une décision claire sur la valeur de sortie, que le réseau se trompe ou non. Or on possède justement un critère permettant d'évaluer la dispersion entre deux courbes: l'évaluation des figures de mérite (cf. équation (4.1)). Il s'agira donc pour nous de comparer de façon systématique les courbes de sortie du réseau de neurones générées pour différentes valeurs des paramètres libres du réseau (énumérés plus haut) et de choisir les valeurs pour lesquelles l'évaluation de la figure de mérite de ces courbes de sortie est maximale. Pour le cas de la température, on a obtenu:

Valeur de température	Figure de mérite de la sortie
0.25	0.112
0.5	0.113
1.0	0.133
2.0	0.131
3.0	0.127

La valeur de température choisie pour notre réseau de neurones sera donc $T = 1.0$.

Maintenant que l'on a fixé la température, on connaît donc chacun des paramètres qui seront utilisés par le réseau dans son algorithme de classification des événements à quatre jets à partir des distributions de variables topologiques. Le tableau (6-I) résume les valeurs choisies pour chacun des paramètres listés plus haut.

Paramètres libres du réseau	Valeurs choisies pour notre analyse
MSTJN(1)	3
MSTJN(2)	1
MSTJN(3)	1
MSTJN(10)	20
MSTJN(11)	20
MSTJN(12)	20
PARJN(1)	0.001
PARJN(2)	0.5
PARJN(3)	1.0

Tableau 6-I. Valeurs choisies des divers paramètres libres fixant la configuration du réseau de neurones que nous utiliserons pour notre analyse.

Il reste encore une étape pour déterminer complètement la procédure de classification du réseau: il faut déterminer quand on devra considérer l'apprentissage comme terminé, c'est-à-dire déterminer sur combien de cycle (epoch) cet apprentissage doit s'étendre. Pour ce faire, on applique la même méthode que celle qu'on a utilisée pour déterminer la température du réseau: on calcule la figure de mérite de la sortie du réseau en fonction du nombre d'epoch et lorsque la courbe de cette fonction ($f(\text{epoch})$) atteindra un plateau, on considérera alors que le réseau ne peut plus apprendre davantage. La figure 6-1 présente cette courbe de l'évolution de la figure de mérite en fonction du nombre d'epoch. Comme elle présente une pente quasiment nulle après 10000 epoch, on effectuera donc 10000 cycles d'entraînement sur le réseau, avant de fixer les poids pour la généralisation.

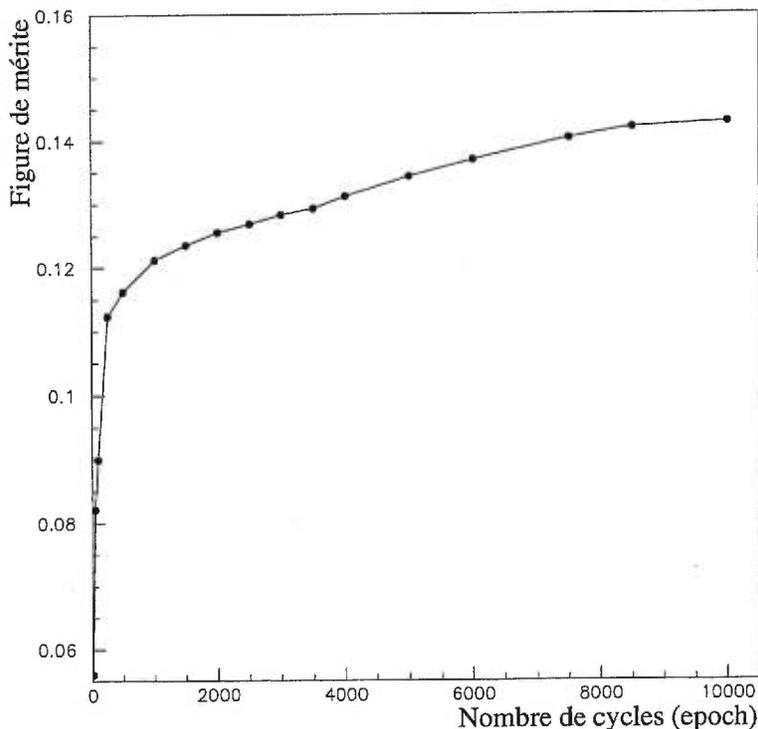


Figure 6-1. Courbe de la variation de la figure de mérite de la sortie du réseau en fonction du nombre de cycles d'entraînement (epoch).

6.2 Analyse de la réponse du réseau

Maintenant que notre réseau de neurones est configuré pour résoudre optimalement notre problème de classification, il ne reste plus qu'à y entrer les distributions (construites à partir de 10000 événements de chaque type) et d'en analyser la sortie. Avant d'analyser cette sortie il faut cependant faire une remarque très importante: les distributions qui ont été fournies au réseau pour son apprentissage ne correspondent pas, pour le cas des variables de jets, aux distributions qui ont été présentées au chapitre 4. En effet, ces distributions discriminent entre un jet de gluon et un jet de quark et non entre un événement à quatre quarks et un événement à deux quarks et deux gluons. Les classes de ces distributions ne sont donc pas celles nécessaires à l'évaluation de R_{4q} , contrairement au cas des variables globales. Il s'agit donc de produire les distributions de ces variables pour chacun des jets à partir des étiquettes de classes qui nous intéressent. Ainsi la distribution de la multiplicité du premier jet ne comparera plus le cas où ce jet provient d'un quark avec le cas où ce jet provient d'un gluon, il comparera plutôt le cas où la multiplicité du premier jet est calculée pour un événement $q\bar{q}gg$ avec celui où elle est calculée pour un événement $q\bar{q}q'\bar{q}'$. Or, comme dans chaque événement $q\bar{q}gg$ il y a deux jets qui proviennent des quarks, la distribution d'une variable de jets pour un ensemble d'événements $q\bar{q}gg$ comprendra donc un mélange de cas où le jet proviendra d'un quark avec ceux où il proviendra d'un gluon, ce mélange étant proportionné par la probabilité théorique que le jet ayant la fraction x_i de l'impulsion totale des quatre jets soit un quark ou un gluon (probabilité calculée à l'aide des équations d'Altarelli-Parisi implantées dans JETSET). Bien sûr, cela fait que le pouvoir de discrimination de ces variables de jet en souffrira beaucoup. Par contre, le fait que l'on tienne compte dans nos distributions de la probabilité adéquate qu'un jet donné provienne d'un quark ou d'un gluon lorsque l'événement est de type $q\bar{q}gg$ nous assure que notre méthode de discrimination des événements à quatre jets est cohérente avec la physique

du Modèle Standard dans lequel cette étude s'inscrit. De plus, ces variables continueront tout de même à constituer des discriminants complétant le pouvoir de séparation des variables globales. Les figures 6-2, 6-3 et 6-4 présentent ces nouvelles distributions.

En utilisant ces distributions (ainsi que celles des variables globales), comme entrées du réseau de neurones dont la configuration a été établie plus tôt, on obtiendra alors la sortie présentée à la figure 6-5.

Analysons maintenant cette figure pour voir comment elle nous permettra d'obtenir l'information voulue, c'est-à-dire une évaluation du nombre d'événements à quatre quarks contenus dans un échantillon d'événements à quatre jets. Bien entendu cette courbe n'étant que la caractérisation de la capacité du réseau à classifier un échantillon inconnu d'événements à quatre jets, elle ne sera pas utilisée directement pour le calcul du nombre d'événements à quatre quarks contenus dans un tel échantillon. En effet, si on construit les distributions des variables sélectionnées à partir d'un échantillon de données OPAL reconstruites en quatre jets, la sortie du réseau sera une combinaison linéaire des deux courbes présentées à la figure 6-5. Alors comment faire, puisque la classification n'est pas parfaite, pour savoir quels événements sont de type $q\bar{q}gg$ et quels sont ceux de type $q\bar{q}q'\bar{q}'$?

Pour cela il faut tout d'abord se fixer une valeur de coupure arbitraire sur les valeurs de sortie du réseau. Ainsi, par ce choix de coupure, on décidera que tous les événements qui ont une valeur de sortie supérieure à la valeur de la coupure seront considérés comme des événements $q\bar{q}q'\bar{q}'$, alors que les autres seront pris comme des événements $q\bar{q}gg$. Bien sûr, les courbes de sortie présentées à la figure 6-5 nous montrent que peu importe la valeur de coupure choisie, il y aura toujours des événements $q\bar{q}gg$ qui passeront cette coupure et des événements $q\bar{q}q'\bar{q}'$ qui ne la passeront pas. Ainsi, le nombre d'événements passant la coupure ne pourra en aucun cas représenter le nombre exact d'événements à quatre quarks

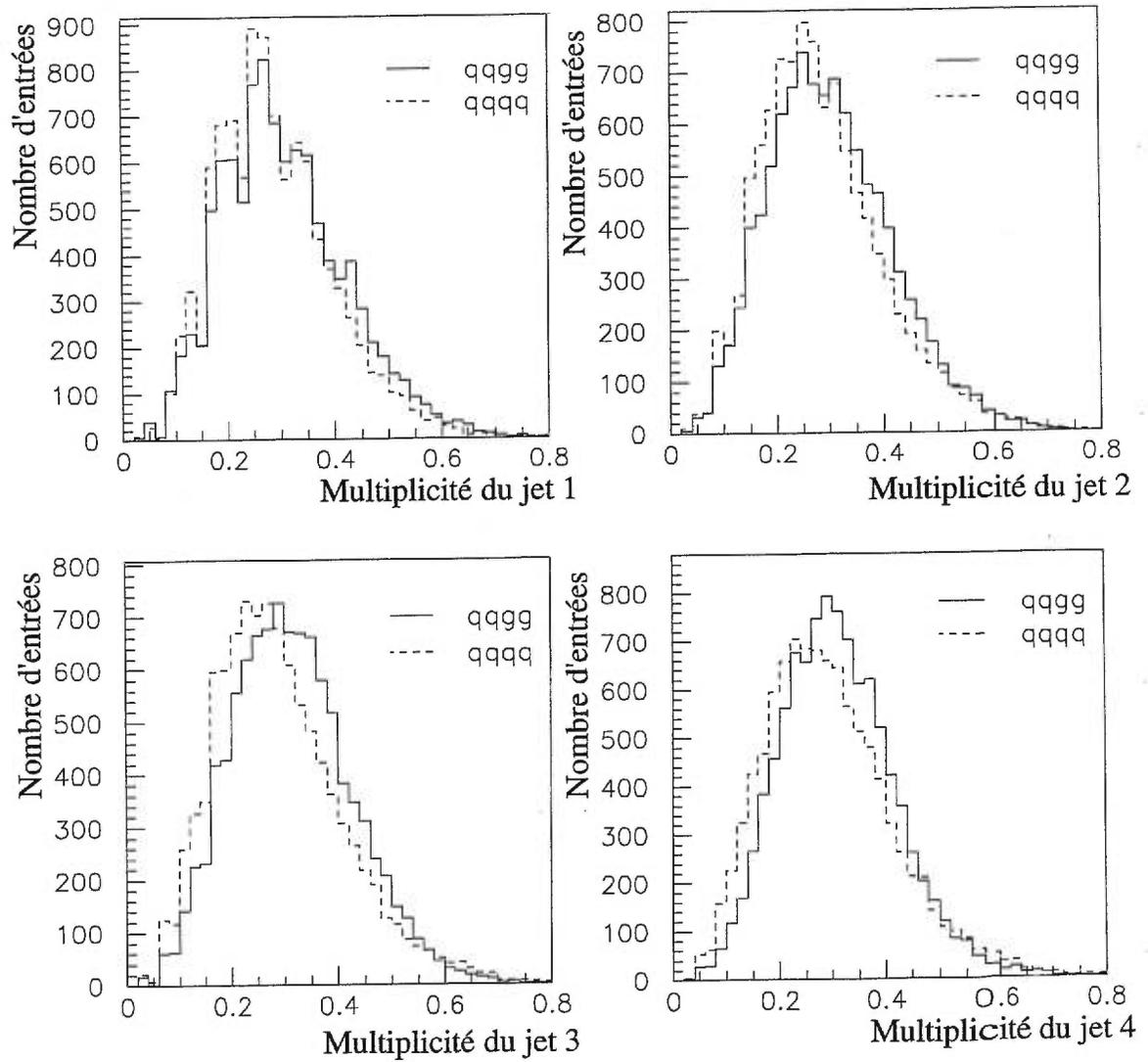


Figure 6-2. Distribution de la multiplicité normalisée pour chacun des quatre jets lorsque l'événement est un $q\bar{q}gg$ (traits pleins) ou lorsque c'est un $q\bar{q}q'\bar{q}'$ (traits pointillés).

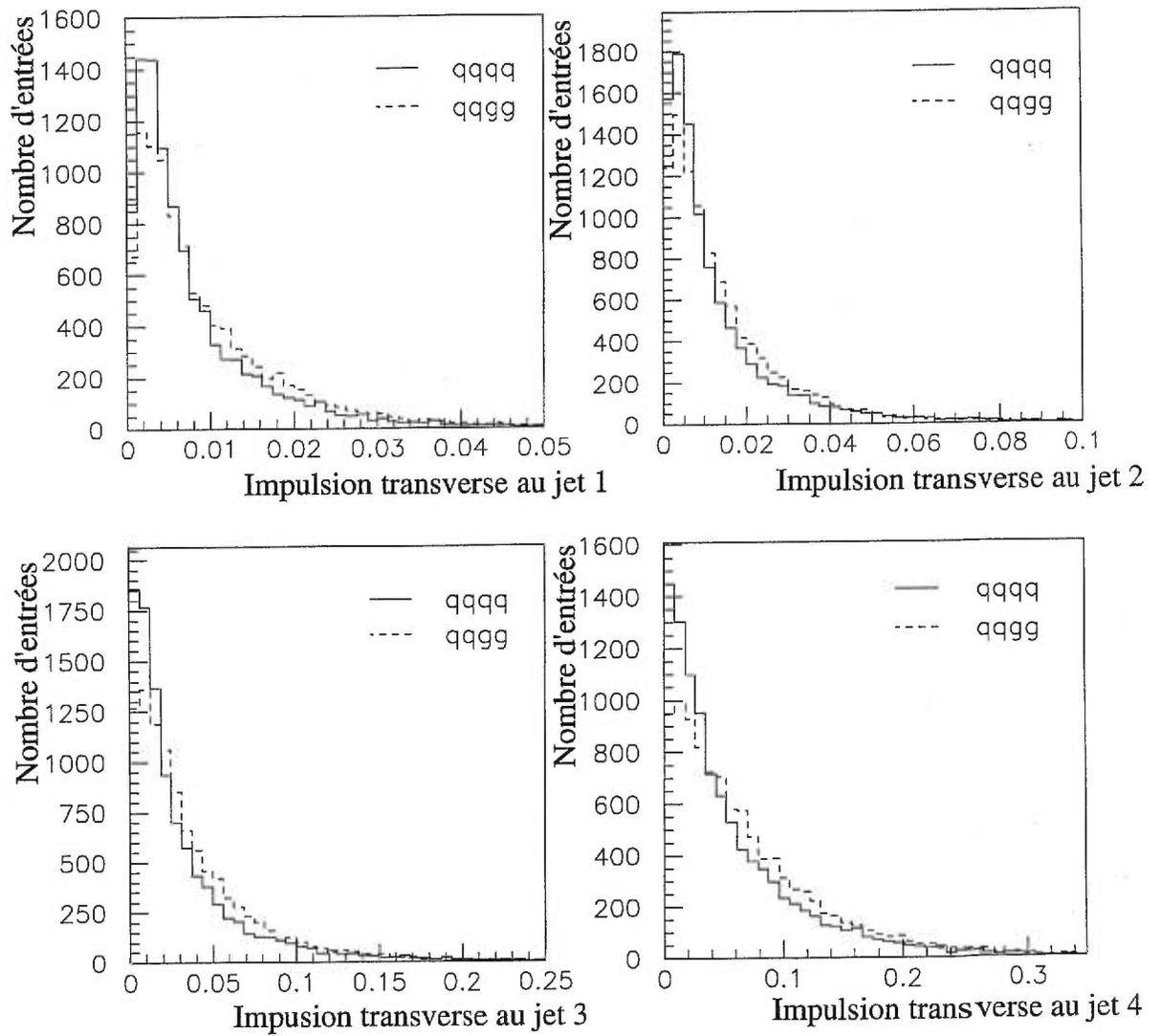


Figure 6-3. Distribution de l'impulsion transverse au carré pour chacun des quatre jets lorsque l'événement est un $q\bar{q}gg$ (traits pointillés) ou lorsque c'est un $q\bar{q}q'\bar{q}'$ (traits pleins).

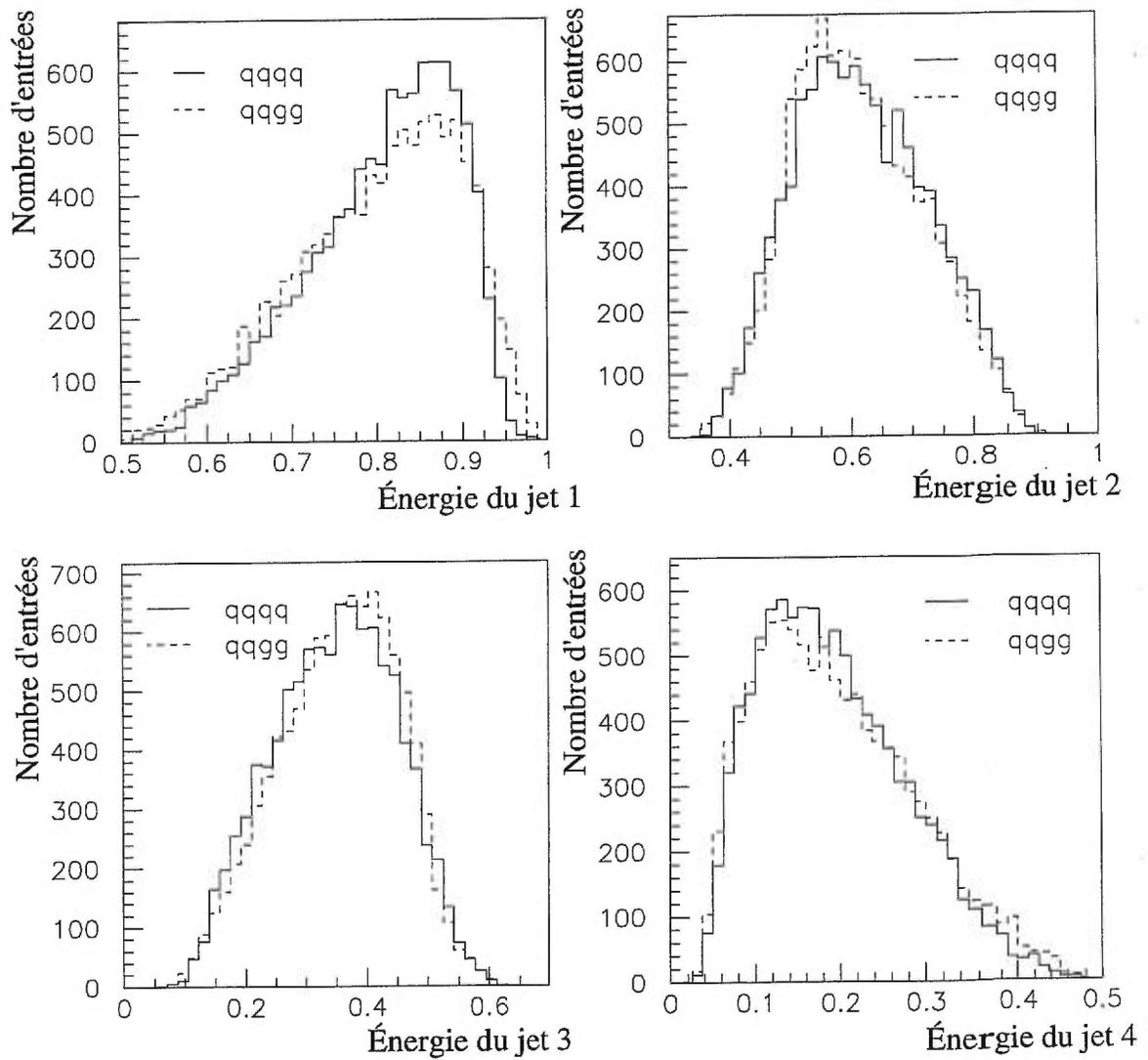


Figure 6-4. Distribution de l'énergie pour chacun des quatre jets lorsque l'événement est un $q\bar{q}gg$ (traits pointillés) ou lorsque c'est un $q\bar{q}q'\bar{q}'$ (traits pleins).

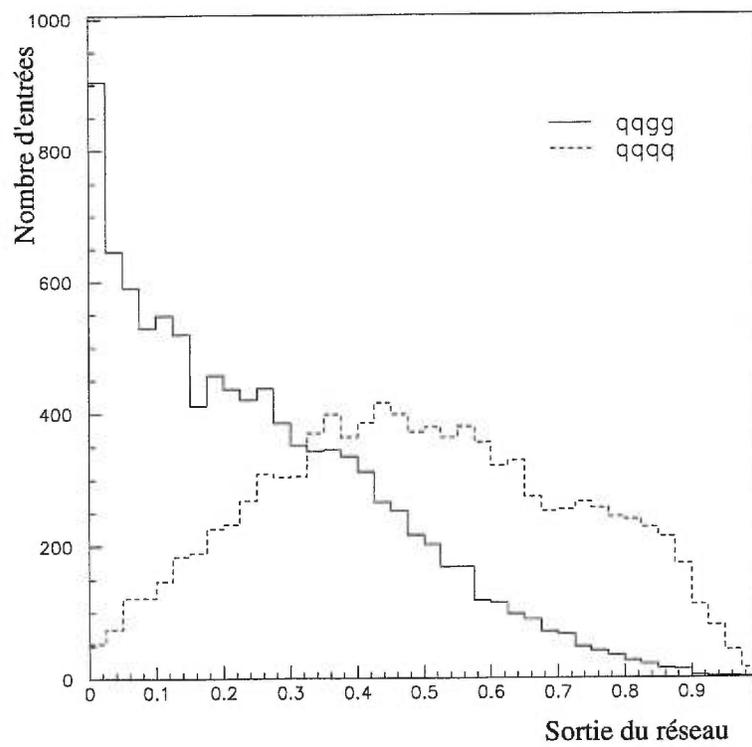


Figure 6-5. Courbes de sortie du réseau de neurones. La courbe en traits pleins correspond à la sortie du réseau lorsqu'on lui présente des événements $q\bar{q}gg$, alors que la courbe en traits pointillés correspond à la réponse du réseau aux événements $qqq\bar{q}'$.

contenus dans l'échantillon analysé par le réseau. Il faut donc pouvoir, à partir de ce nombre N dont la définition est teintée d'arbitraire, évaluer le vrai nombre d'événements à quatre quarks. C'est là que les courbes de sortie nous seront utiles. En effet, nous définirons deux quantités qui seront évaluées à partir de ces courbes et qui nous permettront, pour chaque valeur de coupure possible, de calculer le nombre exacte (avec une incertitude donnée, bien entendu) d'événements à quatre quarks présentés au réseau de neurones.

La première de ces grandeurs est l'efficacité. Elle est définie par [1, 5]:

$$\epsilon_q = \frac{\# \text{ de 4-quarks passant la coupure } \delta}{\# \text{ total de 4-quarks}} = \frac{N^{4q>\delta}}{N^{4q}} \quad (6.1)$$

L'efficacité détermine donc la capacité de la coupure δ à accepter les événements à quatre quarks. La deuxième de ces grandeurs est la pureté, définie par [1, 5]:

$$p_q = \frac{\# \text{ de 4-quarks passant la coupure } \delta}{\# \text{ d'événements passant la coupure}} = \frac{N^{4q>\delta}}{N^{4q>\delta} + N^{2q2g>\delta}} \quad (6.2)$$

La pureté sera donc maximale lorsque le degré de contamination des événements passant la coupure sera minimal.

Pour calculer l'efficacité et la pureté de la classification exécutée par le réseau, il s'agit d'utiliser les courbes de sortie du réseau de neurones présentées à la figure 6-5 et d'y définir une coupure δ permettant d'y déterminer les valeurs de N , N^{4q} , $N^{4q>\delta}$ et $N^{2q2g>\delta}$ servant à définir ϵ_q et p_q . Grâce à ces deux définitions, on pourra ainsi caractériser le traitement des données par le réseau de neurones avec une courbe de l'efficacité en fonction de la pureté. La figure 6-6 présente cette courbe dans le cas précis de notre analyse. Ainsi, si par exemple on passe un échantillon d'événements à quatre jets dans notre réseau de neurones et qu'on compte disons 1000 événements passant une certaine coupure ($N^{>\delta} = 1000$), disons $\delta = 0.55$, sachant que pour cette coupure la pureté est de 81.3%, on saura alors que le véritable nombre d'événements $q\bar{q}q'\bar{q}'$ ayant passés cette coupure sera de 813. De plus comme on sait que l'efficacité pour cette même coupure est de 40.2%, on saura alors que le nombre total d'événements à

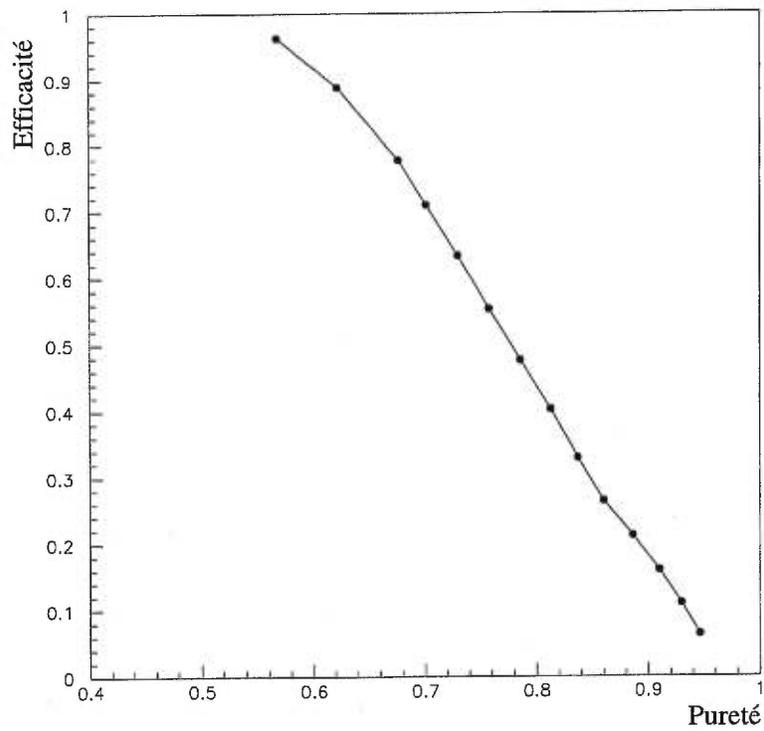


Figure 6-6. Courbe de l'efficacité en fonction de la pureté de la classification des événements à quatre jets par le réseau de neurones.

quatre quarks dans l'échantillon d'événements à quatre jets sera de 2022. On voit donc comment le réseau de neurones, même s'il n'effectue pas une classification parfaite des événements à quatre jets, nous permet quand même d'évaluer le nombre d'événements à quatre quarks contenus dans un échantillon donné: il s'agit simplement de calculer à partir de la réponse du réseau la relation:

$$N_{4q} = N^{>\delta} \cdot \frac{p_q}{\epsilon_q} \quad (6.3)$$

L'erreur sur ces grandeurs ne sera que statistique car elle porte sur un échantillon limité de données Monte Carlo. Ainsi, en supposant des incertitudes de $\Delta A = \sqrt{A}$ pour $N^{>\delta}$, $N^{4q>\delta}$, N^{4q} et $N^{2q2g>\delta}$, on peut donc évaluer l'incertitude de N_{4q} grâce à la formule standard d'évaluation des incertitudes:

$$\frac{\Delta N_{4q}}{N_{4q}} = \sqrt{\left(\frac{\Delta N^{>\delta}}{N^{>\delta}}\right)^2 + \left(\frac{\Delta p_q}{p_q}\right)^2 + \left(\frac{\Delta \epsilon_q}{\epsilon_q}\right)^2} \quad (6.4)$$

avec

$$\frac{\Delta \epsilon_q}{\epsilon_q} = \sqrt{\left(\frac{\Delta N^{4q>\delta}}{N^{4q>\delta}}\right)^2 + \left(\frac{\Delta N^{4q}}{N^{4q}}\right)^2} \quad (6.5)$$

et

$$\frac{\Delta p_q}{p_q} = \sqrt{\left(\frac{\Delta N^{4q>\delta}}{N^{4q>\delta}}\right)^2 + \left(\frac{\Delta B}{B}\right)^2} \quad (6.6)$$

où

$$B = N^{4q>\delta} + N^{2q2g>\delta} \Rightarrow \Delta B = \sqrt{(\Delta N^{4q>\delta})^2 + (\Delta N^{2q2g>\delta})^2}$$

Notre méthode nous permet ainsi d'évaluer $N_{4q} \pm \Delta N_{4q}$ et donc de connaître la valeur de R_{4q} en prenant le rapport de cette valeur de N_{4q} avec celle prédite par la théorie.

Toute cette analyse présuppose que l'on a un échantillon d'événements à quatre jets qui viennent vraiment d'événements à quatre partons. Dans les faits, lorsque l'on reconstruit des événements à quatre jets à partir des données empiriques, on ne reconstruit pas seulement des événements à quatre partons, mais aussi certains événements à deux, trois et cinq partons dont les valeurs

du paramètre de reconstruction (y^{34}) tombent dans l'intervalle des valeurs de ce paramètre pour les événements à quatre partons. En prenant un y_{cut} assez élevé, on peut se débarrasser de la majorité des cinq partons contaminant notre échantillon à quatre jets, mais il reste toujours une importante concentration d'événements à deux et trois partons. Cela fausse à la base notre évaluation de N_{4q} . Il faut donc trouver un moyen de corriger cela. Ce moyen on le possède déjà. En effet, en utilisant exactement le même réseau de neurones que celui que l'on a utilisé précédemment, mais en produisant l'apprentissage à partir des distributions de nos 20 variables topologiques construites pour des événements de type $q\bar{q}gg + q\bar{q}q'\bar{q}'$ et de type $q\bar{q} + q\bar{q}g$ (mal reconstruits en quatre jets), on pourra ainsi séparer les événements à quatre partons de ceux à deux ou trois partons dans notre échantillon d'événements à quatre jets. Comme les courbes de sortie du réseau de neurones (figure 6-7) ont une séparation presque parfaite, on pourra donc utiliser directement la classe des quatre partons produite par le réseau comme entrée du réseau de neurones séparant $q\bar{q}gg$ de $q\bar{q}q'\bar{q}'$. La courbe de pureté vs efficacité donnée à la figure 6-8 nous permettra, quant à elle, d'évaluer l'incertitude sur le nombre d'événements à quatre partons, incertitude qu'il suffira d'ajouter à celle de $N^{>\delta}$. Ainsi pour que l'analyse soit complète, il faut d'abord construire un réseau de neurones pour séparer les vrais événements à quatre jets de ceux qui sont issus de deux ou trois partons mal reconstruits en quatre jets et ensuite passer le produit de cette séparation dans un second réseau de neurones: celui séparant les $q\bar{q}gg$ des $q\bar{q}q'\bar{q}'$. À partir de là, on pourra évaluer N_{4q} adéquatement grâce à la technique présentée plus haut. C'est en cela que consiste notre méthode d'évaluation de R_{4q} , indépendante de celle utilisant des corrélations angulaires entre les jets.

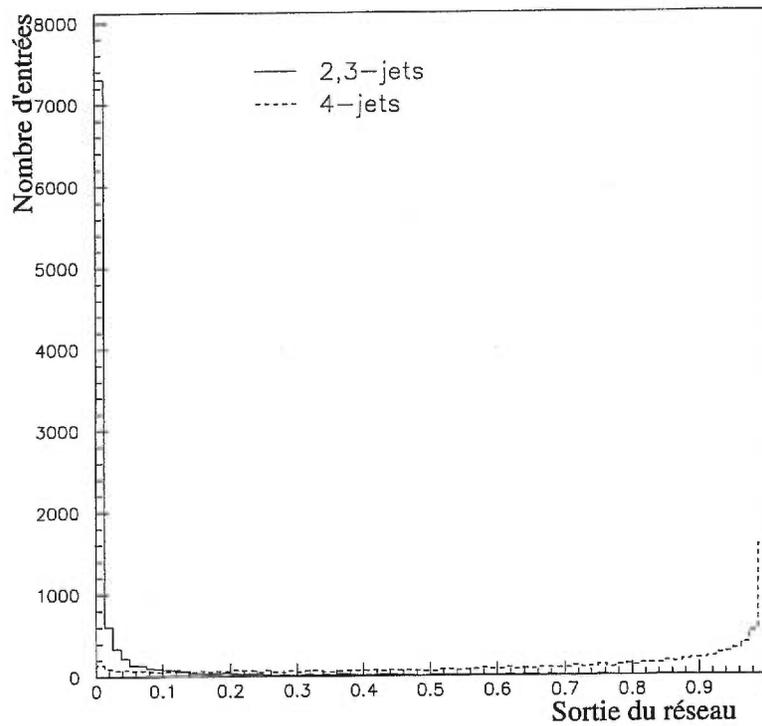


Figure 6-7. Courbes de sortie du réseau de neurones retraçant les événements à quatre partons. La courbe en traits pleins correspond à la sortie du réseau lorsqu'on lui présente des événements $q\bar{q}+q\bar{q}g$, alors que la courbe en traits pointillés correspond à la réponse du réseau aux événements $q\bar{q}gg + q\bar{q}q'\bar{q}'$.

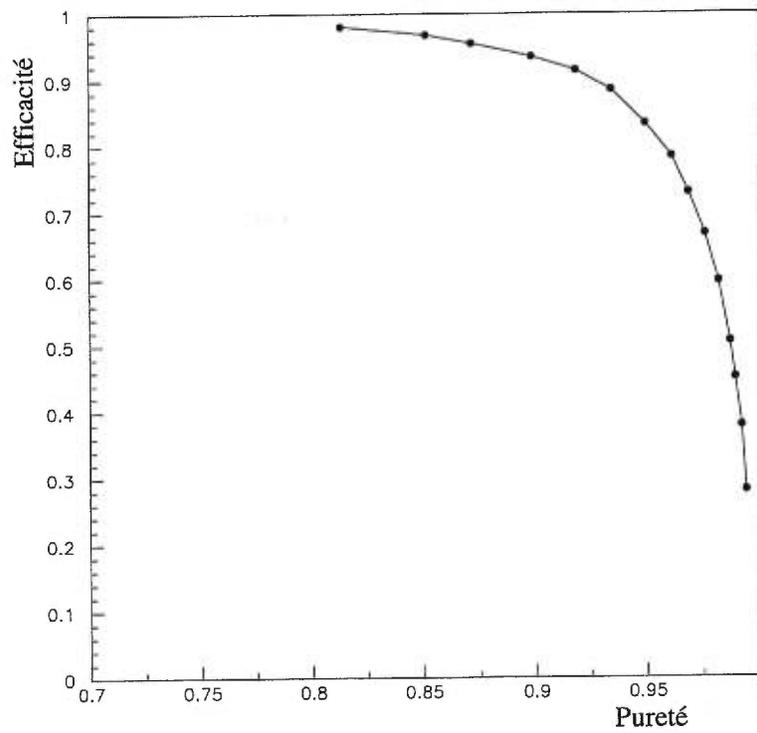


Figure 6-8. Courbe de l'efficacite en fonction de la pureté de la classification des événements à quatre partons par le réseau de neurones.

Conclusion

La réponse que donnera le réseau de neurones à un échantillon d'événements à quatre jets est caractérisée par une courbe de la pureté (p_q) en fonction de l'efficacité (ϵ_q)

Dans notre volonté de construire une méthode de mesure de R_{4q} (donc de N_{4q}) qui soit différente de celles employant des corrélations angulaires entre les jets des événements provenant de la désintégration hadronique du Z^0 au deuxième ordre en α_s , nous avons dans un premier temps (chapitre 4) effectué un choix de variables topologiques discriminant entre les événements de type $q\bar{q}gg$ et ceux du type $q\bar{q}q'\bar{q}'$. Nous avons ensuite configuré un réseau de neurones artificiels de sorte que celui-ci puisse effectuer une classification de ces événements à partir des distributions des variables topologiques choisies, générées par Monte Carlo. L'architecture alors retenue pour ce réseau fut celle d'un perceptron multicouche "feed-forward" à apprentissage par rétro-propagation. En utilisant une courbe de la pureté en fonction de l'efficacité pour caractériser la réponse du réseau de neurones à un échantillon connu d'événements à quatre jets, on a finalement pu extraire le nombre d'événements à quatre quarks contenus dans un échantillon inconnu d'événements à quatre jets, après s'être préalablement assuré que chacun de ces quatre jets provenait d'un parton parent différent. La procédure ainsi développée constitue donc une méthode effective, "théoriquement" complète, permettant de mesurer la valeur de N_{4q} .

Cette méthode n'est cependant pas encore "empiriquement" complète. En effet, les distributions employées pour l'entraînement du réseau de neurones ne considèrent pas les effets du détecteur OPAL. Le réseau présenté n'est donc pas

entraîné pour classifier des événements provenant de données expérimentales. Il faudra, dans une étude ultérieure, produire les distributions des variables sélectionnées non seulement à l'aide de JETSET, pour simuler l'hadronisation, mais également avec GOPAL, pour simuler la réaction du détecteur au calcul des valeurs de ces variables. Il faudra ensuite reconfigurer le réseau de neurones, en tenant compte de ces nouvelles corrections. Bien sûr la méthode d'évaluation de N_{4q} restera complètement la même. Par contre la courbe de la pureté en fonction de l'efficacité sera probablement différente, nous empêchant ainsi d'utiliser ce qui a été présenté ici pour immédiatement évaluer N_{4q} sur les données OPAL. C'est en ce sens que l'on peut dire que la méthode n'est que "théoriquement" complète. Une fois que ce travail aura été fait, il ne restera finalement plus qu'à calculer R_{4q} à partir de la valeur de N_{4q} obtenue des événements OPAL et à voir si cette valeur est compatible avec 1. On pourra ainsi évaluer la possibilité pour de la nouvelle physique ou juger de la fiabilité des méthodes de discrimination par corrélations angulaires.

Même si une mesure ultérieure de R_{4q} s'avérait compatible avec 1, excluant ainsi la possibilité d'existence d'un gluino dans cette gamme de masse (≈ 20 GeV), la méthode développée continuera d'être utile aux diverses analyses en physique des hautes énergies. Elle pourra par exemple servir à raffiner les différents modèles de fragmentation. Elle constitue donc un outil et une façon de faire générale, qu'on a décidé ici d'appliquer à un problème particulier.

Bibliographie

- [1] R. Barlow; Measurement of $\Gamma(Z - b\bar{b})/\Gamma(Z - \text{hadrons})$ using the boosted sphericity product, OPAL Physics Note (PN087), 1992.
- [2] R. Barlow; Jets in high-energy interactions, *Rep. Prog. Phys.*, 56 (1993) 1067.
- [3] B. Diu C. Cohen-Tannoudji and F. Laloe; *Mécanique Quantique*, Harman, 1973.
- [4] T. Rognvaldsson et L. Lonnblad C. Peterson; *JETNET 3.0 -A Versatile Artificial Neural Network Package*, décembre 1993. LU TP 93-29.
- [5] N. Capdevielle; Études systématiques des réseaux de neurones pour l'identification des jets issus de quarks lourds dans les interactions e^+e^- . Mémoire de maîtrise, Université de Montréal, mars 1996.
- [6] D. Davignon; Mesure de la fraction apparente des configurations à 4 quarks dans les désintégrations hadroniques du Z^0 en 4 jets. Mémoire de maîtrise, Université de Montréal, décembre 1999.
- [7] B. Denby; Neural network tutorial for high energy physicists, In *Fermilab-conf-90/94*, mai 1990.
- [8] P. Depommier; Interactions fondamentales (PHY6668), 1999. www.lps.umontreal.ca/~pom/phy6668.html.
- [9] M. Peskin et D. Schroeder; *An Introduction to Quantum Field Theory*, Addison-Wesley, 1995.
- [10] F. Mandl et G. Shaw; *Quantum Field Theory*, John Wiley and Sons, 1984.

- [11] M. Bengtsson et P. M. Zerwas; Four-jet events in e^+e^- annihilation: testing the three-gluon vertex, *Physics Letters B*, 208(2) (1988) 306.
- [12] R. D. Field et R. P. Feynman; A parametrization of the properties of quark jets, *Nuclear Physics B*, 136 (1978) 1.
- [13] G. C. Fox et S. Wolfram; Event shapes in e^+e^- annihilation, *Nuclear Physics B*, 149 (1979) 413.
- [14] D. E. Groom *et al.*; The review of particle physics, *The European Physical Journal C*, C15 (2000) 1.
- [15] S. Kluth *et al.*; Qcd studies with e^+e^- annihilation data at 130 and 136 GeV, OPAL Physics Note (PN203), 1995.
- [16] C. Théoret H. Jeremie, E. Levebvre; Remarks on the resolvability of high-energy particle jets, *Nuclear Instrument and Methods in Physics Research A*, 405 (1988) 126.
- [17] E. Lefebvre et P. Leblanc H. Jeremie; Study of four-jet event from Z^0 decays at LEP using Bengtsson-Zerwas correlation and jet masses, OPAL Physics Note (PN321), 1997.
- [18] J. W. Gary I. M. Dremin; Hadron multiplicities, hep-ph/0004215.
- [19] F. E. Low et So-Young Pi J. F. Donoghue; Tensor analysis of hadronic jets in quantum chromodynamics, *Physical Review D*, 20(11) (1979) 2759.
- [20] G. Kane; *Modern Elementary Particle Physics*, Addison-Wesley, 1993.
- [21] N. K. Kasabov; *Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering*, The MIT Press, 1996.
- [22] T. Khanna; *Foundations of Neural Networks*, Addison-Wesley, 1990.

- [23] E. Lefebvre; *Mesure de la fraction d'événements à quatre quarks dans les désintégrations multihadroniques du Z^0 au LEP*. Thèse de doctorat, Université de Montréal, avril 1998.
- [24] S. L. Lloyd; *The OPAL Primer*, 1994.
- [25] Collaboration OPAL; *The OPAL Detector at LEP*, 1990. CERN-PPE/90-114.
- [26] Collaboration OPAL; *Opal detector performance paper, Draft text*, décembre 1997.
- [27] D. H. Perkins; *Introduction to High Energy Physics*, Addison-Wesley, 1987.
- [28] A. Pich; *Quantum chromodynamics*, Technical report, European School of High-Energy Physics, 1995, CERN95-04.
- [29] H. Przysiezniak; *Étude de la production des mésons D^* sur le pic de résonance du Z^0 observés au LEP avec le détecteur OPAL*. Thèse de doctorat, Université de Montréal, avril 1994.
- [30] R. K. Ellis, D. A. Ross et A. E. Terrano; *The perturbative calculation of jet structure in e^+e^- annihilation*, *Nuclear Physics B*, 178 (1981) 421.
- [31] T. Sjostrand; *PYTHIA 6.0 and JETSET 7.4 Physics and Manual*, août 1995. LU TP 95-20.
- [32] F. Fogelman Soulié; *Réseaux de neurones en physique des hautes énergies*, *Bulletin de la S.F.P.*, décembre 1992.
- [33] R. K. Wangsness; *Electromagnetic Fields*, John Wiley and Son, 1986.
- [34] V. Zacek; *Introduction à la physique subatomique (PHY2601)*, 2000.
www.lps.umontreal.ca/~zacekv/phy2601/contenu.html.