

# Chapitre X

## Calibration en énergie des jets

### Sommaire

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Formation des jets</b>	<b>2</b>
2.1	Gerbe partonique	2
2.2	Hadronisation	3
<b>3</b>	<b>Reconstruction des jets</b>	<b>4</b>
3.1	Algorithmes de regroupement	5
3.2	Identification des jets dans CMS	6
3.3	Saveur des jets	7
<b>4</b>	<b>Calibration en énergie des jets dans CMS</b>	<b>10</b>
4.1	Correction de l'empilement	12
4.2	Correction de la réponse du détecteur en $p_T$ et en $\eta$	14
4.3	Propagation à la MET	14
4.4	Corrections résiduelles	15
4.5	Incertitude sur la correction en énergie des jets	17
4.6	Correction de la résolution en énergie	18
<b>5</b>	<b>Phénoménologie des événements <math>\gamma + \text{jet}</math></b>	<b>19</b>
5.1	Principe des événements $\gamma + \text{jet}$ et réponse balancée	19
5.2	Effets radiatifs et activité additionnelle	20
5.3	Utilisation conjointe de la réponse MPF	22
<b>6</b>	<b>Correction résiduelle absolue en <math>p_T</math> avec les événements <math>\gamma + \text{jet}</math></b>	<b>22</b>
6.1	Sélection des événements	22
6.2	Analyse	22
6.3	Incertitudes	22
6.4	Résultats	22
<b>7</b>	<b>Correction de la résolution en énergie avec les événements <math>\gamma + \text{jet}</math></b>	<b>22</b>
7.1	Définition de la résolution en énergie des jets	22
7.2	Analyse	22
7.3	Incertitudes	22
7.4	Résultats	22
<b>8</b>	<b>Conclusion</b>	<b>23</b>

## 1 Introduction

Dans le chapitre sur le modèle standard, nous avons traité de l'interaction forte et noté que la constante de couplage de cette interaction,  $g_s$ <sup>1</sup>, décroît avec l'énergie. Ainsi se produit à haute énergie le phénomène de « liberté asymptotique ».

L'échelle d'énergie atteinte au LHC, de l'ordre du TeV, permet de réaliser des calculs perturbatifs car dans ce cas  $g_s \rightarrow 0$ . Toutefois, cela n'est vrai qu'au moment de la collision initiale entre les protons.

1. La constante de couplage pour l'interaction forte est souvent notée  $\alpha_s$ . Dans un souci de cohérence à travers ce manuscrit, nous conservons ici la notation  $g_s$  introduite dans le chapitre sur le modèle standard.

En effet, les quarks et gluons issus de cette collision, du fait de l'interaction forte à laquelle ils sont sensibles, vont radier d'autres quarks et gluons. Par conservation de l'énergie, l'énergie portée par une de ces particules diminue au fur et à mesure de ces radiations et  $g_s$  augmente. Arrivé à des énergies de l'ordre de la centaine de MeV, il n'est plus possible de réaliser des calculs perturbatifs.

Le phénomène de confinement de couleur réapparaît alors et les quarks et gluons sont ainsi confinés au sein de hadrons, dont la charge de couleur est nulle. Ce processus est appelé « hadronisation ». Les quarks et les gluons produits lors des collisions de haute énergie se manifestent ainsi sous la forme d'un flux collimé de particules stables, un « jet ».

Le LHC fait se collisionner des protons, aussi la collision a réellement lieu entre les constituants des protons, c'est-à-dire les quarks et les gluons. Lors des collisions de protons, il y a donc une forte probabilité d'interaction par interaction forte. Il en résulte une omniprésence des jets dans les analyses réalisées dans l'expérience CMS. Leur caractérisation est donc un point essentiel pour la collaboration.

Ce chapitre aborde la calibration en énergie des jets qui permet d'en obtenir une caractérisation la plus précise possible. Dans un premier temps, le processus de formation des jets est décrit d'un point de vue théorique. Ensuite, la méthode de reconstruction des jets dans le cadre de l'expérience CMS est présentée, ainsi que le principe de calibration de ces objets physiques. Dans un second temps, une des étapes de la calibration en énergie des jets est discutée plus en détails. La phénoménologie des événements utilisés est présentée, puis les analyses permettant d'obtenir les corrections en énergie et en résolution de l'énergie des jets sont exposées.

## 2 Formation des jets

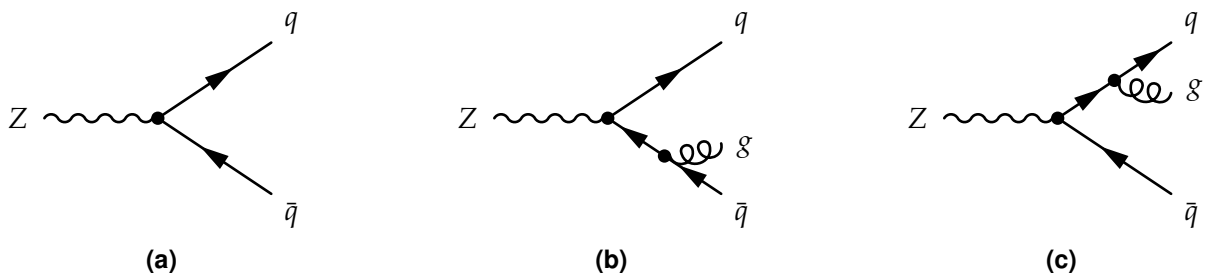
Lorsqu'une particule colorée, c'est-à-dire un quark ou un gluon, est issue de la collision, cette particule possède une haute énergie et  $g_s \ll 1$ . Cette particule colorée radie, par interaction forte, d'autres particules colorées. Par conservation, l'énergie portée par chaque particule colorée ainsi obtenue diminue et par conséquence,  $g_s$  augmente.

Tant que l'échelle d'énergie est suffisamment grande pour que  $g_s \ll 1$ , ce qui correspond à des énergies supérieures à la centaine de MeV, il est possible de réaliser des calculs perturbatifs. La radiation de particules colorées crée ce que s'appelle la « gerbe partonique », ce qui est le sujet de la prochaine section.

Au fur et à mesure des radiations, l'échelle en énergie diminue et en deçà d'une centaine de MeV, il n'est plus possible de réaliser des calculs perturbatifs car  $g_s$  augmente. Des modèles paramétriques sont alors utilisés pour caractériser le phénomène de « hadronisation », sujet de la section suivante.

### 2.1 Gerbe partonique

Lorsqu'une particule colorée est issue d'une collision au LHC, elle se trouve dans un premier temps dans le régime de liberté asymptotique. Elle radie alors d'autres particules colorées. Ainsi, pour un événement  $Z \rightarrow q\bar{q}$  comme celui de la figure X.1a avec deux quarks dans l'état final, il est possible d'obtenir par radiation d'un gluon un état  $q\bar{q}g$  comme ceux illustrés sur les figures X.1b et X.1c, par exemple.



**Figure X.1** – Un boson  $Z$  se désintègre en paire quark-antiquark. Dans les cas des figures X.1b et X.1c, un gluon supplémentaire est radié.

Il est légitime de se demander quelle est la probabilité d'obtenir un état  $q\bar{q}g$  à partir d'un état  $q\bar{q}$ . Des calculs de section efficace permettent d'obtenir [1], pour un état initialement à  $X$  partons dont un parton  $i$  radie un parton  $j$ ,

$$d\sigma_{X+j} \simeq \sigma_X \sum_{i \in \{X\}} \frac{g_s}{2\pi} \frac{d\theta^2}{\theta^2} dz P_{ij}(z) \quad (\text{X.1})$$

où  $\theta$  est l'angle entre le parton radié  $j$  et le parton radiant  $i$ . La grandeur  $P_{ij}(z)$  est la probabilité qu'un parton de type  $i$  radie un parton de type  $j$  emportant une fraction  $z$  de l'énergie initiale de  $i$ , qui s'exprime

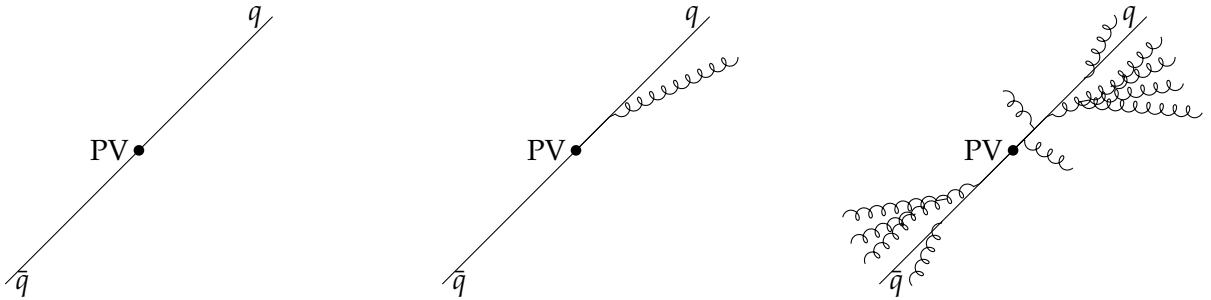
$$P_{qq}(z) = C_F \frac{1+z^2}{1-z}, \quad P_{qg}(z) = C_F \frac{1+(1-z)^2}{z}, \quad (\text{X.2})$$

$$P_{gg}(z) = C_A \frac{z^4 + 1 + (1-z)^4}{z(1-z)}, \quad P_{gq}(z) = T_R(z^2 + (1-z)^2), \quad (\text{X.3})$$

et  $P_{g\bar{q}}(z) = P_{gq}(z)$ , avec  $C_F = \frac{4}{3}$ ,  $C_A = 3$  et  $T_R = \frac{1}{2}$ . La probabilité de radier un parton supplémentaire diverge dans deux cas :

- le parton radié a une énergie faible devant celle du parton radiant, c'est la limite infrarouge ;
- l'angle entre le parton radié et le parton radiant est petit, c'est la limite colinéaire.

Les nouveaux partons ainsi radiés, et les partons initiaux, continuent chacun ce processus jusqu'à ce que le phénomène de confinement de couleur réapparaisse. Nous obtenons alors, pour un parton directement issu de la collision, une gerbe partonique, c'est-à-dire un ensemble collimé de particules colorées, ce qui est illustré sur la figure X.2. Ce sont ces particules qui vont participer au phénomène de hadronisation dû au confinement de couleur.



(a) Deux quarks sont initialement produits, ce qui correspond au diagramme de la figure X.1a.

(b) Un des quarks peut radier un gluon, ce qui correspond au diagramme de la figure X.1c.

(c) Le processus est réitéré, donnant un ensemble de particules colorées.

**Figure X.2** – Illustration de la formation de deux gerbes partoniques à partir d'une paire de quarks.

## 2.2 Hadronisation

Lorsque des particules coloréesradient d'autres particules colorées, la conservation de l'énergie implique que chaque particule, individuellement, perd de l'énergie. Dans le chapitre sur le modèle standard, nous avons vu que  $g_s$  augmente lorsque l'échelle d'énergie diminue. Lorsque l'échelle d'énergie passe en-deçà de quelques centaines de MeV,  $g_s$  diverge. Le phénomène de confinement de couleur réapparaît et la gerbe partonique, constituée de particules colorées, subit le phénomène de hadronisation et produit un flux collimé de hadrons, particules de charge de couleur nulle composées de particules colorées. Certains de ces hadrons peuvent comporter des quarks des deuxième ou troisième génération. Ils sont alors instables et peuvent être amenés à se désintégrer, auquel cas ce sont leurs produits de désintégration qui sont observés dans le détecteur.

Le phénomène de hadronisation ayant lieu lorsque  $g_s \gg 1$ , il n'est pas possible de réaliser des calculs perturbatifs. Afin de décrire ce phénomène, il faut donc avoir recours à des modèles paramétriques. Nous en décrivons ici deux, le modèle des cordes de Lund [2] et le modèle d'agglomération hadronique [3].

### 2.2.1 Modèle des cordes de Lund

Dans le modèle des cordes de Lund [2], les quarks sont reliés en paires  $q\bar{q}$  par des « cordes » de couleur, de tension  $\kappa \simeq 1 \text{ GeV}\cdot\text{fm}^{-1}$ , comme sur la figure X.3b. Les gluons sont décrits comme des nœuds des cordes de couleur.

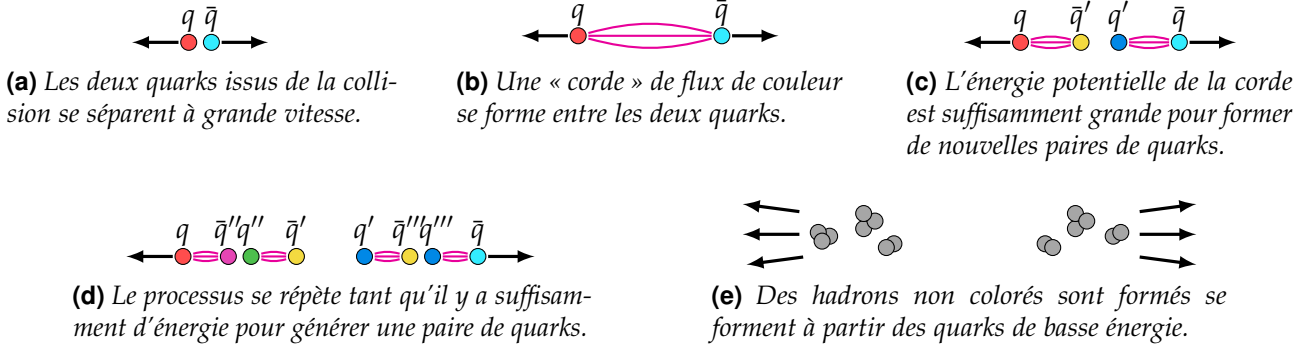


Figure X.3 – Processus de formation de deux jets dans le cadre du modèle des cordes de Lund.

Lorsque deux charges colorées s'éloignent, l'énergie potentielle augmente. Une fois que l'énergie potentielle est suffisamment grande, une nouvelle paire  $q'\bar{q}'$  est créée (fig. X.3c), avec une probabilité proportionnelle à  $\exp(-\frac{\pi}{\kappa}m_{q'})$ ; la probabilité d'obtenir des quarks lourds par ce processus est donc très faible. Le partage de l'énergie entre les paires de quarks est régie par une fonction de partition dont les paramètres sont estimés expérimentalement.

### 2.2.2 Modèle d'agglomération hadronique

Le modèle d'agglomération hadronique [3] repose sur l'hypothèse de conservation des nombres quantiques ainsi que de l'énergie-impulsion entre les partons issus de la gerbe hadronique et les hadrons obtenus après hadronisation.

Dans un premier temps, les gluons de la gerbe partonique se désintègrent en paires  $q\bar{q}$ . Les partons, uniquement des quarks donc, se rassemblent dans un second temps en agglomérats de charge de couleur nulle, c'est le *pré-confinement*. Deux cas de figure se présentent alors :

- la masse de l'agrégat est proche de celle d'un hadron, l'agrégat produit ce hadron;
- la masse de l'agrégat n'est pas proche de celle d'un hadron et son énergie est supérieure à un seuil  $Q_0$ , cet agrégat se désintègre en agrégats plus petits et forme plusieurs hadrons.

Ce processus est illustré sur la figure X.4.

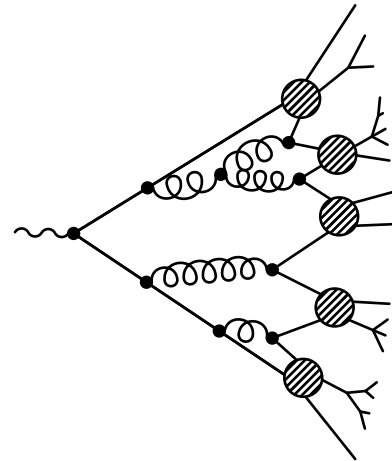


Figure X.4 – Schématisation de l'hadronisation dans le cadre du modèle d'agglomération hadronique.

## 3 Reconstruction des jets

Les particules colorées ne peuvent donc pas être directement observées dans le détecteur. Leur signature expérimentale est un flux collimé de particules stables composé de hadrons, de leptons

et de photons. La présence de hadrons s'explique directement par le processus de hadronisation décrit dans la section précédente. Les leptons proviennent de la désintégration, par interaction faible, des hadrons de saveur lourde, ou plus précisément des quarks de deuxième et troisième génération composant ces hadrons lourds. Les photons sont radiés par les particules électriquement chargées.

Un processus physique comme celui de la figure X.1a produit seulement quelques particules, en l'occurrence deux, et non des ensembles de particules, comme sur la figure X.4 qui pourrait correspondre à l'état effectivement observé pour le processus de la figure X.1a. Afin de pouvoir étudier le processus initial, il est nécessaire de définir une observable décrivant les particules colorées à l'origine de ces flux collimés de particules stables.

Cette observable est un « jet ». À partir des particules identifiées à l'aide de l'algorithme de *Particle Flow* (PF)<sup>2</sup>, un algorithme de regroupement permet d'obtenir la liste des jets de l'événement. Il existe plusieurs algorithmes de regroupement dont le principe est décrit dans la section suivante.

### 3.1 Algorithmes de regroupement

Il existe deux catégories d'algorithmes permettant de regrouper les particules en jets, les algorithmes de cônes et les algorithmes de recombinaison séquentielle. Dans la section 2, nous avons vu que les radiations de partons sont plus importantes pour de basses énergies (limite infrarouge) ou pour un parton radié colinéaire au parton initial (limite colinéaire). Afin de conserver des prédictions de QCD vérifiables sur des jets réels, les algorithmes de regroupement doivent être insensibles à l'ajout d'une particule de basse énergie ou au partage d'une particule en deux particules d'énergies inférieures. C'est ce que l'on appelle l'insensibilité IRC, pour *InfraRed and Colinear*. La plupart des algorithmes de cônes ne sont pas IRC-insensibles, alors que la plupart des algorithmes de recombinaison séquentielle le sont.

#### 3.1.1 Les algorithmes de cônes

Les algorithmes de cônes regroupent toutes les particules ayant une direction  $\vec{p}$  telle que la distance  $\Delta R_{pa}$  à la direction de l'axe du cône  $\vec{a}$  dans le plan  $(\eta, \phi)$ <sup>3</sup> est inférieure à une distance de coupure  $R_c$ , c'est-à-dire si

$$\Delta R_{pa}^2 = (\eta_p - \eta_a)^2 + (\phi_p - \phi_a)^2 < R_c^2. \quad (\text{X.4})$$

Alors, la direction  $\vec{a}$  du cône est redéfinie comme étant la direction moyenne de toutes les particules rassemblées dans ce cône. Ce processus est itéré jusqu'à la stabilisation des cônes. Enfin, les cônes sont séparés en cas de superposition, une particule ne pouvant appartenir qu'à un seul jet.

L'algorithme SIScone [4], *Seedless Infrared Safe Cone*, est un exemple d'algorithme de cônes IRC-insensible. Dans un premier temps, tous les cônes stables possibles sont reconstruits. Ces cônes sont alors fusionnés, les cônes ayant l'impulsion transverse la plus grande absorbant des cônes d'impulsion transverse moindre dont ils contiennent déjà une fraction. Un exemple de reconstruction de jets à l'aide de l'algorithme SIScone est présenté sur la figure X.5.

#### 3.1.2 Les algorithmes de recombinaison séquentielle

Les algorithmes de recombinaison séquentielle commencent par considérer que chaque particule forme un jet d'une seule particule. Puis, à l'aide d'une métrique donnée, la paire de jets les plus proches entre eux fusionne en un seul jet tant que la distance entre eux est en-deçà d'une valeur seuil. Les jets fusionnés donnent la liste des jets de l'événement. Il est également possible de fixer le nombre de jets à déterminer et non la valeur seuil de la distance entre les jets à fusionner.

Plusieurs métriques peuvent être définies, chacune correspondant à un algorithme de recombinaison séquentielle proposant des regroupement différents.

**Algorithme  $k_T$**  La distance  $d_{ij}$  entre deux jets  $i$  et  $j$  est

$$d_{ij} = \min(p_{Ti}^2, p_{Tj}^2) \frac{\Delta R_{ij}^2}{R^2}, \quad \Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2 \quad (\text{X.5})$$

2. L'algorithme de *Particle Flow* est décrit dans la section 4.1 du chapitre « Dispositif expérimental ».

3. Les coordonnées  $\eta$  et  $\phi$  sont définies dans la section 2.1 du chapitre « Dispositif expérimental ».



où  $p_{T_x}$  est l'impulsion transverse du jet  $x$  et  $R$  un paramètre libre. Cet algorithme a l'avantage de regrouper les particules très fidèlement vis-à-vis de la gerbe hadronique, mais les jets obtenus possèdent une géométrie spatiale irrégulière, comme cela se voit sur la figure X.5.

**Algorithme de Cambridge/Aachen** La distance  $d_{ij}$  entre deux jets  $i$  et  $j$  est

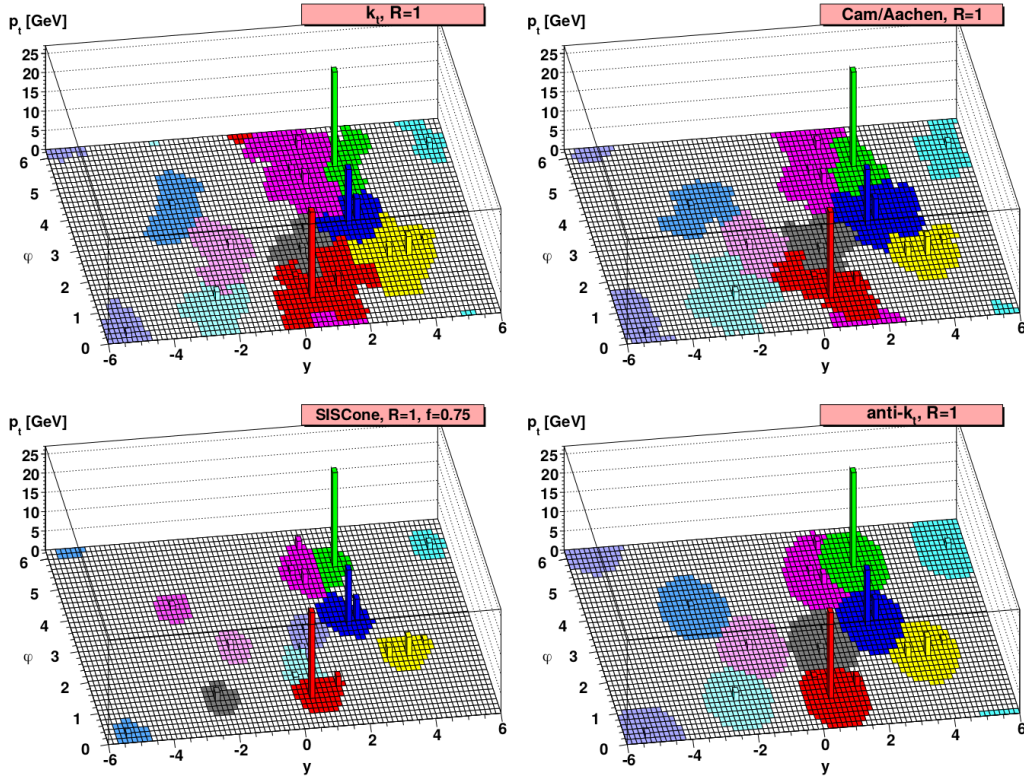
$$d_{ij} = \frac{\Delta R_{ij}^2}{R^2}, \quad \Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2 \quad (\text{X.6})$$

où  $R$  est un paramètre libre. Le regroupement des jets est ainsi uniquement basé sur l'écart angulaire.

**Algorithme anti- $k_T$**  [5] La distance  $d_{ij}$  entre deux jets  $i$  et  $j$  est

$$d_{ij} = \min \left( \frac{1}{p_{T_i}^2}, \frac{1}{p_{T_j}^2} \right) \frac{\Delta R_{ij}^2}{R^2}, \quad \Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2 \quad (\text{X.7})$$

où  $p_{T_x}$  est l'impulsion transverse du jet  $x$  et  $R$  un paramètre libre. Le regroupement des particules se fait ainsi autour des particules de plus haute énergie. Cet algorithme propose un regroupement des particules moins fidèle à la gerbe hadronique, mais produit des jets de forme régulière, comme cela se voit sur la figure X.5.

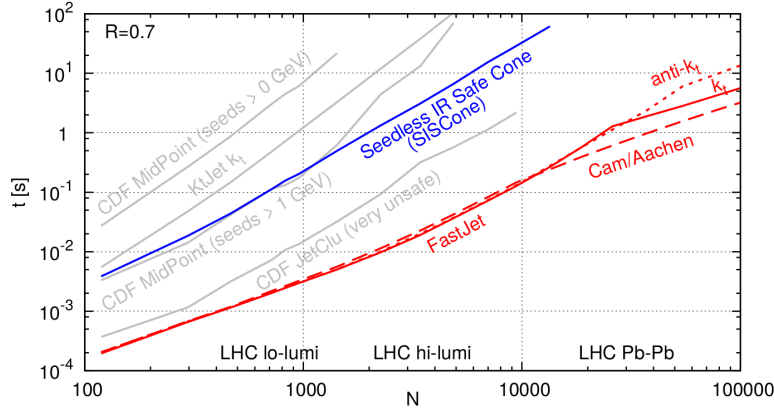


**Figure X.5** – Formes des jets reconstruits à partir de différents algorithmes pour un même événement [5]. En haut à gauche,  $k_T$ ; en haut à droite, C/A; en bas à gauche, SISCONE; en bas à droite, anti- $k_T$ . L'algorithme anti- $k_T$  permet d'obtenir des jets de forme régulière, conique.

Le temps de calcul de ces algorithmes est un enjeu majeur au LHC. Leurs temps d'exécution sont représentés en fonction du nombre d'événements d'empilement sur la figure X.6. L'algorithme anti- $k_T$  se place parmi les algorithmes les plus rapides. Dans les conditions des collisions proton-proton du LHC, il permet le traitement d'un événement en moins d'une milliseconde. C'est cet algorithme de regroupement qui est utilisé dans le cadre de l'expérience CMS.

### 3.2 Identification des jets dans CMS

Les jets ainsi reconstruits à l'aide des algorithmes de recombinaison sont en fait des « candidats » jets. À l'instar des particules individuelles, des critères d'identification leur sont appliqués afin de rejeter le bruit de fond et s'assurer de la qualité des jets ainsi reconstruits.



**Figure X.6** – Temps de recombinaison d’un événement dijet simulé de 50 GeV contenant  $N$  événements d’empilement pour différents algorithmes de reconstruction des jets [6].

Ces critères reposent sur les caractéristiques des candidats jets comme la fraction d’énergie provenant de leurs constituants neutres ou encore le nombre de ces constituants. Ces critères dépendent des années de prise de données et de la pseudo-rapidité du jet, c’est-à-dire de la région du détecteur dans laquelle il se trouve.

Les critères utilisés pour les années 2016, 2017 et 2018, listés page 8, permettent d’obtenir une efficacité d’identification des jets supérieure à 99 % dans chacune des régions en  $\eta$  du détecteur. La réjection du bruit de fond est supérieure à 98 % pour  $|\eta| \leq 3,0$  et supérieure à 36 % pour  $|\eta| > 3,0$ .

### 3.3 Saveur des jets

Pour étudier la physique du processus initial, la connaissance de la particule colorée à l’origine d’un jet ainsi identifié dans le détecteur est une information de choix. Il est impossible de connaître avec certitude cette particule, mais sa nature influe directement sur certaines propriétés des jets, permettant de l’estimer.

#### 3.3.1 Saveur de la particule initiale et caractéristiques des jets

**Le quark  $t$**  possède un temps de vie trop court pour participer à l’hadronisation. Il se désintègre alors par interaction faible en un autre quark, très majoritairement un quark  $b$ , et un boson  $W$ . Le nouveau quark issu de cette désintégration forme alors un jet.

Les autres quarks,  $d$ ,  $u$ ,  $s$ ,  $c$  et  $b$ , sont plus stables que le top et participent à l’hadronisation. Ils se retrouvent alors confinés au sein des hadrons formés.

**Le quark  $b$**  ne forme pas de hadron stable. Il se désintègre en quark  $c$  ou  $u$  selon

$$b \rightarrow cW^-, \quad \bar{b} \rightarrow \bar{c}W^+, \quad b \rightarrow uW^-, \quad \bar{b} \rightarrow \bar{u}W^+. \quad (\text{X.8})$$

Ces désintégrations font intervenir les modules des coefficients  $V_{cb}$  et  $V_{ub}$  de la matrice CKM<sup>4</sup> dont les valeurs sont faibles et sont donc fortement supprimées.

Les hadrons contenant un quark  $b$  ont ainsi une durée de vie  $\tau$  de l’ordre de la picoseconde [7, 8] et peuvent voyager sur une distance de l’ordre du millimètre. Les traces des particules chargées issues de cette nouvelle désintégration proviennent donc d’un vertex secondaire (SV), différent du vertex primaire (PV). Ces traces sont « déplacées ». Pour chacune d’entre elles, il est possible de déterminer le paramètre d’impact (IP) au vertex primaire, dont la valeur est typiquement plus grande que pour des traces provenant du vertex primaire, comme cela est illustré sur la figure X.7.

Le boson  $W$  issu du changement de saveur du quark  $b$  se désintègre leptoniquement dans 30 % des cas [9] selon

$$W^\pm \rightarrow \ell^\pm \nu_\ell. \quad (\text{X.9})$$

4. La matrice CKM est introduite dans la section 2.3.4 du chapitre « Particules, interactions et phénoménologie ».

Propriété du jet à identifier	$ \eta  \leq 2,4$	$2,4 <  \eta  \leq 2,7$	$2,7 <  \eta  \leq 3,0$	$3,0 <  \eta $
Fraction d'énergie				
— hadronique neutre	$< 0,90$	$< 0,90$	$< 0,98$	
— électromagnétique neutre	$< 0,90$	$< 0,90$	$> 0,01$	$< 0,90$
— hadronique chargée	$> 0$			
— électromagnétique chargée	$< 0,99$			
Nombre de constituants				
— neutres	$> 1$	$> 1$	$> 2$	$> 10$
— chargés	$> 0$			

**Tableau X.1** – Critères d'identification des jets à CMS pour l'analyse des données de 2016.

Propriété du jet à identifier	$ \eta  \leq 2,4$	$2,4 <  \eta  \leq 2,7$	$2,7 <  \eta  \leq 3,0$	$3,0 <  \eta $
Fraction d'énergie				
— hadronique neutre	$< 0,90$	$< 0,90$		$> 0,02$
— électromagnétique neutre	$< 0,90$	$< 0,90$	$< 0,99$ et $> 0,02$	$< 0,90$
— hadronique chargée	$> 0$			
Nombre de constituants				
— neutres	$> 1$	$> 1$	$> 2$	$> 10$
— chargés	$> 0$			

**Tableau X.2** – Critères d'identification des jets à CMS pour l'analyse des données de 2017.

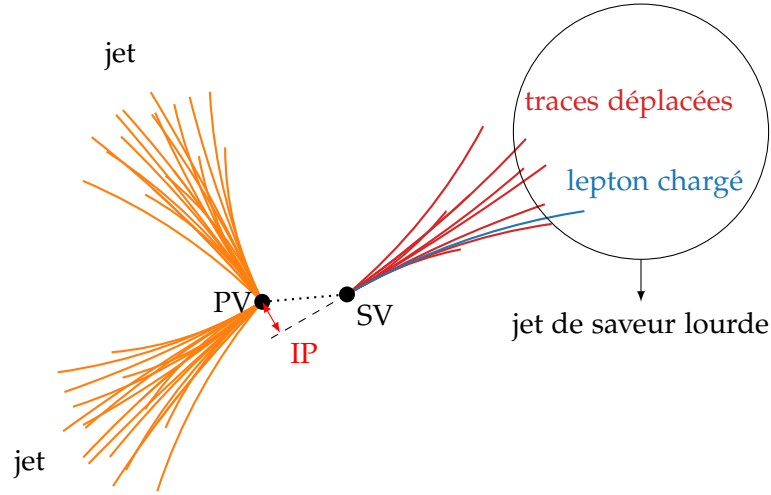
Propriété du jet à identifier	$ \eta  \leq 2,6$	$2,6 <  \eta  \leq 2,7$	$2,7 <  \eta  \leq 3,0$	$3,0 <  \eta  \leq 5,0$
Fraction d'énergie				
— hadronique neutre	$< 0,90$	$< 0,90$		$> 0,02$
— électromagnétique neutre	$< 0,90$	$< 0,99$	$< 0,99$ et $> 0,02$	$< 0,90$
— hadronique chargée	$> 0$			
Nombre de constituants				
— neutres	$> 1$		$> 2$	$> 10$
— chargés	$> 0$	$> 0$		

**Tableau X.3** – Critères d'identification des jets à CMS pour l'analyse des données de 2018.

Propriété du jet à identifier	$ \eta  \leq 2,6$	$2,6 <  \eta  \leq 2,7$	$2,7 <  \eta  \leq 3,0$	$3,0 <  \eta  \leq 5,0$
Fraction d'énergie				
— hadronique neutre	$< 0,90$	$< 0,90$		$> 0,02$
— électromagnétique neutre	$< 0,90$	$< 0,99$	$< 0,99$ et $> 0,02$	$< 0,90$
— hadronique chargée	$> 0$			
Nombre de constituants				
— neutres	$> 1$		$> 2$	$> 10$
— chargés	$> 0$	$> 0$		

**Tableau X.4** – Critères d'identification des jets à CMS pour l'analyse des données de 2017-UL.





**Figure X.7** – Illustration d’un jet de saveur lourde. Les particules composant ce jet proviennent d’un vertex secondaire (SV), différent du vertex primaire (PV) où a lieu la collision entre les protons et la formation du hadron lourd à l’origine du SV. Le paramètre d’impact (IP) est également indiqué.

Le lepton chargé donne une signature caractéristique lors des collisions proton-proton ayant lieu au LHC. Dans les autres cas, le  $W$  se désintègre hadroniquement selon

$$W^{\pm} \rightarrow q\bar{q}. \quad (\text{X.10})$$

Le nombre de constituants du jet, ainsi que le nombre de traces déplacées, est alors plus important.

**Le quark  $c$**  suit le même schéma que le quark  $b$ . Cependant, son canal de désintégration principal est

$$c \rightarrow sW^{+}, \quad \bar{c} \rightarrow \bar{s}W^{-}, \quad (\text{X.11})$$

ce qui fait intervenir le module du coefficient  $V_{cs}$  de la matrice CKM, proche de 1. Les hadrons contenant un quark  $c$  ont ainsi une durée de vie  $\tau$  de inférieure à la picoseconde [8] et il est plus difficile d’identifier les jets issus de quarks  $c$  que ceux issus de quarks  $b$ .

**Les quarks  $d$ ,  $u$  et  $s$**  forment des hadrons étant :

- très instables, par exemple les  $\pi^0$ , dont seuls les produits de désintégration sont observés ;
- faiblement instables, par exemple les  $K^+$ , qui peuvent se propager jusque dans les parties sensibles du détecteur et donc être directement observés ;
- stables, par exemple les protons, qui sont directement observés dans le détecteur.

Dans tous les cas, les traces des particules chargées observées proviennent du PV, lieu de formation du quark initial. Le phénomène décrit précédemment pour les quarks  $b$  et  $c$  n’est donc pas observable. Les jets issus de ces trois types de quarks, les plus légers, sont ainsi regroupés sous la dénomination de « jets légers ».

**Les gluons** portent une charge de couleur plus importante que les quarks<sup>5</sup>. Les jets initiés par des gluons comportent typiquement plus de particules électriquement chargées et sont moins collimés que les jets légers [10].

### 3.3.2 Identification de la saveur des jets

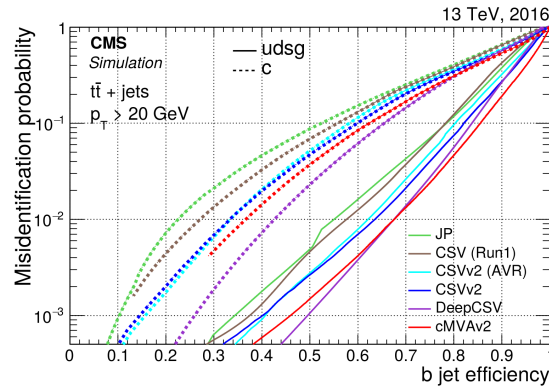
Les jets présentent ainsi des propriétés caractéristiques, selon qu’il s’agisse d’un jet léger (quarks  $d$ ,  $u$  ou  $s$ ), d’un jet lourd (quarks  $c$  ou  $b$ ) ou d’un jet issu d’un gluon. En utilisant ces propriétés, des algorithmes d’identification de la saveur des jets ont été mis au point par la collaboration CMS [11].

5. Les quarks portent une couleur, les antiquarks une anticouleur et les gluons portent une couleur et une anticouleur.

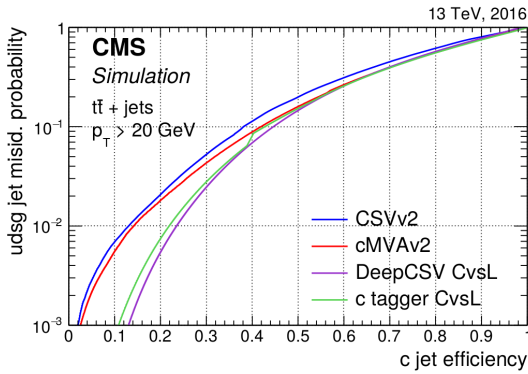
Les avancées récentes dans le domaine du *Deep Learning*, appliquées à l'identification des jets [12], ont permis l'amélioration de ces algorithmes. L'algorithme DEEPCSV [13] a ainsi été développé.

Les variables utilisées dans cet algorithme sont décrites dans la référence [13]. Elles sont traitées par un réseau de neurones profond de quatre couches cachées de 100 nœuds connectés les uns aux autres. Ce réseau est entraîné sur un ensemble d'événements simulés  $t\bar{t}$ , présentant de nombreux jets de quarks  $b$ , et multi-jets à l'aide des bibliothèques KERAS [14] et TENSORFLOW [15].

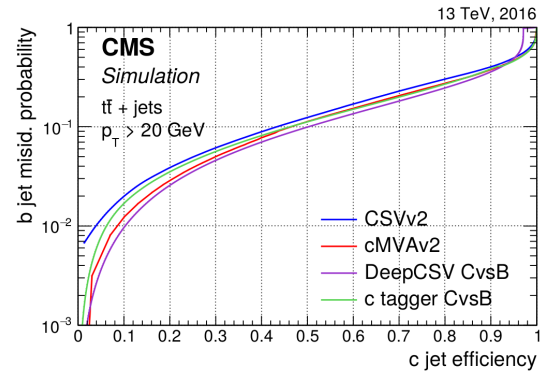
Les performances ainsi obtenues pour l'algorithme DEEPCSV peuvent être comparées à d'autres algorithmes d'identification de la saveur des jets sur la figure X.8. Les algorithmes CMVA2 et DEEPCSV présentent les meilleures performances en termes d'identification des jets de quark  $b$ . Pour le traitement des jets de quark  $c$ , l'algorithme DEEPCSV propose les meilleures performances. Dans les analyses présentées dans ce chapitre et le chapitre 5, c'est cet algorithme qui est utilisé afin d'identifier les jets issus de quarks  $c$  ou  $b$ .



(a) Probabilité de mauvaise identification en tant que jet de quark  $b$  de jets de gluon ou quarks légers (traits pleins) ou de jets de quark  $c$  (pointillés) en fonction de l'efficacité d'identification des jets de quark  $b$ .



(b) Probabilité de mauvaise identification en tant que jet de quark  $c$  de jets de gluon ou quarks légers en fonction de l'efficacité d'identification des jets de quark  $c$ .



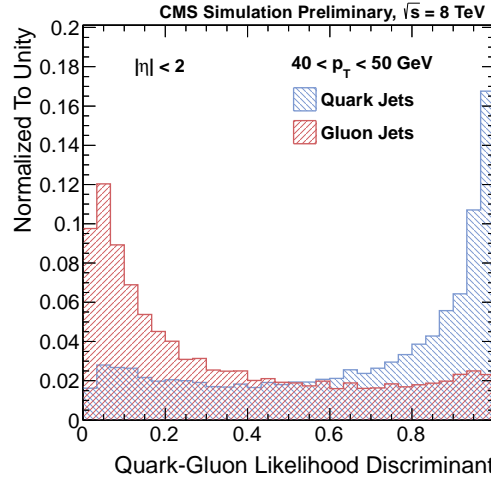
(c) Probabilité de mauvaise identification en tant que jet de quark  $b$  de jets de quark  $c$  en fonction de l'efficacité d'identification des jets de quark  $c$ .

Figure X.8 – Comparaison des performances des algorithmes d'identification de la saveur des jets [13].

La discrimination entre jet léger et jet initié par un gluon peut être réalisée à l'aide d'une fonction de vraisemblance [10] renvoyant un score entre 0 et 1 pour chaque jet, correspondant à la probabilité que ce jet soit issu d'un quark. La densité de probabilité de cette fonction, selon qu'il s'agisse de jets initiés par des gluons ou des quarks, est représentée sur la figure X.9.

## 4 Calibration en énergie des jets dans CMS

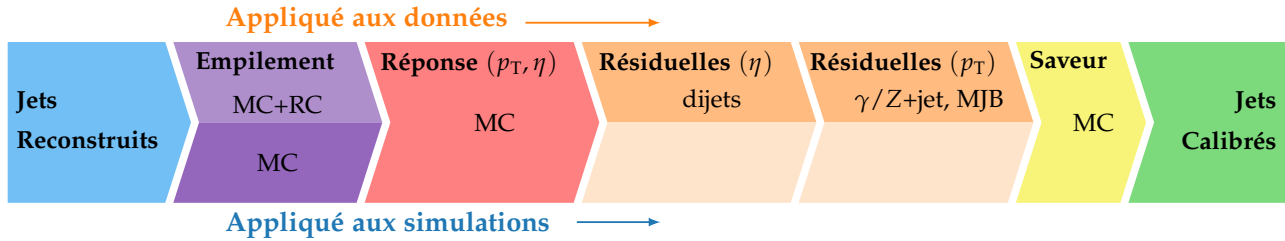
Les jets sont des objets composites, complexes, qu'il est nécessaire de calibrer comme tout autre objet reconstruit. La précision apportée à la mesure des jets est capitale dans de nombreuses analyses,



**Figure X.9** – Densité de probabilité de la fonction de vraisemblance utilisée pour discriminer les jets issus de gluons de ceux issus de quarks [10]. En rouge, pour les jets issus de gluons. En bleu, pour des jets issus de quarks.

où il s’agit d’une source majeure d’incertitude systématique. Les avancées réalisées récemment sur la calibration des jets ont ainsi permis d’améliorer la précision sur la mesure de la section efficace inclusive de production de jets et de la masse du quark  $t$  [16].

À partir des jets reconstruits par les méthodes décrites précédemment, un procédé de *correction de l’énergie des jets* (JEC<sup>6</sup>) est réalisé. Il permet de corriger l’échelle en énergie des jets (JES<sup>7</sup>) ainsi que la résolution sur cette énergie (JER<sup>8</sup>). La collaboration CMS utilise une approche factorisée dans laquelle plusieurs étapes corrigent un effet en particulier et dépendent des étapes précédentes [16]. La figure X.10 résume ces étapes, décrites plus en détails dans les sections qui suivent.



**Figure X.10** – Étapes successives de la JEC pour les données et les simulations [16]. Les corrections des étapes marquées « MC » sont obtenues par l’étude de simulations, celles marquées « RC » par une méthode de cône aléatoire (Random Cone). Les types d’événements utilisés dans les corrections résiduelles sont également indiqués.

Distinguons trois stades ou « niveaux » de connaissance sur les particules.

- Le niveau *particule*, noté  $p_{\text{tcl}}$ , ou niveau « vrai », se réfère aux objets et variables après hadronisation mais avant interaction avec le détecteur. Il s’agit donc des grandeurs recherchées, uniquement accessibles dans les événements simulés.
- Le niveau *reconstruit*, noté  $p_{\text{reco}}$ , correspond aux objets et variables après interaction avec le détecteur et reconstruction par l’algorithme de PF.
- Le niveau *corrigé* ou *calibré*, noté  $p_{\text{corr}}$ , correspond aux objets et variables corrigés, c’est-à-dire ceux du niveau reco auxquels ont été appliquées les corrections.

Définissons également une variable importante pour ce chapitre, la réponse d’un jet,

$$R = \frac{p_{\text{T}}}{p_{\text{T}_{\text{ptcl}}}}. \quad (\text{X.12})$$

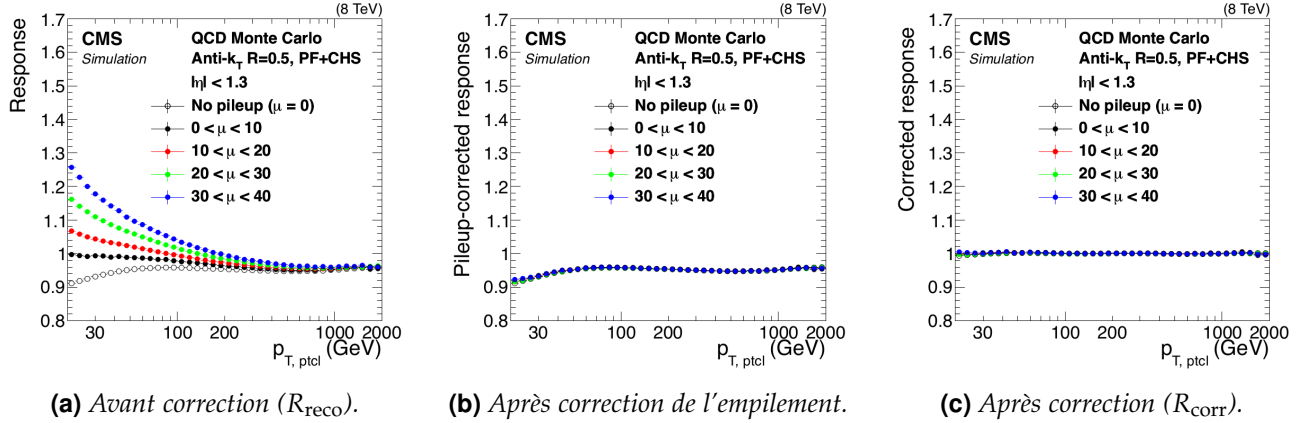
La réponse peut être définie à différents niveaux, et par définition  $R_{\text{ptcl}} = 1$ . Si la JEC est correcte, alors les variables corrigées doivent correspondre aux variables au niveau particule, c’est-à-dire  $R_{\text{corr}} = 1$ .

6. Jet Energy Correction

7. Jet Energy Scale

8. Jet Energy Resolution

Sur la figure X.11 sont représentées les réponses de jets d'événements QCD simulés à différentes étapes de la JEC. Après avoir appliqué toutes les corrections, ce qui correspond à la figure X.11c, la réponse est sensiblement égal à 1, ce qui montre que la JEC est correcte.



**Figure X.11** – Valeur moyenne de la réponse de jets d'événements QCD simulés en fonction de  $p_{T, \text{ptcl}}$  à différentes étapes de la JEC [16] et pour différentes valeurs d'interactions d'empilement  $\mu$ .

Les jets au niveau particule sont construits en appliquant la procédure de recombinaison à toutes les particules de durée de vie  $\tau$  telle que  $c\tau > 1$  cm à l'exception des neutrinos [16]. Les hadrons contenant des quarks  $c$  ou  $b$  ne rentrent pas dans cette catégorie et ce sont donc leurs produits de désintégration qui sont pris en compte pour la recombinaison. Exclure les neutrinos de la recombinaison au niveau particule est une convention adoptée par la collaboration CMS, mais pas de manière universelle en physique des particules. Les neutrinos sont en fait généralement inclus au niveau particule. La réponse des jets étant mesurée dans des événements contenant peu de neutrinos, comme cela est discuté dans la section 5, ce choix n'apporte pratiquement aucune différence à la JEC. L'intérêt de cette convention est de pouvoir définir la réponse des jets d'une manière qui soit accessible expérimentalement et qui réduit significativement les différences de réponse entre jets lourds et jets légers ou de gluons, à cause des neutrinos produits dans les désintégration des quarks lourds.

#### 4.1 Correction de l'empilement

Des contributions additionnelles à l'énergie et à l'impulsion des jets peuvent apparaître du fait de l'empilement<sup>9</sup>. La correction de l'empilement a pour but de soustraire ces contributions et est appliquée dans les données et les événements simulés. Elle permet d'améliorer la résolution du détecteur et d'obtenir une JES plus précise.

L'empilement asynchrone est réduit par l'analyse temporelle des signaux des calorimètres, l'empilement synchrone par la méthode de soustraction des hadrons chargés (CHS, *Charged Hadron Subtraction*), décrite ci-après.

Pour chacun des vertex primaires de l'événement, la somme des impulsions transverses au carré des traces associées au vertex est calculée. Le vertex primaire principal est choisi comme étant le vertex présentant la plus grande valeur de cette somme. Les autres vertex primaires sont considérés comme des vertex d'empilement. Toutes les traces associées aux vertex d'empilement sont retirées de l'événement. La reconstruction des jets est alors réalisée à partir de l'événement « nettoyé ». La procédure CHS permet ainsi de supprimer environ 50 % de l'empilement synchrone, uniquement à l'aide du trajectographe. La correction de l'empilement peut être calculée avec et sans utilisation de la CHS; la JEC est peu modifiée par ce choix. Cependant, l'utilisation de la CHS permet d'améliorer la résolution en  $p_T$  des jets. L'efficacité de reconstruction des vertex d'empilement étant de 30 %, des traces de hadrons chargés non associées à un vertex subsistent. De plus, cette méthode ne permet pas de corriger l'empilement des particules neutres.

La correction de l'empilement résiduel, principalement due aux particules neutres, aux traces non associées à un vertex et à l'empilement asynchrone qui n'a pas pu être corrigé totalement, est déter-

9. Le phénomène d'empilement est décrit dans la section 1.4 du chapitre « Dispositif expérimental ».

minée à l'aide de la méthode de l'aire hybride (*hybrid jet area*). Il s'agit d'une correction paramétrique, appliquée à chaque jet indépendamment, dépendant de :

- la densité en énergie dans le plan  $(\eta, \phi)$  de l'événement contenant ce jet,  $\rho$ ;
- l'aire du jet dans le plan  $(\eta, \phi)$ ,  $A_j$ ;
- la pseudo-rapidité du jet,  $\eta$ ;
- l'impulsion transverse du jet avant application de cette correction et après CHS,  $p_{T_{\text{reco}}}^{\text{CHS}}$ .

La correction  $C_{\text{PU}}$  à appliquer à un jet s'exprime alors

$$C_{\text{PU}}(p_{T_{\text{reco}}}^{\text{CHS}}, \eta, A_j, \rho) = 1 - \frac{[\rho_0(\eta) + \rho \beta(\eta)(1 + \gamma(\eta) \log p_{T_{\text{reco}}}^{\text{CHS}})] A_j}{p_{T_{\text{reco}}}^{\text{CHS}}} \quad (\text{X.13})$$

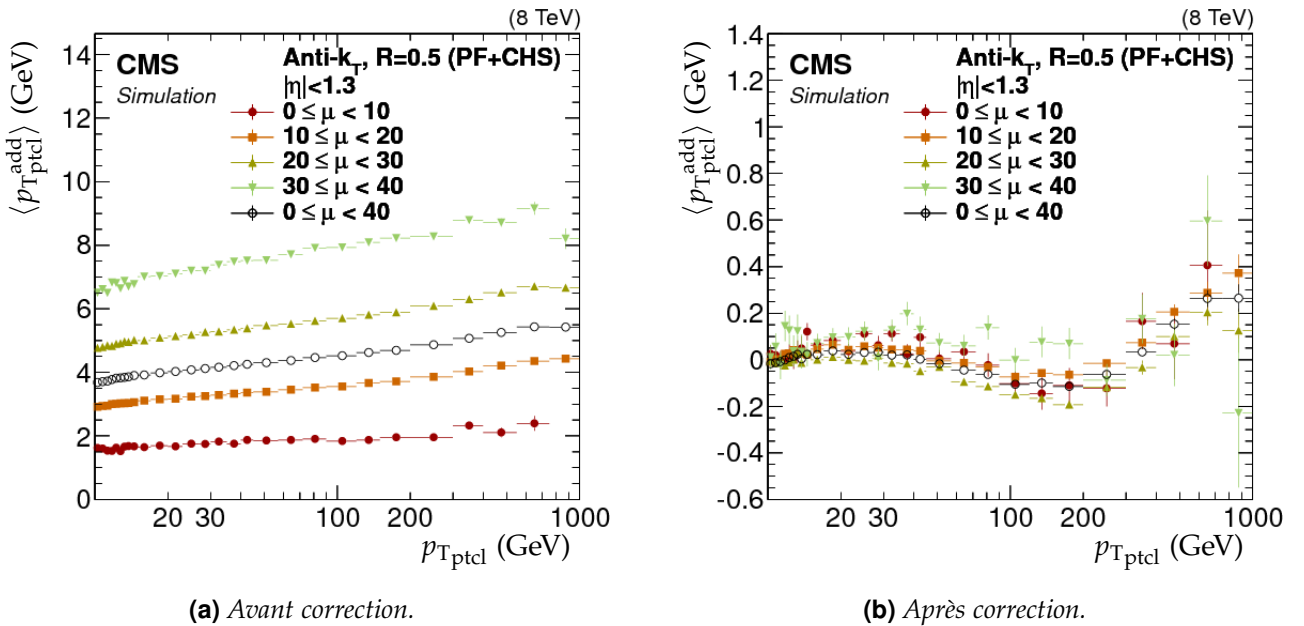
où  $\rho_0(\eta)$ ,  $\beta(\eta)$  et  $\gamma(\eta)$  sont les paramètres de cette correction, dépendants de  $\eta$ . Ils sont déterminés à partir de la contribution additionnelle de l'empilement au niveau particule  $p_{T_{\text{ptcl}}}^{\text{add}}$ , estimée à partir d'événements QCD multijet simulés avec et sans empilement, telle que

$$\langle p_{T_{\text{ptcl}}}^{\text{add}} \rangle(\rho, \eta, p_{T_{\text{reco}}}^{\text{CHS}}) = \langle p_{T_{\text{ptcl}}}^{\text{avec PU}} - p_{T_{\text{ptcl}}}^{\text{sans PU}} \rangle, \quad (\text{X.14})$$

avec  $p_{T_{\text{ptcl}}}^{\text{avec PU}}$  et  $p_{T_{\text{ptcl}}}^{\text{sans PU}}$  les impulsions du jet au niveau particule avec et sans empilement. La contribution additionnelle de l'empilement au niveau particule est alors paramétrée en fonction de  $\rho$ ,  $\eta$ ,  $p_{T_{\text{reco}}}^{\text{CHS}}$  et  $A_j$  afin d'obtenir les paramètres  $\rho_0(\eta)$ ,  $\beta(\eta)$  et  $\gamma(\eta)$  de l'équation (X.13) qui peut se réécrire

$$C_{\text{PU}}(p_{T_{\text{reco}}}^{\text{CHS}}, \eta, A_j, \rho) = 1 - \frac{\langle p_{T_{\text{ptcl}}}^{\text{add}} \rangle}{p_{T_{\text{reco}}}^{\text{CHS}}}. \quad (\text{X.15})$$

La figure X.12 montre  $\langle p_{T_{\text{ptcl}}}^{\text{add}} \rangle$  en fonction de l'impulsion transverse du jet au niveau particule, avant et après application de la correction de l'empilement. Les résultats de la figure X.12b sont cohérents avec l'absence d'énergie supplémentaire due à l'empilement à  $\pm 0,2$  GeV. Dans le cas d'un grand nombre d'interactions d'empilement ( $\mu > 30$ ), un léger effet est visible, lié à une dépendance quadratique en  $\rho$  de la contribution en énergie de l'empilement qui n'est pas modélisée [16].



**Figure X.12** – Contribution additionnelle de l'empilement au niveau particule telle que définie dans l'équation (X.14) pour  $|\eta| < 1,3$  en fonction de l'impulsion du jet au niveau particule pour différentes valeurs du nombre d'interaction d'empilement ( $\mu$ ) [16].

La correction ainsi décrite doit être légèrement adaptée pour pouvoir l'appliquer aux données à cause des biais de simulation du détecteur. Pour cela, un ajustement en fonction de  $\eta$  est déterminé à l'aide de la méthode de cône aléatoire (RC, *Random Cone*). La méthode RC reconstruit les jets à l'aide de cônes dont la direction en  $(\eta, \phi)$  est choisie de manière aléatoire. L'étude est réalisée sur des événements dits de « zéro biais », dans lesquels aucune énergie provenant d'une interaction dure<sup>10</sup>

10. définir ?



n'est présente. Dans ce cas, la valeur moyenne du  $p_T$  des jets reconstruits par la méthode RC permet d'estimer la moyenne de la contribution additionnelle de l'empilement, c'est-à-dire

$$\langle p_T^{\text{add}} \rangle^{\text{RC}} = \langle p_{T\text{cône}} \rangle. \quad (\text{X.16})$$

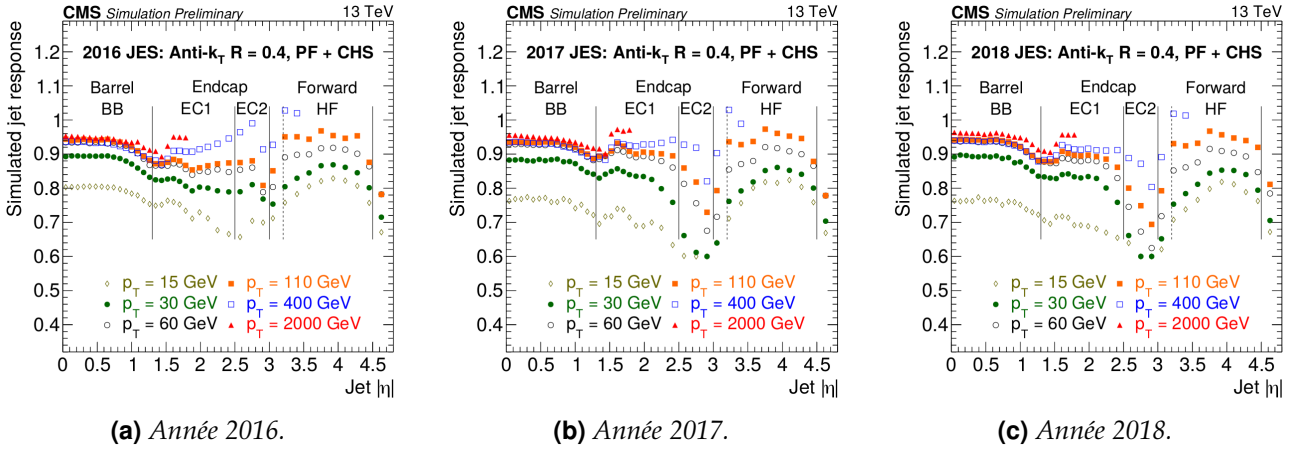
Il est alors possible de définir un facteur d'échelle à appliquer aux paramètres  $\rho_0$  et  $\beta$  de l'équation (X.13) lorsque cette correction est appliquée aux données. Ce facteur d'échelle s'exprime

$$\frac{\langle p_T^{\text{add}} \rangle_{\text{données}}^{\text{RC}}(\eta, \rho_{\text{données}})}{\langle p_T^{\text{add}} \rangle_{\text{simulation}}^{\text{RC}}(\eta, \rho_{\text{simulation}})}. \quad (\text{X.17})$$

La contribution additionnelle de l'empilement est ainsi corrigée dans les simulations et les données.

## 4.2 Correction de la réponse du détecteur en $p_T$ et en $\eta$

La réponse du détecteur CMS à un jet n'est pas uniforme selon la valeur de  $p_T$  et  $\eta$  du jet. La réponse au niveau reconstruit de jets simulés  $R_{\text{reco}}$ , déterminée grâce à une simulation du détecteur CMS basée sur GEANT4 [17], combinée à PYTHIA 6.4 [18] avec les réglages Z2\* [19], est représentée sur la figure X.13 pour les trois années du Run II du LHC. Il apparaît, par exemple, qu'un jet de  $p_T = 30$  GeV nécessite une correction allant de 10 % dans la région centrale  $|\eta| < 0,7$  à plus de 30 % lorsque  $|\eta| \simeq 3$  en 2017 et 2018.



**Figure X.13** – Réponse des jets reconstruits en fonction de  $p_T$  et  $\eta$  lors du Run II [20]. La baisse dans la région  $3,0 < |\eta| < 3,2$  (respectivement  $|\eta| > 4,5$ ) est due à la transition le bouchon (endcap) et la partie avancée (forward) du détecteur (reps. son acceptation). La dégradation au cours du temps du détecteur dans la région « EC2 » s'observe par la baisse de la réponse des jets dans cette région de 2016 à 2017.

Afin de corriger la réponse du détecteur en  $p_T$  et en  $\eta$ , la correction  $C_{\text{Rép}}$  à appliquer s'exprime

$$C_{\text{Rép}}(p_{T\text{reco}}', \eta) = \frac{\langle p_{T\text{ptcl}} \rangle}{\langle p_{T\text{reco}}' \rangle} = \frac{1}{\langle R'_{\text{reco}} \rangle} \quad (\text{X.18})$$

où  $p_{T\text{reco}}'$  est l'impulsion transverse du jet après correction de l'empilement. Les moyennes sont réalisées sur les jets appartenant à la même cellule d'une grille en  $(p_{T\text{ptcl}}, \eta)$  prédéfinie [16].

## 4.3 Propagation à la MET

L'impulsion transverse manquante  $\vec{E}_T^{\text{miss}}$  et sa norme l'énergie transverse manquante  $E_T^{\text{miss}}$ , définies dans la section 4.4 du chapitre « Dispositif expérimental », doivent prendre en compte les corrections apportées aux jets afin de garder une description cohérente de l'événement vis-à-vis de leurs définitions.

La JEC est ainsi propagée à  $\vec{E}_T^{\text{miss}}$  par la correction dite de « type-I »,

$$\vec{E}_{T \text{ type-I}}^{\text{miss}} = \vec{E}_{T \text{ reco}}^{\text{miss}} + \sum_{\substack{\text{jets} \\ p_{T\text{reco}} > 15 \text{ GeV}}} (\vec{p}_{T\text{reco}} - \vec{p}_{T\text{corr}}) - \vec{O}_{\text{RC}} \quad (\text{X.19})$$



où  $\vec{p}_{T\text{corr}}$  correspond à l'impulsion transverse du jet après correction de la réponse du détecteur et  $\vec{O}_{RC}$  la contribution moyenne de l'empilement obtenu par la méthode RC.

#### 4.4 Corrections résiduelles

Les corrections décrites dans les sections précédentes permettent d'obtenir une bonne correction en énergie des jets. Toutefois, des différences dans les réponses des jets, de l'ordre du pourcent, subsistent entre données et simulations. Des corrections résiduelles à appliquer aux données sont ainsi déterminées afin de réduire ces écarts, définies telles que

$$C_{\text{Res}} = \frac{R_{\text{simulations}}}{R_{\text{données}}} . \quad (\text{X.20})$$

Le principe est d'estimer la réponse du jet en s'appuyant sur un objet de référence pouvant être un boson Z (événements Z + jet), un photon (événements  $\gamma$  + jet) ou un autre jet (événements dijets et multijets). Deux méthodes existent et sont utilisées de manière complémentaire :

**la méthode de la balance** estime que l'objet de référence et le jet sont balancés au niveau particule, c'est-à-dire d'impulsion transverse totale nulle, soit

$$\vec{p}_{T\text{ptcl}}^{\text{réf}} + \vec{p}_{T\text{ptcl}}^{\text{jet}} = \vec{0} \Rightarrow p_{T\text{ptcl}}^{\text{jet}} = p_{T\text{ptcl}}^{\text{réf}} . \quad (\text{X.21})$$

Ainsi, au niveau reconstruit <sup>11</sup>,

$$\vec{p}_{T\text{reco}}^{\text{réf}} + \vec{p}_{T\text{reco}}^{\text{jet}} = \vec{0} \Leftrightarrow \vec{p}_{T\text{ptcl}}^{\text{réf}} + R_{\text{bal}} \vec{p}_{T\text{ptcl}}^{\text{jet}} = \vec{0} \quad (\text{X.22})$$

car l'objet de référence, fidèlement reconstruit, permet de considérer

$$\vec{p}_{T\text{reco}}^{\text{réf}} \simeq \vec{p}_{T\text{ptcl}}^{\text{réf}} = \vec{p}_T^{\text{réf}} . \quad (\text{X.23})$$

La réponse balancée  $R_{\text{bal}}$  est alors définie comme

$$R_{\text{bal}}(p_T, \eta) = \frac{p_{T\text{reco}}^{\text{jet}}}{p_T^{\text{réf}}} \quad (\text{X.24})$$

**la méthode « MPF »** (*MET Projection Fraction*) prend en compte l'ensemble de l'activité hadronique de l'événement et considère l'impulsion de recul vis-à-vis de l'objet de référence, c'est-à-dire

$$\vec{p}_{T\text{ptcl}}^{\text{réf}} + \vec{p}_{T\text{ptcl}}^{\text{reco}} = \vec{0} \Rightarrow \vec{p}_{T\text{reco}}^{\text{réf}} + \vec{p}_{T\text{reco}}^{\text{reco}} = -\vec{E}_T^{\text{miss}} \Leftrightarrow \vec{p}_{T\text{reco}}^{\text{réf}} + R_{\text{MPF}} \vec{p}_{T\text{ptcl}}^{\text{reco}} = -\vec{E}_T^{\text{miss}} . \quad (\text{X.25})$$

En appliquant (X.23) à l'équation précédente, il est possible d'écrire

$$\vec{p}_T^{\text{réf}} - R_{\text{MPF}} \vec{p}_T^{\text{réf}} = -\vec{E}_T^{\text{miss}} , \quad (\text{X.26})$$

ce qui permet de définir la réponse MPF  $R_{\text{MPF}}$  comme

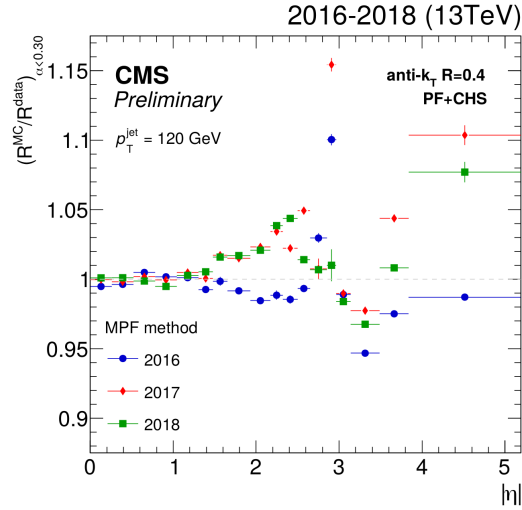
$$R_{\text{MPF}}(p_T, \eta) = 1 + \frac{\vec{p}_T^{\text{réf}} \cdot \vec{E}_T^{\text{miss}}}{|\vec{p}_T^{\text{réf}}|^2} . \quad (\text{X.27})$$

##### 4.4.1 Correction résiduelle relative en $\eta$

La première de ces corrections résiduelles, fonction de  $\eta$ , est obtenue à partir de la comparaison données-simulations sur une sélection d'événements dijets. Son but est de rendre indépendant de  $\eta$  le rapport données sur simulations de la réponse des jets. Cette correction s'appuie sur la bonne reconstruction des jets dans le barillet, c'est pourquoi elle est qualifiée de « relative ».

Lorsqu'un événement présente un premier jet avec  $|\eta| < 1,3$ , c'est-à-dire dans la région de référence du barillet, et un second avec  $|\eta| > 1,3$  et de  $p_T$  similaire, le premier sert d'objet de référence afin de calibrer le second. La correction à appliquer aux données ainsi obtenue est illustrée sur la figure X.14 dans le cas des jets d'impulsion transverse égale à 120 GeV.

11. Le niveau reconstruit prend ici en compte les étapes de correction de l'empilement et de la réponse du détecteur.



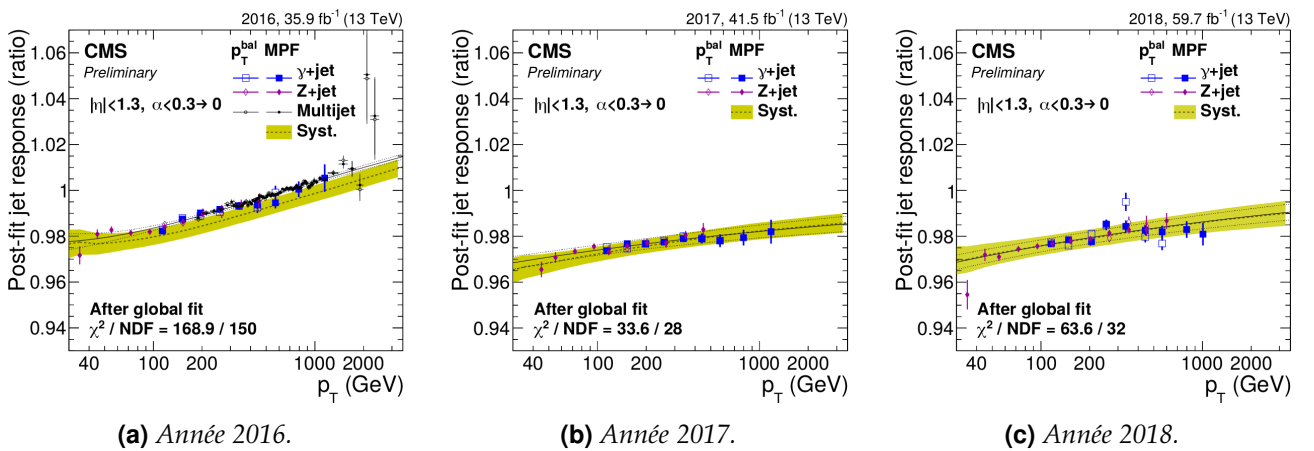
**Figure X.14** – Correction résiduelle relative en  $\eta$  lors du Run II obtenue avec des événements dijets et la méthode MPF [20].

#### 4.4.2 Correction résiduelle absolue en $p_T$

Cette correction, fonction de  $p_T$ , a pour but de rendre indépendant de  $p_T$  le rapport données sur simulations de la réponse des jets. Elle combine, à l'aide d'un ajustement global, les comparaisons données-simulations de plusieurs types d'événements afin de couvrir un large spectre de valeurs de  $p_T$ . Chaque type d'événement est en effet dominant, de par sa statistique, dans une gamme de  $p_T$  donnée :

- événements  $Z + \text{jet}$  : il s'agit d'événements  $Z(\rightarrow \mu^+\mu^-) + \text{jet}$  et  $Z(\rightarrow e^+e^-) + \text{jet}$ , sélectionnés par la présence d'une paire de muons ou d'électrons compatibles avec la désintégration d'un  $Z$ , ils couvrent la région  $p_T < 200$  GeV ;
- événements  $\gamma + \text{jet}$  : sélectionnés dans les données à l'aide d'un déclenchement basé sur la présence d'un photon, ils permettent de traiter la région  $200 \text{ GeV} < p_T < 800$  GeV ;
- événements multijets : ces événements contiennent au moins deux jets dans l'état final et couvrent la région  $p_T > 800$  GeV.

L'objet de référence utilisé pour calibrer le jet, que ce soit un boson  $Z$  ( $Z + \text{jet}$ ), un photon ( $\gamma + \text{jet}$ ) ou un autre jet (multijets), possède une meilleure résolution en énergie. Cette correction s'appuie donc sur un objet de référence qui n'est pas un jet, d'où son qualificatif d'« absolue ». La correction à appliquer aux données ainsi obtenue est illustrée sur la figure X.15 dans le cas des jets de pseudo-rapidité  $|\eta| < 1.3$ .



(a) Année 2016.

(b) Année 2017.

(c) Année 2018.

**Figure X.15** – Correction résiduelle absolue en  $p_T$  pour  $|\eta| < 1,3$  lors du Run II obtenue par ajustement global sur les événements  $\gamma + \text{jet}$ ,  $Z + \text{jet}$  et multijets [20].

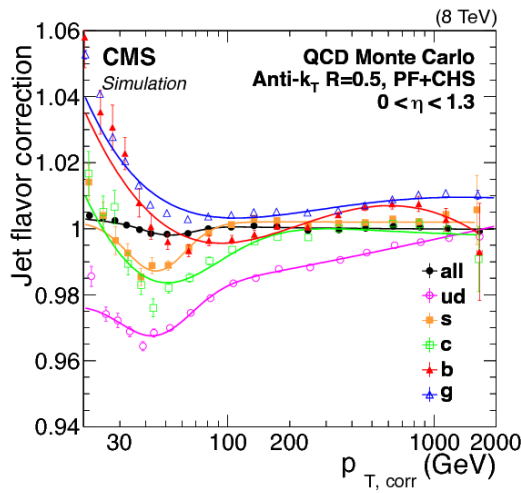
L'analyse des événements  $\gamma + \text{jet}$  utilisés dans ces ajustement globaux pour les années 2018, uti-

lisés dans la figure X.15c, et 2017-UL<sup>12</sup> a fait partie de mon travail de thèse. La phénoménologie de ces événements ainsi que leur analyse sont détaillées dans les sections 5 et 6.

#### 4.4.3 Correction résiduelle de saveur

Il existe une différence de réponse selon la saveur du jet, majoritairement due à la fragmentation en énergie et la composition du jet qui dépend de cette saveur [16]. Par exemple, des particules de bas  $p_T$  se retrouvent hors de la zone d'acceptation du détecteur. Or, des jets initiés par des gluons présentent de nombreuses particules de bas  $p_T$  par rapport aux jets issus de quarks légers. Dans une moindre mesure, les jets lourds possèdent également plus de particules de bas  $p_T$  que les jets de quarks légers suite à la désintégration du hadron lourd<sup>13</sup>. La proportion de particules neutres dans le jet est également un des paramètres affectant le plus sa réponse.

La correction  $C_{\text{Sav}}$  à appliquer aux données et aux simulations, obtenue à l'aide de PYTHIA 6.4 [18] avec les réglages Z2\* [19] sur des événements simulés dijet, Z + jet et  $\gamma$  + jet, est représentée sur la figure X.16. Des corrections allant jusqu'à 4 % sont appliquées.



**Figure X.16** – Correction résiduelle de saveur en fonction de l'impulsion du jet préalablement corrigée par les corrections décrites dans les sections précédentes,  $p_{T,\text{corr}}$ , pour des jets de pseudo-rapidité  $|\eta| < 1,3$  [16].

#### 4.5 Incertitude sur la correction en énergie des jets

Chacune des étapes de la JEC comporte des incertitudes liées aux effets systématiques et, dans une moindre mesure, statistiques. La maîtrise de ces incertitudes est un enjeu important pour de nombreuses analyses de la collaboration CMS où elles constituent une des sources d'incertitudes les plus importantes. Les sources d'incertitudes considérées pour la JEC sont regroupées en six groupes [20] :

**Échelle absolue** ou *Absolute scale* sur les légendes des figures X.17a à X.17f. Cette incertitude prend en compte l'échelle en énergie des objets de référence utilisés pour obtenir la correction résiduelle absolue en  $p_T$  décrite dans la section 4.4.2 ainsi que les effets de l'ISR et du FSR<sup>14</sup>.

**Échelle relative** ou *Relative scale* sur les légendes des figures X.17a à X.17f. Cette incertitude est principalement due à la JER qui s'applique à l'objet de référence dans la correction résiduelle relative en  $\eta$  décrite dans la section 4.4.1 ainsi que les effets de l'ISR et du FSR.

**Empilement** ou *Pileup* sur les légendes des figures X.17a à X.17f. Il s'agit de rendre compte de l'incertitude sur la détermination de la contribution additionnelle de l'empilement au niveau particule. Une incertitude de 5 % sur le rapport données sur simulations de cette contribution, obtenue à

12. La dénomination « UL » signifie *Ultra-Legacy* et correspond à une réinterprétation des données récoltées une fois que la collaboration peut prendre plus de recul sur l'obtention de celles-ci.

13. Le lecteur pourra se référer à la section 3.3 pour plus de détails sur la saveur des jets.

14. L'*Initial State Radiation* et le *Final State Radiation* sont abordés dans la section 5

l'aide de la méthode de cône aléatoire, est considérée. La différence résiduelle entre la contribution obtenue par cône aléatoire et la contribution réelle extraite des simulations est également prise en compte.

**Méthode et jeux de données** ou *Method & sample* sur les légendes des figures X.17a à X.17f. Cette incertitude correspond aux écarts observés entre les méthodes utilisant les réponses  $R_{bal}$  et  $R_{MPF}$  d'une part et entre les analyses utilisant les jeux de données  $Z + \text{jet}$ ,  $\gamma + \text{jet}$  et dijets d'autre part.

**Saveur** ou *Jet flavor* sur les légendes des figures X.17a à X.17f. L'incertitude sur la dépendance en saveur de la réponse des jets dans les simulations est estimée à partir de la différence entre deux générateurs, PYTHIA [18] et HERWIG [21].

**Stabilité temporelle** ou *Time stability* sur les légendes des figures X.17a à X.17f. La JEC est déterminée pour chaque période de prise de donnée chaque année. Les écarts observés entre ces périodes sont inclus dans cette source d'incertitude.

La figure X.17 résume les valeurs de ces incertitudes pour les trois années du Run II. L'incertitude globale sur la JEC est généralement inférieure à 2 %, excepté pour les cas  $p_T \leq 30$  GeV ou  $|\eta| \geq 2$  où elle peut être de l'ordre de 10 %.

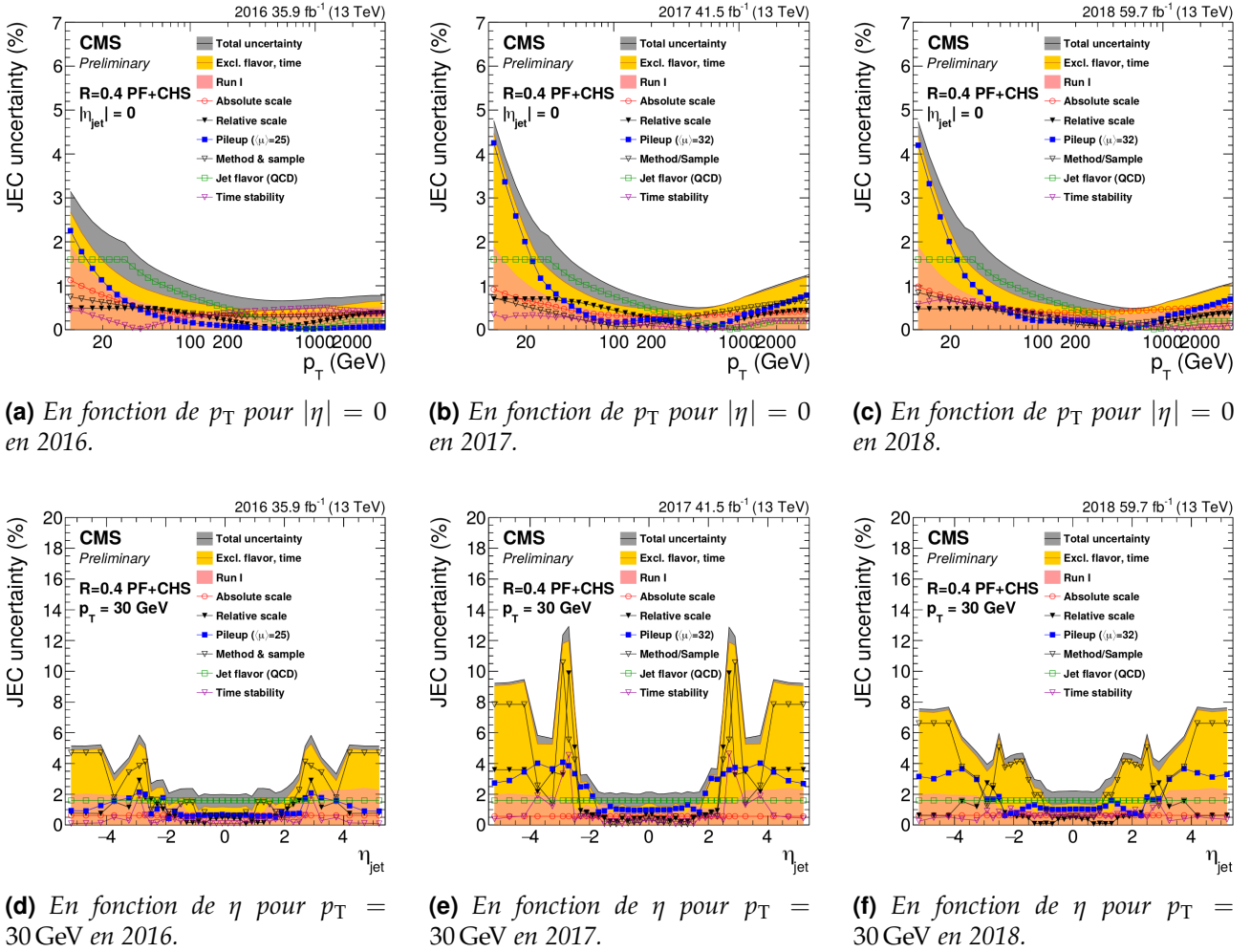


Figure X.17 – Incertitudes relatives sur la JEC lors du Run II [20].

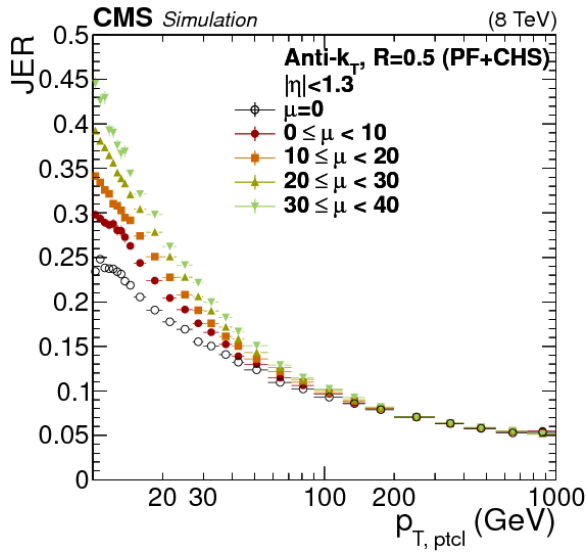
#### 4.6 Correction de la résolution en énergie

La résolution en énergie des jets, notée JER, est de l'ordre de 10 % pour des jets avec  $p_T \geq 50$  GeV et peut atteindre environ 40 % pour des jets de bas  $p_T$  et un empilement important [16]. Cette résolution est donc bien moins bonne que celles d'autres objets physiques tels que les électrons (2 à 5 %), les

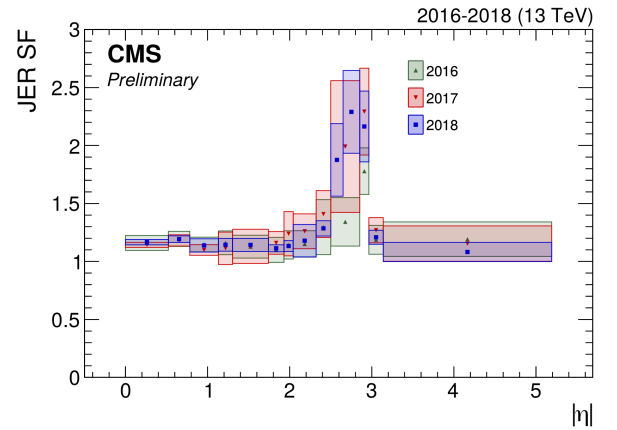
muons (1 à 6 %) et les photons (environ 1 %), ce qui peut introduire des biais importants dans les analyses cherchant des résonances étroites, par exemple.

La JER est définie comme la largeur de la gaussienne obtenue par un ajustement sur la distribution de  $R_{\text{corr}}$  des jets, c'est-à-dire  $p_{T,\text{corr}}/p_{T,\text{ptcl}}$ . Sa mesure est réalisée à l'aide d'événements  $\gamma$  + jet et Z + jet et les résultats obtenus lors du Run I sont présentés sur la figure X.18a. Elle est définie comme une fonction de  $p_{T,\text{ptcl}}$ ,  $\eta$  et  $\mu$ .

La JER observée dans les simulations diffère de celle observée dans les données, elle est légèrement meilleure. Afin de pouvoir réaliser des analyses comparant données et simulations, il est nécessaire d'avoir une JER comparable dans ces deux catégories d'événements. La JER des simulations est ainsi détériorée par un facteur d'échelle (*JER SF*), déterminé à partir d'événements  $\gamma$  + jet et dijet et défini en fonction de  $\eta$ . Les résultats obtenus lors du Run II sont présentés sur la figure X.18b. Le principe est le même que pour les corrections résiduelles décrites dans les sections 4.4.1 et 4.4.2. Au lieu de s'intéresser à la moyenne de la distribution, c'est sa largeur qui est étudiée.



(a) JER en fonction de  $p_T$  dans le barillet de CMS ( $|\eta| < 1.3$ ) pour différentes valeurs d'interactions d'empilement  $\mu$  lors du Run I [16].



(b) Facteurs d'échelle de la résolution en énergie des jets en fonction de  $\eta$  lors du Run II [20].

**Figure X.18** – Résolution en énergie des jets dans les simulations et facteurs d'échelle à leur appliquer.

L'analyse des événements  $\gamma$  + jet utilisés pour les années 2018 et 2017-UL<sup>15</sup> a fait partie de mon travail de thèse. La phénoménologie de ces événements ainsi que leur analyse sont détaillées dans les sections 5 et 7.

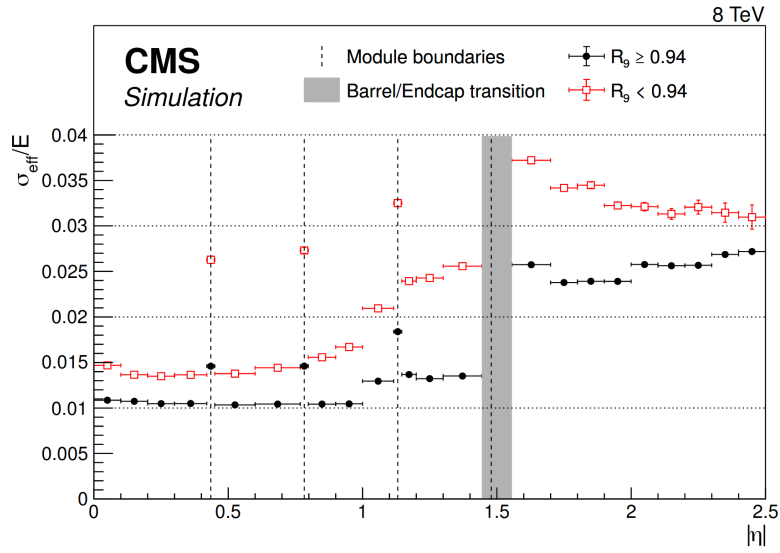
## 5 Phénoménologie des événements $\gamma$ + jet

Les événements  $\gamma$  + jet peuvent être utilisés afin d'obtenir la correction résiduelle absolue en  $p_T$  des jets, introduite dans la section 4.4.2, ainsi que la résolution en énergie des jets. Les analyses correspondantes sont abordées dans les sections 6 et 7.

### 5.1 Principe des événements $\gamma$ + jet et réponse balancée

L'état final d'un événement  $\gamma$  + jet comporte un jet à calibrer d'une part et un photon utilisé comme objet de référence d'autre part. En effet, les performances de reconstruction des photons sont meilleures que celles des jets. Sur la figure X.19, la résolution sur les photons est inférieure à 4 % et de l'ordre du pourcent dans le barillet. Dans le cas des jets, sur la figure X.18a, la résolution minimale est de 5 %. L'utilisation de photons comme objet de référence est donc justifiée.

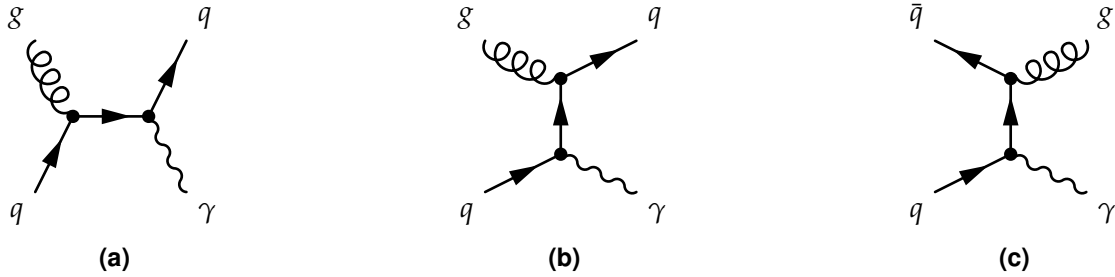
15. La dénomination « UL » signifie *Ultra-Legacy* et correspond à une réinterprétation des données récoltées une fois que la collaboration peut prendre plus de recul sur l'obtention de celles-ci.



**Figure X.19** – Résolution relative en énergie des photons en fonction de  $\eta$  pour des événements simulés  $h \rightarrow \gamma\gamma$  [22]. La variable  $R_9$  est définie dans la section ???.

Des diagrammes de Feynman correspondant à des événements  $\gamma + \text{jet}$  sont présentés sur la figure X.20. Ces événements ne comportent pas de neutrino issu de l'interaction dure<sup>16</sup>, il n'y a donc pas d'énergie transverse manquante due à la physique de ces événements. L'impulsion transverse étant nulle dans l'état initial, par conservation, elle est nulle dans l'état final. Le photon et le jet sont donc balancés, c'est-à-dire

$$\vec{p}_{T\text{ptcl}}^\gamma + \vec{p}_{T\text{ptcl}}^{\text{jet}} = \vec{0} \Rightarrow p_{T\text{ptcl}}^\gamma = p_{T\text{ptcl}}^{\text{jet}}. \quad (\text{X.28})$$



**Figure X.20** – Exemples de diagrammes de Feynman de processus physiques donnant un photon et un jet dans l'état final.

L'impulsion transverse du jet doit donc être égale à celle du photon, objet de référence. Leur ratio au niveau reconstruit doit donc être égal à 1. La bonne résolution en énergie sur les photons permet de considérer que leur impulsion transverse au niveau reconstruit est égale à leur impulsion transverse au niveau particule. Ainsi, la méthode de la balance<sup>17</sup> permet de définir

$$R_{bal} = \frac{p_{T\text{reco}}^{\text{jet}}}{p_T^\gamma}, \quad (\text{X.29})$$

qui doit valoir 1 après correction. Cette méthode est performante pour les événements à un photon et un jet dont la topologie est représentée sur la figure X.21a.

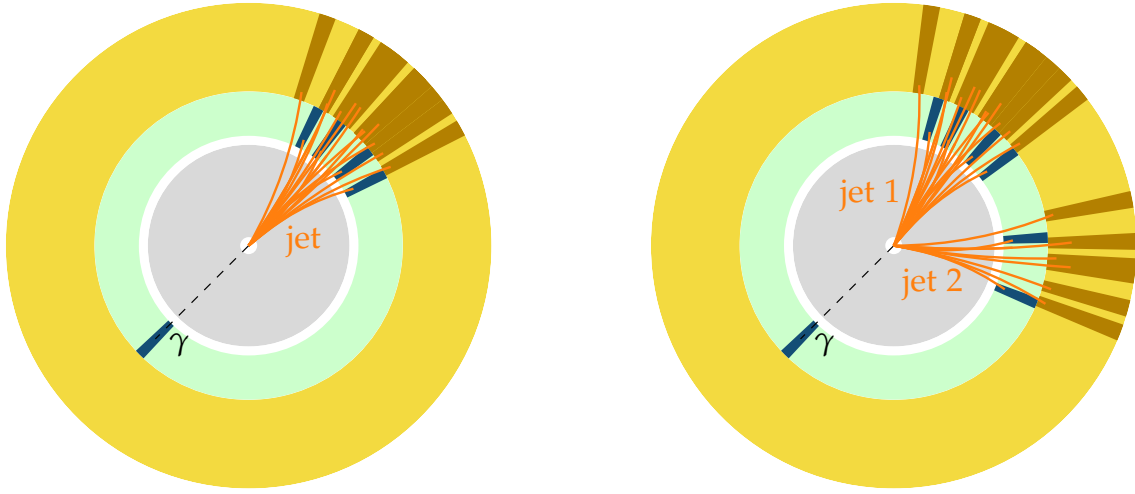
## 5.2 Effets radiatifs et activité additionnelle

Des effets radiatifs peuvent survenir et altérer la topologie des événements  $\gamma + \text{jet}$ . Un photon peut ainsi être radié dans l'état initial (ISR, *Initial State Radiation*) ou dans l'état final (FSR, *Final State*

16. Des neutrinos peuvent apparaître lors de la formation du jet.

17. La méthode de la balance est introduite dans la section 4.4.



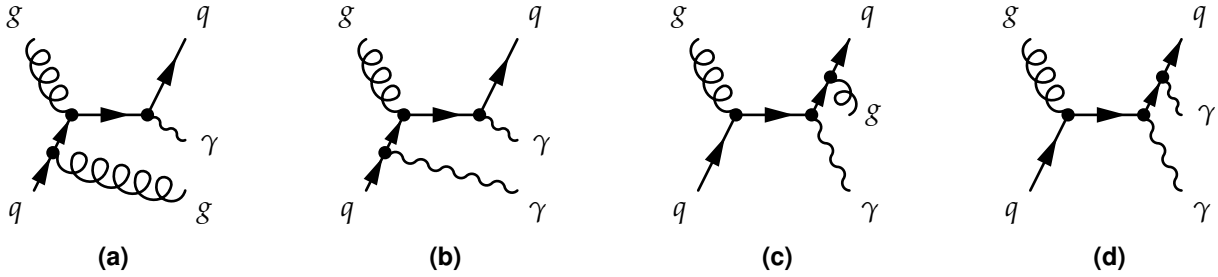


(a) Topologie typique des événements correspondant aux diagrammes de la figure X.20.

(b) Topologie typique des événements correspondant au diagramme de la figure X.22c.

**Figure X.21** – Topologies typiques des événements  $\gamma + \text{jet}$ .

Radiation), ce qui correspond aux diagrammes de Feynman des figures X.22b et X.22d. Un système composé d'un des photons et du jet n'est donc pas balancé dans ce cas. Il est possible de supprimer ce biais en imposant la présence d'un seul photon dans l'événement. La section efficace de production d'événements  $\gamma + \text{jet}$  à 13 TeV est importante [23], il est donc possible de sélectionner de manière stricte les événements afin d'obtenir une bonne pureté tout en conservant une statistique suffisante.



**Figure X.22** – Exemples de diagrammes de Feynman de processus avec ISR (X.22a, X.22b) ou FSR (X.22c, X.22d) donnant des événements avec deux jets (X.22a, X.22c) ou deux photons (X.22b, X.22d) dans l'état final.

L'ISR et la FSR peuvent aussi produire un gluon, ce qui correspond aux diagrammes de Feynman des figures X.22a et X.22c. Plusieurs jets sont alors présents dans l'état final et sont ordonnés par impulsion transverse décroissante. Le cas de la figure X.22a, correspondant à un jet additionnel par ISR, peut être supprimé par une condition sur les directions du photon et du premier jet qui doivent être opposées. Dans le cas du diagramme de la figure X.22c, correspondant à un jet additionnel par FSR, le photon est balancé avec le système des deux jets. La topologie d'un tel événement est illustrée sur la figure X.21b. La réponse balancée est alors considérée entre le photon et le premier jet,

$$R_{bal} = \frac{p_{T\text{reco}}^{\text{jet 1}}}{p_T^\gamma}. \quad (\text{X.30})$$

La présence d'un jet secondaire, comme sur la figure X.21b, crée un déséquilibre dans  $R_{bal}$  dû à la physique de l'événement et non à la JES. Il ne faut donc pas corriger cet effet. Pour cela, il faut pouvoir se ramener au cas où un seul jet est présent, comme dans l'événement de la figure X.21a. L'activité additionnelle liée aux jets supplémentaires est quantifiée par la variable

$$\alpha = \frac{p_{T\text{reco}}^{\text{jet 2}}}{p_T^\gamma}. \quad (\text{X.31})$$

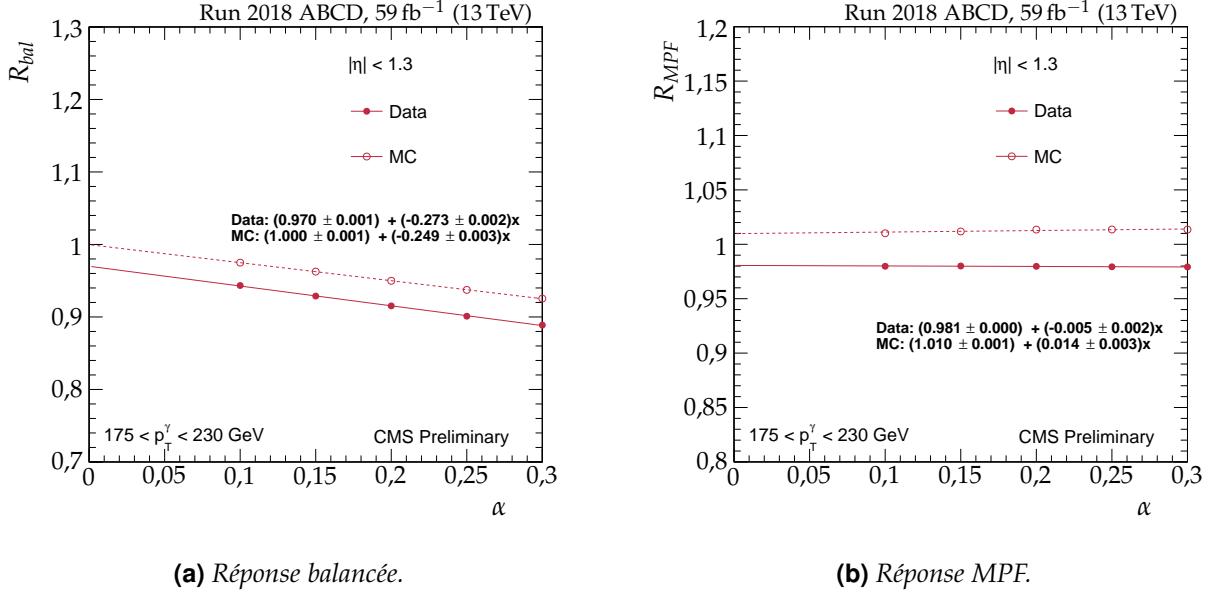
L'analyse des événements  $\gamma + \text{jets}$  est ainsi réalisée à différentes valeurs de  $\alpha$ , puis une extrapolation de  $R_{bal}$  à  $\alpha = 0$  permet d'obtenir le résultat souhaité. Cette procédure est détaillée dans la section 6.

### 5.3 Utilisation conjointe de la réponse MPF

En complément de la réponse balancée, la réponse MPF, définie comme

$$R_{MPF} = 1 + \frac{\vec{p}_T^\gamma \cdot \vec{E}_T^{\text{miss}}}{|\vec{p}_T^\gamma|^2}, \quad (\text{X.32})$$

est également analysée. L'impulsion de toutes les particules présentes étant considérée,  $R_{MPF}$  est moins sensible à l'activité additionnelle que  $R_{bal}$ , ce qui se retrouve sur les résultats de la figure X.23 où ces deux réponses sont représentées en fonction de  $\alpha$ .



**Figure X.23** – Réponses des jets en fonction de  $\alpha$  et extrapolation pour  $|\eta| < 1,3$  et  $175 < p_T^\gamma < 230$  GeV en 2018.

L'utilisation de la réponse MPF nécessite une bonne reconstruction de  $\vec{E}_T^{\text{miss}}$ , ce qui est le cas grâce aux bonnes performances de l'algorithme de PF. Son utilisation conjointe avec la réponse balancée permet d'obtenir des résultats complémentaires. Des écarts significatifs observés entre les deux méthodes indiqueraient ainsi des effets incompris, nécessitant de plus amples investigations. Dans la situation de la figure X.23 par exemple, les rapports des réponses balancée et MPF entre données et simulations valent  $0,970 \pm 0,001$  et  $0,971 \pm 0,001$ , ce qui est tout à fait compatible.

## 6 Correction résiduelle absolue en $p_T$ avec les événements $\gamma + \text{jet}$

### 6.1 Sélection des événements

### 6.2 Analyse

### 6.3 Incertitudes

### 6.4 Résultats

## 7 Correction de la résolution en énergie avec les événements $\gamma + \text{jet}$

### 7.1 Définition de la résolution en énergie des jets

### 7.2 Analyse

### 7.3 Incertitudes

### 7.4 Résultats

## 8 Conclusion

### Références

- [1] G. P. SALAM. *Elements of QCD for hadron colliders*. 2010. arXiv : 1011.5131 [hep-ph]. URL : <https://arxiv.org/pdf/1011.5131.pdf>.
- [2] B. ANDERSSON & coll. « Parton fragmentation and string dynamics » (avr. 1983). URL : <http://cds.cern.ch/record/143980>.
- [3] J.-C. WINTER, F. KRAUSS & G. SOFF. « A modified cluster-hadronisation model ». *The European Physical Journal C* **36.3** (août 2004), p. 381-395. DOI : 10.1140/epjc/s2004-01960-8. URL : <http://dx.doi.org/10.1140/epjc/s2004-01960-8>.
- [4] G. P. SALAM & G. SOYEZ. « A practical seedless infrared-safe cone jet algorithm ». *Journal of High Energy Physics* **05** (mai 2007), p. 86. DOI : 10.1088/1126-6708/2007/05/086. URL : <http://dx.doi.org/10.1088/1126-6708/2007/05/086>.
- [5] M. CACCIARI, G. P. SALAM & G. SOYEZ. « The Anti- $k_T$  jet clustering algorithm ». *Journal of High Energy Physics* **04** (avr. 2008), p. 63. DOI : 10.1088/1126-6708/2008/04/063. arXiv : 0802.1189 [hep-ph].
- [6] G. P. SALAM. « Towards jetography ». *The European Physical Journal C* **67.3-4** (mai 2010), p. 637-686. DOI : 10.1140/epjc/s10052-010-1314-6. URL : <http://dx.doi.org/10.1140/epjc/s10052-010-1314-6>.
- [7] The DØ Collaboration. « Measurement of the  $B_s^0$  lifetime in the exclusive decay channel  $B_s^0 \rightarrow J/\Psi \phi$  ». *Physical Review Letters* **94** (2005).
- [8] G. BELLINI, I. BIGI & P. DORNAN. « Lifetimes of charm and beauty hadrons ». *Physics Reports* **289.1** (1997), p. 1-155. DOI : [https://doi.org/10.1016/S0370-1573\(97\)00005-7](https://doi.org/10.1016/S0370-1573(97)00005-7). URL : <http://www.sciencedirect.com/science/article/pii/S0370157397000057>.
- [9] Particle Data Group. « Review of Particle Physics ». *Phys. Rev.* **D98** (août 2018). DOI : 10.1103/PhysRevD.98.030001.
- [10] The CMS Collaboration. *Performance of quark/gluon discrimination in 8 TeV pp data*. Rapp. tech. CMS-PAS-JME-13-002. Geneva : CERN, 2013. URL : <http://cds.cern.ch/record/1599732>.
- [11] The CMS Collaboration. « Identification of b-quark jets with the CMS experiment ». *Journal of Instrumentation* **8.04** (avr. 2013). DOI : 10.1088/1748-0221/8/04/p04013. URL : <http://dx.doi.org/10.1088/1748-0221/8/04/P04013>.
- [12] D. GUEST & coll. « Jet flavor classification in high-energy physics with deep neural networks ». *Physical Review D* **94.11** (déc. 2016). DOI : 10.1103/physrevd.94.112002. URL : <http://dx.doi.org/10.1103/PhysRevD.94.112002>.
- [13] The CMS Collaboration. « Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV ». *Journal of Instrumentation* **13.05** (mai 2018), P05011-P05011. DOI : 10.1088/1748-0221/13/05/p05011. URL : <http://dx.doi.org/10.1088/1748-0221/13/05/P05011>.
- [14] F. CHOLLET & coll. KERAS. <https://keras.io>. 2015.
- [15] M. ABADI & coll. TENSORFLOW : *Large-scale machine learning on heterogeneous distributed systems*. Software available from tensorflow.org. 2015. URL : <https://www.tensorflow.org/>.
- [16] The CMS Collaboration. « Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV ». *Journal of Instrumentation* **12.02** (fév. 2017). DOI : 10.1088/1748-0221/12/02/p02014. URL : <https://doi.org/10.1088/1748-0221/12/02/p02014>.
- [17] J. ALLISON & coll. « GEANT4 developments and applications ». *IEEE Transactions on Nuclear Science* **53.1** (fév. 2006), p. 270-278. DOI : 10.1109/tns.2006.869826.
- [18] T. SJÖSTRAND, S. MRENNA & P. SKANDS. « PYTHIA 6.4 physics and manual ». *Journal of High Energy Physics* **05** (mai 2006). DOI : 10.1088/1126-6708/2006/05/026. URL : <http://dx.doi.org/10.1088/1126-6708/2006/05/026>.

- [19] The CMS Collaboration. « Event generator tunes obtained from underlying event and multiparton scattering measurements ». *European Physical Journal C* **76.3** (2016), p. 155. DOI : [10.1140/epjc/s10052-016-3988-x](https://doi.org/10.1140/epjc/s10052-016-3988-x). arXiv : [1512.00815](https://arxiv.org/abs/1512.00815) [hep-ex].
- [20] The CMS Collaboration. « Jet energy scale and resolution performance with 13 TeV data collected by CMS in 2016-2018 » (avr. 2020). URL : <https://cds.cern.ch/record/2715872>.
- [21] J. BELLM & coll. « HERWIG 7.0/HERWIG++ 3.0 release note ». *The European Physical Journal C* **76.196** (avr. 2016). DOI : [10.1140/epjc/s10052-016-4018-8](https://doi.org/10.1140/epjc/s10052-016-4018-8).
- [22] CMS Collaboration. « Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at  $\sqrt{s} = 8$  TeV ». *Journal of Instrumentation* **10** (fév. 2015). DOI : [10.1088/1748-0221/10/08/P08010](https://doi.org/10.1088/1748-0221/10/08/P08010). URL : <https://cds.cern.ch/record/1988093>.
- [23] CMS Collaboration. « Measurement of differential cross sections for inclusive isolated-photon and photon+jets production in proton-proton collisions at  $\sqrt{s} = 13$  TeV ». *European Physical Journal C* **79.20** (juil. 2018). DOI : [10.1140/epjc/s10052-018-6482-9](https://doi.org/10.1140/epjc/s10052-018-6482-9). URL : <http://cds.cern.ch/record/2628267>.

