Astronomie gamma depuis le sol et l'espace : premières analyses du réseau hybride HESS-II et recherche de candidats blazars parmi les sources non-identifiées du Fermi-LAT

Julien Lefaucheur

▶ **To cite this version:**

Julien Lefaucheur. Astronomie gamma depuis le sol et l'espace : premières analyses du réseau hybride HESS-II et recherche de candidats blazars parmi les sources non-identifiées du Fermi-LAT. Physique [physics]. UNIVERSITÉ PARIS DIDEROT SORBONNE PARIS CITÉ (Paris 7), 2015. Français. NNT : . tel-01128429v2

HAL Id: tel-01128429

https://theses.hal.science/tel-01128429v2

Submitted on 1 Jun 2015 (v2), last revised 11 Nov 2015 (v4)

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

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

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

UNIVERSITÉ PARIS DIDEROT SORBONNE PARIS CITÉ

ED 560 : Sciences de la Terre et de l'Environnement et Physique de l'Univers, Paris

Thèse de Doctorat

Presentée par

Julien LEFAUCHEUR

pour obtenir le grade de
Docteur en Sciences de l'Université Paris 7

Astronomie γ depuis le sol et l'espace : premières analyses du réseau hybride HESS-II et recherche de candidats blazars parmi les sources non-identifiées du Fermi-LAT

Thèse dirigée par Santiago PITA

Soutenue publiquement le 29 janvier 2015

Membres du jury :

Rapporteurs :	Giovanni LAMANNA	-	LAPP
	Benoît LOTT	-	CENBG
Présidente :	Catherine BOISSON	-	LUTH
Examinateurs :	Frédéric PIRON	-	LUPM
	David WILLIAMS	-	SCIPP
Directeur :	Santiago PITA	-	APC

Laboratoire APC - AstroParticule Cosmologie
UMR 7164 CNRS - UNIVERSITÉ PARIS 7 - CEA - OBSERVATOIRE DE PARIS

[*The knights look onward as Sir Lancelot steps forward to the Bridge keeper*]

Bridgekeeper: STOP! Who would cross the Bridge of Death must answer me these questions three, ere the other side he see.

Sir Lancelot: Ask me the questions, bridgekeeper. I am not afraid.

Bridgekeeper: What... is your name?

Sir Lancelot: My name is Sir Lancelot of Camelot.

Bridgekeeper: What... is your quest?

Sir Lancelot: To seek the Holy Grail.

Bridgekeeper: What... is your favourite colour?

Sir Lancelot: Blue.

Bridgekeeper: [*Suddenly friendly*] Go on. Off you go.

Sir Lancelot: Oh, thank you. Thank you very much. [*crosses the bridge*]

Sir Robin: That's easy. [He approaches the Bridge. The others follow him.]

Bridgekeeper: STOP! Who would cross the Bridge of Death must answer me these questions three, ere the other side he see.

Sir Robin: Ask me the questions, bridgekeeper. I'm not afraid.

Bridgekeeper: What... is your name?

Sir Robin: Sir Robin of Camelot.

Bridgekeeper: What... is your quest?

Sir Robin: To seek the Holy Grail.

Bridgekeeper: What... is the capital of Assyria?

Sir Robin: I don't know that. AAAAAAAAHHH! [*is cast into the gorge by an invisible force*]

[*Arthur and the other knights realise that this will not be as easy as they thought. Galahad then approaches the Bridge.*]

Bridgekeeper: STOP! What is your name?

Galahad: Sir Galahad of Camelot.

Bridgekeeper: What... is your quest?

Galahad: I seek the Grail.

Bridgekeeper: What... is your favourite colour?

Galahad: Blue. No, yellOOOOOOW! [is cast into the gorge]

Bridgekeeper: Hee hee heh [chuckling at Galahad's demise, then turns to Arthur] . STOP! What... is your name?

King Arthur: It is Arthur, King of the Britons.

Bridgekeeper: What... is your quest?

King Arthur: To seek the Holy Grail.

Bridgekeeper: What... is the air-speed velocity of an unladen swallow?

King Arthur: What do you mean? An African or European swallow?

Bridgekeeper: Huh? I... I don't know that. EEEEEEAAAARRGH! [*is cast into the gorge*]

Sir Bedevere: How do you know so much about swallows?

King Arthur: Well, you have to know these things when you're a king, you know.

[*Arthur and Bedevere begin to cross the bridge*]

***Monty Python and the Holy Grail*, 1975**

Table des matières

I Introduction **3**

1 L'astronomie γ de haute et de très haute énergie **5**
1.1 Introduction . . . 5
1.2 Les rayons cosmiques . . . 7
1.2.1 La détection des rayons cosmiques . . . 8
1.2.2 Accélération et confinement des particules chargées . . . 9
1.3 Processus physique de production de γ . . . 10
1.3.1 Processus électromagnétiques . . . 10
1.3.2 Processus hadronique . . . 11
1.4 Les expériences d'astronomie γ de haute énergie . . . 11
1.4.1 Un bref historique . . . 11
1.4.2 Le satellite *Fermi-LAT* . . . 12
1.5 Les expériences d'astronomie γ de très haute énergie . . . 14
1.5.1 Un bref historique . . . 14
1.5.2 La génération actuelle d'imageurs Tcherenkov . . . 15
1.5.3 La relève . . . 17
Bibliographie . . . 18

2 Noyaux actifs de galaxies **21**
2.1 Introduction . . . 21
2.2 Différentes classes d'AGN . . . 21
2.3 Schéma d'unification des AGN . . . 24
2.4 Distribution spectrale des blazars . . . 25
2.5 Séquence des blazars . . . 29
2.6 AGN observés dans le domaine γ . . . 32
2.6.1 AGN observés à haute énergie par *Fermi-LAT* . . . 32
2.6.2 AGN observés à très haute énergie avec les imageurs Tcherenkov au sol . . . 34
Bibliographie . . . 35

3 Analyses multivariées **37**
3.1 Introduction . . . 37
3.2 Exemple introductif . . . 37
3.3 Phases de développement d'une classification binaire . . . 40
3.4 Arbres de décision « boostés » . . . 41
3.4.1 Arbres de décision . . . 41
3.4.2 « Boosting » . . . 43
3.5 Réseaux de neurones . . . 45
3.5.1 Neurone artificiel . . . 45
3.5.2 Du simple neurone au réseau MLP . . . 49
Bibliographie . . . 52

II L'expérience *H.E.S.S.*, analyses et performances en mode hybride **53**

4 L'expérience *H.E.S.S.* **55**
4.1 Imagerie atmosphérique Tcherenkov . . . 55
4.1.1 Développement des gerbes EM et hadroniques . . . 55

4.1.2 L'effet Tcherenkov . . . 57
4.1.3 Détection des gerbes . . . 58
4.2 Le réseau *H.E.S.S.* . . . 60
4.2.1 Le site . . . 60
4.2.2 Télescopes de type HESS-I . . . 61
4.2.3 Le grand télescope pour la phase HESS-II . . . 63
4.3 Calibration des données . . . 64
4.3.1 Principe . . . 64
4.3.2 Estimation des piédestaux P^{HG} et P^{LG} . . . 65
4.3.3 Estimation de γ_e^{ADC} et du rapport HG/LG . . . 65
4.3.4 Estimation du coefficient FF . . . 66
4.3.5 Efficacité optique . . . 67
4.4 Prises de données . . . 67
4.5 Qualité des données . . . 68
Bibliographie . . . 70

5 Méthodes d'analyse dans *H.E.S.S.* 71
5.1 Analyses de bas niveau dans la collaboration *H.E.S.S.* . . . 71
5.1.1 La méthode de Hillas . . . 71
5.1.2 Modèle 3D . . . 75
5.1.3 Modèle ++ . . . 77
5.1.4 X_{eff} . . . 79
5.1.5 ImPACT . . . 82
5.2 L'analyse Paris-MVA . . . 83
5.2.1 Les reconstructions utilisées . . . 83
5.2.2 Reconstruction de l'énergie . . . 84
5.2.3 Paramètres de discrimination γ/fond . . . 85
5.2.4 Discrimination γ/fond . . . 86
5.2.5 Performances . . . 87
5.3 Extraction du signal et estimation du fond . . . 89
5.3.1 Extraction du signal . . . 89
5.3.2 Méthode des régions OFF multiples . . . 90
5.3.3 Méthode de la région en anneau . . . 91
5.4 Reconstruction spectrale et courbes de lumière . . . 91
5.4.1 Fonctions de réponse de l'instrument . . . 91
5.4.2 Reconstruction spectrale par « *forward-folding* » . . . 92
5.4.3 Courbes de lumière . . . 94
5.5 START . . . 94
5.5.1 Gestion des données . . . 94
5.5.2 Gestion des IRF . . . 95
5.5.3 Détermination de l'énergie seuil . . . 96
5.5.4 Représentation d'un spectre . . . 96
5.5.5 Quelques exemples d'applications . . . 96
Bibliographie . . . 101

6 Préparation de l'analyse en mode hybride à base de simulations 103
6.1 Introduction . . . 103
6.2 Simulations . . . 103
6.3 Reconstruction des évènements γ avec CT_5 . . . 104
6.3.1 Les évènements γ vus par le réseau . . . 104
6.4 Modifications/Améliorations dans APC-MVA . . . 107
6.4.1 Pré-sélection des images mesurées . . . 109

6.4.2 Pré-sélection des images projetées . . . 110
6.4.3 Modification du calcul des paramètres discriminants . . . 110
6.4.4 Amélioration de la PSF avec la sélection $\sin\theta_{ij}$ . . . 112
6.5 Construction des analyses préliminaires . . . 115
6.5.1 Préparation de l'analyse . . . 115
6.5.2 Pondération entre les électrons et les protons . . . 116
6.5.3 L'espace des phases de l'analyse . . . 116
6.5.4 Les variables discriminantes . . . 118
6.5.5 Réglages de la méthode BDT et phases d'entraînement . . . 118
6.5.6 Point de fonctionnement . . . 124
6.6 Estimation des performances préliminaires . . . 125
6.6.1 Le calcul de la sensibilité . . . 125
6.6.2 Résultats . . . 128
Bibliographie . . . 132

7 Analyse APC-MVA en mode hybride 133
7.1 Introduction . . . 133
7.2 Simulations . . . 133
7.3 Fond réel . . . 134
7.4 Construction des analyses . . . 135
7.4.1 Développement et caractéristiques des analyses . . . 135
7.4.2 Optimisation des coupures de discrimination γ/fond . . . 136
7.4.3 Lissage des efficacités γ . . . 137
7.5 Configurations d'analyse et performances . . . 138
7.6 Détection des premières sources . . . 140
7.6.1 Nébuleuse du Crabe . . . 140
7.6.2 PKS 2155-304 . . . 141
7.6.3 PG 1553+113 . . . 142
7.6.4 Mkn 501 . . . 143
7.7 Spectres . . . 144
7.7.1 Nébuleuse du Crabe . . . 144
7.7.2 PKS 2155-304 . . . 145
7.7.3 PG 1553+113 . . . 145
7.7.4 Mkn 501 . . . 146
Bibliographie . . . 150

8 Résumé 153

III Recherche de candidats blazars parmi les sources non-identifiées du *Fermi-LAT* du catalogue 2FGL 155

9 Introduction 157
9.1 Contexte et objectifs . . . 157
9.1.1 Identification des blazars parmi les sources non-identifiées du 2FGL . . . 157
9.1.2 Détermination des coordonnées des candidats blazars du 2FGL à l'aide de l'expérience *WISE* . . . 158
9.2 L'expérience *Fermi-LAT* . . . 159
9.2.1 Le satellite *Fermi* . . . 159
9.2.2 Le *LAT* . . . 160
9.2.3 Performances du *LAT* . . . 161
9.2.4 Le catalogue 2FGL . . . 162

9.3 L'expérience *WISE* . . . 165
9.3.1 Le satellite . . . 165
9.3.2 Performances de l'instrument . . . 166
9.3.3 Le catalogue . . . 166
Bibliographie . . . 169

10 Recherche de candidats blazars à partir du catalogue 2FGL 171
10.1 Introduction . . . 171
10.2 Définition des différents échantillons . . . 172
10.2.1 Échantillon « signal » . . . 172
10.2.2 Échantillons de « fond » . . . 172
10.2.3 Échantillon des sources non-identifiées . . . 172
10.3 Variables discriminantes . . . 173
10.3.1 Discrimination et brillance des sources . . . 173
10.3.2 Paramètres spectraux . . . 173
10.3.3 Paramètres de variabilité . . . 177
10.3.4 Paramètres sélectionnés pour la discrimination . . . 180
10.4 Détermination des coupures de discrimination . . . 181
10.4.1 Sélection et réglages des analyses multivariées . . . 181
10.4.2 Méthodologie pour la détermination des performances . . . 182
10.4.3 Performances de l'analyse « extragalactique » . . . 182
10.4.4 Performances de l'analyse « galactique » . . . 183
10.5 Candidats blazars . . . 183
10.A Stabilité des indices de dureté HR_{ij} . . . 193
10.B Paramètres TMVA des analyses BDT et MLP . . . 196
10.C Performances des méthodes individuelles BDT et MLP de l'analyse de la région « extragalactique » . . . 198
10.D Performances des méthodes individuelles BDT et MLP de l'analyse de la région « galactique » . . . 200
Bibliographie . . . 202

11 Recherche de candidats blazars à partir des catalogues 2FGL et *WISE* 203
11.1 Introduction . . . 203
11.2 Statut sur la recherche de contreparties IR des blazars . . . 204
11.2.1 L'émission non-thermique des blazars vue par *WISE* . . . 204
11.2.2 L'approche géométrique et ses résultats . . . 205
11.3 Descriptif de notre approche . . . 206
11.4 Dépendance spatiale du fond infrarouge . . . 207
11.5 Comparaison des variables discriminantes . . . 207
11.5.1 Choix des seuils de sélection . . . 210
11.5.2 Correction de l'extinction dans l'infrarouge . . . 211
11.5.3 Les couleurs . . . 211
11.5.4 Le flux infrarouge des blazars . . . 211
11.5.5 La variabilité . . . 214
11.6 Construction et performances des analyses . . . 214
11.6.1 Définition des différents échantillons . . . 214
11.6.2 Sélection des analyses multivariées . . . 216
11.6.3 Réglages de la méthode MLP . . . 218
11.6.4 Performances . . . 218
11.6.5 Points de fonctionnement de l'analyse . . . 221
11.7 Résultats . . . 223
11.7.1 Sources non-identifiées du 2FGL sans « flag » de reconstruction . . . 224

11.7.2 Sources non-identifiées du 2FGL avec des « flags » de reconstruction 225
11.7.3 Discussion 225
Bibliographie 245

12 Résumé **247**

IV Conclusions et perspectives **249**

13 Conclusion et perspectives **251**

Acronymes

2MASS *The Two Micron All Sky Survey.*
AGILE *Astro-rivelatore Gamma a Immagini LEggero.*
AMS *Alpha Magnetic Spectrometer.*
CAT *Cherenkov Array at Thémis.*
CTA *Cherenkov Telescope Array.*
EGRET *Energetic Gamma-Ray Experiment Telescope.*
Fermi *Fermi Gamma-ray Space Telescope.*
GBM *Glast Burst Monitor.*
H.E.S.S. *High Energy Spectroscopic System.*
HAWC *High-Altitude Water Čerenkov.*
HEGRA *High Energy Gamma Ray Astronomy.*
ISS station spatiale internationale.
JEM-EUSO *Japanese Experiment Module - Extreme Universe Space Observatory.*
LAT *Large Area Telescope.*
MAGIC *Major Atmospheric Gamma-ray Imaging Cherenkov Telescope.*
NVSS *NRAO VLA SKY SURVEY.*
SAS-2 *Small Astronomy Satellite 2.*
VERITAS *Very Energetic Radiation Imaging Telescope Array.*
WISE *Wide-Field Infrared Survey Explorer.*
0FGL *Fermi-LAT Bright Sources List.*
1FGL *First Fermi-LAT catalog.*
2FGL *Fermi-LAT 2-Year Source Catalog.*
2LAC *Second LAT AGN Catalog.*
3FGL *Fermi LAT 3-Year Source Catalog.*

AdaBoost « boosting » adaptif.
ADC convertisseur analogique-numérique.
AGN noyaux actifs de galaxies.
ARS Analog Ring Sampling.

BDT arbres de décisions « boostés ».
BL Lac BL Lacertae.

CMB fond diffus cosmologique.
CORSIKA *COsmic Ray SImulations for KAscade.*

EM électromagnétique.

FPGA réseau de portes programmables.
FSRQ *Flat Spectrum Radio Quasar.*

GZK Greisen-Zatsepin-Kuzmin.

HBL High-frequency peaked BL Lac.
HG haut gain (High Gain).

IACT télescopes d'imagerie atmosphérique Čerenkov.
IBL Intermediate-frequency peaked BL Lac.
IR infrarouge.
IRF fonctions de réponse de l'instrument.

KS Kolmogorov-Smirnof.

LBL Low-frequency peaked BL Lac.
LED diodes électroluminescentes.
LG bas gain (Low Gain).

MC Monte-Carlo.
MLP perceptron multicouche.
MVA analyses multivariées.

NASA *National Aeronautics and Space Administration*.
NN réseaux de neurones.
NSB bruit de fond du ciel.

p.e. photon-électrons.
PM photomultiplicateur.
PSF fonction d'étalement du point.
PWL loi de puissance.
PWN plérions.

RF forêt d'arbres décisionnels.
Roma-BZCAT *Multifrequency catalogue of blazars*.

SAM Swift Analogue Memory.
SNR restes de supernova.
SVM machine à vecteurs de support.

TMVA *Toolkit for Multivariate Data Analysis*.

ULIRG *Ultra Luminous Infrared Galaxies*.
UV ultraviolet.

WBS *Wise γ-ray Blazars Strip*.

Préambule

Ce manuscrit présente les travaux que j'ai effectués durant ma thèse sous la direction de Santiago Pita au laboratoire Astroparticule et Cosmologie (APC). Il y sera question d'astronomie γ de haute énergie (entre $\sim$30 GeV et $\sim$300 GeV) par l'intermédiaire du télescope *Large Area Telescope* (*LAT*) embarqué à bord du satellite *Fermi Gamma-ray Space Telescope* (*Fermi*) et d'astronomie γ de très haute énergie (au-delà de $\sim$100 GeV) via l'expérience *High Energy Spectroscopic System* (*H.E.S.S.*). L'observation du rayonnement gamma permet de sonder les accélérateurs cosmiques, dans lesquels la matière ou le rayonnement sont dans des conditions extrêmes de densité, et de champ gravitationnel ou électromagnétique. À l'aide de ces observations, il devient aujourd'hui possible d'étudier la matière dans des conditions qu'il est impossible de reproduire sur Terre.

Durant ma thèse, j'ai été membre de la collaboration *H.E.S.S.* Cette expérience utilise la technique d'imagerie atmosphérique Tcherenkov afin d'étudier les sources γ émettrices au-delà de la centaine de GeV. On compte aujourd'hui près de 150 sources γ détectées par les différentes expériences d'imagerie Tcherenkov. On cherche aujourd'hui à réduire le seuil en énergie des instruments actuels afin de détecter de nouvelles sources et de sonder de manière plus fine les accélérateurs de particules. Une bonne partie de ce manuscrit est dédiée à cette expérience qui a débuté sa seconde phase (HESS-II) avec l'ajout d'un cinquième grand télescope de 28 m de diamètre au réseau déjà existant, initialement constitué de quatre télescopes de 12 m de diamètre. Prenant pour point de départ l'analyse Paris-MVA initialement développée pour la première phase de *H.E.S.S.*, mon travail a été dans un premier temps de l'adapter afin de prendre en compte le cinquième télescope dans l'analyse. Par la suite, un travail de fond a été réalisé sur la recherche des points de fonctionnement optimaux afin de tirer un maximum de significativité pour des sources ayant des niveaux de flux différents.

Je me suis aussi impliqué dans l'identification des blazars parmi les sources non-identifiées de *Fermi-LAT* du catalogue *Fermi-LAT 2-Year Source Catalog* (2FGL). Les blazars font partie de la classe des noyaux actifs de galaxies (AGN), objets associés à un trou noir supermassif et entre autres à un disque d'accrétion, et dont le jet relativiste est quasi-aligné avec notre ligne de visée. Un important effet Doppler relativiste entraîne un décalage des spectres vers les hautes énergies et l'amplification de leur luminosité. Cette particularité facilite l'étude dans le domaine γ des mécanismes d'accélération au sein des jets relativistes via la variabilité et la forme spectrale mesurée mais permet aussi de contraindre le schéma d'unification des différentes classes d'AGN, qui fait l'hypothèse que les différences observées dans leur caractéristiques sont dues à des biais observationnels. Sur les $\sim$1900 sources détectées par le *LAT*, on compte près de $\sim$1100 blazars et environ $\sim$200 sources galactiques. Les autres sources ($\sim$600) ne possèdent pas de contrepartie à d'autres longueurs d'onde et leur nature reste à déterminer. À basse latitude galactique ($|b| < 10°$), il est estimé qu'une centaine de blazars doivent s'y trouver afin de rendre compte du déficit observé à ces latitudes. À plus haute latitude, on s'attend à ce que les sources non-identifiées soient principalement réparties entre les blazars et les pulsars. Compte tenu du nombre de blazars émetteurs γ actuels, j'ai cherché à augmenter ce nombre en identifiant les potentiels candidats parmi les sources non-identifiées à l'aide de deux analyses indépendantes basées sur des méthodes multivariées.

Dans le Partie I de ce manuscrit, après une brève présentation de l'astronomie γ de haute et de très haute énergie au chapitre 1, nous présentons les AGN au chapitre 2. Nous introduisons au chapitre 3 les méthodes dites multivariées qui seront utilisées tout au long des Parties II et III afin de construire des classifications binaires « signal »/« fond ».

La Partie II du manuscrit est dédiée à l'expérience d'astronomie γ de très haute énergie *H.E.S.S.* Au cours du chapitre 4, nous présentons la méthode d'imagerie atmosphérique Tcherenkov ainsi que le fonctionnement du réseau *H.E.S.S.* Dans le chapitre 5, nous présentons les différentes analyses de bas niveau utilisées dans la collaboration *H.E.S.S.*, ainsi qu'un outil de haut niveau développé au cours de cette thèse permettant l'étude spectrale des sources γ. Dans le chapitre 6, une première

estimation des performances du réseau HESS-II en mode hybride est réalisée à l'aide de simulations Monte-Carlo (MC) pour le signal et pour le fond. Nous présentons au chapitre 7 l'analyse APC-MVA en mode hybride construite à partir de données réelles pour le fond, et les résultats des premières analyses de données sur un lot de sources sélectionnées.

La Partie III du manuscrit est consacrée à la recherche des candidats blazars parmi les sources non-identifiées du catalogue 2FGL détectées par le *Fermi-LAT*. Après une présentation des satellites *Fermi-LAT* et *Wide-Field Infrared Survey Explorer* (*WISE*) au chapitre 9, une classification binaire « blazar »/« non-blazar » est développée au chapitre 10 à partir de variables discriminantes construites à partir des paramètres du catalogue 2FGL. Pour finir, nous cherchons à déterminer la nature des sources non-identifiées du catalogue 2FGL en recherchant des contreparties infrarouges de type blazar spatialement compatibles à l'aide du satellite *WISE*.

Première partie

Introduction

Chapitre 1

L'astronomie γ de haute et de très haute énergie

Sommaire

1.1 Introduction 5
1.2 Les rayons cosmiques 7
1.3 Processus physique de production de γ 10
1.4 Les expériences d'astronomie γ de haute énergie 11
1.5 Les expériences d'astronomie γ de très haute énergie 14
Bibliographie 18

1.1 Introduction

Seule une infime fraction du spectre électromagnétique (EM), le domaine « visible » entre ∼400 nm et ∼800 nm, est naturellement accessible à l'homme (Figure 1.1a). Le reste du spectre électromagnétique, à l'exception d'une partie de l'infrarouge (IR) et du domaine radio entre 50 µm et 10 m, est absorbé par l'atmosphère. Dans sa globalité, le spectre électromagnétique s'étend sur une vingtaine d'ordres de grandeur (Figure 1.1b) .

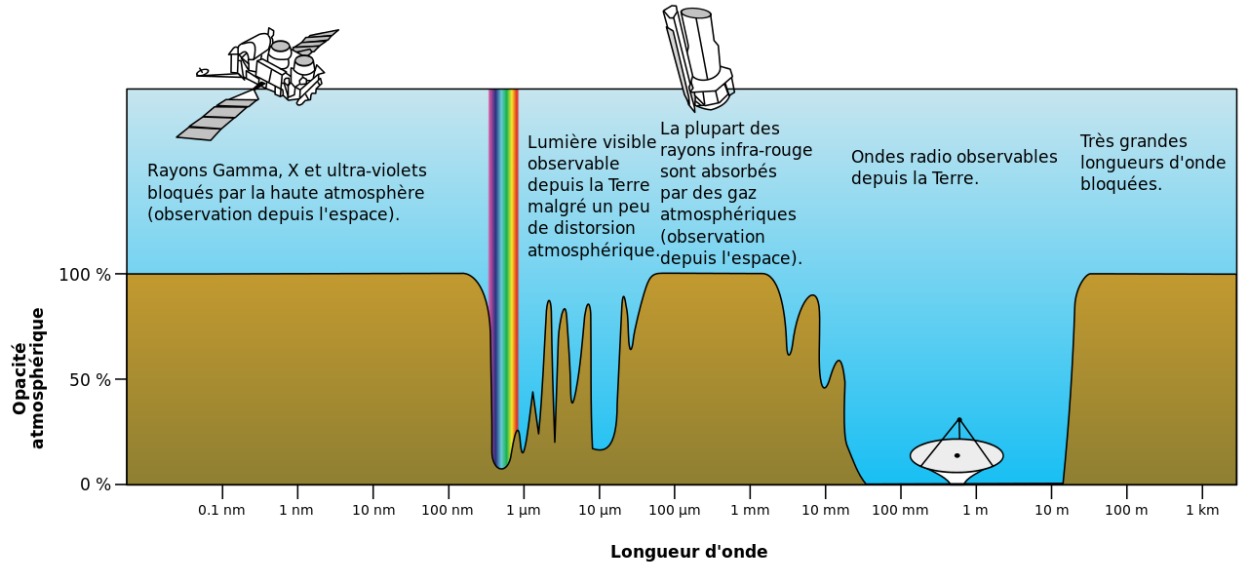
L'astrophysique des hautes énergies est née avec la radioastronomie au lendemain de la deuxième guerre mondiale puis s'est rapidement développée à partir des années 1960 avec les premières fusées, puis les satellites X et γ. Elle est aujourd'hui considérée comme un domaine de recherche à part entière.

L'astronomie γ de haute énergie commence aux alentours de la trentaine de MeV [1], et au-delà de la centaine de GeV [2] débute le domaine de l'astronomie γ de très haute énergie. L'absorption des rayons γ dans l'atmosphère a nécessité le développement de techniques de détection des photons de haute énergie dans l'espace. Aux très hautes énergies, au-delà de la dizaine de GeV, le développement de gerbes atmosphériques induites par les γ incidents et composées de particules ultra-relativistes, permettent via la détection du rayonnement Tcherenkov de ces dernières la caractérisation des γ incidents. L'astronomie γ depuis le sol repose sur la récolte de cette lumière émise dans le visible et dans l'ultraviolet (UV). À ce titre, l'atmosphère fait partie intégrante du dispositif de détection.

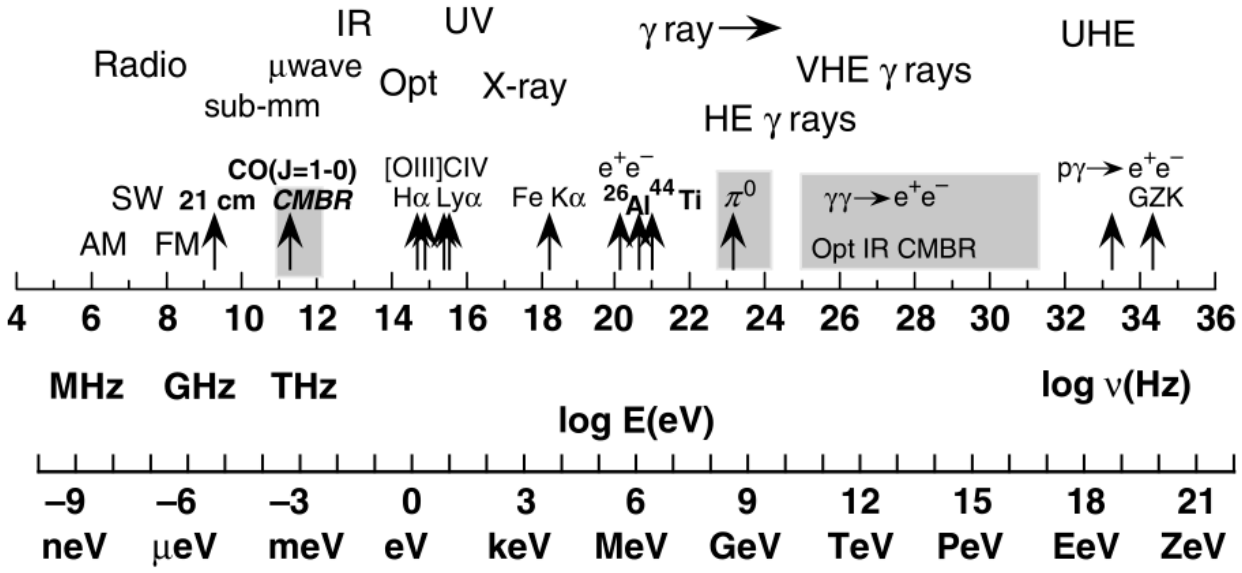
De manière générale, une grande partie des photons optiques ou IR sont émis via des processus thermiques (un objet chauffé émet de la lumière), par exemple par les étoiles. Un photon possédant une énergie de 1 TeV contient environ mille milliards de fois l'énergie d'un photon optique et dans les cas extrêmes, les processus thermiques peuvent produire des photons de l'ordre de ∼1 keV. On associe donc aux photons de haute et de très haute énergie des mécanismes de production non-thermiques auxquels on ne peut pas associer de température effective. Il faut alors faire appel à l'accélération de particules chargées relativistes et à des processus de production de photons par des interactions électromagnétiques ou nucléaires.

La production de particules aussi énergétiques est associée aux phénomènes les plus violents de l'Univers lors desquels la matière et le rayonnement sont dans des conditions extrêmes de champ

1. 1 eV=1,602 176 565 × 10^{-19} J.
2. Les expériences HESS-II puis CTA permettront de descendre ce seuil à quelques dizaines de GeV.



(a)



(b)

FIGURE 1.1: (a) Transmission de l'atmosphère en fonction de la longueur d'onde (Crédits : *National Aeronautics and Space Administration* (NASA)). (b) Spectre électromagnétique s'étendant du domaine radio au domaine γ de très haute énergie en fonction de la fréquence et de l'énergie des photons. (Crédits : [Dermer et Menon, 2009]).

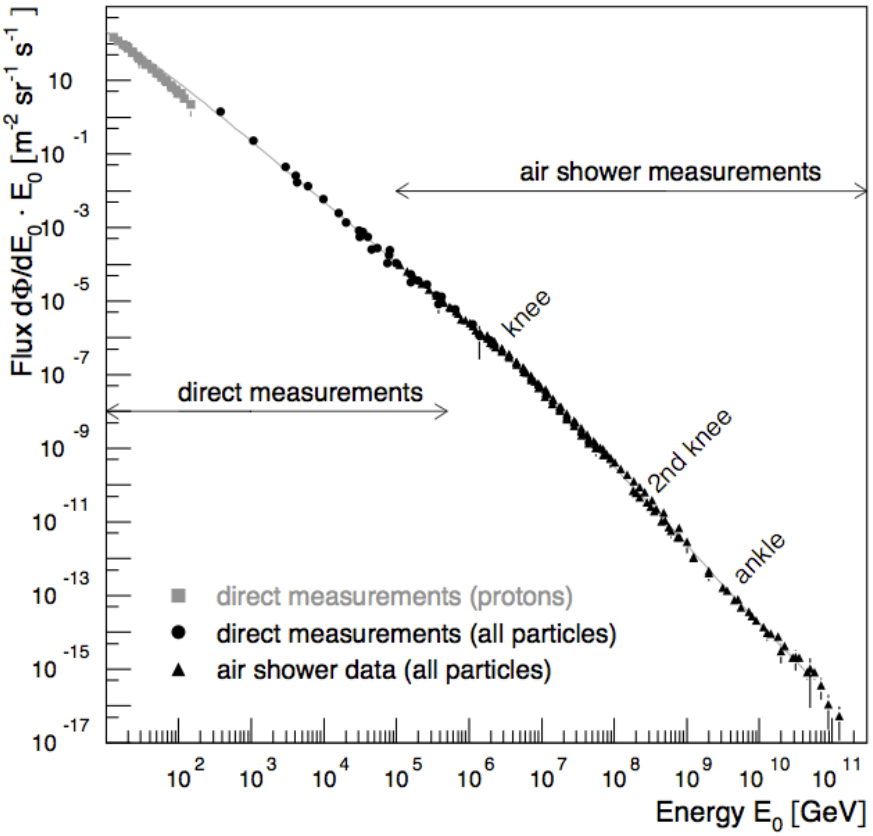


FIGURE 1.2: Spectre différentiel des rayons cosmiques mesuré par diverses expériences (Crédits : [Blümer et al., 2009]).

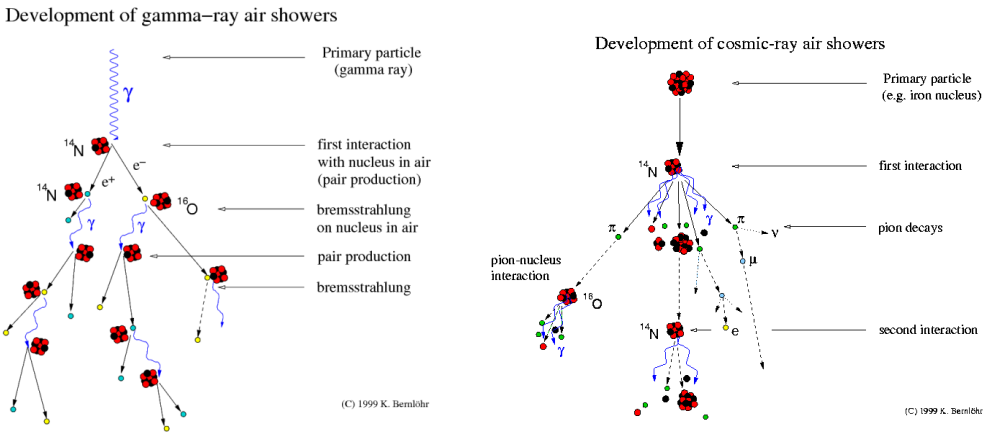
gravitationnel et/ou électromagnétique, lesquelles sont aujourd'hui impossibles à reproduire sur Terre. Par exemple, les pressions colossales s'exerçant au sein d'une étoile à neutrons dont une cuillère à soupe du cœur possèderait une masse de $\sim 10^{15}$ g !

On constate que les photons très énergétiques ont un spectre différentiel décroissant qui généralement suit localement un comportement en loi de puissance. Ceci signifie que, contrairement aux photons optiques, les évènements de haute énergie sont rares, ce qui implique à son tour que le nombre de sources émettant à haute énergie est bien plus faible en proportion que le nombre de sources émettant en optiques ou en IR.

Depuis leur découverte, une des grandes questions de l'astrophysique des hautes énergies concerne l'origine des rayons cosmiques. Mais plus généralement il s'agit de comprendre les processus d'accélération et de rayonnement, et les sites associés dans le cosmos. Nous présentons à la section 1.2 les moyens mis en place pour détecter les rayons cosmiques ainsi qu'un bref aperçu des processus d'accélération des particules chargées. À la section section 1.3, nous présentons les processus physiques permettant de produire des photons γ. Pour finir, nous présentons respectivement aux sections 1.4 et 1.5 les expériences d'astronomie γ de haute et de très haute énergie.

1.2 Les rayons cosmiques

Le rayonnement cosmique est constitué de particules chargées (protons, électrons, etc) ainsi que de particules neutres telles que les photons ou les neutrinos. Dans ce manuscrit, nous désignons par rayons cosmiques les particules chargées (composées à 99 % de hadrons). Après une brève présentation des moyens mis en œuvre afin de détecter les rayons cosmiques, nous discutons rapidement des moyens permettant d'accélérer les particules chargées de très haute énergie. On pourra trouver une discussion très générale sur les rayons cosmiques dans par exemple [Blümer et al., 2009].



(a) (b)

FIGURE 1.3: Principe de développement de gerbes (a) EM induite par un photon de haute énergie, et (b) hadronique produite un noyau de fer de haute énergie (Crédits : http://www.mpi-hd.mpg.de/hfm/%7Ebernlohr/)

1.2.1 La détection des rayons cosmiques

Historiquement, Victor Franz Hess est le premier à avoir déterminé en 1912 de manière indirecte l'existence de particules chargées extra-terrestres. Il fait cette découverte à l'aide de ballons allant jusqu'à une altitude de 5 km en mesurant le taux d'ionisation de l'atmosphère en fonction de l'altitude. On doit au physicien Robert Andrews Millikan en 1927 la dénomination de ces particules par rayons cosmiques. C'est le physicien français Pierre Auger qui en 1938 démontra l'existence de particules « primaires » pouvant posséder des énergies de l'ordre de 10^{15} eV, à travers la détection des grandes gerbes atmosphériques EM et hadroniques (composées de particules dites « secondaire », responsables de l'ionisation de l'atmosphère, voir Figure 1.3) issues de l'interaction avec les noyaux des molécules de l'atmosphère.

On peut observer sur la Figure 1.2 le spectre différentiel des rayons cosmiques mesuré par différentes expériences, qui s'étend de $\sim 10^9$ eV à $\sim 10^{20}$ eV et s'étale sur plus de trente ordres de grandeur en flux. Celui-ci suit une loi de puissance avec deux changements d'indices qui sont visibles à $\sim 10^{15}$ eV (appelé genou) et à $\sim 10^{19}$ eV (appelé cheville). Une mise en évidence d'une coupure dans le spectre des rayons cosmiques au delà de quelques $\sim 10^{19}$ eV est compatible avec la limite Greisen-Zatsepin-Kuzmin (GZK), signant l'interaction des rayons cosmiques de ultra haute énergie avec les photons du fond diffus cosmologique (CMB) [Abraham et al., 2008].

À ce jour, l'étude des rayons cosmiques dans le domaine du GeV est réalisé par l'expérience *Alpha Magnetic Spectrometer* (*AMS*) [Aguilar et al., 2013], véritable détecteur de particules embarqué à bord de la station spatiale internationale (*ISS*) opérationnel depuis 2011. Il sera chargé de mesurer pendant une vingtaine d'années avec une précision jamais égalée les contributions des différentes espèces de particules chargées du spectre des rayons cosmiques. Il a aussi pour mission de sonder l'existence de la matière noire à travers la recherche de signatures dans le spectre des rayons cosmiques, par exemple celles des positrons.

Aux énergie supérieures, de l'ordre de $\sim 10^{15}$ eV, les rayons cosmiques sont étudiés depuis le sol en utilisant l'atmosphère comme un calorimètre. Les produits secondaires détectés au sol permettent de reconstruire la direction et l'énergie de la particule incidente. La rareté des événements (on s'attend à compter une particule de 10^{20} eV par siècle et par km^2) implique l'utilisation de larges surfaces de collection.

Par exemple, l'observatoire Auger installé dans la Pampa argentine utilise 1600 détecteurs espacés de 1,5 km couvrant une surface totale de 3000 km^2. Ces détecteurs permettent d'échantillonner

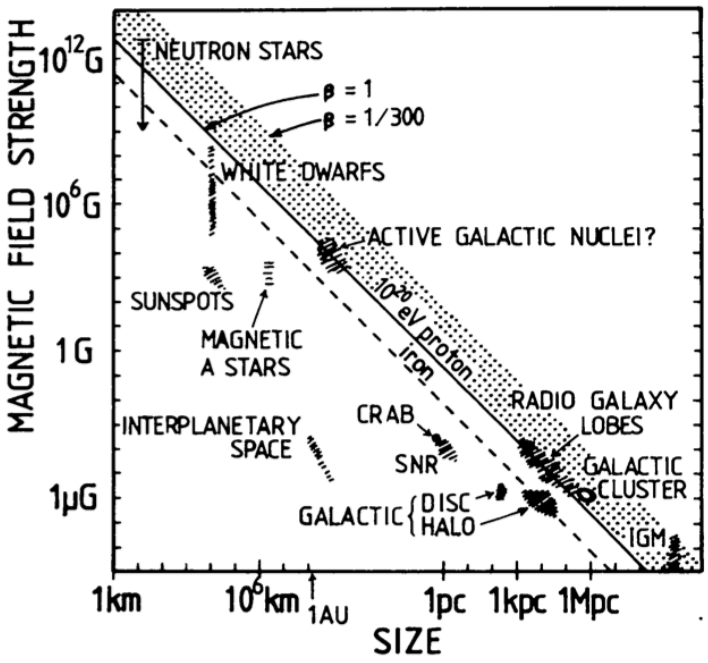


FIGURE 1.4: Diagramme d'Hillas originel [Hillas, 1984]. Champ magnétique en fonction de la taille caractéristique des sources pour différentes classes de sources. Les objets sous la diagonale en train plein (pointillé) ne peuvent pas accélérer des protons (noyaux de fer) de 10^{20} eV.

les particules arrivant au sol et de reconstruire les caractéristiques (énergie, direction et nature) des particules primaires. Une deuxième technique basée sur la détection dans l'UV de l'émission des molécules de l'air excitées par le passage de la gerbe est aussi utilisée. Les expériences *Haverah Park experiment*, *AGASA*, *HiRes* et Auger et MILAGRO ont permis de mettre en évidence un certain degré d'anisotropie dans la direction d'arrivée des rayons cosmiques à grande échelle et d'apporter des éléments de réponse quant à leur composition. On pourra obtenir les informations associées à ces deux thématiques respectivement dans [Abdo et al., 2009b] et [Kampert et Unger, 2012]. Notons qu'une nouvelle technique d'observation depuis l'espace des gerbes initiées par les rayons cosmiques d'ultra haute énergie est actuellement en cours de développement avec l'expérience *Japanese Experiment Module - Extreme Universe Space Observatory (JEM-EUSO)* [Ebisuzaki et al., 2008], qui est prévu pour embarquement à bord de l'*ISS*.

1.2.2 Accélération et confinement des particules chargées

L'accélération de particules chargées peut être liée à la présence d'un champ électrique suivant les lignes de champ magnétique intense (par exemple à la proximité d'une étoile à neutrons), ou à un milieu magnétisé dans lequel se déplace une onde de choc. Historiquement, le physicien Fermi [Fermi, 1949] présenta le premier mécanisme faisant intervenir la déflexion de particules chargées sur des nuages magnétisés en mouvement dans le milieu interstellaire. Le faible gain d'énergie (dû au fait qu'il s'agissait d'un processus du second ordre) amena le physicien Bell [Bell, 1978] (entre autres) à adapter ce principe à des milieux choqués permettant à la fois de produire une distribution des particules suivant une loi de puissance universelle (rendant ainsi compte du spectre des rayons cosmiques mesuré, voir Figure 1.2) et d'obtenir un mécanisme d'accélération efficace (du premier ordre cette fois).

L'énergie maximale que peut atteindre un rayon cosmique dépend de la nature de l'accélérateur et de ses caractéristiques propres. Plus les rayons cosmiques resteront confinés dans la zone d'accélération, et plus le gain en énergie sera important. En négligeant les pertes radiatives, la particule doit avoir un rayon de Larmor $r_L = p/ZqB \approx \mathcal{E}/ZqBc$ plus petit que l'étendue caractéristique L de la source : $L > 2r_L$ [Hillas, 1984]. Ce qui peut s'écrire :

$$\left(\frac{B}{1\,\mathrm{G}}\right)\left(\frac{L}{1\,\mathrm{pc}}\right) > \frac{21}{Z}\left(\frac{\mathcal{E}}{10^{20}\,\mathrm{eV}}\right) \tag{1.1}$$

où B est le champ magnétique, $\mathcal{E}$ l'énergie de la particule et Z sa charge. À titre d'illustration, le champ magnétique moyen de notre galaxie, évalué à 2 nG, n'est pas suffisant pour confiner la trajectoire d'un rayon cosmique d'ultra haute énergie, ce qui renforce l'hypothèse d'une origine extragalactique de ces derniers. La Figure 1.4 est le diagramme originel d'Hillas qui montre différentes sources galactiques et extragalactiques représentées en fonction de leur champ magnétique et de leur taille caractéristique. Un objet sous la ligne pleine ne satisfait pas au critère d'Hillas pour accélérer des protons à 10^{20} eV. Les restes de supernovæ supposés responsables de la production d'une grande fraction des rayons cosmiques ne satisfont pas à ce critère. Pour plus de détails sur le sujet, notamment sur la prise en compte des pertes radiatives des particules, on pourra par exemple se référer à [Ptitsyna et Troitsky, 2010]

1.3 Processus physique de production de γ

Cette partie rappelle les processus d'émission de rayons γ par des particules chargées. Pour une discussion générale sur les différents processus de rayonnement électromagnétique en astrophysique on consultera par exemple l'ouvrage de [Rybicki et Lightman, 1986].

1.3.1 Processus électromagnétiques

1.3.1.1 Effet Compton-inverse

Lors d'interactions avec les photons, les électrons ou positrons relativistes peuvent leur céder une partie de leur énergie via le mécanisme appelé Compton-inverse. L'énergie $\varepsilon_{\max}$ que peut atteindre un photon est maximale lorsque l'électron et le photon ont des quantités de mouvement suivant la même direction mais de sens opposé :

$$\varepsilon_{\max} = \gamma^2 h\nu(1+\beta)^2 \approx 4\gamma^2 h\nu \tag{1.2}$$

Selon l'énergie du photon, deux régimes sont à considérer pour l'expression de la section efficace. En notant $x = h\nu/m_e c^2$ le rapport entre l'énergie initiale du photon et l'énergie de masse de l'électron, le premier régime correspond à des échanges d'énergies faibles ($x \ll 1$). La section efficace de diffusion Compton tend asymptotiquement vers la section efficace de diffusion Thomson :

$$\sigma_{\mathrm{C}} \underset{x\ll 1}{\sim} \sigma_T(1-2x) \tag{1.3}$$

où $\sigma_{\mathrm{T}} = \frac{8\pi}{3}\left(\frac{e^2}{m_e c^2}\right)^2$. Dans le régime dit de Klein-Nishina ($x \gg 1$), où les transferts d'énergie du lepton vers le photon sont importants, elle diminue :

$$\sigma_{\mathrm{C}} \underset{x\gg 1}{\sim} \frac{3}{8x}\sigma_T\left(\frac{1}{2} + \ln 2x\right) \tag{1.4}$$

Dans le cas où un électron est plongé dans une distribution de photons isotrope, la puissance qu'il rayonne est donnée par :

$$P_{\mathrm{IC}} = -\frac{4}{3}\sigma_{\mathrm{T}} c\beta^2\Gamma^2 u_{\mathrm{rad}} \tag{1.5}$$

où u_{rad} est la densité d'énergie radiative. On peut montrer qu'une population d'électrons distribuée selon une loi de puissance $dN \propto E^{-p}dE$ produira des photons suivant une loi de puissance proportionnelle à $I(\nu) \propto \nu^{-\frac{p-1}{2}}$.

1.3.1.2 Rayonnement synchrotron

Un électron ou positron plongé dans un champ magnétique B et dont le vecteur vitesse fait un angle α avec le champ B rayonne une énergie par unité de temps :

$$P_{\mathrm{Sync}} = -\frac{4}{3}\sigma_{\mathrm{T}} c\beta^2\Gamma^2 u_{\mathrm{B}} \tag{1.6}$$

où $u_B = B^2/2\mu_0$ représente la densité d'énergie magnétique. Le fait que la forme des équations 1.5 et 1.6 soit identique indique simplement que le processus électromagnétique sous-jacent fait intervenir une particule chargée et un photon. On peut noter que le rapport $P_{\rm Sync}/P_{\rm IC} = u_B/u_{\rm rad}$ permet d'avoir des informations sur la densité d'énergie de photons en ayant une mesure du champ magnétique, et vice versa. De la même manière que pour le processus Compton-inverse, des électrons distribués selon une loi de puissance $dN \propto E^{-p}dE$ produiront une population de photons suivant une loi de puissance : $I(\nu) \propto B^{\frac{p+1}{2}}\nu^{-\frac{p-1}{2}}$.

1.3.1.3 Création de paires

Deux photons peuvent interagir et créer une paire électron/positron si leurs énergies satisfont la relation :

$$\varepsilon_1\varepsilon_2 \geq \frac{2m_ec^2}{1-\cos\theta} \tag{1.7}$$

où θ est l'angle entre les quantités de mouvement des deux photons ($\theta = 2\pi$ pour une collision frontale). Ce processus est observé de manière indirecte en étudiant le spectre différentiel de noyaux actifs lointains. On constate une absorption des photons de haute énergie qui est due à leur interaction avec des photons de basse énergie qui constituent le rayonnement de fond extragalactique.

1.3.1.4 Bremsstrahlung

Une particule chargée se déplaçant dans le champ électrique des noyaux va être ralentie et émettre un rayonnement de freinage (*Bremsstrahlung* en allemand). Ce mécanisme de production de photons peut créer des X durs ou des γ mous à partir d'une population d'électrons thermiques, mais c'est son importance dans le développement des gerbes EM qui nous intéressera particulièrement (voir le chapitre 4).

1.3.2 Processus hadronique

Les protons individuels ou appartenant à un noyau plus lourd peuvent interagir avec les atomes du milieu interstellaire selon la réaction $pp \rightarrow pp\pi^0$. L'énergie seuil de cette réaction est $\varepsilon_s = 2m_\pi c^2 + \frac{m_\pi^2}{2m_p}c^2 = 280\,\text{MeV}$. Le pion neutre résultant de ce type d'interaction possède une durée de vie de $8{,}4 \times 10^{-17}\,\text{s}$ et se désintègre via l'interaction électromagnétique en $\pi^0 \rightarrow \gamma\gamma$. Le spectre de photons de haute énergie résultant suit la distribution parente de hadrons.

1.4 Les expériences d'astronomie γ de haute énergie

En raison de leur faible surface efficace de détection, les détecteurs spatiaux sont principalement sensibles à l'émission des sources en deçà de 100 GeV. Différentes techniques ont été utilisées afin de détecter et caractériser les rayons γ, la plus récente correspondant au couplage d'un télescope à création de paires avec un calorimètre. Ces détecteurs possèdent un taux élevé de rejet du bruit de fond généré par les rayons cosmiques chargés grâce à l'utilisation de boucliers anti-coïncidence. Après un bref historique des expériences passées, nous présentons les résultats obtenus par le satellite *Fermi*, équipé du *LAT*. Une présentation technique du *Fermi-LAT* sera effectuée au chapitre 9.

1.4.1 Un bref historique

Le premier satellite dédié à l'observation du ciel γ de haute énergie entre 20 MeV et 1 GeV est le *Small Astronomy Satellite 2 (SAS-2)* [Fichtel et al., 1975]. Lancé en 1972 et opérationnel pendant un peu moins de six mois, il possédait une surface de collection de $540\,\text{cm}^2$ et une résolution angulaire de $\sim 3°$. Il a mis en évidence l'émission γ d'un nombre réduit de sources galactiques telles que le pulsar du Crabe, le pulsar de Véla et le pulsar Geminga. À partir de 1975, le satellite *COS-B* équipé

de l'instrument *Gamma-Ray Telescope* [Bignami et al., 1975] surveille le ciel pendant près de sept ans entre 2 keV et 6 GeV. Doté d'une surface efficace de collection de 25 cm^2 et d'une précision angulaire de $\sim$2°, il détecte 25 sources dont la première source extragalactique à ces énergies : le blazar 3C 273 [Swanenburg et al., 1978]. Le *Compton Gamma-Ray Observatory* [Kanbach et al., 1988] actif de 1991 à 2000 était équipé entre autres du détecteur *Energetic Gamma-Ray Experiment Telescope* (*EGRET*). Ce dernier possédait une surface efficace de détection de l'ordre de $\sim$1500 cm^2 et était doté d'une précision angulaire de $\sim$1°. Il a permis d'obtenir la première carte du ciel γ entre 20 MeV et 30 GeV. Un catalogue contenant 271 sources a été publié, dont 60 % ne possédaient pas de contreparties à d'autres longueurs d'onde [Hartman et al., 1999]. La majorité des sources identifiées étaient des sources extragalactiques et l'expérience a permis l'étude fine de 5 pulsars.

Aujourd'hui, deux satellites observant dans le domaine γ de haute énergie sont actifs. Le satellite *Astro-rivelatore Gamma a Immagini LEggero* (*AGILE*), mis en orbite en juin 2007, est équipé du télescope *Grid*. Ce dernier utilise une technologie similaire à celle du *LAT* mais possède une sensibilité équivalente à celle de *EGRET*, de par sa surface de collection ($\sim$500 cm^2) et la faible profondeur de son calorimètre. Néanmoins, il possède un champ de vue six fois plus important que celui d'*EGRET* et une précision angulaire deux fois supérieure. Un catalogue de 47 sources a été publié, composé majoritairement de pulsars et de blazars [Pittori et al., 2009]. Le successeur de *EGRET*, le *LAT* embarqué sur le satellite *Fermi* a permis de révolutionner la vision du ciel γ grâce à sa sensibilité inégalée. La prochaine section lui est dédiée.

1.4.2 Le satellite *Fermi-LAT*

Mis en orbite en juin 2008, en trois mois de fonctionnement le *LAT* a permis la détection de 207 sources avec un haut niveau de significativité [Abdo et al., 2009a]. Avec une surface efficace de collection de $\sim$1 m^2, il est doté d'une sensibilité dix fois plus importante que l'instrument *EGRET*. Son champ de vue >2 sr lui permet de monitorer l'ensemble du ciel et donc d'observer des phénomènes transients et d'étudier la variabilité des sources d'une manière inégalée. À titre d'illustration, on peut observer une image du ciel au delà de 1 GeV vue par le *Fermi-LAT* sur la Figure 1.5. La bande centrale très lumineuse représente l'émission γ du plan galactique. À plus haute latitude galactique, les sources extragalactiques composées essentiellement de noyaux actifs de galaxie dominent le ciel.

À ce jour, l'équipe du *Fermi-LAT* a fourni un catalogue de 1873 sources détectées au delà de 100 MeV pendant une période de deux ans, le catalogue 2FGL [Nolan et al., 2012]. Différentes classes d'objets galactiques ont été détectées par le *Fermi-LAT* : restes de supernovæ, plérions, amas globulaire, systèmes binaires X, et notamment des pulsars. Ces derniers représentent la plus grande fraction des objets galactiques détectés (environ 55 % des $\sim$200 des sources galactiques), et la moitié d'entre eux sont de nouvelles détections [Abdo et al., 2013]. Parmi ces objets, les pulsars dits milliseconde constituent une nouvelle classe d'objets qui n'avaient pas encore été observés à ces énergies. Le nombre important de pulsars γ permet d'avancer de manière significative dans la contrainte des modèles d'émission, afin de cibler notamment dans quelle partie de ces objets l'émission γ est produite [Harding, 2007]. Notons aussi la découverte de l'émission γ des novæ également mise en évidence [The Fermi-LAT Collaboration, 2014].

Les sources extragalactiques, majoritairement des blazars, représentent plus de 80 % des sources associées ou identifiées du catalogue 2FGL, ce qui fait de ce satellite un détecteur de blazars par excellence. On trouvera une discussion à propos des AGN détectés à haute énergie à la section 2.6.1. Notons aussi que le satellite est équipé de l'instrument *Glast Burst Monitor* (*GBM*), facilitant la traque des phénomènes transients de type sursaut γ. On compte à ce jour 35 de ces objets détectés par le *LAT* [Ackermann et al., 2013].

Parmi les sources détectées à haute énergie par le *Fermi-LAT*, 1/3 des sources (575) n'ont pas de contrepartie à d'autres longueurs d'onde. Nous reviendrons sur cette problématique dans la Partie III de ce manuscrit qui est dédiée à l'identification des blazars parmi ces sources non-identifiées.

Le satellite *Fermi* dont le temps de fonctionnement optimal est estimé à 10 ans, continuera à apporter des données indispensables à une meilleure compréhension des phénomènes astrophysiques

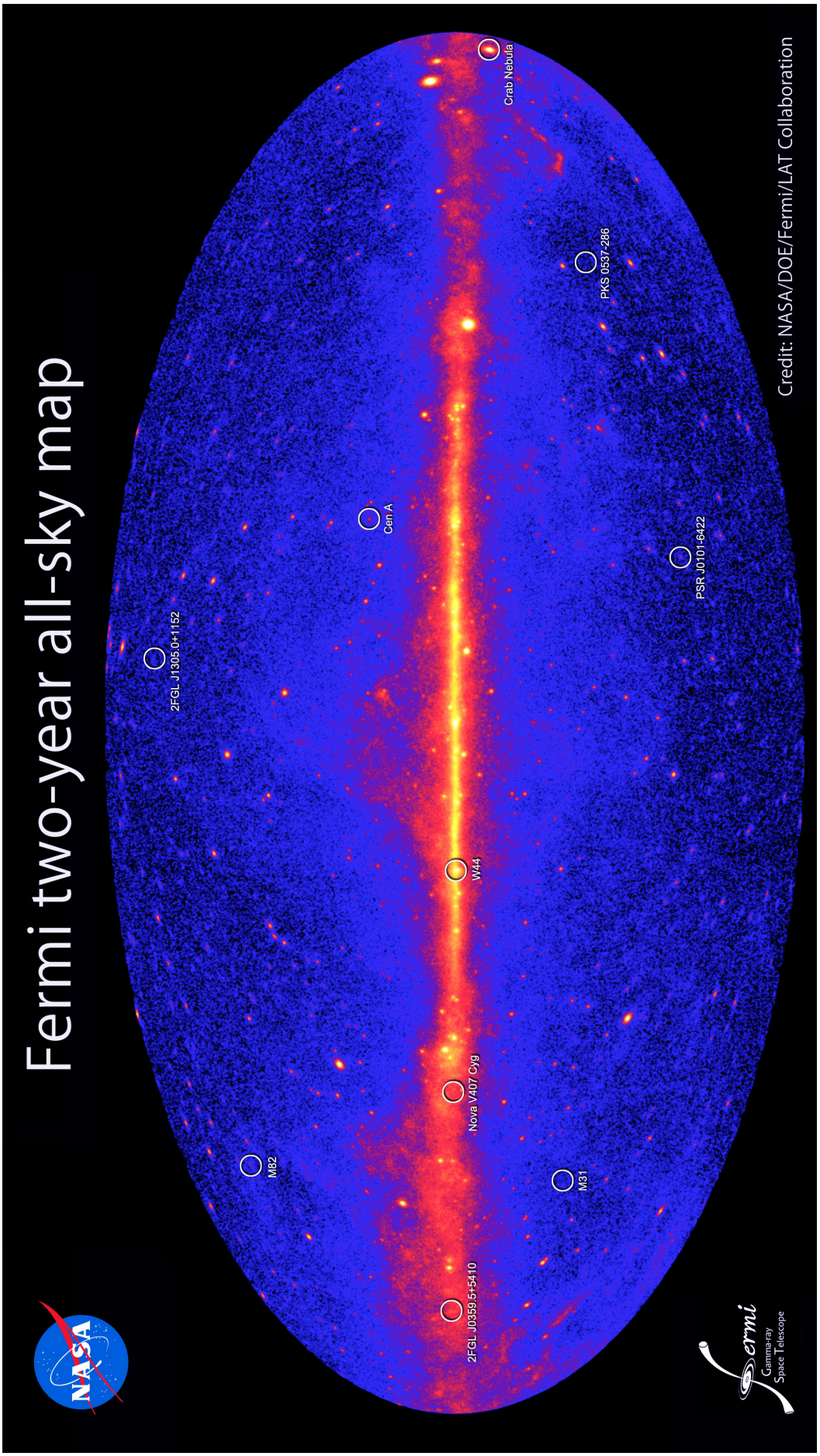


FIGURE 1.5: Vue du ciel γ par l'expérience *Fermi-LAT* en coordonnées galactique (Crédits : NASA).

dans le domaine γ. Il est aussi aujourd'hui d'une très grande utilité aux expériences au sol utilisant des télescopes d'imagerie atmosphérique Tcherenkov, afin de cibler les recherches de nouvelles sources émettrices dans le domaine du TeV.

1.5 Les expériences d'astronomie γ de très haute énergie

Le flux de photons de très haute énergie ($E \geq$100 GeV) suivant souvent en première approximation une loi de puissance, leur rareté rend la détection par les satellites de ces derniers très ardue à cause de leur faible surface de collection. Pour remédier à ce problème, l'utilisation de télescopes au sol se servant de l'atmosphère comme d'un calorimètre afin de détecter les produits secondaires des gerbes EM permet d'augmenter considérablement la surface efficace de détection. Après une brève présentation de l'historique du domaine, nous présentons les trois principales expériences contemporaines utilisant la technique d'imagerie atmosphérique Tcherenkov (présentée en détail au chapitre 4). Pour finir, l'expérience *Cherenkov Telescope Array* (*CTA*) qui représente la prochaine génération d'expériences d'astronomie γ de très haute énergie est présentée.

1.5.1 Un bref historique

Les premiers pas vers la détection du rayonnement γ de très haute énergie ont eu lieu en 1953 [Galbraith et Jelley, 1953] avec un simple photomultiplicateur dans lequel étaient focalisés à l'aide d'un miroir parabolique les photons Tcherenkov produis par des gerbes atmosphériques. Cet appareillage simple permettait de mettre en évidence les gerbes initiées par des rayons cosmiques, mais pas l'émission de γ provenant d'une source, ces derniers étant en effet noyés dans le bruit ambiant.

Il a fallu attendre l'avènement du télescope *Whipple*, construit en 1968, afin d'obtenir la première détection de la nébuleuse Crabe, 21 ans plus tard [Weekes et al., 1989]. En effet, le développement du matériel et des techniques d'analyse nécessaires afin d'isoler le signal ténu des rayons γ dans l'immense bruit généré par les cascades de particules initiées par des rayons cosmiques a nécessité un long parcours. L'utilisation d'une caméra dotée de 37 pixels de 0,25° de diamètre [Cawley et al., 1985], d'un réflecteur principal de 10 m ainsi que d'une analyse permettant de discriminer les images initiées par des γ de celles des rayons cosmiques [Hillas, 1985] a permis de détecter pour la première fois la nébuleuse du Crabe au-delà de 700 GeV. Deux ans plus tard, la première source extragalactique, le blazar Mkn 421, est détecté avec un niveau de détection de 6σ en $\sim$8 h [Punch et al., 1992].

En 1996, l'expérience *Cherenkov Array at Thémis* (*CAT*) [Barrau et al., 1998] utilise un réflecteur de 5 m de diamètre et une caméra dotée d'une fine pixellisation avec 500 photomultiplicateurs de 0,12° de diamètre, et d'une électronique rapide. Ces deux dernières caractéristiques permettent de baisser le seuil en énergie de l'instrument aux alentours de $\sim$250 GeV.

En 1999, pour la première fois, l'expérience *High Energy Gamma Ray Astronomy* (*HEGRA*) [HEGRA Collaboration et al., 1999] utilise plusieurs télescopes en coïncidence afin de détecter les rayons γ. Cette technique, appelée stéréoscopie, permet de mieux contraindre les caractéristiques du photon γ incident mais surtout de réduire de manière très significative le fond hadronique en exploitant l'asymétrie des gerbes qui en résultent. Cinq télescopes constituent le réseau, chacun doté d'un réflecteur de 3 m de diamètre et d'une caméra de 271 photomultiplicateurs de 0,25° de diamètre.

La génération actuelle de télescopes d'imagerie atmosphérique Tcherenkov (voir section 1.5.2) a profité de l'ensemble des points forts de cette ancienne génération.

Notons qu'il existe deux autres méthodes permettant de détecter des γ de haute énergie, sur lesquelles nous ne reviendrons pas par la suite. La première détecte les particules arrivant au sol à l'aide de photomultiplicateurs, de piscines ou de cuves d'eau (*High-Altitude Water Čerenkov* (*HAWC*) [Dingus, 2007] qui prend la suite de *MILAGRO* [Atkins et al., 2000], cette dernière expérience s'est arrêtée en 2008). Cette technique permet d'observer la totalité du ciel jour et nuit. Ses principaux

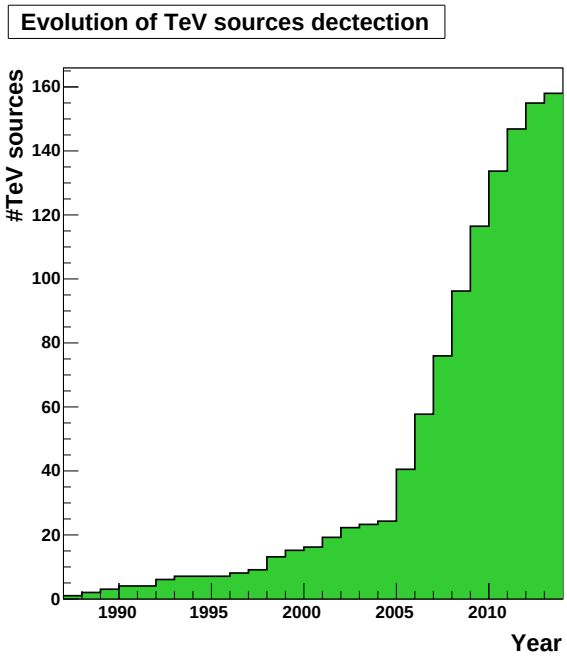


FIGURE 1.6: Évolution du nombre de sources détectées au TeV.

inconvénients sont le faible taux de rejet du fond hadronique et le haut seuil de détection de ∼1 TeV, ce dernier étant dû au fait que les gerbes EM doivent arriver jusqu'au sol. La seconde méthode est l'échantillonnage qui reconstruit le front d'onde Tcherenkov sphérique de la gerbe (à une dizaine de GeV) à l'aide de nombreux miroirs et peut ainsi déterminer les caractéristiques du photon incident (*CELESTE* [de Naurois et CELESTE Collaboration, 2001], *STACEE* [Hanna et al., 2002]).

1.5.2 La génération actuelle d'imageurs Tcherenkov

La génération d'expériences contemporaines utilisant la technique d'imagerie atmosphérique Tcherenkov comprend les trois expériences *H.E.S.S.*, *Very Energetic Radiation Imaging Telescope Array* (*VERITAS*) et *Major Atmospheric Gamma-ray Imaging Cherenkov Telescope* (*MAGIC*). Elles tirent profit des enseignements provenant de la génération précédente sur l'apport de la stéréoscopie et sur la nécessité de l'utilisation de grandes surfaces de miroirs et d'une électronique à la nanoseconde. En 2004, l'expérience *H.E.S.S.*, [Hinton, 2004] située en Namibie, utilisait un réseau composé de quatre télescopes de 13 m, chacun équipé d'une caméra dotée de 960 photomultiplicateurs, pour observer le ciel de l'hémisphère sud. La seconde phase de l'expérience, appelée HESS-II, a été inaugurée en 2012 lors de l'installation d'un cinquième télescope de 28 m de diamètre avec une caméra possédant 2048 pixels, afin de réduire l'énergie seuil de l'instrument (∼50 GeV). Nous reviendrons sur la description de cette expérience au chapitre 4. L'expérience *VERITAS* possède une sensibilité comparable à la première phase de l'expérience *H.E.S.S.*, elle utilise un ensemble de quatre télescopes de 12 m de diamètre dont l'installation s'est terminée en 2007 [Holder et al., 2008]. Située aux États-Unis, elle scrute essentiellement le ciel de l'hémisphère nord. L'expérience *MAGIC* a fait le choix d'utiliser un grand télescope de 17 m afin d'obtenir une énergie seuil de l'ordre de ∼50 GeV dans le but de détecter des sursauts γ [Cortina et al., 2005]. Celui-ci a été installé en 2003 et a été rejoint en 2009 par un second télescope de mêmes dimensions afin de pouvoir bénéficier de la technique de stéréoscopie [Moralejo, 2009]. Cette expérience est située sur l'île de la Palma (Canaries) à une latitude de ∼18°.

On doit à ces trois expériences l'émergence d'une discipline mature. On compte aujourd'hui 158 sources γ émettant dans le domaine du TeV [3]. L'augmentation significative du nombre de sources

3. http://tevcat.uchicago.edu

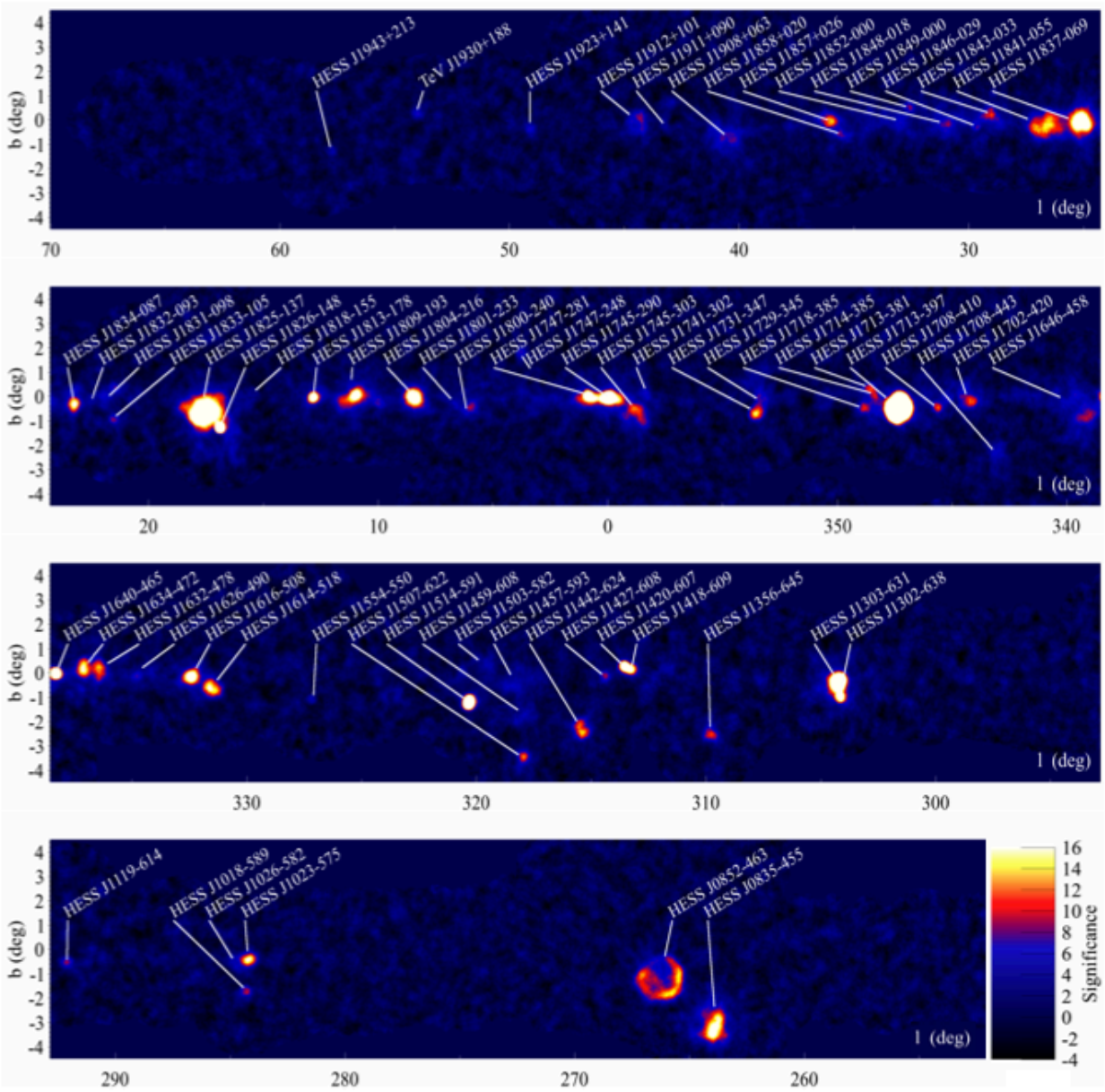


FIGURE 1.7: Plan galactique vue par l'expérience *H.E.S.S.* [Carrigan et al., 2013].

détectées autour de 2005, que l'on peut constater sur la Figure 1.6, correspond aux mises en service de ces trois expériences. Parmi ces découvertes figurent cinq types de sources galactiques, les restes de supernovæ, les plérions, les systèmes binaires, les amas stellaires et l'émission pulsée des pulsars, ainsi que de l'émission diffuse le long du plan galactique. Le relevé du plan galactique par l'expérience *H.E.S.S.* a permis la détection d'une soixantaine de sources [Carrigan et al., 2013]. Une fraction de ces sources (∼30 %) reste aujourd'hui de nature inconnue. L'émission γ pulsée du pulsar de la nébuleuse du Crabe a été détectée par l'expérience *MAGIC* en 2008 [Aliu et al., 2008] et récemment, l'émission du pulsar Véla a été détectée par le cinquième grand télescope du réseau HESS-II [4].

En ce qui concerne les noyaux actifs de galaxie, depuis la première détection de Mkn 421 par l'expérience *Whipple* [Punch et al., 1992], on compte aujourd'hui près de 60 noyaux actifs de galaxie détectés au-delà de 100 GeV, dont la grande majorité appartiennent à la classe des blazars. On trouvera une discussion à propos des AGN détectés à très haute énergie à la section 2.6.2.

1.5.3 La relève

La relève des expériences contemporaines d'imagerie atmosphérique Tcherenkov *H.E.S.S.*, *MAGIC* et *VERITAS* sera assurée par l'observatoire *CTA* [Actis et al., 2011] prévu pour être totalement opérationnel à l'horizon 2020. Deux réseaux hybrides, l'un au nord et l'autre au sud, constitués de différents types de télescopes, permettront d'une part d'observer des évènements de très basse énergie (∼10 GeV) à l'aide de grands télescopes munis de réflecteurs de 24 m de diamètre, mais aussi d'avoir accès aux rayons γ très énergétiques en couvrant une large surface de petits et de moyens télescopes (respectivement 6 m et 12 m de diamètre), tous dotés d'un champ de vue important ($\geq 6°$). Ces caractéristiques vont permettre à l'observatoire *CTA* d'atteindre un gain en sensibilité d'un facteur 5 à 10 à toutes les énergies par rapport aux expériences qui sont aujourd'hui en fonctionnement. Les premières données techniques devraient être disponibles en 2017.

4. http://www2.cnrs.fr/presse/communique/3651.htm

Bibliographie

A. A. Abdo, et al. Fermi/Large Area Telescope Bright Gamma-Ray Source List. ApJS, 183:46–66, July 2009a. doi: 10.1088/0067-0049/183/1/46.

A. A. Abdo, et al. The Large-Scale Cosmic-Ray Anisotropy as Observed with Milagro. ApJ, 698: 2121–2130, June 2009b. doi: 10.1088/0004-637X/698/2/2121.

A. A. Abdo, et al. The Second Fermi Large Area Telescope Catalog of Gamma-Ray Pulsars. ApJS, 208:17, October 2013. doi: 10.1088/0067-0049/208/2/17.

J. Abraham, et al. Observation of the Suppression of the Flux of Cosmic Rays above 41019eV. *Physical Review Letters*, 101(6):061101, August 2008. doi: 10.1103/PhysRevLett.101.061101.

M. Ackermann, et al. The First Fermi-LAT Gamma-Ray Burst Catalog. ApJS, 209:11, November 2013. doi: 10.1088/0067-0049/209/1/11.

M. Actis, et al. Design concepts for the Cherenkov Telescope Array CTA: an advanced facility for ground-based high-energy gamma-ray astronomy. *Experimental Astronomy*, 32:193–316, December 2011. doi: 10.1007/s10686-011-9247-0.

M. Aguilar, et al. First Result from the Alpha Magnetic Spectrometer on the International Space Station: Precision Measurement of the Positron Fraction in Primary Cosmic Rays of 0.5-350 GeV. *Physical Review Letters*, 110(14):141102, April 2013. doi: 10.1103/PhysRevLett.110.141102.

E. Aliu, et al. Observation of Pulsed γ-Rays Above 25 GeV from the Crab Pulsar with MAGIC. *Science*, 322:1221–, November 2008. doi: 10.1126/science.1164718.

R. Atkins, et al. Milagrito, a TeV air-shower array. *Nuclear Instruments and Methods in Physics Research A*, 449:478–499, July 2000. doi: 10.1016/S0168-9002(00)00146-7.

A. Barrau, et al. The CAT imaging telescope for very-high-energy gamma-ray astronomy. *Nuclear Instruments and Methods in Physics Research A*, 416:278–292, October 1998. doi: 10.1016/S0168-9002(98)00749-9.

A. R. Bell. The acceleration of cosmic rays in shock fronts. I. MNRAS, 182:147–156, January 1978.

G. F. Bignami, et al. The COS-B experiment for gamma-ray astronomy. *Space Science Instrumentation*, 1:245–268, August 1975.

J. Blümer, R. Engel, et J. R. Hörandel. Cosmic rays from the knee to the highest energies. *Progress in Particle and Nuclear Physics*, 63:293–338, October 2009. doi: 10.1016/j.ppnp.2009.05.002.

S. Carrigan, et al. The H.E.S.S. Galactic Plane Survey - maps, source catalog and source population. *ArXiv e-prints*, July 2013.

M. F. Cawley, et al. Application of imaging to the atmospheric Cherenkov technique. *International Cosmic Ray Conference*, 3:453–456, August 1985.

J. Cortina, et al. Technical Performance of the MAGIC Telescope. *International Cosmic Ray Conference*, 5:359, 2005.

M. de Naurois et CELESTE Collaboration. Status and Current Sensitivity of the CELESTE Experiment. In F. A. Aharonian et H. J. Völk, editors, *American Institute of Physics Conference Series*, volume 558 of *American Institute of Physics Conference Series*, page 540, April 2001.

C. Dermer et G. Menon. *High energy radiation from black holes: gamma rays, cosmic rays, and neutrino*. Princeton series in astrophysics. Princeton University Press, 2009.

B. L. Dingus. HAWC (High Altitude Water Cherenkov) Observatory for Surveying the TeV Sky. In S. Ritz, P. Michelson, et C. A. Meegan, editors, *The First GLAST Symposium*, volume 921 of *American Institute of Physics Conference Series*, pages 438–439, July 2007. doi: 10.1063/1.2757390.

T. Ebisuzaki, et al. The JEM-EUSO Project: Observing Extremely High Energy Cosmic Rays and Neutrinos from the International Space Station. *Nuclear Physics B Proceedings Supplements*, 175: 237–240, January 2008. doi: 10.1016/j.nuclphysbps.2007.11.005.

E. Fermi. On the Origin of the Cosmic Radiation. *Physical Review*, 75:1169–1174, April 1949. doi: 10.1103/PhysRev.75.1169.

C. E. Fichtel, et al. High-energy gamma-ray results from the second small astronomy satellite. ApJ, 198:163–182, May 1975. doi: 10.1086/153590.

W. Galbraith et J. V. Jelley. Light Pulses from the Night Sky associated with Cosmic Rays. Nature, 171:349–350, February 1953. doi: 10.1038/171349a0.

D. S. Hanna, et al. The STACEE-32 ground based gamma-ray detector. *Nuclear Instruments and Methods in Physics Research A*, 491:126–151, September 2002. doi: 10.1016/S0168-9002(02)01126-9.

A. K. Harding. Pulsar High-Energy Emission From the Polar Cap and Slot Gap. *ArXiv e-prints*, October 2007.

R. C. Hartman, et al. The Third EGRET Catalog of High-Energy Gamma-Ray Sources. ApJS, 123:79–202, July 1999. doi: 10.1086/313231.

HEGRA Collaboration, et al. Performance of the stereoscopic system of the HEGRA imaging air Čerenkov telescopes: Monte Carlo simulations and observations. *Astroparticle Physics*, 10: 275–289, May 1999. doi: 10.1016/S0927-6505(98)00062-0.

A. M. Hillas. The Origin of Ultra-High-Energy Cosmic Rays. ARA&A, 22:425–444, 1984. doi: 10.1146/annurev.aa.22.090184.002233.

A. M. Hillas. Cerenkov light images of EAS produced by primary gamma. In F. C. Jones, editor, *International Cosmic Ray Conference*, volume 3 of *International Cosmic Ray Conference*, pages 445–448, August 1985.

J. A. Hinton. The status of the HESS project. New A Rev., 48:331–337, April 2004. doi: 10.1016/j.newar.2003.12.004.

J. Holder, et al. Status of the VERITAS Observatory. In F. A. Aharonian, W. Hofmann, et F. Rieger, editors, *American Institute of Physics Conference Series*, volume 1085 of *American Institute of Physics Conference Series*, pages 657–660, December 2008. doi: 10.1063/1.3076760.

K.-H. Kampert et M. Unger. Measurements of the cosmic ray composition with air shower experiments. *Astroparticle Physics*, 35:660–678, May 2012. doi: 10.1016/j.astropartphys.2012.02.004.

G. Kanbach, et al. The project EGRET (Energetic Gamma-Ray Experiment Telescope) on NASA's Gamma-Ray Observatory (GRO). Space Sci. Rev., 49:69–84, 1988.

A. Moralejo. Status of Magic-II. *ArXiv e-prints*, December 2009.

P. L. Nolan, et al. Fermi Large Area Telescope Second Source Catalog. ApJS, 199:31, April 2012. doi: 10.1088/0067-0049/199/2/31.

C. Pittori, et al. First AGILE catalog of high-confidence gamma-ray sources. A&A, 506:1563–1574, November 2009. doi: 10.1051/0004-6361/200911783.

K. V. Ptitsyna et S. V. Troitsky. PHYSICS OF OUR DAYS Physical conditions in potential accelerators of ultra-high-energy cosmic rays: updated Hillas plot and radiation-loss constraints. *Physics Uspekhi*, 53:691–701, October 2010. doi: 10.3367/UFNe.0180.201007c.0723.

M. Punch, et al. Detection of TeV photons from the active galaxy Markarian 421. Nature, 358:477, August 1992. doi: 10.1038/358477a0.

G. B. Rybicki et A. P. Lightman. *Radiative Processes in Astrophysics*. Wiley-VCH, June 1986.

B. N. Swanenburg, et al. COS B observation of high-energy gamma radiation from 3C273. Nature, 275:298, September 1978. doi: 10.1038/275298a0.

The Fermi-LAT Collaboration. Fermi Establishes Classical Novae as a Distinct Class of Gamma-Ray Sources. *ArXiv e-prints*, August 2014.

T. C. Weekes, et al. Observation of TeV gamma rays from the Crab nebula using the atmospheric Cerenkov imaging technique. ApJ, 342:379–395, July 1989. doi: 10.1086/167599.

Chapitre 2

Noyaux actifs de galaxies

Sommaire

2.1 Introduction 21
2.2 Différentes classes d'AGN 21
2.3 Schéma d'unification des AGN 24
2.4 Distribution spectrale des blazars 25
2.5 Séquence des blazars 29
2.6 AGN observés dans le domaine γ 32
Bibliographie 35

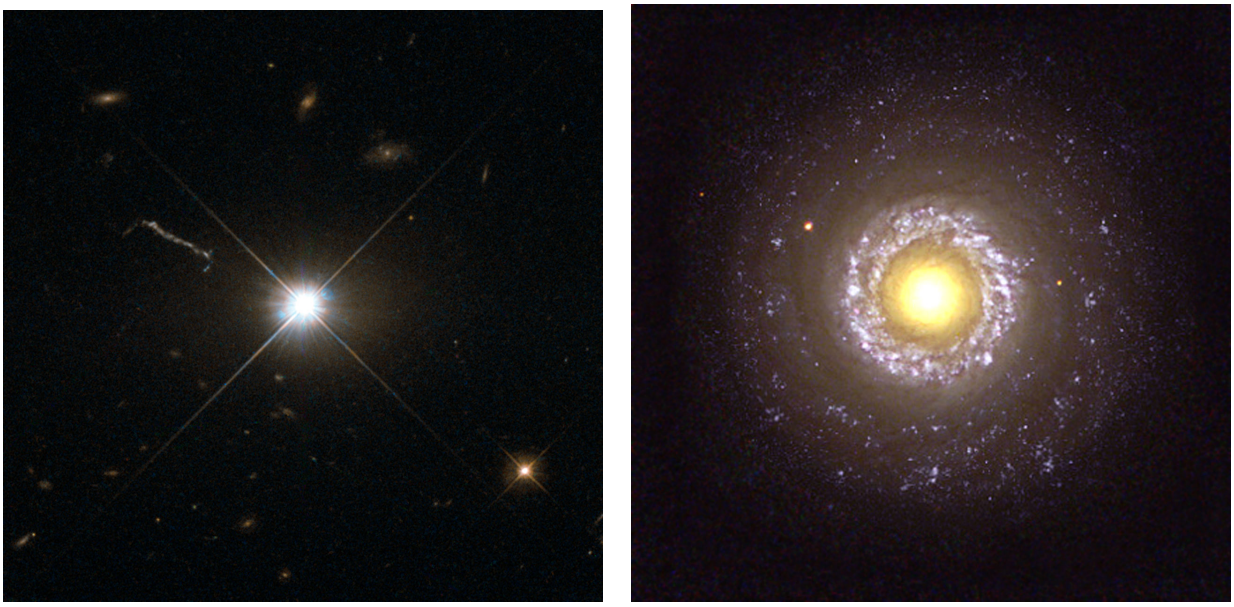
2.1 Introduction

Certaines galaxies présentent en leur cœur un noyau qui montre des phénomènes très énergétiques qui ne peuvent pas être attribués aux étoiles. Ceux-ci peuvent atteindre une luminosité bolométrique (intégrée sur toutes les fréquences) jusqu'à 10 000 fois plus grande que la luminosité typique d'une galaxie ($L \approx 10^{44}\,\mathrm{erg\,s^{-1}}$). Ils ont pour caractéristiques d'émettre dans une large gamme de fréquence (infrarouge, optique, ultraviolet et X), et certains d'entre eux ont la particularité d'émettre dans les domaines radio et γ. Une bonne partie des AGN présentent une forte variabilité ce qui permet de contraindre, avec des arguments de causalité, la taille caractéristique de la région d'émission qui correspond à moins d'un dixième de parsec. Ces différentes observations ont conduit à supposer l'existence d'un trou noir supermassif au cœur des galaxies qui serait alimenté par un disque d'accrétion. En effet l'accrétion de matière par un tel objet est aujourd'hui le seul procédé capable d'expliquer les luminosités colossales concentrées dans un volume aussi faible.

Dans ce chapitre, nous commençons par présenter les principales classes de noyaux actifs de galaxies à la section 2.2 ainsi qu'une présentation rapide du schéma d'unification à la section 2.3. À la section 2.4, une brève présentation de la distribution spectrale des blazars est présentée, avant de faire un tour d'horizon des développements récents autour de la question de l'existence de la séquence des blazars à la section 2.5. On présentera à la section 2.6 un état des lieux sur l'observation des AGN dans le domaine γ.

2.2 Différentes classes d'AGN

La très grande variété des caractéristiques observationnelles des AGN, telles que leur émission et leur morphologie dans le domaine radio ou la présence de raies d'émission dans leur spectre optique, a donné naissance à une classification historique qui peut dérouter. Aujourd'hui l'appellation quasar (contraction de « quasi-stellar radio source ») est tout aussi utilisée que l'appellation AGN, bien qu'ils ne présentent pas tous une émission radio observable. Celle-ci désignait originellement des noyaux actifs de galaxies très brillants en radio et dont l'émission optique surpassait l'émission de la galaxie hôte (Figure 2.1a). Nous passons ici brièvement en revue les principales terminologies utilisées pour les AGN ainsi que leurs caractéristiques observables. Le lecteur désireux d'en savoir plus pourra se référer par exemple à [Krolik, 1999].



(a) (b)

FIGURE 2.1: (a) Quasar 3C 273 situé à une distance d'environ 750 Mpc et (b) galaxie Seyfert NGC 7742 qui se trouve à environ 22 Mpc (Crédits : http://www.spacetelescope.org). Les deux images on été réalisés avec le télescope Hubble.

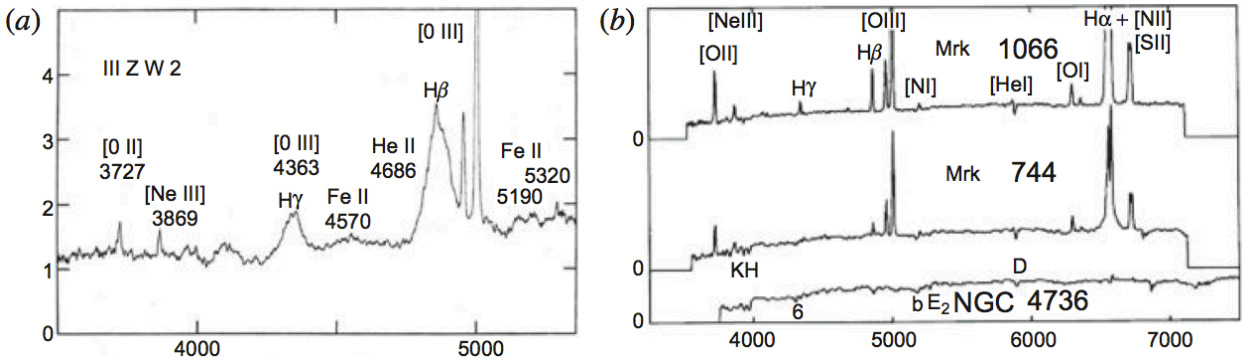
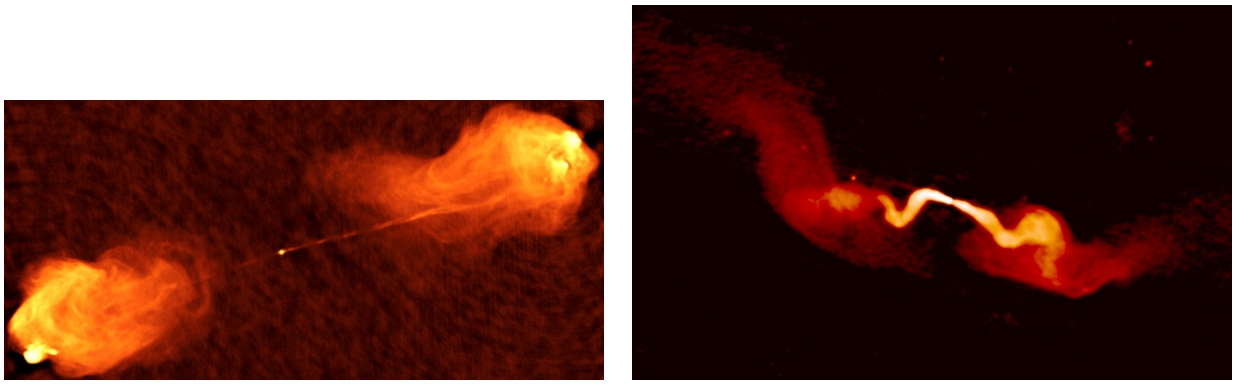


FIGURE 2.2: À gauche, spectre optique de la galaxie de Seyfert 1 III ZW 2 qui présente des raies d'émission larges d'hydrogène. On peut aussi observer des raies étroites correspondant à de l'oxygène et du néon. À droite, les spectres des galaxies Mrk 1066 et Mrk 744 de type Seyfert 2 sont représentés, sur lesquels on peut observer des raies d'émission étroites. Le spectre de la galaxie non-active NGC 4736 est aussi représenté à titre d'indication (Crédits : [Longair, 2011]).



(a) (b)

FIGURE 2.3: (a) Radiogalaxie Cygnus A de type FR II située à environ 230 Mpc et (b) radiogalaxie NGC 383 de type FR I qui se trouve à environ 64 Mpc (Crédits : http://images.nrao.edu/). Ces images ont été obtenues avec le VLA à une fréquence de 5 GHz.

Une première distinction concerne les objets qui sont silencieux ou bruyants dans le domaine radio [1]. Les premiers sont constitués de galaxies de Seyfert et de QSO [2] (« quasi-stellar object »). Les galaxies de Seyfert [Seyfert, 1943] sont pour la plupart des galaxies spirales qui présentent une faible luminosité bolométrique, la galaxie hôte étant généralement détectable en optique (Figure 2.1b). Elles sont différenciées selon que leur spectre d'émission optique présente des raies larges et brillantes (type 1) ou exclusivement des raies étroites et peu brillantes (type 2). À titre d'illustration, on peut trouver sur la gauche de la Figure 2.2 un spectre optique d'une galaxie Seyfert 1, et sur la droite deux spectres de galaxies de Seyfert 2 ainsi que le spectre d'une galaxie non-active donné à titre de comparaison. Les noyaux actifs dénommés QSO sont également silencieux en radio et présentent une forte émission dans le domaine optique ainsi que des raies d'émission larges.

Parmi les objets bruyants en radio, on trouve les radiogalaxies et les blazars. On distingue les radiogalaxies de type FR-I et FR-II [Fanaroff et Riley, 1974] qui sont classées selon leur brillance et la morphologie de leurs jets dans le domaine radio. Les radiogalaxies de type FR-II sont les plus brillantes en radio. Elles présentent des jets relativistes très collimatés qui se terminent en formant de gigantesques structures radio dans lesquelles on peut observer des points chauds signalant l'interaction des jets avec le milieu environnant (Figure 2.3a). Leur spectre optique présente des raies d'émission larges et étroites. A contrario, l'émission radio des radiogalaxies de type FR-I est concentrée près du noyau actif et les jets sont moins collimatés que ceux des radiogalaxies de type FR-II (Figure 2.3b). Seules des raies d'émission étroites sont observées dans leur spectre optique.

Parmi les émetteurs radios, on distingue aussi les blazars [3] qui sont des AGN dont le jet relativiste est quasi-aligné avec notre ligne de visée. Ils bénéficient d'un important effet Doppler relativiste qui se traduit par un décalage des spectres vers les hautes énergies et par l'amplification de leur luminosité. Les blazars sont divisés en deux catégories, les BL Lacertae (BL Lac) et les *Flat Spectrum Radio Quasar* (FSRQ) [4], selon l'absence ou la présence, respectivement, de raies d'émissions larges (EW > 5 Å [5]).

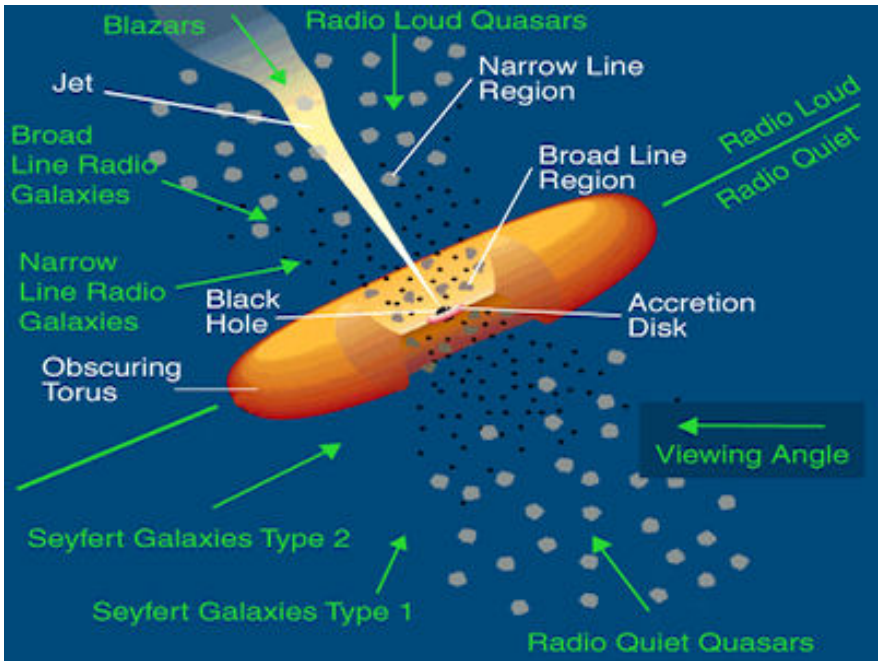


FIGURE 2.4: Schéma de l'unification des différents types d'AGN (Crédits : Marek A. Abramowicz).

2.3 Schéma d'unification des AGN

Certaines caractéristiques observationnelles des AGN semblent communes aux différentes classes, ce qui a conduit à la recherche d'un modèle sous-jacent. Proposé par [Urry et Padovani, 1995], l'unification des différentes classes de noyaux actifs de galaxies part de l'hypothèse que les disparités observées dans leurs caractéristiques sont dues à des biais observationnelles. Ou autrement dit que ce sont les mêmes objets mais observés sous un angle différent. Le modèle d'unification utilise les ingrédients principaux suivants [Ghisellini, 2012] :

— un trou noir supermassif d'une masse comprise entre $10^6\,M_\odot$ et $10^{10}\,M_\odot$

— un disque d'accrétion constituant le réservoir d'énergie principal

— une couronne de rayons X entourant le disque d'accrétion

— un tore de poussière situé à quelques parsecs du trou noir qui intercepte une fraction du rayonnement du disque et qui le réémet dans l'infrarouge

— des nuages de gaz se déplaçant rapidement à une distance du trou noir inférieure au parsec et responsables des raies d'émission larges

— des nuages de gaz avec des vitesses moins importantes à une centaine de parsec du trou noir et responsables des raies d'émission étroites

— pour ∼10 % des AGN, deux jets composés de matière ayant des vitesses d'ensemble relativistes, dont l'orientation signale probablement l'axe de rotation du trou noir

Ce modèle simple permet d'expliquer les différences entre certaines classes d'objets, mais n'explique pas l'existence des deux catégories : radio-émetteur et radio-silencieux. Cette différentiation pourrait être liée aux caractéristiques de l'engin central ou à des conditions d'environnement liées à une

1. Un objet est classé radio-bruyant si le rapport entre son flux radio à 5 GHz et son flux optique est supérieur à 10.
2. Terme inventé à la suite de quasars pour qualifier les quasars à émission radio faible ou indétectable.
3. Le terme blazar vient de l'anglais terme « blaze » qui signifie en français « embrasement » ou « flamboiement ».
4. Aussi appelé « OOV » pour « Optically Violent Variable ».
5. EW est la largeur équivalente, c'est une mesure qualitative de l'intensité d'une raie.

distribution de la puissance des jets non-uniforme, qui pourrait freiner leur développement ou les rendre difficilements détectables.

On peut observer sur la Figure 2.4 une représentation du modèle d'unification des AGN où sont indiqués les composants présentés ci-dessus et les classes de noyaux actifs qui rendent compte de la position de l'observateur. D'après ce modèle, en plus de l'émission continue liée au disque d'accrétion, les objets vus de face (perpendiculairement au disque) vont présenter des raies d'émission larges et étroites. À l'inverse, les objets vus par la tranche présenteront uniquement des raies d'émission étroites (la zone d'émission des raies larges est cachée par le tore). Ils sont respectivement définis comme des AGN de type 1 et de type 2. Ceux pour lesquels la présence de raies n'est pas évidente sont appelés AGN de type 0 (typiquement les BL Lac). On associe généralement comme population parente aux blazars de type BL Lac et FSRQ respectivement les radiogalaxies de type FR-I et FR-II. L'orientation du jet permet de comprendre les différences d'émission entre les radiogalaxies et les blazars, mais ne permet cependant pas d'expliquer les différences intrinsèques entre les objets de type FSRQ et BL Lac.

Avant de présenter la tentative d'unification des blazars proposée par [Fossati et al., 1998] à la section 2.5, nous présentons les grandes lignes de l'émission spectrale des blazars dans la section suivante.

2.4 Distribution spectrale des blazars

La distribution spectrale en énergie (SED) des blazars à la particularité d'être bimodale et de recouvrir une grande partie du spectre électromagnétique (de la radio jusqu'au TeV dans certains cas). On peut observer sur la Figure 2.5 les SED typiques des FSRQ et des BL Lac (νF_ν). Ces derniers sont différenciés selon que la première bosse pique à haute énergie, ils sont appelés High-frequency peaked BL Lac (HBL), ou à basse énergie, ils sont appelés Low-frequency peaked BL Lac (LBL). Il existe aussi une population intermédiaire appelée Intermediate-frequency peaked BL Lac (IBL). Du fait des mesures de polarisation, il est généralement admis que la première bosse correspond à l'émission du rayonnement synchrotron (voir chapitre 1) des électrons relativistes du jet. Lorsque le champ de photons ainsi créé sert de cible aux électrons [6] relativistes via l'interaction Compton-inverse (voir chapitre 1), on parle de processus « Synchrotron Self Compton » (SSC). Celui-ci est considéré comme étant à l'origine de la seconde bosse à plus haute énergie. Dans le cas où un champ de photons externe sert de cible, on parle de processus « External Compton » (EC).

On peut voir sur la Figure 2.6 la distribution spectrale en énergie du blazar de type FSRQ 3C 454.3 obtenue lors de différentes observations au cours de l'année 2009 [7] (dates indiquées sur la figure), ou correspondant à des données d'archives [Bonnoli et al., 2011]. À haute énergie, les données de 2009 ont été obtenues avec le satellite *Fermi*-*LAT*, à plus basse énergie dans le domaine des X avec le satellite SWIFT/XRT, et enfin dans le domaine optique avec le satellite SWIFT/UVOT. Les différentes courbes pleines (chacune d'entre elles correspond à une époque) représentent la modélisation de l'émission du blazar avec les composantes suivantes [8] : disque d'accrétion, émission de la couronne X, émission du tore (traits pointillés sur la figure), émissions SSC et EC (un champ de photons externe est aussi considéré).

Une autre propriété des blazars est leur extrême variabilité observée à toutes les longueurs d'onde. À titre d'illustration, on peut observer sur la Figure 2.7 la courbe du lumière de 3C 454.3, qui représente le flux de la source en fonction du temps, dans le domaine optique (SWIFT/UVOT, en bas), le domaine des X (SWIFT/XRT, au milieu) et dans le domaine γ de haute énergie (*Fermi*-*LAT*, en haut). Les lignes jaunes correspondent aux époques des modélisations spectrales décrites

6. Il existe aussi des modèles hadroniques mais qui peinent à reproduire la variabilité rapide observée dans les blazars.

7. Durant la première semaine de décembre 2009, ce blazar était l'objet le plus brillant de tout le ciel γ.

8. On pourra trouver une très bonne introduction à tout ce qui est lié à la modélisation des processus d'émission des AGN dans [Ghisellini, 2012].

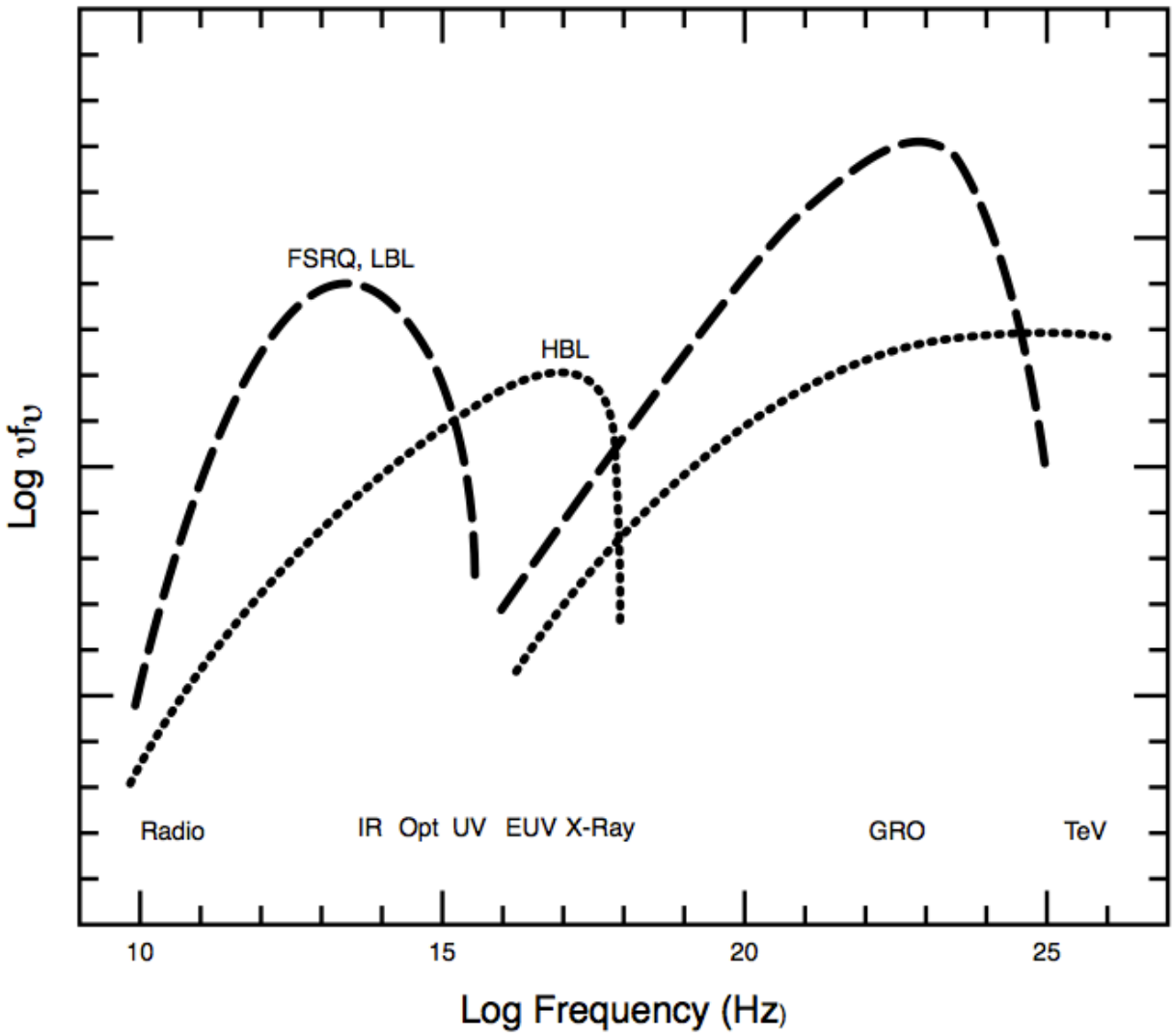


FIGURE 2.5: Distributions spectrales en énergie pour trois types de blazars explicités dans le texte(Crédits : [Urry, 1998]).

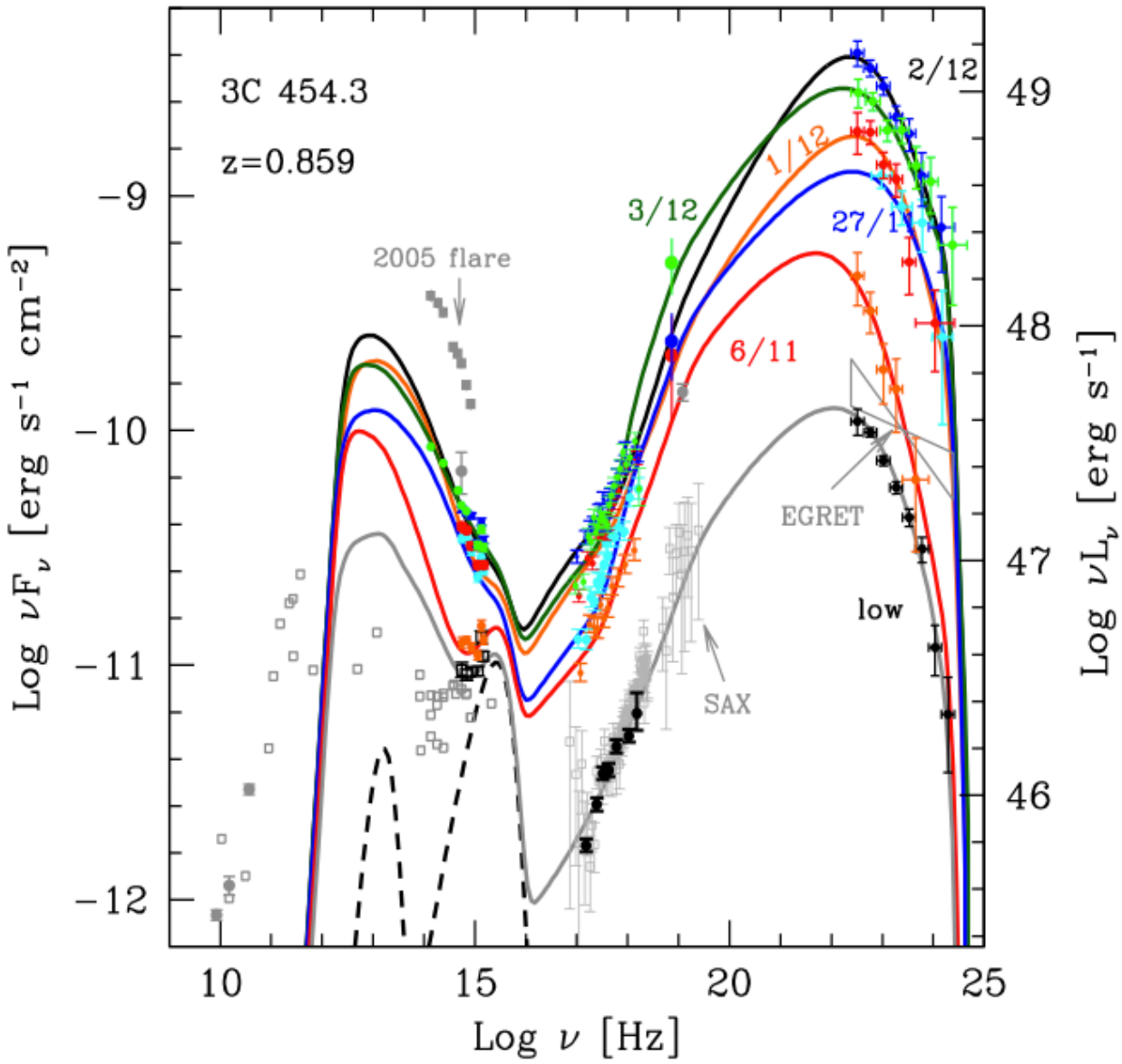


FIGURE 2.6: Distribution spectrale en énergie du FSRQ 3C 454.3 pour différentes époques (Crédits : [Bonnoli et al., 2011]). Les différentes époques correspondant à l'année 2009 sont indiquées sur la figure, et les données grisées correspondent à des données d'archives. Les différentes courbes correspondent à des modèles qui incluent les composantes suivantes : disque d'accrétion, émission de la couronne X, émission du tore (traits pointillés), processus EC et SSC (voir texte).

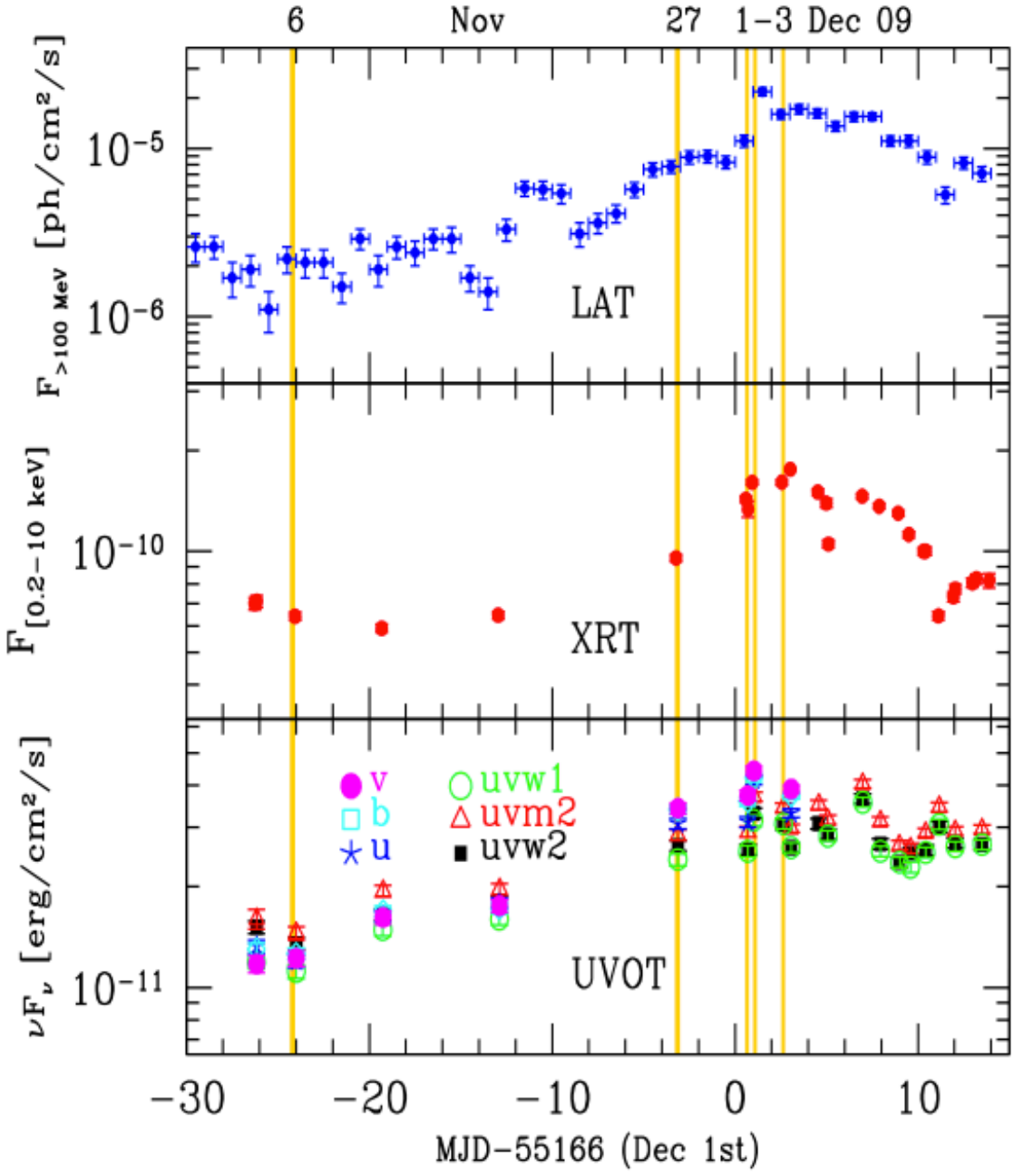


FIGURE 2.7: Courbe de lumière du blazar 3C 454.3 simultanée en optique, X et γ (Crédits : [Bonnoli et al., 2011]). Les différentes sous-figures correspondent au domaine optique (SWIFT/UVOT, en bas), au domaine des X (SWIFT/XRT, au milieu) et au domaine γ de haute énergie (*Fermi*-*LAT*, en haut). Les lignes jaunes correspondent aux époques des modélisations spectrales représentées sur la Figure 2.6.

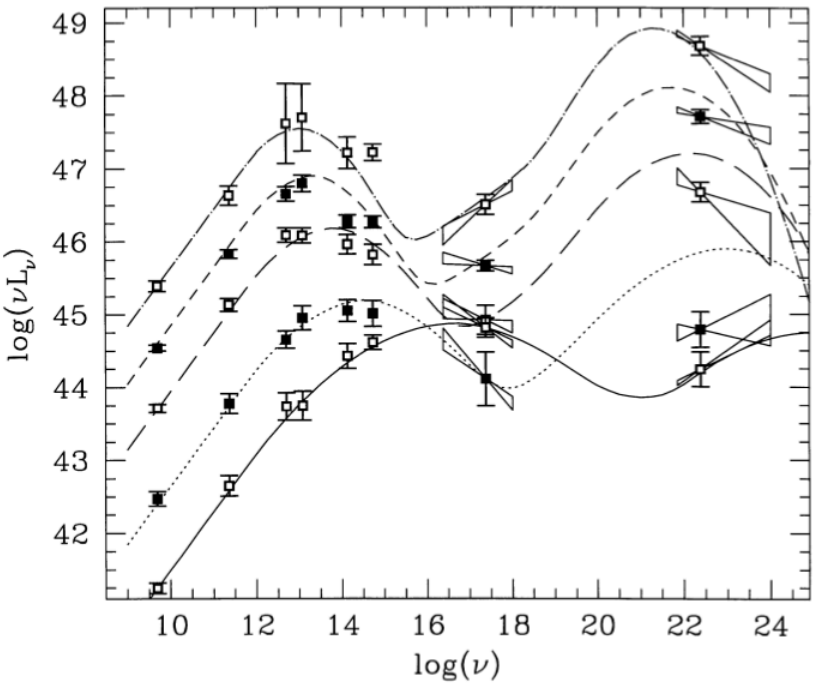


FIGURE 2.8: Séquence spectrale des blazars (Crédits : [Fossati et al., 1998]).

plus haut qui sont représentées sur la Figure 2.6.

2.5 Séquence des blazars

À l'aide d'un échantillon de 126 blazars, qu'il séparent en différents groupes en fonction de leur luminosité bolométrique, [Fossati et al., 1998] superposent les distributions spectrales moyennes obtenues. Ils observent que les positions des pics d'émission synchrotron et Compton-inverse sont corrélées avec la luminosité (Figure 2.8). Ils observent également que les FSRQ sont plus brillants, avec un maximum d'émission synchrotron situé entre les domaines radio et infrarouge, et que les BL Lac (de type HBL) possèdent une luminosité bolométrique moins importante et un pic d'émission synchrotron situé dans le domaine des X. Considérant que le modèle d'émission attribue la première bosse au rayonnement synchrotron des électrons accélérés dans les jets et la seconde bosse à un processus Compton-inverse, [Ghisellini et al., 1998] propose que les FSRQ possèdent un champ de photons externe et un champ magnétique plus importants que les BL Lac ce qui contribue à expliquer le fait que les électrons subissent des pertes d'énergie radiatives plus élevées dans les jets. Néanmoins, l'existence de FSRQ dont le pic synchrotron se trouve dans le domaine des X [Padovani et al., 2002] ainsi que le faible nombre de blazars utilisés dans cette étude (pouvant conduire à des biais de sélection) soulèvent des critiques quant à la validité de cette séquence [Padovani, 2007, Giommi et al., 2012b,a].

Plus récemment, [Ghisellini et Tavecchio, 2008] proposent que les différentes caractéristiques entre BL Lac et FSRQ s'expliquent par des différences sur la masse du trou noir central et le taux d'accrétion associé. Dans leur modèle, l'engin central des FSRQ possède un taux d'accrétion très important et entraîne la formation d'un disque qui fournit un champ de photons externe et un champ magnétique important, qui entraîne à son tour des pertes radiatives dans les jets. Les BL Lac posséderaient un taux d'accrétion très inférieur à la limite d'Eddington [9] ce qui limiterait les

9. Limite qui correspond à un équilibre entre la gravitation et la pression de radiation des photons sur le disque. Elle correspond à la valeur maximale du taux d'accrétion.

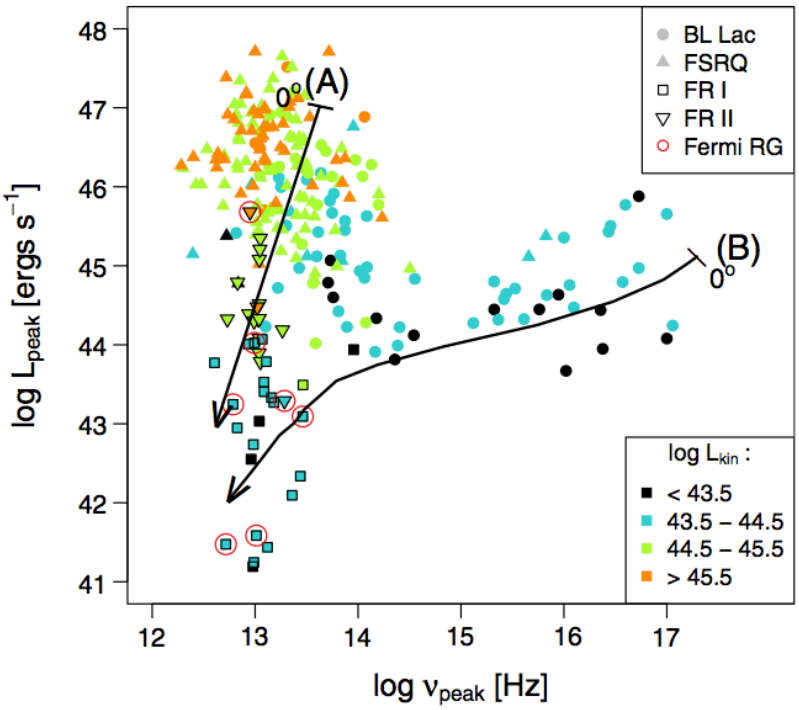


FIGURE 2.9: Évolution de la luminosité du pic synchrotron pour les BL Lac (ronds), les FSRQ (triangles) et les radiogalaxies FR-I (carrés) et FR-II (triangles inversés) (Crédits : [Meyer et al., 2011]).

pertes radiatives des électrons.

Un peu plus tard, [Meyer et al., 2011] proposent une séquence d'évolution selon l'alignement du jet avec notre direction de visée, à l'aide d'un échantillon de 216 blazars et de 41 radiogalaxies. Sur la Figure 2.9 sont représentées des BL Lac (ronds), des FSRQ (triangles) et des radiogalaxies FR-I (carrés) et FR-II (triangles inversés), selon leur luminosité radio intégrée et l'énergie du maximum d'émission synchrotron. Deux chemins sont envisagés, A et B. Le chemin A décrit l'évolution d'un objet puissant pour un non-alignement progressif du jet (FSRQ vers radiogalaxie FR-II). Le chemin B montre l'évolution d'un objet peu puissant pour un non-alignement progressif du jet (BL Lac vers radiogalaxie FR-I).

Parallèlement à ces développements, [Giommi et al., 2012a] étudient l'influence des biais de sélection dans les domaines radio et X sur la séquence des blazars proposée par [Fossati et al., 1998]. Ils proposent un modèle permettant de classer les blazars selon leur type (FSRQ, BL Lac). Un certain nombre d'ingrédients est utilisé tels qu'une distribution unique des luminosités, une distribution des largeurs des raies d'émission, une émission dans le jet selon le processus SSC, une émission thermique due au disque d'accrétion, et l'émission de la galaxie hôte (on pourra trouver plus détails dans la publication). Après optimisation et validation des simulations sur des données réelles, [Giommi et al., 2012a] étudient la corrélation entre la luminosité à 5 GHz et la position du maximum du pic synchrotron des blazars simulés. On peut observer sur la Figure 2.10 des FSRQ (cercles ouverts) et des BL Lac (carrés ouverts) simulés et correspondant à des objets sélectionnables en radio (rouge) et en X (vert) selon les critères utilisés par [Fossati et al., 1998]. Les cercles pleins représentent des BL Lac dont la largeur équivalente (EW) est inférieure à 2 Å, et qui dans la réalité ne pourraient pas avoir une mesure de redshift et seraient donc impossibles à placer sur la séquence des blazars de [Fossati et al., 1998]. [Giommi et al., 2012a] prédisent donc l'existence d'une population de blazars de haute luminosité bolométrique avec un pic synchrotron situé à haute énergie dont le redshift peut difficilement être déterminé spectroscopiquement dans le visible. Sur la base de leur

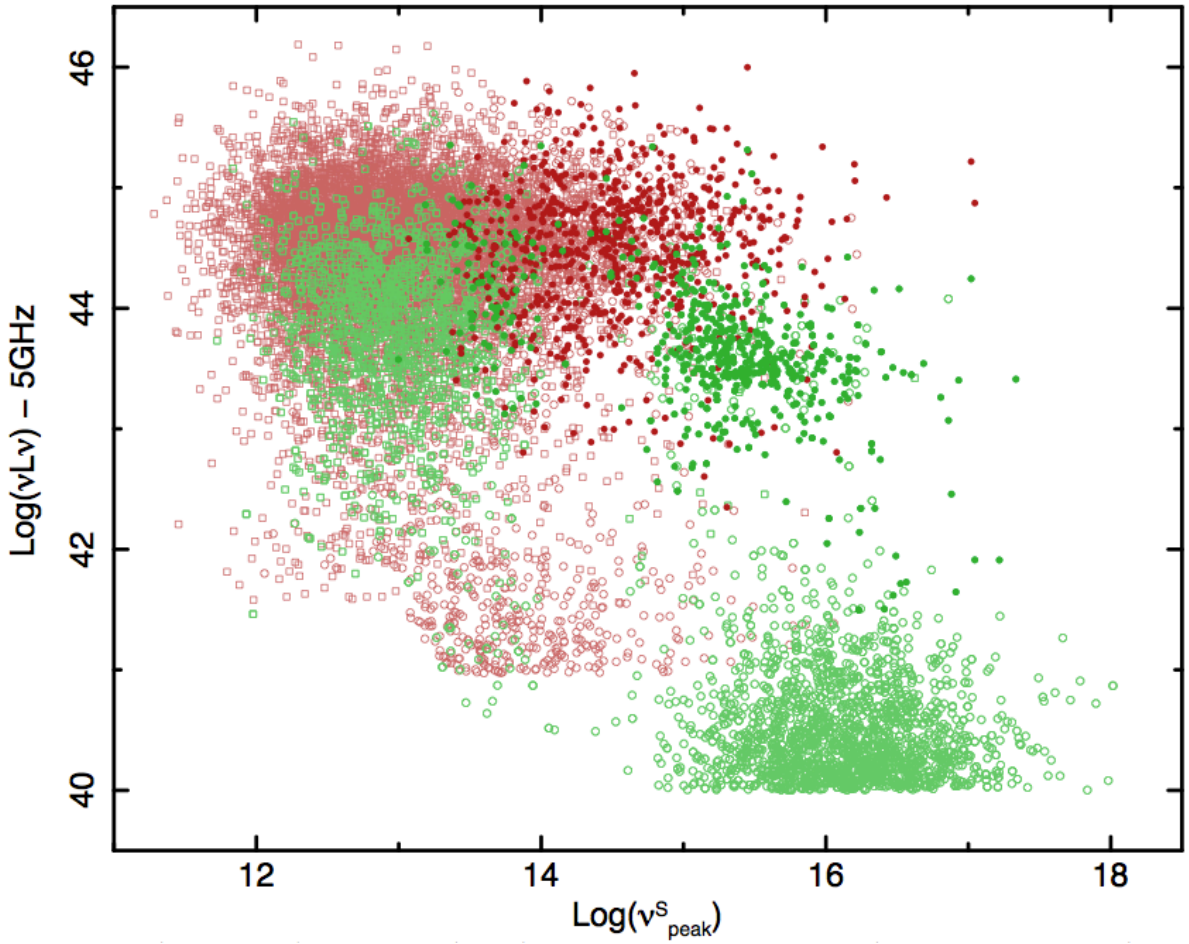
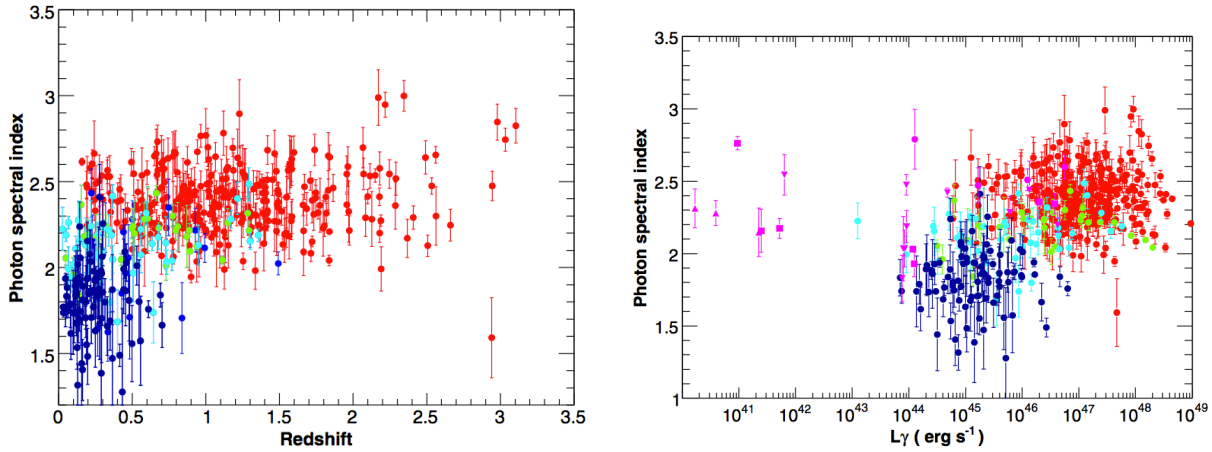


FIGURE 2.10: Luminosité en fonction de la position du maximum du pic synchrotron pour des blazars simulés et sélectionnés de manières différentes (Crédits : [Giommi et al., 2012a]). Les FSRQ et les BL Lac, sélectionné selon la procédure de [Fossati et al., 1998] sont respectivement représentés par des carrés ouverts et des cercles ouverts. Les blazars sélectionnés avec un critère sur le flux radio sont représentés en rouge, et ceux sélectionnés avec un critère sur le flux dans le domaine des X sont représentés en vert. Les cercles pleins représentent des BL Lac qui possèdent une EW < 2 Å et qui ne pouvaient pas être placés sur la séquence des blazars (absence de redshift).



(a) (b)

FIGURE 2.11: (a) Indice spectral en fonction du redshift pour les classes d'AGN : FSRQ (rouge), LBL (vert), IBL (bleu ciel), LBL (bleu), non-blazar (cyan) (Crédits : [Ackermann et al., 2011]). (b) Indice spectral en fonction de la luminosité γ pour ces mêmes AGN.

modèle, ils concluent que l'anti-corrélation observée par [Fossati et al., 1998] est due à un effet de sélection. Ils proposent par ailleurs que les blazars de haute luminosité qui ont des raies d'émission fortement diluées par l'émission non-thermique, qui sont aujourd'hui considérés comme des BL Lac, soient plutôt considérés comme des FSRQ.

2.6 AGN observés dans le domaine γ

Cette section est consacrée aux AGN observés en γ de haute énergie (de ∼30 MeV à ∼300 GeV) par le satellite *Fermi-LAT* ainsi qu'à ceux étudiés à très haute énergie (au-delà d'une centaine de GeV) par les imageurs Tcherenkov atmosphériques.

2.6.1 AGN observés à haute énergie par *Fermi-LAT*

Le télescope *LAT* embarqué à bord du satellite *Fermi* est opérationnel depuis juin 2008. Un catalogue de 1873 sources détectées avec un test statistique supérieur à 25 [10] lors des deux premières années de service de l'expérience a été rendu public par la collaboration *Fermi* : le 2FGL. Parmi ces sources, on compte 1092 AGN et 196 sources galactiques, qui sont essentiellement des pulsars et des nébuleuses à vent de pulsar. Le reste des sources, qui représente ∼30 % de l'échantillon total, est constitué d'objets non-identifiés dont la nature reste à déterminer.

Dans l'échantillon de 886 AGN sélectionnés composant le catalogue *Second LAT AGN Catalog* (2LAC) [11] [Ackermann et al., 2011], les blazars représentent 97,3 % des sources. Parmi ces blazars on en compte 310 de type FSRQ (36 %), 395 de type BL Lac (46 %) et 157 de type indéterminé (18 %). Parmi les BL Lac on compte respectivement 61, 81 et 160 objets de type LBL, IBL et HBL.

Comme on peut le voir sur la Figure 2.11a, la distribution des redshifts des BL Lac pique autour de $z \sim 0{,}2 - 0{,}3$ (seuls 44 % des BL Lac ont une mesure de redshift [Shaw et al., 2013]) alors que la distribution des redshifts des FSRQ s'étale jusqu'à $z = 3{,}10$ avec une moyenne autour de 1. Les indices spectraux mesurés, qui sont en moyenne plus durs pour les BL Lac que pour les FSRQ, et la

10. Le test statistique est défini comme TS $= 2(\log \mathcal{L}(\text{source}) - \log \mathcal{L}(\text{pas de source}))$ où $\mathcal{L}$ est la fonction de vraisemblance selon un modèle avec ou sans source. Un test statistique de 25 correspond à un niveau de détection de $4{,}05\,\sigma$.

11. Le catalogue 2LAC est inclus dans le 2FGL et contient uniquement des AGN. Les nombres donnés dans ce paragraphe correspondent à la sélection « Clean Sample ».

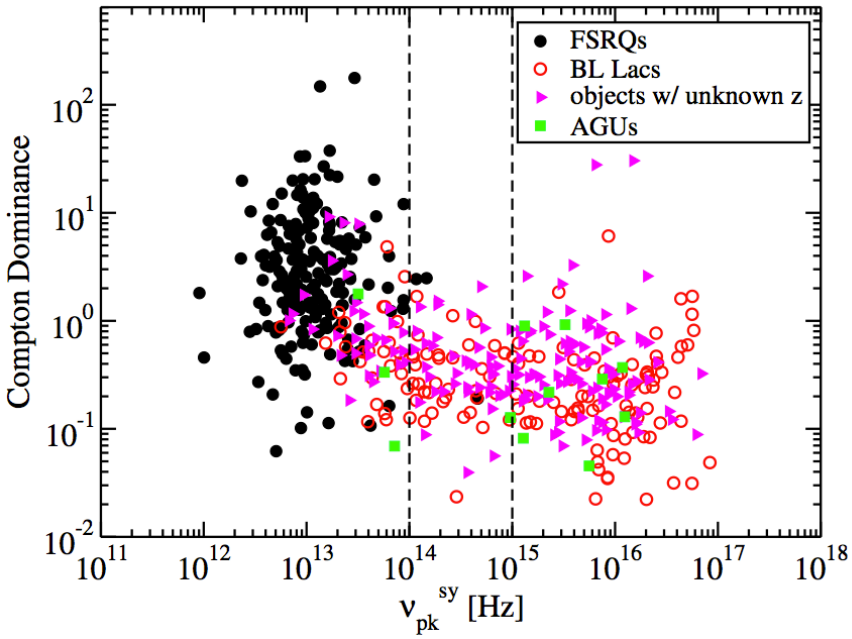


FIGURE 2.12: Rapport de dominance Compton et pic du maximum d'émission synchrotron pour différents types de blazars : BL Lac (ronds rouges), FSRQ (ronds noirs), type indéterminé (carrés verts) et blazars sans mesure de redshift (triangle mauve). Les traits noirs en pointillés représente la limite des objets intermédiaires définis dans le catalogue 2LAC. Crédits : [Finke, 2013].

plus grande sensibilité du *LAT* à haute énergie, participent à la légère asymétrie entre les nombres de BL Lac et de FSRQ détectés. Aucune corrélation n'est observée entre les indices spectraux des FSRQ et leur redshift alors qu'une diminution de la dureté des spectres des BL Lac en fonction du redshift est observée.

La Figure 2.11 montre une forte corrélation entre l'indice spectral et la luminosité γ, ce qui est compatible avec la séquence des blazars proposée par [Fossati et al., 1998]. Néanmoins, l'absence de mesure de redshift pour une large fraction des BL Lac, qui implique que l'on ne peut pas les placer sur ce diagramme, incite à la prudence sur le caractère physique d'une telle corrélation.

Suite à l'étude de [Giommi et al., 2012a] (section 2.5), [Giommi et al., 2013] élargissent leur modélisation au domaine γ en se basant sur les données du catalogue 2LAC. À l'aide d'une sélection de blazars en radio et en γ utilisée entre autres pour déterminer les relations reliant le rapport de dominance Compton et l'indice spectral avec la position du maximum d'émission synchrotron, [Giommi et al., 2013] comparent leur prédiction basées sur les simulations de blazars avec les données des catalogues. Ils reproduisent la prévalence des BL Lac dans la population des blazars détectés par le *Fermi*-*LAT*, les différences observationnelles entre les FSRQ et les BL Lac (indice spectral, position du pic synchrotron, valeur moyenne des redshits), ainsi que la fraction de BL Lac sans redshift mesurable. Leur étude prédit également que ces derniers possèdent une distribution de redshifts proche de celle des FSRQ (comprise entre 0,5 et 2).

Récemment, [Finke, 2013] étudie la corrélation entre le rapport de dominance Compton (rapport entre les luminosités des pics Compton et synchrotron, indépendant du redshift) et la position du maximum du pic synchrotron en utilisant l'échantillon de 886 blazars détectés par le satellite *Fermi*-*LAT* (« clean sample ») du catalogue 2LAC (Figure 2.12). Une modélisation simple est utilisée avec les trois ingrédients suivants : le champ magnétique des régions d'émission, un champ de photons externe et l'angle entre la direction du jet et la ligne de visée. L'étude montre que les sources possédant un faible champ magnétique et un faible champ de photons externe présentent un pic d'émission synchrotron à plus haute énergie, du fait des faibles pertes radiatives des électrons

(typiquement HBL). Si le champ magnétique et le champ de photons externe augmentent, les pertes radiatives des électrons sont alors plus importantes, la position du pic synchrotron diminue avec la fréquence, et à partir d'un certain niveau un régime de type FSRQ est atteint. Notons que dans cette étude la présence d'un champ de photons externe est indispensable afin d'obtenir les corrélations entre le paramètre de dominance Compton et le pic du maximum d'émission synchrotron, et ce particulièrement pour les FSRQ.

Compte tenu du nombre de blazars actuels et des questions en suspens liées au schéma d'unification et en particulier à la validité de la séquence des blazars, une étude qui se focalise sur la recherche de nouveaux candidats blazars a été conduite. La Partie III de cette thèse sera dédiée à la recherche de candidats blazars parmi les 575 sources non-identifiées détectées dans le domaine γ de haute énergie à l'aide du satellite *Fermi-LAT*. À basse latitude galactique ($|b| < 10°$), il est estimé qu'une centaine de blazars doivent s'y trouver afin de rendre compte du déficit observé à ces latitudes. À plus haute latitude, on s'attend à ce que les sources non-identifiées soient principalement réparties entre les blazars et les pulsars. Une première partie de ce travail (chapitre 10) sera consacrée au développement d'une analyse multivariée basée sur une sélection soignée de variables discriminantes afin de rechercher les candidats blazars parmi les sources non-identifiées du catalogue 2FGL. Ce travail sera complété par le développement d'une analyse basée sur une approche multivariée qui sera dédiée à l'identification des contreparties IR de type blazar des sources non-identifiées du catalogue 2FGL. Ce dernier point permettra par la même occasion d'obtenir des coordonnées de précision astrométrique pour les sources qui seront potentiellement associées avec une contrepartie dans l'infrarouge facilitant ainsi le suivi multi-longueur d'onde.

2.6.2 AGN observés à très haute énergie avec les imageurs Tcherenkov au sol

L'étude du rayonnement de très haute énergie émis par les AGN est un des enjeux scientifiques majeurs de l'astronomie γ de très haute énergie. En observant plus d'une centaine d'AGN sélectionnés dans le domaine des X, une émission au TeV a été détectée pour 58 d'entre eux. Cette population est dominé par les blazars (55/58) et en particulier par les BL Lac (52/55) dont la plupart sont situés à des redshifts inférieurs à $\sim$0,2. En plus des blazars, on compte aussi trois radiogalaxies détectées à très haute énergie, NGC 1275, Centaurus A et M 87.

Les observations réalisées avec les détecteurs Tcherenkov de la génération actuelle ont permis la caractérisation spectrale et parfois celle de la variabilité temporelle pour la plupart de ces blazars. Un cas d'étude remarquable est celui des sursauts de PKS 2155-304 détectés en juillet 2006, qui a permis de mettre en évidence une variabilité échantillonnée à la minute [Aharonian et al., 2007], ainsi qu'une étude spectrale et temporelle de la source [Aharonian et al., 2009, Katarzyński et al., 2008, H.E.S.S. Collaboration, 2010, H.E.S.S. Collaboration et al., 2012b]. Les distributions spectrales en énergie des BL Lac sont généralement expliquées dans le cadre d'un modèle leptonique à une zone via le processus SSC. Bien que les approches de type SSC soient largement utilisées, des études récentes réalisées avec l'expérience *H.E.S.S.* sur des blazars de type HBL défient sérieusement ces approches. En effet, elles échouent à reproduire les caractéristiques spectrales de certaines sources [H.E.S.S. Collaboration et al., 2012c,a] ou les corrélations observées à différentes longueurs d'onde lors d'états hauts [H.E.S.S. Collaboration et al., 2011].

Au-delà de l'étude de leurs propriétés, la détection de nombreux BL Lac à différents redshifts a permis de détecter avec un haut niveau de confiance la signature de l'interaction des photons de très haute énergie avec le rayonnement du fond extragalactique [H.E.S.S. Collaboration et al., 2013]. La détection de blazars lointains émetteurs au TeV est aussi motivée par la recherche d'une éventuelle violation de l'invariance de Lorentz. Certains modèles de gravité quantique prédisent qu'à une certaine énergie caractéristique, de l'ordre de l'énergie de Planck, la vitesse de la lumière serait dépendante de l'énergie. L'étude, pour des sources montrant une variation rapide du flux, de délais temporels entre les courbes de lumière correspondant à différentes bandes en énergie permet de contraindre (limites inférieures) l'énergie caractéristique à laquelle ces phénomènes interviendraient. Des études sur les blazars Mkn 421 [Biller et al., 1999], Mkn 501 [MAGIC Collaboration et al., 2008]

et PKS 2155-304 [H.E.S.S. Collaboration, 2011] ont été conduites. On notera que les meilleures contraintes actuelles ont été réalisées avec des sursauts γ détectés par le *Fermi-LAT* [Vasileiou et al., 2013].

La seconde phase de l'expérience *H.E.S.S.* (HESS-II) permettra un gain en sensibilité, une baisse de l'énergie seuil du réseau et permettra une meilleure connexion entre les domaines γ de haute et de très haute énergie. Elle permettra aussi une meilleure caractérisation de la forme spectrale des blazars observés et donnera accès à une population de ces sources à plus haut redshift. Dans la Partie II de cette thèse, nous présenterons le développement d'une méthode d'analyse de type multivariée destinée à une exploitation optimale de HESS-II en mode hybride.

Bibliographie

M. Ackermann, et al. The Second Catalog of Active Galactic Nuclei Detected by the Fermi Large Area Telescope. ApJ, 743:171, December 2011. doi: 10.1088/0004-637X/743/2/171.

F. Aharonian, et al. An Exceptional Very High Energy Gamma-Ray Flare of PKS 2155-304. ApJ, 664:L71–L74, August 2007. doi: 10.1086/520635.

F. Aharonian, et al. Simultaneous multiwavelength observations of the second exceptional γ-ray flare of PKS 2155-304 in July 2006. *Astron. Astrophys.*, 502:749–770, August 2009. doi: 10.1051/0004-6361/200912128.

S. D. Biller, et al. Limits to Quantum Gravity Effects on Energy Dependence of the Speed of Light from Observations of TeV Flares in Active Galaxies. *Physical Review Letters*, 83:2108–2111, September 1999. doi: 10.1103/PhysRevLett.83.2108.

G. Bonnoli, et al. The γ-ray brightest days of the blazar 3C 454.3. MNRAS, 410:368–380, January 2011. doi: 10.1111/j.1365-2966.2010.17450.x.

B. L. Fanaroff et J. M. Riley. The morphology of extragalactic radio sources of high and low luminosity. MNRAS, 167:31P–36P, May 1974.

J. D. Finke. Compton Dominance and the Blazar Sequence. ApJ, 763:134, February 2013. doi: 10.1088/0004-637X/763/2/134.

G. Fossati, et al. A unifying view of the spectral energy distributions of blazars. MNRAS, 299: 433–448, September 1998. doi: 10.1046/j.1365-8711.1998.01828.x.

G. Ghisellini. Radiative Processes in High Energy Astrophysics. *ArXiv e-prints*, February 2012.

G. Ghisellini et F. Tavecchio. The blazar sequence: a new perspective. MNRAS, 387:1669–1680, July 2008. doi: 10.1111/j.1365-2966.2008.13360.x.

G. Ghisellini, et al. A theoretical unifying scheme for gamma-ray bright blazars. MNRAS, 301: 451–468, December 1998. doi: 10.1046/j.1365-8711.1998.02032.x.

P. Giommi, et al. A simplified view of blazars: clearing the fog around long-standing selection effects. MNRAS, 420:2899–2911, March 2012a. doi: 10.1111/j.1365-2966.2011.20044.x.

P. Giommi, et al. Simultaneous Planck, Swift, and Fermi observations of X-ray and γ-ray selected blazars. A&A, 541:A160, May 2012b. doi: 10.1051/0004-6361/201117825.

P. Giommi, P. Padovani, et G. Polenta. A simplified view of blazars: the γ-ray case. MNRAS, 431: 1914–1922, May 2013. doi: 10.1093/mnras/stt305.

H.E.S.S. Collaboration. VHE γ-ray emission of PKS 2155-304: spectral and temporal variability. A&A, 520:A83, September 2010. doi: 10.1051/0004-6361/201014484.

H.E.S.S. Collaboration. Search for Lorentz Invariance breaking with a likelihood fit of the PKS 2155-304 flare data taken on MJD 53944. *Astroparticle Physics*, 34:738–747, April 2011. doi: 10.1016/j.astropartphys.2011.01.007.

H.E.S.S. Collaboration, et al. Simultaneous multi-wavelength campaign on PKS 2005-489 in a high state. A&A, 533:A110, September 2011. doi: 10.1051/0004-6361/201016170.

H.E.S.S. Collaboration, et al. Discovery of hard-spectrum γ-ray emission from the BL Lacertae object 1ES 0414+009. A&A, 538:A103, February 2012a. doi: 10.1051/0004-6361/201118406.

H.E.S.S. Collaboration, et al. A multiwavelength view of the flaring state of PKS 2155-304 in 2006. A&A, 539:A149, March 2012b. doi: 10.1051/0004-6361/201117509.

H.E.S.S. Collaboration, et al. Discovery of VHE γ-ray emission and multi-wavelength observations of the BL Lacertae object 1RXS J101015.9-311909. A&A, 542:A94, June 2012c. doi: 10.1051/0004-6361/201218910.

H.E.S.S. Collaboration, et al. Measurement of the extragalactic background light imprint on the spectra of the brightest blazars observed with H.E.S.S. A&A, 550:A4, February 2013. doi: 10.1051/0004-6361/201220355.

K. Katarzyński, et al. Modelling rapid TeV variability of PKS2155-304. MNRAS, 390:371–376, October 2008. doi: 10.1111/j.1365-2966.2008.13753.x.

J. H. Krolik. *Active galactic nuclei : from the central black hole to the galactic environment*. 1999.

M. S. Longair. *High Energy Astrophysics*. February 2011.

MAGIC Collaboration, et al. Probing quantum gravity using photons from a flare of the active galactic nucleus Markarian 501 observed by the MAGIC telescope. *Physics Letters B*, 668:253–257, October 2008. doi: 10.1016/j.physletb.2008.08.053.

E. T. Meyer, et al. From the Blazar Sequence to the Blazar Envelope: Revisiting the Relativistic Jet Dichotomy in Radio-loud Active Galactic Nuclei. ApJ, 740:98, October 2011. doi: 10.1088/0004-637X/740/2/98.

P. Padovani. The blazar sequence: validity and predictions. Ap&SS, 309:63–71, June 2007. doi: 10.1007/s10509-007-9455-2.

P. Padovani, et al. BeppoSAX Observations of Synchrotron X-Ray Emission from Radio Quasars. ApJ, 581:895–911, December 2002. doi: 10.1086/344406.

C. K. Seyfert. Nuclear Emission in Spiral Nebulae. ApJ, 97:28, January 1943. doi: 10.1086/144488.

M. S. Shaw, et al. Spectroscopy of the Largest Ever γ-Ray-selected BL Lac Sample. ApJ, 764:135, February 2013. doi: 10.1088/0004-637X/764/2/135.

C. M. Urry. Multiwavelength properties of blazars. *Advances in Space Research*, 21:89–100, 1998. doi: 10.1016/S0273-1177(97)00619-4.

C. M. Urry et P. Padovani. Unified Schemes for Radio-Loud Active Galactic Nuclei. PASP, 107:803, September 1995. doi: 10.1086/133630.

V. Vasileiou, et al. Constraints on Lorentz invariance violation from Fermi-Large Area Telescope observations of gamma-ray bursts. Phys. Rev. D, 87(12):122001, June 2013. doi: 10.1103/PhysRevD.87.122001.

Chapitre 3

Analyses multivariées

Sommaire

3.1 Introduction 37
3.2 Exemple introductif 37
3.3 Phases de développement d'une classification binaire 40
3.4 Arbres de décision « boostés » 41
3.5 Réseaux de neurones 45
Bibliographie 52

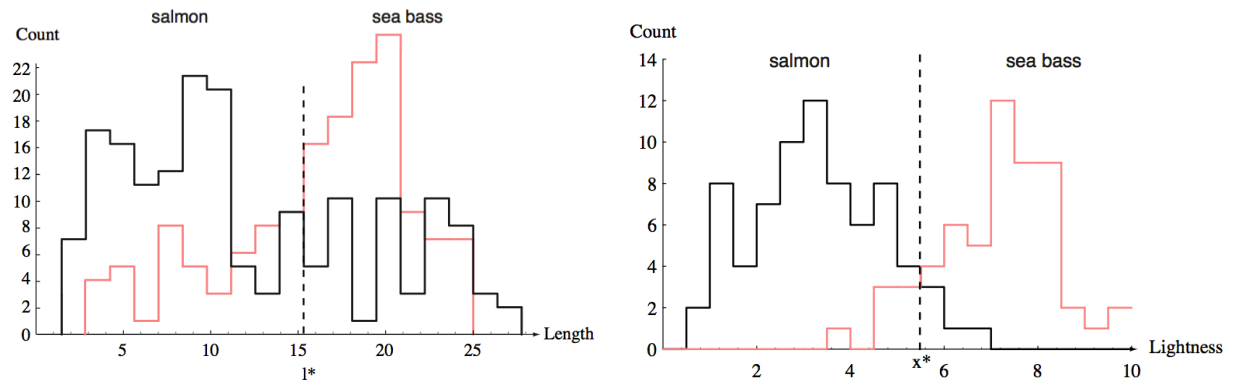
3.1 Introduction

Les méthodes d'analyses multivariées (MVA) sont aujourd'hui largement utilisées dans les domaines de la physique, des sciences sociales, médicales ou économiques. Y font aussi référence les appellations « data mining » et « machine-learning ». Ces techniques sont utilisées afin de résoudre deux classes de problèmes que sont la régression et la classification. Après une phase d'apprentissage utilisant des populations d'évènements possédant des caractéristiques connues, les méthodes MVA permettent pour un problème de régression de prédire les attributs d'un évènement, et pour un problème de classification d'associer une catégorie à un évènement. On cherchera par exemple dans le premier cas à déterminer l'énergie d'un photon γ incident, et dans le second cas à déterminer si une particule incidente est un photon γ ou un hadron (classification binaire). La spécificité des méthodes multivariées réside dans leur capacité à gérer des problèmes dans lesquels plusieurs variables sont présentes. Un chapitre leur est consacré car elles seront utilisées dans la Partie II et la Partie III pour des problèmes de classifications binaires. Nous nous focaliserons exclusivement sur ce type de problème.

Après la présentation à la section 3.2 d'un exemple permettant d'illustrer un problème type qui peut être résolu à l'aide d'une analyse multivariée, les étapes de la construction d'une analyse utilisant une méthode MVA sont présentées à la section 3.3. Aux sections 3.4 et 3.5 nous présentons deux méthodes multivariées qui seront utilisées par la suite (Partie II et Partie III), les arbres de décisions « boostés » (BDT) et les réseaux de neurones de type perceptron multicouche (MLP).

3.2 Exemple introductif

Afin d'illustrer la démarche de résolution d'un problème à l'aide de techniques MVA, étudions un exemple de classification binaire tiré de [Duda et al., 2004] dans lequel une usine d'emballage de poissons, qui reçoit des saumons et des bars, souhaite automatiser sa procédure d'acheminement pour l'une ou l'autre espèce de poisson jusqu'à une plateforme de traitement dédiée selon leur espèce. Équipé d'un appareil photo, le chef de projet immortalise des individus des deux populations et note des différences entre les caractéristiques des deux populations. La largeur, la longueur, la clarté, le nombre d'écailles, la position des yeux par rapport à la bouche, etc. semblent être des traits qui sont en moyenne différents pour chacune des deux espèces. Ces différences expriment le fait que les saumons et les bars possèdent des modèles sous-jacents distincts ou, autrement dit, des descriptions différentes.



(a) (b)

FIGURE 3.1: Distributions de (a) la longueur et de (b) la luminosité des bars (rouge) et des saumons (noir) (Crédits : [Duda et al., 2004]). Les valeurs $l^\star$ et $x^\star$ représentent les coupures pour lesquelles les erreurs globales de mauvaise classification sont minimales.

L'approche va consister à supposer l'existence de deux classes de poissons, les saumons et les bars, et à associer une des deux classes à chaque évènement de type inconnu en se basant sur ses caractéristiques. Pour cela, supposons que la première partie de l'acheminement est équipée d'une bande transporteuse sur laquelle les poissons passent chacun à leur tour, qu'un système optique permet d'obtenir l'image d'un individu sur laquelle le fond a été soustrait, et qu'une chaîne d'analyse permet d'obtenir les caractéristiques de chaque poisson telle que la longueur ou la luminosité. La personne en charge, observant que les saumons sont en moyenne moins longs que les bars, décide de choisir deux lots de poissons bien identifiés, l'un exclusivement composé de saumons et l'autre exclusivement composé de bars, afin de comparer leurs longueurs respectives. Sur la Figure 3.1a, on peut observer la distribution des longueurs, que l'on notera l, des saumons (en noir) et des bars (en rouge). Les bars sont bien en moyenne plus longs que les saumons, mais le pouvoir de séparation « saumon/bar » de cette variable est faible. À titre d'indication, la valeur $l^\star$ représente la valeur pour laquelle le taux de mauvaise classification globale est minimal. L'analyste se penche maintenant sur les distributions de la luminosité des poissons, que l'on notera x, qui sont représentées sur la Figure 3.1b. Cette fois, une séparation nette est visible entre la luminosité des bars et celles des saumons, ces derniers étant moins lumineux. De la même manière que pour la longueur, la valeur $x^\star$ qui permet d'obtenir le taux de mauvaise classification minimal est indiquée.

En possession d'une variable ayant un important pouvoir de séparation entre les différentes espèces de poissons, se pose maintenant la question du choix de la valeur à attribuer à $x^\star$. En effet, les clients peuvent accepter d'avoir un certain taux de contamination de saumon dans les caisses de bars mais pas le contraire. Dans ce cas, le propriétaire a tout intérêt à choisir une valeur plus faible de $x^\star$ que celle indiquée sur la Figure 3.1b qui conduira à un taux de rejet plus important des bars au détriment d'une perte de saumons plus importante, afin de répondre aux désidérata de la clientèle.

On peut tenter de rajouter une seconde observable dans le but d'obtenir un pouvoir discriminant supplémentaire, en utilisant par exemple le fait que les saumons sont en moyenne légèrement plus fins que les bars. La Figure 3.2 représente la valeur de l'épaisseur en fonction de la luminosité des saumons (en noir) et des bars (en rouge), avec différentes limites de séparation entre les deux espèces (Figures 3.2a, 3.2a et 3.2c). La séparation linéaire en traits pleins visible sur la Figure 3.2a permet de réduire le taux global de mauvaise classification. Une autre option serait d'utiliser la séparation représentée sur la Figure 3.2b qui permet de séparer de manière optimale les bars et les saumons pour ces poissons identifiés mais qui entraînera des performances médiocres lors du traitement de nouveaux poissons, sa marge d'adaptation étant beaucoup trop réduite. Dans ce dernier cas, on

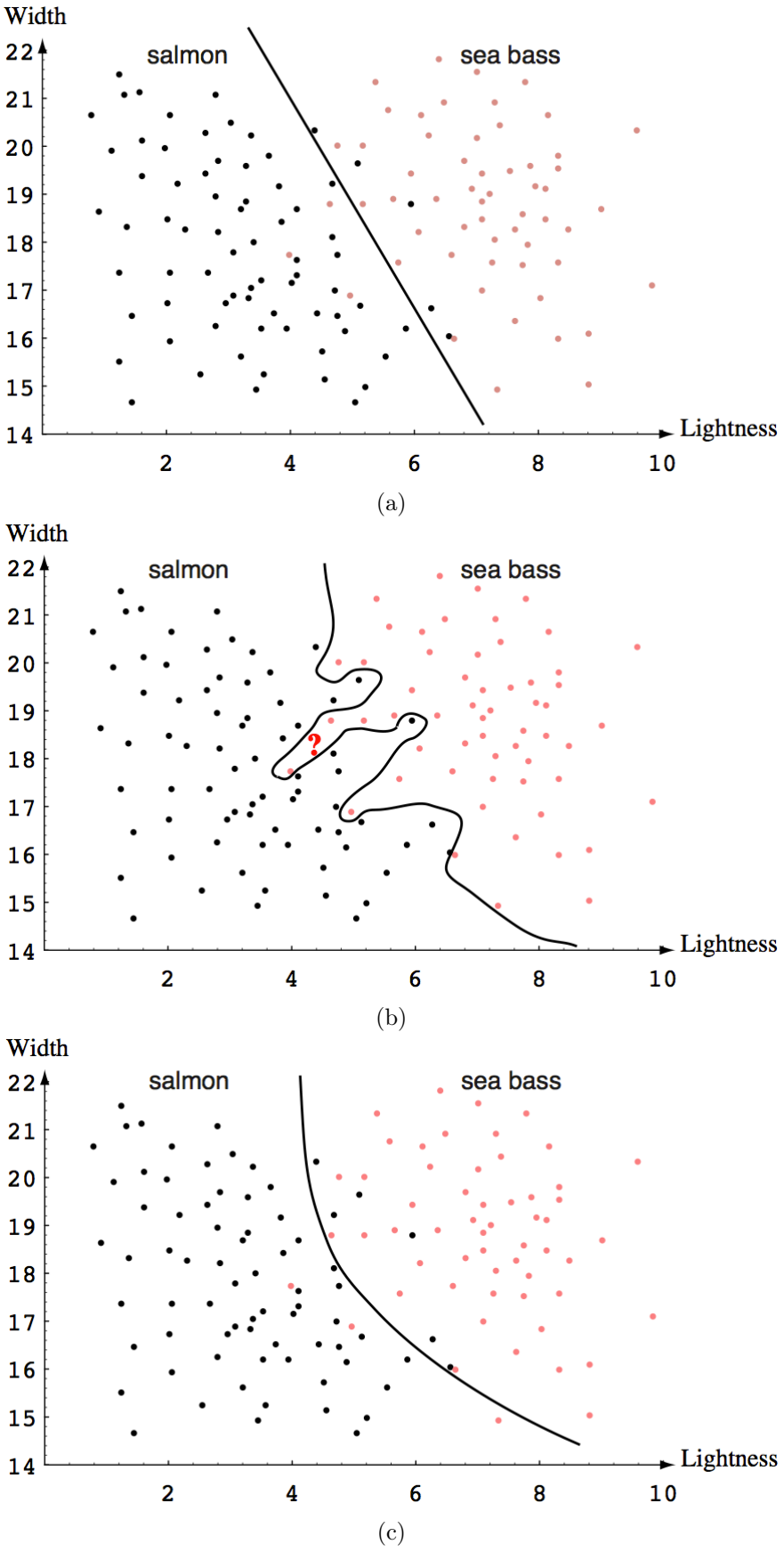
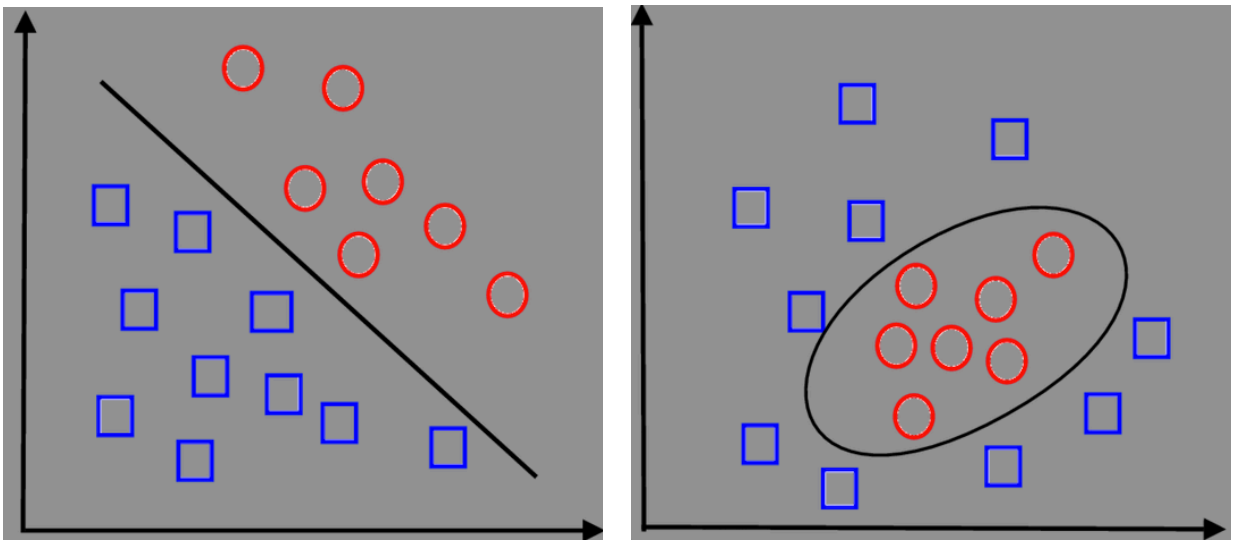


FIGURE 3.2: Choix de la séparation optimale (Crédits : [Duda et al., 2004]).



(a) (b)

FIGURE 3.3: Données séparables de manière linéaire et non-linéaire (Crédits : C. Raymond).

parle alors de « surentraînement » (en anglais « overtraining » ou « overfitting »). Une troisième option serait d'utiliser une séparation non-linéaire, représentée sur la Figure 3.2c.

La force des méthodes multivariées consiste à résoudre des problèmes de ce type, avec potentiellement un grand nombre de variables en jeu, en les réduisant à des problème à une dimension. Une méthode de classification binaire basée sur une méthode MVA attribuera donc aux poissons une valeur unique qui rendra compte de leur nature « bar » ou « saumon ».

3.3 Phases de développement d'une classification binaire

La construction d'une classification binaire « signal/fond » utilisant des méthodes multivariées repose sur le choix de variables discriminantes sélectionnées en fonction de leur pouvoir de séparation « signal/fond ». L'objectif d'une méthode multivariée est de trouver une relation entre les variables, qui permettent d'attribuer à chaque évènement une valeur unique qui rendra compte de sa nature « signal » ou « fond ».

Le choix de la méthode multivariée est étroitement lié au variables discriminantes, car celles-ci possèdent leur propre spécificité concernant les performances optimales qu'il est possible d'atteindre selon les relations qui lient les variables utilisées. En effet, le fait que le « signal » et le « fond » soit séparable de manière linéaire (cas représenté sur la Figure 3.3a) ou au contraire qu'une séparation non-linéaire (cas représenté sur la Figure 3.3b) soit nécessaire dicte le choix de la méthode multivariée. De plus, certaines méthodes sont sensibles au fait que les variables discriminantes sont linéairement corrélées. Il est parfois nécessaire de décorréler les variables ou de rendre gaussienne des variables afin d'avoir de bonnes performances.

La construction de la classification contient deux étapes qui nécessitent l'utilisation de deux lots distincts d'évènements, à savoir le lot d'entraînement et le lot de test. Chacun de ces lots est composé d'évènements de « signal » et de « fond ». Le lot d'entraînement est utilisé lors de la phase d'apprentissage du classifieur. Cette phase d'apprentissage est propre à chacune des méthodes multivariées (on s'intéressera par la suite à cette phase pour les arbres de décision « boostés » et les réseaux de neurones de type MLP) et consiste à attribuer à chaque évènement du lot d'entraînement une valeur unique, que l'on notera ζ, qui rend compte de sa nature « signal » ou « fond ». Une fois la phase d'entraînement réalisée, le lot test est utilisé d'une part pour vérifier le comportement du classifieur sur de nouveaux évènements, et d'autre part pour estimer les performances de la

classification. Dans le premier cas les distributions ζ des évènements « signal » et de « fond » sont comparées pour chacun des lots d'entraînement et de test. Si des différences significatives sont observées c'est que le classifieur n'a pas réussi s'adapter à de nouveaux évènements et on est dans un cas de surentraînement ou de sous entraînement. Les performances sont obtenues sur ce même lot, en calculant pour chaque valeur de ζ les efficacités de reconstruction du « signal », notées ε_{Sig}, et les efficacités de reconstruction du « fond », notées ε_{Bkg}. Ces quantités sont calculées comme le rapport du nombre d'évènements ayant une valeur supérieure ou égale à ζ et le nombre total d'évènements pour les échantillons d'évènements de « signal » et de « fond ».

Une fois le classifieur entraîné et les performances calculées, il est nécessaire de choisir un point de fonctionnement, une valeur $\zeta^\star$, au-delà de laquelle tous les évènements seront considérés comme « signal », avec un certain niveau de contamination, et en deçà comme du « fond ». Celle-ci dépend du problème considéré, selon qu'on veuille obtenir un rapport signal sur bruit maximal, un taux de contamination minimal ou une efficacité de reconstruction du signal donnée.

Dans la suite, nous présenterons les phases d'apprentissage de deux méthodes multivariées qui permettent de séparer le « signal » du « fond » de manière non-linéaire.

3.4 Arbres de décision « boostés »

Dans cette partie, nous présentons la méthode BDT afin de résoudre un problème de classification binaire « signal/fond ». Nous commencerons par introduire les arbres de décision puis la technique de « boosting ». On considérera par la suite que les lots d'entraînement de « signal » et de « fond » ont un nombre d'évènements équivalent (en pratique ces nombres diffèrent et on associe des poids aux évènements afin d'avoir des lots comportant le même nombre d'évènements).

3.4.1 Arbres de décision

Un arbre de décision permet de classer des données métriques ou non-métriques à l'aide de questions à réponse binaire. Il se présente sous la forme d'une succession de nœuds au niveau desquels un critère de sélection permet de séparer les données. Dans le cas d'une classification binaire « signal/fond », l'arbre attribuera une valeur ± 1 à des régions de l'espace des phases des variables discriminantes, selon que la majorité des évènements utilisés lors de la phase d'entraînement dans cette région soient du signal ou du fond. De nouveaux évènements pourront être traités avec l'arbre de décision généré, auxquels ils leur seront attribués une valeur $\zeta = \pm 1$ selon qu'il soient considérés comme du signal $(+1)$ ou comme du fond (-1).

3.4.1.1 Développement

On notera respectivement S et B les classes des évènements « signal » et « fond », et x_n les n variables discriminantes associées aux évènements $\{S_i\}$ et $\{B_i\}$. La construction d'un arbre se déroule de la manière suivante. Au premier nœud (la racine de l'arbre), pour chaque variable on classe les évènements selon une relation d'ordre :

$$\begin{gathered} x_1^{S_3} \leq x_1^{S_1} \leq \cdots \leq x_1^{B_2} \leq x_1^{B_3} \\ x_2^{B_4} \leq x_2^{S_4} \leq \cdots \leq x_2^{S_3} \leq x_2^{S_{10}} \\ \cdots \leq \cdots \leq \cdots \\ x_n^{B_7} \leq x_n^{B_9} \leq \cdots \leq x_n^{S_1} \leq x_n^{B_4} \end{gathered}$$

On sélectionne ensuite la variable x_k et la valeur associée $x_k^\star$ qui permettent d'obtenir un pouvoir de séparation maximal entre les évènements de « signal » et de « fond » (on reviendra sur ce critère un peu plus bas, section 3.4.1.2). Deux nouveaux nœuds sont créés dans lesquels les évènements

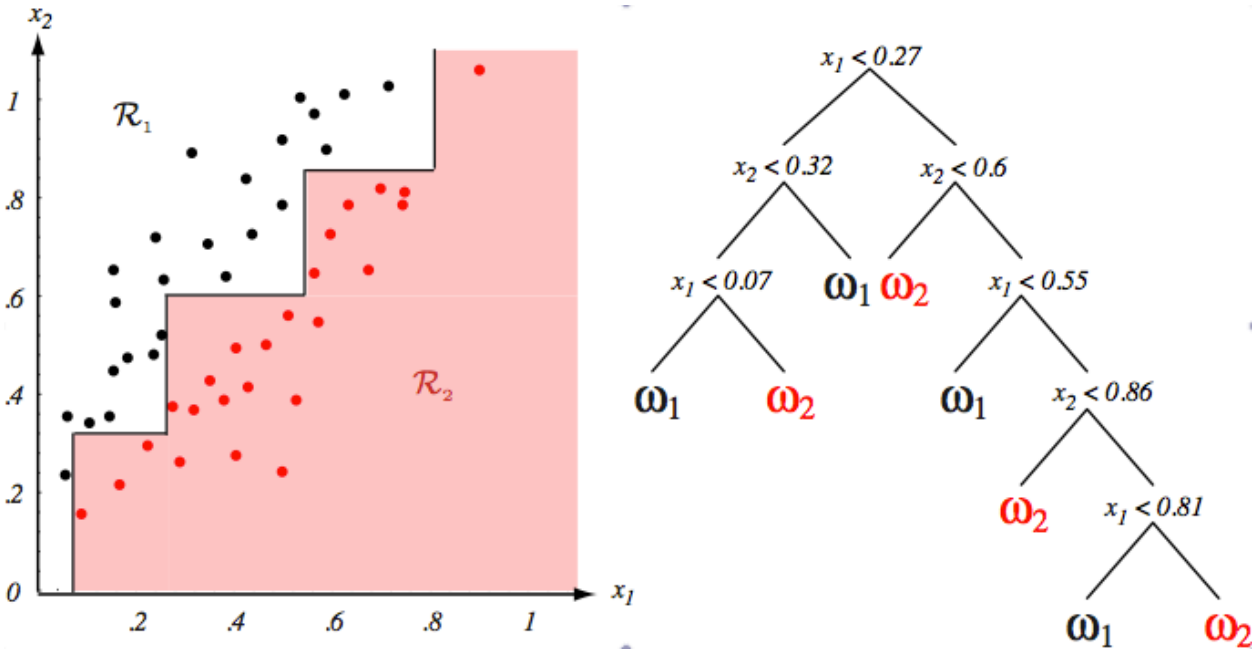


FIGURE 3.4: Exemple d'arbre de décision (Crédits :Duda et al. [2004]).

sont dirigés selon que leur valeur est plus grande ou plus petite que $x_k^\star$. La procédure de sélection de la variable qui permet d'atteindre le pouvoir de séparation maximal est appliquée dans chaque nœud, puis la procédure se répète jusqu'à ce qu'un critère d'arrêt du développement de l'arbre soit vérifié (on y reviendra plus bas, section 3.4.1.3).

À titre d'illustration, à gauche de la Figure 3.4 on pourra trouver la distribution de deux variables x_1 et x_2 pour des évènements de « signal » (en noir) et de « fond » (en rouge). À droite de la Figure 3.4 est représenté le développement d'un arbre de décision qui permet d'obtenir la séparation, symbolisée par une ligne noire sur la figure de gauche, entre le signal et le fond.

3.4.1.2 Pouvoir de séparation

Optimiser le pouvoir de séparation dans un nœud revient à maximiser la pureté de sa population. On définira respectivement les puretés du signal et du fond par : $p_S = \frac{N_S}{N_S+N_B}$ et $p_B = \frac{N_B}{N_S+N_B}$. Nous utiliserons par la suite la notion d'impureté, notée $I(N)$ pour le nœud N, afin d'obtenir le critère de séparation optimale. Celle-ci est maximale lorsque la population du nœud est constituée à part égale d'évènements de classe S et B, et minimale lorsque sont présents exclusivement des évènements de même classe (elle est par définition symétrique entre S et B). Trois fonctions d'impureté sont couramment utilisées. La première fonction est le taux de mauvaises classifications défini comme $I(N) = 1 - \max(p_i)$ où p_i représente la fraction d'évènements de catégorie $i = \{S, B\}$. La seconde, qui rappelle la forme de l'entropie, est définie par $I(N) = -\sum_i p_i \ln p_i$. Enfin la troisième fonction est appelée l'indice de Gini, elle est définie comme $\sum_{i\neq j}\sum_j p_i p_j = 2p_s(1-p_s)$. Ces trois fonctions sont représentées sur la Figure 3.5. La première fonction est la plus piquée lorsque les populations sont présentes en quantité égales. Elle est la moins utilisée car sa dérivée présente une discontinuité, ce qui pose problème pour la recherche d'un optimum. Restent les fonctions « entropie » et l'indice de Gini, cette dernière est la plus couramment utilisée avec les arbres de décision.

Choisir le meilleur critère de sélection dans le nœud N revient à choisir la variable x_n pour laquelle la baisse de l'impureté est maximale. Elle est définit comme :

$$\Delta I(N) = I(N) - (P_G I(N_G) + P_D I(N_D)) \tag{3.1}$$

où les indices G et D correspondent respectivement les deux nœuds suivants (à gauche et à droite).

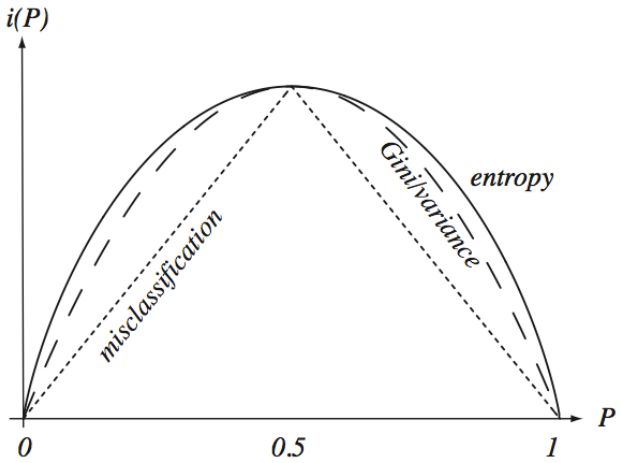


FIGURE 3.5: Les trois fonctions d'impureté décrites dans le texte normalisées par leur intégrale (Crédits : Duda et al. [2004]).

Les quantités P_G et P_D sont les fractions d'évènements de classe S qui iront respectivement dans les nœuds N_G et N_D.

3.4.1.3 Feuille et arrêt du développement

Une feuille est un nœud terminal de l'arbre auquel est associé un label qui permet de classer un évènement comme « signal » ou « fond ». Laisser l'arbre se développer jusqu'à atteindre dans chaque feuille le maximum de pureté reviendrait dans le pire des cas à avoir un chemin pour chaque évènement des lots d'entraînement. C'est un cas de surentraînement pour lequel l'arbre connaît trop spécifiquement les évènements ayant servi à la construction de l'arbre. Dans le cas où l'arrêt du développement de l'arbre est prématuré, le quadrillage de l'espace des phases serait au contraire trop grossier et les performances finales seraient réduites. Un critère pour l'arrêt du développement de l'arbre est donc nécessaire.

L'un des critères les plus utilisés consiste à créer une feuille lorsque le nombre d'évènements dans un nœud représente une certaine fraction du nombre d'évènements total du lot d'entraînement (typiquement 5 %). Cela permet d'avoir un découpage de l'espace des variables discriminantes très fin lorsque le nombre d'évènements est grand, et au contraire un découpage grossier lorsque le nombre d'évènements est faible. Une feuille aura un label « signal » ou « fond » selon le nombre d'évènements de classe majoritaire.

3.4.2 « Boosting »

Originellement développée pour combiner les décisions d'arbres de décision ayant de faibles performances individuelles afin de former une décision solide, la technique de « boosting » peut en théorie être utilisée pour augmenter les performances de n'importe quel classifieur et lutter contre le surentraînement.

3.4.2.1 Principe

Le premier algorithme montrant la faisabilité de la méthode [Shapire, 1990] combinait la réponse de 3 arbres de décision pour former une décision unique. En 1995, [Freund, 1995] apporta des améliorations à cet algorithme et en 1996 émergea le premier algorithme de « boosting » adaptif (AdaBoost) [Freund et Schapire, 1996]. Dans cette partie, l'algorithme AdaBoost est présenté en s'inspirant de l'introduction au « boosting » de [Freund et Schapire, 1999].

Supposons qu'une personne ayant décidé de faire du pari sur les matchs de football lors de la prochaine coupe du monde un art à part entière, décide d'écrire un programme lui permettant d'optimiser ses chances de prédictions. Amateur ou amatrice de football possédant quelques notions de stratégie footballistique, notre individu sera capable d'extraire une collection de règles simples, comme parier sur l'équipe ayant la première place dans le championnat, pour l'équipe ayant en son sein un joueur de qualité rare et exceptionnelle ou, si le climat est aride, pour l'équipe ayant gagné le plus de matchs pendant les périodes climatiques sèches. Ces règles sont simples et imprécises, mais prises individuellement on s'attend à ce que sur le long terme chacune d'entre elles permette de prédire l'issue d'un match au moins un peu mieux qu'une décision tirée à pile ou face. Notre personnage récupère maintenant des informations sur la préparation et la composition des différentes équipes et est en possession d'un grand nombre de règles supplémentaires. Voulant tirer avantage des règles acquises, il est confronté au problème suivant : comment les combiner afin de fournir une prédiction précise ?

Dans le cas de l'utilisation du « boosting » avec des arbres de décision on parle de forêt d'arbres de décision. Les premiers arbres générés auront un poids associé plus important que les autres et permettront de reconnaître le gros des évènements « signal » ou « fond ». Les autres, qui lors de leur génération se concentreront sur les évènements difficiles, auront des taux de mauvaises classifications importants (légèrement inférieur à 1/2) mais seront spécialisés dans l'identification des évènements « hors-norme ». Un faible poids leur est associé mais la combinaison de plusieurs décisions faibles pourra former une décision précise.

3.4.2.2 « Booster » des arbres de décision

Le principe d'AdaBoost est de construire une forêt d'arbres de décision T_k avec un ensemble C_k d'évènements i caractérisé par leurs variables $\mathbf{x}^i$ et leurs labels $y_i = \{\pm 1\}$. À chaque création d'un nouvel arbre T_k, la technique associe aux évènements mal classés un poids supérieur à celui des évènements bien classés. L'utilisation de ces évènements « hors-normes » sera alors privilégiée pour le développement de l'arbre de décision suivant. La décision finale sera la somme des décisions de chaque arbre pondérée par un facteur dépendant du taux de mauvaise classification pour chaque arbre, noté ε_k. En notant $W_k(i)$ la distribution des poids utilisées pour la fabrication de l'arbre T_k et $\mathcal{D}$ l'ensemble des couples $\{\mathbf{x}^i, y_i\}$ l'algorithme s'écrit :

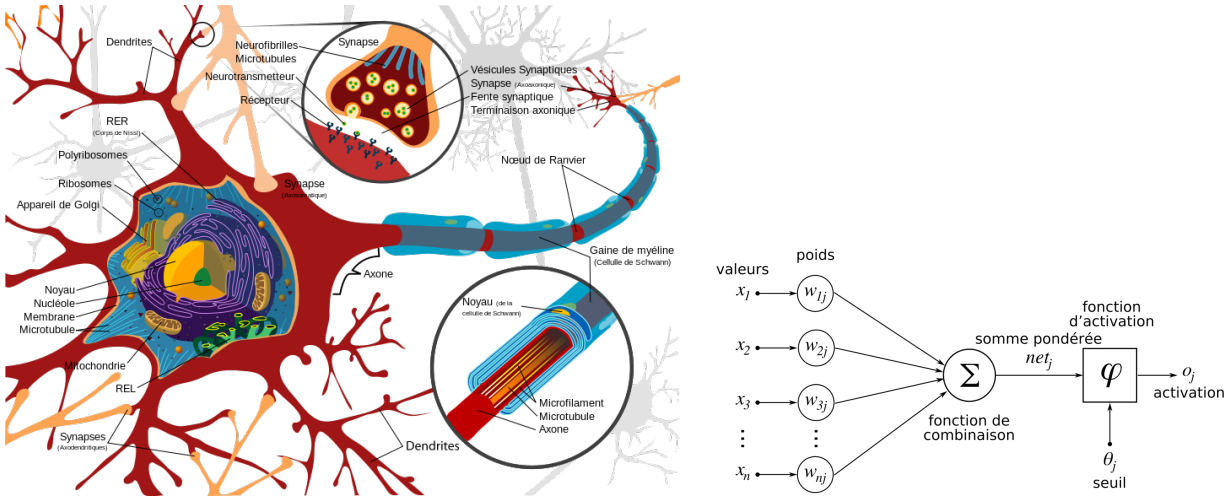
Algorithme 1 AdaBoost

$k \leftarrow 1$
tant que $k \leq k_{\max}$ **faire**
  Entraîner un arbre de décision T_k en utilisant $\mathcal{D}$ suivant $W_k(i)$
  $\varepsilon_k \leftarrow$ Taux d'erreur de mauvaises classifications de T_k mesuré dans $\mathcal{D}$ avec $W_k(i)$
  $\alpha_k \leftarrow \frac{1}{2}\ln[(1-\varepsilon_k)\varepsilon_k]$
  $W_{k+1}(i) \leftarrow \frac{W_k(i)}{Z_k} \times \begin{cases} \exp(-\alpha_k) & \text{si } \zeta_k(\mathbf{x}^i) = y_i \text{ (correctement classé)} \\ \exp(+\alpha_k) & \text{si } \zeta_k(\mathbf{x}^i) \neq y_i \text{ (incorrectement classé)} \end{cases}$
  $k \leftarrow k+1$
fin
return T_k et α_k pour $k = \{1, 2, \cdots, k_{\max}\}$

où Z_k représente une constante de normalisation pour la distribution $W_k(i)$ et $\zeta_k(\mathbf{x}^i)$ la décision de classification de l'arbre T_k pour l'évènement i. La décision finale est la somme des décisions pondérées par le paramètre α_k :

$$\zeta(\mathbf{x}) = \sum_{k=1}^{k_{\max}} \alpha_k \zeta_k(\mathbf{x}) \tag{3.2}$$

De larges valeurs pour $\zeta(\mathbf{x})$ correspondent à des évènements « signal ». En exprimant l'erreur de classification de l'arbre k comme $\varepsilon = 1/2 - \gamma_k$, γ_k mesurant l'écart de la décision k par rapport à



(a) (b)

FIGURE 3.6: À gauche schéma d'un neurone biologique et à droite, représentation d'un modèle de neurone. (Crédits : fr.wikipedia.org).

une décision aléatoire, on peut montrer [Freund et Schapire, 1997] que le taux global de mauvaises classifications peut s'écrire :

$$\mathcal{E} = \prod_{k=1}^{k_{\max}} 2\sqrt{\varepsilon_k(1-\varepsilon_k)} = \prod_{k=1}^{k_{\max}} \sqrt{1-4\gamma_k} \leq \exp\left(-2\sum_{k=1}^{k_{\max}} \gamma_k^2\right) \tag{3.3}$$

et peut être rendu aussi petit que l'on veut en choisissant une grande valeur de $k_{\max}$. Afin de régler la vitesse de l'apprentissage, le paramètre β est introduit tel que, $\alpha \to \alpha^\beta$. En choisissant β plus petit que l'unité, la vitesse de convergence sera réduite et permettra ainsi l'utilisation d'un plus grand nombre d'arbres de décision.

3.4.2.3 Utilisation dans le manuscrit

La technique des arbres de décision « boostés » sera d'abord utilisée dans la Partie II lors des chapitres 6 et 7 afin de discriminer les photons γ des évènements de fond (hadrons et électrons). Nous l'utiliserons aussi au cours de la Partie II dans le chapitre 10 lors de la recherche des candidats blazars parmi les sources non-identifiées *Fermi-LAT* du catalogue 2FGL.

3.5 Réseaux de neurones

Nous commencerons par présenter le fonctionnement d'un neurone artificiel qui permet de résoudre des problèmes dans lesquels les données sont linéairement séparables, puis celui d'un réseau de neurones de type MLP, qui utilise des neurones disposés en couches, afin de résoudre des problèmes de classification pour lesquels les données ne sont pas séparables de façon linéaire.

3.5.1 Neurone artificiel

3.5.1.1 Principe

Un neurone artificiel est un modèle simplifié de neurone biologique. Schématiquement (Figure 3.6a), un neurone biologique est une cellule contenant un noyau cellulaire et dotée de deux prolongements, l'axone et les dendrites. Le premier est un prolongement unique véhiculant l'influx nerveux et permettant de se connecter aux dent-rites d'un autre neurone. Les secondes ont une

structure ramifiée et chacune d'entre elles véhicule l'influx nerveux mais cette fois jusqu'au corps cellulaire. Si l'excitation nerveuse portée par les dendrites dépasse un certain seuil, alors une décision est prise et un influx nerveux est propagé le long de l'axone.

À partir de cette modélisation très simpliste, on peut dégager trois propriétés fondamentales d'un neurone. Tout d'abord, le poids synaptique [1] associé à chaque dendrite. Des dendrites différentes auront un poids plus ou moins important afin d'influencer la décision relative à l'envoi d'un signal par le neurone. La seconde est la combinaison du signal et du poids de chaque dendrite. La troisième est la décision : un signal est envoyé ou non. La construction d'un neurone artificiel afin de créer une classification binaire permettant de séparer des données linéairement séparables est basée sur ces trois propriétés. L'objectif est de trouver les poids qui vont permettre de minimiser les mauvaises classifications lors d'une procédure itérative d'apprentissage décrite par la suite.

Considérons N évènements ayant un ensemble de n variables discriminantes x_n^i et un label $y^i = \{0, 1\}$, où $i \in [1, N]$. Afin de construire une classification de ces évènements, on utilisera un neurone ayant n entrées et une valeur de sortie que l'on notera ζ (Figure 3.6b). À chaque entrée sera associée une variable discriminante et un poids synaptique w_n. On définit l'activation du neurone a pour l'évènement i comme la combinaison linéaire : $a^i = \sum_{k=1}^{n} w_k x_k^i$. À cette quantité est associée l'activité du neurone, notée $f(a)$, qui peut prendre différentes formes comme par exemple, une fonction logistique $f(a) = 1/(1 + e^{-a})$, une fonction tangente hyperbolique, ou une fonction porte. Celles-ci ont la particularité d'être comprises entre 0 et 1 afin de pouvoir correspondre aux labels des évènements. La plupart du temps, on utilisera une entrée supplémentaire, que l'on appellera biais, à laquelle on attribuera une valeur $x_0 = 1$ et un poids w_0, et qui permet simplement d'effectuer une translation de la fonction d'activation du neurone.

On introduit une fonction mesurant le taux de mauvaises classifications commises par le neurone : $\mathcal{E} = 1/2 \sum_i^n (\zeta_i - y^i)^2$ où ζ_i est la sortie du neurone pour l'évènement i, qui dépend des poids $(w_0 \cdots w_n)$ par l'intermédiaire de l'activité f du neurone. L'objectif est de trouver le minimum de la fonction $\mathcal{E}$ à l'aide d'une phase d'apprentissage. Dans le cas d'une utilisation séquentiel des évènements, lorsque l'un d'eux passe dans le neurone, les poids sont modifiés de manière à se rapprocher du minimum de la fonction d'erreur de mauvaises classifications. Le processus est itératif, les poids modifiés sont réinjectés en entrée du neurone et le processus se répète jusqu'à ce qu'un critère de convergence soit validé. En utilisant par exemple un algorithme de type rétropropagation du gradient (voir plus bas à la section 3.5.2.2), le passage d'un évènement induit une variation des poids telle que :

$$\Delta \mathbf{w} = -\eta \nabla_{\mathbf{w}} \mathcal{E} \tag{3.4}$$

où $\mathbf{w}$ représente l'ensemble des poids w_0 et w_n, et où η permet de régler la vitesse de l'apprentissage.

3.5.1.2 Exemple

Afin d'illustrer le fonctionnement d'un neurone dans une classification binaire avec deux variables discriminantes x_1 et x_2, intéressons-nous à un exemple tiré de [MacKay, 2003]. L'espace vectoriel des poids (w_1, w_2) est représenté sur la Figure 3.7 où on peut y trouver la fonction d'activation du neurone, pour quelques couples (w_1, w_2) en fonction de x_1 et x_2, représentée ici par une fonction logistique :

$$f(\mathbf{x}, \mathbf{w}) = \frac{1}{1 + \exp(-w_1 x_1 - w_2 x_2)} \tag{3.5}$$

On remarquera que plus l'intensité du vecteur $\mathbf{w} = (w_1, w_2)$ est importante et plus la pente de l'activité du neurone est abrupte. L'objectif est d'obtenir le couple (w_1, w_2) permettant d'obtenir le taux de mauvaises classifications le plus bas.

On peut trouver sur la Figure 3.8 différentes étapes de la phase d'apprentissage d'un tel neurone. Le neurone est doté de deux poids w_1 et w_2 et d'un biais w_0, afin de séparer linéairement des évènements « signal » (carrés) des évènements de « fond » (croix) qui sont représentés sur la sous-figure

1. Le poids synaptique module la communication entre deux neurones.

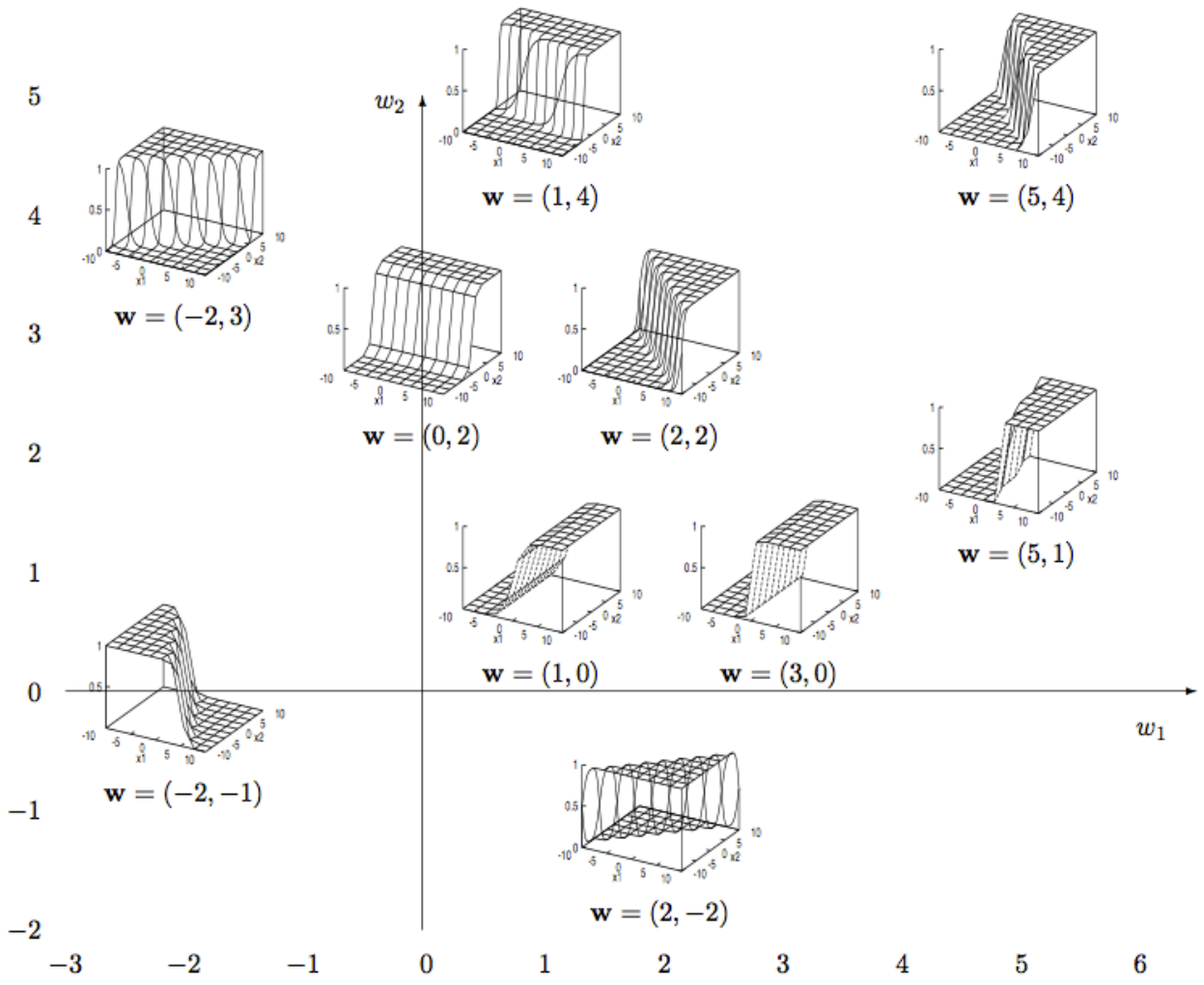


FIGURE 3.7: Espace des phases des poids (w_1, w_2) pour lequel est représenté la sortie d'une neurone (Crédits : MacKay [2003]).

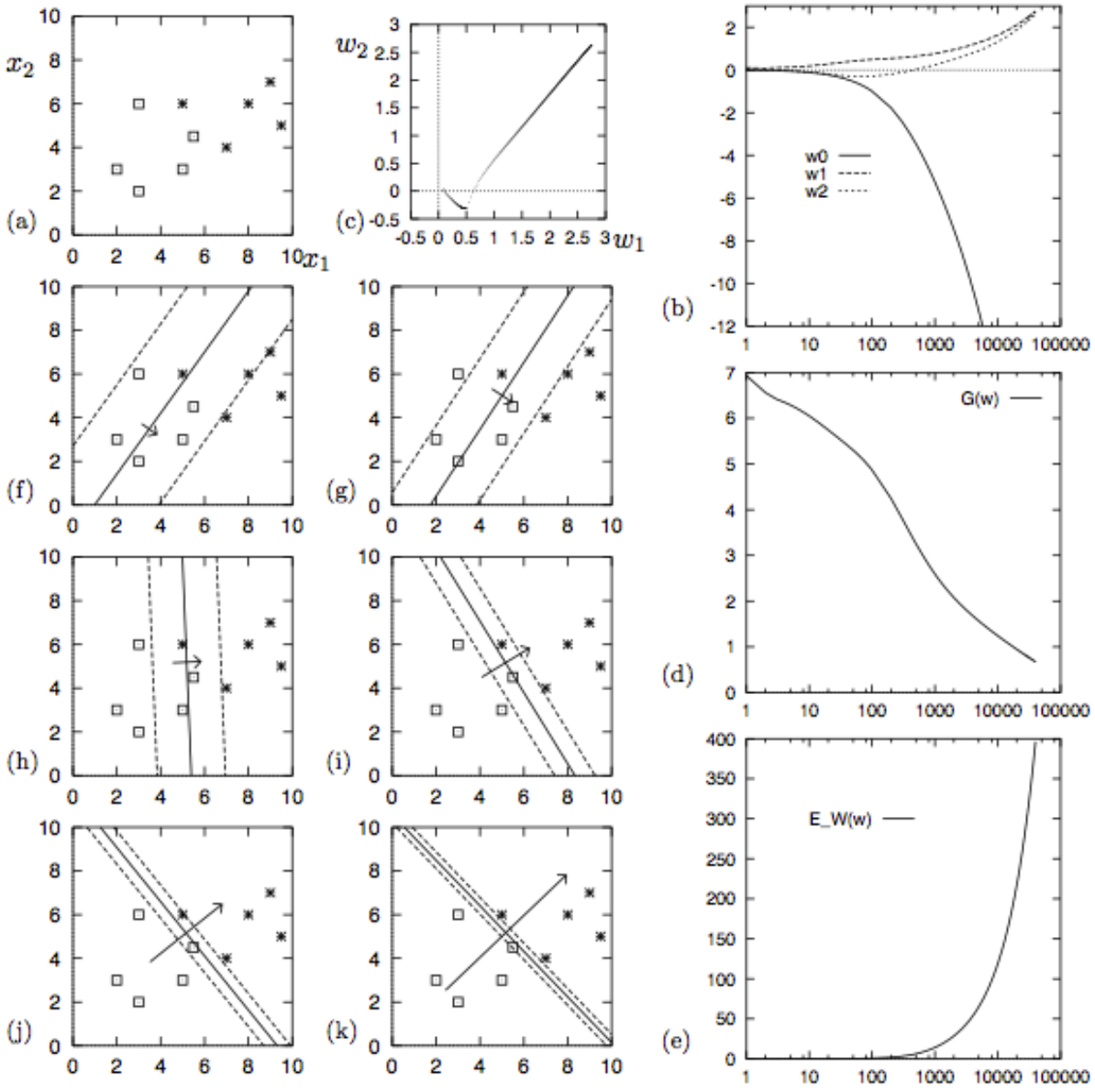
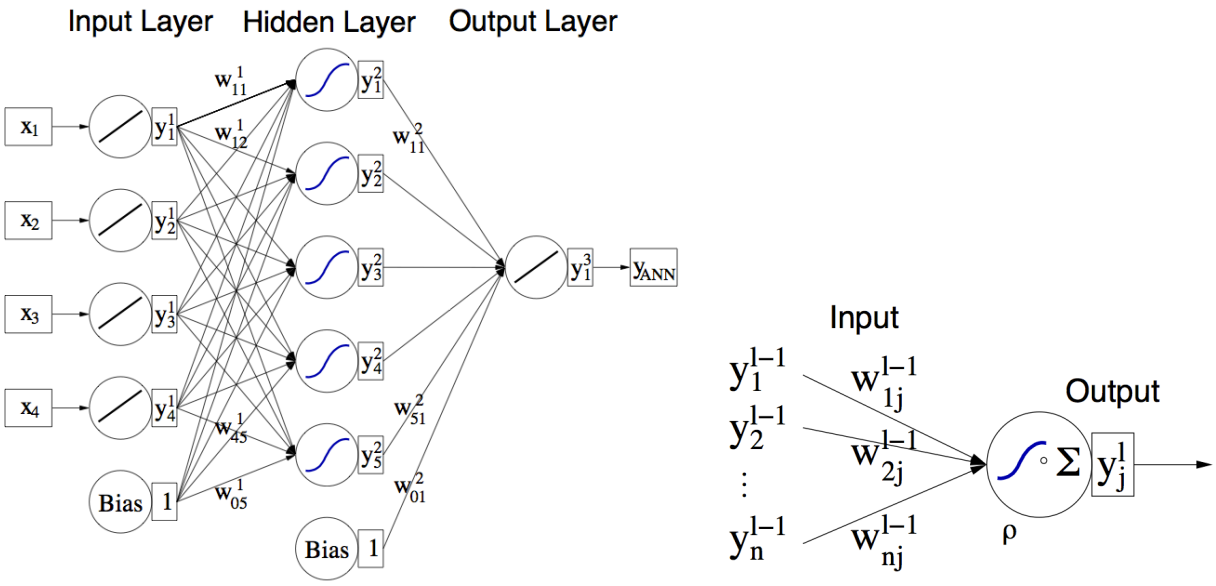


FIGURE 3.8: Illustration de la phase d'apprentissage d'un neurone (Crédits : MacKay [2003]). (a) Données utilisées lors de la phase d'apprentissage. (b) Évolution des poids w_0, w_1 et w_2 selon le nombre d'itérations. (c) Évolution des poids w_0 et w_1 et w_2 dans leur espace. (d) Évolution de la fonction d'erreurs de mauvaises classifications selon le nombre d'itérations. (e) Évolution de la norme du vecteur poids en fonction du nombre d'itérations. (f-k) La fonction permettant de séparer les données en fonction de 30, 80, 500, 3000, 10 000 et 40 000 itérations (trait plein). Y sont aussi représentés un vecteur proportionnel à (w_1, w_2) et des contours correspondant à une variation de la fonction de séparation de $a = \pm 1$ (traits pointillés).



(a) (b)

FIGURE 3.9: (a) Représentation d'un réseau de neurones MLP constitué d'une couche de neurones d'entrée, d'une couche de neurones cachée et d'un neurone de sortie. (b) Un neurone appartenant à la couche l, avec n entrées, portant toutes un poids w_{ij}^{l-1}, et une sortie (Crédits : [Hoecker et al., 2007])

(a). Dans cette exemple, l'évolution des poids en fonction du nombre d'itérations est représentée sur la sous-figure (b). On peut trouver l'évolution des poids w_1 et w_2 dans leur espace sur la sous-figure (c). Le taux de mauvaises classifications est représenté en fonction du nombre d'itérations sur la sous-figure (d), et on peut trouver l'évolution de la norme du vecteur (w_1, w_2) sur la sous-figure (e). Sur les sous-figures (f-k), on peut trouver la séparation linéaire entre les deux classes d'évènements pour les nombres d'itérations suivants : 30, 80, 500, 3000, 10 000 et 40 000, auxquels correspondent les poids représentés sur la sous-figure (b). On constate que les données sont de mieux en mieux séparées en fonction du nombre d'itérations.

Néanmoins, on constate que les poids divergent à partir de quelques milliers d'itérations ce qui correspond à un cas de surentraînement : le neurone « colle » aux données. Différentes méthodes existent afin de lutter contre le surentraînement, comme la régularisation, qui consiste à introduire un terme supplémentaire dans la fonction d'erreur $\mathcal{E} \to \mathcal{E} + \mathcal{R}$, qui est généralement une fonction de la norme du vecteur $\mathbf{w}$, afin d'empêcher la divergence des poids. Dans cette thèse, nous utiliserons une méthode qui permet de contrôler la convergence avec les échantillons d'entraînement et de test (section 3.5.2.3).

3.5.2 Du simple neurone au réseau MLP

3.5.2.1 La fonction à ajuster

Un neurone seul ne permet pas de résoudre des problèmes non-linéaires. Pour y remédier, le réseau de neurones MLP [Rumelhart et al., 1986] est constitué d'au moins une couche de neurones intermédiaires, dites couches « cachées ». Celles-ci sont situées entre la couche d'entrée, qui est constituée d'autant de neurones qu'il y a de variables dans le problème de classification, et la couche de sortie qui composée d'un neurone unique dans un problème de classification binaire. Chaque neurone n'est autorisé à avoir que des connections directes avec les neurones composant la couche suivante (Figure 3.9). Dans le cas d'un réseau constitué d'une unique couche cachée de N_c neurones avec une fonction d'activation de type tangente hyperbolique avec N_{var} variables

discriminantes, et d'un neurone de sortie ayant une fonction d'activation linéaire, la sortie du réseau est :

$$\zeta = \sum_{j=1}^{N_c} y_j^{(2)} w_{j1}^{(2)} = \sum_{j=1}^{N_c} \tanh\left(\sum_{i=1}^{N_{\text{var}}} x_i w_{ij}^{(1)}\right) w_{j1}^{(2)} \tag{3.6}$$

L'entraînement du réseau est réalisé avec N évènements ayant a caractéristiques $\mathbf{x}_a = (x_1, \cdots, x_{\text{var}})_a$ et un label $y_a = \{0, 1\}$, valant $+1$ pour les évènements « signal ». La fonction erreur à minimiser peut s'écrire :

$$\mathcal{E}(\mathbf{x}, \mathbf{w}) = \sum_{a=1}^{N} \mathcal{E}_a(\mathbf{x}_a, \mathbf{w}) = \frac{1}{2} \sum_{a=1}^{N} (\zeta - y_a)^2 \tag{3.7}$$

où $\mathbf{w}$ représente l'ensemble des poids du réseau.

3.5.2.2 Algorithme d'apprentissage : rétropropagation du gradient

Dans cette thèse, nous utiliserons l'algorithme de rétropropagation du gradient afin d'obtenir les poids permettant de minimiser l'erreur totale. Celui-ci permet d'ajuster les poids selon la direction pour laquelle l'erreur est minimale en utilisant exclusivement les dérivées premières de la fonction d'erreur. À chaque itération k les poids sont augmentés d'un facteur :

$$\Delta \mathbf{w} = \mathbf{w}^{k+1} - \mathbf{w}^k = -\eta \nabla_{\mathbf{w}} \mathcal{E} \tag{3.8}$$

où η représente le taux d'apprentissage, et permet d'ajuster la vitesse pour atteindre le maximum. Dans le cas d'un algorithme séquentiel (en anglais *on-line*), lors duquel chaque poids est augmenté à chaque passage d'un évènement dans le réseau [2], les poids connectés au neurone de sortie sont augmentés de :

$$\Delta w_{jl}^{(2)} = -\eta \frac{\partial \mathcal{E}_a}{\partial w_{jl}^2} = -\eta (\zeta_a - y_a) y_{j,a}^{(2)} \tag{3.9}$$

et ceux directement connectés à la couche de neurones cachée, sont à leur tour augmentés de [3] :

$$\Delta w_{ij}^{(1)} = -\eta \frac{\partial \mathcal{E}_a}{\partial w_{ij}^{(1)}} = -\eta (\zeta_a - y_a) y_{j,a}^{(2)} (1 - y_{j,a}^{(2)}) w_{j1}^{(2)} x_{i,a} \tag{3.10}$$

On définit un cycle lorsque l'ensemble des évènements composant l'échantillon d'entraînement a été utilisé.

On signalera qu'un second algorithme est très largement utilisé, l'algorithme « BFGS » [4], qui utilise les dérivées secondes afin d'obtenir l'optimum de la fonction erreur ce qui permet de réduire le nombre d'itérations, au détriment d'un temps de calcul légèrement plus long par itération.

3.5.2.3 Arrêt de l'entraînement

Dans cette thèse nous utiliserons une technique implémentée dans la librairie *Toolkit for Multivariate Data Analysis* (TMVA) afin d'arrêter la phase d'apprentissage pour éviter d'arriver dans un régime de surentraînement. Celle-ci consiste à comparer les taux de mauvaises classifications sur les lots d'entraînement et de test tous les N_{cycle}. L'entraînement est arrêté lorsque lorsque les erreurs sur les deux lots convergent tous les deux.

2. dans le cas où l'on cumule tous les évènements pour incrémenter les poids, on parle d'algorithme *batch*.
3. en utilisant une dérivation en chaîne et : $\tanh' x = \tanh x(1 - \tanh x)$
4. Le nom vient des quatre créateurs de l'algorithme, Broyden, Fletcher, Goldfarb et Shannon. Voir [Hoecker et al., 2007] pour plus d'informations.

3.5.2.4 Utilisation dans le manuscrit

Nous l'utiliserons au cours de la Partie II dans le chapitre 10 lors de la recherche des candidats blazars parmi les sources non-identifiées *Fermi-LAT* du catalogue 2FGL. Dans cette même partie, cette fois au chapitre 11, nous utiliserons un réseau de neurones de type MLP pour construire une classification binaire afin d'identifier les contreparties infrarouges de type blazar des sources non-identifiées du catalogue 2FGL.

Bibliographie

R. E. Duda, P. E. Hart, et D. G. Stork. *Pattern Classification*. Wiley, 2004.

Yoav Freund. Boosting a weak learning algorithm by majority. *Information and Computation*, 2 (121):256–295, 1995.

Yoav Freund et Robert E. Schapire. Experiment with a new boosting algorithm. In *Machine Learning: Proceedings of the Thirteenth International Conference*, 1996.

Yoav Freund et Robert E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 1997.

Yoav Freund et Robert E. Schapire. A short introduction to boostin. In *In Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, pages 1401–1406. Morgan Kaufmann, 1999.

A. Hoecker, et al. TMVA - Toolkit for Multivariate Data Analysis. *ArXiv Physics e-prints*, March 2007.

David J. C. MacKay. *Information Theory, Inference, and Learning Algorithms*. 2003.

David E. Rumelhart, Geoffrey E. Hinton, et Ronald J. Williams. Learning representations by back-propagating errors. *Nature*, 1986.

Robert E. Shapire. The strength of weak learnability. *Machine Learning*, 5:197–227, 1990.

Deuxième partie

L'expérience *H.E.S.S.*, analyses et performances en mode hybride

Chapitre 4

L'expérience *H.E.S.S.*

Sommaire

4.1 Imagerie atmosphérique Tcherenkov 55
4.2 Le réseau *H.E.S.S.* 60
4.3 Calibration des données 64
4.4 Prises de données 67
4.5 Qualité des données 68
Bibliographie 70

4.1 Imagerie atmosphérique Tcherenkov

La technique d'imagerie atmosphérique Tcherenkov utilise l'atmosphère comme un calorimètre et permet de faire ressortir les gerbes atmosphériques initiées par des rayons γ de très haute énergie parmi les gerbes – beaucoup plus nombreuses – dues aux interactions entre les rayons cosmiques chargés et l'atmosphère. Un photon γ interagit via le processus de création de paires avec le champ électrique d'un atome de l'atmosphère ce qui déclenche le développement d'une cascade de particules. Les particules chargées relativistes de la gerbe se déplaçant plus rapidement que la vitesse de la lumière c dans le milieu, les atomes émettent une émission de photons Tcherenkov [Jackson, 1998]. Ce sont ces photons qui sont captés par des télescopes d'imagerie atmosphérique Čerenkov (IACT) au sol qui permettent d'avoir une image de la gerbe et d'obtenir les caractéristiques de la particule incidente.

4.1.1 Développement des gerbes EM et hadroniques

Deux types de gerbes nous intéresseront : les gerbes électromagnétiques initiées par un photon ou un électron de haute énergie et les gerbes hadroniques ayant pour particule primaire un proton ou un noyau du rayonnement cosmique.

Les trois processus physiques dominants dans le développement de la gerbe EM sont la création de paires e^+e^-, le rayonnement de freinage des électrons et les pertes d'énergie des électrons par ionisation. Dans le cas des gerbes initiées par un photon, celui-ci interagit avec le champ EM d'un atome de l'atmosphère pour donner une paire électron-positron. Ces derniers, tout comme la particule primaire dans le cas des gerbes initiées par un électron, émettent à leur tour des photons de haute énergie par effet Bremsstrahlung. Le processus se répète jusqu'à ce que les particules aient une énergie trop faible pour produire des paires électron-positron ou du rayonnement de freinage. On définit E_c comme l'énergie critique en dessous de laquelle les pertes d'énergie des électrons par ionisation des molécules de l'air deviennent dominantes (dans l'air : $E_c \approx 85\,\mathrm{MeV}$), ce qui freine le développement de la gerbe.

De manière schématique, considérons le modèle simplifié d'Heitler [Heitler, 1954] qui permet d'obtenir les ordres de grandeurs du développement d'une gerbe EM. Considérons X_0, la longueur d'interaction dans une couche de matériau de numéro atomique Z et de nombre de masse A définie par [Beringer et al., 2012], comme :

$$X_0 = \frac{716{,}4\,\mathrm{g\,cm^{-2}}\,A}{Z(Z+1)\ln(287/\sqrt{Z})} \tag{4.1}$$

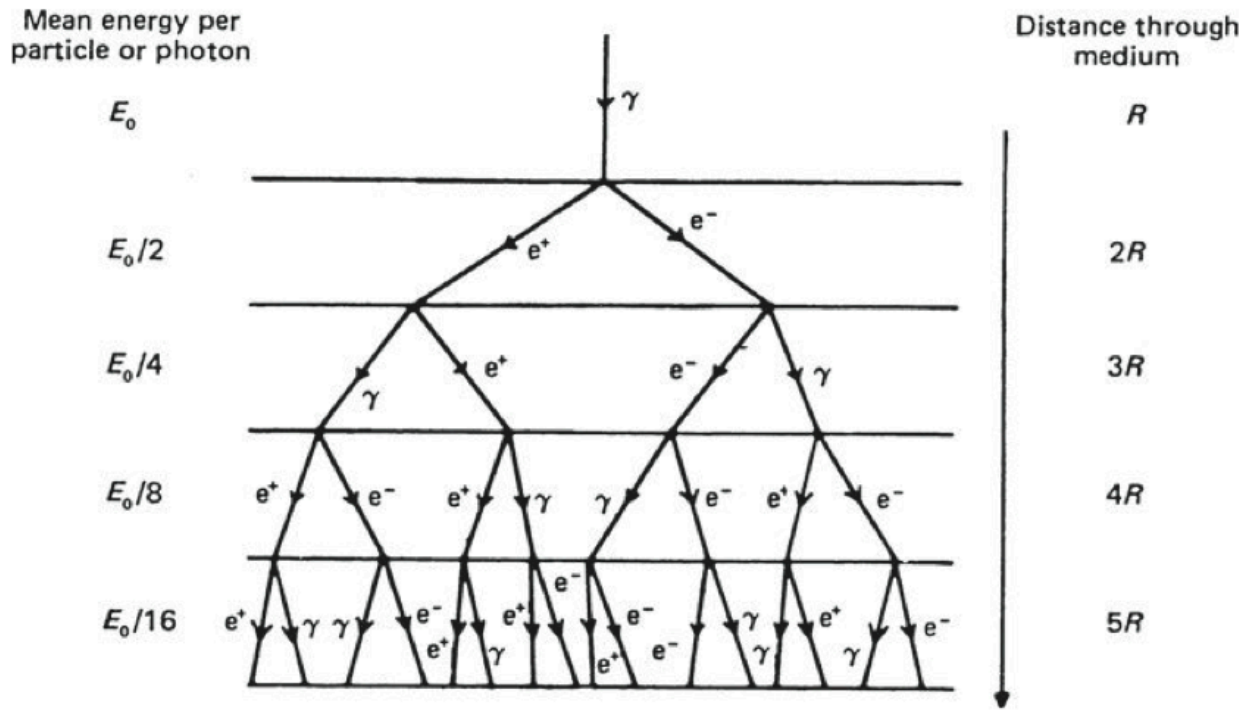


FIGURE 4.1: Modèle d'Heitler d'une gerbe électromagnétique (Crédits : [Longair, 2011]).

Il s'agit de la longueur caractéristique pour laquelle un électron relativiste ne garde en moyenne que $1/e$ de son énergie initiale, le reste étant cédé au milieu par rayonnement Bremsstrahlung (dans l'air : $X_0 \approx 37\,\mathrm{g\,cm^{-2}}$). Elle représente aussi 7/9 du libre parcours moyen d'un photon de haute énergie avant qu'il n'interagisse via le processus de création de paires. L'approximation d'Heitler consiste à supposer que pour une même épaisseur d'atmosphère traversée, que l'on notera $X = \int_0^z \rho(z')dz'$ [1], la perte d'énergie moyenne des électrons ou des photons sera la même :

$$E(X) = E_0 \exp\left(-\frac{X}{X_0}\right) \tag{4.2}$$

Ainsi, au bout d'une distance parcourue $R = X_0 \ln 2$, un électron cédera la moitié de son énergie à un photon par rayonnement de freinage et un photon cédera l'intégralité de son énergie, également répartie, à une paire e^+e^-. Le développement de la gerbe est illustré sur la Figure 4.1. On en déduit la profondeur à laquelle le maximum de gerbe est atteint, $X_{\max} = X_0 \ln \frac{E_0}{E_c}$, correspondant à une altitude du maximum de gerbe de :

$$z_{\max} = z_0 \ln\left(\frac{\rho_0 z_0}{X_0 \ln(E_0/E_c)}\right) \tag{4.3}$$

Dans le cas des gerbes hadroniques, la collision d'un hadron avec un composant de l'atmosphère initie le développement d'une gerbe qui, contrairement au cas EM, aura plusieurs composantes et un développement transverse plus important. Les processus hadroniques sont à l'origine de création de nucléons, de noyaux plus légers que la particule incidente si celle-ci est un noyau lourd, des mésons K et des pions π. Ces derniers produisent des composantes EM par l'intermédiaire de la désintégration $\pi^0 \to \gamma\gamma$. Une composante muonique est aussi présente du fait de la désintégration des pions et des mésons chargés. De la même manière que pour les gerbes EM, il est possible de construire un modèle simple qui permet d'obtenir les ordres de grandeurs du développement de la gerbe. Une discussion détaillée peut être trouvée dans [Matthews, 2005].

1. le profil de masse volumique de l'atmosphère peut être approximé par : $\rho(z) = \rho_0 e^{-z/z_0}$ où $\rho_0 \approx 1{,}2 \times 10^3\,\mathrm{g\,cm^{-3}}$

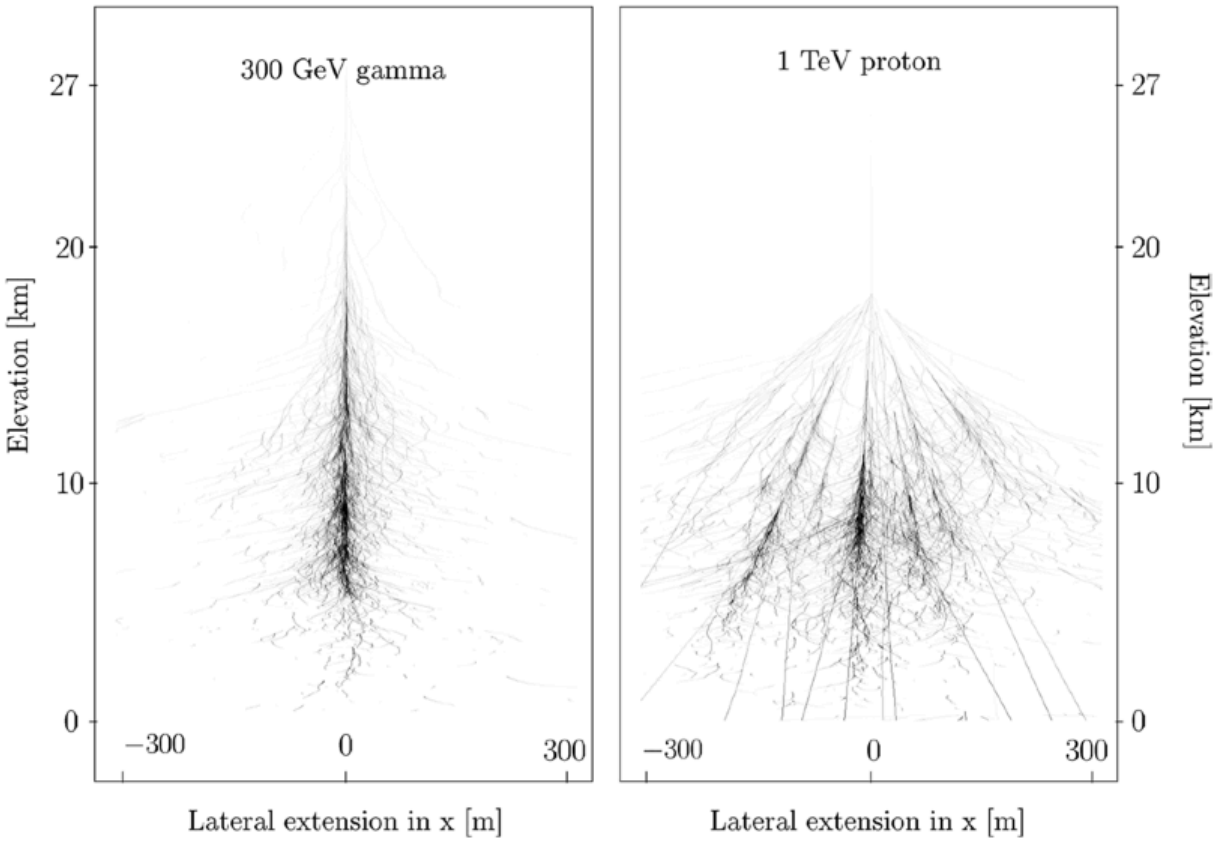


FIGURE 4.2: Simulations d'une gerbe EM de 300 GeV et d'une gerbe hadronique initiée par un proton de 1 TeV (Crédits : [Aharonian et al., 2008]).

La Figure 4.2 représente une gerbe EM produite par un photon γ de 300 GeV ainsi qu'une gerbe initiée par un proton de 1 TeV, toutes deux calculées par simulation. À l'inverse des gerbes EM dans lesquelles il existe une dispersion transversale relativement faible ($\sim$1 m), les gerbes hadroniques ont un important transfert de quantité de mouvement transverse. Cette propriété est fondamentale pour la discrimination γ/hadrons (chapitre 5).

4.1.2 L'effet Tcherenkov

Lors de la traversée de l'atmosphère par une particule chargée relativiste, celle-ci peut perturber le cortège électronique des atomes la constituant et induire une émission de photons Tcherenkov dans l'optique et dans l'UV. Le nombre de photons Tcherenkov émis par unité de longueur et par intervalle de longueur d'onde $[\lambda,\ \lambda + \mathrm{d}\lambda]$ au passage d'une particule relativiste de vitesse v et de charge Z s'écrit :

$$\frac{\mathrm{d}^2 N}{\mathrm{d}x \mathrm{d}\lambda} = \frac{2\pi\alpha Z^2}{\lambda^2}\left(1 - \frac{c}{vn(\lambda)}\right) \tag{4.4}$$

où $\alpha = \frac{e^2}{4\pi\varepsilon_0 \hbar c} \approx 1/137$ est la constante de structure fine et $n(\lambda)$ l'indice de réfraction de l'air pour la longueur d'onde λ, pourvu que la relation $v > c/n(z)$ soit vérifiée. Cette relation exprime le fait que les fronts d'onde émis par le milieu produisent des interférences constructives. Dans le cas contraire, aucun rayonnement n'est émis. La Figure 4.3 représente deux cas : le premier à gauche pour lequel les interférences sont destructives ($v \leq c/n(\lambda)$), et le second à droite, qui permet d'avoir des interférences constructives et donc une émission de photons. Par analogie avec le bang

et $z_0 \approx 8{,}5$ km.

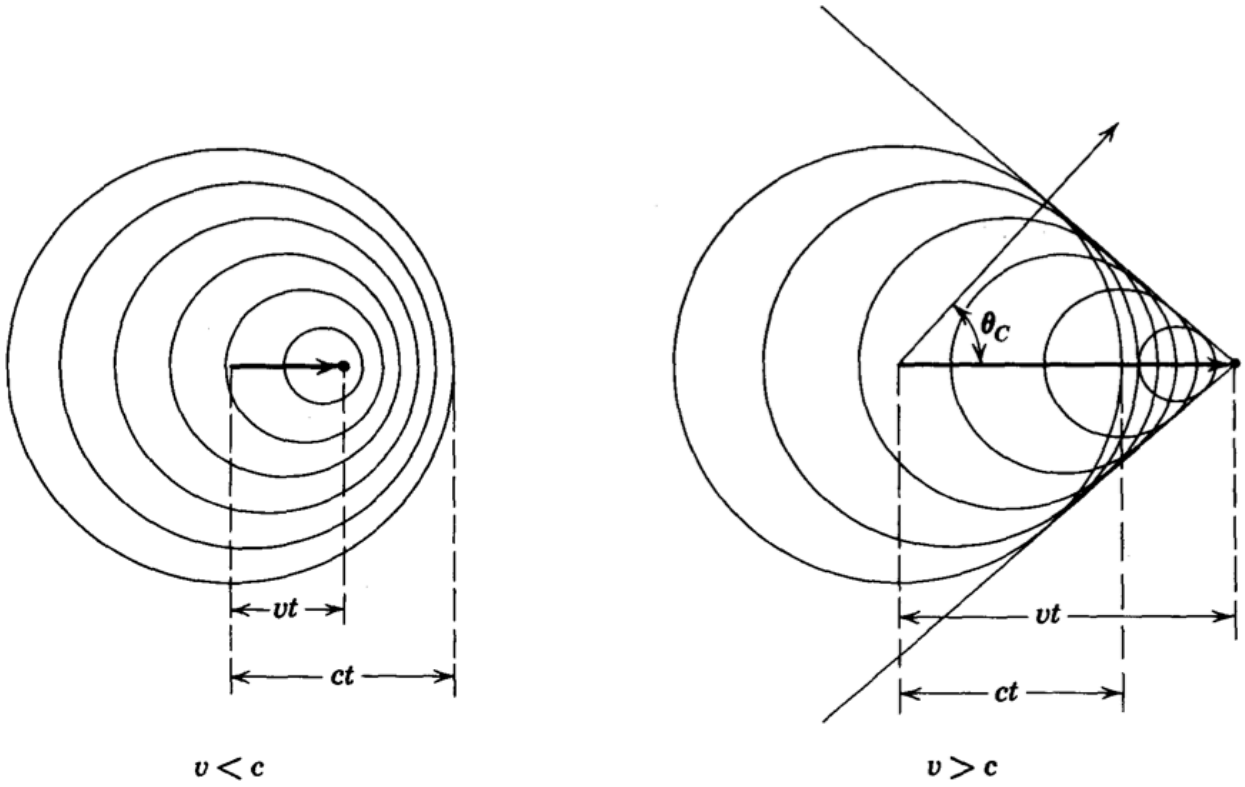


FIGURE 4.3: Principe de l'effet Tcherenkov (Crédits : [Jackson, 1998]).

supersonique – lorsqu'un avion dépasse la vitesse du son dans l'air par exemple – l'émission de photons Tcherenkov s'effectue dans un cône ayant un angle d'ouverture θ_c :

$$\cos\theta_c(z) = \frac{c}{vn(z)} \tag{4.5}$$

où z représente l'altitude de la particule. En première approximation, et en considérant que l'air est sec, la dépendance de l'indice de réfraction à la densité s'exprime selon : $n(z) - 1 \approx 2{,}9 \times 10^{-4} \exp(-z/h_0)$ avec $h_0 \approx 7{,}3\,\text{km}$. L'angle θ_c va donc décroître en fonction de l'altitude. À titre d'exemple, il est de l'ordre de 0,66° (0,74°) pour une altitude de 10 km (8 km), ce qui donne en première approximation, au niveau de la mer, une tache de lumière de densité constante sur une région circulaire de ∼100 m de rayon (à gauche de la Figure 4.4). Au-delà de ce rayon, la diffusion multiple des électrons secondaires de la gerbe entraîne une décroissance exponentielle de la densité de photons. À droite de la Figure 4.4 est représentée la densité moyenne de photons émis le long du développement de gerbes EM de 100 GeV.

4.1.3 Détection des gerbes

La technique utilisée par les IACT consiste à focaliser à l'aide de grands miroirs les photons Tcherenkov dans le visible sur des caméras constituées de photomultiplicateur (PM) et dotées d'une électronique rapide qui transforme la lumière en signal électrique mesurable. Cela équivaut à prendre une photo à très haute vitesse d'obturation afin d'isoler le signal Tcherenkov très tenu du bruit de fond du ciel. L'image de la gerbe obtenue servira de base à la reconstruction de la direction et de l'énergie de la particule incidente (voir chapitre 5).

En considérant un télescope doté d'un miroir de surface A_m, de pixels de taille angulaire Ω_{pix}, d'un temps d'intégration du signal τ et en notant ε le produit de l'efficacité quantique des PM et de la transmittance de l'atmosphère, le rapport signal sur bruit correspondant à la détection d'une

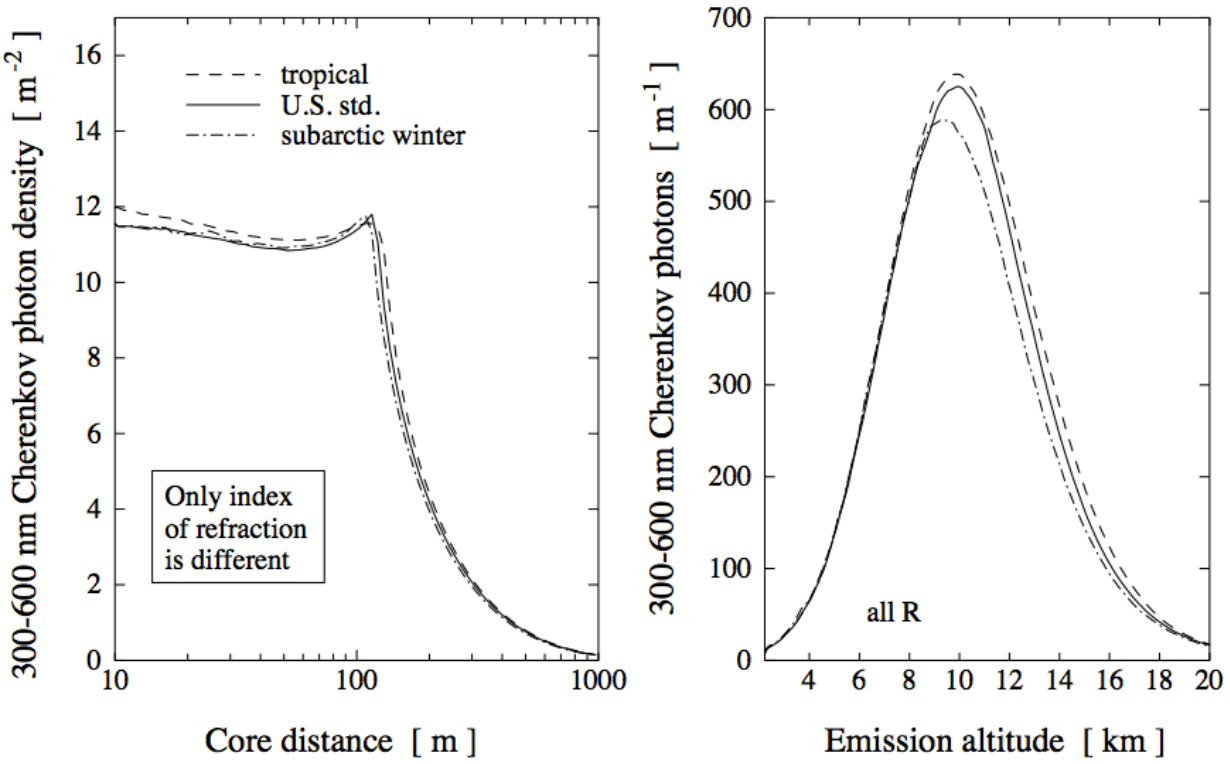


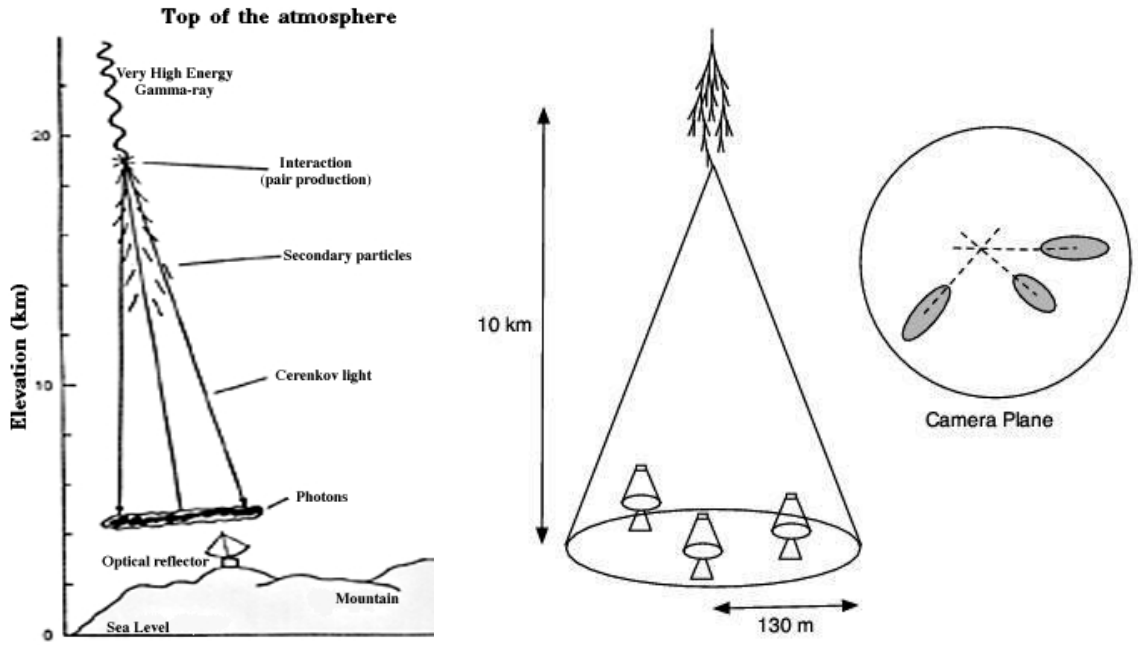
FIGURE 4.4: Densité de photons Tcherenkov moyenne en fonction du paramètre d'impact (avec absorption de l'atmosphère) et de l'altitude le long l'axe de gerbe EM (sans absorption atmosphérique) pour différent modèle d'atmosphère. 2000 gerbes γ de 100 GeV ont été simulées au zénith pour une altitude d'observation de 2000 m (Crédits : [Bernlöhr, 2000]).

gerbe dans un pixel s'écrit :

$$N_\sigma = \frac{S}{\sqrt{B}} = \frac{\Phi_c A_m \Delta t_c \varepsilon}{\sqrt{\Phi_{NSB} A_m \Omega_{pix} \varepsilon \tau}} \propto \left(\frac{A_m \varepsilon}{\Phi_{NSB} \Omega_{pix} \tau} \right)^{\frac{1}{2}} \tag{4.6}$$

où Φ_c représente le flux de photons Tcherenkov émis par une gerbe EM initiée par un γ incident, Δt_c le temps moyen du flash Tcherenkov et enfin Φ_{NSB} le flux du bruit de fond du ciel (NSB) (par unité d'angle solide). Le flash Tcherenkov étant de l'ordre de quelques nanosecondes, le choix de τ est un compromis entre intégration du signal et minimisation du bruit de fond. Le choix de l'ouverture angulaire Ω_{pix} des pixels est déterminée par la dimension de la largeur de la gerbe vue du sol ($\sim 0{,}1°$ pour une gerbe de $\sim$100 GeV au zénith). Réduire Ω_{pix} permet d'augmenter le taux de détection des gerbes EM. Enfin, l'utilisation de grandes surfaces de collection diminue l'énergie seuil de détection, celle-ci d'autant plus petite que le rapport $S/\sqrt{B}$ est grand. Notons aussi que l'observation du ciel lors des nuits sans lune est nécessaire afin de réduire le NSB. Notons également que cette contrainte, ainsi que celle d'un ciel sans nuages, réduit d'environ 10 % le temps utile d'observation.

La technique de stéréoscopie, qui utilise plusieurs IACT, permet d'obtenir des images simultanées d'une même gerbe sous des angles différents (Figure 4.5b). Ces différentes images donnent la possibilité de mieux contraindre les caractéristiques des gerbes électromagnétiques, à savoir la direction et l'énergie du photon incident, et d'exploiter au mieux l'asymétrie axiale propre au développement des gerbes initiées par des hadrons afin de réduire le bruit de fond généré par celles-ci (chapitre 5).



(a) (b)

FIGURE 4.5: (a) Principes de l'imagerie atmosphérique Tcherenkov (Crédits : NASA) et (b) de la stéréoscopie (Crédits : http://www.hermanusastronomy.co.za/).

4.2 Le réseau *H.E.S.S.*

Dans cette section nous présentons le réseau de télescopes *H.E.S.S.*. Après une présentation du site sur lequel est installé l'expérience, les caractéristiques et le fonctionnement des télescopes moyens de type HESS-I et celles du nouveau télescope (CT_5) sont détaillés.

4.2.1 Le site

L'expérience *H.E.S.S.* est située en Namibie dans la région de Khomas aux coordonnées : S23° E16°, à 1800 m d'altitude. Ce lieu a été choisi pour la qualité de son ciel et son climat désertique, bien adapté pour des observations astronomiques, ainsi que pour sa faible densité de population (2,6 hab./km^2). En outre, cette localisation possède un champ magnétique plus faible que la moyenne terrestre (d'un facteur ∼2,5), limitant ainsi la déflexion des particules chargées dans l'atmosphère, et permet l'observation des parties centrales de la Galaxie dans des conditions optimales (à faible angle zénithal).

Le réseau est constitué de quatre télescopes de taille moyenne (12 m de diamètre), de type HESS-I, et d'un cinquième télescope (28 m de diamètre), récemment installé, que l'on appellera CT_5 (Figure 4.6). Les quatre télescopes moyens sont situés sur les coins d'un carré de 120 m de coté, les télescopes formant un axe nord-sud et un axe est-ouest, et le cinquième télescope est placé au centre du carré. La configuration initiale des télescopes de type HESS-I a été pensée afin de récolter de manière optimale les photons dans le cône de lumière Tcherenkov généré par une gerbe EM et limiter le taux de déclenchement en stéréoscopie des évènements muonique [Guy, 2003]. Le placement du cinquième télescope au centre du réseau permet de réduire le seuil en énergie et d'améliorer la sensibilité.

En plus des télescopes, une station météorologique permet de mesurer la température, la pression, la vitesse et la direction du vent ainsi que le taux d'humidité de l'air. Enfin un radiomètre permet de monitorer la progression des nuages dans le champ de vue et un lidar mesure la quantité

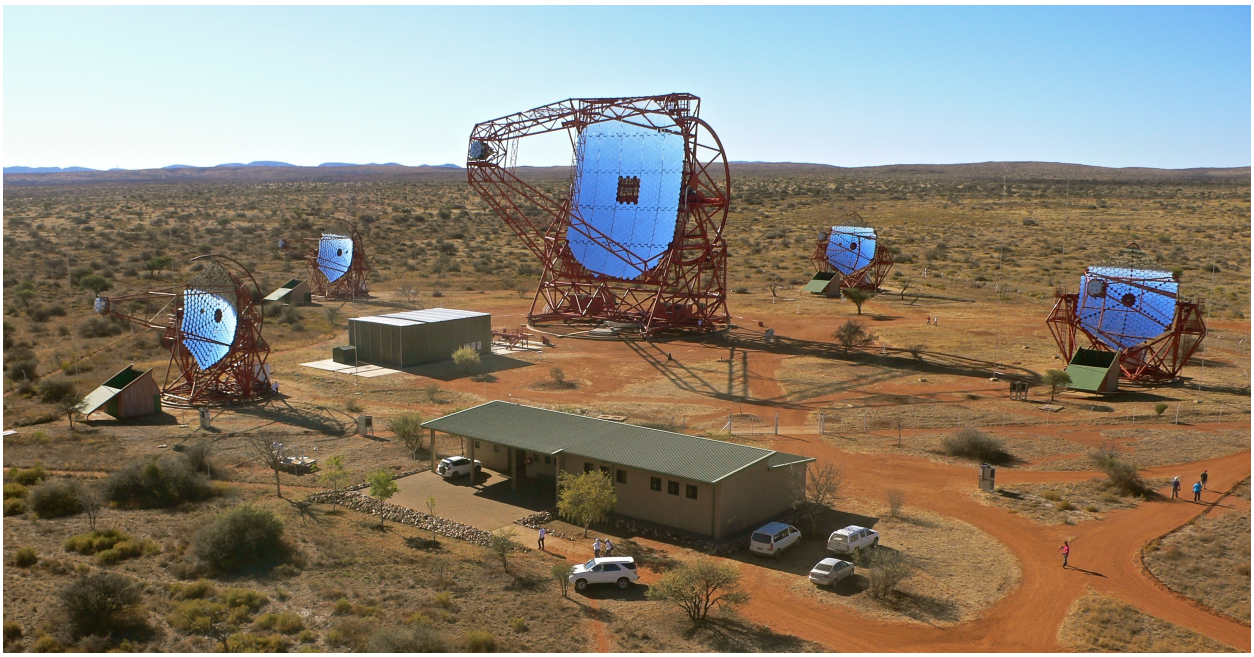


FIGURE 4.6: Réseau de télescopes de l'expérience *H.E.S.S.* (Crédits : S. Klepser). La salle de contrôle est située dans le bâtiment au premier plan.

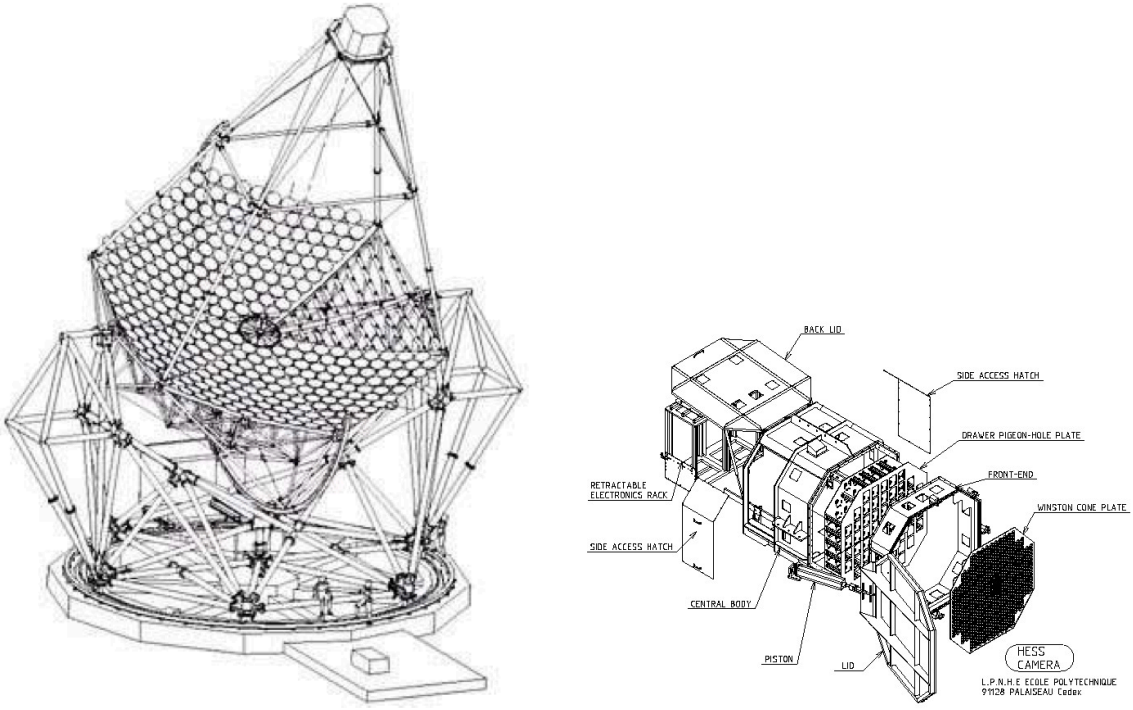
d'aérosols en fonction de l'altitude.

4.2.2 Télescopes de type HESS-I

Les quatre télescopes de type HESS-I sont constitués d'une monture altazimutale actionnée par plusieurs moteurs qui mettent en mouvement des roues en contact avec des rails en azimut et en altitude permettant d'atteindre une vitesse en azimut de $200\,^\circ\,\mathrm{min}^{-1}$ et une vitesse en élévation de $100\,^\circ\,\mathrm{min}^{-1}$. Un ensemble de 382 miroirs circulaires aluminisés de 60 cm de diamètre constitue le réflecteur principal (Figure 4.7a). Au début de l'expérience, chaque miroir avait une réflectivité d'au moins 80 % entre 300 et 600 nm. Celle-ci a diminué avec le temps, et une campagne de ré-aluminisation a été entreprise entre avril 2010 et octobre 2011 pour leur rendre leur réflectivité d'origine. Chaque miroir est attaché à un support motorisé qui permet de régler son orientation lors de l'installation ou après des opérations de maintenance. Les miroirs sont assemblés selon la technique Davies-Cotton [Lewis, 1990] qui permet de limiter les effets d'aberrations hors-axe tout en assurant un étalement en temps limité (isochronisme). Ils sont positionnés selon une sphère de 15 m de rayon, longueur qui correspond à la longueur focale du télescope.

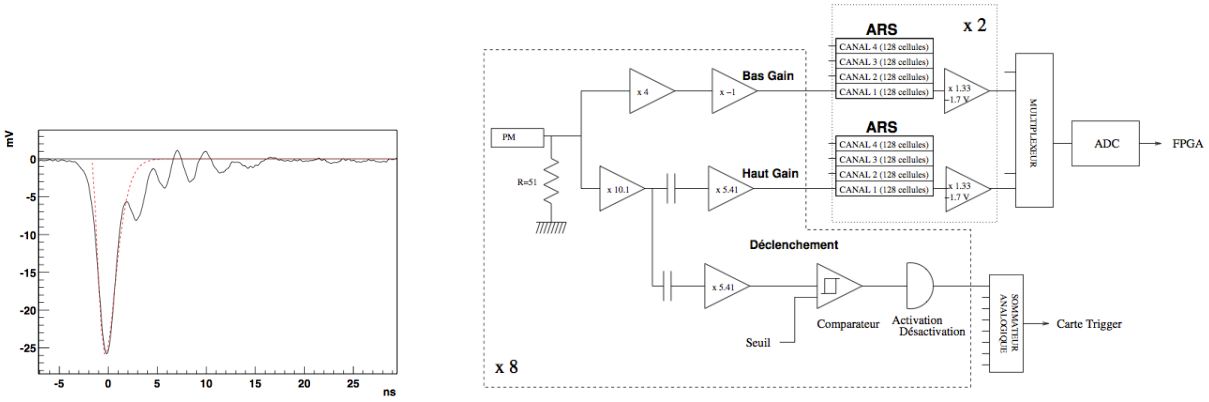
La caméra [Vincent et al., 2003], représentée en vue éclatée sur la Figure 4.7b, est placée à la distance focale du miroir. Elle est constituée de 960 PM construits par Photonis (modèle XP-2960), d'ouverture angulaire de 0,16° qui sont assemblés selon une structure en nid d'abeille pour un champ de vue total de 5°. Des cônes de Winston compensent la perte de lumière due aux espaces morts entre les pixels et permettent de réduire la collecte des photons parasites qui sont réfléchis par le détecteur ou par le sol, en particulier lors d'observations à faible élévation. La fenêtre d'ouverture des PM est en verre borosilicate et transmet les photons à partir de 300 nm. Ceux-ci sont ensuite convertis en photoélectrons par une photocathode en bialkali qui possède une efficacité quantique de l'ordre de 25 % au pic. Des dynodes au nombre de 8 permettent d'amplifier le signal qui est ensuite transmis au niveau de l'anode. On peut observer sur la Figure 4.8a le signal moyen correspondant à un photoélectron en fonction du temps, mesuré à l'aide d'un oscilloscope à la sortie d'un PM.

Les PM sont arrangés par groupes de 16 dans des tiroirs qui contiennent leur électronique. Des cartes analogiques chargées de lire et de déclencher l'enregistrement du signal sont placées à



(a) (b)

FIGURE 4.7: (a) Télescope de type HESS-I (Crédits : [Bernlöhr et al., 2003]) et (b) sa caméra (Crédits : https://www.mpi-hd.mpg.de/hfm/HESS/).



(a) (b)

FIGURE 4.8: (a) Signal moyen d'un photoélectron mesuré au niveau de l'anode d'un PM avec un oscilloscope en fonction du temps. (b) Schéma d'une carte analogique avec la voie de déclenchement et les voies de lecture d'un PM (Crédits : [Guy, 2003]).

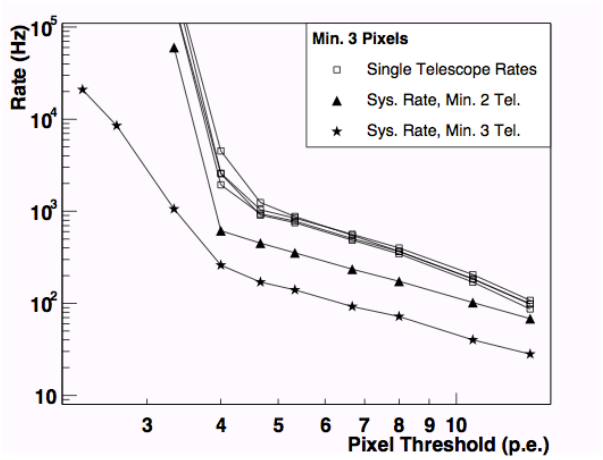


FIGURE 4.9: Taux de déclenchement en fonction de Q_0 des télescopes de type HESS-I pour $N_1 = 3$ (prise de données réelles, Crédits : [Funk et al., 2004]).

l'arrière des tiroirs (Figure 4.8b). Deux voies d'acquisition, l'une de haut gain et l'autre de bas gain (section 4.3), transmettent le signal de chaque PM à deux mémoires Analog Ring Sampling (ARS). Initialement développées pour l'expérience ANTARES, chaque mémoire ARS est constituée de 128 cellules dans lesquelles est échantillonné le signal d'un PM à un taux de 1 GHz. Lorsque la caméra reçoit l'ordre d'enregistrer un évènement, la mémoire analogique est lue pendant une fenêtre temporelle de 16 ns (correspondant à 16 cellules) et le contenu est numérisé à l'aide d'un convertisseur analogique-numérique (ADC) sur une carte munie d'un réseau de portes programmables (FPGA). Un système de déclenchement local équipe chaque caméra, composé de deux niveaux, appelés L_0 et L_1, qui permet d'influer sur la décision d'enregistrer un évènement et de diminuer le temps mort de la caméra. Le premier niveau, L_0, consiste à demander un signal dans un PM correspondant à une charge (en photoélectrons) supérieure à Q_0. Du fait de l'extension typique d'une gerbe EM de $\sim 1°$, un évènement γ induira un signal significatif dans plusieurs PM voisins, à l'inverse du bruit de fond du ciel, qui éclairera, de manière aléatoire et sans structure spatiale, l'ensemble de la caméra. Le second niveau de déclenchement, L_1, requiert d'avoir au minimum N_1 photomultiplicateurs signalant une charge au-delà de Q_0 dans un secteur – les caméras sont divisées en 38 secteurs de 64 PM, avec un recouvrement entre les secteurs – et qui coïncident temporellement. Ce travail est réalisé par un comparateur sur les cartes analogiques (Figure 4.8b). Sur la Figure 4.9 sont représentés les taux de déclenchement obtenus lors d'observations réelles pour des télescopes individuels de type HESS-I (carrés), ainsi que pour les coïncidences de deux (triangles) et trois télescopes (étoiles) en fonction du seuil en charge Q_0 pour $N_1 = 3$. On peut y observer une lente décroissance du taux de déclenchement pour les seuils $Q_0 \geq 5$ p.e., et en deçà une augmentation brutale du taux de déclenchement dû au NSB et aux coïncidences fortuites entre télescopes qui en résultent.

Dans le cas où le critère L_1 est vérifié, un ordre est envoyé à l'ensemble des tiroirs pour suspendre l'acquisition des données et un signal est envoyé au système de déclenchement central. Si dans une fenêtre temporelle de 80 ns le système d'acquisition central reçoit d'au moins deux caméras un signal correspondant à la validation du critère L_1 de déclenchement local, l'ordre est envoyé de récupérer les données. Le temps mort final est de l'ordre de 10 %. On pourra trouver une description détaillée du système de déclenchement du réseau *H.E.S.S.* avant l'ajout du cinquième télescope dans [Funk et al., 2004].

4.2.3 Le grand télescope pour la phase HESS-II

Le cinquième télescope, inauguré le 28 septembre 2012, est aujourd'hui le plus grand télescope Tcherenkov jamais construit au monde. Lorsque le télescope pointe vers le zénith, il atteint une

hauteur comparable à un immeuble de vingt étages pour un poids total de 580 tonnes. La monture altazimutale a été développée pour permettre de repositionner le cinquième télescope rapidement dans le cas où une alerte multi-longueur d'onde est donnée pour un phénomène transitoire de type sursaut γ. Le même principe que pour les télescopes moyens est utilisé ce qui permet à CT_5 de se repositionner avec une vitesse en élévation et en azimut de $100\,^\circ\,\mathrm{min}^{-1}$.

Afin de préserver l'isochronisme des temps d'arrivée des photons dans le plan focal, une structure parabolique a été choisie pour le réflecteur principal. Utiliser cette information – qui n'était pas disponible pour les télescopes moyens – peut conduire à une réduction significative du fond hadronique et à une amélioration de la reconstruction des gerbes [Aliu et al., 2009]. Le réflecteur du cinquième télescope est composé de 875 miroirs hexagonaux, de distance entre deux cotés parallèles de 90 cm, et possède une distance focale de 36 m, distance à laquelle est suspendue la caméra. Le réflecteur principal possède une surface totale de collecte de $614\,\mathrm{m}^2$.

La caméra du cinquième télescope [Bolmont et al., 2014] a une granularité bien plus fine que les quatre premières. En effet, elle est dotée de 2048 pixels, eux aussi groupés par tiroirs de 16 PM, chacun possédant une ouverture angulaire de 0,07°. Le champ de vue est de 3,2°. L'électronique d'enregistrement du signal suit la même logique que pour les caméras de type HESS-I. Pour des questions de capacité et de temps mort, les mémoires ARS ont été remplacées par des mémoires Swift Analogue Memory (SAM) qui permettent d'échantillonner le signal dans 256 cellules. Cela permet de conserver un temps mort similaire à celui du réseau composé uniquement des télescopes de taille moyenne, et ce malgré un taux de déclenchement pour CT_5 d'un ordre de grandeur supérieur. L'information temporelle enregistrée est le temps correspondant au maximum de l'amplitude du signal mesurée à la sortie d'un PM, il est calculé par des FPGA.

Pour la phase II de *H.E.S.S.* la stratégie de déclenchement du réseau HESS-I a été adaptée pour enregistrer les évènements où CT_5 déclenche seul. Cela correspond au mode mono, qui permet d'étudier des évènements de très basse énergie. Un seuil de déclenchement supplémentaire, noté L_2, équipe la cinquième caméra et permettra dans un futur proche de faire une analyse bas niveau, comme le calcul des premiers moments d'Hillas (chapitre 5, section 5.1), afin de baisser le seuil de CT_5 tout en gardant un temps mort acceptable. Il n'est à ce jour pas utilisé.

4.3 Calibration des données

La calibration des données consiste à convertir les données brutes mesurées en informations exploitables qui seront utilisées lors de l'analyse. Cette partie présente les grandes lignes de la calibration, une discussion détaillée est donnée dans [Aharonian et al., 2004]. La dernière section présente la méthode permettant d'estimer l'efficacité optique globale des télescopes.

4.3.1 Principe

À la sortie de chaque PM le signal analogique est amplifié dans une voie haut gain (High Gain) (HG) et une voie bas gain (Low Gain) (LG), puis enregistré dans les ARS ou les SAM (respectivement pour les télescopes de taille moyenne et CT_5), et est digitalisé à l'aide d'un ADC. La calibration consiste à convertir le nombre de coups d'ADC dans chaque voie en une quantité physique, le nombre de photon-électrons (p.e.). Celui-ci est donnée par :

$$A^{HG} = \frac{ADC^{HG} - P^{HG}}{\gamma_e^{ADC}} \times FF \tag{4.7}$$

$$A^{LG} = \frac{ADC^{LG} - P^{LG}}{\gamma_e^{ADC}} \times HG/LG \times FF \tag{4.8}$$

où ADC^X représente la mesure en coups d'ADC dans les voies haut gain et bas gain ($X = \{HG, LG\}$), P^X la position de la ligne de base correspondante (le piédestal), γ_e^{ADC} est le gain dans la voie haut gain, HG/LG le rapport des amplifications entre les voies haut gain et bas gain,

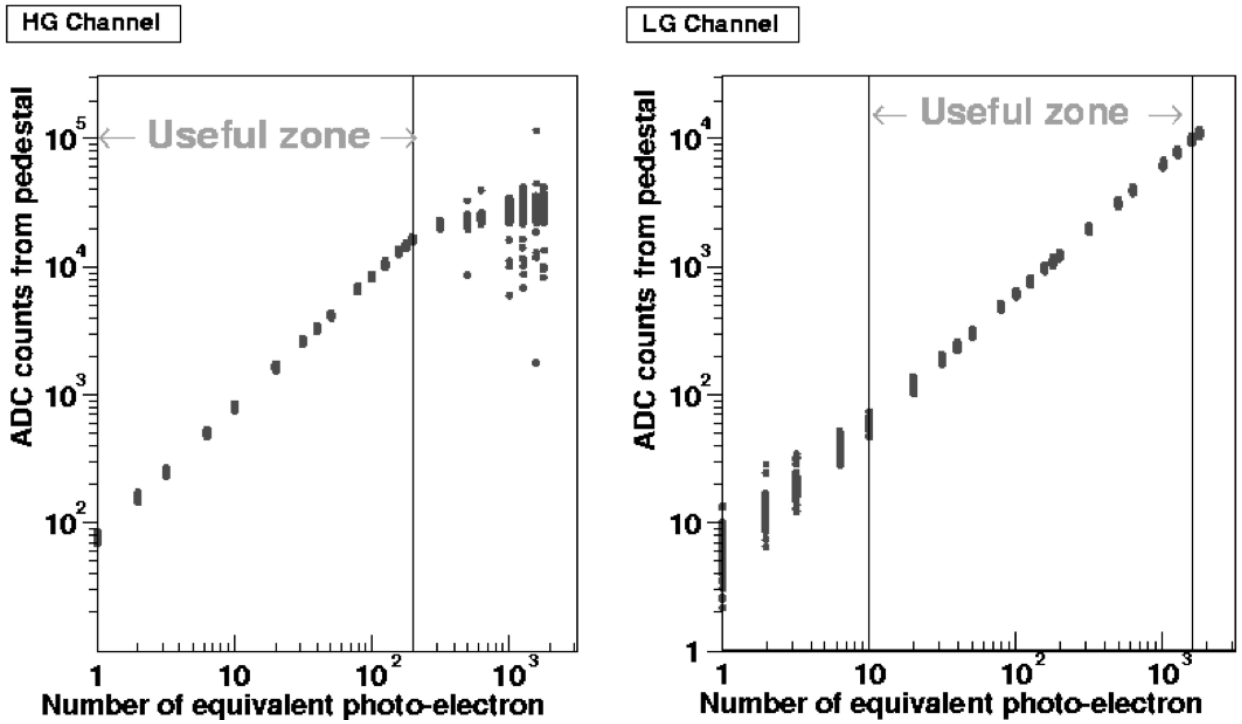


FIGURE 4.10: Nombre de coups ADC mesuré en fonction du nombre de photoélectrons pour les haut gains et bas gains pour une caméra de type HESS-I (Crédits : [Aharonian et al., 2004]).

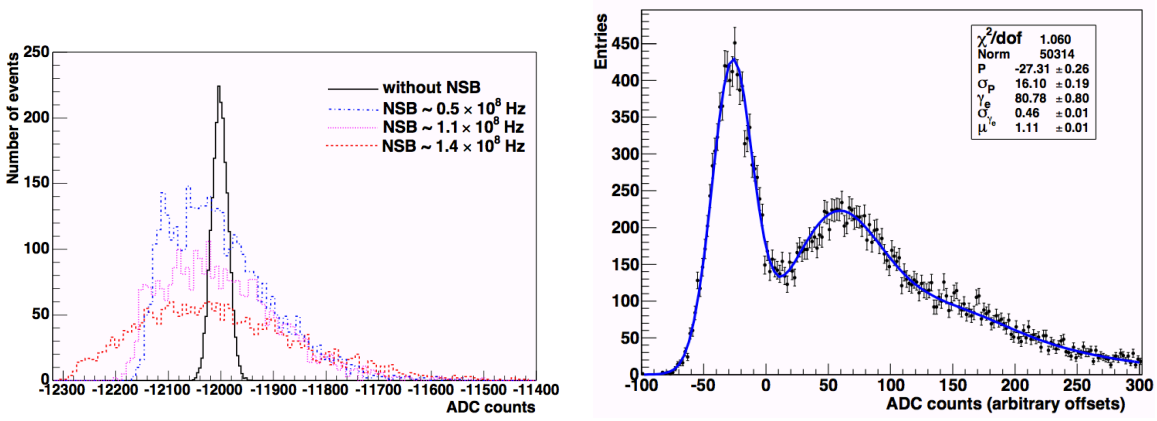
et enfin FF est le coefficient de *flat-field*. Du fait des régimes différents de linéarité entre le nombre de coups d'ADC et le nombre de photoélectrons pour les deux voies d'amplification (Figure 4.10), jusqu'à une valeur 150 p.e. on utilise la valeur A^{HG}, et au-delà de 200 p.e., la valeur de A^{LG}. Dans la zone intermédiaire une moyenne pondérée est utilisée.

4.3.2 Estimation des piédestaux P^{HG} et P^{LG}

Les piédestaux sont définis comme la valeur moyenne du nombre de coups d'ADC dans un pixel en l'absence de lumière Tcherenkov. Lorsque le capot de la caméra est fermé, le bruit de l'électronique de la caméra crée une distribution de coups d'ADC gaussienne, dont la largeur dépend de la température de l'électronique (trait noir sur la Figure 4.11a). En observation réelle, le NSB déforme et élargit cette distribution de référence mais la valeur moyenne reste stable pour des taux de NSB proches de 100 MHz (distributions rose et rouge sur la Figure 4.11a). Les piédestaux sont estimés dans les deux voies d'amplification toutes les minutes lors de la prise de données réelles.

4.3.3 Estimation de γ_e^{ADC} et du rapport HG/LG

La mesure du gain γ_e^{ADC} dans la voie haut gain est réalisée lors d'une procédure dédiée. Toutes les deux nuits chaque caméra est illuminée par un système composé d'une diodes électroluminescentes (LED) pulsée à 70 Hz et d'un diffuseur situé à ∼2 m, système qui est réglé de telle manière que chaque pixel puisse recevoir fréquemment un photon unique, ce qui permettra la mesure du gain au photon-électron. Lors de cette étape, les caméras sont dans leurs abris respectifs afin de limiter la contamination due au NSB. Les distributions des charges résultantes par l'illumination de n



(a) (b)

FIGURE 4.11: (a) Piédestaux pour différents niveaux de NSB et (b) exemple de l'ajustement du gain d'un pixel dans la voie haut gain pour un télescope de type HESS-I (Crédits : [Aharonian et al., 2004]).

photon-électrons d'intensité μ sont ajustées selon :

$$G(x) = N \left(\frac{\exp(-\mu)}{\sqrt{2\pi}\sigma_P} \exp\left[-\frac{1}{2} \left(\frac{x-P}{\sigma_P} \right)^2 \right] + \kappa \sum_{n=1}^{m \gg 1} \frac{\exp(-\mu)}{\sqrt{2\pi}\sigma_{\mathrm{T}}} \frac{\mu^n}{n!} \exp\left[-\frac{1}{2} \left(\frac{x-(P+n\gamma_e^{ADC})}{\sigma_{\mathrm{T}}} \right)^2 \right] \right) \tag{4.9}$$

où x représente le nombre de coups d'ADC. Le premier terme est une modélisation du bruit de l'électronique par une distribution gaussienne de largeur σ_P avec une position moyenne de la ligne de base P. La somme des termes correspond au nombre de photon-électrons distribué selon une loi de Poisson, chacun contribuant à une distribution de lumière suivant une gaussienne de largeur $\sigma_{\mathrm{T}} = \sqrt{\sigma_P^2 + n\sigma_{\gamma_e}^2}$ et de moyenne $P + n\gamma_e^{ADC}$. Les deux quantités N et κ sont des constantes de normalisation. Du fait du caractère poissonien du nombre de photon-électrons, il est possible d'ajuster le paramètre κ pour contrôler le bon déroulement de l'ajustement, pour lequel on attend une valeur proche de l'unité. Un exemple d'ajustement pour un pixel est illustré sur la Figure 4.11b. Le premier pic à gauche correspond au piédestal et le reste de la distribution correspond à la somme des contribution de un ou plusieurs photon-électrons.

Afin de pouvoir convertir le nombre de coups d'ADC en photon-électrons dans la voie bas gain, il est nécessaire de déterminer le rapport moyen des gains HG/LG. Dans la zone intermédiaire entre 30 et 150 p.e., la variation des haut gains et des bas gains en fonction du nombre de photon-électrons est linéaire (Figure 4.10). Les signaux mesurés lors d'observations réelles dans ce régime permettent d'obtenir une estimation du rapport des gains HG/LG pour chaque pixel.

4.3.4 Estimation du coefficient FF

Afin de prendre en compte la dispersion des efficacités de collection de lumière entre les différents pixels, il est nécessaire de déterminer pour chacun d'entre eux le facteur de correction FF. Dans ce but, une mesure de l'uniformité de la collection de lumière est effectuée lors d'observations spécifiques au cours desquelles les caméras sont éclairées uniformément à l'aide d'une LED pour les télescopes de taille moyenne et à l'aide d'un laser pour CT_5. Pour chaque pixel, l'écart à la moyenne ainsi mesuré correspond à l'inverse du coefficient FF recherché (équation 4.7).

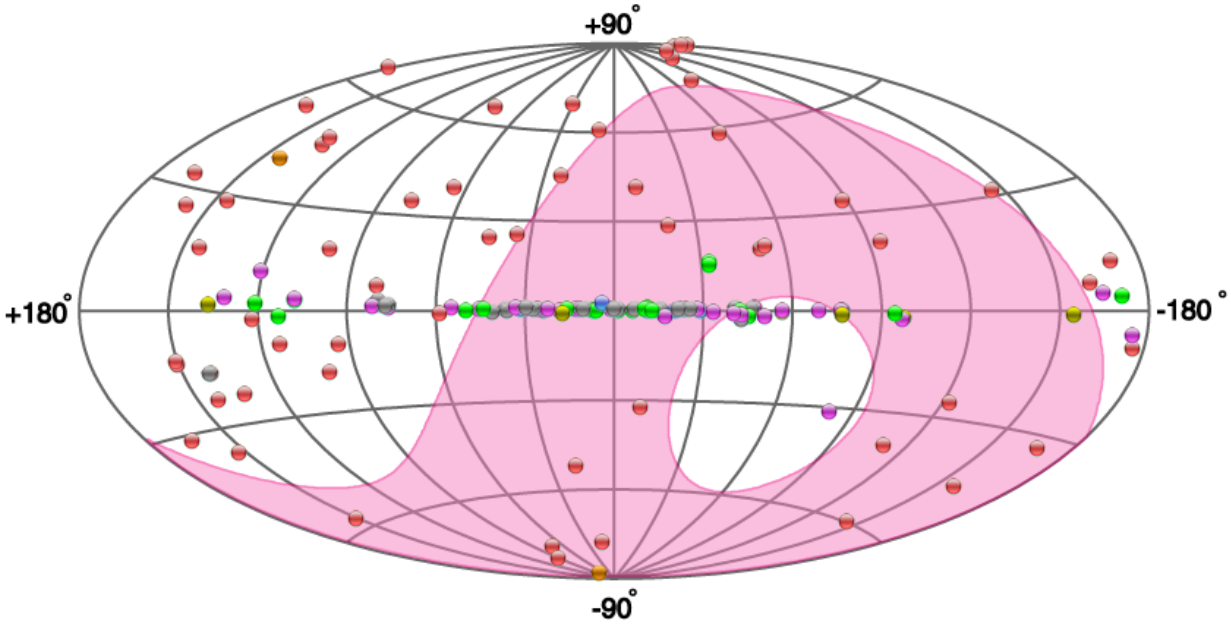


FIGURE 4.12: Région du ciel vue par l'expérience *H.E.S.S.* et sources détectées au TeV (Crédits : http://tevcat.uchicago.edu).

4.3.5 Efficacité optique

L'efficacité optique des télescopes prend en compte la perte de lumière entre la gerbe et le signal mesuré du fait de la réflectivité des miroirs, des espaces et des angles morts, de la transmission des cônes de Winston et enfin de l'efficacité quantique des PM. Cette quantité varie en fonction des télescopes et aussi en fonction du temps (vieillissement, changement de composants ou ré-aluminisation des miroirs). Au niveau de l'analyse (détermination de l'énergie, reconstruction spectrale, etc.), il est nécessaire d'avoir une estimation de l'efficacité optique de chaque télescope, afin d'utiliser les fonctions de réponse de l'instrument (IRF) qui correspondent à l'état de l'instrument (voir section 5.4 chapitre 5).

La méthode utilisée pour estimer l'efficacité optique de chacun des télescopes utilise une comparaison entre les prédictions d'un modèle de muon et des images de muons atmosphériques obtenues lors de prises de données réelles. Ces images sont identifiables facilement du fait de leur forme particulière en anneaux ou arcs de cercle. Cette méthode est détaillée dans [Chalme-Calvet et al., 2014].

4.4 Prises de données

La prise de données s'effectue lors de nuits sans lune, afin de limiter le niveau de bruit de fond du ciel, et par temps clair. Une observation correspond le plus souvent à une période de 28 min, appelée « run », afin de limiter une variation trop grande d'angle zénithal. Dans le cas où une intempérie se présente (foudre lointaine, nuage dans le champ de visée d'un télescope signalé par une baisse du taux de trigger) la prise de donnée est interrompue jusqu'à ce que la situation s'améliore. Afin d'avoir une calibration robuste, des « runs » dédiés à la calibration sont également pris, et ce de manière fréquente. Un « run » d'observation avec uniquement les télescopes de type HESS-I correspond à un volume de données d'environ ∼2 Go. L'ajout de CT_5 peut augmenter ce volume jusqu'à ∼50 Go. Tous les mois, les données sont envoyées par avion de Namibie vers la France et l'Allemagne afin d'être calibrées et analysées.

La partie du ciel accessible à l'expérience *H.E.S.S.* est représentée sur la Figure 4.12. Un comité

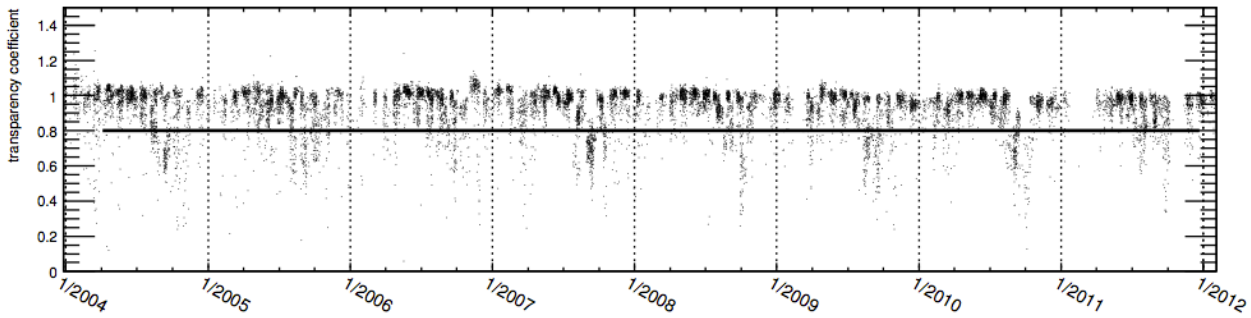


FIGURE 4.13: Évolution du paramètre de transparence de l'atmosphère entre 2004 et 2012 (Crédits : [Hahn et al., 2012]).

est chargé de définir le planning d'observation à partir de propositions justifiées émises par les membres de la collaboration ou occasionnellement par des personnes extérieures. À titre d'exemple, lors de l'année 2015, environ 1000 h d'observations effectives sont prévues dont 1/5 sont réservées aux campagnes d'observations multi-longueur d'onde et simultanées avec d'autres expériences. Les observations sont partagées entre les différents groupes de travail : galactique, extragalactique, astroparticule et calibration. Le temps d'observation demandé par l'ensemble de la collaboration représente ∼4700 h, ce qui équivaut à une pression de l'ordre de cinq.

4.5 Qualité des données

Sélectionner les données selon des critères de qualité est nécessaire afin de limiter les systématiques lors la phase de reconstruction des évènements ou de l'estimation du niveau de flux des sources.

Un premier critère concerne la qualité du fonctionnement des caméras, et en particulier le nombre de « pixels cassés ». En effet, lors d'une acquisition de données un pixel peut être temporairement désactivé lors du passage d'une étoile dans son champ de vue, ou complètement désactivé pour cause de dysfonctionnement. D'autres pixels peuvent être éliminés lors de la phase de calibration. Afin de limiter les systématiques, on exclut généralement de l'analyse les télescopes ayant une fraction de « pixels cassés » supérieure à 15 %.

Un second critère concerne la qualité de l'atmosphère. Les nuages et les aérosols absorbent les photons Tcherenkov, et entraînent des taux de déclenchement plus bas que la moyenne, ainsi qu'une variabilité temporelle anormale. Il est donc nécessaire d'avoir un moyen qui permette d'exclure de l'analyse ces « runs » problématiques.

Une méthode permettant d'identifier ces « runs » est présentée dans [Hahn et al., 2012]. Elle utilise une quantité permettant de s'affranchir des systématiques liées à l'instrument afin de juger de la qualité de l'atmosphère. En supposant que le taux de trigger R d'un télescope (corrigé selon l'angle zénithal de l'observation) est directement proportionnel, à une déviation près Δ, au spectre intégré des rayons cosmiques, on peut écrire : $R \propto E_0^{-1,7+\Delta}$ où E_0 est l'énergie seuil de déclenchement du télescope. D'autre part, au premier ordre, l'énergie seuil est supposée être inversement proportionnelle à l'efficacité de reconstruction des muons μ, au gain moyen des pixels g et à un coefficient caractérisant la transparence de l'atmosphère η : $E_0 \propto (\eta\mu g)^{-1}$. On peut donc définir la quantité t pour chaque télescope, proportionnelle à η, comme : $\eta \propto \frac{^{1,7-\Delta}\sqrt{R}}{\mu g} \equiv t$. Le critère de qualité pour chaque observation est défini selon le nombre de télescopes N participant à la prise de données : $T = k_N \sum_i^N t_i$ où k_N normalise la distribution des quantités T à l'unité. La Figure 4.13 représente le paramètre T pour l'ensemble des données prises entre janvier 2004 et décembre 2012 ainsi que la coupure de qualité symbolisée par une ligne noire. Cette méthode a été utilisée pour sélectionner les données pour les analyses n'incluant pas CT_5 qui ont été effectuées durant cette

thèse (voir par exemple section 5.5 au chapitre 5).

La méthode de [Hahn et al., 2012] est actuellement en cours d'adaptation pour les « runs » incluant CT_5. Pour les résultats présentés au chapitre 7, qui incluent le cinquième télescope, une sélection de qualité plus rudimentaire sera utilisée. Elles consiste à comparer le taux de déclenchement des différents télescopes, corrigés de leur dépendance en angle zénithal, à une valeur de référence préalablement déterminée. Cette méthode permet d'exclure les « runs » présentant des taux de déclenchement inférieurs à un certain seuil. Elle permet aussi dans certains cas d'identifier des dysfonctionnements pour un télescope en particulier, qui se traduisent par un taux de déclenchement inférieur à celui des autres. Le télescope est alors exclu de l'analyse sans pour autant que l'on perde le « run ». On sélectionnera aussi les télescopes qui possèdent moins de 10 % de « pixels cassés ».

Bibliographie

F. Aharonian, et al. Calibration of cameras of the H.E.S.S. detector. *Astroparticle Physics*, 22: 109–125, November 2004. doi: 10.1016/j.astropartphys.2004.06.006.

F. Aharonian, et al. High energy astrophysics with ground-based gamma ray detectors. *Reports on Progress in Physics*, 71(9):096901, September 2008. doi: 10.1088/0034-4885/71/9/096901.

E. Aliu, et al. Improving the performance of the single-dish Cherenkov telescope MAGIC through the use of signal timing. *Astroparticle Physics*, 30:293–305, January 2009. doi: 10.1016/j.astropartphys.2008.10.003.

J. Beringer, et al. Review of particle physics. *Phys. Rev. D*, 86:010001, Jul 2012. doi: 10.1103/PhysRevD.86.010001. URL http://link.aps.org/doi/10.1103/PhysRevD.86.010001.

K. Bernlöhr. Impact of atmospheric parameters on the atmospheric Cherenkov technique*. *Astroparticle Physics*, 12:255–268, January 2000. doi: 10.1016/S0927-6505(99)00093-6.

K. Bernlöhr, et al. The optical system of the H.E.S.S. imaging atmospheric Cherenkov telescopes. Part I: layout and components of the system. *Astroparticle Physics*, 20:111–128, November 2003. doi: 10.1016/S0927-6505(03)00171-3.

J. Bolmont, et al. The camera of the fifth H.E.S.S. telescope. Part I: System description. *Nuclear Instruments and Methods in Physics Research A*, 761:46–57, October 2014. doi: 10.1016/j.nima.2014.05.093.

R. Chalme-Calvet, et al. Muon efficiency of the H.E.S.S. telescope. *ArXiv e-prints*, March 2014.

S. Funk, et al. The trigger system of the H.E.S.S. telescope array. *Astroparticle Physics*, 22:285–296, November 2004. doi: 10.1016/j.astropartphys.2004.08.001.

J. Guy. *Premiers résultats de l'expérience HESS et étude du potentiel de détection de matière noire supersymétrique*. PhD thesis, UNIVERSITÉ PARIS 6, 2003.

J. Hahn, et al. Ensuring long-term stability of data quality selection for H.E.S.S. under challenging atmospheric conditions. In F. A. Aharonian, W. Hofmann, et F. M. Rieger, editors, *American Institute of Physics Conference Series*, volume 1505 of *American Institute of Physics Conference Series*, pages 721–724, December 2012. doi: 10.1063/1.4772361.

W. Heitler. *Quantum theory of radiation*. 1954.

John D Jackson. *Classical Electrodynamics*. Wiley, third edition, 1998.

D.A. Lewis. Optical characteristics of the whipple observatory tev gamma-ray imaging telescope. *Experimental Astronomy*, 1(4):213–226, 1990. ISSN 0922-6435. doi: 10.1007/BF00428516. URL http://dx.doi.org/10.1007/BF00428516.

M. S. Longair. *High Energy Astrophysics*. February 2011.

J. Matthews. A Heitler model of extensive air showers. *Astroparticle Physics*, 22:387–397, January 2005. doi: 10.1016/j.astropartphys.2004.09.003.

P. Vincent, et al. Performance of the H.E.S.S. Cameras. *International Cosmic Ray Conference*, 5: 2887, July 2003.

Chapitre 5

Méthodes d'analyse dans *H.E.S.S.*

Sommaire

5.1 Analyses de bas niveau dans la collaboration *H.E.S.S.* 71
5.2 L'analyse Paris-MVA 83
5.3 Extraction du signal et estimation du fond 89
5.4 Reconstruction spectrale et courbes de lumière 91
5.5 START 94
Bibliographie 101

Dans ce chapitre nous commençons par présenter brièvement (section 5.1) les différentes analyses de bas niveau utilisées dans la collaboration *H.E.S.S.*. La présentation de chaque analyse est constituée d'une première partie consacrée à la reconstruction des propriétés des gerbes, et d'une seconde partie correspondant à la discrimination γ/fond. Une attention particulière sera dédiée à l'analyse en mode stéréo Paris-MVA, présentée à la section 5.2 [1]. Nous présenterons ensuite les différentes méthodes utilisées pour soustraire le fond résiduel (section 5.3), et les analyses de haut niveau, qui concernent la reconstruction de la forme spectrale, des courbes de lumière et des cartes du ciel (section 5.4). Enfin, nous présenterons le logiciel START, développé au cours de cette thèse pour les analyses spectrales.

5.1 Analyses de bas niveau utilisées par la collaboration *H.E.S.S.*

Après la calibration, on dispose pour chaque évènement et pour chaque télescope de la charge mesurée en photoélectrons de chaque pixel. Avant de commencer la reconstruction des évènements à proprement parler, un nettoyage des images brutes est réalisé afin de minimiser la contamination par le bruit de fond de ciel. Le nettoyage le plus utilisé consiste à appliquer un double seuil : garder un pixel s'il a une charge supérieure ou égale q_1 et si au moins l'un de ses voisins à une charge supérieure ou égale à q_2. On parle par exemple de « nettoyage 510 » pour les seuils $q_1 = 5$ et $q_2 = 10$.

5.1.1 La méthode de Hillas

Cette méthode, qui est basée sur les paramètres de Hillas, est très utilisée pour sa robustesse et sa simplicité. Les performances qu'il est possible d'atteindre avec cette méthode sont néanmoins limitées par rapport aux autres analyses présentées ci-après. Elle est adaptée pour l'analyse de sources ayant un niveau de flux supérieur à $\sim$10 % de la nébuleuse du Crabe. Depuis le début de l'expérience *H.E.S.S.* de nombreux résultats ont été publiés avec cette méthode ([Aharonian et al., 2006a], etc.).

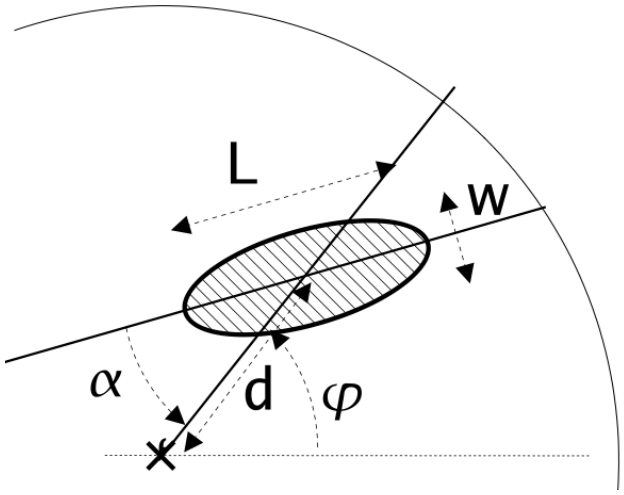


FIGURE 5.1: Paramètres de Hillas (Crédits : [de Naurois, 2006]).

5.1.1.1 Reconstruction des propriétés des gerbes

Introduite par [MacKeown et al., 1983] et généralement appelée la paramétrisation de Hillas [Hillas, 1985], la méthode des ellipses consiste à paramétrer les images de la gerbe enregistrées par les caméras par des gaussiennes bidimensionnelles (des ellipses). Les paramètres, illustrés sur la Figure 5.1, sont :

— la largeur w de l'ellipse

— la longueur l de l'ellipse

— la charge totale Q (en p.e.) contenue dans l'ellipse

— l'angle φ entre l'axe reliant le centre de la caméra et le barycentre des charges et un axe de référence de la caméra

— la distance nominale d entre le barycentre des charges et le centre de la caméra

— l'angle α entre l'axe principal de l'ellipse et l'axe reliant le barycentre de l'ellipse au centre de la caméra

À l'aide de la charge q_i de chaque pixel i à la position (x_i, y_i) il est possible d'exprimer les paramètres de Hillas en introduisant les premiers et seconds moments des charges [de Naurois, 2012] :

$$\langle x \rangle = \frac{\sum_i q_i x_i}{\sum_i q_i}, \ \langle y \rangle = \frac{\sum_i q_i y_i}{\sum_i q_i} \tag{5.1}$$

$$\langle x^2 \rangle = \frac{\sum_i q_i x_i^2}{\sum_i q_i}, \ \langle y^2 \rangle = \frac{\sum_i q_i y_i^2}{\sum_i q_i}, \ \langle xy \rangle = \frac{\sum_i q_i x_i y_i}{\sum_i q_i} \tag{5.2}$$

ainsi que les variances associées :

$$\sigma_{x^2} = \langle x^2 \rangle - \langle x \rangle^2, \ \sigma_{y^2} = \langle y^2 \rangle - \langle y \rangle^2, \ \sigma_{xy} = \langle xy \rangle - \langle x \rangle \langle y \rangle \tag{5.3}$$

1. Ce chapitre se focalise exclusivement sur les analyses développées pour la phase I de *H.E.S.S.*. Les développements pour l'exploitation du réseau avec le nouveau grand télescope CT_5 seront présentés dans les deux chapitres suivants.

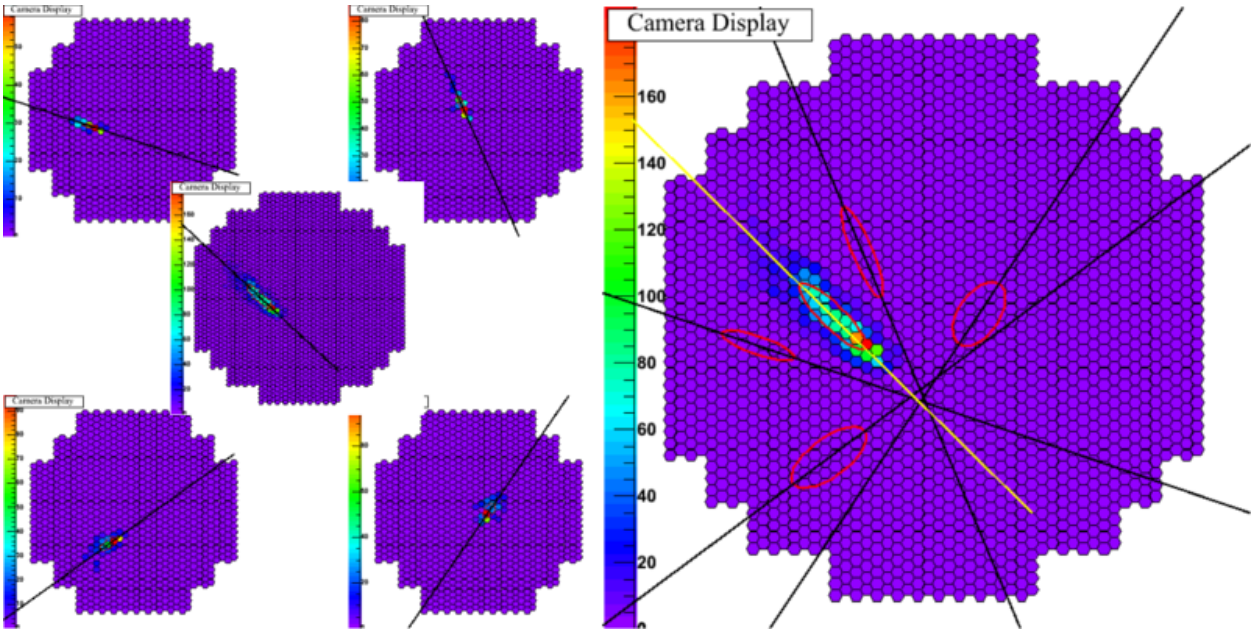


FIGURE 5.2: Principe de la stéréoscopie. À gauche, intensités et paramétrisation de Hillas d'une gerbe de γ simulée, vue dans chacune des 5 caméras. À droite, projection de l'ensemble des paramétrisations de Hillas dans CT5. L'intensité vue par la cinquième caméra y est représentée.

Les paramètres de Hillas s'expriment alors simplement selon:

$$\delta = \sqrt{\langle x^2 \rangle + \langle y^2 \rangle} \tag{5.4}$$

$$l = \sigma_{x^2} + \sigma_{y^2} + z \tag{5.5}$$

$$w = \sigma_{x^2} + \sigma_{y^2} - z \tag{5.6}$$

$$\alpha = \arcsin \frac{b}{\delta} \tag{5.7}$$

$$\tag{5.8}$$

où z et b sont définis par :

$$z = \sqrt{\chi^2 + 4\sigma_{xy}} \tag{5.9}$$

$$b = \sqrt{\frac{(1+\chi/z)\langle x \rangle^2 + (1-\chi/z)\langle y \rangle^2 - 2\sigma_{xy}\langle x \rangle \langle y \rangle}{2}} \tag{5.10}$$

$$\tag{5.11}$$

et où $\chi = \sigma_{x^2} - \sigma_{y^2}$.

Le grand axe de l'ellipse pointe vers la direction de la source, mais celle-ci est dégénérée, la source pouvant se trouver sur n'importe quel point de l'axe principal. La technique de stéréoscopie qui utilise plusieurs images d'un même évènement permet de lever cette dégénérescence (Figure 5.2). Dans la collaboration *H.E.S.S.*, la méthode de reconstruction est basée sur la détermination du barycentre des intersections entre l'axe principal des différentes images. À chaque intersection (donc à chaque paire de télescopes) un poids $\omega^{i,j}$ est attribué tel que $\omega^{i,j} = \omega_e^{i,j} \omega_Q^{i,j} \omega_\theta^{i,j}$.

Le terme $\omega_e^{i,j}$ représente une pondération sur l'ellipticité $e = l/w$ des paires d'images :

$$\omega_e^{i,j} = \frac{1}{\frac{1}{e_i} + \frac{1}{e_j}} \tag{5.12}$$

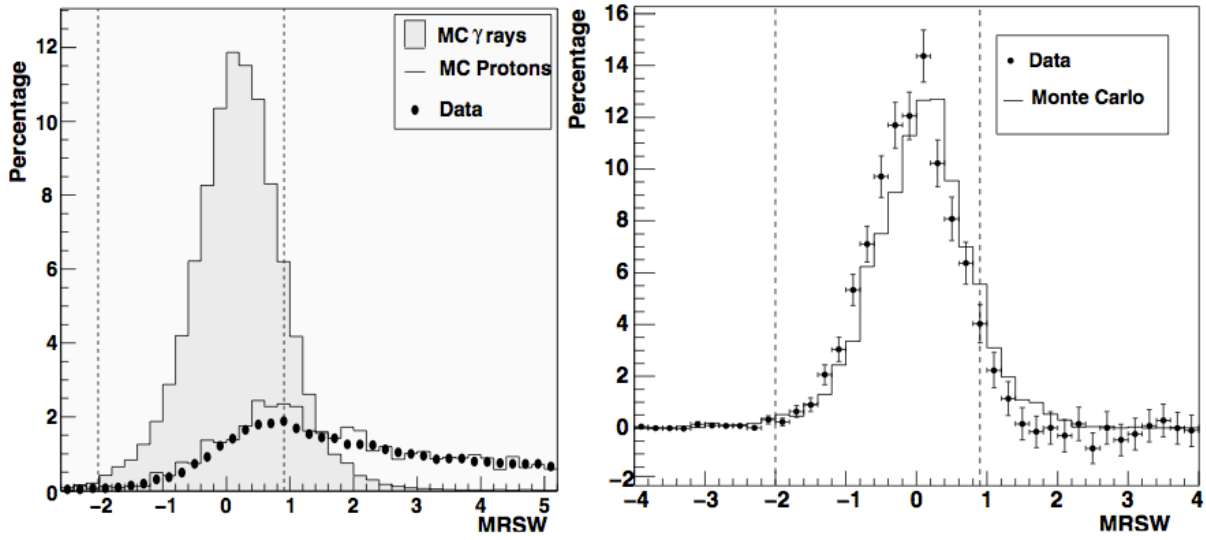


FIGURE 5.3: Distributions du paramètre MSCW (Crédits : [Aharonian et al., 2006a]). À gauche, des γ simulés (distribution remplie), des hadrons simulés (traits pleins) et des évènements obtenus lors d'observations réelles hors-champ. À droite, des γ simulés (traits pleins) et des évènements provenant de la nébuleuse du Crabe après coupures de pré-sélection. Les coupures optimales sur le paramètre MSCW sont indiquées par des lignes verticales.

Celui-ci minimise l'influence des images à faible ellipticité, pour lesquelles la détermination des axes principaux est plus difficile.

Une pondération est associée à la charge vue dans chaque télescope :

$$\omega_Q^{i,j} = \frac{1}{\frac{1}{Q_i} + \frac{1}{Q_j}} \tag{5.13}$$

Ce poids permet de donner un plus grande importance aux paires contenant des images avec une grande charge pour lesquelles il est généralement plus simple de bien reconstruire les axes principaux.

Enfin, le dernier poids ω_θ représente une pondération sur l'angle entre les deux axes principaux :

$$\omega_\theta^{i,j} = \sin\theta_{i,j} \tag{5.14}$$

Il privilégie les paires de télescopes formant un grand angle ce qui réduit l'erreur sur la localisation de l'intersection.

Une fois la direction connue, le pied de gerbe R et la hauteur du maximum de gerbe $H_{\max}$ sont reconstruits géométriquement de manière similaire. On pourra par exemple se référer à la méthode I présentée dans [Hofmann, 2000].

5.1.1.2 Discrimination

Une méthode de discrimination γ/fond largement répandue est basée sur le fait que, pour un évènement d'énergie donnée, la charge mesurée dans un télescope, les longueurs et les largeurs des images, dépendent du pied de gerbe mesuré par rapport au télescope. Des tables sont construites avec des γ simulés pour tirer profit du fait que les gerbes γ ont des largeurs et des longueurs plus resserrées que pour les hadrons. Ces tables contiennent les longueurs $\langle l \rangle_{R_i,Q_i}$ et les largeurs $\langle w \rangle_{R_i,Q_i}$ moyennes en fonction des pieds de gerbe reconstruits R_i par rapport au télescope i et des charges mesurées Q_i. Afin de tirer profit de la stéréoscopie, on introduit les paramètres de longueurs et de largeurs réduites MSCL et MSCW (en anglais *Mean SCaled Length* et *Mean SCaled Width*) qui

sont définis par :

$$\mathrm{MSCL} = \frac{1}{N_{\mathrm{Tels}}} \sum_{i}^{N_{\mathrm{Tels}}} \frac{l_i - \langle l \rangle_{R_i,Q_i}}{\sigma^{l}_{R_i,Q_i}} \tag{5.15}$$

$$\mathrm{MSCW} = \frac{1}{N_{\mathrm{Tels}}} \sum_{i}^{N_{\mathrm{Tels}}} \frac{w_i - \langle w \rangle_{R_i,Q_i}}{\sigma^{w}_{R_i,Q_i}} \tag{5.16}$$

où les quantités $\sigma^{l}_{R_i,Q_i}$ et $\sigma^{w}_{R_i,Q_i}$ sont les écarts-types des distributions des longueurs et des largeurs. Les paramètres MSCL et MSCW sont définis de manière à ce que leur distributions soient centrées sur zéro pour des évènements γ. Dans le cas des hadrons, des déviations par rapport aux valeurs moyennes des largeurs et des longueurs sont attendues. Cette technique a été largement utilisée pour les publications de la collaboration *H.E.S.S.* pendant les premières années de prises de données. À titre d'exemple, on peut observer sur la Figure 5.3 la distribution du paramètre MSCW pour différents lots d'évènements, extraite de la première publication sur la détection de la nébuleuse du Crabe et des méthodes implémentées dans *H.E.S.S.* [Aharonian et al., 2006a]. À gauche sont représentées les distributions pour des γ simulés (distribution remplie), des hadrons simulés (traits pleins) et des évènements obtenus lors d'observations réelles hors-champ. À droite, on trouve les mêmes γ simulés (traits pleins) et la distribution des évènements provenant de la nébuleuse du Crabe après coupures de pré-sélection. Les coupures optimales sur les paramètres MSCL et MSCW (représentées par des lignes verticales à droite de la Figure 5.3 pour le paramètre MSCW) ont été déterminées en maximisant le rapport signal sur bruit en utilisant des γ simulés selon une loi de puissance d'indice $\Gamma = 2{,}69$ et du fond réel.

5.1.2 Modèle 3D

La reconstruction par un modèle de la photosphère en trois dimensions, dite modèle 3D, a été proposée par [Lemoine-Goumard et al., 2006] afin de mieux exploiter les symétries des gerbes EM et les corrélations des images mesurées entre les différents télescopes. Elle est basée sur un modèle physique du développement de la gerbe EM qui suppose l'existence de plusieurs symétries.

Cette méthode permet d'obtenir une précision angulaire qui au zénith varie selon l'énergie entre 0,04° et 0,1°. Sa résolution en énergie, qui varie entre 15 et 20 %, est similaire à celle obtenue avec une approche standard. Enfin, elle permet d'obtenir une meilleure sensibilité pour les multiplicités supérieures ou égales à trois. Cette méthode a été utilisée dans plusieurs publications ([Aharonian et al., 2006b], etc.)

5.1.2.1 Reconstruction des propriétés des gerbes

Cette approche consiste à modéliser la photosphère Tcherenkov d'une gerbe EM par une gaussienne tridimensionnelle. Il est supposé que la distribution angulaire des photons Tcherenkov est indépendante de l'énergie de la particule incidente ainsi que de l'altitude du point d'émission.

Le développement longitudinal de la gerbe est modélisé par une gaussienne unidimensionnelle de largeur σ_{L}. Cette approximation est totalement en accord avec les simulations pour une altitude inférieure à $\sim$12 km. Au-delà de cette limite, la modélisation sous-estime le nombre de photons Tcherenkov, mais ceux-ci ne représentent que quelques pourcents du nombre total de photons émis. Le développement latéral de la gerbe est modélisé par une gaussienne dimensionnelle de largeur unique σ_{T}. Une symétrie azimutale est donc supposée dans son développement.

En notant ε l'angle entre la direction d'un photon Tcherenkov et l'axe de la gerbe EM et η l'angle maximum pour l'émission Tcherenkov (vérifiant $\cos\eta = 1/n(z)$, voir chapitre 4) la probabilité d'émission d'un photons Tcherenkov est modélisée par :

$$I(\varepsilon) = \begin{cases} K & \text{si } \varepsilon < \eta \\ K\frac{\eta}{\varepsilon} \exp\left(-\frac{\varepsilon-\eta}{4\eta}\right) & \text{si } \varepsilon \geq \eta \end{cases} \tag{5.17}$$

où K est une constante de normalisation $K = (9\pi\eta^2)^{-1}$. [Lemoine-Goumard et al., 2006] ont montré qu'une bonne approximation pouvait être réalisée en prenant une valeur pour η qui ne dépend que de l'angle zénithal θ : $\eta = 15\,\text{mrad} \times \cos\theta$. La valeur choisie correspond à un angle d'ouverture Tcherenkov à l'altitude du maximum d'une gerbe initiée par un γ de 1 TeV.

Dans ce modèle, les différents paramètres caractérisant la gerbe sont au nombre de huit :

— deux angles θ_0 et ϕ_0 définissant la direction de l'axe de la gerbe dans le référentiel du système de télescope

— les coordonnées du paramètre d'impact (x_0, y_0) dans le référentiel du système de télescopes

— la position de la hauteur du maximum de gerbe $H_{\max}$

— les deux largeurs σ_{L} et σ_{T}

— le nombre total de photons Tcherenkov N_{c} composant la photosphère Tcherenkov

L'obtention des paramètres physiques de la gerbe s'effectue selon une procédure de recherche du maximum d'une fonction de vraisemblance, en comparant les charges mesurées et les charges prédites par le modèle 3D dans chaque pixel. La charge prédite dans un pixel est donnée par l'intégrale :

$$q_{\text{th}} = S_{\text{Tel}} \Delta w_{\text{pixel}} I(\varepsilon) \cos\theta \int_0^\infty n_{\text{c}}(r) dr \tag{5.18}$$

où $n_{\text{c}}(r)$ représente la densité de photons Tcherenkov contenue dans la gerbe. L'intégration se fait sur la distance r entre le point d'émission d'un photon et le télescope de surface S_{Tel}, pour lequel un pixel couvre un angle solide Δw_{pixel} et est orienté selon une direction (θ, ϕ). On pourra trouver le détail du calcul du nombre de photons arrivant sur un pixel, en fonction des paramètres de la gerbe, dans l'annexe A de [Lemoine-Goumard et al., 2006]. La probabilité de mesurer dans un pixel i une charge comprise entre q et $q + \mathrm{d}q$, sachant que l'on s'attend à avoir une charge q_{th} s'exprime selon :

$$P_i = \sum_{n=0}^{\infty} \frac{q_{\text{th}}^n e^{-q_{\text{th}}}}{n!} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(s-n)^2}{2\sigma^2}\right) \mathrm{d}q \tag{5.19}$$

Dans cette formule, on prend en compte la distribution poissonienne sur le nombre de photoélectrons et la résolution des PM ($\sigma = 0{,}4$). Une fonction de vraisemblance est construite afin d'obtenir les paramètres physique de la gerbe : $\mathcal{L} = \prod_i P_i$. L'ajustement est initialisé à partir de paramètres obtenus avec une reconstruction de type Hillas (section 5.1.1).

5.1.2.2 Discrimination

Le modèle 3D étant un modèle physique de gerbe EM, un premier niveau de discrimination consiste à exclure tous les évènements qui ne convergent pas vers un optimum de la fonction de vraisemblance. Sur des simulations γ, le taux de convergence est de l'ordre de ∼95 %, et sur des évènements de fond obtenus lors d'observations réelles, celui-ci est de ∼30 %. Après ce premier niveau de discrimination, deux variables sont utilisées afin de réduire le fond résiduel.

La première variable correspond à la hauteur du maximum de la gerbe D_s reconstruite par le modèle 3D, exprimée en termes d'épaisseur d'atmosphère (en g cm^{-2}). À gauche de la Figure 5.4 le paramètre D_s est représenté en fonction du nombre de nombre de photons Tcherenkov constituant la photosphère des gerbes, lui aussi reconstruit par le modèle 3D, pour des γ simulés à quatre énergie fixes (de 200 GeV à 5 TeV) et pour un angle zénithal de 0°. À droite de cette figure, on peut trouver ce même paramètre pour des hadrons obtenus lors d'observations réelles à 16° d'angle zénithal. Dans le cas des hadrons, à l'inverse des γ, la distribution ne présente pas de corrélation en fonction du nombre de photons Tcherenkov de la photosphère. Une région « physique » (en pointillé sur les deux figures) est définie qui permet de doubler le rejet des hadrons au détriment d'une perte de ∼10 % des évènements γ :

$$61{,}5(\ln N_c - 13{,}5 + T(\theta)) \leq D_s \leq 61{,}5(\ln N_c - 10{,}0 + T(\theta)) + 300 \tag{5.20}$$

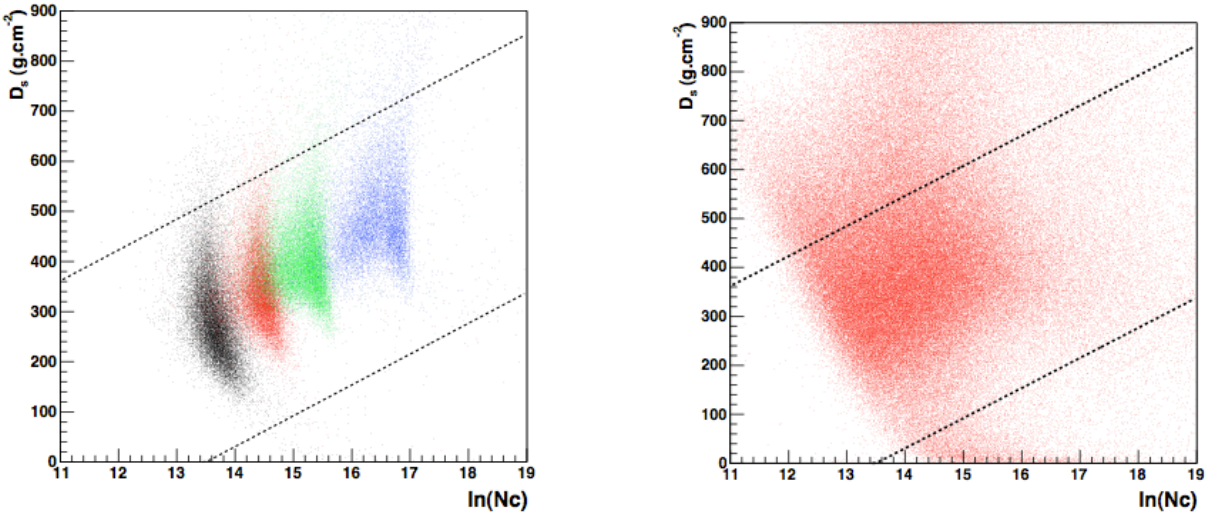


FIGURE 5.4: Hauteur du maximum de gerbe exprimée en épaisseur d'atmosphère ajustée par le modèle 3D en fonction du nombre de photons Tcherenkov constituant la photosphère pour (à gauche) des gerbes γ simulées au zénith pour quatre énergies fixes (de gauche à droite : 0,2, 0,5, 1 et 5 TeV) et (à droite) du fond hadronique obtenu lors d'observations réelles (Crédits : [Lemoine-Goumard, 2006]).

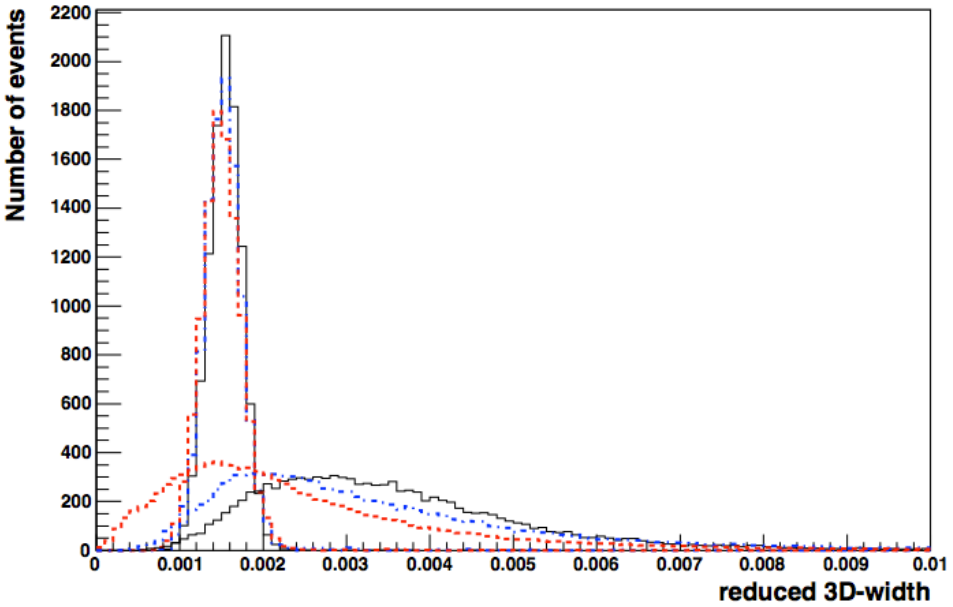


FIGURE 5.5: Distributions de la largeur réduite ω pour des γ simulés à 1 TeV et à 18° d'angle zénithal (pics centrés à $1{,}5 \times 10^{-3}$) et pour du fond hadronique obtenus lors d'observation réelles (distributions plus étalées). Les couleurs rouge, bleu et noire représentent respectivement des évènements ayant des multiplicités 2, 3 et 4 (Crédits : [Lemoine-Goumard, 2006]).

où $T(\theta) = 3{,}28(1 - \cos\theta)$ représente une correction qui dépend de l'angle zénithal d'observation.

La seconde variable de discrimination, la largeur réduite ω, est une quantité décrivant le développement latéral de la gerbe construite à partir de la largeur transverse ajustée par le modèle 3D. Elle est reliée à un paramètre physique (la largeur de la gerbe) et est en ce sens plus robuste que des variables liées à la qualité d'un ajustement (par exemple le « *goodness of fit* »). Elle est normalisée par la hauteur du maximum de gerbe pour s'affranchir des variations d'angle zénithal et de l'énergie des γ :

$$\omega = \frac{\sigma_{\mathrm{T}}\rho(z_{\max})}{D_s} \tag{5.21}$$

où $\rho(z_{\max})$ correspond à la masse volumique de l'air à l'altitude du maximum de gerbe. Sur la Figure 5.5 on peut observer les distributions normalisées de ω pour des γ simulés à 1 TeV pour un angle zénithal de 18° qui piquent autour de la valeur $1{,}5 \times 10^{-3}$. Les hadrons ont quant à eux des distributions plus étalées. Les couleurs rouge, bleu et noire représentent respectivement des évènements ayant des multiplicités 2, 3 et 4. On observe un pouvoir discriminant important qui varie selon la multiplicité des évènements.

5.1.3 Modèle ++

Le modèle ++ [de Naurois et Rolland, 2009] est une technique de reconstruction inspirée de la méthode développé dans l'expérience *CAT* [Le Bohec et al., 1998]. Elle consiste à déterminer les paramètres des gerbes incidentes à l'aide d'une procédure de recherche du maximum d'une fonction de vraisemblance en comparant des images réelles et des images issues d'un modèle semi-analytique de gerbe EM. Cette approche a été fortement optimisée pour *H.E.S.S.* et constitue aujourd'hui l'une des analyses les plus sensibles de la collaboration. Du fait de l'utilisation d'un paramètre basé sur la valeur du maximum de la fonction de vraisemblance, appelée « *Shower Goodness* », (section 5.1.3.2) et de sa sensibilité au niveau de bruit du fond de ciel, elle nécessite une très grande maîtrise de l'instrument.

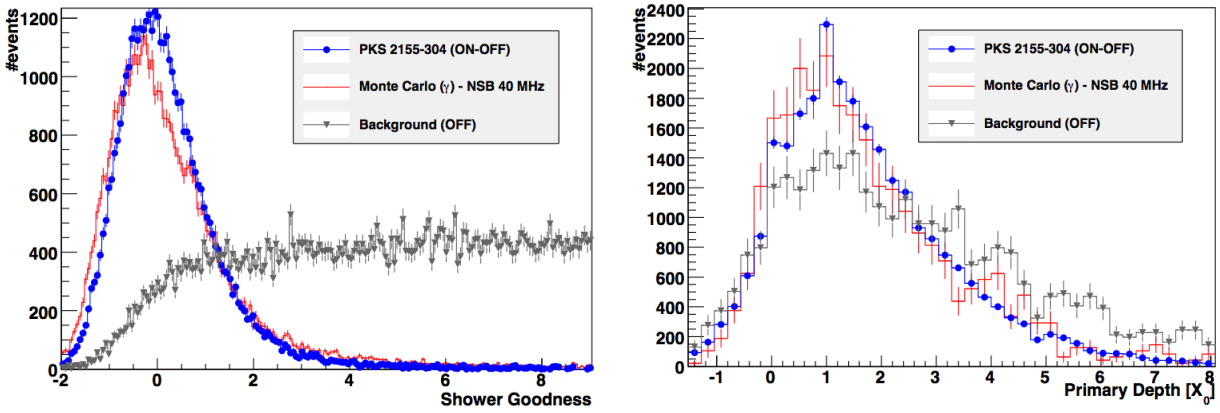
Cette approche permet d'atteindre une précision angulaire entre 0,1° et 0,04° en fonction de l'énergie et de l'angle zénithal. La résolution en énergie est en deçà de 15 % à toute les énergies et meilleure que 10 % à partir de 500 GeV. Enfin, cette méthode permet de gagner un facteur deux en sensibilité par rapport à une analyse standard. De nombreux résultats ont été publiés avec cette analyse ces dernières années ([Abramowski et al., 2012b], etc.).

5.1.3.1 Reconstruction des propriétés des gerbes

Le modèle ++ est un modèle semi-analytique de gerbe EM basé sur la paramétrisation dans le plan focal de la distribution des photons Tcherenkov émis par celle-ci. En comparant, à l'aide d'une procédure de maximisation de vraisemblance, les intensités observées et attendues dans chaque caméra, la méthode permet de reconstruire la direction, le paramètre d'impact, la hauteur de première interaction ainsi que l'énergie de la particule incidente. Des abaques contenant des images de gerbes dans le plan d'une caméra sont générées à l'aide d'un modèle semi-analytique, pour différents angles zénithaux, paramètres d'impacts, énergies et hauteurs de la première interaction. Celles-ci servent à la comparaison avec les données réelles. Aucun nettoyage des images n'est réalisé.

La densité de probabilité qui correspond à l'observation dans un pixel d'un signal de s photoélectrons pour une valeur attendue μ est obtenue en convoluant une distribution de Poisson représentant n photoélectrons avec la résolution d'un PM. Cette dernière est modélisée selon une distribution gaussienne de largeur $\sqrt{\sigma_p^2 + n\sigma_\gamma^2}$ où σ_p est la largeur du piédestal (section 4.3 du chapitre 4) et σ_γ la résolution moyenne d'un PM. La densité de probabilité s'écrit :

$$P(s|\mu, \sigma_p, \sigma_\gamma) = \sum_n \frac{\mu^n e^{-\mu}}{n!\sqrt{2\pi(\sigma_p^2 + n\sigma_\gamma^2)}} \exp\left(-\frac{(s-n)^2}{2(\sigma_p^2 + n\sigma_\gamma^2)}\right) \tag{5.22}$$



(a) (b)

FIGURE 5.6: Distributions des paramètres de (a) *Shower Goodness* et (b) de la hauteur de première interaction (Crédits : [de Naurois et Rolland, 2009]).

Une fonction de vraisemblance est construite pour chaque télescope permettant d'obtenir les paramètres de la gerbe:

$$\mathcal{L} = \prod_{\text{pixel } i} \mathcal{L}_i = \prod_{\text{pixel } i} P(s_i|\mu_i, \sigma_p, \sigma_\gamma) \tag{5.23}$$

5.1.3.2 Discrimination

Cette analyse utilise principalement deux paramètres pour la discrimination γ/fond. Le premier, le *Shower Goodness* (SG), est défini comme étant la somme normalisée sur les pixels composant l'image principale, des différences entre les vraisemblances d'un pixel avec les paramètres mesurés et les paramètres prédits :

$$\text{SG} = \frac{\sum_{\text{pixel } i} \ln \mathcal{L}(s_i, \mu_i) - \langle \ln \mathcal{L} \rangle|_{\mu_i}}{\sqrt{2 \times \text{NdF}}} \tag{5.24}$$

où NdF est le nombre de degrés de liberté (nombre de pixels - 6). Ce paramètre est très sensible aux différents niveaux de NSB et des tables dédiées sont utilisées pour corriger les valeurs moyennes et les largeurs des piédestaux dans chaque pixel. On peut observer sur la Figure 5.6a la distribution de ce paramètre pour des γ simulés dans des conditions correspondant aux observations de la source PKS 2155-304 (en rouge), pour le signal issu de ces mêmes observations (en bleu), et enfin pour des évènements hors-source (en gris).

Le second paramètre utilisé pour la discrimination est la hauteur de première interaction reconstruite par le modèle ++. Sa distribution est représentée sur la Figure 5.6b avec le même code couleur que celui utilisé pour la Figure 5.6a.

5.1.4 X_{eff}

[Dubois et al., 2009] ont proposé une approche basée sur une combinaison des résultats de trois méthodes, la méthode de Hillas, le modèle 3D et l'analyse « model » [de Naurois et al., 2003, de Naurois, 2006]. Cette dernière provient de l'expérience *CAT* et a été adaptée afin d'être utilisée par la collaboration *H.E.S.S.* [2].

Par comparaison avec l'analyse modèle ++, cette méthode permet d'obtenir une meilleure précision angulaire et une meilleure résolution en énergie au-delà de ∼2 TeV. Elle a entre autres été

2. Les principales différences avec le modèle ++ concernent la phase de discrimination qui utilisait aussi des paramètres discriminants construits à partir des paramètres de Hillas ainsi que l'utilisation de tous pixels pour le calcul du paramètre SG.

appliquée afin de caractériser l'émission non-thermique dans la région de Vela X [Abramowski et al., 2012a] et plus récemment l'émission au TeV résultante de l'interaction entre la supernova 349.7+0.2 et un nuage moléculaire [H. E. S. S. Collaboration et al., 2014b]. L'analyse X_{eff} a aussi été utilisée afin de rechercher la signature de particules de matière noire dans la galaxie naine sphéroïdale du Sagittaire [H. E. S. S. Collaboration et al., 2014a].

5.1.4.1 Reconstruction des évènements γ

L'estimation de la direction incidente ou de l'énergie d'un évènement, que l'on notera de manière indifférenciée $\hat{x}$, est déterminée à partir d'une combinaison linéaire des valeurs fournies par les trois reconstructions :

$$\hat{x} = \sum_{i=1}^{3} w_i x_i \tag{5.25}$$

où les poids w_i vérifient :

$$\sum_{i=1}^{3} w_i = 1 \tag{5.26}$$

Afin d'obtenir un estimateur du paramètre $\hat{x}$ qui présente un biais de reconstruction minimal, les poids sont obtenus en minimisant sa variance :

$$\sigma^2(\hat{x}) = \sum_{i=1}^{3}\sum_{j=1}^{3} E_{ij} w_i w_j \tag{5.27}$$

où les éléments de matrice E_{ij} sont estimés à partir des simulations de γ.

5.1.4.2 Discrimination

La méthode de discrimination γ/fond est basée sur l'utilisation des fonctions de densité de probabilité associées à quatre variables discriminantes :

— les longueurs et largeurs réduites

— le *Shower Goodness* de l'analyse « model »

— la largeur réduite du modèle 3D

La méthode utilise le paramètre de discrimination final $X_{\text{eff}}(\{\chi_i\})$, qui peut être vu comme la probabilité qu'un évènement soit du fond, définie comme :

$$X_{\text{eff}}(\{\chi_i\}) = \frac{\eta H(\{\chi_i\})}{(1-\eta)G(\{\chi_i\}) + \eta H(\{\chi_i\})} \tag{5.28}$$

où le paramètre η représente la fraction d'évènements de fond (après coupures de pré-sélection) :

$$\eta = \frac{N_{\text{Bkg}}}{N_{\text{Bkg}} + N_{\text{Sig}}} \tag{5.29}$$

où $\{\chi_i\}$ est l'ensemble des variables de discrimination utilisées, et G et H sont les fonctions de densité de probabilité associées aux évènements γ (N_{Sig}) et aux hadrons (N_{Bkg}), respectivement. Les corrélations entre les variables sont négligées ce qui permet d'exprimer les deux fonctions de densités de probabilité comme $G(\{\chi_i\}) = \sum_i g_i(\chi_i)$ et $H(\{\chi_i\}) = \sum_i h_i(\chi_i)$, où les notations g_i et h_i font référence à chacune des fonctions de densité de probabilité ajustées sur les simulations pour le signal, et sur les données réelles pour le fond. Les densités de probabilité associées à chacune des variables discriminantes sont illustrées sur la Figure 5.7, en vert pour les γ et en noir pour le fond. Sur la Figure 5.8 sont représentés, le rapport signal sur bruit (à gauche) et la significativité (à droite), tous les deux en fonction de la valeur de X_{eff} pour des données de la nébuleuse du Crabe,

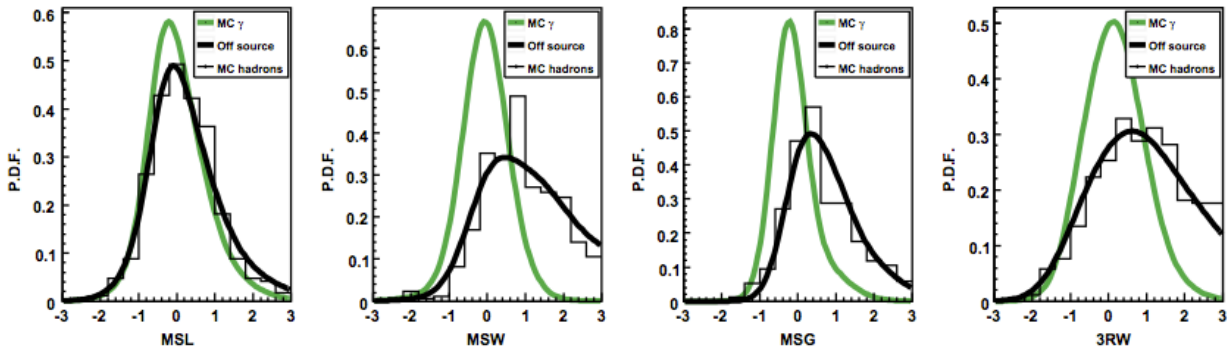


FIGURE 5.7: Distributions des fonctions de densité de probabilité des quatre variables discriminantes de l'analyse X_{eff}, en vert pour les γ et en noir pour le fond (histogrammes pour les données réelles et courbes pour les simulations) (Crédits : [Dubois et al., 2009]). Les paramètre MSL, MSR, MSG et 3RW tels que définis par [Dubois et al., 2009], ont été redéfinis dans ce chapitre comme : MSCL, MSCW, SG et w.

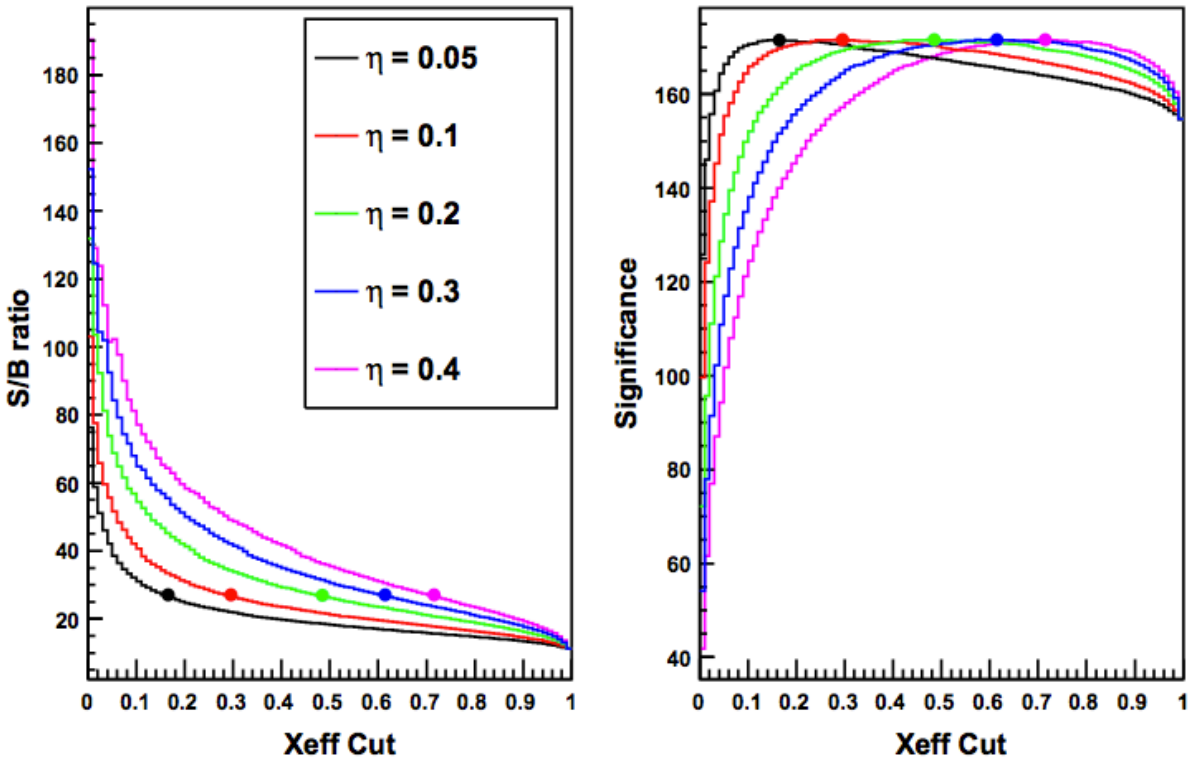


FIGURE 5.8: Distributions du rapport signal sur bruit et de la significativité en fonction du paramètre X_{eff}, pour différentes valeurs de η et pour des données de la nébuleuse du Crabe. Les points correspondent avec la valeur du paramètre X_{eff} pour laquelle le maximum de significativité est atteint (Crédits : [Dubois et al., 2009]).

et ce pour différentes valeurs de η. Les points indiquent pour chaque courbe la valeur de X_{eff} qui maximise la significativité. Ces figures montrent que la valeur η n'a d'incidence que sur la valeur de discrimination X_{eff} et pas sur le maximum de significativité qu'il est possible d'atteindre. Des résultats similaires ont été obtenus avec des sources plus faibles et ont conduit à l'utilisation d'une valeur unique de discrimination finale pour deux valeurs (et différentes coupures de pré-sélection) de η : $\eta = 0{,}1$ pour les sources ayant un niveau de flux intégré supérieur à 10 % de la nébuleuse du Crabe et $\eta = 0{,}5$ pour des sources moins brillantes.

Aujourd'hui, l'analyse X_{eff} a été révisée afin de remplacer l'analyse « model » par l'analyse modèle ++. Une publication est en cours de préparation [3].

5.1.5 ImPACT

Cette approche récente et très prometteuse est basée sur des abaques en deux dimensions pour la reconstruction et sur une approche multivariée pour la discrimination. Elle permet d'atteindre une sensibilité équivalente à celle du modèle ++.

5.1.5.1 Reconstruction des évènements γ

Comme dans le cas de l'analyse modèle ++, ImPACT (*Image Pixel-wise fit for Atmospheric Cherenkov Telescopes*) utilise des abaques [Parsons et Hinton, 2014] pour la reconstruction des γ incidents. La principale différence vient du fait qu'il n'y a pas de modèle analytique ici, les distributions des intensités attendues proviennent directement de simulations MC massives. Cette approche permet un meilleur traitement des fluctuations des gerbes EM, des images tronquées en bord de caméra ainsi que de la fenêtre d'intégration finie du signal qui implique une perte d'information pour les événements à grand pied de gerbe et de très haute énergie. La procédure de maximisation de la fonction vraisemblance est la même que celle présentée à la section 5.1.3.

5.1.5.2 Discrimination

Contrairement à l'analyse modèle ++, cette analyse n'utilise pas le paramètre SG présenté précédemment, mais une technique de discrimination basée sur une analyse multivariée développée dans [Ohm et al., 2009]. Cette technique consiste à utiliser la méthode à base d'arbres de décision boostés avec six paramètres discriminants :

- MSCL et MSCW : les largeurs et longueurs réduites définies à la section 5.1.1
- MRSLO et MRSWO : les largeurs et longueurs réduites utilisant des tables produites avec des évènements hadroniques
- X_{max} : la hauteur du maximum de gerbe reconstruite par ImPACT
- $\Delta E/E$: la dispersion en énergie reconstruite

L'espace des phases est divisé en cinq intervalles d'énergie reconstruite (entre ∼100 GeV et ∼100 TeV) et en six intervalles d'angles zénithaux (par pas de 10° entre 0° et 60°), chacun des intervalles ayant une coupure de discrimination finale dédiée. L'analyse est optimisée pour deux catégories de sources, celles ayant un indice $\Gamma = 2{,}6$ avec une flux intégré de 10 % de la nébuleuse du Crabe au-delà de 200 GeV, et celles ayant un niveau de flux de ∼1 % de celui de la nébuleuse du Crabe (dix fois plus faible) avec un indice spectral de 2,0.

3. En attendant, on pourra trouver les détails de la nouvelle analyse X_{eff} dans une note interne : https://hess-confluence.desy.de/confluence/download/attachments/3184196/Official_Xeffv2_note.pdf.

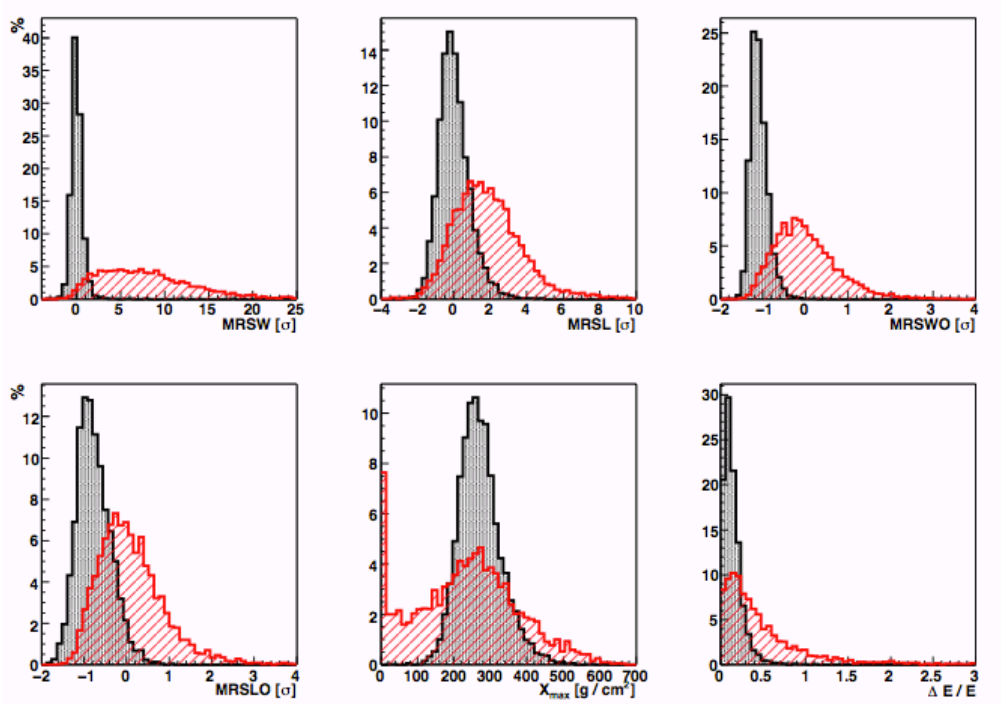


FIGURE 5.9: Paramètres discriminants de l'analyse ImPACT pour des évènements ayant une énergie entre 500 GeV et 1 TeV, et un angle zénithal entre 15° et 25° pour les γ (noir) et du fond (rouge). Voir texte pour la définition des variables. À noter que les paramètres reconstruits représentés sur cette figure proviennent de la méthode de Hillas (Crédits : [Ohm et al., 2009]).

5.2 L'analyse Paris-MVA

5.2.0.3 Introduction

L'analyse Paris-MVA [Becherini et al., 2011] a été développée et optimisée pour l'étude de sources faibles (∼1 % de la nébuleuse du Crabe) avec *H.E.S.S.* en utilisant une approche non-paramétrique. Afin d'avoir une analyse robuste, en particulier vis-à-vis des variations du niveau de NSB, une présélection a permis de retenir huit paramètres discriminants. Cinq de ces paramètres proviennent des approches standard et du modèle 3D, et trois d'entre eux ont été inventés afin d'améliorer le pouvoir de séparation γ/fond. Les paramètres de type « *Shower Goodness* » (section 5.1.3) n'ont pas été pris en compte du fait de leur sensibilité aux conditions d'observation qui nécessite une calibration très détaillée. La méthode a été optimisée pour différents angles zénithaux et différents offsets [4] d'observation de manière à pouvoir être utilisée pour l'analyse de l'ensemble des sources observées par *H.E.S.S.*.

Dans ce qui suit nous présentons les méthodes de reconstruction, puis la méthode d'estimation de l'énergie des évènements. Après une discussion sur les paramètres discriminants sélectionnés, nous décrivons la méthode de discrimination γ/fond.

5.2.1 Les reconstructions utilisées

L'analyse Paris-MVA utilise deux méthodes de reconstruction. Elles serviront d'une part à obtenir les caractéristiques des γ incidents, comme la direction d'arrivée, la hauteur du maximum de gerbe ou le pied de gerbe, et d'autre part à construire certaines des variables discriminantes qui seront utilisées par la suite.

4. L'offset est une décalage entre la position de la source et le centre du champ de vue.

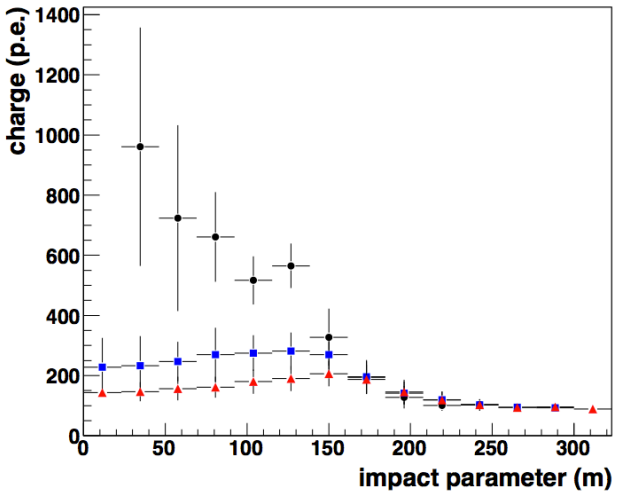


FIGURE 5.10: Charges Q en fonction du paramètre d'impact pour des γ simulés à trois angles zénithaux différents : 0° (points noirs), 32° (carrés bleus), 41° (triangles rouges) à énergie, offset, efficacité et maximum de la gerbe fixés (Crédits : [Becherini et al., 2011]).

La reconstruction géométrique de la gerbe est effectuée à l'aide de la méthode des ellipses présentée à la section 5.1.1. Un « nettoyage 510 » est préalablement effectué afin de minimiser la contamination du NSB. Les paramètres de Hillas sont ensuite calculés sur les images nettoyées et une pré-sélection est effectuée avant d'obtenir les caractéristiques géométriques de la gerbe. Cette pré-sélection consiste à exclure les images dont le barycentre des charges est trop proche du bord de la caméra (on choisit une distance maximale correspondant à 80 % du rayon d'une caméra) ainsi que celles ayant une charge totale trop faible suivant la configuration d'analyse dédiée à un certain type de sources (section 5.2.5). Enfin, la reconstruction de la gerbe est effectuée en ne gardant que les évènements ayant au moins deux télescopes passant ces coupures de pré-sélection selon la méthode exposée à la section 5.1.1. La direction de la gerbe, sa hauteur du maximum et enfin son pied de gerbe sont ainsi reconstruits.

La seconde reconstruction utilisée est celle qui provient de l'ajustement du modèle 3D. Les paramètres géométriques de la gerbe précédemment obtenus sont utilisés comme valeurs initiales pour l'ajustement. Celui-ci permet d'obtenir une direction et une énergie dans l'hypothèse où la gerbe est initiée par photon γ et permet aussi de caractériser la photosphère Tcherenkov qui fournit directement trois variables discriminantes, la largeur réduite, son erreur et la hauteur du maximum de gerbe (section 5.1.2). Afin de préserver le maximum d'information pour la discrimination [5] le lot de pixels retenu lors de l'ajustement est plus large que celui utilisé pour la reconstruction géométrique du paragraphe précédent. Il consiste, à partir d'un « nettoyage 57 », à ajouter les pixels se trouvant à 4σ de l'ellipse de Hillas obtenue avec ce même nettoyage. La photosphère ainsi ajustée est ensuite utilisée pour construire d'autres variables discriminantes. En effet, il est possible de tirer profit du manque de symétrie azimutale des gerbes hadroniques en exploitant plus en avant le modèle 3D de la photosphère, qui lui suppose cette symétrie (il s'agit d'un modèle de gerbe EM). Pour ce faire, la photosphère est projetée dans le plan focal des tous les télescopes, puis les images ainsi obtenues sont comparées aux images réelles à travers leurs paramètres de Hillas, en l'occurrence, la direction et l'énergie.

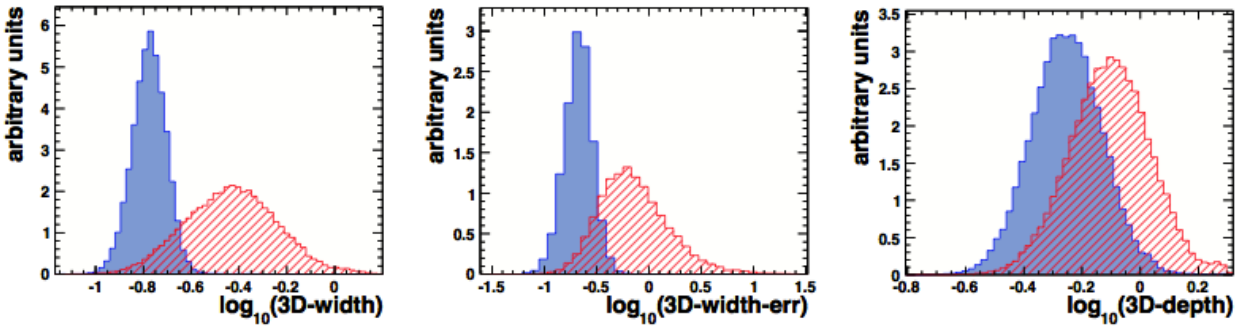


FIGURE 5.11: Paramètres provenant du modèle 3D utilisés dans la discrimination de l'analyse Paris-MVA (Crédits : [Becherini et al., 2011]). Les distributions des paramètres 3D-width et 3D-width-error sont représentées pour des évènements γ (bleu) et hadroniques (rouge) ayant des énergies entre (1 TeV et 2 TeV). Le paramètre 3D-depth est représenté pour des énergies entre 100 GeV et 300 GeV.

5.2.2 Reconstruction de l'énergie

L'estimation de l'énergie d'un évènement γ est réalisée sur la base de tables, générées à partir de simulations d'évènements γ à énergies fixes, qui contiennent le profil des charges mesurées dans un télescope en fonction du pied de gerbe par rapport à celui-ci (Figure 5.10). L'inversion de ces tables permet d'obtenir une relation entre la charge mesurée dans une caméra et l'énergie de l'évènement, et ce pour différents pieds de gerbe. Cette relation est ajustée pour différents paramètres d'observation, l'angle zénithal, l'offset et l'efficacité optique. Il est donc possible, à partir du pied de gerbe reconstruit, de déterminer une énergie pour chaque télescope dans l'hypothèse où il s'agit d'un rayon γ. Les différences entre les conditions d'observation et celles correspondant aux simulations sont prises en compte via un jeu d'interpolations. L'énergie finale $\widetilde{E}$ attribuée à l'évènement sera une moyenne des énergies estimées dans chaque télescope, pondérées par leurs charges.

5.2.3 Paramètres de discrimination γ/fond

L'analyse Paris-MVA utilise un total de huit variables discriminantes. Parmi ces variables, deux proviennent de la discrimination « à la Hillas » présentées à la section 5.1.1 : les MSCL et MSCW. Ces paramètres mesurent les déviations entre les longueurs et les largeurs des images réelles et celles provenant directement de simulations de γ. Trois paramètres provenant du modèle 3D sont aussi utilisés. Le premier est la largeur réduite ω présentée à la section 5.1.2, que l'on appellera ici 3D-width, et qui permet d'obtenir un pouvoir discriminant à toutes les énergies (à gauche de la Figure 5.11). Le second paramètre, appelé 3D-width-error, est défini comme l'erreur de la largeur ajustée dérivée par MINOS [6] et est utilisé parce que l'ajustement des évènements hadroniques par le modèle 3D (qui, rappelons-le, est un modèle de gerbe électromagnétique) produit en moyenne des largeurs moins bien contraintes que pour les évènements γ (au centre la Figure 5.11). Le troisième paramètre, appelé 3D-depth, correspond au paramètre D_s défini à la section 5.1.2 et est normalisé par la limite haute de l'équation 5.20 :

$$\text{3D-depth} = \log_{10}\left(\frac{D_s}{61{,}5(\ln N_c - 10{,}0 + T(\theta)) + 300}\right) \tag{5.30}$$

Il possède un pouvoir de discrimination γ/fond très important à basse énergie (à droite de la Figure 5.11).

5. Il s'agit surtout de ne pas perdre l'information sur les irrégularités des gerbes hadroniques (ce qu'on appelle « la barbe de hadrons ») comparées aux images issues des gerbes EM beaucoup plus régulières sur les bords.
6. http://seal.web.cern.ch/seal/documents/minuit/mnerror.pdf

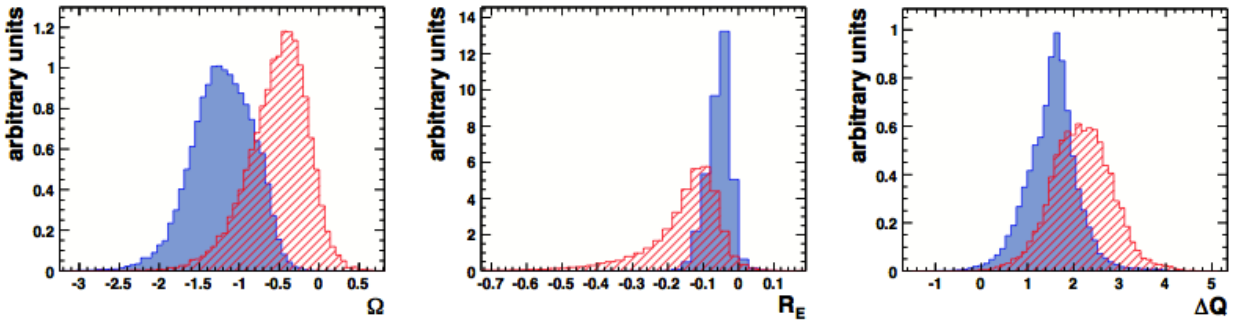


FIGURE 5.12: Nouveaux paramètres créés spécifiquement pour l'analyse Paris-MVA (Crédits : [Becherini et al., 2011]). Les distributions des paramètres Ω, R_{E} et ΔQ sont représentées pour des évènements γ (bleu) et hadroniques (rouge) ayant des énergies entre 1 TeV et 2 TeV pour des angles zénithaux entre 10° et 20°.

Le lot de paramètres utilisés pour la discrimination est complété par l'ajout de trois variables créés spécifiquement pour l'analyse Paris-MVA [Becherini et al., 2011]. Tout d'abord le paramètre Ω, qui exploite les différences entre la direction reconstruite à partir des images mesurées et la direction reconstruite à partir des images issues de la projection sur toutes les caméras de la photosphère du modèle 3D :

$$\Omega = \log_{10}[\mathrm{acos}\,(\vec{\theta}^{\mathrm{M3D}}_{\mathrm{Hillas}} \cdot \vec{\theta}_{\mathrm{Hillas}})] \tag{5.31}$$

La distribution de Ω est illustrée à gauche de la Figure 5.12 pour les hadrons (rouge) et les γ (bleu). Étant donné qu'on impose aux hadrons un modèle EM à symétrie azimutale, la direction basée sur les images issues de la projection de la photosphère sur les caméras s'écarte souvent de celle construite directement à partir des images réelles des gerbes hadroniques.

Ensuite, le paramètre R_{E} qui mesure le rapport entre l'énergie estimée ($\widetilde{E}$) à partir des images mesurées et celles issues de la photosphère $\widetilde{E}_{\mathrm{M3D}}$:

$$R_{\mathrm{E}} = \log_{10} \frac{\widetilde{E}}{\widetilde{E}_{\mathrm{M3D}}} \tag{5.32}$$

On peut observer au centre de la Figure 5.12 qu'une queue de distribution est présente pour les évènements hadroniques, alors que la distribution des évènements γ tend à être plus piquée.

Enfin, le dernier paramètre ΔQ est basé sur les différences entre les charges mesurées par chaque télescope Q^i et la charge $Q^{*,i}$ estimée à partir de l'énergie reconstruite $\widetilde{E}$ (à l'aide d'une inversion des tables de calibration en énergie) :

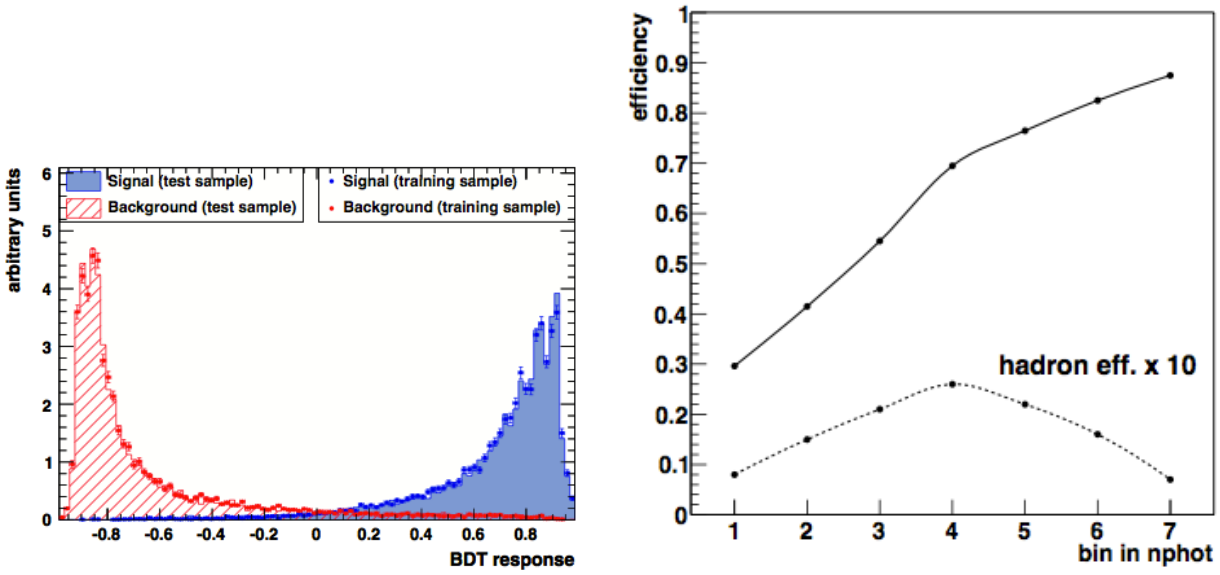
$$\Delta Q = \log_{10}\left(\frac{1}{N_{\mathrm{Tels}}} \sum_{i}^{N_{\mathrm{Tels}}} \frac{(Q^i - Q^{*,i})^2}{Q_{tot}}\right) \tag{5.33}$$

De grands écarts de charge sont attendus pour les évènements mal reconstruits (à droite de la Figure 5.12).

5.2.4 Discrimination γ/fond

La stratégie adoptée pour exploiter au mieux le pouvoir discriminant des paramètres présentés au paragraphe précédent consiste à utiliser une méthode multivariée non-paramétrique. Pour ce problème de classification binaire, la méthode d'analyse multivariée BDT (chapitre 3) a été choisie à la fois pour ses performances et pour sa capacité à gérer des variables dont le pouvoir discriminant varie selon les conditions d'observation [7].

7. Certaines variables peuvent par exemple être très discriminantes à haute énergie et peu discriminantes à basse énergie.



(a) (b)

FIGURE 5.13: (a) Exemple de distribution du paramètre ζ des évènements γ (bleu) et de fond (rouge), de multiplicité deux, ayant une énergie comprise entre ∼2 et ∼10 TeV et un angle zénithal entre 10 et 20°, pour les lots d'entraînement (trait plein) et de test (point). (b) Exemple de gabarit d'efficacité de reconstruction γ et efficacités de reconstruction du fond associées en fonction de l'intervalle en N_c au zénith et pour la multiplicité deux (Crédits : [Becherini et al., 2011]).

La classification γ/fond se déroule de la manière suivante. Tout d'abord, une forêt d'arbres de décision est générée à partir d'un lot d'évènements γ simulés et d'évènements de fond, le lot d'entraînement. Cette forêt est constituée d'arbres de grande profondeur et la méthode de « *pruning* » est utilisée afin de couper les branches qui sont statistiquement faibles [8]. Lors de cette phase, où les huit variables discriminantes précédemment introduites sont utilisées, chaque arbre de décision attribuera une note aux évènements (+1 pour des évènements « γ-like » et −1 pour les autres). Les décisions des différents arbres sont combinées afin d'attribuer une note unique, que l'on notera ζ, rendant compte de la nature « γ-like » des évènements. À l'aide d'un second lot d'évènements, le lot test, les performances du classificateur sont estimées en fonction d'une coupure sur la distribution de ζ : les efficacités de reconstruction du signal et du fond. Une fois la valeur finale de discrimination $\zeta^\star$ fixée, ce qui peut être réalisé de différentes manières, la nature « γ-like » de nouveaux évènements, obtenus par exemple lors d'observations réelles, peut être évaluée.

On peut trouver un exemple de distribution du paramètre ζ sur la Figure 5.13a pour des évènements γ (bleu) et de fond (rouge), correspondant à des conditions d'observation particulières, pour un lot ayant servi à la génération d'une forêt d'arbres de décision (traits) et pour un lot test (points).

Les distributions des variables discriminantes, ainsi que leur pouvoir discriminant relatif, diffèrent selon l'énergie, l'angle zénithal et la multiplicité des évènements. Afin de prendre en compte ces variations et optimiser le pouvoir de séparation γ/fond, un entraînement dédié est réalisé pour différents intervalles de cet espace des phases. Le découpage de ce dernier est effectué selon l'angle zénithal (8 intervalles) la multiplicité (un intervalle de multiplicité deux et un intervalle au-delà) et le nombre de photons composant la photosphère de la gerbe ajustée par le modèle 3D (7 intervalles, ce qui revient à prendre en compte la dépendance en énergie).

La coupure finale de discrimination $\zeta^\star$, qui est différente pour chaque intervalle de l'espace des phases, est déterminée à l'aide d'un gabarit des efficacités de reconstruction γ. Celui-ci est

8. On pourra trouver les réglages techniques de la méthode dans [Becherini et al., 2011].

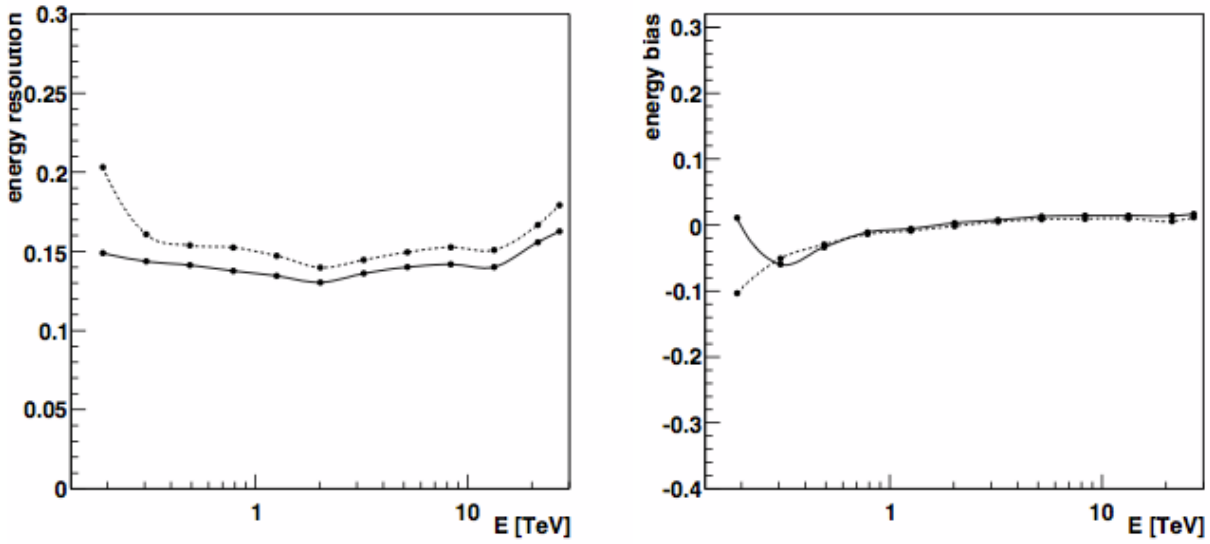


FIGURE 5.14: (a) Résolution et (b) biais en énergie de l'analyse Paris-MVA, estimées au zénith, pour deux configurations d'analyse qui sélectionnent les images avec un seuil en charge de 40 p.e. (traits pointillés) et de 150 p.e. (traits pleins) [Becherini et al., 2011].

déterminé à travers un processus itératif pour cinq configurations [9] adaptées à différents types de sources (indice et brillance). Un exemple de gabarit des efficacités de reconstruction γ en fonction l'intervalle en N_c est donné sur la Figure 5.13b pour l'intervalle en multiplicité deux et pour un angle zénithal de 0° (avec un seuil en charge pour la sélection des images de 110 p.e.).

5.2.5 Performances

Cinq configurations d'analyse ont été développées, qui diffèrent par le seuil minimal $Q_{\min}$ en charge des images sélectionnées. Choisir un $Q_{\min}$ bas permet d'augmenter de manière significative la surface de collection et donc de réduire l'énergie seuil de l'instrument. On privilégiera ce type de configuration pour l'analyse des sources faibles à indice spectral mou. À l'inverse, pour des études de morphologie de sources où une grande précision angulaire est requise, on augmentera le seuil en charge des images sélectionnés afin d'améliorer la reconstruction de la direction des évènements, au détriment d'une perte de surface de collection. Les cinq configurations développées sélectionnent les images selon des seuils en charge minimale $Q_{\min}$ de : 40, 60, 80, 110 et 150 p.e. Elles seront respectivement dénommées dans la suite par les qualificatifs : « elm », « fir », « ash », « teak » et « ipe » [10].

Sur la Figure 5.14, on peut observer une illustration des performances de la méthode d'estimation de l'énergie. À gauche de la figure est représentée la résolution en énergie (son écart-type) en fonction de l'énergie vraie pour des conditions d'observation au zénith. Elle est de l'ordre de ∼15 % à toutes les énergies, sauf près du seuil pour l'analyse « elm ». À droite de cette même figure est représenté le biais associé. Il est meilleur que 5 % au-delà de 300 GeV.

À gauche de la Figure 5.15 est représentée la précision angulaire à 68 % en fonction de l'angle zénithal d'observation, pour un offset de 0,5°, et pour les configurations « elm » (ligne pointillés) et « ipe » (ligne pleine). Elle vaut respectivement ∼0,1° et ∼0,08° pour les configurations « elm » et « ipe » en deçà de 40° d'angle zénithal. Pour rappel, cette précision angulaire correspond à celle d'une reconstruction géométrique de type « Hillas ».

La surface efficace de collection en fonction de l'énergie vraie est illustrée à droite de la Fi-

9. Les cinq configurations sélectionnent les images selon des seuils en charge différents : 40, 60, 80, 110 et 150 p.e.
10. Correspondant en anglais aux noms d'arbres à bois de plus en plus durs.

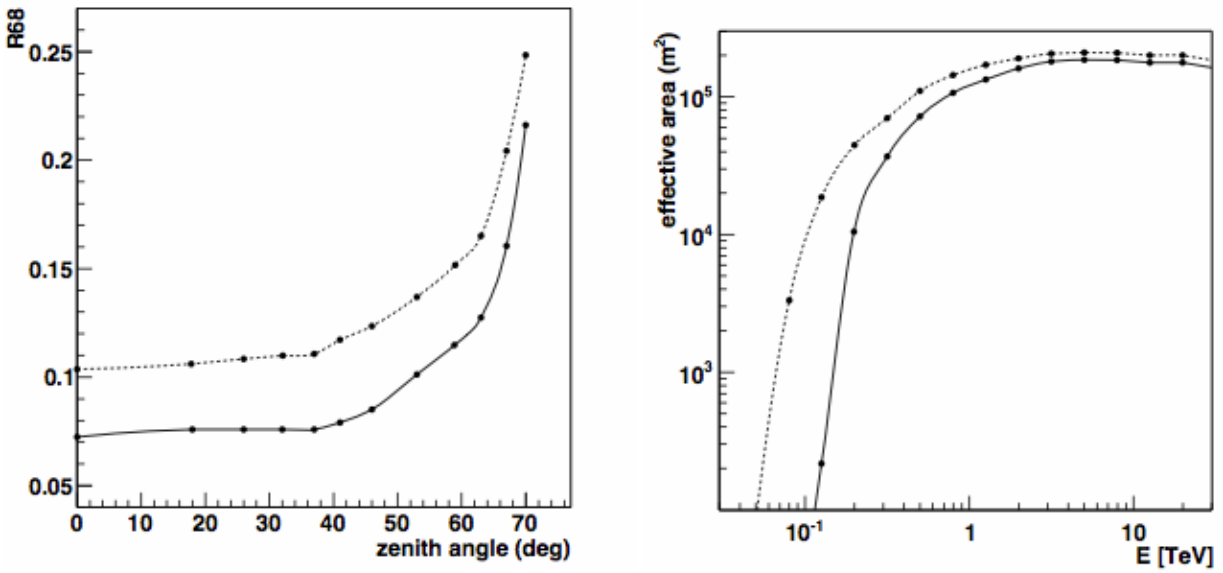


FIGURE 5.15: (a) Précision angulaire à 68 % en fonction de l'angle zénithal pour un offset de 0,5° et (b) surface efficace de collection en fonction de l'énergie vraie de l'analyse Paris-MVA à un angle zénithal de 18°, pour deux configurations d'analyse qui sélectionnent les images avec un seuil en charge de 40 p.e. (traits pointillés) et de 150 p.e. (traits pleins) [Becherini et al., 2011].

gure 5.15, pour un angle zénithal de 18°, pour une efficacité optique nominale, et pour les deux configurations précédemment citées. La configuration « elm » permet dans ces conditions d'atteindre un seuil en énergie (section 5.5.3) de l'ordre de ∼150 GeV, alors que le seuil énergie de la configuration « ipe » atteint ∼300 GeV.

On pourra se référer à la section 7 de la publication [Becherini et al., 2011] pour des comparaisons entre une analyse standard et l'analyse Paris-MVA sur une liste de sources détectées par *H.E.S.S.*, dont les résultats ont été précédemment publiés. L'analyse Paris-MVA est aujourd'hui largement utilisée dans la collaboration *H.E.S.S.*, en tant que première ou seconde analyse (respectivement pour la publication ou pour le contrôle des résultats). Elle a été utilisée comme première analyse pour les publications récentes sur la découverte des BL Lac 1RXS J101015.9-311909 [H.E.S.S. Collaboration et al., 2012] et KUV 00311-1938 [Becherini et al., 2012].

5.3 Extraction du signal et estimation du fond

Après avoir appliqué les algorithmes de reconstruction et les coupures de discrimination à des données obtenues lors d'observations réelles, l'évaluation d'un éventuel signal et de son niveau de significativité (section 5.3.1) nécessite d'estimer le fond résiduel. Ce fond est constitué majoritairement de hadrons, mais aussi d'électrons qui possèdent des caractéristiques proches de celles des évènements γ et qui passent donc à travers les coupures de sélection finales. Selon les analyses de haut niveau qui seront conduites, l'estimation de la forme spectrale de la source (section 5.4) ou la production d'une carte du ciel, deux techniques sont fréquemment utilisées pour l'extraction du fond résiduel (pour une discussion générale voir [Berge et al., 2007]).

5.3.1 Extraction du signal

En règle générale, pour analyser une source ponctuelle ou de faible extension les observations sont effectuées en mode « *wobble* », ce qui consiste à obtenir une image de la source dans les caméras

légèrement hors axe [11]. Une région circulaire d'une certaine extension, généralement de l'ordre de la précision angulaire à 68 %, est définie autour de la position de l'objet que l'on veut analyser (région ON), ainsi que plusieurs régions OFF dans lesquelles aucune source γ n'est présente. Les quantités N_{ON} et N_{OFF} qui sont respectivement les nombres d'évènements qui passent les coupures de sélection (« γ-like ») dans les régions ON et dans les régions OFF, sont alors estimées. On appellera excès le nombre d'évènements « γ-like » dans la région ON soustrait du bruit de fond résiduel et défini de manière simple par :

$$N_\gamma = N_{\mathrm{ON}} - \alpha N_{\mathrm{OFF}} \tag{5.34}$$

où α est le rapport des exposition entre la région ON et les régions OFF.

Une première estimation de la significativité d'un signal provenant d'une source γ consiste à comparer l'excès N_γ avec une estimation de l'erreur sur sa mesure (son écart-type) :

$$N_\sigma = \frac{N_\gamma}{\sigma(N_\gamma)} = \frac{N_{\mathrm{ON}} - \alpha N_{\mathrm{OFF}}}{\sqrt{N_{\mathrm{ON}} + \alpha^2 N_{\mathrm{OFF}}}} \tag{5.35}$$

Néanmoins, pour les cas où la statistique est faible ainsi que pour ceux où α diffère de 1, [Li et Ma, 1983] ont développé un estimateur de la significativité basé sur un test de rapport de maximum de vraisemblance entre une hypothèse d'existence de signal et une hypothèse d'absence de signal. Ce rapport de vraisemblance est noté λ et la quantité $-2\ln\lambda$ se comporte comme un χ^2 lorsque N_{ON} et N_{OFF} sont assez grands pour avoir un comportement gaussien. On définit le nombre de déviations standard N_σ par :

$$\begin{aligned} N_\sigma = \sqrt{-2\ln\lambda} = \sqrt{2}\Bigg\{ & N_{\mathrm{ON}} \ln\left[\frac{1+\alpha}{\alpha}\left(\frac{N_{\mathrm{ON}}}{N_{\mathrm{ON}}+N_{\mathrm{OFF}}}\right)\right] \\ & + N_{\mathrm{OFF}} \ln\left[(1+\alpha)\left(\frac{N_{\mathrm{OFF}}}{N_{\mathrm{ON}}+N_{\mathrm{OFF}}}\right)\right]\Bigg\}^{1/2} \end{aligned} \tag{5.36}$$

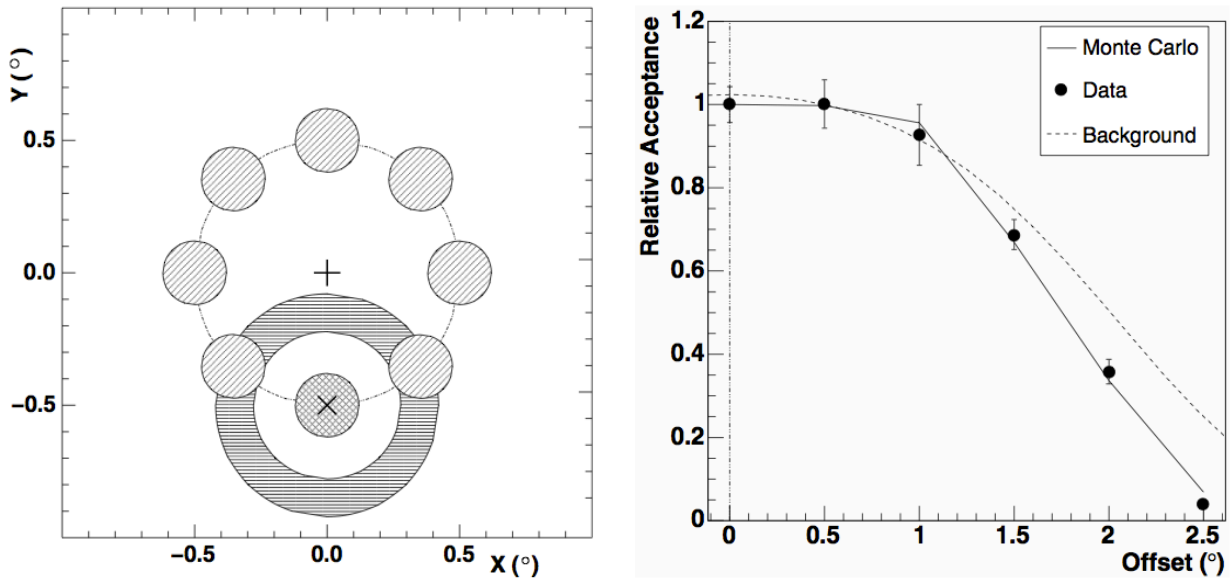
5.3.2 Méthode des régions OFF multiples

Utilisée pour la première fois par la collaboration *HEGRA* [Daum et al., 1997], la méthode des régions OFF multiples a été développée pour estimer le fond résiduel et reconstruire la forme spectral des sources. En supposant que l'acceptance radiale est de symétrie azimutale, on définit un nombre maximal de régions OFF de même taille que la région ON et dont les centres sont à la même distance radiale que celle-ci. À titre d'illustration, sur la Figure 5.16a le centre de la caméra est indiqué par le signe « + » et la position d'observation qui définit le centre de la région ON est indiquée par une croix. Y sont aussi représentées sept régions OFF. Lors de la création des régions OFF, les régions d'exclusions, qui sont définies par la position et l'étendue (ou un rayon de sûreté) des sources γ déjà détectées, sont prises en compte. La normalisation α entre les régions ON et OFF est définie comme étant le rapport de leurs surfaces respectives.

5.3.3 Méthode de la région en anneau

La méthode de la région en anneau est exclusivement utilisée pour la production de cartes du ciel. Elle consiste à estimer l'excès dans chaque pixel représentant une direction d'observation en évaluant son fond sous-jacent à l'aide d'un anneau centré sur celui-ci (voir l'anneau sur la Figure 5.16b, la croix représente la position du pixel considéré). Il est alors nécessaire de prendre en compte la différence d'acceptance radiale entre le pixel central, pour lequel on évalue le fond, et les pixels formant l'anneau. À titre d'exemple, on peut observer sur la Figure 5.16b l'acceptance radiale d'une caméra de type HESS-I pour du fond obtenu lors d'observation réelles (ligne pointillée). Sont également montrées sur cette figure, à titre de comparaison, les acceptances radiales obtenues à partir de γ simulés (ligne pleine) et à partir d'observations réelles d'une source γ.

11. Par exemple pour une source extragalactique on utilise un décalage de $\delta = \pm 0{,}5°$.



(a) (b)

FIGURE 5.16: (a) Illustration des régions utilisées pour extraire le signal d'une source γ. (b) Comparaison des acceptances radiales pour une caméra de type HESS-I pour des γ simulés (ligne pleine) et obtenus lors d'observations réelles après soustraction du fond (points), et pour du fond réel (ligne pointillée). (Crédits : [Aharonian et al., 2006a]).

5.4 Reconstruction spectrale et courbes de lumière

La reconstruction spectrale et la détermination des courbes de lumière nécessitent des IRF. Celles-ci sont générées à partir des simulations pour différentes configurations d'analyse et pour différentes conditions d'observation.

5.4.1 Fonctions de réponse de l'instrument

Les surfaces efficaces de collection et les résolutions en énergie sont générées à partir de γ simulés à énergies fixes. Ces simulations sont générées pour différentes conditions d'observation selon :

- 15 bandes en angle zénithal, que l'on note θ, entre 0° et 76° par pas de 0,05 en $\cos\theta$
- 6 bandes en offset de 0° et 3° par pas de 0,5°
- 6 bandes d'efficacité optique, notée ε et qui correspond à une fraction de l'efficacité optique nominale, variant de 50 % à 100 % par pas de 10 %
- 2 bandes en azimuth correspondant à 0° (vers le sud) et 180° (vers le nord)

Les énergie fixes sont définies entre 20 GeV et 125 TeV selon un pas permettant d'obtenir 5 intervalles par décade en logarithme.

La surface efficace à l'énergie vraie E, que l'on notera $\mathcal{A}(E;\theta,\delta,\varepsilon)$, est donnée par le rapport entre le nombre d'évènements passant les coupures de discrimination finales et le nombre d'évènements simulés, multiplié par la surface de tirage des simulations. Un exemple de surface efficace en fonction de l'énergie vraie est donnée sur la Figure 5.17 pour la configuration « elm » avec $\theta = 18°$, $\delta = 0{,}5°$ et $\varepsilon = 70\,\%$.

La résolution en énergie et son biais associé, que l'on notera respectivement $\mathcal{R}(E;\theta,\delta,\varepsilon)$ et $\mathcal{B}(E;\theta,\delta,\varepsilon)$ à l'énergie vraie E, sont obtenues à partir de la distribution de $\ln(\widetilde{E}/E)$ [12] qui est

12. Des intervalles d'énergie logarithmique sont utilisés pour évaluer le spectre, et le rapport $\ln(\widetilde{E}/E)$ représente bien l'effet attendu de migration pour cette représentation du spectre.

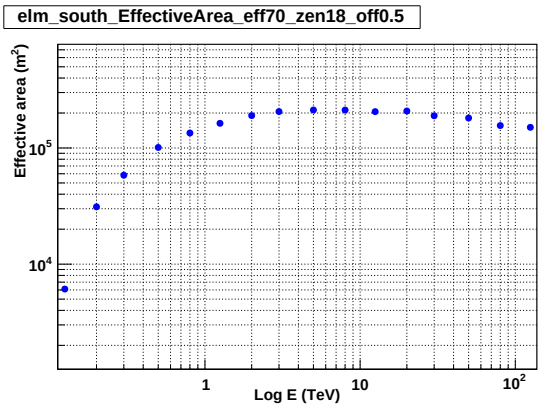
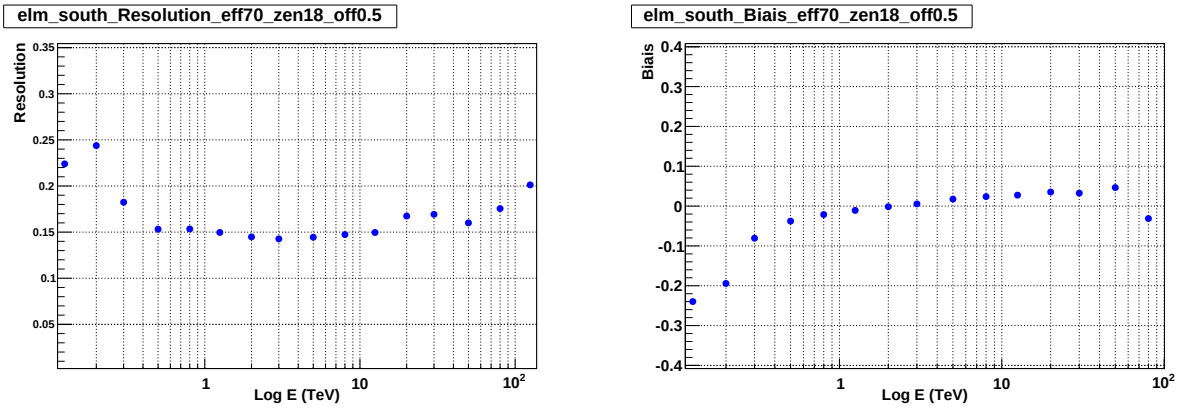


FIGURE 5.17: Surface de collection en fonction de l'énergie vraie pour la configuration « elm » avec $\theta = 18°$, $\delta = 0{,}5°$ et $\varepsilon = 70\,\%$



(a) (b)

FIGURE 5.18: (a) Résolution et (b) biais en énergie en fonction de l'énergie vraie pour la configuration « elm » avec $\theta = 18°$, $\delta = 0{,}5°$ et $\varepsilon = 70\,\%$

ajustée par une fonction gaussienne. Le biais est représenté par sa moyenne et la résolution par son écart-type. On pourra trouver respectivement sur les Figures 5.18a et 5.18b la variation en fonction de l'énergie de cette résolution et de son biais, obtenues pour les mêmes conditions que pour la Figure 5.17. On constate que la résolution en énergie est de l'ordre de 20 à 25 % en dessous du seuil en énergie de l'instrument pour ces conditions d'observation (qui est ici de ∼200 GeV) et atteint ∼15 % entre ∼500 GeV et ∼5 TeV. Quant au biais, il est de l'ordre de 20 % en dessous du seuil, et est inférieur à 5 % au-delà de 400 GeV. Notons que si les performances présentées ici pourraient clairement faire l'objet d'une optimisation, elles sont cependant suffisantes pour l'étude de sources ne présentant pas de structures spectrales variant rapidement avec l'énergie. Ces biais et résolution sont pris en compte de manière adaptée lors de la reconstruction spectrale des sources (voir ci-dessous).

5.4.2 Reconstruction spectrale par « *forward-folding* »

L'approche par « *forward-folding* » permet à la fois de prendre en compte la résolution en énergie de l'instrument (y compris les biais) et de traiter de manière adaptée les statistiques poissonniennes. Elle nécessite de faire une hypothèse sur la forme théorique de son spectre, que l'on notera $\phi(E;\{\Lambda\})$ et qui dépend d'un ensemble de paramètres $\{\Lambda\}$. Les lois spectrales auxquelles on fera référence dans cette thèse sont les suivantes :

— la loi de puissance :

$$\phi(E) = \phi_0 \left(\frac{E}{E_{\text{Ref}}}\right)^{-\Gamma}$$

qui permet de représenter l'émission γ de la majorité des AGN et une bonne partie des sources galactiques.

— la loi de puissance avec une coupure exponentielle :

$$\phi(E) = \phi_0 \left(\frac{E}{E_{\text{Ref}}}\right)^{-\Gamma} e^{-\beta E}$$

où l'énergie de coupure est donnée par β^{-1}.

— la loi de puissance parabolique :

$$\phi(E) = \phi_0 \left(\frac{E}{E_{\text{Ref}}}\right)^{-\Gamma-\beta \ln E}$$

qui correspond à une parabole dans une représentation νF_ν (ou $E^2\phi(E)$)

— la loi de puissance brisée continue :

$$\phi(E) = \phi_0 \left(\frac{E}{E_{\text{Ref}}}\right)^{-\Gamma_1} \left(1 + \left(\frac{E}{E_{\text{cut}}}\right)^{\frac{\Gamma_2-\Gamma_1}{\beta}}\right)^{-\beta}$$

qui est principalement utilisée pour caractériser l'émission pulsée des pulsars détectés par le *Fermi-LAT*.

Pour une hypothèse donnée $\phi(E;\{\Lambda\})$, l'ajustement de ses paramètres $\{\Lambda\}$ est effectuée avec une procédure de recherche du maximum d'une fonction de vraisemblance [Piron, 2000], lors de laquelle l'excès mesuré pendant un temps d'observation T_{Obs} et l'excès attendu, que l'on notera S^{th}, seront comparés. La détermination de cet excès S^{th} dans un intervalle d'énergie reconstruite i entre $\widetilde{E}_i$ et $\widetilde{E}_{i+1}$, prend en compte les IRF et s'exprime selon :

$$S_i^{\text{th}} = T_{\text{Obs}} \int_0^\infty \mathrm{d}E\, \phi(E;\{\Lambda\}) \mathcal{A}(E;\theta,\delta,\varepsilon) \int_{\widetilde{E}_i}^{\widetilde{E}_{i+1}} \mathrm{d}\widetilde{E}\, \mathcal{R}(\widetilde{E},E;\theta,\delta,\varepsilon) \tag{5.37}$$

où la fonction $\mathcal{R}(\widetilde{E}, E; \theta, \delta, \varepsilon)$ représente la densité de probabilité de mesurer une énergie reconstruite $\widetilde{E}$ pour une énergie vraie E donnée et $\mathcal{A}(E; \theta, \delta, \varepsilon)$ est la surface de collection présentée à la section précédente.

Considérons N intervalles d'énergie reconstruite $[\widetilde{E}_i, \widetilde{E}_{i+1}]$, que l'on dénotera par les indices i ($i \in [1, N]$), dans lesquels on a compté $N_{\mathrm{ON},i}$ évènements dans la région ON et $N_{\mathrm{OFF},i}$ évènements dans les régions OFF. On notera de la manière suivante les quantités attendues compte tenu de l'hypothèse considérée : $N^{\mathrm{th}}_{\mathrm{ON},i}$ et $N^{\mathrm{th}}_{\mathrm{OFF},i}$ qui vérifient $S^{\mathrm{th}}_i = N^{\mathrm{th}}_{\mathrm{ON},i} - \alpha N^{\mathrm{th}}_{\mathrm{OFF},i}$, où α représente le rapport des expositions entre les régions ON et OFF. La probabilité de compter $N_{\mathrm{ON},i}$ et $N_{\mathrm{OFF},i}$ dans l'intervalle d'énergie reconstruite i sachant que l'on s'attend à obtenir $N^{\mathrm{th}}_{\mathrm{ON},i}$ et $N^{\mathrm{th}}_{\mathrm{OFF},i}$ est un produit de deux lois de Poisson. On peut l'écrire :

$$\begin{aligned} P_i \equiv P_i(N_{\mathrm{ON},i}, N_{\mathrm{OFF},i} | N^{\mathrm{th}}_{\mathrm{ON},i}, N^{\mathrm{th}}_{\mathrm{OFF},i}) = & \frac{(N^{\mathrm{th}}_{\mathrm{ON},i})^{N_{\mathrm{ON},i}}}{N_{\mathrm{ON},i}\,!} \exp(-N^{\mathrm{th}}_{\mathrm{ON},i}) \\ & \times \frac{(N^{\mathrm{th}}_{\mathrm{OFF},i})^{N_{\mathrm{OFF},i}}}{N_{\mathrm{OFF},i}\,!} \exp(-N^{\mathrm{th}}_{\mathrm{OFF},i}) \end{aligned} \tag{5.38}$$

Les quantités $N^{\mathrm{th}}_{\mathrm{OFF},i}$, inconnues du fait de l'absence d'un modèle pour le fond, sont estimées en maximisant la fonction de vraisemblance, associée à la probabilité ci-dessus, par rapport à $N^{\mathrm{th}}_{\mathrm{OFF},i}$, sachant que cette équation $\partial\mathcal{L}/\partial N^{\mathrm{th}}_{\mathrm{OFF},i} = 0$ peut être résolue de manière analytique. Une fois l'expression de $N^{\mathrm{th}}_{\mathrm{OFF},i}$ déterminée, on l'injecte dans l'équation 5.38 et on construit la fonction de vraisemblance suivante :

$$\mathcal{L}(\{\Lambda\}, N^{\mathrm{th}}_{\mathrm{OFF},i}) = \prod_{i=1}^{N} P_i \tag{5.39}$$

Le maximum de cette fonction sera recherché numériquement afin d'estimer les paramètres $\{\Lambda\}$. Dans le cas général, plusieurs observations sont combinées afin de déterminer la forme spectrale d'une source γ et il est alors nécessaire de prendre en compte pour chaque observation les fonctions d'instruments adéquates (en fonction de θ, δ et ε). Le produit sur les intervalles d'énergie de l'équation équation 5.39 est alors remplacé par autant de produits qu'il y a de lots de données.

En testant deux hypothèses $\mathcal{H}_0$ et $\mathcal{H}_1$ de forme spectrale différentes, et appartenant à des familles paramétriques continues, on peut quantifier laquelle des hypothèses est privilégiée à l'aide du rapport des fonctions de vraisemblance : $\lambda_c = -2 \ln \frac{\mathcal{L}_{H_0}}{\mathcal{L}_{H_1}}$. Ce dernier se comporte asymptotiquement comme un χ^2 à un degré de liberté. En prenant comme hypothèse une loi de puissance pour $\mathcal{H}_0$ et par exemple une loi de puissance parabolique pour $\mathcal{H}_1$, il est possible de mettre en évidence la présence de courbure dans un spectre.

5.4.3 Courbes de lumière

Une courbe de lumière, qui permet de suivre la variabilité temporelle d'une source, représente le flux intégré au-delà d'une énergie donnée en fonction du temps. La détermination du niveau de flux intégré d'une source dans un intervalle de temps Δt suit la méthode décrite dans [H.E.S.S. Collaboration, 2010]. En supposant que seul le niveau de flux varie, on peut exprimer le flux intégré au-delà de l'énergie E_{min} en fonction du rapport entre l'excès mesuré et l'excès théorique :

$$\begin{aligned} \Phi(E \geq E_{\mathrm{min}}) = & \frac{N_{\mathrm{ON}} - \alpha N_{\mathrm{OFF}}}{\Delta t \int_0^\infty \mathrm{d}E\, \phi(E; \{\Lambda\}) \mathcal{A}(E; \theta, \delta, \varepsilon) \int_{E_{\mathrm{min}}}^\infty \mathrm{d}\widetilde{E}\, \mathcal{R}(\widetilde{E}, E; \theta, \delta, \varepsilon)} \\ & \times \int_{E_{\mathrm{min}}}^\infty \mathrm{d}E\, \phi(E) \end{aligned} \tag{5.40}$$

Le résultat dépend peu de la forme spectrale utilisée et dans le cas où cette dernière n'est pas déterminée, il est tout de même possible d'estimer un niveau de flux en supposant une loi spectrale simple et en prenant une énergie minimale qui n'est pas trop proche du seuil de l'instrument.

5.5 START

Dans cette section nous présentons un outil d'analyse de haut niveau développé au cours de cette thèse afin de reconstruire par méthode de « *forward-folding* » (section 5.4.2) la forme spectrale des sources γ et d'en faire les courbes de lumière. Cet outil a été conçu dans un langage orienté objet de manière à répondre aux critères suivants : permettre une utilisation souple selon les besoins de l'utilisateur (ajout d'hypothèses spectrales supplémentaires, ajustement multi-composantes, gestion et visualisation des fonctions d'instrument), permettre une adaptation facile pour différents types d'instruments (HESS-I, HESS-II et plus tard *CTA*), et enfin assurer un temps d'exécution rapide. Ce dernier point est assuré par une gestion des données astucieuse et par l'utilisation de méthodes rapides et robustes pour le calcul du maximum de la fonction de vraisemblance [13] et des intégrales de l'équation 5.37 [14]. START [15] a été conçu de manière à pouvoir être utilisé par les membres de la collaboration *H.E.S.S.*, indépendamment du cadre d'analyse de bas niveau choisi [16]. Dans un soucis de portabilité, il est indépendant du logiciel SASH, qui est au cœur des analyses « *on-line* » et de bas niveau de la collaboration *H.E.S.S.* Il est testé en Allemagne comme une piste pour les analyses de haut niveau de *CTA*. Cet outil est aujourd'hui utilisé par plusieurs collègues travaillant sur l'analyse des données de la collaboration *H.E.S.S.*, en particulier à l'APC et au MPIK (Heidelberg).

5.5.1 Gestion des données

Deux classes de stockage sont utilisées afin de gérer les données obtenues après application des coupures de sélection finales. La première, que l'on appellera « Band », représente généralement un « run » d'observation et contient les caractéristiques de celui-ci telles que l'angle zénithal, l'offset, l'efficacité optique, le temps effectif d'observation, etc. On y enregistre également les IRF moyennées (section 5.5.2) correspondant aux paramètres de l'observation : surface efficace de collection, résolution en énergie et biais, le tout en fonction de l'énergie vraie. Notons qu'il est possible de réaliser un découpage des données différent (plus petit ou plus grand qu'un run), ce qui est utile par exemple lorsque la variation d'angle zénithal (et donc d'IRF) au cours du « run » paraît trop importante. L'utilisateur ayant découpé préalablement les données en un certain nombre d'intervalles en énergie, les informations relatives à chacun de ces intervalles sont stockées dans une classe dénommée « EnergyBin », elle-même membre de la classe « Band ».

Une instance de la classe « EnergyBin » contient les caractéristiques de l'intervalle en énergie reconstruite $[\widetilde{E}_{\rm min}\ ,\ \widetilde{E}_{\rm max}]$, telles que le nombre d'évènements dans la région ON et dans la région OFF, l'excès et la significativité associés. Après la phase de maximisation de la vraisemblance, ces instances seront complétées avec les valeurs théoriques attendues pour le nombre de ON et de OFF. La classe « EnergyBin » contient également une partie liée aux IRF correspondant à la seconde intégrale de l'équation 5.37 (voir ci-dessous).

5.5.2 Gestion des IRF

Les IRF correspondant à l'ensemble des caractéristiques d'une observation donnée sont calculées par interpolation linéaire entre les paramètres MC des IRF. Les fonctions moyennées qui en résultent, et qui sont déterminées pour chacune des énergies fixes, sont étendues à l'ensemble de la gamme en énergie vraie avec la méthode d'interpolation « Akima » [Akima, 1970]. Celle-ci permet d'obtenir des fonctions lisses pour faciliter la procédure d'intégration sur l'énergie vraie de l'équation 5.37. Sur la Figure 5.19, on peut trouver l'exemple d'une surface efficace interpolée avec la méthode

13. http://seal.web.cern.ch/seal/MathLibs/Minuit2/html/index.html
14. http://www.gnu.org/software/gsl/
15. Le lien suivant permet de trouver des informations relatives à START (accessible par les membres de la collaboration *H.E.S.S.*) : https://hess-confluence.desy.de/confluence/display/HESS/HESS+Open+Source+Tools+-+HOWTO+get+started+with+START.
16. Il est fréquemment utilisé pour les analyses réalisées avec HAP (H.E.S.S. Analysis Program) et a été adapté pour les sorties de ParisAnalysis.

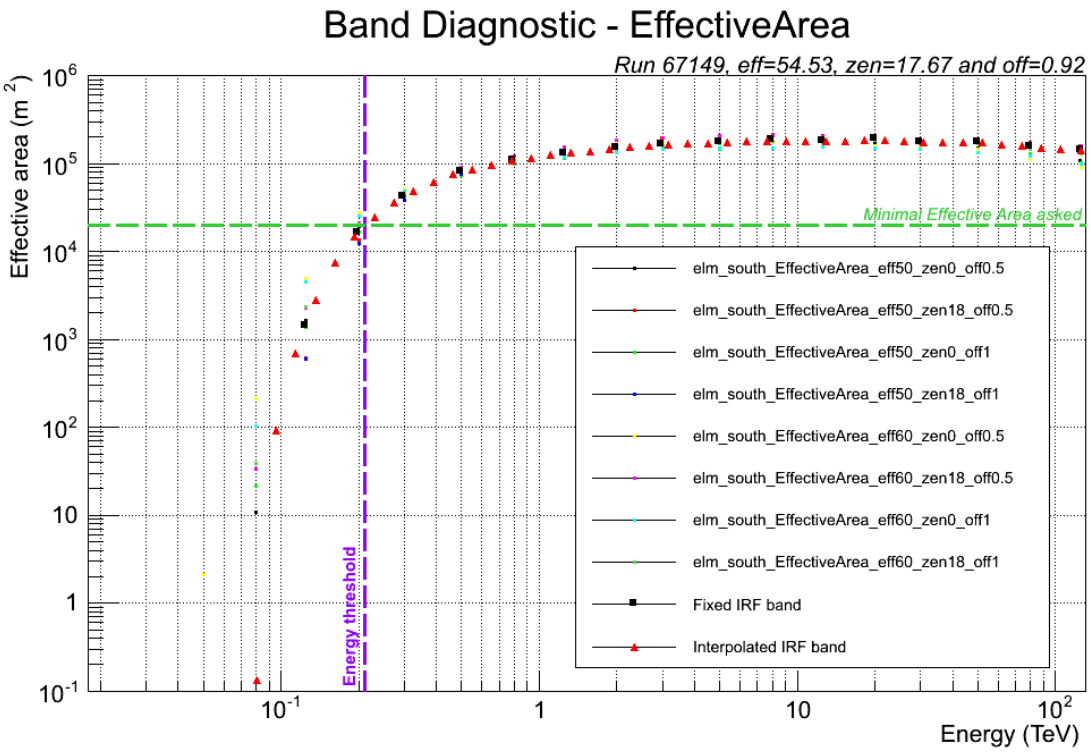


FIGURE 5.19: Surface efficace associée à un « run » d'observation.

« Akima », représentée par des triangles rouges, correspondant à un « run » d'observation. Les plus petits points représentent les surfaces efficaces générées avec des paramètres MC qui encadrent les caractéristiques d'observation du « run » (ici $\theta = 17{,}7°$, $\delta = 0{,}9°$ et $\varepsilon = 54{,}5\,\%$). Les points noirs correspondent à l'interpolation linéaire de la surface pour chaque énergie fixe.

Afin d'optimiser le calcul de l'intégrale d'évènements attendus (équation 5.37) dans chaque intervalle d'énergie reconstruite, représenté par un objet « EnergyBin », on lui associe la fonction $\kappa(E)$ qui dépend de l'énergie vraie :

$$\kappa(E) = \int_{\widetilde{E}_{\min}}^{\widetilde{E}_{\max}} \mathrm{d}\widetilde{E}\, \mathcal{R}(\widetilde{E}, E; \theta, \delta, \varepsilon) \tag{5.41}$$

Celle-ci est construite à partir de la résolution en énergie $\mathcal{R}(\widetilde{E}, E; \theta, \delta, \varepsilon)$ et correspond à la probabilité de mesurer une énergie reconstruite comprise entre $\widetilde{E}_{\min}$ et $\widetilde{E}_{\max}$ pour un évènement ayant une énergie vraie E (seconde intégrale sur l'énergie reconstruite de l'équation 5.37). Cette fonction est calculée une fois pour toute pour chaque intervalle et son adresse mémoire est stockée dans l'instance de l'« EnergyBin » correspondante. Cela induit un gain significatif en temps de calcul.

5.5.3 Détermination de l'énergie seuil

Lors de la phase de recherche du maximum de vraisemblance, il est nécessaire de déterminer pour chaque observation (ou « Band ») une énergie seuil. Celle-ci sera l'énergie minimale à partir de laquelle les données seront utilisées pour le calcul de la fonction de vraisemblance. Il n'existe pas de définition formelle de l'énergie seuil. Son rôle est de limiter les systématiques liées à basse énergie. Elle peut être obtenue en convoluant la surface efficace de collection d'une « Band » avec un spectre de source typique et en prenant l'énergie qui correspond à un certain pourcentage du maximum de la fonction qui en résulte. Ici, afin de se prémunir des effets liés à d'éventuelles systématiques liées aux fonctions d'instrument à faible énergie, on utilisera deux critères pour déterminer l'énergie seuil

d'une « Band ». On demandera à ce que, d'une part la surface efficace de collection soit supérieure à une valeur minimale, et d'autre part à ce que le biais en énergie soit inférieur à deux fois la résolution en énergie. Pour une configuration d'analyse classique comme la méthode de Hillas, un seuil de surface efficace de collection de 4 ha est utilisé, et dans le cas d'une analyse multivariée on descendra ce seuil jusqu'à 2 ou 3 ha. On peut observer sur la Figure 5.19, symbolisée par une ligne verticale violette, l'énergie seuil d'une « Band » correspondant à une observation à $\theta = 17{,}7°$, $\delta = 0{,}9°$ et $\varepsilon = 54{,}5\,\%$. Un minimum de 2 ha pour la surface efficace de collection a été demandé (courbe verte). Durant la phase de maximisation de la fonction de vraisemblance, tous les intervalles d'énergie dont la borne inférieure se trouve sous l'énergie seuil sont exclus de l'analyse.

5.5.4 Représentation d'un spectre

La procédure de maximisation de la vraisemblance permet d'estimer les paramètres de l'hypothèse spectrale testée, ainsi que la matrice de covariance associée. On associe généralement au spectre un intervalle de confiance à 68 % (voir les exemples à la section suivante) calculé à l'aide de la matrice de covariance. On y associe aussi des points de flux et des résidus qui sont évalués en énergie reconstruite à partir de la différence entre le signal mesuré et le signal attendu compte tenu du résultat de l'ajustement spectral. Ces derniers ne sont pas directement issus de la procédure de maximisation de la fonction de vraisemblance, ils permettent simplement de juger de l'écart entre la mesure et la prédiction. Seuls les paramètres ajustés $\{\Lambda\}$ ont une importance physique. Les erreurs associées au flux et aux résidus sont calculés avec la méthode présentée dans [Rolke et al., 2005]. Elle permet de rendre compte des statistiques du signal et du fond mesurés qui sont le plus souvent poissonniennes.

5.5.5 Quelques exemples d'applications

À titre d'exemple le spectre de la nébuleuse du Crabe obtenu à partir de 43,8 h d'observation ($N_\sigma \approx 168$) est représenté sur la Figure 5.20 entre 370 GeV et 50 TeV. On retrouve une loi de puissance avec une coupure exponentielle similaire à celle présentée dans [Aharonian et al., 2006a]. Un autre exemple est montré sur la Figure 5.21 qui correspond à la courbe de lumière du BL Lac PKS 2155-304 lors du célèbre sursaut de 2006. Elle a été générée dans les conditions décrites dans [Aharonian et al., 2007]. De la même manière que pour les résidus, le traitement des erreurs est réalisé avec la méthode de « Rolke ».

START a fréquemment été utilisé avec l'analyse Paris-MVA au cours de cette thèse, notamment pour assurer le « round-up », tâche de service dont j'ai eu la responsabilité pendant un an et qui consiste à présenter mensuellement à la collaboration *H.E.S.S.* le résultat d'une analyse systématique des observations des AGN. Les analyses ont été conduites avec les configurations « elm » et « fir », celles-ci étant adaptées respectivement pour l'étude des AGN peu brillants ayant un spectre mou et un spectre dur. Nous montrons ci-dessous deux exemples, l'un correspondant au BL Lac Mkn 501 et l'autre à la radiogalaxie M 87.

La Figure 5.22 montre des résultats relatifs à Mkn 501. Tout d'abord, la courbe de lumière qui correspond à trois périodes d'observation, juin 2004, juillet 2007 et juin 2012 est montrée sur la Figure 5.22a. La dernière série d'observations a fait partie d'une campagne multi-longueur d'onde qui a été déclenchée par l'expérience *MAGIC*. On peut observer sur la Figure 5.22b l'évolution de la significativité en fonction du temps d'observation cumulé. Un spectre moyen y est associé (Figure 5.22c) pour une hypothèse de loi de puissance parabolique (la valeur $\lambda_c \approx 2{,}0$ ne permet pas de conclure sur la présence de courbure dans le spectre).

Nous terminons cette série d'exemples avec la radiogalaxie M 87, initialement détectée par le réseau HESS-I en 2005 lors d'un état haut [Beilicke et al., 2005]. On dispose aujourd'hui de plus de $\sim$107 h de données. Cette source est peu brillante, son flux est de l'ordre du pourcent de la nébuleuse de Crabe, et possède un indice relativement dur ($\Gamma \sim 2{,}5$). Ces caractéristiques sont adaptées pour la configuration « fir ». Après analyse de l'ensemble des données disponibles, un excès de $\sim$550

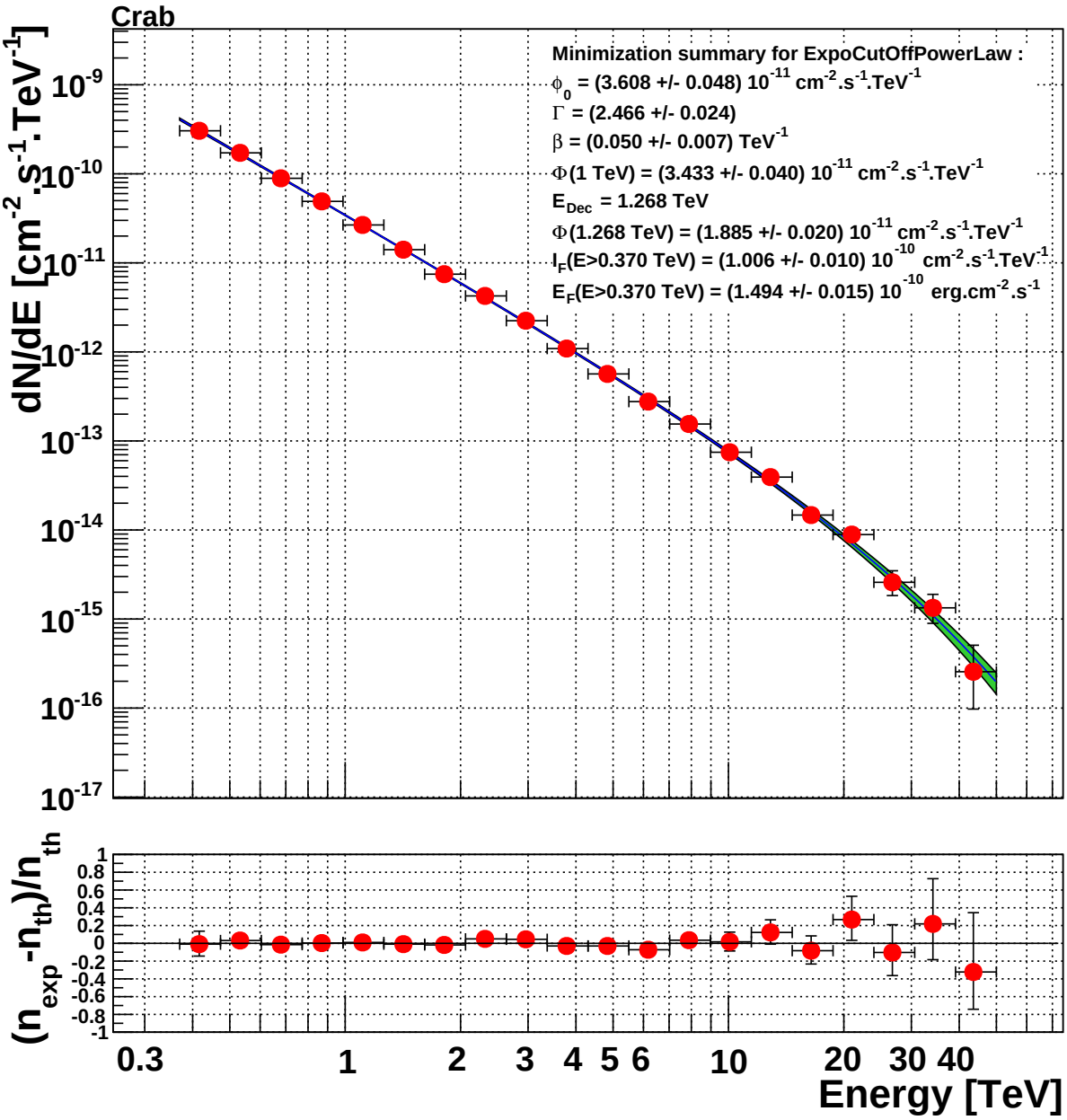


FIGURE 5.20: Spectre de la nébuleuse du Crabe avec une hypothèse de loi de puissance avec une coupure exponentielle généré avec START pour un temps d'observation effectif de 43,8 h ($N_\sigma \approx 168$).

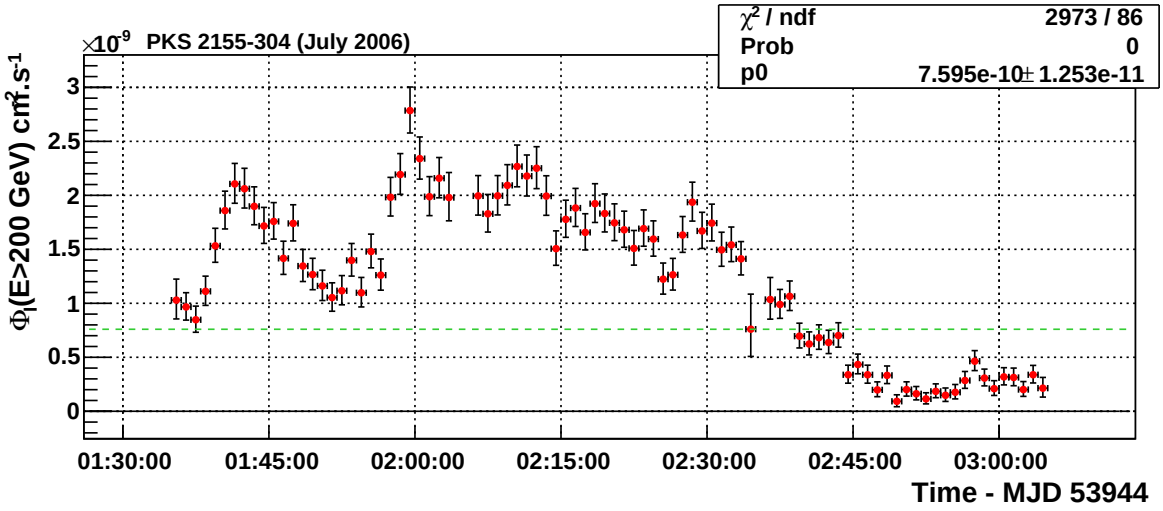
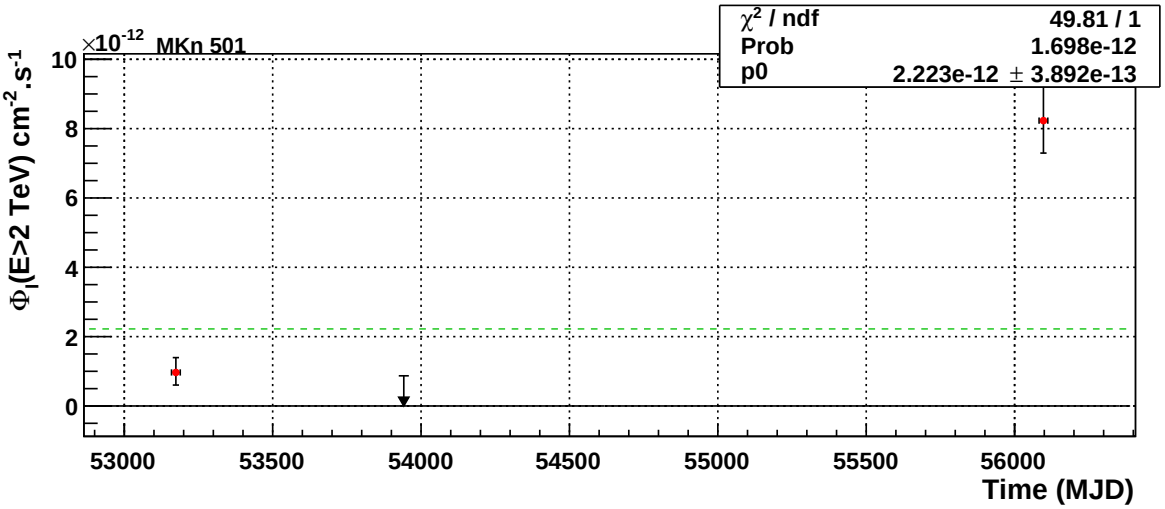
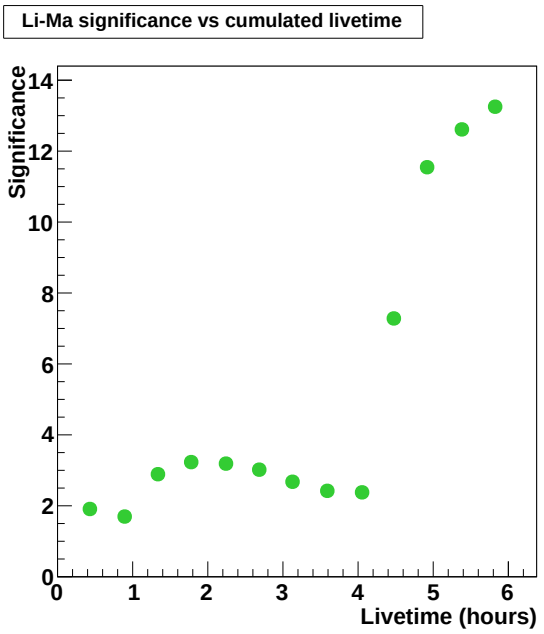


FIGURE 5.21: Exemple de courbe de lumière de la source PKS 2155-304 générée avec START. État haut observé lors de la nuit du 28 juillet 2006.

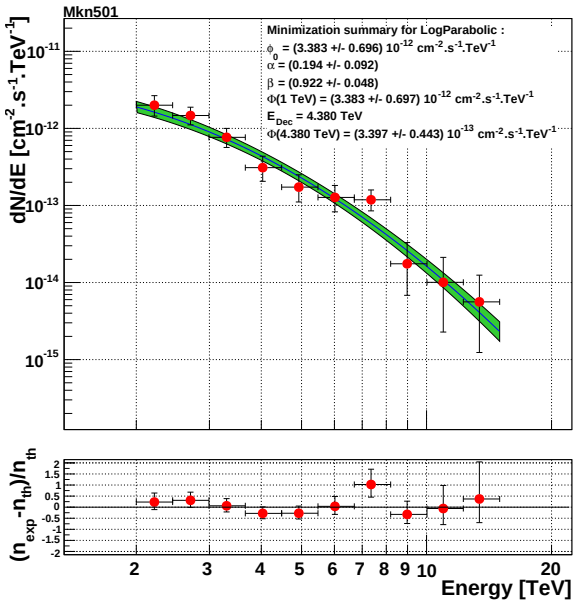
γ est obtenu pour une significativité de 17σ. La courbe de lumière, dont l'intervalle de temps Δt correspond à une année, est représentée sur la Figure 5.23a. On notera l'état haut de M 87 en 2005 qui correspond à la première détection, citée plus haut. Enfin, les Figures 5.23b et 5.23c représentent respectivement les spectres différentiels de M 87 pour une hypothèse de loi de puissance et une hypothèse de loi de puissance parabolique. Le rapport $\lambda \approx 17$ indique la présence d'une courbure significative dans le spectre.



(a) Courbe de lumière.

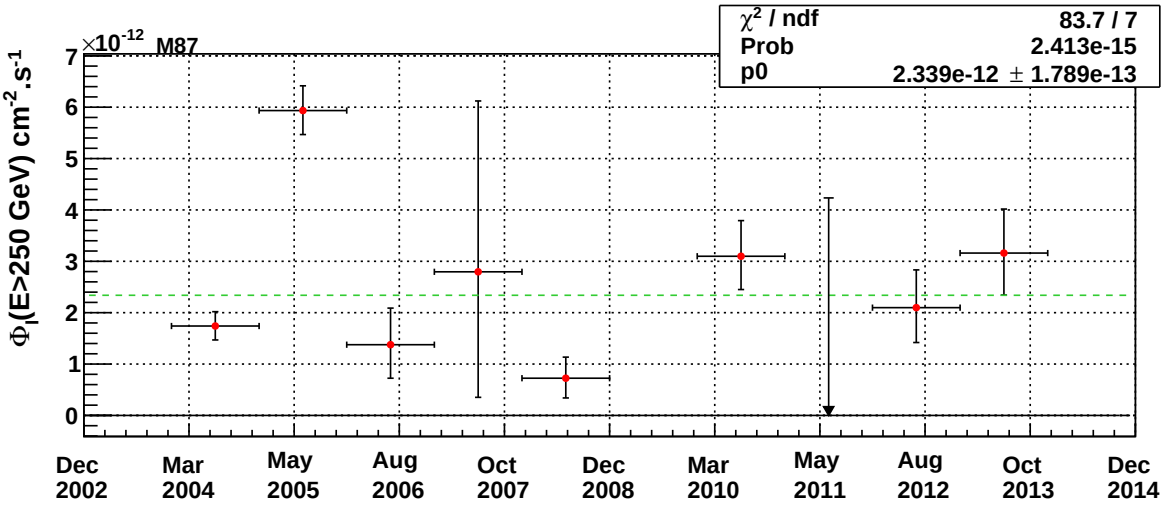


(b) Évolution de la significativité en fonction du temps d'observation effectif.

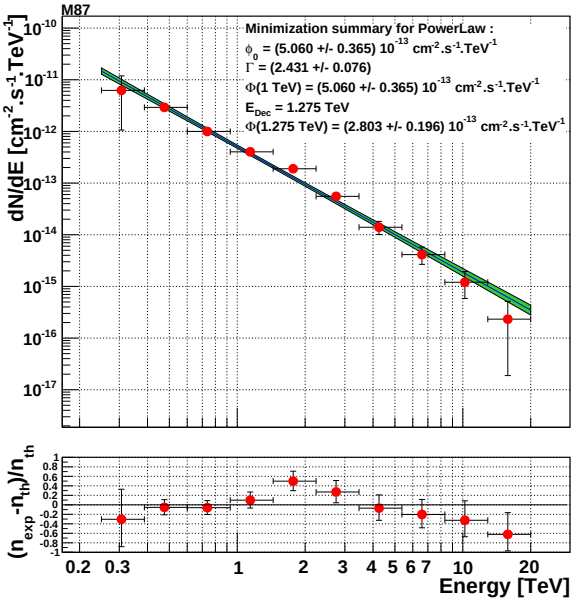


(c) Spectre moyenné sur l'ensemble des données pour une hypothèse de loi de puissance parabolique.

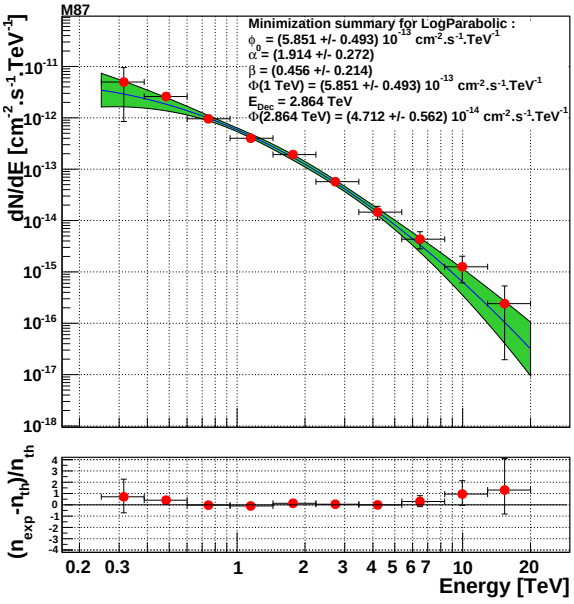
FIGURE 5.22: Activité de la source Mkn 501 de juin 2004 à juin 2012.



(a) Courbe de lumière.



(b) Simple loi de puissance.



(c) Loi de puissance parabolique.

FIGURE 5.23: Courbe de lumière et spectres de la radiogalaxie M 87 pour un temps d'observation effectif de 107 h ($N_\sigma = 17$) et différentes hypothèses de lois spectrales.

Bibliographie

A. Abramowski, et al. Probing the extent of the non-thermal emission from the Vela X region at TeV energies with H.E.S.S. A&A, 548:A38, December 2012a. doi: 10.1051/0004-6361/201219919.

A. Abramowski, et al. Spectral Analysis and Interpretation of the γ-Ray Emission from the Starburst Galaxy NGC 253. ApJ, 757:158, October 2012b. doi: 10.1088/0004-637X/757/2/158.

F. Aharonian, et al. Observations of the Crab nebula with HESS. A&A, 457:899–915, October 2006a. doi: 10.1051/0004-6361:20065351.

F. Aharonian, et al. Discovery of very high energy γ-ray emission from the BL Lacertae object H 2356-309 with the HESS Cherenkov telescopes. A&A, 455:461–466, August 2006b. doi: 10.1051/0004-6361:20054732.

F. Aharonian, et al. An Exceptional Very High Energy Gamma-Ray Flare of PKS 2155-304. ApJ, 664:L71–L74, August 2007. doi: 10.1086/520635.

H. Akima. A new method of interpolation and smooth curve fitting based on local procedures. *Journal of the Association for Computing Machinery*, 17(4):589–602, October 1970.

Y. Becherini, et al. A new analysis strategy for detection of faint γ-ray sources with Imaging Atmospheric Cherenkov Telescopes. *Astroparticle Physics*, 34:858–870, July 2011. doi: 10.1016/j.astropartphys.2011.03.005.

Y. Becherini, et al. Discovery of VHE γ-ray emission from the very distant BL Lac KUV 00311-1938 by H.E.S.S. In F. A. Aharonian, W. Hofmann, et F. M. Rieger, editors, *American Institute of Physics Conference Series*, volume 1505 of *American Institute of Physics Conference Series*, pages 490–493, December 2012. doi: 10.1063/1.4772304.

M. Beilicke, et al. Observation of the giant radio galaxy M87 at TeV energies with the H.E.S.S. Cherenkov telescopes. *International Cosmic Ray Conference*, 4:299, 2005.

D. Berge, S. Funk, et J. Hinton. Background modelling in very-high-energy γ-ray astronomy. A&A, 466:1219–1229, May 2007. doi: 10.1051/0004-6361:20066674.

A. Daum, et al. First results on the performance of the HEGRA IACT array. *Astroparticle Physics*, 8:1–11, December 1997. doi: 10.1016/S0927-6505(97)00031-5.

M. de Naurois. Analysis methods for Atmospheric Cerenkov Telescopes. *ArXiv Astrophysics e-prints*, July 2006.

M. de Naurois. L'astronomie gamma de très haute énergie de h.e.s.s. à cta. ouverture d'une nouvelle fenêtre astronomique sur l'univers non thermique., 2012.

M. de Naurois et L. Rolland. A high performance likelihood reconstruction of γ-rays for imaging atmospheric Cherenkov telescopes. *Astroparticle Physics*, 32:231–252, December 2009. doi: 10.1016/j.astropartphys.2009.09.001.

M. de Naurois, et al. Application of an Analysis Method Based on a Semi-Analytical Shower Model to the First H.E.S.S. Telescope. *International Cosmic Ray Conference*, 5:2907, July 2003.

F. Dubois, G. Lamanna, et A. Jacholkowska. A multivariate analysis approach for the imaging atmospheric Cherenkov telescopes system H.E.S.S. *Astroparticle Physics*, 32:73–88, September 2009. doi: 10.1016/j.astropartphys.2009.06.003.

H. E. S. S. Collaboration, et al. Search for dark matter annihilation signatures in H.E.S.S. observations of Dwarf Spheroidal Galaxies. *ArXiv e-prints*, October 2014a.

H. E. S. S. Collaboration, et al. H.E.S.S. detection of TeV emission from the interaction region between the supernova remnant G349.7+0.2 and a molecular cloud. *ArXiv e-prints*, December 2014b.

H.E.S.S. Collaboration. VHE γ-ray emission of PKS 2155-304: spectral and temporal variability. A&A, 520:A83, September 2010. doi: 10.1051/0004-6361/201014484.

H.E.S.S. Collaboration, et al. Discovery of VHE γ-ray emission and multi-wavelength observations of the BL Lacertae object 1RXS J101015.9-311909. A&A, 542:A94, June 2012. doi: 10.1051/0004-6361/201218910.

A. M. Hillas. Cerenkov light images of EAS produced by primary gamma. In F. C. Jones, editor, *International Cosmic Ray Conference*, volume 3 of *International Cosmic Ray Conference*, pages 445–448, August 1985.

W. Hofmann. Data analysis techniques for stereo IACT systems. In B. L. Dingus, M. H. Salamon, et D. B. Kieda, editors, *American Institute of Physics Conference Series*, volume 515 of *American Institute of Physics Conference Series*, pages 318–322, June 2000. doi: 10.1063/1.1291386.

S. Le Bohec, et al. A new analysis method for very high definition imaging atmospheric Cherenkov telescopes as applied to the CAT telescope. *Nuclear Instruments and Methods in Physics Research A*, 416:425–437, October 1998. doi: 10.1016/S0168-9002(98)00750-5.

M. Lemoine-Goumard, B. Degrange, et M. Tluczykont. Selection and 3D-reconstruction of gamma-ray-induced air showers with a stereoscopic system of atmospheric Cherenkov telescopes. *Astroparticle Physics*, 25:195–211, April 2006. doi: 10.1016/j.astropartphys.2006.01.005.

Marianne Lemoine-Goumard. Stéréoscopie de gerbes de gamma avec les télescopes h.e.s.s. : premières images de vestiges de supernovæ au tev. Master's thesis, Ecole Polytechnique, 2006.

T.-P. Li et Y.-Q. Ma. Analysis methods for results in gamma-ray astronomy. ApJ, 272:317–324, September 1983. doi: 10.1086/161295.

P. K. MacKeown, et al. Primary Cosmic Ray Spectrum above 10**12 EV from Cerenkov Light Images. *International Cosmic Ray Conference*, 9:175, August 1983.

S. Ohm, C. van Eldik, et K. Egberts. γ/hadron separation in very-high-energy γ-ray astronomy using a multivariate analysis method. *Astroparticle Physics*, 31:383–391, June 2009. doi: 10.1016/j.astropartphys.2009.04.001.

R. D. Parsons et J. A. Hinton. A Monte Carlo template based analysis for air-Cherenkov arrays. *Astroparticle Physics*, 56:26–34, April 2014. doi: 10.1016/j.astropartphys.2014.03.002.

F. Piron. *Etudes des propriétés spectrales et de la variabilité de l'émission gamma supérieure à 250 GeV des noyaux actifs de galaxies de type blazar observés dans le cadre de l'expérience CAT*. PhD thesis, Ecole Polytechnique, 2000.

W. A. Rolke, A. M. López, et J. Conrad. Limits and confidence intervals in the presence of nuisance parameters. *Nuclear Instruments and Methods in Physics Research A*, 551:493–503, October 2005. doi: 10.1016/j.nima.2005.05.068.

Chapitre 6

Préparation de l'analyse en mode hybride à base de simulations

Sommaire

6.1 Introduction 103
6.2 Simulations 103
6.3 Reconstruction des évènements γ avec CT_5 104
6.4 Modifications/Améliorations dans APC-MVA 107
6.5 Construction des analyses préliminaires 115
6.6 Estimation des performances préliminaires 125
Bibliographie 132

6.1 Introduction

En septembre 2012 le nouveau télescope (section 4.2.3 du chapitre 4) a été inauguré. À partir de cette date, une phase relativement longue de mise en fonctionnement et de compréhension du nouvel instrument a débuté, lors de laquelle un réglage fin des paramètres de CT_5 a été entrepris. En octobre 2013, le développement des premières chaînes de calibration incluant le cinquième télescope a permis l'analyse préliminaire des données obtenues après mars 2013.

Durant cette phase de mise en fonctionnement, une estimation des performances préliminaires du réseau de télescopes HESS-II en mode hybride a été entreprise à l'aide de simulations de γ pour le signal et des simulations de hadrons et d'électrons pour le fond. Ce travail indispensable a permis d'une part de préparer le développement d'une analyse robuste incluant CT_5 dans le réseau dès lors qu'un nombre conséquent de données réelles calibrées serait disponible, mais aussi d'obtenir des sensibilités préliminaires de l'instrument.

La section 6.2 présente les différentes simulations utilisées tout au long de ce chapitre pour construire une analyse préliminaire. Vient ensuite à la section 6.3 une discussion sur ce qu'apporte CT_5 en termes de reconstruction pour différentes classes d'évènements. Les améliorations ou modifications de l'analyse à quatre télescopes Paris-MVA, qui conduiront à la nouvelle appellation APC-MVA pour l'analyse en mode hybride, sont présentées à la section 6.4. La construction de l'analyse en elle-même est détaillée à la section 6.5 et l'estimation des performances préliminaires est présentée à la section 6.6.

6.2 Simulations

Pour déterminer les performances du réseau HESS-II de façon anticipée, des simulations massives de γ, protons et électrons ont été utilisées. Celles-ci sont nécessaires pour sonder les fluctuations dans le développement des gerbes atmosphériques. Nous avons utilisé les simulations mises à disposition par la partie allemande de la collaboration *H.E.S.S.* car elle proposait la statistique la plus fournie. Ces simulations ont été générées avec le logiciel *COsmic Ray SImulations for KAscade* (CORSIKA) [1] [Heck et al., 1998], qui a été utilisé par plusieurs générations d'expériences d'astronomie Tcherenkov.

1. https://web.ikp.kit.edu/corsika/

	γ	Protons	Électrons
N_{Sim}	5×10^8	2×10^8	$3{,}7 \times 10^9$
Indice spectral	-2	-2	-2
Angle zénithal (°)	20	20	20
Offset (°)	0,5	Diffus	Diffus
θ_Ω(°)	0	5	5
R_{max}(m)	600	1000	1000
Énergie (TeV)	0,005–10	0,01–50	0,005–20
NSB (MHz)	100		

TABLE 6.1: Caractéristiques des simulations utilisées.

L'estimation avec CORSIKA de la production de lumière des gerbes initiées par des photons γ et des électrons est très bien contrainte et les comparaisons avec un générateur indépendant (KASCADE [Kertzman et Sembroski, 1994] [2]) montrent un très bon accord. Pour la simulation des gerbes initiées par des hadrons, les systématiques liées à la production de lumière Tcherenkov en fonction des modèles d'interaction choisis on été récemment étudiés dans le cadre de la préparation de *CTA* [Bernlöhr et al., 2013]. Ce travail montre que les systématiques liées à la production de lumière par les gerbes hadroniques du fait du choix des différents modèles sont inférieures à 5 %.

La simulation des télescopes a été réalisée avec le logiciel sim_hessarray [3]. Ce dernier est largement utilisé et montre un accord remarquable avec le logiciel SMASH [Guy, 2003], lui aussi utilisé par la collaboration *H.E.S.S.*. L'ensemble de la chaîne de simulation (gerbes avec CORSIKA et détecteur avec sim_hessarray) a été validé avec la phase I de *H.E.S.S.*, en particulier en comparant les données simulées avec le signal extrait en provenance de sources particulièrement fortes (la nébuleuse du Crabe, les sursauts de certains noyaux actifs de galaxies). Cependant, notons que la simulation du télescope CT_5 a été réalisée avant la mise en fonctionnement de ce dernier. Bien qu'elle reste à valider, les premiers résultats obtenus en mode mono sur le pulsar de Véla (plus de 10σ, avec une estimation du fond très bien maîtrisée) montrent un bon accord entre simulations et données réelles.

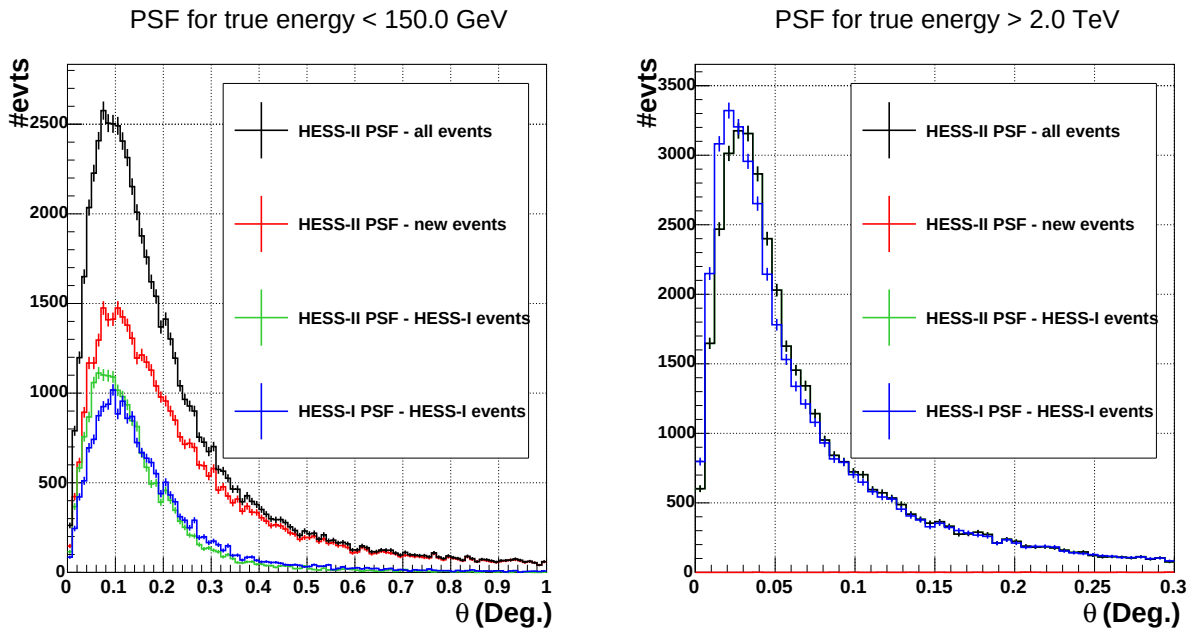
Toutes les simulations de gerbes ont été générées selon une loi de puissance d'indice -2 et un angle zénithal de 20°. Les photons ont été tirés suivant une direction de provenance unique, correspondant à l'émission d'une source ponctuelle (à un offset unique de 0,5°), tandis que les électrons et les protons ont été générés de manière diffuse avec un angle de tirage θ_Ω de 5°. Quant au pied de gerbe maximal des évènements, il est de 600 m pour les γ et de 1 km pour les électrons et les hadrons. Le NSB est fixé à 100 MHz, une valeur moyenne des taux mesurés en dehors du plan galactique. L'efficacité optique des quatre télescopes de type HESS-I a été fixée à 70 % de sa valeur nominale et celle du cinquième télescope à 80 %. Les caractéristiques des simulations sont résumées dans la Table 6.1.

6.3 Reconstruction des évènements γ en mode hybride

6.3.1 Les évènements γ vus par le réseau

L'ajout d'un nouveau télescope possédant une surface de miroir bien plus importante (un seuil de déclenchement escompté de ∼10–20 GeV) et un champ de vue réduit en comparaison des quatre autres télescopes implique l'apparition d'une nouvelle population d'évènements vus par le réseau. Si on considère les évènements qui déclenchent le réseau de télescopes avec une multiplicité minimale de deux, on peut définir les classes d'évènements suivantes :

2. On reparlera de KASCADE au chapitre 7.
3. http://www.mpi-hd.mpg.de/hfm/%7bernlohr/HESS/Software/sim_hessarray/



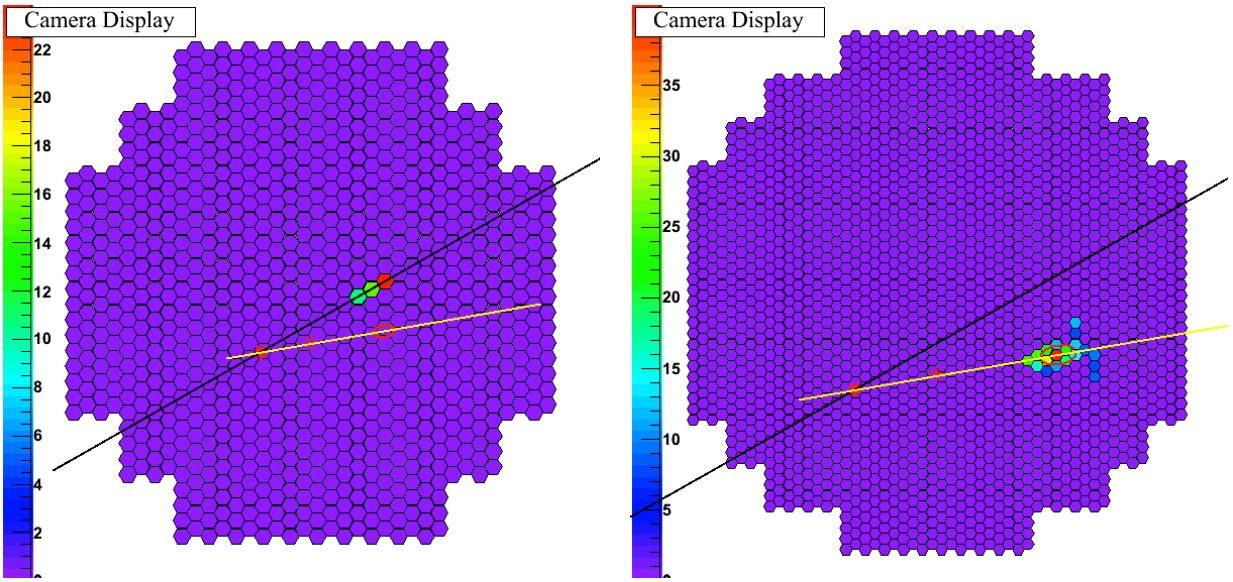
(a) (b)

FIGURE 6.1: Précision angulaire obtenue avec la méthode des ellipses pour différents lots d'évènements vus par le réseau de 5 télescopes, ayant une énergie vraie (a) inférieure à 150 GeV et (b) supérieure à 2 TeV. Les distributions verte et rouge représentent les deux populations définies dans le texte : les évènements déjà observés par le réseau HESS-I et les évènements de multiplicité deux ayant déclenché CT_5. La distribution noire représente la somme de ces deux lots d'évènements. Quant à la distribution bleue, elle représente la précision angulaire obtenue avec les évènements déjà détectés en mode stéréo pour lesquelles on a supprimé CT_5 de la reconstruction. Sur la sous-figure (b), les évènements de multiplicité deux ayant déclenché CT_5 représentent une fraction négligeable (<0,1 %) de l'ensemble des évènements vus par le réseau (la courbe verte et la courbe noire sont quasiment confondues).

— celle composée par les évènements qui auraient été détectés par le réseau HESS-I, et pour lesquels on obtient dans la grande majorité des cas une image supplémentaire dans CT_5. Ces évènements sont de multiplicité au moins égale à trois 3 lorsque le cinquième télescope déclenche.

— celle composée par les évènements qui ne déclenchent que CT_5 et un télescope de type HESS-I. Il s'agit de nouveaux évènements, qui n'auraient pas été détectés en mode stéréo sans la présence de CT_5.

Les Figures 6.1a et 6.1b représentent les distributions de la précision angulaire obtenue avec des simulations de γ et avec la méthode des ellipses, respectivement pour des évènements de basse énergie ($\leq$150 GeV) et de haute énergie ($\geq$2 TeV) pour les deux classes d'évènements définies ci-dessus. Sur chacune des figures, la distribution de la précision angulaire obtenue avec HESS-II et associée aux évènements déjà observés avec HESS-I est en vert, celle associée aux nouveaux évènements de multiplicité 2 est en rouge, et enfin celle correspondant à l'ensemble des évènements (somme des deux précédentes) est en noir. À titre de comparaison, la fonction d'étalement du point (PSF) de la première population, correspondant aux évènements déjà détectés en stéréo, et calculée sans CT_5, est représentée en bleu. Les coupures de pré-sélection (section 6.4.1) appliquées pour obtenir ces distributions consistent à sélectionner des images qui possèdent, après « nettoyage 57 », une charge supérieure à 40 p.e., un nombre de pixels au moins égal à trois et une distance nominale inférieure à 80 % du rayon des caméras.

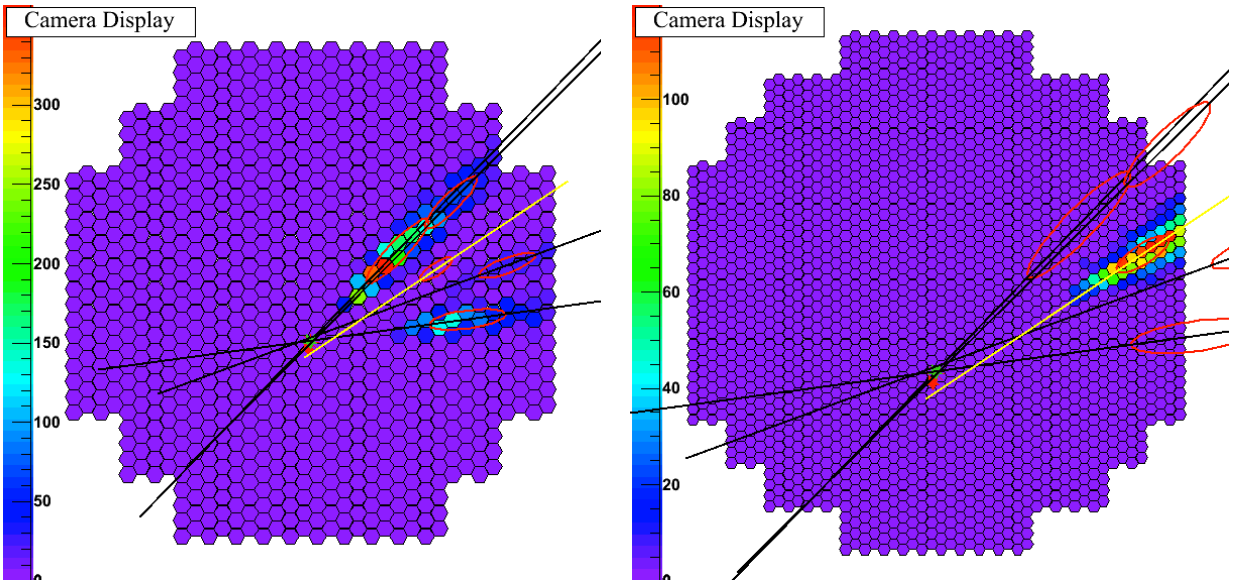


(a) (b)

FIGURE 6.2: Évènement γ de 100 GeV vu par CT5 (b) et un télescope de type HESS-I (a).

À très basse énergie, la population de γ est majoritairement constituée (à $\sim$60 %) d'évènements de multiplicité 2 qui déclenchent CT_5 et un télescope de type HESS-I. On peut en trouver un exemple sur la Figure 6.2. À gauche, la Figure 6.2a représente une caméra de type HESS-I dans laquelle les intensités des charges dans les pixels pour les images passant la pré-sélection sont représentées ainsi que la paramétrisation de Hillas associée (axe principal en noir, l'ellipse est confondue avec le trait noir). L'ellipse de CT_5 (en rouge) et son axe principal (en jaune) sont superposés dans cette même caméra de type HESS-I. À droite, on trouve la cinquième caméra avec pour ce même évènement les intensités des charges dans ses pixels pour l'image passant la pré-sélection ainsi que son ellipse de Hillas et son axe principal. Comme précédemment, sont aussi superposés l'axe principal et l'ellipse de l'image de la caméra de type HESS-I. L'astérisque et l'étoile rouges représentent respectivement la position de la direction vraie et celle de la direction reconstruite. Du fait de la différence en surface entre le grand et les télescopes de taille moyenne, les évènements de ce type possèdent une information conséquente dans CT_5 mais très peu de pixels actifs dans un autre télescope ce qui entraîne un élargissement de la distribution de la PSF (comparer les distributions verte et rouge de la Figure 6.1a). On constate que CT_5 permet à très basse énergie d'améliorer la reconstruction angulaire des évènements déjà détectés et que l'émergence d'une nouvelle classe d'évènements à deux télescopes, entraîne un élargissement de la PSF totale (distribution noire sur la Figure 6.1a). Il en résultera également une augmentation sensible de la surface efficace de collection.

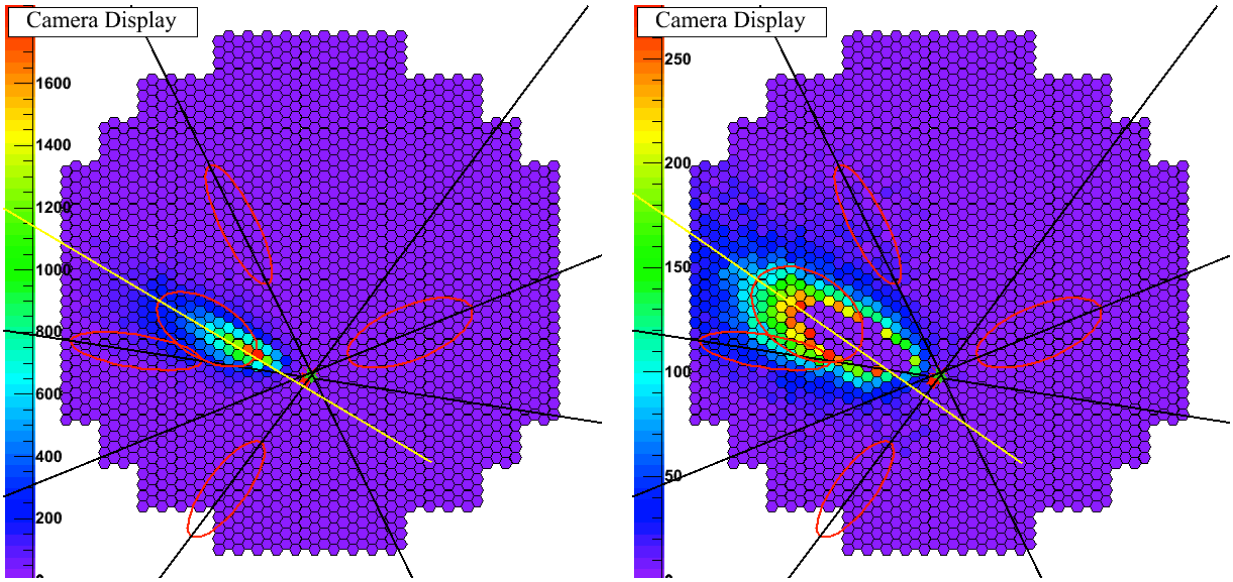
À plus haute énergie (Figure 6.1b), on constate tout d'abord que la nouvelle population d'évènements (distribution rouge) représente une fraction négligeable de l'ensemble des évènements vu par le réseau ($< 0{,}1$ %). En effet, il n'est possible d'avoir des évènements de multiplicité deux avec des gerbes de très haute énergie que si leur pied de gerbe est lointain. Or, du fait du champ de vue réduit de la cinquième caméra, l'image d'une telle gerbe dans celle-ci est fortement coupée, ce qui conduit à un barycentre des charges proche du bord de la caméra. Une telle image ne sera pas présélectionnée pour la reconstruction de la géométrie de la gerbe (section 6.4.1). La très grande majorité des évènements auraient donc été détectés par le réseau HESS-I. On constate également sur cette même figure, une légère dégradation de la PSF lorsque CT_5 participe à la reconstruction de la géométrie de la gerbe (comparer les distributions noire et bleue sur la Figure 6.1b). Afin d'en comprendre la raison, nous avons inspecté les images d'un grand nombre d'évènements participant à cette dégradation. Il résulte de cette étude que la principale raison de cette dégradation est la



(a) (b)

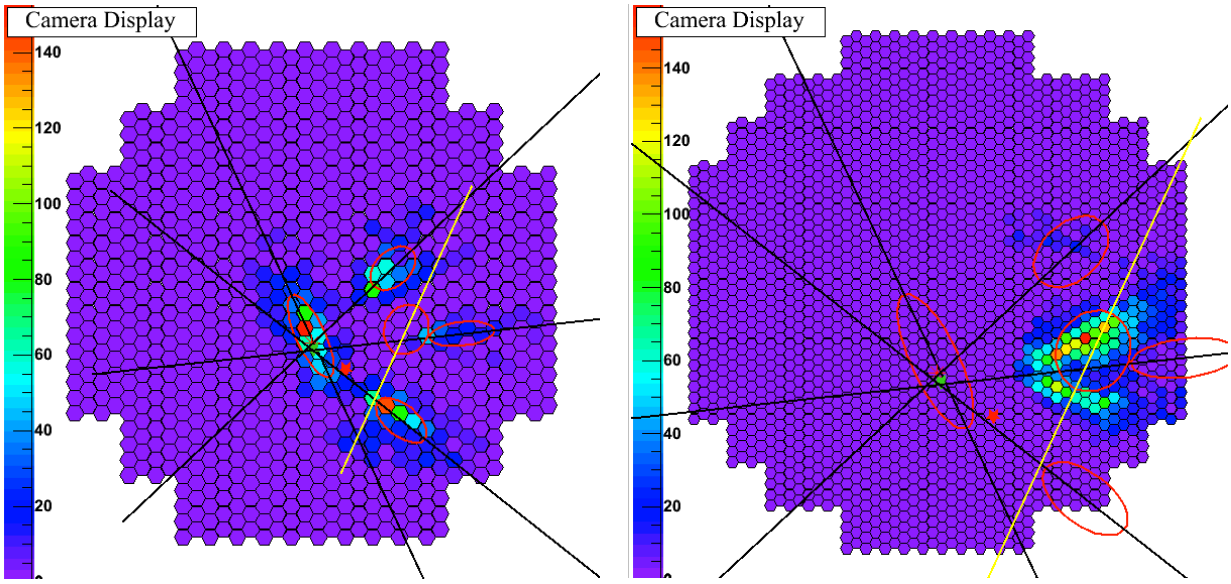
FIGURE 6.3: Évènement γ de 4,61 TeV. Illustration d'une image tronquée dans CT_5 conduisant à une dégradation de la PSF.

troncation des images de manière asymétrique dans la caméra de CT_5. À titre d'exemple, on peut observer sur la Figure 6.3 les images associées à un évènement γ, de ∼4,6 TeV et de multiplicité cinq, projetées dans une caméra de type HESS-I (Figure 6.3a) et dans la cinquième caméra (Figure 6.3b). La même représentation déjà utilisée sur la Figure 6.2 est appliquée et on a ajouté la position de la direction reconstruite sans CT_5 qui est représentée par une étoile verte. Lorsque l'image est tronquée de manière asymétrique par un bord de la caméra, la détermination de l'axe principal de l'ellipse est biaisée, ce qui conduit à une mauvaise estimation de la direction du γ. Un second exemple est donné sur la Figure 6.4 pour un évènement γ de 7,27 TeV. À gauche, la Figure 6.4a représente la cinquième caméra ainsi que l'intensité des charges de ses pixels, l'axe principal et l'ellipse associés. On y trouvera aussi les ellipses et les axes principaux des images observées dans les télescopes moyens. À droite, on a cette fois représenté l'intensité des pixels de CT_5 qui possèdent au maximum 15 % de la charge du pixel le plus chaud ainsi que l'ellipse et l'axe principal déterminés à partir de ces charges. On constate la présence d'un halo (queue de la distribution des charges) qui est tronqué de manière asymétrique et qui conduit, comme précédemment, à une mauvaise estimation de l'axe principal. De manière plus anecdotique, du fait de la grande surface de collection de CT_5 il est possible d'observer des détails fins de la gerbe, comme par exemple des sous-gerbes dues à des fluctuations du développement de la gerbe au niveau des premières interactions, comme illustré pour un évènement γ de 1,12 TeV sur la Figure 6.5. On observe dès le début du développement de la gerbe (Figure 6.5b) une sous-gerbe qui s'écarte de la gerbe principale et qui conduit à une image qui ne présente l'habituelle symétrie axiale des gerbes leptoniques. On constate que cette sous-gerbe est à peine visible dans télescopes moyens et n'affecte pas la reconstruction de la direction de l'évènement. Pour finir, un dernier exemple correspondant à un évènement γ de 6 TeV rendant compte de la grande sensibilité de CT_5 est illustré sur la Figure 6.6 (même représentation que sur la Figure 6.4) qui révèle la présence d'un muon dans l'image de CT_5 lorsque l'on ne garde que les pixels ayant au maximum 15 % de la charge du pixel le plus intense (l'image est aussi tronquée de manière asymétrique).



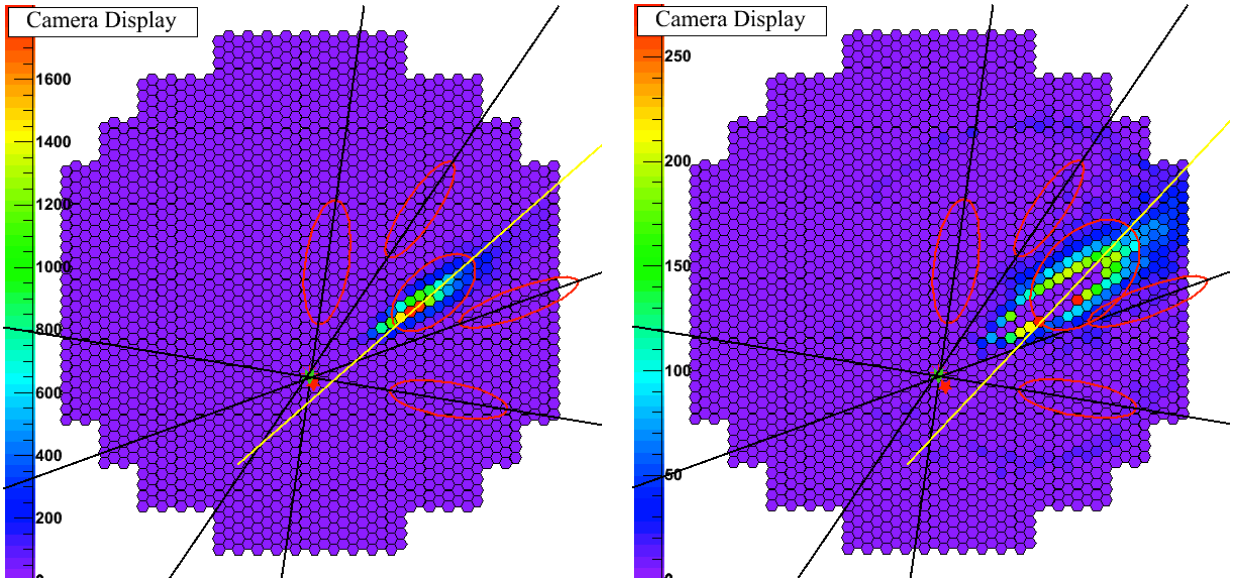
(a) (b)

FIGURE 6.4: Évènement γ de 7,27 TeV. Illustration de la présence d'un halo dans une image de CT_5 conduisant à une dégradation de la PSF. L'image de droite contient les pixels ayant au plus 15 % de la charge du pixel le plus chaud dans l'image entière.



(a) (b)

FIGURE 6.5: Évènement γ de 1,12 TeV présentant une fluctuation dans le développement de la gerbe observée par CT5 mais qui n'est pas accessible aux caméras de type HESS-I.



(a) (b)

FIGURE 6.6: Évènement γ de 6,02 TeV. Illustration de la grande sensibilité de CT_5 conduisant à une dégradation de la PSF. L'image de droite contient les pixels ayant au plus 15 % de la charge du pixel le plus chaud dans l'image entière et révèle la présence d'un muon qui n'est pas détecté par les autres télescopes.

6.4 Modifications/Améliorations dans APC-MVA

Un certain nombre de modifications ou améliorations ont été apportées à l'analyse Paris-MVA pour prendre en charge l'arrivée du cinquième télescope. Ces modifications ont été appliquées pour rendre le calcul de certains paramètres discriminants plus cohérent avec l'ensemble de la chaîne d'analyse, mais aussi pour résoudre le problème de la dégradation de la PSF à haute énergie liée à l'ajout de CT_5, et par la même occasion pour améliorer la précision angulaire à toutes les énergies. Avant de décrire ces modifications, nous rappelons brièvement la définition des coupures de pré-sélection qui sont appliquées en début d'analyse.

6.4.1 Pré-sélection des images mesurées

La pré-sélection permet d'exclure de l'analyse, si nécessaire, certaines images afin d'améliorer la reconstruction géométrique de la gerbe avec la méthode des ellipses. Pour qu'une image participe à la reconstruction, elle doit valider trois critères après nettoyage :

— avoir une charge minimale supérieure à $Q^{\min}$

— avoir un barycentre des charges situé en deçà d'un distance $L_{\mathrm{D}}^{\max}$ par rapport au centre de la caméra

— avoir un nombre de pixels supérieur à $N_{\mathrm{pix}}^{\min}$

Le premier critère permet de sélectionner les images avec une charge suffisante pour permettre une bonne détermination des paramètres de Hillas. Il permet aussi de se prémunir contre une contamination trop forte des charges mesurées par le bruit de fond du ciel. La coupure en charge diffère selon le type de source étudié. Pour les sources ayant un indice très mou, pour lesquelles il est important d'être efficace à basse énergie, on acceptera dans l'analyse des images ayant une charge totale de 40 p.e., ce qui aura pour effet de baisser le seuil en énergie. Dans le cas des sources ayant un spectre dur on utilisera un seuil à 60 ou à 80 p.e. afin de prendre en compte l'existence d'une

population plus importante à haute énergie [4].

Le second critère $L_{\mathrm{D}}^{\mathrm{max}}$ permet d'exclure les images fortement tronquées par le bord du champ de vue d'une caméra. Pour les caméras de type HESS-I, une coupure de 0,525 m était utilisée qui correspond à 80 % de son rayon. On utilisera le même principe pour les images dans CT_5, en ne gardant que celles dont le barycentre est au maximum à 0,80 m du centre de la caméra (80 % du rayon).

Enfin, le troisième critère permet de sélectionner les images possédant au minimum 4 pixels après nettoyage. Il s'agit d'un compromis entre la maximisation de la surface de collection (garder le plus de télescopes possibles afin de récupérer des évènements de multiplicité deux) et l'optimisation de la précision angulaire.

Après ce niveau de pré-sélection, chaque évènement aura sa propre liste de télescopes avec laquelle on reconstruira la géométrie avec la méthode des ellipses et avec laquelle on calculera les paramètres discriminants.

6.4.2 Pré-sélection des images projetées

Le modèle 3D permet d'ajuster une photosphère représentant l'émission de la gerbe qui permet, par projection sur les télescopes, d'obtenir des images prédites [5]. La comparaison de ces dernières avec les images mesurées permet, grâce à l'exploitation de leurs différences, d'augmenter le pouvoir de séparation γ/fond. C'est sur ces différences qu'on été construites les paramètres discriminants Ω_{57} et R_{E} (section 5.2.3 du chapitre 5).

Une modification par rapport à la procédure présentée dans [Becherini et al., 2011] consiste à appliquer pour les images projetées une procédure analogue à la pré-sélection des images mesurées. Un nouveau niveau de pré-sélection a été introduit, la pré-sélection « modèle 3D », afin d'une part d'avoir une cohérence entre le traitement des images expérimentales et le traitement des images prédites, et d'autre part pour un meilleur contrôle du calcul des paramètres discriminants (voir sous-section suivante). De la même manière que pour la pré-sélection des images mesurées, chaque évènement aura sa propre liste de télescopes contenant des images prédites présélectionnées.

Pour illustrer cette démarche, à gauche de la Figure 6.7a sont représentées, pour un proton de 4,0 TeV, les images mesurées passant la pré-sélection et projetées dans une caméra unique de type HESS-I, et à droite les prédictions obtenues à partie de la projection de la photosphère et passant la pré-sélection « modèle 3D ». Les étoiles rouge, verte et bleue représentent respectivement les directions reconstruites avec les images mesurées passant la pré-sélection, les images projetées à partir du modèle et passant la pré-sélection « modèle 3D » et la direction provenant directement de l'ajustement du modèle 3D. L'astérisque rouge représente la direction vraie du proton (on ne s'attend pas à la reconstruire correctement). De même, la Figure 6.7b représente ce même évènement mais cette fois représenté dans le champ de vue de la cinquième caméra. On constate une différence entre les directions obtenues à partir des images mesurées et des images prédites, ainsi qu'un nombre de télescopes différent (3 pour la pré-sélection et 5 pour la pré-sélection modèle 3D dans ce cas particulier). Ce type de différences sera exploité pour la discrimination γ/fond.

6.4.3 Modification du calcul des paramètres discriminants

L'ajout du niveau de pré-sélection dit « modèle 3D » va permettre de rendre plus cohérent le calcul de certains paramètres discriminants, notamment pour les deux variables R_{E} et Ω_{57} qui exploitent les différences entre les prédictions du modèle 3D et les données expérimentales. En effet, l'analyse Paris-MVA utilisait des images mesurées présélectionnées pour le calcul de ces paramètres

4. Il s'agit généralement de sources galactiques, qu'il est éventuellement possible de résoudre, et pour lesquelles une amélioration de la précision angulaire peut être obtenue avec un Q^{min} grand (entraînant une augmentation du seuil en énergie).

5. Les télescopes actifs lors d'une prise de données dont les images ne passent pas la pré-sélection mais pour lesquelles le modèle 3D prédit un signal significatif sont appelés les « dark-télescopes ».

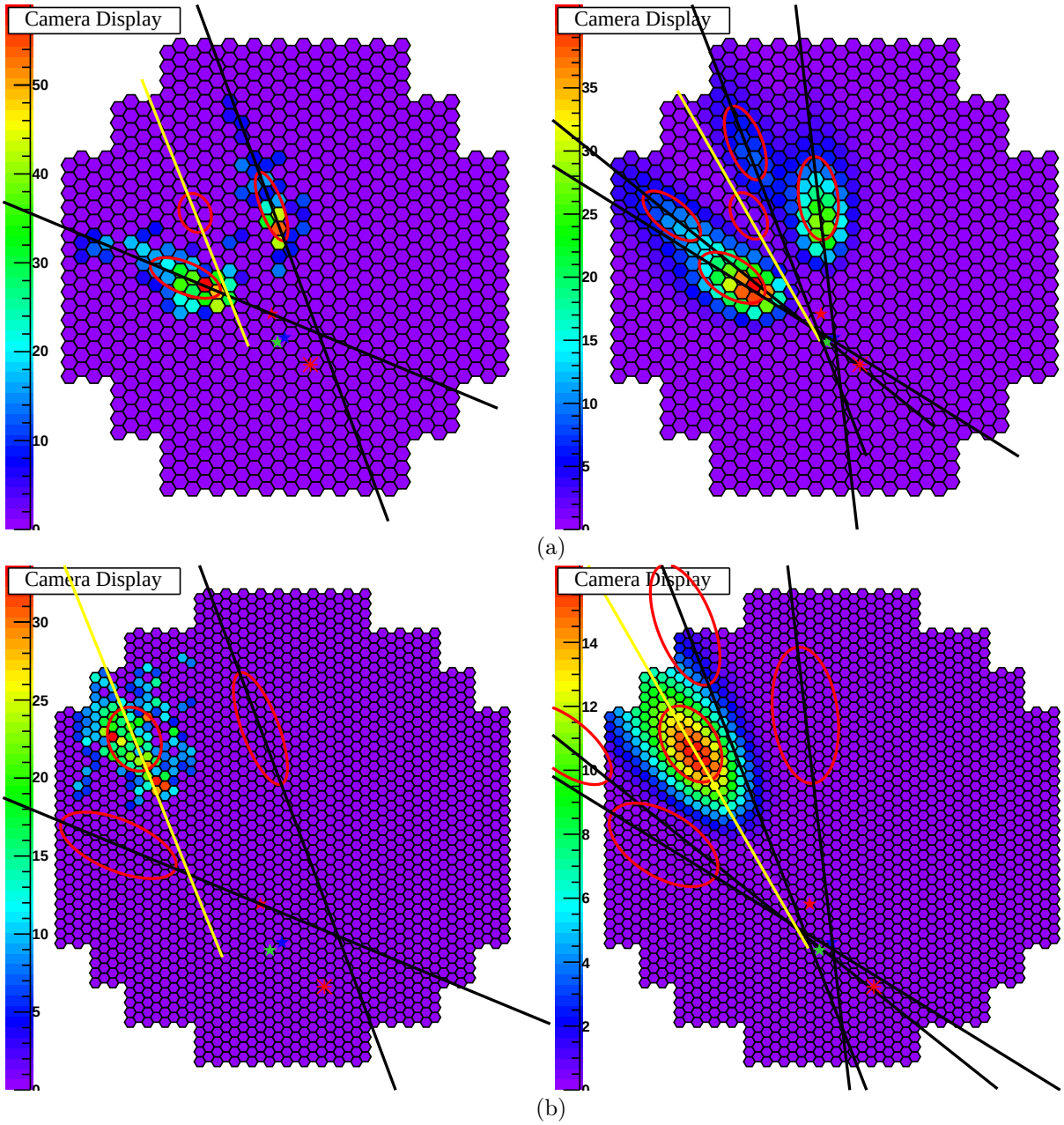


FIGURE 6.7: Illustration de la pré-sélection « modèle 3D » avec un proton simulé de 4,0 TeV. (a) À gauche, images mesurées passant la pré-sélection et projetées dans une caméra unique de type HESS-I. À droite, prédictions obtenues à partie de la projection de la photosphère et passant la pré-sélection « modèle 3D ». Les étoiles rouge, verte et bleue représentent respectivement les directions reconstruites avec les images mesurées passant la pré-sélection, les images projetées à partir du modèle et passant la pré-sélection « modèle 3D » et la direction provenant directement de l'ajustement du modèle 3D. L'astérisque rouge représente la direction vraie du proton. (b) Même chose que précédemment mais cette fois dans le référentiel de la cinquième caméra.

discriminants tandis qu'aucune pré-sélection n'était effectuée sur les images prédites à partir du modèle 3D. Les tables en énergie étant générées avec un niveau de pré-sélection, il était plus cohérent d'en faire de même sur les images prédites. De la même manière, la direction reconstruite avec les paramètres de Hillas à partir des prédictions du modèle 3D, utilisée pour le calcul de Ω_{57}, sera estimée en utilisant la liste de télescopes qui passent la pré-sélection « modèle 3D ». Cela permet d'éliminer de la reconstruction de la direction, les images avec très peu de charges ou en bord de caméra qui induisent une détermination biaisée de l'axe principal et donc de la direction.

Le dernier paramètre qui est calculé de manière différente est ΔQ (section 5.2.3 du chapitre 5), qui mesure la dispersion entre les charges mesurées et les charges déduites, en inversant les tables de calibration en énergie, à partir de l'énergie mesurée et des paramètres de la gerbe (pied de gerbe par rapport à chaque télescope et hauteur du maximum de gerbe). Pour estimer ce paramètre, seuls les télescopes sélectionnés au niveau de la pré-sélection (des images mesurées) seront utilisés. Ce calcul était au préalable réalisé avec tous les télescopes actifs lors du run.

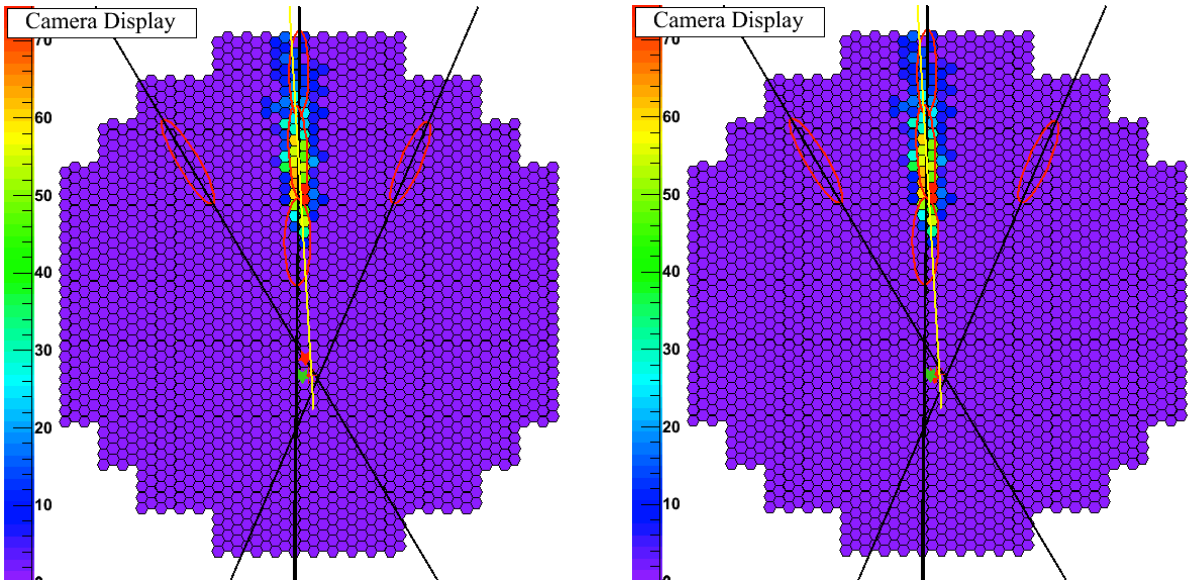
6.4.4 Amélioration de la PSF avec la sélection $\sin\theta_{ij}$

Avec l'ajout de CT_5, un élargissement global de la PSF a été constaté (section 6.3 ; Figure 6.1). Après une étude au cas par cas d'évènements problématiques, pour lesquels une reconstruction avec CT_5 dégrade la précision angulaire, il s'est avéré que lors de la phase de reconstruction de la direction, lorsque plusieurs intersections sont en jeu et que CT_5 est présent au niveau de la pré-sélection, les charges importantes des images de la cinquième caméra (de l'ordre d'un facteur 6 à pied de gerbe égal du fait de la différence de la taille des réflecteurs) biaisent le rôle de la relation explicitée au chapitre 5 à la section 5.1.1 : $\omega^p = \omega_e^p \omega_Q^p \omega_\theta^p$. Cette relation pondère chaque paire p de télescopes participant à la reconstruction géométrique de la gerbe, de telle manière que celles formant un petit angle – qui auront tendance à détériorer la PSF – auront une influence moindre dans la reconstruction de la direction du fait de sa dépendance en $\sin\theta$. Cependant, pour les paires qui impliquent CT_5, la grande surface de collection de ce télescope conduit à des charges généralement plus importantes qui vont augmenter l'importance dans la reconstruction des paires dont les axes forment de petits angles.

À titre d'exemple, on peut observer sur la Figure 6.8a un évènement vu par 5 télescopes dont les paramètres de Hillas (le grand axe de l'ellipse de CT_5 est en jaune) de tous les télescopes sont projetés dans la caméra de CT_5. Les étoiles rouge et verte désignent respectivement la direction reconstruite avec et sans CT_5, et l'astérisque rouge la direction vraie du rayon γ. Bien que le grand axe de CT_5 pointe vers la direction vraie de l'évènement, la reconstruction de la direction en mode hybride, donc en incluant CT_5, est moins bonne qu'en mode stéréo. En ajoutant CT_5 dans la reconstruction, on passe d'une distance entre les directions vraie et reconstruite de 0,04° à 0,10°. Ceci est dû à la présence de trois images qui sont quasiment alignées, donc aux grandes erreurs associées aux positions des intersections de leurs axes, et à la sur-pondération de certaines paires du fait de la présence d'une charge dans CT_5.

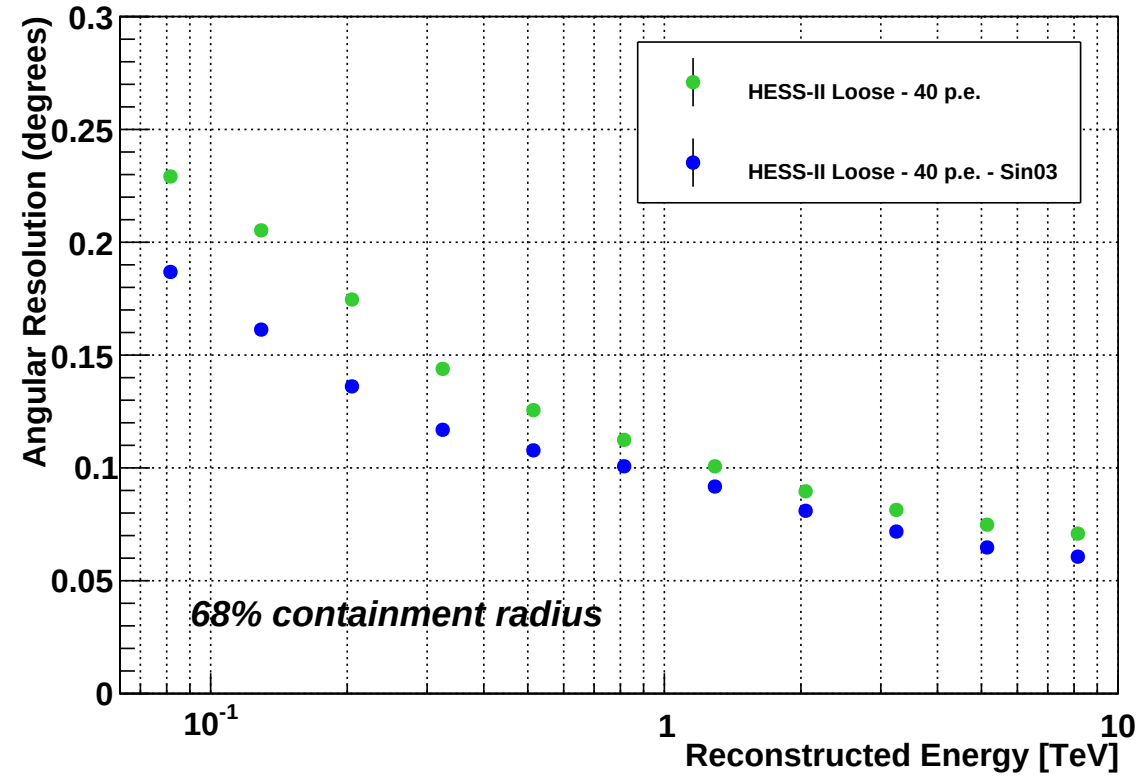
Nous proposons de remédier à ce problème en mettant en place une coupure, que l'on appellera $\sin\theta_{ij}$, afin de rejeter dans la reconstruction de la direction de la gerbe les paires de télescopes dont l'angle entre les axes des images est plus petit qu'une valeur θ_{ij}, tout en préservant CT_5 dans la plupart des cas. L'étude a été menée sur un lot de γ simulés en appliquant des coupures de sélection finale correspondant à une configuration « *loose* »[6]. Le résultat de notre étude nous a conduit à choisir une valeur de $\sin\theta_{ij} = 0{,}3$ qui correspond à un compromis entre une amélioration de la précision angulaire et une perte acceptable d'évènements (de multiplicité deux). L'effet de la coupure $\sin\theta_{ij}$ est représenté sur la Figure 6.8b, pour l'évènement problématique précédemment présenté sur la Figure 6.8a. En appliquant celle-ci, on gagne un facteur cinq en précision angulaire, et un facteur deux lorsque l'on exclut CT_5 de la reconstruction.

6. Cette configuration sélectionne les images qui ont une charge totale supérieure ou égale à 40 p.e. et applique des coupures de discrimination sur les paramètre MSCL et MSCW (voir les coupures « *spectrum* » dans [Aharonian et al., 2006]).

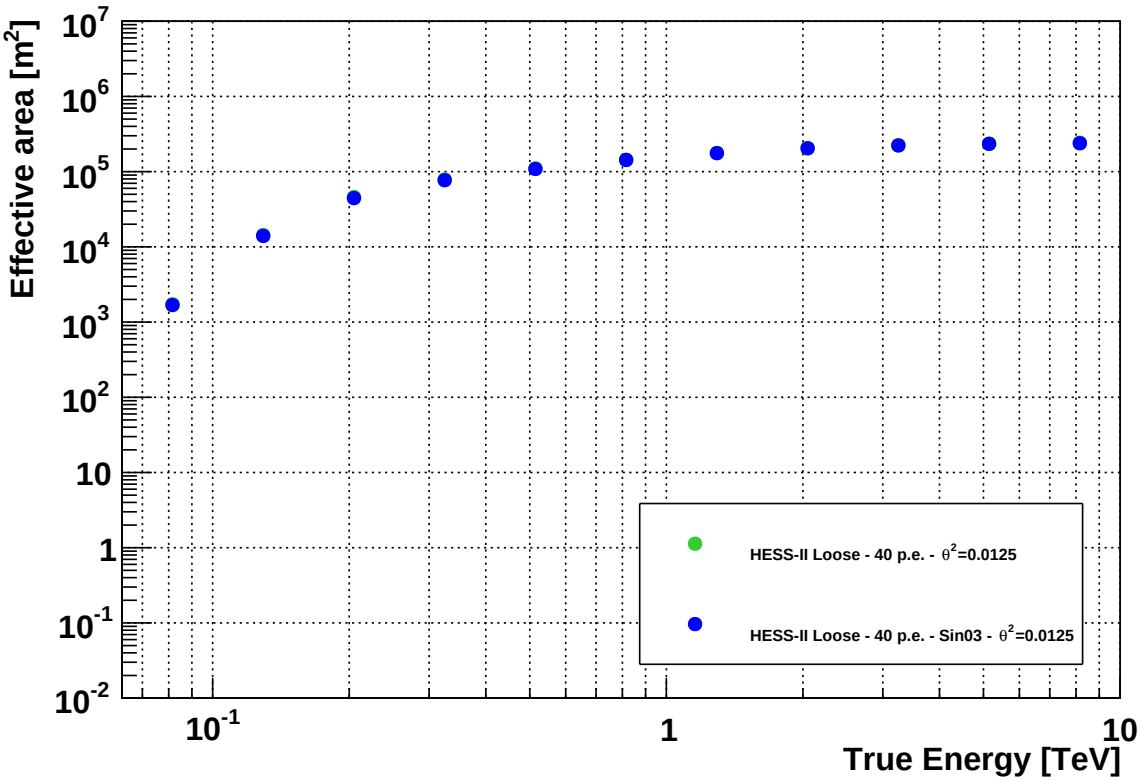


(a) (b)

FIGURE 6.8: Illustration de la coupure $\sin\theta_{ij}$ avec un évènement γ de 803 GeV pour lequel trois images sont alignées. (a) Ellipses de Hillas projetées dans la cinquième caméra et intensité de l'image de CT_5. Les étoiles rouge et verte désignent respectivement la direction reconstruite avec et sans CT_5. L'astérisque rouge représente la direction vraie. (b) Même chose mais avec la coupure $\sin\theta_{ij} = 0{,}3$, permettant de gagner pour cet évènement un facteur 5 en précision angulaire lorsque CT_5 est présent, et un facteur 2 lorsque CT_5 ne participe pas à la reconstruction. Les intersections des trois images alignées ne sont pas prises en compte lorsque la coupure $\sin\theta_{ij} = 0{,}3$ est appliquée.



(a)



(b)

FIGURE 6.9: Mise en évidence de l'effet de la coupure $\sin\theta_{ij} = 0{,}3$. (a) Résolutions angulaires en fonction de l'énergie reconstruite et (b) surfaces efficaces en fonction de l'énergie vraie. Les points verts et bleus représentent respectivement une configuration qui n'utilise pas la coupure $\sin\theta_{ij}$ et une configuration où $\sin\theta_{ij} = 0{,}3$ est utilisée.

Les Figures 6.9a et 6.9b représentent les résolutions angulaires à 68 % en fonction de l'énergie reconstruite et la surface efficace en fonction de l'énergie vraie, pour deux configurations qui utilisent respectivement $\sin\theta_{ij} = 0$ (en vert, la coupure n'est pas appliquée) et $\sin\theta_{ij} = 0{,}3$ (en bleu). On constate que la coupure $\sin\theta_{ij}$ apporte un gain en résolution angulaire à toutes les énergies. Ce gain est de l'ordre de ∼20 % à basse énergie (typiquement en dessous de 200 GeV) et de ∼15 % au-delà de 1 TeV. Fait remarquable, la perte d'évènements γ est en moyenne inférieure à 3 % en deçà de ∼600 GeV et devient négligeable au-delà (inférieure à 1 %). Ces résultats nous ont conduit à inclure cette coupure dans l'analyse hybride présentée à la section suivante.

6.5 Construction des analyses préliminaires

Cette section présente la construction, à partir de simulations de γ, protons et électrons, d'une analyse de type « elm » avec laquelle sera estimée une sensibilité préliminaire pour le réseau HESS-II en mode hybride. Après une présentation générale des différentes étapes de l'analyse, nous décrivons la pondération qui sera appliquée aux électrons et aux hadrons, afin de prendre en compte les différences entre les spectres simulés et les spectre mesurés, lors de la phase d'entraînement. Ensuite, le découpage de l'espace des phases qui permet de séparer le traitement des évènements selon leurs caractéristiques, ainsi que les distributions des variables discriminantes seront discutés. La méthode d'optimisation des coupures de discrimination finale et les réglages de la méthode multivariée seront aussi présentés. Enfin, nous terminerons avec une discussion sur les résultats des entraînements. Pour estimer le gain de performances par rapport à une analyse en mode stéréo sans CT_5, nous développons une analyse stéréo en parallèle.

6.5.1 Préparation de l'analyse

Dans l'optique d'estimer des performances pour des sources faibles et extragalactiques, la pré-sélection des images simulées (voir section 6.4.1) est effectuée avec les valeurs suivantes pour les différentes coupures :

- $Q^{\text{min}} = 40$ p.e.
- $L_{\text{D}}^{\text{max}} = 0{,}8$ m pour CT_5 et $L_{\text{D}}^{\text{max}} = 0{,}525$ m pour les caméras de type HESS-I
- $N_{\text{pix}}^{\text{min}} = 4$
- $\sin\theta_{ij} = 0{,}3$ pour la configuration en mode hybride

Seuls les évènements pour lesquels l'ajustement du modèle 3D converge seront gardés. La même pré-sélection sera utilisée pour les images provenant de la projection dans le plan focal des télescopes de la photosphère ajustée par le modèle 3D.

La première étape consiste à générer des tables de longueurs et de largeurs qui permettront plus tard de construire les paramètres réduits (MSCL et MSCW). Ces tables sont générées avec un « nettoyage 57 » et seuls les évènements passant la pré-sélection et reconstruits dans une région circulaire autour de la position de la source simulée de 0,2° ont été gardés. La même procédure a été appliquée pour la génération des tables de calibration en énergie.

La chaîne d'analyse utilisée est composée de quatre étapes. Dans un premier temps, après avoir obtenu des intensités nettoyées et les paramètres de Hillas, la pré-sélection des images est appliquée pour reconstruire la géométrie de la gerbe avec la méthode des ellipses. Ensuite, les paramètres du modèle 3D sont ajustés en prenant comme point de départ pour la procédure de minimisation les paramètres de la gerbe précédemment obtenus. Après cela, les paramètres de Hillas sont calculés à partir des images prédites par le modèle 3D et la géométrie de la gerbe est reconstruite avec les images qui passent la pré-sélection « modèle 3D ».

Après avoir calculé les paramètres discriminants deux voies sont possibles. La première conduit au stockage des évènements pour les utiliser plus tard lors des phases d'entraînement et de test afin de déterminer les coupures de discrimination γ/fond. La seconde voie mène à l'analyse de données

réelles ou à l'estimation des performances, grâce à l'application des coupures de discrimination qui ont été préalablement obtenues. Lors de cette deuxième phase, une coupure angulaire de $\theta^2 = 0{,}0125$ est appliquée par rapport à la position d'observation.

Notons que pour la génération des tables et la phase d'entraînement, seule 50 % de la statistique disponible pour les γ, les hadrons et les électrons sera utilisée. Le reste de la statistique servira à la détermination des performances.

6.5.2 Pondération entre les électrons et les protons

Le fond est composé d'électrons et de protons simulés selon une loi de puissance d'indice -2. Il est nécessaire de pondérer chaque évènement pour rendre compte de la réalité des flux mesurés par diverses expériences. La pondération entre les hadrons et les électrons sera utilisée d'une part pour la visualisation des paramètres discriminants, et d'autre part lors des phases d'entraînement et de test.

Cette pondération pour la particule k ($k = \{e, p\}$) s'exprime comme :

$$\widetilde{w}_k(E) = \frac{\phi_k^{\mathrm{exp}}(E)}{\phi_k^{\mathrm{exp}}(1\,\mathrm{TeV})\left(\frac{E}{1\,\mathrm{TeV}}\right)^{\Gamma_k^{\mathrm{Sim}}}} \frac{T_0}{T_k^{\mathrm{Sim}}} \tag{6.1}$$

où T_0 est un temps qui peut être arbitrairement choisi, car seule la pondération relative est importante lors de la phase d'entraînement, ϕ_k^{exp} représente le flux différentiel expérimental et T_k^{Sim} est un temps d'observation virtuel qui permet de rendre compte du nombre réel d'évènements simulés. Chaque temps T_k^{Sim} est calculé de la manière suivante :

$$T_k^{\mathrm{Sim}} = \frac{N_k^{\mathrm{Sim}}}{\phi_k^{\mathrm{exp}}(1\,\mathrm{TeV})\mathcal{A}^k\Omega_k \int_{E_{\mathrm{min}}}^{E_{\mathrm{max}}} E^{\Gamma_k^{\mathrm{Sim}}}\mathrm{d}E} \tag{6.2}$$

où N_k^{Sim} et $\mathcal{A}^k$ représentent respectivement les nombre totaux d'évènements simulés entre les énergies E_{min} et E_{max}, et la surface sur laquelle les simulations ont été générées ($\mathcal{A}^k = \pi R_k^{\mathrm{Sim},2}$). On pourra trouver les valeurs des différents paramètres des simulations pour les protons et les électrons dans la Table 6.1.

Les flux expérimentaux pour les protons et les électrons qui ont été utilisés proviennent de [Bernlöhr et al., 2013] :

$$\phi_p^{\mathrm{exp}} = 0{,}096\left(\frac{E}{1\,\mathrm{TeV}}\right)^{-2{,}70}\,\mathrm{m}^{-2}\,\mathrm{s}^{-1}\,\mathrm{TeV}^{-1}\,\mathrm{sr}^{-1} \tag{6.3}$$

$$\begin{aligned}\phi_e^{\mathrm{exp}} = 6{,}85\times10^{-5}\left(\frac{E}{1\,\mathrm{TeV}}\right)^{-3{,}21} + \Bigg[&\frac{3{,}19\times10^{-3}}{7{,}76\times10^{-2}(E/1\,\mathrm{TeV})\sqrt{2\pi}} \\ &\times\exp\Big(-\frac{1}{2}\left(\log\frac{(E/1\,\mathrm{TeV})}{1{,}07\times10^{-2}}\right)^2(7{,}76\times10^{-2})^2\Big)\Bigg] \\ &\mathrm{m}^{-2}\,\mathrm{s}^{-1}\,\mathrm{TeV}^{-1}\,\mathrm{sr}^{-1}\end{aligned} \tag{6.4}$$

6.5.3 L'espace des phases de l'analyse

Pour augmenter le pouvoir de séparation γ/fond, le traitement d'un évènement sera différencié selon le nombre de télescopes qui passent les critères de pré-sélection et selon l'énergie reconstruite de la particule incidente. En lieu et place de l'énergie reconstruite nous utiliserons, pour le découpage de l'espace des phases, le nombre de photons Tcherenkov N_c composant la photosphère des gerbes reconstruites par le modèle 3D. De manière générale on découpe aussi l'espace des phases en bandes d'angle zénithal, mais ici l'étude se focalise sur un angle zénithal unique de 20°.

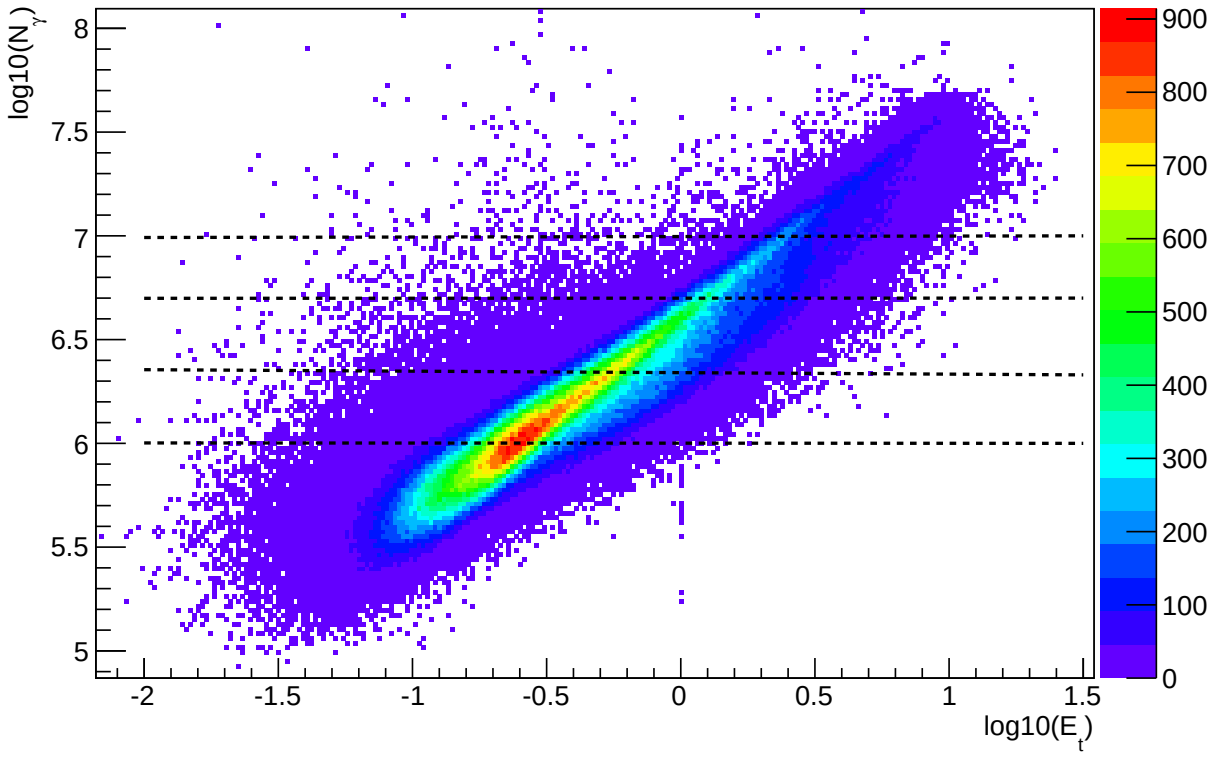


FIGURE 6.10: Nombre de photons Tcherenkov en fonction de l'énergie reconstruite pour des γ. Le découpage en N_c est représenté par les lignes horizontales.

Le découpage de l'espace des phases est guidé par la statistique disponible des données simulées après la phase de pré-sélection, en particulier des protons et des électrons, qui sont ici le facteur limitant. Afin d'obtenir des nombres d'évènements conséquents pour les phases d'entraînement et de test, les évènements de fond sont sélectionnés jusqu'à un offset de 1°, distance limite pour laquelle aucune variation significative des paramètres discriminants n'a été observée. Compte tenu de cela, nous avons découpé l'espace des phases en deux intervalles de multiplicité (deux et supérieure à deux) et cinq intervalles en N_c. On peut observer sur la Figure 6.10 le découpage choisi en N_c qui est représenté par des lignes horizontales. À basse énergie (premier intervalle en N_c), nous disposons de ∼8000 évènements de fond. Dans les autres intervalles, la statistique totale est de l'ordre de ∼5000 évènements de fond, à l'exception des deux intervalles de plus haute énergie et de faible multiplicité pour lesquels le nombre d'évènements de fond est plus faible, de l'ordre de 2000 évènements. Quant aux photons γ, leur nombre est très supérieur à celui du fond dans tous les intervalles (typiquement d'un facteur 10).

6.5.4 Les variables discriminantes

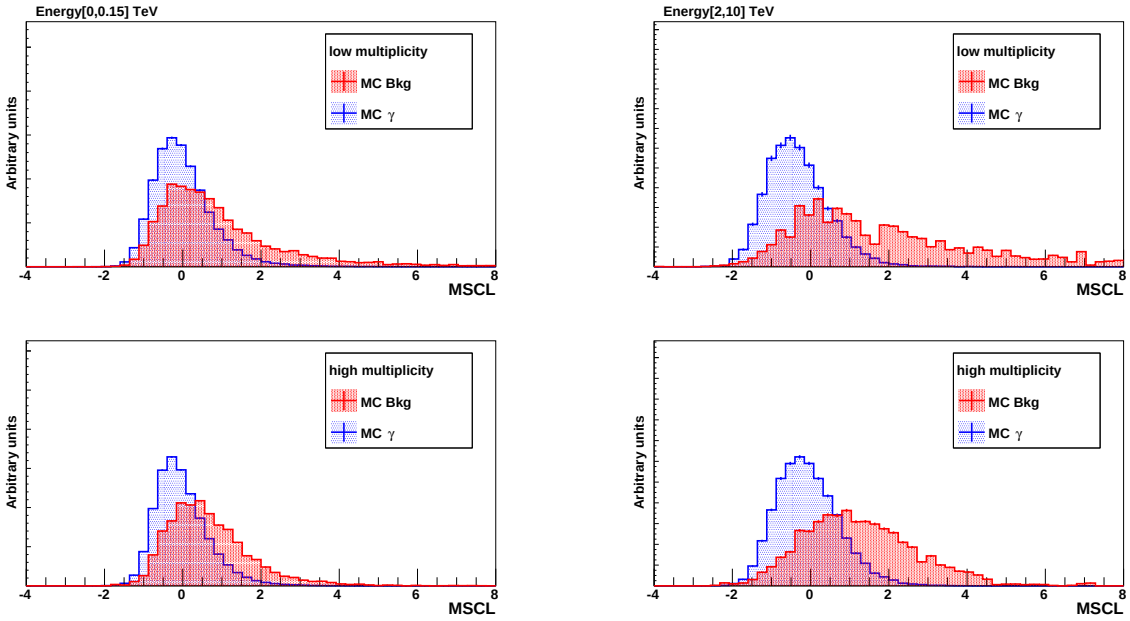
Les variables discriminantes utilisées dans cette étude en mode hybride ont été présentées au chapitre 5. Les distributions de ces huit variables sont représentées sur les Figures 6.11 à 6.18. Chacune d'entre elles est séparée en figures. Les sous-figures de gauche, labellisées (a), représentent la distribution d'un paramètre pour des évènements γ (bleu) et de fond (rouge) appartenant au premier intervalle en N_c qui représente, au premier ordre, une énergie reconstruite pour les γ inférieure à ∼150 GeV. Les sous-figures de droite, labellisées (b), représentent des évènements appartenant au dernier intervalle en N_c ce qui représente, au premier ordre, des γ ayant des énergies au-delà de 2 TeV. Dans les deux cas, les sous-figures du dessus représentent les évènements de multiplicité deux, et celles du dessous les évènements de multiplicité strictement supérieure à deux. Tous les évènements représentés ont passé la pré-sélection et la pré-sélection sur les images mesurées et la pré-sélection « modèle 3D » sur les images prédites (voir section 6.4). Les γ sélectionnés ont une distance entre la position de la source simulée et leur direction d'arrivée reconstruite inférieure à 0,2°. Quant aux électrons et aux protons, ils ont un offset maximal de 1,0°.

On remarquera que la tendance générale du pouvoir discriminant pour une variable donnée dépend fortement de l'énergie et de la multiplicité des évènements. Hormis le cas particulier des deux variables ΔQ et 3d-depth, on notera que comme attendu le pouvoir de discrimination augmente en fonction de l'énergie et de la multiplicité des évènements car plus les images sont nombreuses et contiennent de l'information, et plus il est facile de différencier les γ du fond.

Notons le cas particulier de la variable ΔQ, pour laquelle une nouvelle population est visible sur la Figure 6.16a, qui n'est pas présente en l'absence de CT_5. Elle concerne des évènements hadroniques de multiplicité deux et de basses énergies reconstruites qui ont de grandes valeurs de ΔQ. Ces évènements ont une hauteur de maximum de gerbe reconstruite très basse et ont une charge détectée près du seuil des caméras de type HESS-I, et de l'ordre de 80 p.e. pour CT_5. Ces observations permettent de situer le pied de gerbe de ces hadrons aux alentours d'un des télescopes moyens du réseau. Ils possèdent une énergie reconstruite très faible, de l'ordre de 20 GeV, qui induisent une très grande différence entre la charge mesurée dans CT_5 et la charge prédite à l'aide des tables de calibration en énergie des γ, ce qui conduit à une très grande valeur de ΔQ. Ces mêmes évènements sont visibles sur la Figure 6.15a, ils peuplent la droite de la distribution rouge du paramètre 3d-depth (une grande valeur de 3d-depth équivaut à un petit $H_{\max}$).

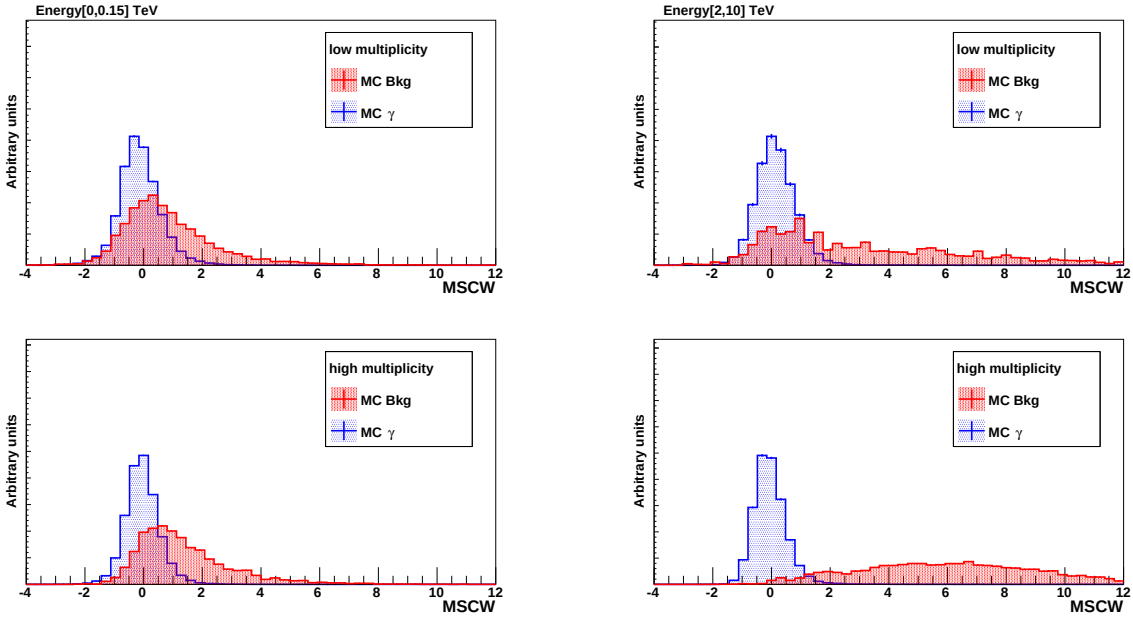
6.5.5 Réglages de la méthode BDT et phases d'entraînement

La méthode multivariée BDT (déjà utilisée dans Paris-MVA) a été choisie pour ce problème de classification binaire γ/fond (voir chapitre 3). Sa rapidité d'exécution, sa capacité à gérer des variables qui peuvent être peu discriminantes dans certains intervalles de l'espace des phases, ainsi que sa capacité à traiter les corrélations entre les variables, font d'elle une méthode non-paramétrique



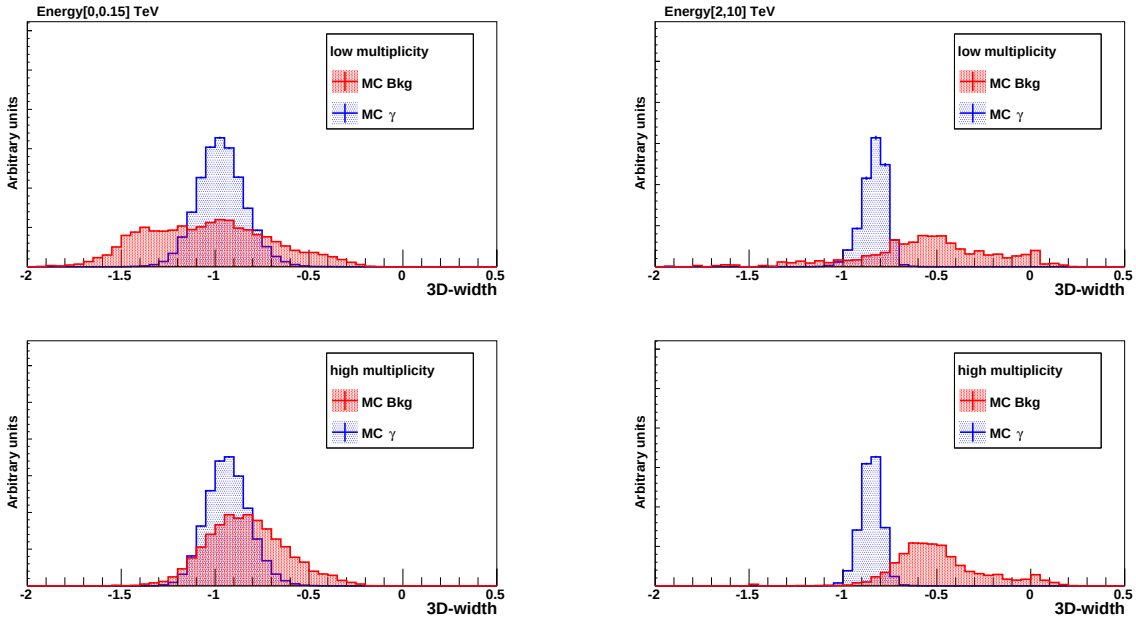
(a) (b)

FIGURE 6.11: Distributions du paramètre MSCL pour des évènements ayant (a) des énergies reconstruites inférieures à 150 GeV et (b) supérieures à 2 TeV. Distributions des γ (bleu) et du fond (rouge) pour des évènements de multiplicité 2 (haut) et supérieure à 2 (bas).



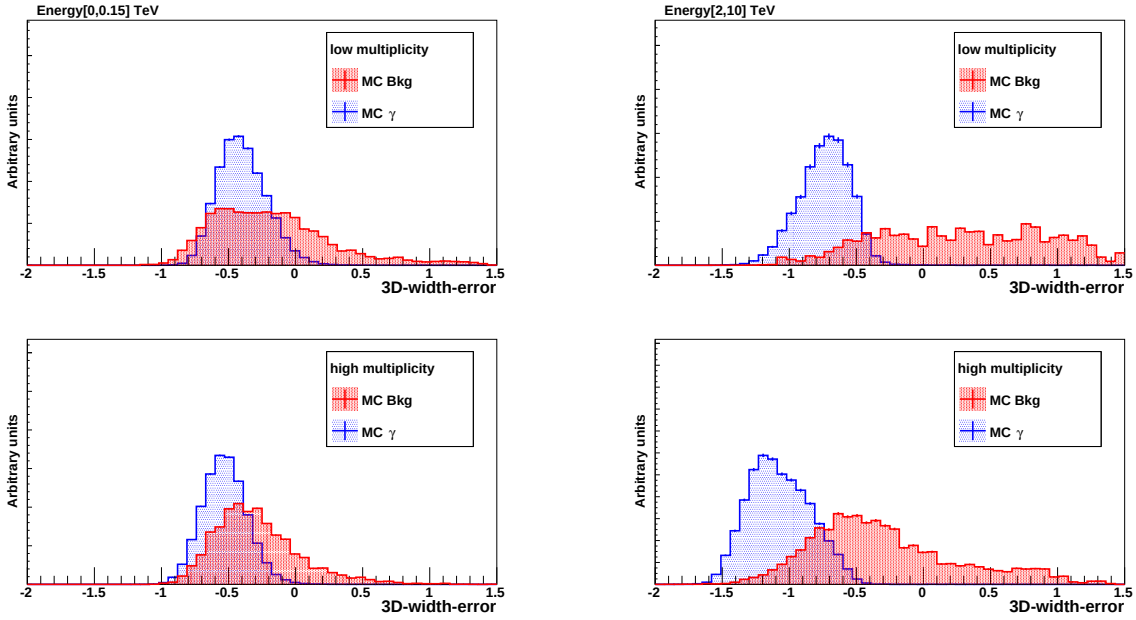
(a) (b)

FIGURE 6.12: Distributions du paramètre MSCW pour des évènements ayant (a) des énergies reconstruites inférieures à 150 GeV et (b) supérieures à 2 TeV. Distributions des γ (bleu) et du fond (rouge) pour des évènements de multiplicité 2 (haut) et supérieure à 2 (bas).



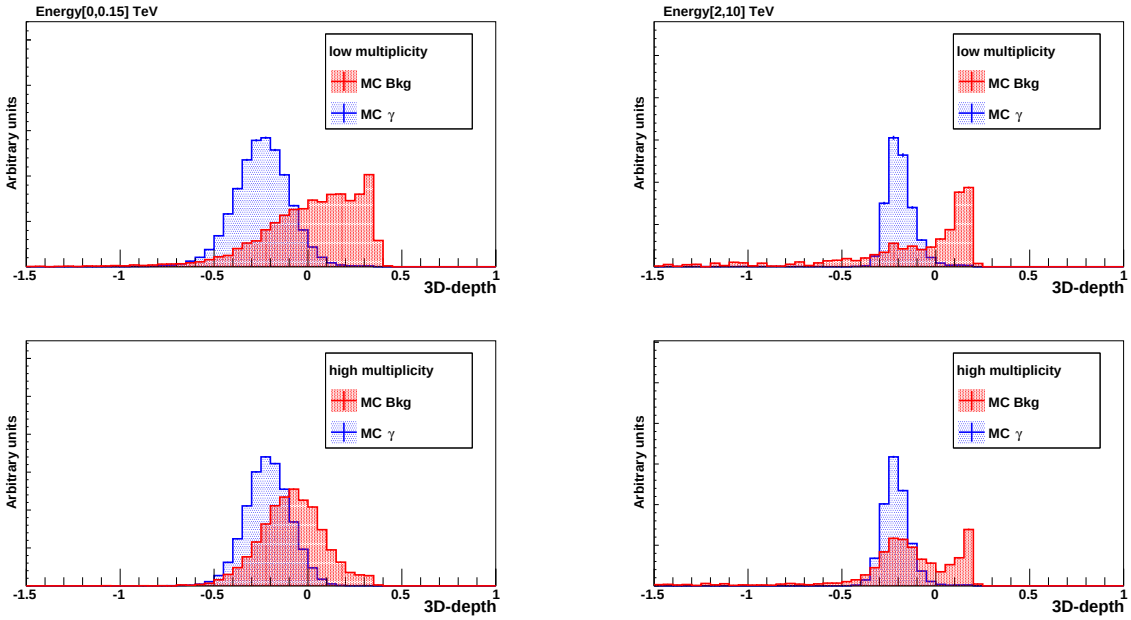
(a) (b)

FIGURE 6.13: Distributions du paramètre 3D-width pour des évènements ayant (a) des énergies reconstruites inférieures à 150 GeV et (b) supérieures à 2 TeV. Distributions des γ (bleu) et du fond (rouge) pour des évènements de multiplicité 2 (haut) et supérieure à 2 (bas).



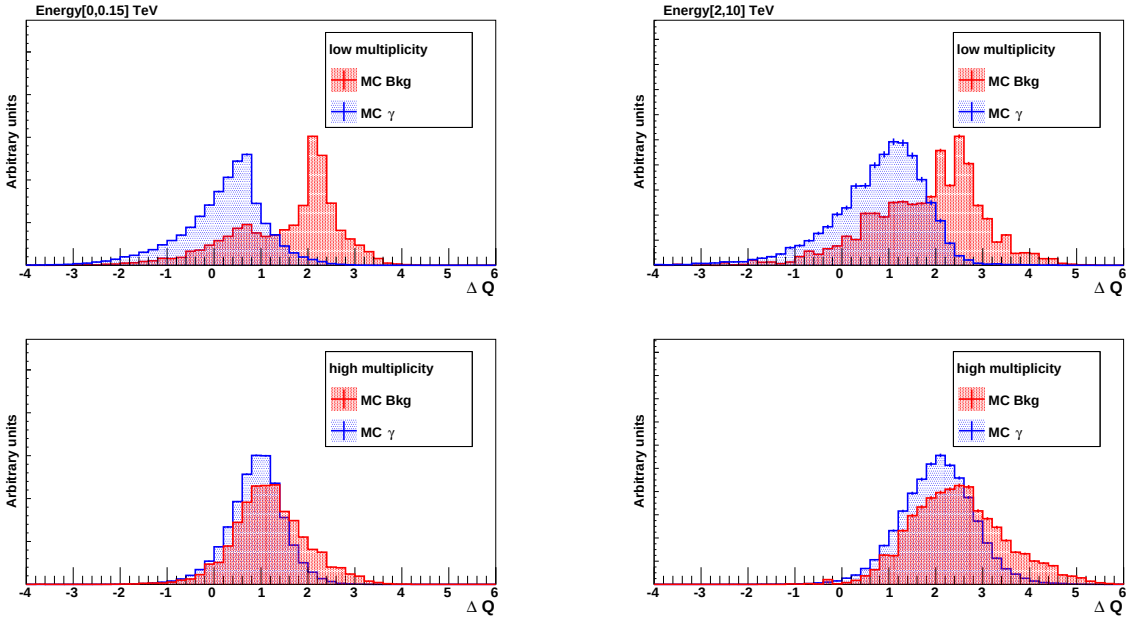
(a) (b)

FIGURE 6.14: Distributions du paramètre 3D-width-error pour des évènements ayant (a) des énergies reconstruites inférieures à 150 GeV et (b) supérieures à 2 TeV. Distributions des γ (bleu) et du fond (rouge) pour des évènements de multiplicité 2 (haut) et supérieure à 2 (bas).



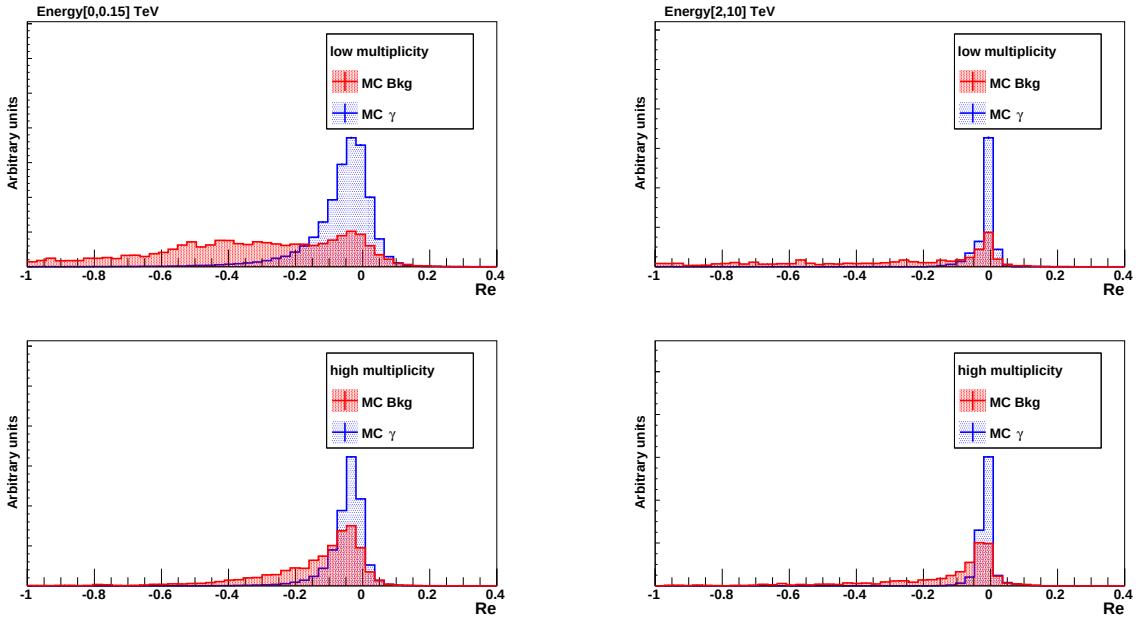
(a) (b)

FIGURE 6.15: Distributions du paramètre 3D-depth pour des évènements ayant (a) des énergies reconstruites inférieures à 150 GeV et (b) supérieures à 2 TeV. Distributions des γ (bleu) et du fond (rouge) pour des évènements de multiplicité 2 (haut) et supérieure à 2 (bas).



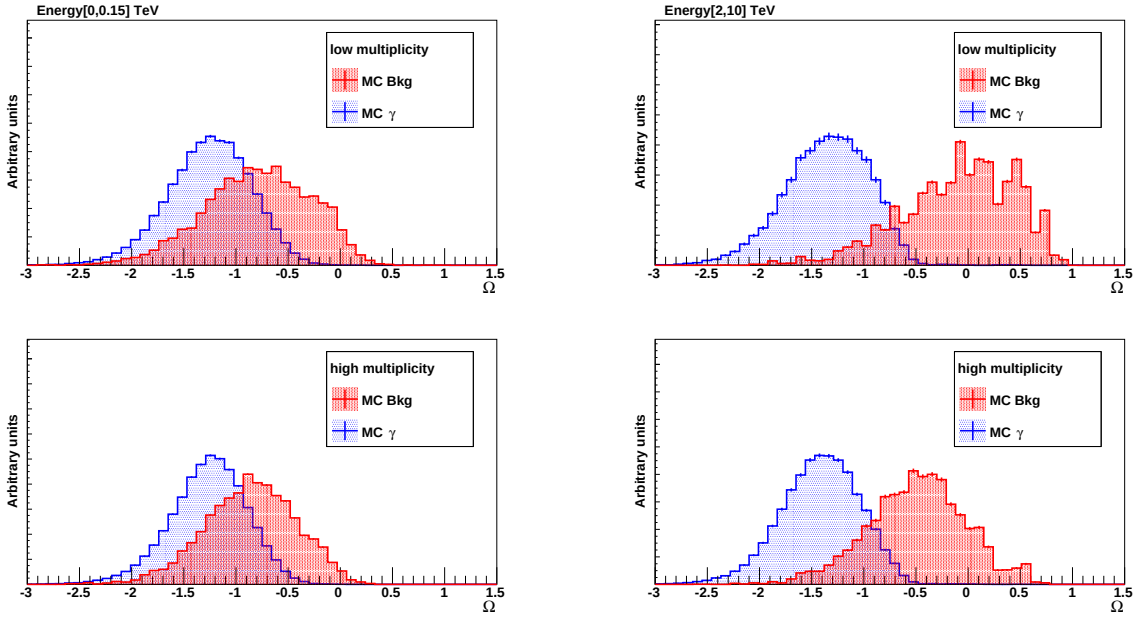
(a) (b)

FIGURE 6.16: Distributions du paramètre ΔQ pour des évènements ayant (a) des énergies reconstruites inférieures à 150 GeV et (b) supérieures à 2 TeV. Distributions des γ (bleu) et du fond (rouge) pour des évènements de multiplicité 2 (haut) et supérieure à 2 (bas).



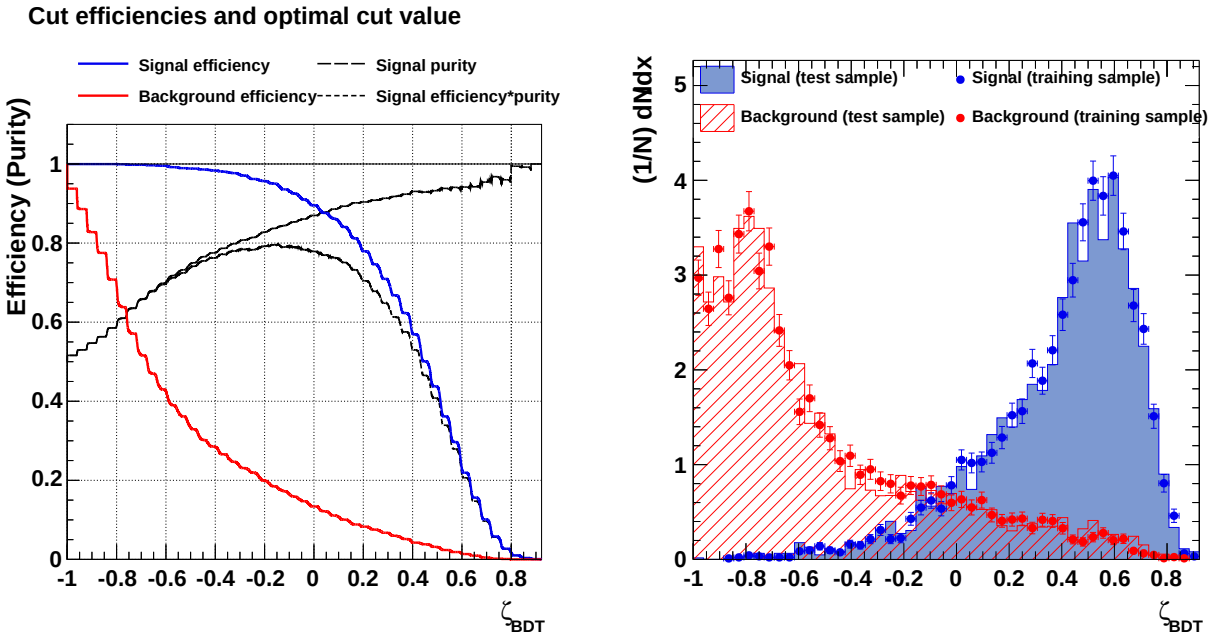
(a) (b)

FIGURE 6.17: Distributions du paramètre R_{E} pour des évènements ayant (a) des énergies reconstruites inférieures à 150 GeV et (b) supérieures à 2 TeV. Distributions des γ (bleu) et du fond (rouge) pour des évènements de multiplicité 2 (haut) et supérieure à 2 (bas).



(a) (b)

FIGURE 6.18: Distributions du paramètre Ω pour des évènements ayant (a) des énergies reconstruites inférieures à 150 GeV et (b) supérieures à 2 TeV. Distributions des γ (bleu) et du fond (rouge) pour des évènements de multiplicité 2 (haut) et supérieure à 2 (bas).



(a) (b)

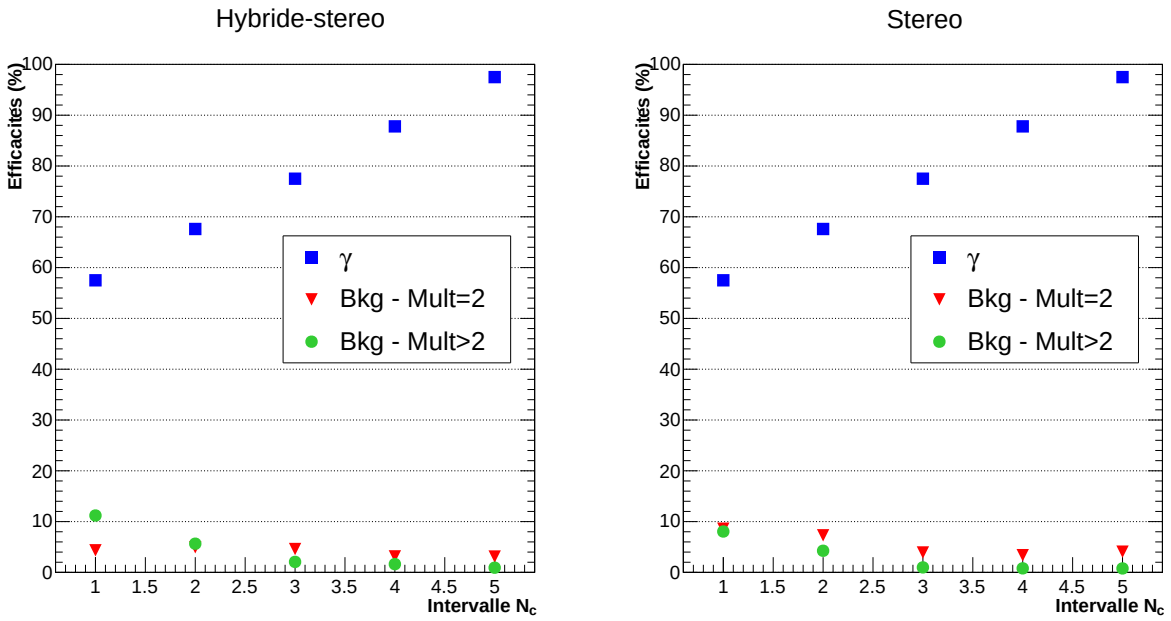
FIGURE 6.19: (a) Efficacités de reconstruction du signal (en bleu) et du fond (en rouge) en fonction de ζ pour le premier intervalle en N_c et de multiplicité 2. (b) Distributions de ζ pour les évènements γ (en rouge), et de fond (en bleu) pour les lots d'entraînements (traits pleins), et de test (points), pour ce même intervalle.

particulièrement adaptée pour cette analyse. Le réglage de la méthode suit la procédure explicitée dans [Becherini et al., 2011]. L'algorithme AdaBoost est utilisé avec des arbres d'une profondeur maximale de 20 nœuds. Pour un intervalle donné de l'espace des phases, le nombre d'arbres générés N_t qui composent la forêt dépend de la statistique disponible lors de la phase d'entraînement : $N_t = \min(\max(n_t, 10), 200)$ avec $n_t = (N_{\text{evt}}/100)$. Dans le cas présent, le nombre d'arbres est compris entre dix et cinquante. Notons que les développeurs [7] de TMVA recommandent d'utiliser des arbres courts d'une profondeur de deux ou trois nœuds et de générer une forêt avec un grand nombre d'arbres. Cette méthode sera appliquée au chapitre suivant, mais les essais réalisés n'ont pas montré de différences significatives dans les résultats entre les deux approches. On pourra trouver l'ensemble des réglages de la méthode dans [Becherini et al., 2011].

L'entraînement associé à chaque intervalle de l'espace des phases est réalisé en utilisant une fraction des évènements, que l'on appellera le lot d'entraînement. Chaque arbre de décision attribuera une valeur $+1$ pour les évènements « γ-like » et -1 aux autres évènements. La valeur de discrimination ζ est une combinaison des décisions de tous les arbres (voir chapitre 3 pour plus de détails). Dans un second temps, le reste de la statistique, que l'on appellera le lot test, est utilisé pour déterminer les efficacités de sélection [8] des évènements γ, que l'on notera ε_{Sig}, et du fond, que l'on notera ε_{Bkg}, en fonction de la valeur ζ. À titre d'exemple, la Figure 6.19a représente les quantités ε_{Sig} (en bleu) et ε_{Bkg} (en rouge), toutes deux estimées sur le lot test, en fonction de ζ pour le premier intervalle de N_c de multiplicité deux. De plus, cette phase permet de mettre en évidence d'éventuels effets liés au sur-entraînement en comparant les distributions du paramètre ζ des évènements signal et fond entre les lots d'entraînement et de test. Pour quantifier de possibles différences, on utilisera le test de Kolmogorov-Smirnof (KS) (voir chapitre 3). Un exemple de com-

7. Section 7 de : http://tmva.sourceforge.net/docu/TMVAUsersGuide.pdf

8. Pour rappel, ces efficacités sont évaluées sur l'ensemble des évènements passant les coupures de pré-sélection y compris la convergence du modèle 3D.



(a) (b)

FIGURE 6.20: Efficacités de reconstruction γ et de fond pour les différents intervalles de l'espace des phases des analyses (a) hybride et (b) stéréo. Les efficacités de fond pour les multiplicités deux sont symbolisées par des triangles rouges, et celles de multiplicités supérieures par des ronds verts. Elles sont déduites des valeurs de $\varepsilon_{\mathrm{Sig}}$ (voir texte).

paraison entre les distributions des lots de test (points) et d'entraînement (traits pleins) est donné sur la Figure 6.19b, pour le premier intervalle en N_c et de multiplicité deux, sur laquelle les couleurs bleues et rouges représentent respectivement les évènements γ et de fond.

6.5.6 Point de fonctionnement

Choisir un point de fonctionnement dans un intervalle de l'espace des phases revient à choisir une valeur pour la coupure de discrimination finale, que l'on notera $\zeta^\star$ [9], à laquelle sont associées des efficacités de reconstruction du signal et du fond données. On peut à l'inverse fixer les valeurs de $\varepsilon_{\mathrm{Sig}}$, ce qui permet de veiller au comportement de ce paramètre en fonction de l'énergie – important pour éviter les discontinuités dans les surfaces de collection – et en déduire les valeurs de $\zeta^\star$ et $\varepsilon_{\mathrm{Bkg}}$. Dans cette étude, nous nous sommes inspirés d'un gabarit des efficacités de reconstruction γ obtenu avec la méthode décrite dans [Becherini et al., 2011] afin de fixer le point de fonctionnement. Ce gabarit qui a été obtenu pour une analyse en mode stéréo est adapté pour la détection des sources ayant un niveau de flux intégré de l'ordre de 1 % de la nébuleuse du Crabe et un spectre relativement mou (indice spectral de $\Gamma = -3$). On sélectionnera la valeur $\zeta^\star$ qui correspond à l'efficacité de reconstruction γ donnée par le gabarit (courbes bleues sur la Figure 6.20). Cette méthode n'est pas optimale mais elle permet néanmoins d'estimer le gain en sensibilité apporté par l'ajout du cinquième télescope au réseau HESS-I et d'avoir une estimation préliminaire de la sensibilité du réseau HESS-II en mode hybride. Nous présenterons au prochain chapitre une nouvelle méthode d'optimisation lors de la mise en place d'une analyse utilisant des données réelles pour le fond.

Les Figures 6.20a et 6.20b représentent les efficacités de reconstruction pour les γ, indépendantes de la multiplicité, et pour le fond (triangles rouges et cercles verts respectivement pour les multiplicités deux et supérieures à deux) en fonction de l'intervalle en N_c, respectivement pour les analyses

9. Pour rappel, tous les évènements ayant une valeur $\zeta < \zeta^\star$ ne seront pas considérés comme des évènements « γ-like ».

en mode hybride et stéréo. Une adaptation a été nécessaire par rapport au gabarit présenté dans [Becherini et al., 2011] du fait des différences dans le découpage de l'espace des phases. À basse énergie, notamment pour le premier intervalle en N_c, et pour la multiplicité deux, on constate un pouvoir de séparation γ/fond plus important lorsque CT_5 est présent. Ceci est dû au fort pouvoir discriminant de la variable ΔQ qui a été discuté à la section 6.5.4. À plus haute énergie, toujours pour la multiplicité deux, ε_{Bkg} est de l'ordre de 5 % pour les deux analyses. Pour les intervalles de haute multiplicité, l'efficacité de rejet du fond croît lentement en fonction de l'énergie dans les deux analyses. Elle est de ∼90 % en deçà de ∼150 GeV et elle est supérieure à 99 % au-delà de ∼1 TeV.

À titre d'illustration, on peut trouver sur les Figures 6.21 et 6.22 les distributions du paramètre ζ pour des évènements γ (bleu) et des évènements de fond (rouge), respectivement pour le premier intervalle en N_c de l'espace des phases, et pour les quatre autres intervalles. Est aussi indiquée par une flèche rouge, pour chaque intervalle de l'espace des phases, la valeur finale de discrimination $\zeta^\star$ correspondante.

6.6 Estimation des performances préliminaires

Le calcul des performances d'un instrument permet entre autres d'estimer le temps d'observation requis pour extraire un signal significatif d'une source ayant un certain niveau de flux. Il permet aussi de quantifier de manière précise le gain ou la perte de sensibilité apporté par des modifications dans la chaîne d'analyse, au niveau de la reconstruction ou de la discrimination. Dans cette section nous présentons les performances préliminaires obtenues avec les simulations présentées à la section 6.2 pour le réseau HESS-II en mode hybride et, à titre de comparaison, pour le réseau HESS-I en mode stéréo. Nous commencerons par détailler la méthode d'estimation de la sensibilité différentielle.

6.6.1 Le calcul de la sensibilité

Pour estimer la sensibilité différentielle, nous avons à disposition des simulations de γ qui composent le signal et des simulations de hadrons et d'électrons qui composent le fond. Le produit final que nous souhaitons obtenir est le flux différentiel minimal en fonction de l'énergie reconstruite que l'on va pouvoir détecter en un temps d'observation T_{Obs}. Cette procédure ne peut pas être conduite en énergie vraie, du fait de la difficulté de reconstruire celle-ci dans le cas des protons.

Afin d'estimer le flux minimal qui peut être détecté dans un intervalle en énergie reconstruite, il est nécessaire de se mettre dans des conditions réalistes d'observation (voir section 5.3.2 du chapitre 5), à savoir créer une région ON ($\theta^2_{\text{ON}} = 0{,}0125°$) et créer des régions OFF (on prendra cinq régions de même offset et de même taille angulaire que la région ON). Afin de déterminer le nombre d'évènements de fond attendus dans les régions définies ci-dessus, nous considérons dans un premier temps – et ce pour faire cette estimation sur une statistique convenable – les évènements de fond sélectionnés jusqu'à 1° par rapport au centre de la caméra (comme pour les phases d'entraînement) et passant les coupures finales de discrimination. Nous apportons à ce nombre une première correction qui consiste à prendre en compte les spectres expérimentaux des électrons et des protons ainsi que le temps T_{Obs} pour lequel nous souhaitons évaluer la sensibilité (voir section 6.5.2).

En notant N_{bkg} le nombre d'évènements de fond ainsi corrigé sélectionné dans une région de 1° de rayon, on peut estimer le nombre d'évènements de fond attendus dans les régions OFF selon la transformation :

$$N_{\text{bkg}} \rightarrow \left(\frac{0.0125}{1}\right) \frac{N_{\text{bkg}}}{\alpha} \tag{6.5}$$

où $\alpha = 1/5$ est le rapport entre les angles solides de la région ON et des régions OFF.

Compte tenu du nombre d'évènements de fond ainsi obtenu, il faut ensuite estimer le nombre de γ, que l'on notera N_γ, nécessaire pour obtenir un niveau de détection de 5σ. Pour cela, l'estimation de la significativité de [Li et Ma, 1983] définie au chapitre 5 à l'équation 5.36 est utilisée, où le

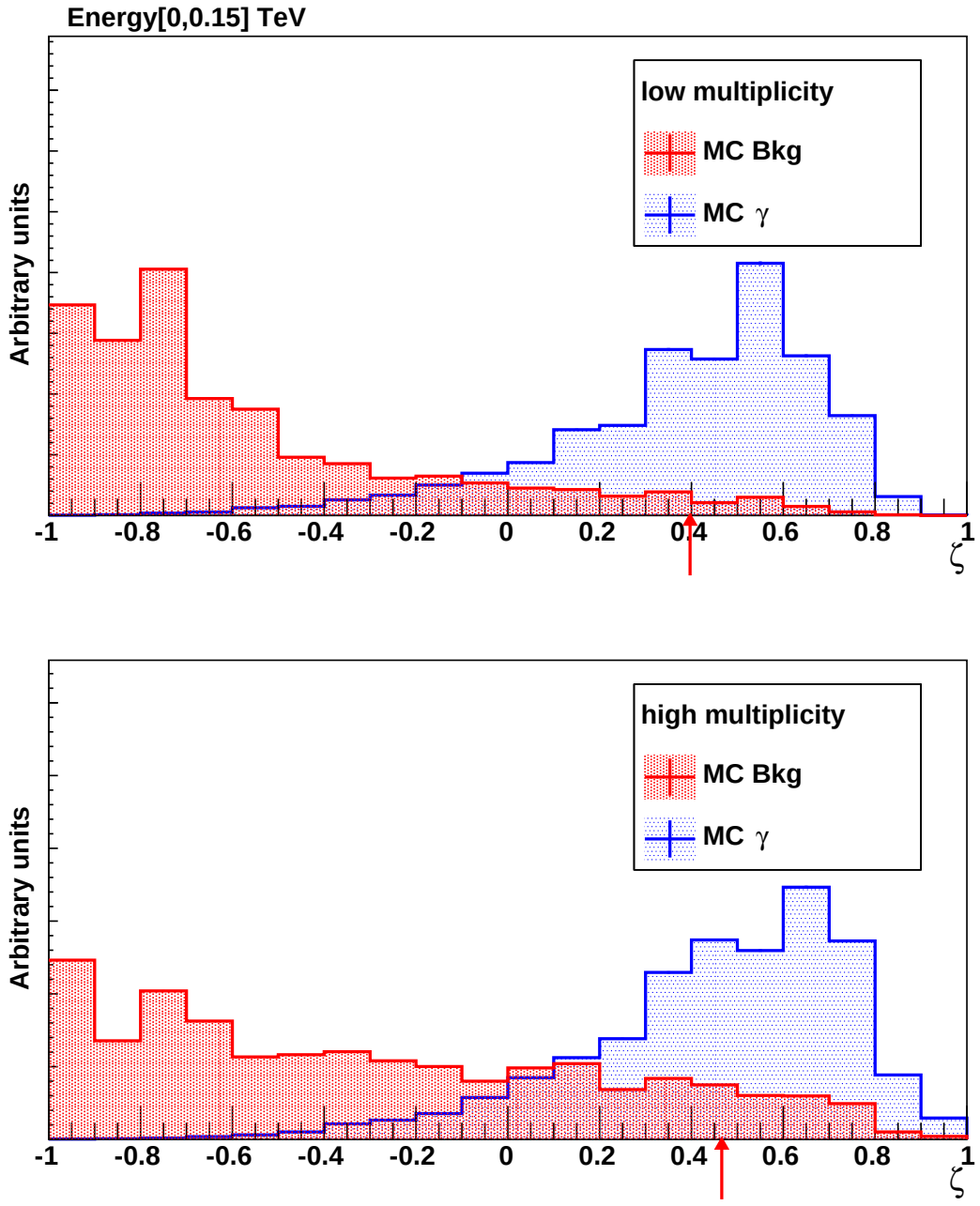


FIGURE 6.21: Distributions du paramètre ζ pour des évènements γ simulés (bleu) et des évènements de fond (rouge) pour le premier intervalle en N_c (multiplicité deux en haut et multiplicité supérieure à deux en bas). La valeur finale de discrimination $\zeta^\star$ est indiquée par une flèche rouge.

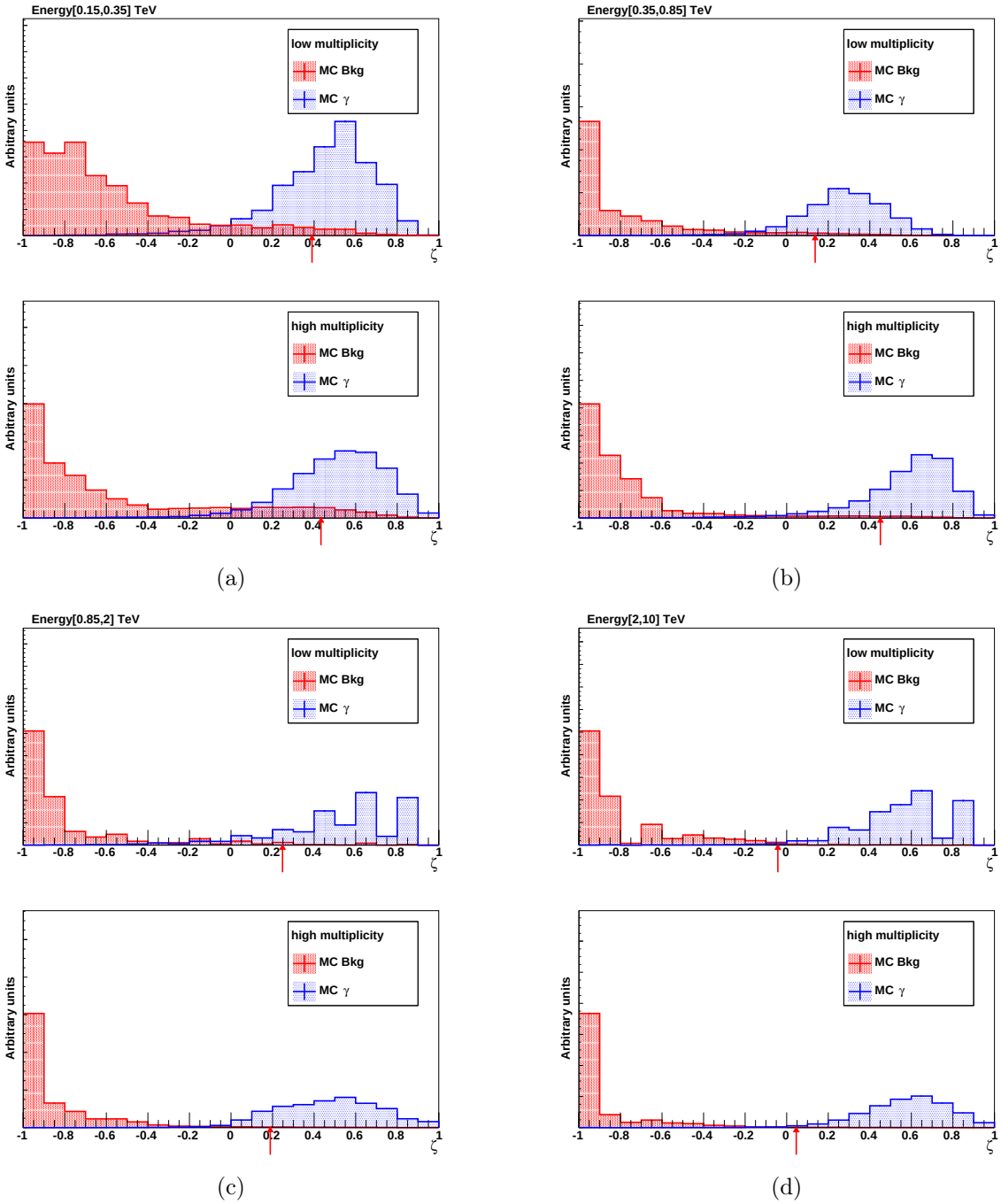


FIGURE 6.22: Distributions du paramètre ζ pour des évènements γ simulés (bleu) et des évènements de fond (rouge) pour les intervalles 2 et 3 en N_c (a et b) et les intervalles 4 et 5 en N_c (c et d). Pour chaque intervalle, les évènements de multiplicité deux sont représentés en haut et les évènements de multiplicité supérieure à deux sont représentés en bas. La valeur finale de discrimination $\zeta^\star$ est indiquée par une flèche rouge.

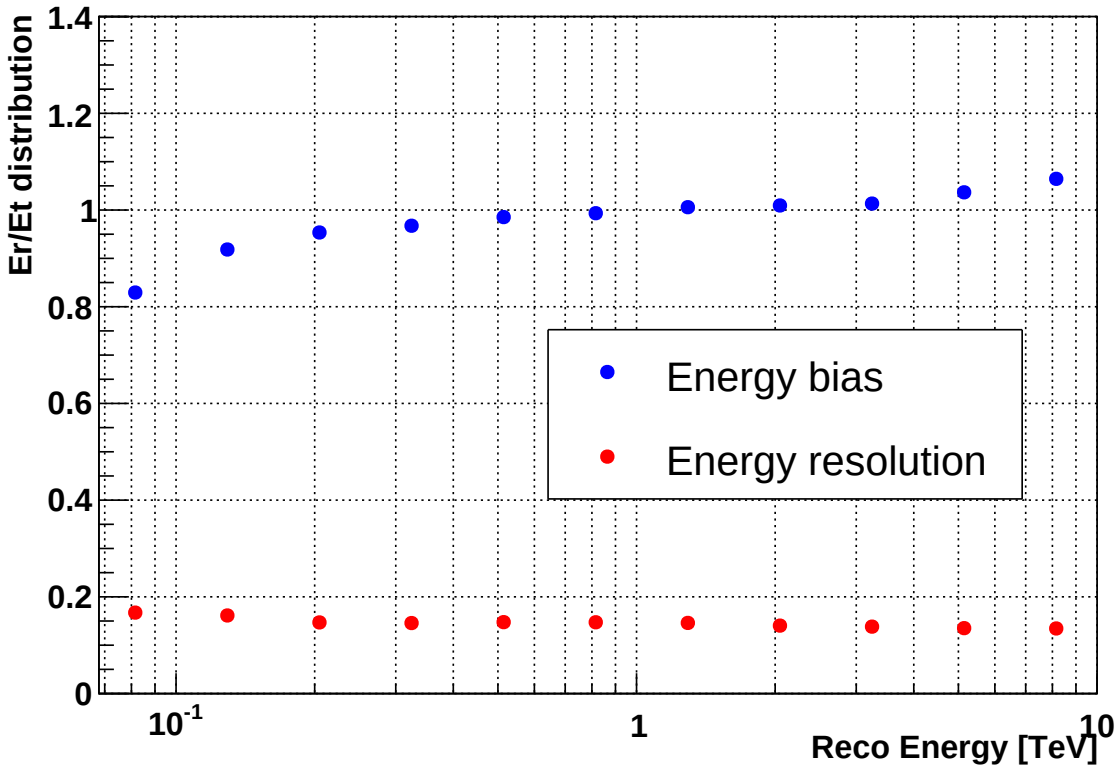


FIGURE 6.23: Résolution (rouge) et biais (bleu) en énergie en fonction de l'énergie reconstruite en mode hybride, obtenus à l'aide de l'ajustement d'une gaussienne.

nombre total d'évènements dans la région ON vaut :

$$N_{\mathrm{ON}} = N_\gamma + \alpha N_{\mathrm{bkg}} \tag{6.6}$$

Une contrainte supplémentaire est appliquée afin de limiter les effets systématiques liées à la soustraction du fond. Dans le cas où N_γ est inférieur à 5 % du nombre d'évènements de fond dans la région ON le nombre de γ est fixé à cette valeur afin d'obtenir une estimation du flux conservative.

Une fois le nombre de γ estimé, il est possible de construire un flux différentiel selon :

$$\Phi(\widetilde{E}) = \frac{N_\gamma}{T_{\mathrm{Obs}} \mathcal{A}(\bar{E}) \Delta\widetilde{E}} \tag{6.7}$$

où $\Delta\widetilde{E}$ est la largeur de l'intervalle en énergie reconstruite et $\mathcal{A}(\bar{E})$ est la surface efficace pour l'énergie vraie $\bar{E}$ qui correspond à la moyenne des énergies vraies des évènements formant l'intervalle en énergie reconstruite de largeur $\Delta\widetilde{E}$. Cette énergie moyenne $\bar{E}$ permet de prendre en compte les déversements dus à la méthode de reconstruction de l'énergie. En effet, on peut observer sur la Figure 6.23 que la méthode utilisée pour estimer l'énergie des rayons γ possède un biais négatif de reconstruction vers les basses énergies. Ne pas prendre en compte cette correction reviendrait à biaiser le nombre de photons de basse énergie et par conséquent à biaiser aussi le flux qui entre dans le calcul de la sensibilité différentielle.

6.6.2 Résultats

Les résultats présentés ci-dessous correspondent à une analyse hybride menée d'abord sur un lot d'évènements différent de celui utilisé pour les phases d'entraînement et de test, auquel nous avons

ajouté 10 % de ces derniers afin d'augmenter la statistique. Nous présentons également, à titre de comparaison, les résultats d'une analyse stéréo menée sur les mêmes évènements après avoir retiré CT_5. Les performances sont représentées sur les Figures 6.24 et 6.25 pour les configurations hybride en rouge et stéréo en noir.

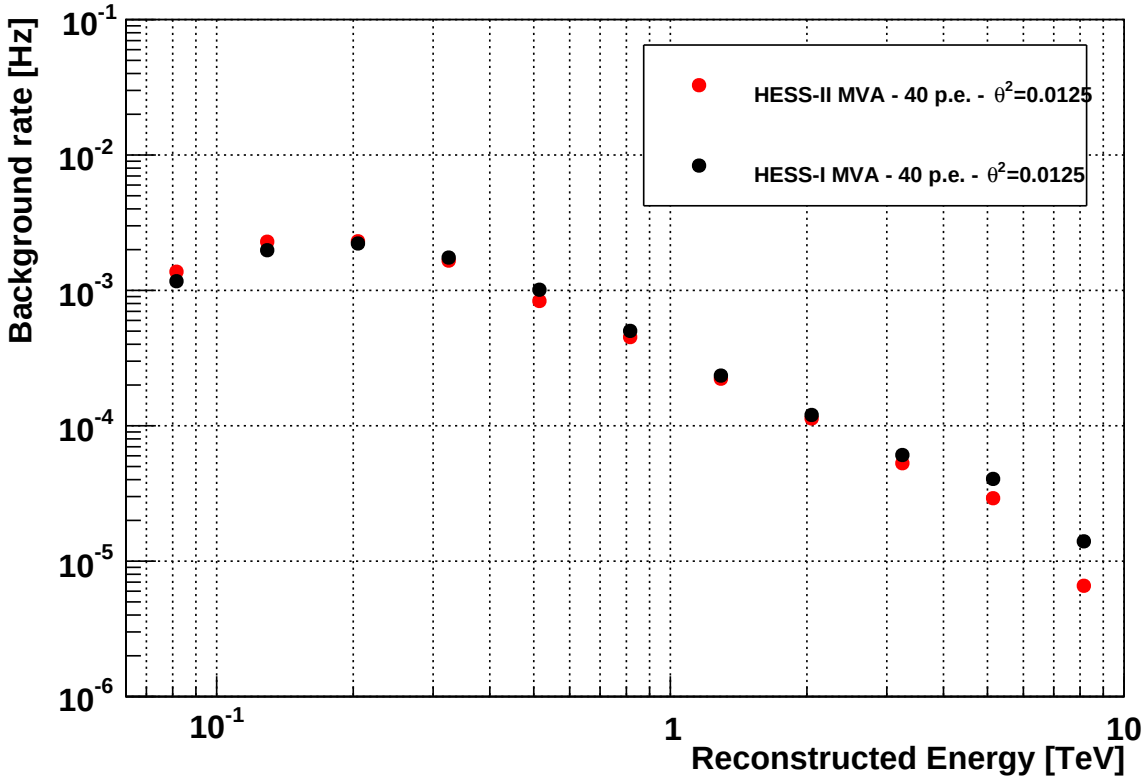
La Figure 6.24a représente le taux de fond (électrons et protons) passant les coupures de sélection finales en fonction de l'énergie reconstruite. Au-delà de ∼3 TeV, l'écart significatif entre les taux de fond des deux analyses s'explique par la présence de la coupure $\sin\theta_{ij}$ en mode hybride pour les télescopes moyens. Les évènements de haute énergie possédant majoritairement de grands pieds de gerbe, les images dans les caméras auront tendance à être alignées. Le cinquième télescope ne participant pas à la reconstruction de ces évènements à cause de son champ de vue réduit par rapport aux télescopes moyens, cette diminution du taux de fond est principalement due à la coupure $\sin\theta_{ij}$, qui n'est appliquée que sur l'analyse hybride. À plus basse énergie (≤200 GeV), l'analyse hybride détecte un taux de fond plus important que l'analyse stéréo, de l'ordre de 20 % à 80 GeV, du fait de l'arrivée de la nouvelle population d'évènements de faible multiplicité (section 6.3).

Les surfaces de collection des deux types de réseaux sont représentées sur la Figure 6.24b (même code couleurs que précédemment). Un gain de surface de collection est visible en deçà de 800 GeV qui est à attribuer à la grande surface de collection de CT_5. À 80 GeV, le gain en surface efficace dû au cinquième télescope est de 300 %. On peut observer au-delà de 1 TeV une légère diminution de la surface efficace qui correspond à une perte d'évènements inférieure à 1 %. Une fois de plus, cette différence est à attribuer à la coupure $\sin\theta_{ij}$ qui est compatible avec les résultats de la section 6.4.4.

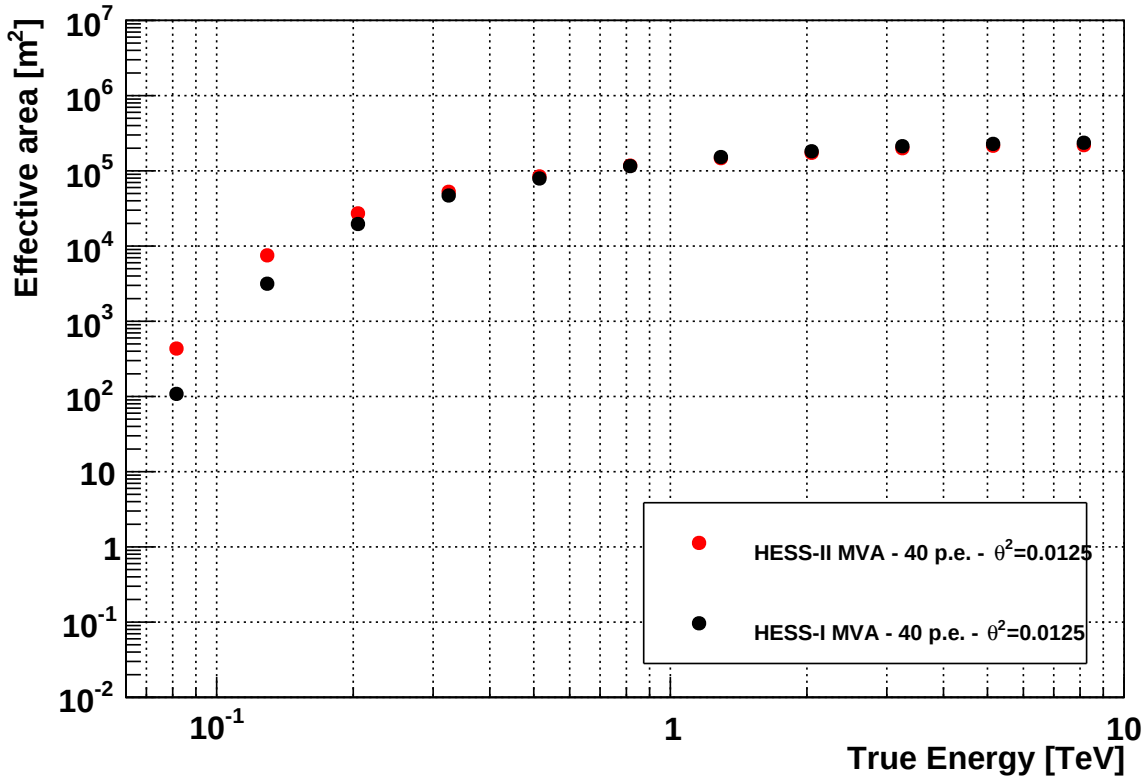
Sur la Figure 6.25a est représentée la précision angulaire à 68 %. Au-delà de 2 TeV, un gain de PSF de ∼5 % est obtenu avec la configuration hybride que l'on peut une fois de plus attribuer à la coupure $\sin\theta_{ij}$, pour les raisons déjà évoquées plus haut. En dessous de ∼400 GeV on observe une légère amélioration de la PSF qui est le résultat de deux effets contraires : l'amélioration de l'imagerie liée à la grande surface de collection de CT_5 et l'arrivée massive d'évènements de multiplicité deux, observés par CT_5 et un télescope moyen, dont la reconstruction de la direction par la méthode des ellipses est difficile.

La sensibilité différentielle est représentée sur la Figure 6.25b. L'ajout du cinquième télescope entraîne une amélioration significative de la sensibilité en dessous de 800 GeV, avec un gain maximal de 30 % à 125 GeV. La sensibilité à 80 GeV de la configuration stéréo n'est pas représentée à cause de sa faible surface de collection, de l'ordre de 100 m^2. Au-delà de ∼1 TeV l'apport de CT_5 en termes de sensibilité est négligeable. Les deux points au-delà de 5 TeV qui ont une sensibilité différentielle plus importante pour la configuration hybride expriment le fait que la coupure $\sin\theta_{ij}$, initialement introduite afin d'améliorer la PSF, supprime aussi une fraction non négligeable d'évènements hadroniques à très grand pied de gerbe (les images dans les caméras ont tendance à être très alignées).

Les résultats présentés ici ont été totalement dérivés à partir de simulations pour les γ et pour le fond composé de protons et d'électrons. Notons que ces simulations ont été générées avant la mise en fonctionnement de l'instrument CT_5 et que par conséquent elles n'ont pas bénéficié d'un retour sur expérience issu des premières périodes d'observation (une nouvelle production est actuellement en préparation).

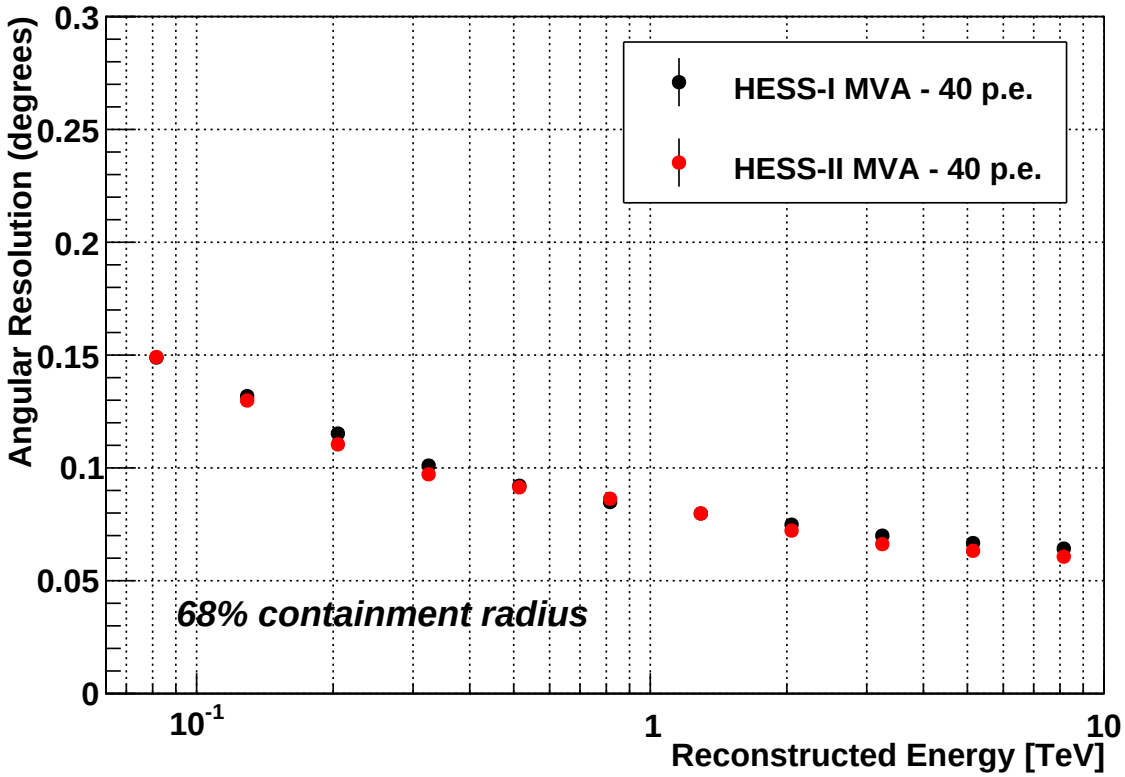


(a) Taux de fond en fonction de l'énergie reconstruite.

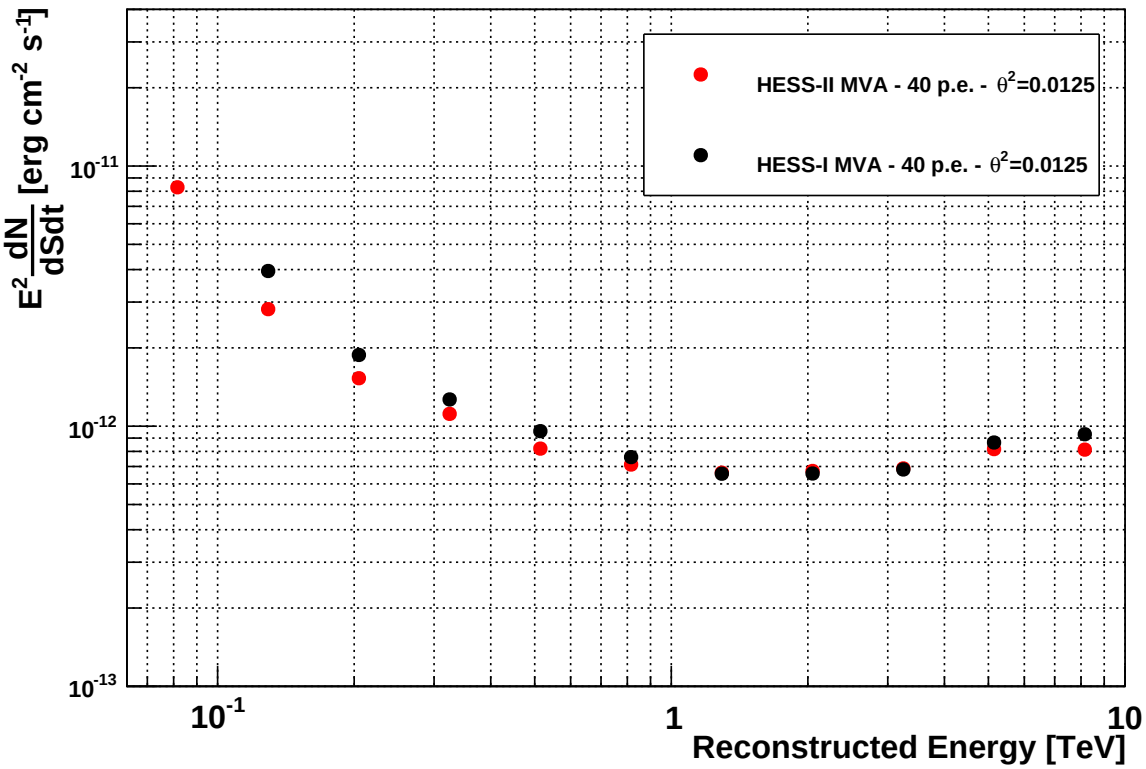


(b) Surface efficace γ en fonction de l'énergie vraie.

FIGURE 6.24: Comparaison du (a) taux de fond et (b) de la surface efficace entre les analyses en mode hybride (rouge) et stéréo (noir).



(a) PSF en fonction de l'énergie reconstruite.



(b) Sensibilité en fonction de l'énergie reconstruite ($T_{\mathrm{Obs}} = 50\,\mathrm{h}$).

FIGURE 6.25: Comparaison de (a) la précision angulaire et (b) de la sensibilité différentielle ($T_{\mathrm{Obs}} = 50\,\mathrm{h}$) entre les analyses en mode hybride (rouge) et stéréo (noir).

Bibliographie

F. Aharonian, et al. Evidence for VHE γ-ray emission from the distant BL Lac PG 1553+113. A&A, 448:L19–L23, March 2006. doi: 10.1051/0004-6361:200600010.

Y. Becherini, et al. A new analysis strategy for detection of faint γ-ray sources with Imaging Atmospheric Cherenkov Telescopes. *Astroparticle Physics*, 34:858–870, July 2011. doi: 10.1016/j.astropartphys.2011.03.005.

K. Bernlöhr, et al. Monte Carlo design studies for the Cherenkov Telescope Array. *Astroparticle Physics*, 43:171–188, March 2013. doi: 10.1016/j.astropartphys.2012.10.002.

J. Guy. *Premiers résultats de l'expérience HESS et étude du potentiel de détection de matière noire supersymétrique.* PhD thesis, UNIVERSITÉ PARIS 6, 2003.

D. Heck, et al. *CORSIKA: a Monte Carlo code to simulate extensive air showers.* February 1998.

M. P. Kertzman et G. H. Sembroski. Computer simulation methods for investigating the detection characteristics of TeV air Cherenkov telescopes. *Nuclear Instruments and Methods in Physics Research A*, 343:629–643, April 1994. doi: 10.1016/0168-9002(94)90247-X.

T.-P. Li et Y.-Q. Ma. Analysis methods for results in gamma-ray astronomy. ApJ, 272:317–324, September 1983. doi: 10.1086/161295.

Chapitre 7

Analyse APC-MVA en mode hybride

Sommaire

7.1 Introduction . . . 133
7.2 Simulations . . . 133
7.3 Fond réel . . . 134
7.4 Construction des analyses . . . 135
7.5 Configurations d'analyse et performances . . . 138
7.6 Détection des premières sources . . . 140
7.7 Spectres . . . 144
Bibliographie . . . 150

7.1 Introduction

Dans le chapitre précédent, nous avons présenté une analyse préliminaire en mode hybride développée à l'aide de simulations MC de γ, hadrons et électrons, afin d'estimer le gain en performance induit par l'ajout du cinquième télescope. Nous présentons ici l'analyse APC-MVA en mode hybride construite à partir de données réelles pour le fond, et les résultats des premières analyses de données.

Nous commencerons par introduire les simulations et les données de fond utilisées pour la construction de l'analyse respectivement à la section 7.2 et à la section 7.3. Nous présenterons ensuite à la section 7.4 la construction des analyses et en particulier une nouvelle méthode permettant d'optimiser le pouvoir de séparation γ/fond pour un type de source donné. S'en suivra à la section 7.5 une présentation des configurations d'analyse adaptées à l'étude de sources de nature différentes, et à la section 7.6 une présentation des résultats obtenus sur les premières sources observées par le réseau HESS-II. Ensuite, nous présenterons à la section 7.7 l'analyse spectrale préliminaire de ces premières données.

7.2 Simulations

Nous présentons ici les différents jeux de simulations γ utilisées à différents niveaux de la construction de l'analyse.

Un premier jeu de simulations de γ a été utilisé pour la phase d'entraînement afin de déterminer les coupures finales de discrimination γ/fond, pour la génération des IRF, ainsi que pour analyser les premières données en mode hybride. Ces simulations correspondent à la production utilisée au chapitre 6 car elle est la seule, à ce jour, à posséder une statistique suffisante pour la phase d'entraînement. Elle a été augmentée par rapport au chapitre précédent pour avoir à disposition différents angles zénithaux (0, 20, 30, 40, 45, 50 et 55°), mais elle n'est disponible que pour un seul offset de 0,5°. Des mesures de l'efficacité optique récentes, après la période de ré-aluminisation des miroirs des télescopes de type HESS-I, donnent pour ceux-ci des valeurs comprises entre 55 % et 65 % de l'efficacité optique nominale. Pour CT_5 cette valeur est de ∼75 %. Les simulations ayant été générées avec des efficacités optiques de 70 % pour les télescopes de type HESS-I et de 80 % pour CT_5, il faudra faire attention à des biais systématiques dans l'estimation des surfaces de collection, en particulier près du seuil.

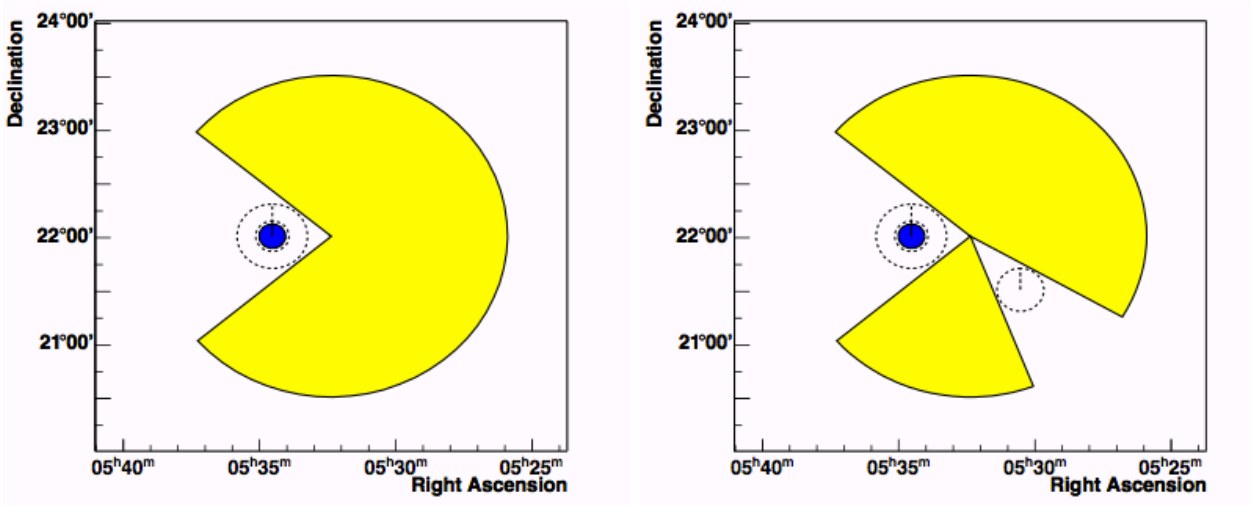


FIGURE 7.1: Illustration de la méthode permettant d'obtenir du fond réel (Crédits : [Marandon, 2010]). À gauche, la source est indiquée en bleu et la région d'exclusion de la source par des pointillés en noir. La zone jaune représente la partie du champ de vue sur laquelle sont intégrés les évènements de fond. À gauche, une zone d'exclusion correspondant à une source dans le champ de vue est ajoutée.

Dans la chaîne d'analyse sur laquelle nous travaillons, l'évaluation de l'énergie se fait à l'aide de tables dédiées qui sont produites à des énergies fixes, et l'évaluation des paramètres réduits (MSCL et MSCW, voir section 5) à partir de tables produites en utilisant des simulations générées en loi de puissance. Pour des raisons pratiques et du fait d'une équivalence supposée pour la simulation de CT_5 entre sim_hessarray et SMASH [Guy, 2003], nous avons utilisé pour ces paramètres des tables précédemment générées avec des simulations basées sur les logiciels KASKADE et SMASH. Ces tables ont été générées pour différents angles zénithaux, entre 0° et 70° par pas correspondant à 0,05 en $\cos\theta$, pour différents offsets compris entre 0° et 2,5° (par pas de 0,5°), et pour des efficacités optiques correspondant à 52,5 % et 78,0 % de la valeur nominale, respectivement pour les télescopes de taille moyenne et CT_5. Nous avons vérifié un bon comportement général sur l'estimation des paramètres réduits associés à des γ simulés avec CORSIKA et sim_hessarray en utilisant les tables produites avec KASCADE et SMASH. Pour l'énergie, les biais effectifs obtenus lors de l'utilisation d'une chaîne de simulation (avec des efficacités optiques données) avec les tables de calibration en énergie issues de l'autre chaîne (avec des efficacités optiques différentes, en particulier pour les télescopes de taille moyenne) n'ont que peu d'incidence sur la détermination de la sensibilité de l'instrument, puisque la surface de collection utilisée (voir équation 6.7) est estimée en fonction de l'énergie vraie moyenne associée à un intervalle en énergie mesurée, et qu'une méthode ad hoc a été mise en place pour établir proprement cette relation. Ceci est moins vrai pour les spectres préliminaires qui seront montrés à la fin de ce chapitre pour illustrer les résultats, car l'estimation de la vraisemblance utilise les IRF issues de la chaîne KASKADE+SMASH, parmi lesquels il y a les biais sur l'estimation de l'énergie. Ceci conduit à une prise en compte des biais qui peut être quelque peu erronée, surtout près du seuil. On reviendra sur ce point à la fin du chapitre.

7.3 Fond réel

La spécificité de ce chapitre vient de l'utilisation de fond réel pour la phase d'optimisation des coupures de discrimination γ/fond. Afin de caractériser de manière fine l'espace des phases des paramètres discriminants, l'utilisation d'un nombre important d'évènements de fond est essentiel. Ceux-ci ont été obtenus lors d'observations de champs de vue comportant dans la plupart des cas au moins une source. Un total de ∼125 h d'observation, correspondant à 311 runs passant les critères

de qualité explicitée à la section 4.5, ont été utilisées afin d'obtenir des évènements de fond. Afin d'éviter une contamination de l'échantillon par des γ, les sources situées dans des régions confuses dans lesquelles règne une forte émission diffuse γ, telles que Vela ou Sgr A*, ont été exclues de la sélection.

Après la calibration des données, les évènements ont été reconstruits avec la méthode des ellipses et le modèle 3D avec les réglages présentés dans la section 7.4.1. Leurs paramètres discriminants ont été calculés et stockés afin de pouvoir les utiliser ultérieurement lors des phases d'entraînement. La méthode d'extraction du fond est illustrée sur la Figure 7.1. Une région d'exclusion, représentée en pointillés noirs sur la figure, est définie autour de la source représentée en bleu. On intègre le fond pendant toute la durée de l'observation sur la région jaune (dite « *Pac-Man* »). Son rayon est déterminé en s'assurant que les évènements de fond qu'elle contient ont des variables discriminantes qui se comportent de manière similaire à celles du fond se trouvant à 0,5° d'offset (car on considère une source à 0,5° d'offset). Dans le cas où plusieurs sources sont présentes dans le champ de vue, le découpage est répété (à droite de la Figure 7.1). Les évènements de fond ont été intégrés jusqu'à un offset de 1,2° et des régions d'exclusion en anneau de rayon extérieur de 0,3° centrées sur la position des sources γ ont été utilisées.

7.4 Construction des analyses

Dans cette section, nous présentons la chaîne d'analyse permettant l'entraînement des arbres de décision boostés et la recherche des points de fonctionnement optimaux pour différents types de sources.

7.4.1 Développement et caractéristiques des analyses

Les analyses développées par la suite utiliseront, afin de présélectionner les images qui serviront lors de la reconstruction géométrique des évènements et pour le calcul des paramètres discriminants, les mêmes réglages que ceux présentés à la section 6.4.1. Pour rappel, ces réglages sont :

- $Q^{\text{min}} = 40$ p.e.
- $L_{\text{D}}^{\text{max}} = 0{,}8$ m pour CT_5 et $L_{\text{D}}^{\text{max}} = 0{,}525$ m pour les caméras de type HESS-I
- $N_{\text{pix}}^{\text{min}} = 4$
- $\sin\theta_{ij} = 0{,}3$

Seuls les évènements pour lesquels l'ajustement du modèle 3D converge seront gardés. La même présélection sera utilisée pour les images provenant de la projection, dans le plan focal des télescopes, de la photosphère ajustée par le modèle 3D (voir section 6.4.2). On trouvera une description et une discussion détaillée de la préparation de l'analyse à la section 6.5.1.

Cette analyse (de la calibration jusqu'à la détermination des variables discriminantes) sera appliquée aux simulations γ pour le lot signal, avec une coupure angulaire de $\theta_{\text{ON}} = 0{,}15°$ qui correspond à la précision angulaire autour de $\sim$100 GeV déterminée au chapitre précédent pour la configuration hybride. L'analyse sera aussi appliquée au fond réel, sélectionné selon la procédure décrite à la section précédente.

L'optimisation du pouvoir de séparation γ/fond nécessite un découpage de l'espace des phases qui est effectué selon la multiplicité, l'angle zénithal et le nombre N_c de photons reconstruits par le modèle 3D de la photosphère Tcherenkov des évènements. On séparera le traitement des évènements selon leur multiplicité (deux et supérieure à deux), selon leur angle zénithal (sept intervalles suivant les spécificités des simulations détaillées à la section 7.2), et enfin selon le nombre de photons Tcherenkov N_c qui composent la photosphère des gerbes reconstruites par le modèle 3D. Ce dernier découpage a été réalisé pour chaque angle zénithal et multiplicité donnés en fonction de la statistique disponible pour l'entraînement. L'algorithme suivant a été choisi. S'il est possible d'avoir au moins cinq intervalles N_c, chacun ayant au moins 50 000 évènements γ et de fond passant les coupures de

pré-sélection, on utilisera autant d'intervalles qu'il est possible de construire. Dans le cas contraire, selon la statistique disponible, on définira cinq intervalles en N_c avec au moins 40 000, 30 000, 20 000 ou 10 000 évènements. Si ce dernier critère n'est pas vérifié, on utilisera un unique intervalle en N_c avec au moins 5000 évènements. Après application de cette procédure, l'espace des phases est découpé en 61 intervalles.

On utilisera la méthode d'analyse multivariée BDT pour la séparation des γ et du fond avec des réglages différents de ceux présentés au chapitre précédent. La méthode de « *pruning* » ne sera pas utilisée. On préférera l'utilisation d'arbres de faibles profondeurs (trois nœuds au maximum) qui composeront une forêt d'arbres de décision constituée d'un grand nombre d'individus (300). Aucune différence significative dans les résultats n'a été observée par rapport aux réglages précédents, cependant les créateurs de la librairie TMVA recommandent fortement cette solution [1].

7.4.2 Optimisation des coupures de discrimination γ/fond

Le découpage de l'espace des phases étant défini pour l'entraînement et la méthode BDT étant réglée, il convient maintenant d'obtenir pour chacun des intervalles la valeur finale de discrimination γ/fond, que l'on notera $\zeta^\star$. Une nouvelle méthode a été développée afin de choisir la valeur de $\zeta^\star$ qui conduira à obtenir le niveau de significativité maximal d'une source type, définie par un niveau de flux $\phi(E_{\rm Ref}) \equiv \phi_0$, à une énergie $E_{\rm Ref}$, et par un indice spectral Γ. On notera par la suite N_γ le nombre total d'évènements γ simulés passant les coupures de pré-sélection dans un intervalle de l'espace des phases et obtenu dans une région circulaire de rayon $\theta_{\rm ON}$. De la même manière, le nombre d'évènements de fond sélectionnés dans la région « Pac-Man » de rayon caractéristique $\theta_{\rm OFF}$, sera noté $N_{\rm OFF}$.

Afin de déterminer $\zeta^\star$, on utilisera l'estimation de la significativité donnée par l'équation 5.35. Le point critique est la pondération correcte à attribuer aux évènements γ issus des simulations appartenant à l'intervalle de l'espace des phases étudié, sachant que les évènements de fond ont été intégrés pendant un temps $T^{\rm eff}_{\rm Obs}$. Connaissant ce temps d'observation $T^{\rm eff}_{\rm Obs}$, on peut attribuer à chaque évènement γ simulé un poids, que l'on notera $w_\gamma(E)$ et qui dépend de son énergie vraie E :

$$w_\gamma(E) = \frac{\phi_0 \left(\frac{E}{E_{\rm Ref}}\right)^{-\Gamma}}{\phi_0^{\rm simu} \left(\frac{E}{E_{\rm Ref}}\right)^{-\Gamma^{\rm simu}}} \tag{7.1}$$

où $\phi_0^{\rm simu}$ est déterminé à partir du nombre de γ simulés $N^{\rm simu}$ entre les énergies $E^{\rm min}$ et $E^{\rm max}$:

$$\phi_0^{\rm simu} = \frac{N^{\rm simu}}{T^{\rm eff}_{\rm Obs} \mathcal{A}^\gamma \int_{E^{\rm min}}^{E^{\rm max}} dE \left(\frac{E}{E_{\rm Ref}}\right)^{-\Gamma^{\rm simu}}} \tag{7.2}$$

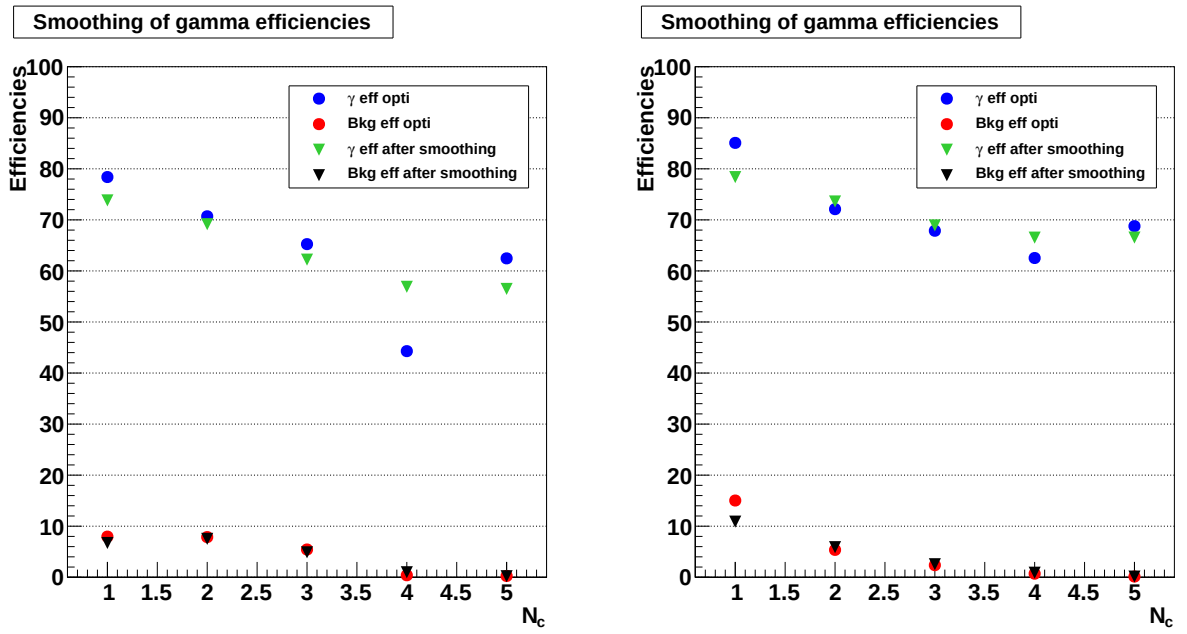
où $\mathcal{A}^\gamma$ est la surface sur laquelle les γ ont été générés. On peut alors exprimer le nombre total d'évènements attendus pendant un temps $T^{\rm eff}_{\rm Obs}$, que l'on notera S, émis par une source définie par les paramètres ϕ_0 et Γ selon :

$$S = \sum_{i=1}^{N_\gamma} (1 \times w_\gamma(E)) \tag{7.3}$$

Le nombre total d'évènements de fond $N_{\rm OFF}$ doit lui aussi être corrigé. Afin de mimer l'analyse d'une source réelle, nous définissons cinq régions ($\alpha = 1/5$) OFF de même taille angulaire que la région ON. Le nombre d'évènements de fond attendus, que l'on notera B, peut alors être estimé suivant :

$$B = \sum_{i=1}^{N} \frac{1}{\alpha} \left(\frac{2\pi(1-\cos\theta_{\rm ON})}{2\pi\eta_i(1-\cos\theta_{\rm OFF})} \right) N^i_{\rm OFF} \tag{7.4}$$

1. On pourra lire le guide d'utilisateur de la librairie TMVA ou se référer à la discussion ci-après : http://ehc.ac/p/tmva/mailman/message/33042540/



(a) Multiplicité deux.

(b) Multiplicité supérieure à deux.

FIGURE 7.2: Illustration du lissage des efficacités γ pour la configuration optimisée pour un flux de l'ordre de 1 % du Crabe à 300 GeV, pour un angle zénithal compris entre 20 et 30°.

où η_i est le rapport entre la surface « Pac-Man » et la surface circulaire de rayon θ_{OFF} pour le « run » d'observation i ($i = [1; N]$).

En utilisant le fait que $N_{\mathrm{ON}} = S + \alpha N_{\mathrm{OFF}}$, le niveau de significativité N_σ peut être obtenu en réécrivant l'équation 5.35 en fonction de ζ :

$$N_\sigma(\zeta) = \frac{\varepsilon_{\mathrm{Sig}}(\zeta) S}{\sqrt{\varepsilon_{\mathrm{Sig}}(\zeta) S + (\alpha+1)\varepsilon_{\mathrm{Bkg}}(\zeta) B}} \tag{7.5}$$

où $\varepsilon_{\mathrm{Sig}}(\zeta)$ et $\varepsilon_{\mathrm{Bkg}}(\zeta)$ représentent respectivement les efficacités de reconstruction des évènements γ et de fond. La valeur finale de discrimination $\zeta^\star$ vérifie :

$$\left.\frac{\mathrm{d}N_\sigma(\zeta)}{\mathrm{d}\zeta}\right|_{\zeta=\zeta^\star} = 0 \tag{7.6}$$

7.4.3 Lissage des efficacités γ

Afin d'avoir des surfaces efficaces de collection ne présentant pas de variation brutale en fonction de l'énergie [2] nous avons utilisé une méthode permettant de lisser les efficacités de reconstruction γ en fonction des intervalles en N_c, pour un angle zénithal et une multiplicité donnés. En lieu et place d'une méthode d'ajustement avec un polynôme, une méthode de régression non-paramétrique [3] a été utilisée afin d'obtenir une tendance de l'efficacité de reconstruction γ en fonction de l'intervalle en N_c.

On peut observer sur la Figure 7.2 deux exemples de lissage des efficacités γ en fonction de l'intervalle en N_c, pour la multiplicité deux (à gauche) et les multiplicités supérieures à deux (à droite). Ces intervalles correspondent à des angles zénithaux compris entre 20 et 30°. Les efficacités γ et de fond dérivées à partir du maximum de significativité dans un intervalle de l'espace des

2. Cela peut arriver si pour un angle zénithal et une multiplicité donnés, les efficacités de reconstruction γ ne présentent pas un comportement continu en fonction de l'intervalle en N_c

3. La méthode de Narayada-Watson [Nadaraya, 1964].

Configurations	E_{Ref}	$\phi_0(E_{\text{Ref}})$ [cm^{-2} s^{-1} TeV^{-1}]	Γ
Brillante	1 TeV	$3{,}45 \times 10^{-11}$	2,6
Intermédiaire	1 TeV	$2{,}0 \times 10^{-12}$	3,3
Faible	300 GeV	$8{,}0 \times 10^{-12}$	3,5

TABLE 7.1: Paramètres spectraux utilisés pour le développement des différentes configurations.

phases sont respectivement représentées par des points bleus et rouges. Les triangles verts et noirs représentent les nouveaux points de fonctionnement après lissage des efficacités γ. On peut observer un cas problématique sur la Figure 7.2a (quatrième intervalle de l'espace des phases), qui présente une efficacité de reconstruction γ en deçà de la tendance observée. L'application de la méthode de lissage permet de retrouver un comportement continu qui n'entraînera pas de variation brusque de la surface efficace de collection finale.

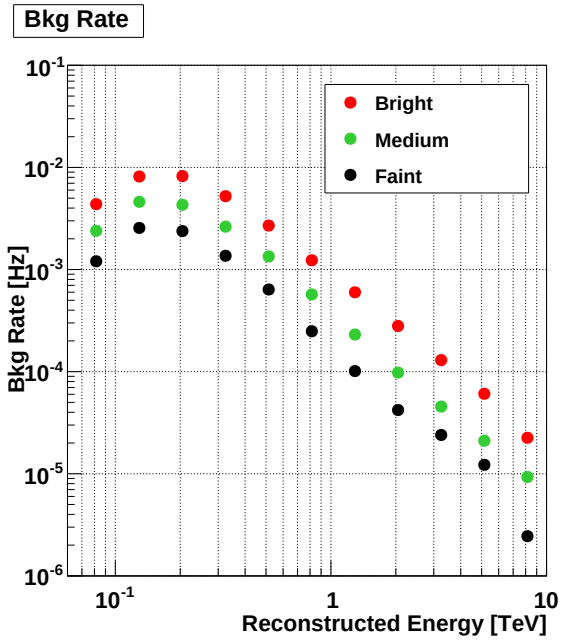
7.5 Configurations d'analyse et performances

Afin de préparer l'analyse des futures données, l'approche présentée dans la section 7.4.2 va être optimisée de manière indépendante pour trois types de sources différentes. Celles-ci correspondent à des sources fortes (on optimisera sur le spectre de la nébuleuse du Crabe), à des sources d'intensité moyenne (on optimisera sur le spectre du BL Lac PKS 2155-304 dans son état bas) et à des sources faibles (on optimisera sur un spectre en loi de puissance normalisé à 1 % du flux de la nébuleuse du Crabe à 300 GeV). Les caractéristiques des configurations ainsi définies sont résumées dans la Table 7.1.

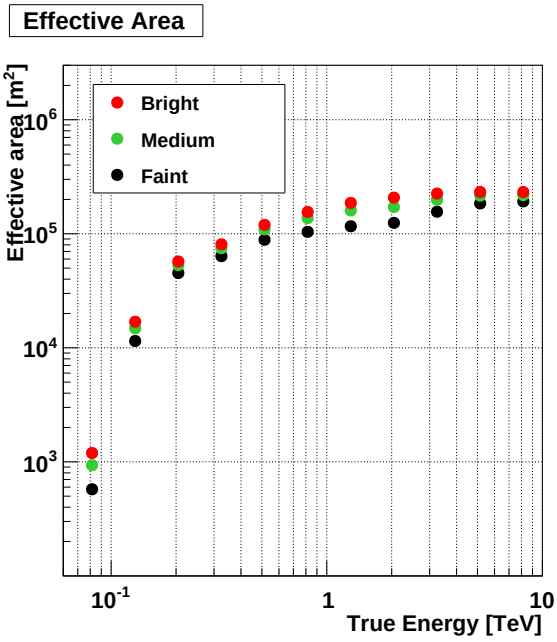
Pour chacune de ces trois configurations, la recherche des coupures optimales de discrimination est faite dans chacun des 61 intervalles de l'espace des phases. Pour chaque bande en angle zénithal et en multiplicité, les efficacités γ sont lissées comme expliqué à la section 7.4.3. Un exemple de l'évolution des efficacités ε_{Sig} et ε_{Bkg} en fonction de N_c est illustré sur la Figure 7.2.

À partir des efficacités ainsi obtenues, nous avons déterminé pour chacune des configurations ses performances (Figure 7.3). En haut à gauche est représenté le taux de fond en fonction de l'énergie reconstruite des évènements (en Hz) pour les trois configurations : en rouge pour les sources brillantes, en vert pour les sources intermédiaires et en noir pour les sources faibles. Sur la figure de droite, la surface efficace de collection en fonction de l'énergie vraie est représentée pour ces mêmes configurations. Comme attendu, les deux configurations pour l'étude des sources brillantes laissent passer plus de fond mais permettent de récolter plus d'évènements γ que la configuration dédiée à l'étude des sources faibles. Notons que pour des questions liées à la faible statistique disponible pour les données de fond, les taux de fond ici présentés ont été évalués sur un lot d'évènements dont une fraction importante (jusqu'à 60 % pour certains intervalles de l'espace des phases) peut avoir été utilisée lors de la phase d'entraînement. Les taux de fond peuvent ainsi avoir été légèrement sous-évalués et il conviendra de les re-déterminer lorsque davantage de fond sera disponible.

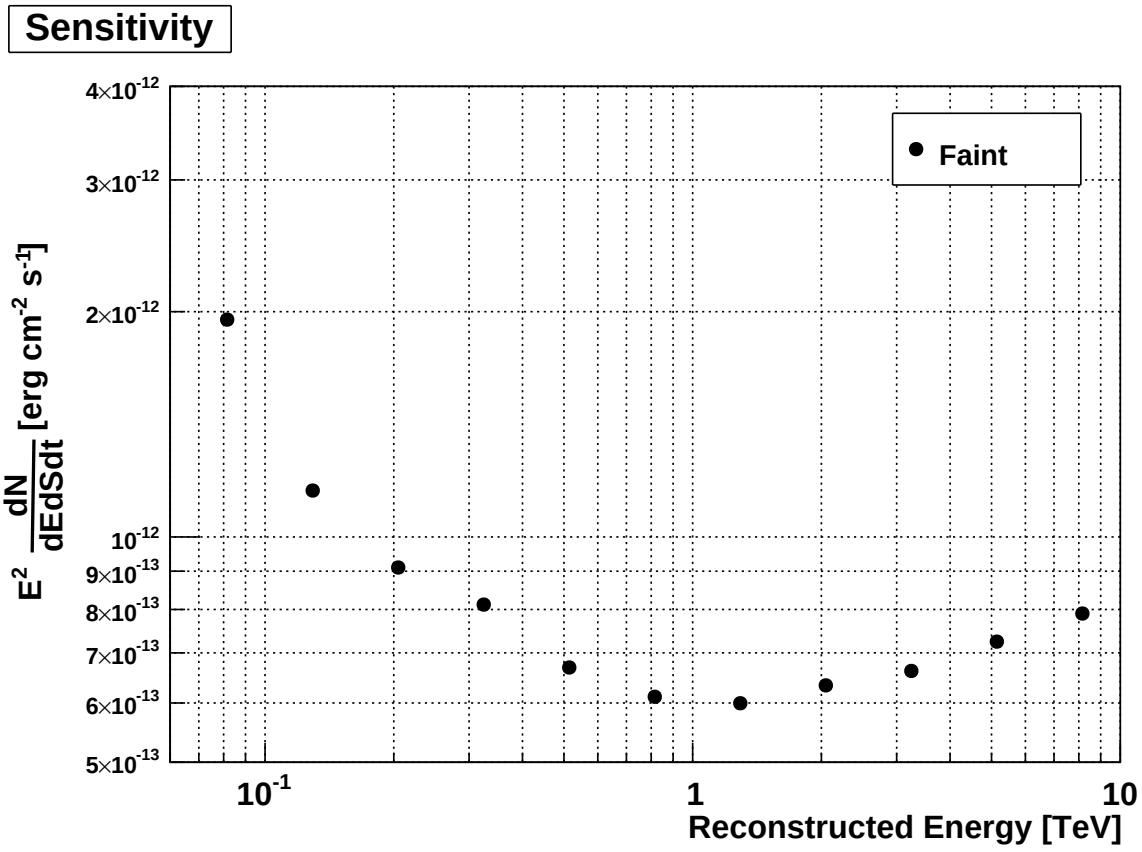
La connaissance du taux de fond et de la surface de collection permet de déterminer la sensibilité de l'instrument, c'est-à-dire le flux minimal que l'on doit pouvoir détecter dans des conditions données. Ce flux minimal est naturellement évalué avec la configuration dédiée aux sources faibles. Celui-ci est représenté sur la Figure 7.3c pour un temps d'observation de 50 h et pour une détection à 5σ dans chaque intervalle d'énergie reconstruite (5 par décade). On constate que l'on obtient une sensibilité légèrement meilleure que le pourcent du Crabe en dessous de $\sim$600 GeV et qu'entre 1 et 10 TeV on obtient une sensibilité qui croît en fonction de l'énergie et qui est comprise entre 1 % et 8 % du flux de la nébuleuse du Crabe.



(a)



(b)



(c)

FIGURE 7.3: (a) Taux de fond en fonction de l'énergie reconstruite pour les trois configurations. (b) Surface efficace de collection en fonction de l'énergie vraie pour les trois configurations. (c) Sensibilité différentielle en fonction de l'énergie reconstruite de la configuration dédiée à l'étude des sources faibles pour un temps d'observation de 50 h et pour une détection à 5σ dans chaque intervalle d'énergie reconstruite.

	MJD	$T_{\mathrm{Obs}}^{\mathrm{eff}}$ (h)	$\langle\theta\rangle$ (°)	N_{ON}	N_{OFF}	α	Excès	N_σ	$N_\sigma/\sqrt{t(\mathrm{h})}$
Crabe	56596-56602	9,10	46,5	5108	6828	0,081	4554	106,0	35,1
PKS 2155-304	56405-56600	23,4	13,3	3718	17 528	0,104	1898	36,5	7,6
PG 1553+113	56596-56602	11,8	35,2	479	1916	0,125	240	12,6	3,7
Mkn501 (I)	56827-56867	7,4	63,8	337	128	0,125	322	31,4	11,5
Mkn501 (II)	56831-56831	1,8	63,5	245	47	0,125	239	28,8	21,6

TABLE 7.2: Résultats des observations de *H.E.S.S.* sur une sélection de sources. Les jours correspondant au premier et au dernier « run » d'observation utilisés pour l'analyse sont indiqués dans MJD. On trouvera aussi l'angle zénithal moyen d'observation, et le niveau de significativité par racine d'heure. Les autres quantités ont été définies dans cette thèse. Deux lots de données différents sont utilisés pour l'analyse de Mkn 501.

7.6 Détection des premières sources

Nous présentons ci-dessous les résultats de l'analyse des premières observations avec HESS-II pour quatre sources présentant des niveaux de brillance et des caractéristiques spectrales différents : la nébuleuse du Crabe, Mkn 501, PKS 2155-304 et PG 1553+113.

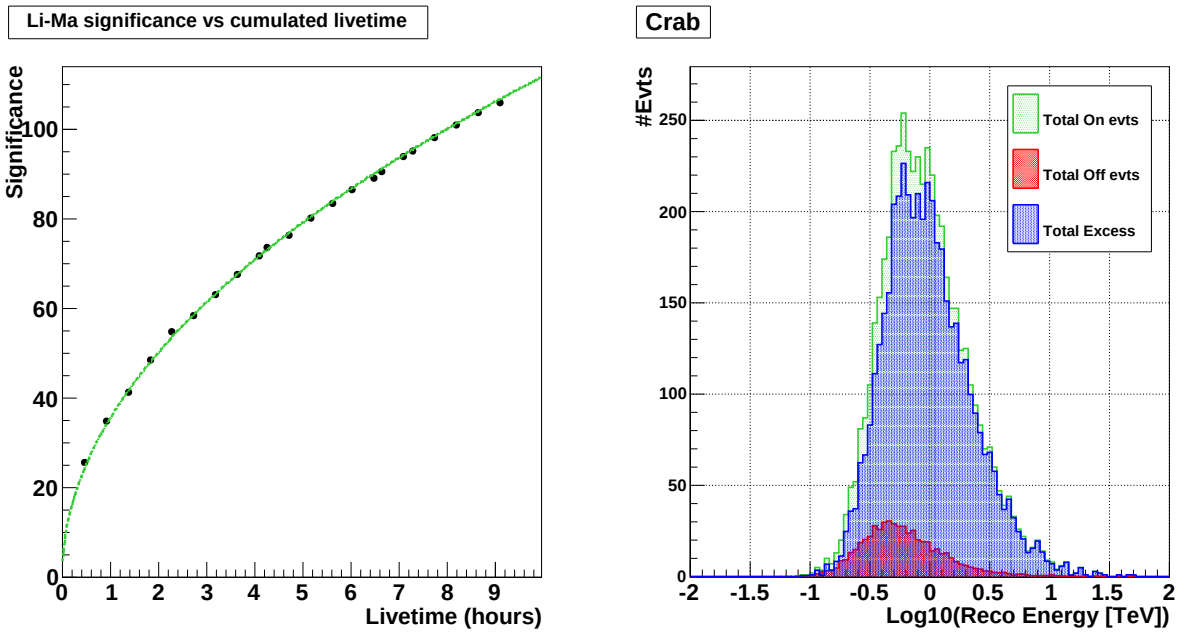
Après calibration des données, la chaîne d'analyse décrite en détail à la section 6.5.1 et dont les réglages ont été précisés aux sections 7.4 et 7.5 a été utilisée. S'agissant ici d'un signal sur des données réelles, la méthode des OFF multiples a été appliquée (section 5.3.2). La sélection en qualité des « runs » a été réalisée de la même manière qu'à la section 7.3. Les résultats en termes d'extraction du signal de chacune des analyses sont résumés dans la Table 7.2.

7.6.1 Nébuleuse du Crabe

Située à une distance de ∼2 kpc, la nébuleuse du Crabe est une nébuleuse à vent de pulsar utilisée comme source de référence en astronomie de très haute énergie du fait de sa très grande brillance. Du domaine radio jusqu'à quelques GeV l'émission de la source est très bien reproduite par l'émission synchrotron d'électrons relativistes. À plus haute énergie, l'interaction des électrons relativistes avec les photons produits par rayonnement synchrotron (processus SSC) devient le processus dominant. Utilisée généralement comme source de calibration pour son absence de variabilité, la découverte de sursauts par les satellites *AGILE* [Tavani et al., 2011] et *Fermi-LAT* [Abdo et al., 2011] a surpris la communauté. Dans les cas extrêmes, une variation d'au moins un facteur 10 a été relevée dans le flux de la source en moins d'une journée entre 100 MeV et 1 GeV. Ces observations couplées au fait qu'aucune variation du signal pulsé n'a été observée conduisent à situer la zone d'émission au-delà du cylindre de lumière (on pourra consulter la discussion récente de [Aliu et al., 2014] ainsi que les références associées). Dans le domaine γ de très haute énergie, aucune variation significative du flux n'a été observée par les expériences *MAGIC* et *VERITAS*. Les pulsations du pulsar de la nébuleuse du Crabe ont été détectées par *MAGIC* au-delà de 25 GeV [Aliu et al., 2008] et par *VERITAS* au-delà de 100 GeV [VERITAS Collaboration et al., 2011]. La récente détection d'un signal pulsé significatif entre 400 GeV et ∼2 TeV par *MAGIC* en 300 h [4], laisse à penser que sa détection par *H.E.S.S.* (bien que la source soit visible au mieux avec un angle zénithal de ∼45°) pourrait être possible si un temps d'observation conséquent y était consacré.

La configuration d'analyse optimisée pour l'étude des sources brillantes a été appliquée à la nébuleuse du Crabe, et a permis d'obtenir un niveau de détection de 106σ en 9,1 h ($35{,}1\,\sigma\sqrt{\mathrm{h}^{-1}}$), correspondant à un excès de 4554 évènements. Un taux très important de $(8{,}3 \pm 0{,}1)\,\gamma\,\mathrm{min}^{-1}$ est mesuré, pour un taux d'évènements de fond de 1,0 par minute. À titre de comparaison, l'analyse « *loose* » utilisée dans la publication [Aharonian et al., 2006] permettait d'obtenir un taux de γ par minute de $12{,}2 \pm 0{,}2$ pour un taux de significativité horaire de $23{,}1\,\sigma\sqrt{\mathrm{h}^{-1}}$ (avec une coupure angulaire plus lâche). L'évolution de la significativité en fonction du temps d'observation est représentée

4. http://fermi.gsfc.nasa.gov/science/mtgs/symposia/2014/abstracts/185



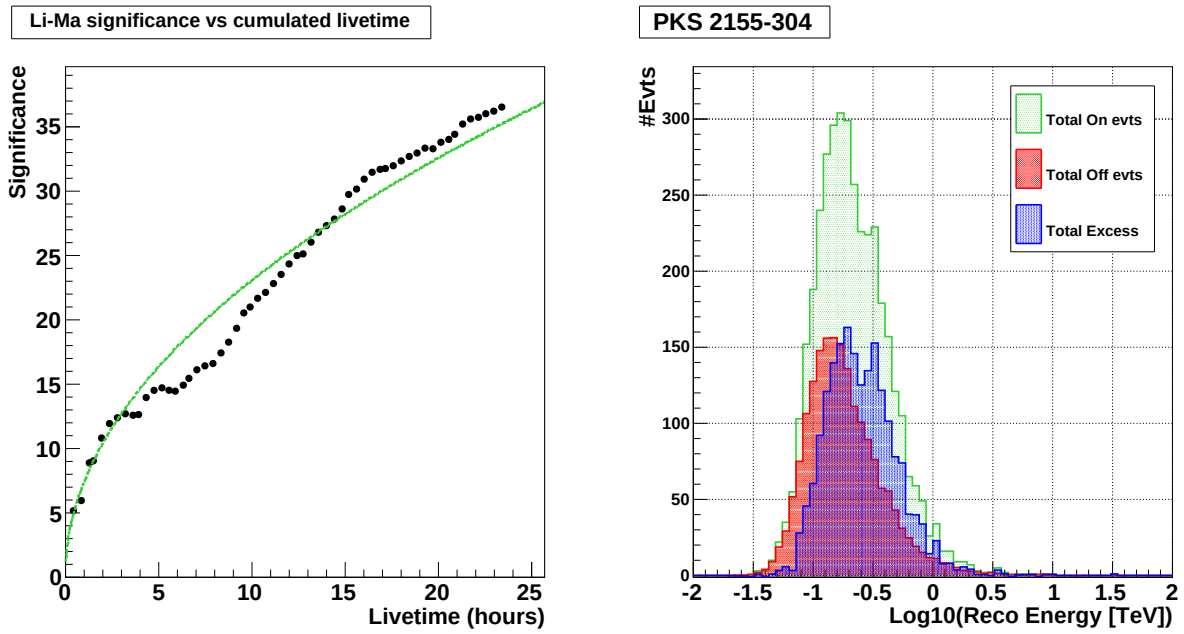
(a) (b)

FIGURE 7.4: Détection de la nébuleuse du Crabe. (a) Significativité en fonction du temps d'observation cumulé. (b) Distributions de l'énergie reconstruite des évènements de la région ON (vert), des évènements des régions OFF (après normalisation à la région ON, rouge) et de l'excès mesuré (bleu).

sur la Figure 7.4a (points noirs) ainsi qu'un ajustement de celle-ci par une fonction en racine carré du temps (courbe verte). On peut observer sur la Figure 7.4b la distribution de l'énergie reconstruite des évènements de la région ON (vert), du fond résiduel (rouge) et enfin de l'excès (en bleu). Un excès significatif est mesuré dès ∼200 GeV, et ce malgré un angle zénithal moyen d'observation de ∼45°.

7.6.2 PKS 2155-304

La source PKS 2155-304 est un BL Lac situé à une distance correspondant à $z = 0{,}116$ et est fréquemment observé par le réseau *H.E.S.S.* Elle est relativement forte, avec un flux moyen de son état bas estimé entre 2005 et 2007 à ∼15 % du flux de la nébuleuse du Crabe [H.E.S.S. Collaboration, 2010]. Son état haut détecté lors de la nuit du 28 juillet 2006 a permis de mettre en évidence la variabilité la plus courte jamais mesurée pour un AGN, de l'ordre de la minute [Aharonian et al., 2007]. Lors de ce sursaut, une modélisation de type SSC dépendante du temps a été réalisée, rendant compte de la courbe de lumière observée [Katarzyński et al., 2008]. Suite au premier sursaut de la nuit du 28 elle a fait l'objet de nombreuses études pour caractériser son comportement multi-longueur d'onde [Aharonian et al., 2009a, H.E.S.S. Collaboration et al., 2012]. Peu après le lancement du satellite *Fermi-LAT*, une campagne simultanée lors d'un état bas de PKS 2155-304 regroupant les expériences *H.E.S.S.*, *Fermi-LAT*, SWIFT/XRT, XRTE et ATOM a permis d'étudier la variabilité entre différentes longueurs d'onde, et a mis en évidence la difficulté des modèles SSC à une zone à reproduire les différentes corrélations observées [Aharonian et al., 2009b]. Elle est également utilisée comme sonde pour différentes questions de physique, telles que la détermination de la densité du rayonnement du fond extragalactique infrarouge [H.E.S.S. Collaboration et al., 2013] ou la recherche d'une éventuelle violation de l'invariance de Lorentz. Dans le deuxième cas, une éventuelle variation de la vitesse de la lumière en fonction de l'énergie est recherchée en étudiant les délais temporels dans les courbes lumière entre différentes bandes en énergie. Du fait de la statistique très importante



(a) (b)

FIGURE 7.5: Détection du BL Lac PKS 2155-304. (a) Significativité en fonction du temps d'observation cumulé. (b) Distributions de l'énergie reconstruite des évènements de la région ON (vert), des évènements des régions OFF (après normalisation à la région ON, rouge) et de l'excès mesuré (bleu).

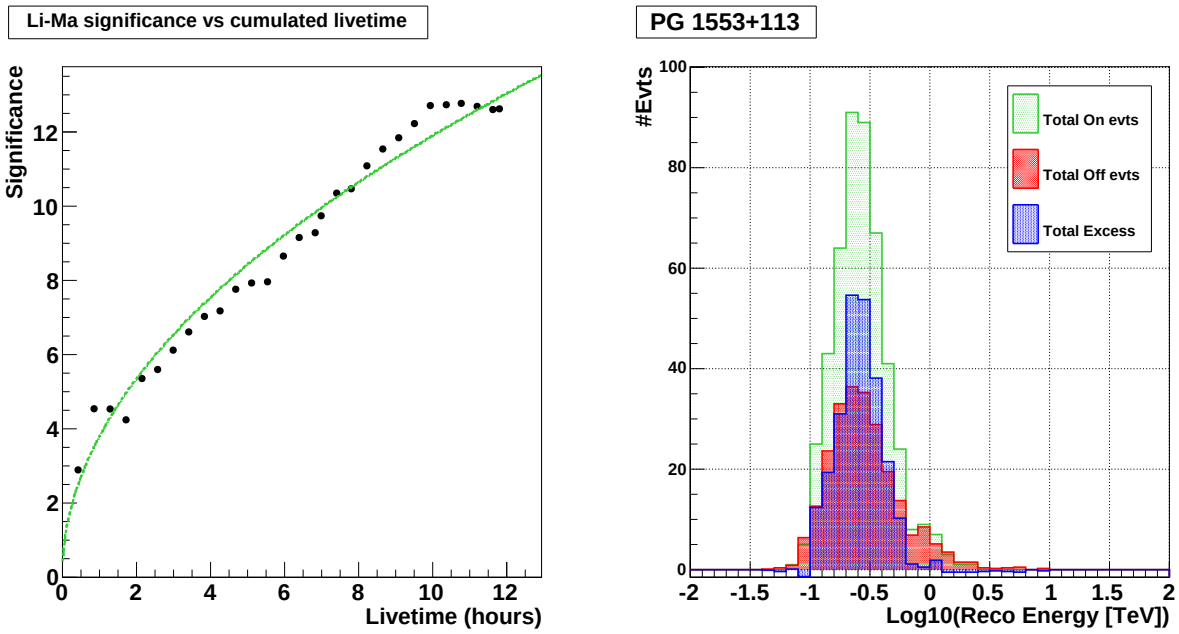
de photons γ récoltés lors de l'état haut de PKS 2155-304 en juillet 2006, la limite supérieure la plus contraignante avec des blazars a été obtenue [Aharonian et al., 2008a, H.E.S.S. Collaboration, 2011]. PKS 2155-304 est aussi utilisé afin de rechercher d'hypothétiques particules telles que les axions, initialement proposés pour résoudre certains problèmes liés aux violations de symétries dans les lois de la physique, en cherchant d'éventuelles irrégularités dans son spectre [Brun et al., 2013].

Le BL Lac PKS 2155-304 est détecté avec HESS-II en 23,4 h avec une significativité de 36,5 σ (7,6 $\sqrt{\text{h}^{-1}}$) avec la configuration d'analyse « intermédiaire » (optimisée pour 15 % du Crabe). L'évolution de la significativité en fonction du temps montrée sur la Figure 7.5a est caractéristique d'une source variable. On peut voir sur la Figure 7.5b que l'excès mesuré est significatif à partir de ~80 GeV [5]. Cet excès correspond à ~1900 évènements pour un taux de $(1{,}35 \pm 0{,}03)\,\gamma\,\text{min}^{-1}$ (et un taux de fond de $(1{,}30 \pm 0{,}01)\,\text{min}^{-1}$). Ce taux de γ est plus faible que celui mesuré avec HESS-I entre 2005 et 2007, ce qui semble indiquer une diminution du flux de la source. On reviendra sur ce point à la section 7.7.2.

7.6.3 PG 1553+113

Le second blazar analysé est le BL Lac PG 1553+113 qui a montré, lors d'une précédente observation, un flux au-delà de 300 GeV correspondant à ~3,4 % du flux de la nébuleuse du Crabe [Aharonian et al., 2008b]. Cette source possède un spectre très mou, typique des AGN distants, et est de ce fait détectable principalement à basse énergie. Malgré de nombreuses campagnes d'observation en optique, aucune mesure de redshift n'a été possible ce qui laisse supposer que son continuum optique est noyé par l'émission non-thermique du jet. La meilleure contrainte est donnée par les méthodes spectroscopiques avec la détection solide avec HST de deux raies d'émission $\text{Ly}\alpha + \text{O}\,VI$ correspondant à un système intermédiaire situé à un redshift de $z = 0{,}395$ [Danforth et al., 2010]. Des

5. PKS 2155-304 passe à 7° du zénith sur le site de l'expérience *H.E.S.S.*.



(a) (b)

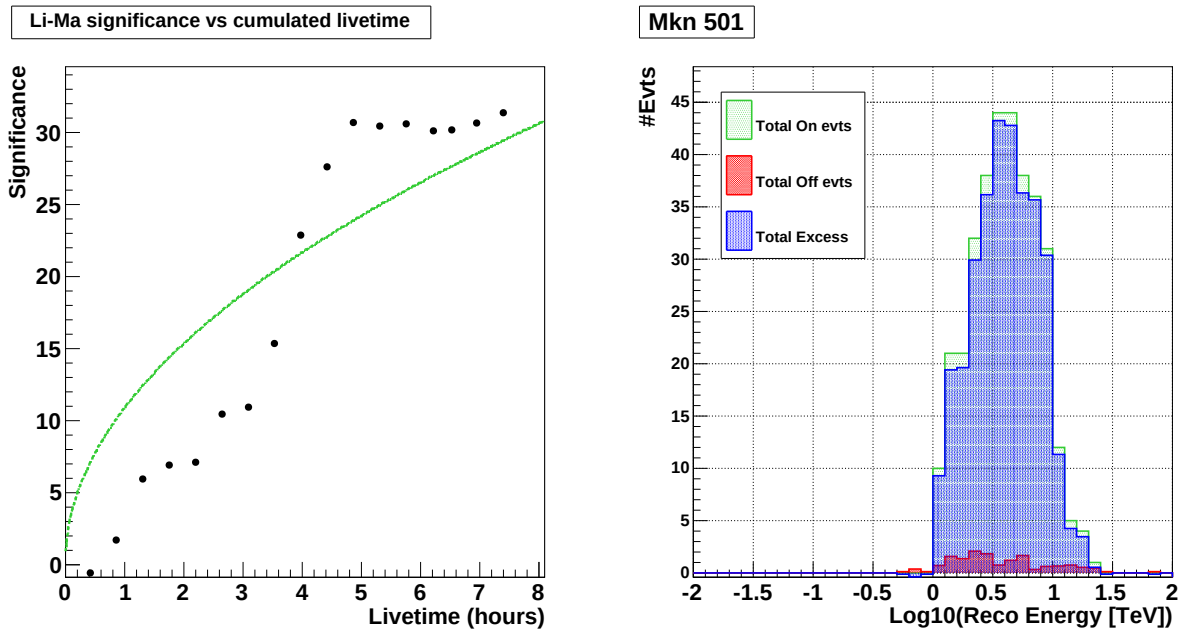
FIGURE 7.6: Détection du BL Lac PG 1553+113. (a) Significativité en fonction du temps d'observation cumulé. (b) Distributions de l'énergie reconstruite des évènements de la région ON (vert), des évènements des régions OFF (après normalisation à la région ON, rouge) et de l'excès mesuré (bleu).

contraintes de redshift sont également obtenues par des méthodes indirectes conduisant à $z < 0{,}67$ [Prandini et al., 2009], $z < 0{,}64$ [Sanchez et al., 2013], $z < 0{,}58$ [Danforth et al., 2010], $z \sim 0{,}49$ (publication de la collaboration *H.E.S.S.*, accepté pour publication) et $z \sim 0{,}36$ [Aleksić et al., 2014]. Dans ce dernier travail, la distribution spectrale en énergie est modélisée à l'aide d'une approche SSC.

La source a été analysée avec la configuration dédiée aux sources faibles. Elle est détectée en 11,8 h avec un taux horaire de significativité de $3{,}7\,\sigma\sqrt{\mathrm{h}^{-1}}$, un taux de γ de $(0{,}34 \pm 0{,}03)\,\gamma\,\mathrm{min}^{-1}$ et un taux de fond de $(0{,}19 \pm 0{,}01)\,\mathrm{min}^{-1}$. La distribution de l'excès est représentée en bleu sur la Figure 7.6b, celui-ci est concentré en dessous de 630 GeV et centré vers 250 GeV. On peut observer sur la Figure 7.6a l'évolution de la significativité en fonction du temps d'observation, elle s'écarte légèrement de la tendance attendue pour une source stable.

7.6.4 Mkn 501

Mkn 501 est un BL Lac très proche ($z = 0{,}034$) qui a montré en 1997 des périodes de variabilité temporelle importantes lors desquelles il a atteint des flux un ordre de grandeur supérieurs à celui de la nébuleuse Crabe [Protheroe et al., 1997, Djannati-Atai et al., 1999, Aharonian et al., 1997]. Les spectres obtenus lors de ces sursauts ont été décrits avec une modélisation SSC qui permet de rendre compte des corrélations observées entre la variabilité dans le domaine des X (BeppoSAX) et dans le domaine γ de très haute énergie [Djannati-Atai et al., 1999]. Une corrélation entre la dureté du spectre et le niveau de flux a été mise en évidence. En 2005, une variation correspondant au doublement du flux en $\sim$2 min a été observée par le télescope *MAGIC* [Albert et al., 2007] permettant de rechercher une éventuelle violation de l'invariance de Lorentz [MAGIC Collaboration et al., 2008]. Du fait de sa position dans le ciel, le réseau de télescopes *H.E.S.S.* ne peut observer cet objet qu'à très haut angle zénithal (les observations présentées ici ont été réalisées à $\theta \approx 63°$) ce qui permet de caractériser son émission au-delà de 2 TeV dans de bonnes conditions.



(a) (b)

FIGURE 7.7: Détection du BL Lac Mkn 501. (a) Significativité en fonction du temps d'observation cumulé. (b) Distributions de l'énergie reconstruite des évènements de la région ON (vert), des évènements des régions OFF (après normalisation à la région ON, rouge) et de l'excès mesuré (bleu).

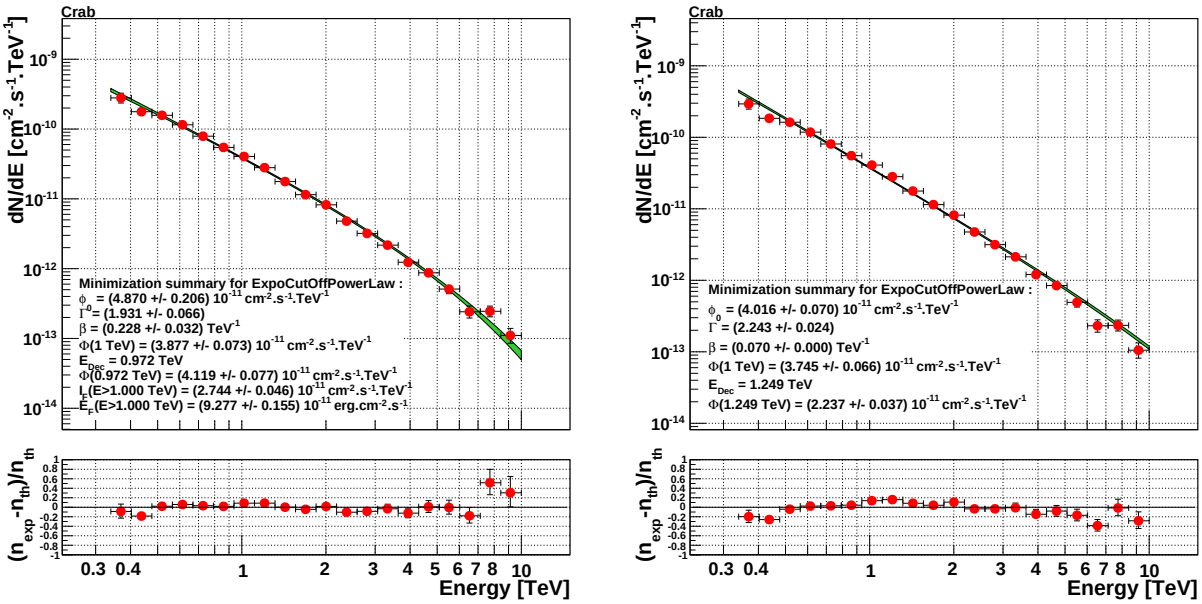
L'évolution de la significativité en fonction du temps d'observation de Mkn 501, obtenue avec l'analyse utilisée pour les sources brillantes, est représentée sur la Figure 7.7a. On peut observer une période lors de laquelle la source est particulièrement active (correspondant à quatre « runs » d'observation à la date 56831 MJD, dénotés par « II » dans la Table 7.2). Lors de cette nuit, un excès de ∼290 évènements a été mesuré en 1,8 h d'observation, ce qui correspond à un taux de $(2{,}25 \pm 0{,}15)\,\gamma\,\mathrm{min}^{-1}$ et à une significativité horaire de $\sim 21{,}6\,\sigma\sqrt{\mathrm{h}^{-1}}$. On peut observer sur la Figure 7.7b les distribution de l'excès (bleu) et le fond résiduel négligeable (correspondant à un taux mesuré de $(0{,}05 \pm 0{,}01)\,\mathrm{min}^{-1}$).

7.7 Spectres

Dans cette section nous présentons les analyses spectrales préliminaires pour les sources précedemment étudiées (la nébuleuse du Crabe, PKS 2155-304, PG 1553+113 et Mkn 501). Les résultats ci-dessous sont à prendre avec précaution car les IRF utilisées ont été produites, pour les télesopes de taille moyenne, avec des efficacités optiques plus élevées (70 %) que celles déterminées expérimentalement (∼60 %). On utilisera un seuil en énergie plus haut (déterminé par une surface de collection minimale de 4 ha) afin de minimiser l'effet de ces systématiques sur les résultats. Le cas de Mkn 501 fera l'objet d'un traitement particulier.

7.7.1 Nébuleuse du Crabe

L'évaluation du spectre de la nébuleuse de Crabe nécessite l'utilisation d'IRF générées à 45° et 50° d'angle zénithal. Or, les simulations à 45° (50°) d'angle zénithal sont générées jusqu'à une énergie maximale de 20 TeV (30 TeV). Cela a une conséquence sur l'évaluation du signal attendu dans les intervalles en énergie reconstruite proches, qu'il ne sera pas possible d'estimer correctement au-delà



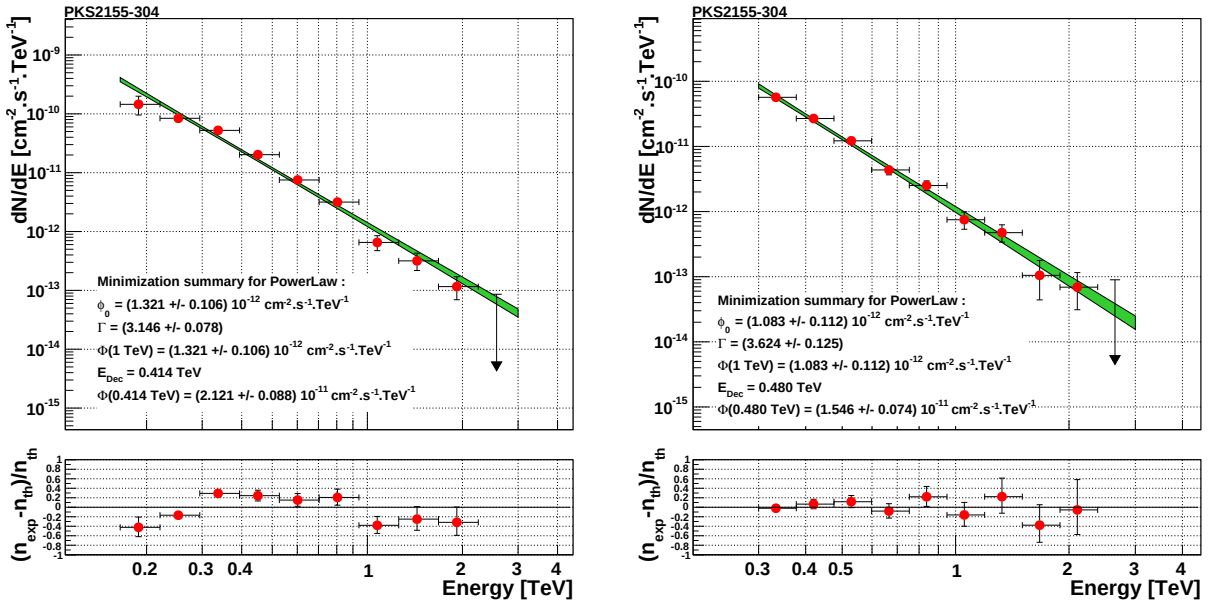
(a) (b)

FIGURE 7.8: Spectres différentiels de la nébuleuse du Crabe (a) pour une hypothèse de loi de puissance avec une coupure exponentielle ($E_{\mathrm{Ref}} = 1\,\mathrm{TeV}$) (b) idem, avec l'énergie de coupure exponentielle fixée à 14,3 TeV.

de 10 TeV. Pour cette raison, on limitera ci-dessous notre ajustement du spectre de la nébuleuse du Crabe à une énergie maximale de ∼10 TeV, et ce bien que le signal au-delà de cette énergie soit très significatif ($\sim 10\,\sigma$). Le spectre est très bien reproduit avec une hypothèse en loi de puissance avec une coupure exponentielle, hypothèse qui est largement préférée à une simple loi de puissance ($\lambda \approx 62$, voir section 5.4.2). À cause de la contrainte sur l'énergie maximale de l'ajustement, l'énergie caractéristique de la coupure exponentielle, qui a été mesurée à $(14{,}3 \pm 2{,}1)$ TeV par [Aharonian et al., 2006], est tirée vers le bas : $E = (4{,}4 \pm 0{,}5)$ TeV (Figure 7.8a). La normalisation ϕ_0 est plus importante de 30 % et l'indice est inférieur de 0,5 (20 %). Néanmoins, en fixant l'énergie de coupure à 14,3 TeV (Figure 7.8b) on obtient une normalisation de $\phi_0 = (4{,}02 \pm 0{,}07) \times 10^{-11}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{TeV}^{-1}$ qui s'écarte de la valeur de [Aharonian et al., 2006] de seulement 7 % et un indice spectral légèrement inférieur de 0,15 (6 %) à celui obtenu par *H.E.S.S.* en 2006. Ce résultat est tout à fait remarquable compte tenu de la compréhension encore incomplète de l'instrument ainsi, que des efficacités optiques préliminaires utilisées pour la production des IRF qui s'écartent des valeurs obtenues par calibration à base de muons.

7.7.2 PKS 2155-304

Le spectre de PKS 2155-304 ajusté par une loi de puissance est montré sur les Figures 7.9a et 7.9b. Sur la première, pour laquelle l'énergie seuil a été déterminée avec une surface minimale de 4 ha, les résidus montrent clairement une courbure dans le spectre, avec en particulier des valeurs négatives pour les deux premiers points à basse énergie. En augmentant le seuil en énergie à 300 GeV l'ajustement par une loi de puissance paraît beaucoup plus satisfaisant, comme montré sur la Figure 7.9b. La courbure en énergie dans le premier cas s'explique probablement par l'utilisation de surfaces efficaces de collection que l'on sait susceptibles d'avoir des biais près du seuil (voir discussion en section 7.2). Pour l'ajustement utilisant une énergie seuil de 300 GeV, l'indice spectral obtenu de $\Gamma = 3{,}62 \pm 0{,}13$ est très proche de celui déterminé lors de la caractérisation de l'état bas de la source entre 2005 et 2007 [H.E.S.S. Collaboration, 2010], qui indiquait $\Gamma = 3{,}5 \pm 0{,}1$. Le flux



(a) (b)

FIGURE 7.9: Spectres différentiels de PKS 2155-304 ajustés par une loi de puissance (a) pour un énergie seuil de 165 GeV et (b) pour une énergie seuil de 300 GeV ($E_{\mathrm{Ref}} = 1$ TeV).

estimé est plus bas de 40 % par rapport à celui mesuré pour l'état bas entre 2005 et 2007, ce qui confirme une tendance observée par *H.E.S.S.* ces dernières années (publication de la collaboration *H.E.S.S.*, en préparation).

7.7.3 PG 1553+113

Le spectre de PG 1553+113 a été obtenu avec la configuration dédiée aux sources faibles. La Figure 7.10 montre l'ajustement pour une hypothèse en loi de puissance. Les paramètres ajustés sont $\phi_0(300\,\mathrm{GeV}) = (3{,}8 \pm 0{,}5) \times 10^{-11}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{TeV}^{-1}$ et $\Gamma = 4{,}4 \pm 0{,}4$. Ces deux valeurs sont tout à fait compatibles avec les flux mesurés à diverses époques par l'expérience *H.E.S.S.* (selon les périodes entre 2,6 et 5,9 $\times 10^{-11}$ cm^{-2} s^{-1} TeV^{-1}, avec des indices entre 4 et 4,5, voir [Aharonian et al., 2008b]). Contrairement à ce qui a été observé dans le cas de la nébuleuse du Crabe et de PKS 2155-304, l'ajustement est stable en imposant une surface minimale de 2, 3 ou 4 ha. Le spectre représenté a été déterminé avec une surface minimale de 2 ha.

7.7.4 Mkn 501

Reconstruire le spectre de Mkn 501 avec les IRF générées à partir des simulations CORSIKA/sim_hessarray est délicat. En effet, l'angle zénithal maximal auquel ces simulations ont été générées est de 55°, alors que les données de Mkn 501 ont été obtenues avec un angle zénithal moyen d'observation de 63°. Nous nous sommes donc repliés sur les simulations KASKADE/SMASH, tout en sachant que leur compatibilité avec des évènements entraînés sur des γ simulés avec CORSIKA/sim_hessarray n'a pas été démontrée. Les IRF associées aux simulations KASKADE/SMASH ont été générées jusqu'à un angle zénithal de 70°. Une comparaison faite sur le spectre de la nébuleuse du Crabe montre qu'avec ces IRF on obtient une normalisation majorée de 30 % pour un indice spectral similaire et une énergie de coupure de ∼10 TeV. La différence de normalisation peut être expliquée en grande partie par les différences entre efficacités optiques utilisées dans les simulations sim_hessarray et SMASH pour les télescopes de taille moyenne.

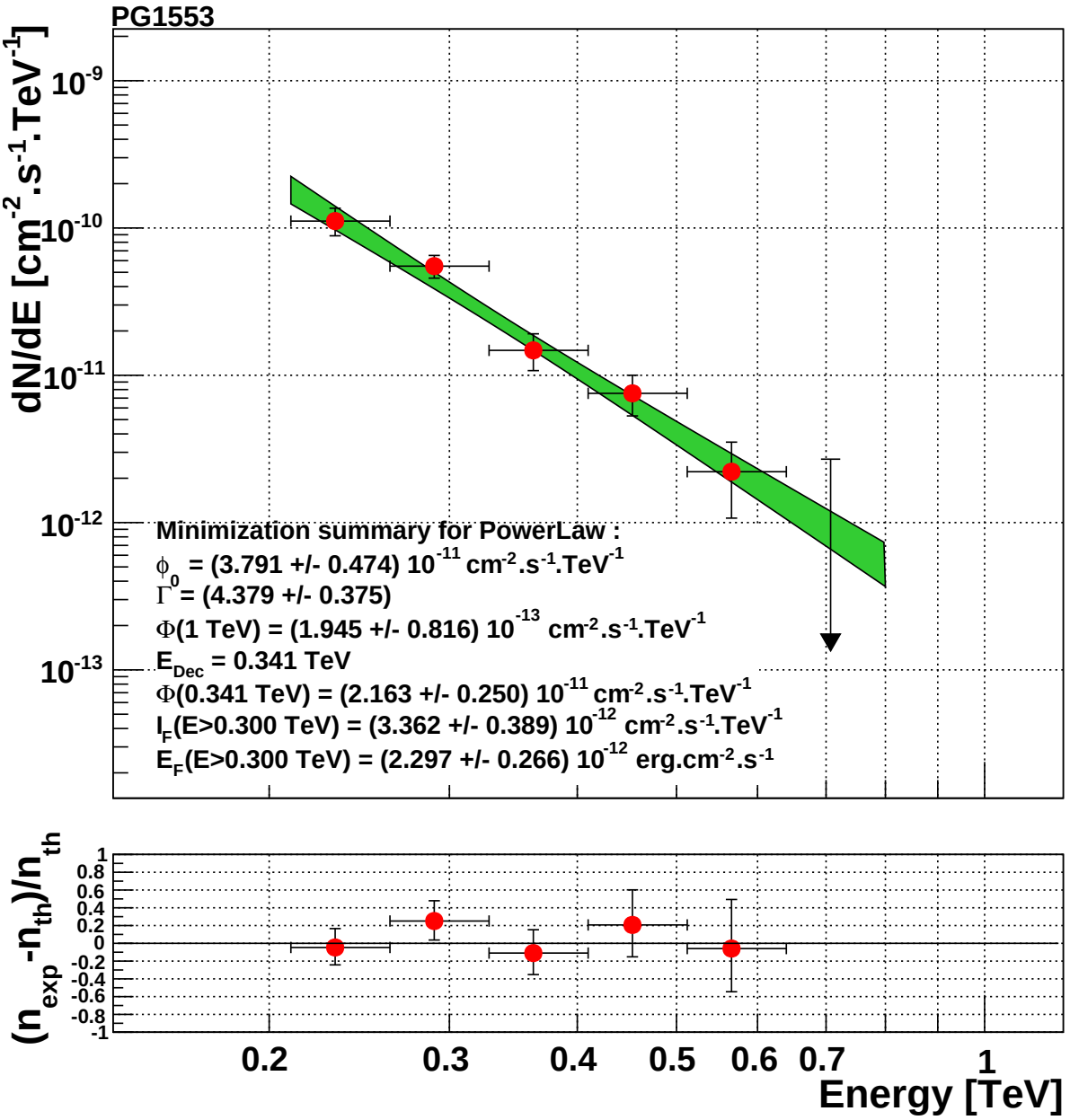


FIGURE 7.10: Spectre différentiel de PG 1553+113 obtenu avec une surface minimale de 2 ha et pour une hypothèse de loi de puissance ($E_{\mathrm{Ref}} = 300\,\mathrm{GeV}$).

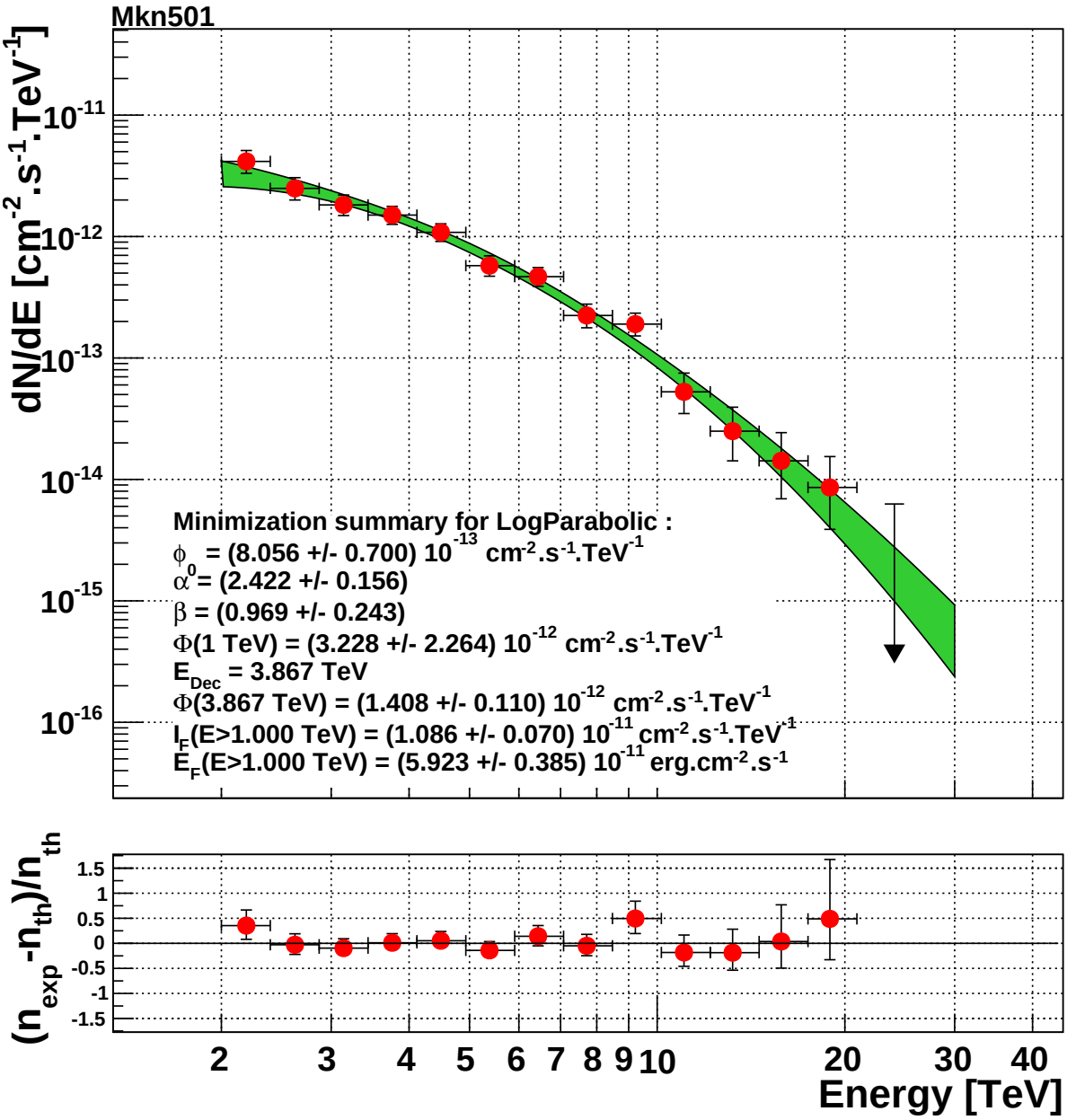


FIGURE 7.11: Spectre différentiel de Mkn 501 obtenu avec une surface minimale de 4 ha et pour une hypothèse de loi de puissance ($E_{\mathrm{Ref}} = 5\,\mathrm{TeV}$).

La Figure 7.11 représente le spectre de Mkn 501 ajusté selon une loi de puissance parabolique ($E_{\mathrm{Ref}} = 5\,\mathrm{TeV}$), qui semble être privilégiée par rapport à l'hypothèse de loi de puissance ($\lambda \approx 25$). L'énergie correspondant au maximum du pic d'émission en νF_ν ($E_{\mathrm{max}} = E_{\mathrm{Ref}} 10^{(2-\alpha)/2\beta}$) est déterminée à $E_{\mathrm{max}} = (3{,}03 \pm 0{,}48)\,\mathrm{TeV}$. Cette estimation est très fragile par rapport aux systématiques, car notre spectre démarre en haut du pic, et que tout ceci se passe près du seuil, là où les systématiques sont plus difficiles à maîtriser. Pour rappel, les expériences CAT et MAGIC ont également observé des spectres en loi de puissance parabolique lors de différents sursauts, avec des pics culminant (en représentation νF_ν) entre 250 GeV et 1 TeV ([Djannati-Atai et al., 1999] et [Albert et al., 2007]). Le niveau de flux au maximum du pic d'émission pour les données de juin 2014 ($\nu F_\nu \approx 5{,}4 \times 10^{-11}\,\mathrm{erg\,cm^{-2}\,s^{-1}}$) correspond au niveau de flux le plus bas des sursauts observés en 1997 par l'expérience CAT ($\nu F_\nu \approx 4{,}9 \times 10^{-11}\,\mathrm{erg\,cm^{-2}\,s^{-1}}$).

Bibliographie

A. A. Abdo, et al. Gamma-Ray Flares from the Crab Nebula. *Science*, 331:739–, February 2011. doi: 10.1126/science.1199705.

F. Aharonian, et al. Measurement of the flux, spectrum, and variability of TeV γ-rays from MKN 501 during a state of high activity. A&A, 327:L5–L8, November 1997.

F. Aharonian, et al. Observations of the Crab nebula with HESS. A&A, 457:899–915, October 2006. doi: 10.1051/0004-6361:20065351.

F. Aharonian, et al. An Exceptional Very High Energy Gamma-Ray Flare of PKS 2155-304. ApJ, 664:L71–L74, August 2007. doi: 10.1086/520635.

F. Aharonian, et al. Limits on an Energy Dependence of the Speed of Light from a Flare of the Active Galaxy PKS 2155-304. *Physical Review Letters*, 101(17):170402, October 2008a. doi: 10.1103/PhysRevLett.101.170402.

F. Aharonian, et al. HESS observations and VLT spectroscopy of PG 1553+113. A&A, 477:481–489, January 2008b. doi: 10.1051/0004-6361:20078603.

F. Aharonian, et al. Simultaneous multiwavelength observations of the second exceptional γ-ray flare of PKS 2155-304 in July 2006. *Astron. Astrophys.*, 502:749–770, August 2009a. doi: 10.1051/0004-6361/200912128.

F. Aharonian, et al. Simultaneous Observations of PKS 2155-304 with HESS, Fermi, RXTE, and Atom: Spectral Energy Distributions and Variability in a Low State. ApJ, 696:L150–L155, May 2009b. doi: 10.1088/0004-637X/696/2/L150.

J. Albert, et al. Variable Very High Energy γ-Ray Emission from Markarian 501. ApJ, 669:862–883, November 2007. doi: 10.1086/521382.

J. Aleksić, et al. Probing the very-high-energy gamma-ray spectral curvature in the blazar PG 1553+113 with the MAGIC telescopes. *ArXiv e-prints*, August 2014.

E. Aliu, et al. Observation of Pulsed γ-Rays Above 25 GeV from the Crab Pulsar with MAGIC. *Science*, 322:1221–, November 2008. doi: 10.1126/science.1164718.

E. Aliu, et al. A Search for Enhanced Very High Energy Gamma-Ray Emission from the 2013 March Crab Nebula Flare. ApJ, 781:L11, January 2014. doi: 10.1088/2041-8205/781/1/L11.

P. Brun, D. Wouters, et for the H. E. S. S. collaboration. Constraints on axion-like particles with H.E.S.S. from observations of PKS 2155-304. *ArXiv e-prints*, July 2013.

C. W. Danforth, et al. HST/COS Observations of the Ly alpha Forest toward the BL Lac Object 1ES1553+113. *ArXiv e-prints*, May 2010.

A. Djannati-Atai, et al. Very High Energy Gamma-ray spectral properties of MKN 501 from CAT Čerenkov telescope observations in 1997. A&A, 350:17–24, October 1999.

J. Guy. *Premiers résultats de l'expérience HESS et étude du potentiel de détection de matière noire supersymétrique*. PhD thesis, UNIVERSITÉ PARIS 6, 2003.

H.E.S.S. Collaboration. VHE γ-ray emission of PKS 2155-304: spectral and temporal variability. A&A, 520:A83, September 2010. doi: 10.1051/0004-6361/201014484.

H.E.S.S. Collaboration. Search for Lorentz Invariance breaking with a likelihood fit of the PKS 2155-304 flare data taken on MJD 53944. *Astroparticle Physics*, 34:738–747, April 2011. doi: 10.1016/j.astropartphys.2011.01.007.

H.E.S.S. Collaboration, et al. A multiwavelength view of the flaring state of PKS 2155-304 in 2006. A&A, 539:A149, March 2012. doi: 10.1051/0004-6361/201117509.

H.E.S.S. Collaboration, et al. Measurement of the extragalactic background light imprint on the spectra of the brightest blazars observed with H.E.S.S. A&A, 550:A4, February 2013. doi: 10.1051/0004-6361/201220355.

K. Katarzyński, et al. Modelling rapid TeV variability of PKS2155-304. MNRAS, 390:371–376, October 2008. doi: 10.1111/j.1365-2966.2008.13753.x.

MAGIC Collaboration, et al. Probing quantum gravity using photons from a flare of the active galactic nucleus Markarian 501 observed by the MAGIC telescope. *Physics Letters B*, 668:253–257, October 2008. doi: 10.1016/j.physletb.2008.08.053.

V. Marandon. *Nébuleuses de pulsars : sondage profond de la Galaxie au TeV et Nébuleuses de pulsars : sondage profond de la Galaxie au TeV et études multi-longueurs d'onde.* PhD thesis, 2010.

E. A. Nadaraya. On estimating regression. *Theory Probab. Appl.*, 9(1):141–142, 1964.

E. Prandini, et al. A new upper limit on the redshift of PG 1553+113 from observations with the MAGIC Telescope. *ArXiv e-prints*, July 2009.

R. J. Protheroe, et al. Very high energy gamma rays from Markarian 501. *ArXiv Astrophysics e-prints*, October 1997.

D. A. Sanchez, S. Fegan, et B. Giebels. Evidence for a cosmological effect in γ-ray spectra of BL Lacertae. A&A, 554:A75, June 2013. doi: 10.1051/0004-6361/201220631.

M. Tavani, et al. Discovery of Powerful Gamma-Ray Flares from the Crab Nebula. *Science*, 331: 736–, February 2011. doi: 10.1126/science.1200083.

VERITAS Collaboration, et al. Detection of Pulsed Gamma Rays Above 100 GeV from the Crab Pulsar. *Science*, 334:69–, October 2011. doi: 10.1126/science.1208192.

Chapitre 8

Résumé

Nous avons présenté dans cette partie du manuscrit le fonctionnement de l'expérience *H.E.S.S.*, les méthodes de calibration et les méthodes d'analyse utilisées pour l'expérience HESS-I par différentes équipes de la collaboration. Nous avons mis l'accent sur la méthode d'analyse à base d'arbres de décision « boostés » qui a servi de point de départ aux développements présentés dans les chapitres 6 et 7. Par la même occasion, nous avons aussi présenté START, logiciel de haut niveau dédié à l'étude spectrale des sources par « *forward-folding* » et entièrement développé au cours de cette thèse. START se caractérise par sa modularité (développé en C++), sa portabilité (facile à adapter pour les différentes chaînes d'analyse de *H.E.S.S.*), sa facilité d'utilisation (nombreuses options et documentation disponible), et sa rapidité d'exécution.

Une partie importante du travail présenté ici concerne le développement d'une analyse multivariée à base d'arbres de décision « boostés » pour la nouvelle phase du réseau de télescopes *H.E.S.S.* en mode hybride. Prenant pour point de départ l'analyse Paris-MVA initialement développée pour *H.E.S.S.*, ce travail a été d'abord réalisé de manière exploratoire à base de simulations MC, ce qui a permis de sonder le pouvoir discriminant des différentes variables, d'examiner leur comportement en présence du nouveau télescope CT_5, et de donner une estimation de la sensibilité préliminaire de l'instrument en mode hybride.

Dans un deuxième temps, après accumulation de plusieurs centaines d'heures d'observation de bonne qualité, il a été possible de refaire ce travail à partir des données de fond réel. Nous avons alors optimisé la détermination des points de fonctionnement de la méthode BDT (pour les 61 intervalles de l'espace des phases pour lesquels nous avons réalisé des entraînements indépendants), et ce pour trois configurations représentant des types de sources de niveau de flux et d'indice spectral différents. Les résultats obtenus pour l'extraction du signal des premières sources observées (la nébuleuse du Crabe, PKS 2155-304, PG 1553+113 et Mkn 501) sont tout à fait remarquables, avec par exemple $35{,}1\,\sigma\sqrt{\mathrm{h}^{-1}}$ pour la nébuleuse du Crabe ou $3{,}7\,\sigma\sqrt{\mathrm{h}^{-1}}$ pour une source très faible comme PG 1553+113. Les spectres montrés à titre préliminaire sont conformes à ce qui a déjà été mesuré pour ces différentes sources. Les erreurs systématiques sont cependant importantes du fait de la nature préliminaire des simulations MC et de l'écart du modèle numérique de l'instrument avec la réalité (notamment en ce qui concerne les efficacités optiques).

Troisième partie

Recherche de candidats blazars parmi les sources non-identifiées du *Fermi-LAT* du catalogue 2FGL

Chapitre 9

Introduction

Sommaire

9.1 Contexte et objectifs **157**
9.2 L'expérience *Fermi-LAT* **159**
9.3 L'expérience *WISE* **165**
Bibliographie **169**

9.1 Contexte et objectifs

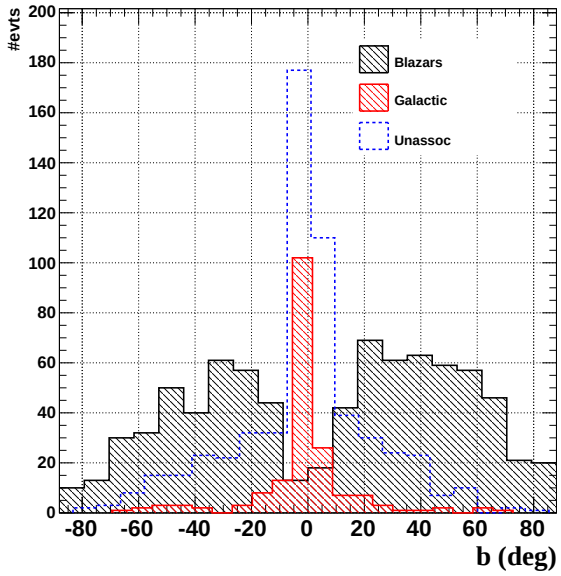
Le télescope *Fermi-LAT* est opérationnel depuis juin 2008 et observe l'intégralité du ciel entre 30 MeV et quelques centaines de GeV. Un catalogue de 1873 sources détectées lors des deux premières années avec un niveau de significativité supérieur à 4σ a été publié par la collaboration *Fermi* : le 2FGL. Parmi ces sources, on compte 1092 AGN et 196 sources galactiques, qui sont essentiellement des pulsars et des plérions. Le reste des 575 sources sont de nature non-identifiée.

Plusieurs auteurs suggèrent qu'un grand nombre des sources non-identifiées sont des AGN. En observant la distribution de la latitude des sources galactiques, des blazars, et des sources non-identifiées, on constate une diminution significative du nombre de blazars aux faibles latitudes (Figure 9.1) qui pourrait être compensée par un certain nombre de sources non-identifiées. À plus haute latitude les deux classes de sources majoritaires sont les blazars et les pulsars, ce qui suggère que les sources non-identifiées se répartissent entre ces deux classes.

La détermination des candidats blazars parmi les sources non-identifiées du catalogue 2FGL permettra d'augmenter de manière significative l'échantillon de blazars détectés à haute énergie. L'utilité est multiple. Cela servira d'une part à augmenter le nombre de FSRQ et de BL Lac qui pourront servir à affiner les études qui sont aujourd'hui conduites sur le schéma d'unification des AGN et en particulier sur la séquence des blazars (voir section 2.6.1), et d'autre part à augmenter le nombre de blazars qui pourront servir de cibles futurs aux détecteurs au sol d'imagerie atmosphérique Tcherenkov. Ces potentiels candidats blazars émetteurs à très haute énergie pourront alors être utilisés afin de mieux contraindre les modèles d'émission et de variabilité des AGN ainsi que pour chercher des signes de physique « au-delà du modèle standard » (voir section 2.6.2).

9.1.1 Identification des blazars parmi les sources non-identifiées du 2FGL

Le catalogue 2FGL contient un nombre important de paramètres permettant de mettre en évidence des différences, d'un point de vue statistique, entre la forme spectrale ou la variabilité temporelle des blazars et celle des sources galactiques. L'utilisation de techniques d'MVA est un moyen d'obtenir un certain nombre de candidats blazars en utilisant un choix approprié de variables discriminantes (chapitre 3). Certains auteurs ([Mirabal et al., 2012], [Ferrara et al., 2012] et [Doert et Errando, 2013]) se sont attelés à la tâche d'obtenir les meilleurs candidats blazars parmi les sources non-identifiées, chacun ayant son propre choix de paramètres discriminants. Néanmoins, toutes ces approches ont en commun l'utilisation des paramètres dits de « hardness ratio » ou indices de dureté, construits à partir du 2FGL en considérant des différences de flux entre bandes en énergie connexes. Dans le cas où une source n'a pas le niveau de détection requis pour estimer un flux dans



(a)

FIGURE 9.1: Latitude galactique des blazars (noir) et des sources galactiques identifiées (rouge), et des sources non-identifiées (bleu) *Fermi*.

une des cinq bandes en énergie fournies par le 2FGL, c'est une limite supérieure à 2σ sur le flux qui est donnée dans le catalogue et utilisée par ces auteurs en lieu et place du flux. L'impact de l'utilisation de manière indifférenciée des flux ou des limites supérieures dans le calcul des indices de dureté n'est pas discuté dans leurs publications respectives. Ainsi, si le pouvoir discriminant de ces paramètres semble en premier lieu prometteur, nous montrerons qu'il est difficile à contrôler, car il est fonction de la brillance de la source et donc de la présence éventuelle de limites supérieures dans les différentes bandes en énergie (chapitre 10, section 10.3.2). Le calcul des performances final d'une analyse MVA qui utilise des paramètres dont le pouvoir discriminant dépend ainsi du niveau de flux des sources est alors biaisé.

Une première partie de ce travail (chapitre 10) sera consacrée au développement d'une analyse multivariée basée sur une sélection soignée de variables discriminantes pour rechercher les candidats blazars parmi les sources non-identifiées du catalogue 2FGL. Le traitement des sources non-identifiées de haute et de basse latitude galactique sera différencié afin de prendre en compte, lors de l'entraînement de la méthode, des différences constatées dans les populations qui constituent le fond.

9.1.2 Détermination des coordonnées des candidats blazars du 2FGL à l'aide de l'expérience *WISE*

La PSF du *Fermi-LAT* rend la recherche de contreparties à d'autres longueurs d'onde ardue, l'erreur de localisation d'une source pouvant varier de quelques centièmes de degrés pour les sources les plus brillantes à $\sim$0,5° pour les plus faibles. À titre d'exemple, la Figure 9.2 représente en bleu les sources détectées par le satellite *WISE* en IR (en bleu) dans la région de la source 2FGL J1801.7+4405 (position radio et γ respectivement en rouge et vert). En 2011, [Massaro et al., 2011] propose de rechercher les contreparties IR des sources *Fermi* non-identifiées et se trouvant à haute latitude galactique à l'aide de l'expérience *WISE* (section 9.3). Cette expérience a observé l'intégralité du ciel dans quatre filtres micro-onde pendant plus d'une année et a permis la détection de $\sim$750 millions d'objets, avec une précision sur la détermination des coordonnées bien inférieure

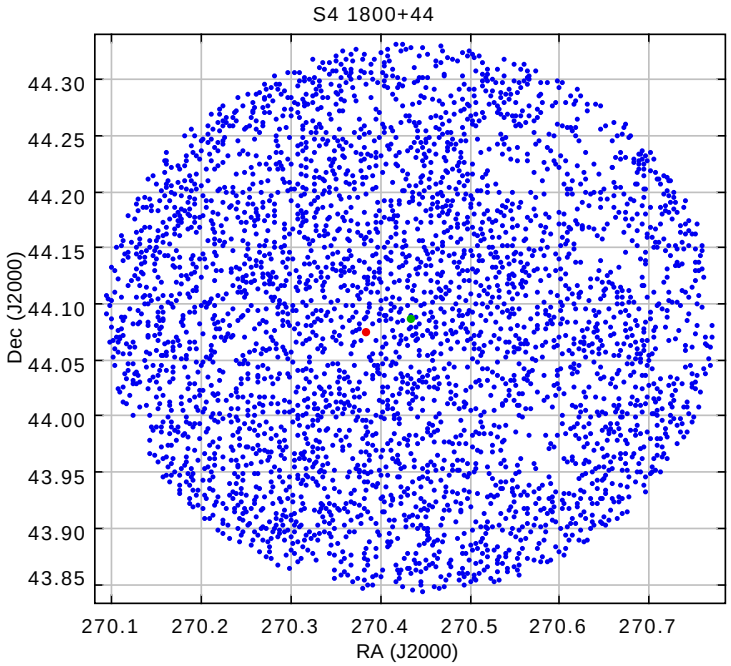


FIGURE 9.2: Exemple de contreparties infrarouges dans la région de la position d'une source *Fermi* à un niveau de confiance de 99,9 %. Position radio de la source S4 1800+44 (rouge), position *Fermi* en (vert) et positions des contreparties *WISE* (bleu).

à l'arcseconde. La méthode d'association proposée par [Massaro et al., 2011] part du principe que l'émission des blazars dans l'IR est dominée par des processus non-thermiques, ce qui n'est pas le cas des étoiles et des galaxies non actives. [Massaro et al., 2011] observe alors que les blazars occupent un emplacement privilégié dans un diagramme couleur-couleur (la couleur est une différence de magnitude entre deux filtres) ce qui facilite leur identification. À l'aide d'une approche géométrique d'analyse en composantes principales utilisant un lot de blazars identifié dont les coordonnées sont connues avec précision, [Massaro et al., 2013] montre la faisabilité de la méthode en l'appliquant à l'échantillon des sources non-identifiées. Cela permet alors d'obtenir pour chacune des sources non-identifiées du 2FGL un certain nombre de candidats blazars « IR » et de faciliter le suivi multi-longueur d'onde pour confirmer ou réfuter la nature de l'objet.

Nous compléterons le travail présenté au chapitre 10 par le développement d'une méthode alternative à celle de utilisée par [Massaro et al., 2013], basée sur une approche multivariée, consacrée à l'identification des contreparties IR des candidats blazars. Cette méthode sera construite à partir des magnitudes mesurées par les différents filtres de l'expérience *WISE* et optimisée de manière différenciée selon la position des sources étudiées par rapport au plan galactique, prenant ainsi en compte les différences de structure liées aux populations constituant le fond.

9.2 L'expérience *Fermi-LAT*

9.2.1 Le satellite *Fermi*

Mis sur orbite le 11 juin 2008, le satellite *Fermi* – anciennement dénommé GLAST pour *Gamma-ray Large Area Telescope* – transporte les télescope *LAT* et *GBM*. Le premier est un télescope à création de paires e^+e^-, doté d'un champ de vue de $\sim$2 sr, chargé d'observer l'ensemble du ciel entre une vingtaine de MeV et une centaine de GeV [Atwood et al., 2009]. Le *GBM* a lui pour objectif de déterminer l'origine de l'émission d'objets transitoires (des sursauts γ) entre $\sim$10 keV et $\sim$30 MeV et de transmettre l'information au *LAT* pour qu'il puisse se repositionner afin d'observer les sursauts

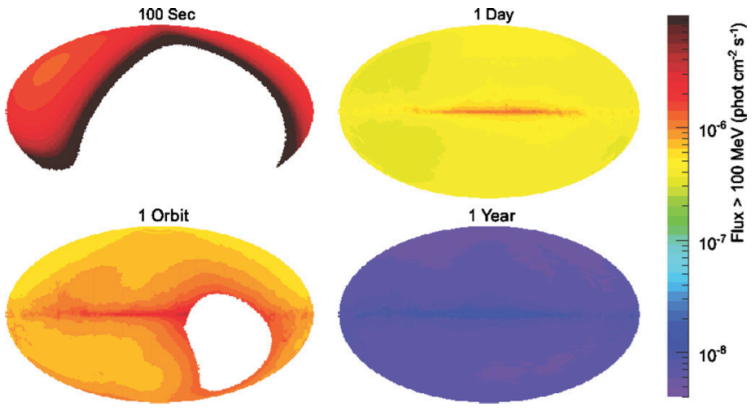
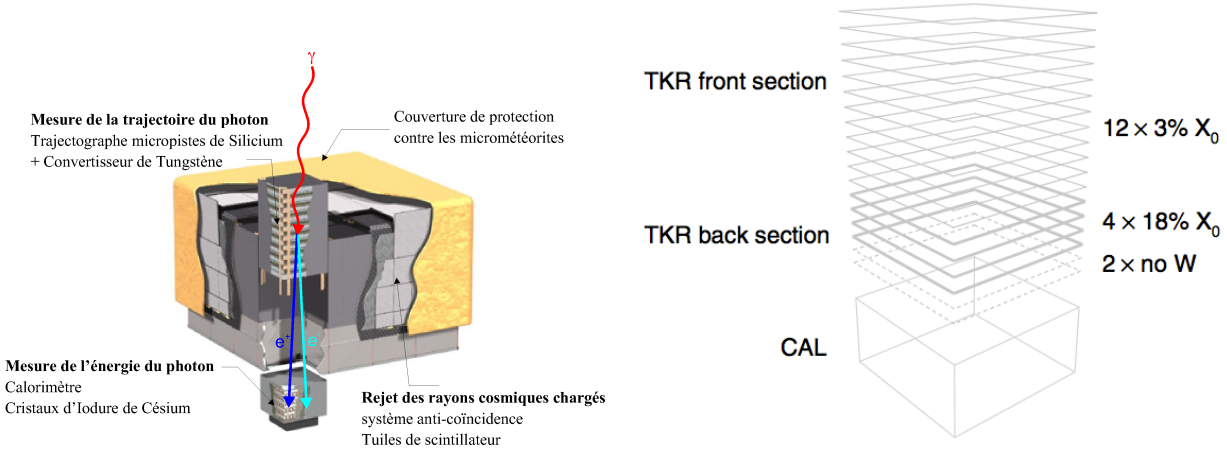


FIGURE 9.3: Sensibilité et couverture du ciel du *LAT* pour différentes échelles de temps (Crédits : [Atwood et al., 2009]).



(a) (b)

FIGURE 9.4: (a) Vues schématique du *LAT* (Crédits : `http://www.cenbg.in2p3.fr/`), et (b) une des 16 tours constituant le détecteur (Crédits : [Ackermann et al., 2012]).

brillants [Meegan et al., 2009]. Prévu pour être opérationnel pendant au moins 5 ans, le temps de vie optimal de la mission est estimé à 10 ans. Le satellite *Fermi* suit une orbite basse quasi-circulaire à une altitude de ∼565 km avec une inclinaison de 25,6° par rapport au plan équatorial et effectue une orbite en 96,5 min. Son axe de rotation suit un mouvement de précession avec une période 53,4 jours et les panneaux solaires font toujours face au Soleil. La couverture du ciel et la sensibilité intégrée pour différentes échelles de temps sont représentées sur la Figure 9.3. Dans cette partie sont brièvement présentés le principe de fonctionnement du *LAT*, les méthodes de reconstruction et de discrimination associées ainsi que ses performances.

9.2.2 Le *LAT*

Le *LAT* est constitué de trois parties, chacune dédiée à une tâche spécifique (Figure 9.4a) : le suivi et la reconstruction de l'énergie des rayons γ sont effectués respectivement avec un trajectographe et un calorimètre, et le rejet des rayons cosmiques chargés est assurée par un bouclier anti-coïncidence. Les deux premiers modules sont divisés en 16 tours disposées selon une grille 4×4. En plus de ces trois sous-parties un système d'acquisition permet de sélectionner les candidats γ dont les données sont ensuite transmises au sol.

Chacune des 16 tours du trajectographe dispose de 18 paires de détecteurs (Figure 9.4b) qui sont constituées d'une alternance d'un matériau semi-conducteur à base de silicium et de convertisseurs de tungstène. Un espace de 2 mm sépare les couches de silicium. Chaque couche sensible permet d'obtenir les coordonnées des paires e^+e^- qui résultent de l'interaction du γ incident avec une des plaques de tungstène. Les quatre dernières couches de tungstène sont quant à elles plus épaisses, et les deux dernières paires de détecteurs ne sont pas séparées par du tungstène. Aux basses énergies, la résolution angulaire est limitée par la diffusion multiple des paires e^+e^-, ainsi que par l'émission Bremsstrahlung de ces dernières. Un évènement déclenche l'enregistrement du signal s'il active au moins trois paires de couches de silicium. Des IRF spécifiques sont dédiées aux évènements déclenchant les plaques fines (*front*) et épaisses (*back*) du détecteur. L'utilisation de cette technologie dans le cas d'une expérience spatiale a considérablement amélioré la résolution angulaire par rapport au détecteur de la génération précédente (*EGRET*).

Dans chacun des 16 modules, une matrice en trois dimensions de cristaux de iodure de césium compose le calorimètre. Celle-ci est arrangée selon 8 couches comportant chacune 12 cristaux. Au passage d'une particule chargée, cette dernière cède son énergie au calorimètre sous forme de photons (réaction de scintillation). Deux photodiodes sont attachées à chaque cristal. La lumière est alors collectée par des photodiodes de deux types, chacune ayant une réponse énergétique différente (2 MeV-1,6 GeV et 100 MeV-70 GeV). Le profil de la gerbe est alors estimé en mesurant l'énergie dans chacun de cristaux. L'énergie de la particule initiale est reconstruite après la calibration des données selon un algorithme exposé dans [Ackermann et al., 2012]. En plus de la mesure de l'énergie des particules incidentes, le développement du signal tout au long du calorimètre apporte des informations sur la direction initiale du γ incident et augmente la discrimination γ/rayons cosmiques.

Le bouclier anti-coïncidence est constitué de 89 tuiles d'un scintillateur plastique recouvrant entièrement le champ de vue du télescope. Un rayon cosmique incident va céder dans le matériau une partie de son énergie qui va être récoltée par des photomultiplicateurs puis enregistrée. Si la particule parvient jusqu'au détecteur, la coïncidence des temps permet de ne pas enregistrer l'évènement. Élément indispensable du télescope, ce dispositif permet de rejeter plus de 99,97 % du fond composé par les particules chargées.

9.2.3 Performances du *LAT*

Les techniques d'analyse utilisées par la collaboration *Fermi-LAT* lors de la publication du catalogue 2FGL sont exposées dans le papier [Ackermann et al., 2012]. Ce sont les performances produites avec la version *Pass 7_ V6* des méthodes d'analyse qui sont brièvement présentées [1] ici (Figures 9.5a-9.5f).

La surface efficace en fonction de l'énergie pour des photons ayant une incidence normale ($\theta \sim 0$) est représentée sur la Figure 9.5a, pour les évènements de type *front* (rouge) ou *back* (bleu). La somme des deux contributions est donnée par la courbe noire. La dégradation de la surface en fonction de l'angle d'incidence θ des photons est illustrée sur la Figure 9.5b pour des γ de 10 GeV.

Sur la Figure 9.5c, on peut voir la résolution en énergie. Dans le domaine du GeV, la résolution est inférieure à ∼10 %. Près du seuil de l'instrument, à une centaine de MeV, elle est de 15 % à 20 %. Sa dépendance en fonction de l'angle d'incidence est illustrée sur la figure Figure 9.5d pour des photons de 10 GeV.

La résolution angulaire calculée en ne gardant que 68 % (95 %) des meilleurs évènements reconstruits est illustrée sur la Figure 9.5e en traits pleins (pointillés) pour des photons en incidence normale. Le *LAT* permet d'obtenir une résolution angulaire (68 %) de l'ordre de 0,2° au-delà de la dizaine de GeV qui se dégrade rapidement aux basses énergies ($\propto E^{-1}$) pour atteindre ∼5° à 100 MeV. L'effet majeur responsable de cette dégradation déjà exposée plus haut est la diffusion multiple des paires e^+e^- au sein du trajectographe.

1. http://www.slac.stanford.edu/exp/glast/groups/canda/archive/pass7v6/lat_Performance.htm

Finalement, la sensibilité différentielle estimée pour une source ponctuelle ayant un indice différentiel de -2 et observée pendant 3 ans est représentée sur la Figure 9.5f pour différentes latitudes galactiques ($b = 0°$ en noir, 30° en rouge et 90° en bleu). En raison de l'émission diffuse dans le plan galactique, détecter une source ayant un niveau de flux donné demandera un temps d'observation plus court aux hautes latitudes qu'aux basses latitudes. Pour les hautes latitudes, le point de fonctionnement optimal se situe autour du GeV pour un flux minimal détectable de $\sim 10^{-12}$erg cm^{-2} s^{-1} en trois ans. Dans la région du centre galactique, à 10 GeV, le seuil de détection est quant à lui de $\sim 10^{-11}$erg cm^{-2} s^{-1}.

9.2.4 Le catalogue 2FGL

Le catalogue 2FGL [Nolan et al., 2012], successeur des *First Fermi-LAT catalog* (1FGL) [Abdo et al., 2010] et du *Fermi-LAT Bright Sources List* (0FGL) [Abdo et al., 2009], regroupe l'ensemble des sources détectées avec une significativité supérieure à $\sim 4\sigma$ (TS $\geq$ 25), au bout de deux ans d'observation du ciel. Un total de 1873 sources est contenu dans le catalogue. L'ensemble des sources du catalogue 2FGL est récapitulé, classées par type, dans la Table 9.1. On distingue trois classes de sources : les sources identifiées, associées et non-associées (ou non-identifiées).

9.2.4.1 Les sources identifiées

Une source est dite identifiée si elle valide au moins un des trois critères suivants. Le premier est un critère d'émission périodique dans le temps. Il concerne exclusivement les pulsars et les systèmes binaire X, pour lesquels on peut détecter une émission γ pulsée avec une probabilité qu'elle soit fortuite de 10^{-6}. Une morphologie de source spatialement étendue visible à d'autres longueurs d'onde constitue un second critère. Cela concerne quelques restes de supernova (SNR), plérions (PWN) ainsi que les lobes de la radiogalaxie Centaurus A. Le dernier critère concerne la corrélation entre la variabilité du flux γ d'une source dans au moins deux bandes en énergie et la variabilité de la source dans au moins une autre gamme de longueur d'onde. Les sources qui varient de manière corrélée à différentes longueurs d'onde sont principalement des AGN. Au total, on dénombre 127 sources identifiées. Notons le cas de la nébuleuse du Crabe pour lequel aucun des trois critères précédents n'est validé mais qui par sa forme spectrale reconnaissable et compatible avec ce qui est mesuré à d'autres longueurs d'onde est tout de même considéré comme une source identifiée.

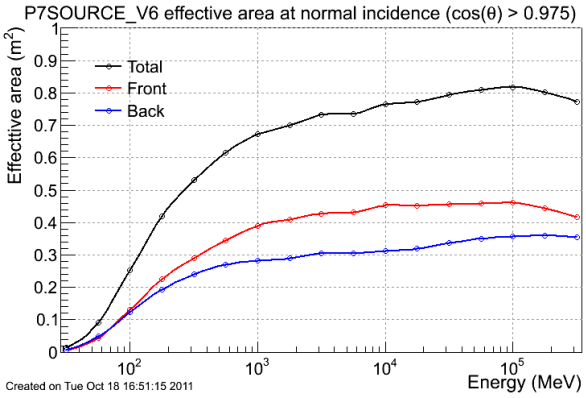
9.2.4.2 Les sources associées

Lorsque la méthode d'identification échoue, une procédure d'association déjà utilisée pour le 1FGL qui consiste à chercher les contreparties spatiales à diverses longueurs d'onde dans différents catalogues est appliquée. À chaque contrepartie potentielle d'une des sources du catalogue est associée une probabilité rendant compte de la solidité de l'association. Une probabilité d'au moins 80 % est requise pour considérer une association. Le taux de fausses associations donné par la méthode (puis validé par Monte-Carlo) est de 5 %. Au total, 1141 sources possèdent au moins une contrepartie autre que le GeV (*EGRET*).

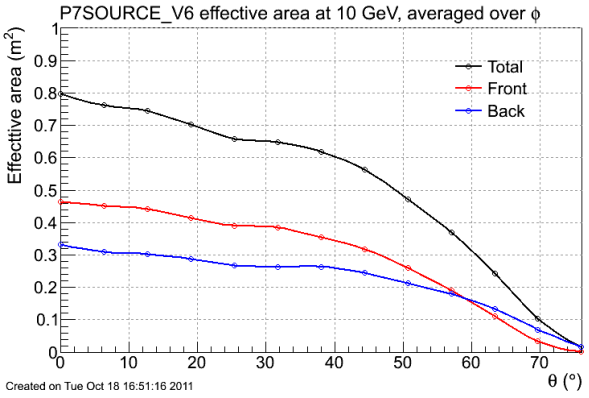
On notera qu'en plus de la procédure basée sur la coïncidence spatiale décrite ci-dessus, les AGN ont bénéficié de l'utilisation de deux méthodes d'association supplémentaires et ce, afin de s'affranchir des biais de sélection dus à la limite en flux des catalogues radios. Ces deux méthodes additionnelles utilisent les propriétés physiques des contreparties telles que leur densité et leur forme spectrale en radio. Elles sont toutes les deux décrites en détails dans [Ackermann et al., 2011] et permettent d'atteindre un taux de fausses associations pour les AGN de moins de 2 %.

9.2.4.3 Les sources non-identifiées

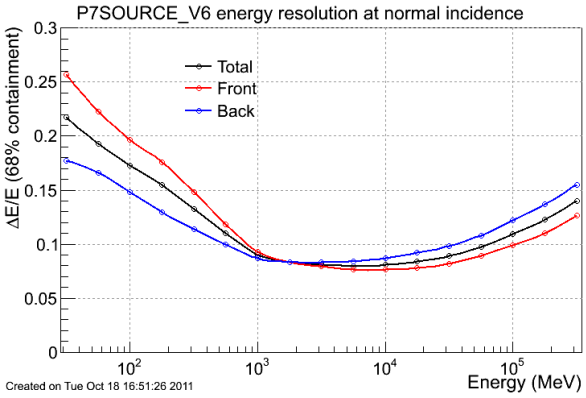
La dernière classe de sources est constituée de sources n'ayant pu être ni identifiées, ni associées. Elle compte 575 membres dont la nature reste à déterminer. Environ 50 % de ces sources présentent



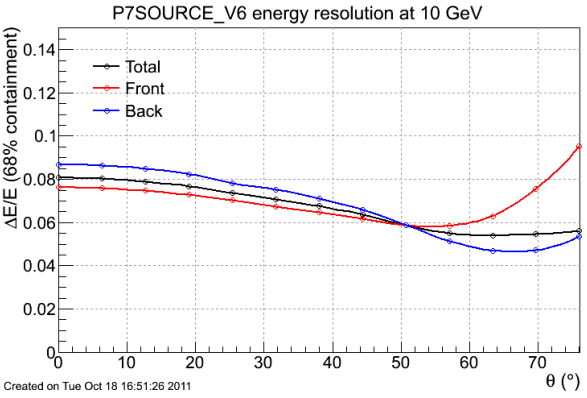
(a) Surface efficace en fonction de E



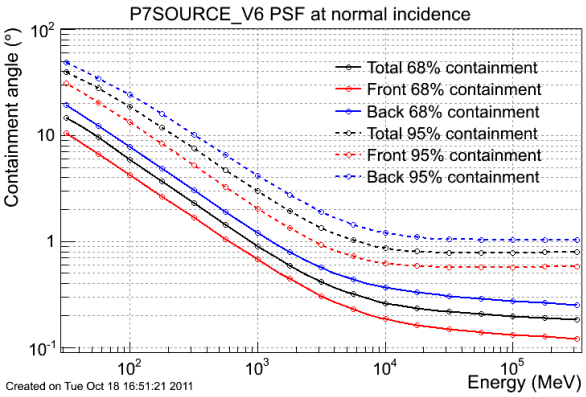
(b) Surface efficace en fonction de θ



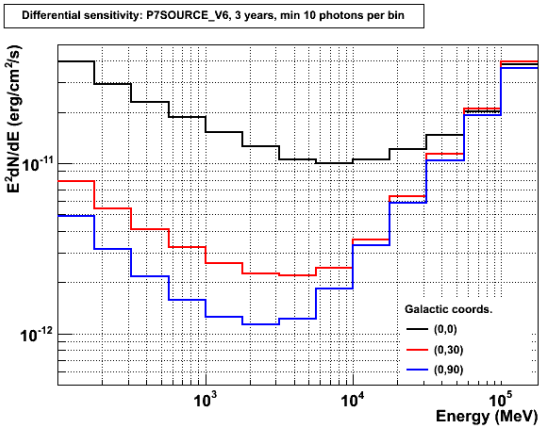
(c) $\Delta E/E$ en fonction de E



(d) $\Delta E/E$ en fonction de θ



(e) Résolution angulaire



(f) Sensibilité pour $T_{\mathrm{Obs}} = 3$ ans

FIGURE 9.5: Performances du *Fermi-LAT*. Explications dans le texte.

Classification	Identifiées	Associées
Sources galactiques		
Pulsars avec pulsation *Fermi*	83	...
Pulsars sans pulsation *Fermi*	...	25
Plérions	3	...
Rémanents de supernovæ	6	4
Rémanents de supernovæ/Plérions	...	58
Amas globulaires	0	11
Systèmes binaires X	4	0
Novæ	1	0
Incertaine	...	1
Total	196	
Sources extragalactiques		
BL Lac	7	429
FSRQ	17	353
Galaxies actives (hors blazars)	1	10
Radiogalaxie	2	10
Galaxies de Seyfert	1	5
Blazars de type incertain	0	257
Galaxies non-actives	2	4
Galaxies à haut taux de formation d'étoiles	0	4
Total	1102	
Sources non-identifiées		
Non associées	...	575

TABLE 9.1: Sources présentes dans le catalogue 2FGL.

un « flag » d'analyse qui indique un problème de reconstruction. Dans la plupart des cas, ces problèmes concernent des variations du flux ou de la localisation des sources selon le type de modèle d'émission diffuse utilisé, ou des difficultés rencontrées lors de l'ajustement spectral ou de position (plus d'informations dans la publication : [Nolan et al., 2012]).

9.2.4.4 L'après 2FGL

La prochaine version du catalogue, le *Fermi LAT 3-Year Source Catalog* (3FGL), qui contiendra les sources détectées avec 4 ans de données, n'est pour le moment pas publique[2]. On sait tout de même que la proportion de sources non-identifiées est du même ordre de grandeur que dans les catalogues 1FGL et 2FGL [Ballet et Burnett, 2013].

9.3 L'expérience *WISE*

WISE est une expérience spatiale dédiée à l'observation de l'intégralité du ciel dans l'infrarouge moyen à l'aide de 4 filtres micrométriques que l'on dénotera W_1, W_2, W_3 et W_4, correspondant respectivement aux longueurs d'onde 3,4 µm, 4,6 µm, 12 µm et 22 µm [Mainzer et al., 2005, Liu et al., 2008, Wright et al., 2010]. Successeur du projet au sol *The Two Micron All Sky Survey (2MASS)* [Skrutskie et al., 2006], *WISE* a pour principal objectif de fournir une image du ciel IR permettant de mener des études de population sur des sources extragalactique comme les *Ultra Luminous Infrared Galaxies* (ULIRG), de mieux comprendre les origines de la formation stellaire en détectant des naines brunes, ou même, la détection des géocroiseurs dans le système solaire.

2. Une première version du catalogue a été publiée le 8 janvier 2015 : [The Fermi-LAT Collaboration, 2015].

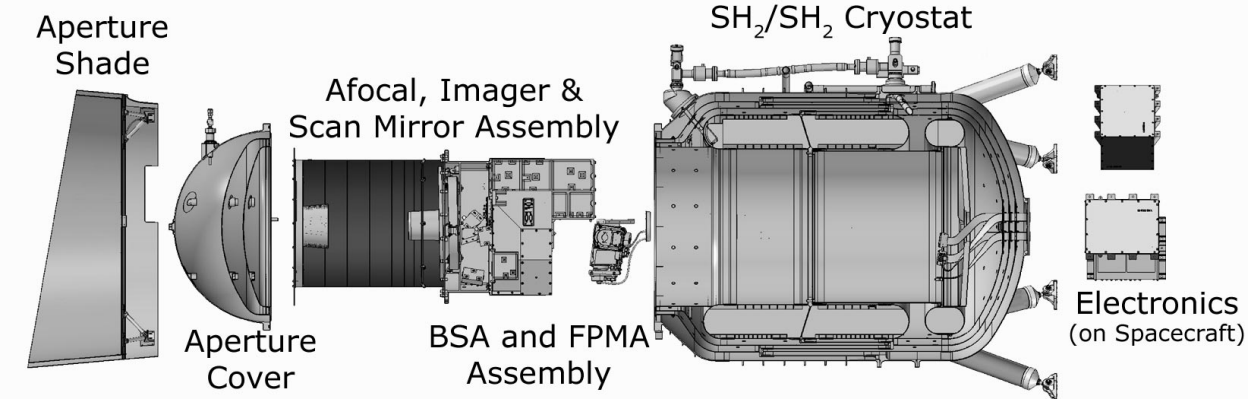


FIGURE 9.6: Charge utile du satellite *WISE* (Crédits : http://wise2.ipac.caltech.edu/).

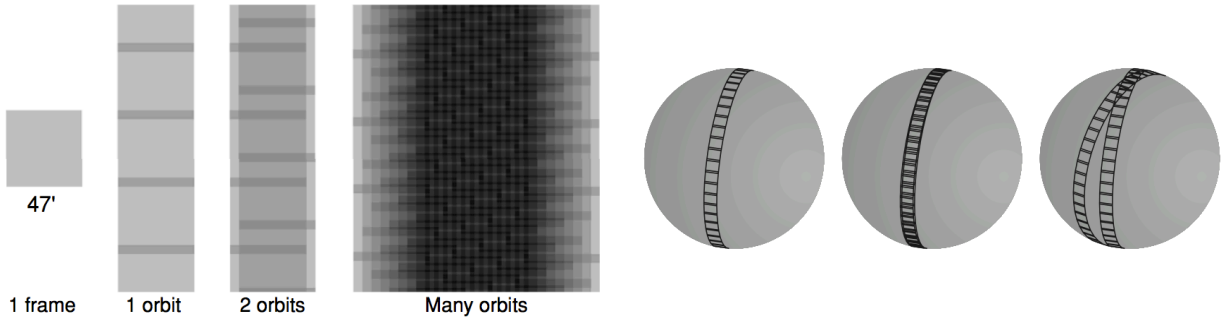
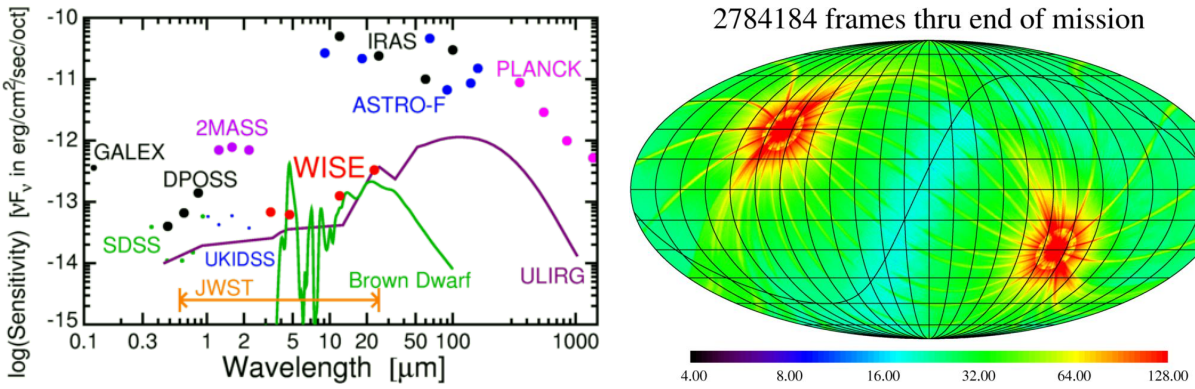


FIGURE 9.7: Stratégie de balayage du satellite *WISE*. À gauche, une image simple et les images prises lors d'une orbite, de deux orbites et d'une multitude d'orbites. À droite, succession des images prises lors d'une orbite, de deux orbites consécutives et de deux orbites séparées de 20 jours (Crédits : [Wright et al., 2010]).

9.3.1 Le satellite

Le satellite (figure 9.6) est doté d'un système optique composé d'une lentille de 40 cm, d'un miroir à balayage pour compenser le mouvement orbital, et de quatre plans focaux, chacun constitué de 1024 pixels conférant au télescope un champ de vue total de $47'$. Les détecteurs des filtres W_1 et W_2 utilisent un matériau à base de HgCdTe, et ceux des filtres W_3 et W_4 de plus basse fréquence sont constitués de SiAs. Le point de fonctionnement des deux détecteurs W_1 et W_2 est autour $\sim$30 K, tandis que les filtres de plus haute longueur d'onde sont refroidis à une température plus basse de $\sim$8 K. Le système optique est lui-même conservé à une température de $\sim$17 K. Un cryostat à base d'hydrogène solide assure le refroidissement des différents composants.

La trajectoire du satellite est suit orbite polaire de basse altitude, la distance moyenne du satellite à la Terre étant de $\sim$540 km. Le satellite effectue des grands cercles passant par les pôles nord et sud de l'écliptique, croisant le plan de l'écliptique aux longitudes équatoriales $l = \pm 90°$. Six mois sont nécessaires pour avoir une couverture totale du ciel. Le déplacement constant du télescope est compensé par le miroir à balayage qui permet d'avoir un « ciel constant » pour les 4 différents filtres. La stratégie d'observation de l'intégralité du ciel est représentée sur la Figure 9.7. La prise d'une image de $47'$ dure $\sim$11 s : 9 s sont consacrées au temps de pose, 1 s à la lecture des données et la dernière seconde au changement de la ligne de visée de $42'$. Deux images consécutives se chevauchent de 10 %.



(a) (b)

FIGURE 9.8: (a) Sensibilité différentielle (Crédits : http://www.astro.ucla.edu/%7Ewright/WISE/) et (b) acceptance (Crédits : http://wise2.ipac.caltech.edu/docs/release/allwise/expsup/index.html) du satellite *WISE*.

9.3.2 Performances de l'instrument

La résolution angulaire des différents filtres est d'environ 6″ pour les trois premiers et atteint 12″ à 22 µm. L'erreur de localisation est quant à elle de 0,01″ pour les filtres W_1 et W_2, de 0,09″ pour W_3 et enfin de 0,13″ pour W_4. Sur la Figure 9.8a sont représentées la sensibilité différentielle de l'instrument ainsi que celles de diverse expériences. On trouvera aussi les spectres typiques de naines brunes et de ULIRG pour comparaison. La première phase de l'expérience (*Full Cryogenic survey*), qui débuta le 7 janvier 2010 et se termina le 6 août 2010, a permis d'observer le ciel 1,2 fois. Après l'évaporation de l'hydrogène solide dans un des compartiments du cryostat, la seconde phase (*3-Bands Cryogenic survey*) qui dura environ 1 mois a permis de couvrir 30 % du ciel avec les filtres W_1, W_2 et W_3 avec une sensibilité proche de leurs valeurs initiales pour les deux premiers, et une sensibilité réduite pour le troisième. Le 20 septembre 2010, après l'épuisement des réserves d'hydrogène, la troisième phase (*NEOWISE Post-Cryo phase*) a permis d'observer 70 % du ciel à 3,4 µm et 4,6 µm jusqu'en février 2011. Durant cette période, les détecteurs HgCdTe des filtres W_1 et W_2 ont atteint l'équilibre thermique avec leur environnement (∼73 K) sans perte significative de sensibilité. Par contre, les filtres W_3 et W_4 sont arrivés à saturation et aucune information utile n'a pu être extraite. L'acceptance combinée des quatre filtres pour l'ensemble des différentes phases de l'expérience est représentée sur la Figure 9.8b.

9.3.3 Le catalogue

Le catalogue *WISE All-Sky Release Source Catalog*[3] comprend 563 921 584 objets détectés avec un rapport de signal sur bruit de plus de 5σ dans au moins l'un des quatre filtres. Les données utilisées dans cette thèse correspondent à la distribution « *All-Sky Data Release Products*[4] » qui comprend l'ensemble les données obtenues entre janvier et août 2010. L'immense quantité de données fournie par l'expérience *WISE* permet l'étude d'un grand pan de la zoologie astrophysique.

Les mesures de magnitudes relatives $m_i = -2.5 \log_{10} F_S^i + \text{Cste}$ à l'intensité dans le même filtre d'une source de référence (de magnitude nulle dans le filtre) permettent de construire trois couleurs indépendantes $c_1 = m_1 - m_2$, $c_2 = m_2 - m_3$ et $c_3 = m_3 - m_4$. La figure 9.9 représente la couleur c_1 en fonction de c_2 pour différents modèles d'objets astrophysiques observables par le télescope. Différences qui se traduisent par une place particulière dans le diagramme selon la classe de la source. Sur la figure, on peut observer que les objets les plus « rouges » dans c_1, qui sont plus

3. http://wise2.ipac.caltech.edu/docs/release/allsky/expsup/sec2_2.html
4. http://wise2.ipac.caltech.edu/docs/release/allsky/expsup/index.html

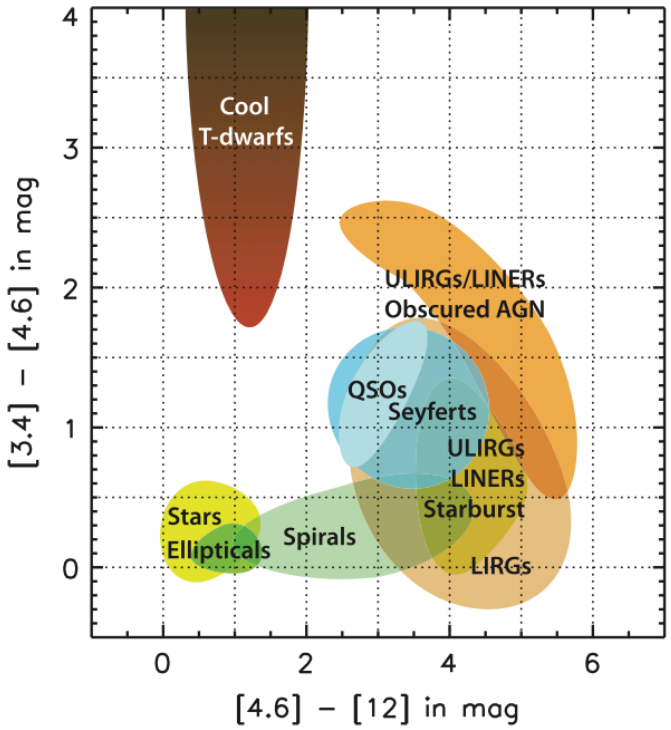


FIGURE 9.9: Zoologie astrophysique dans le diagramme couleurs $c_1(c_2)$ (Crédits : [Wright et al., 2010]).

brillants à 4,6 µm qu'à 3,4 µm, sont les naines brunes. Dans c_2 les objets les plus rouges sont les ULIRG, les AGN et les galaxies à haut taux de formation d'étoiles. Quant aux étoiles et les galaxies elliptiques, elles ont des couleurs c_1 et c_2 plus proches de zéro.

Bibliographie

A. A. Abdo, et al. Fermi/Large Area Telescope Bright Gamma-Ray Source List. ApJS, 183:46–66, July 2009. doi: 10.1088/0067-0049/183/1/46.

A. A. Abdo, et al. Fermi Large Area Telescope First Source Catalog. ApJS, 188:405–436, June 2010. doi: 10.1088/0067-0049/188/2/405.

M. Ackermann, et al. The Second Catalog of Active Galactic Nuclei Detected by the Fermi Large Area Telescope. ApJ, 743:171, December 2011. doi: 10.1088/0004-637X/743/2/171.

M. Ackermann, et al. The Fermi Large Area Telescope on Orbit: Event Classification, Instrument Response Functions, and Calibration. ApJS, 203:4, November 2012. doi: 10.1088/0067-0049/203/1/4.

W. B. Atwood, et al. The Large Area Telescope on the Fermi Gamma-Ray Space Telescope Mission. ApJ, 697:1071–1102, June 2009. doi: 10.1088/0004-637X/697/2/1071.

J. Ballet et T. H. Burnett. The fermi lat third source catalog. The Astroparticles Physics Conference. ICRC, 2013.

M. Doert et M. Errando. Search for gamma-ray-emitting active galactic nuclei in the Fermi-LAT unassociated sample using machine learning. *ArXiv e-prints*, December 2013.

E. C. Ferrara, et al. Fermi's Mystery Sources: Methods for Classification and Association. *ArXiv e-prints*, June 2012.

F. Liu, et al. Development of the Wide-field Infrared Survey Explorer (WISE) mission. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7017 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, July 2008. doi: 10.1117/12.790087.

A. K. Mainzer, et al. Preliminary design of the Wide-Field Infrared Survey Explorer (WISE). In H. A. MacEwen, editor, *UV/Optical/IR Space Telescopes: Innovative Technologies and Concepts II*, volume 5899 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 262–273, August 2005. doi: 10.1117/12.611774.

F. Massaro, et al. Identification of the Infrared Non-thermal Emission in Blazars. ApJ, 740:L48, October 2011. doi: 10.1088/2041-8205/740/2/L48.

F. Massaro, et al. Unveiling the Nature of Unidentified Gamma-Ray Sources. II. Radio, Infrared, and Optical Counterparts of the Gamma-Ray Blazar Candidates. ApJS, 206:13, June 2013. doi: 10.1088/0067-0049/206/2/13.

C. Meegan, et al. THE fermi gamma-ray burst monitor. ApJ, 702:791, September 2009. doi: 10.1088/0004-637X/702/1/791.

N. Mirabal, et al. Fermi's sibyl: mining the gamma-ray sky for dark matter subhaloes. MNRAS, 424:L64–L68, jul 2012.

P. L. Nolan, et al. Fermi Large Area Telescope Second Source Catalog. ApJS, 199:31, April 2012. doi: 10.1088/0067-0049/199/2/31.

M. F. Skrutskie, et al. The Two Micron All Sky Survey (2MASS). AJ, 131:1163–1183, February 2006. doi: 10.1086/498708.

The Fermi-LAT Collaboration. Fermi Large Area Telescope Third Source Catalog. *ArXiv e-prints*, January 2015.

E. L. Wright, et al. The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance. AJ, 140:1868, December 2010. doi: 10.1088/0004-6256/140/6/1868.

Chapitre 10

Recherche de candidats blazars à partir du catalogue 2FGL

Sommaire

10.1 Introduction . . . 171
10.2 Définition des différents échantillons . . . 172
10.3 Variables discriminantes . . . 173
10.4 Détermination des coupures de discrimination . . . 181
10.5 Candidats blazars . . . 183
10.A Stabilité des indices de dureté HR_{ij} . . . 193
10.B Paramètres TMVA des analyses BDT et MLP . . . 196
10.C Performances des méthodes individuelles BDT et MLP de l'analyse de la région « extragalactique » . . . 198
10.D Performances des méthodes individuelles BDT et MLP de l'analyse de la région « galactique » . . . 200
Bibliographie . . . 202

10.1 Introduction

Dans ce chapitre nous présentons une analyse multivariée combinant deux techniques MVA, développée afin d'identifier les candidats blazars parmi les 575 sources non-identifiées du catalogue 2FGL [Nolan et al., 2012]. Aux hautes latitudes ($|b| \geq 10°$) la grande majorité des sources identifiées du catalogue 2FGL sont des blazars, auxquels s'ajoute une cinquantaine de pulsars proches, ce qui laisse penser qu'un certain nombre de candidats blazars pourraient se trouver parmi les source non-identifiées à ces latitudes. De plus, la distribution en latitude galactique des sources identifiées (Figure 9.1) met en évidence un déficit de blazars au niveau du plan galactique, ce qui suggère encore une fois un potentiel de détection de nouveaux candidats blazars parmi les sources non-identifiées.

Un certain nombre d'auteurs ([Ackermann et al., 2012], [Ferrara et al., 2012], [Mirabal et al., 2012] et [Doert et Errando, 2013]) ont récemment utilisé des paramètres du catalogue 2FGL dans ce but. En effet, le catalogue 2FGL contient des paramètres décrivant les caractéristiques spectrales et la variabilité temporelle des sources qui permettent en les utilisant directement ou en créant de nouvelles quantités à partir de ces derniers, de mettre en évidence des différences de comportement selon le type de source (par exemple, une émission variable pour les blazars ou une courbure du spectre pour les pulsars). Par la suite nous passerons en revue les variables les plus importantes utilisées dans ces différentes études, nous montrerons les limites de certaines d'entre elles et nous en proposerons une nouvelle sélection qui sera utilisée pour construire une classification « blazar/non-blazar ». Il sera montré qu'il est utile d'optimiser l'approche différemment selon que l'on travaille près du plan galactique ou au contraire à de hautes latitudes.

Ce travail s'inspire de la démarche présentée par la collaboration *Fermi* dans [Ackermann et al., 2012] et [Ferrara et al., 2012] pour les catalogues 1FGL [Abdo et al., 2010] et 2FGL, ainsi que de l'analyse présentée par [Doert et Errando, 2013] pour le catalogue 2FGL. Il reprend également la suggestion de [Mirabal et al., 2012] de diviser l'espace des phases pour le traitement des sources

non-identifiées dans le plan et hors du plan galactique. Cependant, il se distingue par un soin particulier porté au choix des variables de discrimination et par le soucis d'avoir une bonne maîtrise des performances.

La section 10.2 présente les différents échantillons de sources utilisés pour la dérivation des coupures discriminantes ainsi que l'échantillon des sources non-identifiées. S'en suit à la section 10.3 une discussion sur les variables discriminantes utilisées par les auteurs précédemment cités ainsi que notre choix final de variables pour l'étude présentée ici. Les analyses multivariées et leurs performances respectives seront présentées à la section 10.4. Pour finir, la liste finale de candidats blazars obtenue au terme de cette étude sera présentée à la section 10.5.

10.2 Définition des différents échantillons

Construire une classification binaire « blazar »/« non-blazar » nécessite d'avoir des échantillons de sources identifiées comme blazars (« signal ») et non-blazars (« fond ») pour déterminer les coupures de discrimination optimales (chapitre 3). Nous présenterons ci-dessous brièvement les échantillons « signal » et « fond », ainsi que l'échantillon de sources non-identifiées. Ces échantillons sont constitués exclusivement de sources ponctuelles, puisque les blazars vus par le *Fermi-LAT* le sont.

De plus, la collaboration *Fermi-LAT* met en garde contre la fiabilité de la position ou des paramètres spectraux de 162 sources (les « c-sources ») qui se trouvent dans des régions particulières du plan galactique et dans lesquelles les systématiques liées au choix du modèle de fond diffus sont parfois mal maîtrisées. À ces sources s'ajoute un certain nombre de sources pour lesquelles au moins un problème a été rencontré lors de la procédure d'ajustement des paramètres spectraux ou de position, et qui sont signalées par un « flag » dans le catalogue 2FGL (pour plus d'informations voir la publication [Nolan et al., 2012]). Il en résulte qu'il pourrait exister des sources non-identifiées qui seraient des détections fortuites. La présentation des candidats blazars à la section 10.5 sera donc différenciée selon qu'un problème de reconstruction est signalé ou non dans le catalogue.

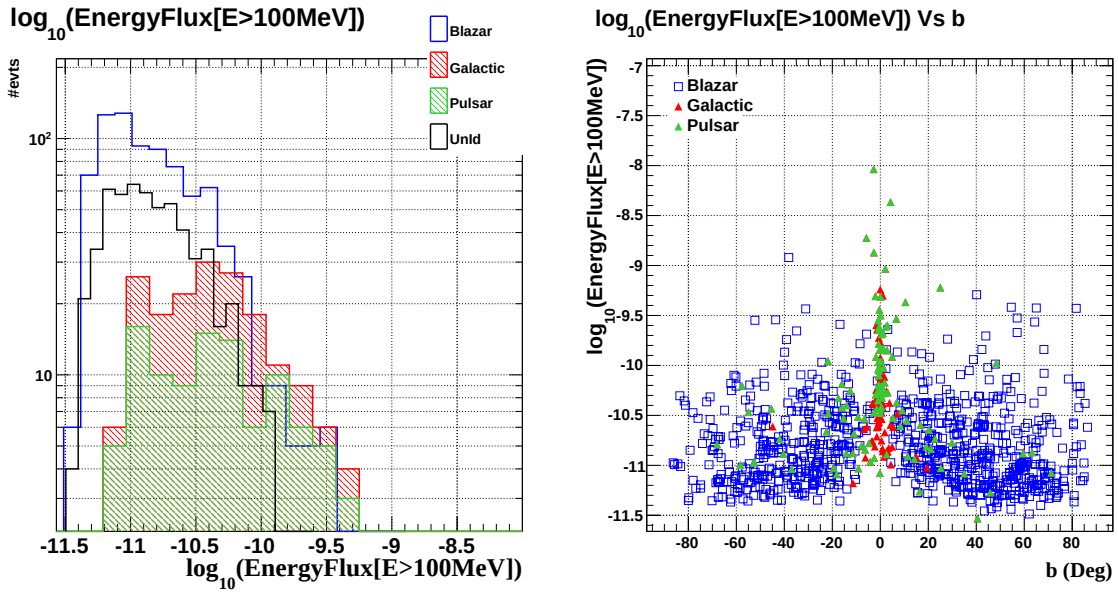
10.2.1 Échantillon « signal »

Un total de 806 blazars identifiés composés à 46 % de FSRQ et à 54 % de BL Lac constitue l'échantillon « signal ». À haute latitude ($|b|$ >10°) on compte 770 blazars identifiés (96 % de l'échantillon total) et aux plus faibles latitudes 36 (4 %).

10.2.2 Échantillons de « fond »

Les 187 sources γ ponctuelles identifiées ou associées de notre Galaxie sont majoritairement situées dans le plan galactique. Elles constitueront l'échantillon de fond pour l'étude dédiée à l'identification des candidats blazars aux faibles latitudes galactiques ($|b| \leq 10°$). Les pulsars constituent la classe dominante des sources galactiques (58 %) suivis par les SNR (33 %). Ces dernières sont exclusivement distribuées autour du centre galactique ($|b| < 6{,}0°$ et $l \in [-100, 100]$).

Hors du plan, les sources identifiées ou associées sont principalement des pulsars et des blazars (41 % des pulsars ont des latitudes galactiques supérieures à 10°). On cherchera à augmenter le pouvoir de séparation blazar/non-blazar en s'affranchissant de la dégradation du pouvoir discriminant principalement due aux restes de supernova. L'échantillon de fond pour l'étude hors du plan galactique ($|b|$ >10°) sera donc exclusivement constitué des 108 pulsars présents dans le catalogue 2FGL (le pulsar PSR J1730-3350 dont l'extension n'est pas déterminée est inclue dans les différents échantillons de fond).



(a) (b)

FIGURE 10.1: Propriétés des sources identifiées et non-identifiées. (a) Distribution du flux d'énergie entre 0,1 et 100 GeV des blazars (bleu), des sources galactiques (rouge), des pulsars (vert) et de l'ensemble des source non-identifiées (noir). (a) Flux d'énergie en fonction de la latitude galactique pour les blazars (bleu), les sources galactiques (rouge) et les pulsars (vert).

10.2.3 Échantillon des sources non-identifiées

L'échantillon des sources non-identifiées pour lequel seront recherchés les candidats blazars représente un total de 575 sources. Environ 53 % de ces sources ont une latitude galactique $|b| \leq 10°$ et 73 % d'entre elles sont signalées comme ayant au moins un problème de reconstruction. À plus haute latitude, la fraction de sources ayant un problème de reconstruction est de 26 %.

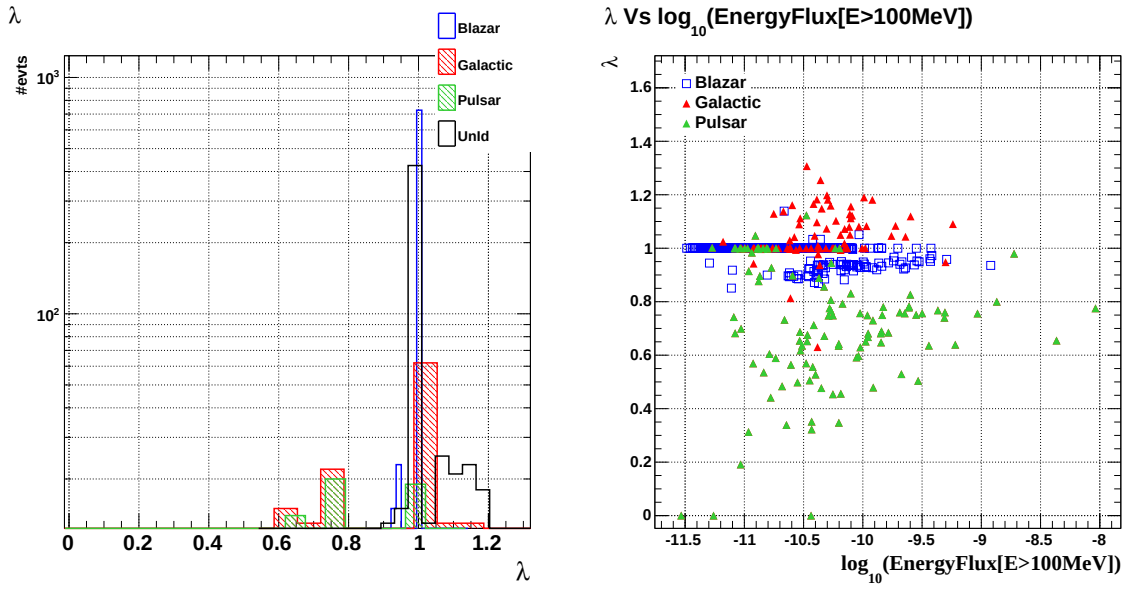
10.3 Variables discriminantes

10.3.1 Discrimination et brillance des sources

Deux classes de variables sont particulièrement intéressantes pour discriminer les blazars des autres classes de sources : les paramètres spectraux et les paramètres décrivant la variabilité de la source. On s'attend en effet à ce que les blazars soient variables, contrairement aux pulsars (on ne parle pas de la pulsation des sources), aux SNR et aux plérions. Autre fait remarquable, les pulsars présentent généralement une courbure dans le spectre représentée par une loi de puissance brisée ou une coupure exponentielle, contrairement aux blazars pour lesquelles une simple loi de puissance ou une loi de puissance parabolique rend généralement très bien compte de la forme spectrale.

Dans la mesure où il est plus aisé de caractériser de manière fine les caractéristiques spectrales et la variabilité des sources ayant un haut niveau de détection, il n'est pas étonnant de constater que le pouvoir discriminant de ces variables dépend du niveau de flux des sources. D'ailleurs ce comportement est un peu différent pour les sources dans le plan galactique qui, à significativité de détection égale, sont en moyenne plus brillantes que celles à haute latitude du fait du fort taux de contamination dû à l'émission diffuse dans certaines régions du plan.

Les sources non-identifiées étant en moyenne plus proches du seuil de détection du *LAT* que les sources galactiques et les blazars identifiées (Figure 10.1a), il est impératif d'utiliser des variables dont le pouvoir discriminant varie le moins possible avec le flux des sources. Sans cette précaution



(a) (b)

FIGURE 10.2: (a) Distribution de paramètre λ et (b) ses corrélations avec le flux en énergie ($E > 100\,\mathrm{MeV}$).

l'estimation des performances d'une analyse multivariée serait biaisée.

10.3.2 Paramètres spectraux

Les Figures 10.2 à 10.4 représentent un choix potentiel de variables spectrales discriminantes construites à partir des paramètres du catalogue 2FGL qui ont déjà été utilisés dans au moins une des études citées plus haut. Pour chacune d'entre elles, on trouvera à gauche la distribution du paramètre pour les blazars (bleu), les pulsars (vert), les autres sources galactiques (rouge), et les sources non-identifiées (noir). À droite, on trouvera la corrélation entre le dit paramètre et le flux en énergie au-delà de 100 MeV pour les blazars, les pulsars et les autres sources galactiques (donc pour les échantillons « signal » et « fond »).

Le catalogue 2FGL met à disposition deux indices spectraux déterminés à l'aide d'une minimisation entre 100 MeV et 100 GeV. Le premier, noté α, représente l'indice spectral d'une loi de puissance (PWL) tandis que le second, noté β, représente l'indice de l'hypothèse spectrale privilégiée, déterminée via un rapport de vraisemblance : α pour une PWL, l'indice à l'énergie de décorrélation pour une PWL parabolique et l'indice pour une PWL avec une coupure exponentielle. Ce dernier a été utilisé dans toutes les études citées précédemment. Le rapport λ défini comme $\lambda = \beta/\alpha$ a un pouvoir discriminant comparable à tous les niveaux de flux (Figures (10.2a)-(10.2b)), qui représente le fait qu'une courbure spectrale est plus courante pour les sources galactiques (notamment pour les pulsars).

Le catalogue fournit les flux intégrés dans cinq bandes en énergie (0,1–0,3 GeV, 0,3–1,0 GeV ; 1–3 GeV ; 3–10 GeV ; 10–100 GeV) avec lesquels ont peut construire 4 indices de dureté (en anglais *hardness ratios*) [Ackermann et al., 2012] :

$$\mathrm{HR}_{ij} = \frac{\Phi_j \langle E_j \rangle - \Phi_i \langle E_i \rangle}{\Phi_i \langle E_i \rangle + \Phi_j \langle E_j \rangle} \tag{10.1}$$

où Φ_i représente le flux intégré dans la bande i et $\langle E_i \rangle$ l'énergie moyenne. Dans le cas où la détermination du flux dans une bande est impossible, la valeur d'une limite supérieure à 2σ est utilisée par [Ackermann et al., 2012], ainsi que dans les publications qui ont suivi ([Ferrara et al., 2012],

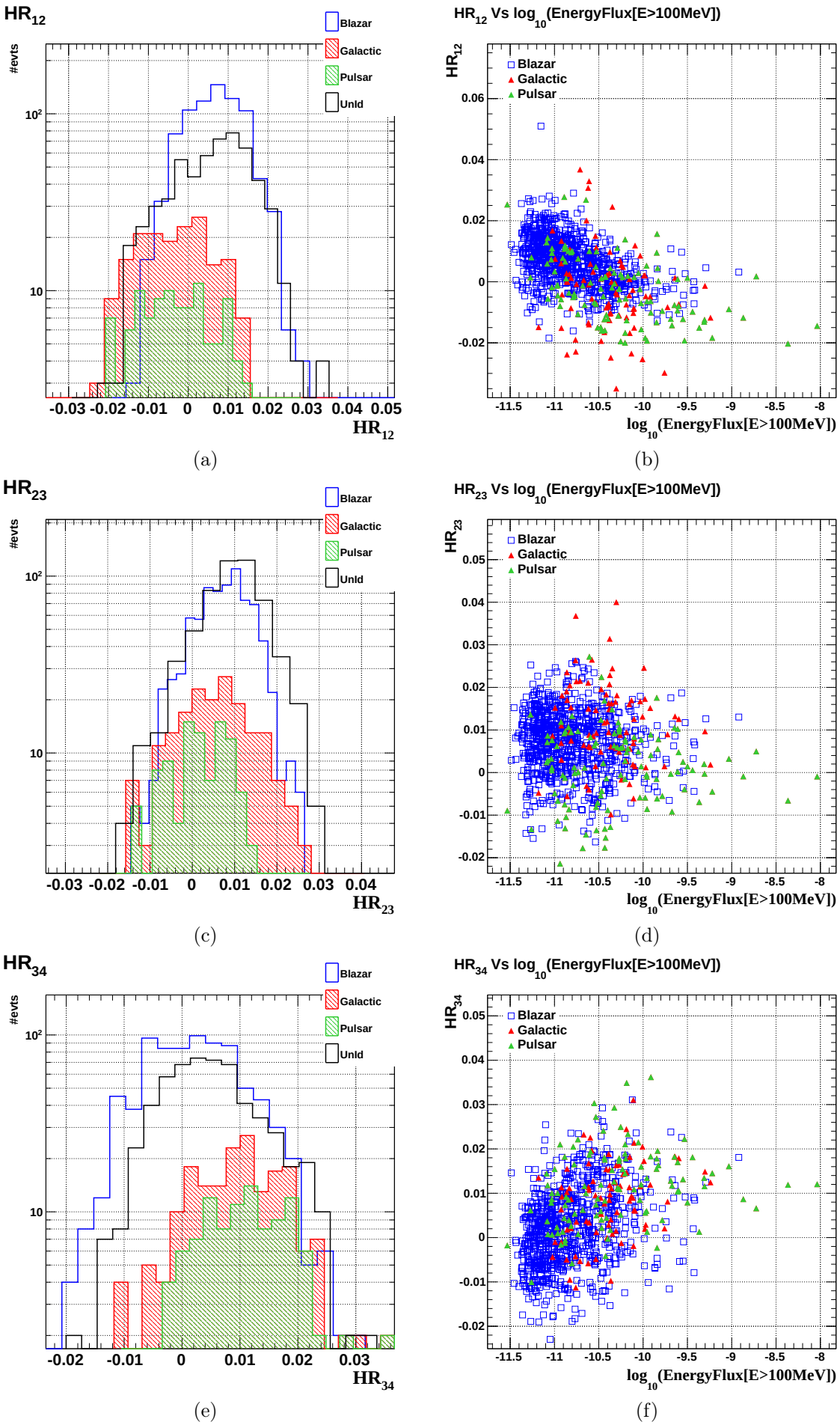


FIGURE 10.3: (a), (c), (e) Distributions des paramètres HR_{12}, HR_{23} HR_{34} et (b), (d) et (f) leurs corrélations avec le flux en énergie ($E > 100\,\text{MeV}$).

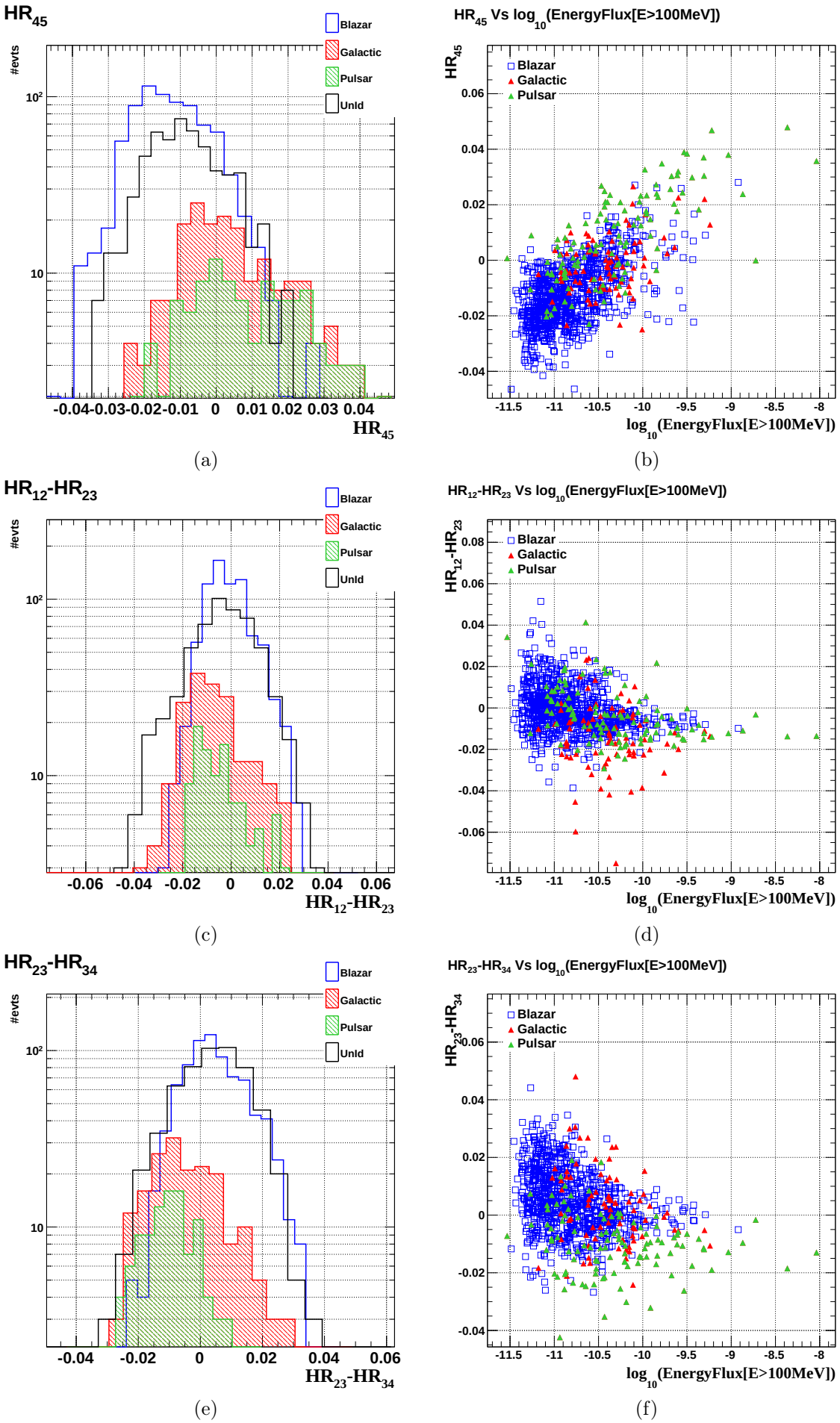


FIGURE 10.4: (a), (c), (e) Distributions des paramètres HR_{45}, $HR_{12} - HR_{23}$ $HR_{23} - HR_{34}$ et (b), (d) et (f) leurs corrélations avec le flux en énergie ($E > 100\,\mathrm{MeV}$).

[Mirabal et al., 2012] et [Doert et Errando, 2013]), en lieu et place du flux dans le calcul des HR_{ij}. Au premier abord, les 4 indices HR_{12}, HR_{23}, HR_{34} et HR_{45} ont un pouvoir discriminant prometteur (Figures 10.3a, 10.3c, 10.3e et 10.4a), bien que HR_{45} présente une certaine dépendance de ce pouvoir discriminant au niveau du flux des sources (Figure 10.4b). Les 4 *hardness ratios* ont été utilisés par [Ferrara et al., 2012] et [Doert et Errando, 2013], tandis que [Mirabal et al., 2012] a utilisé le rapport de flux connexes dans deux bandes en énergie. Notez que dans ce travail nous utilisons une définition légèrement différente des paramètres HR_{ij}, le logarithme en base 10 du flux sera utilisé à la place du flux.

Cependant, le rôle des limites supérieures dans le calcul des indices complique la maîtrise du pouvoir discriminant. Dans l'Appendice 10.A, les distributions des *hardness ratios* et du flux en énergie des sources sont représentées pour les sources galactiques (rouge) et les blazars (bleu), selon qu'il existe (non hachuré) ou non (hachuré), au moins une limite supérieure pour le calcul des paramètres. Naturellement, on remarque que pour chaque paramètre HR_{ij}, l'absence ou la présence de limites supérieures est corrélée avec la brillance de la source. Les blazars et les sources galactiques ne possédant pas de limite supérieure dans le calcul des indices sont plus brillants, et vice versa (Figures 10.10b, 10.10d, 10.11b et 10.11d). Pour les blazars et les sources galactiques, la comparaison des distributions du paramètre HR_{12} (Figure 10.10a) montre que la présence de limites supérieures tire de manière significative les distributions vers les spectres durs ($HR_{12} > 0$). En ce qui concerne l'indice HR_{23} (Figure 10.10c), les deux distributions (avec ou sans limites supérieures) sont relativement stables dans le cas des blazars. Par contre, elles diffèrent notablement en fonction de la brillance de la source pour les sources galactiques, les sources ayant au moins une limite supérieure ont des spectres exclusivement durs. Dans le cas du paramètre HR_{34}, la présence de limites supérieures décale les distributions des sources galactiques et des blazars vers les indices négatifs. Quant au dernier indice HR_{45}, on peut voir sur la Figure 10.11c que les sources galactiques dont l'indice contient des limites supérieures tendent à avoir un spectre plus mou. De plus, le pouvoir discriminant du coté des indices positifs de la distribution est exclusivement effectif pour les sources galactiques brillantes.

À l'aide des indices HR_{ij} on peut construire une quantité mesurant la courbure du spectre (*hardness slope*) introduite par [Ackermann et al., 2012], en faisant la différence entre deux indices de dureté correspondant à deux bandes en énergies connexes (par exemple $HR_{23} - HR_{34}$). Les Figures 10.4c, 10.4e et 10.5a représentent les distributions de trois paramètres de courbure construits de cette manière. Les deux premiers paramètres $HR_{12} - HR_{23}$ et $HR_{23} - HR_{34}$ présentent un pouvoir discriminant indépendant du niveau du flux de la source. Par contre, le pouvoir de discrimination de $HR_{34} - HR_{45}$ est corrélé avec la brillance de la source (Figure 10.5b). La présence de limites supérieures dans les indices complique l'étude de leur stabilité selon la brillance de la source et donc la compréhension de leur distribution. Seul le paramètre de courbure $HR_{23} - HR_{34}$ parmi les 3 existants a été utilisé dans les études de [Ferrara et al., 2012] et de [Doert et Errando, 2013].

Le catalogue 2FGL donne aussi accès à la significativité de la courbure σ_c définie comme le rapport de vraisemblance entre les hypothèses loi de puissance et une loi présentant une courbure du spectre (PWL parabolique ou PWL avec une coupure exponentielle) (Figures 10.5c et 10.5d). Cette variable possède un pouvoir discriminant conséquent mais est manifestement très dépendante du niveau de détection de la source. Elle a néanmoins été utilisée par [Mirabal et al., 2012].

Notons que [Doert et Errando, 2013] ont introduit et utilisé la courbure normalisée $\widetilde{\sigma}_c$ définie par $\widetilde{\sigma}_c = \sigma_c/\sigma$. Elle présente un pouvoir discriminant plus important et permet de réduire la dépendance de ce dernier au flux de la source (Figures 10.5e et 10.5f). Il persiste cependant une dépendance résiduelle, le pouvoir de séparation diminuant légèrement pour les sources très brillantes.

10.3.3 Paramètres de variabilité

Pour caractériser la variabilité d'une source, le catalogue 2FGL donne accès à l'indice de variabilité TS_{var} défini comme le rapport de vraisemblance entre l'hypothèse d'une source variable et l'hypothèse d'une variabilité nulle (flux constant) sur les deux ans d'intégration du signal, répartis

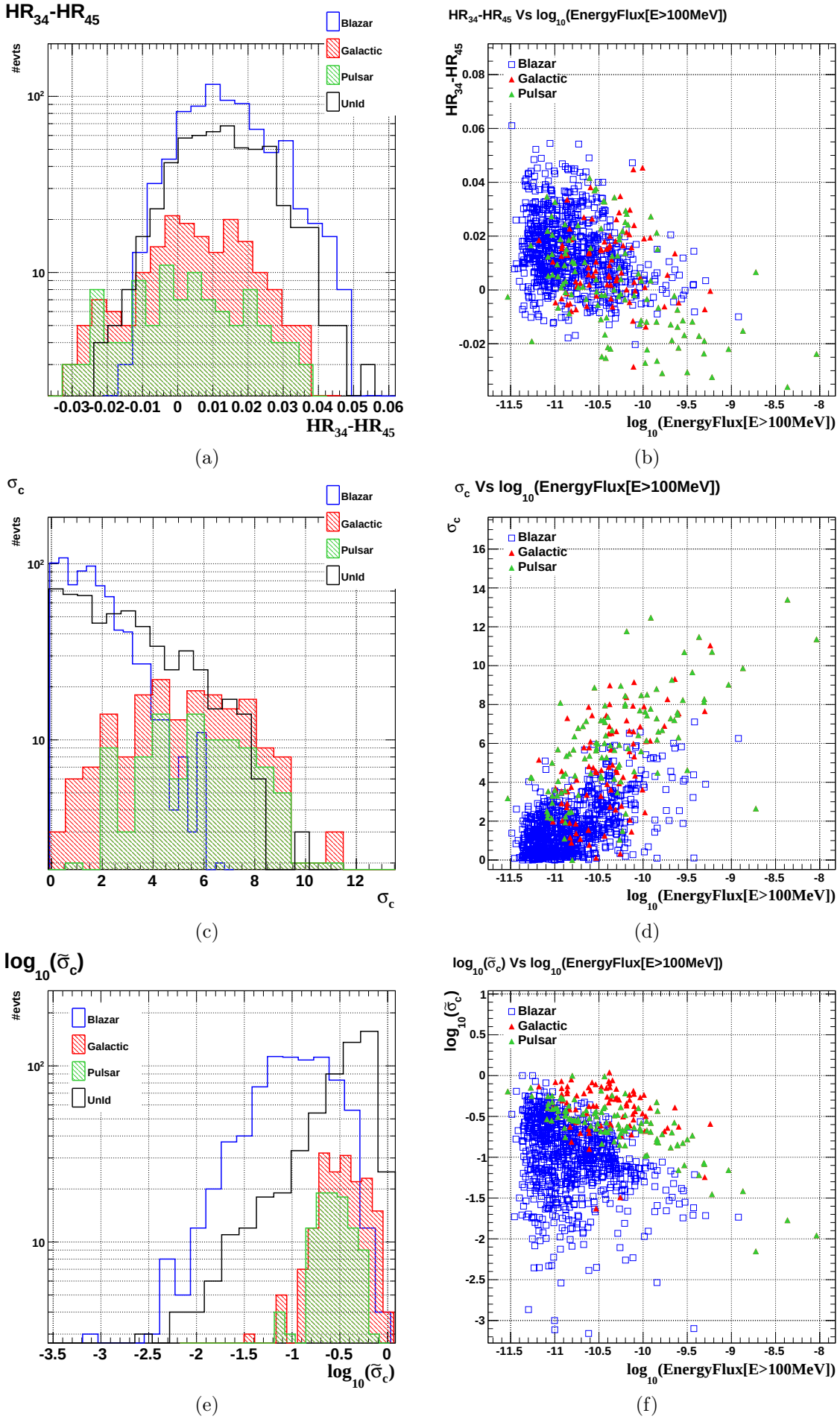


FIGURE 10.5: (a), (c), (e) Distributions des paramètres $HR_{34} - HR_{45}$, σ_c et $\widetilde{\sigma}_c$ (b), (d) et (f) leurs corrélations avec le flux en énergie ($E > 100\,\mathrm{MeV}$).

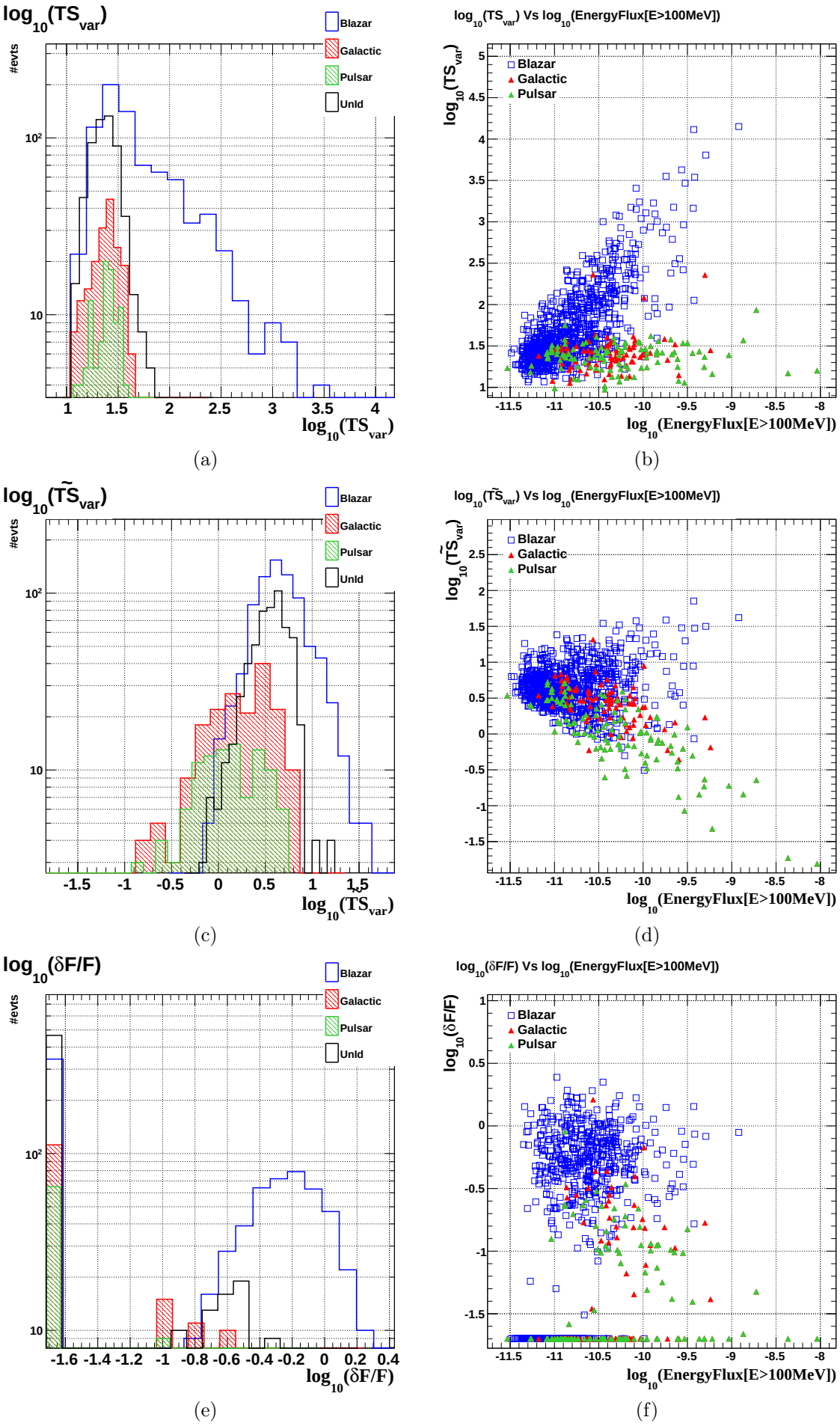


FIGURE 10.6: (a), (c), (e) Distributions des paramètres $\mathrm{TS_{var}}$, $\widetilde{\mathrm{TS}}_{\mathrm{var}}$ et $\delta F/F$ (b), (d) et (f) leurs corrélations avec le flux en énergie ($E > 100\,\mathrm{MeV}$).

en 24 intervalles de temps. On peut observer sur la Figure 10.6b son très grand pouvoir discriminant et une importante corrélation entre TS_{var} et le flux d'énergie pour les blazars, qui est inexistante pour les sources galactiques, et qui mène à un pouvoir de séparation qui est fonction de la brillance des sources. L'étude de [Mirabal et al., 2012] utilise ce paramètre.

Pour réduire la dépendance du pouvoir discriminant de TS_{var} à la brillance de la source, [Doert et Errando, 2013] normalisent TS_{var} par sa significativité de détection : $\widetilde{TS}_{var} = TS_{var}/\sigma$. On note tout de même qu'un certain niveau de corrélation du pouvoir discriminant de $\widetilde{TS}_{var}$ avec le flux est toujours présent (Figure 10.6d).

Enfin, le catalogue 2FGL met à disposition le flux intégré d'une source et son erreur associée pour chaque mois composant les 2 ans d'observation du *Fermi-LAT*. Ces informations sont utilisées par [Ackermann et al., 2012] en introduisant la variabilité fractionnelle, construite de manière à s'affranchir de la brillance des sources, en comparant la variance de l'excès par rapport aux fluctuations statistiques et systématiques attendues :

$$\delta F/F = \sqrt{\frac{\sum_i (\Phi_i - \langle\Phi\rangle)^2}{(N_{\rm int}-1)\langle\Phi\rangle^2} - \frac{\sum_i \sigma_i^2}{N_{\rm int}\langle\Phi\rangle^2} - f_{\rm rel}^2} \tag{10.2}$$

où N_{int} est le nombre d'intervalles de temps (24), Φ_i et σ_i le flux et son écart type statistique pour l'intervalle i, $\langle\Phi\rangle$ la moyenne des flux intégrés dans les 24 intervalles, et f_{rel} l'erreur systématique sur la mesure du flux (2 % pour le catalogue 2FGL). Si la source n'est pas détectée dans un intervalle, une limite supérieure à 2σ est calculée, et l'erreur associée au flux est remplacée dans le catalogue 2FGL par $\sigma_i = 0.5(\Phi_i^{UL} - \Phi_i)$. Dans le cas où l'argument sous la racine de l'équation 10.2 est nul ou si $\delta F/F$ est inférieur à f_{rel}, on imposera $\delta F/F = f_{rel}$. La distribution de $\delta F/F$ ainsi que sa dépendance à la brillance des sources sont représentées sur les Figures 10.6e et 10.6f. Respectivement 58 % et 41 % des blazars et des sources galactiques ont une variabilité fractionnelle supérieure à l'erreur systématique. En considérant l'ensemble des sources non-identifiées, 19 % d'entre elles ont une variabilité fractionnelle supérieure à f_{rel}. Une légère corrélation entre le flux des sources galactiques et le paramètre $\delta F/F$ est visible.

10.3.4 Paramètres sélectionnés pour la discrimination

Après avoir passé en revue l'ensemble des paramètres disponibles qui caractérisent la variabilité et la forme spectrale des différents échantillons de sources, la sélection d'un paramètre comme variable discriminante à été validée à l'aide de deux critères : l'absence de dépendance ou la faible dépendance du pouvoir de séparation par rapport au flux des sources et la maîtrise de sa distribution.

Utiliser un paramètre dont le pouvoir discriminant augmente radicalement avec la brillance de la source dans une analyse multivariée conduit à une estimation des performances qui serait surestimée. En effet, les sources non-identifiées sont de manière générale moins brillantes que les sources utilisées lors de la phase d'entraînement, ainsi que lors de la phase d'estimation du taux de fausses associations, que l'on notera ε_{Bkg}, et du taux d'identifications des sources de type blazars, que l'on notera ε_{Sig}. En prenant en compte des sources dotées de paramètres très discriminants et non représentatifs de l'échantillon des sources non-identifiées, ε_{Bkg} serait sous-estimé et ε_{Sig} surestimé. Pour cette raison, les paramètres HR_{45}, TS_{var} et σ_c ont été exclus de l'analyse.

Les trois premiers indices de dureté, bien qu'ils présentent de prime abord un pouvoir de séparation intéressant et relativement indépendant du niveau de détection des sources, souffrent comme nous l'avons montré à la section 10.3.2 d'une dépendance de leurs distributions à la présence ou non de limites supérieures dans le calcul des HR_{ij}, ce qui complique la maîtrise de leur pouvoir discriminant. L'absence de contrôle que procurent ces paramètres sur les performances finales nous ont donc conduit à n'utiliser aucun des *hardness ratios* et des *hardness slopes*.

Au final, nous retiendrons pour ce travail un nombre limité de paramètres pour construire les analyses multivariées. C'est tout d'abord le cas du paramètre spectral λ dont les valeurs sont proches de 1 pour la majorité des blazars, alors qu'elles s'étalent de manière plus importante pour

les sources galactiques. On utilisera aussi la variable normalisée $\widetilde{\sigma}_c$ qui permet d'obtenir un pouvoir discriminant important et qui possède en plus un pouvoir de discrimination qui diminue avec la brillance des sources. Cela signifie que les performances finales seront donc légèrement biaisées, mais de manière conservative ($\varepsilon_{\mathrm{Sig}}$ devrait être en réalité légèrement plus élevé et $\varepsilon_{\mathrm{Bkg}}$ légèrement moins important que ce que nous estimerons sur les lots de sources bien identifiées). Les paramètres sélectionnés qui caractérisent la variabilité des sources sont au nombre de deux. Le premier $\delta F/F$ permet de déterminer de manière significative la variabilité d'une fraction non négligeable des sources composant les différents échantillons. Enfin, le dernier paramètre sélectionné $\widetilde{\mathrm{TS}}_{\mathrm{var}}$ présente un pouvoir discriminant qui dépend peu du flux des sources dans la région correspondant au flux des sources non-identifiées, même si ce pouvoir discriminant augmente pour des flux plus importants. Ce paramètre joue un rôle important dans les performances obtenues ici.

10.4 Détermination des coupures de discrimination

10.4.1 Sélection et réglages des analyses multivariées

10.4.1.1 La sélection

Une fois les paramètres de discrimination sélectionnés, il convient de choisir une méthode MVA appropriée au problème, ici un problème de classification binaire. La construction d'une analyse se déroule en deux étapes. Lors de la première phase, une fraction des échantillons « signal » et « fond », qu'on appellera le lot d'entraînement, est utilisée pour construire un paramètre de discrimination ζ qui maximise les différences entre ces deux échantillons. La seconde phase consiste à utiliser le reste de la statistique, appelé le lot test, pour dériver les quantités $\varepsilon_{\mathrm{Sig}}$ et $\varepsilon_{\mathrm{Bkg}}$ en fonction de ζ. Cette phase permet également de détecter la présence éventuelle d'effets liés à un sur-entraînement et permet plus généralement de vérifier le bon déroulement de l'analyse à l'aide d'un test statistique de KS qui compare la compatibilité des distributions des valeurs ζ entre les lots d'entraînement et de test, et ce pour les échantillons « signal » et « fond ».

La librairie TMVA propose un certain nombre de méthodes multivariées adaptées à ce type de problème, en particulier les méthodes forêt d'arbres décisionnels (RF), réseaux de neurones (NN) de type MLP, machine à vecteurs de support (SVM) et BDT couramment utilisées dans différents domaines. Après un certain nombre de réglages (voir sous-sections ci-dessous) et d'entraînements [1] pour chacune des méthodes, nous avons pu constater que les quatre méthodes permettent d'obtenir des performances comparables, sans que l'une d'entre elle ne surpasse les autres.

Dans l'optique d'avoir une discrimination robuste, deux méthodes multivariées dont les décisions vont être combinées ont été sélectionnées pour construire les analyses dédiées à chacune des régions « galactique » et « extragalactique ». Nous avons le choix entre deux approches : minimisation d'une fonction d'erreur pour MLP et SVM, et utilisation d'arbres de décision pour BDT et RF. Notre sélection s'est portée sur les méthodes BDT et MLP privilégiant ainsi l'utilisation d'approches de nature différente. Le choix de BDT à la place de RF et de MLP à la place de SVM relève de l'arbitraire, les performances étant comparables pour ce problème particulier.

À titre de comparaison, l'étude [Ferrara et al., 2012] développée dans [Ackermann et al., 2012] utilise la combinaison des méthodes de classification de RF et de régression logistique, celle de [Doert et Errando, 2013] la combinaison de RF et d'un réseau de neurones, et enfin [Mirabal et al., 2012] la décision simple de l'analyse RF.

10.4.1.2 Réglages de la méthode BDT

Pour atteindre avec la méthode BDT des performances similaires à celles obtenues avec les autres techniques multivariées précédemment citées, il est nécessaire d'appliquer l'algorithme AdaBoost

1. Ce travail a été effectué en aveugle, tous les essais ayant été exclusivement effectués sur les lots de sources identifiées.

avec un ré-échantillonnage des évènements utilisés pour la génération d'un arbre de décision à l'aide de l'option TMVA *UseBaggedBoost*. À chaque itération, seule une fraction des évènements sont choisis aléatoirement parmi les échantillons de « signal » et « fond ». Chaque évènement est sélectionné X fois, où X représente un nombre aléatoire généré à partir d'une loi de probabilité de Poisson de paramètre 0,6. Cette valeur est proposée par défaut par TMVA, la stabilité des résultats basée sur des valeurs un peu différentes a été testée. L'utilisation de la méthode BDT avec l'algorithme de *bagging* permet de réduire les possibles fluctuations statistiques dont a l'air de souffrir l'algorithme AdaBoost originel pour ce problème particulier.

Différents réglages de la méthode BDT ont été testés et nous avons choisi une configuration qui optimise les performances sur les lots de test. La profondeur maximale d'un arbre est fixée à 3 et on génère une forêt d'arbres contenant 400 individus pour les deux analyses « galactique » et « extragalactique ». En ce qui concerne la fraction d'évènements minimale que doit contenir un nœud pour développer une feuille, celle-ci a été fixée à 5 %. Enfin, la détermination de la granularité optimale de l'espace des phases a été laissée à TMVA. Un taux d'apprentissage de $\beta = 0{,}2$ a été utilisé. L'ensemble des paramètres TMVA utilisés sont résumés dans la Table 10.4 de l'Appendice 10.B.

10.4.1.3 Réglages de la méthode MLP

Les échantillons signal et fond étant de dimensions différentes, il est nécessaire de normaliser les évènements de manière à ce que les différents lots aient le même nombre total d'évènements (ce qui est naturellement fait par la méthode BDT). Dans la librairie TMVA, cette normalisation s'effectue avec l'option *NormMode=EqualNumEvents*. Aucune différence n'a été observée dans les performances du réseau de neurones selon que le nombre de couches de neurones cachées le constituant est égal ou supérieur à l'unité. On préférera utiliser le modèle le plus simple, c'est-à-dire un réseau de neurones constitué d'une seule couche de neurones intermédiaire de $N_{\rm var} + 5$ neurones (option par défaut de TMVA). La minimisation de la fonction de mauvaises classifications s'effectue avec l'algorithme de *Back Propagation* dans lequel les poids sont modifiés à chaque passage d'un évènement dans le réseau. La régularisation des poids n'a pas été nécessaire. Les valeurs standards pour le taux d'apprentissage et sa décroissance ont été utilisées. L'ensemble des paramètres est résumé dans l'Appendice 10.B (Table 10.5).

10.4.2 Méthodologie pour la détermination des performances

Les paramètres des méthodes BDT et MLP étant réglés, il convient d'estimer leurs performances individuelles ainsi que les performances obtenues en couplant leurs décisions.Une approche simple a été considérée ici, qui consiste à se donner pour chacune des méthodes BDT et MLP un niveau de contamination maximal de $\varepsilon_{\rm Bkg} \approx 10\,\%$ pour obtenir la valeur du paramètre de discrimination finale $\zeta^\star$ ainsi que le taux d'identification des sources de type blazar $\varepsilon_{\rm Sig}$ associé. L'entraînement des méthodes BDT et MLP est réalisé en parallèle avec les mêmes lots d'évènements, correspondant à un tirage aléatoire de 50 % des sources parmi les échantillons de sources associées ou identifiées. Le reste de la statistique, le lot test, est ensuite utilisé pour obtenir les quantités $\varepsilon_{\rm Bkg}$ et $\varepsilon_{\rm Sig}$ en fonction de ζ pour chacune des deux méthodes. Différents lots d'entraînement peuvent conduire à des performances légèrement différentes, ce qui nous a conduit à réaliser 200 entraînements pour déterminer les valeurs moyennes des taux $\varepsilon_{\rm Bkg}$ et $\varepsilon_{\rm Sig}$ ainsi qu'une estimation de l'écart-type de ces dernières.

Les performances des analyses finales qui couplent les décisions BDT et MLP sont estimées de la même manière. Lors de chaque entraînement, le taux d'identification des sources de type blazar et de le taux de fausses associations sont calculés en se basant sur les sources qui sont reconnues comme candidats blazars par les deux méthodes multivariées. Ce choix entraîne que la contamination de l'échantillon sera plus faible que les ∼10 % initialement imposés au détriment d'une légère détérioration du taux d'identification des blazars des méthodes individuelles BDT et MLP. Une fois les performances déterminées des analyses « galactique » et « extragalactique », on

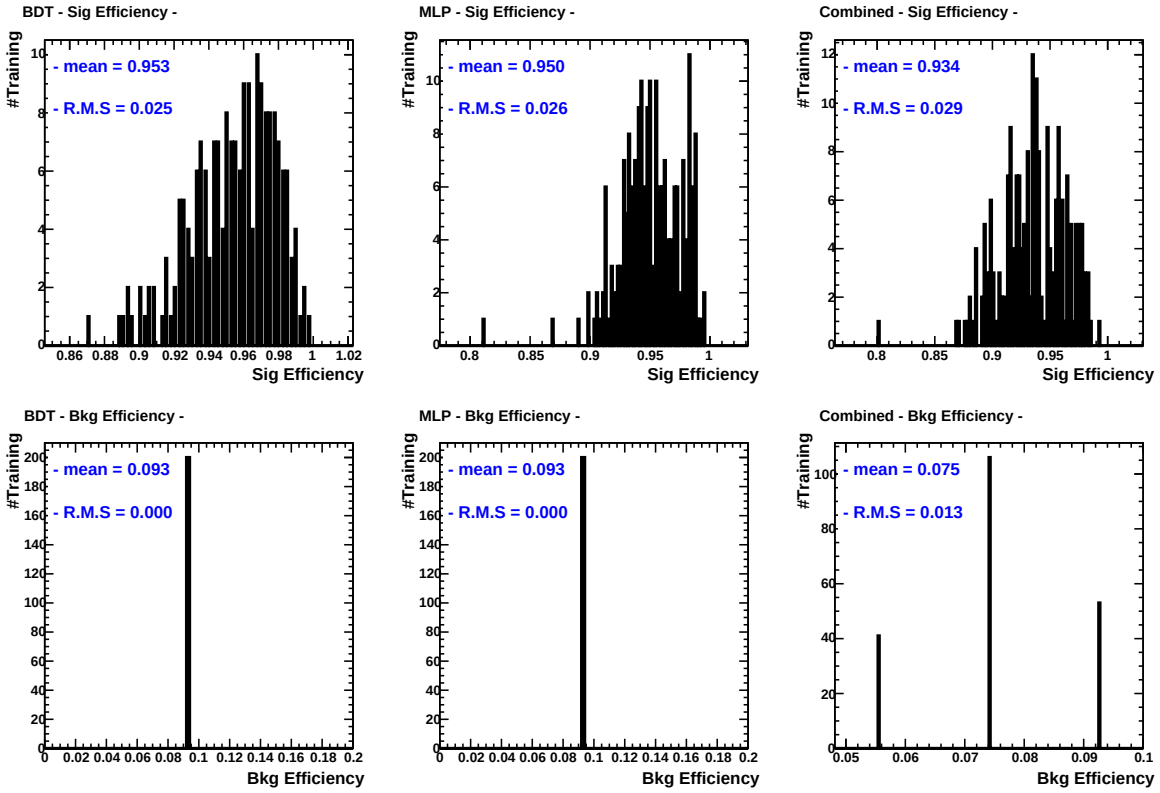


FIGURE 10.7: Performances de l'analyse « extragalactique ». Sont représentées respectivement à gauche, au milieu et à droite, les distributions des quantités ε_{Sig} et ε_{Bkg} de la méthode BDT, MLP et du couplage des deux méthodes multivariées, pour 200 entraînements.

choisira pour chacune d'entre elles des entraînements dédiés avec les méthodes BDT et MLP en accord avec les performances moyennes estimées des méthodes individuelles.

10.4.3 Performances de l'analyse « extragalactique »

La Figure 10.7 représente les distributions pour l'analyse « extragalactique » des taux d'identifications des sources de type blazar et de fausses associations des méthodes BDT et MLP, ainsi que les performances finales qui couple les deux décisions des 200 entraînements effectués. Pour chaque méthode la valeur moyenne et l'écart-type des distributions sont indiqués. Comme indiqué précédemment, les performances des méthodes BDT et MLP sont comparables. Avec la méthode BDT le taux d'identification des blazars est en moyenne de $(95{,}3 \pm 2{,}5)\,\%$ pour un taux de mauvaises classifications fixé à $10\,\%$ [2]. De la même manière, la méthode MLP permet d'obtenir un taux d'identification des blazars moyen de $\varepsilon_{\text{Sig}} = (95{,}3 \pm 2{,}6)\,\%$ pour $\varepsilon_{\text{Bkg}} \sim 10\,\%$. En combinant les deux décisions, on obtient les performances finales moyennes de l'analyse « extragalactique » qui sont : $\varepsilon_{\text{Sig}} = (93{,}4 \pm 2{,}9)\,\%$ et $\varepsilon_{\text{Bkg}} = (7{,}5 \pm 1{,}3)\,\%$. Le choix du couplage des décisions est justifié par la réduction du taux de fausses associations de $\sim$20 % au détriment d'une légère baisse de $\sim$2 % du taux d'identification des sources de type blazar.

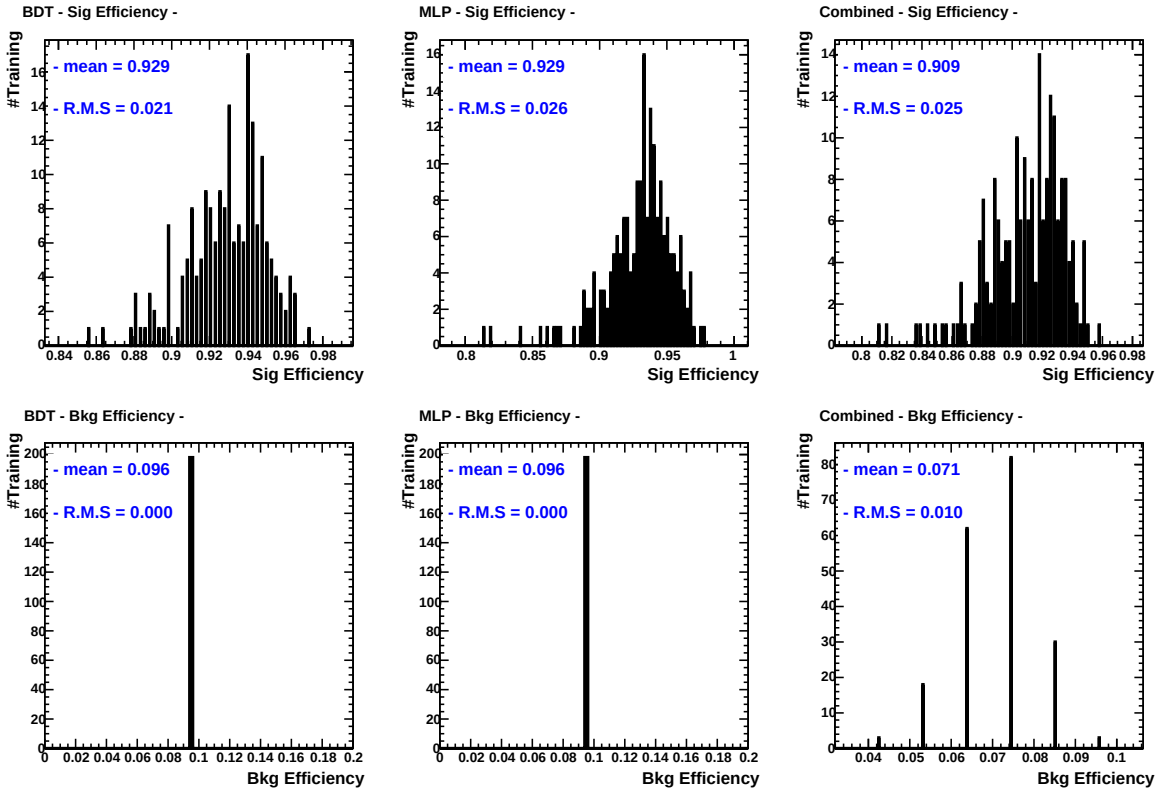


FIGURE 10.8: Performances de l'analyse « galactique ». Sont représentées respectivement à gauche, au milieu et à droite, les distributions des quantités $\varepsilon_{\mathrm{Sig}}$ et $\varepsilon_{\mathrm{Bkg}}$ de la méthode BDT, MLP et du couplages des deux méthodes multivariées, pour 200 entraînements.

10.4.4 Performances de l'analyse « galactique »

De la même manière, les performances de l'analyse « galactique » sont illustrées sur la Figure 10.8. Les méthodes BDT et MLP permettent toutes les deux d'atteindre des taux d'identification de blazars équivalents qui sont respectivement de $(92{,}9 \pm 2{,}1)\,\%$ et de $(92{,}9 \pm 2{,}6)\,\%$ pour un taux de fausses associations fixé à $\sim$10 %. La combinaison des deux décisions réduit légèrement les taux d'identification des sources de type blazar de $\sim$2 % et de fausses associations de $\sim$25 % : $\varepsilon_{\mathrm{Sig}} = (90{,}9 \pm 2{,}5)\,\%$ et $\varepsilon_{\mathrm{Bkg}} = (7{,}1 \pm 1{,}0)\,\%$.

10.5 Candidats blazars

Les coupures $\zeta^{\star}$ ainsi que les performances des analyses associées sont récapitulées dans la Table 10.1. On pourra trouver les performances individuelles des méthodes BDT et MLP pour les analyses des régions extragalactique et galactique respectivement dans l'Appendice 10.C et l'Appendice 10.D. Une fois les phases d'entraînement terminées et les choix techniques fixés, les analyses utilisant les décisions BDT et MLP combinées ont été appliquées une fois unique à l'échantillon des sources non-identifiées selon leur latitude galactique. Les candidats blazars obtenus parmi les sources non-identifiées du catalogue 2FGL sont présentés dans la Table 10.2 et la Table 10.3, selon qu'ils possèdent ou ne possèdent pas de mises en gardes quant à la qualité de la reconstruction (« flags »). À titre d'indication, les candidats blazars obtenus dans les études de [Mirabal et al.,

2. Le taux de mauvaise classification est légèrement différent des 10 % annoncés car la statistique utilisée ne permet pas d'obtenir ce nombre « rond ». Il est de $\varepsilon_{\mathrm{Bkg}} \approx 9{,}3\,\%$ pour l'analyse « extragalactique » et de $\varepsilon_{\mathrm{Bkg}} \approx 9{,}6\,\%$ pour l'analyse « galactique ».

		BDT	MLP	Combinée
Analyse galactique	$\varepsilon_{\mathrm{Sig}}$	$(92{,}9 \pm 2{,}1)\,\%$	$(92{,}9 \pm 2{,}6)\,\%$	$(90{,}9 \pm 2{,}5)\,\%$
	$\varepsilon_{\mathrm{Bkg}}$	$\sim 10\,\%$	$\sim 10\,\%$	$(7{,}1 \pm 1{,}0)\,\%$
	$\zeta^{\star}$	0,1612	0,4633	$\zeta^{\star}_{\mathrm{BDT}}$ et $\zeta^{\star}_{\mathrm{MLP}}$
Analyse extragalactique	$\varepsilon_{\mathrm{Sig}}$	$(95{,}3 \pm 2{,}5)\,\%$	$(95{,}3 \pm 2{,}6)\,\%$	$(93{,}4 \pm 2{,}9)\,\%$
	$\varepsilon_{\mathrm{Bkg}}$	$\sim 10\,\%$	$\sim 10\,\%$	$(7{,}5 \pm 1{,}3)\,\%$
	$\zeta^{\star}$	0,4246	0,4355	$\zeta^{\star}_{\mathrm{BDT}}$ et $\zeta^{\star}_{\mathrm{MLP}}$

TABLE 10.1: Efficacités de reconstruction des différentes analyses.

2012] et [Doert et Errando, 2013] sont respectivement signalés avec un exposant « † » et « * ». L'ensemble des candidats blazars de notre étude sont représentés sur la carte de la Figure 10.9.

Un total de 248 candidats blazars sont obtenus dont 179 se situent à une latitude galactique $|b| > 10°$, et dont 69 possèdent une latitude $|b| \leq 10°$. Parmi les candidats blazars de haute latitude 22,9 % sont signalés dans le catalogue 2FGL comme ayant un possible problème de reconstruction. Aux basses latitudes, cette fraction augmente jusqu'à 56,5 %.

Dans l'étude de [Mirabal et al., 2012], un échantillon de 216 candidats blazars de haute latitude est proposé et parmi eux, 173 font partie de notre sélection (96,6 %). [Doert et Errando, 2013], qui ne séparent pas le traitement des sources selon leur latitude, obtiennent 159 candidats blazars de haute latitude dont 143 (90,0 %) et 156 (98,1 %) ont été respectivement désignés comme blazars potentiels par notre étude, et celle de [Mirabal et al., 2012].

À basse latitude galactique, [Doert et Errando, 2013] proposent un total de 72 candidats blazars dont 49 (68,5 %) font partie de notre sélection. Cette différence provient de la classification des sources possédant un « flags » de reconstruction, pour lesquelles [Doert et Errando, 2013] proposent un total de 48 candidats dont seulement 27 (56,3 %) sont identifiés comme tels par notre analyse, alors que pour les sources proposées qui ne sont pas signalées comme pouvant présenter des problèmes de reconstruction, 91,7 % de nos candidats sont communs à leur sélection. Un total de 20 candidats blazars de basse latitude sont proposés exclusivement par notre étude.

D'après [Ackermann et al., 2011], près de 200 AGN sont attendus à basse latitude galactique. Les candidats proposés de manière indépendante dans cette étude et celle de [Doert et Errando, 2013] pourraient contribuer à une fraction non négligeable du déficit de blazars observé. Toutefois, il est indispensable de combiner des observations multi-longueur d'onde afin de déterminer la nature exacte des sources non-identifiées du catalogue 2FGL. Au prochain chapitre nous présentons une méthode permettant d'identifier les contreparties infrarouges de type blazar des sources non-identifiées du catalogue 2FGL afin, entre autres, d'obtenir des coordonnées de précision astrométrique pour faciliter le suivi multi-longueur d'onde.

Notons que dans ce travail nous nous sommes exclusivement intéressés à obtenir les candidats blazars parmi les sources non-identifiées mais pas à leur nature (typiquement FSRQ ou BL Lac). Cette identification est posssible, comme suggéré dans [Hassan et al., 2013] qui se sont intéressés aux blazars de type indéterminé du catalogue 2FGL, en utilisation les différences spectrales des deux populations. En effet, les FSRQ ont tendance à avoir un indice spectral plus mou que les BL Lac (voir section 2.6.1 et notamment la Figure 2.11a), et leur maximum du pic d'émission Compton-inverse se situe à plus basse énergie que pour les BL Lac, ce qui peut se traduire au premier ordre par une énergie pivot plus importante pour les BL Lac que pour les FSRQ. C'est une piste qui pourra être explorée dans le futur.

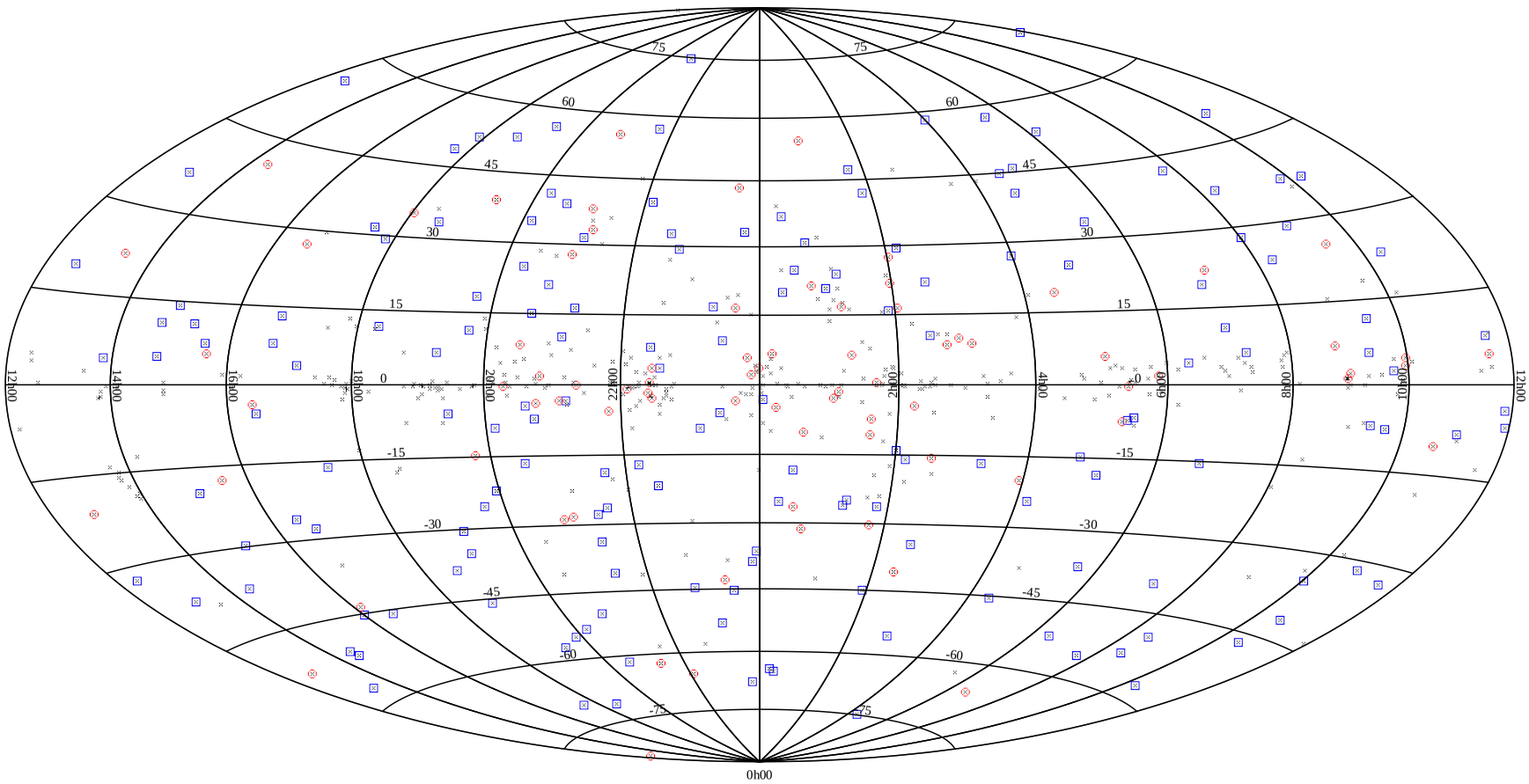


FIGURE 10.9: Projection d'Aïtoff des coordonnées galactiques des candidats blazars. Les points noirs représentent l'échantillon des sources non-identifiées, les carrés bleus et les cercles rouges représentent respectivement les candidats blazars proposés dans notre étude sans et avec *flag* de reconstruction.

TABLE 10.2: Liste de candidats blazars parmi les sources non-identifiées ne présentant pas de problèmes de reconstruction. Les exposants † et * représentent respectivement les candidats blazars de [Mirabal et al., 2012] et de [Doert et Errando, 2013].

Nom 2FGL	l (°)	b (°)	Valeur BDT	Valeur MLP
Candidats blazars de haute latitude				
2FGL J0004.2+2208†*	108.7321	-39.4298	0.6642	0.7203
2FGL J0031.0+0724†*	114.0951	-55.1084	0.7570	0.8257
2FGL J0032.7-5521†*	308.5234	-61.5683	0.9197	0.8750
2FGL J0048.8-6347†*	303.4062	-53.3304	0.5761	0.5032
2FGL J0102.2+0943†*	127.3523	-53.0533	0.7887	1.0347
2FGL J0103.8+1324†*	127.5563	-49.3598	0.7966	0.7976
2FGL J0116.6-6153†*	297.7491	-54.9862	0.8395	0.9991
2FGL J0133.4-4408†*	279.2349	-71.0022	0.7974	0.9723
2FGL J0143.6-5844†*	290.4679	-57.1018	0.6887	1.0391
2FGL J0158.4+0107†*	155.3271	-57.4715	0.8230	0.9460
2FGL J0158.6+8558†*	124.2008	23.2625	0.5801	0.6052
2FGL J0200.4-4105†*	261.9103	-70.0983	0.8400	0.9358
2FGL J0226.1+0943†*	158.1029	-46.5823	0.9197	0.9742
2FGL J0227.7+2249†*	150.2321	-34.8878	0.8251	1.0320
2FGL J0239.5+1324†*	159.2411	-41.7125	0.8630	0.9569
2FGL J0305.0-1602†*	200.1507	-57.1463	0.7104	0.6980
2FGL J0312.8+2013†*	162.5069	-31.5688	0.8425	0.9896
2FGL J0316.1-6434†*	281.4998	-46.0811	0.7824	1.0516
2FGL J0338.2+1306†*	173.4705	-32.9289	0.7375	0.7165
2FGL J0340.7-2421†	218.5077	-51.9486	0.7600	0.9545
2FGL J0345.2-2356†*	218.2661	-50.8498	0.7647	0.8786
2FGL J0409.5+0509†*	186.6135	-32.2991	0.6715	0.8080
2FGL J0409.8-0357†*	195.8670	-37.3670	0.8205	1.1171
2FGL J0416.0-4355†*	249.1967	-45.9214	0.7138	0.7338
2FGL J0418.9+6636	141.5288	11.5586	0.6083	0.6011
2FGL J0420.9-3743†*	240.3007	-45.1367	0.8540	1.0110
2FGL J0427.2-6705†	279.1883	-38.5948	0.7275	0.8290
2FGL J0438.0-7331†*	286.0880	-35.1684	0.7947	0.9615
2FGL J0439.8-1858†*	216.9204	-37.2483	0.8536	1.0458
2FGL J0523.3-2530†	228.2277	-29.8440	0.6249	0.6043
2FGL J0540.1-7554†*	287.3096	-30.5674	0.6326	0.5750
2FGL J0555.9-4348†*	250.4051	-28.0177	0.8245	0.9623
2FGL J0600.8-1949	225.7017	-19.7512	0.6167	0.6035
2FGL J0602.7-4011†*	246.7764	-25.9564	0.8823	0.8725
2FGL J0644.6+6034†*	155.0673	22.5985	0.7583	1.0546
2FGL J0719.2-5000†	261.2891	-16.2730	0.7171	0.8106
2FGL J0723.9+2901†*	189.4818	19.3557	0.9145	1.0527
2FGL J0737.5-8246†*	295.0864	-25.4673	0.4963	0.5184
2FGL J0745.5+7910†	134.9955	29.1433	0.5663	0.5546
2FGL J0746.0-0222†*	221.3923	11.0232	0.7506	0.6978
2FGL J0802.6-0940†*	229.9299	11.0521	0.8015	0.9921
2FGL J0803.2-0339†*	224.6500	14.1571	0.8600	1.0699
2FGL J0844.9+6214†*	153.7525	36.9023	0.7975	0.9493

... la table continue sur la page suivante...

... suite de la page précédente...

Nom 2FGL	l (°)	b (°)	Valeur BDT	Valeur MLP
2FGL J0846.0+2820†⋆	196.3363	36.3046	0.8425	1.0481
2FGL J0859.4-2532†⋆	251.1170	13.2460	0.8373	0.8649
2FGL J0900.9+6736†⋆	146.7126	37.0986	0.8057	0.9619
2FGL J1013.6+3434†⋆	190.5801	55.5448	0.6783	0.6139
2FGL J1016.1+5600†⋆	155.9615	49.9704	0.8104	0.9206
2FGL J1016.4-4244†⋆	274.9602	11.5552	0.6776	0.8183
2FGL J1029.5-2022†⋆	263.3426	31.3770	0.8425	1.0232
2FGL J1032.9-8401†⋆	299.2670	-22.1611	0.7295	0.8394
2FGL J1038.2-2423†	267.9450	29.2901	0.7725	0.8210
2FGL J1115.0-0701†⋆	265.1089	48.6365	0.7856	0.8968
2FGL J1123.3-2527†	279.0435	33.2875	0.4743	0.4571
2FGL J1129.0-0532†⋆	268.5090	51.7418	0.7849	1.0063
2FGL J1129.5+3758†⋆	175.5391	69.6858	0.8799	1.0576
2FGL J1156.7-0751†	280.2592	52.5963	0.8245	0.9536
2FGL J1207.3-5055†	295.8735	11.3355	0.4451	0.4872
2FGL J1214.1-4410†	295.8997	18.1940	0.4360	0.5157
2FGL J1221.4-0633†⋆	289.6797	55.5446	0.8674	1.0096
2FGL J1223.3+7954†⋆	124.4696	37.1341	0.7529	0.8990
2FGL J1249.5-2811†⋆	302.4171	34.6863	0.8704	0.9945
2FGL J1254.2-2203†⋆	303.8040	40.8149	0.7548	0.9535
2FGL J1259.8-3749†⋆	304.7698	25.0105	0.8400	0.9447
2FGL J1305.0+1152†	315.4513	74.4091	0.4476	0.4461
2FGL J1312.9-2351†⋆	309.2479	38.7479	0.8704	0.9916
2FGL J1328.5-4728†⋆	309.4160	14.9322	0.7716	1.0370
2FGL J1335.3-4058†⋆	311.8031	21.1335	0.8852	1.0470
2FGL J1340.5-0412†⋆	325.5120	56.4968	0.8363	1.0467
2FGL J1347.0-2956†⋆	317.0465	31.3978	0.8630	1.0512
2FGL J1410.4+7411†⋆	115.8394	41.8251	0.7863	0.7569
2FGL J1417.5-4404†⋆	318.8502	16.1163	0.8003	0.7889
2FGL J1417.7-5028†⋆	316.6814	10.0837	0.7073	0.5830
2FGL J1423.9-7842†⋆	307.5334	-16.6787	0.6795	1.0183
2FGL J1424.2-1752†⋆	332.0803	39.6833	0.7613	1.0082
2FGL J1458.5-2121†⋆	338.5325	32.5786	0.6483	0.6280
2FGL J1502.1+5548†⋆	92.7043	52.9506	0.9145	1.0788
2FGL J1502.4+4804†⋆	81.2493	57.0711	0.8565	1.0487
2FGL J1513.9-2256†⋆	340.8583	29.2376	0.7845	0.9904
2FGL J1517.2+3645†⋆	59.5796	57.8180	0.7126	0.8354
2FGL J1544.5-1126†	356.1697	32.9943	0.6713	0.7667
2FGL J1548.3+1453†⋆	25.5855	47.1833	0.8799	1.0505
2FGL J1553.5-0324†⋆	5.3849	36.5299	0.8461	0.9902
2FGL J1612.0+1403†	27.7744	41.6115	0.4896	0.4762
2FGL J1614.8+4703†⋆	73.6862	45.7729	0.7158	0.5994
2FGL J1617.5-2657†	349.6935	16.6634	0.6437	0.7288
2FGL J1622.8-0314†⋆	10.7075	30.7081	0.4417	0.5800
2FGL J1623.2+4328†	68.4149	44.6540	0.6685	0.7599
2FGL J1636.6-0841†⋆	7.8667	24.7854	0.5159	0.6452
2FGL J1646.7-1333†⋆	5.1597	19.9420	0.8425	1.0404

... la table continue sur la page suivante...

... suite de la page précédente...

Nom 2FGL	l (°)	b (°)	Valeur BDT	Valeur MLP
2FGL J1647.0+4351†⋆	68.8245	40.3399	0.6618	0.8280
2FGL J1659.2-0142†⋆	17.6614	23.8590	0.7507	0.8794
2FGL J1704.3+1235†⋆	32.5125	29.3907	0.7741	0.9904
2FGL J1704.6-0529†⋆	14.9195	20.7527	0.5922	0.5383
2FGL J1730.8+5427†⋆	82.1707	33.3065	0.7683	0.9513
2FGL J1738.9+8716†⋆	119.9981	27.9462	0.9145	0.9004
2FGL J1741.0+1347†⋆	37.7637	21.7269	0.7144	0.7306
2FGL J1741.1-6750†	325.2361	-18.7240	0.7627	0.8790
2FGL J1747.6+0324†⋆	28.6337	15.7236	0.7163	0.7158
2FGL J1748.8+3418†⋆	59.6266	27.0840	0.8551	0.9484
2FGL J1757.5-6028†⋆	333.0312	-17.1277	0.4983	0.5018
2FGL J1816.5+4511	72.8544	24.7412	0.5161	0.4375
2FGL J1821.8+0830†	37.2240	10.4175	0.5788	0.5387
2FGL J1842.3-5839†⋆	336.9656	-21.7841	0.7600	0.9363
2FGL J1902.7-7053†⋆	324.3424	-26.4372	0.7973	0.7322
2FGL J1917.0-3027†⋆	7.5063	-18.4683	0.5992	0.5470
2FGL J1919.5-7324†⋆	321.6372	-27.9095	0.8425	0.9924
2FGL J1936.5-0855†	29.9749	-14.0551	0.4286	0.4371
2FGL J1942.9-3528†⋆	4.2758	-25.2260	0.8814	1.0489
2FGL J1947.8-0739†	32.4260	-16.0143	0.5082	0.4820
2FGL J2002.8-2150†⋆	20.0940	-25.0235	0.7095	0.7271
2FGL J2004.6+7004†⋆	102.8615	19.4740	0.6825	0.8534
2FGL J2006.5-2256†⋆	19.2747	-26.2081	0.8251	1.0512
2FGL J2017.5-1618†	27.2219	-26.2065	0.6344	0.6311
2FGL J2021.5+0632†	49.6142	-16.6707	0.6073	0.5797
2FGL J2034.7-4201†⋆	359.0881	-36.3920	0.8645	0.9386
2FGL J2042.8-7317†⋆	320.7349	-33.7966	0.6669	0.7072
2FGL J2046.2-4259†⋆	358.0896	-38.5800	0.6873	0.8347
2FGL J2103.5-1112†⋆	37.8635	-34.3708	0.6062	0.5751
2FGL J2107.6+2506†	71.9899	-14.9377	0.6190	0.5738
2FGL J2115.4+1213†⋆	62.5960	-24.5037	0.6034	0.6924
2FGL J2125.0-4632†⋆	352.9940	-45.3380	0.9080	1.0637
2FGL J2131.0-5417†⋆	341.9804	-44.7125	0.9197	1.0977
2FGL J2132.5+2605†⋆	76.6906	-18.4386	0.7548	1.0269
2FGL J2133.9+6645†⋆	105.1655	10.9634	0.7323	0.7899
2FGL J2134.6-2130†⋆	28.9523	-45.0824	0.8163	1.0539
2FGL J2200.1-6931†⋆	321.2820	-40.9605	0.8257	1.1050
2FGL J2213.7-4754†⋆	348.0122	-53.0535	0.5358	0.4632
2FGL J2227.8+0051†	66.0731	-45.5626	0.4313	0.4396
2FGL J2228.6-1633	43.2155	-55.3634	0.8111	1.5222
2FGL J2246.3+1549†	83.9724	-37.4080	0.8081	0.9870
2FGL J2251.8+4211†⋆	100.4034	-15.3949	0.5940	0.4494
2FGL J2257.9-3646†⋆	3.8991	-64.1856	0.7381	0.7009
2FGL J2300.0-3553†⋆	5.6487	-64.7983	0.7044	0.8042
2FGL J2319.3-3830†⋆	356.5112	-67.6441	0.7167	0.7513
2FGL J2347.2+0707†⋆	96.2152	-52.3854	0.7959	0.7825
2FGL J2354.2-6615†⋆	311.8028	-49.8693	0.5082	0.4829

... la table continue sur la page suivante...

... suite de la page précédente...

Nom 2FGL	l (°)	b (°)	Valeur BDT	Valeur MLP
2FGL J2358.4-1811†*	66.3697	-74.8794	0.6252	0.6625
Candidats blazars de basse latitude				
2FGL J0248.5+5131*	140.7906	-7.2427	0.3693	0.6266
2FGL J0307.4+4915*	144.5498	-7.8247	0.2951	0.5639
2FGL J0332.1+6309*	139.9551	5.7256	0.4546	0.7285
2FGL J0353.2+5653	145.8674	2.3468	0.4877	0.7856
2FGL J0431.5+3622	164.7155	-8.0591	0.5419	0.6634
2FGL J0516.7+2634*	178.4937	-6.6314	0.4933	0.7365
2FGL J0524.1+2843*	177.6363	-4.0778	0.3816	0.6489
2FGL J0600.9+3839*	173.1898	7.6409	0.4877	0.8127
2FGL J0658.4+0633*	207.8824	4.5373	0.4021	0.4693
2FGL J0725.8-0549*	222.0594	4.9349	0.5523	0.9003
2FGL J0737.1-3235	246.8587	-5.5622	0.3120	0.6344
2FGL J0758.8-1448*	233.9430	7.6475	0.2742	0.4802
2FGL J0838.8-2828	250.5994	7.8157	0.1781	0.5236
2FGL J0841.3-3556*	256.8943	3.7197	0.4148	0.8915
2FGL J1058.7-6621*	291.9610	-5.9257	0.5016	0.9343
2FGL J1117.2-5341*	289.1588	6.6378	0.3841	0.4675
2FGL J1240.6-7151	302.0828	-9.0121	0.3154	0.6224
2FGL J1353.5-6640*	309.0328	-4.5478	0.4107	0.7512
2FGL J1422.3-6841*	311.0643	-7.2716	0.3924	0.8171
2FGL J1507.0-6223*	317.9355	-3.5423	0.5024	0.9102
2FGL J1601.1-4220*	336.2693	7.9352	0.3536	0.7282
2FGL J1626.4-4408	338.3651	3.4437	0.1677	0.4856
2FGL J1649.2-3004*	351.9892	9.3849	0.2673	0.7586
2FGL J1748.9-3923*	351.3839	-5.9786	0.5480	0.9269
2FGL J1753.8-4446	347.1190	-9.4458	0.3144	0.5728
2FGL J1759.4-2954*	0.6888	-3.0870	0.3536	0.6279
2FGL J2107.8+3652*	81.0513	-7.2082	0.3692	0.7565
2FGL J2110.3+3822*	82.4953	-6.5747	0.4920	0.8190
2FGL J2114.1+5440*	94.8386	4.1211	0.5185	0.6033
2FGL J2231.0+6512	108.8958	6.2302	0.3967	0.5172

TABLE 10.3: Liste de candidats blazars parmi les sources non-identifiées présentant des problèmes de reconstruction. Les exposants † et * représentent respectivement les candidats blazars de [Mirabal et al., 2012] et [Doert et Errando, 2013].

Nom 2FGL	l (°)	b (°)	Flags	Valeur BDT	Valeur MLP
Candidats blazars de haute latitude					
2FGL J0014.3-0509†*	99.3738	-66.3123	16	0.6167	0.6239
2FGL J0124.6-2322†	188.1351	-81.6110	2	0.6939	0.8295
2FGL J0312.5-0914†*	191.4521	-52.2199	2	0.7659	0.6673
2FGL J0404.0+3843†*	159.1360	-10.2482	8	0.7440	1.0439
2FGL J0428.0-3845†*	241.8369	-43.7789	18	0.8698	1.1229
2FGL J0458.4+0654†*	192.7578	-21.3142	8	0.9197	1.1824
2FGL J0620.8-2556†*	233.5211	-17.8056	2	0.6990	0.8376
2FGL J0641.2+7315*	141.4806	25.1687	2	0.9145	0.9650
2FGL J0753.2+1937†*	201.4517	22.1237	1024	0.9016	1.1001
2FGL J0923.5+1508†*	215.8530	40.4225	1024	0.9197	0.8837

... la table continue sur la page suivante...

... suite de la page précédente...

Nom 2FGL	l (°)	b (°)	Flags	Valeur BDT	Valeur MLP
2FGL J0937.9-1434†*	248.4889	27.2129	128	0.5714	0.6711
2FGL J1059.9-2051†*	270.7325	34.9411	8	0.8425	1.0392
2FGL J1105.4-7622†	296.6910	-14.8003	8	0.7791	0.8274
2FGL J1208.6-2257†*	290.2458	38.8587	256	0.7906	0.8983
2FGL J1315.6-0730†*	313.4041	54.8687	1024	0.9145	1.1041
2FGL J1339.2-2348†*	316.7930	37.7690	8	0.7887	1.0201
2FGL J1345.8-3356†*	315.6480	27.5840	1	0.7507	0.8888
2FGL J1351.1-2749†*	318.7302	33.2001	8	0.6752	0.9013
2FGL J1506.9+1052†*	12.5061	54.2448	16	0.7112	0.8088
2FGL J1511.8-0513†*	354.6205	43.1208	16	0.8227	0.8438
2FGL J1632.6-2328c†	354.7687	16.4484	33	0.5432	0.5601
2FGL J1656.4-0738†*	11.8108	21.3081	8	0.8496	1.0196
2FGL J1708.4+1003c†*	30.3684	27.4160	32	0.7578	0.6556
2FGL J1724.9-0508c†*	17.9894	16.5969	56	0.8667	0.9511
2FGL J1727.6+0647†*	29.4432	21.7033	136	0.7448	0.8162
2FGL J1749.1+0515*	30.5323	16.2378	24	0.5560	0.4700
2FGL J1837.9+3821†*	67.0889	18.9110	8	0.6788	0.7416
2FGL J1846.6-2519†	9.6274	-10.2577	8	0.7450	0.6564
2FGL J1913.8-1237†	24.0804	-10.6218	8	0.5823	0.5464
2FGL J1925.7-7836c†*	315.7484	-28.2464	33	0.8852	1.0509
2FGL J1942.7-8049c†	313.1553	-28.7933	34	0.7030	0.5542
2FGL J1944.3+7325†	105.4684	22.2658	16	0.6343	0.7492
2FGL J1952.6-3252†*	7.6892	-26.4029	384	0.5823	0.4465
2FGL J1955.9-0241†	37.9686	-15.5668	1	0.4516	0.4580
2FGL J2017.4-3215†*	9.9396	-31.2738	1	0.7465	0.7912
2FGL J2031.4-1842†*	26.1034	-30.1720	1416	0.7882	1.0369
2FGL J2053.2+1212c†*	59.1143	-20.1544	32	0.8246	1.0492
2FGL J2112.3-4832†*	350.5898	-42.9712	1	0.8342	1.0263
2FGL J2124.0-1513†*	35.8134	-40.5694	1024	0.9197	0.9754
2FGL J2343.3-4752†*	331.0491	-65.3720	2	0.8163	1.0465
2FGL J2356.0-5256†*	320.9426	-62.2108	2	0.8360	1.1017
Candidats blazars de basse latitude					
2FGL J0222.7+6820*	131.2290	6.9570	1	0.4434	0.7890
2FGL J0225.9+6154c	133.8176	1.0464	32	0.4306	0.8008
2FGL J0233.9+6238c*	134.4024	2.0675	288	0.2347	0.8447
2FGL J0415.2+5518*	149.2098	3.1901	1	0.5295	0.8512
2FGL J0423.4+5612	149.3997	4.6488	256	0.5328	0.9681
2FGL J0547.4+3722	172.9691	4.7054	8	0.3719	0.7547
2FGL J0742.7-3113	246.2396	-3.8719	8	0.1979	0.6240
2FGL J0753.2-1634	234.7805	5.5821	256	0.5363	1.0469
2FGL J1303.7-6316c*	304.3188	-0.4298	32	0.5035	0.8259
2FGL J1324.4-5411*	307.8065	8.3625	24	0.5980	1.0222
2FGL J1407.6-5937c*	312.4389	1.8301	40	0.4265	0.8101
2FGL J1415.7-6520*	311.5798	-3.8906	8	0.4642	0.8753
2FGL J1456.7-6247c*	316.7153	-3.3205	32	0.3668	0.5135
2FGL J1510.9-5808c	320.5003	-0.1066	56	0.3487	0.4636
2FGL J1614.9-5212*	331.3270	-0.9311	2	0.3574	0.7000
2FGL J1618.0-5825*	327.3001	-5.7037	8	0.4482	1.1627
2FGL J1619.6-4509*	336.7955	3.5857	1	0.4401	0.9136
2FGL J1630.2-4752*	336.1270	0.3853	2	0.3747	0.6036
2FGL J1639.8-4921c*	336.1011	-1.7899	32	0.4725	0.9096
2FGL J1648.1-4930*	336.8797	-2.9073	24	0.5191	0.8925
2FGL J1717.3-2809*	357.2549	5.6153	10	0.3385	0.7179
2FGL J1727.8-2308*	2.7609	6.4705	8	0.3174	0.4688
2FGL J1731.9-2703c*	359.9755	3.5486	40	0.2810	0.5504
2FGL J1733.2-2913c*	358.3116	2.1391	34	0.2250	0.4743
2FGL J1747.2-3507	354.8876	-3.5091	136	0.2504	0.5159

... la table continue sur la page suivante...

... suite de la page précédente...

Nom 2FGL	l (°)	b (°)	Flags	Valeur BDT	Valeur MLP
2FGL J1805.0-0845*	19.7957	6.1610	8	0.3976	0.5604
2FGL J1813.6-2821	3.5594	-5.0423	8	0.4207	0.8869
2FGL J1827.4-1445c	17.1017	-1.5054	32	0.3820	0.4983
2FGL J1830.4-1634*	15.8154	-2.9840	8	0.5073	0.9161
2FGL J1835.4+1036*	40.6338	8.3366	8	0.4737	0.9611
2FGL J1835.4+1349*	43.5617	9.7445	9	0.4248	0.8221
2FGL J1835.5-0649	25.0503	0.3894	2	0.3469	0.6286
2FGL J1844.3+1548*	46.3161	8.6779	4	0.4338	0.7753
2FGL J1902.3-1106*	24.2267	-7.4346	9	0.2770	0.5826
2FGL J1908.8-0132	33.5334	-4.5729	512	0.4303	0.7596
2FGL J1958.6+4020*	75.7736	5.6982	1	0.3686	0.7629
2FGL J2040.1+4105c	80.8294	-0.3938	56	0.6184	0.7445
2FGL J2055.8+4754*	87.8906	1.7284	264	0.3666	0.8294
2FGL J2104.9+3555*	79.9438	-7.4079	9	0.5117	1.0054

10.A Stabilité des indices de dureté HR_{ij}

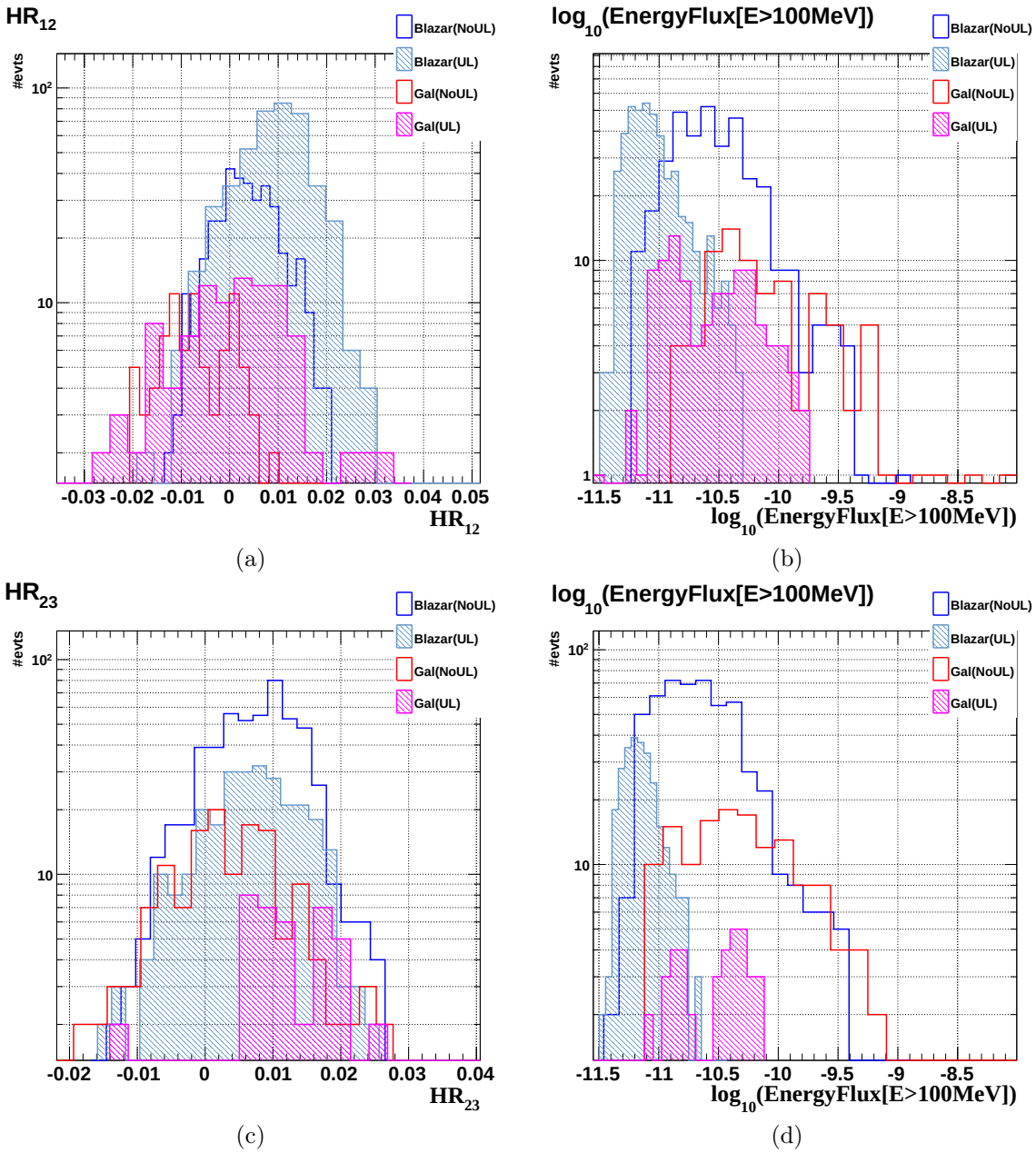


FIGURE 10.10: Stabilité des indices de duretés HR_{12} et HR_{23} selon la présence de limite supérieur dans le calcul. À gauche distributions des paramètres HR_{12} et HR_{23} (bleu=blazars sans limite supérieure, bleu hachuré=blazars avec au moins une limite supérieure, rouge=sources galactiques sans limite supérieure, magenta hachuré=sources galactiques avec au moins une limite supérieure) et à droite distributions du flux en énergie des blazars et des sources galactiques (même code couleur).

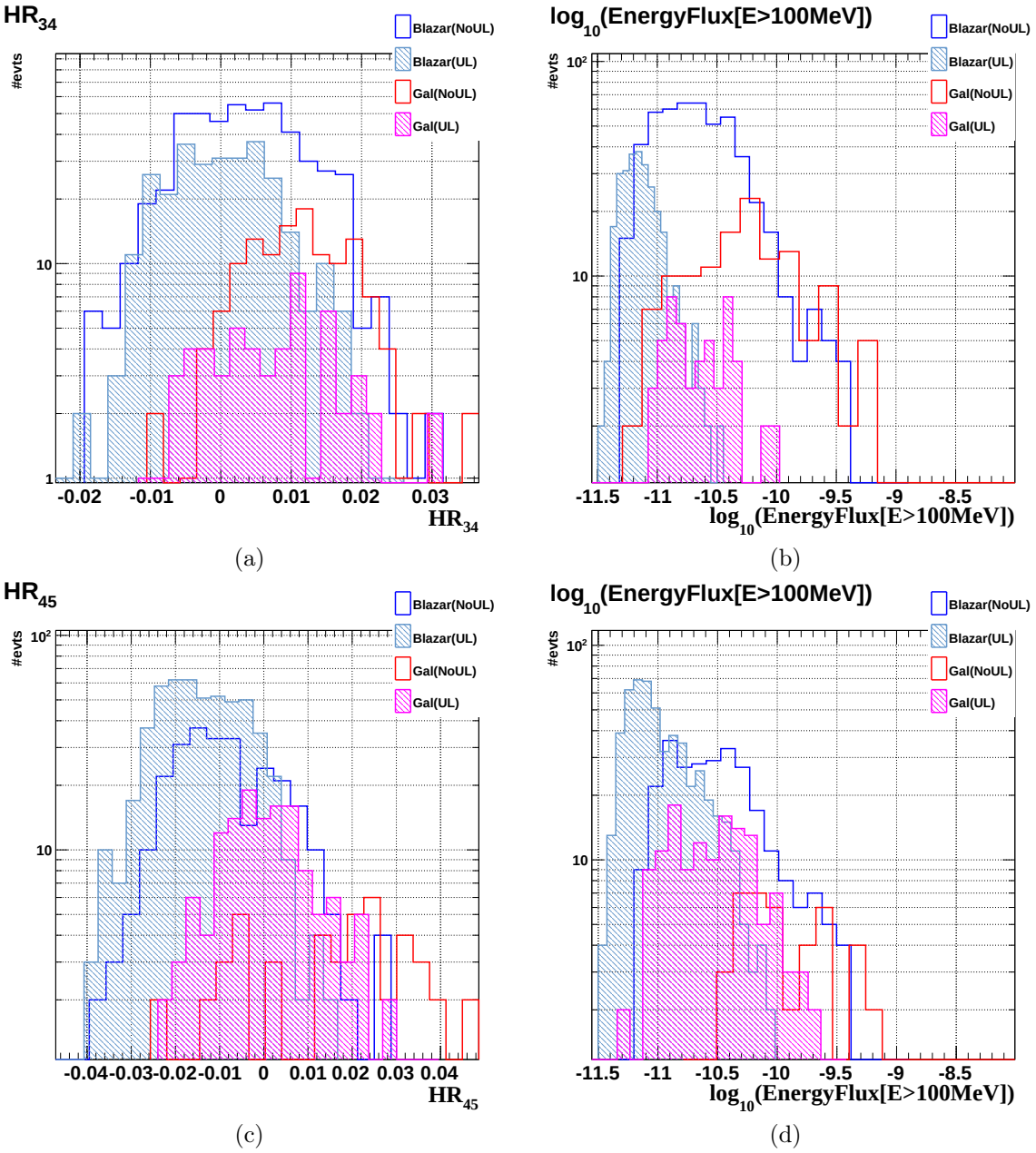


FIGURE 10.11: Stabilité des indices de duretés HR_{34} et HR_{45} selon la présence de limite supérieur dans le calcul. À gauche distributions des paramètres HR_{34} et HR_{45} (bleu=blazars sans limite supérieure, bleu hachuré=blazars avec au moins une limite supérieure, rouge=sources galactiques sans limite supérieure, magenta hachuré=sources galactiques avec au moins une limite supérieure) et à droite distributions du flux en énergie des blazars et des sources galactiques (même code couleur).

10.B Paramètres TMVA des analyses BDT et MLP

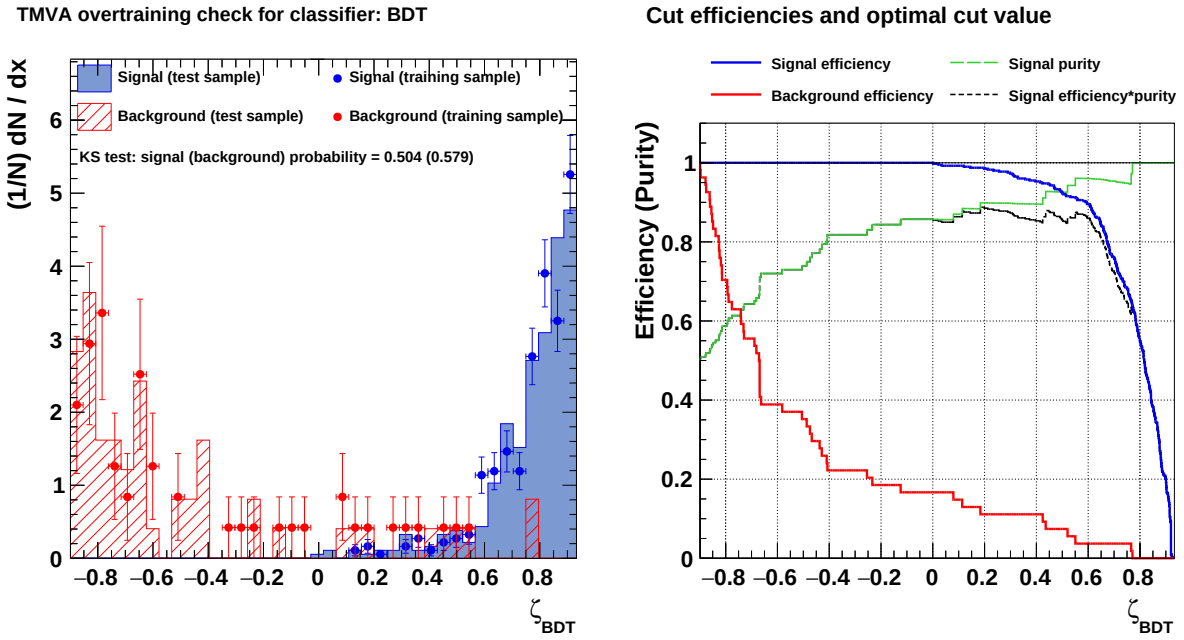
Option	Valeur	Description
BoostType	*AdaBoost*	Algorithme de *boosting*
NTrees	400	Nombre d'arbres maximum.
MaxDepth	3	Profondeur maximale des arbres.
MinNodeSize	5,0 %	Fraction minimale d'événements dans un nœud pour faire une feuille.
UseBaggedBoost	*True*	À chaque itération, un ensemble d'événements choisis aléatoirement est utilisé pour la génération de l'arbre
BaggedSampleFraction	0,6	Fraction d'événements utilisée pour le *bagging*
AdaBoostBeta	0,2	Taux d'apprentissage.
NCuts	−1	Granularité de l'espace des phases.

TABLE 10.4: Paramètres de l'analyse BDT dans TMVA.

Option	Valeur	Description
HiddenLayers	N+5	Architecture du réseau
TrainingMethod	BP	Algorithme d'apprentissage
BPMode	*sequential*	Méthode d'apprentissage
NCycles	2000	Nombre de cycle après lequel l'entraînement est arrêté
ConvergenceTests	150	Arrête l'entraînement après ce nombre d'itérations sans amélioration
LearningRate	0,02	Taux d'apprentissage
DecayRate	0,01	Taux de décroissance du taux d'apprentissage
TestRate	5	Nombre de cycles après lesquelles effectuer le test de convergence
Sampling	1	Fraction d'événements choisie aléatoirement utilisée pendant *TestRate* cycles
SamplingEpoch	1	Fraction du nombre du cycles, après laquelle tous les événements sont utilisés pour l'entraînement
UseRegulator	*False*	Utilisation de la régularisation

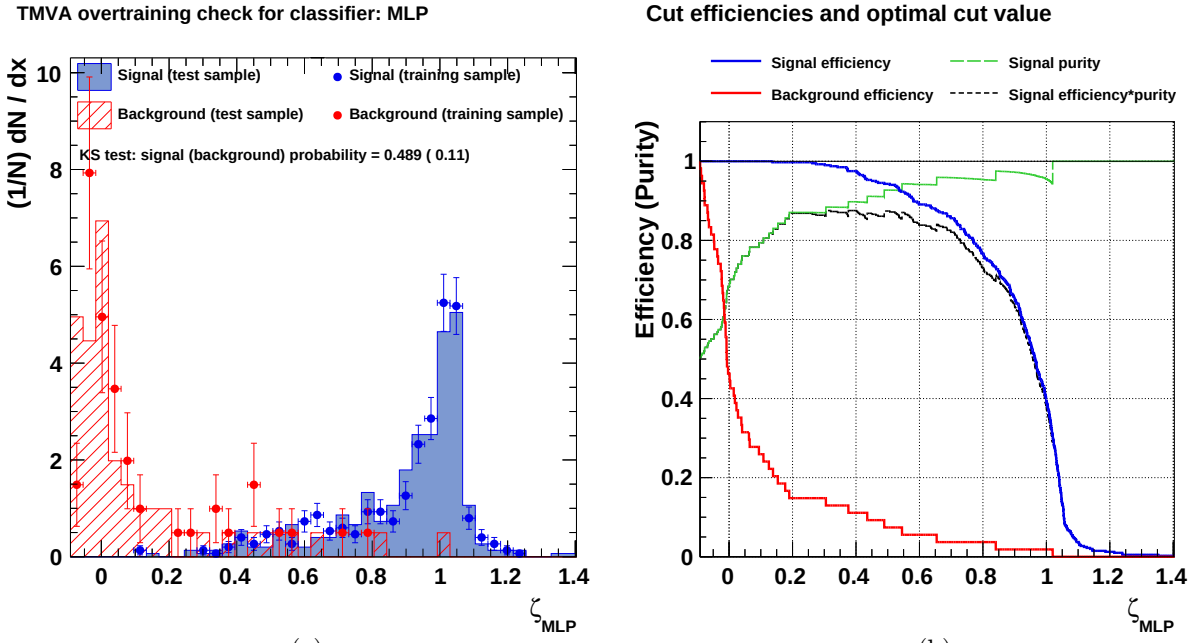
TABLE 10.5: Paramètres de l'analyse MLP dans TMVA.

10.C Performances des méthodes individuelles BDT et MLP de l'analyse de la région « extragalactique »



(a) (b)

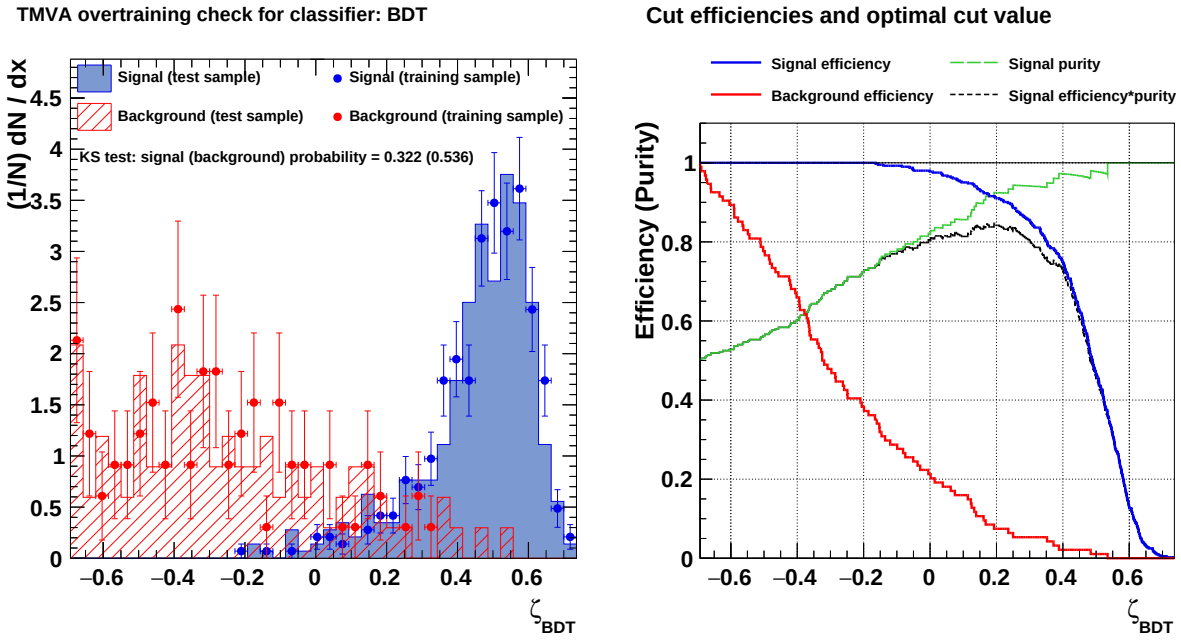
FIGURE 10.12: (a) Efficacités de reconstruction du signal (bleu) et du fond (rouge) en fonction de ζ de la méthode BDT pour l'analyse « extragalactique ». (b) Distributions du paramètre ζ des évènements signal (bleu) et des évènements de fond (rouge) pour les lots d'entraînement (traits pleins) et de test (points).



(a) (b)

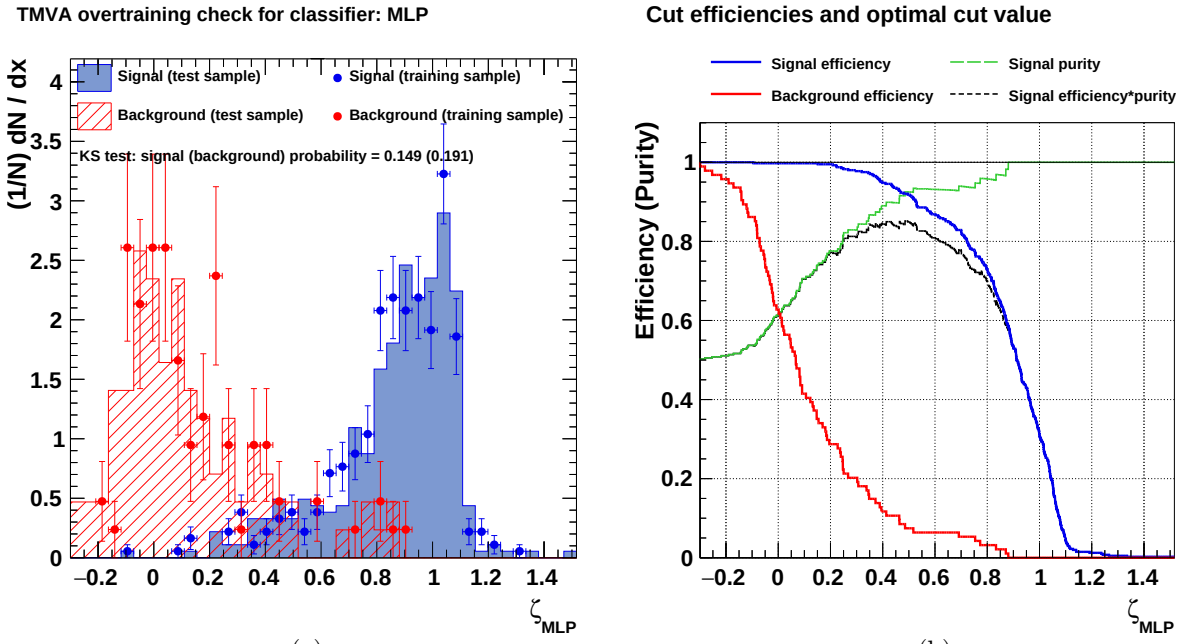
FIGURE 10.13: (a) Efficacités de reconstruction du signal (bleu) et du fond (rouge) en fonction de ζ de la méthode MLP pour l'analyse « extragalactique ». (b) Distributions du paramètre ζ des évènements signal (bleu) et des évènements de fond (rouge) pour les lots d'entraînement (traits pleins) et de test (points).

10.D Performances des méthodes individuelles BDT et MLP de l'analyse de la région « galactique »



(a) (b)

FIGURE 10.14: (a) Efficacités de reconstruction du signal (bleu) et du fond (rouge) en fonction de ζ de la méthode BDT pour l'analyse « galactique ». (b) Distributions du paramètre ζ des évènements signal (bleu) et des évènements de fond (rouge) pour les lots d'entraînement (traits pleins) et de test (points).



(a) (b)

FIGURE 10.15: (a) Efficacités de reconstruction du signal (bleu) et du fond (rouge) en fonction de ζ de la méthode MLP pour l'analyse « galactique ». (b) Distributions du paramètre ζ des évènements signal (bleu) et des évènements de fond (rouge) pour les lots d'entraînement (traits pleins) et de test (points).

Bibliographie

A. A. Abdo, et al. Fermi Large Area Telescope First Source Catalog. ApJS, 188:405–436, June 2010. doi: 10.1088/0067-0049/188/2/405.

M. Ackermann, et al. The Second Catalog of Active Galactic Nuclei Detected by the Fermi Large Area Telescope. ApJ, 743:171, December 2011. doi: 10.1088/0004-637X/743/2/171.

M. Ackermann, et al. A Statistical Approach to Recognizing Source Classes for Unassociated Sources in the First Fermi-LAT Catalog. ApJ, 753:83, July 2012. doi: 10.1088/0004-637X/753/1/83.

M. Doert et M. Errando. Search for gamma-ray-emitting active galactic nuclei in the Fermi-LAT unassociated sample using machine learning. *ArXiv e-prints*, December 2013.

E. C. Ferrara, et al. Fermi's Mystery Sources: Methods for Classification and Association. *ArXiv e-prints*, June 2012.

T. Hassan, et al. Gamma-ray active galactic nucleus type through machine-learning algorithms. MNRAS, 428:220–225, January 2013. doi: 10.1093/mnras/sts022.

N. Mirabal, et al. Fermi's sibyl: mining the gamma-ray sky for dark matter subhaloes. MNRAS, 424:L64–L68, jul 2012.

P. L. Nolan, et al. Fermi Large Area Telescope Second Source Catalog. ApJS, 199:31, April 2012. doi: 10.1088/0067-0049/199/2/31.

Chapitre 11

Recherche de candidats blazars à partir des catalogues 2FGL et *WISE*

Sommaire

11.1 Introduction **203**
11.2 Statut sur la recherche de contreparties IR des blazars **204**
11.3 Descriptif de notre approche **206**
11.4 Dépendance spatiale du fond infrarouge **207**
11.5 Comparaison des variables discriminantes **207**
11.6 Construction et performances des analyses **214**
11.7 Résultats **223**
Bibliographie **245**

11.1 Introduction

Dans le chapitre précédent, une méthode permettant d'obtenir un nombre important de candidats blazars parmi les sources non-identifiées du catalogue 2FGL a été présentée. Pour confirmer ou réfuter la nature des candidats blazars il est nécessaire pour chaque source d'obtenir des données dans des domaines spectraux différents. Malheureusement, l'erreur de localisation associée aux sources *Fermi-LAT* varie de quelques centièmes de degré à ∼0,5°, selon leur brillance, ce qui complique la recherche de contreparties à des longueurs d'onde différentes.

À l'aide des données IR de l'expérience *WISE*, dont l'erreur de localisation est inférieure à l'arc-seconde, [Massaro et al., 2011] observe que les blazars occupent un emplacement distinct des autres sources « thermiques » dans un diagramme couleur-couleur. En utilisant une méthode géométrique basée sur une analyse en composantes principales [D'Abrusco et al., 2013], [Massaro et al., 2013] montre la possibilité d'identifier de potentielles contreparties IR de type « blazar » pour les sources non-identifiées du catalogue 2FGL.

L'objectif de ce chapitre est de rechercher à l'aide d'une méthode de classification binaire les contreparties infrarouges de type blazar qui signalent la présence d'un blazar parmi les sources non-identifiées du catalogue 2FGL. Nous exploiterons les différences statistiques de paramètres construits à partir des mesures de magnitudes fournies par l'expérience *WISE* entre les contreparties IR d'un lot identifié de blazars et le fond infrarouge, que l'on définit ici comme les sources IR obtenues dans des régions ne comprenant aucun blazar des catalogues *Multifrequency catalogue of blazars* (Roma-BZCAT) [Massaro et al., 2009] et 2FGL [Nolan et al., 2012].

Ce travail se distingue de l'approche géométrique utilisée par [Massaro et al., 2013], non seulement par l'utilisation d'une méthode multivariée, mais aussi par une estimation du taux de fausses identifications dû aux sources du fond IR cité ci-dessus et l'utilisation du fait que les blazars ont un moyen très efficace de production de photons IR, ce qui les distingue de la plupart des autres sources.

Après un rappel de la méthodologie utilisée par [Massaro et al., 2013] pour la recherche de contreparties blazars IR des sources non-identifiées du *Fermi-LAT* à la section 11.2, nous présentons notre démarche à la section 11.3. La dépendance du fond infrarouge est étudiée à la section 11.4, et la

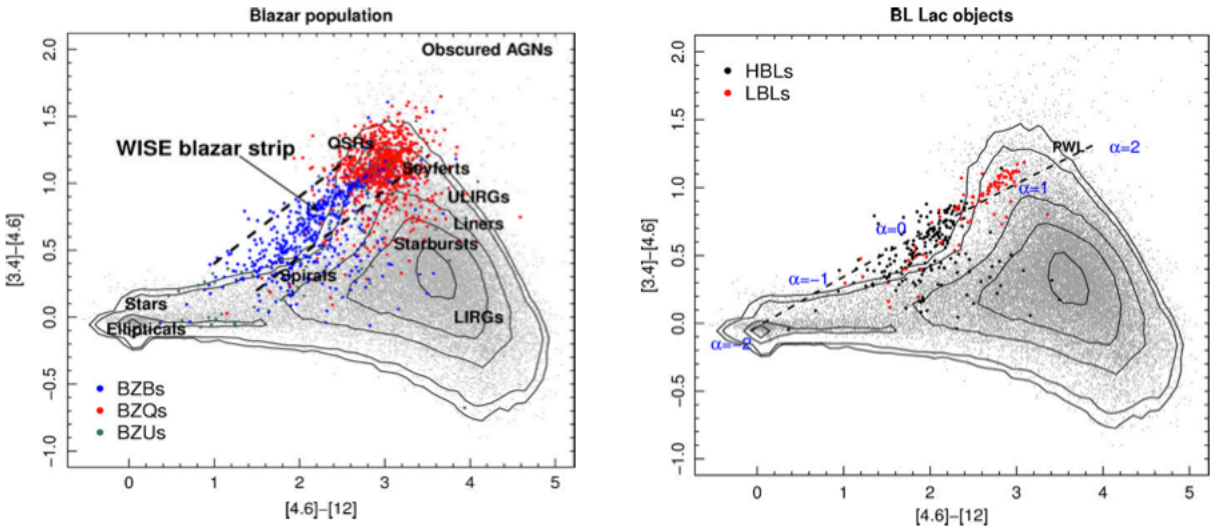


FIGURE 11.1: Diagrammes couleur-couleur $c_1(c_2)$ (Crédits : [Massaro et al., 2011]). À gauche sont représentés 1365 blazars, les BL Lac en bleu, les FSRQ en rouge et ceux de type incertain en vert. Les points gris représentent le fond IR pris dans une région de haute latitude galactique de 3° de rayon. À droite, la courbe noire représente c_1 en fonction de c_2 pour différents indices de loi de puissance (indiqués en bleu). Les HBL et les LBL sont respectivement de couleur noire et rouge.

comparaison des variables discriminantes est présentée à la section 11.5. La construction des analyses et les performances associées sont présentées à la section 11.6 et les résultats à la section 11.7. Nous discuterons des différences avec les résultats de [Massaro et al., 2013].

11.2 Statut sur la recherche de contreparties IR des blazars

Cette section présente un état des lieux des travaux déjà existants concernant la recherche, avec l'expérience *WISE*, de contreparties IR de type blazar parmi les sources non-identifiées du catalogue 2FGL. Nous commencerons par présenter l'idée principale [Massaro et al., 2011] sur laquelle se basent ces travaux, puis nous présenterons la méthode développée dans [D'Abrusco et al., 2013] qui permet d'identifier les contreparties des blazars les plus prometteuses. Enfin, nous terminerons par un bilan de leurs résultats qui ont été présentés dans [Massaro et al., 2013].

11.2.1 L'émission non-thermique des blazars vue par *WISE*

L'idée sous-jacente de la recherche de contreparties IR des blazars proposée par [Massaro et al., 2011] repose sur la particularité de ces derniers d'émettre des photons dans le domaine infrarouge via des processus non-thermiques, en l'occurrence par rayonnement synchrotron. La forme spectrale de ce rayonnement contraint les magnitudes observables dans les différents filtres de telle manière que les différentes couleurs (différences entre deux magnitudes) prennent des valeurs particulières qui les distinguent de la majorité des autres sources. Ainsi, sur la gauche de la Figure 11.1, issue de cet article, sont représentées en gris deux couleurs c_1 et c_2 pour l'ensemble des sources contenues dans une région de ciel de haute latitude de 3° de rayon. Les blazars de type BL Lac, FSRQ, ainsi que les blazars de type incertain sont respectivement de couleurs bleue, rouge et verte. Les classes des différents types de sources sont également indiqués. Les blazars proviennent du croisement entre les catalogues Roma-BZCAT et *WISE*. La méthode d'association consiste à rechercher les contreparties IR dans une région circulaire de 2,4″ autour de la position fournie par le catalogue Roma-BZCAT, en considérant une erreur de localisation pour les catalogues Roma-BZCAT et *WISE*

valant respectivement 1″ et 1,4″. On remarque qu'une grande fraction des blazars occupent une région bien définie sur ce diagramme, relativement éloignée de la majorité des autres sources IR, et délimitée par les deux lignes pointillées noires à gauche de la Figure 11.1. Cette région est appelée la *Wise γ-ray Blazars Strip* (WBS) par [Massaro et al., 2011].

Au premier ordre, les spécificités de la WBS peuvent s'expliquer en supposant que les photons IR émis par des blazars suivent une loi de puissance de la forme $\nu F_\nu \propto \nu^{-\alpha}$. On peut alors écrire la couleur c_1, qui est la différence de magnitudes entre les filtres W_1 et W_2, comme :

$$c_1 = m_1 - m_2 = +2{,}5\alpha \log\left(\frac{\lambda_2}{\lambda_1}\right) + \text{Cste} \tag{11.1}$$

où la constante est liée aux flux de la source de référence dans les filtres W_1 et W_2. En procédant de la même manière pour la couleur c_2, on peut obtenir l'emplacement des blazars dans le diagramme couleur-couleur pour différentes valeur de l'indice spectral α. À droite de la Figure 11.1 est représentée en ligne pointillée noire la couleur c_1 en fonction de c_2 pour des valeurs de l'indice $\alpha \in [-2, 2]$, ainsi qu'un lot de HBL en noir et un lot de LBL en rouge, tous deux provenant du catalogue Roma-BZCAT. On peut noter sur les deux figures que les FSRQ et les LBL ont tendance à avoir un indice plus mou que les HBL, ce qui rend compte de la position de la bosse synchrotron de ces objets.

En conclusion, [Massaro et al., 2011] a montré que l'émission non-thermique IR des blazars possède des particularités qui la distinguent des autres types de sources. C'est cette propriété qui est exploitée pour identifier les contreparties IR de type blazar des sources non-identifiées du catalogue 2FGL par [D'Abrusco et al., 2013] et [Massaro et al., 2013].

11.2.2 L'approche géométrique et ses résultats

L'approche géométrique développée dans [D'Abrusco et al., 2013] utilise un lot de blazars qui est commun aux catalogues Roma-BZCAT et 2FGL pour caractériser l'émission IR qui leur est propre. Partant du constat que les blazars sont localisés d'une manière particulière dans l'espace des couleurs c_1, c_2 et c_3, une analyse en composantes principales ($\mathrm{PC}_1, \mathrm{PC}_2, \mathrm{PC}_3$) a permis de les situer dans une région de cet espace, région appelée le *locus*. [D'Abrusco et al., 2013] ont caractérisé la distribution des blazars dans cette région, et ont classifié ceux-ci suivant différentes catégories, ce qui leur a permis par la suite d'étudier la compatibilité des contreparties IR des sources non-identifiées avec ces classifications.

La méthode nécessite un lot de blazars dont la position est connue avec précision. Le catalogue Roma-BZCAT contient 3149 blazars avec leurs positions déterminées avec une précision généralement bien meilleure que 5″. Pour associer les blazars de ce catalogue avec la ou les contreparties IR du catalogue *WISE*, la recherche s'effectue dans une région circulaire de rayon de 3,3″ autour de la position des blazars. Cette valeur de 3,3″ correspond au rayon à partir duquel le nombre de sources *WISE* associé avec des blazars du catalogue Roma-BZCAT devient plus petit que le nombre de sources associées de manière fortuite. La méthode permet d'obtenir un échantillon de 610 blazars qui appartiennent aux deux catalogues Roma-BZCAT et 2FGL. Pour chacune de ces sources, une contrepartie unique IR a été trouvée.

À l'aide des contreparties identifiées en IR pour ces blazars, trois régions cylindriques dans l'espace des composantes principales ($\mathrm{PC}_1, \mathrm{PC}_2, \mathrm{PC}_3$) sont construites à partir des couleurs c_1, c_2 et c_3. Les deux premières régions, que l'on désignera respectivement par les labels b et q, correspondent à l'emplacement privilégié des blazars de type BL Lac et FSRQ. La troisième, que l'on désignera par le label m, correspond à une population mixte pour laquelle la distinction entre les types BL Lac et FSRQ est floue. Pour rechercher les contreparties IR de type blazar dans une région du ciel, les sources IR sont représentées dans l'espace de composantes principales par un ellipsoïde d'axes principaux (σ_1,σ_2,σ_3), calculés à partir des erreurs sur les trois couleurs. Une source est donc représentée par 6 points qui correspondent aux coordonnées de $\pm\sigma_i$ (avec $i = \{1, 2, 3\}$). De manière quantitative, une note $s_{\mathrm{b,q,m}}$ (*the score*) est attribuée à chacune des contreparties selon sa

compatibilité avec chacune des régions b, q ou m. Cette quantité est définie comme $s_{\mathrm{b,q,m}} = \frac{1}{6}\omega n$ où n représente le nombre de points dans l'une des régions b, q ou m, et où ω est le rapport entre le volume de l'ellipsoïde de la source dans l'espace de composantes principales et le plus grand volume de l'ellipsoïde appartenant au *locus*. Les associations IR sont classées en trois catégories selon la valeur du paramètre $s_{\mathrm{b,q,m}}$:

— Classe A^{m} : $s_{\mathrm{b,q,m}} \geq s^{90\%}_{\mathrm{b,q,m}}$

— Classe B^{m} : $s^{60\%}_{\mathrm{b,q,m}} \leq s_{\mathrm{b,q,m}} \leq s^{90\%}_{\mathrm{b,q,m}}$

— Classe C^{m} : $s^{30\%}_{\mathrm{b,q,m}} \leq s_{\mathrm{b,q,m}} \leq s^{60\%}_{\mathrm{b,q,m}}$

où $s^{X\%}_{\mathrm{b,q,m}}$ représente le $X^{\mathrm{ième}}$ centile de la distribution du *score* des blazars appartenant au *locus* dans les régions b, q ou m. Les sources qui possèdent une note qui ne correspond à aucun intervalle ne sont pas considérées comme une contrepartie IR de type blazar.

Afin d'estimer les performances de la méthode géométrique, [D'Abrusco et al., 2013] l'appliquent aux blazars ayant servi à la création du locus en les traitant comme s'ils avaient été des sources non-identifiées, c'est-à-dire en analysant l'ensemble de leurs contreparties infrarouges situées dans leurs boîtes d'erreurs respectives (qui correspondent à un niveau de confiance sur leur localisation de 95 %). Avec l'ensemble de ces blazars identifiés, ils déterminent que pour ∼80 % d'entre eux, au moins une contrepartie IR sélectionnée par l'analyse est celle qui coïncide avec la position du catalogue Roma-BZCAT. L'approche géométrique ne permettant pas d'estimer de manière direct le fond résiduel, ce dernier est estimé indirectement en sondant l'environnement proche des blazars identifiés qui est représenté par une région annulaire autour de la zone de recherche des candidats blazars IR. Les contreparties de cette région qui sont sélectionnées et auxquelles une classe a été attribuée qui est au moins égale à celle de la meilleure association obtenue pour la source non-identifiée, permettent d'avoir une idée de la qualité de l'association. On pourra regarder la dernière colonne de la Table 8 à la page 15 de [D'Abrusco et al., 2013] pour avoir plus de détails.

[Massaro et al., 2013] appliquent la méthode décrite plus haut aux sources non-identifiées du catalogue 2FGL afin d'identifier celles qui ont des *scores* compatibles avec une nature de type blazar, ainsi qu'aux blazars déjà identifiés mais de type incertain, et ce afin de déterminer s'il s'agit de BL Lac ou de FSRQ. Les contreparties IR de type blazar sont recherchées dans une région circulaire de rayon θ_{95}, et pour chacune des contreparties IR, le paramètre $s_{\mathrm{b,q,m}}$ est calculé. Sur l'ensemble des 575 sources non-identifiées du catalogue 2FGL, leur méthode leur permet d'obtenir un total de 146 candidats de type blazar répartis parmi 110 sources γ. Pour les sources non-identifiées n'étant pas signalées comme ayant des problèmes de reconstruction ils obtiennent respectivement 2, 12 et 61 contreparties de classe A^{m}, B^{m} et C^{m}. Pour les sources non-identifiées qui possède un avertissement sur la qualité de la fiabilité de la reconstruction, les nombres de contreparties de classes A^{m}, B^{m} et C^{m} sont respectivement de 8, 20 et 43. L'estimation indirecte du fond résiduel pour chacune des sources associées, qui avait précédemment était appliquée aux blazars identifiés, n'est pas présentée par [Massaro et al., 2013].

11.3 Descriptif de notre approche

Contrairement à la méthode développée par [D'Abrusco et al., 2013], notre approche utilise, afin de rechercher les contreparties IR de type blazar des sources non-identifiées, des différences statistiques de paramètres construits à partir des mesures de magnitudes par le satellite *WISE* entre un lot de blazars identifiés et différents lots de fond IR. L'utilisation de différents lots permettra de spécialiser les analyses en fonction des caractéristiques du fond afin d'obtenir une classification binaire des sources infrarouges comme « blazars » ou « non-blazars ».

Le lot de blazars contient 558 sources appartenant à la fois aux catalogues 2FGL et Roma-BZCAT et dont la contrepartie IR est située à moins de 3,3″ de la position fournie par le catalogue Roma-BZCAT. Quant aux fond IR, il sera montré qu'il est sujet à de grandes variations selon la région du ciel observée. En effet, les sources qui le constituent sont majoritairement des étoiles et

des galaxies non-actives qui sont dans le premier cas beaucoup plus nombreuses à basse latitude galactique. Le fond sera donc estimé de manière différente suivant différentes régions du ciel. Nous utiliserons une méthode multivariée pour la construction de la classification binaire « blazar »/« non-blazar ».

À la fin de ce chapitre, nous fournirons la liste des contreparties infrarouges de type blazar associées à chacune des sources non-identifiées du catalogue 2FGL. L'utilisation de fond pour la construction de l'analyse nous permettra d'associer à chacune des contreparties une estimation du fond résiduel qui rendra compte de la robustesse de l'association.

11.4 Dépendance spatiale du fond infrarouge

Dans l'optique d'optimiser le pouvoir de séparation « blazar »/« non-blazar », une étude sur la variation des trois couleurs du fond IR en fonction de la position spatiale de la zone d'observation a été entreprise. Nous commencerons par nous intéresser à la variation spatiale du fond IR de haute latitude galactique ($|b| > 10°$) puis à celle de basse latitude. Pour cette étude, nous sélectionnerons les sources IR qui sont détectées dans au moins l'un des quatre filtres avec un rapport signal sur bruit supérieur ou égal à 5. Nous appellerons « zone extragalactique » la région correspondant à $|b| > 10°$. Celle-ci comporte principalement des galaxies non-actives mais aussi une proportion d'étoiles de notre Galaxie, qui augmente lorsque l'on s'approche du plan galactique. Pour étudier la variété de ce fond, 10 régions d'un rayon de 0,7° choisies aléatoirement de part et d'autre du plan galactique et ne comportant pas de source des catalogues 2FGL et Roma-BZCAT ont été utilisées. La Figure 11.2 montre que les distributions des trois couleurs sont relativement stables, à l'exception des régions de latitude $|b| \leq 40°$, pour lesquelles les distributions de c_1 et dans un moindre mesure de c_2, montrent un léger décalage vers les faibles valeurs. En comparant les distributions de c_1 avec les contributions galactiques de la sous-section suivante, on constate qu'une contribution galactique contamine ces régions du fait de leur proximité au plan galactique.

De la même manière, nous appellerons « zone galactique » la région correspondant à $|b| \leq 10°$. Nous étudions la variation de la distribution des trois couleurs dans cette région en fonction de la longitude galactique, en découpant le plan galactique en tranches de 20° de longitude. Pour chacune de ces tranches, les sources IR ont été sélectionnées aléatoirement parmi 30 sous-régions d'un rayon de 0,1° ne comportant pas de source détectée par le *Fermi-LAT* ni de source appartenant au catalogue Roma-BZCAT. On constate sur la Figure 11.3 que la couleur c_1 est relativement stable, alors que c_2, et dans une moindre mesure c_3, sont plus touchées par la variation de la longitude. Cela s'explique par le fait qu'en dehors du centre galactique les contributions des étoiles et du fond extragalactique s'additionnent. On peut observer sur la Figure 11.3b une composante extragalactique vers les grandes valeurs de c_2 (comparer avec la Figure 11.2b).

Cette étude met en évidence des différences marquées entre les distributions des paramètres discriminants pour différentes régions du fond. Elle nous conduit à définir pour la suite trois régions de l'espace pour lesquelles nous développerons des analyses indépendantes :

— la région « extragalactique » ($|b| > 10°$), très majoritairement composée de sources qui se trouvent en dehors de notre Galaxie

— la région « centre galactique » que l'on définira comme la région correspondant à $|b| \leq 10°$ et $l \in [-100°; 100°]$, peuplée en majorité de sources appartenant à notre Galaxie mais avec une composante visible similaire à celle observée dans la région extragalactique

— la région « anti-centre galactique » que l'on définira comme la région correspondant à $|b| \leq 10°$ et $|l| > 100°$ et qui contient une population mixte de sources de nature galactique et extragalactique.

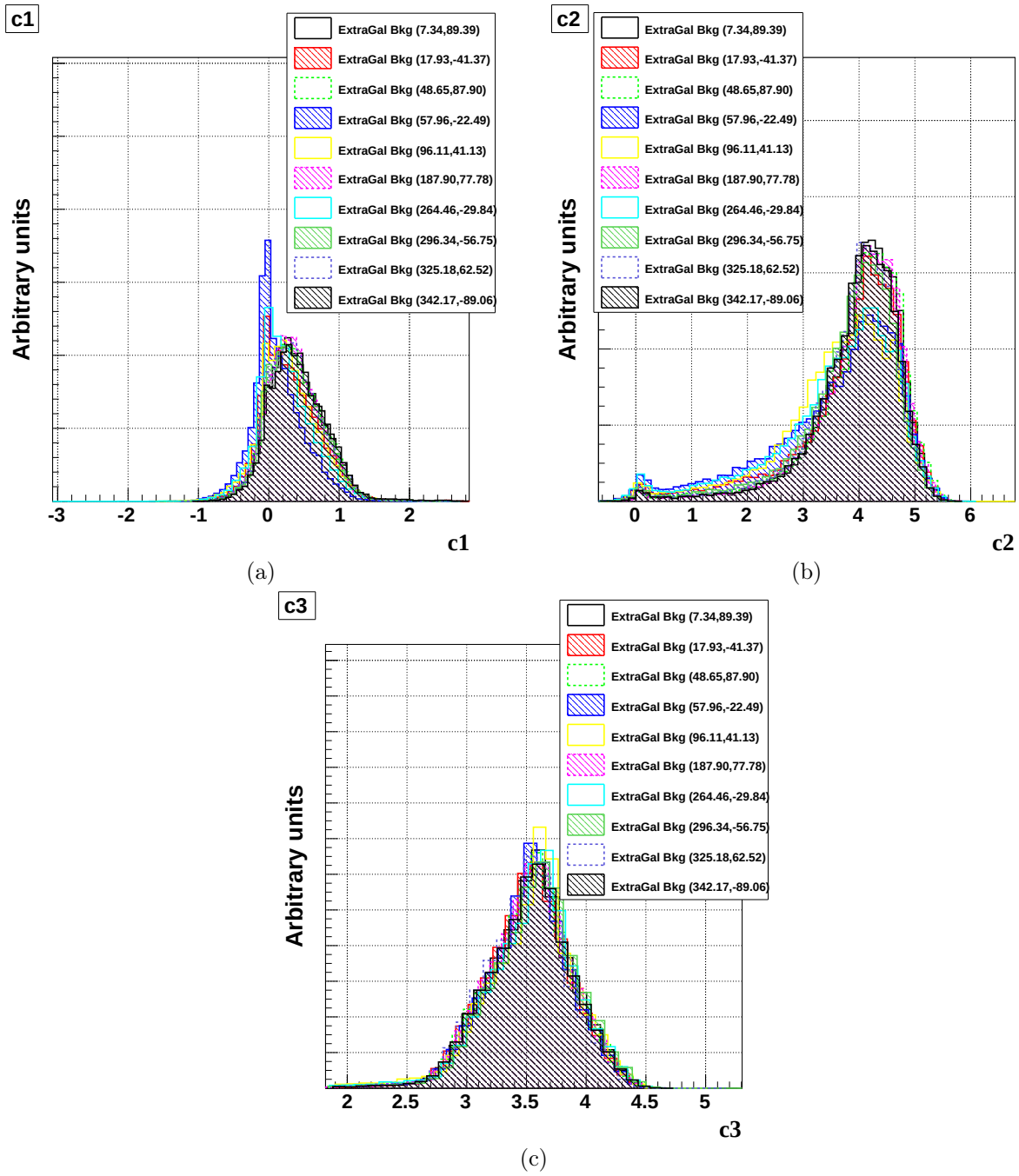


FIGURE 11.2: Variation spatiale des distributions des couleurs c_1, c_2 et c_3 pour le fond infrarouge extragalactique. Les différents histogrammes représentent les contributions infrarouges de 10 régions extra-galactiques d'un rayon de 0,7° tirées aléatoirement. Les coordonnées galactiques du centre de chaque région sont indiquées.

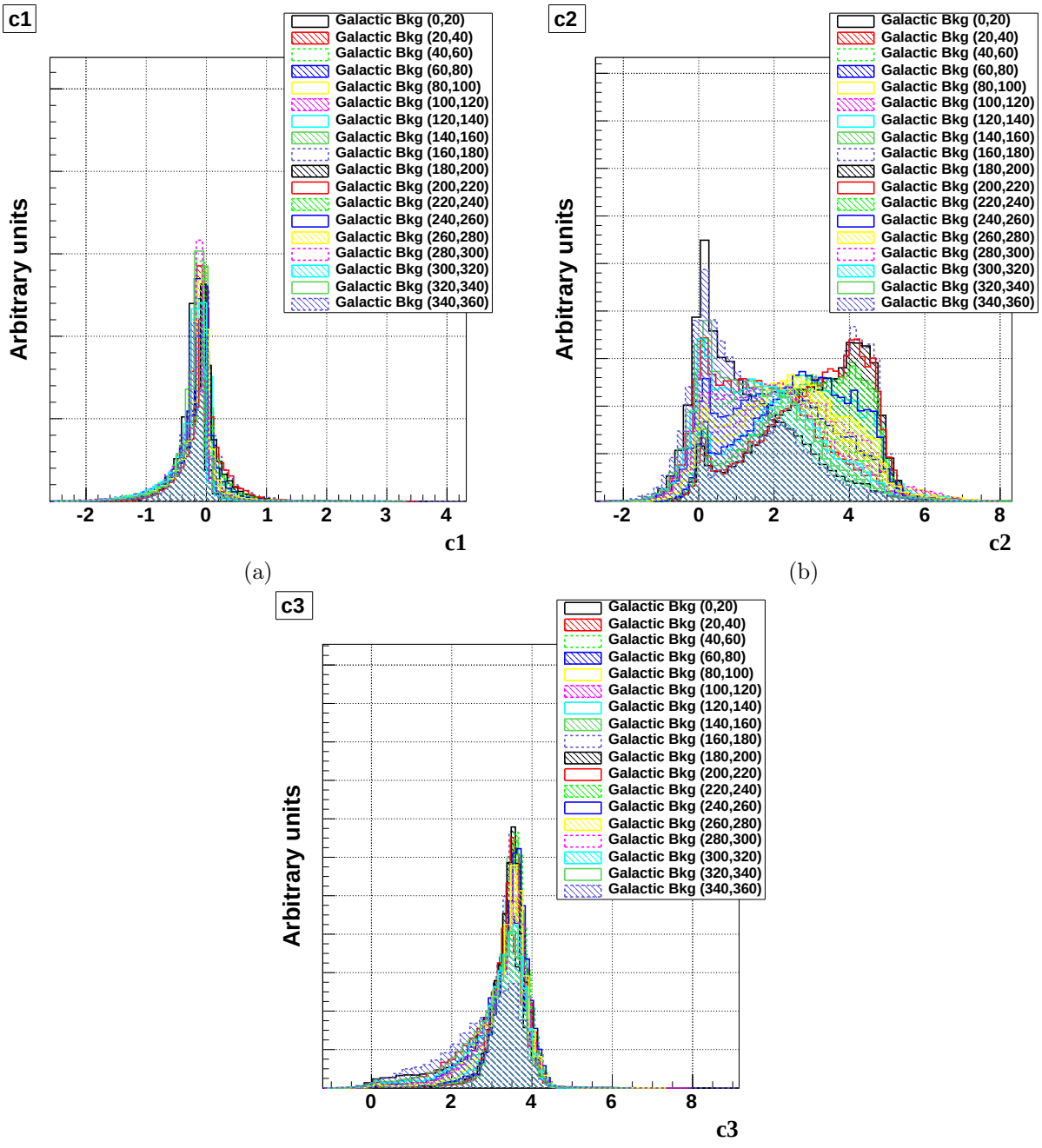


FIGURE 11.3: Variation spatiale des distributions des couleurs c_1, c_2 et c_3 pour le fond infrarouge galactique. Les différents histogrammes représentent les contributions infrarouges de 18 régions du centre galactique correspondant à autant de tranches de largeur de 20° en longitude. Les bornes en longitude de chaque tranche sont indiquées.

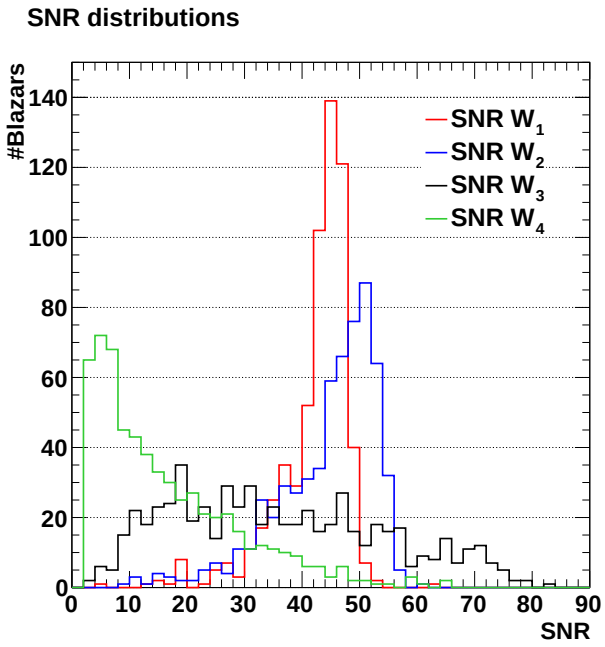


FIGURE 11.4: Distributions des rapports de signal sur bruit dans les quatre filtres de l'échantillon « signal ».

11.5 Comparaison des variables discriminantes

Nous présentons dans cette section les distributions des trois couleurs pour notre lot de blazars identifiés ainsi que pour différents types de fonds IR. Une estimation du flux intégré des blazars sera elle aussi présentée ainsi que des arguments qui appuieront son utilisation dans une analyse multivariée. Pour conclure nous discuterons rapidement d'un indice, qui appartient au catalogue *WISE*, caractérisant la variabilité des sources IR.

11.5.1 Choix des seuils de sélection

Il est nécessaire d'appliquer une sélection sur le niveau de signal sur bruit (SNR) associé à chacun des filtres. En effet, lorsque $SNR \leq 2$ dans l'un des filtres la magnitude associée n'est pas mesurée mais remplacée par une limite supérieure, ce qui pose problème pour la détermination des couleurs. C'est le cas d'une partie importante des 558 millions de sources détectées par *WISE* (voir sur la Figure 9.2 la densité typique de contreparties infrarouges possible dans la région d'une source *Fermi*-*LAT* identifiée). À contrario, comme on peut le voir sur la Figure 11.4, les contreparties IR des 610 blazars du *Fermi*-*LAT* sont très bien mesurées par les quatre filtres de *WISE* (sauf pour une faible fraction pour laquelle le filtre W_4 détecte les blazars plus près du seuil). En s'inspirant de ces distributions, on imposera pour la suite qu'une source ait un rapport signal sur bruit supérieur ou égal à 5 dans les trois premiers filtres, et supérieur ou égal à 3 dans le dernier filtre. Cela permet de se débarrasser d'une partie très importante des sources qui constituent le fond, qui n'ont pas un niveau d'émission semblable à celui des blazars *Fermi*-*LAT* et qui ne disposent pas d'une détection ferme dans chacun des quatre filtres. Cette sélection à pour effet d'éliminer 99 %, 93 % et 98 % des sources, respectivement pour les régions « extragalactique », « centre galactique » et « anti-centre galactique ». Moins de 5 % des blazars ne passent pas ce premier niveau de sélection.

Un second niveau de sélection sera appliqué aux sources IR qui concerne la fiabilité de la mesure photométrique. À cet effet, le catalogue *WISE* met à disposition pour chacun des filtres un indice, noté Ccf_W, qui peut prendre les valeurs suivantes (classées par ordre de priorité) :

- D,d : fausse détection ou contamination liée à la présence d'un pic de diffraction
- P,p : fausse détection ou contamination liée à un problème de rémanence dû à une source brillante précédemment détectée
- H,h : fausse détection ou contamination liée à la présence du halo d'une source brillante proche
- O,o : fausse détection ou contamination liée à des réflexions internes dans le montage optique
- 0 : aucun problème.

Afin d'avoir des mesures de magnitudes de qualité, on exclura les sources ayant des indices différents de « 0 », « o » ou « h ». Cette sélection entraîne l'élimination supplémentaire de 12 %, 53 %, et 33 % des sources, respectivement pour les régions « extragalactique », « centre galactique » et « anti-centre galactique ».

Du fait de ces sélections, le nombre de contreparties attendu dans la boîte d'erreur d'une source fermi non-identifiées, s'étalera de quelques unités à quelques dizaines pour les sources de haute latitude, et de quelques dizaines à quelque centaines pour les sources de basse latitude.

11.5.2 Correction de l'extinction dans l'infrarouge

Lors de la recherche de contreparties infrarouges de type blazar dans la région d'une source non-identifiée du *Fermi-LAT*, on supposera que chacune des sources infrarouges a une origine extragalactique. Du fait de la présence de grains de poussière dans le milieu interstellaire qui absorbent et diffusent les photons il est nécessaire de corriger les magnitudes des sources délivrées par le catalogue *WISE* de cette extinction. Ces corrections dépendent de la position de la source considérée et du filtre dans lequel a été mesuré la magnitude. Elles augmentent significativement lorsque l'on s'approche du plan galactique.

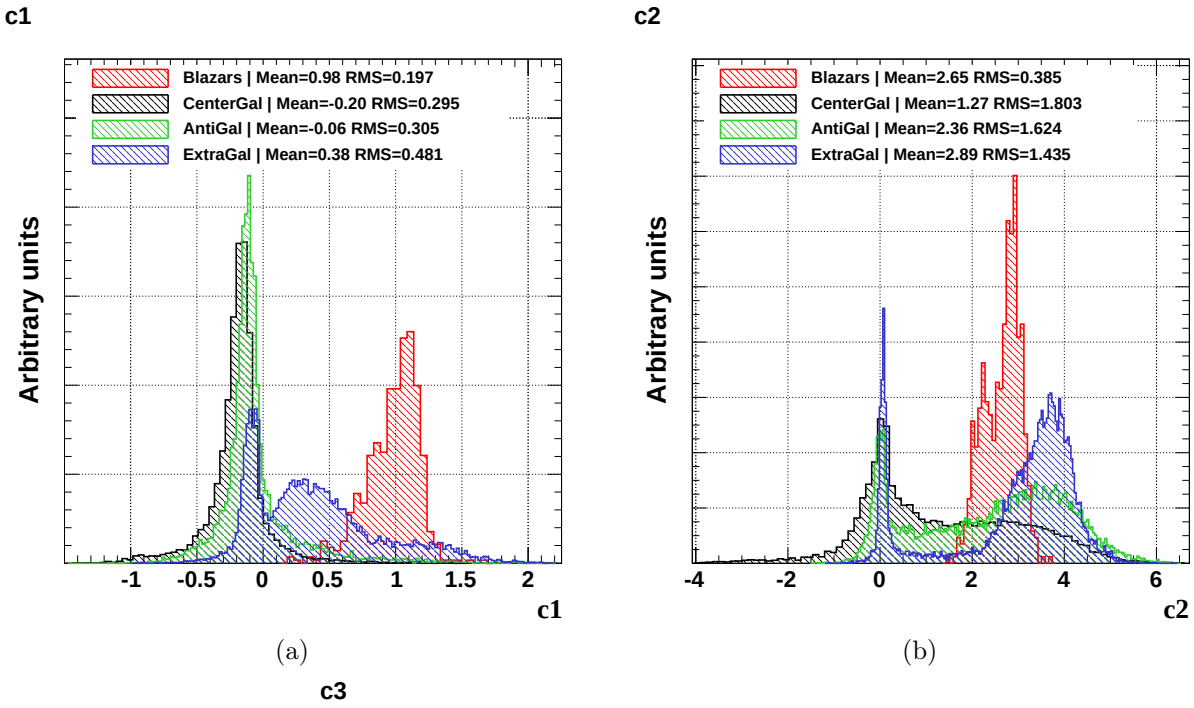
On utilisera par la suite le service proposé par l'IRSA afin d'obtenir les corrections à apporter pour les filtres W_1 et W_2. Les corrections dans les filtres de plus haute longueur d'onde sont négligeables. On pourra se référer au site Internet de l'IRSA [1] afin d'obtenir une description détaillée et les articles de références sur les méthodes utilisées. L'absorption en magnitude est calculée à partir de cartes de colonnes de densité d'hydrogène à 100 µm, qui ont été obtenues avec le satellite IRAS et avec l'instrument DIRBE embarqué à bord du satellite COBE. L'extrapolation dans la bande passante des filtres W_1 et W_2 est réalisée à l'aide de la méthode décrite dans [Indebetouw et al., 2005].

L'absorption des photons dans l'infrarouge suivant une loi de puissance en fonction de la longueur d'onde [2], la correction dans W_1 sera plus élevée que celle dans W_2. Cela a pour conséquences de décaler vers les faibles valeurs les couleurs c_1 et c_2, notamment pour les régions de basse latitude galactique pour lesquelles l'absorption est importante. Par la suite, on appliquera de manière systématique les corrections de l'extinction à toutes les sources avec lesquelles nous travaillerons.

11.5.3 Les couleurs

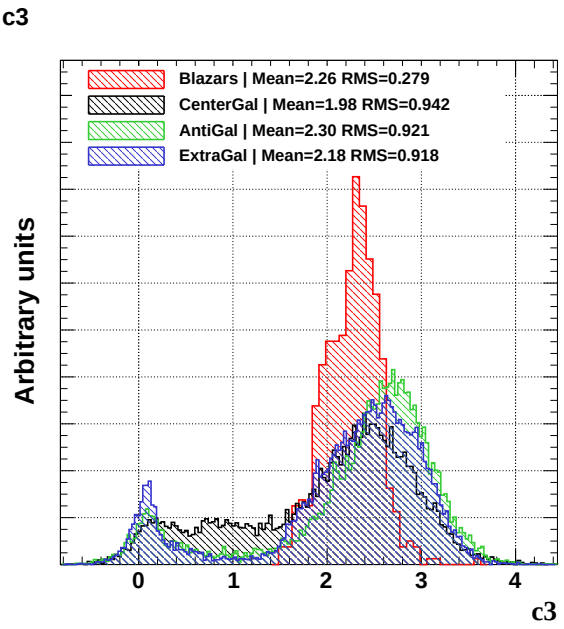
Les distributions des trois couleurs c_1, c_2 et c_3 ont été construites à partir des mesures de magnitudes dans les quatre filtres du satellite *WISE*. Les sources considérées sont celles qui passent les critères de sélection définis à la section 11.5.1. Les magnitudes ont été corrigées de l'extinction infrarouge selon la procédure précédemment explicitée. Ces distributions sont représentées sur la Figure 11.5, en rouge pour le lot de blazars identifiés, et en bleu, noir ou vert pour des lots de fond correspondant respectivement aux régions « extragalactique », « centre galactique » et « anti-centre galactique », qui ont été définis à la section 11.4. Cette figure confirme que les distributions des couleurs présentent un fort pouvoir de discrimination blazar/fond, mais elle montre aussi que le pouvoir de séparation varie de manière significative selon la région du ciel observée. Ces deux observations montrent la pertinence de découper l'espace des phases en longitude et en latitude

1. http://irsa.ipac.caltech.edu/applications/DUST/
2. Entre 1 µm et ∼8 µm, l'absorption est proportionnelle à $\lambda^{-\beta}$ où β est de l'ordre 1,8 [Draine, 2003].



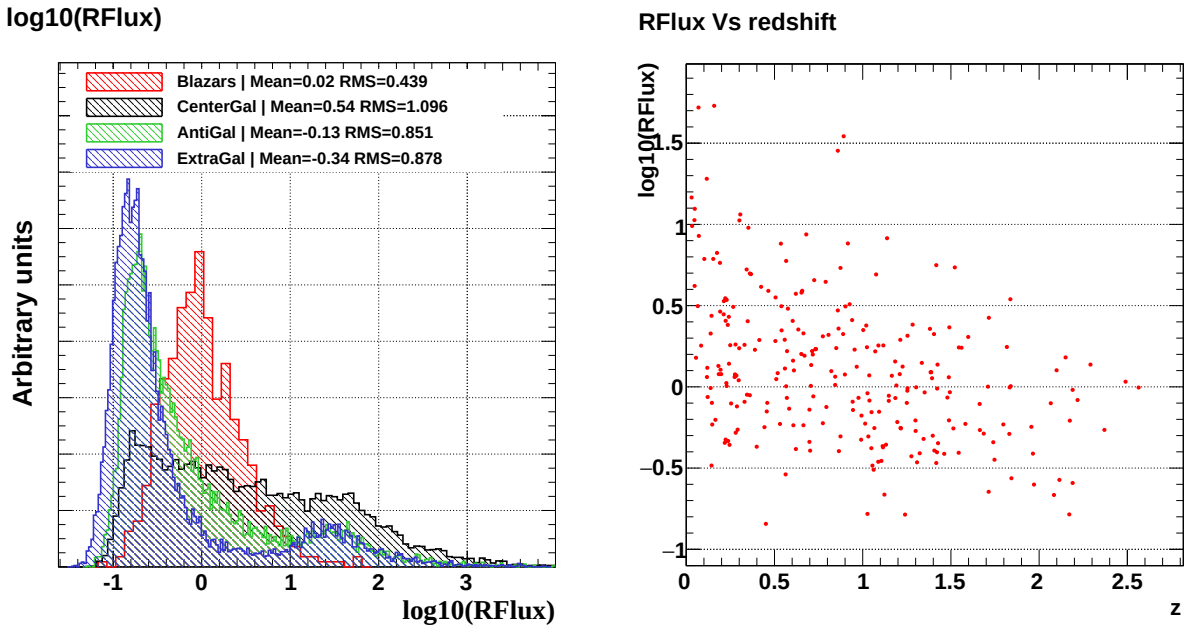
(a)

(b)



(c)

FIGURE 11.5: Distributions des couleurs c_1, c_2 et c_3 obtenues après correction de l'extinction infrarouge (voir texte) pour les sources possédant un niveau de signal sur bruit supérieur à 5 dans les trois premiers filtres et supérieur à 3 dans le dernier filtre. Les différents histogrammes représentent les blazars identifiés (rouge), le fond « extragalactique » (bleu), le fond « centre galactique » (noir) et le fond « anti-centre galactique » (vert).



(a) (b)

FIGURE 11.6: (a) Distributions du paramètre RFlux obtenues après correction de l'extinction infrarouge (voir texte) pour les sources possédant un niveau de signal sur bruit supérieur à 5 dans les trois premiers filtres et supérieur à 3 dans le dernier filtre. Les différents histogrammes représentent les blazars identifiés (rouge), le fond « extragalactique » (bleu), le fond « centre galactique » (noir) et le fond « anti-centre galactique » (vert). (b) Paramètre RFlux en fonction du redshift pour les blazars identifiés dont le redshift est connu.

galactique afin de spécialiser la recherche des contreparties infrarouges de type blazar en fonction du type de fond.

11.5.4 Le flux infrarouge des blazars

À l'aide d'un échantillon de blazars du catalogue 2FGL détectés en IR par le satellite *WISE*, [D'Abrusco et al., 2012] proposent d'obtenir un paramètre représentatif du flux intégré des blazars dans l'infrarouge en sommant les contributions des flux intégrés dans les quatre filtres pour étudier le paramètre de dominance Compton, défini comme le rapport entre les flux intégrés de la bosse inverse Compton (domaine γ) et de la bosse synchrotron (domaine infrarouge). Utiliser le paramètre de dominance Compton afin d'augmenter le pouvoir de séparation « blazar »/« non-blazar » n'est cependant pas envisageable car il faudrait pour cela avoir en possession un flux γ à associer à toutes les sources IR. On peut observer sur la Figure 11.6a la distribution de ce flux infrarouge, calculé selon la méthode proposée par [D'Abrusco et al., 2012], et divisé par une valeur de référence qui correspond au flux moyen des blazars, quantité que l'on appellera par la suite RFlux. Y sont représentés les blazars et les différents types de fond précédemment définis (même code couleur). Bien que ce paramètre possède un pouvoir de séparation « blazar »/« non-blazar » évident, notamment entre le signal et le fond « extragalactique », il est à priori délicat d'utiliser un flux comme variable discriminante dans une analyse multivariée, ce dernier étant fonction de la distance à la source. Néanmoins, on peut observer sur la Figure 11.6b le paramètre RFlux en fonction du redshift pour 249 blazars identifiés possédant une mesure de redshift, et on constate une légère dépendance du niveau du flux moyen des blazars au redshift qui reste tout de même relativement faible par comparaison avec le gros de la distribution. On montrera par la suite que l'utilisation de cette variable accroît le pouvoir de séparation de l'analyse.

Le paramètre RFlux est calculé de la manière suivante :

$$\mathrm{RFlux} = \frac{1}{F_{\mathrm{Ref}}} \sum_{i=1}^{4} F_i = \sum_{i=1}^{4} (F^{\star}_{i,0}/f_c) \times 10^{-m_i/2{,}5} \tag{11.2}$$

où $F^{\star}_{i,0}$ est la densité de flux de la source de référence dans le filtre i, et F_{Ref} est le flux moyen du lot de blazars identifiés ($F_{\mathrm{Ref}} \approx 7{,}41 \times 10^{-12}\,\mathrm{erg\,cm^{-2}\,s^{-1}}$). Cette quantité ayant été calculée pour un indice spectral de -2, elle nécessite en principe l'application d'une correction f_c qui dépend de l'indice (la couleur) de la source. Cette correction, qui est donnée dans [Wright et al., 2010], est de l'ordre de 1 % pour les filtres W_1, W_2 et W_4 et peut atteindre quelques pourcent pour W_3 qui possède une bande passante plus importante. Néanmoins, au vue des faibles magnitudes que cela représente, nous utiliserons par la suite $f_c = 1$.

11.5.5 La variabilité

Le catalogue *WISE* met à disposition une mesure de variabilité par filtre pour chaque source. Il utilise pour cela différentes observations d'une source dans le temps, dont le nombre dépend à la fois du filtre et de la région observée (exposition). Cet indice est calculé s'il existe au moins 6 mesures de qualité dans un filtre et prend des valeurs comprises en 0 et 10. Les valeurs rendent comptent d'une probabilité croissante de variabilité. Ce paramètre présente en principe un pouvoir de séparation blazar/fond intéressant. Nous ne l'utiliserons pas car la non-homogénéité de l'acceptance de *WISE* complique son utilisation dans une analyse multivariée. Néanmoins, pour chaque contrepartie infrarouge associée à une source non-identifiée du catalogue 2FGL, on indiquera dans les résultats finaux l'absence ou la présence de variabilité d'une source IR dans les filtres W_1 et W_2, ou le cas échéant l'absence d'information.

11.6 Construction et performances des analyses

Cette section a pour objectif d'introduire l'analyse dédiée à la recherche des contreparties IR des sources non-identifiées *Fermi-LAT*. Après avoir défini les différents échantillons de sources (évènements) qui seront utilisés pour les phases d'entraînements et de test, nous discuterons de notre choix d'utiliser pour ce problème particulier de classification binaire « blazar/non-blazar », un réseau de neurones, ainsi que les réglages associés. Enfin, les performances et le point de fonctionnement de l'analyse seront discutés.

11.6.1 Définition des différents échantillons

Toutes les sources infrarouges qui constituent les différents échantillons détaillés ci-dessous ont été sélectionnées selon les seuils et les critères de qualité définis à la section 11.5.1.

11.6.1.1 Échantillon « signal »

L'échantillon « signal », construit à partir des blazars de la WBS, est constitué de 558 blazars appartenant à la fois aux catalogues Roma-BZCAT et 2FGL. Pour chacune des sources, la contrepartie IR a été obtenue dans une région circulaire d'un rayon de 3,3″ centrée autour de la position fournie par le catalogue Roma-BZCAT. Les distributions des variables discriminantes pour cet échantillon ont été présentées sur les Figures 11.5 et 11.6a (en rouge).

11.6.1.2 Échantillons de « fond »

La section 11.4 nous a permis de mettre en évidence des différences marquées entre les distributions des paramètres discriminants selon que les sources sont de nature galactique et extragalactique.

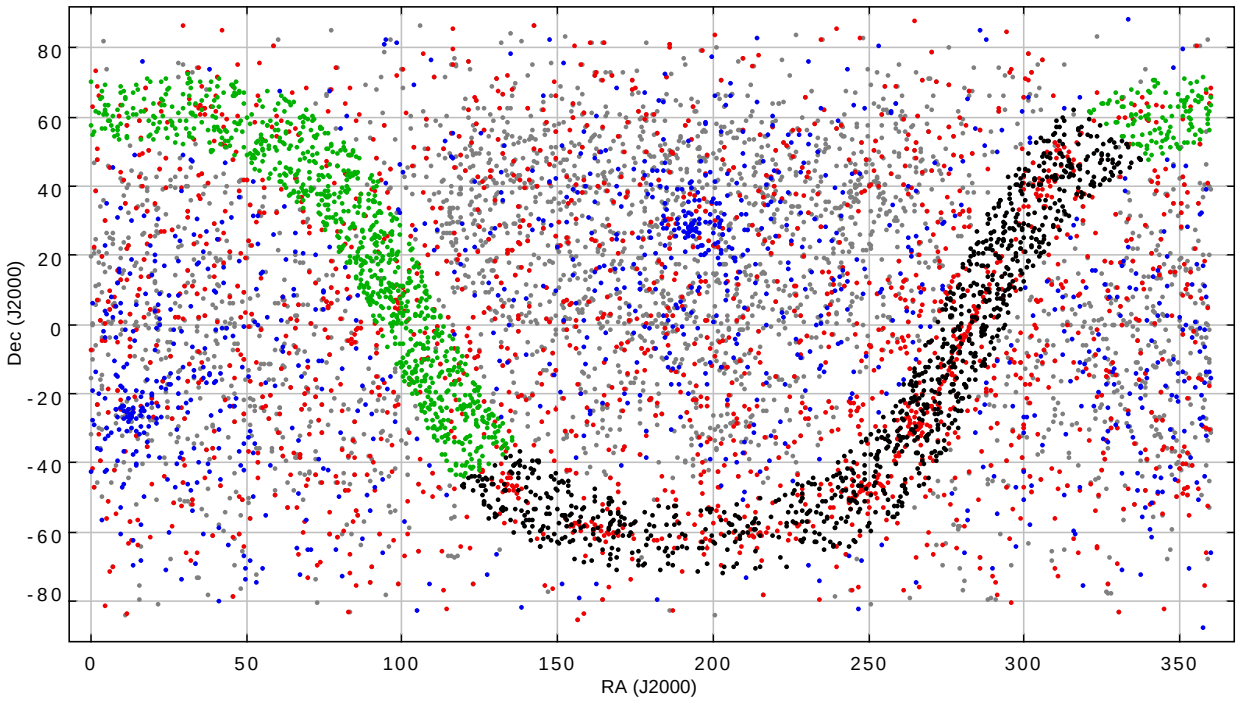


FIGURE 11.7: Coordonnées angulaires dans le système équatorial des positions des sous-régions utilisées pour la construction des échantillons de fond « extragalactique » (bleu), « centre galactique » (noir) et « anti-centre galactique » (vert). Y sont aussi représentées les coordonnées des sources appartenant aux catalogues 2FGL (rouge) et Roma-BZCAT (gris).

Cette étude nous a conduit à séparer le travail suivant les trois régions précedemment définies : « extragalactique », « centre galactique » et « anti-centre galactique ».

Afin de constituer les échantillons qui serviront à l'entraînement pour ces trois régions, nous avons tiré aléatoirement pour chacune d'entre elles 1000 sous-régions circulaires choisies de manière à ce qu'aucune source appartenant aux catalogues 2FGL et Roma-BZCAT ne soit présente. Les rayons de ces sous-régions sont de 0,08°, 0,05° et 0,07°, respectivement pour les échantillons « extragalactique », « centre galactique » et « anti-centre galactique », et ont été choisis de manière à avoir un nombre total de sources dans chacun des trois échantillons supérieur à 10 000. Ces nombres sont respectivement de ∼15 000, ∼13 000 et ∼11 500. Les positions des sous-régions utilisées pour construire ces échantillons sont représentées respectivement en bleu, vert et noir sur la Figure 11.7.

11.6.2 Sélection des analyses multivariées

Pour ce problème de classification, nos échantillons de « signal » et de « fond » présentent des nombres d'évènements très différents. En effet, notre sélection de contreparties IR de blazars est limitée par les correspondances entre les catalogues 2FGL et Roma-BZCAT ce qui conduit à un échantillon composé de 558 évènements. À l'inverse, le fond IR est abondant car le catalogue *WISE* contient plusieurs centaines de millions de sources et les régions ne comportant pas de sources des catalogues 2FGL et Roma-BZCAT sont nombreuses. Nous sommes donc dans un problème où le fond, d'un point de vue statistique, est très bien contraint, au contraire des évènements « signal » pour lesquels nous possédons une faible statistique.

Le choix de la méthode multivariée a reposé sur sa capacité à fournir des résultats stables, dans ce cas de figure avec une statistique limitée pour le signal. Dans une phase très préliminaire, nous avons retiré 20 blazars de l'échantillon « signal », réalisé les phases d'entraînement et de test avec différentes méthodes d'analyse multivariées, et tenté de retrouver les contreparties IR

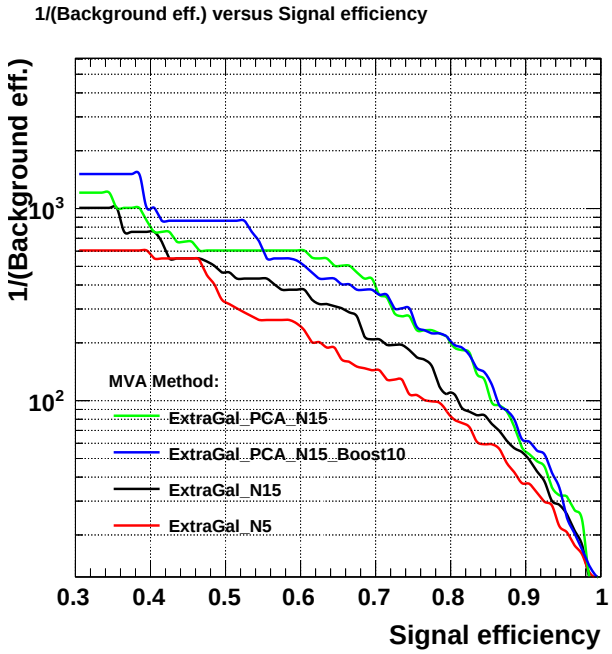


FIGURE 11.8: Illustration de la variation des performances de la méthode MLP pour l'analyse « extragalactique » en fonction de différents réglages. N5 et N15 représentent respectivement l'utilisation d'une couche cachée de 5 et 15 neurones, PCA désigne la réalisation d'une transformation dans l'espace des composantes principales préalablement à l'entraînement, et Boost10 désigne l'utilisation de la technique de « boosting » avec un nombre de « boosts » égal à 10.

compatibles avec la position donnée par le catalogue Roma-BZCAT. Cet exercice, qui s'est avéré très encourageant, nous a conduit à nous intéresser à un nombre limité de méthodes non-paramétriques. Parmi les méthodes multivariées testées, seuls le réseau de neurones de type MLP et la technique SVM permettent d'obtenir des performances stables et reproductibles, avec un avantage significatif pour la méthode MLP. Les techniques basées sur des arbres de décision telles que BDT et RF semblent donner des performances intéressantes, mais le réglage optimal qui permet d'avoir une stabilité dans les performances n'a pas été trouvé. Nous nous focaliserons par la suite sur la méthode MLP qui semble être la plus performante, et elle sera utilisée dans le cadre de la librairie TMVA.

11.6.3 Réglages de la méthode MLP

Les réglages qui ont permis d'optimiser la méthode MLP sont présentés dans la Table 11.1. La méthodologie utilisée afin d'obtenir ces réglages a consisté à comparer les performances du réseau de neurones avec un jeu de paramètres fixés, tout en faisant varier les valeurs de l'un d'entre eux. Ces paramètres sont entre autres la transformation à appliquer aux variables discriminantes (décorrélation, etc.), le nombre de neurones sur une unique couche cachée, le nombre de couches cachées, et l'utilisation de la méthode de « boosting ».

Les trois couleurs c_1, c_2 et c_3 étant linéairement corrélées, il s'est avéré que l'application d'une décomposition en composantes principales (PCA) des variables discriminantes avant la phase d'entraînement est une étape qui permet d'améliorer les performances de manière très significative. Pour le choix de l'architecture du réseau, nous avons commencé par utiliser une unique couche de neurones cachée. En faisant varier le nombre de neurones sur cette couche entre $N_{\rm var} + N_{\rm var}$ [3] et $N_{\rm var} + 30$, un optimum a été obtenu pour la configuration $N_{\rm var} + 15$. Complexifier la structure du réseau en ajoutant une seconde couche cachée de neurones a conduit dans le meilleur des cas à

3. $N_{\rm var}$ représente le nombre de variables discriminantes utilisées.

Option	Valeur	Description
HiddenLayers	N+15	Architecture du réseau
TrainingMethod	BP	Algorithme d'apprentissage
BPMode	*sequential*	Méthode d'apprentissage
NCycles	20 000	Nombre de cycles avant arrêt de l'entraînement
ConvergenceTests	20	Arrête l'entraînement après ce nombre d'itérations sans amélioration
LearningRate	0,02	Taux d'apprentissage
DecayRate	0,01	Taux de décroissance du taux d'apprentissage
TestRate	7	Nombre de cycles après lequel effectuer le test de convergence
Sampling	1,0	Fraction d'événements choisis aléatoirement utilisée pendant *TestRate* cycles
SamplingEpoch	1,0	Fraction du nombre de cycles après lequel tous les événements sont utilisés pour l'entraînement
UseRegulator	*False*	Utilisation de la régularisation
Boost_Type	*AdaBoost*	Méthode de « boosting »
Boost_Num	10	Nombre de « boosts »
VarTransform	PCA	Transformation à appliquer aux variables discriminantes avant la phase d'entraînement

TABLE 11.1: Paramètres TMVA de la méthode MLP utilisée pour la recherche des contreparties IR des sources non-identifiées du *Fermi-LAT*.

	Classement	Importance
Analyse « extragalactique »	c_1	34,4
	c_3	3,69
	c_2	2,65
	RFlux	2,11
	Classement	Importance
Analyse « centre galactique »	c_1	56,8
	RFlux	6,77
	c_2	4,06
	c_3	3,91
	Classement	Importance
Analyse « anti-centre galactique »	c_1	36,4
	c_3	7,62
	c_2	4,78
	RFlux	1,71

TABLE 11.2: Classement selon l'importance des variables discriminantes des différentes analyses. Voir formule 85 dans : http://tmva.sourceforge.net/docu/TMVAUsersGuide.pdf.

des performances équivalentes à celles précédemment obtenues avec une couche unique. Concernant l'utilisation de la technique du « boosting », son intérêt est double. Cette technique permet généralement d'améliorer de manière significative les performances et aussi d'éviter une spécialisation trop importante d'un classifieur conduisant à une mauvaise adaptation du réseau lors du traitement de nouveaux évènements (sur-entraînement). Ici, les tests ont été conduits en effectuant un nombre de « boosts » compris entre 1 et 100. À partir de dix « boosts », une légère amélioration des performances est obtenue.

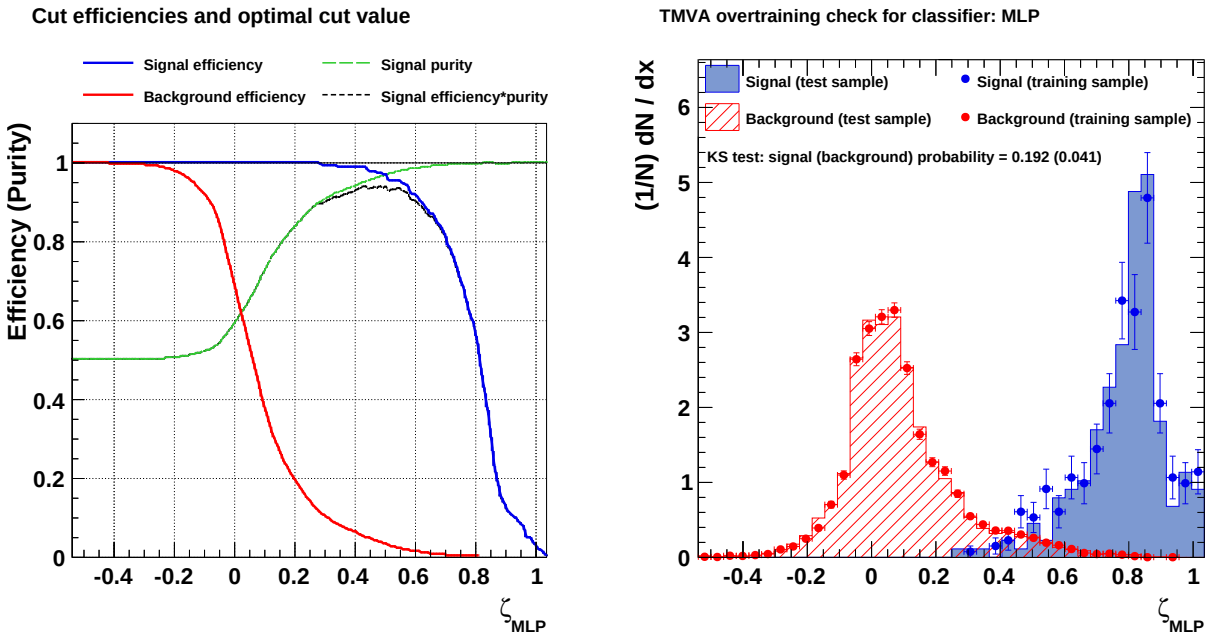
À titre d'illustration, on peut trouver sur la Figure 11.8 l'inverse de l'efficacité de reconstruction du fond en fonction de l'efficacité de reconstruction du signal pour différents réglages du réseau de neurones. Les courbes rouge et noire illustrent le gain de performance entre deux réseaux possédant une unique couche cachée constituée respectivement de $N_{\text{var}} + 5$ et $N_{\text{var}} + 15$ neurones. La courbe verte représente les performances de cette dernière configuration avec un changement de base pour les variables discriminantes, réalisé avec une décomposition en composantes principales. La courbe bleue représente cette dernière configuration avec l'usage de 10 « boosts ».

Au final, nous avons choisi d'effectuer une décomposition en composantes principales sur les variables discriminantes, de constituer le réseau avec une unique couche cachée composée de $N_{\text{var}}+15$ neurones et d'effectuer 10 étapes de « boosting ». On pourra trouver l'importance [4] relative des variables discriminantes estimée lors de la phase d'entraînement pour chacune des trois analyses dans la Table 11.2.

11.6.4 Performances

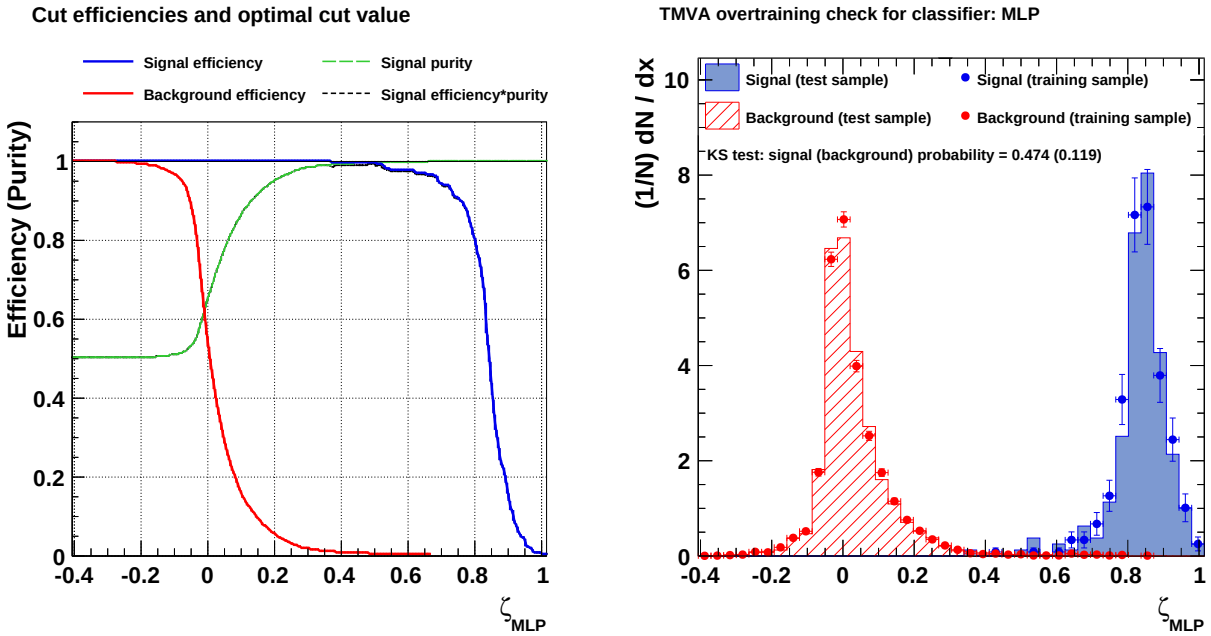
Les entraînements des réseaux de neurones sont effectués avec 60 % des évènements des échantillons « signal » et « fond », le reste des événements étant utilisé afin de dériver les performances des analyses. La stabilité des résultats a été éprouvée en tirant aléatoirement et à plusieurs reprises différentes réalisation des lots d'entraînement et de test parmi les sources des échantillons « signal » et « fond ». Les performances sont représentées à gauche des Figures 11.9, 11.10 et 11.11, respectivement pour les analyses « extragalactique », « centre galactique » et « anti-centre galactique ». Les efficacités de reconstruction du signal (en bleu), notées ε_{Sig}, et du fond (en rouge), notées ε_{Bkg}, sont représentées en fonction du paramètre ζ. Pour l'analyse « centre galactique », et dans une moindre

4. On trouvera la définition de l'importance d'une variable pour un réseau de neurones dans le documentation de TMVA (formule 85) : http://tmva.sourceforge.net/docu/TMVAUsersGuide.pdf.



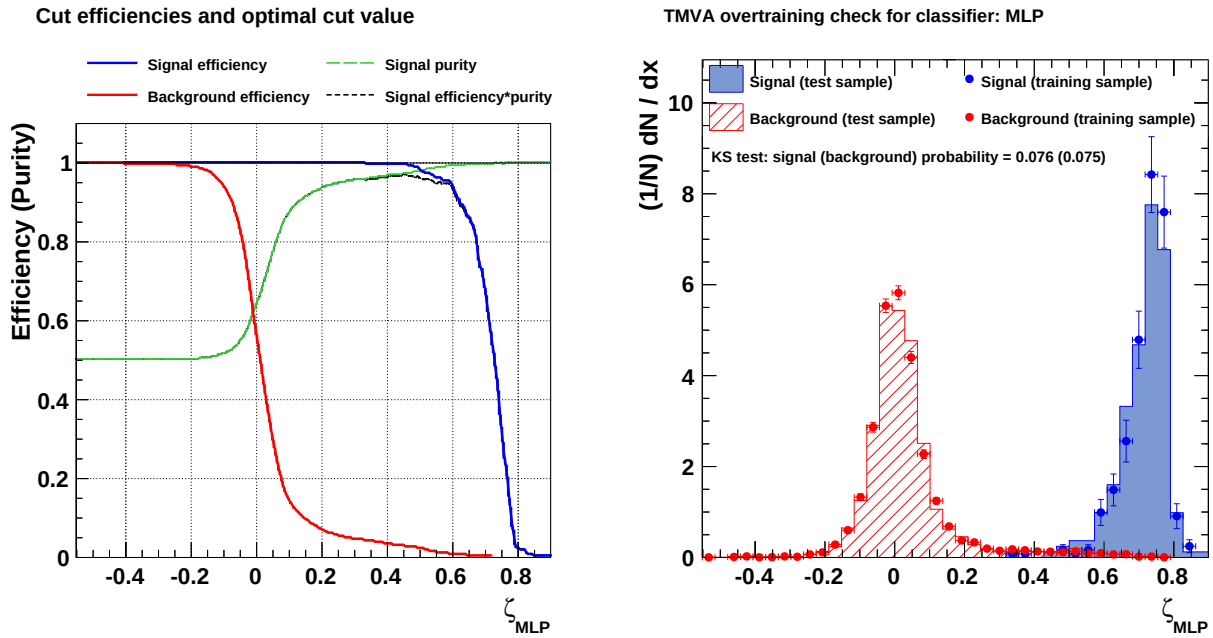
(a) (b)

FIGURE 11.9: (a) Efficacités de reconstruction du signal (bleu) et du fond (rouge) en fonction de ζ pour l'analyse « extragalactique ». (b) Distributions du paramètre ζ des évènements signal (bleu) et des évènements de fond (rouge) pour les lots d'entraînement (traits pleins) et de test (points).



(a) (b)

FIGURE 11.10: (a) Efficacités de reconstruction du signal (bleu) et du fond (rouge) en fonction de ζ pour l'analyse « centre galactique ». (b) Distributions du paramètre ζ des évènements signal (bleu) et des évènements de fond (rouge) pour les lots d'entraînement (traits pleins) et de test (points).



(a) (b)

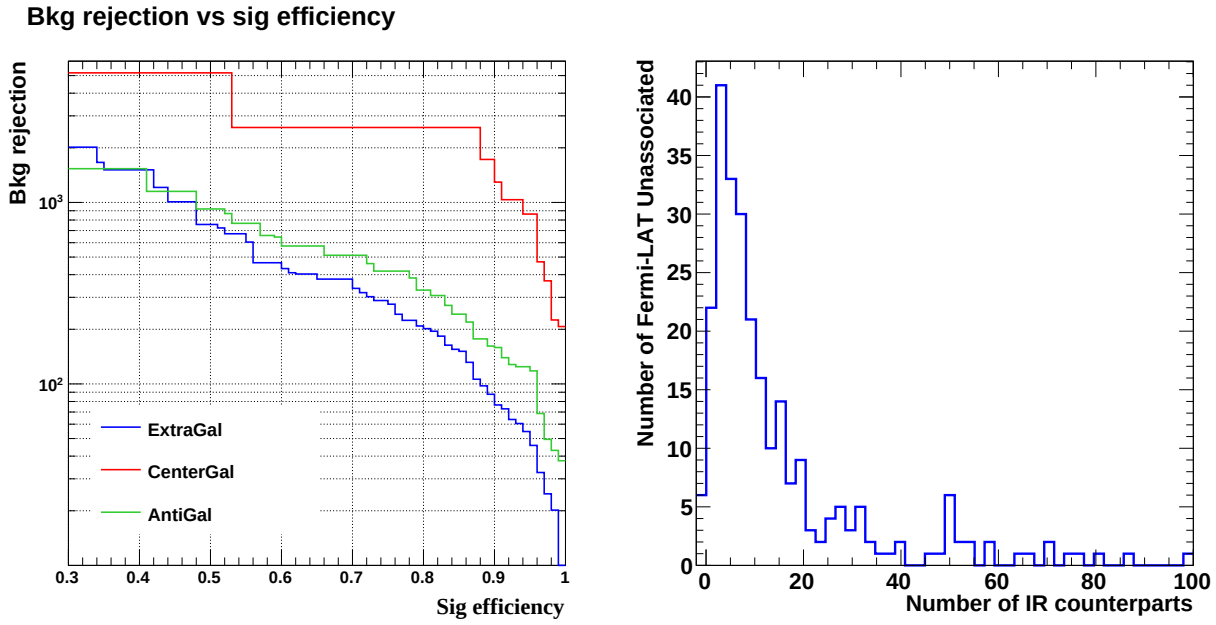
FIGURE 11.11: (a) Efficacités de reconstruction du signal (bleu) et du fond (rouge) en fonction de ζ pour l'analyse « anti-centre galactique ». (b) Distributions du paramètre ζ des évènements signal (bleu) et des évènements de fond (rouge) pour les lots d'entraînement (traits pleins) et de test (points).

mesure pour l'analyse « anti-centre galactique », le fort pouvoir de séparation signal/fond de la couleur c_1 permet d'obtenir un taux de rejet du fond très important tout en gardant une grande efficacité d'identification du signal. À titre d'illustration, pour l'analyse d'une source « galactique », lorsque $\varepsilon_{\mathrm{Sig}} \approx 90\,\%$, $\varepsilon_{\mathrm{Bkg}}$ est inférieure à 0,1 %. Pour cette même efficacité de reconstruction et dans le cas de l'analyse « extragalactique », $\varepsilon_{\mathrm{Bkg}}$ est de l'ordre de 1,2 %. À droite des figures précédentes, les distributions du paramètre ζ associées sont représentées pour les évènements signal (bleu) et pour les évènements de fond (rouge), et ce pour les lots d'entraînement (traits pleins) et de tests (points). Ces figures permettent de vérifier le très bon accord entre la réponse des classifieurs pour les lots d'entraînement et de test. On trouvera sur la Figure 11.12a une comparaison des performances des trois analyses qui sont représentées par $\varepsilon_{\mathrm{Bkg}}^{-1}$ en fonction de $\varepsilon_{\mathrm{Sig}}$.

11.6.5 Points de fonctionnement de l'analyse

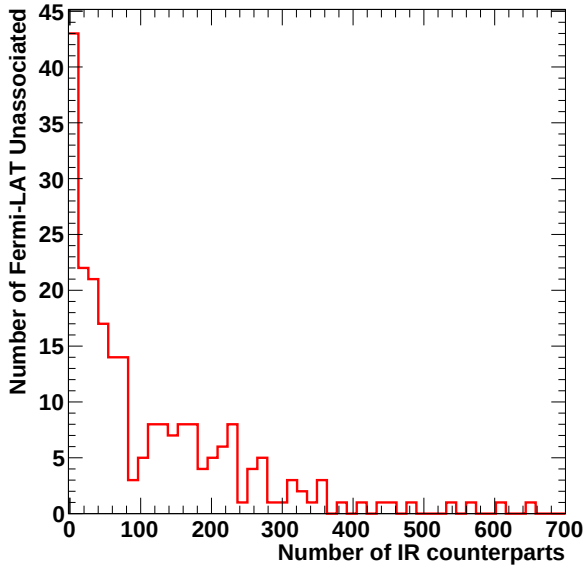
Les performances des analyses étant établies, il convient maintenant d'y associer des points de fonctionnement qui permettront, lors de l'analyse d'une source non-identifiée, un compromis entre l'efficacité de reconstruction du signal et le niveau de contamination du fond. Il est important de noter ici que, pour chaque source non-identifiée du *Fermi-LAT*, un point de fonctionnement correspond à un nombre de fausses identifications $\eta = \varepsilon_{\mathrm{Bkg}} \times N_c$, où N_c représente le nombre de contreparties infrarouges dans la boîte d'erreur de cette source. Les distributions de N_c pour les sources non-identifiées sont représentées sur les Figures 11.12b, 11.12c et 11.12d, selon qu'elles se trouvent respectivement dans les régions « extragalactique », « centre galactique » et « anti-centre galactique ». Les valeurs médianes correspondantes sont de ∼8,5, ∼69 et ∼57. En complément de ce qui est visible sur ces figures, notons qu'une dizaine de sources non-identifiées de la région « extragalactique » ont entre 100 et 400 contreparties, et qu'une quinzaine de sources de la région « anti-centre galactique » ont entre 200 et 1000 contreparties.

Afin de déterminer si une des contreparties infrarouges se trouvant dans la boîte d'erreur d'une

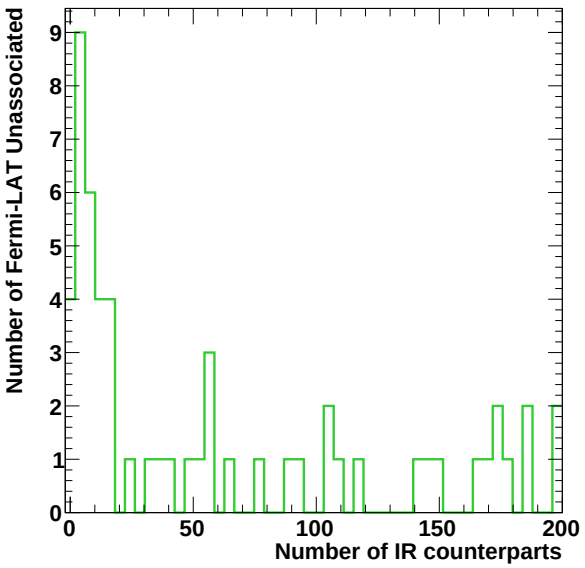


(a)

(b)



(c)



(d)

FIGURE 11.12: (a) Performances des analyses « extragalactique » (bleu), « centre galactique » (rouge) et « anti-centre galactique » (vert). Nombre de contreparties infrarouges dans la boîte d'erreur des sources non-identifiées pour les régions (b) « extragalactique », (c) « centre galactique » et (d) « anti-centre galactique ».

		Régime I	Régime II	Régime III	Régime IV
Analyse « extragalactique »	$\varepsilon_{\mathrm{Sig}}$	40,0 %	60,0 %	80,0 %	90,0 %
	$\varepsilon_{\mathrm{Bkg}}$	$7{,}0 \times 10^{-4}$	$2{,}1 \times 10^{-3}$	$5{,}0 \times 10^{-3}$	$1{,}12 \times 10^{-2}$
	$\zeta^{\star}$	0,8335	0,7912	0,7045	0,6216
Analyse « centre galactique »	$\varepsilon_{\mathrm{Sig}}$	40,0 %	60,0 %	80,0 %	90,0 %
	$\varepsilon_{\mathrm{Bkg}}$	$1{,}9 \times 10^{-4}$	$3{,}9 \times 10^{-4}$	$3{,}9 \times 10^{-4}$	$7{,}70 \times 10^{-4}$
	$\zeta^{\star}$	0,8544	0,8335	0,8016	0,7584
Analyse « anti-centre galactique »	$\varepsilon_{\mathrm{Sig}}$	40,0 %	60,0 %	80,0 %	90,0 %
	$\varepsilon_{\mathrm{Bkg}}$	$7{,}0 \times 10^{-4}$	$1{,}7 \times 10^{-3}$	$3{,}0 \times 10^{-3}$	$6{,}30 \times 10^{-3}$
	$\zeta^{\star}$	0,7420	0,7122	0,6734	0,6219

TABLE 11.3: Efficacités de reconstruction des différents régimes pour les différentes analyses et valeurs finales de discrimination associées.

source non-identifiée est potentiellement un blazar, nous allons identifier le régime de fonctionnement qui lui convient, auquel nous associerons une valeur de l'efficacité de reconstruction du fond $\varepsilon_{\mathrm{Bkg}}$. Nous définissons pour chacune des analyses les quatre régimes suivants :

— régime I : contreparties dont la valeur de ζ est plus grande que la valeur $\zeta^{\star}_{\mathrm{I}}$ correspondant à l'efficacité de reconstruction $\varepsilon^{\mathrm{I}}_{\mathrm{Sig}} = 40\,\%$

— régime II : contreparties dont la valeur de ζ est comprise entre les valeurs $\zeta^{\star}_{\mathrm{II}}$ et $\zeta^{\star}_{\mathrm{I}}$, $\zeta^{\star}_{\mathrm{II}}$ correspondant à $\varepsilon^{\mathrm{II}}_{\mathrm{Sig}} = 60\,\%$

— régime III : contreparties dont la valeur de ζ est comprise entre les valeurs $\zeta^{\star}_{\mathrm{III}}$ et $\zeta^{\star}_{\mathrm{II}}$, $\zeta^{\star}_{\mathrm{III}}$ correspondant à $\varepsilon^{\mathrm{III}}_{\mathrm{Sig}} = 80\,\%$

— régime IV : contreparties dont la valeur de ζ est comprise entre les valeurs $\zeta^{\star}_{\mathrm{IV}}$ et $\zeta^{\star}_{\mathrm{III}}$, $\zeta^{\star}_{\mathrm{IV}}$ correspondant à $\varepsilon^{\mathrm{IV}}_{\mathrm{Sig}} = 90\,\%$

Une contrepartie infrarouge correspondant à un de ces régimes est vue par le réseau de neurones comme un possible candidat blazar. À chacun de ces régimes on associera une valeur de $\varepsilon^{i}_{\mathrm{Bkg}}$ ($i = \{\mathrm{I}, \mathrm{II}, \mathrm{III}, \mathrm{IV}\}$) correspondant aux points de fonctionnement illustrés sur la Figure 11.12a. On pourra trouver pour chaque régime et pour chacune des trois analyses, les valeurs de $\varepsilon^{i}_{\mathrm{Sig}}$, $\varepsilon^{i}_{\mathrm{Bkg}}$ et de $\zeta^{\star}_{i}$ dans la Table 11.3.

Chaque source non-identifiée du *Fermi-LAT* a un nombre de contreparties N_c qui lui est propre. Chacune de ces contreparties correspondant à un des régimes I, II, II ou IV définis ci-dessus, il est possible de lui attribuer une estimation du nombre de fausses identifications attendues, $\eta^{i} = \varepsilon^{i}_{\mathrm{Bkg}} \times N_c$. À partir de cette estimation, nous définissons quatre classes A, B, C et D qui répartissent les contreparties selon les valeurs de η :

— Classe A : $\eta < 1/100$

— Classe B : $1/100 \leq \eta < 5/100$

— Classe C : $5/100 \leq \eta < 10/100$

— Classe D : $10/100 \leq \eta < 50/100$

Une source infrarouge pour laquelle le nombre moyen de fausses associations est supérieur à 1/2 ne sera pas considérée.

11.7 Résultats

Les analyses multivariées ont été réglées en aveugle sans jamais tester l'application des différents entraînements sur les sources non-identifiées. C'est seulement une fois que les points de fonctionnement ont été déterminés et que les classes de candidats blazars ont été définies que nous avons appliqué la méthode sur les sources non-identifiées du catalogue 2FGL. Pour chacune d'entre elles, l'ensemble des sources IR qui sont présentes dans la région circulaire de rayon θ_{95} correspondant à un niveau de confiance de 95 % sur leur localisation et qui passent les seuils de sélection définis à la section 11.5.1 ont été analysées selon la procédure explicitée à la section précédente. À titre de

contrôle, pour chaque source non-identifiée nous avons estimé le nombre de fausses identifications en appliquant la méthode dans une région annulaire de rayon inférieur θ_{95} et de rayon extérieur de $\sqrt{2}\theta_{95}$, centrée sur la position des sources non-identifiées (qui est définie de manière à ce que l'angle solide soit identique à celui de la boîte d'erreur de la source non-identifiée). On notera $\widetilde{N}_W$ le nombre de contreparties infrarouges dans cette région et $\widetilde{N}_c$ le nombre parmi celles-ci qui seront classées A, B, C ou D lors de notre recherche de candidats blazars.

Notre méthode permet d'obtenir 130 contreparties IR de type blazar réparties parmi 104 sources non-identifiées du catalogue 2FGL. On en compte respectivement 21, 35, 26 et 48 de classe A, B, C et D, ce qui permet de leur associer un nombre de fausses identifications meilleur que 0.01, 0.05, 0.1 et 0.5, respectivement.

Les résultats obtenus pour chacune des sources non-identifiées du catalogue 2FGL sont présentés dans la Table 11.6 et la Table 11.7, selon qu'elles possèdent ou ne possèdent pas de mises en gardes (« flags ») quant à la qualité de la reconstruction. On y trouvera les informations suivantes :

- le nom de la source non-identifiée, suivi de l'exposant ♯ si la source est considérée comme un candidat blazar à l'issue de notre étude présentée au chapitre 10, et suivie des exposants ⋆ et † respectivement si que la source est considérée comme candidat blazar à l'issue des études présentées par [Doert et Errando, 2013] et [Mirabal et al., 2012]. Un dernier exposant (un entier) représente le nombre de candidats blazars IR obtenus par [Massaro et al., 2013].
- le nombre de contreparties IR N_c classées A, B, C ou D, et le nombre total de contreparties IR dans la boîte d'erreur de la source non-identifiée N_W.
- le nom dans le catalogue *WISE* des contreparties IR classées A, B, C ou D. L'exposant « ∘ » informe si la contrepartie est également obtenue par [Massaro et al., 2013], et dans le cas contraire cela est signalé par l'exposant « × ».
- les coordonnées des contreparties IR classées A, B, C ou D provenant du catalogue *WISE*.
- la distance δ entre la position des contreparties IR classées A, B, C ou D et celle de la source non-identifiée.
- les rapports signal sur bruit dans les quatre filtres pour les contreparties IR classées A, B, C ou D notés SNR_1, SNR_2, SNR_3 et SNR_4.
- les valeurs de c_1, c_2, c_3 et de RFlux des contreparties IR classées A, B, C ou D.
- l'indice Ccf_W qui pour chaque filtre signale la qualité de la photométrie des contreparties IR classées A, B, C ou D.
- une indication de variabilité dans les filtres W_1 et W_2 pour les contreparties IR classées A, B, C ou D qui est signalée par le symbole « ∘ », et dans le cas contraire par le symbole « × ».
- la valeur ζ_{MLP} attribuée par le réseau de neurones aux contreparties IR classées A, B, C ou D.
- le nombre moyen de fausses identifications attendues, η, pour les contreparties IR classées A, B, C ou D.
- la classe de l'association des contreparties IR sélectionnées.
- le nombre de contreparties IR $\widetilde{N}_c$ classées A, B, C ou D et le nombre total de contreparties IR $\widetilde{N}_W$ dans la région de fond autour de la source non-identifiée. On indiquera entre parenthèses la classe de la meilleure association dans la région de fond de la source non-identifiée.

11.7.1 Sources non-identifiées du 2FGL sans « flag » de reconstruction

Concernant les sources non-identifiées du catalogue 2FGL qui ne possèdent pas de mise en garde sur la qualité de la reconstruction, nous obtenons dans la région « extragalactique » 44 candidats blazars IR pour 39 d'entre elles, sur un total de 201 sources (1,1 contreparties par source en moyenne). Parmi ces contreparties on en compte respectivement 13, 10, 7 et 14 de classe A, B, C et D. En étudiant l'environnement proche (région annulaire de contrôle du fond) des sources non-identifiées

qui sont associées avec une contrepartie infrarouge, on constate que l'on obtient pour deux d'entre elles au moins une source IR possédant une classe égale ou supérieure à leurs meilleures associations. On compte 27 sources non-identifiées qui correspondent à des candidats blazars proposés conjointement dans les études menées par [Doert et Errando, 2013], [Mirabal et al., 2012] et dans celle menée au chapitre précédent, pour lesquelles au moins une contrepartie de type blazar a été trouvée. Ces résultats sont à comparer avec ceux obtenus par [Massaro et al., 2013], qui obtient 61 contreparties de type blazar pour 49 sources non-identifiées (1,2 contreparties par source), dont 34 sont des candidats blazars proposés de manière conjointe par les trois études précédemment citées. Les contreparties IR de type blazar qui sont communes à cette étude et à celle de [Massaro et al., 2013] sont au nombre de 30.

Pour les 56 sources non-identifiées de la région « anti-centre galactique », un total de 11 contreparties infrarouges de type blazar (respectivement 3, 6 et 2 de classe B, C et D) sont obtenues pour 9 d'entre elles. Deux des sources non-identifiées pour lesquelles au moins une contrepartie de type blazar est obtenue possèdent une contrepartie mieux classée dans leur environnement proche. Sur les 10 candidats blazars proposés conjointement par [Doert et Errando, 2013] et dans l'étude du chapitre précédent, nous obtenons une contrepartie de type blazar pour 3 d'entre elles. [Massaro et al., 2013] proposent 7 contreparties IR pour 6 sources non-identifiées, dont deux d'entre elles correspondent à des candidats blazars proposés conjointement par les deux études précédemment citées. Dans la région « anti-centre galactique », 3 contreparties de type blazar sont communes à notre étude et à celle de [Massaro et al., 2013].

Sur les 27 sources non-identifiées de la région « centre galactique », une contrepartie unique est obtenue pour 6 d'entre elles (aucune contrepartie IR mieux classée n'est trouvée dans leur environnement proche). Parmi elles, on en dénombre deux de classe A, deux de classe B, une de classe C et une de classe D. L'une de ces sources est un candidat blazar proposé par [Doert et Errando, 2013] et par l'étude du chapitre 10. [Massaro et al., 2013] propose lui aussi une contrepartie infrarouge pour cette source (différente de la notre), et obtient un total de 6 contreparties pour 6 sources non-identifiées. Dans cette région, une seule contrepartie de type blazar est commune à cette étude et à celle de [Massaro et al., 2013].

On trouvera un récapitulatif des résultats pour les sources non-identifiées sans « flag » de reconstruction dans la Table 11.4.

11.7.2 Sources non-identifiées du 2FGL avec des « flags » de reconstruction

Parmi les 71 sources non-identifiées de la région « extragalactique » signalées comme pouvant présenter un problème de reconstruction, on obtient 19 contreparties IR de type blazar pour 16 d'entre elles (1 de classe A, 6 de classe B, 3 de classe C et 9 de classe D). Pour trois de ces sources non-identifiées, une contrepartie de classe au moins équivalente à la meilleure association dans leur boîte d'erreur respective est trouvée dans leur environnement proche (région annulaire de contrôle du fond). Sur les 31 candidats blazars proposés conjointement par les études de [Doert et Errando, 2013], [Mirabal et al., 2012] et celle du chapitre précédent, 7 d'entre elles sont associées avec de potentielles contreparties infrarouges. [Massaro et al., 2013] propose des contreparties de type blazar pour huit de ces sources, et obtient un total de 22 contreparties pour 17 sources non-identifiées. On compte 11 contreparties IR de type blazar communes à cette étude et à celle de [Massaro et al., 2013].

Dans la région « anti-centre galactique », nous proposons 7 contreparties IR de classe D réparties entre 5 sources non-identifiées (sur un total de 42). Deux des sources non-identifiées possèdent une source de classe supérieure ou égale à celle obtenue dans leurs boîtes d'erreurs. Aucune contreparties n'est associées aux trois candidats blazars commun proposés par l'étude de [Doert et Errando, 2013] et de celle du chapitre précédent. L'étude de [Massaro et al., 2013] propose 21 contreparties pour 10 sources non identifiées, dont une pour les candidats blazars communs aux deux études citées précédemment. Dans cette région, une seule contrepartie IR de type blazar est commune à cette étude et à l'étude de [Massaro et al., 2013].

Sur les 178 sources non-identifiées de la région « centre galactique » un total de 43 contreparties infrarouges est obtenus pour 29 d'entre elles (1,5 contreparties par source). On en compte respectivement 4, 14, 9 et 15 de classe A, B, C et D. Parmi les 29 sources non-identifiées associées avec une contrepartie dans l'infrarouge, dix d'entre elles ont au moins une source IR dans leur environnement proche qui est de classe au moins égale à celle de leur meilleure association. Pour six candidats blazars communs à l'étude de [Doert et Errando, 2013] et à l'étude du chapitre précédent, au moins une contrepartie IR de type blazar est proposée. Pour ces candidats blazars, [Massaro et al., 2013] en propose six, et pour l'ensemble des sources non-identifiées, ils obtiennent 28 contreparties pour 21 d'entre elles (1,3 contreparties par source). Dans la région « centre galactique », 8 contreparties de type blazar sont communes à notre étude et à celle de [Massaro et al., 2013].

On pourra se référer à la Table 11.5 pour un résumé des résultats pour les sources non-identifiées avec un « flags » de reconstruction.

11.7.3 Discussion

Dans la région « extragalactique », nous avons trouvé 63 candidats blazars IR classés A, B, C ou D pour un total de 55 (parmi 272) sources non identifiées du catalogue 2FGL. L'excès de 8 contreparties est en très bon accord avec le nombre de fausses identifications attendu (à partir de la définition des classes A, B, C et D) qui prévoit une fourchette entre 3 et 13. Il est aussi en bon accord avec les résultats obtenus sur la région annulaire de contrôle de chaque source, qui donne un total de 5 sources classées A, B, C ou D.

Parmi les 98 sources non-identifiées de la région « anti-centre galactique », nous avons trouvé 18 candidats blazars IR pour 14 d'entre elles. L'excès de 4 contreparties est en très bon accord avec le nombre de fausses identifications attendu, compris entre 1 et 5. Il est aussi en bon accord avec les résultats obtenus sur la région annulaire de contrôle de chaque source, qui donne un total de 4 sources classées A, B, C ou D.

Dans la région « centre galactique », 49 candidats blazars IR classés A, B, C ou D sont obtenus pour 35 sources non-identifiées (parmi 205). L'excès de 14 contreparties est légèrement supérieur que le nombre de fausses identifications attendu estimé entre 2 et 10. Il est similaire aux résultats obtenus sur la région annulaire de contrôle de chaque source, qui donne un total de 10 sources classées A, B, C ou D.

Les comparaisons ci-dessus nous montrent que nous avons une bonne maîtrise du nombre de fausses identifications inhérentes à la méthode. La légère différence constatée pour la région « centre galactique » illustre la complexité de cette région.

Des différences existent entre nos candidats blazars IR et ceux issus du travail de [Massaro et al., 2013]. Elles sont de plusieurs ordres. D'une part, certains candidats prévus par [Massaro et al., 2013] ont des rapports signal sur bruit dans un ou plusieurs filtres inférieures au seuils que nous nous sommes fixés (voir section 11.5.1), et d'autres ont une photométrie qui est contaminée par des effets (liés à des problèmes instrumentaux ou du fait de la présence de sources brillantes) signalés dans le catalogue *WISE* comme potentiellement problématiques (voir section 11.5.1). Nous avons fait le choix d'éliminer ces sources de notre échantillon en amont des phases d'entraînement et d'application des coupures, et cela conduit naturellement à un déficit de certains candidats blazars par rapport à [Massaro et al., 2013]. D'autre part, l'utilisation du paramètre Rflux comme paramètre discriminant conduit à éliminer certains candidats de [Massaro et al., 2013] qui ont par ailleurs des caractéristiques de type blazar sur les 3 couleurs. Une autre raison est liée à la correction de l'extinction infrarouge, qui est de fait légèrement différente entre ce qui est fait par [Massaro et al., 2013] et ce que nous faisons dans ce travail (voir section 11.5.2). Aucun détail n'est donné sur l'implémentation de cette correction dans [Massaro et al., 2013], mais les valeurs des couleurs fournies après correction de l'extinction diffèrent légèrement des nôtres. Cela conduit, selon les cas, à l'élimination de certains candidats pour [Massaro et al., 2013] ou pour nous. Finalement, nous faisons le constat qu'un certain nombre de sources que nous considérons comme candidats blazars ne sont pas des candidats de [Massaro et al., 2013], et ce malgré des valeurs de ζ montrant une

forte compatibilité avec des caractéristiques des blazars et une très faible probabilité d'association à une source du fond. Nous interprétons cette dernière différence comme un bénéfice lié à notre méthode qui ne fait pas d'hypothèse sur la géométrie de la distribution des blazars dans l'espace des paramètres. La révision récente par [Massaro et al., 2014] du statut d'un certain nombre de ces candidats à partir d'une étude basée uniquement sur les couleurs c_1 et c_2 et des données radio du *NRAO VLA SKY SURVEY (NVSS)* va dans notre sens.

Nous fournissons ci-dessous et à titre d'illustration, des exemples de différences d'associations pour quelques sources non-identifiées entre cette étude et celle de [Massaro et al., 2013] :

- **2FGL J0004.2+2208**
  - *Wise J000429.68+215736.2* de classe C est traitée dans le régime IV. Elle possède une couleur c_1 en queue de distribution et des paramètres c_2, c_3 et RFlux totalement en accord avec le gros des la distribution des blazars. La valeur de c_1 peut expliquer pourquoi [Massaro et al., 2013] ne la sélectionne pas.
- **2FGL J0039.1+4331**
  - *Wise J003858.27+432947.0* de classe C^m a une valeur $\zeta \approx 0{,}57$ et se trouve légèrement sous la valeur seuil du régime IV.
- **2FGL J0103.8+1324**
  - *Wise J010345.73+132345.4* de classe A est traitée dans le régime I. Ses paramètres de couleurs et le paramètre RFlux sont dans le cœur de la distribution des blazars. On notera qu'elle est obtenue dans l'étude de [Massaro et al., 2014], basée sur les deux couleurs c_1 et c_2 et en utilisant les données radio du *NVSS* [Condon et al., 1998].
- **2FGL 0158.4+0107**
  - *Wise J015910.05+010514.7* de classe C^m n'a pas été sélectionnée car elle possède un indice de qualité photométrique « P » dans le troisième filtre.
  - *Wise J015836.25+010632.1* de classe C^m a une valeur $\zeta \approx 0{,}58$ et se trouve légèrement sous la valeur seuil du régime IV.
  - *Wise J015757.45+011547.8* de classe C^m a une valeur $\zeta \approx 0{,}57$ et se trouve légèrement sous la valeur seuil du régime IV.
- **2FGL J0858.3-4333**
  - *Wise J085839.22-432642.7* de classe C^m a une valeur $\zeta \approx 0{,}53$ et se trouve sous la valeur seuil du régime IV. Les deux premières couleurs de cette source utilisées par [Massaro et al., 2013] sont $c_1 = 0.74$ et $c_2 = 1.82$, et celles déterminées après correction de l'extinction dans notre étude (section 11.5.2) sont $c_1 = 0.97$ et $c_2 = 1.78$. La valeur de c_1 explique le fait que cette source ne soit pas sélectionnée comme un candidat blazar par notre étude. Une différence dans l'application de la correction de l'extinction infrarouge est à l'origine du désaccord.
- **2FGL J0928.8-3530**
  - *Wise J092849.83-352948.9* de classe B est traitée dans le régime I. Ses paramètres de couleurs et le paramètre RFlux sont dans le cœur de la distribution des blazars. On notera qu'elle est obtenue dans l'étude de [Massaro et al., 2014], basée sur les deux couleurs c_1 et c_2 et en utilisant les données radio du *NVSS*.
- **2FGL J1518.4-5233**
  - *Wise J151807.34-523431.5* de classe A est traitée dans le régime II. Ses paramètres c_1, c_2 et RFlux sont dans le cœur de la distribution des blazars. Néanmoins, elle possède une couleur c_3 en marge de la distribution ($c_3 \approx 1.1$) ce qui explique pourquoi [Massaro et al., 2013] ne la retient pas.
- **2FGL J1641.8-5319**

  - *Wise J164059.60-532258.5* de classe C est traitée dans le régime III. Ses paramètres c_1, c_2, c_3 et RFlux sont dans le cœur de la distribution des blazars.

— **2FGL J1742.5-3323**

  - *Wise J174201.11-332607.9* de classe B^m a une valeur $\zeta \approx -0{,}12$ et est clairement considérée comme du fond IR. Ceci est dû à la grande valeur de Rflux $\approx 2{,}3$ par rapport à la distribution de ce paramètre pour les blazars. On notera que le flux γ intégré au-delà de 100 GeV de la source 2FGL J0858.3-4333 est compatible avec le cœur de la distribution du flux γ des blazars du catalogue 2FGL

— **2FGL J1745.6+0203** possède une contrepartie commune à l'étude de [Massaro et al., 2013] (classe B^m) et à celle présentée ici (de classe D). Ici, la contrepartie de classe A^m obtenue dans [Massaro et al., 2013] n'est pas sélectionnée.

  - *Wise J174507.82+015442.5* de classe A^m a une valeur $\zeta = 0{,}6216$ et elle est traitée dans le régime IV. Néanmoins, le taux de contamination associé est de $\eta \approx 0{,}59$, ce qui l'exclut de la sélection.

— **2FGL J1759.2-3853**

  - *Wise J175903.29-384739.5* de classe A^m a une valeur $\zeta \approx -0{,}15$ et est clairement considérée comme du fond IR. Ceci est dû à la grande valeur de Rflux $\approx 2{,}17$ par rapport à la distribution de ce paramètre pour les blazars. On notera que le flux γ intégré au-delà de 100 GeV de la source 2FGL J1759.2-3853 est compatible avec le cœur de la distribution du flux γ des blazars du catalogue 2FGL.

— **2FGL J1808.3-3356**

  - *Wise J180825.25-335615.1* de classe B est traitée dans le régime II. Ses paramètres c_1, c_2, c_3 et RFlux sont dans le cœur de la distribution des blazars.

— **2FGL J1904.8-0705**

  - *Wise J1904444.57-070740.1* de classe C^m n'a pas été sélectionnée car elle possède un indice de qualité photométrique « p » dans le troisième filtre.

— **2FGL J1950.3+1223**

  - *Wise J195014.51+123119.6* de classe A est traitée dans le régime I. Ses paramètres c_1, c_2, c_3 et RFlux sont dans le cœur de la distribution des blazars.

— **2FGL J2114.1+5440** est associée par notre méthode avec une source de classe D, mais [Massaro et al., 2013] en propose une différente.

  - *Wise J211508.92+544815.7* de classe C^m n'a pas été sélectionnée car elle possède un indice de qualité photométrique « d » dans le deuxième filtre.

	« extragalactique » cette étude	« extragalactique » [Massaro et al., 2013]	« anti-centre galactique » cette étude	« anti-centre galactique » [Massaro et al., 2013]	« centre galactique » cette étude	« centre galactique » [Massaro et al., 2013]
N_s	39	49	9	6	6	6
$N_c^{\rm tot}$	44	61	11	7	6	6
$N_s^{\rm tot}(\cap)$	30		3		1	
$\widetilde{N}_c^{\rm tot}$	2	×	2	×	0	×
$N_b(\cup)$	37	48	6	4	1	1
$N_b(\cap)$	27	34	3	2	1	1

TABLE 11.4: Résumé des résultats obtenus pour les sources non-identifiées ne possédant pas de problème de reconstruction pour l'étude présentée dans ce chapitre et celle de [Massaro et al., 2013]. Les rangées N_s et $N_c^{\rm tot}$ correspondent respectivement au nombre de sources non-identifiées pour lesquelles au moins une contrepartie infrarouge de type blazar est trouvée et au nombre total de contreparties infrarouges de type blazar obtenues. La rangée « $N_s^{\rm tot}(\cap)$ » représente le nombre de contreparties de type blazar commun aux deux études. $\widetilde{N}_c^{\rm tot}$ est le nombre de sources non-identifiées pour lesquelles au moins une contrepartie de type blazar est trouvée, mais qui possède au moins une contrepartie de type blazar dans leur environnement proche aussi bien classée que celles obtenues dans leur boîte d'erreur. Le nombre « $N_b(\cup)$ » (« $N_b(\cap)$ ») correspond au nombre de candidats blazars en prenant l'union (l'intersection) des propositions de [Doert et Errando, 2013], [Mirabal et al., 2012], et de l'étude du chapitre 10, pour lesquels au moins un contrepartie de type blazar est obtenue.

	« extragalactique » cette étude	« extragalactique » [Massaro et al., 2013]	« anti-centre galactique » cette étude	« anti-centre galactique » [Massaro et al., 2013]	« centre galactique » cette étude	« centre galactique » [Massaro et al., 2013]
N_s	16	17	5	10	29	21
$N_s^{\rm tot}$	19	22	7	21	43	28
$N_s^{\rm tot}(\cap)$	11		1		8	
$\widetilde{N}_c^{\rm tot}$	3	×	2	×	10	×
$N_b^{\rm tot}(\cup)$	12	11	3	4	11	8
$N_b^{\rm tot}(\cap)$	7	8	0	1	6	6

TABLE 11.5: Résumé des résultats obtenus pour les sources non-identifiées possédant un problème de reconstruction pour l'étude présentée dans ce chapitre et celle de [Massaro et al., 2013]. Les notations sont les mêmes que celle de la Table 11.4.

TABLE 11.6: Liste des contreparties infrarouges de type blazar des sources non-identifiées ne présentant pas de problème de reconstruction. On trouvera la définition des colonnes au début de la section 11.7.

Nom 2FGL	N_c/N_W	Wise Id	RA (°)	DEC (°)	δ ($''$)	SNR_1	SNR_2	SNR_3	SNR_4	c_1	c_2	c_3	RFlux	Ccf_W	Var_W	ζ_{MLP}	η	Classe	$\tilde{N}_c/\tilde{N}_W$
Sources non-identifiées « extragalactique » (201)																			
J0004.2+2208$^{\sharp\dagger\star,0}$	1/5	J000429.68+215736.2$^{\times}$	1.1237	21.9601	674.5	38.2	41.7	32.5	13.2	1.51	3.11	2.34	-0.12	0000	×	0.6440	0.0587	C	0/9
J0031.0+0724$^{\sharp\dagger\star,0}$	0/3																		0/3
J0032.7-5521$^{\sharp\dagger\star,0}$	0/1																		0/1
J0039.1+4331^{†,1}	0/10																		0/8
J0048.8-6347$^{\sharp\dagger\star,0}$	0/7																		0/3
J0102.2+0943$^{\sharp\dagger\star,0}$	0/3																		0/7
J0103.8+1324$^{\sharp\dagger\star,0}$	1/4	J010345.73+132345.4$^{\times}$	15.9406	13.3960	48.0	40.5	39.1	12.0	3.8	0.68	2.02	2.26	-0.48	0000	×	0.8561	0.0026	A	0/5
J0106.5+4854^{0}	0/0																		0/4
J0116.6-6153$^{\sharp\dagger\star,1}$	1/7	J011619.59-615343.5°	19.0816	-61.8954	160.7	40.6	40.8	21.0	4.4	0.85	2.34	1.85	-0.40	0000	×	0.9158	0.0046	A	0/4
J0129.4+2618^{†,0}	0/30																		0/78
J0133.4-4408$^{\sharp\dagger\star,2}$	1/6	J013306.35-441421.3°	23.2765	-44.2393	416.8	43.4	44.6	22.6	5.0	0.83	2.25	1.92	-0.29	0000	×	0.9449	0.0040	A	0/10
J0143.6-5844$^{\sharp\dagger\star,1}$	1/5	J014347.39-584551.3°	25.9475	-58.7643	90.5	45.3	45.6	19.7	3.8	0.68	1.93	1.89	-0.42	0000	×	0.9931	0.0033	A	0/1
J0158.4+0107$^{\sharp\dagger\star,3}$	0/28																		0/15
J0158.6+8558$^{\sharp\dagger\star,2}$	1/20	J014935.30+860115.3°	27.3971	86.0209	603.4	44.1	47.0	27.6	6.5	0.70	2.31	2.00	-0.24	0000	×	0.9389	0.0132	B	1/20 (C)
J0200.4-4105$^{\sharp\dagger\star,0}$	0/4																		1/3 (B)
J0221.2+2516^{†*,0}	0/13																		0/14
J0226.1+0943$^{\sharp\dagger\star,0}$	0/4																		1/7 (C)
J0227.7+2249$^{\sharp\dagger\star,1}$	1/4	J022744.35+224834.3°	36.9348	22.8095	73.2	46.5	50.3	41.2	15.1	0.95	2.58	2.09	0.05	0000	○	0.8120	0.0086	A	0/3
J0239.5+1324$^{\sharp\dagger\star,0}$	0/4																		0/6
J0251.0+2557^{†,0}	0/3																		0/5
J0305.0-1602$^{\sharp\dagger\star,0}$	0/5																		0/4
J0312.8+2013$^{\sharp\dagger\star,0}$	0/3																		0/3
J0316.1-6434$^{\sharp\dagger\star,1}$	1/5	J031614.31-643731.4°	49.0596	-64.6254	210.4	44.0	45.3	20.6	4.3	0.74	2.09	1.84	-0.41	0000	×	1.0149	0.0033	A	0/6
J0318.0+0255^{0}	0/5																		0/2
J0336.0+7504^{0}	0/2																		0/4
J0338.2+1306$^{\sharp\dagger\star,0}$	0/1																		0/2
J0340.7-2421$^{\sharp\dagger,0}$	0/7																		0/5
J0345.2-2356$^{\sharp\dagger\star,0}$	0/6																		0/4
J0409.5+0509$^{\sharp\dagger\star,0}$	0/11																		0/9
J0409.8-0357$^{\sharp\dagger\star,1}$	1/3	J040946.57-040003.4°	62.4441	-4.0010	144.1	44.5	49.7	32.3	9.6	0.89	2.36	1.94	0.03	0000	×	0.8591	0.0020	A	0/6
J0414.9-0855^{†*,1}	0/6																		0/9
J0416.0-4355$^{\sharp\dagger\star,1}$	1/16	J041605.81-435514.6°	64.0242	-43.9207	47.2	44.2	46.3	36.3	14.3	1.11	2.89	2.44	-0.18	0000	×	0.7954	0.0344	B	0/18
J0418.9+6636$^{\sharp,0}$	0/3																		0/3
J0420.9-3743$^{\sharp\dagger\star,0}$	0/4																		0/5
J0427.2-6705$^{\sharp\dagger,0}$	0/30																		0/28
J0438.0-7331$^{\sharp\dagger\star,0}$	0/7																		0/6
J0439.8-1858$^{\sharp\dagger\star,0}$	0/0																		0/0
J0515.0-4411^{†*,0}	1/49	J051607.99-441827.6$^{\times}$	79.0333	-44.3077	819.8	47.0	51.9	49.2	13.8	1.05	2.18	1.73	0.01	0000	×	0.7865	0.2349	D	1/30 (D)

... la table continue sur la page suivante...

... suite de la page précédente...

Nom 2FGL	N_c/N_W	Wise Id	RA (°)	DEC (°)	δ (″)	SNR_1	SNR_2	SNR_3	SNR_4	c_1	c_2	c_3	RFlux	Ccf_W	Var_W	$\hat{\varsigma}_{MLP}$	η	Classe	$\tilde{N}_c/\tilde{N}_W$
J0523.3-2530$^{\sharp,0}$	0/2																		0/1
J0533.9+6759^{0}	0/8																		0/2
J0539.3-0323^{†,*,0}	0/118																		0/131
J0540.1-7554$^{\sharp,\star,0}$	0/18																		1/23 (D)
J0545.6+6018^{†,0}	0/6																		1/3 (B)
J0555.9-4348$^{\sharp,\star,2}$	2/25	J055618.74-435146.1°	89.0781	-43.8628	338.1	43.6	47.1	29.8	8.5	0.912	492.14	-0.34	0000	$\times$	0.7219	0.1199	D	0/18
		J055518.54-435224.0$^{\times}$	88.8273	-43.8734	460.7	32.9	44.6	24.8	4.4	1.132	221.70	-0.44	0000	$\times$	0.7649	0.1199	D	
J0600.8-1949$^{\sharp,0}$	0/15																		1/12 (C)
J0602.7-4011$^{\sharp,\star,1}$	1/16	J060237.10-401453.2°	90.6546	-40.2481	239.5	46.2	46.9	42.6	18.4	0.972	472.35	-0.01	0000	$\times$	0.6575	0.1878	D	1/11 (A)
J0644.6+6034$^{\sharp,\star,1}$	0/4																		0/1
J0719.2-5000$^{\sharp,0}$	0/8																		0/5
J0723.9+2901$^{\sharp,\star,1}$	1/5	J072354.83+285929.9°	110.9785	28.9917	153.4	26.5	38.9	26.0	10.8	1.142	892.40	-0.20	0000	$\times$	0.7814	0.0240	B	0/5
J0737.5-8246$^{\sharp,\star,0}$	0/10																		0/6
J0745.5+7910$^{\sharp,0}$	0/7																		0/8
J0746.0-0222^{†,*,1}	1/15	J074627.03-022549.3°	116.6126	-2.4304	391.9	42.5	44.0	17.8	3.7	0.672	091.99	-0.27	hh00	$\times$	1.0021	0.0099	A	1/8 (B)
J0756.3-6433^{†,*,1}	1/4	J075624.60-643030.6°	119.1025	-64.5085	200.1	44.1	46.1	20.0	6.2	0.892	212.11	-0.56	0000	$\circ$	0.7798	0.0192	B	0/4
J0802.6-0940$^{\sharp,\star,0}$	0/10																		1/7 (B)
J0802.7-5615^{0}	0/38																		0/15
J0803.2-0339$^{\sharp,\star,0}$	0/2																		0/2
J0807.0-6511^{†,*,1}	0/59																		0/115
J0843.6+6715^{†,0}	0/1																		0/1
J0844.9+6214$^{\sharp,\star,1}$	1/22	J084406.81+621458.6°	131.0284	62.2496	337.9	46.7	45.2	25.5	8.4	0.702	072.23	-0.12	0000	$\times$	0.9236	0.0145	B	0/18
J0846.0+2820$^{\sharp,\star,0}$	1/27	J084529.01+282057.6$^{\times}$	131.3709	28.3494	444.2	34.6	32.1	20.0	6.6	1.333	052.30	-0.41	0000	$\times$	0.6356	0.3169	D	0/30
J0859.4-2532$^{\sharp,\star,0}$	0/28																		0/18
J0900.9+6736$^{\sharp,\star,1}$	0/13																		0/10
J0928.8-3530^{†,*,0}	1/3	J092849.83-352948.9$^{\times}$	142.2076	-35.4969	53.2	40.7	43.2	24.7	8.9	0.962	592.24	-0.23	0000	$\times$	0.7256	0.0144	B	0/3
J0952.7-3717^{†,0}	0/22																		0/25
J0953.6-1504^{0}	0/7																		0/3
J0955.0-3949^{1}	1/12	J095458.30-394655.0°	148.7429	-39.7820	160.4	41.3	42.4	25.6	6.1	0.792	282.03	-0.19	0000	?	0.9150	0.0079	A	1/4 (A)
J1012.1-4236^{†,0}	0/16																		0/20
J1013.6+3434$^{\sharp,\star,1}$	1/12	J101256.54+343648.8°	153.2356	34.6136	515.6	39.2	36.4	16.7	4.2	0.892	622.09	-0.44	0000	$\times$	0.6902	0.1408	D	0/5
J1016.1+5600$^{\sharp,\star,1}$	0/21																		0/27
J1016.4-4244$^{\sharp,\star,0}$	0/14																		0/5
J1029.5-2022$^{\sharp,\star,1}$	0/8																		0/6
J1032.9-8401$^{\sharp,\star,1}$	0/6																		0/4
J1038.2-2423^{†,1}	2/6	J103748.10-242845.5$^{\times}$	159.4505	-24.4793	488.4	40.0	35.1	7.5	3.1	0.531	792.62	-0.58	0000	$\times$	0.6331	0.0704	C	0/6
		J103754.92-242544.5°	159.4788	-24.4291	301.7	36.3	39.0	31.4	12.1	1.233	262.58	-0.22	0000	$\times$	0.6724	0.0704	C	
J1115.0-0701$^{\sharp,\star,0}$	0/2																		0/1
J1120.0-2204^{0}	0/2																		0/2
J1123.3-2527$^{\sharp,0}$	1/29	J112325.38-252857.0$^{\times}$	170.8558	-25.4825	83.5	47.1	49.1	45.2	23.1	0.842	482.49	0.24	0000	$\times$	0.7011	0.3404	D	0/17
J1129.0-0532$^{\sharp,\star,0}$	0/2																		0/4

... la table continue sur la page suivante...

... suite de la page précédente...

Nom 2FGL	N_c/N_W	Wise Id	RA (°)	DEC (°)	δ ($''$)	SNR_1	SNR_2	SNR_3	SNR_4	c_1	c_2	c_3	RFlux	Ccf_W	Var_W	ζ_{MLP}	η	Classe	$\widetilde{N}_c/\widetilde{N}_W$
J1129.5+3758$^{\sharp\dagger\star,0}$	0/3																		0/5
J1156.7-0751$^{\sharp\dagger,0}$	1/7	J115655.55-075450.3$^{\times}$	179.2315	-7.9140	275.4	43.8	35.3	12.0	3.1	1.01	2.15	2.04	-0.41	0000	×	0.7373	0.0336	B	0/13
J1207.3-5055$^{\sharp\dagger,2}$	0/51																		1/31 (C)
J1214.1-4410$^{\sharp\dagger,0}$	0/9																		1/11 (D)
J1221.4-0633$^{\sharp\dagger\star,0}$	0/2																		0/4
J1223.3+7954$^{\sharp\dagger\star,0}$	1/18	J122358.17+795327.8$^{\times}$	185.9924	79.8911	115.4	41.3	38.5	10.4	3.0	0.47	1.90	2.23	-0.63	0000	×	0.6944	0.2113	D	0/14
J1226.0+2953^{0}	0/1																		0/2
J1227.7-4853^{0}	0/5																		0/4
J1231.3-5112^{0}	0/20																		0/22
J1249.5-2811$^{\sharp\dagger\star,0}$	0/6																		0/2
J1254.2-2203$^{\sharp\dagger\star,1}$	1/9	J125422.47-220413.6°	193.5937	-22.0705	100.7	42.9	36.8	15.4	3.1	0.67	2.31	1.77	-0.51	0000	○	0.8731	0.0060	A	0/10
J1259.8-3749$^{\sharp\dagger\star,1}$	1/8	J125949.80-374858.1°	194.9575	-37.8161	61.1	43.0	38.7	14.3	3.3	0.71	2.10	1.99	-0.54	0000	○	0.9280	0.0053	A	0/10
J1305.0+1152$^{\sharp\dagger,0}$	0/12																		0/22
J1311.7-3429^{0}	0/1																		0/0
J1312.9-2351$^{\sharp\dagger\star,0}$	0/23																		0/9
J1328.5-4728$^{\sharp\dagger\star,0}$	1/3	J132840.61-472749.2$^{\times}$	202.1692	-47.4637	58.0	37.8	37.6	15.4	5.1	0.62	2.05	2.36	-0.51	0000	○	0.7826	0.0144	B	0/3
J1335.3-4058$^{\sharp\dagger\star,0}$	0/17																		1/11 (D)
J1340.5-0412$^{\sharp\dagger\star,1}$	0/12																		0/10
J1346.0-2605^{†,1}	0/15																		0/20
J1347.0-2956$^{\sharp\dagger\star,1}$	1/8	J134706.89-295842.3°	206.7787	-29.9784	134.1	44.4	45.5	20.5	4.8	0.79	2.10	1.91	-0.27	0000	○	1.0032	0.0053	A	0/9
J1400.2-2412^{†*,0}	0/9																		0/10
J1400.7-1438^{†,0}	0/6																		0/9
J1410.4+7411$^{\sharp\dagger\star,0}$	0/4																		0/0
J1417.5-4404$^{\sharp\dagger\star,0}$	0/14																		0/24
J1417.7-5028$^{\sharp\dagger\star,0}$	0/16																		0/28
J1423.9-7842$^{\sharp\dagger\star,0}$	0/8																		0/13
J1424.2-1752$^{\sharp\dagger\star,0}$	1/3	J142412.36-175008.3$^{\times}$	216.0515	-17.8356	158.9	46.5	49.1	21.6	3.1	0.32	1.72	1.49	-0.10	h000	×	0.6439	0.0352	B	0/3
J1446.8-4701^{0}	0/3																		0/14
J1458.5-2121$^{\sharp\dagger\star,0}$	0/10																		0/8
J1502.1+5548$^{\sharp\dagger\star,0}$	0/11																		0/13
J1502.4+4804$^{\sharp\dagger\star,0}$	0/20																		2/13 (A)
J1513.5-2546^{†,0}	0/2																		0/9
J1513.9-2256$^{\sharp\dagger\star,0}$	0/4																		0/4
J1517.2+3645$^{\sharp\dagger\star,1}$	2/18	J151752.12+364125.5$^{\times}$	229.4672	36.6904	508.8	42.3	42.0	32.2	9.8	0.95	3.05	2.19	-0.48	0000	×	0.6287	0.2113	D	0/28
		J151649.26+365022.9°	229.2053	36.8397	421.0	44.4	48.2	32.9	9.5	0.95	2.63	2.07	-0.39	0000	○	0.7137	0.0863	C	
J1539.2-3325^{0}	0/52																		0/79
J1543.7-0241^{0}	0/16																		1/19 (D)
J1544.1-2554^{†,0}	0/54																		0/56
J1544.5-1126$^{\sharp\dagger,0}$	0/3																		0/4
J1548.3+1453$^{\sharp\dagger\star,0}$	0/3																		0/1
J1553.5-0324$^{\sharp\dagger\star,0}$	0/6																		0/5

... la table continue sur la page suivante...

... suite de la page précédente...

Nom 2FGL	N_c/N_W	Wise Id	RA (°)	DEC (°)	δ (″)	SNR_1	SNR_2	SNR_3	SNR_4	c_1	c_2	c_3	RFlux	Ccf_W	Var_W	ζ_{MLP}	η	Classe	$\tilde{N}_c/\tilde{N}_W$
J1612.0+1403[‡†],1	0/13																		0/22
J1614.8+4703[‡†⋆],4	2/48	J161434.67+470420.1°	243.6445	47.0723	178.1	46.7	53.3	65.5	33.4	1.083	122.19		0.24	0000	×	0.8148	0.1032	D	0/33
		J161541.22+471111.8°	243.9218	47.1866	701.4	47.8	51.8	38.4	12.2	0.742	282.28		-0.20	h0o0	×	0.8152	0.1032	D	
J1617.5-2657[‡†],0	0/70																		0/45
J1622.8-0314[‡†⋆],1	0/6																		0/5
J1623.2+4328[‡†],0	0/34																		0/22
J1627.8+3219[†⋆],1	0/7																		0/2
J1631.0-1050[†],0	0/50																		1/43 (B)
J1631.6-2819[0]	0/286																		0/163
J1636.6-0841[‡†⋆],0	0/40																		0/36
J1646.7-1333[‡†⋆],0	0/5																		0/1
J1647.0+4351[‡†⋆],1	0/16																		0/14
J1653.6-0159[0]	0/1																		0/0
J1659.2-0142[‡†⋆],0	0/6																		0/11
J1704.3+1235[‡†⋆],0	0/8																		0/0
J1704.6-0529[‡†⋆],0	0/5																		0/6
J1709.0-0821[0]	0/20																		0/36
J1710.0-0323[0]	0/77																		0/43
J1721.0+0711[†⋆],0	0/4																		0/8
J1722.5-0420[0]	0/9																		0/9
J1730.6-0353[1]	1/31	J173052.86-035247.1°	262.7203	-3.8798	223.9	38.5	46.2	39.6	15.8	1.222	922.28		-0.02	0000	?	0.7973	0.0666	C	0/9
J1730.8+5427[‡†⋆],0	0/39																		0/56
J1738.9+8716[‡†⋆],0	0/5																		0/6
J1741.0+1347[‡†⋆],0	0/7																		0/7
J1741.1-6750[‡†],0	0/8																		0/8
J1744.1-7620[0]	0/4																		0/7
J1745.6+0203[†],2	1/53	J174526.95+020532.7°	266.3623	2.0924	208.9	46.8	49.3	40.5	17.4	0.992	572.27		0.10	0000	×	0.7393	0.2541	D	0/32
J1747.6+0324[‡†⋆],0	0/9																		0/20
J1748.8+3418[‡†⋆],0	0/12																		0/18
J1757.5-6028[‡†],0	0/0																		0/5
J1802.8-6706[†],0	0/11																		1/12 (D)
J1805.8+0612[0]	0/7																		0/4
J1816.5+4511[4],0	0/0																		0/6
J1821.8+0830[‡†],0	0/10																		0/18
J1824.5+1013[†],0	0/9																		0/15
J1828.7+3231[†],0	0/16																		0/12
J1832.2-6502[0]	0/9																		0/5
J1842.3+2740[†],1	0/13																		0/4
J1842.3-5839[‡†⋆],0	0/3																		0/5
J1902.7-7053[‡†⋆],0	0/6																		0/4
J1917.0-3027[‡†⋆],0	0/4																		0/3

... la table continue sur la page suivante...

... suite de la page précédente...

Nom 2FGL	N_c/N_W	Wise Id	RA (°)	DEC (°)	δ ($''$)	SNR_1	SNR_2	SNR_3	SNR_4	c_1	c_2	c_3	RFlux	Ccf_W	Var_W	ζ_{MLP}	η	Classe	$\widetilde{N}_c/\widetilde{N}_W$
J1919.5-7324$^{\sharp\dagger\star,0}$	0/17																		0/19
J1924.9-1036^{†,1}	1/12	J192501.63-104316.3°	291.2568	-10.7212	410.6	31.2	32.9	28.1	13.4	1.26	3.25	2.66	-0.19	0000	×	0.6561	0.1408	D	0/4
J1936.5-0855$^{\sharp\dagger,0}$	0/18																		2/33 (D)
J1942.9-3528$^{\sharp\dagger\star,0}$	0/13																		0/38
J1946.4-5402^{0}	0/5																		0/2
J1946.7-1118^{†*,0}	1/31	J194617.40-111447.6$^{\times}$	296.5725	-11.2466	434.4	44.2	35.0	13.4	3.5	0.32	2.06	1.98	-0.25	h000	×	0.7542	0.1486	D	0/31
J1947.8-0739$^{\sharp\dagger,0}$	0/5																		0/11
J1949.4-1457^{†,0}	0/59																		0/13
J2002.8-2150$^{\sharp\dagger\star,0}$	0/7																		0/10
J2004.6+7004$^{\sharp\dagger\star,2}$	0/33																		0/30
J2006.5-2256$^{\sharp\dagger\star,0}$	0/25																		0/5
J2017.5-1618$^{\sharp\dagger,0}$	0/2																		0/0
J2021.5+0632$^{\sharp\dagger,2}$	2/7	J202155.45+062913.7°	305.4811	6.4871	377.2	44.8	47.8	22.4	6.8	0.81	2.09	1.78	-0.02	0000	?	0.9892	0.0046	A	0/7
		J202154.66+062908.7°	305.4778	6.4858	370.0	36.7	37.8	22.2	11.9	0.97	2.67	2.47	-0.14	0000	?	0.7349	0.0336	B	
J2034.7-4201$^{\sharp\dagger\star,0}$	0/3																		0/0
J2042.8-7317$^{\sharp\dagger\star,0}$	0/6																		0/6
J2044.4-4757^{0}	0/5																		0/12
J2046.2-4259$^{\sharp\dagger\star,0}$	0/7																		0/4
J2103.5-1112$^{\sharp\dagger\star,0}$	0/0																		0/1
J2107.6+2506$^{\sharp\dagger,0}$	0/15																		1/11 (C)
J2112.5-3042^{0}	0/1																		0/0
J2115.4+1213$^{\sharp\dagger\star,0}$	0/4																		0/5
J2125.0-4632$^{\sharp\dagger\star,0}$	0/9																		0/10
J2131.0-5417$^{\sharp\dagger\star,0}$	0/9																		0/12
J2132.5+2605$^{\sharp\dagger\star,0}$	0/14																		0/9
J2133.5-6431^{0}	0/4																		0/1
J2133.9+6645$^{\sharp\dagger\star,1}$	1/80	J213349.21+664704.3°	323.4551	66.7845	133.8	42.6	44.3	20.9	5.0	0.75	2.11	1.86	-0.43	0o00	∘	1.0050	0.0529	C	1/159 (D)
J2134.6-2130$^{\sharp\dagger\star,1}$	0/1																		0/4
J2200.1-6931$^{\sharp\dagger\star,0}$	0/6																		0/0
J2212.6+0702^{0}	0/5																		0/5
J2213.7-4754$^{\sharp\dagger\star,0}$	0/7																		0/1
J2227.8+0051$^{\sharp\dagger,0}$	0/11																		1/12 (C)
J2228.6-1633$^{\sharp,0}$	0/4																		0/5
J2246.3+1549$^{\sharp\dagger,0}$	0/3																		0/4
J2251.8+4211$^{\sharp\dagger\star,0}$	0/18																		1/39 (D)
J2257.9-3646$^{\sharp\dagger\star,0}$	0/3																		1/7 (C)
J2259.0-8254^{†,0}	0/8																		1/16 (B)
J2300.0-3553$^{\sharp\dagger\star,1}$	0/4																		1/7 (C)
J2319.3-3830$^{\sharp\dagger\star,1}$	1/8	J232000.11-383511.4°	350.0005	-38.5865	541.1	33.8	32.2	17.4	7.2	1.13	3.06	2.60	-0.56	0000	×	0.6339	0.0939	C	0/4
J2332.5-5535^{0}	0/10																		0/5
J2339.6-0532^{0}	0/2																		0/4

... la table continue sur la page suivante...

... suite de la page précédente...

Nom 2FGL	N_c/N_W	Wise Id	RA (°)	DEC (°)	δ ($''$)	SNR_1	SNR_2	SNR_3	SNR_4	c_1	c_2	c_3	RFlux	Ccf_W	Var_W	ζ_{MLP}	η	Classe	$\widetilde{N}_c/\widetilde{N}_W$
J2347.2+0707$^{\sharp\dagger\star,0}$	0/2																		0/0
J2351.6-7558^{†*,0}	0/3																		0/6
J2354.2-6615$^{\sharp\dagger\star,0}$	0/10																		0/11
J2358.4-1811$^{\sharp\dagger\star,1}$	0/3																		0/4
Sources non-identifiées « anti-centre galactique » (56)																			
J0002.7+6220^{0}	0/185																		0/185
J0212.1+5318^{0}	0/9																		0/7
J0224.0+6204^{0}	0/14																		0/16
J0237.9+5238^{*,0}	1/9	J023749.24+523931.8$^{\times}$	39.4552	52.6589	117.6	30.2	26.4	14.4	4.3	1.03	3.02	2.27	-0.69	hh00	×	0.6222	0.0567	C	0/4
J0248.5+5131$^{\sharp\star,0}$	1/3	J024834.26+513117.9$^{\times}$	42.1428	51.5216	28.6	33.5	33.5	9.4	3.6	0.91	1.98	2.59	-0.48	0000	?	0.6252	0.0189	B	0/5
J0307.4+4915$^{\sharp\star,0}$	0/3																		0/1
J0332.1+6309$^{\sharp\star,1}$	1/149	J033153.90+630814.1°	52.9746	63.1373	114.5	43.4	46.2	35.8	10.2	0.88	2.35	1.96	-0.20	0000	×	0.7259	0.2590	D	0/108
J0340.5+5307^{0}	0/12																		0/24
J0353.2+5653$^{\sharp,0}$	0/25																		0/86
J0359.5+5410^{0}	0/5																		0/15
J0426.7+5434^{0}	0/106																		0/94
J0431.5+3622$^{\sharp,1}$	0/7																		0/8
J0516.7+2634$^{\sharp\star,0}$	0/0																		0/5
J0524.1+2843$^{\sharp\star,0}$	0/10																		0/8
J0600.9+3839$^{\sharp\star,0}$	0/3																		0/2
J0605.3+3758^{0}	1/9	J060512.91+375149.0$^{\times}$	91.3038	37.8636	435.1	18.0	21.2	8.8	3.1	1.32	2.98	2.40	-0.78	h000	×	0.6340	0.0567	C	0/6
J0608.3+2037^{0}	0/75																		0/83
J0658.4+0633$^{\sharp\star,0}$	1/4	J065809.92+063810.1$^{\times}$	104.5414	6.6362	385.7	40.4	41.2	35.4	11.8	1.36	3.04	1.98	-0.09	0000	×	0.7088	0.0122	B	0/7
J0713.5-0952^{1}	1/397	J071223.28-094536.3°	108.0970	-9.7601	1118.3	37.9	39.7	29.0	14.7	1.25	3.11	2.27	-0.02	0000	×	0.7499	0.2587	D	1/132 (C)
J0725.8-0549$^{\sharp\star,0}$	0/1																		0/3
J0737.1-3235$^{\sharp,0}$	1/6	J073738.91-323256.2$^{\times}$	114.4122	-32.5490	385.2	38.6	38.9	34.3	14.1	1.03	3.09	2.59	-0.19	0000	?	0.6605	0.0378	B	0/8
J0744.1-2523^{2}	2/14	J074406.14-252154.1$^{\times}$	116.0256	-25.3650	142.9	44.5	48.8	66.1	34.8	1.22	2.36	2.07	0.82	0h00	×	0.6579	0.0882	C	0/14
		J074359.96-252220.5$^{\times}$	115.9999	-25.3724	183.5	40.3	39.7	26.1	12.2	0.88	2.63	2.62	-0.14	hh00	×	0.6550	0.0882	C	
J0758.8-1448$^{\sharp\star,0}$	0/18																		0/19
J0838.8-2828$^{\sharp,1}$	2/12	J083842.77-282830.9°	129.6782	-28.4753	112.8	33.0	28.2	13.6	4.1	1.05	2.73	2.24	-0.68	0000	×	0.6550	0.0756	C	1/5 (A)
		J083833.01-282410.7$^{\times}$	129.6376	-28.4030	356.5	15.9	18.9	9.5	3.2	1.38	3.19	2.52	-0.87	hh00	×	0.6246	0.0756	C	
J0841.3-3556$^{\sharp\star,1}$	0/7																		0/13
J2231.0+6512$^{\sharp,0}$	0/320																		0/265
J2249.1+5758^{0}	0/184																		0/172
Sources non-identifiées « centre galactique » (27)																			
J0854.7-4501^{0}	0/7																		0/10
J0858.3-4333^{1}	0/68																		0/53
J0934.0-6231^{0}	0/10																		0/9
J1033.5-5032^{*,0}	0/9																		0/5
J1036.1-6722^{0}	0/6																		0/33
J1056.2-6021^{0}	0/11																		0/10

... la table continue sur la page suivante...

... suite de la page précédente...

Nom 2FGL	N_c/N_W	Wise Id	RA (°)	DEC (°)	δ ($''$)	SNR_1	SNR_2	SNR_3	SNR_4	c_1	c_2	c_3	RFlux	Ccf_W	Var_W	ζ_{MLP}	η	Classe	$\widetilde{N}_c/\widetilde{N}_W$
J1058.7-6621$^{\sharp\star,0}$	0/269																		0/337
J1117.2-5341$^{\sharp\star,0}$	0/48																		0/65
J1208.5-6240^{0}	0/26																		0/55
J1240.6-7151$^{\sharp,0}$	0/30																		0/13
J1306.2-6044^{0}	0/34																		0/27
J1353.5-6640$^{\sharp\star,0}$	0/22																		0/11
J1404.0-5244^{1}	1/46	J140313.11-524839.5°	210.8046	-52.8110	524.0	30.4	33.8	16.7	7.8	1.27	2.97	2.63	-0.39	0000	?	0.8498	0.0178	B	0/67
J1422.3-6841$^{\sharp\star,0}$	0/67																		1/86 (B)
J1507.0-6223$^{\sharp\star,0}$	0/207																		0/170
J1518.4-5233^{0}	1/12	J151807.34-523431.5$^{\times}$	229.5306	-52.5754	184.8	42.5	43.0	27.9	3.6	0.56	2.09	1.08	0.03	0000	$\times$	0.8382	0.0046	A	0/31
J1536.4-4949^{0}	0/0																		0/0
J1601.1-4220$^{\sharp\star,0}$	0/162																		0/109
J1617.3-5336^{0}	0/262																		0/284
J1620.8-4928^{0}	0/1																		0/1
J1622.8-5006^{0}	0/2																		0/1
J1624.1-4040^{0}	0/1																		0/16
J1626.4-4408$^{\sharp,0}$	0/246																		0/161
J1639.8-5145^{0}	0/13																		1/42 (B)
J1641.8-5319^{0}	1/230	J164059.60-532258.5$^{\times}$	250.2484	-53.3829	510.6	10.5	18.7	17.2	7.2	1.15	2.75	2.67	-0.49	hh00	?	0.8269	0.0889	C	0/175
J1643.3-4928^{0}	0/442																		0/442
J1649.2-3004$^{\sharp\star,0}$	0/43																		0/73
J1702.5-5654^{0}	0/37																		0/13
J1730.6-2409^{0}	0/71																		0/118
J1742.5-3323^{1}	0/172																		0/186
J1743.2-2304^{0}	0/251																		1/200 (C)
J1746.5-3238^{0}	0/67																		0/49
J1748.9-3923$^{\sharp\star,0}$	0/36																		0/49
J1753.8-4446$^{\sharp,0}$	0/8																		0/8
J1759.2-3853^{1}	0/26																		0/22
J1759.4-2954$^{\sharp\star,0}$	0/213																		0/242
J1808.3-3356^{0}	1/48	J180825.25-335615.1$^{\times}$	272.1052	-33.9375	41.0	24.8	31.7	13.4	5.5	0.65	2.24	2.29	-0.16	00o0	$\circ$	0.8411	0.0185	B	0/41
J1819.3-1523^{0}	0/114																		0/57
J1820.6-3219^{0}	0/105																		0/110
J1830.9-3132^{0}	0/11																		0/9
J1839.0-0539^{0}	0/2																		0/4
J1901.1+0427^{0}	0/78																		0/107
J1904.8-0705^{1}	0/9																		0/17
J1906.5+0720^{0}	0/29																		0/42
J1950.3+1223^{0}	1/31	J195014.51+123119.6$^{\times}$	297.5605	12.5221	461.3	25.1	29.3	20.7	9.0	1.13	2.83	2.85	-0.28	0000	?	0.8561	0.0060	A	0/30
J2018.0+3626^{0}	0/25																		0/20
J2028.3+3332^{0}	0/37																		0/18
J2030.7+4417^{0}	0/37																		0/54

... la table continue sur la page suivante...

... suite de la page précédente...

Nom 2FGL	N_c/N_W	Wise Id	RA (°)	DEC (°)	δ (″)	SNR_1	SNR_2	SNR_3	SNR_4	c_1	c_2	c_3	RFlux	Ccf_W	Var_W	ζ_{MLP}	η	Classe	$\tilde{N}_c/\tilde{N}_W$
J2034.9+3632[0]	0/55																		0/54
J2041.2+4735[0]	0/112																		0/84
J2107.8+3652[d⋆,0]	0/6																		0/1
J2110.3+3822[d⋆,0]	0/6																		0/10
J2111.3+4605[0]	0/28																		0/13
J2114.1+5440[d⋆,1]	1/276	J211240.12+543747.1[×]	318.1672	54.6298	790.2	8.4	21.0	18.0	6.4	1.913	3.522	2.43	-0.65	0000	?	0.7614	0.2133	D	0/208
J2117.5+3730[0]	0/8																		0/5
J2139.8+4714[0]	0/14																		1/42 (B)

TABLE 11.7: Liste des contreparties infrarouges de type blazar des sources non-identifiées qui peuvent présenter des problèmes de reconstruction. On trouvera la définition des colonnes au début de la section 11.7.

Nom 2FGL	N_c/N_W	Wise Id	RA (°)	DEC (°)	δ ($''$)	SNR_1	SNR_2	SNR_3	SNR_4	c_1	c_2	c_3	RFlux	Ccf_W	Var_W	ζ_{MLP}	η	Classe	$\widetilde{N}_c/\widetilde{N}_W$
Sources non-identifiées « extragalactique » (71)																			
J0014.3-0509$^{\sharp\dagger\star,0}$	0/6																		0/2
J0124.6-2322$^{\sharp\dagger,0}$	0/10																		0/5
J0312.5-0914$^{\sharp\dagger\star,0}$	0/132																		0/35
J0341.8+3148c^{†,3}	1/74	J034153.27+315019.1$^{\times}$	55.4720	31.8387	118.1	42.3	43.4	16.4	8.2	0.79	1.86	2.51	0.12	0000	×	0.7322	0.3548	D	0/90
J0404.0+3843$^{\sharp\dagger\star,0}$	0/372																		0/141
J0428.0-3845$^{\sharp\dagger\star,0}$	0/20																		0/23
J0440.5+2554c^{1}	1/209	J044116.28+260309.2$^{\times}$	70.3179	26.0526	774.0	39.3	33.9	8.2	3.5	0.51	1.77	2.30	-0.32	0000	×	0.8289	0.4492	D	0/52
J0458.4+0654$^{\sharp\dagger\star,0}$	0/10																		0/7
J0534.8-0548c^{†*,0}	1/152	J053449.80-054454.9$^{\times}$	83.7075	-5.7486	231.0	38.5	27.6	7.7	15.5	0.12	1.84	5.09	0.37	000h	×	0.8055	0.3267	D	1/66 (D)
J0534.9-0450c^{0}	0/87																		3/74 (D)
J0538.5-0534c^{0}	0/233																		0/306
J0541.8-0203c^{0}	0/11																		0/35
J0543.2-0120c^{1}	1/65	J054324.78-011545.6°	85.8533	-1.2627	302.9	46.0	48.1	28.3	10.8	0.74	1.65	1.86	0.42	0000	×	0.7461	0.3116	D	0/86
J0547.1+0020c^{0}	0/26																		0/16
J0547.5-0141c^{1}	0/286																		0/530
J0607.5-0618c^{1}	0/49																		2/99 (D)
J0620.8-2556$^{\sharp\dagger\star,1}$	0/19																		1/24 (B)
J0641.2+7315$^{\sharp\star,0}$	0/11																		1/4 (B)
J0753.2+1937$^{\sharp\dagger\star,0}$	0/6																		1/22 (D)
J0923.5+1508$^{\sharp\dagger\star,0}$	0/20																		0/17
J0937.9-1434$^{\sharp\dagger\star,0}$	0/2																		0/6
J1059.9-2051$^{\sharp\dagger\star,1}$	1/8	J110025.72-205333.4°	165.1072	-20.8926	410.2	43.4	45.5	27.6	7.9	0.90	2.46	1.99	-0.15	0000	×	0.8318	0.0172	B	0/9
J1105.4-7622$^{\sharp\dagger,0}$	2/12	J110409.06-762719.3$^{\times}$	166.0378	-76.4554	413.0	47.9	55.4	77.3	82.4	0.70	2.40	2.12	1.83	0000	×	0.7581	0.0575	C	0/10
		J110632.74-762521.0$^{\times}$	166.6364	-76.4225	293.5	46.8	48.3	15.2	6.0	0.50	1.48	2.24	-0.15	0h00	×	0.7196	0.0575	C	
J1208.6-2257$^{\sharp\dagger\star,1}$	1/46	J120816.33-224921.9°	182.0681	-22.8228	561.3	43.2	42.7	24.2	6.1	0.99	2.25	2.00	-0.19	0000	×	0.7530	0.2205	D	1/52 (D)
J1315.6-0730$^{\sharp\dagger\star,2}$	2/4	J131543.62-073659.0°	198.9318	-7.6164	388.5	40.7	41.9	18.3	5.4	0.94	2.39	2.15	-0.42	0000	×	0.6809	0.0469	B	0/2
		J131552.98-073301.9°	198.9708	-7.5505	247.6	44.7	50.7	39.3	13.8	0.87	2.27	2.04	0.08	0000	○	0.8274	0.0086	A	
J1329.5-3448^{†*,0}	0/19																		0/15
J1339.2-2348$^{\sharp\dagger\star,0}$	0/7																		1/8 (A)
J1345.8-3356$^{\sharp\dagger\star,1}$	1/15	J134543.05-335643.3°	206.4294	-33.9454	94.3	42.7	38.3	19.3	5.6	0.83	2.31	2.09	-0.41	0000	×	0.8142	0.0322	B	0/18
J1351.1-2749$^{\sharp\dagger\star,0}$	0/7																		0/7
J1407.4-2948^{†,1}	1/71	J140829.80-300423.5$^{\times}$	212.1242	-30.0732	1245.0	47.0	47.6	38.4	12.3	0.60	2.54	2.08	-0.03	0000	×	0.8449	0.0469	B	0/43
J1506.9+1052$^{\sharp\dagger\star,0}$	0/9																		0/5
J1511.8-0513$^{\sharp\dagger\star,0}$	1/3	J151148.57-051346.8$^{\times}$	227.9524	-5.2297	56.7	42.7	36.9	8.4	3.0	0.67	1.96	2.33	-0.58	0000	○	0.7847	0.0144	B	0/2
J1602.4+2308^{0}	0/5																		0/3
J1620.5-2320c^{†,0}	0/27																		0/49
J1624.2-2124^{1}	0/535																		0/441
J1625.2-0020^{0}	0/2																		0/1

... la table continue sur la page suivante...

... suite de la page précédente...

Nom 2FGL	N_c/N_W	Wise Id	RA (°)	DEC (°)	δ ($''$)	SNR_1	SNR_2	SNR_3	SNR_4	c_1	c_2	c_3	RFlux	Ccf_W	Var_W	ζ_{MLP}	η	Classe	$\widetilde{N}_c/\widetilde{N}_W$
J1630.3+3732^{0}	0/4																		0/6
J1632.6-2328c$^{\sharp\dagger,1}$	1/31	J163306.53-233207.3°	248.2772	-23.5354	462.4	44.8	46.6	35.8	13.2	1.17	2.68	2.43	0.23	00o0	×	0.7016	0.3639	D	0/19
J1645.7-2148c^{0}	0/101																		0/58
J1656.4-0738$^{\sharp\dagger\star,0}$	0/224																		0/166
J1708.4+1003c$^{\sharp\dagger\star,0}$	0/8																		1/3 (B)
J1716.6-0526c^{1}	1/11	J171717.66-052520.1°	259.3236	-5.4223	552.9	40.9	45.6	41.5	21.1	1.31	2.95	2.37	0.15	0000	×	0.7848	0.0527	C	0/13
J1721.5-0718c^{0}	1/31	J172128.94-070955.1$^{\times}$	260.3706	-7.1653	489.5	46.9	47.8	53.7	25.8	0.41	2.39	2.21	0.50	0000	×	0.7894	0.1486	D	0/13
J1724.9-0508c$^{\sharp\dagger\star,0}$	0/13																		0/26
J1727.6+0647$^{\sharp\dagger\star,0}$	0/9																		0/13
J1729.5-0854^{0}	0/36																		0/70
J1749.1+0515$^{\sharp\star,0}$	0/11																		0/13
J1827.6+1149^{†*,0}	0/15																		0/23
J1837.9+3821$^{\sharp\dagger\star,3}$	2/51	J183837.16+381900.5°	279.6549	38.3168	515.2	39.3	39.2	15.3	3.3	0.90	2.14	1.84	-0.62	0000	×	0.9056	0.0337	B	0/44
		J183656.31+382232.8°	279.2346	38.3758	692.7	41.0	42.4	32.6	13.8	1.15	2.97	2.53	-0.38	0000	×	0.7158	0.2445	D	
J1846.6-2519$^{\sharp\dagger,0}$	0/54																		0/76
J1858.3-2218^{0}	0/4																		0/12
J1904.9-3720c^{0}	0/32																		0/28
J1913.8-1237$^{\sharp\dagger,0}$	0/99																		1/50 (D)
J1925.7-7836c$^{\sharp\dagger\star,0}$	0/27																		0/23
J1942.7-8049c$^{\sharp\dagger,0}$	0/19																		1/44 (D)
J1944.3+7325$^{\sharp\dagger,1}$	0/50																		0/38
J1952.6-3252$^{\sharp\dagger\star,0}$	0/350																		0/209
J1955.9-0241$^{\sharp\dagger,0}$	0/16																		0/10
J2006.2-0929^{†,0}	0/66																		0/56
J2006.9-1734^{†,0}	0/11																		0/9
J2009.1-0339^{†,0}	0/26																		0/83
J2009.2-1505^{†,0}	0/4																		0/3
J2017.4-3215$^{\sharp\dagger\star,0}$	0/10																		0/9
J2031.4-1842$^{\sharp\dagger\star,0}$	0/23																		0/25
J2039.8-5620^{0}	0/0																		0/2
J2053.2+1212c$^{\sharp\dagger\star,0}$	0/2																		0/5
J2112.3-4832$^{\sharp\dagger\star,0}$	0/8																		0/5
J2124.0-1513$^{\sharp\dagger\star,1}$	1/13	J212423.63-152558.2°	321.0985	-15.4328	799.6	26.0	23.6	11.2	4.6	1.15	3.15	2.56	-0.61	0000	×	0.6254	0.1526	D	3/19 (B)
J2251.1-4927^{†*,0}	0/5																		0/4
J2343.3-4752$^{\sharp\dagger\star,0}$	0/5																		0/3
J2356.0-5256$^{\sharp\dagger\star,0}$	0/8																		0/7
Sources non-identifiées « anti-centre galactique » (42)																			
J0007.7+6825c^{0}	0/144																		0/134
J0038.8+6259^{0}	0/4																		0/7
J0221.3+6025c^{0}	0/58																		0/88
J0222.7+6820$^{\sharp\star,0}$	0/2																		0/5

... la table continue sur la page suivante...

... suite de la page précédente...

Nom 2FGL	N_c/N_W	Wise Id	RA (°)	DEC (°)	δ ($''$)	SNR_1	SNR_2	SNR_3	SNR_4	c_1	c_2	c_3	RFlux	Ccf_W	Var_W	ζ_{MLP}	η	Classe	$\widetilde{N}_c/\widetilde{N}_W$
J0225.9+6154c$^{\sharp,0}$	0/16																		0/4
J0227.2+6029c^{*,0}	0/53																		0/49
J0233.9+6238c$^{\sharp\star,2}$	0/308																		0/269
J0253.9+5908^{*,0}	0/34																		0/107
J0308.7+5954^{0}	0/484																		0/415
J0330.3+5816c^{0}	2/142	J033111.67+581834.7$^{\times}$	52.7986	58.3097	429.9	29.8	23.3	6.0	4.9	0.46	1.67	3.18	-0.48	0000	×	0.6989	0.4319	D	0/185
		J033108.18+581810.6$^{\times}$	52.7841	58.3030	396.3	25.1	23.6	5.3	3.7	0.83	1.94	2.74	-0.59	0000	×	0.7060	0.4319	D	
J0404.6+5822^{*,0}	0/169																		0/215
J0415.2+5518$^{\sharp\star,0}$	0/88																		0/71
J0423.4+5612$^{\sharp,1}$	1/37	J042430.44+561525.8°	66.1269	56.2572	562.2	24.3	24.4	8.6	4.7	1.27	2.98	2.47	-0.70	0000	×	0.6520	0.2331	D	0/23
J0430.2+3508c^{0}	0/5																		0/13
J0526.6+2248^{*,0}	0/1																		0/6
J0529.3+3821^{0}	0/174																		0/92
J0540.3+3549c^{0}	0/104																		1/150 (C)
J0547.4+3722$^{\sharp,0}$	0/6																		0/10
J0616.6+2425^{2}	0/231																		2/73 (D)
J0621.2+2508^{0}	0/15																		0/42
J0631.7+0428^{1}	0/1019																		0/892
J0634.3+0356c^{2}	0/362																		3/275 (D)
J0637.0+0416c^{4}	0/94																		2/135 (D)
J0641.1+1006c^{0}	0/111																		0/153
J0642.9+0319^{0}	0/49																		0/21
J0647.7+0032^{0}	0/18																		0/108
J0705.3-1043c^{0}	0/850																		0/681
J0708.5-1020c^{*,6}	2/164	J070801.08-102428.9$^{\times}$	107.0045	-10.4080	477.8	39.1	42.4	27.6	7.1	1.03	2.05	1.70	0.14	0000	?	0.6865	0.4988	D	1/173 (D)
		J070753.94-101634.2$^{\times}$	106.9748	-10.2762	594.8	46.2	49.4	37.0	13.9	0.61	2.05	1.58	0.44	0000	×	0.6930	0.4988	D	
J0742.7-3113$^{\sharp,1}$	0/55																		0/123
J0748.5-2204^{1}	0/216																		0/160
J0753.2-1634$^{\sharp,0}$	1/57	J075350.16-164735.4$^{\times}$	118.4590	-16.7932	922.1	19.5	24.1	16.7	5.8	1.59	3.23	2.64	-0.63	0000	×	0.6274	0.3591	D	3/71 (D)
J0758.0-2615c^{0}	0/213																		0/147
J2201.2+5926^{1}	0/522																		0/812
J2210.1+5913^{0}	0/63																		0/59
J2219.6+5850^{0}	0/176																		0/191
J2221.0+6307^{0}	0/204																		0/174
J2237.2+6316^{*,0}	0/41																		0/71
J2250.7+6305c^{0}	0/197																		0/113
J2257.5+6222c^{0}	0/118																		0/88
J2353.3+6643c^{0}	0/199																		0/210
J2359.4+6751c^{0}	1/172	J235952.54+675100.0$^{\times}$	359.9690	67.8500	155.6	37.1	33.3	5.4	5.9	0.53	1.44	3.76	-0.50	hh00	×	0.7541	0.1121	D	0/147
J2359.6+6543c^{0}	0/218																		0/248
Sources non-identifiées « centre galactique » (178)																			
J0846.7-4053^{0}	0/180																		0/173

... la table continue sur la page suivante...

... suite de la page précédente...

Nom 2FGL	N_c/N_W	Wise Id	RA (°)	DEC (°)	δ ($''$)	SNR_1	SNR_2	SNR_3	SNR_4	c_1	c_2	c_3	RFlux	Ccf_W	Var_W	ζ_{MLP}	η	Classe	$\widetilde{N}_c/\widetilde{N}_W$
J0848.7-4324^{0}	1/163	J084759.26-432549.7$^{\times}$	131.9969	-43.4305	505.7	25.9	28.7	13.6	6.3	0.85	2.68	2.06	-0.16	0000	○	0.7990	0.1259	D	1/203 (C)
J0850.1-4846^{0}	0/54																		0/86
J0858.0-4815^{0}	1/224	J085816.12-482046.4$^{\times}$	134.5672	-48.3462	340.1	43.3	44.7	29.4	7.2	0.50	2.34	1.73	0.11	hh00	×	0.7901	0.1731	D	5/184 (B)
J0900.5-4441c^{1}	0/32																		1/75 (B)
J0901.7-4655^{0}	0/4																		0/62
J0902.8-4741c^{*,0}	0/9																		0/6
J0904.0-4823c^{0}	0/159																		0/275
J0907.9-4716c^{0}	0/52																		0/72
J0914.1-4756^{1}	1/114	J091422.93-475246.9$^{\times}$	138.5955	-47.8797	241.3	42.1	50.8	52.9	51.8	1.60	3.39	2.89	0.83	0000	×	0.9182	0.0220	B	1/112 (B)
J0922.2-5214c^{1}	3/230	J092115.08-521423.4$^{\times}$	140.3128	-52.2398	540.0	38.2	47.4	53.4	31.4	1.45	3.30	2.47	0.41	0000	?	0.8601	0.0444	B	1/91 (B)
		J092144.59-521514.5$^{\times}$	140.4358	-52.2540	271.0	25.8	32.0	22.6	21.5	1.18	3.20	2.84	-0.12	0000	×	0.8071	0.0889	C	
		J092132.69-521449.6$^{\times}$	140.3862	-52.2471	378.2	39.1	43.1	20.4	22.4	1.00	2.46	2.77	0.03	0000	○	0.8006	0.1777	D	
J1010.7-5643c^{0}	0/223																		3/184 (B)
J1020.0-6029^{0}	1/81	J102115.25-604152.8$^{\times}$	155.3135	-60.6980	905.6	23.6	41.2	27.4	19.5	1.69	3.16	2.51	-0.10	00h0	×	0.8252	0.0313	B	0/439
J1023.5-5749c^{0}	0/0																		0/0
J1027.4-5730c^{1}	0/80																		1/40 (B)
J1030.4-6015^{0}	0/188																		0/252
J1036.4-5828c^{*,0}	0/118																		0/96
J1038.6-5850c^{*,0}	0/32																		0/32
J1046.8-6005c^{0}	0/9																		0/17
J1050.3-5922c^{0}	0/22																		0/31
J1056.0-5853^{0}	0/137																		2/119 (C)
J1059.3-6118c^{*,0}	2/231	J105913.33-612129.8$^{\times}$	164.8056	-61.3583	167.4	10.0	33.8	60.4	20.8	1.66	3.46	3.50	1.50	0h0h	○	1.1026	0.0446	B	0/254
		J105853.69-610737.9$^{\times}$	164.7237	-61.1272	692.1	9.2	14.7	43.5	24.4	0.82	3.54	4.12	1.65	00hh	○	0.8640	0.0446	B	
J1102.1-6308c^{0}	0/317																		0/179
J1104.7-6036^{0}	0/19																		0/10
J1105.6-6114^{0}	0/146																		0/143
J1118.9-6027c^{0}	0/177																		0/95
J1138.8-6233c^{0}	2/408	J113727.04-623502.4$^{\times}$	174.3627	-62.5840	570.6	6.7	17.4	9.8	3.9	1.07	2.82	2.22	-0.18	0000	○	0.8282	0.1576	D	2/396 (C)
		J114011.94-623536.6$^{\times}$	175.0498	-62.5935	589.1	39.0	51.0	65.4	40.5	1.14	3.30	2.07	1.35	hh00	○	0.9172	0.0788	C	
J1203.6-6243c^{0}	0/28																		0/46
J1231.1-6512^{0}	0/158																		0/139
J1236.1-6155^{0}	0/12																		0/40
J1243.9-6232^{0}	0/20																		0/41
J1248.6-5510^{*,1}	1/231	J124946.07-550758.6°	192.4420	-55.1330	603.0	36.6	32.5	26.4	9.6	0.99	2.71	2.52	-0.38	0000	?	0.8301	0.0892	C	0/395
J1255.8-5828^{2}	1/209	J125646.33-581306.8$^{\times}$	194.1931	-58.2186	1008.5	34.2	35.8	21.9	6.1	1.07	2.39	2.22	-0.49	0000	×	0.8014	0.1615	D	1/359 (C)
J1303.7-6316c$^{\sharp\star,0}$	0/63																		1/82 (B)
J1309.6-6230c^{0}	0/109																		0/99
J1312.0-6458^{0}	0/3115																		1/1156 (D)
J1317.2-6304^{0}	0/80																		0/89
J1320.1-5756^{*,0}	0/120																		0/91
J1324.4-5411$^{\sharp\star,1}$	1/45	J132415.49-541104.4°	201.0646	-54.1846	126.2	31.1	29.7	17.4	6.4	1.37	2.88	2.46	-0.66	00o0	×	0.7655	0.0348	B	0/129

... la table continue sur la page suivante...

... suite de la page précédente...

Nom 2FGL	N_c/N_W	Wise Id	RA (°)	DEC (°)	δ ($''$)	SNR_1	SNR_2	SNR_3	SNR_4	c_1	c_2	c_3	RFlux	Ccf_W	Var_W	ζ_{MLP}	η	Classe	$\widetilde{N}_c/\widetilde{N}_W$
J1329.7-6108^{0}	0/48																		0/44
J1335.4-5658^{*,0}	1/459	J133451.11-565517.2$^{\times}$	203.7130	-56.9215	364.2	41.8	46.5	39.0	10.9	0.76	2.14	1.57	-0.02	0000	∘	0.9636	0.0887	C	0/288
J1349.9-6222^{0}	0/54																		0/45
J1358.8-6027c^{0}	0/147																		0/126
J1405.5-6121^{0}	0/5																		0/13
J1407.6-5937c$^{\sharp\star,0}$	0/259																		0/329
J1414.1-5450^{0}	0/62																		2/202 (C)
J1415.7-6520$^{\sharp\star,0}$	0/75																		0/49
J1422.5-6137c^{0}	0/19																		0/20
J1427.6-6048c^{*,0}	0/38																		0/65
J1437.2-5211^{0}	0/28																		0/75
J1446.6-5753^{0}	0/55																		0/118
J1456.7-6247c$^{\sharp\star,0}$	0/123																		1/198 (C)
J1503.9-5800c^{0}	0/75																		0/65
J1510.9-5808c$^{\sharp,0}$	0/32																		0/69
J1512.5-6247c^{1}	2/357	J151214.19-624421.7$^{\times}$	228.0591	-62.7394	222.0	9.4	16.8	10.9	5.5	1.58	2.98	2.60	-0.48	0000	?	0.7998	0.2758	D	1/346 (D)
		J151137.47-624033.7$^{\times}$	227.9062	-62.6760	561.5	38.8	45.0	28.9	8.4	0.88	1.99	1.70	0.21	0000	×	0.8016	0.2758	D	
J1528.0-5841^{0}	0/58																		1/98 (C)
J1539.3-4636^{0}	0/1191																		1/820 (D)
J1551.0-4636^{0}	0/90																		0/143
J1551.3-5333c^{0}	0/167																		0/119
J1552.8-4824^{0}	0/212																		1/286 (D)
J1554.4-5317c^{0}	0/8																		0/10
J1610.1-4808^{0}	0/319																		0/308
J1614.9-5212$^{\sharp\star,0}$	0/8																		0/54
J1615.2-5138^{0}	0/20																		0/11
J1617.6-4219^{*,0}	1/104	J161800.70-421216.6$^{\times}$	244.5029	-42.2046	500.7	7.2	14.3	8.7	7.6	2.11	3.59	3.12	-0.52	0h00	×	0.8788	0.0201	B	1/198 (D)
J1618.0-5825$^{\sharp\star,0}$	0/11																		1/23 (A)
J1619.0-4650^{0}	0/1774																		0/1473
J1619.6-4509$^{\sharp\star,0}$	0/212																		0/264
J1619.7-5040c^{0}	0/9																		0/13
J1620.6-5111c^{0}	0/15																		0/83
J1624.0-4941c^{0}	0/76																		0/62
J1630.1-4615^{0}	0/138																		0/126
J1630.2-4752$^{\sharp\star,0}$	0/53																		0/43
J1632.4-4753c^{0}	0/7																		0/7
J1634.4-4743c^{0}	0/5																		0/7
J1636.3-4740c^{0}	0/5																		0/7
J1638.0-4703c^{0}	0/12																		0/12
J1639.7-5504^{*,0}	0/142																		0/99
J1639.8-4921c$^{\sharp\star,1}$	0/335																		0/393
J1648.1-4930$^{\sharp\star,0}$	0/222																		0/216

... la table continue sur la page suivante...

... suite de la page précédente...

Nom 2FGL	N_c/N_W	Wise Id	RA (°)	DEC (°)	δ ($''$)	SNR_1	SNR_2	SNR_3	SNR_4	c_1	c_2	c_3	RFlux	Ccf_W	Var_W	ζ_{MLP}	η	Classe	$\widetilde{N}_c/\widetilde{N}_W$
J1650.6-4603c^{0}	0/26																		0/12
J1651.8-4439c^{0}	0/286																		0/357
J1652.5-4351c^{0}	0/48																		0/39
J1653.9-4627c^{0}	0/77																		0/62
J1657.5-4652^{0}	0/47																		0/49
J1658.1-4743^{0}	0/120																		0/140
J1700.8-4912^{*,0}	0/133																		1/228 (B)
J1704.9-4618^{0}	0/560																		1/542 (D)
J1710.5-5020^{0}	0/57																		1/79 (B)
J1717.3-2809$^{\sharp\star,0}$	0/128																		1/147 (B)
J1726.6-3545^{0}	0/158																		0/94
J1727.8-2308$^{\sharp\star,0}$	0/10																		0/41
J1728.0-2737c^{0}	0/188																		0/204
J1731.9-2703c$^{\sharp\star,0}$	0/257																		0/232
J1733.2-2913c$^{\sharp\star,0}$	0/272																		0/295
J1733.4-2812c^{*,0}	0/184																		0/189
J1734.7-2533^{0}	1/174	J173432.87-253453.7$^{\times}$	263.6370	-25.5816	180.3	47.1	60.6	84.1	86.1	1.53	2.80	1.94	1.85	0000	$\times$	0.9432	0.0336	B	0/148
J1739.6-2726^{0}	0/195																		1/358 (D)
J1742.0-2540c^{*,0}	0/83																		0/78
J1743.9-3039c^{0}	0/67																		0/43
J1746.6-2851c^{0}	0/0																		0/0
J1747.2-3507$^{\sharp,1}$	0/128																		0/188
J1747.3-2825c^{0}	0/3																		0/3
J1748.6-2913^{0}	0/35																		0/31
J1749.7-3134c^{2}	0/180																		0/132
J1754.1-2930^{1}	0/203																		0/199
J1754.4-2538c^{0}	0/39																		0/36
J1758.8-2402c^{0}	0/103																		0/96
J1759.5-0521^{*,0}	0/42																		0/28
J1800.8-2400^{0}	0/1																		0/12
J1803.3-2148^{0}	0/1																		0/1
J1805.0-0845$^{\sharp\star,0}$	0/9																		0/6
J1807.7-0419^{0}	0/54																		0/58
J1808.5-2037c^{0}	0/72																		0/57
J1811.3-2421^{0}	0/363																		1/347 (C)
J1813.6-2821$^{\sharp,0}$	0/55																		0/56
J1813.7-1139c^{0}	0/328																		1/308 (C)
J1814.1-1735c^{0}	0/11																		0/23
J1817.6-1651c^{0}	0/154																		0/150
J1823.1-1338c^{0}	0/6																		0/5
J1827.4-0846^{0}	0/17																		0/40
J1827.4-1445c$^{\sharp,0}$	0/79																		0/76

... la table continue sur la page suivante...

... suite de la page précédente...

Nom 2FGL	N_c/N_W	Wise Id	RA (°)	DEC (°)	δ ($''$)	SNR_1	SNR_2	SNR_3	SNR_4	c_1	c_2	c_3	RFlux	Ccf_W	Var_W	ζ_{MLP}	η	Classe	$\widetilde{N}_c/\widetilde{N}_W$
J1829.1-0340c^{0}	1/66	J182905.44-034245.4$^{\times}$	277.2727	-3.7126	164.3	34.3	52.5	16.7	26.5	2.63	0.93	3.28	0.53	0h00	×	0.7854	0.0510	C	0/61
J1829.8-0204c^{3}	3/78	J182954.36-015823.2$^{\times}$	277.4765	-1.9731	383.9	26.0	44.8	8.9	9.9	2.67	1.40	2.92	0.01	0000	×	0.9966	0.0151	B	0/95
		J182954.05-015714.5$^{\times}$	277.4752	-1.9540	450.1	29.9	23.4	8.2	3.3	0.76	2.09	2.24	-0.32	0000	×	0.8666	0.0151	B	
		J182952.33-015745.0$^{\times}$	277.4681	-1.9625	415.8	8.9	38.1	14.0	18.4	3.63	1.58	3.85	0.01	0000	×	1.2788	0.0151	B	
J1830.4-1634$^{\sharp\star,0}$	0/76																		0/77
J1831.2-1518^{0}	0/606																		1/575 (D)
J1832.0-0200^{0}	0/22																		0/38
J1833.1-0437c^{0}	1/224	J183319.77-044210.4$^{\times}$	278.3324	-4.7029	290.2	13.3	23.4	7.9	11.2	0.90	2.71	3.81	-0.11	00hh	?	0.7941	0.1731	D	0/268
J1835.4+1036$^{\sharp\star,1}$	2/24	J183551.92+103056.8°	278.9663	10.5158	510.4	43.4	45.0	40.4	18.4	0.97	2.82	2.56	0.06	0000	×	0.8063	0.0093	A	0/12
		J183507.09+103152.0$^{\times}$	278.7796	10.5311	398.3	15.0	18.0	9.3	4.4	1.52	3.11	2.67	-0.80	0000	×	0.7588	0.0185	B	
J1835.4+1349$^{\sharp\star,2}$	3/15	J183535.34+134848.8°	278.8973	13.8136	141.1	43.1	42.6	27.2	7.5	0.69	2.14	2.10	-0.07	0000	?	0.9096	0.0029	A	1/12 (A)
		J183539.21+135054.9$^{\times}$	278.9134	13.8486	206.8	38.0	40.5	24.0	12.3	1.18	2.70	2.73	-0.26	0h00	∘	0.8470	0.0058	A	
		J183522.00+135733.9°	278.8417	13.9594	484.1	37.2	43.4	27.8	8.1	1.16	2.62	2.20	-0.26	0hh0	?	0.8490	0.0058	A	
J1835.5-0649$^{\sharp,0}$	0/51																		0/47
J1837.3-0700c^{0}	0/14																		0/14
J1839.0-0102^{0}	0/162																		0/89
J1839.3-0558c^{0}	0/13																		0/16
J1842.8-0359c^{0}	0/6																		0/19
J1843.7-0312c^{0}	0/150																		0/127
J1844.3+1548$^{\sharp\star,1}$	1/4	J184425.36+154645.9°	281.1057	15.7794	150.7	47.5	49.3	38.9	14.4	0.84	2.30	2.06	0.13	0000	×	0.8983	0.0008	A	0/4
J1844.3-0343c^{0}	0/23																		0/20
J1844.9-1116^{1}	0/156																		0/165
J1847.2-0236^{0}	0/39																		0/23
J1848.2-0139c^{0}	0/17																		0/14
J1849.9-0125c^{0}	0/33																		0/50
J1852.8+0156c^{0}	0/61																		0/60
J1856.2+0450c^{0}	0/135																		0/170
J1857.2+0055c^{0}	0/41																		0/36
J1857.6+0211^{1}	0/30																		0/27
J1857.8+0355c^{0}	0/352																		0/290
J1858.5+0129c^{0}	0/96																		0/182
J1859.3+0312c^{0}	0/124																		0/135
J1902.3-1106$^{\sharp\star,0}$	0/18																		0/29
J1908.8-0132$^{\sharp,0}$	0/183																		0/149
J1914.0+1436^{0}	0/273																		0/120
J1914.4+0951c^{0}	0/74																		0/131
J1921.1+1436c^{0}	0/180																		0/110
J1921.3+0131^{0}	0/36																		0/20
J1923.4+2013^{0}	0/477																		0/548
J1924.8+1724c^{0}	0/202																		0/177
J1931.8+1325^{0}	0/314																		1/286 (D)
J1949.7+2405^{0}	2/151	J194941.90+235811.2$^{\times}$	297.4246	23.9698	419.2	29.8	30.5	8.7	3.0	0.35	2.04	1.64	-0.18	0000	×	0.7684	0.1167	D	0/211

... la table continue sur la page suivante...

... suite de la page précédente...

Nom 2FGL	N_c/N_W	Wise Id	RA (°)	DEC (°)	δ (″)	SNR_1	SNR_2	SNR_3	SNR_4	c_1	c_2	c_3	RFlux	Ccf_W	Var_W	$\hat{\varsigma}_{MLP}$	η	Classe	$\widetilde{N}_c/\widetilde{N}_W$
		J194909.75+241023.4×	297.2907	24.1732	574.9	17.0	30.2	45.3	38.2	1.204	2.273	3.50	0.71	0000	?	0.7652	0.1167	D	
J1958.6+4020♯⋆,1	2/331	J195842.28+401125.8°	299.6762	40.1905	577.2	29.5	32.6	21.8	8.5	1.142	2.822	2.41	-0.51	0000	×	0.8091	0.1279	D	0/226
		J195854.15+401210.1×	299.7257	40.2028	567.7	36.0	32.8	16.4	4.9	0.982	2.692	2.09	-0.56	0000	×	0.7620	0.2557	D	
J1959.9+3336c⁰	0/277																		0/175
J2004.4+3339c⁰	0/1																		0/1
J2009.8+2747⁰	1/656	J200856.79+280348.1×	302.2367	28.0634	1192.5	41.9	45.5	29.6	7.2	0.552	1.51	1.70	0.20	0h00	×	0.8599	0.1267	D	4/1186 (D)
J2012.4+3955c⁰	0/134																		0/183
J2013.8+4115c⁰	0/105																		0/111
J2020.0+4159⁰	0/219																		0/185
J2033.6+3927⁰	0/152																		0/153
J2036.0+4224c⋆,0	1/169	J203543.72+422222.6×	308.9322	42.3730	279.1	15.2	18.3	19.3	16.7	0.452	1.374	4.49	0.91	00hh	◦	0.7672	0.1306	D	0/177
J2038.0+4145c⁰	0/533																		1/442 (D)
J2040.1+4105c♯,0	0/230																		0/236
J2042.0+4252c¹	1/306	J204157.90+424757.3×	310.4913	42.7993	266.8	26.0	34.9	34.1	12.1	0.772	1.55	1.50	0.83	000h	◦	0.9308	0.0591	C	0/419
J2047.9+4536c⁰	0/383																		0/420
J2051.8+5054⁰	0/51																		0/42
J2055.8+4754♯⋆,0	2/57	J205520.68+474851.3×	313.8362	47.8143	466.7	19.0	26.8	17.6	5.0	1.452	1.78	2.12	-0.48	0000	◦	0.7594	0.0440	B	0/251
		J205526.79+475806.8×	313.8616	47.9686	332.3	34.5	38.5	18.2	4.6	0.691	1.99	2.20	0.00	0000	?	0.8546	0.0110	B	
J2103.3+4357c⁰	1/201	J210318.73+435343.6×	315.8281	43.8955	240.0	11.1	25.2	6.4	8.0	1.683	1.41	2.75	-0.60	0h00	◦	0.8299	0.0777	C	2/142 (C)
J2104.9+3555♯⋆,0	0/7																		0/15
J2107.9+5207c⁰	2/151	J210816.67+520645.0×	317.0695	52.1125	184.9	13.9	28.0	25.8	10.4	1.633	1.11	2.27	-0.31	0000	◦	0.7889	0.1167	D	1/144 (C)
		J210842.27+520732.5×	317.1761	52.1257	419.8	41.7	46.8	26.0	19.2	1.112	1.63	2.81	0.10	0h00	×	0.8079	0.0583	C	
J2128.7+5824³	1/125	J212810.83+583336.6°	322.0452	58.5602	608.4	32.8	36.2	20.6	6.9	1.322	1.82	2.22	-0.54	0000	×	0.7748	0.0966	C	2/198 (C)

Bibliographie

J. J. Condon, et al. The NRAO VLA Sky Survey. AJ, 115:1693–1716, May 1998. doi: 10.1086/300337.

R. D'Abrusco, et al. Infrared Colors of the Gamma-Ray-detected Blazars. ApJ, 748:68, March 2012. doi: 10.1088/0004-637X/748/1/68.

R. D'Abrusco, et al. Unveiling the Nature of Unidentified Gamma-Ray Sources. I. A New Method for the Association of Gamma-Ray Blazars. ApJS, 206:12, June 2013. doi: 10.1088/0067-0049/206/2/12.

M. Doert et M. Errando. Search for gamma-ray-emitting active galactic nuclei in the Fermi-LAT unassociated sample using machine learning. *ArXiv e-prints*, December 2013.

B. T. Draine. Interstellar Dust Grains. ARA&A, 41:241–289, 2003. doi: 10.1146/annurev.astro.41.011802.094840.

R. Indebetouw, et al. The Wavelength Dependence of Interstellar Extinction from 1.25 to 8.0 μm Using GLIMPSE Data. ApJ, 619:931–938, February 2005. doi: 10.1086/426679.

E. Massaro, et al. Roma-BZCAT: a multifrequency catalogue of blazars. A&A, 495:691–696, February 2009. doi: 10.1051/0004-6361:200810161.

F. Massaro, et al. Identification of the Infrared Non-thermal Emission in Blazars. ApJ, 740:L48, October 2011. doi: 10.1088/2041-8205/740/2/L48.

F. Massaro, et al. Unveiling the Nature of Unidentified Gamma-Ray Sources. II. Radio, Infrared, and Optical Counterparts of the Gamma-Ray Blazar Candidates. ApJS, 206:13, June 2013. doi: 10.1088/0067-0049/206/2/13.

F. Massaro, et al. Unveiling the nature of the unidentified gamma-ray sources V: analysis of the radio candidates with the kernel density estimation. *ArXiv e-prints*, April 2014.

N. Mirabal, et al. Fermi's sibyl: mining the gamma-ray sky for dark matter subhaloes. MNRAS, 424:L64–L68, jul 2012.

P. L. Nolan, et al. Fermi Large Area Telescope Second Source Catalog. ApJS, 199:31, April 2012. doi: 10.1088/0067-0049/199/2/31.

E. L. Wright, et al. The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance. AJ, 140:1868, December 2010. doi: 10.1088/0004-6256/140/6/1868.

Chapitre 12

Résumé

Dans cette dernière partie du manuscrit, après avoir présenté le fonctionnement des satellites *Fermi*-*LAT* et *WISE*, nous avons présenté deux analyses indépendantes qui permettent d'identifier des candidats blazars parmi les sources non-identifiées du catalogue 2FGL.

La première approche utilise des variables discriminantes construites à partir des paramètres du catalogue 2FGL afin de créer une classification binaire blazar/non-blazar. Une attention particulière a été portée au choix des variables discriminantes utilisées, afin de minimiser l'impact de leurs corrélations avec la brillance des sources dans l'estimation des performances finales. L'analyse, développée à partir de la combinaison des décisions d'une forêt d'arbres de décision « boostés » et d'un réseau de neurones de type MLP, est optimisée selon la latitude galactique (basse ou haute) des sources non-identifiées. Afin d'entraîner et d'estimer les performances des classifieurs nous avons utilisé un lot « signal » composé de blazars identifiés et deux lots de « fond » respectivement composés de pulsars pour le traitement des sources de haute latitude, et de l'ensemble des sources galactiques pour le traitement des sources non-identifiées de basse latitude. Les résultats obtenus à haute latitude sont en très bon accord avec des études similaires déjà existantes. À basse latitude, notre méthode permet d'élargir d'une vingtaine d'unités le nombre de candidats blazars tout en confirmant les candidats déjà proposés par ailleurs.

La seconde approche est basée sur l'observation que les blazars occupent une place spécifique dans des diagrammes couleur-couleur, construits à partir des magnitudes du satellite *WISE*. Cette étude, qui consiste en une classification binaire blazar/non-blazar, se démarque de prime abord de celles déjà existantes par une maîtrise du taux de fausses identifications, par le fait qu'elle ne fait pas d'hypothèse sur la géométrie de l'espace des variables discriminantes, et par l'ajout d'une variable discriminante supplémentaire (RFlux). Cette analyse est basée sur une méthode non-paramétrique de type réseau de neurones « boostés » et a été développée à l'aide d'un échantillon composé de blazars IR identifiés provenant d'une association entre les catalogues 2FGL et Roma-BZCAT, et à l'aide d'échantillons de fond constitués de sources IR provenant de différentes régions du ciel dans lesquelles aucune source des catalogues 2FGL et Roma-BZCAT n'est présente. Une étude sur la dépendance spatiale des trois couleurs du fond IR nous a conduit à séparer le traitement des sources non-identifiées selon trois régions distinctes, « extragalactique », « centre galactique » et « anti-centre galactique », afin d'optimiser le pouvoir de séparation blazar/non-blazar. L'application de notre méthode aux sources non-identifiées du catalogue 2FGL permet d'obtenir 130 contreparties IR de type blazar réparties parmi 104 sources non-identifiées. On en compte respectivement 21, 35, 26 et 48 de classe A, B, C et D, ce qui permet de leur associer un nombre attendu de fausses identifications meilleur que 0,01, 0,05, 0,1 et 0,5 respectivement.

Quatrième partie

Conclusions et perspectives

Chapitre 13

Conclusion et perspectives

Ce manuscrit traite de l'astronomie γ de très haute énergie depuis le sol (au-delà de $\sim$100 GeV) et de l'astronomie γ de haute énergie depuis l'espace (entre $\sim$30 GeV et $\sim$300 GeV). La première partie traite de l'astronomie γ, des noyaux actifs de galaxie et en particulier des blazars, ainsi que de l'utilisation des méthodes multivariées pour la construction de classifications binaires « signal »/« fond ». La seconde partie est dédiée au développement d'une analyse multivariée pour le réseau hybride HESS-II et à l'étude des premières données. La troisième partie est consacrée à la recherche de blazars parmi les sources non-identifiées du catalogue 2FGL du *Fermi*-*LAT*.

*
* *

Ces dix dernières années, l'astronomie γ de très haute énergie a connu un essor considérable grâce à la mise en fonctionnement des expériences d'imagerie atmosphérique Tcherenkov *H.E.S.S.*, *MAGIC* et *VERITAS*. Elles ont mis en évidence l'existence de près de 160 sources émettrices dans le domaine du TeV qui sont remarquables par leur diversité : restes de supernovæ jeunes, restes de supernovæ âgés en interaction avec le milieu environnant, plérions, systèmes binaires, émission pulsée des pulsars, émission diffuse le long du plan galactique, et bien sûr les AGN, essentiellement des blazars et quelques radiogalaxies. Bien que les approches de type SSC soient largement utilisées afin d'expliquer la distribution spectrale en énergie des blazars, des études récentes conduites avec l'expérience *H.E.S.S.* sur des blazars de type HBL défient sérieusement ces approches. En effet, elles échouent à reproduire les caractéristiques spectrales de certaines sources ou les corrélations observées à différentes longueurs d'onde lors d'états hauts. Récemment, pour la première fois une mesure de la densité du fond infrarouge entre 0,3 µm et 0,7 µm a été réalisée à l'aide de blazars brillants détectés par l'expérience *H.E.S.S.* Les blazars permettent aussi d'apporter des contraintes à des problèmes de physique fondamentale tels que la violation de l'invariance de Lorentz ou le couplage aux photons de particules hypothétiques de type axion. Avec l'ajout du grand télescope au réseau *H.E.S.S.*, la seconde phase de l'expérience permettra de gagner en sensibilité et de réduire l'énergie seuil de l'instrument, ce qui permettra à son tour un meilleur recouvrement entre les domaines γ de haute énergie et de très haute énergie, une meilleure caractérisation spectrale des blazars observés et ce qui donnera accès à une population à plus haut redshift.

Une partie importante du travail présenté dans ce manuscrit est consacrée au développement d'une méthode d'analyse de type multivariée à base d'arbres de décision « boostés » destinée à une exploitation optimale du réseau hybride HESS-II. Prenant pour point de départ l'analyse Paris-MVA initialement développée pour *H.E.S.S.*, ce travail a d'abord été réalisé de manière exploratoire à base de simulations MC ce qui a permis de mesurer le gain lié à l'ajout du cinquième télescope et de donner une première estimation de la sensibilité de l'instrument en mode hybride. Après accumulation de plusieurs centaines d'heures d'observation de bonne qualité, il a été possible de refaire ce travail à partir des données de fond réel. Nous avons alors optimisé la détermination des points de fonctionnement de la méthode BDT, et ce pour trois configurations représentant des types de sources de niveau de flux et d'indice spectral différents. Ce travail est une première étape pour aboutir à une exploitation optimale du réseau HESS-II en mode hybride. L'application de l'analyse à un lot de sources sélectionnées (la nébuleuse du Crabe, PKS 2155-304, PG 1553+113 et Mkn 501) montre des résultats remarquables tels que une détection à $35{,}1\,\sigma\sqrt{\mathrm{h}^{-1}}$ pour la nébuleuse du Crabe ou à $3{,}7\,\sigma\sqrt{\mathrm{h}^{-1}}$ pour une source faible comme PG 1553+113. Les spectres montrés à titre préliminaire sont conformes à ce qui a déjà été mesuré pour ces différentes sources mais les systématiques

sont néanmoins très importantes du fait du statut préliminaire des simulations utilisées (notamment en ce qui concerne les efficacités optiques des différents télescopes).

Les points sur lesquels il conviendra de travailler dans un futur proche sont les suivants :

— la génération de nouvelles simulations plus réalistes qui utiliseront notamment des efficacités optiques correspondant à celles mesurées expérimentalement pour les télescopes de taille moyenne, et qui permettront de prendre en compte leur variation avec le temps

— sur cette base, la reproduction des tables nécessaires pour la reconstruction et la discrimination, l'entraînement des arbres de décision « boostés » pour chaque intervalle de l'espace des phases et la recherche de son point de fonctionnement optimal, et enfin la génération des fonctions de réponse de l'instrument

— la prise en compte des aberrations du miroir de CT_5 dans le modèle 3D, plus importantes du fait de sa structure parabolique que pour les télescopes de taille moyenne

— l'amélioration de la reconstruction de la direction des évènements à basse énergie à partir d'une exploitation différente des informations provenant de CT_5 qui pourrait s'inspirer des méthodes développées pour le mode mono

— l'exploitation des temps relatifs associés aux signaux intégrés dans chaque pixel de la caméra de CT_5 pour améliorer la reconstruction et pour la recherche de nouveaux paramètres discriminants

*
* *

Depuis sa mise sur orbite en 2008, le satellite *Fermi*-*LAT* a révolutionné notre vision du ciel γ de haute énergie. Au bout de deux ans de fonctionnement du *LAT*, la collaboration a fourni un catalogue de 1873 sources détectées au delà de 100 MeV : des restes de supernovæ (avec une première mise en évidence de l'accélération de hadrons en leur sein), des plérions, des amas globulaires, des systèmes binaires X, des pulsars, des novæ, des sursauts γ et une majorité d'AGN, en particulier des blazars. En plus de leur utilisation comme sonde pour différentes questions de physique (rayonnement du fond extragalactique, invariance de Lorentz), la détection et la caractérisation spectrale d'un grand nombre de blazars est très utile pour les études sur les caractéristiques de BL Lac, FSRQ et sur l'unification de ces deux populations. En particulier, la validité de la séquence des blazars est toujours un objet de discussion et nécessite la poursuite des efforts, aussi bien sur le terrain des modèles que sur celui de la détection d'un nombre plus important d'objets. Nous contribuons à cet effort en proposant des méthodes basées sur les approches multivariées pour la caractérisation de candidats blazars parmi les nombreuses sources non-identifiées du catalogue 2FGL (1/3 de la population, soit 575 sources).

Nous avons tout d'abord développé une classification binaire « blazar »/« non-blazar » en utilisant des variables discriminantes construites à partir des paramètres du catalogue 2FGL et choisies de manière à minimiser la dépendance du pouvoir de séparation avec le flux des sources. Un traitement différencié à été appliqué selon que les sources non-identifiées se trouvent à haute ou à basse latitude galactique afin de prendre en compte la nature différente du fond dans ces deux régions. Nous avons obtenu un total de 248 candidats blazars, dont 179 se trouvent à haute latitude galactique et 69 à basse latitude galactique. Ces résultats sont en bon accord avec les études existantes et permettent de proposer 20 nouveaux candidats blazars de basse latitude galactique. La seconde approche a eu pour objectif d'exploiter le catalogue de plus de cinq cent millions de sources infrarouges détectées par le satellite *WISE* qui a observé l'intégralité du ciel pendant un an dans des conditions stables. Afin de rechercher les contreparties infrarouges de type blazar spatialement compatibles avec des sources non-identifiées du catalogue 2FGL nous sommes partis de la mise en évidence par Massaro et al. de la spécificité des blazars dans l'espace des couleurs. Nous avons construit ici aussi une classification binaire « blazar »/« non-blazar » de type réseau de neurones, qui se démarque des approches existantes par une maîtrise du taux de fausses identifications, par le fait qu'elle ne

fait pas d'hypothèse sur la géométrie de l'espace des variables discriminantes, et par l'ajout d'une variable discriminante supplémentaire. Ce travail a été fait en découpant le ciel en trois régions distinctes afin de prendre en compte les différences de nature entre les sources composant le fond. Nous avons obtenu 130 contreparties IR de type blazar réparties parmi 104 sources non-identifiées. On en compte 21, 35, 26 et 48 respectivement de classe A, B, C et D, ce qui permet de leur associer un nombre attendu de fausses identifications meilleur que 0,01, 0,05, 0,1 et 0,5 respectivement.

Il reste cependant des points sur lesquels il conviendra de travailler dans un futur proche :

— l'exploitation des données multi-longueur d'onde pour une étude approfondie des candidats blazars proposés dans cette étude

— l'exploitation des caractéristiques spectrales qui distinguent les blazars de type FSRQ et BL Lac afin de déterminer le type des candidats blazars

— l'étude des effets de sélection des blazars liés à l'utilisation du filtre W_4 (celui de plus basse énergie) du satellite *WISE*

*
* *

Ces travaux verront une application immédiate suite à la publication récente [1] du catalogue 3FGL qui correspond à quatre années de données du *Fermi-LAT* et qui comprend près de 1000 sources non-identifiées. Après l'exploitation des données multi-longueur d'onde offertes par diverses expériences afin de confirmer la nature blazar des candidats proposés ici, il sera possible, en se servant des courbes de sensibilité tels que celles produites ici pour HESS-II, de déterminer les objets qui peuvent être détectés par les imageurs atmosphérique Tcherenkov au sol pour augmenter de manière significative l'échantillon de ∼60 de blazars aujourd'hui détectés à très haute énergie. Enfin, les développements réalisés dans ce manuscrit pour l'exploitation du premier réseau hybride HESS-II pourraient avoir une application dans le cadre du futur observatoire *CTA* qui sera composé d'un nombre important de télescopes de trois types différents.

1. Une première version du catalogue a été publiée le 8 janvier 2015.

Table des figures

1.1 Spectre électromagnétique et opacité de l'atmosphère . . . 6
1.2 Spectre des rayons cosmiques . . . 7
1.3 Gerbes atmosphériques électromagnétique et hadronique . . . 8
1.4 Diagramme d'Hillas . . . 9
1.5 Vu du ciel γ par l'expérience *Fermi-LAT* . . . 13
1.6 Évolution du nombre de sources détectées au TeV . . . 15
1.7 Plan galactique vue par l'expérience *H.E.S.S.* . . . 16

2.2 Spectres optiques de Seyfert 1 et 2 . . . 22
2.4 Schéma de l'unification des différents types d'AGN . . . 24
2.6 Distribution spectrale en énergie du blazar 3C 454.3 pour différentes époques . . . 27
2.7 Courbe de lumière du blazar 3C 454.3 simultanée en optique, X et γ . . . 28

3.1 Distributions de la longueur et de la luminosité des bars et des saumons. . . . 38
3.2 Choix de la séparation optimale . . . 39
3.3 Données séparables de manière linéaire et non-linéaire . . . 40
3.4 Exemple d'arbre de décision . . . 42
3.5 Fonctions d'impuretés. . . . 43
3.6 Schéma d'un neurone biologique et représentation d'un modèle simple de neurone . . . 45
3.7 Espace des phases des poids (w_1, w_2) pour lequel est représenté la sortie d'une neurone 47
3.8 Illustration de la phase d'apprentissage d'un neurone . . . 48
3.9 Représentation d'un réseau de neurones MLP . . . 49

4.1 Modèle d'Heitler d'une gerbe électromagnétique . . . 56
4.2 Simulations d'une gerbe EM et d'une gerbe hadronique . . . 57
4.3 Principe de l'effet Tcherenkov . . . 58
4.4 Densité de photons Tcherenkov moyenne en fonction du paramètre d'impact et de l'altitude le long l'axe de gerbe EM . . . 59
4.5 Principes de l'imagerie atmosphérique Tcherenkov et de la stéréoscopie . . . 60
4.6 Réseau de télescopes de l'expérience *H.E.S.S.* . . . 61
4.7 Télescope HESS-I et sa caméra . . . 62
4.8 Signal moyen d'un photoélectron mesuré au niveau de l'anode et schéma d'une carte analogique avec la voie de déclenchement et les voies de lecture d'un photomultiplicateur 62
4.9 Taux de déclenchement en fonction Q_0 des télescopes de type HESS-I . . . 63
4.10 Nombre de coups ADC en fonction du nombre de photoélectrons . . . 65
4.11 Piédestaux pour différents niveaux de NSB et ajustement du gain dans la voie haut gain . . . 66
4.12 Région du ciel vue par l'expérience *H.E.S.S.* . . . 67
4.13 Évolution du paramètre de transparence de l'atmosphère entre 2004 et 2012 . . . 68

5.1 Paramètres de Hillas . . . 72
5.2 Principe de la stéréoscopie . . . 73
5.3 Paramètres de Hillas . . . 74
5.4 Hauteur du maximum de gerbe ajustée par le modèle 3D en fonction du nombre de photons Tcherenkov constituant la photosphère des gerbes électromagnétiques et hadroniques . . . 76
5.5 Distributions de la largeur réduite ω . . . 77

5.6 Distributions des paramètres de *Shower Goodness* et de la hauteur de première interaction . . . 79
5.7 Distributions des fonctions de densité de probabilité des quatre variables discriminantes de l'analyse X_{eff} . . . 80
5.8 Distributions du rapport signal sur bruit et de la significativité en fonction du paramètre X_{eff}, pour différentes valeurs de η et pour des données de la nébuleuse du Crabe . . . 81
5.9 Paramètres discriminants de l'analyse ImPACT . . . 82
5.10 Charges Q en fonction du paramètre d'impact pour des γ simulés pour différents angles zénithaux. . . . 84
5.11 Paramètres provenant du modèle 3D utilisés dans la discrimination de l'analyse Paris-MVA . . . 85
5.12 Nouveaux paramètres créés spécifiquement pour l'analyse Paris-MVA . . . 86
5.13 Exemple de distribution du paramètre ζ des évènements γ et de fond pour les lots d'entraînement et de test, et exemple de gabarit des efficacités de reconstruction . . . 87
5.14 Résolution en énergie de l'analyse Paris-MVA pour deux configurations d'analyse . . . 88
5.15 Précision angulaire et surface efficace de collection de l'analyse Paris-MVA pour deux configurations d'analyse . . . 88
5.16 Acceptance radiale et régions couramment utilisées pour l'extraction du signal . . . 90
5.17 Surface de collection en fonction de l'énergie vraie pour la configuration « elm » avec $\theta = 18°$, $\delta = 0{,}5°$ et $\varepsilon = 70\,\%$ . . . 91
5.18 Résolution et biais en énergie en fonction de l'énergie vraie pour la configuration « elm » avec $\theta = 18°$, $\delta = 0{,}5°$ et $\varepsilon = 70\,\%$ . . . 92
5.19 Surface efficace associée à un « run » d'observation . . . 95
5.20 Spectre de la nébuleuse du Crabe avec une hypothèse de loi de puissance avec une coupure exponentielle généré avec START . . . 97
5.21 Exemple de courbe de lumière de la source PKS 2155-304, lors de l'état haut détecté en 2006, générée avec START . . . 98
5.22 Activité de la source Mkn 501 de juin 2004 à juin 2012 . . . 99
5.23 Courbe de lumière et spectres de la radiogalaxie M 87 . . . 100

6.1 Précision angulaire pour différents lots d'évènements vue par le réseau de 5 télescopes 105
6.2 Évènement γ de 100 GeV . . . 106
6.3 Évènement γ de 4,61 TeV . . . 107
6.4 Évènement γ de 7,27 TeV . . . 108
6.5 Évènement γ de 1,12 TeV . . . 108
6.6 Évènement γ de 6,02 TeV . . . 109
6.7 Illustration de la pré-sélection « modèle 3D » . . . 111
6.8 Illustration de la coupure $\sin\theta_{ij}$ avec un évènement γ de 803 GeV . . . 113
6.9 Mise en évidence de l'effet de la coupure $\sin\theta_{ij} = 0{,}3$ . . . 114
6.10 Nombre de photons Tcherenkov de la photosphère des gerbes en fonction de l'énergie reconstruite pour des γ . . . 117
6.11 Distributions du paramètre MSCL . . . 119
6.12 Distributions du paramètre MSCW . . . 119
6.13 Distributions du paramètre 3D-width . . . 120
6.14 Distributions du paramètre 3D-width-error . . . 120
6.15 Distributions du paramètre 3D-depth . . . 121
6.16 Distributions du paramètre ΔQ . . . 121
6.17 Distributions du paramètre R_{E} . . . 122
6.18 Distributions du paramètre Ω . . . 122
6.19 Performances et contrôle du sur-entraînement pour le premier intervalle N_c de multiplicité 2 . . . 123

6.20 Efficacités de reconstruction γ et de fond pour les différents intervalles de l'espace des phases . . . 124
6.21 Distribution du paramètre ζ pour des évènements *gamma* et de fond simulés pour le premier intervalle en N_c . . . 126
6.22 Distribution du paramètre ζ pour le fond simulé et réel pour les quatre dernier intervalles en N_c . . . 127
6.23 Résolution et biais en énergie en fonction de l'énergie reconstruite de l'analyse en mode hybride. . . . 128
6.24 Comparaison du taux de fond et de la surface efficace entre les analyses en mode hybride et stéréo . . . 129
6.25 Comparaison de la précision angulaire et de la sensibilité différentielle entre les analyses en mode hybride et stéréo . . . 131

7.1 Illustration de la méthode permettant d'obtenir du fond réel . . . 134
7.2 Illustration du lissage des efficacités γ . . . 137
7.3 Performances . . . 139
7.4 Détection de la nébuleuse du Crabe . . . 141
7.5 Détection de PKS 2155-304 . . . 142
7.6 Détection de PG 1553+113 . . . 143
7.7 Détection de Mkn 501 . . . 144
7.8 Spectres de la nébuleuse du Crabe . . . 145
7.9 Spectres de PKS 2155-304 . . . 146
7.10 Spectre de PG 1553+113 . . . 147
7.11 Spectre de Mkn 501 . . . 148

9.1 Latitude galactique des blazars et des sources galactiques identifiées, et non-identifiées *Fermi* . . . 158
9.2 Exemple de contreparties infrarouges dans la région de la position d'une source *Fermi* à un niveau de confiance de 99,9 % . . . 159
9.3 Sensibilité et couverture du ciel du *LAT* . . . 160
9.4 Vue schématique du LAT et de ses principaux constituants . . . 160
9.5 Performances du Fermi-Lat . . . 163
9.6 Charge utile du satellite *WISE* . . . 165
9.7 Stratégie de balayage du satellite *WISE*. . . . 166
9.8 Acceptance et sensibilité différentielle du satellite *WISE*. . . . 167
9.9 Zoologie astrophysique dans le diagramme couleurs $c_1(c_2)$. . . . 167

10.1 Propriétés des différents échantillons . . . 173
10.2 Paramètres spectraux *Fermi* . . . 174
10.3 Paramètres spectraux *Fermi* . . . 175
10.4 Paramètres spectraux *Fermi* . . . 176
10.5 Paramètres spectraux *Fermi* . . . 178
10.6 Paramètres spectraux *Fermi* . . . 179
10.7 Efficacités de reconstruction combinées de l'analyse extragalactique . . . 183
10.8 Efficacités de reconstruction combinées de l'analyse galactique . . . 184
10.9 Projection d'Aïtoff des coordonnées galactiques des candidats blazars . . . 186
10.10 Stabilité des indices de duretés HR_{12} et HR_{23} selon la présence de limite supérieur dans le calcul . . . 194
10.11 Stabilité des indices de duretés HR_{34} et HR_{45} selon la présence de limite supérieur dans le calcul . . . 195
10.12 Performances de la méthode BDT de l'analyse extragalactique . . . 199
10.13 Performances de la méthode MLP de l'analyse extragalactique . . . 199

10.14Performances de la méthode BDT de l'analyse galactique 201
10.15Performances de la méthode MLP de l'analyse galactique 201

11.1 Diagrammes couleur-couleur $c_1(c_2)$ pour un ensemble de blazars sélectionné et du fond infrarouge extragalactique. 204
11.2 Variation spatiale du fond infrarouge extragalactique selon la région du ciel observée 208
11.3 Variation spatiale du fond infrarouge galactique selon la région du ciel observée . . . 209
11.4 Distributions des rapports de signal sur bruit dans les quatre filtres de l'échantillon « signal » 210
11.5 Distributions des couleurs c_1, c_2 et c_3 pour les blazars et différents fond IR 212
11.6 Distributions du paramètre RFlux pour les blazars et différents fond IR, ainsi que sa dépendance en fonction du redshifts des blazars 213
11.7 Positions des sources 2FGL, Roma-BZCAT et des régions de fond utilisées pour la construction des échantillons de fond 215
11.8 Illustration du choix du réglage optimal de la méthode MLP 216
11.9 Performances de la méthode BDT de l'analyse « extragalactique » 219
11.10Performances de la méthode BDT de l'analyse « centre galactique » 220
11.11Performances de la méthode BDT de l'analyse « anti-centre galactique » 220
11.12Performances et nombre de contreparties infrarouges dans la boîte d'erreur des sources non-identifiées 222

Liste des tableaux

6.1 Caractéristiques des simulations utilisées. 104

7.1 Paramètres spectraux utilisés pour le développement des configurations 138
7.2 Résultats des observations de *H.E.S.S.* sur une sélection de sources 140

9.1 Sources présentes dans le catalogue 2FGL 164

10.1 Efficacités de reconstruction des différentes analyses 184
10.2 Liste de candidats blazars parmi les sources non-identifiées ne présentant pas de problèmes de reconstruction. 187
10.3 Liste de candidats blazars parmi les sources non-identifiées présentant des problèmes de reconstruction. 190
10.4 Paramètres de l'analyse BDT dans TMVA 196
10.5 Paramètres de l'analyse MLP dans TMVA 197

11.1 Paramètres TMVA de la méthode MLP utilisée pour la recherche des contreparties IR des sources non-identifiées du *Fermi-LAT* 217
11.2 Classement selon l'importance des variables discriminantes des différentes analyses . 219
11.3 Efficacités de reconstruction des différents régimes et valeurs finales de discrimination associées 223
11.4 Table qui récapitule les résultats pour les sources non-identifiées sans « flag » 228
11.5 Table qui récapitule les résultats pour les sources non-identifiées avec un « flag » . . 228
11.6 Liste des contreparties infrarouge de type blazar des sources non-identifiées ne présentant pas de problème de reconstruction. 229
11.7 Liste des contreparties infrarouge de type blazar des sources non-identifiées qui peuvent présenter des problèmes de reconstruction. 237

Remerciements/salutations :

Je tiens à remercier ou à saluer ici une grande partie des personnes qui ont contribué de près ou de moins près à ce que le projet aboutisse.

Un grand merci à mon directeur de thèse Santiago pour sa disponibilité, son sens critique, sa rigueur et son œil de physicien qui ont permis entre que l'aventure se déroule dans des conditions optimales et d'améliorer de manière très significative la qualité du manuscrit.

Merci aux directeurs Pierre Binétruy et Stavros Katsanevas de m'avoir accueilli au sein du laboratoire APC, d'avoir accepté de remplacer mon ordinateur disparu et d'avoir fait en sorte que la fin de ma thèse de déroule dans de bonnes conditions.

Merci à Catherine Boisson, David Williams, Frédéric Piron, Giovanni Lamanna et Benoît Lott d'avoir accepté d'être membre de mon jury de thèse. Merci en particulier à Giovanni et Benoît, d'avoir accepté la tâche de rapporteur ainsi que le calendrier légèrement serré qui a été imposé (joyeux Noël !).

Merci à Arache, Bruno, Régis et Anne pour leur disponibilité et pour les échanges qui ont permis d'améliorer la qualité des travaux présentés dans ce manuscrit. Bon trajet de thèse à Thomas et à Léa ! Une pensée amicale à Michael et à Yvonne.

Un grand merci à Vincent sans qui la seconde partie de ma thèse aurait été effectuée sans données H.E.S.S. calibrées. Merci à Jean-Philippe d'avoir pris du temps pour m'initier à l'analyse Fermi-LAT et SWIFT/XRT. Merci à Catherine pour les nombreux échanges « Fermi/Wise ». Merci aussi à François et à David de m'avoir impliqué dans l'analyse du blazar PG 1553+113.

Salut à la Team APC primitive, Matthieu, Julian, Romain et Benjamin, ainsi qu'à Flavien, Pierre et Guillaume avec qui j'ai vraiment bien vécu pendant une bonne partie de cette thèse.

Salut aux éléments des différentes équipes, Camill, Pierre, Éléonore, Arthur, Marco, Lucìa, Étienne, Hannah, Vincent, Olivier, Jordan, Julien, Benoît, Astrid et Mehdi sans qui l'aventure, je fais référence à la vie, serait probablement largement beaucoup moins drôle.

Salut à la fratrie, Jérôme, Matthieu, Marie et Clarisse, compagnons d'aventures inaltérables depuis plus d'années que je n'ai de doigts, ainsi qu'à Laùra, Nòra, Véronique, Benoît.

Salut à mes parents pour leur soutien infaillible de tout temps ainsi qu'à l'ensemble de ma famille.

Salut aux absents.

Résumé : Ce manuscrit traite d'astronomie γ de haute énergie (entre $\sim$30 GeV et $\sim$300 GeV) par l'intermédiaire du satellite *Fermi-LAT* et d'astronomie γ de très haute énergie (au-delà de $\sim$100 GeV) via l'expérience *H.E.S.S.* La seconde phase de l'expérience *H.E.S.S.* a débuté en juillet 2012 avec l'ajout d'un cinquième télescope de 28 m de diamètre au réseau initialement constitué de quatre télescopes d'imagerie atmosphérique Tcherenkov de 12 m de diamètre. Dans la première partie de cette thèse, nous présentons le développement d'une analyse en mode hybride basée sur une méthode multivariée dédiée à la détection et à l'étude de sources ayant des caractéristiques spectrales différentes ainsi que les premiers résultats des analyses de données réelles. La seconde partie est dédiée à la recherche des candidats blazars parmi les sources non identifiées du *Fermi-LAT* du catalogue 2FGL. Un premier travail est consacré à l'identification des candidats blazars à l'aide de méthodes multivariées utilisant des variables discriminantes construites à partir des paramètres du catalogue 2FGL. Dans un second temps, à l'aide du catalogue de sources infrarouges obtenues par le satellite *WISE* et d'une méthode non-paramétrique nous recherchons les contreparties de type blazars des sources non-identifiées.
Mots clés : astronomie gamma, noyaux actifs de galaxie, blazars, H.E.S.S., Fermi-LAT, WISE, sources non-identifiées, analyse multivariées

Abstract: This manuscript is about high energy γ-ray astronomy (between $\sim$30 GeV and $\sim$300 GeV) with the *Fermi-LAT* satellite and very high energy γ-ray astronomy (above $\sim$100 GeV) via the *H.E.S.S.* experiment. The second phase of the *H.E.S.S.* experiment began in July 2012 with the inauguration of a fifth 28 m diameter telescope added to the intial array composed of four 12 m diameter imaging atmospheric Cherenkov telescopes. In the first part of this thesis, we present the development of an analysis in hybrid mode based on a multivariate method dedicated to detect and study sources with different spectral shapes and the first analysis results on real data. The second part is dedicated to the research of blazar candidates among the *Fermi-LAT* unidentified sources of the 2FGL catalog. A first development is based on a multivariate approach using discriminant parameters built with the 2FGL catalog parameters. A second development is done with the use of the *WISE* satellite catalog and a non-parametric technic in order to find the blazar-like infrared counterparts of the unidentified sources of the 2FGL catalog.
Key words: gamma-ray astronomy, active galactic nuclei, blazars, H.E.S.S., Fermi-LAT, WISE, unidentified sources, multivariate analysis