

UNIVERSITÉ CLAUDE BERNARD – LYON 1
INSTITUT DE PHYSIQUE DES DEUX INFINIS DE LYON

Mémoire de thèse

présenté par

AURÉLIEN CARLE

Recherche de brisure de la symétrie de
Lorentz dans le secteur du quark top
avec l'expérience CMS au LHC

le

devant le jury composé de

Mme.	Sabine	Crépé-Renaudin	Rapportrice
M.	Christophe	Delaere	Rapporteur
Mme.	Corinne	Augier	Présidente du jury
M.	Christophe	Le Poncin-Lafitte	Examinateur
M.	Stéphane	Perries	Directeur de thèse
M.	Nicolas	Chanon	co-Directeur de thèse



TABLE DES MATIÈRES

1 Le dispositif expérimental	1
1.1 Large Hadron Collider.	2
1.1.1 Accélération de protons	2
1.1.2 Les expériences	3
1.1.3 Luminosité.	4
1.1.4 Empilement	5
1.2 L'expérience CMS (<i>Compact Muon Solenoid</i>).	6
1.2.1 Système de coordonnées	7
1.2.2 Aimant solénoïdal	8
1.2.3 Le trajectographe : pixels et pistes de silicium	9
1.2.4 Calorimètre électromagnétique.	9
1.2.5 Calorimètre hadronique.	14
1.2.6 Chambres à muons	15
1.2.7 Système de déclenchement	17
1.3 Mesure de la luminosité	18
2 Simulation et reconstruction des événements	21
2.1 Génération des événements	21
2.1.1 L'événement dur.	21
2.1.2 Densité de probabilité partonique (PDF)	22
2.1.3 De l'événement dur aux particules dans le détecteur	24
2.1.4 Simulation du détecteur.	25
2.2 Reconstruction des événements	26
2.2.1 L'algorithme <i>Particle-Flow</i> (PF)	26
2.2.2 Isolation des leptons et des photons	32
2.2.3 Reconstruction des jets	33
2.2.4 Étiquetage des jets ISSUS DE QUARKS <i>b</i>	36
2.2.5 Énergie transverse manquante	37
2.3 Conclusion	38

1

LE DISPOSITIF EXPÉRIMENTAL

1.1	Large Hadron Collider	2
1.1.1	Accélération de protons	2
1.1.2	Les expériences	3
1.1.3	Luminosité	4
1.1.4	Empilement.	5
1.2	L'expérience CMS (<i>Compact Muon Solenoid</i>)	6
1.2.1	Système de coordonnées	7
1.2.2	Aimant solénoïdal	8
1.2.3	Le trajectographe : pixels et pistes de silicium	9
1.2.4	Calorimètre électromagnétique	9
1.2.5	Calorimètre hadronique	14
1.2.6	Chambres à muons.	15
1.2.7	Système de déclenchement	17
1.3	Mesure de la luminosité	18

La particule élémentaire la plus massive du Modèle Standard est le quark top, avec une masse d'environ 173 GeV. Inexistant à l'état stable dans la nature, il est nécessaire de le produire pour l'étudier. Pour cela, il faut être en mesure de générer des particules beaucoup plus énergétiques. La stratégie consiste à utiliser des accélérateurs qui permettent de collisionner des particules à grande énergie et ainsi de produire de nouvelles particules issues de la collision.

On distingue deux grandes classes d'accélérateurs de particules :

LES ACCÉLÉRATEURS CIRCULAIRES Dans ces dispositifs les particules accélèrent le long d'un anneau. En permettant de nombreuses rotations, les particules peuvent atteindre de très hautes énergies. Cependant, lors d'une rotation, les particules chargées sont soumises à l'effet synchrotron. Il s'agit d'un rayonnement dont la valeur est inversement proportionnel à rm^4 avec r le rayon de courbure et m la masse de la particule accélérée. Autrement dit, plus le rayon de courbure et la masse sont élevées, moins le rayonnement synchrotron amputera de l'énergie aux particules à collisionner. Cela implique l'utilisation de particules lourdes comme des protons si l'on souhaite minimiser les pertes d'énergie. Les protons n'étant pas élémentaires, leurs structures internes va générer lors d'une collision plusieurs phénomènes parasites impactant la précision des mesures.

LES ACCÉLÉRATEURS LINÉAIRES Comme leur nom l'indique, ils accélèrent les particules en ligne droite. Due à la trajectoire finie des particules accélérées, leur énergie à la collision sont limitées. Cependant, l'absence de rayonnement synchrotron permet d'utiliser des particules élémentaires de basse masse. En utilisant des particules comme l'électron, on assure des signatures extrêmement claires dans le détecteur, ce qui permet d'effectuer des mesures de précisions.

Le LHC (*Large Hadron Collider*) [1] est un collisionneur circulaire p-p. Il est donc un excellent dispositif pour pouvoir étudier des phénomènes physiques en relation avec le quark top. Plusieurs détecteurs lui sont associés. Cette thèse utilise les données mesurées au LHC par le détecteur CMS *Compact Muon Solenoid* [2]. Ce chapitre a pour but de présenter en détails l'accélérateur ainsi que le détecteur.

1.1 LARGE HADRON COLLIDER

Le grand collisionneur de hadron est le plus puissant accélérateur de particules du monde avec 27 km de circonférence. Il est le successeur du LEP (*Large Electron Positron collider*) et il se situe dans le même tunnel à plus de 100 m sous terre. Géographiquement, le LHC est situé à la frontière Franco-Suisse au CERN.

Composé de deux anneaux, le LHC permet d'accélérer des protons à des énergies de 6.5 TeV pour le Run II et 7 TeV pour le Run-III, grâce à 16 cavités radio-fréquence. La trajectoire circulaire des paquets de protons est maintenue par les 9 593 électroaimants supraconducteurs (dont 1 232 aimants dipolaires). Ils sont refroidis à une température de 1.8 K grâce à de l'hélium superfluide et génèrent ainsi un champ magnétique nominal de 8.33 T.

1.1.1 ACCÉLÉRATION DE PROTONS

Les protons sont graduellement accélérés avant d'être injectés dans le LHC. Les différentes zones et dispositifs de la chaîne d'accélération du CERN sont illustrés dans la figure FIG.1.1. Pour une collision p-p, la chaîne est la suivante :

LE DUOPLASMATRON PROTON-ION SOURCE Les protons sont obtenus par ionisation de dihydrogène pur, directement issu d'une bouteille. Ils sont ensuite ionisés sous forme de plasma, séparés des électrons et enfin paquetés.

LE LINAC 2 (1978) Arrivés dans *Linac 2*, les protons sont accélérés jusqu'à 50 MeV (31.4 % de c).

LE PROTON SYNCHROTON BOOSTER (PSB – 1972) L'accélérateur circulaire PSB donne aux protons une énergie de 1.4 GeV (91.6 % de c).

LE PROTON SYNCHROTON (PS – 1959) Dans cet accélérateur, les protons atteignent une énergie de 25 GeV (99.93 % de c).

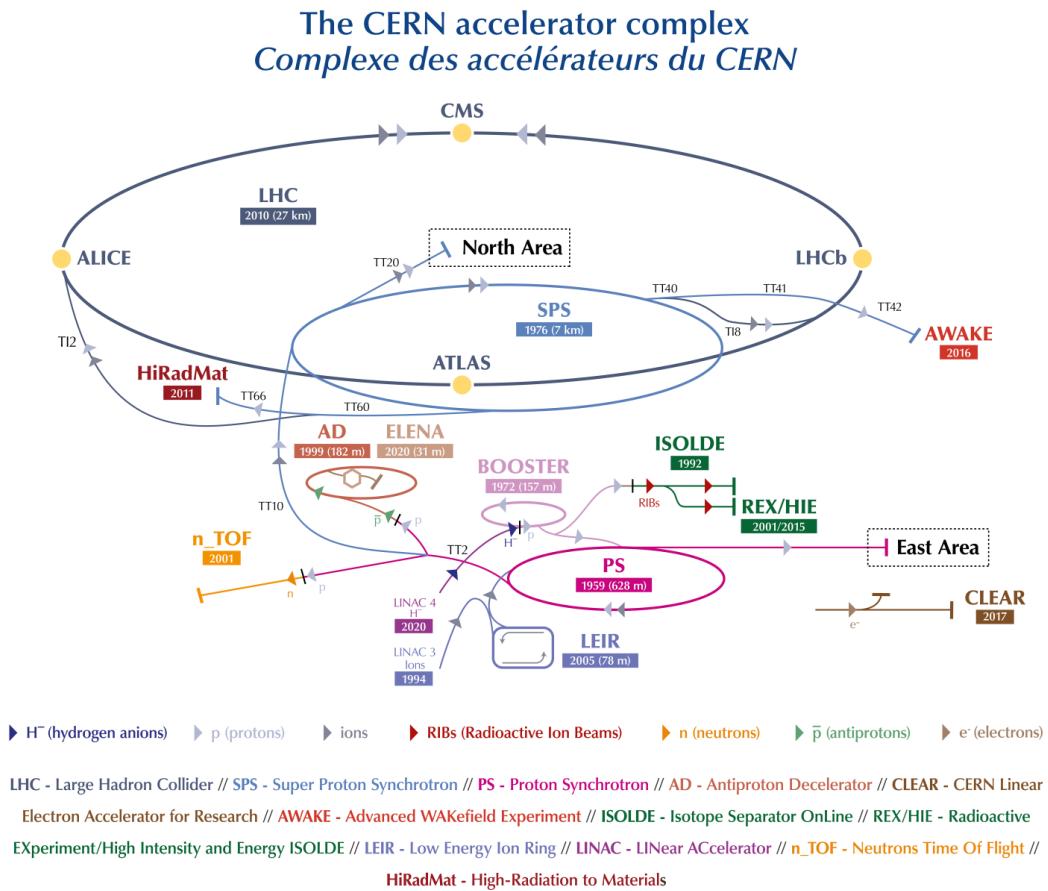


FIG. 1.1: Le complexe d'accélérateurs du CERN.

LE SUPER PROTON SYNCHROTON (SPS – 1976) Finalement, le faisceau subit une dernière étape d'accélération dans le SPS, atteignant 450 GeV (99.9998 %). Les paquets de protons sont ensuite injectés dans le LHC. Ils seront accélérés pour atteindre l'énergie nominale de 6.5 TeV (99.999 999 0 %) par faisceau pour le Run-II.

1.1.2 LES EXPÉRIENCES

Le LHC étant un collisionneur circulaire, il est possible de faire se croiser les faisceaux de protons à plusieurs endroits. Plusieurs dispositifs expérimentaux sont installés, aux croisements. Le LHC est divisé en huit sections (ou points) et compte quatre croisements avec quatre expériences majeures : ALICE [3], ATLAS [4], CMS [5] et LHCb [6].

A LARGE ION COLLIDER EXPERIMENT (ALICE) Au point 2, cette expérience est principalement dédiée à l'étude du déconfinement de la matière nucléaire, le plasma de

quark et gluon. Les données qu'elle utilise sont celles issues des collisions d'ions lourds (Pb-Pb ou p-Pb), mais les collisions p-p sont aussi utilisées afin de calibrer le détecteur.

A TOROIDAL LHC APPARATUS (ATLAS) / COMPACT MUON SOLENOID (CMS) Respectivement au point 1 et 5, ce sont les deux expériences généralistes du LHC. En effet, le programme de physique de ATLAS et CMS est très vaste, et couvre la recherche du boson de Higgs et de nouvelle physique, les mesures de précisions du Modèle Standard, ainsi que la recherche de candidats matière noire. Souvent mises en concurrence, ces expériences sont pourtant complémentaires. Ainsi, on a pu voir le 4 juillet 2012 ces deux expériences annoncer conjointement la découverte d'une particule compatible avec le boson de Higgs [7, 8], chacune confirmant ainsi les résultats de l'autre.

LARGE HADRON COLLIDER BEAUTY (LHCb) Au point 8, c'est la dernière expérience majeure du LHC, principalement dédiée aux mesures de précision du Modèle Standard ainsi qu'à l'étude de la violation de la symétrie CP, grâce à l'étude poussée du quark b . La collaboration LHCb a d'ailleurs annoncé récemment avoir observé pour la première fois la violation de symétrie CP dans le système B_s [9], telle que prévue par le Modèle Standard. Cette récente découverte permet de contraindre encore plus fortement certains modèles de nouvelle physique.

Il existe 3 autres expériences au LHC, installées à proximité des points de croisement des faisceaux : LHCf [10], MoEDAL [11] et TOTEM [12].

LARGE HADRON COLLIDER FORWARD (LHCf) Situé à environ 140 m de part et d'autre d'ATLAS, ce détecteur étudie les particules créées à très petits angles, principalement afin de simuler la production de rayons cosmiques de très haute intensité en laboratoire.

MONOPOLE AND EXOTICS DETECTOR AT THE LHC (MOEDAL) Située juste en aval de LHCb, MOeDAL traque les monopôles magnétiques, grâce à un détecteur spécialement conçu pour ce rôle.

TOTAL ELASTIC AND DIFFRACTIVE CROSS SECTION MEASUREMENT (TOTEM) Destinée à la mesure précise de la luminosité du LHC, cette expérience, située dans la même caverne que CMS, étudie les particules créées à très petits angles. Elle peut ainsi mesurer la section efficace élastique des collisions p-p.

1.1.3 LUMINOSITÉ

La luminosité instantanée est une variable clé dans un accélérateur de particules. Exprimée en unités $\text{cm}^{-2} \text{s}^{-1}$ (ou plus régulièrement fb^{-1}), elle représente le taux de collisions au point d'interaction. Elle est exprimée au LHC en fonction de diverses variables caractéristiques tel que la forme des paquets de protons, leur énergie,... etc.

La luminosité instantanée est la luminosité évaluée pendant une période infinitésimale donnée. Elle vaut :

$$\mathcal{L}_{\text{inst}} = \frac{\gamma f n_p N_p^2}{4\pi \epsilon_n \beta^*} = \frac{f n_p N_p^2}{4\pi \sigma_x \sigma_y} \quad (1.1)$$

où γ est le boost de Lorentz, f est la fréquence de rotation des paquets, n_p est le nombre de paquets de protons circulant dans le LHC, N_p est le nombre de protons par paquet, ϵ_n est l'émittance transversale (mesure du parallélisme du faisceau). Pour favoriser les collisions, les deux faisceaux sont resserrés à l'approche du point de collision. β^* mesure la distance entre le point d'interaction et le point où le faisceau double de largeur. Les $\sigma_{x,y}$ sont les tailles transversales du faisceau au point d'interaction. Les valeurs des différents paramètres du faisceau sont indiquées dans le tableau 1.1 pour diverses scenarii.

Caractéristiques	LHC Run II
Énergie par proton (TeV)	6.5
\mathcal{L} (fb $^{-1}$)	163.5 (FIG.1.2a)
f (Hz)	11 246
Intervalle entre croisements de paquet (ns)	25
Nombre de paquets	2808
Protons par paquet (10^{11})	1.2
β^* (m)	0.55

Tab. 1.1: Les différents paramètres du faisceau p-p du LHC [13].

A partir de la luminosité instantanée, on peut déduire le nombre d'événements observés N , d'un processus étudié i , par l'équation :

$$N_i = \int \mathcal{L}_{\text{inst}} \sigma_i dt \quad (1.2)$$

avec σ_i la section efficace du processus i .

1.1.4 EMPILEMENT

Les protons se déplaçant par paquets, plusieurs interactions p-p peuvent se produire lors d'un même croisement de faisceaux. La collision principale, centre d'intérêt des analyses est nommée « événement dur ». Les autres collisions (élastiques ou inélastiques) parasitent l'événement dur. Cet effet est communément appelé empilement (*pile-up* (PU) en anglais). Pour le Run-II, le nombre moyen d'interactions par croisement $\langle \mu \rangle$ est d'environ 34.

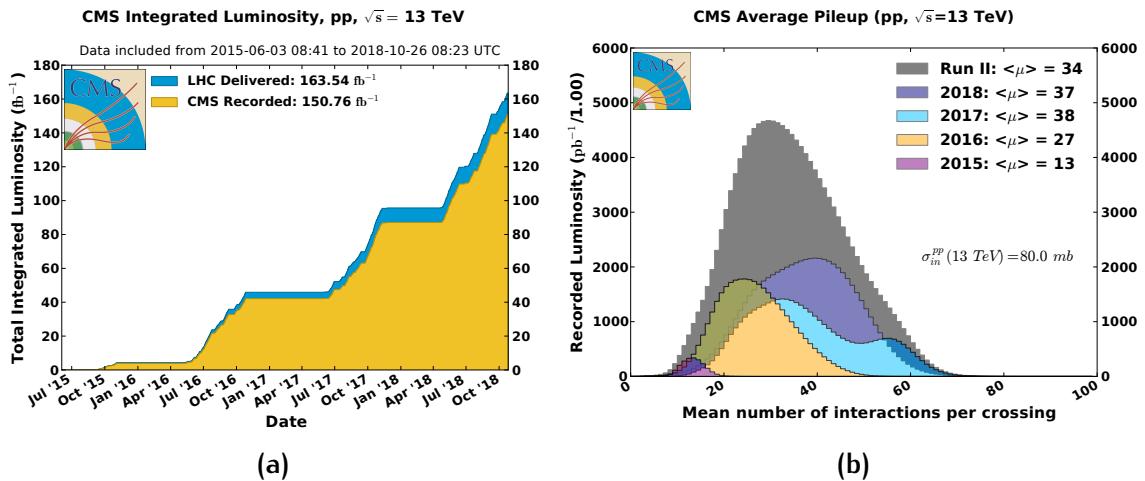


FIG. 1.2: (a) Luminosité intégrée collectée par l’expérience CMS au cours du Run II du LHC.
(b) Nombre moyen d’interactions par croisement de faisceaux observées pendant la prise de données de différentes années du Run II.

EMPILEMENT SYNCHRONE Il est produit en même temps que l’évènement dur. A chaque collision correspond un vertex, qui est le lieu de rencontre de deux protons. Le vertex principal (dur) est celui pour lequel la somme quadratique d’impulsion transverse (p_T) des particules chargées reconstruites dans le détecteur est la plus grande. Les autres sont considérées comme vertex d’empilement. Les particules provenant de vertex d’empilement laissent dans le détecteur des bruits parasites.

EMPILEMENT ASYNCHRONE Il est principalement causé par le temps de réponse des détecteurs. Cet empilement est du au recouvrement entre des évènements en cours de traitement, les suivants et les précédents.

L’empilement se traduit directement par une augmentation du nombre de vertex d’interactions reconstruits par le détecteur. On présente FIG.1.2b le nombre d’interactions moyen par croisement de faisceau, pour différentes années. En absence de d’empilement, une seule interaction serait enregistrée.

1.2 L’EXPÉRIENCE CMS (*COMPACT MUON SOLENOID*)

Le *compact muon solenoid* est une expérience du LHC installée sur le territoire français près de Cessy dans l’Ain. Les motivations étant diverses, de l’étude de la brisure de symétrie électrofaible à la recherche de nouvelle physique (supersymétrie, matière noire, ...) on qualifie cette expérience de généraliste.

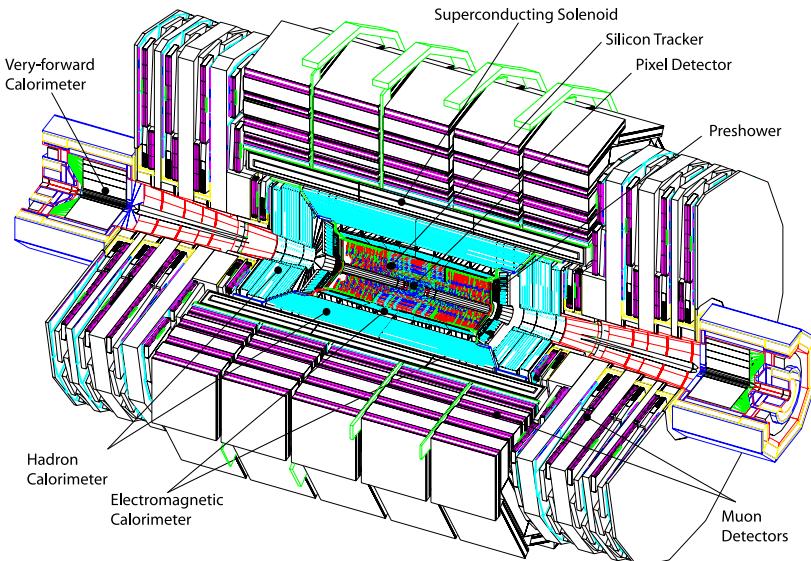


FIG. 1.3: Vue en perspective du détecteur CMS et de ses sous-détecteurs [13].

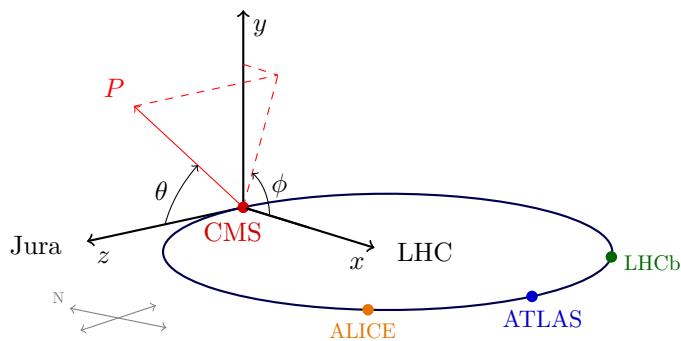


FIG. 1.4: Système de coordonnées du détecteur CMS par rapport à l'anneau du LHC [13].

Une vue éclatée du détecteur CMS est représentée sur la figure FIG.1.3. Ce détecteur est composé de plusieurs sous-parties possédant chacune une spécialité de détection pour différents éléments (électrons, muons, jets, ...).

1.2.1 SYSTÈME DE COORDONNÉES

CMS (figure FIG.1.3) est situé au point 5 du LHC. Il s'agit d'un détecteur compact, cylindrique mesurant 28.7 m de long avec un rayon de 7.5 m, pour un poids de 14 000 t. Il est composé de plusieurs sous-détecteurs, organisés en couches successives de détections. Pour décrire plus en détails ces sous-détecteurs, il est nécessaire de définir un système de référence, où l'origine est donnée par le centre du détecteur. L'axe x pointe vers le centre de l'anneau du LHC, et l'axe z est tangent à la direction du faisceau. L'axe y , perpendiculaire aux deux autres axes, pointe vers le haut (voir figure FIG.1.4).

L'angle azimutal $\phi \in [-\pi, \pi]$ est mesuré dans le plan (x, y) , à partir de l'axe x . L'angle θ , lui, est défini à partir de l'axe z dans le plan transverse (y, z) . Il est plus commode d'utiliser la pseudo-rapidité η plutôt que θ , puisque la densité de production de particules est constante suivant η , mais pas suivant θ . De plus, pour des particules ultra-relativiste la pseudo-rapidité tend vers la rapidité qui est une quantité additive. On définit la pseudo-rapidité par

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right] = \frac{1}{2} \ln \left(\frac{|\vec{p}| + p_z}{|\vec{p}| - p_z} \right) \quad (1.3)$$

où p_z est la projection de l'impulsion le long de l'axe du faisceau.

Du fait de la structure interne des protons, lors d'une collision, l'impulsion totale selon l'axe du faisceau est inconnue. De plus, le détecteur n'est pas hermétique et certaines particules (dont les neutrinos) échappent à la détection ce qui implique que l'on ne peut pas déduire par reconstruction, non plus, l'impulsion p_z . Pour contrer ce problème, la stratégie est d'utiliser le fait que la somme vectorielle des impulsions sera nulle dans le plan de transverse. Ainsi on introduit des variables relatives à ce plan transverse (x, y) , l'impulsion transverse p_T et l'énergie transverse E_T . On a ainsi

$$p_T = \sqrt{p_x^2 + p_y^2} = \frac{|\vec{p}|}{\cosh \eta}$$

$$E_T = E \sin \theta = \frac{E}{\cosh \eta}$$

Également, on peut définir la distance angulaire ΔR entre deux particules i et j tel que :

$$\Delta R_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2} = \sqrt{\Delta \eta^2(i, j) + \Delta \phi^2(i, j)} \quad (1.4)$$

Cette dernière quantité est très utile pour l'estimation de l'isolation d'une particule.

1.2.2 AIMANT SOLÉNOÏDAL

Le solénoïde supraconducteur est un élément central de CMS. Il fournit un champ magnétique de 3.8 T. De forme cylindrique de rayon 6 m, long de 12.5 m, il est parcouru par un courant électrique de 19.14 kA. Afin de maintenir son état supraconducteur, l'aimant est refroidi par un système de cryostat fonctionnant à l'hélium superfluide. Il est ainsi constamment maintenu à une température de 1.9 K.

On a déjà vu précédemment l'intérêt d'un puissant champ magnétique. En effet, la trajectoire d'une particule chargée se courbe en présence d'un champ magnétique, avec un rayon de courbure proportionnel à son impulsion :

$$r = \frac{|\vec{p}|}{qB}$$

avec r le rayon de courbure de la trajectoire, p l'impulsion, q la charge de la particule et B l'intensité du champ magnétique. Plus r est petit, plus mesurer la courbure de

la trajectoire d'une particule est aisée, puisque la courbure de la trajectoire est plus prononcée. Afin d'assurer de bonnes performances sur l'identification des particules, il est nécessaire d'utiliser un champ magnétique puissant pour minimiser r .

1.2.3 LE TRAJECTOGRAPHE : PIXELS ET PISTES DE SILICIUM

Le trajectographe est l'un des composants les plus importants de CMS. En plus de la reconstruction des vertex d'interactions, il est chargé de reconstruire la trajectoire courbée des particules chargées, et ainsi d'en déduire leur impulsion. Il s'agit d'un cylindre d'une longueur de 5.5 m et de 1.1 m de rayon. Le trajectographe est conçu pour couvrir une région angulaire entre $0 < |\eta| < 2.5$, est composé de deux sous-parties.

Le détecteur à pixels est la partie la plus interne du trajectographe, il est composé de quatre couches de détection [14] situées au plus proche du point de collision (à $r = 2.9$ cm, $r = 6.8$ cm, $r = 10.9$ cm et $r = 16$ cm), et de trois disques (bouchons) disposés à chaque extrémité. On dénombre 66×10^6 pixels, mesurant chacun $100 \times 150 \mu\text{m}^2$, pour une superficie totale de détection d'environ 1 m^2 , compose ce détecteur à silicium. La résolution en p_T , obtenue à partir des algorithmes de reconstructions des traces du trajectographe, est présentée dans la figure Fig.1.5a.

Le détecteur à micropistes, englobant le détecteur à pixels, il est composé de 10 couches en silicium. Cela représente environ 10 millions de micropistes. Quatre sous-ensembles sont dénombrables :

LE TIB (*Tracker Inner Barrel*) Composée de quatre couches de micropistes, c'est la partie interne du tonneau du trajectographe,

LE TID (*Tracker Inner enDcaps*) Avec trois couches de détection, cette partie constitue les bouchons internes, et complète le TIB.

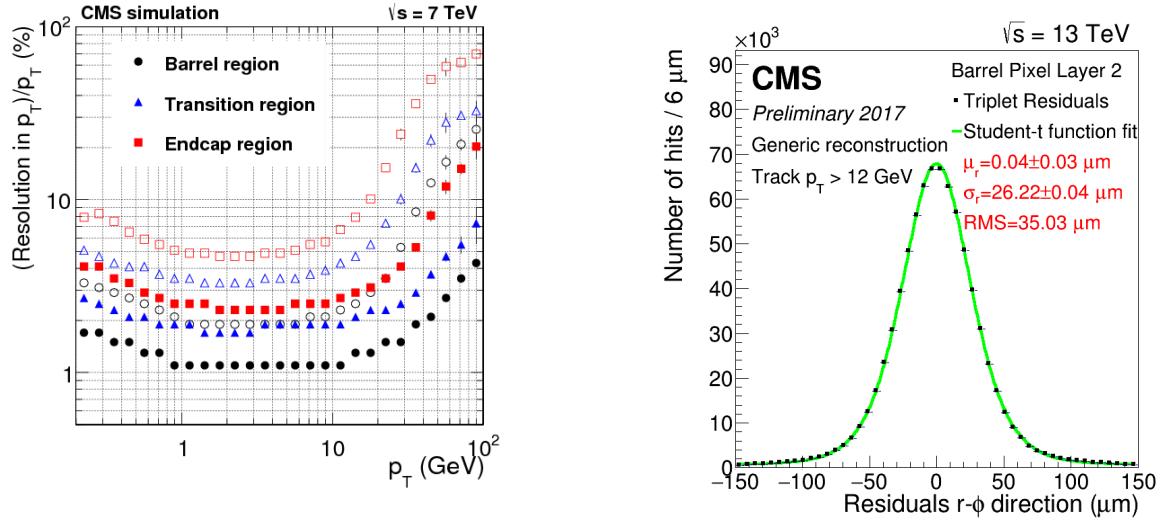
LE TOB (*Tracker Outer Barrel*) Cette partie externe est constituée de six couches de micropistes. Elle entoure le TIB.

LE TEC (*Tracker EndCaps*) Composée de neuf couches, elle forme les deux bouchons externes et complète le TOB.

L'agencement du trajectographe, avec les détecteurs à pixels et à micropistes, est présenté en détail figure 1.6.

1.2.4 CALORIMÈTRE ÉLECTROMAGNÉTIQUE

Enrobant le trajectographe, le calorimètre électromagnétique (ECAL) capte et mesure l'énergie des électrons et des photons. Il s'agit d'un détecteur à scintillation formé



(a) Résolution en p_T des traces reconstruites [15].

(b) Résolution du détecteur à pixels pour la seconde couche [13].

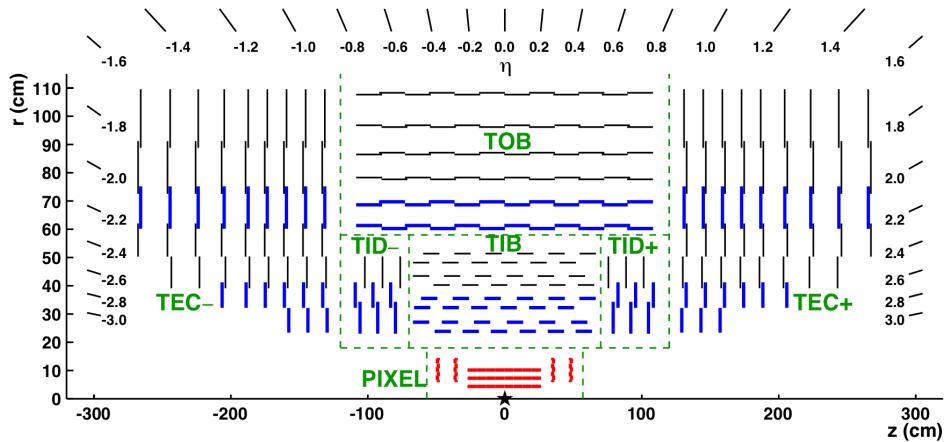


FIG. 1.6: Vue détaillée du trajectographe de l'expérience CMS [13].

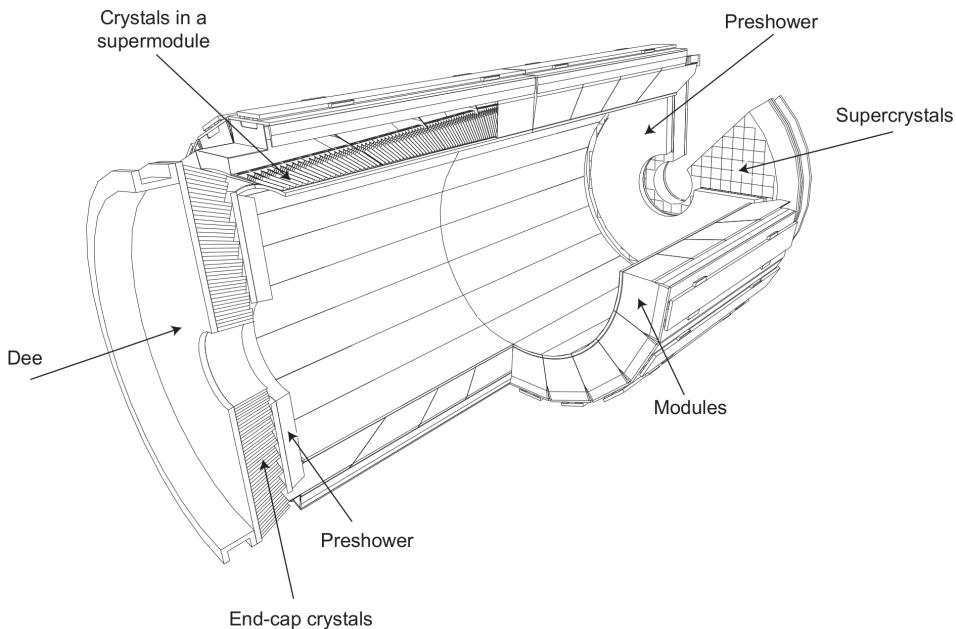


FIG. 1.7: Vue en perspective du calorimètre électromagnétique et de ses sous-éléments [13].

de cristaux de tungstate de plomb ($PbWO_4$). Comme le trajectographe il est composé en deux parties : le tonneau (EB, pour *ECAL Barrel*) couvrant la région angulaire $0 < |\eta| < 1.479$ avec 61200 cristaux ; et le bouchon (EE, pour *ECAL Endcap*) couvrant la région angulaire $1.479 < |\eta| < 3$ avec 15000 cristaux au total. La figure FIG.1.7 présente une vue schématique de la structure du ECAL. Il est à noter que le ECAL est impliqué également dans le déclenchement (voir section 1.2.7).

1.2.4.1 Les cristaux

Ces cristaux ont été spécialement conçus pour avoir une faible longueur de radiation, c'est à dire qu'ils sont capables de capter les photons et les électrons sur une distance relativement courte ($X_0 = 0.89$ cm). Le $PbWO_4$ possède également un faible rayon de molière (2.19 cm)¹. La conséquence directe est la haute densité des cristaux (8.29 g cm^{-3}) du calorimètre. On a également la possibilité d'utiliser des cristaux ayant des dimensions très réduites. Plus précisément, 21.8 mm × 21.8 mm de surface dans le tonneau, et 24.7 mm × 24.7 mm de surface dans les bouchons, ce qui implique une très grande granularité en η et ϕ .

Cependant, il existe un inconvénient majeur qu'est la forte sensibilité de la réponse à la température ($\sim 2\% \text{ }^{\circ}\text{C}^{-1}$). Cela impose un système de refroidissement pour maintenir une température très stable dans ECAL ($\pm 0.05 \text{ }^{\circ}\text{C}$).

Lors du passage d'un électron ou d'un photon, de la lumière est émise en quantité proportionnelle à l'énergie déposée dans le cristal, il s'agit d'un phénomène de scin-

¹ Le rayon de Molière définit la taille du cylindre qui captera 90% de l'énergie de la gerbe électromagnétique.

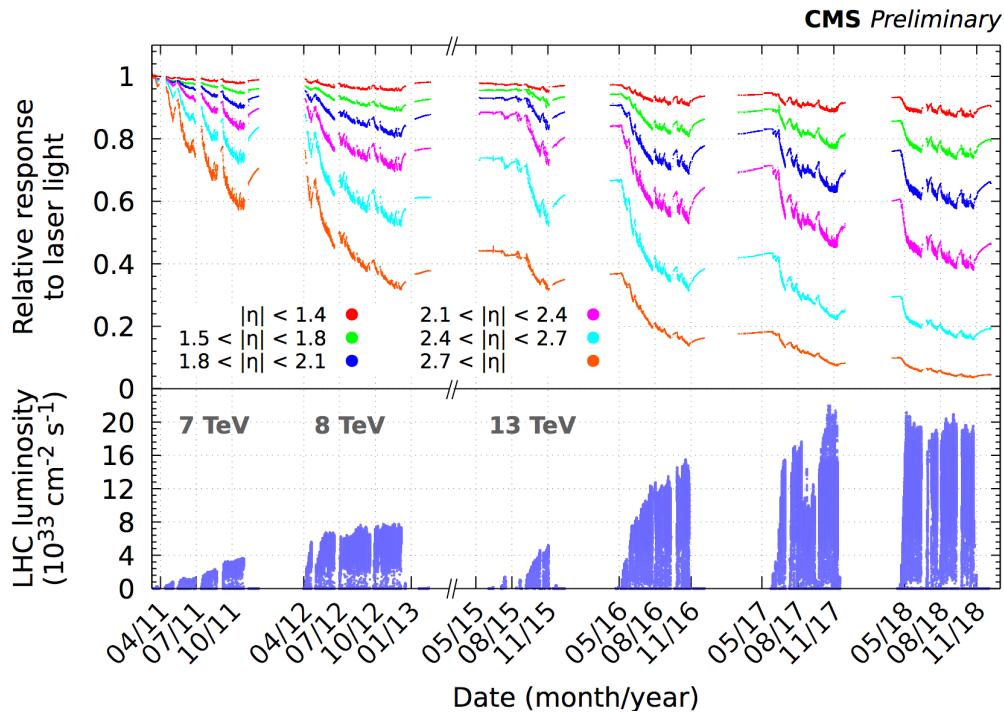


FIG. 1.8: Réponse relative des cristaux du ECAL en fonction du temps, pour différentes zones en $|\eta|$ (haut) et luminosité instantanée du LHC (bas) [16].

tillation. Des photodétecteurs sont associés à chaque cristal pour capter cette lumière. Le champ magnétique intense de 3.8 T et les très fortes radiations, ont imposé le choix de deux technologies différentes : les photodiodes à avalanches pour le tonneau (APD), et les phototriodes à vide pour les bouchons (VPT).

1.2.4.2 *Le détecteur à pied de gerbe*

A cause de la granularité insuffisante des cristaux dans les bouchons, certains hadrons neutres, tels que les π^0 , se désintégrant en deux photons, peuvent être interprétés comme un photon par le calorimètre.

Ainsi, pour faciliter la discrimination des photons issus d'évènements durs et ceux issus de la désintégration d'hadrons neutres, il a été décidé de placer un détecteur à pied de gerbe en amont du calorimètre électromagnétique, dans les bouchons. Ce détecteur permet d'initier la gerbe électromagnétique avant d'entrer dans le ECAL, grâce à une plaque de plomb et d'aluminium. Couvrant une région en $|\eta|$ entre 1.653 et 2.6 avec une surface de détection en silicium de 8 m^2 , il possède une granularité bien plus importante que le ECAL, avec des bandes de détection de 2 mm de large. Grâce à ce dispositif, il est, par exemple, possible d'améliorer la distinction des photons issus d'un π^0 boosté.

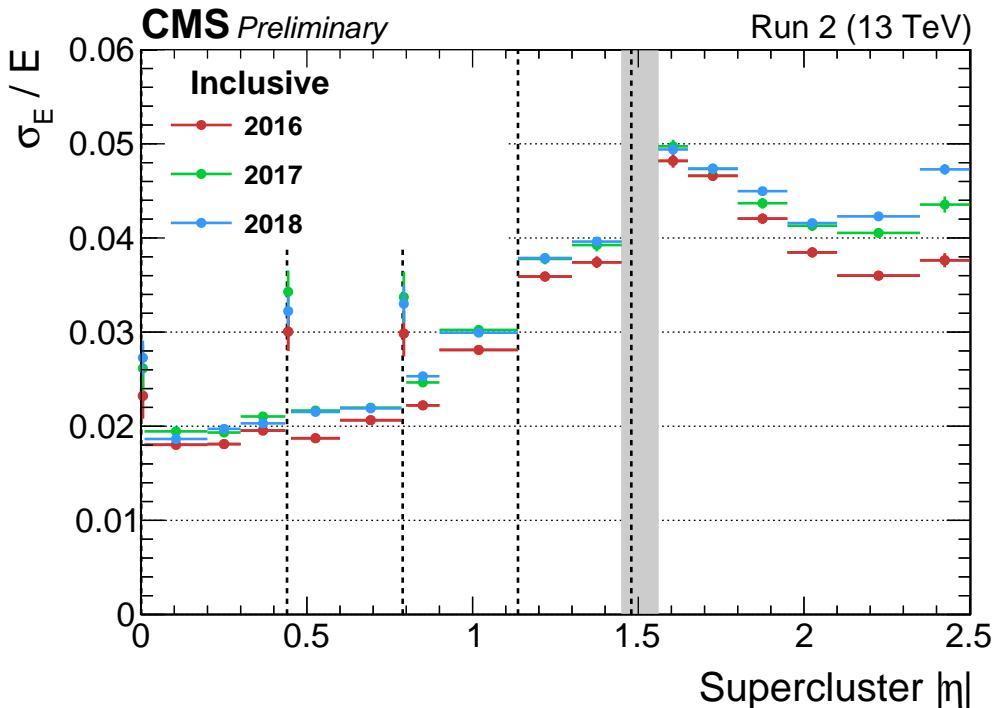


FIG. 1.9: Résolution relative du ECAL lors du Run II en fonction de la position angulaire η du super-cluster des électrons issus des désintégrations $Z \rightarrow e^- e^+$ [17].

1.2.4.3 Performances

Une caractéristique impactant les cristaux du ECAL est la sensibilité aux radiations. Cette sensibilité se manifeste par la perte de transparence des cristaux au fil du temps. Afin de mesurer cet effet, un système de contrôle a été mis en place. Un balayage laser est envoyé sur tous les cristaux afin de mesurer leur transparence. On peut voir dans la figure FIG.1.8 la réponse relative des cristaux du ECAL en fonction de la date de prise de données. On constate clairement la diminution de la réponse relative. Des corrections sont déployées pour corriger cet effet.

La résolution σ du ECAL est paramétrée par :

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

où S est le terme stochastique, dû aux fluctuations dans l'étalement latéral de la gerbe électronique, N le terme de bruit des composants électroniques, et C un terme constant qui prend en compte les erreurs de calibration. On peut voir figure FIG.1.9 l'évolution de la résolution relative en fonction de la position angulaire η du super-cluster d'énergie. La paramétrisation, lors de tests sur faisceaux en 2006, fut évaluée à

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{0.12}{E}\right)^2 + (0.30\%)^2$$

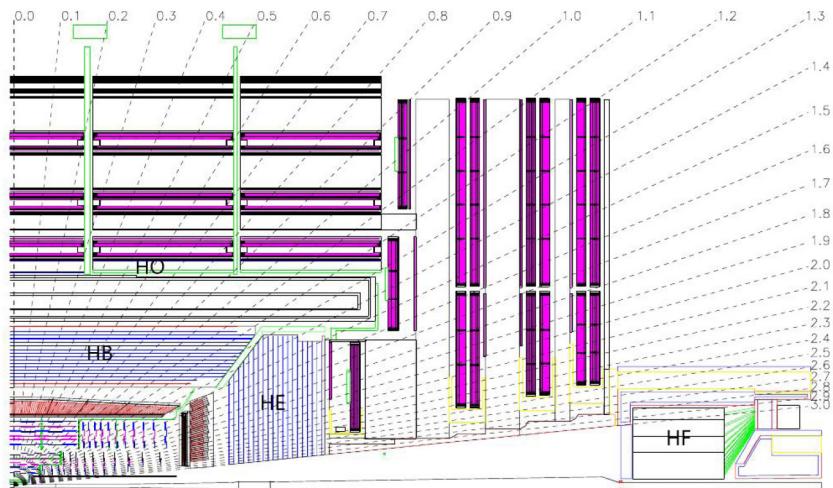


FIG. 1.10: Vue longitudinale du calorimètre hadronique [13].

1.2.5 CALORIMÈTRE HADRONIQUE

Le calorimètre hadronique (HCAL, pour *Hadronic CALorimeter*) permet de mesurer l'énergie des hadrons neutres et chargés. Une succession de couches d'absorbants et de scintillateurs permet de déterminer la trajectoire et l'énergie d'une particule incidente. Il est constitué de trois parties : une partie tonneau (HB et HO) et bouchon (HE) de manière analogue aux ECAL et d'une partie périphérique au détecteur (HF). Une vue longitudinale du HCAL est visible à la figure Fig.1.10. Il est à noter que le HCAL est impliqué également dans le déclenchement (voir section 1.2.7).

Pour comprendre les caractéristiques intrinsèques du HCAL on définit la longueur d'interaction nucléaire λ_0 . Caractéristique des matériaux, elle est la longueur après laquelle 36.8 % ($1/e$) des hadrons seront absorbés par le milieu. L'absorbant retenu pour le tonneau et les bouchons du HCAL est le laiton avec $\lambda_0 = 16.42$ cm. Cet alliage a été choisi pour ses propriétés paramagnétiques, dues au placement du HCAL à l'intérieur de l'aimant. Le laiton est également quasi transparent pour les muons, ce qui permet de conserver la quasi totalité de leur impulsion en traversant le détecteur. Chaque couche d'absorbant mesure 5 cm d'épaisseur dans le tonneau, et 8 cm dans les bouchons. Entre chaque couche d'absorbant se trouve une couche de scintillateur plastique de 3.7 mm d'épaisseur.

Le tonneau couvre une zone angulaire entre $0 \leq |\eta| < 1.48$, et est composé de deux parties différentes. La première partie, le HB, mesure 89 cm de profondeur, soit seulement $5.82 \lambda_0^2$ ². Cette longueur est insuffisante pour absorber la totalité de la gerbe hadronique. Une deuxième partie, le HO, a donc été ajoutée, située à l'extérieur du solénoïde. L'aimant sert alors d'absorbeur, et le HO détecte les gerbes longues ou tardives. Les bouchons couvrent une zone angulaire entre $1.48 \leq |\eta| < 3$. La longueur

² La longueur d'interaction nucléaire, λ_0 , est une longueur caractéristique des matériaux. C'est la longueur après laquelle 36.8 % ($1/e$) des hadrons sont absorbés par le milieu. Pour le laiton, on a $\lambda_0 = 16.42$ cm.

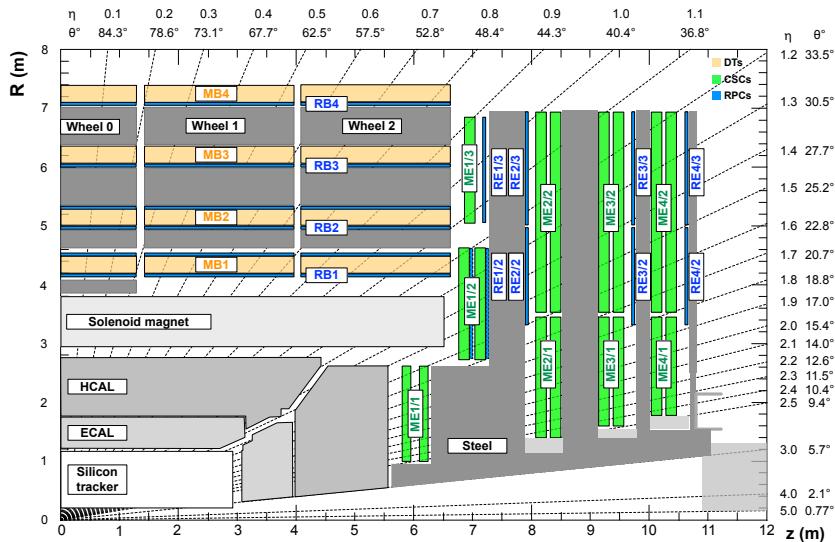


FIG. 1.11: Vue longitudinale des chambres à muons [13].

totale du calorimètre est, en incluant le ECAL, d'environ $10 \lambda_0$, suffisante pour arrêter la gerbe hadronique.

La dernière partie, le HF, couvre une zone angulaire entre $2.9 \leq |\eta| < 5.2$. C'est un cylindre de 130 cm de rayon, situé à 11.2 m du point d'interaction. D'une longueur de 165 cm ($\sim 10 \lambda_0$), il est constitué d'un absorbeur en acier, dans lequel sont introduites des fibres optiques en quartz. Les particules chargées entrant dans le milieu émettent des photons par effet Cherenkov. Ils sont collectés par les fibres puis amplifiés par photomultiplicateurs.

1.2.6 CHAMBRES À MUONS

Le détecteur à muons joue un rôle central dans CMS, puisqu'il couvre trois fonctions principales : identifier les muons, mesurer leur impulsion, et assurer une partie du déclenchement (voir 1.2.7). Il s'agit de la partie externe du détecteur. Les autres types de particules ayant déjà déposé leur énergie dans les autres sous-détecteurs de CMS, seules les muons pourront l'atteindre. Son éloignement au trajectographe, rend la mesure de l'impulsion de ces particules beaucoup plus facile avec leur temps de vie $2.2 \mu\text{s}$.

Il existe trois types de détecteurs à gaz différents dans les chambres, dont l'agencement est présenté à la figure FIG.1.11.

LES TUBES À DÉRIVE (DT) Installés dans le tonneau (voir figure 1.12), ils couvrent une zone angulaire $|\eta| < 1.2$. Il y a 250 stations de DT dans le détecteur.

LES CHAMBRES À PISTES CATHODIQUES (CSC) Installés dans les bouchons, elles sont conçues pour supporter les variations du champ magnétique et les fortes radiations. 540 de ces modules sont installés, sur une zone angulaire $0.9 < |\eta| < 2.4$.

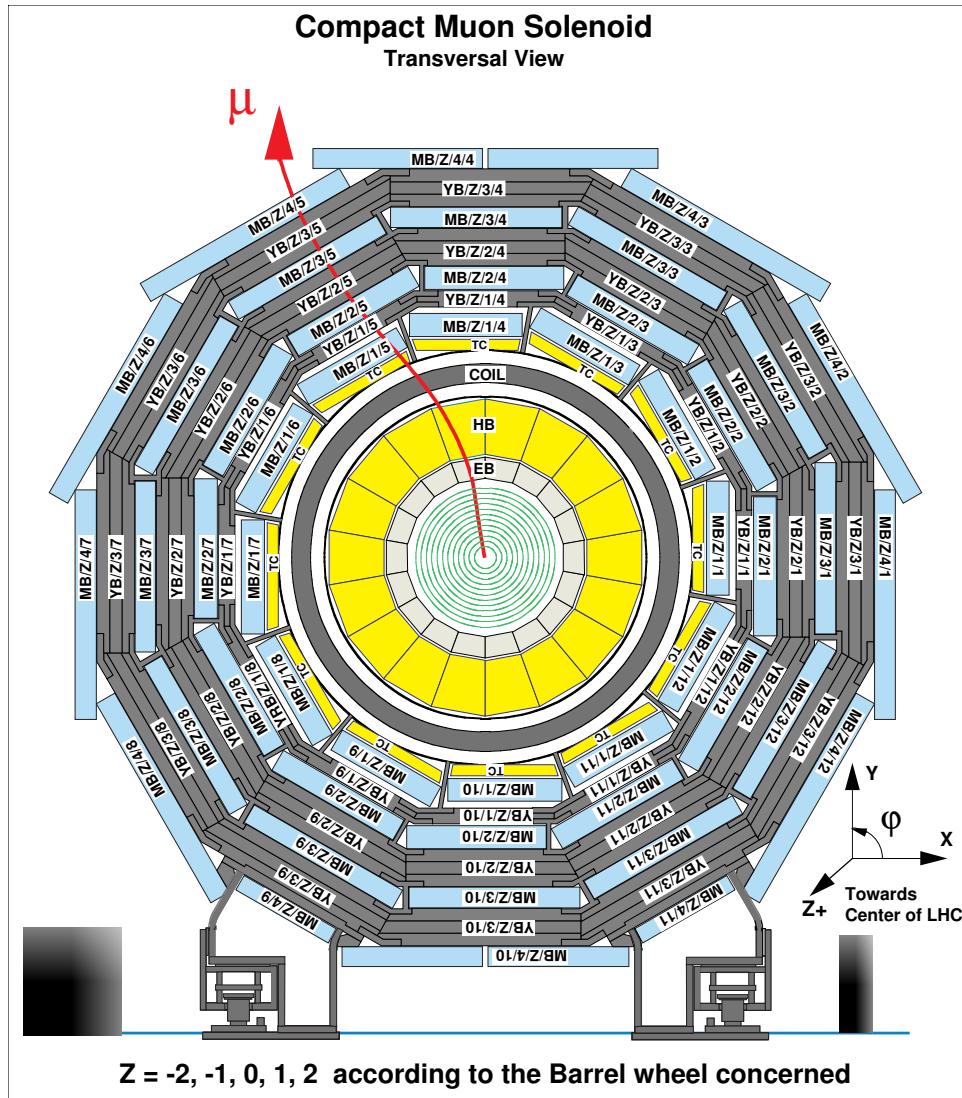


FIG. 1.12: Agencement des tubes à dérives dans le tonneau de CMS (vue transverse). Chaque rectangle bleu correspond à un tube à dérive.

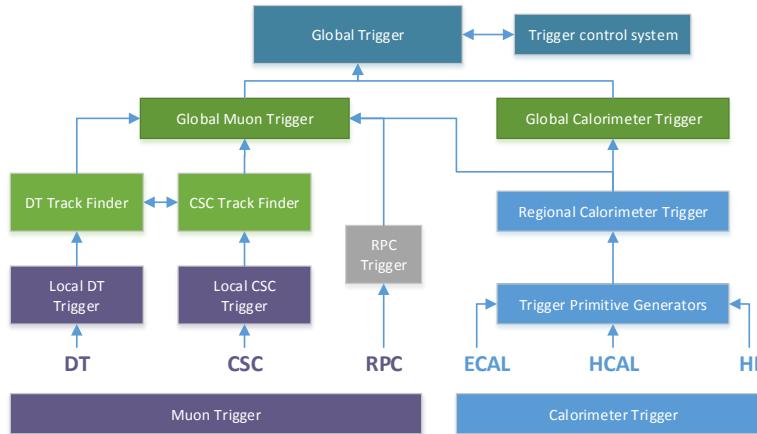


FIG. 1.13: L'architecture du déclencheur de niveau 1

LES DÉTECTEURS À PLAQUES RÉSISTIVES (RPC) Installés à la fois dans le tonneau et dans les bouchons, ils permettent d'assurer une partie du déclenchement grâce à leur temps de réponse inférieur à 25 ns.

1.2.7 SYSTÈME DE DÉCLENCHEMENT

Pendant une prise de données, les paquets de protons se croisent toutes les 25 ns, soit une fréquence de collisions de 40 MHz. Or seul un flux de $\sim 0.4 \text{ Gio s}^{-1}$ peut être enregistré sur disque. En estimant un événement à environ 2 Mio, on obtient alors un débit de données de 80 Tio s^{-1} , bien trop élevé pour être traité en temps réel. La stratégie est de sélectionner en ligne les événements que l'on veut garder et ignorer ceux qui présentent peu d'intérêt pour les analyses de la collaboration.

Pour cela, on utilise un système de déclenchement (*trigger*), à deux niveaux : le niveau 1 (L1) et le HLT (*High Level Trigger*, pour déclenchement de haut niveau). Les figures Fig.1.13 et Fig.1.14 présentent respectivement des déclenchement L1 et HLT.

LE L1 En utilisant directement les informations provenant du détecteur à muons et des calorimètres, ce premier niveau de déclenchement réduit le taux d'événements de 40 MHz à 100 kHz. La décision de garder ou non l'événement est prise en moins de 3.2 μs . Cette électronique est support de 128 algorithmes de déclenchement tournant en parallèle.

LE HLT Le taux d'événements doit être réduit à environ 100 Hz [18] afin d'être enregistré en temps réel. Les HLT, chargés d'effectuer cette sélection, utilisent une ferme d'ordinateurs, installée à proximité du détecteur. Cette ferme réalise une reconstruction rapide afin d'obtenir une description de l'événement en terme d'objets physiques (photon, muon, électrons et jets) plutôt qu'en terme de signaux électriques. Les conditions de sélections sont hautement configurables (choix des

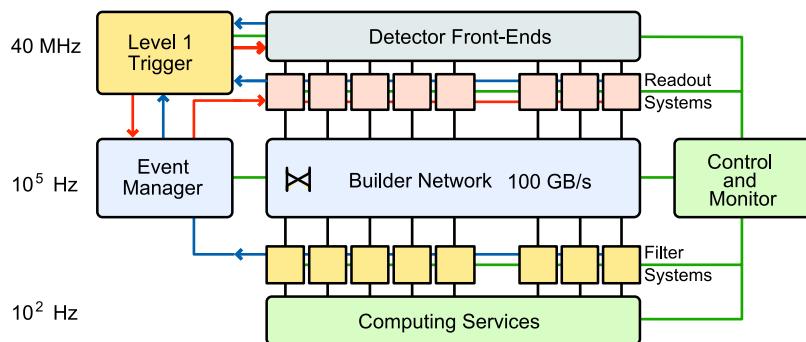


FIG. 1.14: Schéma de la chaîne d'acquisition de données de CMS

particules présentes, seuil en énergie, ...) ainsi de nombreux HLT différents sont utilisés. Le temps moyen alloué au HLT pour prendre une décision est d'environ 50 ms.

Une fois accepté, l'événement est stocké de façon définitive dans un centre de stockage à haute redondance, le *Tier-0*.

1.3 MESURE DE LA LUMINOSITÉ

La mesure de la luminosité est un aspect important pour les analyses physiques. Pour l'analyse contenue dans cette thèse, la compréhension de cette mesure est d'autant plus importante qu'elle est à la fois l'incertitude systématique dominante mais également une quantité corrélée au temps. A CMS la luminosité est mesurée avec une combinaison de cinq luminomètres principaux :

PIXEL LUMINOSITY TELESCOPE (PLT) Il s'agit d'un sous-détecteur dédié à la mesure de la luminosité en utilisant des capteurs en pixel de silicium. Les quarante-huit capteurs sont répartis en seize piles de trois capteurs appelées « télescopes ». Le PLT mesure le taux de « triples mesures », c'est à dire le taux de signaux enregistrés sur les trois capteurs d'un même télescope.

FAST BEAM CONDITIONS MONITOR (BCM1F) Ce sous-détecteur est composé de dix capteurs en silicium, dix capteurs en diamant polycristallins (pCVD) et quatre capteurs en diamant monocristal (sCVD). Cette structure est dotée d'une lecture rapide à 6.25 ns qui permet l'enregistrement de la luminosité en temps réel.

FORWARD CALORIMETER (HF) Le HF joue un rôle clé dans CMS puisqu'il permet de mesurer en temps réel la luminosité instantanée reçue par CMS. Deux méthodes sont utilisées. La première utilise le nombre de tour sans signal pour en déduire le nombre moyen d'interactions par croisements de faisceaux. La deuxième méthode tire parti de la relation linéaire qui existe entre l'énergie transverse moyenne par tour et la luminosité.

PIXEL DETECTOR Ce détecteur à pixels contient 65×10^6 pixels, ce qui lui permet de mesurer les trajectoires des particules issues de la collision avec une extrême précision. C'est également le détecteur le plus proche du faisceau avec des couches cylindriques à 4 cm, 7 cm et 11 cm et des bouchons à chaque extrémité. Il est essentiel pour reconstruire les trajectoires des particules à très courte durée de vie.

RAMSES Depuis 2017, ce sous-détecteur fait partie des systèmes de protection et de surveillance de l'environnement du LHC, mais peut également être utilisé pour fournir une mesure de luminosité. Le capteur est une chambre d'ionisation cylindrique en plastique remplie de 3 L d'air à pression atmosphérique. Les parois sont recouvertes de 4 mm de graphite PE. Il détecte principalement les photons dans la gamme d'énergie de 50 keV à 7 MeV.

Afin de calibrer les méthodes de détections, le LHC effectue régulièrement des scans Van Der Meer [19] pilotés par l'équation :

$$\sigma = \iint \frac{\mu(\Delta x, \Delta y)}{N_1 N_2} d\Delta x d\Delta y \quad (1.5)$$

avec σ la section efficace visible, N_1 et N_2 le nombre de protons dans chaque faisceau. L'intégrande est donné par :

$$\frac{\mu(\Delta x, \Delta y)}{N_1 N_2} = \sigma \frac{\mathcal{L}_{\text{inst}}}{N_1 N_2} \quad (1.6)$$

avec $\mathcal{L}_{\text{inst}}$ la luminosité instantanée. En pratique, les scans permettent de déterminer précisément la forme horizontale et verticale des faisceaux, en décalant l'un des faisceaux selon l'axe x ou y . Ils permettent en plus de déterminer la section efficace inélastique pp visible par les luminomètres ce qui permet de calibrer les mesures de luminosité.

Durant ma thèse, j'ai contribué à la calibration de la luminosité enregistrée en 2018, avec étude de l'incertitude sur le bruit de fond du détecteur. Nous avons utilisé la méthode du comptage d'agrégat de pixels (ou PCC pour *Pixel Cluster Counting*). Cette méthode consiste simplement en un décompte du nombre de clusters de pixels activés au passage d'une particule.

Pour estimer le bruit de fond, il faut effectuer un PCC sur une zone temporelle d'arrêt de l'activité du LHC. Cette zone est identifiable car elle consiste en une longue pause entre deux scans Van Der Meer. L'étude s'est portée sur deux scans à longues séparations effectués le 30 juin et le 1er juillet 2018 où les deux faisceaux étaient séparés par une distance de 6σ , où σ est la largeur du paquet de protons. Le nombre moyen de cluster est mesuré à $\langle n_{\text{cluster}} \rangle = 0.169 \pm 0.003$ (stat).

J'ai proposé une incertitude systématique, en examinant la variation du nombre moyen de clusters en fonction du temps, en utilisant les deux scans à longues séparations. Ce procédé a révélé une stabilité contenue dans l'incertitude systématique absolue de $\sigma_{\text{syst}} = 0.010$ et visible dans la figure Fig.1.15.

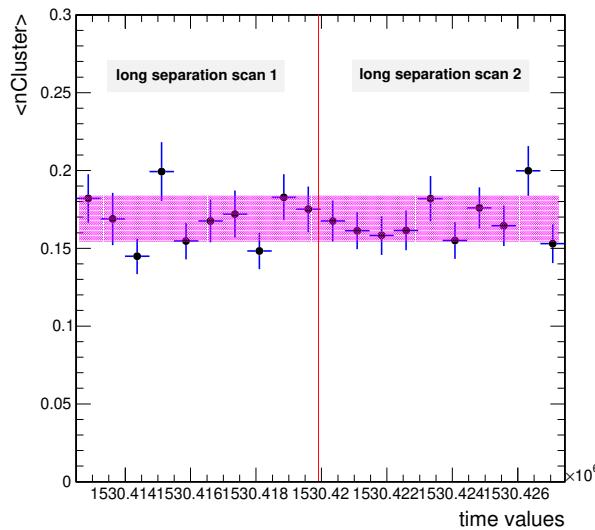


FIG. 1.15: Incertitude systématique du comptage de clusters de pixels avec $\langle n_{\text{cluster}} \rangle$ fonction du temps, pour les deux scans à longue séparation.

Le bruit de fond estimé par PCC est de

$$\langle n_{\text{cluster}} \rangle = 0.169 \pm 0.011 \text{ (stat + syst)} \quad (1.7)$$

où l'incertitude totale a été estimée par sommation en quadrature.

Des mesures similaires des bruits de fond sont effectuées avec les autres luminomètres par le groupe Luminosité de CMS. Ces mesures permettent de soustraire le taux de bruit de fond dans l'estimation des sections efficaces visibles, et sont donc importantes pour la calibration. Cette étude a contribué à la note publique de CMS [20] pour l'estimation de la luminosité sur les données de 2018.

CONCLUSION

Lors de ce chapitre, ont été présentés le LHC et le détecteur CMS. Leur conception a pour but d'enrichir nos connaissances du Modèle Standard grâce à des mesures de précision et d'offrir de nouvelles perspectives de recherche de nouvelles physiques au-delà Modèle Standard. Pour assurer une analyse physique fonctionnelle, la première étape est la reconstruction les événements d'une collision. Il s'agit de la transformation de signaux électriques captés par les différents sous-détecteurs de CMS en particules et autres objets physiques tels que des électrons, des muons et des gerbes hadroniques. Les méthodes utilisées par la CMS sont décrites plus en détail dans le chapitre suivant.

2

SIMULATION ET RECONSTRUCTION DES ÉVÈNEMENTS

2.1	Génération des événements	21
2.1.1	L'événement dur	21
2.1.2	Densité de probabilité partonique (PDF)	22
2.1.3	De l'événement dur aux particules dans le détecteur.	24
2.1.4	Simulation du détecteur	25
2.2	Reconstruction des événements	26
2.2.1	L'algorithme <i>Particle-Flow</i> (PF).	26
2.2.2	Isolation des leptons et des photons.	32
2.2.3	Reconstruction des jets.	33
2.2.4	Étiquetage des jets ISSUS DE QUARKS <i>b</i>	36
2.2.5	Énergie transverse manquante.	37
2.3	Conclusion	38

2.1 GÉNÉRATION DES ÉVÈNEMENTS

Les tests de validité d'une proposition théorique passent par la confrontation expérimentale. Dans le cas de la physique des particules, la stratégie consiste en la comparaison des données mesurées par CMS avec des simulations idéales prédictes par le Modèle Standard. La première étape est de simuler la collision entre deux protons au LHC grâce à des outils appelés générateurs d'évènements Monte-Carlo (MC). La second étape consiste à simuler l'interaction des particules créées avec le détecteur. Enfin, dû à une impossibilité de simuler parfaitement les collisions et les interactions, une phase de correction des évènements MC est nécessaire. Ce chapitre présente les différentes étapes sus-nommées.

2.1.1 L'ÉVÉNEMENT DUR

L'évènement fondamental est la collision entre deux protons. Dans cette collision ce sont leurs constituants, les partons (quarks et gluons), qui vont interagir (voir figure Fig.2.1). Chacun d'entre eux emporte une fraction x de l'impulsion totale du proton incident. Les observables physiques (sections efficaces, ...) sont extraites en utilisant des développements perturbatifs.

Le premier ordre (LO, *Leading Order*) de la QCD perturbative est implémenté dans les programmes de simulations tels que `MadGraph_aMC@NLO` [21]. La complexité des

ordres supérieurs de calculs implique qu'un petit nombre seulement de processus sont disponibles et implémentés dans des programmes particuliers comme Powheg [22].

La section efficace partonique $\sigma_{ij} \rightarrow X$, avec $i, j = g, q$ les composants de l'état initial et X un état final choisi, peut s'exprimer de la façon suivante :

$$\sigma_X = \sum_{ij} \int_0^1 dx_i dx_j \int f_i(x_i, \mu_F^2) f_j(x_j, \mu_F^2) d\sigma_{ij \rightarrow X}(\mu_F, \mu_R) \quad (2.1)$$

Sachant

$$\begin{aligned} d\sigma_{ij \rightarrow X}(\mu_F, \mu_R) &= \frac{1}{2} |\mathcal{M}_{ij \rightarrow X}|^2 (\Phi_N, \mu_F, \mu_R) d\Phi_N \\ d\Phi_N &= \int \prod_{k=1}^N \frac{d^3 p_k}{(2\pi)^3 2E_k} \delta^4 \left(p_i + p_j - \sum_{k=1}^N p_k \right) \end{aligned} \quad (2.2)$$

où x_i la fraction d'énergie emportée par la particule i , $f_i(x_i, \mu_F^2)$ sa densité de probabilité partonique (plus de détails dans la section 2.1.2), μ_F l'échelle de factorisation, μ_R l'échelle de renormalisation, $d\Phi_N$ l'élément infinitésimal de l'espace des phases et $|\mathcal{M}_{ij \rightarrow X}|$ l'élément de matrice de transition entre l'état initial i et l'état final f . Le générateur suit cette distribution d'évènements donnée par les formules de sections efficaces, on obtient alors, en sortie, les 4-vecteurs impulsions de chaque particule produite lors de l'événement dur pour chaque évènement simulé.

Bien que les processus de bases soit générés à cette étape, les particules produites peuvent avoir des caractéristiques supplémentaires nécessitant une couche supplémentaire de génération. Par exemple, les particules soumises à l'interaction forte vont former des gerbes partoniques puis s'hadroniser (voir FIG.2.1), ces deux processus sont générés ultérieurement.

2.1.2 DENSITÉ DE PROBABILITÉ PARTONIQUE (PDF)

Les hadrons peuvent être vus comme une soupe statistique de partons avec, en permanence, création et annihilation de couple quark-antiquark. Autrement dit, quand l'on parle de protons composés de deux quarks up et d'un down, on commet un abus de langage. En effet, la QCD implique que si l'on pouvait prendre en photo, un grand nombre de fois, un proton alors, statistiquement, nous observerions qu'il est composé la plupart du temps de deux quarks up et d'un down. Au vu du nombre de collisions au LHC, la présence d'autres quarks dans les protons est possible. Pour le calcul des sections efficaces des processus faisant intervenir des hadrons on utilise les densités de probabilité partonique $f_i(x_i, \mu_F^2)$ (ou PDF pour *Parton distribution function*). Ces fonctions ne peuvent être calculées analytiquement à cause de l'échelle d'énergie empêchant le calcul perturbatif et valable uniquement dans le cas ultra-relativiste. Ces fonctions sont obtenues expérimentalement à partir d'ajustement sur les données. Les expériences de diffusion inélastique des leptons sur les hadrons, dont fait partie par exemple l'expérience HERA [23] (collisionneur e^\pm -p du laboratoire DESY à Hambourg en Allemagne) jouent un rôle majeur dans l'ajustement des PDF.

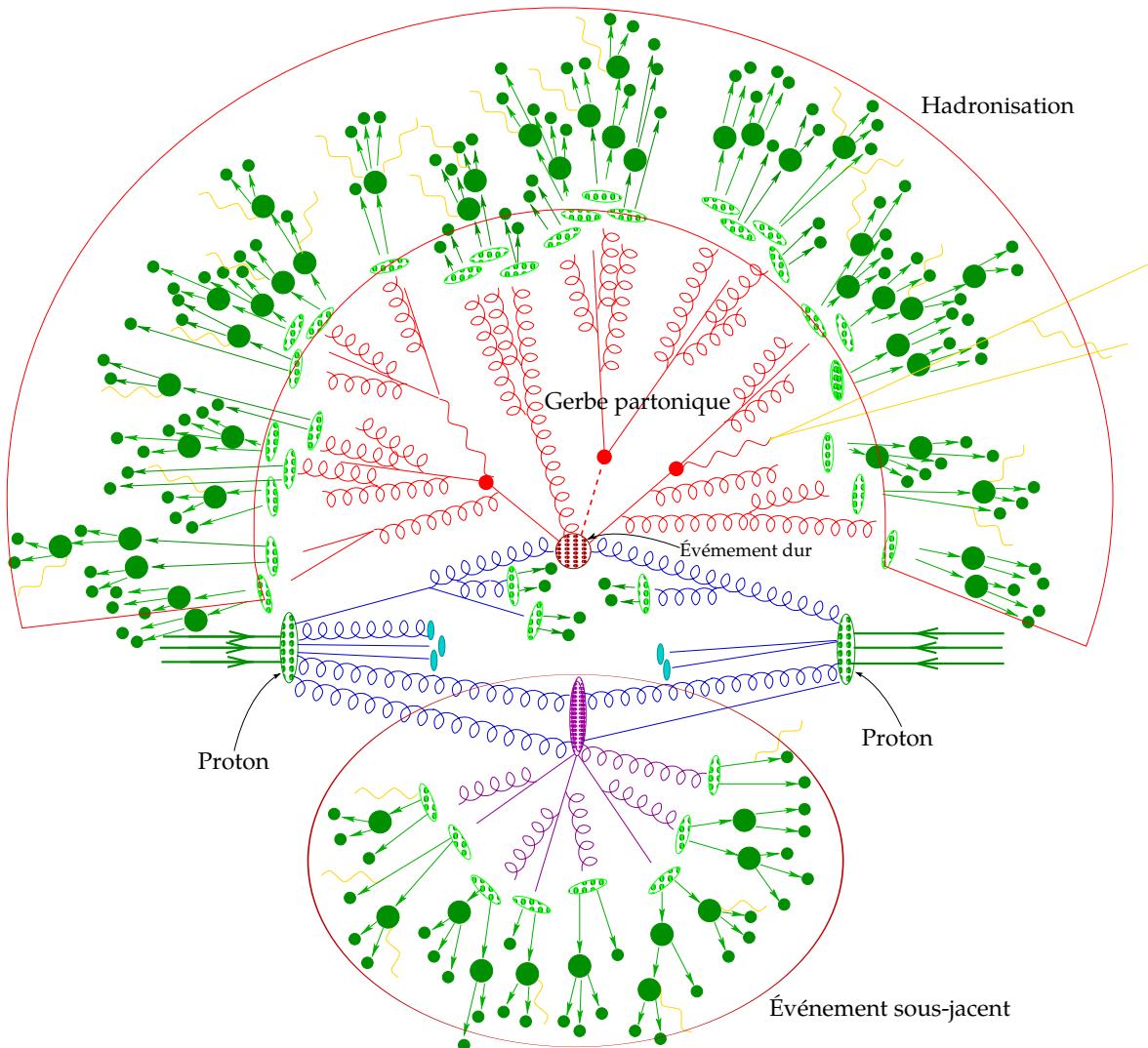


FIG. 2.1: Représentation graphique d'une collision de deux protons simulée à l'aide d'un générateur.

Les PDF dépendent de la variable de Bjorken x qui représente la fraction d'énergie emportée par le parton considéré et l'échelle en énergie du processus Q^2 (transfert d'impulsion). En toute rigueur, les PDF sont évaluées à une échelle d'énergie donnée dite échelle de factorisation μ_F^2 qu'on choisit de l'ordre de grandeur de l'échelle d'énergie du processus étudié : $\mu_F \equiv Q$. Pour les déterminer, on fait le choix d'une paramétrisation à une échelle d'énergie Q_0 donnée. La forme la plus générale est :

$$x f(x, Q_0^2) = A_f x^{a_f} (1-x)^{b_f} I_f(x) \quad (2.3)$$

L'annulation des PDF pour $x \rightarrow 1$ est assurée par le terme $(1-x)^{b_f}$. a_f et b_f sont les paramètres à déterminer à partir des données expérimentales. Cette forme de PDF est attendue avec la règle de comptage des quarks [24]. La dépendance en Q^2 est obtenue à l'aide des équations d'évolution DGLAP (Dokshitzer–Gribov–Lipatov–Altarelli–Parisi) [25–27]. Elles sont essentielles pour prédire à partir des valeurs collectées à bas Q^2 , les PDF à hautes valeurs de Q^2 non explorées par les données mais nécessaires au LHC.

Les PDF utilisées pour produire les simulations ont été fournies par le groupe CTEQ [28] (voir la figure Fig.2.2) et NNPDF [29]. De plus le groupe de travail PDF4LHC [30] effectue des études comparatives des différentes PDF et des prévisions au LHC. Il fournit également une méthode standard pour l'estimation des incertitudes liés aux PDF au LHC grâce à une combinaison des résultats des différents groupes.

Le PDF du gluon est très important à $Q \sim m_t$, c'est à dire que ce sont des gluons qui seront majoritairement représentés dans les collisions. Ceci explique la plus grande proportion, au LHC, de processus $t\bar{t}$ produit par fusion de gluons par rapport à une production par annihilation $q\bar{q}$. Ce résultat déjà discuté dans le chapitre ??.

2.1.3 DE L'ÉVÉNEMENT DUR AUX PARTICULES DANS LE DÉTECTEUR

Les particules colorées provenant de l'événement dur génèrent une gerbe partonique. Il s'agit du phénomène de radiation en cascade de particules et de gluons secondaires provenant de gluons de haute énergie. Une fois atteinte une certaine échelle d'énergie, fixée par le générateur, les gluons restants sont forcés de se désintégrer.

Finalement l'hadronisation intervient et ce sont des hadrons neutres du point de vue de l'interaction forte qui sont produits. Ce phénomène se produit environ 5×10^{-24} s après la production de la particule. Il n'existe pas encore de théorie décrivant correctement l'hadronisation, il faut faire appel à des modèles phénoménologiques.

A l'échelle d'énergie où ils se produisent, la QCD devient non-perturbative, ainsi ses effets sont très complexes à simuler. Il existe néanmoins un ensemble de programmes qui permettent de simuler la chaîne d'hadronisation. On peut citer, par exemple, Pythia [32] et Herwig [33]. Chacun utilise des algorithmes spécialisés afin de simuler au mieux le développement de la gerbe partonique et de la chaîne d'hadronisation.

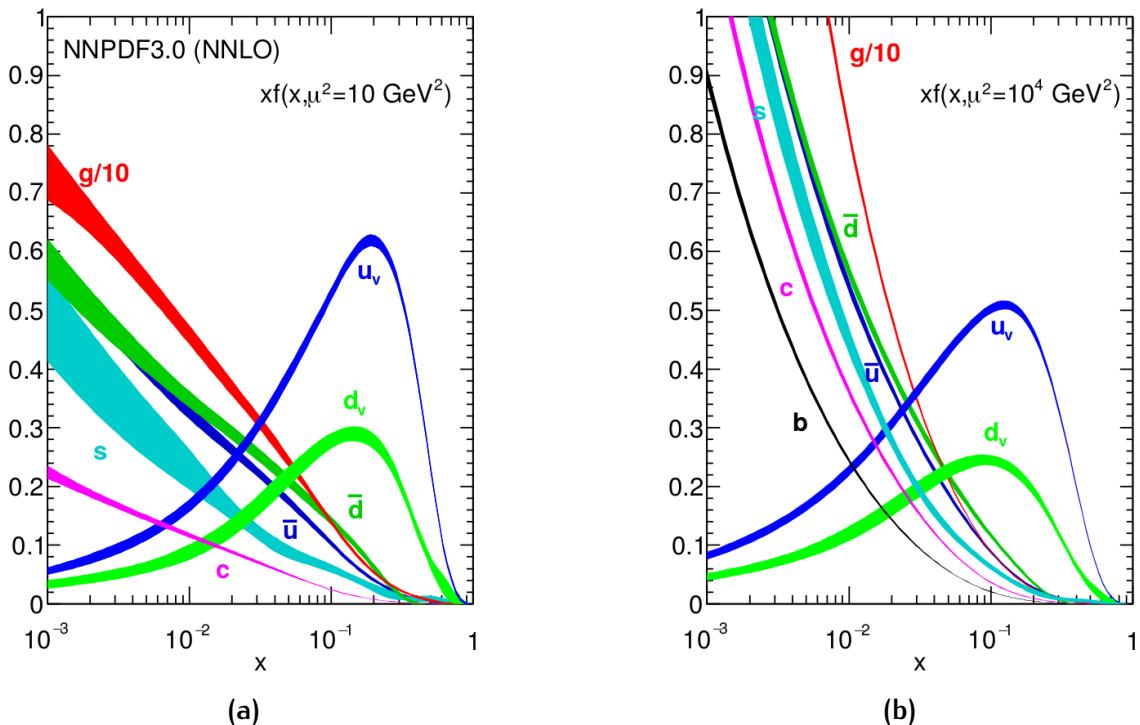


FIG. 2.2: Fonctions de densité partonique pour une échelle en énergie (a) $\mu^2 (= Q^2) = 10 \text{ GeV}$ et (b) $\mu^2 (= Q^2) = 10 \times 10^4 \text{ GeV}$ [31].

2.1.4 SIMULATION DU DÉTECTEUR

A ce stade, les différentes étapes de générations ont produit une liste de particules stables qui peuvent atteindre le détecteur. Il faut maintenant simuler leurs interactions avec CMS. Cette simulation se compose de :

- ▷ La modélisation du détecteur.
- ▷ La modélisation des interactions entre particules générées et la matière des sous-détecteurs de CMS (trajectographe, ECAL, HCAL, Chambre à muons)
- ▷ Les courbures des trajectoires des particules chargées par le solénoïde.
- ▷ La modélisation de l'empilement.

Cette simulation repose sur le logiciel Geant 4 [34]. Sa précision va jusqu'à la simulation du câblage interne du détecteur.

2.2 RECONSTRUCTION DES ÉVÈNEMENTS

La simulation du détecteur fournit en sortie les signaux électriques des sous-détecteurs, à l'instar d'une vraie collision. La reconstruction d'un évènement est donc identique lors de la simulation ou d'une prise de données réelle.

2.2.1 L'ALGORITHME *PARTICLE-FLOW* (PF)

CMS utilise un algorithme dédié pour reconstruire les objets physiques : l'algorithme du *particle-flow* (PF) [35–39]. Cet algorithme combine les informations provenant des sous-détecteurs de CMS pour reconstruire les photons, les électrons, les muons, les hadrons neutres et chargés. L'algorithme PF utilise notamment les traces des particules chargées dans le trajectographe, ainsi que les dépôts énergétiques des différents calorimètres. Les particules sont issues d'une interprétation des signatures combinées entre elles.

Il doit être noté que l'algorithme possède un taux de fausse identification non nul. Il est ainsi primordial d'avoir la plus haute efficacité de reconstruction possible. Pour maximiser cette efficacité les sous détecteurs sont pilotés par des algorithmes de trajectographie et de calorimetrie spécialisés.

2.2.1.1 Trajectographie itérative

Le trajectographe offre une meilleure résolution que les calorimètres pour la mesure de l'impulsion des hadrons. Comme près de $\frac{2}{3}$ de l'énergie d'un jet est due à la présence de hadrons chargés, le trajectographe joue un rôle majeur dans la performance du *particle-flow*.

Une procédure itérative [40], basée sur un filtre de Kalman (KF), est utilisé pour reconstruire les trajectoires des particules en minimisant le taux de faux. La graine de cette itération est donnée par la reconstruction induite des *hits* des deux premières couches du détecteur à pixels, ainsi que la position du vertex primaire. La trajectoire est ensuite complétée grâce aux *hits* des autres couches du trajectographe. Les contraintes de cette première étape sont fortes ce qui conduit à un très faible taux de faux.

Les hits associés à une trajectoire de manière non-ambiguë sont supprimés de la liste des *hits* disponibles. De manière récursive, l'algorithme réitère la construction de trajectoire en relâchant les contraintes à mesure que la liste des *hits* rapetisse. Ainsi, l'efficacité de reconstruction est améliorée tout en gardant un taux de faux très faible grâce à la réduction du nombre de combinaisons.

Suite à la troisième itération, on obtient une efficacité de reconstruction de 99.5 % pour les muons isolés et de plus de 90 % pour les hadrons chargés. Au-delà, la contrainte sur le point d'interaction est relâchée. Sont ainsi reconstruit les trajectoires des hadrons chargés secondaires (hadrons provenant de vertex d'interaction déplacé).

Pour prendre un exemple de reconstruction, une particule chargée de $p_T = 150 \text{ MeV}$ avec un vertex de production éloigné de plus de 50 cm de l'axe du faisceau, si elle ne laisse que trois *hits*, est reconstruite avec une efficacité de 99 %.

2.2.1.2 Vertex d'interaction : reconstruction et empilement

L'étape suivante est la reconstruction des vertex qui utilisent les traces précédemment reconstruites. On distingue deux étapes :

- ▷ Si la différence entre les coordonnées z au point le plus proche de l'axe du faisceau est inférieure à 2 mm, alors les traces sont regroupées en « candidats vertex».
- ▷ La position du vertex d'interaction est calculée par un algorithme d'ajustement [41] qui itère sur tous les agrégats de traces. L'algorithme calcule également des indicateurs sur la qualité de l'ajustement (χ^2) en assignant aux traces de l'agrégat un poids variant entre 0 et 1, suivant leur proximité avec l'axe du faisceau.

La présence de l'empilement parasite l'évènement dur. Pour outre-passé ce bruit de fond, on identifie le vertex primaire comme celui dont la somme des impulsions transverses en quadrature est la plus grande. Les autres vertex sont utilisés afin de réduire l'impact de l'empilement sur les analyses de physiques. En effet, un hadron chargé peut être identifié comme issu de la collision principale ou non, selon le vertex d'interaction auquel sa trace est attachée. On peut alors supprimer du processus de reconstruction tout hadrons chargés provenant d'un vertex secondaire que l'on étiquette comme de l'empilement.

2.2.1.3 Agglomération calorimétrique

Le troisième algorithme est l'algorithme d'agglomération calorimétrique de CMS. Il est décomposé en quatre procédures :

1. Mesurer l'énergie des particules neutres (photons, hadrons neutres).
2. Séparer les dépôts d'énergies des particules neutres de ceux des particules chargées.
3. Reconstruire les photons, les électrons et le rayonnement Bremsstrahlung associé.
4. Aider à la mesure de l'énergie des hadrons chargés pour lesquels la trajectoire n'a pas été bien reconstruite (hadrons de bas p_T).

Cet algorithme est appliqué de façon séparée dans chaque sous-calorimètre (tonneau du ECAL, tonneau du HCAL, bouchons du ECAL et HCAL, ...) en trois phases.

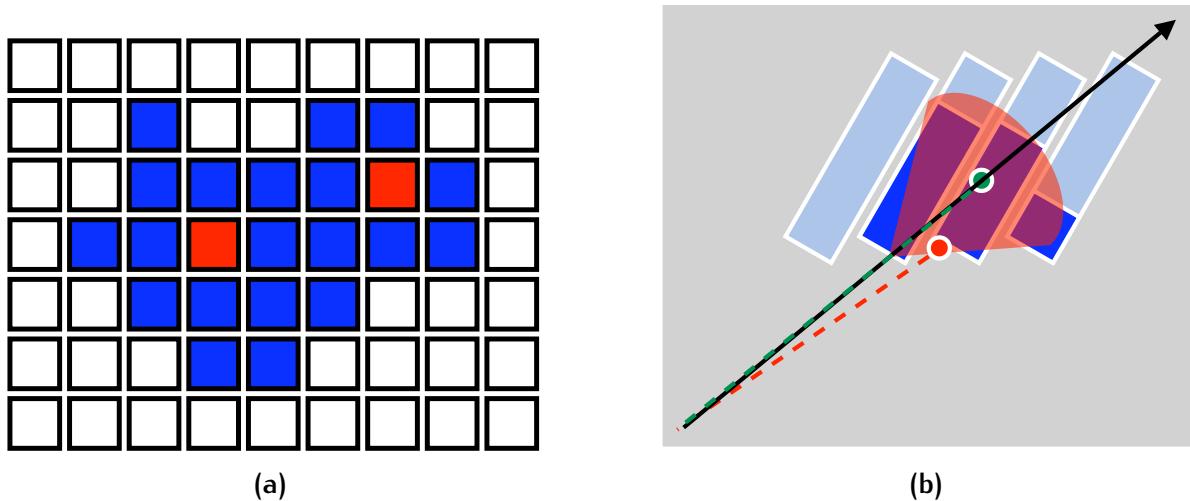


FIG. 2.3: Agglomérat topologique contenant deux graines (a) et détermination de la profondeur du dépôt d'énergie (b). En rouge, la position déterminée sans tenir compte de la profondeur biaisée $|\eta|$. En vert, la profondeur est correctement estimée.

IDENTIFICATION DES GRAINES Il s'agit de repérer les cellules calorimétriques où l'énergie dépasse un certain seuil. Ce seuil est fixé par majoration du bruit de fond du détecteur. Les graines contraignent les cellules voisines à ne pas devenir elles même des graines.

DEFINITION D'UN AGGLOMÉRAT TOPOLOGIQUE Ces agrégat est construit en agrégeant aux graines toutes les cellules qui leur sont adjacentes et possédant une énergie au dessus d'un seuil, fixé comme deux fois la déviation standard du bruit électronique dans le ECAL (80 MeV dans le tonneau, ~ 300 MeV dans les bouchons) et à 800 MeV dans le HCAL. Il peut y avoir plusieurs graines dans un même agglomérat topologique. Dans ce cas, le partage de l'énergie entre chaque graine est effectué selon la distance entre la cellule et la graine, en considérant que les gerbes électromagnétiques déposent leur énergie selon un profil gaussien, dont la largeur ne dépend pas de l'énergie. On peut voir une illustration dans la figure Fig.2.3a.

IDENTIFICATION DE LA POSITION DES AGGLOMÉRATS La position de l'agrégat est déterminée à partir de la graine et des 4 ou 8 cellules voisines, à l'aide de la formule :

$$X = \frac{\sum_i w_i X_i}{\sum_i w_i}, \text{ avec } w_i = \ln \frac{E_i}{E_{th}}$$

avec $X = x, y$ ou z , X_i la position de la cellule i , E_i l'énergie de la cellule i , et E_{th} le seuil d'énergie. Si l'agrégat ne compte qu'une seule graine, toutes les cellules de l'agrégat sont utilisées pour calculer la position. Afin de ne pas introduire de biais en η lors de la détermination de la position, on estime la profondeur du maximum de la gerbe électronique par la formule

$$p = a(b + \ln E)$$

où p est la profondeur, a et b sont des constantes qui dépendent de η , et E l'énergie totale de l'agrégat. On peut voir figure FIG.2.3b l'effet de cette procédure sur la détermination de la position angulaire de l'agrégat. Une vue géométrique est donnée dans la figure FIG.2.3b.

2.2.1.4 L'algorithme de liaison

La dernière étape est la création du lien entre les traces et les agrégats reconstruits. La subtilité provient du fait qu'une particule peut déposer de l'énergie dans les divers sous-détecteurs. Il faut également éliminer les possibilités de double comptage. Ce dernier algorithme produit des « blocs » d'éléments qui vont ensuite servir la reconstruction des particules.

Dans un premier temps on extrapole les traces reconstruites dans le trajectographe. On reconstruit la profondeur longitudinale de la gerbe :

- ▷ Dans le calorimètre électromagnétique, jusqu'à une profondeur correspondant au maximum attendu d'une gerbe électromagnétique.
- ▷ Dans le calorimètre hadronique, jusqu'à une profondeur d'une interaction nucléaire (λ_0 , voir ??).

Si l'extrapolation passe dans la zone délimitée par l'agrégat alors on relie la trace à l'agrégat. On définit la distance du lien par la distance $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ dans le plan (η, ϕ) . Les électrons émettent un rayonnement Bremsstrahlung, il perdent donc de l'énergie. Pour collecter l'énergie perdue, des trajectoires tangentes à la trajectoire principale sont extrapolées dans le ECAL. De la même manière on peut construire un lien entre les deux calorimètres (ECAL et HCAL) par recouvrement de zone. La distance du lien est donnée par ΔR dans le plan (η, ϕ) . Finalement, le spectromètre à muons reconstruit aussi des traces. Lorsque l'ajustement entre les traces des chambres à muons et celles du trajectographe donne un χ^2 acceptable, un lien est fait. Le muon est dit global.

2.2.1.5 Reconstruction des muons

Durant les étapes dites de reconstruction des particules, la reconstruction du muon[42] est la première étape et ce, même avant la reconstruction de l'évènement grâce à l'algorithme du *particle-flow*. Les muons laissent des traces dans le trajectographe et dans les chambres à muons, on a avec eux deux types de reconstructions.

RECONSTRUCTION DES MUONS GLOBAUX Cette reconstruction orientée pour les muons de impulsion transverse ($p_T \gtrsim 200 \text{ GeV}$) consiste en l'interpolation d'un lien entre une trace dans le spectromètre à muons et une trace correspondante, par extrapolation, imprimée dans le trajectographe. L'algorithme de lien trajectoire présenté dans la section 2.2.1.4 fonctionne avec un algorithme similaire. Cette méthode permet d'améliorer la résolution de l'impulsion par rapport à celle utilisant uniquement la trace du trajectographe.

RECONSTRUCTION DES TRACKER MUONS¹

Dans le cas de faibles impulsions ($p_T \lesssim 5 \text{ GeV}$), cette méthode est préférée. L'idée est de ne considérer que les traces présentes dans le trajectographe comme possibles candidats muons. En prenant en compte les pertes d'énergie et l'incertitude due aux multiples diffusions, ces muons sont peu énergétiques. On va donc exiger qu'au minimum une petite trace dans le spectromètre à muons, composée de quelques *hits* de DT ou CSC. Dans le contexte des faibles impulsions, cette approche est plus performante que la reconstruction des muons globaux.

Seul 1 % des muons ne sont pas reconstruits par l'une ou les deux méthodes précédentes. Les traces orphelines dans les chambres à muons sont rangés dans la troisième catégorie des muons *standalone*. Les trois catégories de muons sont ensuite regroupées dans une liste de candidats muons. Les candidats muons globaux et *tracker* muons qui partagent la même trace du trajectographe sont fusionnés en un seul candidat. Une sélection est effectuée sur les candidats muons reconstruits avec l'algorithme standard afin d'identifier les muons *particle-flow*. A CMS, on classe les selections dans trois familles différentes : « isolé », « *pf-tight* » et « *pf-loose* ».

MUON ISOLÉ Les muons sont considérés isolés si, dans un cône de taille $R = 0.3$ centré sur le muon, la somme du p_T des traces et des agrégats calorimétriques est inférieure à 10 % de l'impulsion du muon. En demandant aux muons d'être également globaux, une grande efficacité est obtenue dû à la propreté de l'environnement immédiat autour du muon, l'effet du *particle-flow* est limité.

MUON PF-TIGHT Cette sélection est optimisée pour identifier les muons au sein des jets. Elle requiert un certain nombre de *hits* dans les traces du spectromètre à muons. Est exigée également une compatibilité entre les dépôts d'énergies dans les calorimètres et ceux obtenus par des simulations.

MUON PF-LOOSE Cette sélection relâche la contrainte sur le nombre de *hits*, et supprime la contrainte sur les dépôts d'énergie.

On retire ensuite de la liste des blocs *particle-flow* disponibles, chaque muon reconstruit et on procède ensuite à la reconstruction des électrons.

2.2.1.6 Reconstruction des électrons

Les électrons sont beaucoup plus légers que les muons. Ainsi la radiation Bremsstrahlung est amplifiée par un facteur $(m_\mu/m_e)^4$, ce qui met en défaut l'algorithme de reconstruction des traces de CMS (voir 2.2.1.1). En effet, l'algorithme n'est pas optimisé pour gérer les changements abrupts de trajectoire des électrons, ce qui peut arriver par émission de photon Bremsstrahlung. Uniquement si l'émission est un photon de

¹ Il s'agit de muons qui possède des traces dans le trajectographe mais uniquement un segment de trace dans les chambres.

basse énergie, on peut espérer une reconstruction mais au prix d'une grande incertitude et d'un χ^2 grand. CMS a développé un algorithme dédié à la résolution de ce problème [35, 38, 39]. L'algorithme GSF (pour *Gaussian Sum Filter*) est capable de suivre les brusques changements de trajectoire grâce à son grand nombre de paramètres libres (plus d'une dizaine) et procure ainsi une bien meilleure estimation de l'impulsion des électrons. En contre-partie, cet algorithme est très gourmand en ressource (environ 200 ms par trace), et ne peut donc être utilisé que sur un nombre réduit de traces. Doit donc être introduite une pré-sélection.

Dans l'hypothèse d'un Bremsstrahlung négligeable, l'algorithme GSF reconstruit une trace et s'associe avec un bloc *particle-flow*. Si $E/p \simeq 1$, la trace est sélectionnée.

Au contraire pour un rayonnement Bremsstrahlung trop important, sont distingués deux effets :

- ▷ L'algorithme de reconnaissance de traces échoue lors d'un changement abrupt de trajectoire. La trace résultante contient alors un petit nombre de *hits*.
- ▷ L'algorithme de reconnaissance de traces reconstruit une trace, mais le χ^2 associé est grand.

Une sélection est appliquée en utilisant le nombre de *hits* et le χ^2 . Après une interpolation GSF de seulement 5 paramètres libres (donc moins contraignante), le χ^2 est associé avec le rapport χ_{KF}^2/χ_{GSF}^2 , le nombre de *hits* et le facteur de qualité du lien entre le dépôt d'énergie entre le ECAL et la trace en entrée d'un algorithme d'arbres de décision boosté.

Le candidat est considéré comme un électron si la sortie de l'algorithme est supérieure à une valeur donnée [39]. On retire ensuite les électrons reconstruits de la liste des blocs *particle-flow* disponibles pour la suite de l'algorithme.

2.2.1.7 Reconstruction des photons et hadrons neutres

A ce stade, il ne reste plus que les blocs *particle-flow* correspondants aux hadrons (chargés ou neutres), ainsi qu'aux photons. En préambule de la reconstruction, les liens créés par l'algorithme de liaison sont simplifiés. Si une trace est liée à plusieurs agrégats dans le HCAL, seul le lien ayant la plus petite distance (ΔR , voir 2.2.1.4) est gardé. On procède de façon identique pour les liens dans le ECAL.

L'énergie des agrégats calorimétriques liés à une trace peut être très inférieure à l'impulsion de celle-ci, dans moins de 0.3 % des cas. L'hypothèse de la présence d'un muon, ou d'une fausse trace est privilégié si cette différence est supérieure à 3σ . Un algorithme spécialisé de reconstruction des muons permet d'identifier ces muons. La trace est considérée comme fausse et est supprimée des blocs *particle-flow* s'il n'y pas de présence de muon. Chaque trace restante donne lieu à la création de hadrons chargés *particle-flow*, dont l'impulsion et l'énergie sont mesurées directement depuis la trace. La reconstruction des jets possède sa propre section 2.2.3, plus détaillée.

Finalement, il ne reste que les blocs *particle-flow* qui n'ont aucune trace liée. Les photons *particle-flow* sont reconstruit à partir des agrégats dans le ECAL et hadrons neutres

particle-flow avec les agrégats du HCAL.

Pour alléger les notations, les particules *particle-flow* seront simplement appelées particules.

2.2.2 ISOLATION DES LEPTONS ET DES PHOTONS

L'isolation d'un leptons (muons et électrons) et des photons est une variable importante dans une analyse. Une particule est dite isolée si dans un cône de rayon donné autour d'elle, l'activité détecteur est faible. L'isolation relative du lepton est définie par la relation

$$I_\ell = \frac{\sum p_T^{\text{hadrons chargés}} + \sum p_T^{\text{hadrons neutres}} + \sum p_T^{\text{photons}}}{p_T^\ell}$$

où les sommes portent sur les particules contenues dans un cône de rayon ΔR centré autour du lepton. Ainsi, plus I tend vers 0, plus la particule est isolée. Suivant la saveur du lepton différentes tailles de cônes sont utilisées :

- ▷ $\Delta R = 0.4$ pour les muons
- ▷ $\Delta R = 0.3$ pour les électrons

Pour le calcul de l'isolation, il faut considérer les particules issues de l'interaction principale seulement et éliminer celles provenant de l'empilement. La méthode, appelée CHS, élimine les hadrons chargés issus de l'empilement. Cependant les hadrons neutres étant plus difficilement traitables, on doit appliquer des corrections supplémentaires. Dans le cas des muons, on applique les corrections « $\Delta\beta$ » et les corrections de « surface effective» pour les électrons et les photons.

2.2.2.1 Corrections $\Delta\beta$

Une des conséquences de la symétrie d'isospin associée à l'interaction forte est l'apparition d'une relation, dans un premier temps empiriquement découverte, entre l'énergie combinée des hadrons neutres et des photons issus de l'empilement et la moitié de l'énergie totale des hadrons chargés du de l'empilement.

$$\sum p_T^{\text{hadrons neutres empilement}} + \sum p_T^{\text{photons empilement}} \simeq 0.5 \sum p_T^{\text{hadrons chargés empilement}} \quad (2.4)$$

En soustrayant donc de cette quantité du calcul d'isolation et en vérifiant la positivité de l'énergie de la partie neutre, on obtient :

$$I_\mu^{\text{corrige}} = \frac{\sum p_T^{\text{hadrons chargés}} + \overbrace{\left(\sum p_T^{\text{hadrons neutres}} + \sum p_T^{\text{photons}} - 0.5 \sum p_T^{\text{hadrons chargés pile-up}} \right)}^{=0 \text{ si négatif}}}{p_T^\ell}$$

On considère un muon comme isolé si $I_\mu^{\text{corrige}} < 15\%$.

2.2.2.2 Corrections à la surface effective

La stratégie, pour cette corrections, est tout d'abord de considérer la partie neutre de l'empilement correspondant à $\rho \times EA$ où ρ est la densité d'énergie par unité de surface due à l'empilement dans l'évènement calculé avec l'algorithme FastJet [43] et EA la surface effective :

$$EA = \frac{\alpha_{I,N_{\text{vtx}}}}{\alpha_{\rho,N_{\text{vtx}}}} \quad (2.5)$$

où $\alpha_{I,N_{\text{vtx}}}$ est la pente de l'isolation définie par l'équation (2.4) en fonction du nombre N_{vtx} de vertex primaires et de $\alpha_{\rho,N_{\text{vtx}}}$ la distribution de ρ en fonction du nombre de vertex N_{vtx} .

L'isolation corrigée est donc

$$I_e^{\text{corrigé}} = \frac{\sum p_T^{\text{hadrons chargés}} + \overbrace{\left(\sum p_T^{\text{hadrons neutres}} + \sum p_T^{\text{photons}} - \rho EA \right)}^{=0 \text{ si négatif}}}{p_T^\ell}$$

On considère un électron comme isolé si $I_e^{\text{corrigé}} < 6\%$. Cette valeur est plus faible que pour les muons en raison de la taille du cône plus faible également.

2.2.3 RECONSTRUCTION DES JETS

Le confinement lié à l'interaction forte implique que les quarks et gluons produits lors des collisions ne peuvent pas être observés directement. Lors de leur propagation, ils vont se combiner avec d'autres quarks pour former des hadrons. Cette succession d'hadronisations va induire la production d'une gerbe de hadrons. Formée principalement de hadrons, mais aussi de leptons, cette gerbe est appelée *jet*.

On dispose d'une description globale de l'événement, sous forme d'une liste de particules, grâce à l'algorithme de *particle-flow*. On va cependant utiliser des algorithmes spécialisés qui vont parcourir la liste des particules, et former des agrégats suivant certains critères de distance entre les particules, différents selon les algorithmes. A CMS, deux types d'algorithmes sont principalement utilisés : l'algorithme *anti- k_T* [44] et l'algorithme *Cambridge-Aachen* (C-A) [45]. L'algorithme utilisé par défaut par la collaboration CMS est *anti- k_T* . La figure Fig.2.5 montre la composition d'un jet reconstruit avec l'algorithme *anti- k_T* . Lorsque que l'on veut étudier la sous-structure au sein d'un jet, c'est le dernier que l'on va utiliser. Par exemple, dans le cas de la désintégration d'un boson W de très grande impulsion, les deux quarks peuvent être produits de façon très colinéaire, entraînant la reconstruction d'un seul gros jet, au lieu de deux.

En premier lieu, on supprime de la liste des particules les leptons isolés (voir 2.2.2) pour ne pas les agréger au sein d'un jet et produire un double comptage.

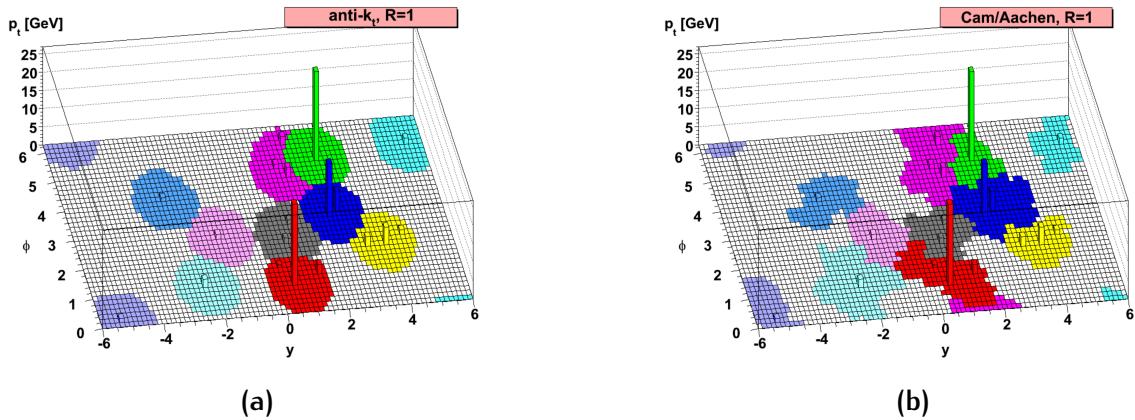


FIG. 2.4: Reconstruction des jets dans le plan (η, ϕ) effectuée sur le même événement par l'algorithme anti- k_T (a) et par l'algorithme C-A (b), pour une distance $R = 1$ [44].

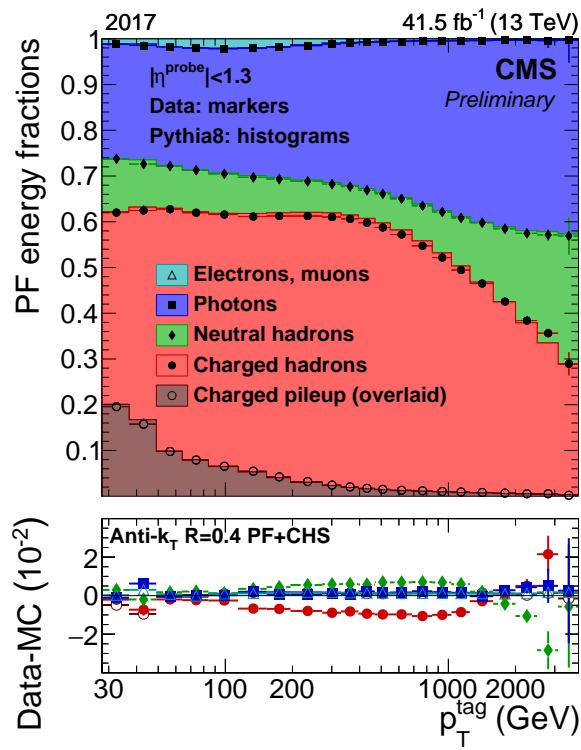


FIG. 2.5: Composition des jets reconstruits à l'aide de l'algorithme $anti-k_T$ en fonction de p_T du jet avec un $\Delta R = 0.4$ dans des évènements dijet sur l'année 2017.

Les deux algorithmes itèrent sur la collection de particules, et tentent de construire un jet en associant les particules deux à deux. Pour ce faire, ils utilisent les quantités suivantes :

$$d_{ij} = \min(k_{T,i}^n, k_{T,j}^n) \frac{\Delta R_{ij}^2}{R^2}$$

$$d_{iB} = k_{T,i}^n$$

où $k_{T,i}$ est l'impulsion transverse de la particule i par rapport à l'axe du faisceau, ΔR_{ij} la distance entre i et j dans le plan (η, ϕ) , et R une distance, choisie par l'utilisateur, symbolisant la largeur du jet. Pour l'algorithme *anti- k_T* , on a $n = -2$, tandis que pour C-A, on a $n = 1$. d_{iB} est un estimateur de la distance entre la particule i et le faisceau.

La première étape consiste à estimer la valeur minimum d_{min} entre tous les d_{ij} et d_{iB} . Si $d_{ij} = d_{min}$ alors les 4-impulsions des objets i et j sont sommées dans un nouvel objet. Par récursion, on complète la liste. Si le minimum est obtenu pour d_{iB} l'objet i est considéré comme un jet et est enlevé de la liste des particules. L'algorithme s'arrête lorsque toutes les particules au-dessus du seuil p_T minimum ont été agglomérées dans un jet.

La différence entre ces deux algorithmes réside dans la fonction de la distance entre particules à grouper. Pour *anti- k_T* , le poids de chaque paire est proportionnel à $\min(1/k_{T,i}^2, 1/k_{T,j}^2)$, ce qui revient à fusionner les particules de grandes impulsions les plus proches en premier. Pour C-A, le poids est uniquement proportionnel à la distance entre les paires : les particules les plus proches sont fusionnées en premier. On peut voir dans la figure Fig.2.4 un exemple de reconstruction des jets sur un même événement par les deux algorithmes.

2.2.3.1 Calibration des jets

Nos algorithmes de reconstruction constituent notre meilleure approximation expérimentale des quarks et des gluons. Il est possible que des particules créées lors de l'hadronisation ne soient pas correctement agglomérées dans le jet. Cela peut être dû au fait que ces particules dévient trop de la trajectoire initiale du quark. De plus, bien que les détecteurs soient étalonnés pour reconstruire au mieux l'énergie des particules, la calibration reste estimée en moyenne et ainsi elle impacte la mesure de l'énergie des jets. Les jets doivent donc être étalonnés, c'est-à-dire que leur énergie doit être évaluée au mieux pour correspondre aux simulations avant d'être utilisés dans des analyses physiques. La résolution en p_T des jets est étudiée après calibration des jets sur des événements dijet (voir Fig.2.5), $Z + \text{jets}$. et $\gamma + \text{jets}$. Des facteurs correctifs sont calculés et fournis centralement au sein de la collaboration, pour tenir compte des différences entre les données et les simulations observées sur les mesures de la résolution².

² La résolution des jets obtenus dans les simulations est meilleure que celle des jets reconstruits sur les données.

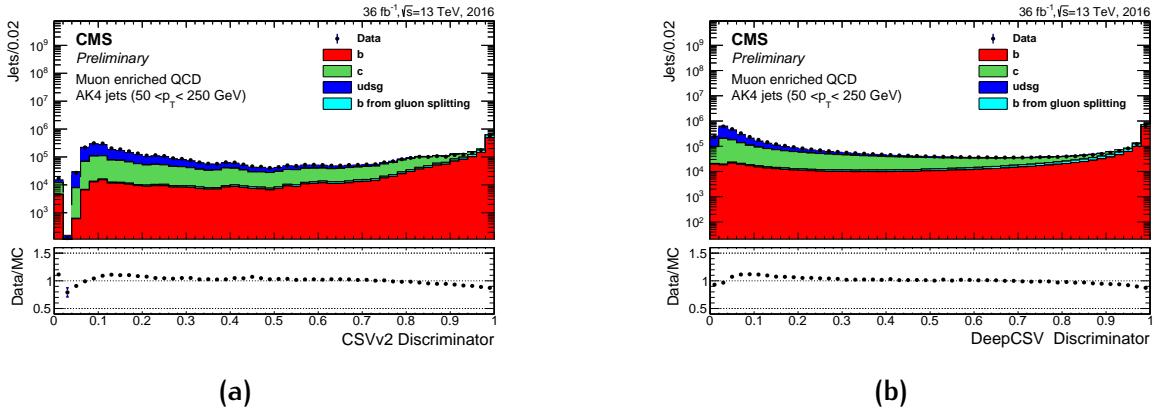


FIG. 2.6: Distribution des algorithmes de CSVv2 (a) et de DeepCSV (b).

2.2.4 ÉTIQUETAGE DES JETS ISSUS DE QUARKS b

A CMS, plusieurs algorithmes sont disponibles, permettant d'identifier les jets formés par un quark b (b -tagging, pour étiquetage des b). Nous allons présenter dans cette partie deux algorithmes.

L'algorithme CSVv2 (pour *Combined Secondary Vertex*) permet d'obtenir la probabilité qu'un jet soit de saveur lourde. Cet algorithme consiste en un réseau de neurones dit perception multicouche avec plusieurs couches cachées et un nombre de nœuds correspondant à deux fois le nombre de variables [46]. Les variables entraînées par le réseau de neurones sont liées au vertex secondaire. On peut noter la masse corrigée du vertex secondaire, nombre de traces liées au vertex secondaire, La liste exhaustive des variables est décrite [47]. Le réseau est entraîné sur des événements multijets. Ils seront ensuite répartis en trois catégories :

- ▷ Un ou plusieurs vertex secondaires ont été reconstruits.
- ▷ Un vertex secondaire n'a pas été reconstruit mais le jet possède au moins deux traces avec un paramètre d'impact 2D d'une signification suffisante ainsi qu'une masse invariante supérieure de 50 MeV à la masse du K_S^0 .
- ▷ Le reste des événements.

La figure FIG.2.6a, montre la distribution du discriminant permettant l'identification des jets issus de b . Elle est en moyenne plus proche de 1. Cette distribution pour les jets de c n'a pas de maxima clair et celle des jets légers est proche de 0. Cet algorithme performant permet une bonne discrimination des jets b par rapport aux saveurs légères mais moins par rapport à des jets de c .

Récemment les avancées de l'apprentissage profond (*deep learning* [48]) ont permis l'amélioration des algorithmes. L'algorithme DeepCSV est la version avancée du CSVv2. Cette version comporte quatre couches cachées de 100 nœuds interconnectés. L'entraînement est réalisé sur un mélange $t\bar{t}$ et multijet. L'architecture du réseau de neurone profond est écrite avec les librairies Kera [49] et Tensor-Flow [50]. Le discriminant utilisé pour les jets issus b est présenté à la figure FIG.2.6b.

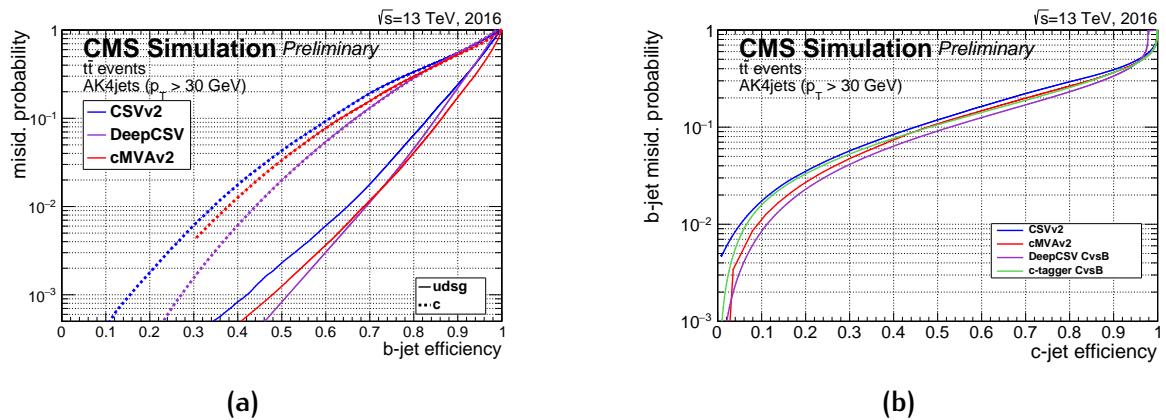


FIG. 2.7: Probabilité de mauvaises identifications en fonction de l'efficacité de sélection des jets de b (a), des jets de c (b) pour différents algorithmes [47].

La figure Fig.2.7a montre la probabilité d'une mauvaise identification d'un jet léger ou d'un c comme un jet de b en fonction de l'efficacité de sélection des jets b pour les différents algorithmes utilisés dans la collaboration. Cette figure permet de bien choisir son algorithme en fonction des besoins de sa propre analyse.

Les algorithmes d'identification sont entraînés sur des échantillons simulés pour obtenir les meilleures performances possibles et ne pas introduire de biais provenant de différences entre les données et les simulations. Les corrections sont extraites en utilisant la méthode dite de *tag and prob* [47] et un ajustement itératif [51]. Ces méthodes sont appliquées sur des événements de données Z + jets. Le jet dont la saveur est estimée à l'aide de l'algorithme est utilisé comme étiquette et les autres jets comme des sondes. La distribution du discriminant dans les simulations est normalisée à celle observée dans les données.

2.2.5 ÉNERGIE TRANSVERSE MANQUANTE

Par rapport au repère conventionnel de CMS, les collisions se déroulent le long de l'axe z . L'impulsion dans le plan (x, y) est donc nulle. Après reconstruction de l'événement, le bilan doit rester nul en vertu de la conservation de l'impulsion. Un bilan non nul signifie que des particules n'ont pas été correctement reconstruites. Les neutrinos interagissent très peu avec la matière, il ne sont par conséquent jamais détectés et font parties de ces particules non reconstruites.

Pour quantifier cette perte d'énergie, on introduit le vecteur d'énergie transverse manquante \cancel{p}_T et l'énergie transverse manquante (MET), sa norme, noté \cancel{E}_T . A l'aide de l'algorithme de *particle-flow*, on définit le vecteur énergie transverse manquante comme l'opposé de la somme vectorielle des impulsions transverses de toutes les particules de l'événement. La résolution sur la mesure de la \cancel{E}_T est très importante, puisque de nombreux modèles de nouvelle physique prédisent des particules qui interagissent peu voire pas avec la matière : l'énergie manquante est alors le seul moyen de pouvoir les mettre en évidence. Dans l'analyse de cette thèse, une résolution précise de la \cancel{E}_T

ne sera pas utilisée car le canal $t\bar{t} \rightarrow b\ell^+\nu_\ell + \bar{b}\ell^-\bar{\nu}_\ell$ a déjà un très bon rapport signal sur bruit de fond.

2.3 CONCLUSION

Dans ce chapitre nous avons vu qu'il faut mettre en place une chaîne de simulation complète et performante pour confronter les données aux prédictions théoriques. La première étape de la chaîne est la réalisation d'une simulation physique. Elle comprend les phénomènes de collisions entre partons, radiations de gluons, hadronisation, désintégration des particules instables, l'événement sous-jacent et les collisions multiples. De plus, une simulation précise du détecteur CMS est nécessaire pour simuler les réponses des différents sous-détecteurs aux particules produites.

Les signaux électriques fournis par le détecteur sont convertis en objets physiques. Le processus de reconstruction est commun aux simulations et aux données. Pour améliorer la résolution des différents objets reconstruits, on utilise l'algorithme *particle-flow* qui combine les informations fournies par les différents sous-détecteurs. On a également vu que plusieurs outils utilisent les résultats du *particle-flow*. Par exemple, l'algorithme de reconstruction des jets qui est central dans l'environnement hadronique qu'est le LHC. Les jets présentent une moins bonne résolution que les muons ou les photons, ainsi d'importantes méthodes de calibrations sont déployées pour une mesure de l'énergie plus précise.

BIBLIOGRAPHIE

- [1] O. S. Brüning, et al. *LHC Design Report*. CERN Yellow Reports : Monographs. CERN, Geneva, 2004. doi :10.5170/CERN-2004-003-V-1. URL <http://cds.cern.ch/record/782076>.
- [2] S. Chatrchyan et al. (CMS). «The CMS Experiment at the CERN LHC». *JINST*, 3 :S08004, 2008. doi :10.1088/1748-0221/3/08/S08004.
- [3] K. Aamodt et al. (ALICE Collaboration). «The ALICE experiment at the CERN LHC». *JINST*, 3 :S08002, 2008.
- [4] G. Aad, et al. «Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC». *Physics Letters B*, 716 :1–29, 2012. doi :10.1016/j.physletb.2012.08.020. [1207.7214](#).
- [5] S. Chatrchyan, et al. «Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC». *Physics Letters B*, 716 :30–61, 2012. doi :10.1016/j.physletb.2012.08.021. [1207.7235](#).
- [6] J. Alves, A. Augusto et al. (LHCb Collaboration). «The LHCb Detector at the LHC». *JINST*, 3 :S08005, 2008.
- [7] G. Aad et al. (ATLAS Collaboration). «Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC». *Phys.Lett.*, B716 :1–29, 2012.
- [8] S. Chatrchyan et al. (CMS Collaboration). «Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC». *Phys.Lett.*, B716 :30–61, 2012.
- [9] R. Aaij et al. (LHCb collaboration). «First observation of CP violation in the decays of B_s^0 mesons». *Phys.Rev.Lett.*, 110 :221601, 2013.
- [10] O. Adriani et al. (LHCf Collaboration). «The LHCf detector at the CERN Large Hadron Collider». *JINST*, 3 :S08006, 2008.
- [11] J. Pinfold, et al. (MoEDAL Collaboration). «Technical Design Report of the MoEDAL Experiment». Rapport technique CERN-LHCC-2009-006. MoEDAL-TDR-001, CERN, Geneva, 2009.
- [12] G. Anelli et al. (TOTEM Collaboration). «The TOTEM experiment at the CERN Large Hadron Collider». *JINST*, 3 :S08007, 2008.

- [13] A. Dominguez, et al. «CMS Technical Design Report for the Pixel Detector Upgrade». Rapport technique, 2012. Additional contacts : Jeffrey Spalding, Fermilab, Jeffrey.Spalding@cern.ch Didier Contardo, Universite Claude Bernard-Lyon I, didier.claude.contardo@cern.ch, URL <https://cds.cern.ch/record/1481838>.
- [14] T. T. G. of the CMS Collaboration. « The cms phase-1 pixel detector upgrade». 2020. [2012.14304](https://cds.cern.ch/record/14304).
- [15] T. C. Collaboration. « Description and performance of track and primary-vertex reconstruction with the CMS tracker». *Journal of Instrumentation*, 9(10) :P10009–P10009, 2014. doi :10.1088/1748-0221/9/10/p10009. URL <https://doi.org/10.1088/1748-0221/9/10/p10009>.
- [16] CMS Collaboration. «CMS ECAL Response to Laser Light». 2019. URL <https://cds.cern.ch/record/2668200>.
- [17] J. Mans, et al. «CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter». Rapport technique, 2012. Additional contact persons : Jeffrey Spalding, Fermilab, spalding@cern.ch, Didier Contardo, Universite Claude Bernard-Lyon I, contardo@cern.ch, URL <https://cds.cern.ch/record/1481837>.
- [18] V. Khachatryan, et al. (CMS Collaboration). «The CMS trigger system. The CMS trigger system». *JINST*, 12 :P01020. 122 p, 2016. doi :10.1088/1748-0221/12/01/P01020. Replaced with the published version. Added the journal reference and DOI. All the figures and tables can be found at <http://cms-results.web.cern.ch/cms-results/public-results/publications/TRG-12-001/index.html>, [1609.02366](https://cds.cern.ch/record/1609.02366), URL <https://cds.cern.ch/record/2212926>.
- [19] V. Balagura. « Van der meer scan luminosity measurement and beam-beam correction». *The European Physical Journal C*, 81(1), 2021. ISSN 1434-6052. doi :10.1140/epjc/s10052-021-08837-y. URL <http://dx.doi.org/10.1140/epjc/s10052-021-08837-y>.
- [20] CMS Collaboration. «CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV». Rapport technique, CERN, Geneva, 2019. URL <https://cds.cern.ch/record/2676164>.
- [21] J. Alwall, et al. «The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations». *Journal of High Energy Physics*, 7 :79, 2014. doi :10.1007/JHEP07(2014)079. [1405.0301](https://cds.cern.ch/record/1405.0301).
- [22] S. Alioli, et al. «A general framework for implementing NLO calculations in shower Monte Carlo programs : the POWHEG BOX». *JHEP*, 1006 :043, 2010.

- [23] F. Willeke et H. Germany. « Hera and the next generation of lepton-ion colliders». *EPAC 2006 - Contributions to the Proceedings*, 2006.
- [24] S. J. Brodsky et G. R. Farrar. «Scaling Laws at Large Transverse Momentum». *Phys. Rev. Lett.*, 31 :1153–1156, 1973. doi :10.1103/PhysRevLett.31.1153.
- [25] G. Altarelli et G. Parisi. «Asymptotic freedom in parton language». *Nuclear Physics B*, 126(2) :298–318, 1977. ISSN 0550-3213. doi :[https://doi.org/10.1016/0550-3213\(77\)90384-4](https://doi.org/10.1016/0550-3213(77)90384-4). URL <https://www.sciencedirect.com/science/article/pii/0550321377903844>.
- [26] Y. L. Dokshitzer. «Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics». *Sov. Phys. JETP*, 46 :641–653, 1977.
- [27] V. N. Gribov et L. N. Lipatov. «Deep inelastic e p scattering in perturbation theory». *Sov. J. Nucl. Phys.*, 15 :438–450, 1972.
- [28] J. Owens, et al. «Global parton distributions with nuclear and finite- Q^2 corrections». *Phys.Rev.*, D87(9) :094012, 2013.
- [29] R. D. Ball, et al. «Parton distributions for the lhc run ii». *Journal of High Energy Physics*, 2015(4) :40, 2015. ISSN 1029-8479. doi :10.1007/JHEP04(2015)040. URL [https://doi.org/10.1007/JHEP04\(2015\)040](https://doi.org/10.1007/JHEP04(2015)040).
- [30] J. Butterworth, et al. «Pdf4lhc recommendations for lhc run ii». *Journal of Physics G : Nuclear and Particle Physics*, 43(2) :023001, 2016. ISSN 1361-6471. doi :10.1088/0954-3899/43/2/023001. URL <http://dx.doi.org/10.1088/0954-3899/43/2/023001>.
- [31] L. Del Debbio. «Parton distributions in the LHC era». Dans «European Physical Journal Web of Conferences», tome 175 de *European Physical Journal Web of Conferences*, page 01006. 2018. doi :10.1051/epjconf/201817501006.
- [32] T. Sjostrand, et al. «PYTHIA 6.2 : Physics and manual». 2001.
- [33] G. Corcella, et al. «HERWIG 6 : An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)». *JHEP*, 0101 :010, 2001.
- [34] S. Agostinelli, et al. «Geant4 — a simulation toolkit». *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3) :250 – 303, 2003. ISSN 0168-9002.
- [35] S. Chatrchyan et al. (CMS Collaboration). «Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET». Rapport technique CMS-PAS-PFT-09-001, CERN, 2009. Geneva, 2009.

- [36] S. Chatrchyan et al. (CMS Collaboration). «Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector». Rapport technique CMS-PAS-PFT-10-001, Geneva, 2010.
- [37] S. Chatrchyan et al. (CMS Collaboration). «Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV». Rapport technique CMS-PAS-PFT-10-002, CERN, Geneva, 2010.
- [38] S. Chatrchyan et al. (CMS Collaboration). «Particle-flow commissioning with muons and electrons from J/ Ψ and W events at 7 TeV». Rapport technique CMS-PAS-PFT-10-003, CERN, Geneva, 2010.
- [39] F. Beaudette, et al. «Electron Reconstruction within the Particle Flow Algorithm (note privée)». Rapport technique CMS-NOTE-2010-034, CERN, Geneva, 2010.
- [40] W. Adam, et al. «Track Reconstruction in the CMS tracker». Rapport technique CMS-NOTE-2006-041, CERN, Geneva, 2006.
- [41] R. Frühwirth, et al. «Adaptive Vertex Fitting». Rapport technique CMS-NOTE-2007-008, CERN, Geneva, 2007.
- [42] S. Chatrchyan et al. (CMS Collaboration). «Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV». *JINST*, 7 :P10002, 2012.
- [43] M. Cacciari, et al. « Fastjet user manual». *The European Physical Journal C*, 72(3), 2012. ISSN 1434-6052. doi :10.1140/epjc/s10052-012-1896-2. URL <http://dx.doi.org/10.1140/epjc/s10052-012-1896-2>.
- [44] M. Cacciari, et al. « The anti- k t jet clustering algorithm». *Journal of High Energy Physics*, 2008(04) :063, 2008.
- [45] Y. Dokshitzer, et al. « Better jet clustering algorithms». *Journal of High Energy Physics*, 1997(08) :001, 1997.
- [46] W. S. Sarle. « Neural networks and statistical models». 1994.
- [47] A. Sirunyan, et al. « Identification of heavy-flavour jets with the cms detector in pp collisions at 13 tev». *Journal of Instrumentation*, 13(05) :P05011–P05011, 2018. ISSN 1748-0221. doi :10.1088/1748-0221/13/05/p05011. URL <http://dx.doi.org/10.1088/1748-0221/13/05/P05011>.
- [48] D. Guest, et al. « Jet flavor classification in high-energy physics with deep neural networks». *Physical Review D*, 94(11), 2016. ISSN 2470-0029. doi :10.1103/physrevd.94.112002. URL <http://dx.doi.org/10.1103/PhysRevD.94.112002>.
- [49] F. Chollet et al. « Keras». <https://github.com/fchollet/keras>, 2015.

- [50] M. Abadi, et al. « Tensorflow : Large-scale machine learning on heterogeneous distributed systems». 2016. [1603.04467](#).
- [51] V. Khachatryan, et al. « Search for the associated production of the higgs boson with a top-quark pair». *Journal of High Energy Physics*, 2014(9), 2014. ISSN 1029-8479. doi :10.1007/jhep09(2014)087. URL [http://dx.doi.org/10.1007/JHEP09\(2014\)087](http://dx.doi.org/10.1007/JHEP09(2014)087).