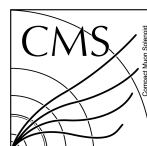

Une méthode de calibration non paramétrique pour les calorimètres de CMS.

Résumé :

DANS LE DÉTECTEUR CMS, L'ÉNERGIE DES HADRONS NEUTRES EST DÉTERMINÉE À PARTIR DE L'ÉNERGIE MESURÉE DANS LES CALORIMÈTRES ÉLECTROMAGNÉTIQUES (E_{ecal}) ET HADRINIQUES (E_{hcal}). UNE CALIBRATION EST CEPENDANT NÉCESSAIRE POUR ESTIMER L'ÉNERGIE VRAIE DU HADRON NEUTRE À PARTIR DE E_{ecal} ET E_{hcal} . DANS UN PREMIER TEMPS, J'AI UTILISÉ COMME CALIBRATION UNE FONCTION LINÉAIRE DE E_{ecal} ET E_{hcal} . ENSUITE, AFIN DE DÉCRIRE LA NON LINÉARITÉ DE LA MESURE DE L'ÉNERGIE, J'AI INVENTÉ UNE NOUVELLE MÉTHODE DE CALIBRATION NON PARAMÉTRIQUE.



Mots clefs : *Calibration, Modélisation, Physique des particules*

Stage encadré par :

Colin Bernet colin.bernet@cern.ch

Bâtiment Paul Dirac

4, Rue Enrico Fermi

69622 Villeurbanne Cedex

Tél. : +33 (0) 4 72 44 84 57

Table des matières

1	Introduction	2
2	Production de l'échantillon	3
3	Calibration par régression linéaire	4
4	Méthode non paramétrique binnée	5
5	Méthodes basées sur les plus proches voisins	5
5.1	Moyenne pondérée	5
5.2	Nettoyage gaussien	6
5.2.1	Principe général de l'algorithme	6
5.2.2	Résultat de la calibration	6
5.3	Fit gaussien	6
5.3.1	Principe général de l'algorithme	6
5.3.2	Résultat de la calibration	6
6	Comparaison des méthodes	7
6.1	Méthodes des plus proches voisins	7
7	Annexes	7
7.1	Fonctions utiles du programme	7

1 Introduction

Après avoir permis la découverte expérimentale du boson de Higgs en 2012, les expériences généralistes ATLAS [1] et CMS [2] installées sur le LHC du CERN, sont toujours en place dans l'optique de découvrir de la nouvelle physique au-delà du modèle standard.

Les détecteurs ATLAS et CMS sont basés sur les mêmes principes : cylindriques, ils sont constitués d'un ensemble de sous-détecteurs disposés en couches concentriques autour du point d'interaction. Les informations provenant de ces sous-détecteurs sont combinées pour déterminer le type, l'énergie et la direction des particules de l'état final de la collision, pour pouvoir mesurer les propriétés de celle-ci, et par exemple déterminer si une particule instable encore inconnue a été produite.

Nous allons nous intéresser plus spécifiquement au détecteur CMS [3]. Celui-ci dispose :

- d'un champ magnétique, pour courber la trajectoire des particules chargées ;
- d'un trajectographe, pour reconstruire la trajectoire des particules chargées, et ainsi obtenir la charge et l'impulsion ;
- d'un calorimètre électromagnétique (ECAL) [4], constitué d'un cristal de tungstate de plomb, permettant de collecter les dépôts d'énergie des particules, principalement électrons et photons, mais aussi hadrons chargés et neutres) ;
- d'un calorimètre hadronique (HCAL) [5], composé de plusieurs couches d'absorbeur en laiton et de carreaux scintillateurs en plastique, avec une segmentation grossière. La résolution du HCAL pour la mesure de l'énergie E d'un hadron est de l'ordre de $100\% \sqrt{(E/\text{GeV})}$;
- de chambres à muons, qui permettent l'identification de ces particules, les seules à pouvoir y parvenir.

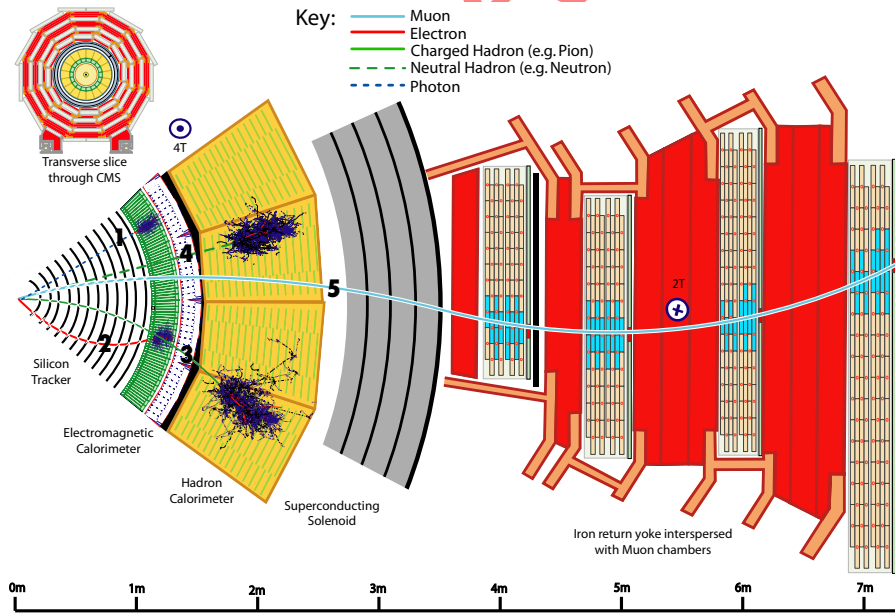


FIGURE 1 – Une esquisse des interactions spécifiques des particules dans une tranche transversale du détecteur CMS.

Détaillons alors le comportement des particules :

- photons (exemple 1 dans la Fig. 1) :
 - déposent leur énergie dans ECAL ;
- e^+, e^- (exemple 2 dans la Fig. 1) :
 - produisent une trace dans le trajectographe ;

- déposent leur énergie dans ECAL ;
- hadrons chargés (exemple 3 dans la Fig. 1) :
 - produisent une trace dans le trajectographe ;
 - déposent en minorité des cas leur énergie dans ECAL ;
 - déposent leur énergie dans HCAL ;
 - fin de course dans HCAL ;
- hadrons neutres (exemple 4 dans la Fig. 1) :
 - déposent leur énergie dans le HCAL ;
 - déposent en minorité des cas leur énergie dans ECAL ;
 - fin de course dans HCAL ;
- μ^+, μ^- (exemple 5 dans la Fig. 1) :
 - produisent une trace dans le trajectographe ;
 - traversent ECAL, HCAL ;
 - chambre à muons.

À noter que dans notre étude seuls les hadrons neutres nous intéressent.

Le connaissance des dépôts d'énergies et du comportement des particules dans les différentes parties du détecteur nous permette de reconnaître et distinguer les particules, cette opération s'appelle le *Particle Flow (PF)*. Cependant, il est aussi nécessaire d'estimer l'énergie des particules (E_{true}) à l'aide d'une calibration des calorimètres. En effet, ces derniers ne présentent pas une réponse linéaire et la somme des énergies dans les calorimètres ne correspond pas à l'énergie de la particule. Cette énergie de calibration sera notée E_{calib} .

En première approximation, nous déterminerons l'énergie calibrée par une fonction linéaire de l'énergie lue dans le ECAL (énergie notée par la suite E_{ecal}) et de celle lue dans le HCAL (notée par la suite H_{ecal}). Cette méthode sera présentée dans la section 3.

Ensuite, ce rapport présente de nouvelles techniques de calibration qui permettent de prendre en compte la non-linéarité des calorimètres. Ces techniques seront présentées dans les sections 4 et 5. Enfin, dans la section 6 nous comparerons ces méthodes.

2 Production de l'échantillon

- paragraphe de Colin sur la simulation
- passer de Root à python -> 1 ligne pour dire que j'ai fait un script
- > plot du nage de point
- séparation $\text{ecal} \neq 0$, $\text{ecal} = 0$:

On séparera et traitera différemment les événements qui ont $E_{\text{ecal}} = 0$. Ces événements sont liés à des particules qui ont interagité avec le détecteur hadronique mais pas avec le détecteur électromagnétique (cf Fig.1). Cette séparation se justifie par le fait que modéliser les dépôts d'énergie dans les deux calorimètres pour en conclure ce qui se passe dans le cas particulier où il n'y a des dépôts que dans un amène un biais. Ainsi, à chaque "création" de calibration, on créera en fait deux modèles.

-> image illustrant la séparation : $\text{ecal} + \text{hcal} < 150$

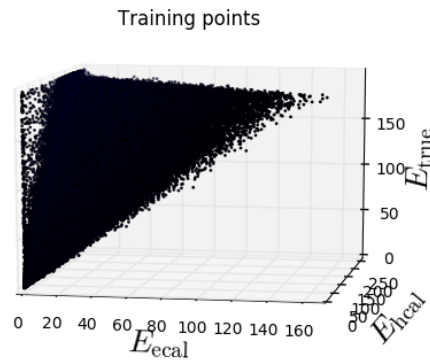


FIGURE 2 – Le nuage de points modélisé

3 Calibration par régression linéaire

- Un première calibration, non binnée

$$E_{\text{true}} = a_1 E_{\text{ecal}} + a_2 E_{\text{hcal}} + b \quad (1)$$

- trouver la meilleure a_1, a_2, b
- pas d'incertitude
- méthode des moindres carrés, ref scikitLearn [6]
- on limite les points de la calibration, \lim_{\min} , \lim_{\max}

- lancer la calibration, doc GitHub [7]

paramères : $\lim = 150, \lim_{\min} = 20, \lim_{\max} = 80$

régression obtenue :

for ecal == 0 :

ecalib = $0.960326144365 \text{ hcal} + 5.92100531606$

for ecal != 0 :

ecalib = $1.34578479311 \text{ ecal} + 0.987111240421 \text{ hcal} + 8.27380815395$

-> plot surface, courbe de calibration

- on calibre un second jeu de données

-> plot ecalib/etrue

- que vaut la calibration ?

-> mise en évidence des non linéarités

- origine de la courbe de fit

- comment fait-on un fit

-scipy [8]

expliquer :

- barre d'erreur

- minimisation du χ^2

- un bon χ^2 réduit ?

- quand nous ferons un fit gaussien ce sera toujours le même principe

4 Méthode non paramétrique binnée

- principe de l'algo
- méthode binnée
- illustration :

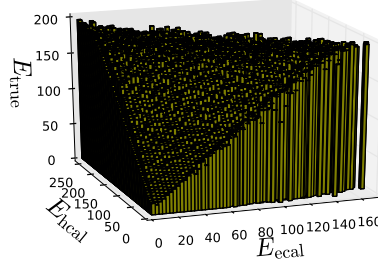


FIGURE 3 – Le nuage de points modélisé par des legos

- > Résultat de la calibration - lancer la calibration ref Git [9]
- paramètres
- > plot surface, courbe de calibration
- on calibre un second jeu de données
- > plot ecalib/etrue
- biais lié aux bins

5 Méthodes basées sur les plus proches voisins

5.1 Moyenne pondérée

- Principe général de l'algorithme
- trouver les plus proches voisins : scikitlearn [10], algorithmes de recherches optimisés

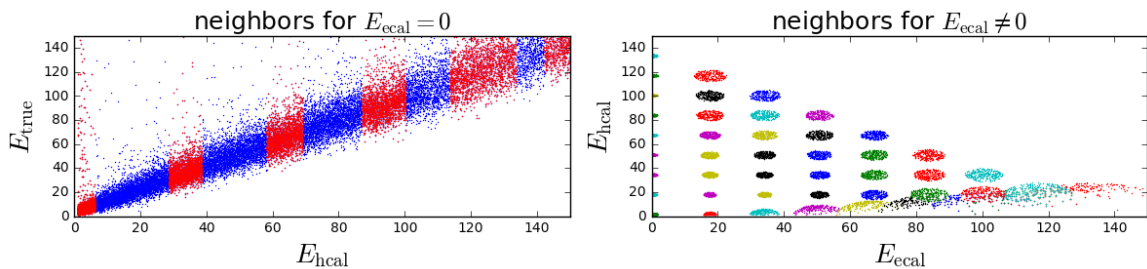


FIGURE 4 – $n_{voisins} = 2000$ pour $E_{cal} = 0$, $n_{voisins} = 250$ pour $E_{cal} \neq 0$

- moyenne pondérée par une gaussienne
 - pas assez lissé, des particules à fort E_{true} mal détectées perturbent la calibration, il faut les enlever
 - Résultat de la calibration
 - créer la calibration, doc git [11] - paramètres :
- $\lim = 150, n_{neighbors, E_{cal}=0} = 2000, n_{neighbors, E_{cal} \neq 0} = 250, \lim = 150, \sigma = 5$

weights = 'gaussian'
 -> plot surface, courbe de calibration
 - on calibre un second jeu de données
 -> plot ecalib/etrue

5.2 Nettoyage gaussien

5.2.1 Principe général de l'algorithme

- on enlève les points éloignés du coeur de la distribution
- principe de l'algo
- interpolation [12]

5.2.2 Résultat de la calibration

- doc git [13] - paramètres :
 lim = 150, $n_{\text{neighbors}, E_{\text{cal}}=0} = 2000$, $n_{\text{neighbors}, E_{\text{cal}} \neq 0} = 250$, lim = 150, $\sigma = 5$
 weights = 'gaussian'
 algorithm = 'auto'
 energystep = 1
 kind = 'cubic'
 cut = 2
 -> plot surface, courbe de calibration
 - on calibre un second jeu de données
 -> plot ecalib/etrue
 - mise en évidence des non linéarités

5.3 Fit gaussien

5.3.1 Principe général de l'algorithme

- principe de l'algo
- Ecalib = μ

5.3.2 Résultat de la calibration

- doc git [14] - paramètres :
 lim = 150, $n_{\text{neighbors}, E_{\text{cal}}=0} = 2000$, $n_{\text{neighbors}, E_{\text{cal}} \neq 0} = 250$, lim = 150, $\sigma = 5$
 algorithm = 'auto'
 kind = 'cubic'
 energystep_{ecal_eq₀} = 1
 energystep_{ecal_ne_q0} = 5
 - > *plotsurface, courbedecalibration*
 - *oncalibreunsecondjeudedonnes*
 - > *plotecalib/etrue*

– *mise en évidence des nonlinéarités*

6 Comparaison des méthodes

6.1 Méthodes des plus proches voisins

7 Annexes

7.1 Fonctions utiles du programme

DRAFT

Références

- [1] ATLAS Collaboration. 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) :1 – 29, 2012.
- [2] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1) :30 – 61, 2012.
- [3] CMS Collaboration. The CMS experiment at the cern LHC. *Journal of Instrumentation*, 3(08) :S08004, 2008. <http://stacks.iop.org/1748-0221/3/i=08/a=S08004>.
- [4] CMS Collaboration. The CMS electromagnetic calorimeter project : technical design report. *Technical Design Report CMS*, 1997. <https://cds.cern.ch/record/349375>.
- [5] CMS Collaboration. CMS : The hadron calorimeter technical design report. *Technical Design Report CMS*, 1997. <https://cds.cern.ch/record/349375>.
- [6] Scikit Learn. *Generalized Linear Models*. scikit-learn.org.
- [7] Samuel Niang. *Classe LinearRegression*.
- [8] Scipy. *Optimization*. docs.scipy.org.
- [9] Samuel Niang. *Classe CalibrationLego*. github.com/sniang/particle_flow_calibration.
- [10] Scikit Learn. *Nearest Neighbors*. scikit-learn.org.
- [11] Samuel Niang. *Classe KNN*. github.com/sniang/particle_flow_calibration.
- [12] Scipy. *Interpolation*. docs.scipy.org.
- [13] Samuel Niang. *Classe KNNGC*. github.com/sniang/particle_flow_calibration.
- [14] Samuel Niang. *Classe KNNGF*. github.com/sniang/particle_flow_calibration.