Period Search and Forecasting Techniques in Gamma-ray Active Galactic Nuclei with a Time-Domain Formalism

Héctor Rueda

université PARIS-SACLAY

Period Search and Forecasting Techniques in Gamma-ray Active Galactic Nuclei with a Time-Domain Formalism

Utilisation d'un formalisme du domaine temporel pour la recherche de périodes et la prédiction de l'évolution temporelle des noyaux actifs de galaxie

Thèse de doctorat de l'université Paris-Saclay

École doctorale n°576: Particules, Hadrons, Énergie et Noyau: Instrumentation, Imagerie, Cosmos et Simulation (PHENIICS)
Spécialité de doctorat: Science des Astroparticules et Cosmologie
Graduate School: Physique. Référent: Faculté des sciences d'Orsay

Thèse préparée dans l'unité de recherche Département de Physique des Particules (Université Paris-Saclay, CEA), sous la direction de Jean-François GLICENSTEIN, directeur de recherche, le co-encadrement de François BRUN, chargé de recherche

Thèse soutenue à Paris-Saclay, le 28 Septembre 2023, par

Héctor RUEDA RICARTE

Composition du jury

Membres du jury avec voix délibérative

Tiina SUOMIJÄRVI Professeure, Université Paris-Saclay (IJCLab)	Présidente
Elina LINDFORS Assistant professor (HDR), Université de Turku (FINCA)	Rapportrice & Examinatrice
David SÁNCHEZ Chargé de recherche (HDR), Université Savoie Mont Blanc (LAPP)	Rapporteur & Examinateur
Juan-Abel BARRIO Professeur, Université Complutense de Madrid (IPARCOS)	Examinateur
Nathalie BESSON Directrice de recherche, CEA-Saclay (IRFU/DPhP)	Examinatrice
Matteo CERRUTI Maître de conférences, Université Paris Cité (APC)	Examinateur

THESE DE DOCTORAT
NNT : 2023UPASP111

université PARIS-SACLAY

ÉCOLE DOCTORALE
Particules, hadrons, énergie et noyau: instrumentation, imagerie, cosmos et simulation (PHENIICS)

Titre: Utilisation d'un formalisme du domaine temporel pour la recherche de périodes et la prédiction de l'évolution temporelle des noyaux actifs de galaxie

Mots clés: Physique des astroparticules, Astronomie des rayons gamma, Noyaux actifs de galaxies, Recherche de périodes, Analyse des séries temporelles, Techniques de prédiction

Résumé: Cette thèse aborde le problème de l'identification des trous noirs binaires (BBH) en développant et en appliquant des algorithmes de recherche de périodicité aux courbes de lumière des Noyaux Actifs de Galaxie (AGN) à haute énergie (HE). La découverte récente des ondes gravitationnelles (GW) et la détection correspondante de BBH supermassifs (SMBBH) ont rendu ce problème particulièrement pertinent. La détection de périodicités dans les flux AGN peut fournir des indices sur les candidats SMBBH. Pour ce faire, nous devons extraire la période de l'émission en la séparant du bruit stochastique.

Le manuscrit débute par une discussion des différents modèles physiques proposés pour expliquer l'émission périodique des AGN, suivie d'une revue de la littérature sur les sources de rayons gamma périodiques détectées. Deux projets liés aux développements de télescopes à imagerie Tcherenkov atmosphérique (IACTs) sont présentés. Le premier se concentre sur les observations à très grands angles zénithaux avec le système stéréoscopique à haute énergie (H.E.S.S.). Ces observations augmentent la gamme des angles zénithaux utilisés, permettant la détection des rayons gamma les plus énergiques et l'augmentation du temps disponible pour suivre les alertes telles que celles des GW.

Le deuxième projet consiste à calibrer le système de datation des événements de la caméra NectarCAM pour le Cherenkov Telescope Array (CTA), la prochaine génération IACTs. L'objectif principal de la thèse est de développer et d'appliquer des algorithmes de recherche de périodicité aux courbes de lumière des AGN. Une approche dans le domaine temporel décrit le bruit stochastique présent dans l'émission des AGN, en utilisant des modèles de bruit auto-régressifs (AR) pour les données régulièrement espacées et des modèles AR en temps continu (CAR) pour les données irrégulièrement espacées. Un algorithme Markov Chain Monte Carlo (MCMC) est développé pour ajuster les paramètres décrivant les termes périodiques, les dérives linéaires et les points de flux corrélés dus au bruit coloré. La théorie de l'information est utilisée pour comparer l'adéquation aux données observées des différents modèles MCMC, en pénalisant la complexité du modèle. Une méthode de recherche de périodicité dans le domaine spectral est employée pour la validation croisée de nos résultats. Enfin, les composantes stochastiques sont étudiées plus en profondeur en appliquant une analyse de la densité spectrale de puissance (PSD).

Nous appliquons l'algorithme aux données à échantillonnage régulier de 27 candidats AGN périodiques de l'instrument Fermi-LAT. Grâce à ces résultats, nous pouvons étudier la variabilité de la période, fournissant des contraintes pour les modèles physiques. Ensuite, nous utilisons le Fermi-LAT Light Curve Repository, où plus de 1500 courbes de lumière d'AGNs sont disponibles. Certaines de ces sources ont des limites supérieures de flux, qui sont traitées comme des points manquants pour cette analyse. Ainsi, nous disposons d'un échantillon de données irrégulièrement espacées à analyser en utilisant les modèles CAR. Plusieurs nouvelles sources périodiques ont été trouvées. Nous appliquons également l'algorithme aux données de H.E.S.S., dont l'échantillonnage en temps est plus irrégulier. Bien que l'évaluation statistique des périodes dans ce type de données soit difficile, il est possible d'étudier le bruit corrélé et la PSD des sources. Enfin, les paramètres de sortie du MCMC sont utilisés pour prédire le comportement du flux de HE, permettant ainsi d'optimiser les stratégies d'observation. Les techniques de prévision sont validées sur un ensemble de données séparé avant d'être appliquées pour prédire les flux des AGNs.

La thèse se termine par un résumé des résultats du projet, des implications et par des suggestions pour les travaux futurs qui permettraient d'améliorer l'analyse, notre compréhension de la physique impliquée et d'explorer de nouvelles lignes de recherche.

université PARIS-SACLAY
ÉCOLE DOCTORALE
Particules, hadrons, énergie et noyau: instrumentation, imagerie, cosmos et simulation (PHENIICS)

Title: Period Search and Forecasting Techniques in Gamma-ray Active Galactic Nuclei with a Time - Domain Formalism

Keywords: Astroparticle physics, Gamma-ray astronomy, Active Galactic Nuclei, Period search, Time Series analysis, Forecasting techniques

Abstract: This thesis addresses the problem of identifying binary black holes (BBH) by developing and applying period search algorithms to high-energy (HE) Active Galactic Nuclei (AGN) light curves (LCs). The recent discovery of gravitational waves (GW) and the expected supermassive BBH (SMBBH) detection make this problem particularly relevant. Detecting periodic modulations in AGN fluxes can provide clues to potential SMBBH candidates. To achieve this, we need to accurately extract the period from the emission while separating it from other components, especially stochastic noise.

We begin with a discussion of different physical models proposed to explain periodic emission in AGN systems, followed by a literature review of gamma-ray (γ) sources where periodicity or quasi-periodic oscillations have been observed. Two projects related to the development of Imaging Atmospheric Cherenkov Telescopes (IACTs) are presented. The first project focuses on Very Large Zenith Angles observations with the High Energy Stereoscopic System (H.E.S.S.). This improves the possible range of observations zenith angle, which allows for the detection of the highest energy spectrum of the γ-ray sky and increases the amount of time available to follow-up alerts such as for GW. The second project involves calibrating the timing system of the NectarCAM camera for the Cherenkov Telescope Array (CTA), the next generation of IACTs.

The main focus of the thesis is on developing and applying periodicity search algorithms to HE AGN LCs. A time-domain approach is employed to describe the stochastic colored noise present in AGN emission, using auto-regressive (AR) noise models for regularly spaced data and continuous-time AR (CAR) models for irregularly sampled data. A Markov Chain Monte Carlo (MCMC) algorithm is developed to retrieve parameters describing periodic terms, linear drifts, and correlated flux points due to colored noise. Information theory is utilized to compare different MCMC models, considering the fit to the observed data and penalizing model complexity. A spectral-domain periodicity search method based on Lomb-Scargle Periodograms is introduced for cross-validation. Finally, the stochastic components of each source are deeper studied by applying a Power Spectral Density (PSD) analysis.

The algorithm is applied to various γ-ray data sets. First, regularly sampled data from 27 Fermi-LAT periodic AGN candidates are analyzed using AR models. With these results, we can investigate the variability of the period and amplitude of oscillating terms, providing constraints for different physical models. Next, we use the Fermi-LAT Light Curve Repository, where more than 1500 variable γ-ray AGN LCs are available. Some sources from this dataset have flux upper limits, which are treated as missing points for this analysis. Thus, we have a sample of irregularly spaced data to analyze using the CAR models for the stochastic component. Here, several periodic sources are found, among which more than half are new candidates. We also apply the algorithm to H.E.S.S. data, which has significant flux gaps and irregular sampling. Although statistically assessing periodic oscillation in this type of data is challenging, information about correlated noise and the PSD of the H.E.S.S. sources can be obtained. Finally, the MCMC output parameters are used to predict future HE flux behavior, allowing for optimized observation strategies. The forecasting techniques are validated on a separate data set before being applied to predict the timing and amplitude of periodic flux variations in AGNs, particularly relevant for SMBBH.

The thesis concludes with a summary of the project's findings, implications, and suggestions for future work to improve the analysis, enhance our understanding of the physics involved, and explore new lines of research.

Contents

List of Figures

List of Tables

List of Acronyms

1FLE Fermi-LAT First Low Energy Catalog

4FGL Fermi-LAT Fourth Source Catalog

AD Anderson & Darling

ADAF Advection-Dominated Accretion Flow

ADC Analog-to-Digital Converter

AGN Active Galactic Nuclei

AIC Akaike Information Criterion

AR Autoregressive

ARMA Autoregressive Moving Average

ASIC Application-Specific Integrated Circuit

ATel Astronomer's Telegram

BBH Binary Black Hole

BCP Bayesian Change Point

BFP Bayes Factor Periodogram

BSMBH Binary Supermassive Black Hole

CAR Continuous Autoregressive

CARMA Continuous Autoregressive Moving Average

CTA Cherenkov Telescope Array

DAQ Data Acquisition

DEC Declination

DTB Digital Trigger Backplane

EAS Extensive Air Shower

EM Electromagnetic

FE Forecast Error

FEB Front-End Board

FFCLS Flat-Field Calibration Light Source

FPGA Field-Programmable Gate Array

FPM Focal Plane Module

GR Gelman & Rubin

GRV Gaussian Random Variable

GTI Good Time Interval

HAWC High-Altitude Water Cherenkov Observatory

HDI Highest Density Interval

HE High Energy

H.E.S.S. High Energy Stereoscopic System

HEGS H.E.S.S. Extragalactic Survey

HGPS H.E.S.S. Galactic Plane Survey

HVPA High Voltage and pre-Amplification

IACT Imaging Atmospheric Cherenkov Telescope

IR Infrared

ISCO Innermost Stable Circular Orbit

KL Kullback-Leiber

L0 Local Trigger

L1 Level 1 Camera Trigger

L2 Level 2 Camera Trigger

L1A Level 1 Accept

LAT Large Area Telescope

LC Ligh Curve

LCR Light Curve Repository

LED Light-Emitting Diode

LHAASO Large High Altitude Air Shower Observatory

lnBF logarithm of the Bayes Factor

LSP Lomb-Scargle Periodogram

LST Large-sized Telescope

MA Moving Average

MAGIC Major Atmospheric Gamma-ray Imaging Cherenkov

MAPE Mean Absolute Percentage Error

MCMC Markov Chain Monte Carlo

ME Mean Error

MILAGRO Milagro Gamma Ray Observatory

MJD Modified Julian Day

MLE Maximised log-Likelihood Estimator

MST Medium-sized Telescope

NSB Night Sky Background

OU Ornstein-Uhlenbeck

PI Prediction Interval

PMT Photomultiplier Tube

PSD Power Spectral Density

PSF Point Spread Function

PSR Pulsar

PWN Pulsar Wind Nebulae

QPO Quasi-Periodic Oscillations

RA Right Ascension

RMS Root Mean Squared

RMSE Root Mean Squared Error

ROI Region of Interest

SED Spectral Energy Distribution

SMBH Supermassive Black Hole

SNR Supernova Remnants

SST Small-sized Telescope

TIB Trigger Interface Board

TiCkS Timing and Clock-Stamping

TOM Time of Maximum

ToO Target of Opportunity

TS Test Statistic

UCTS Unified Clock and Time-Stamping

UHE Ultra High Energy

UL Upper Limits

UV Ultraviolet

VERITAS Very Energetic Radiation Imaging Telescope Array System

VHE Very High Energy

VLZA Very Large Zenith Angle

WCD Water Cherenkov Detector

WN White Noise

$\mathbf{wPD_{HS}}$ Weighted Prediction Distance of the High State

Synthese

L'émission de photons de haute énergie des Noyaux Actifs de Galaxie (AGN) sont parmi les phénomènes les plus énergétiques de l'univers. Les rayons gamma (γ) de haute énergie (HE) révèlent des informations cruciales sur les environnements extrêmes entourant les trous noirs supermassifs (SMBH). Bien que l'émission de lumière des AGN soit chaotique, des observations récentes ont révélé des motifs temporels périodiques dans certaines courbes de lumière (LC) de haute énergie.

Il est important de comprendre l'origine de ces signaux périodiques, car cela peut nous éclairer sur certains processus physiques fondamentaux des AGN. Différents modèles physiques peuvent expliquer les périodicités des LC, incluant les mouvements orbitaux, la précession, et les instabilités d'accrétion. Leur étude peut améliorer la connaissance de la dynamique de leur moteur central, de la géométrie du jet, et de l'interaction entre les phénomènes d'accrétion et d'écoulement. Un modèle plausible implique des trous noirs binaires supermassifs (BSMBH), qui sont susceptibles de se former lors des fusions de galaxies. Ces systèmes pourraient être responsables de la précession apparente des jets radio et présenteraient une modulation périodique détectable de leurs flux. L'intérêt pour les BSMBH a été ravivé par la détection d'ondes gravitationnelles provenant de fusions de BH de masse stellaire par les collaborations LIGO et VIRGO, et par la perspective d'observer des fusions de BH massifs avec les instruments à venir. Les trous noirs sont identifiés efficacement par leur émission radio ou HE, et l'une des façons les plus simples de rechercher des candidats BSMBH est de détecter des périodicités dans leurs LCs.

Cette thèse présente une étude complète des techniques de recherche de périodicité appliquées aux AGNs à rayons γ HE avec une nouvelle approche dans le domaine temporel. Nous utilisons des données provenant d'observatoires tels que Fermi-LAT et H.E.S.S. pour identifier et caractériser les signaux variables et périodiques dans la population des AGNs. Grâce à ces méthodes, nous visons à identifier les candidats présentant des signaux périodiques significatifs et à étudier leur corrélation avec d'autres propriétés physiques.

La première partie de la thèse est consacrée à une *Introduction à l'astrophysique des hautes énergies*, fournissant un contexte théorique et expérimental essentiel pour encadrer cette étude. Le chapitre 1 commence par un aperçu historique de la découverte des particules à haute énergie et de l'émergence de l'astronomie à haute énergie. Le chapitre se penche ensuite sur les processus physiques qui sous-tendent l'émission de rayons γ, donnant un aperçu des sources et des régions d'où proviennent ces photons. Enfin, nous explorons les populations de sources galactiques et extragalactiques du ciel HE, en décrivant leurs principales caractéristiques et l'origine de leur émission de rayons γ. Dans le chapitre 2, nous présentons les différentes techniques d'observations des rayons gamma et des particules depuis l'espace et depuis le sol, avec leurs avantages et leurs limites. Nous décrivons également brièvement les principales caractéristiques des détecteurs de rayons γ en fonc-

tionnement qui ont été pris en compte dans l'analyse de cette thèse.

Dans cette thèse, j'ai mené plusieurs projets dans le cadre des collaborations gérant les réseaux de télescopes Cherenkov atmosphériques imageurs (IACTs) H.E.S.S. et CTA. Ces projets sont présentés dans la partie *Contributions aux IACTs*.

Dans le chapitre 3, nous testons les performances des observations H.E.S.S. à très grand angle zénithal (VLZA) à travers la détection et l'analyse de la source HESS J2019+368 observée à basse altitude. Les observations VLZA deviennent cruciales lorsque nous tentons de détecter l'émission de rayons γ au-delà de 100 TeV avec les IACT. La faisabilité de cette technique est démontrée par la détection et l'analyse de la région MGRO J2019+37 avec H.E.S.S. Les résultats obtenus avec les réponses actuelles des détecteurs indiquent que les observations VLZA peuvent prolonger le temps d'observation disponible pour les phénomènes transitoires. Cependant, des améliorations des analyses sont encore possibles en termes de résolution angulaire, de correction atmosphérique et de séparation gamma-hadron. Les observations VLZA avec la génération actuelle d'IACT servent également de précurseurs pour les futurs télescopes visant des objectifs scientifiques au-delà de la détection des rayons γ, tels que la détection des neutrinos UHE et les études sur la composition de la masse des rayons cosmiques.

Le chapitre 4 présente les tests sur les performances en temps de la NectarCAM, une caméra qui sera déployée dans un type de télescope CTA. Ces performances seront cruciales pour réduire le bruit de l'image, améliorer le nettoyage de l'image et faire la distinction entre les photons de rayons γ et le bruit de fond des rayons cosmiques. Le système de chronométrage NectarCAM a été évalué et testé dans un environnement contrôlé au CEA Paris-Saclay. Il a été démontré que la caméra répondait aux exigences temporelles de CTA, ce qui est crucial pour l'analyse précise des données collectées. Les expériences sur la précision temporelle d'un pixel seul montrent que la caméra atteint une résolution temporelle inférieure à 1 ns pour des intensités lumineuses supérieures à 10 photons. L'étalonnage du temps de transit du PMT réduit efficacement les décalages temporels entre les pixels, améliorant ainsi la précision temporelle globale de la caméra. Après étalonnage, la précision temporelle globale des pixels répond à l'exigence du CTA de 2 ns pour la valeur efficace de la distribution de la différence de temps mesuré entre deux pixels éclairés simultanément. En outre, la précision temporelle du système de déclenchement de la caméra a été évaluée à l'aide d'une source laser, et les résultats démontrent qu'elle est meilleure que 0.5 ns, ce qui est meilleur que les exigences de CTA. Ces résultats garantissent la fiabilité et la précision du système de datation de NectarCAM pour les futures observations du CTA.

La partie centrale de cette thèse, *Recherche de Périodicité*, a été le développement, le test et l'application d'un algorithme de recherche de périodicité Monte-Carlo Markov Chain (MCMC) utilisant une approche dans le domaine temporel pour des données de rayons γ. La composante stochastique est traitée séparément des composantes déterministes. Elle est décrite par un modèle de bruit autorégressif (AR) pour les données régulières ou par un modèle autorégressif en temps continu (CAR) pour les données échantillonnées de manière irrégulière.

L'utilisation d'une approche dans le domaine temporel offre plusieurs avantages. L'un d'entre eux est la possibilité de séparer les différentes composantes d'une LC, telles que le signal périodique et le bruit stochastique. Nous avons montré que l'ajout d'une com-

posante périodique peut améliorer de manière significative l'adéquation aux données par rapport aux modèles de bruit AR/CAR uniquement. De même, l'ajout d'une composante AR/CAR améliore généralement l'adéquation aux données par rapport aux modèles de bruit blanc (WN). Nous avons également prouvé la capacité de notre méthode à détecter les harmoniques des oscillations et à analyser l'évolution des périodes et des amplitudes dans le temps.

Nous avons ainsi contribué à la détection et à l'analyse de nouveaux candidats AGN périodiques dans les observations Fermi-LAT, et nous avons réalisé des analyses de variabilité et de corrélation pour le ciel extragalactique avec H.E.S.S.. Cette thèse met l'accent sur l'importance de l'analyse multi-longueur d'onde et sur le potentiel d'analyse conjointe avec des données d'origines multiples. Alors que les avantages de l'approche du domaine temporel sont évidents, l'analyse des données de H.E.S.S. met également en évidence les limites de la méthode de recherche sur les données des télescopes terrestres. Les observations irrégulières et limitées obtenues par des télescopes comme H.E.S.S. posent des problèmes par rapport aux capacités de balayage continues de Fermi-LAT. Néanmoins, malgré ces limitations, l'analyse dans le domaine temporel peut encore fournir des informations précieuses et contribuer à notre compréhension des mécanismes physiques à l'origine de l'émission HE des sources astrophysiques. Enfin, nous avons étudié la capacité de l'algorithme à prédire l'évolution future de l'émission gamma. La méthode de prédiction est présentée avec une procédure d'évaluation des prédictions basée sur les données passées déjà observées. Les valeurs d'émission futures sont comparées aux intervalles de confiance basés sur les postérieurs bayésiens du MCMC pour chaque source. Les résultats de cette technique peuvent contribuer à notre compréhension du comportement à long terme des flux AGN et fournir un outil précieux pour la planification des observations et l'étude des phénomènes périodiques.

Avant le travail analytique, le chapitre 5 fournit une vue d'ensemble des modèles théoriques pour l'émission périodique des AGNs, en se concentrant sur le rôle des BSMBHs comme candidats potentiels et avec une discussion plus détaillée des modèles géométriques. Enfin, le chapitre présente une revue complète de la littérature sur les AGN périodiques émettant en rayons γ.

Le chapitre 6 présente le développement d'un algorithme de Monte Carlo par chaîne de Markov (MCMC) utilisé pour estimer les paramètres déterministes et ceux du bruit corrélé dans les LC des AGN à rayons γ HE. Pour commencer, nous avons analysé les caractéristiques des données temporelles et sélectionné les modèles théoriques qui décrivent le mieux les séries temporelles. Nous avons choisi les modèles AR pour la description du bruit stochastique des données régulières, et les modèles CAR pour la description des données échantillonnées de manière irrégulière. Différentes composantes déterministes, telles que les tendances linéaires, saisonnières et périodiques, ont été ajoutées aux modèles. Ensuite, nous avons intégré les différents modèles dans un algorithme d'optimisation par chaîne de MCMC utilisant le langage de programmation R. Ce faisant, nous avons récupéré les paramètres des modèles qui décrivaient le mieux les différentes séries temporelles. La qualité de l'algorithme MCMC a été testée à l'aide de diagnostics testant la précision et efficacité de la convergence du MCMC. L'objectif est d'établir le modèle qui décrit le plus précisément chaque série de données. Nous comparons les modèles à l'aide de la théorie de l'information, ce qui nous permet d'évaluer la dispersion de l'ajustement et la complexité du modèle. La qualité de l'ajustement a été testée en effectuant différents tests statistiques sur les résidus. L'introduction d'une méthode de recherche de périodicité

dans le domaine spectral basée sur les périodogrammes de Lomb-Scargle (LSP) permet une validation croisée efficace.

Dans le chapitre 7, la méthodologie est appliquée à un sous-échantillon d'AGN détectés par le satellite Fermi-LAT, comprenant des sources brillantes avec des LC régulièrement échantillonnées. Seules 14 des 27 sources de cette analyse se sont révélées périodiques, alors que la présence de caractéristiques AR dans les LCs des AGNs pourrait expliquer les oscillations restantes non détectées. La méthode sépare efficacement les composantes stochastiques et périodiques pendant l'ajustement des courbes de lumière. Alternativement, les différentes gammes d'énergie employées pour récupérer les LCs pourraient expliquer cette divergence entre ce travail et les travaux publiés d'autres équipes. L'inclusion d'une composante périodique dans les modèles pour les trois sources les plus significatives améliore l'ajustement des données par rapport aux modèles qui ne prennent en compte que le bruit AR. L'étude examine également les variations de période et d'amplitude en fonction du temps de PG 1553+113 et PKS 2155-304, fournissant un aperçu du comportement de ces sources. Les résultats soutiennent l'applicabilité de divers modèles physiques et éclairent la nature des variations périodiques de flux dans les AGN.

Dans le chapitre 8, nous avons étendu le formalisme utilisé pour la recherche de périodes en introduisant une méthode basée sur le CAR capable d'analyser des données échantillonnées de façon irrégulière. L'utilisation de modèles CAR pour la composante stochastique permet d'inclure les limites supérieures de flux (UL) en tant que points de données manquants dans l'analyse. Une fois encore, cette méthode sépare efficacement les composantes stochastiques et périodiques au cours du processus d'ajustement de la LC. En appliquant cet algorithme à l'ensemble de la population AGN du LCR Fermi-LAT, nous avons découvert plusieurs nouvelles sources périodiques et confirmé des oscillations dans d'autres. La validation croisée avec la méthode spectrale du logiciel Agatha a donné des résultats similaires. En comparant les résultats de périodicité obtenus avec les modèles CAR et les modèles AR précédents, des signaux périodiques cohérents ont été trouvés pour les AGN apparaissant dans les deux analyses. Cependant, des contraintes plus strictes sur la convergence MCMC ont affecté la récupération de modèles périodiques pour certaines sources. La stabilité temporelle de la périodicité est également étudiée dans ce formalisme, ce qui a permis de trouver une diminution drastique de la période de PKS 2155-304. Ce résultat peut être précieux pour relier la source à un modèle théorique et pour contraindre les paramètres physiques décrivant les mécanismes d'émission. Ce chapitre sert de point de départ au chapitre suivant, dans lequel nous avons l'intention d'appliquer des techniques similaires pour comprendre la variabilité des AGN dans le domaine des très hautes énergies (VHE) en utilisant les données H.E.S.S. Les modèles CAR sont fondamentaux pour cette prochaine étape, car H.E.S.S. et d'autres IACTs ont des LCs avec des caractéristiques différentes de celles de Fermi-LAT en raison de leurs conditions d'observation et de leurs performances.

Dans le chapitre 9, la méthodologie est appliquée aux données H.E.S.S., qui présentent des erreurs significatives dans les flux et un échantillonnage irrégulier. Malgré les difficultés liées à l'évaluation statistique des oscillations périodiques dans ces données, des informations précieuses sur le bruit corrélé et la variabilité de la source sont obtenues à partir de leur densité spectrale de puissance (DSP). La DSP représente la variabilité de l'émission du blazar et fournit des informations sur la contribution de la puissance de la variabilité aux différentes échelles d'une série temporelle, représentée dans le domaine des fréquences ($\nu \propto 1/t$). Ici, nous utilisons la DSP pour modéliser le comportement stochas-

tique du bruit de l'émission du blazar après avoir soustrait les tendances déterministes. Nous montrons les résultats pour différents échantillonnages de données et l'effet sur la puissance du bruit pour différentes échelles de temps, une information précieuse pour comprendre les caractéristiques de variabilité des LCs des AGNs. Nous avons également effectué une analyse conjointe de la périodicité Fermi-LAT et H.E.S.S. pour PKS 2155-304. Avec une corrélation non négligeable entre les flux HE et VHE, l'algorithme est capable de retrouver de manière significative et efficace la valeur de la période obtenue dans les chapitres précédents et dans la littérature. Nous avons ainsi mis en évidence les limites des télescopes terrestres à rayons γ pour la recherche de périodes dans le domaine temporel. Néanmoins, avec un échantillonnage de données suffisant, nous avons obtenu des informations précieuses sur le comportement stochastique des sources du ciel extragalactique de H.E.S.S.

Finalement, le chapitre 10 présente une méthode de prévision basée sur l'évaluation par le MCMC des paramètres des séries temporelles analysées afin de produire et d'évaluer les valeurs futures prédites. Tout d'abord, nous avons testé le pouvoir de prévision des flux AGN de Fermi-LAT, en incluant des informations sur leur comportement stochastique grâce au modèle de bruit CAR(1), en incluant leurs tendances déterministes linéaires et périodiques, fondamentales pour la compréhension de la prévision. Nous avons utilisé des simulations pour générer des informations sur les mesures d'erreur attendues, basées sur l'acquisition de connaissances à partir de l'ajustement MCMC des LC des sources. Ensuite, nous avons proposé des valeurs prévisionnelles futures pour toutes les sources analysées, ainsi que les dates oú les flux émis sont les plus faibles et les plus élevés. Les résultats démontrent le pouvoir prédictif de la méthode et son application potentielle dans l'optimisation des stratégies d'observation et la réalisation d'observations multi-longueurs d'onde ou multi-messagers. L'évaluation de la prédiction donne un aperçu de la stabilité temporelle du comportement périodique et permet d'estimer les futurs états d'émission élevés et faibles.

Pour clore cette thèse, dans la partie *Conclusions*, nous discutons des différents résultats obtenus tout au long de la thèse et de leurs implications pour la compréhension des processus astrophysiques sous-jacents. Nous proposons également des suggestions pour de futures directions de recherche afin d'avancer et d'améliorer les techniques de recherche périodique et toutes les possibilités d'applications qu'elles pourraient avoir.

Cette thèse peut servir de point de départ pour d'autres travaux dans ce domaine. Les prochaines étapes possibles comprennent l'analyse de résolutions d'échantillonnage plus fines, l'utilisation de modèles AR/CAR d'ordre supérieur, l'inclusion d'ULs comme contraintes, et l'étude des corrélations entre les paramètres de sortie MCMC et les variables physiques connues des AGNs. L'algorithme développé dans cette thèse a des applications polyvalentes au-delà de l'étude des AGN à rayons γ. Il peut être utilisé pour étudier l'émission de sources galactiques, telles que les binaires ou le Centre Galactique, afin de découvrir ou d'améliorer notre connaissance de leur comportement périodique, en tenant compte du bruit corrélé et des évolutions temporelles déterministes. De plus, l'analyse peut être étendue pour inclure des données provenant d'autres longueurs d'onde, telles que des observations optiques ou radio, afin d'avoir une compréhension plus large des différents mécanismes astrophysiques. En outre, la thèse reconnaît l'importance des futures installations telles que le CTA, qui fourniront des observations plus complètes et plus détaillées pour l'analyse dans le domaine temporel.

Preamble

Active Galactic Nuclei (AGNs) are among the most energetic phenomena in the universe, emitting high-energy (HE) gamma-rays (γ) that reveal crucial insights into the extreme environments surrounding supermassive black holes (SMBH). While the overall emission from AGNs is highly variable, recent observations have revealed periodic temporal patterns in some HE γ-ray light curves (LC). Understanding these periodic signals' origin is essential as it can shed light on the fundamental physical processes in AGNs. Such periodicities can be explained by various models, including orbital motions, precession, and accretion instabilities, and their investigation can improve the knowledge of their central engine dynamics, jet geometry, and the interplay between accretion and outflow phenomena.

One plausible model involves binary supermassive black holes (BSMBH), which are likely to form during galaxy mergers. These binary systems could be responsible for the apparent precession of radio jets and would show detectable periodic modulation of their fluxes. The interest in BSMBH has been renewed by the detection of gravitational waves from stellar-mass BH merger events by the LIGO and VIRGO collaboration and the prospect of observation of massive BH merger events at upcoming instruments. Black hole are most effectively identified by their radio or HE emission, and one of the easiest ways to search for BSMBH candidates is to detect periodicities in their LCs.

This thesis presents a comprehensive study of periodicity search techniques applied to HE γ-ray AGNs with a novel time-domain approach. We utilize data from advanced observatories such as Fermi-LAT and H.E.S.S. to identify and characterize variable and periodic signals in the AGNs population. Through these methods, we aim to identify candidates with significant periodic signals and investigate their correlation with other physical properties.

The first part of the thesis is dedicated to an *Introduction to High-Energy Astrophysics*, providing essential theoretical and experimental background to frame this study. Chapter 1 begins with a historical overview of the discovery of HE particles and the emergence of HE astronomy. The chapter then delves into the physical processes underlying γ-ray emission, providing insights into the sources and regions where these photons originate. Finally, we explore the galactic and extragalactic population of sources in the HE sky, describing their main characteristics and γ-ray emission origins. In Chapter 2 we present the various techniques, with their capabilities and limitations, for observing astrophysical HE photons and particles from space and from the ground. We also provide a description of the main characteristics of the working γ-ray detectors considered in the analysis of this thesis.

As a member of the H.E.S.S. and CTA collaborations of Imaging Atmospheric Cherenkov Telescopes (IACTs) arrays, different projects for this thesis have been conducted. These are presented in the *Contributions to IACTs* part. In Chapter 3, we test the performance of very large zenith angles (VLZA) H.E.S.S. observations through the detection and analysis of the low-altitude source HESS J2019+368. Chapter 4 presents the tests on the timing capabilities of the NectarCAM, a camera that will be deployed in a type of CTA telescope. These timing abilities will be crucial for reducing image noise, improving image cleaning, and distinguishing between γ-ray photons and cosmic-ray background.

The core part of this thesis, *Periodicity Search*, is dedicated to the creation, testing and application of algorithms for finding and predicting deterministic trends in AGN LCs using a time-domain approach. The stochastic behavior is separated from the deterministic trends, and it is described by an autoregressive (AR) noise model for regular data or continuous-time autoregressive (CAR) model for irregularly sampled data. Before the analytical work, Chapter 5 provides an overview of theoretical models for the oscillating emission of AGNs, focusing on the role of BSMBHs as potential candidates and with a more detailed discussion of geometrical models. Finally, the chapter presents a comprehensive review of the literature on periodic γ-ray AGNs. Chapter 6 presents the development of a Markov Chain Monte Carlo (MCMC) algorithm used for retrieving parameters related to periodic terms, linear drifts, and the AR/CAR correlated noise in the LCs of HE γ-ray AGNs. Information theory is employed to compare various MCMC models, taking into account the goodness of fit to the observed data while penalizing excessive model complexity. The introduction of a spectral-domain periodicity search method based on Lomb-Scargle Periodograms (LSP) allows for effective cross-validation. In Chapter 7, the methodology is applied to a sub-sample of Fermi-LAT AGNs, comprising bright sources with regularly sampled LCs. The results from this analysis reveal periodic signals in certain sources. In Chapter 8, the analysis is extended to include Fermi-LAT LCs with irregular sampling from the Fermi-LAT Light Curve Repository. The use of CAR models for the stochastic component allows for the inclusion of flux upper limits (UL) as missing data points in the analysis. The LCR dataset significantly increases the sample size compared to previous work and numerous new periodic sources are identified. For both Fermi-LAT data sets, the temporal stability of these results is examined, providing constrains to the different proposed physical models . Additionally, in Chapter 9 the methodology is applied to H.E.S.S. data, which presents significant flux gaps and irregular sampling. Despite the challenges of statistically assessing periodic oscillations in such data, valuable information about correlated noise and the source variability is obtained from their Power Spectral Density (PSD). Finally, in Chapter 10 the MCMC output parameters are used to predict future HE flux behavior, allowing for optimized observing strategies. These prediction techniques are evaluated using training and validation data sets before being applied to predict specific features and temporal variations in the data.

To close this thesis, we bring the *Conclusions* part, where we discuss the various results obtained throughout the thesis and their implications for understanding the underlying astrophysical processes. We also propose suggestions for future research directions to advance and improve the periodic search techniques and all the application possibilities it has.

Introduction to High-Energy Astrophysics

Chapter 1

γ-ray astronomy

This chapter discusses the knowledge gained from observing γ-rays and related particles in the Universe. After a brief historical overview on the discovery of HE particles and the birth of HE astronomy in Section 1.1, we explain the physical processes underlying the emission of these photons (Section 1.2), which helps us to understand the regions where γ-rays originate. We then move on to the examination of the different types of astrophysical sources that emit HE photons and their distribution in the sky. In particular, in Section 1.3.1 we examine galactic sources, which are found in the Milky Way and provide insights into γ-ray emission in our immediate cosmic neighborhood. In Section 1.3.2, we shift our focus to extragalactic sources located beyond our Galaxy. By comprehensively studying both galactic and extragalactic sources of γ-rays, we gain a deeper understanding of the diverse astrophysical processes associated with γ-ray emission throughout the Universe.

1.1 Brief overview of the birth of γ-ray astronomy

In 1895, Roentgen discovered a type of spontaneous radiation of unknown origin, which he called X-rays (L'Annunziata, 2007). Shortly after, Marie and Pierre Curie conducted a thorough investigation into this new phenomenon and made significant contributions to our understanding of radioactivity. During their research, Henri Becquerel discovered a distinct component of radioactivity, and a few years later, Ernest Rutherford identified a type of radiation that was less able to penetrate matter. Rutherford named these emissions alpha and beta radiation, using the first Greek letters.

In the year 1900, while studying the radioactivity of radium salts, Paul Villard noticed a different kind of radiation that was even more energetic and penetrating than alpha and beta radiation (Gerward & Rassat, 2000). Following the established convention, Rutherford named this new type of radiation gamma radiation in 1903. However, it took some time to fully comprehend its nature. Initially, Rutherford thought that gamma radiation consisted of extremely fast beta particles. Yet, unlike the other types of radiation, gamma radiation did not deflect when exposed to a magnetic field, indicating its lack of electric charge. The understanding of gamma radiation as an electromagnetic (EM) wave began to unfold when William H. Bragg demonstrated in 1910 that this radiation could ionize gas, similar to X-rays. Finally, Rutherford and his colleague Edward Andrade, in 1914, conducted diffraction experiments using a crystal and were able to measure the wavelength of gamma radiation, which turned out to be shorter than that of X-rays, providing

strong evidence for its wave-like nature.

HE astrophysics (Longair, 2011) has a rich historical background that dates back to the early 20th century. One significant milestone in this field was the cosmic ray discovery made by Victor Hess in 1912 (Hess, 2018). During a series of balloon flights, Hess observed that the intensity of ionizing radiation increased with altitude, indicating the presence of HE particles originating from outer space. This groundbreaking finding revolutionized our understanding of cosmic rays and opened up a new realm of exploration in astrophysics. However, the detection and study of cosmic rays posed one main challenge. Since cosmic rays are charged particles, their paths are influenced by magnetic fields, making it difficult to trace their sources accurately.

γ-ray astronomy is fundamental in the study of cosmic rays. γ-rays, are the shortest wavelength EM radiation then the most energetic type of light. Its energy spectrum ranges from 100 keV photons to the observational UL. Thus, γ-rays can provide valuable insights into the origin and acceleration mechanisms of cosmic rays. By detecting γ-rays emitted during HE astrophysical processes researchers can indirectly study the cosmic ray populations and the physical processes associated with them. Over the years, a century later from cosmic ray discovery, HE astrophysics has been devoted to the study of the cosmos by observing and analyzing this type of radiation.

1.2 Physical Mechanisms

The EM radiation can be classified as thermal or non-thermal depending on its production mechanism (Diehl, 2001). Thermal processes occur due to the collective motion of a large number of particles interacting electromagnetically within a body. Astronomical light sources as stars can be approximated as black bodies in thermal equilibrium, emitting radiation whose spectrum depends solely on the temperature of the body, as described by Planck's law:

$$I(\nu) = \frac{8\pi h\nu^3}{c^3} \cdot \frac{1}{e^{\frac{h\nu}{kT}} - 1}$$

This equation represents the energy density of radiation at frequency ν. As the temperature of the body increases, the distribution of radiation shifts to higher energies, resulting in the emission of more energetic photons. Wien's law relates the wavelength at which the peak of the black body spectrum occurs to its temperature: $\lambda_{\text{peak}} T = 2.898 \times 10^{-3}$ m $\cdot$ K. For instance, in the visible spectrum, the peak occurs at a temperature of approximately 6000 K. To obtain photons in the γ-ray region (1 MeV), a thermal source with a temperature of around 2×10^9 K would be required, which exceeds the temperature of any known source in the Universe.

From the above, it can be deduced that γ-rays, due to their frequency range and energy, must arise from non-thermal processes in HE environments involving relativistic particles. Examples of such environments include AGN, pulsars (PSRs), supernova remnants (SNR), among others, which will be further discussed in Sections 1.3.1 and 1.3.2.

We will now delve into the main non-thermal processes for the production of gamma radiation in astrophysical environments (Degrange & Fontaine, 2015), distinguishing between leptonic processes, involving electrons and positrons, and hadronic processes, involving nucleons and mesons.

1.2.1 Leptonic Processes

Synchrotron radiation

Synchrotron radiation (Fig. 1.1, left panel) is produced when a fast-moving charged particle, such as an electron, interacts with a strong magnetic field, causing the particle to orbit around the direction of the field. In environments with strong magnetic fields, such as those surrounding neutron stars, synchrotron photons can even extend into the lower energy range of γ-rays.

Bremsstrahlung radiation

Bremsstrahlung is another important phenomenon involving the interaction of particles with strong EM fields (Fig. 1.1, left panel). It occurs when an electron is deflected by an electric field, often produced by a charged particle such as an atomic nucleus. As a result of this acceleration, the electron emits HE radiation of the order of MeV.

Inverse Compton scattering

Inverse Compton scattering (Fig. 1.1, center panel) refers to the phenomenon in which HE electrons and positrons travelling at relativistic speeds collide with photons present in the surrounding medium. These photons can come from various sources, such as stellar radiation or synchrotron radiation at lower energy levels. During the collision, a significant fraction of the electron's energy is transferred to the photon, resulting in the production of a new photon of higher energy, typically in the X-ray or γ-ray range.

Pair annihilation

When a particle meets its antiparticle, both are annihilated (Fig. 1.1, right panel), resulting in the production of EM radiation. In the specific case of an electron and a positron $e^+e^- \rightarrow 2\gamma$, this annihilation process results in the emission of two photons, each with a minimum energy equal to the rest mass of the particles, which is 0.511 MeV. These photons are therefore in the γ-ray energy range.

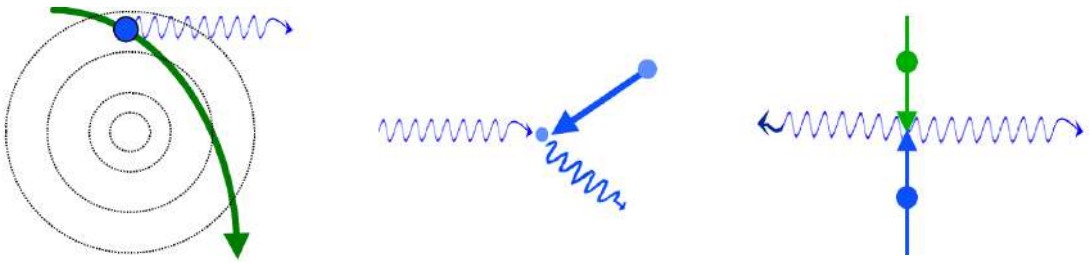

Figure 1.1. (Diehl, 2001) Schematic description of the γ-ray leptonic production mechanisms. [*Left Panel*] Interactions in strong EM fields. [*Center Panel*] Inverse Compton scattering [*Right Panel*] Pair annihilation.

1.2.2 Hadronic Processes

Decay of Neutral Pions

In the denser regions of the interstellar medium, strong interaction events occur where HE protons and nuclei collide with interstellar gas nuclei, producing neutral and charged pions:

$$p + \text{nucleus} \rightarrow \pi^0 + \pi^{\pm}$$

The charged pions decay by two processes into an electron/positron and two neutrinos:

$$\pi^+ \rightarrow \mu^+ + \nu_\mu; \pi^- \rightarrow \mu^- + \bar{\nu}_\mu$$

while the neutral pion quickly decays into a pair of γ-rays:

$$\pi^0 \rightarrow \gamma + \gamma$$

Each γ-ray has a peak in the energy distribution at about 70 MeV, which is half the rest mass of the pion. Observing this energy spectrum therefore gives us information about the most energetic collisions (>135 MeV) between nucleons that have occurred in the Universe.

Nuclear de-excitation

The strong interaction plays a crucial role in binding neutrons and protons together in the atomic nucleus, overcoming the repulsive Coulomb force between them. Similar to electrons in atomic orbitals, the energy states of the nucleus are quantized. The energy difference between an excited state and the ground state of the nucleus is typically in the MeV range. Consequently, any energetic interaction involving the excitation of the nucleus will result in the emission of a γ-ray as the nucleus returns to its ground state (Fig. 1.2). Because of the extreme physical conditions required for such processes, γ-ray emissions are often associated with nucleosynthesis events such as supernovae or interactions involving cosmic rays.

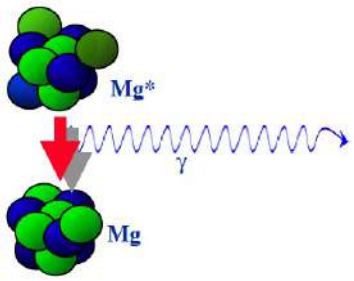

Figure 1.2. (Diehl, 2001) Schematic description of the γ-ray hadronic production mechanism of atomic nuclei de-excitation.

1.3 Astrophysical sources

The Fermi-LAT Fourth Source Catalog (4FGL, Abdollahi et al. (2020)) is a comprehensive catalog of γ-ray sources detected by the Large Area Telescope (LAT) on board NASA's

Fermi Gamma-ray Space Telescope, launched in June 2008 (see Section 2.1.2). The 4FGL third data release (DR3, Abdollahi et al. (2022)) is based on 12 years of Fermi-LAT observations in the energy range from 50 MeV to 1 TeV, containing 6658 sources of which 2157 are unassociated. The 4FGL catalog provides information on the position, spectral properties, and flux characteristics and variability for every detected γ-ray source across the entire sky. As 4FGL sources are included based on the statistical significance of their detection over the entire time period, it does not contain transient sources which are detectable only over a short duration, including Gamma-ray Bursts (GRBs, Ajello et al. (2019)), solar flares (Ackermann et al., 2014), and most novae (Ackermann et al., 2014). In Figure 1.3, the entire sky distribution of 4FGL galactic and extragalactic sources is shown. Along this thesis project, in Chapters 7, 8, 9 and 10 we make an extensive use of the extragalactic information provided by this catalog.

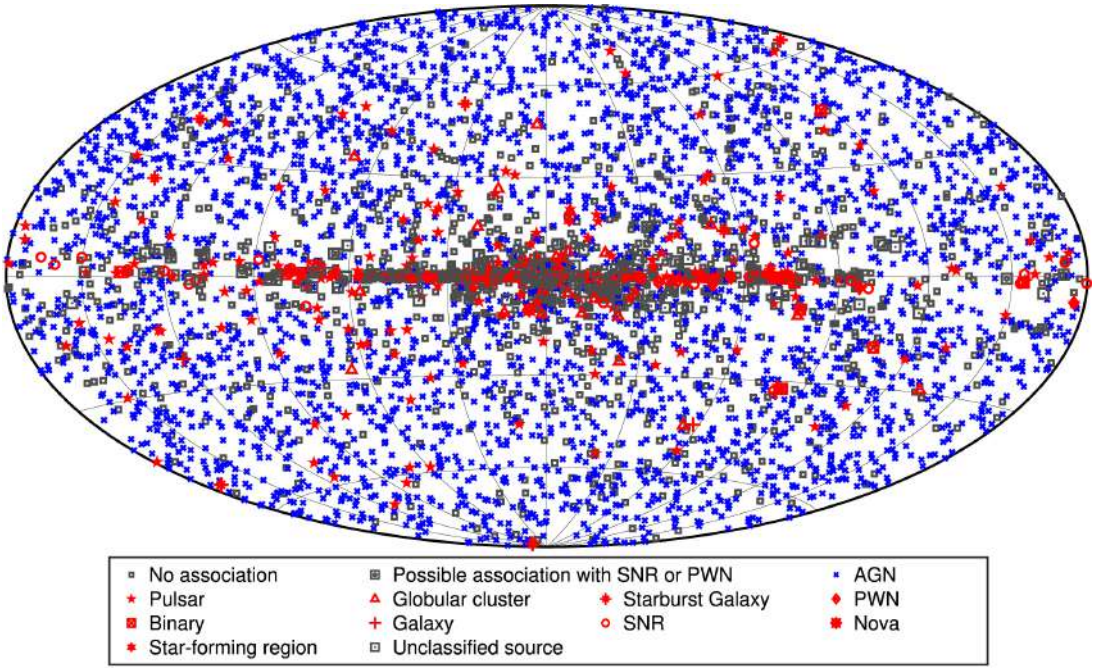

Figure 1.3. (Abdollahi et al., 2020) Full sky map showing the association class of the 4FGL sources.

Hereafter, we list the different 4FGL galactic (Section 1.3.1) and extragalactic (Section 1.3.2) sources, providing a short description and the number of every kind of object. These objects are linked to violent astrophysical environments thus they accelerate particles to extreme speeds, generating non-thermal radiation, including γ-rays, through the mechanisms described in the previous section.

1.3.1 Galactic Sources

From the more than 4501 associated HE sources, only 12 % have a galactic origin, most of which located in the galactic plane (galactic latitude $|b| < 10°$). As we can see form the list below, the most violent γ-ray emitting phenomena hosted by our galaxy are associated with objects in their last stage of evolution.

- **4 Novae (NOV).** A nova is a sudden, transient astronomical event that occurs in the surfaces of white dwarf stars due to mass accretion from the companion in a

binary system. As the accreted material builds up on the surface of the white dwarf, it undergoes a thermonuclear explosion, causing a dramatic increase in brightness. The γ-ray emission can be explained by the interaction of shock-accelerated particles in the explosion with the dense wind of the red giant companion (Tatischeff & Hernanz, 2007).

- **137 Pulsars (PSR).** A PSR is a rapidly rotating neutron star that emits beams of EM radiation. PSRs are dense remnants of massive stars that have exploded as supernovae. They have strong magnetic fields and emit radiation across different wavelengths. Rotations at very high frequencies creates the conditions for HE emission by interacting particles.

- **155 Milisecond Pulsars (MSP).** Millisecond PSRs are a specific type of PSR with extremely rapid rotation periods, typically < 30 ms, that have been spun up by accretion from a binary companion. These objects now represent more than half of the PSRs 4FGL population.

- **43 Supernova remnants (SNR).** Is the nebula structure formed after the massive star explosion as a supernova. These explosive events release an enormous amount of energy and create shockwaves that interact with the surrounding interstellar medium, emitting γ-rays.

- **19 Pulsar wind nebulae (PWN).** Is a supernova remnant fed by a wind of ultrarelativistic particles that is emitted by the magnetic poles of a PSR. γ-rays are generated when the PSR particles interact with those in the remnant shock wave material.

- **114 Supernova remnant / Pulsar wind nebula (SPP).** Indicates potential association with SNR or PWN.

- **35 Globular cluster (GLC).** Is a dense, spherical collection of stars that typically contains hundreds of thousands to millions of stars, with ages of $\sim 10^{10}$ years, forming one of the oldest constituents of the Milky Way (Abdo et al., 2010a). Possible sources of γ-rays in GLCs include MSPs.

- **5 Star-forming region (SFR).** Star-forming regions are areas in galaxies where new stars are being born from the gravitational collapse of molecular clouds. These regions are characterized by intense radiation and energetic processes, making them potential sources of γ-ray emission.

- **7 Binary (BIN).** Linked with known X-rays binaries, these gravitationally bound systems consist of one star, generally in the main sequence, whose material is being accreted by a very compact object, either a neutron star or a black hole. γ rays are believed to be produced in these systems by particles accelerated either within a relativistic jet, or by a PSR wind colliding with the stellar wind and/or the outflowing equatorial disk of the massive star.

- **11 High-mass binary (HMB).** These binary systems are comprised of a massive star, with more than 10 solar masses, and either a neutron star or a black hole.

- **8 Low-mass binary (LMB).** In these systems, the companion star has a mass similar to or less than the Sun.
- **Galactic Center.** It corresponds to the γ-ray observations of Sagittarius A* (4FGL J1745.6-2859), the SMBH in the center of the Milky Way.

1.3.2 Extragalactic Sources

The remaining 88 % of the associates objects in 4FGL is formed by sources of extragalactic origin. Most of these are categorized as AGN classification which is formed by different sub-classes.

AGNs are distant galaxies whose main EM emission is not coming from its stars glow but from its active nuclei. This is explained by the existence of a SMBH, between 10^6 and 10^9 solar masses in the center of the galaxy. When accreting the spiraling disk of material, surrounded by obscuring dust torus, converts the gravitational energy into broad band EM energy in the form of a jet, only occurring in the 10% of the AGNs population. In Figure 1.4, we see illustration of the AGN components along with the description of the AGN unification model (Netzer, 2015), where the underlying physics of the galaxy is the same and observed differences between various AGN classes are due to differences in their orientation with respect to our line of sight (Urry & Padovani, 1995).

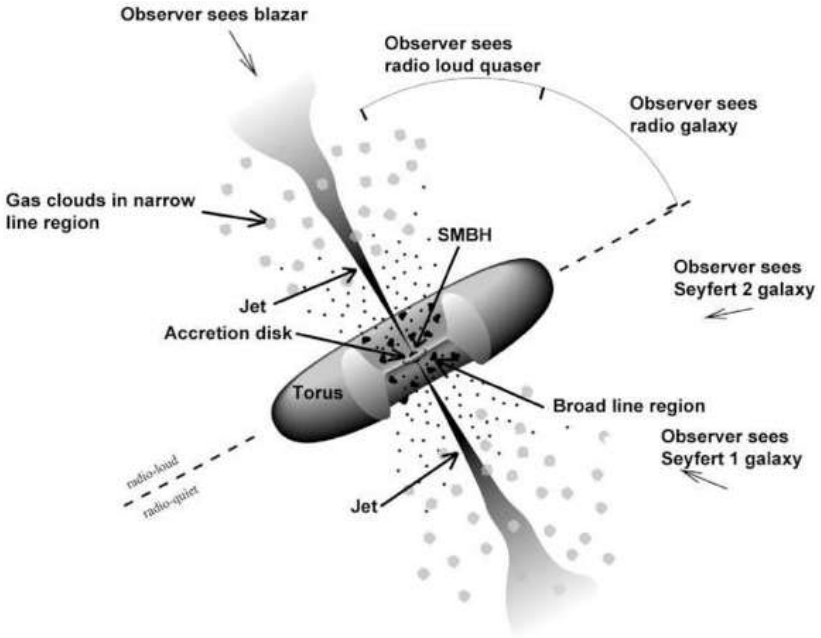

Figure 1.4. (Urry & Padovani, 1995) Schematic representation of the AGN unified model.

Most of the AGNs in the 4FGL catalog are related to blazar types, in which the jets are viewed under a small angle with the line of sight, so we see Doppler boosted emission from these jets. Blazars are known to show variability in all wavebands on times scales from minutes to years, thus they are the main focus of analysis within this thesis.

The typical broadband Spectral Energy Distribution (SED) of a blazar is represented in Figure 1.5, extending from radio to Very High Energy (VHE) photons (E $>$ 200 GeV). We identify a characteristic two-hump shape formed by two different origin peaks: the

first corresponds to the synchrotron frequency peak (ν^S_{peak}) and reflects the maximum energy at which particles can be accelerated in the AGN; the second is usually related to inverse Compton scattering processes processes. With this, the population of blazars can be classified according to the synchrotron frequency peak they exhibit: High-frequency synchrotron peak (HSP) if $\nu^S_{\text{peak}} > 10^{15}$ Hz; Intermediate-frequency synchrotron peak (ISP) if 10^{15} Hz $> \nu^S_{\text{peak}} > 10^{14}$ Hz; and Low-frequency synchrotron peak (LSP) if $\nu^S_{\text{peak}} < 10^{14}$ Hz.

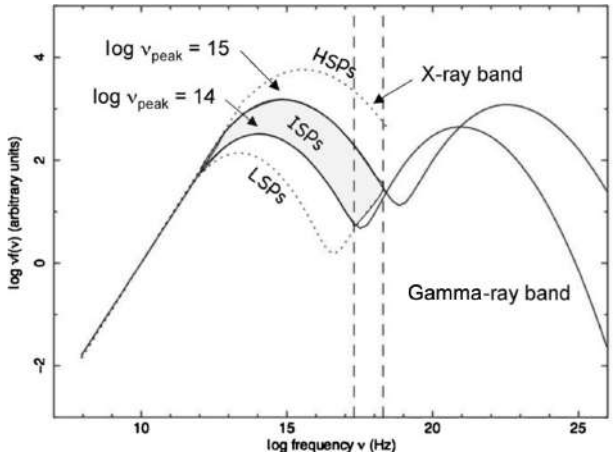

Figure 1.5. (Abdo et al., 2010b) Typical SED of a blazar and illustration of the different types based on the ν^S_{peak}.

The distribution of the 4FGL according to the associated extragalactic object class is as follows:

- **1457 BL Lac (BLL).** Bl Lacertae objects are a subclass of blazars. Their name is derived from their prototype, Bl Lacertae. They are characterized by a spectrum that lacks optical emission or absorption lines, unlike other types of blazars. This characteristic makes them easy to identify, but it also makes it difficult to determine their redshift.

- **794 Flat-Spectrum Radio Quasar (FSRQ).** FSRQs are a class of blazars that show broad emission lines in the visible part of the spectrum. The width of these lines is explained by the rapid rotation of the black hole, which creates a Doppler effect on the radiation emitted by the surrounding material.

- **45 Radio galaxy (RDG).** RDG are a class of AGNs identified for their strong radio emission. These emissions typically have a jet-like structure, where material is ejected from the central core and forms two lobes on opposite sides, perpendicular to the plane of the galaxy, as the observer angle is perpendicular to the jet (see Figure 1.4). This material is composed of particles moving at relativistic speeds that emit synchrotron radiation as they decelerate.

- **2 Steep-Spectrum Radio Quasar (SSRQ).** A SSRQ is a type of AGN that exhibits a steep radio spectrum, meaning that its radio emission decreases rapidly with increasing frequency. Their jets are viewed at larger angles than blazars, hence, the beaming effects of jets should not be severe (Gu & Ai, 2011).

- **5 Compact Steep Spectrum Radio Source (CSS).** CSS are a class of AGN characterized by the presence of radio jets that extend only on galactic scales, typically up to 20 kiloparsecs in size. These sources are commonly found in massive early-type galaxies that have little ongoing star formation (Gordon et al., 2023).
- **9 Nonblazar active galaxy (AGN).**
- **1492 Blazar candidate of uncertain type (BCU).**
- **2 Seyfert galaxy (SEY).** Seyfert galaxies are lower-luminosity AGNs, characterized by strong, highly-ionization emission lines. They can be classified into two types based on spectroscopic observations. Type 1 Seyfert galaxies have broad emission lines, indicating the presence of rapidly expanding hot gas in the central region at velocities of thousands of kilometers per second. In contrast, Type 2 Seyferts have strong emission lines at more moderate velocities.
- **8 Narrow-line Seyfert 1 (NLSY1).** NLSY1 galaxies are a class of AGN that share the properties of Type 1 Seyfert galaxies, but with unique characteristics, including the narrowest Balmer lines, the strongest Fe II emission, and extreme properties in X-rays.
- **8 Starburst galaxy (SBG).** These galaxies contain regions of high star formation rates, thought to be the result of mergers and tidal interactions between gas-rich galaxies. These interactions create denser regions of gas in which young and massive stars form. As a result, these galaxies undergo violent and unstable processes in these dense regions, leading to the production of HE γ-rays.
- **6 Normal galaxy (GAL).** The γ-ray emission can also be linked to star formation regions.
- **134 Unknown (UNK).**

Chapter 2

γ-ray observatories

Understanding the underlying production mechanisms described above is essential to improving the capabilities of telescopes designed to detect this radiation. The photons have HE and deep penetration, rendering traditional reflector-based focusing methods ineffective. Consequently, the methods used to capture these photons are adapted to the processes by which they are produced. As we have observed, these production mechanisms involve HE non-thermal phenomena similar to those found in particle physics. Accordingly, the detectors used in this field of astronomy have analogous characteristics to those used in large-scale particle accelerator projects.

It is essential to consider the opacity of the atmosphere at different wavelengths of the EM spectrum. As shown in Figure 2.1, the atmosphere is transparent to visible light and radio waves between 1 mm and 10 m. However, it becomes completely opaque to wavelengths shorter than visible light, including ultraviolet (UV) and γ-rays. As a result, direct observations of gamma radiation from the Earth's surface are not possible, and it is necessary to use detectors on board space satellites.

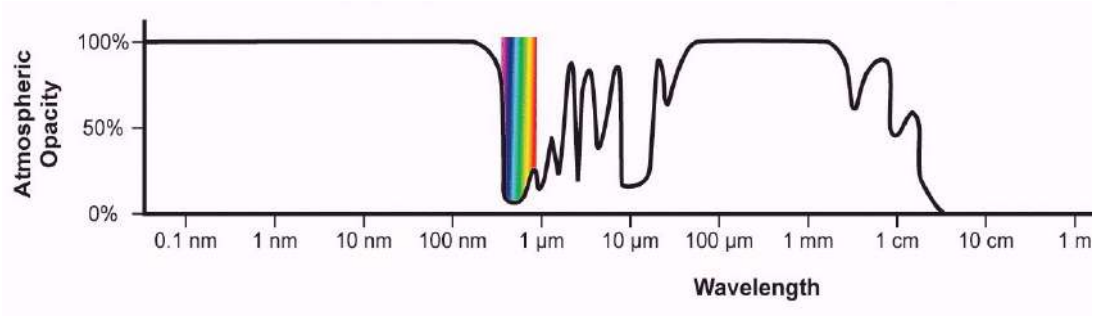

Figure 2.1. Opacity of the atmosphere to EM radiation at different wavelengths.

There are two types of γ-ray detectors: space-based (Section 2.1) and ground-based (Section 2.2). These detectors complement each other and cover a wide range of energies and fluxes. Space-based detectors are suitable for measuring γ-rays in the MeV to mid-GeV energy range, while ground-based detectors are needed for higher energies. Space-based satellites are able to directly detect the incident γ-ray photons while ground-based telescopes rely on the detection of the extensive air showers (EAS) Cherenkov light produced when VHE γ-rays interact with the atmosphere. Also, in Section 2.3, we make a short review of the Water Cherenkov detectors (WCD).

2.1 Space telescopes

2.1.1 Direct detection methods

The dominant interaction process between photons and detectors in the energy range between 1 and 30 MeV is Compton scattering (Fig. 2.2, left). The γ-ray first collides with an electron in an outer scintillation layer, causing the photon to scatter. The scattered photon is then completely absorbed by a second scintillation layer. By analyzing the information recorded in the intermediate photo-electric layers, it is possible to determine the energy and incoming direction of the γ-ray.

For γ-rays with energies above 100 MeV, pair production is the primary detection process (Fig. 2.2, right). These detectors typically consist of several layers. In the initial layers, the γ-ray is converted into an electron-positron pair, which is detected and tracked by subsequent layers. Finally, the pair is absorbed by a calorimeter. By reconstructing the path of the particles, the energy and direction of arrival of the photon can be determined.

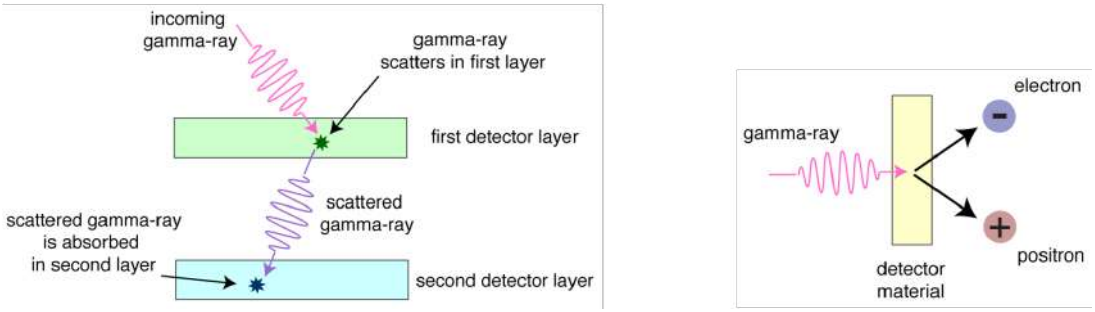

Figure 2.2. (NASA's Imagine the Universe) Illustration of the space-based direct detection methods. [*Left Panel*] Compton scattering detector. [*Right Panel*] Pair-production detector.

2.1.2 The Fermi Large Area Telescope

The Fermi Gamma-ray Space Telescope is a telescope for γ-rays launched by NASA on 11 June 2008 at a 565 km altitude orbit. The LAT (Atwood et al., 2009) is one of the telescope instruments, devoted to the detection of γ-rays within the energy range from 20 MeV to more than 300 GeV. It measures the direction, energy and arrival time of the HE photons detected over the large isotropic background of energetic charged cosmic rays. The Fermi satellite also includes the Gamma-ray Burst Monitor (GBM, Meegan et al. (2009)), aiming to extend the energy range over which GRBs are observed.

The Fermi-LAT is a pair-conversion telescope consisting of a precision converter-tracker and a calorimeter, each with 4×4 modules, as illustrated in Figure 2.3. The incident photon passes freely through an anti-coincidence detector of 89 scintillating tiles. At the same time, charged particles produce a flash as they pass through, allowing them to be distinguished with an efficiency of 0.9997. The photon continues until it interacts with an atom in one of the 16 thin tungsten converter foils, creating an electron-positron pair. These foils are interspersed with 18 layers of silicon tracking material, which allows the trajectory of the electron-positron pair to be reconstructed within the detector by recording the electrical pulses produced by the ionization of the material as the particles pass through. Finally, the particles deposit their remaining energy in a calorimeter consisting

of 8 stacked layers of cesium iodide scintillating crystals. The combined information from the anti-coincidence detector, the conversion-tracking foils and the calorimeter allows the energy, time and direction of incidence of the detected beam to be determined.

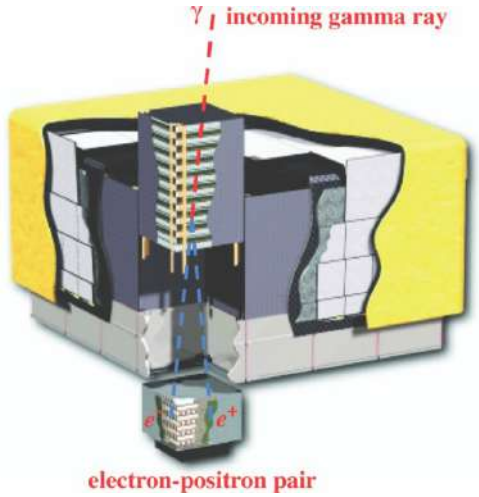

Figure 2.3. Atwood et al. (2009) Schematic diagram of the LAT components and operation.

The Point Spread Function (PSF) of the LAT refers to the spatial distribution of γ-ray events detected by the telescope. It characterizes the instrument's ability to accurately determine the position of a γ-ray source in the sky, and describes how the detected events are spread out around the true position of the source. The 68% containment radius PSF is $\sim 5^\circ$ at 100 MeV and 0.8° at 1 GeV.

The improved Pass 8 (Bruel et al., 2018) is the latest reprocessed data release of the Fermi LAT, which includes improvements and updates to the instrument's calibration and data analysis techniques. It provides better background rejection techniques, allowing for improved sensitivity and angular resolution compared to previous data releases. The Pass 8 data includes more accurate modeling of the instrument response, including the PSF, resulting in better characterization of the sources detected by the LAT. For Pass 8, the energy resolution is $<$10% between 1 GeV and 100 GeV. Below 1 GeV the energy resolution degrades by $\sim$20% at 100 MeV and $\sim$28% at 30 MeV.

2.2 Imaging Atmospheric Cherenkov Telescopes

IACTs are ground-based observatories designed to detect VHE photons from celestial sources. They are based on the phenomena known as Cherenkov radiation shower or EAS, a process induced in the atmosphere by a primary particle, either a γ or a cosmic ray (de Naurois & Mazin, 2015).

When a VHE γ-ray photon or a cosmic ray enters the Earth's atmosphere and interacts with a nucleus or molecule, typically at an altitude of 10-20 km, it undergoes a process where it is transformed into an electron-positron pair. These newly created charged particles continue to interact with atmospheric particles through processes like Bremsstrahlung and Compton scattering, causing them to lose energy and emit additional VHE photons. This cascade of interactions is repeated multiple times, resulting in a shower or cascade of VHE photons and particles. In Figure 2.4 we can see the different cascade shapes depending on the primary particle and the energy of incidence. Cosmic ray showers width

are broadened as compared to γ-ray EAS. The distinct characteristics of shower morphology, combined with the ability to reconstruct the direction of the primary particle, enable IACTs to effectively differentiate γ-ray photons from the isotropic background of cosmic rays. This discrimination is crucial in the detection and identification of the astrophysical γ-ray signals.

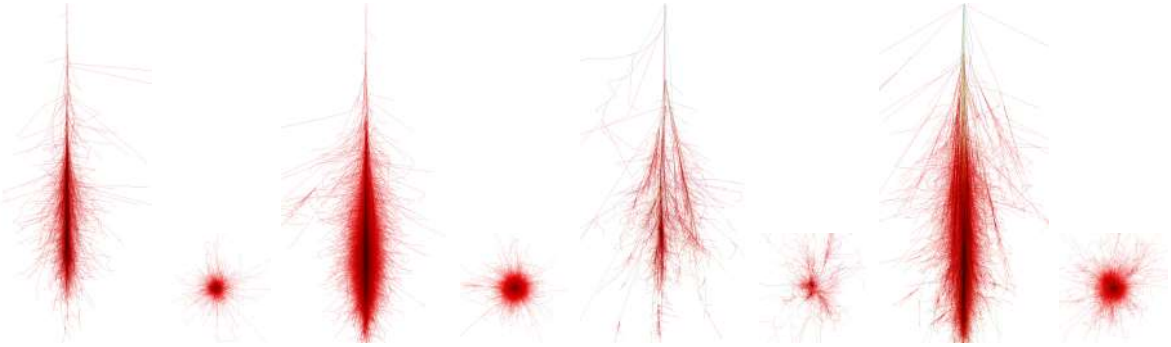

Figure 2.4. (Heck et al., 2012) EAS xz and xy projections of [*Left panel*] Photon shower of 100 GeV. [*Left-Center panel*] Photon shower of 1 TeV. [*Right-Center panel*] Proton shower of 100 GeV. [*Right panel*] Proton shower of 1 TeV.

As these charged particles travel through the atmosphere at speeds greater than the speed of light in that medium (v > c/n, where n is the refractive index), they induce an asymmetric polarization of air molecules (Fig. 2.5, left), generating a shock wave that leads to the emission of Cherenkov photons (Fig. 2.5, right). These Cherenkov photons cover a range of wavelengths from infrared (IR) to UV.

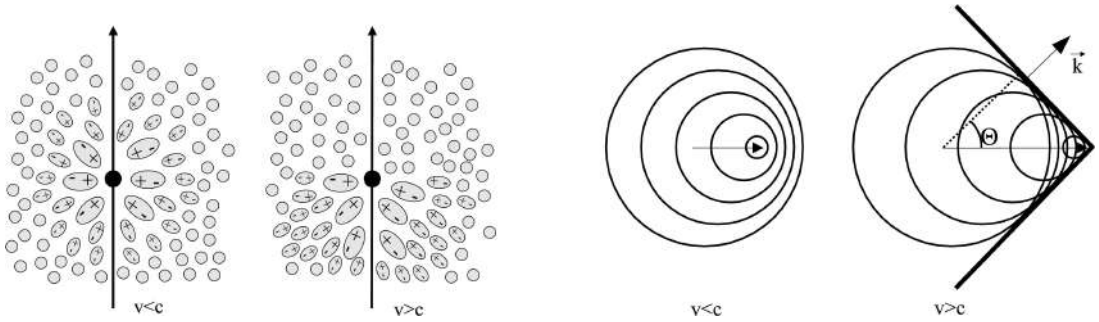

Figure 2.5. (de Naurois & Mazin, 2015) [*Left panel*] Polarisation of the medium by a passing relativistic particle. [*Right panel*] Cherenkov wave front formation.

The Cherenkov light is emitted in a cone shape with a certain angle, and the base of this cone, known as the Cherenkov light pool, can reach surfaces of $\sim 10^5\ m^2$. The IACTs, with their large effective area, then detect the Cherenkov photon pools from each shower, which arrive in a short pulse of a few ns duration (Fig. 2.6). These telescopes are equipped with large mirror surfaces that use large parabolic segmented reflectors to reflect and focus the light into their focal plane, forming an image of the shower. In the focal plane of the IACT camera we find an array of fast photomultiplier tubes (PMTs), or pixels, which digitise the image when the signal crosses the energy threshold to trigger recording, coupled to GHz sampling electronics. By measuring the properties of the Cherenkov light, such as its arrival time and intensity, IACTs can reconstruct the energy and direction of the incoming γ-rays. As in the example in Figure 2.6, the most effective

recording condition is as an array of telescopes in stereoscopic mode. If two or more of these detectors record the same event, then a stereo image of the shower can be obtained. With this, a larger area is covered, allowing a better reconstruction of the event and improving the angular resolution.

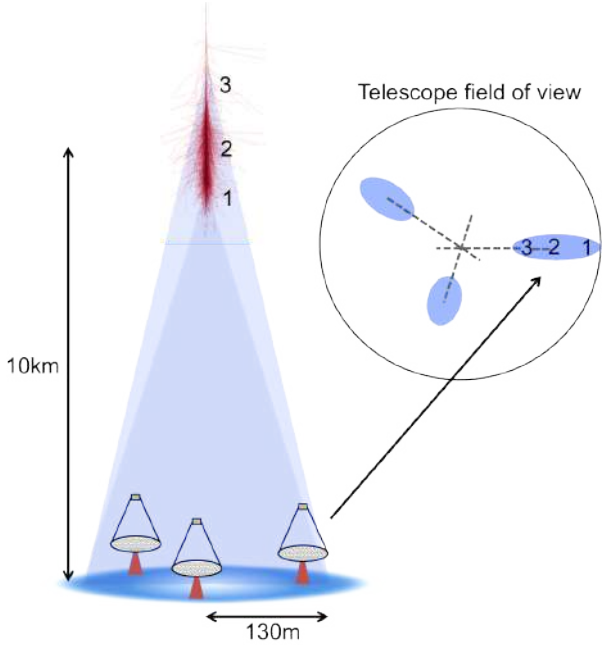

Figure 2.6. (Holder, 2015) Schematic representation of the stereoscopic imaging technique.

The current working generation of IACTs is designed to have the highest sensitivity in the VHE energy band (50 GeV - 50 TeV), with maximum sensitivity from 100 GeV to 10 TeV. They consist of: the Very Energetic Radiation Imaging Telescope Array System (VERITAS, Weekes et al. (2002)), located in Arizona, consisting of four 110 m^2 telescopes in the ∼50 GeV to 50 TeV range, operational since January 2007; the Major Atmospheric Gamma-ray Imaging Cherenkov (MAGIC, Mirzoyan (2004)), consisting of two 17 m diameter telescopes in the ∼ 50 GeV to 30 TeV range, operational since 2009; and the High Energy Stereoscopic System (H.E.S.S.), which is part of this work and will be discussed in the next section. The future generation of IACTS, under construction is the Cherenkov Telescope Array (CTA), described in Section 2.2.2. In Chapter 4 we describe the work done in this thesis for the performance of the NectarCAM, one of the Cherenkov cameras of the CTA.

These VHE observatories effectively complement the Fermi-LAT space telescope. Fermi-LAT covers a larger area of the sky in the HE range, providing a large number of targets for observation by the IACTs, which is crucial due to the limited field of view of these telescopes.

2.2.1 High Energy Stereoscopic System

H.E.S.S. (Hinton & HESS Collaboration, 2004) is located in Namibia observing the Southern Hemisphere sky. It is the IACT with the largest field-of-view and the only one in the Southern hemisphere able to observe the Inner Galaxy and the Galactic Center. The initial four HESS telescopes CT1-4 of the Phase I, completed in 2004, are arranged in the

form of a square having a side length of 120 m, to provide multiple stereoscopic views of air showers. Each telescope of Phase I has a diameter of 13 m, with a total mirror area of 108 m^2 per telescope, 960 PMTs to resolve image details and a field-of-view of about 5°. Starting in 2012, in Phase II of the project, a single huge dish with about 600 m^2 mirror area was added at the center of the array (see Fig. 2.7), increasing the energy coverage, sensitivity and angular resolution of the instrument. The camera of the new telescope contains 2048 pixels.

Figure 2.7. The H.E.S.S. Array of IACTs.

Telescopes are triggered when a minimum number of adjacent pixels in the camera detect a certain amount of light, typically set at 3 adjacent pixels within a 1.3 ns time window, with a threshold of approximately 5 photoelectrons. To reduce Night Sky Background (NSB) contamination, the threshold is increased on nights with higher NSB, such as partial moonlight.

The central triggering system, which operates as the second level of triggering, collects information from each telescope trigger. Its main task is to allow the stereoscopic reconstruction of events by taking into account the coincidence of different arrival times (Balzer et al., 2014). The central trigger decision depends on the operating mode: stereo mode records events triggered by at least two CT1-4 telescopes; mono mode records only CT5 events; and hybrid mode is a combination of stereo and mono modes, recording either with a CT5 trigger or with triggers from more than one CT1-4 telescope. After the Analog-to-Digital Converter (ADC) converts the PMT electrical charge into digital counts for each pixel, the calibration system filters out the NSB photons, allowing the EAS-emitted Cherenkov light to be estimated (Aharonian et al., 2004).

After data acquisition and calibration, the properties of the shower images, such as the direction and energy of the primary particle, are reconstructed. A classical analytical method called Hillas (1985) parameterization is commonly used. This technique relies on the moments of the shower image recorded by the cameras. γ-ray showers often have an elliptical shape, and the Hillas parameters, which include width, length, center of gravity, ellipse orientation, and angular distance between the telescope pointing position and the expected target position, are used to characterize the shower. The major axis of this ellipse indicates the direction of the incoming shower and can be reconstructed with higher accuracy when multiple telescopes image the same shower.

The second approach, called Model++ (de Naurois & Rolland, 2009), uses a semi-analytical shower modeling technique. This technique involves a χ^2 test that compares

raw Cherenkov camera pixel images of a photon-induced atmospheric particle shower with predictions from a semi-analytical model. The analysis chain uses a simulation of the Cherenkov light distribution on the camera and compares it to the measured light distribution in each pixel. The model includes parameters for the longitudinal evolution of the shower, the behavior of the charged particles, the energy and position of the initiating electrons/positrons, the rate and spatial distribution of the Cherenkov photons, and atmospheric factors such as absorption and conditions. In addition, the simulation accounts for instrumental features such as collection efficiency, mirror reflectivity, telescope geometry, photoelectron-to-ADC conversion, response function, integration window, and trigger systems. This reconstruction technique provides a more accurate directional and energy reconstruction of the photon-induced shower compared to Hillas parameterization, resulting in better gamma efficiency, especially at low energies, and improved background rejection. Various data cuts can be imposed in Model++ analysis. A standard cut is used to eliminate 95% of the background events while retaining 70% of the photons. On the other hand, a loose cut can be applied to increase the efficiency of the γ-ray, but also leading to an increased background contamination.

2.2.2 Cherenkov Telescope Array

Based on the technology of current ground-based detectors, the CTA is the major future facility of the IACT under construction (Acharya et al., 2013). It will utilize the imaging atmospheric Cherenkov technique to explore the VHE universe through the detection of γ-rays spanning a range from 20 GeV to 300 TeV. The observatory will consist of several tens of telescopes distributed across two sites: La Palma in Spain, to cover the northern hemisphere sky, and Paranal in Chile, to cover the southern hemisphere. The CTA is designed to improve the capabilities of existing observatories by providing a more extensive collection area and broader energy coverage.

The array will have three size classes of telescopes, illustrated in Figure 2.8:

- **Small-sized telescopes (SST).** The SSTs compact telescopes with a primary mirror based on 18 hexagonal segments with a total aperture of 4.3 meters. The primary function of this reflecting surface is to collect and concentrate the Cherenkov light emitted by γ-rays and direct it to the camera. Inside the camera, PMTs convert the captured light into electrical signals that are then processed by dedicated electronics for further analysis. They are specifically designed to observe the γ-ray sky in the energy range of 5 TeV to 300 TeV. To increase the detection capability for the most energetic events, the SSTs have smaller reflectors and a larger field of view compared to the other telescopes in the CTA. This design optimization aims to maximize the sensitivity of the SSTs to higher energy photons. SSTs will only operate in the southern site, with a large number of telescopes $\sim$ 35 deployed over a large area, covering several square kilometers.

- **Medium-sized telescopes (MST).** The MSTs have a reflector of 11.5 meters and are optimized to detect γ-rays in its core energy range, from about 150 GeV to 5 TeV. The planned configuration consist in 14 telescopes in Chile and 9 in La Palma. MSTs can mount two different cameras: the FlashCAM and the NectarCAM, whose timing capabilities studies are presented in Chapter 4.

- **Large-sized telescopes (LST).** LSTs are the largest type of telescopes within the CTA. These telescopes are characterized by their 45 meters height with a parabolic reflecting surface of 23 meters in diameter. The uniqueness of LSTs lies in their low energy sensitivity, covering the energy range from 20 to 150 GeV, which is closer to the threshold of space experiments such as Fermi-LAT. Since low-energy γ-rays produce a relatively small amount of Cherenkov light, telescopes with large collecting areas are essential for obtaining clear images of these faint signals. As part of the CTA project, four LSTs will be installed in the northern hemisphere to complement the capabilities of the other telescope types in the array.

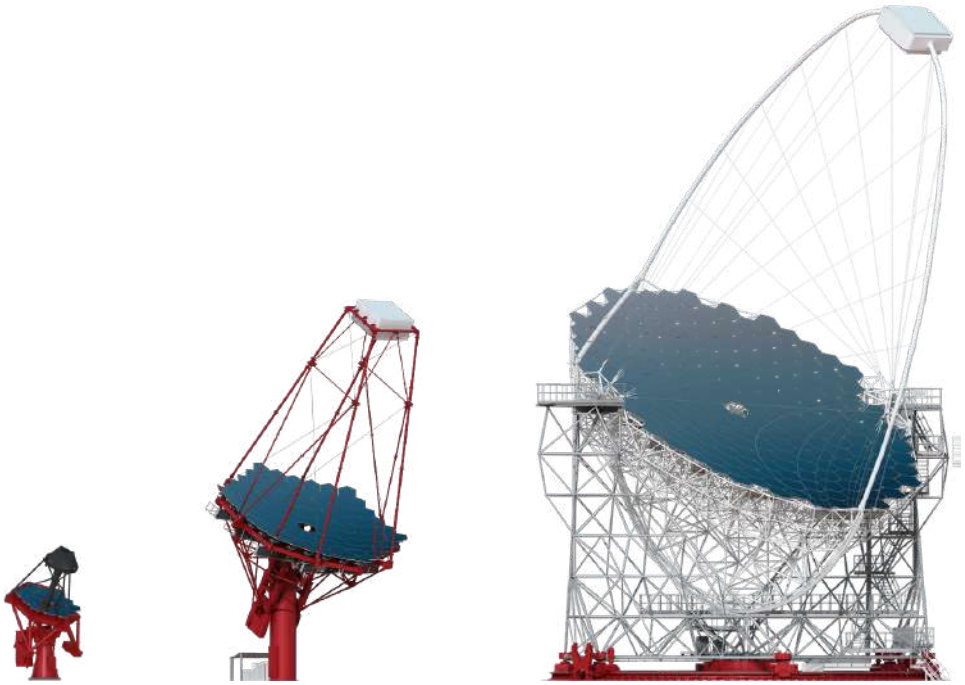

Figure 2.8. (G. Pérez, IAC, SMM) Schematic representation of the three size classes of telescopes of the CTA: SST, MST and LST.

2.3 Water Cherenkov detectors

WCDs offer an alternative approach for the detection of cosmic and γ-rays using the EAS technique operating at a higher energy threshold than IACTs. Instead of directly detecting the Cherenkov light emitted in the atmosphere, WCDs observe the cascade of secondary particles that reach the ground as a result of the EAS induced by γ-rays. This is achieved by employing water tanks on the ground that generate Cherenkov light when traversed by the particles, detected by a series of PMTs.

WCDs are often deployed in the form of arrays, encompassing a large area ranging from 10^4 to 10^5 square meters, with individual detectors spaced a few meters apart. While WCDs have a higher energy threshold and relatively lower angular resolution compared to IACTs, they provide continuous monitoring of the entire sky above the detector. Although they may be less sensitive to point sources, the nearly 100% duty cycle allows WCDs to

capture a wide range of cosmic and γ-ray events, making these detectors well suited to studying transient phenomena.

Milagro Gamma Ray Observatory (MILAGRO, Yodh (1996)), located in the United States, was a pioneering WCD experiment designed to observe large air showers and search for very energetic γ-rays. Located at an altitude of 2530 m, it operated from 2000 to 2008 and played a pivotal role in the field of γ-ray astronomy. MILAGRO's design, with its large water pool and large surface area, enabled the detection of secondary particles produced in air showers and provided a cost-effective solution for observing γ rays in the TeV energy range.

The High-Altitude Water Cherenkov Observatory (HAWC, Westerhoff (2014)) is a γ-ray observatory using high duty cycle, wide field of view WCDs. Located in Mexico, HAWC is capable of detecting γ-rays in the energy range from below 100 GeV to above 100 TeV. The observatory consists of an array of 300 WCDs and is designed to achieve more than an order of magnitude greater sensitivity than its predecessor, MILAGRO. An example of a HAWC water tank is shown in the left panel of Figure 2.9.

Completed in 2021, the Large High Altitude Air Shower Observatory (LHAASO, di Sciascio & Lhaaso Collaboration (2016)) in China is a state-of-the-art observatory dedicated to the study of γ-rays and cosmic rays. It uses a hybrid array that combines several techniques for TeV observations, as shown in the Figure 2.9 right panel. The observatory consists of several key components, including a 1.3 km^2 array known as KM2A, which is composed of EM particle (ED) and muon (MD) detectors. There is also a WCD array (WCDA) with a total active area of 78,000 m^2. In addition, LHAASO has 18 wide field-of-view air Cherenkov telescopes (WFCTA) and a newly proposed electron-neutron detector array (ENDA) covering an area of 10,000 m^2. Together, these components enable LHAASO to make comprehensive and precise observations of VHE particles and photons.

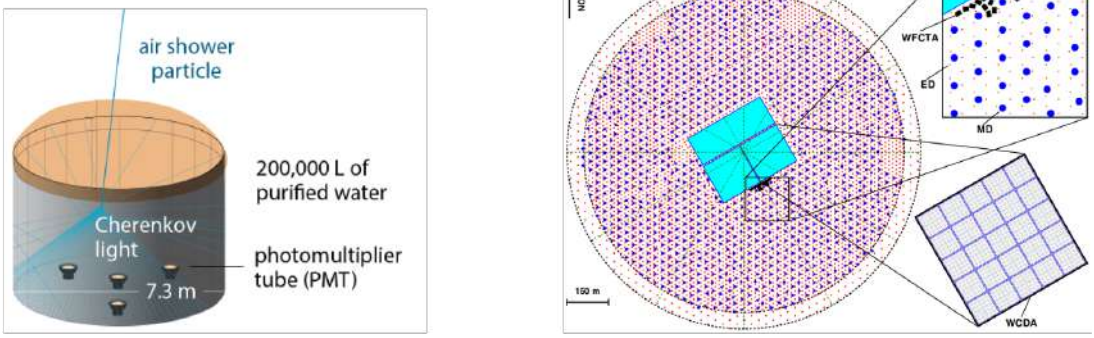

Figure 2.9. [*Left panel*] (HAWC/WIPAC) Schematic representation of the HAWC WCD [*Right panel*] (Liu & Lhaaso Collaboration, 2022) ' Schematic representation of the LHAASO experiment

Contributions to IACTs

Chapter 3

Testing the capabilities of Very Large Zenith Angle detection for H.E.S.S.

This chapter is related to an analysis cross-check made with Paris Analysis for the H.E.S.S. detection of the Very Large Zenith Angle (VLZA) ($\theta > 60°$) low-altitude source MGRO J2019+37. This analysis allows us to test the performance and discuss the limitations of such type of observations for the H.E.S.S. telescopes.

3.1 Why VLZA

Observations conducted at VLZA have the potential to enhance the sensitivity of IACTs to the highest energies since it increases the effective collection area of the instrument. This extends the energy coverage of HESS and enhances its sensitivity to detect rare and VHE phenomena. The capability of VLZA was demonstrated on different sources, as Markarian 421 by H.E.S.S. (Aharonian et al., 2005), the Crab Nebula by MAGIC (Acciari et al., 2020a) and the Galactic Center by MAGIC (Acciari et al., 2020b) and VERITAS (Adams et al., 2021).

The use of VLZA to detect ultra-high-energy (UHE) γ-ray photons using IACTs was proposed by Sommers & Elbert (1987). This approach takes advantage of the increased propagation lengths and projection effects (proportional to $1/\cos\theta$) at larger zenith angles, resulting in a larger Cherenkov light pool. The increase in propagation length from around $\sim$ 10 km to about $\sim$ 100 km leads to a significant expansion of the Cherenkov light pool, from approximately $\sim 0.0001\text{km}^2$ for vertical air-showers to about 1km^2 at zenith angles of about 80° . However, this approach has certain disadvantages, including reduced angular resolution, an increased energy threshold by an order of magnitude, and additional systematic uncertainties due to the increased mass of the atmosphere traversed by the Cherenkov light.

One of the most promising topics related to VLZA observations is PeVatrons. A PeVatron is an astrophysical source or environment capable of accelerating particles up to energies in the peta-electronvolt range (PeV, 10^{15} electronvolts). The particle has to be boosted by extremely powerful cosmic accelerators to reach these energies, with the leading candidates being SMBH, SNR, PSR, and AGNs. Thus, PeVatrons are believed to be responsible for the production of Galaxy VHE cosmic rays observed on Earth.

With H.E.S.S., the detection of PeVatrons candidates up to 100 tera-electronvolts (TeV, 10^{13} electronvolts), was initially achieved by Abramowski et al. (2016) and Abdalla et al. (2021). In this works (see Figure 3.1), they studied the emission of the Galactic Center expected to come from the SMBH Sagittarius A*, and the unidentified Galactic source HESS J1702-420, respectively.

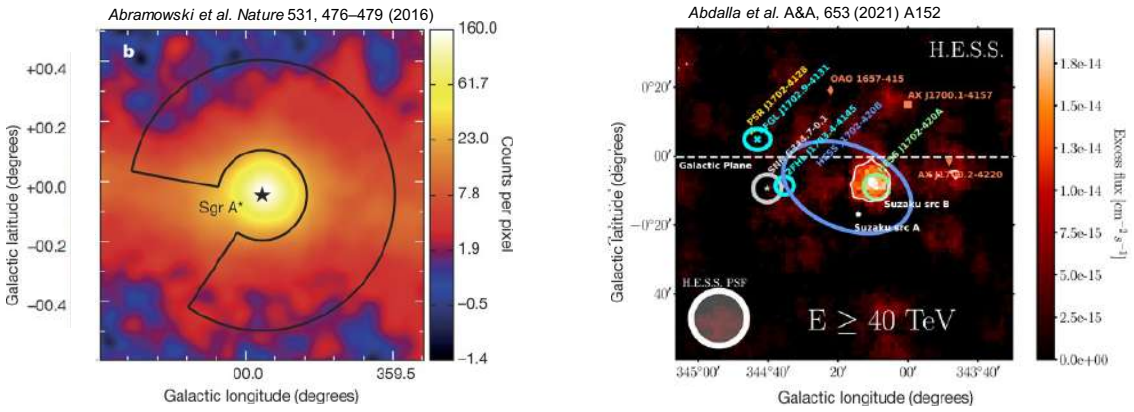

Figure 3.1. [*Left Panel*] H.E.S.S. γ-ray count map of the Galactic Center region. The color scale shows the number of γ-rays per pixel. [*Right Panel*] H.E.S.S. γ-ray flux map of the HESS J1702-420 region, computed with the Ring Background Method above 40 TeV.

For this study, we have selected MGRO J2019+37, an extended region with an unidentified origin possibly related to a PWN. This source was firstly discovered by the MILAGRO γ-ray observatory (Abdo et al., 2007a,b) located toward the rich star formation region Cygnus-X of the Galactic plane, being at the time the second brightest TeV source in the northern hemisphere, after the Crab Nebula.

VERITAS has also detected MGRO J2019+37 (VER J2019+407) as a bright extended source ($\sim 1°$) located near the pulsar PSR J2021+3651 and the star formation region Sh 2-104. They retrieve an spectrum in the 1–30 TeV range well described by a power-law with a photon index of 1.75 ± 0.3. In Figure 3.2, we show the maps of Aliu et al. (2014) and Abeysekara et al. (2018), where they used 70 and 300 observation hours, respectively.

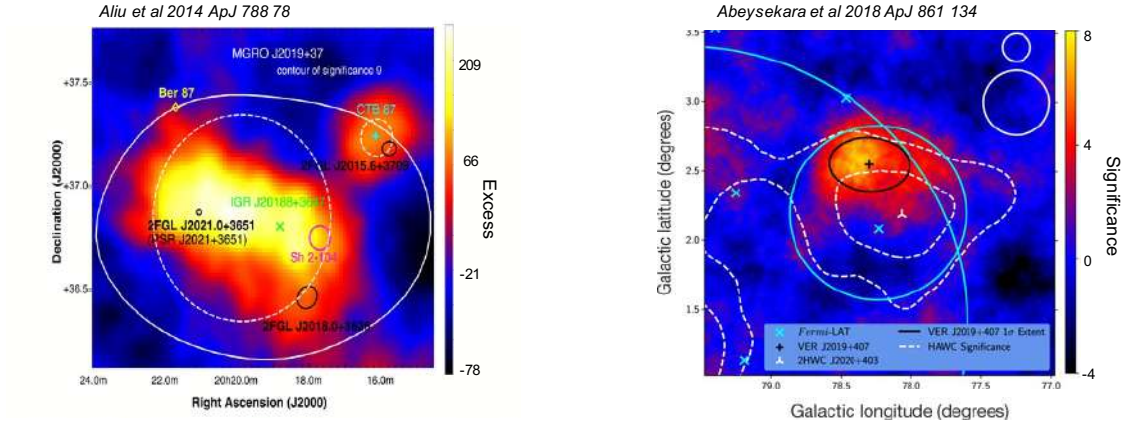

Figure 3.2. [*Left Panel*] VERITAS γ-ray excess map of the VER J2019+407 region above 600 GeV. The color bar indicates the number of excess events within an extended integration radius of 0.23°. [*Right Panel*] VERITAS γ-ray significance map of the VER J2019+407 region integrated with an extended integration radius of 0.23°.

WCDs are valuable for PeVatron astrophysics, as they are designed to detect gamma and cosmic-rays in the highest energy range. Abeysekara et al. (2020) for the HAWC collaboration detected a total of nine PeVatron candidates above 56 TeV, using $\sim$ 1000 days of data from 2105 to 2018. MGRO J2019+37 is one of the three sources with significant signal above 100 TeV, as represented in the left panel of Figure 3.3. The right panel shows the source analysis by Albert et al. (2021) for HAWC, where they resolve the emission into two sources. HAWC J2016+371 associated with low significance with the evolved supernova remnant CTB 87 and HAWC J2019+368 associated with PSR J2021+3651 and its X-ray PWN, the Dragonfly nebula.

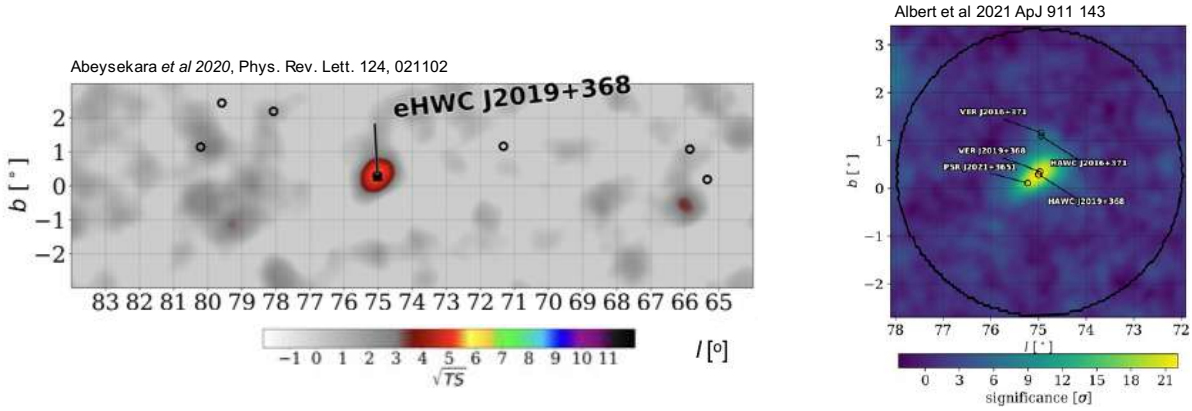

Figure 3.3. [*Left Panel*] HAWC significance map of MGRO J2019+37 in the Galactic plane for an estimated energy > 100 TeV emission. For comparison, black open circles show sources from the 2HWC HAWC Observatory Gamma-Ray Catalog (Abeysekara et al., 2017). [*Right Panel*] HAWC significance map with relevant sources labeled assuming an $\alpha = -2.7$ spectrum.

For the LHAASO collaboration, Cao et al. (2021) presented 12 UHE sources through the detection of $\sim$ 500 photons with energies above 100 TeV and up to 1.4 PeV. In this work, they confirmed all HAWC sources in the overlap region with energy above 100 TeV, including MGRO J2019+37.

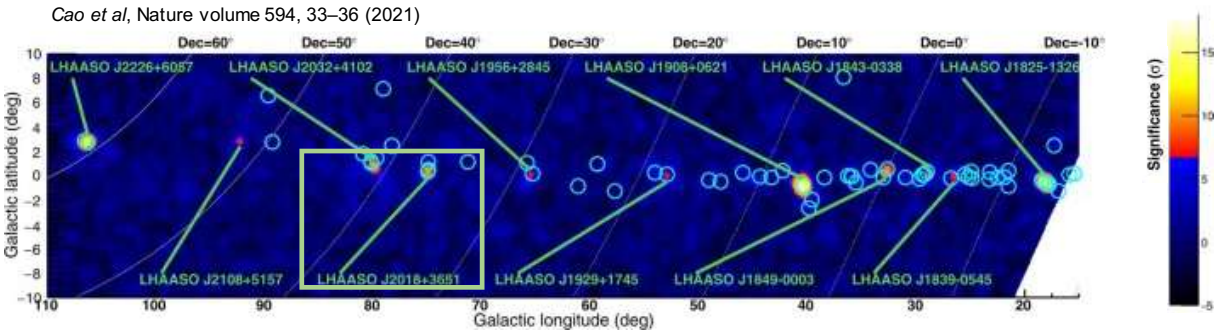

Figure 3.4. LHAASO sky significance map at energies above 100 TeV. The green box remarks the position of MGRO J2019+37.

3.2 H.E.S.S. Detection

The choice for H.E.S.S. is motivated by its visibility only under VLZA, sufficient brightness to be detected in a reasonable time, and its extended dimensions. Given that H.E.S.S. cameras were recently upgraded (Ashton et al., 2020; Bi et al., 2022), these observations are the first ones to validate the VLZA performance of the array in its actual configuration.

MGRO J2019+37 (HESS J2019+368) was observed for about 40 hours in 2020 with zenith angles of $60° < \theta < 67°$ using all five H.E.S.S. telescopes both under dark and moderate moonlight conditions. Since we focus on the VHE detection, in the current work we analyzed the data from four 12-m telescopes using the *standard cuts* of the semi-analytical *Model ++*, which was cross-checked and validated with an independent event calibration and reconstruction analysis (Aharonian et al., 2006). Both pipelines show the detection of the source with high significance exceeding 5 sigma. Significance is calculated according to the Li & Ma (1983) definition.

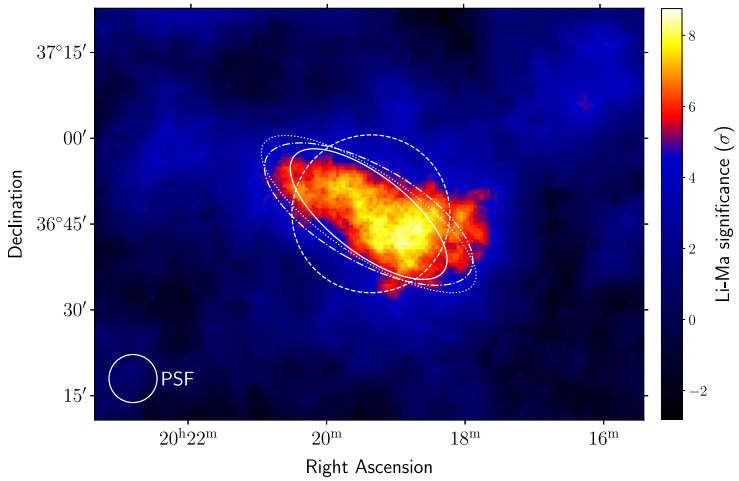

Figure 3.5. Significance map of MGRO J2019+37 in ICRS frame obtained with oversampling of 0.14°, corresponding to 68% containment radius of PSF at these zenith angles. The map includes 2D-Gaussian fit contours of HESS J2019+368 (solid) as well as VER J2019+368 (Abeysekara et al., 2018) (dash-dotted) and HAWC J2019+368 (Albert et al., 2021) (dotted) contours. The dashed circle indicates the spectra extraction region used in Fig. 3.6.

Fig. 3.5 shows the significance map obtained by H.E.S.S. At such high zenith angles, the PSF is significantly degraded and has a 68% containment radius of about 0.14°. We fitted the source with 2D Gaussian and obtained the following results:

$$\text{Right Ascension (RA)} = (304.87 \pm 0.03)°$$
$$\text{Declination (DEC)} = (36.80 \pm 0.01)°$$
$$\sigma_x = (0.27 \pm 0.02)°$$
$$\sigma_y = (0.11 \pm 0.01)°$$
$$\text{and a rotatoin of } 18° \pm 4°$$

It is worth noting that there is still room for improvement since H.E.S.S. analysis is not optimized for VLZA.

For the energy reconstruction, we have chosen the safe threshold of 3 TeV accessible by both reconstruction pipelines and spectrum extraction region from Abeysekara et al.

(2018). The simple power-law fit gives the following results:

$$\text{Flux normalization of } \Phi = (9.42 \pm 0.90) \times 10^{-15} \text{ cm}^{-2} \text{ s}^{-1} \text{ TeV}^{-1}$$
$$\text{at } E = 10.03\,TeV$$
$$\text{with index } \Gamma = 2.25 \pm 0.10$$

The comparison with VERITAS (Abeysekara et al., 2018) and HAWC (Albert et al., 2021) spectra is presented in Fig. 3.6. The H.E.S.S. and VERITAS points are directly comparable since the extraction regions are identical in accordance with VERITAS paper Abeysekara et al. (2018). HAWC points are normalized by the factor of 2.71 for this comparison as proposed in their original paper Albert et al. (2021) after investigation of differences between spectra reconstruction used by HAWC and IACTs. It can be seen that the spectral points are in satisfactory agreement with each other within statistical uncertainties. It is important to note that H.E.S.S. systematical uncertainties for this analysis are larger than those using data obtained with lower zenith angles due to the increased mass of the atmosphere traversed by Cherenkov photons. Using the aerosol data (Holch et al., 2022) for the observation period, we estimated that the Cherenkov photon loss not accounted for by standard analysis is about 10% averaging over the dataset. In the current analysis, we do not apply these corrections and quote the additional systematic uncertainty of this order of magnitude.

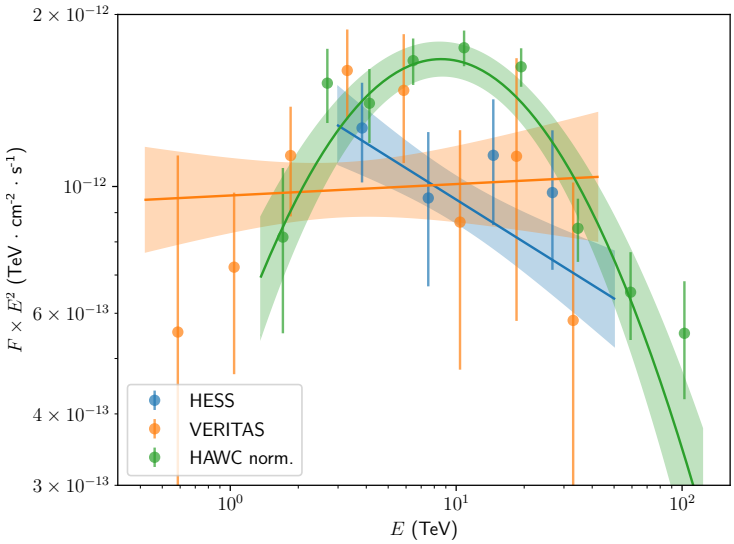

Figure 3.6. Comparison of MGRO J2019+37 spectra. Butterfly plots correspond to best fits as in original works. H.E.S.S. points are fitted with a power-law function. The error bars indicate statistical uncertainties only.

3.3 Discussion

VLZA observations become crucial as we attempt to detect γ-ray emission beyond 100 TeV with IACTs. The feasibility of this technique was shown by different telescopes detecting different sources, galactic and extra-galactic, point-like, and extended ones. This work shows that we obtained satisfactory results using modern cameras and up-to-date detector responses without additional effort. This indicates that this technique could also extend

the observation time available for transient phenomena Abdalla et al. (2019) although most of the transient are detected at lower energies.

The analysis of observations of MGRO J2019+37 region with H.E.S.S. still has room for significant improvement in different directions: improving the angular resolution using more sophisticated methods Parsons & Hinton (2014); Lypova et al. (2021); Abdalla et al. (2021); implementing improved atmospheric correction at VLZA; improving gamma-hadron separation using all five H.E.S.S. telescopes Olivera-Nieto et al. (2021).

VLZA observations with the current generation of IACTs can also serve as a pathfinder not only for CTA Acharya et al. (2018), but also for next-generation telescopes aimed at the scientific goals beyond γ-ray detection, e.g., ultra-HE neutrino detection Ahnen et al. (2018), cosmic-ray mass composition Neronov et al. (2016), and other studies accessible only by measuring highly inclined air-showers.

H.E.S.S. acknowledgments[1]

[1] https://www.mpi-hd.mpg.de/hfm/HESS/pages/publications/auxiliary/HESS-Acknowledgements-2021.html

Chapter 4

The NectarCAM Timing System for CTA

Within this thesis, there is a contribution to testing the timing performances of the NectarCAM (Glicenstein, 2016), a Cherenkov camera that will be installed on MSTs of the northern array of the CTA. This chapter describes the laboratory characterization and calibration of the NectarCAM before deployment, with the fulfillment of the technical CTA timing requirements. These timing abilities will be crucial for reducing image noise, improving image cleaning, and distinguishing between γ-ray photons and cosmic-ray backgrounds. It is also fundamental to enable coincidence identification with neighboring telescopes for stereoscopic observations.

First, we describe the camera and the setup to perform the tests. Then, we describe the different timing accuracy and systematic uncertainties for single pixels and the entire camera. Also, we study the camera trigger timing accuracy (Section 4.6), whose analysis is the main contribution to the NectarCAM done for this thesis. The results of this work are reflected in a publication (Bradascio et al., 2023), and a conference proceeding (Bradascio et al., 2022).

4.1 Describing the camera

The NectarCAM architecture is illustrated in Figure 4.1. NectarCAM employs a modular design comprising 265 modules, each including 7 seven Hamamatsu PMTs with Winston cone light concentrators (Hénault et al., 2013). The PMTs, contained in the camera's focal plane module (FPM) (Tsiahina et al., 2021), detect the Cherenkov light and convert it into an electric signal. The electric signal is then preamplified in the high voltage and pre-amplification (HVPA) boards and split into low and high gain channels. Then, the read-out of the signals is performed in the front-end board (FEB) by the NECTAr chips (Delagnes et al., 2011), where they are digitized and passed through the local trigger (L0) channel. The NECTAr chips store the digitized signal, sampled at 1 GHz, acting as a circular buffer that holds 1 microsecond of data until a camera trigger occurs.

The trigger occurs when enough light is detected within a short period (few ns) in a compact region of the focal plane (Schwanke et al., 2015). First, a neighboring group of pixels is triggered (L0), followed by the camera trigger (level 1 (L1) and level 2 (L2)) through the combination of digitized signals at the digital trigger system. The L0 trigger

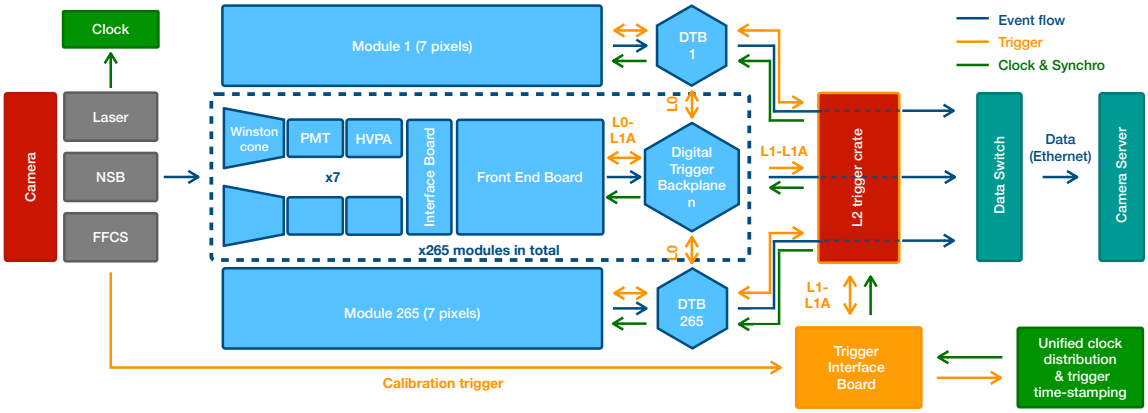

Figure 4.1. Schematic illustration of the signal and trigger chain of NectarCAM. The main detection body of the camera is shown in blue. The camera server is not part of the camera body. The light sources used in the darkroom for the verifications described in this paper are also shown on the left in grey.

is processed in each FEB using application-specific integrated circuits (ASICs). The resulting L0 digital signal is processed in the digital trigger backplane (DTB) using the six neighboring modules, creating a 37-pixel region (see Figure 4.2). With this, the L1 signal is formed, ensuring the trigger homogeneity throughout the camera. Each 265 DTBs Level 1 Accept (L1A) signals are combined in an OR operation to generate the camera trigger (L2) if some configuration conditions are given. The accepted L2 signals are sent to the trigger interface board (TIB) (Tejedor et al., 2022), which returns the signal to the FEBs. Then, the NECTAr chip read-out is stopped, and the 60 ns waveform for every pixel is digitized and stored.

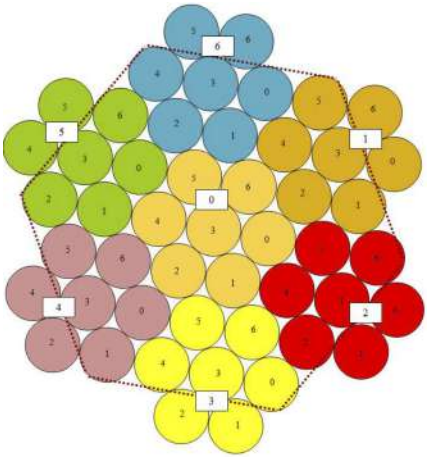

Figure 4.2. Schematic illustration of the 37 neighboring pixels where the L1 trigger is formed by combing the L0 signals. Seven signals are formed in the central cluster (marked with 0) and five signals in each of the six surrounding clusters (cluster 1 to 6). Each circle represents one PMT, numbered from 0 to 6. PMTs with the same color are part of the same module. The dashed black line encloses the 37 pixels used to form the L1 trigger.

To stamp the time of the event, the TIB also sends the trigger signal to the Unified Clock and Time-Stamping (UCTS) module, which incorporates the Timing and Clock-Stamping (TiCkS) board (Champion et al., 2018). When the TiCkS board receives camera trigger signals from the TIB, it stores the absolute time-stamp, trigger-class, and other additional information on the event.

4.2 Test setup

The first NectarCAM camera has been integrated at the CEA Paris-Saclay test facility, as illustrated in Figure 4.3. The facility includes a thermalized dark room with camera services and calibration light sources, connected to a control room housing the data acquisition (DAQ) (Hoffmann et al., 2017), storage, and control systems. The camera comprises readout electronics with 1855 pixels distributed in 265 modules and is connected to a dedicated camera server via 10 Gb/s optical links. A separate camera slow control server handles all devices from the control room.

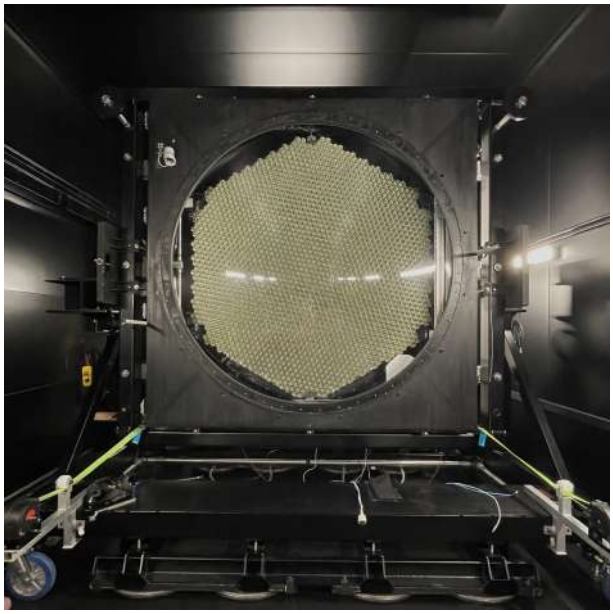

Figure 4.3. NectarCAM camera with entrance window in the dark room of CEA Paris-Saclay. It has a size of $2.9 \times 2.8 \times 1.5$ m and weighs about two tons. The full camera is equipped with 1855 PMTs.

Three light sources, at a distance of 12 m from the center of NectarCAM, were used for camera evaluation: a flat-field calibration light source (FFCLS) with 13 light-emitting diodes (LED); a continuous NSB; and a laser source. The FFCLS emits pulsed light at 390 nm wavelength, reproducing the maximum of the received UV Cherenkov spectrum. The 519 nm LED NSB source reproduces the median photoelectron rate of the typical night sky spectrum. The laser source provides uniform illumination at 373 nm wavelength and is used for precise timing measurements.

The time of maximum (TOM) represents the reconstructed arrival time of a signal from a PMT. The PMT signal is sampled every nanosecond within a 60-nanosecond window. By analyzing the waveform after subtracting the pedestal, the TOM is determined by finding the position of the pulse's maximum.

4.3 Single pixel timing precision

In this section, we evaluate the systematic timing uncertainty of each pixel in the NectarCAM. In order to meet the requirements of the CTA, the NectarCAM camera needs to achieve a single pixel timing precision better than 1 ns for a light illumination above 20 photons (equivalent to 5 photoelectrons).

We conduct experiments by illuminating the camera with a laser source, generating uniform light at frequencies of 1 kHz and intensities ranging from 8.0 nW to 20 nW. The TOM of each photon pulse is measured for every pixel using two methods: identifying the largest peak position and fitting a Gaussian curve to the peak. Both methods yield consistent results within 50 ps. The timing uncertainty of each pixel is determined by calculating the TOM distribution's Root Mean Square (RMS).

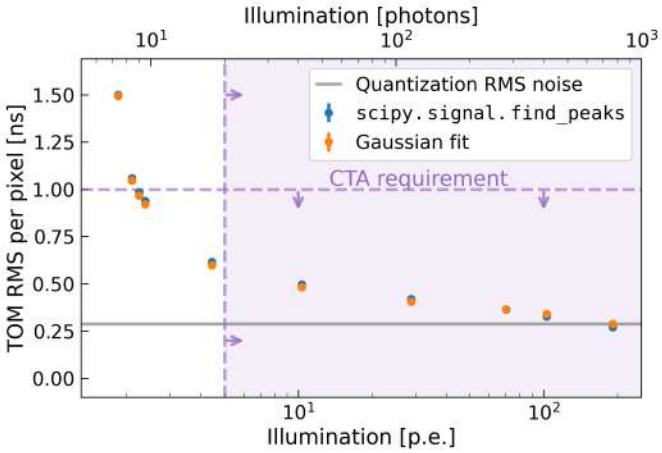

Figure 4.4. Timing precision per pixel (in ns) as a function of the charge of the illumination signal (in photoelectrons and photons on the bottom and top of the horizontal axis, respectively). The timing resolution is given by the mean of the TOM RMS distribution over all the 1855 pixels. Both methods are shown (in blue and orange). The gray solid line shows the quantization (RMS) noise given by $\frac{1}{\sqrt{12}}$ ns. The dashed violet lines and arrows show the 1 ns requirement limit to be valid between 20 (5) and 2000 (500) photons (p.e.) (violet area).

Figure 4.4 displays the weighted mean of the TOM RMS values for all pixels as a function of the illumination charge computed from both methods. It shows that for light intensities above approximately 10 photons, both methods achieve a time resolution of less than 1 ns. This meets the CTA requirement, stating that the RMS uncertainty on the mean relative reconstructed arrival time in each camera pixel should not exceed 1 ns for photon amplitudes ranging from 20 to 2000 photons per pixel. This is represented by the violet area. However, at a light illumination of 800 photons, the pixel time precision reaches its limit (gray dashed line), corresponding to the quantified RMS noise of $1/\sqrt{12}$ ns.

4.4 PMT transit time correction

Prior to estimating the camera's overall timing precision, a relative calibration of the TOM is necessary to synchronize the 1855 pixels in the camera. Two systematic effects cause a timing offset between pixels. Timing offsets are influenced by variations in the arrival times of the Level 1 Accept (L1A) trigger signal sent back from the TIB due to Field-Programmable Gate Array (FPGA) jitter in the Digital Timing Boards (Tavernier et al., 2020). Calibrating the dispersion in L1A delay involves adjusting delays in the digital backplanes and can be performed with an ordinary data taking run.

Another effect is the PMT transit time, which can only be addressed after analysis. The PMT transit time, which refers to the transfer time of the electron avalanche in the

PMT (Leo, 1994) and depends on the high voltage applied to the dynodes, introduces delays in the camera. In order to maintain a nominal gain of 40000 across the camera's detection plane, different voltages are selected for each pixel, resulting in varying delays that typically degrade trigger performance if not corrected for. Particularly for γ-ray energies below 100 GeV (Tejedor et al., 2013), the drop in telescope performance in terms of effective area becomes noticeable.

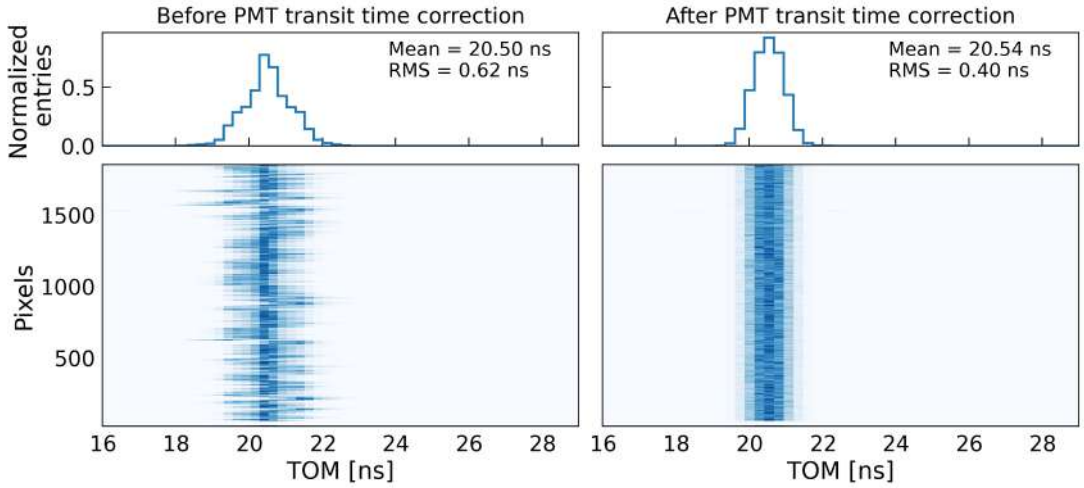

Figure 4.5. Mean TOM distribution for all pixels [*Left Panel*] before and [*Right Panel*] after the PMT transit time correction for a light illumination of $\sim$ 70 p.e. and for a uniform high voltage of 1000 V for all the pixels.

Figure 4.5 left panel illustrates the impact of transit time spread between pixels. It shows that the mean TOM distributions for all pixels are not aligned but exhibit a time shift. To correct this effect, all pixels were set to voltages ranging from 700 V to 1100 V and illuminated with various intensities from 13 LEDs. The dependence of each pixel's TOM as a function of their nominal voltage was evaluated by performing a least-squares fit. Then, the PMT transit time effect was corrected by shifting the TOM values to align with the fit value at 1000 V for each pixel. The right panel of Figure 4.5 depicts the synchronized mean TOM distributions after the correction, where the PMT transit time RMS is significantly reduced.

4.5 Global pixel timing precision

The CTA requirement for the RMS of the TOM difference distribution for any two simultaneously illuminated pixels is 2 ns. To evaluate the timing resolution of the camera after the PMT transit time correction, we illuminate the camera with a $\sim$ 20 p.e. uniform light using the laser source.

For all pixel pairs, the TOM difference (Δt_{TOM}) is calculated with and without PMT transit time correction. Figure 4.6 shows the mean distribution of Δt_{TOM} for each pair of pixels, showing significant improvement after calibration. We see that the RMS of the Δt_{TOM} distribution is reduced from 623.60 ps to 256.30 ps after the correction, meeting the 2 ns CTA requirement.

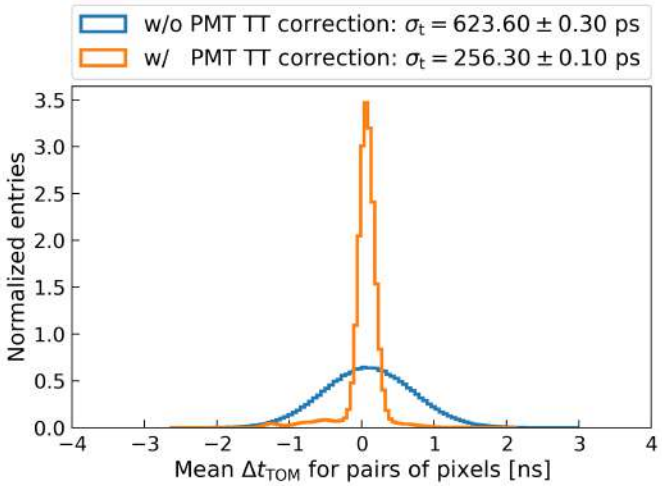

Figure 4.6. Normalized distribution of the mean difference between the TOM value for each couple of pixels over all the events. The distributions with and without PMT transit time correction are shown in orange and blue, respectively. The standard deviation σ_t of the two distributions is shown in the legend and the error is only statistical.

4.6 Camera trigger timing accuracy

As we mentioned, the TiCkS board of the UCTS module is responsible for time-stamping the events captured by the camera. In order to ensure effective coordination and synchronization of observations across the entire CTA array, achieving a high level of timing accuracy is crucial. We illuminate the entire camera using a laser source at 1 kHz with intensities ranging from 20 p.e. to 191 p.e. The start times of the laser flashes were recorded with an external TiCkS board. It also records a date every time a signal triggers the camera. With this configuration, two different analyses are employed to quantify the RMS of the time-stamp distribution.

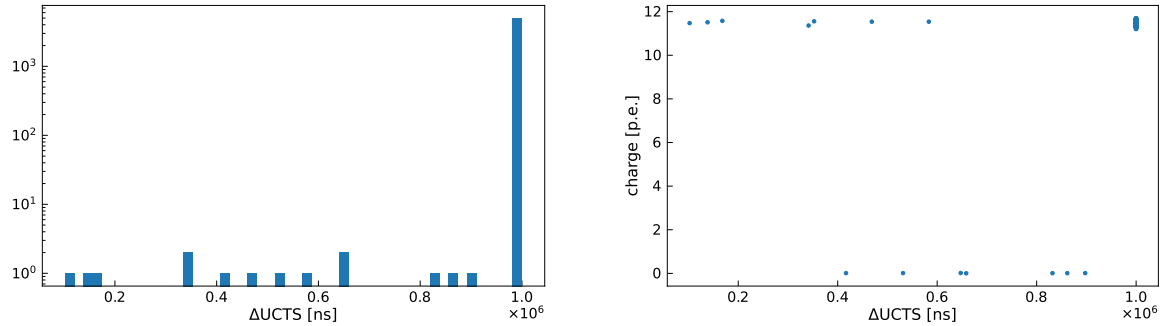

Figure 4.7. [*Left Panel*] Histogram of the $\Delta UCTS$ between two consecutive events for the 5000 events selected in the FFCLS run. [*Right Panel*] Mean camera charge as a function of $\Delta UCTS$.

The first analysis evaluates the time difference distribution between two consecutive events. Given the 1 kHz frequency of the flash light source, every trigger is expected to occur every 1×10^6 ns, this is, $\Delta\text{UCTS} \sim 1 \times 10^6$ ns. This frequency corresponds to the ΔUCTS peak in the top left histogram of Figure 4.7. The ΔUCTS values below the peak correspond to triggers out of sync with the flash, thus not expected nor desirable for technical requirements. In order to discard either a source or a trigger problem, the

pixel charge is analyzed, inferring that these triggers are produced from an external source (e.g., muons). In the right panel of Figure 4.7, the mean camera charge for the i_{th} event is represented as a function of ΔUCTS. This result shows that events in the $\sim 1 \times 10^6$ ns peak are produced by the light source, an external source produces half of the values below the peak, and the other half is produced by the light source after the external source trigger. In Figure 4.8, we show an example of the camera pixel charge triggered by the external source and triggered by the flash.

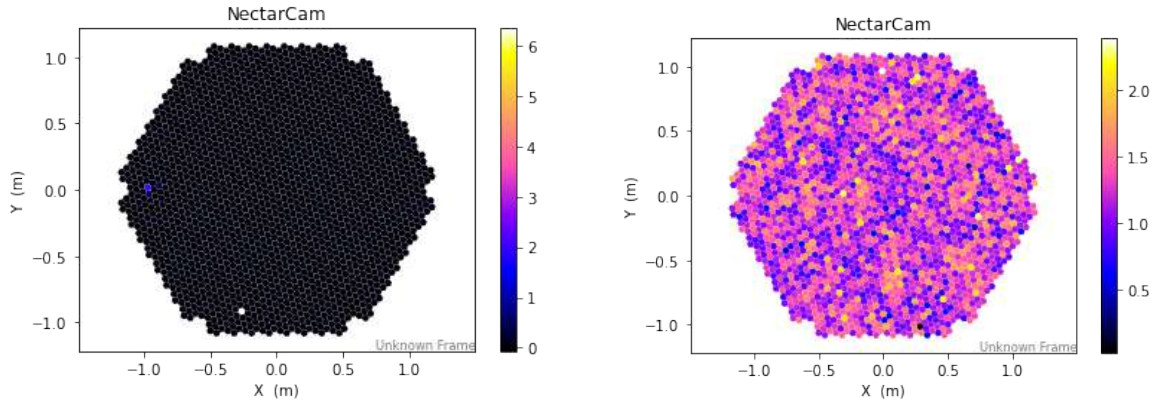

Figure 4.8. [*Left Panel*] Out of flash trigger charge [*Right Panel*] Flash trigger charge.

Thus, external trigger events can be filtered by mean charge value, and the camera trigger timing accuracy can be analyzed. Figure 4.9 represents the filtered ΔUCTS histogram, from which we fit a Gaussian distribution obtaining the RMS uncertainty.

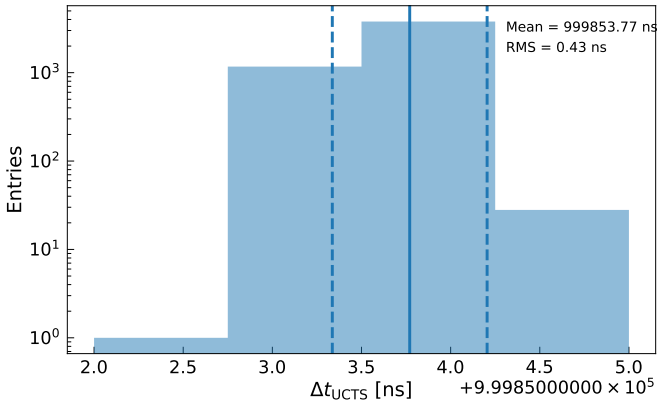

Figure 4.9. Filtered histogram of the UCTS difference with mean and standard deviation.

An alternative method consists in comparing the laser flashes time stamp with the light arrival UCTS in the detection plane of the camera ($\Delta t_{\mathrm{TiCkS}} = t_{\mathrm{UCTS}} - t_{\mathrm{laser}}$). Figure 4.10 presents the RMS values obtained for each configuration. The trigger timing accuracy achieved with both methods is consistent and better than 0.5 ns, below the CTA requirements.

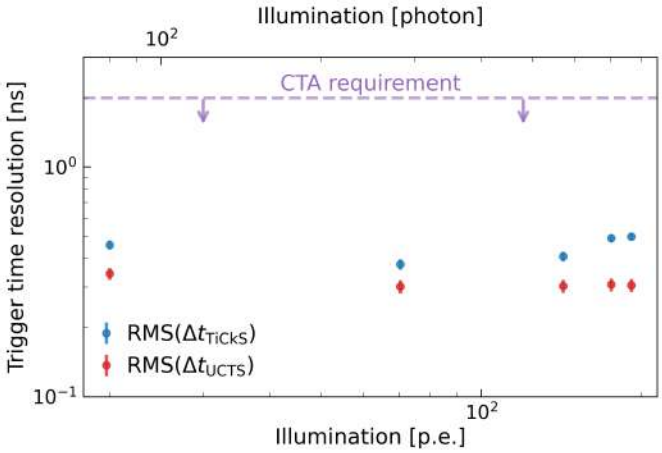

Figure 4.10. Camera trigger precision in ns as a function of the charge of the illumination signal (in photons and photoelectrons on the bottom and top of the horizontal axis, respectively) using two different methods. The CTA requirement of 2 ns for the camera timing resolution is also shown by the violet dashed line.

4.7 Discussion

The NectarCAM timing system has been evaluated and tested in a controlled environment at the CEA Paris-Saclay facility. The camera has been demonstrated to fulfill the CTA timing requirements, which is crucial for accurately analyzing the data collected. The experiments on single-pixel timing precision show that the camera achieves a time resolution of less than 1 ns for light intensities above 10 photons. The PMT transit time calibration effectively reduces the timing offsets between pixels, improving the overall timing precision of the camera. After calibration, the global pixel timing precision satisfies the CTA requirement of 2 ns for the RMS of the TOM difference distribution between any two simultaneously illuminated pixels. Additionally, the camera trigger timing accuracy has been evaluated using a laser source, and the results demonstrate that the trigger timing is better than 0.5 ns, surpassing the CTA requirements.

CTA acknowledgments[1]

[1] https://www.cta-observatory.org/consortium_acknowledgments/

Periodicity Search

Chapter 5

Theory in AGN Periodic Systems

In this Chapter, we describe the different theoretical models that explain the oscillating emission of blazars (Section 5.1), among which BSMBH are natural candidates to explain the observed periodicity in the LCs. Then, in Section 5.1.1, we give a more detailed explanation of the so-called geometrical models for periodic emission. Finally, in Section 5.2, we review the literature on periodic γ-ray AGNs that has emerged since the Fermi-LAT launch in 2008.

5.1 Physical Models

Models explaining the periodicity of HE emission of blazars involve a variety of mechanism from jet precession (Caproni et al. (2013)) to periodic changes in the disk accretion flow (Gracia et al., 2003) transmitted to the jet. An important class of models are geometric models (Rieger, 2004), reviewed in detail in the next section. In geometric models, the periodicity in the observed emission is due to a change in the viewing angle of the jet components, usually related to binary system of SMBHs. In non-geometric models, the periodic change in flux is explained by the variability in the jet itself, involving a single SMBH system.

The lighthouse model of Camenzind & Krockenberger (1992) explains quasi-periodic oscillations (QPOs) with periods of a few dozens days by the rotation of plasma bubbles around the central axis of the jet for a single SMBH. In that model, the observed period increases with time and the period amplitude is also time-dependent (see Fig. 5.1).

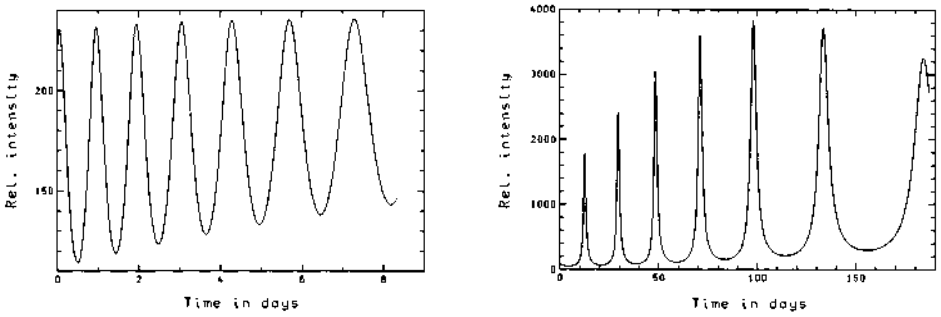

Figure 5.1. (Camenzind & Krockenberger, 1992) Theoretical LCs with a QPO flare as derived from the lighthouse model for [*Left Panel*] 3C 273. [*Right Panel*] Bl Lac object.

In Zhou et al. (2018), they link a similar model to a γ-ray 34.5 days QPO from the blazar PKS 2247-131. The helical structure driven by a coiled magnetic field produces that the density inhomogenities blobs follow the magnetic lines along the jet, as represented in Figure 5.2. With this model, the modulated oscillation is caused by the periodic change of the observer viewing angle.

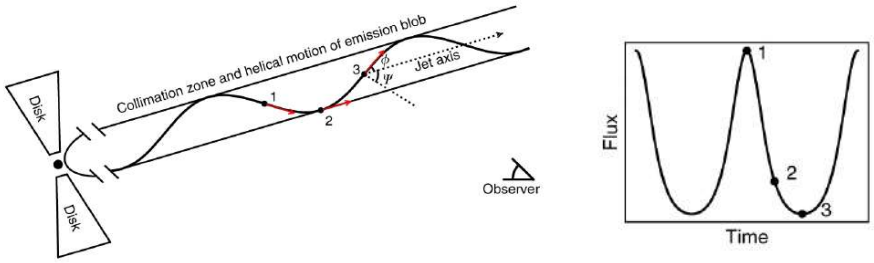

Figure 5.2. (Zhou et al., 2018) Helical jet model that explains AGN γ-ray periodic emission. The scheme shows three different oscillation status related to the pitch and inclination angles (ϕ, Ψ) between the jet and the line of sight.

Another possible mechanism is based on the instability of the boundary (at transition radius r_{tr}) between an outer thin disk (Gracia et al., 2003) or a torus (Zanotti et al., 2003) and an inner radiatively inefficient flow or 'advection-dominated accretion flow' (ADAF). In the model of Gracia et al. (2003), the boundary is slowly moving outward and the period of oscillations increases with time.

Sobacchi et al. (2017) have modelled the HE LC of PG 1553+113 with a geometrical model based on a BSMBH system in which one of the black hole has a precessing jet. In this model, the periodicity is due to the orbital motion of the BSMBH and not expected to change significantly with time while the flux amplitude is related to geometrical properties of the jet, which can change smoothly with time.

Tavani et al. (2018) have interpreted the PG 1553+113 2.2 year periodicity of its HE LC with a model where blazar jets of a BSMBH are periodically perturbed by magneto-gravitational stresses (Cavaliere et al., 2017). In this model, the smaller black hole stresses periodically the jet launched by the heavier one, triggering synchrotron emission and inverse Compton scattering in the GeV energy range. The observed emission could come either from a single jet or from the 2 jets of the BSMBH system. In the latter case, Tavani et al. (2018) predict a stable period with smooth amplitude changes from cycle to cycle, while in the former case these changes are erratic.

5.1.1 Geometrical models

In the HE or VHE domain, it is very likely that the emission comes from the jet. Following Rieger (2004), an important class of models are purely geometric. In this models, the emission is not variable, but the viewing angle of the jet changes with time. The periodic emission could come from:

1. **Orbital motion in a binary black hole (BBH).**

   The period of this motion is given by the Keplerian period:

$$P_k = \frac{2\pi}{\Omega_k} \tag{5.1}$$

with $\Omega_k = \sqrt{G\,[M+m]/d^3}$, where M and m are the primary and companion mass, respectively, and d is the separation of the binary.

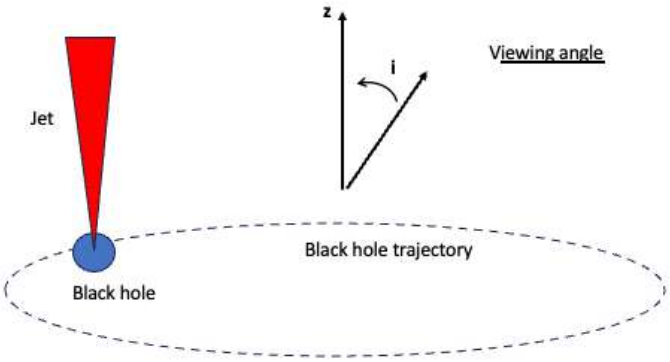

Figure 5.3. Schematic description of the orbital motion in a BBH system model. The black hole with the jet moves around the center of mass of the binary system.

2. **Jet precession.**

The precession of the jet is related to the accretion disk precession of the BH with mass M (Romero et al., 2003). The accretion disk has a radius $r_{\rm d}$. The disk angular velocity is $\omega_{\rm d} = (GM/r_{\rm d}^3)^{1/2}$. Then, the ratio between the orbital and the precessing periods can be related through $\omega_{\rm d}$:

$$\frac{P_{\rm k}}{P_{\rm p}} = \frac{3}{4}\frac{m}{M}\kappa^{3/2}\left(\frac{M}{m+M}\right)^{1/2}\cos\theta, \tag{5.2}$$

where $P_{\rm p} = \frac{2\pi}{\Omega_p}$ and $\kappa = r_{\rm d}/d$. Since $\kappa < 1$, normally $P_{\rm k}/P_{\rm p} < 1$.

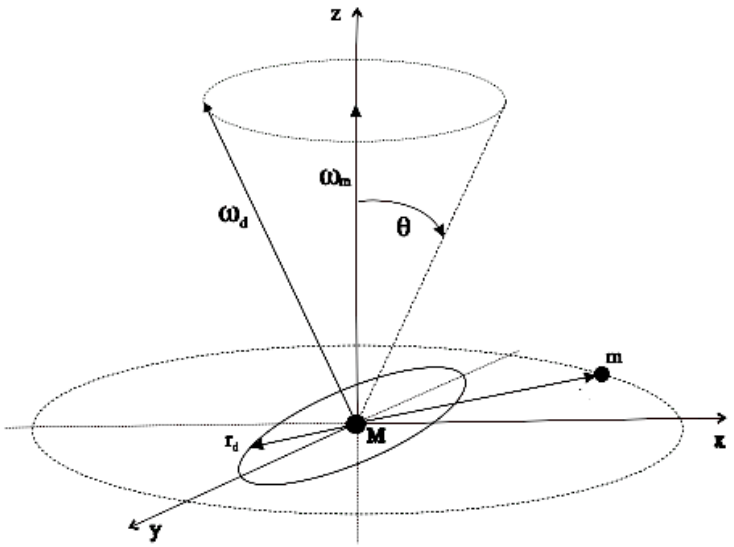

Figure 5.4. (Romero et al., 2003) Schematic description of the driven precessing accretion disk in a BBH system model.

3. **Blob rotation in a jet.**

   This is the case for a bulk flowing plasma close to the innermost stable circular orbit (ISCO) of a BH, consisting of the accretion disk and developing jet which is strongly influenced by the magnetic field as in Zhou et al. (2018) and Mohan & Mangalam (2015), represented in Figure 5.2 and 5.5.

   In this model, the period $P(r_{\rm ISCO})$, depends on $r_{\rm ISCO}$, the radial scale at which the material is injected.

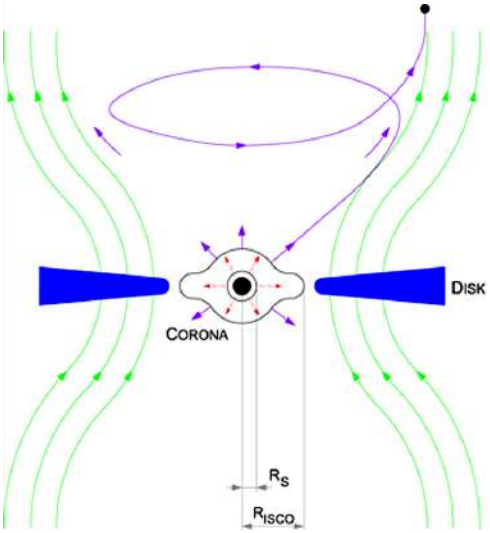

Figure 5.5. Schematic description of the blob rotation in a jet. We see the blob trajectories within the jet influenced by the strong magnetic fields. The black hole has a Schwarzschild radius $R_S = 2GM/c^2$. The $R_{\rm ISCO}$ indicates the bulk injection scale.

In the 2 last cases, the periodic change may not be driven by a companion.

Connection between observed period and true period

Following Rieger (2004), we consider an emitting component which moves relativistically along an idealized helical path at small inclination toward the observer. Denote by $\vec{i} = (\sin i,\, 0,\, \cos i)$ the normalized vector pointing toward the observer and let $\vec{x}_b(t) = (R\,\cos\Omega\, t,\, R\,\sin\Omega\, t\,,\, v_z\, t)$ be the position vector of the emission region, with v_z the outflow velocity along the z-axis.

1. For orbital-driven helical motion, $R \equiv M\, d/(m + M)$, $\Omega \equiv \Omega_k$, and i is the angle between the direction of the total angular momentum and the line-of-sight.

2. Precessional driving may be modelled by setting $R \equiv R(t) = v_z\, t\, \tan\theta$, with $\theta \leq i$ the half opening angle of the precession cone, $\Omega \equiv \Omega_p$ and with i the angle between the cone axis and line of sight.

3. Finally, for a well-collimated, internally rotating jet flow, R may be identified with the radial scale r_0 at which the knot is injected, $\Omega \equiv \Omega(r_0)$ and i is the angle between the jet axis and line of sight.

In general, we measure a peak in the observed flux each time the velocity vector of the component points closest toward the observer, say at point A and point B. In the frame fixed in the center of the galaxy, the component moves from point A to point B within the time span $P = 2\,\pi/\Omega$. Note that the frame fixed in the center of the galaxy is not moving (except for redshift) compared to the observer frame. The observed period is determined by measuring the arrival times of light pulses emitted at A and B. This period however, will usually be much shorter than the actual physical period P, as the travel distance for a pulse emitted at B is much shorter than the one for a pulse emitted at A, due to the relativistic motion of the component in the direction of the observer. The observed difference in arrival times depends on the projection of the velocity vector of the knot $\vec{\beta}_b(t) = \dot{\vec{x}}_b(t)/c$ on the direction $\vec{i}$ to the observer. Hence, for an infinitesimal time interval $\mathrm{d}t$ we have

$$\mathrm{d}t_{\mathrm{obs}} = (1+z)\,[1 - \vec{\beta}_b(t)\cdot\vec{i}\,]\,\mathrm{d}t = (1+z)\;[1 - \beta_b(t)\;\cos\psi(t)\,]\;\mathrm{d}t\,, \tag{5.3}$$

where the factor $(1+z)$ accounts for a possible redshift dependence and $\cos\psi(t) = \vec{i}\cdot\dot{\vec{x}}_b/(|\vec{i}|\,|\dot{\vec{x}}_b|)$. For the relation between the observed period P_{obs} and P one thus obtains

$$P_{\mathrm{obs}} = (1+z)\int_0^P\;[1 - \beta_b(t)\;\cos\psi(t)\,]\;\mathrm{d}t\;. \tag{5.4}$$

1. In the case of orbital-driven helical motion, one has

$$\vec{\beta}_b(t) \;\; = (\tfrac{-R\Omega}{c}\sin\Omega\,t,\;\tfrac{R\Omega}{c}\cos\Omega\,t\,,\;\tfrac{v_z}{c}) \tag{5.5}$$

$$\vec{\beta}_b(t)\dot{\vec{i}} \;\; = \tfrac{-R\Omega}{c}\sin i\sin\Omega\,t + \cos i\tfrac{v_z}{c} \tag{5.6}$$

   so that one finds by integrating over one period that

$$\boxed{P_{\mathrm{obs}} = (1+z)\;\left[1 - \frac{v_z}{c}\;\cos i\right]\;P_k\,,} \tag{5.7}$$

   assuming that the observed emission is dominated by radiation from a single jet.

2. In the case of precessional driving one has

$$\begin{aligned}\vec{\beta}_b(t) = (&\frac{-v_z(\tan\theta)t\Omega}{c}\sin\Omega\,t + \frac{v_z(\tan\theta)}{c}\cos\Omega\,t,\\ &\frac{v_z(\tan\theta)t\Omega}{c}\cos\Omega\,t + \frac{v_z(\tan\theta)}{c}\sin\Omega\,t,\;\frac{v_z}{c})\end{aligned} \tag{5.8}$$

$$\vec{\beta}_b(t)\dot{\vec{i}} \;\; = \tfrac{-v_z(\tan\theta)t\Omega}{c}\sin i\sin(\Omega\,t) + \tfrac{v_z\tan\theta}{c}\sin i\cos(\Omega\,t) + \cos i\tfrac{v_z}{c} \tag{5.9}$$

   Compared to the previous case, the period integral will contain the contribution

$$\Delta P_{\mathrm{obs}} \;\; = (1+z)\textstyle\int_{t_0}^{P_p+t_0}\tfrac{-v_z(\tan\theta)t\Omega}{c}\sin i\sin(\Omega\,t)dt \tag{5.10}$$

$$= -v_z(1+z)\tfrac{(\tan\theta)P_p}{c}\sin i\cos(\Omega\,t_0) \tag{5.11}$$

   so that the relation of P_{obs} to P_p is

$$\boxed{P_{\mathrm{obs}} = (1+z)\;\left[1 - \frac{v_z}{c}\;\cos i\;-\frac{v_z\tan(\theta)\cos(\Omega\,t_0)}{c}\sin i\right]\;P_p\,.} \tag{5.12}$$

3. in the rotational case, one has simply:

$$\boxed{P_{\rm obs} = (1+z)\,\left[1 - \frac{v_z}{c}\,\cos i\right]\,P(r_0)\,.} \tag{5.13}$$

Approximate formulae

v_z can be written as a function of the bulk Lorentz factor γ_b by

$$v_z/c = (1 - 1/\gamma_b^2)^{1/2} \simeq 1 - (1/2)\frac{1}{\gamma_b^2} \tag{5.14}$$

with the second equality being valid in the limit $\gamma_b \gg 1$. One has $\gamma_b \simeq (5-15)$ for typical blazars. For a blazar, one has $i \sim \frac{1}{\gamma_b}$. Then

$$\sin i \quad = i + O(i^3) \tag{5.15}$$
$$\cos i \quad = 1 - \tfrac{i^2}{2} + O(i^3) \tag{5.16}$$
$$\frac{v_z}{c}\,\cos i \quad = 1 - \tfrac{1}{\gamma_b^2} \tag{5.17}$$

The relation between P_k and $P_{\rm obs}$ becomes:

$$\boxed{P_{\rm obs} \simeq (1+z)\frac{1}{\gamma_b^2}P_k\,,} \tag{5.18}$$

In the first mechanism, the observed period is smaller than the true period by a factor typically of 100. The typical separation is

$$\begin{aligned} d &\simeq 3.6\ 10^{17}{\rm cm}\,(P_k/200{\rm yr})^{2/3}\,\left((M+m)/5\,10^8 M_\odot\right)^{1/3} \\ &\simeq 0.1{\rm pc}\,(P_k/200{\rm yr})^{2/3}\,\left((M+m)/5\,10^8 M_\odot\right)^{1/3} \end{aligned} \tag{5.19}$$

From here, an interesting quantity to calculate is the lifetime of the binary. The binary decays by emission of gravitational radiation on a timescale (Shapiro & Teukolsky, 1983)

$$\tau_{\rm gr} = \frac{5}{256}\frac{c^5}{G^3}\frac{d^4}{Mm(M+m)} \tag{5.20}$$

Introducing the Schwarschild radii of M and m, and using seconds as length units

$$r_S^M \quad = 2\tfrac{GM}{c^2} = 5000{\rm s}\left(\tfrac{M}{5\,10^8 M_\odot}\right) \tag{5.21}$$
$$r_S^m \quad = 2\tfrac{Gm}{c^2} = 5000{\rm s}\left(\tfrac{m}{5\,10^8 M_\odot}\right), \tag{5.22}$$

one has

$$c\tau_{\rm gr} = \frac{5}{32}(\frac{d}{10^7})^4\frac{1}{r_S^M r_S^m (r_S^M + r_S^m)}10^{28}{\rm s} \tag{5.23}$$

and finally

$$c\tau_{\rm gr} = 6\ 10^{15}{\rm s} = 6\ 10^7{\rm yr} \tag{5.24}$$

5.2 Literature review on AGN periodicities in γ-rays

Fermi-LAT's continuous and extensive monitoring of the γ-ray sky has established it as a key instrument for the study of time-domain astronomy. Its high duty cycle and long-term observations allow the identification and systematic monitoring of many transient and variable sources. Since its launch, several studies have explored different methods to account for periodic or QPOs in HE sources, mainly in the AGN LCs.

The first evidence of a periodic oscillation for γ-ray sources was provided by Sandrinelli et al. (2014) for PKS 2155-304 and by Ackermann et al. (2015) for PG 1553+113. In the first paper, they confirm an optical oscillation of 315 days and detect a significant Fermi-LAT signal at $\sim$ 630 days, twice the value for the optical LC. In the second, they report a low-significance period of 796 days using $\sim$ 6.9 years of Fermi-LAT data. Correlations in the radio and optical LCs supported this oscillation.

Another multi-wavelength QPO search was performed by Sandrinelli et al. (2016a) on year-like timescales in the LCs of six blazars, comparing the HE emission with the optical-near-IR bands. With $\sim$ 6 years of Fermi-LAT observations, they found a periodicity of 642 days on PKS 2155-304, close to previous results. Zhang et al. and collaborators analyzed the possible QPO in the blazars PKS 2155-304, PKS 0426-380 and PKS 0301-243 in three different papers (2017a; 2017b; 2017c) using $\sim$ 8 years of Fermi-LAT data. The three similar works detect periodicities of 635, 1223 and 767 days at significance levels of $\sim$ 4.9, 3.6 and 5σ, respectively.

The first systematic search for cyclic γ-ray sources from days to years was performed by Prokhorov & Moraghan (2017). In the long term, they found evidence for QPO from four blazars (listed in Table 5.1), three of which are considered potential candidates for BSMBH systems due to their high redshift, where major galaxy mergers are more likely to occur. Peñil et al. (2020) performed another systematic search for periodic Fermi-LAT AGNs using ten different techniques. Their analysis presents nine new periodic candidates with significance greater than 4σ.

All the detections mentioned above and the following periodicity results (Sandrinelli et al., 2017; Li et al., 2018; Sandrinelli et al., 2018; Tavani et al., 2018; Chevalier et al., 2019; Ait Benkhali et al., 2020; Zhang et al., 2021; Ren et al., 2023) were performed with spectral-based methods such as LSP (see Section 6.4) or Fourier transform analyses such as the Weighted Wavelet Z Transform (Foster, 1996) or the Continuous Wavelet Transforms (Torrence & Compo, 1998).

The other class of searches is performed with time-domain methods, namely Gaussian Processes (Covino et al., 2020; Zhang et al., 2021) and autoregressive moving average (ARMA) or continuous-time autoregressive moving average (CARMA) models (Kelly et al., 2014; Yang et al., 2021). In most of these works, the periodic component is included as a periodic kernel in a stochastic LC.

Table 5.1 lists the reviewed literature results of periodic γ-ray Fermi-LAT sources.

In summary, different methods have been developed to analyze the available HE data in search of periodic AGNs. Several periodic sources have been identified, with good agreement between the two main types of methods. The results of this work have improved our understanding of the physical origin of these emissions. However, the main limitation of most previous research is the lack of separate detailed identification and analysis of all the components responsible for AGN emission. In the next chapter, we

develop a method adapted to a time-domain approach, which is able to clearly separate each emission component, whether stochastic or deterministic.

Table 5.1. Literature γ-ray periodic sources. For each 4FGL source, the list indicates: Association; source type; detected period in literature and reference.

4FGL Name	Association	Type	Period [days]	Reference
J0043.8+3425	GB6 J0043+3426	fsrq	657	16
J0102.8+5825	TXS 0059+581	fsrq	767	16
J0108.6+0134	4C +01.02	bll	$268_{\pm 55}$	16
J0158.5+0133	4C +01.28	bll	445	8
J0210.7-5101	PKS 0208-512	fsrq	949	16
J0211.2+1051	CGRaBS J0211+1051	bll	621	16
J0245.9-4650	PKS 0244-470	fsrq	225	20
J0252.8-2218	PKS 0250-225	fsrq	438	16
J0303.4-2407	PKS 0301-243	bll	$767_{\pm 109}$, $752_{\pm 89}$, 730,	7, 14, 16
J0428.6-3756	PKS 0426–380	bll	$1223_{\pm 247}$, $937_{\pm 153}$, 1241	6, 14, 16
J0449.4-4350	PKS 0447-439	bll	$937_{\pm 153}$, 913	14, 16
J0457.0-2324	QSO J0457-2324	fsrq	949	16
J0501.2-0158	PKS 0458-02	fsrq	621	16
J0521.7+2112	RX J0521.7+2112	bll	1022	16
J0522.9-3628	PKS 0521-365	bll	402	18
J0538.8-4405	PKS 0537-441	bll	280, 273, $286_{\pm 73}$	4, 17, 19
J0721.9+7120	PKS 0716+71	bll	346, $344_{\pm 16}$, 1022	8, 10, 11
J0808.2-0751	QSO B0805-077	fsrq	658	8
J0811.4+0146	QSO B0808+019	bll	1570	16
J0818.2+4222	QSO B0814+42	bll	803	16
J1048.4+7143	S5 1044+71	fsrq	$116_{\pm 33}$	19
J1146.9+3958	B2 1144+40	fsrq	1205	16
J1248.3+5820	QSO B1246+586	bll	803	16
J1303.0+2434	MG2 J130304+2434	bll	730	16
J1454.4+5124	TXS 1452+516	bll	767	16
J1522.1+3144	B2 1520+31	bll	$176_{\pm 48}$	19
J1555.7+1111	PG 1553+113	bll	$796_{\pm 29}$, 798, $780_{\pm 63}$, 803 $808_{\pm 102}$, 790, 803, 792	2, 8, 11, 12 14, 15, 16, 17
J1649.4+5235	87GB 164812.2+524023	bll	986	16
J1903.2+5540	TXS 1902+556	bll	1387	16
J2056.2-4714	PKS 2052-47	fsrq	637, 620	8, 16
J2158.8-3013	PKS 2155-304	bll	630, $642_{\pm 59}$, $635_{\pm 47}$, 644, $620_{\pm 41}$ $685_{\pm 9}$, $625_{\pm 70}$, 610, 621,	1, 3, 5, 8, 11 13, 14, 15, 16
J2202.7+4216	BL Lacertae	bll	698, $680_{\pm 35}$, $680_{\pm 35}$	8, 9, 11
J2250.0-1250	PKS 2247-131	bcu	$217_{\pm 38}$	19
J2258.1-2759	VSOP J2258-2758	fsrq	475	16

References: (1) Sandrinelli et al. (2014) (2) Ackermann et al. (2015) (3) Sandrinelli et al. (2016a) (4) Sandrinelli et al. (2016b) (5) Zhang et al. (2017a) (6) Zhang et al. (2017b) (7) Zhang et al. (2017c) (8) Prokhorov & Moraghan (2017) (9) Sandrinelli et al. (2017) (10) Li et al. (2018) (11) Sandrinelli et al. (2018) (12) Tavani et al. (2018) (13) Chevalier et al. (2019) (14) Ait Benkhali et al. (2020) (15) Covino et al. (2020) (16) Peñil et al. (2020) (17) Yang et al. (2021) (18) Zhang et al. (2021) (19) Ren et al. (2023) (20) Das et al. (2023)

Chapter 6

Building a periodicity search time-series algorithm

The core of this thesis project is to build, test and apply periodicity search algorithms to time-series data, specifically to HE astrophysics LCs. To do so, we use a time-domain approach that describes the stochastic components with an AR or CAR noise model.

In our method (Rueda et al., 2022), the periodicity is included as part of the deterministic component of the emitted flux, modeled separately from the stochastic component, as explained in Section 6.1. The description and characterization of the stochastic behavior are crucial to understanding the emission's physical origin. Then, removing all the deterministic trends from the flux allows us to analyze the stochastic component assuming that these are stationary, i.e., the noise properties do not change over time. The assumption of stationarity may be broken during an AGN outburst. Section 6.1.2 is dedicated to the theory of colored-noise AR/CAR models, which are known to describe the AGNs stochastic emission better (Moreno et al., 2019).

Then, Section 6.2 describes all the steps in creating a MCMC algorithm to search for the desired parameters that conform our models. Here, the model has a distinction for regularly and irregularly sampled data, using AR or CAR noise components, respectively. Also, in Section 6.2.2, we include a review of the MCMC chains' quality.

In Section 6.3, from the different MCMC models' outcomes, the objective is to establish the model that accurately describes the data. To do so, we compare the models using Information Theory, enabling us to evaluate the dispersion of the data and the model complexity.

An alternative spectral-domain periodicity search is introduced in Section 6.4 as cross-checking. These kind of methods are based on LSP (Lomb, 1976; Scargle, 1982; VanderPlas, 2018) and its possible variants. Here, we use the public domain Agatha program by Feng et al. (2017). Compared to these spectral methods, with our algorithm, the amplitude of the periodic terms is determined, allowing the physical analysis and the study of the evolution with time, as done in Sections 7.3 and 8.3.

Finally, in Section 6.5, we use the MCMC sampling parameters to determine the PSD of the time series analyzed. In Tarnopolski et al. (2020), they evaluated existing spectral and time-domain methods to obtain PSD from an AGN sample, only considering pure stochastic models in the LCs.

6.1 Describing the AGN model

6.1.1 Overview

A model that clearly separates the deterministic and the stochastic component in any astrophysical flux description is fundamental to understanding the emission's origin. In the AGNs field, it helps to describe the observed periodicity better and, therefore, is related to the theory of BBH exposed in Chapter 5.

One possibility for a periodic pattern in the LC is a damped oscillation that originates internally in the system, such as from the disk emission. It could be fitted as a periodic kernel in the noise component. Alternatively, an external forcing, like BBH, could also produce a strictly periodic pattern detected in our algorithm as a deterministic part. It is also possible for a combination of both factors to contribute to the periodicity. That is why our method allows for a comprehensive description of both internal oscillations and external influences.

Now, before motivating the choice of the physical model employed in the description of the AGN LCs, a set of considerations have to be taken into account:

- The noise on AGN LC is known to have a log-normal amplitude distribution (see Ait Benkhali et al. (2020)). In this project, we work with the logarithm of the flux (designed below as "lflux").
- The LC baselines are not stable and exhibit drifts. In this project, baseline drifts are taken as a linear drift of lflux with time.
- The noise on LCs could be additive, multiplicative or a combination of both. Noise on a time scale of a few weeks is modelled as additive coloured noise in this work.
- The detected periods may not be stable in time. The period instability could be an artifact of a multiplicative source of noise. In Sections 7.3 and 8.3, we analyze data in different short intervals over all the observation time to consider a possible drift of the lflux period.

The model lflux LC $\phi(t)$ is the sum of a mean lflux $\bar{\phi}$ with a linear trend, a periodic term, and a stochastic component ϵ. The periodic term can be written as:

$$\phi_P(t) = \sum_j (A_j \cos(\omega_j t) + B_j \sin(\omega_j t)) \tag{6.1}$$

The A_j, B_j and ω_j are assumed to take constant values inside each analysis time interval. Then, for the full LC model, the time evolution of the flux obeys:

$$\phi(t) = \bar{\phi} + Ct + \phi_P + \epsilon(t) = \underbrace{\bar{\phi}}_{\text{mean}} + \underbrace{Ct_n}_{\text{linear}} + \sum_j \underbrace{(A_j \cos(\omega_j t_n) + B_j \sin(\omega_j t_n))}_{\text{Periodic}} + \underbrace{\epsilon(t)}_{\text{stochastic}} \tag{6.2}$$

6.1.2 Choice of the stochastic models

In this section, we use additive colored noise in the explanation of the stochastic behavior of the AGNs. Colored or correlated noise refers to noise with a frequency-dependent PSD

(see Section 6.5), i.e., the signal is not constant across all frequencies and it is usually described by a power law. In opposition, for White Noise (WN) processes the PSD intensity is equal at all frequencies.

In our model, we assume that the evolution in time of the random component $\epsilon(t)$ obeys a stochastic equation of the AR(p) or CAR(p) statistical type of order p (Brockwell & Davis, 1991). In these models, the noise part of the current value is correlated with the$p$ lagged values. The order p describes the memory length of the autocorrelation and the model's parameters describe the magnitude of the lagged values influence.

Instead of AR (or CAR) models, it would be more general to use AR(p)MA(l) models of order (p, l), where the lagged l error terms are also used to describe the present value. However, AR models can be more easily related to physical properties of the AGN system, such as correlation time-scales or eigen-frequencies (e.g, equation 6.5).

The popular Ornstein-Uhlenbeck (OU) (1930) process for AR(1)/CAR(1) models satisfies:

$$\dot{\epsilon} = -\frac{\epsilon}{\tau} + \frac{\sqrt{D}}{\tau}\zeta(t) \tag{6.3}$$

where τ is the correlation time of noise, D is a constant, ζ is WN, and the dot is the time derivation. The noise term can also be generalized for models involving higher order derivatives such as the CARMA models, described below. With this, the time evolution of the flux satisfies:

$$\phi(t) = \bar{\phi} + Ct + \textstyle\sum_j (A_j \cos(\omega_j t) + B_j \sin(\omega_j t)) + \epsilon(t) \tag{6.4}$$

$$\dot{\epsilon} = -\tfrac{\epsilon}{\tau} + \tfrac{\sqrt{D}}{\tau}\zeta(t) \tag{6.5}$$

Equation 6.5 can be formally integrated to

$$\epsilon(t) = \exp\{((s-t)/\tau)\}\epsilon(s) + w(t-s),\ t \geq s \tag{6.6}$$

where s is the previous time value and $w(t-s)$ is a Gaussian Random Variable (GRV):

$$w(t-s) \sim N(0, D \times (1 - \exp(-2(t-s)/\tau)) \tag{6.7}$$

Below, we describe the generalization of Equations 6.4 and 6.6 for regularly and irregularly sampled data.

AR model

In an AR model, the current value is expressed as a linear function of the previous values of the same variable, along with an error term.

AR(1)

If the data is uniformly sampled in time, the correlation time τ can be transformed to a parameter in discrete time, then equation 6.6 is integrated into an AR(1) model:

$$\epsilon(t_n) = \beta_1 \epsilon(t_{n-1}) + w(t_n - t_{n-1}) \tag{6.8}$$

where β_1 is the AR(1) coefficient of the first lagged value, and the WN $w(t_n - t_{n-1})$ is described as a Gaussian with zero mean and constant variance

AR(p)

Equation 6.8 can be generalized to account for the correlation to the p order as an AR(p) model:

$$\epsilon(t_n) = \sum_{j=1}^{k} \beta_j \epsilon(t_n - j) + w(\delta t) \tag{6.9}$$

with $(\delta t = t_n - t_j)$.

Considering equation 6.9, the evolution in time of ϕ and ϵ is now described by:

$$\phi(t_n) = \bar{\phi} + \textstyle\sum_j (A_j \cos(\omega_j t_n) + B_j \sin(\omega_j t_n)) + C t_n + \epsilon(t_n) \tag{6.10}$$

$$\epsilon(t_n) = \textstyle\sum_{j=1}^{k} \beta_j \epsilon(t_{n-j}) + w(\delta t) \tag{6.11}$$

Now, in order to eliminate the correlation between flux measurements at consecutive times, we can remove ϵ from equation 6.10 using an auxiliary variable z defined by the linear combination:

$$z = \phi(t_n) - \sum_{j=1}^{k} \beta_j \phi(t_{n-j}). \tag{6.12}$$

The evolution of z is described by an equation similar to equation 6.9, except for the ϵ term which is now replaced by a GRV.

$$z = \bar{\phi}' + \sum_j (A'_j \cos(\omega_j t_n) + B'_j \sin(\omega_j t_n)) + C' t_n + w(\delta t) \tag{6.13}$$

The physical $\bar{\phi}, A_j, B_j, C$ variables are obtained from the $\bar{\phi}', A'_j, B'_j, C'$ variables by the transformation given in appendix A.

The model of equation 6.13 is closely related to ARMA models with "exogenous covariates" (Feigelson et al., 2018), while its extension to irregular spacing would be a variant of space state models (Durbin & Koopman, 2001). It is, however, not of common use in HE astrophysics. As emphasized in the introduction, it has the advantage of clearly separating the periodic and the stochastic part of the signal. Compared to only ARMA/CARMA-based approaches such as those from Goyal et al. (2018) and Yang et al. (2021), it has the potential of identifying multiple periods such as harmonics, giving crucial clues on the underlying physical mechanism of the flux oscillation.

CAR model

In a CAR model, the time series data exhibits dependence among its values over time, in a continuous-time domain. The qualitative description is similar to an AR model, but with the influence of the past values decaying over time depending on the τ_p parameters.

CAR(1)

In the CAR(1) model, ϵ satisfies the OU model in Equation 6.3, integrated to Equation 6.6.

Equations 6.4 and 6.6 are sufficient to describe the co-evolution of $\phi(t)$ and $\epsilon(t)$. However, in the CAR(1) model, it is possible to completely eliminate the ϵ variable. The evolution of $\phi(t)$ is then described by equation

$$\begin{aligned}\phi(t) - A(t-s)\phi(s) = \bar{\phi}(1 - A(t-s)) + C(t - A(t-s)s) + w(t-s) \\ + \sum_j (A_j(\cos(\omega_j t) - A(t-s)\cos(\omega_j s)) + B_j(\sin(\omega_j t) - A(t-s)(\sin(\omega_j s))\end{aligned} \tag{6.14}$$

with $A(t) = \exp\{\frac{-t}{\tau}\}$.

Before generalizing to a CAR(p) model, we provide the solutions for a CAR(2) model.

CAR(2)

In the CAR(2) noise model, the variable ϵ obeys:

$$\ddot{\epsilon} - (\xi_1 + \xi_2)\dot{\epsilon} + \xi_1\xi_2\epsilon(t) = \frac{\sqrt{D}}{\tau^{3/2}}\zeta(t) \tag{6.15}$$

Equation 6.15 is equivalent to a matrix equation, namely:

$$\frac{d}{dt}\begin{pmatrix}\epsilon \\ \dot{\epsilon}\end{pmatrix} + \begin{pmatrix}0 & -1 \\ \xi_1\xi_2 & -(\xi_1+\xi_2)\end{pmatrix}\begin{pmatrix}\epsilon \\ \dot{\epsilon}\end{pmatrix} = \frac{\sqrt{D}}{\tau^{3/2}}\begin{pmatrix}0 \\ \zeta\end{pmatrix} \tag{6.16}$$

Taking the initial values of ϵ and $\dot{\epsilon}$ at time t_0, the solution of equation 6.16 is:

$$\epsilon(t) = k_1\exp\{(\xi_1(t-t_0))\} + k_2\exp\{(\xi_2(t-t_0))\} + N_1(t-t_0,\xi_1,\xi_2) \tag{6.17}$$

$$\dot{\epsilon}(t) = k_1\xi_1\exp\{(\xi_1(t-t_0))\} + k_2\xi_2\exp\{(\xi_2(t-t_0))\} + N_2(t-t_0,\xi_1,\xi_2) \tag{6.18}$$

where N_1, N_2 are components of a Gaussian vector, the explicit expressions of which is given in the Appendix B. The values of k_1 and k_2 are given by:

$$k_1 = \frac{\xi_2\epsilon(t_0)}{\xi_2-\xi_1} - \frac{\dot{\epsilon}(t_0)}{\xi_2-\xi_1} \tag{6.19}$$

$$k_2 = -\frac{\xi_1\epsilon(t_0)}{\xi_2-\xi_1} + \frac{\dot{\epsilon}(t_0)}{\xi_2-\xi_1}. \tag{6.20}$$

ϵ and $\dot{\epsilon}(t)$ play the role of hidden variables. Note that ξ_1 and ξ_2 are equivalent to the τ parameter in the CAR(1) model, and they could behave as complex damped oscillations (Appendix B). Their time evolution between t_0 and t is entirely determined by equations 6.17 and 6.18. Once the evolution of $\epsilon(t)$ and $\dot{\epsilon}(t)$ is known, the time evolution of $\phi(t)$ is obtained from equation 6.10.

CAR(p)

For a generalized CAR(p) model from Equation 6.15. , the evolution of ϵ is described by equation:

$$\epsilon^{(p)} - (\sum\xi_k)\epsilon^{(p-1)} + (\sum_{j<k}(\xi_j\xi_k))\epsilon^{(p-2)} + .. + (-1)^p(\Pi\xi_k)\epsilon(t) = \frac{\sqrt{D}}{\tau^{(p+1/2)}}\zeta(t) \tag{6.21}$$

Introducing the p-component vectors:

$$\mathbf{e} = \begin{pmatrix} \epsilon(t) \\ \dot{\epsilon}(t) \\ \vdots \\ \epsilon^{(p-1)(t)} \end{pmatrix}, \quad \mathbf{r} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \frac{\sqrt{D}}{\tau^{(p+1/2)}}\zeta(t) \end{pmatrix} \tag{6.22}$$

and the matrix:

$$M = \begin{pmatrix} 0 & -1 & 0 & \dots \\ 0 & 0 & -1 & \dots \\ \dots & \dots & \dots & \dots \\ -(\sum \xi_k) & (\sum_{j<k}(\xi_j \xi_k)) & \dots & (-1)^p(\Pi \xi_k) \end{pmatrix} \tag{6.23}$$

Equation 6.21 can be written as:

$$\frac{d\mathbf{e}}{dt} + M\mathbf{e}(t) = \mathbf{r}(t) \tag{6.24}$$

which is the vector form of the OU model.

The solution of equation 6.24 is:

$$\mathbf{e}(t) = V \begin{pmatrix} k_1 \exp\{\xi_1(t-t_0)\} \\ k_2 \exp\{\xi_2(t-t_0)\} \\ \dots \\ k_p \exp\{\xi_p(t-t_0)\} \end{pmatrix} + \mathbf{N} \tag{6.25}$$

where V is the Vandermonde matrix:

$$V = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \xi_1 & \xi_2 & \dots & \xi_p \\ \dots & \dots & \dots & \dots \\ \xi_1^{p-1} & \xi_2^{p-1} & \dots & \xi_p^{p-1} \end{pmatrix} \tag{6.26}$$

and $\mathbf{N}$ is a Gaussian vector whose components are given in Appendix B.

The initial conditions at $t = t_0$ fix the values of the constants k_i to

$$\begin{pmatrix} k_1 \\ k_2 \\ \dots \\ k_p \end{pmatrix} = V^{-1}\mathbf{e}(t_0) \tag{6.27}$$

For the low values of p, the covariance matrix of the Gaussian vector can be explicitly computed using equation B.17. However, modeling noise by a CAR(p) with $p > 2$ is not easy to interpret physically.

6.2 Integrating the models in an MCMC algorithm to search for the parameters

When trying to analyze and describe complex astrophysical data as AGN LCs, multiple physical processes and data-taking systematics can be responsible for such observed emission. That is why we need to test different parametric models in order to find the best description of the observed data.

The search for the parameters θ that build the different possible hierarchical models is done by Bayesian inference using MCMC (see the box below) samplers. MCMC sampling is a powerful statistical technique used to explore complex parameter spaces and estimate the probability distribution of parameters in a model. It works by constructing a Markov chain (Frigessi & Heidergott, 2011) that explores the parameter space, generating a sequence of representative parameter samples.

The models have been implemented in the R version of JAGS (Plummer, 2012). JAGS is a MCMC based on the Gibbs sampling algorithm. The output of the program is a set of posterior probability distributions, one for each parameter included in the model computed. These resulting samples facilitate the estimation of various statistical properties of the parameters, such as the mean, variance and credible intervals.

MCMC

The MCMC sampling procedure starts with a prior probability distribution for each parameter. By Bayes theorem, the posterior probability distribution of the parameters is the product of the likelihood (Eq. 6.37) and the prior distributions. Iteratively, the algorithm generates a new set of parameter values by proposing a move based on the current parameter probability. The acceptance or rejection of the proposed move is determined using the Gibbs sampling algorithm (Casella & George, 1992), which considers the likelihood of the proposed parameter values given the available data. If the move is accepted, the chain transitions to the new parameter values; otherwise, it remains at the current values. This iterative process persists for a large number of iterations, enabling the chain to explore the parameter space systematically and gradually converge toward the desired target distribution.

As the chain advances, the generated samples become increasingly representative of the target distribution, and the correlation between consecutive samples diminishes. After sufficient iterations, the chain reaches a stationary state where the distribution of samples converges to the final posterior distribution.

6.2.1 MCMC models and parameters

The different stochastic models are composed of a mean value $\bar{\phi}$, an AR/CAR term, and a WN component $N(0, \sigma)$ for AR, or GRV $w(t-s)$ for CAR. For each stochastic model, the deterministic components are added to perform six different MCMC fits: pure noise, linear, sinusoidal, harmonic, linear + sinusoidal, linear + harmonic. This gives 18 models computed for AR models and the 12 models for CAR, as we are only working with order $p_{\text{AR}} = 2$, and $p_{\text{CAR}} = 1$.

The following list indicates the parameters and the mathematical description of each stochastic model and deterministic component to fit.

Stochastic model AR:

- White Noise $[\bar{\phi}, \sigma]$: $\phi(t_n) = \bar{\phi} + N(0, \sigma)$ (6.28)

- AR(1) $[\bar{\phi},\ \sigma,\ \beta_1]$: $\phi(t_n) = \bar{\phi} + \beta_1\phi(t_{n-1}) + N(0,\sigma)$ (6.29)
- AR(2) $[\bar{\phi},\ \sigma,\ \beta_1,\ \beta_2]$: $\phi(t_n) = \bar{\phi} + \beta_1\phi(t_{n-1}) + \beta_2\phi(t_{n-2}) + N(0,\sigma)$ (6.30)

Stochastic model CAR:

- GRV $[\bar{\phi},\ \tau,\ D]$: $\bar{\phi}(1-\exp\{-(t-s)/\tau\}) + N(0, D(1-\exp\{-2(t-s)/\tau\}))$ (6.31)
- CAR(1) $[\bar{\phi},\ \tau,\ D]$: $\exp\{-(t-s)/\tau\}\phi(s) + \text{GRV}$ (6.32)

Deterministic component:

- Linear $[C]$: Ct_n (6.33)
- Sinusoidal $[A_j,\ B_j,\ \omega_j]$: $\sum_j (A_j\cos(\omega_j t_n) + B_j\sin(\omega_j t_n))$ (6.34)
- Harmonic $[A_j,\ B_j,\ A'_j,\ B'_j,\ \omega_j]$: $\sum_j (A_j\cos(\omega_j t_n) + B_j\sin(\omega_j t_n)$ $+A'_j\cos(2\omega_j t_n) + B'_j\sin(2\omega_j t_n))$ (6.35)

The number of models' parameters are between 2, for a pure WN model, and 9, for a CAR(2) model with a linear term, a sinusoidal term and its harmonic. Taking the example of an AR(2) model with a single period and no linear term, the conditional probability of obtaining a lflux $\phi(t_n)$ is

$$\begin{aligned} &P(\phi(t_n)|\phi(t_{n-1}),\phi(t_{n-2}),\bar{\phi},\beta_1,\beta_1,C,\omega_i,A_i,B_i,\sigma..) = \\ &N(\bar{\phi} + \beta_1\phi(t_{n-1}) + \beta_2\phi(t_{n-2}) + A\cos(\omega t_n) + B\sin(\omega t_n),\sigma_1) \end{aligned} \tag{6.36}$$

The likelihood for the parameters $\bar{\phi},\beta_1,\beta_1,C,\omega_i,A_i,B_i,\sigma$ is (Robert (2007), section 4.5)

$$\begin{aligned} &L(\bar{\phi},\beta_1,\beta_1,C,\omega_i,A_i,B_i,\sigma..) = \\ &\tfrac{1}{\sigma^{\mathrm{N}}}\prod_0^{\mathrm{N}}\exp\left\{-\tfrac{(\phi(t_n)-(\bar{\phi}+\beta_1\phi(t_{n-1})+\beta_2\phi(t_{n-2})+A\cos(\omega t_n)+B\sin(\omega t_n)))^2}{2\sigma_1^2}\right\} \end{aligned} \tag{6.37}$$

where N is the number of measurements and $\sigma_1^2 = \sigma^2 + \text{err}(t_n)^2$ takes into account the lflux measurement error at time t_n.

The prior distribution of the MCMC parameters are selected to be as vague and non-informative as possible for the data sample analyzed. This allows to minimize the influence and bias on the parameters posterior inference. The MCMC configuration for each data set analyzed is specified at each Chapter (7, 8, 9).

In order to avoid working with numbers in different orders of magnitude, which may result in lower efficiency for the MCMC sampling, data are standardized by being re-scaled to their mean ($\bar{x}$) and std (s_x).:

$$\phi(t)_{st} = \frac{\phi(t) - \bar{\phi}(t)}{s_{\phi(t)}} \tag{6.38}$$

$$t_{st} = \frac{t - \bar{t}}{s_t} \tag{6.39}$$

After the completion of the MCMC, the parameter outputs are transformed back to the original scale.

6.2.2 Quality of the generated Markov Chains

Different MCMC diagnostics (Kruschke, 2015) are used to to check whether the quality of a sample generated with an MCMC algorithm is sufficient to provide an accurate approximation of the target distribution.

Representativeness is a crucial quality feature of the MCMC. It ensures that the samples generated by the MCMC algorithm accurately represent the target distribution. To assess representativeness, we check for chains' convergence with the Gelman & Rubin diagnostic (GR, see the box below), and we examine the autocorrelation of the samples. These diagnostics help to determine if the MCMC chain has explored the parameter space sufficiently and reached a stationary distribution without a significant influence from the arbitrary initial value.

GR diagnostic

The GR diagnostic, also known as the R-hat statistic, is a specific tool used to assess convergence and mixing in MCMC chains. It compares the variation between multiple chains, initialized with different starting values, to the variation within each individual chain. If the chains have mixed well and reached convergence, the between-chain variation should be similar to the within-chain variation. A value close to 1 indicates convergence, while values larger than 1 suggest that the chains have not yet mixed well and further iterations may be necessary.

Accuracy is another essential quality feature of MCMC. It refers to the ability of the algorithm to provide accurate estimates of the target distribution's parameters. Accurate estimation requires the MCMC chain to converge to the actual values of the parameters, minimizing bias and error. To assess accuracy, it is checked that the parameter estimates are relatively similar for independent MCMC runs using different seed states.

Efficiency is a quality feature that evaluates the computational performance of the MCMC algorithm. It measures how quickly the algorithm can generate a sufficient number of independent samples from the target distribution. Efficient MCMC methods minimize the number of iterations needed to achieve convergence and provide reliable parameter estimates.

6.3 Model Selection using Information Criteria

6.3.1 Akaike Information Criterion

Information criteria are introduced in model selection problems as a form of quantitative explanation of the best model's goodness of fit compared to the others or to compare every possible model simultaneously.

Shannon (1948) Entropy is a fundamental concept in information theory. It is a measure of the uncertainty in a probability distribution and reflects the complexity of its description. Given a random variable $X \rightarrow x_1, ..., x_n$ from a known probability distribu-

tion $P(x_i)$, for each possible outcome i, the entropy is:

$$H(X) = -\sum_{i}^{n} P(x_i) \log P(x_i) \tag{6.40}$$

$H(X)$ can be defined as the information content of the outcome i. So, if the probability $P(x_i)$ maximizes the entropy, we will need more information to represent a more complex model.

From Shannon Entropy, we can derive the concept of Kullback-Leiber (KL) divergence, operating as a relative entropy. KL is the statistical measure of model approximation between a probability distribution $P(x_i)$ representing the true description of data, and the approximating model $Q(x_i)$:

$$D_{KL}(P \parallel Q) = \sum_{i}^{n} P(x_i) \log \frac{P(x_i)}{Q(x_i)} \tag{6.41}$$

$D_{KL}(P \parallel Q)$ can also be defined as the information lost when $Q(x_i)$ is used to describe $P(x_i)$. KL is not a distance as it is asymetric $D_{KL}(P \parallel Q) \neq D_{KL}(Q \parallel P)$

In order to ilustrate these concepts, the following example is given:

Binary Entropy Function as a Coin Flip

Here, we want to determine the coin flip entropy, knowing the outcome probability. Now, being the possible outcome heads (1) or tails (0), $X \to \{0, 1\}$ and the probability $P(X = 1) = p$, $P(X = 0) = 1 - p$, the entropy is:

$$H(X) = H_b(p) = -p \log_2 p - (1 - p) \log_2 (1 - p) \tag{6.42}$$

In the case of a fair coin, the probability of each outcome is equal ($p = 0.5$) and the entropy is maximum ($H_b(p = 0.5) = 1$). This means that the uncertainty on next outcome result is maximum. If we have a biased coin, one side is more likely to come up ($p \neq 1-p$) and the outcome uncertainty or surprise is reduced. For $p = 0.8$, the entropy is lower ($H_b(p = 0.8) = 0.7219 < 1$). Finally, for a one-sided coin, only one outcome is possible ($p = 1$) and there is no entropy or uncertainty in the proccess ($H_b(p = 1) = 0$). These examples are represented in the left curve on Figure 6.1.

Now, we want to compare two coin flips configurations by determining the KL divergence between them. Here, $P(X)$ is a fair coin, then $P(X = 0) = P(X = 1) = 0.5$ and $Q(X)$ is any other coin configuration $Q(X = 0) = 1 - Q(X = 1) = q$. Then, the divergence:

$$D_{KL}(P \parallel Q) = \log 0.5 - 0.5 \log(q(1 - q)) \tag{6.43}$$

When the coin $Q(X)$ is closer to a fair coin P(X), the distance between the two models is reduded, while the divergence is maximum when $Q(X)$ is a completely biased coin. In the right pannel on Figure 6.1, the KL divergence for any possible configuration of the second coin q is represented.

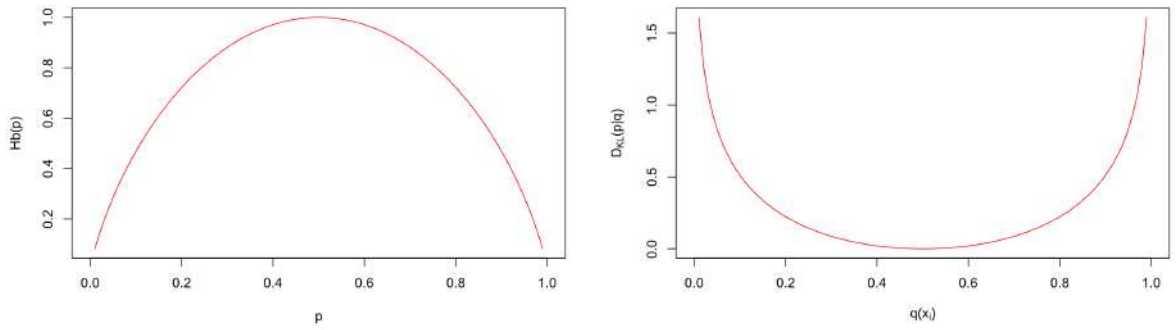

Figure 6.1. Example of a coin fllip [*Left Panel*] Entropy [*Right Panel*] KL divergence.

The Akaike Information Criterion (AIC) is an effective, general tool for selecting a parsimonious model which describes and infers empirical data based on KL distance and Fisher's Maximized log-Likelihood Estimator (MLE).

The MLE estimates, for a set of random variables $x_1, ..., x_n$, the most likely parameters Θ to describe the data, from a sample space of the parameters Ω through a likelihood function $L(\Theta)$:

$$L(\Theta) = \prod_{j=1}^{n} Pr(X_j \in A_j \mid \Theta) \tag{6.44}$$

meaning that we search the value of Θ so that L is maximized. The likelihood statistic $L(\Theta)$ represents the likelihood or probability of obtaining the observed data, given a particular model input. As an example:

Homoscedastic normal distribution

For a normally distributed sample of random variables, the probability density distribution is:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\right\} \tag{6.45}$$

with mean μ and standard deviationσ. In the homoscedasticity assumption, i.e., the variance iis constant for all the random variables, we search the value of the mean μ that maximizes the likelihood function:

$$L(\mu_j) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{\left(-\frac{(x_{ij}-\mu_i)^2}{2\sigma^2}\right)\right\} \tag{6.46}$$

Now, to better handle Equation 6.46, we can apply the natural logarithm function on both sides to convert the multiplication into a summation:

$$\log(L(\mu_j)) = \log\left[\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left\{\left(-\sum_{i=1}^{n}\frac{(x_{ij}-\mu_i)^2}{2\sigma^2}\right)\right\}\right] \tag{6.47}$$

$$\log(L(\mu_j)) = n\log\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_{ij}-\mu_i)^2 \tag{6.48}$$

Therefore, maximizing L means to minimize:

$$\sum_{i=1}^{n}(x_{ij} - \mu_i)^2 \tag{6.49}$$

Which for the normality scenario is the same as the least square method.

Akaike (1974) proposed a way to estimate divergence based on the MLE, the AIC. AIC is used as a measure of the information lost when the fitted model is used to approximate the process that generates the empirical data:

$$AIC = -2\log\mathcal{L}(\theta|y) + 2K \tag{6.50}$$

where $\log\mathcal{L}(\theta|y)$ is the log-likelihood of the model given the data y, K is the number of model's parameters (defined as θ) and operates as a penalty to model complexity. Thus, AIC is an effective tool for selecting a simple model that describes and infers empirical data, avoiding overfitting and underfitting. In our MCMC pipeline, the assessment is done using the following:

$$AIC = D(\theta|y)_{min} + 2K \tag{6.51}$$

where $D(\theta|y)_{min} = -2\log\mathcal{L}(\theta|y)$ is the minimum deviance of the MCMC posterior sample.

For each LC, within all the possible MCMC implementations, the one with the minimum AIC is selected. Now, for model statistical assessment, AIC does not carry much information as it is on a relative scale and depends on sample size. However, what matters in model assessment is ΔAIC, the difference between AIC values over multiple nested models. Given a full model F and a reduce model R:

$$\Delta AIC = AIC_R - AIC_F = -2\log(\mathcal{L}(\theta_0)/\mathcal{L}(\theta)) - 2k = \Lambda - 2k \tag{6.52}$$

where Λ is the likelihood ratio test statistic, $\mathcal{L}(\theta_0)$ and $\mathcal{L}(\theta)$ are the MLE under the null (R) and alternative (F) hypothesis, respectively, and k is the difference of parameters between models.

From this definition, a p_{value} can be computed, which shows the probability of obtaining the value Λ under the null hypothesis conditions. As stated in Efron & Hastie (2016), chapter 13 (see also Murtaugh (2014)), the relationship between p_{value} and ΔAIC can be drawn as:

$$p_{value} = Pr(\chi^2_k > \Lambda) = Pr(\chi^2_k > \Delta AIC + 2k) \tag{6.53}$$

where $\Lambda = \Delta AIC + 2k$ follows a χ^2 distribution with k degrees of freedom. In the following, we use this relation to assess a p_{value} value when comparing nested models.

6.3.2 Best model's goodness of fit

After fitting data with all the available models and selecting the best one through the AIC information, it is necessary to evaluate the performance of the best regression. It is possible to fall into model misspecification when, within all the statistical models available, they all fail to represent the distribution that generated the observed data.

With our fit, we intend to describe the trends, periodicities and correlated noise in the time series. By subtracting all these estimated components, we expect that the

fit's residuals show no apparent deviation from stationarity. This is, the residuals are independent and identically distributed (i.i.d.) random variable following a Gaussian distribution.

To do so, as suggested by Feigelson et al. (2018), it is helpful to check for residual normality by performing an Anderson & Darling (1954) test (AD test, see the box below). The AD test is a statistical test used to assess whether a given dataset follows a specific distribution. If applied for a normal distribution it is able to detect departures from normality, rejecting the hypothesis of normality when the test p-value is below a given significance level, which is usually set to 0.05.

AD test

The AD test measures the discrepancy between the empirical distribution of the data and the expected distribution by comparing the observed and the theoretical cumulative distribution function. The test statistic is based on the squared differences between the observed and expected values, with higher values indicating a larger deviation from the expected distribution. The critical values of the AD test can be used to determine the significance of the discrepancy.

6.4 Using spectral method as cross-check

In the preceding sections, we have described a time-domain MCMC method to identify periodicities in LCs. To verify our detections, we compare the periods identified through our method with an alternative approach based on spectral analysis. The public domain Agatha program (Feng et al., 2017) utilizes the information retrieved from correlated noise models to produce different LSP (see the box below) for astronomical time-series data and to assess the statistical significance of the identified periodic components.

LSP

The LSP is a widely used technique in time series analysis to identify periodicities in unevenly sampled data. It provides a way to estimate the PSD of a signal, allowing the detection of periodicities even in the presence of irregular or sparse data. The periodogram computes the signal's power at different frequencies, highlighting significant peaks that correspond to potential periods:

$$P(\nu) = \frac{1}{N_v \cdot \sigma} \left\{ \frac{\left(\sum_{i=1}^{N_v} (y_i - \bar{y}) \cdot \cos\left[2\pi\nu(t_i - \varphi)\right]\right)^2}{\sum_{i=1}^{N_v} \cos^2\left[2\pi\nu(t_i - \varphi)\right]} + \frac{\left(\sum_{i=1}^{N_v} (y_i - \bar{y}) \cdot \sin\left[2\pi\nu(t_i - \varphi)\right]\right)^2}{\sum_{i=1}^{N_v} \sin^2\left[2\pi\nu(t_i - \varphi)\right]} \right\}$$

where N_v represents the number of data points, σ is the variance of the data, $\bar{y}$ is the mean value of the data, φ is the phase offset, and the summations are performed over all data points. The numerator of the formula represents the power contribution from the periodic signal at frequency ν, while the denominators account for the total power and provide a normalization factor.

First, using a noise model selection tool (Feng et al., 2016) Agatha retrieves the best order k of a moving average (MA) stochastic model, including $k = 0$ corresponding to

WN.Then, it simultaneously fits the retrieved noise model with the deterministic components, accounting for linear trends and periodic sinusoids. The program includes four variations of the LSP that were tested with simulated LCs. Here, we use the Bayes factor periodogram (BFP), which utilizes a maximum likelihood estimation for all the components. The significance of the computed periodograms is determined by evaluating the logarithm of the Bayes Factor (lnBF), which represents the maximum likelihood ratio between the periodic and noise models for every frequency, thus estimating the significance of the identified period.

Agatha also provides a *moving periodogram*, i.e., a periodogram calculated in different time windows, which is a helpful tool for finding long-term changes in periods, as studied in Sections 7.3 and 8.3. The authors of Agatha recommend using their program in combination with an MCMC search to refine the results. As explained in the previous section, the priors of our MCMC search are not based on Agatha results, which are only used as cross-checks.

6.5 Determining the PSD from the stochastic MCMC parameters

In the context of AGNs HE astrophysics, PSD analysis is a powerful tool for studying the temporal variability of astrophysical sources. PSD characterizes power distribution across different frequencies in a time series, referred to as frequency-domain or spectral analysis. PSD provides insights into the underlying physical processes responsible for the observed fluctuations, and the presence of specific features in the analysis can provide clues about the jet emission mechanisms.

This section explains the use of the estimated MCMC parameters for the stochastic models to obtain the PSD of the LCs analyzed. Hereafter, we present the PSD computation for AR and CAR models, including the relative errors.

AR model

The spectral density for an AR(p) model is given by:

$$\rho_\nu = \frac{\sigma^2}{2\pi} \frac{1}{\left(1 + \sum_{k=1}^{p} \beta_k \exp\{(-ik\omega\delta t)\}\right)\left(1 + \sum_{k=1}^{p} \beta_k \exp\{(+ik\omega\delta t)\}\right)} \tag{6.54}$$

AR(1)

For AR(1), equation 6.54 gives

$$\rho_\nu = \frac{\sigma^2}{2\pi} \frac{1}{\left(1 + \beta_1 \exp\{(-i\omega\delta t)\}\right)\left(1 + \beta_1 \exp\{(+i\omega\delta t)\}\right)} = \frac{\sigma^2}{2\pi} \frac{1}{\left(1 + {\beta_1}^2 + 2\beta_1 \cos\omega\delta t\right)} \tag{6.55}$$

as in the Example 4.1.4 of Brockwell & Davis (2016). The PSD error is:

$$\frac{\delta\rho_\nu}{\rho_\nu} = 2\delta\beta_1 \frac{(\cos\omega\delta t + \beta_1)}{\left(1 + {\beta_1}^2 + 2\beta_1 \cos\omega\delta t\right)} \tag{6.56}$$

AR(2)

For AR(2), equation 6.54 gives

$$\rho_\nu = \frac{\sigma^2}{2\pi} \frac{1}{(1+\beta_1 \exp\{(-i\omega\delta t)\}+\beta_2 \exp\{(-2i\omega\delta t)\})(1+\beta_1 \exp\{(+i\omega\delta t)\}+\beta_2 \exp\{(+2i\omega\delta t)\})} \tag{6.57}$$

$$= \frac{\sigma^2}{2\pi} \frac{1}{\left(1+{\beta_1}^2+{\beta_2}^2-2\beta_2+\cos(\omega\delta t)(\beta_1\beta_2+\beta_1)+4\beta_2\cos^2(\omega\delta t)\right)} \tag{6.58}$$

as in the Example 4.4.1 of Brockwell & Davis (2016). The error on ρ_ν is given by:

$$\frac{\delta\rho_\nu}{\rho_\nu} = 2\frac{(\delta\beta_1(\beta_1+(\beta_2+1)\cos(\omega\delta t)+\delta\beta_2(\beta_2-1+\beta_1\cos(\omega\delta t)+2\cos^2(\omega\delta t)))}{\left(1+{\beta_1}^2+{\beta_2}^2-2\beta_2+\cos(\omega\delta t)(\beta_1\beta_2+\beta_1)+4\beta_2\cos^2(\omega\delta t)\right)} \tag{6.59}$$

CAR Model

For a CAR model, the PSD is given by:

$$\rho_\nu = \frac{D^2}{2\pi} \frac{1}{(1+\sum_{k=1}^{p}\tau_k(-i\omega)^k)\,(1+\sum_{k=1}^{p}\tau_k(i\omega)^k)} \tag{6.60}$$

as in the CAR limit for the Equation 2.13 of Brockwell & Marquardt (2005).

CAR(1)

The results mimic exactly those of AR(1):

$$\rho_\nu = \frac{D^2}{2\pi} \frac{1}{(1+{\tau_1}^2\omega^2)} \tag{6.61}$$

which is a Lorentzian profile. The relative error is

$$\frac{\delta\rho_\nu}{\rho_\nu} = 2\delta\tau_1 \frac{(\tau_1\omega^2)}{(1+{\tau_1}^2\omega^2)} \tag{6.62}$$

CAR(2)

The spectral density for CAR(2) is easily found to be:

$$\rho_\nu = \frac{D^2}{2\pi} \frac{1}{((1-\tau_2\omega^2)^2+{\tau_1}^2\omega^2)} \tag{6.63}$$

The relative error is

$$\frac{\delta\rho_\nu}{\rho_\nu} = 2\omega^2 \frac{(-\delta\tau_2(1-\tau_2\omega^2)+\tau_1\delta\tau_1)}{((1-\tau_2\omega^2)^2+{\tau_1}^2\omega^2)} \tag{6.64}$$

Chapter 7

Applying the algorithm to Fermi-LAT regularly sampled data

This chapter is related to Sections 3 and 4 of the publication resulting from this work (Rueda et al., 2022). In this project, we analyze regularly sampled data as we only apply AR models for the stochastic component. In Section 7.1, we analyze a sample of several Fermi-LAT periodic AGN candidates and explain how the LCs are generated for these candidates. The results obtained from the MCMC search are discussed in Section 7.2, while the temporal stability of these results is examined in Section 7.3.

7.1 Source selection and Data Analysis

Here, we focus on a subset of the AGN population within the 4FGL. The catalog covers a complete survey of the sky in the 50 MeV - 1 TeV energy range and includes a total of 5064 sources, of which 3207 are identified as AGNs. Among the AGNs, 3137 are classified as blazars, 42 as radio galaxies, and 28 as other types of AGNs (see Section 1.3.2 for a detailed description). Previous research on period detection in γ-ray astronomy, reviewed in Section 5.2, influenced our selection of AGNs for this study. This work was performed in 2021, so only publications until this date are considered. Information on the properties of the selected sources and the results of previous studies on their periodicity can be found in Table 7.1.

The analysis of each source was performed using Enrico, a community-developed Python package to conduct Fermi-LAT analysis (Sanchez & Deil, 2013), which consists in a simplified complete analysis chain based on the FermiTools[1]. The version of the Fermitools is 2.0.8 via the conda repository. The data are obtained from Fermi-LAT Data Server[2] introducing Astroquery (Ginsburg et al., 2019) in our pipeline. The spacecraft file is downloaded for the whole mission time range using Enrico. More than 12 years of Pass 8 LAT data (Atwood et al., 2013) are used, using 145 time bins from the start of the mission (54700 Modified Julian Day (MJD)) until the end of March 2021 (59303 MJD) to build the LCs. We do so to form LCs with a binning of ~ 31 days. A monthly binning allows us to account for the long-term correlated noise in order to look for yearly periods. The spectral model is set to be a PowerLaw2, defined as a simple power law with the

[1] https://github.com/fermi-lat/Fermitools-conda

[2] https://fermi.gsfc.nasa.gov/cgi-bin/ssc/LAT/LATDataQuery.cgi

Table 7.1. List of Fermi-LAT AGN sample with 4FGL Name and Association, source type, detected period in literature and reference.

4FGL Name	Association	Type	Period [days]	Reference
J0043.8+3425	GB6 J0043+3426	fsrq	657	3
J0102.8+5825	TXS 0059+581	fsrq	767	3
J0158.5+0133	4C +01.28	bll	445	4
J0210.7-5101	PKS 0208-512	fsrq	949	3
J0211.2+1051	CGRaBS J0211+1051	bll	621	3
J0252.8-2218	PKS 0250-225	fsrq	438	3
J0303.4-2407	PKS 0301-243	bll	730, $766_{\pm 109}$	3, 9
J0428.6-3756	QSO B0426-380	bll	1241, $1223_{\pm 248}$	3, 8
J0449.4-4350	PKS 0447-439	bll	913	3
J0457.0-2324	QSO J0457-2324	fsrq	949	3
J0501.2-0158	PKS 0458-02	fsrq	621	3
J0521.7+2112	RX J0521.7+2112	bll	1022	3
J0721.9+7120	PKS 0716+71	bll	1022, 346	3, 4
J0808.2-0751	QSO B0805-077	fsrq	658	4
J0811.4+0146	QSO B0808+019	bll	1570	3
J0818.2+4222	QSO B0814+42	bll	803	3
J1146.9+3958	B2 1144+40	fsrq	1205	3
J1248.3+5820	QSO B1246+586	bll	803	3
J1303.0+2434	MG2 J130304+2434	bll	730	3
J1454.4+5124	TXS 1452+516	bll	767	3
J1555.7+1111	PG 1553+113	bll	790, 803, 798, $780_{\pm 63}$, 803	2, 3, 4, 5, 6
J1649.4+5235	87GB 164812.2+524023	bll	986	3
J1903.2+5540	TXS 1902+556	bll	1387	3
J2056.2-4714	PMN J2056-4714	fsrq	620, 637	3, 4
J2158.8-3013	PKS 2155-304	bll	685, 610, 621, 644, $620_{\pm 41}$, $635_{\pm 47}$	1, 2, 3, 4, 5, 7
J2202.7+4216	BL Lacertae	bll	698, $680_{\pm 35}$	4, 5
J2258.1-2759	VSOP J2258-2758	fsrq	475	3

References:(1) Chevalier et al. (2019) (2) Covino et al. (2020) (3) Peñil et al. (2020) (4) Prokhorov & Moraghan (2017) (5) Sandrinelli et al. (2018) (6) Tavani et al. (2018) (7) Zhang et al. (2017a) (8) Zhang et al. (2017b) (9) Zhang et al. (2017c)

integral number of counts between two energies as the normalization.

The LCs are obtained by running the entire chain into time bins using `gtlike`, which computes a binned likelihood analysis to find the best-fit model parameters. In astronomy parameter estimation, the use of this method was first described by Cash (1979), and specifically, by Mattox et al. (1996) in γ-ray astronomy. The likelihood analysis is set to find the best-fitting model parameters and provides information about the source spectrum, flux, errors, and other characteristics.

All processes and configurations needed before the application of `gtlike` are automatically performed by the Enrico software and are summarised in the following list:

- `gtselect`. Performs selection cuts on event data files. This helps to select events that are most likely identified as photons in point sources analysis. To do so, we use `evclass=128` to keep only SOURCE class events and `evtype=3`, which corresponds to standard analysis in Pass 8. We use a region of interest (ROI) of 10 degrees centred in the analysed source coordinates [RA, DEC]. Events with a zenith angle greater than 100° are rejected to reduce the γ-ray contamination of the Earth limb.

- `gtmktime`. Creates Good Time Intervals (GTIs) in which the satellite works in a

standard data-taking mode based on the spacecraft data file variables. As a standard recommendation, high-quality data is selected using `(DATA_QUAL>0)&&(LAT_CONFIG==1)`.

- `gtpsf`. Calculates the effective PSF, as a function of energy at a given source location, averaged over an observation.

- `gtbin`. Bins the LAT event lists in time, energy, and/or space into a format suitable for further analysis. Using CCUBE, a counts cube is created, i.e., a three-dimensional map of the number of γ-ray counts detected in each energy bin and for each pixel in the sky. In our analysis, the energy binning is between 1 GeV and 300 GeV.

- `enrico_xml`. Generates an XML file source model containing information about the source's spatial and spectral parameters, which serves as an input to `gtlike`. The background emission is modelled adopting the Galactic diffuse emission model (`gll_iem_v07.fits`), as well as the Extragalactic isotropic diffuse emission model (`iso_P8R3_SOURCE_V3_v1.txt`) which includes the cosmic-ray background.

- `gtltcube`. Creates a cube covering the entire sky of the integrated livetime. This carries information about the accumulated time during which the LAT actively takes event data and depends on the angle between the direction to a source and the instrument z-axis.

- `gtexpcube2`. Generates a binned exposure map by applying the livetime cube at each position in the ROI. In order to account for the contribution from all the sources surrounding the target that could affect the exposure map, sources up to 10 degrees outside the ROI are considered.

- `gtsrcmaps`. Convolves source model components with instrument response. It takes as input each source spectrum in the XML model, the counts map created with `gtbin`, and the binned exposure map created with `gtexpcube2`, convolving the exposure with the effective PSF.

Finally, `gtlike` is applied using the `MINUIT` fit optimizer (James & Roos, 1975), that explores the parameter space to find the minimum of the objective function. Within each time bin, the pipeline generates a LC point unless the test statistic (TS, see the box below) is under 9, in which case an UL is derived. As justified in Section 6.1.1, the logarithm of the flux is employed before applying the algorithm.

TS

The TS used to assess the significance of a potential signal is typically the likelihood ratio TS. In the Fermi-LAT analysis, he TS is calculated as the difference in log-likelihood values between two competing hypotheses: the null hypothesis (H0), that assumes that there is no source present at a specific position, and the alternative hypothesis (H1), that considers the presence of a source at that position:

$$\mathrm{TS} = 2\left(\log L(\mathrm{H1}) - \log L(\mathrm{H0})\right)$$

where $L(\mathrm{H1})$ is the likelihood under the alternative hypothesis, and $L(\mathrm{H0})$ is the likelihood under the null hypothesis.
The TS is often converted to a detection significance, usually expressed as the number of standard deviations (sigma) above the background level. In Fermi-LAT analysis, a TS value of around 25 is commonly considered as a 5-sigma detection, indicating a high confidence in the presence of a source.

In the Figure 7.1 we show examples for different LCs computed with Enrico.

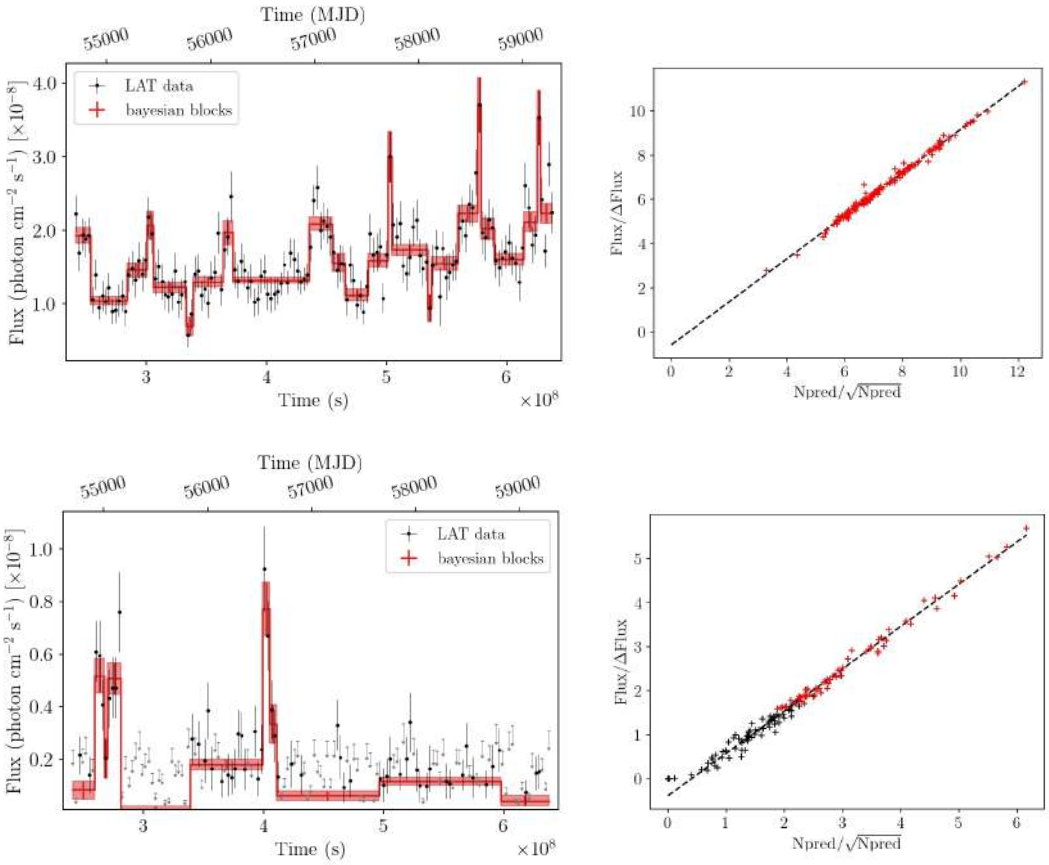

Figure 7.1. [*Top Panels*] Example of 4FGL J1555.7+1111 for a regularly sampled LC computed with Enrico. [*Bottom Panels*] Example of 4FGL J1303.0+2434 for a irregularly sampled LC computed with Enrico. [*Left Panels*] Flux LAT points and Bayesian blocks. [*Right Panels*] $N_{pred}/\sqrt{N_{pred}}$ vs Flux/ΔFlux.

In the Top Panels, we show the example of a regularly sampled LC for J1555.7+1111. On the other hand, the bottom panel show an irregularly sampled LC (J1303.0+2434) where most points are computed as UL. To check the results, in the right panels, we show a plot $N_{pred}/\sqrt{N_{pred}}$ vs Flux/ΔFlux. As suggested by Sanchez & Deil (2013), a good correlation between the variables implies a good computation of the errors. In the LCs plots, a Bayesian block analysis is included.

As we are considering only uniformly sampled data, LCs with flux gaps, as for J1303.0+2434, are removed for the periodicity study. Thus, the final AGN sample is limited to the sources included in the following section.

7.2 Detected periodic sources

After obtaining the LCs of all the sources in the AGN sample after the analysis chain explained in the previous section, the MCMC algorithm is applied following the procedures in Chapter 6 using the following configuration.

First, as we only work with regularly sampled data, we use AR models for the stochastic terms. Both stochastic and deterministic terms are explained in detail in Section 6.2.1. To account for observational uncertainties in the Fermi-LAT data, each model includes a normal distribution $N(0, \epsilon_{obs})$ with standard deviation ϵ_{obs} corresponding to the measurement systematic 1σ errors in flux.

As the data is standardized, we can easily set the scale for the model parameters' prior distributions. We draw the prior for the offset ($\bar{\phi}$), the amplitude terms (A, B), the slope (C), and the AR terms (β_1, β_2) as a normal distribution around 0 with a standard deviation of 2. The WN parameter prior (σ) is uniformly distributed between $[1 \times 10^{-3}, 1 \times 10^{3}]$.

The posterior distribution is limited for the period parameter between a minimum and a maximum value. In the original scale, the lower value is set to 500 days to avoid the MCMC chains from being stuck in a possible artificial period[3] of 1 year and close values. The upper limit is set slightly above half of the data time span ($\sim$ 2200 days), so the period detection can be representative. The prior distribution is a normal centered in the middle of the space drawn for this parameter ($\sim$ 1350 days), with a standard deviation of $\sim$ 800 days.

For each sampling, three independent chains were run, using 8000 iterations with a burning length of 4000 samples.

7.2.1 Results of the MCMC search

For all sources, a periodic model is preferred, and the periodic signal is assessed through the computation of the p_{value} by its relationship with ΔAIC (eq.6.53). The results are shown in Table 7.2.

All AGNs analyzed present a correlated colored noise, depicted from the AR(N) terms. Including these components over WN is fundamental in terms of significance regarding the model selection procedure presented in Section 6.3. The $\Delta\text{AIC}_{\text{models}} = \text{AIC}_{\text{WN}} - \text{AIC}_{\text{AR}}$ are between 10 and 60 for all sources, which is more than sufficient to reject a WN model over a colored noise one. Nonetheless, it is essential to remark that using an AR noise

[3] https://fermi.gsfc.nasa.gov/ssc/data/analysis/LAT_caveats_temporal.html

Table 7.2. MCMC fit results of the AGN Fermi-LAT sample. For each source, the list indicates: the best model in terms of AIC; the AIC value; the period mean and standard deviation in days; the period 95% HDI in days; the ΔAIC between the periodic and the noise model; the p-value computed from ΔAIC. * indicates specific prior assumption and ** indicates inferior posterior convergence for period (see Section 7.2.2). The results are sorted by MCMC period detection significance and a separation at $p_{value} = 5 \times 10^{-3}$ is made.

4FGL Name	Best Model	AIC	Period	Period $HDI_{95\%}$	ΔAIC	p_{value}
J1555.7+1111	AR(1) + lin + sin	298.03	774 ± 10	755 - 793	28.98	1.2×10^{-7}
J2158.8-3013	AR(1) + lin + sin	340.07	614 ± 16	589 - 642	9.85	1.2×10^{-3}
J1903.2+5540	AR(1) + lin + sin	381.92	1120 ± 95	1040 - 1230	8.7	2.1×10^{-3}
J0303.4-2407	AR(1) + lin + harm	323.83	821 ± 40	761 - 870	5.58	8.2×10^{-3}
J0521.7+2112	AR(2) + sin	309.1	1136 ± 128	990 - 1280	5.52	9.2×10^{-3}
J1248.3+5820*	AR(2) + sin	383.73	2048 ± 169	1800 - 2350	4.29	1.6×10^{-2}
J0211.2+1051	AR(1) + harm	301.31	1398 ± 122	1190 - 1630	3.62	1.8×10^{-2}
J0449.4-4350**	AR(1) + lin + sin	296.29	746 ± 229	505 - 1030	3.6	2.2×10^{-2}
J0457.0-2324*	AR(1) + sin	293.61	1300 ± 153	975 - 1590	2.11	4.4×10^{-2}
J2202.7+4216*	AR(1) + lin + sin	261.19	1799 ± 219	1430 - 2250	2.01	4.6×10^{-2}
J0721.9+7120**	AR(1) + sin	321.08	987 ± 220	574 - 1520	1.94	4.7×10^{-2}
J0818.2+4222**	AR(2) + sin	360.61	955 ± 356	501 - 1790	1.68	5.3×10^{-2}
J0428.6-3756*	AR(1) + lin + sin	288.65	1310 ± 175	889 - 1650	0.94	7.4×10^{-2}
J0210.7-5101**	AR(2) + lin + sin	210.93	1080 ± 351	502 - 1640	0.68	8.3×10^{-2}

model decreases the significance of periodic signals compared to WN models. This means that, in most sources analyzed, $\Delta AIC_{WN} = AIC_{WN\ only} - AIC_{WN\ periodic}$ is greater than $\Delta AIC_{AR} = AIC_{AR\ only} - AIC_{AR\ periodic}$. Considering only WN models could lead to the detection of non-robust periodicities, which can be related to AR features in the LCs.

The most significant periodic signal is that for PG 1553+113, with a p_{value} of 1.2×10^{-7} denoting solid evidence of the source's periodic behaviour. The period found of 774 days($\sim$ 2.1 years) is compatible with previous periodicity studies on the source. Besides, the LC shows a clear linear trend $C = 9 \times 10^{-5}$ and a small noise/AR term $\beta_1 = 0.22$. Including a linear trend in the fit of the source's behaviour is also fundamental in terms of the significance of periodicity. In the fit without a linear trend, the same period has a much lower $p'_{value} = 3.4 \times 10^{-4}$ and a much higher noise term $\beta'_1 = 0.49$.

The next sources with the most significant periodic signal are PKS 2155-304 and TXS 1902+556, with p_{values} below 0.005. For TXS 1902+556, a period of 1120 days ($\sim$ 3.1 years) is identified. This value is not in agreement with previous literature periodic analysis, Peñil et al. (2020), where a period of $\sim$ 3.8 years is found in the low significance level ($> 2.5\sigma$). On the other hand, for PKS 2155-304, the fitted periodic signal 614 days ($\sim$ 1.7 years) is compatible with the literature. Both sources also show a linear trend and AR noise behaviour. The descriptions above for the three most significant periodic source's MCMC fits are shown in Figure 7.2.

Two sources, PKS 0301-243 and RX J0521.7+2112, are below a p_{value} of 0.01, with a period of 821 days ($\sim$ 2.3 years) and 1136 days ($\sim$ 3.1 years), respectively. The first value is higher and incompatible with literature, where a period of $\sim$ 2 years is found in the high significance level ($> 3\sigma$). A literature compatible high significant period of $\sim$ 2.8 years is found for RX J0521.7+2112. This source does not show a linear trend, and it is dominated by an AR component of second order with $\beta_1 = 0.38$ and $\beta_2 = 0.24$. PKS 0301-243 shows a particular sinusoidal behaviour, a principal periodic component of 821 days ($\sim$ 2.3 years) with a second harmonic. The result of this source is shown in

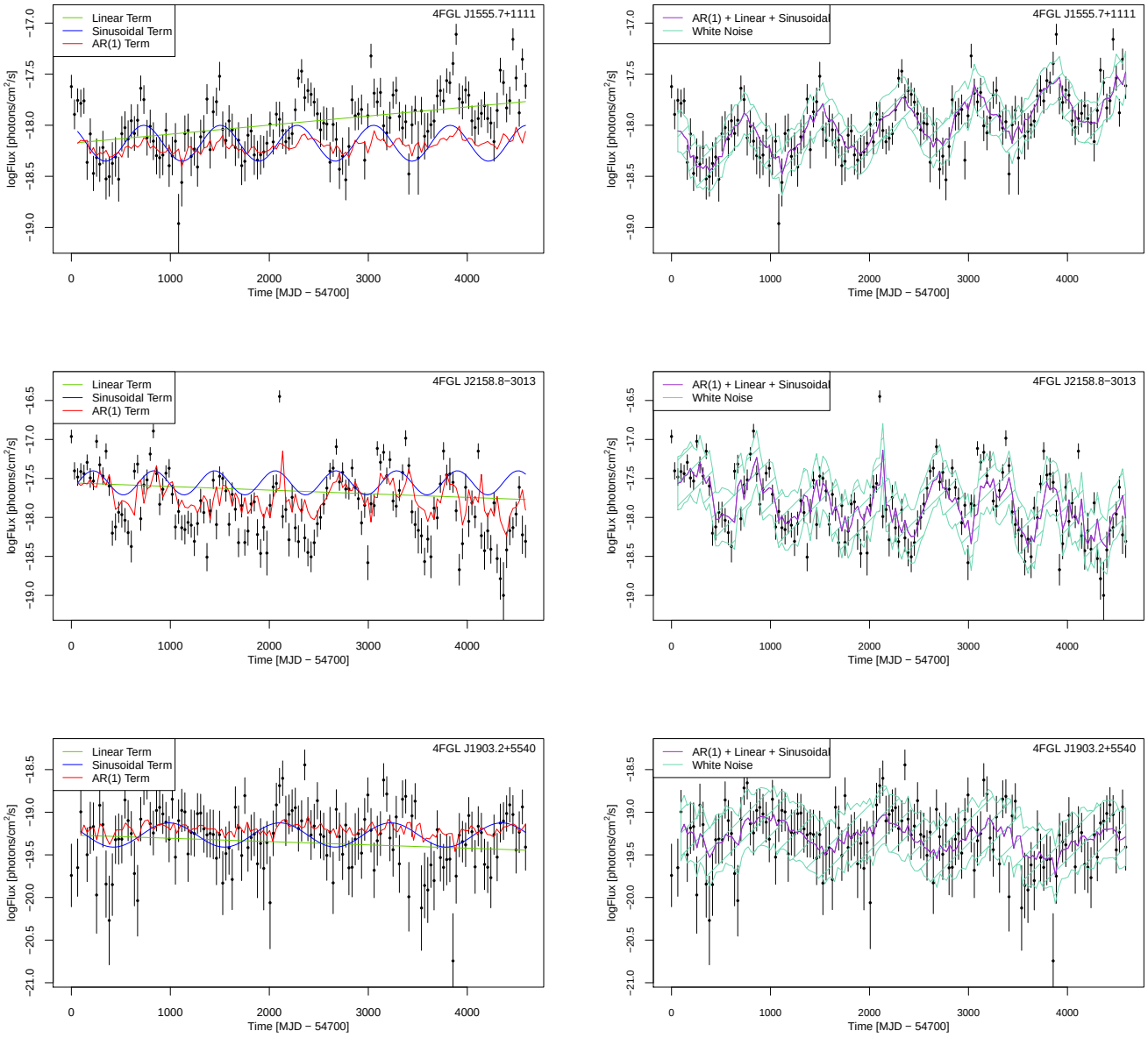

Figure 7.2. MCMC fit for the three sources with the most significant periodic signals. [*Left Panel*] Separated fitted components of the model. Green line indicates the linear trend. Blue line indicates the sinusoidal term. Red line indicates the stochastic term [*Right Panel*] General fit and WN σ component.

Figure 7.3. For CGRaBS J0211+1051, a harmonic oscillation of 1398 days ($\sim$ 3.8 years) is also found with lower significance. The appearance of a harmonic component in the source LC will be discussed in Section 7.4.

The periodicity significance of the remaining AGNs in Table 7.2 is low ($p_{value} > 0.01$), and the MCMC performance is not as good as for the high significance cases. As a result, the standard deviation of the period parameter increases as well as the 95% Highest Density Intervals (HDI), which, for some of the sources, is stuck either in the lower or the upper posterior limit. Also, the colored noise terms are much higher, all above 0.5.

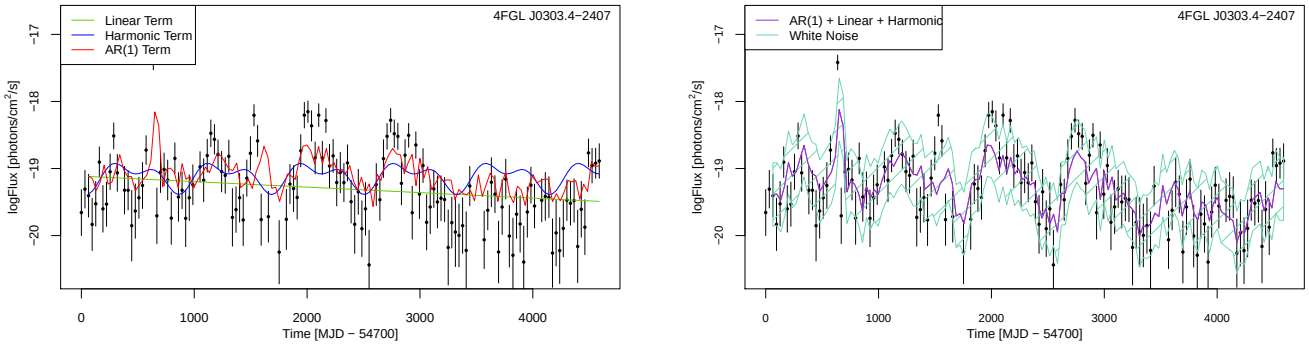

Figure 7.3. MCMC fit for PKS 0301-243. [*Left Panel*] Separated fitted components of the model. Green line indicates the linear trend. Blue line indicates the sinusoidal harmonic term. Red line indicates the stochastic term. [*Right Panel*] General fit and WN σ component.

Spectral cross-check using Agatha

After completing the time-series analysis through the MCMC fits, a spectral analysis, presented in Section 6.4, is performed as a cross-check. Agatha's approach to noise differs from our models but is closely related. They compute the periodogram considering correlated noise described by a MA component of order k. Using the Agatha dedicated noise model selection tool, we get the same AR order k of correlated noise as the one retrieved in the MCMC fit. The results of the Agatha BFP analysis using a MA(k) stochastic model can be found in Table 7.3.

Table 7.3. Agatha cross-check comparison of the AGN Fermi-LAT sample. For each source, the list indicates: the MCMC period mean and standard deviation in days; the Agatha period mean and standard deviation; the Agatha lnBF. The results are sorted by MCMC period detection significance.

4FGL Name	$\text{Period}_{\text{MCMC}}$	$\text{Period}_{\text{AGATHA}}$	lnBF
J1555.7+1111	774 ± 10	771 ± 29	10.9
J2158.8-3013	614 ± 16	615 ± 26	6.3
J1903.2+5540	1120 ± 95	1163 ± 55	0.6
J0303.4-2407	821 ± 40	773 ± 26	-1.2
J0521.7+2112	1136 ± 128	1139 ± 73	11
J1248.3+5820	2048 ± 169	2039 ± 133	3
J0211.2+1051	1398 ± 122	1446 ± 59	3
J0449.4-4350	746 ± 229	669 ± 14	7.2
J0457.0-2324	1300 ± 153	1330 ± 59	7.9
J2202.7+4216	1799 ± 219	1763 ± 89	0.1
J0721.9+7120	987 ± 220	1011 ± 96	7.3
J0818.2+4222	955 ± 356	1333 ± 21	2.1
J0428.6-3756	1310 ± 175	1262 ± 114	13.2
J0210.7-5101	1080 ± 351	1025 ± 13	3.9

As an example, BFPs of the sources represented in Figure 7.2 are shown in Figure 7.4.

Of the 14 periodic LCs analyzed, the Agatha periods for all sources but for 4FGL J0818.2+4222 and 4FGL J0303.4-2407 are compatible with the MCMC results. Nonetheless, the two most significant signals of 4FGL J0818.2+4222 are 1337 days and 805 days, with very close lnBF values of 2.1 and 1.6, respectively. The second most significant signal at 805 days is compatible with the MCMC result at 867 days. For 4FGL J0303.4-2407,

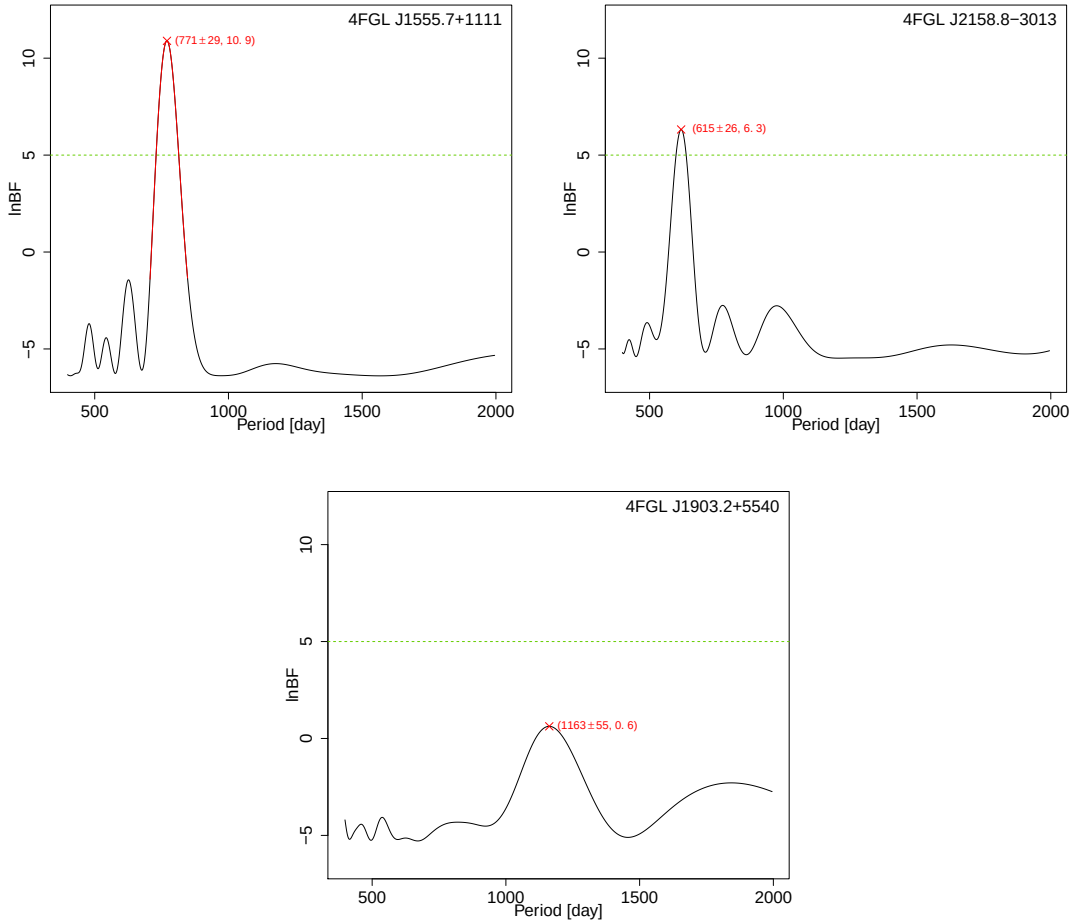

Figure 7.4. Agatha results for the three sources with most significant MCMC sinusoidal signals. The green dashed line indicates the 5 lnBF significance as prescribed by Feng et al. (2017).

a lnBF=-1.2 means that a noise model is favored over the periodic model at this given period. This might be due to the presence of a harmonic signal in the MCMC fit, which is obtained with Agatha as a double-peaked periodogram corresponding to each oscillation order. 4FGL J1903.2+5540 and 4FGL J2202.7+421 have a lnBF=0.6 and lnBF=0.1, respectively, meaning that the periodic model is poorly favored over the noise model. All the remaining sources present compatible periods with an important level of significance, i.e., lnBF$\geq$ 3.

As will be studied in Section 7.3, the time stability of the period and the variation of the amplitude of the periodic modulation as a function of the period are two important clues of the physical mechanism causing the oscillations. The local period and amplitudes can be obtained by building a spectrogram with time windows of the LC.

7.2.2 Systematics of the MCMC search

Period prior dependence

For the sources marked with *, in Table 7.2 we take the period values from a good posterior distribution obtained using specific priors that differ from the general prior

discussed in Section 6.2. For the sources marked with **, the MCMC chains exhibit poorer convergence for the period parameter, resulting in posterior distributions that deviate from the expected symmetric Gaussian shape.

Regarding the high-significance sources, changes in the prior distribution have negligible influence on the MCMC sampling, and a vague prior is sufficient for the MCMC fit. The posterior distributions follow approximately symmetric Gaussians, from which we obtain the parameters as the mean value with standard deviation. This is illustrated in the left panel of Figure 7.5 for PG 1553+113.

The center-left panel of Figure 7.5 demonstrates an example of poorer MCMC sampling where the chains fail to converge, resulting in a wider posterior distribution that extends to lower and higher values around the peak. Nevertheless, the peak of the distribution remains close to its mean value. This situation applies to low-significance sources marked with ** in Table 7.2, where altering the prior distribution does not significantly impact the output.

For low-significance sources marked with *, using different priors may yield improved posterior results. As shown in the center-right panel of Figure 7.5, when employing a general prior, the chain convergence becomes problematic. Consequently, the HDI and standard deviation are larger, and the mean of the distribution deviates from the peak value. However, by centering the prior distribution around the highest value and reducing its standard deviation, the posterior distribution becomes closer to a symmetric Gaussian, as observed in the right panel of Figure 7.5. This demonstrates a strong dependence between the posterior and the choice of the prior distribution. Therefore, the acceptance of these results is lower and can be correlated with their inferior significance.

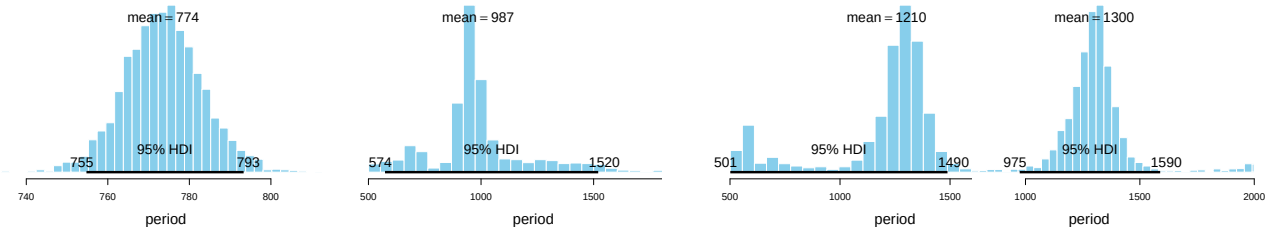

Figure 7.5. Examples of posterior distributions for the period parameter. [*Left panel*] 4FGL J1555.7+1111 using general prior. [*Center-left panel*] 4FGL J0721.9+7120 using general prior. [*Center-right panel*] 4FGL J0457.0-2324 using general prior. [*Right panel*] 4FGL J0457.0-2324 using specific prior.

Tests for normality

As suggested in Section 6.3.2, we have tested the results of the LC fits for normality by performing an AD test on the residuals. We also fitted a Gaussian distribution $N(\text{mean}, \sigma)$. The results are presented in Table 7.4; an example is illustrated in Figure 7.6.

The p-value from the AD test allows us to reject the hypothesis of normality if its value is equal to or less than 0.05. For all sources except 4FGL 1903.2+5540, the p-value exceeds 0.05. However, for 4FGL 1903.2+5540, we observe some outliers on the left side of the residual distribution, indicating a failure of the normality test, as shown in Figure 7.6.

By analyzing the Gaussian distribution fit, we can draw several conclusions. As expected from well-fitted data, the mean value of the residuals is centered close to 0. Addi-

tionally, the standard deviation values N_σ align with those obtained from the MCMC fit WN term $\sigma_{\rm MCMC}$. An example is depicted in Figure 7.7.

Table 7.4. Test for normality of the AGN Fermi-LAT sample. For each source, the list indicates: the p-value of the AD test; Gaussian distribution fit mean $N_{\rm mean}$ and standard deviation N_σ; the WN σ component in the MCMC fit.

4FGL Name	$p_{valueAD}$	N_{mean}	N_σ	$\sigma_{\rm MCMC}$
J1555.7+1111	0.89	-1×10^{-4}	0.2	0.2
J2158.8-3013	0.36	9×10^{-4}	0.33	0.34
J1903.2+5540	9×10^{-4}	-1.8×10^{-3}	0.33	0.34
J0303.4-2407	0.38	7×10^{-3}	0.45	0.47
J0521.7+2112	0.66	9.7×10^{-3}	0.41	0.43
J1248.3+5820	0.43	5×10^{-4}	0.3	0.31
J0211.2+1051	0.21	7.3×10^{-3}	0.46	0.46
J0449.4-4350	0.29	1.4×10^{-3}	0.33	0.34
J2202.7+4216	0.29	-2.6×10^{-3}	0.53	0.55
J0818.2+4222	0.324	8.7×10^{-2}	0.38	0.38
J0721.9+7120	0.4	7×10^{-4}	0.52	0.53
J0457.0-2324	0.3	1×10^{-2}	0.55	0.56
J0428.6-3756	0.06	4×10^{-3}	0.51	0.52
J0210.7-5101	0.86	7.8×10^{-2}	0.59	0.61

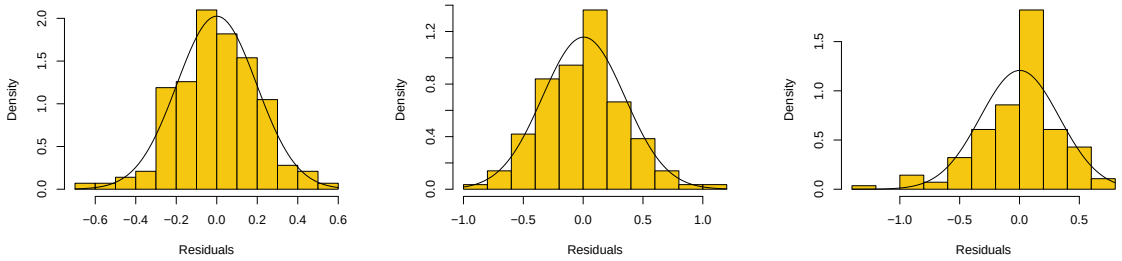

Figure 7.6. Residuals histogram for the three sources with most significant periodic signals. Black line shows a Gaussian distribution fit with $N_{\rm mean}$ and N_σ parameters given in Table 7.4. [*Left panel*] 4FGL J1555.7+1111 [*Center panel*] 4FGL 2158.8-3013 [*Right panel*] 4FGL J1903.2+5540.

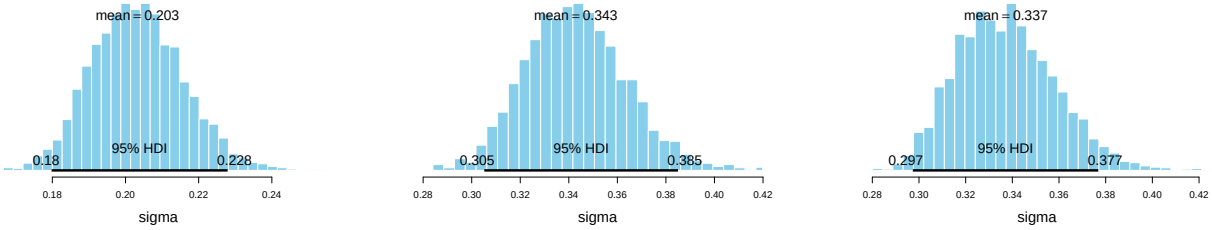

Figure 7.7. Examples of posterior distributions for the sigma parameter for the three sources with most significant periodic signals. [*Left panel*] 4FGL J1555.7+1111 [*Center panel*] 4FGL 2158.8-3013 [*Right panel*] 4FGL J1903.2+5540.

Correlation between parameters

For every source analyzed, the correlation between MCMC parameters is not as important as to affect the efficiency of the sampling chains. There are correlations with values ~ 0.6 between periodic parameters (period, A and B) and between AR parameters (β_1, β_2) with the others in the model. In the Figure 7.8 the corner plot for PG 1553+113 is shown, useful to visualize the pairwise correlations between model parameters.

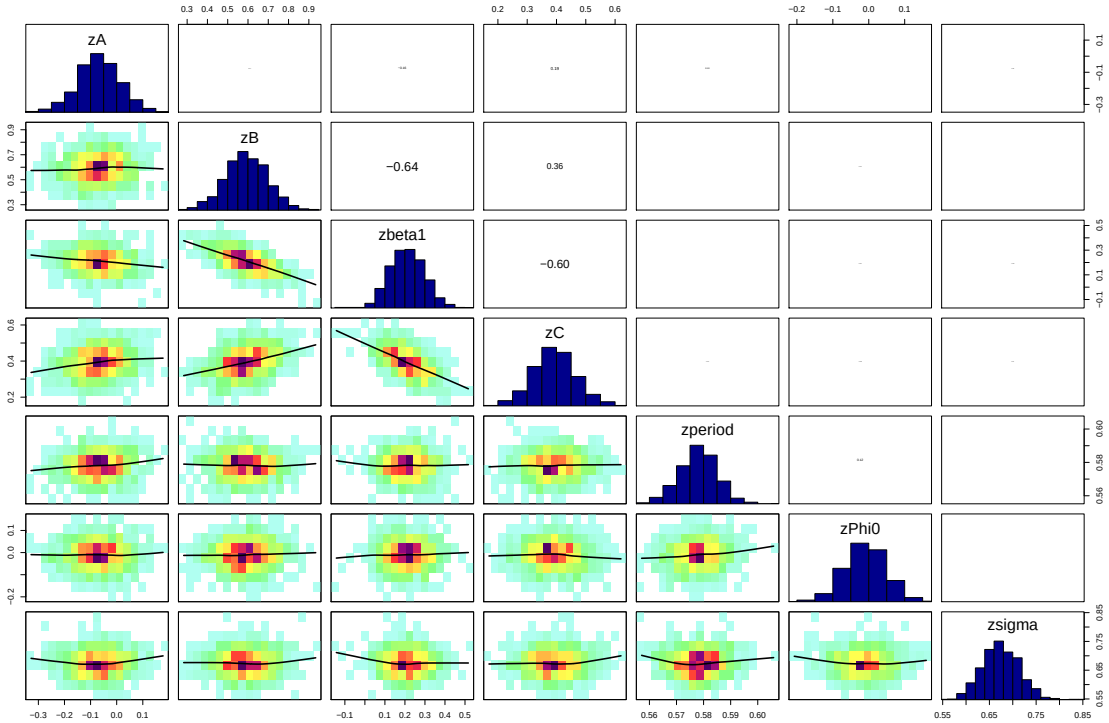

Figure 7.8. Example of a corner plot for 4FGL J1555.7+1111. The diagonal plots show the posterior distributions for every standardize parameter (marked as z). The 2D maps below the diagonal show the density of the posterior distribution with one parameter in each axis. The values above the diagonal show the Pearson correlation coefficients between parameters. See Section 6.2.1 for a parameters' description.

7.3 Periodicity temporal stability

For this analysis, several time windows are selected to search for periodicity change in the LCs time span. Each time window has a width of 0.4 the full LC time span, this is, 0.4×4620 days ~ 1848 days (~ 5 years). Thus, only the most significant periodic sources with periods smaller than 900 days are studied, i.e., 4FGL J1555.7+1111 and 4FGL J2158.8-3013. A total of 5 time windows are computed, each centred in 920, 1607, 2294, 2981, and 3669 days from the start, respectively. The MCMC fits are applied at each time window for every source included. The results are shown in Table 7.5 and Figure 7.9 shows the LCs analysed with the fitted MCMC components at each time window.
Then, an Agatha moving periodogram is performed as a cross-check, as explained in Section 6.4. The results are shown in Table 7.6. Figure 7.10 shows the temporal variation of the sources period for MCMC and Agatha analysis.

Table 7.5. MCMC fit results of the AGN time windows. For each source and window, the list indicates: the best model in terms of AIC; the AIC value; the period mean and standard deviation in days; the period 95% HDI in days; the ΔAIC between the periodic and the noise model; the p-value computed from ΔAIC.

4FGL Name		Best Model	AIC	$\text{Period}_{\text{MCMC}}$	$\text{Period HDI}_{95\%}$	ΔAIC	p-value
J1555.7+1111	1	AR(1) + sin	131.5	735 ± 32	671 - 798	14.32	1.5×10^{-4}
	2	AR(1) + lin + sin	125.26	838 ± 56	732 - 945	8.14	2.7×10^{-3}
	3	AR(1) + lin + sin	132.44	835 ± 74	701 - 984	5.39	9.8×10^{-3}
	4	AR(1) + lin + sin	128.86	758 ± 31	698 - 821	14.2	1.5×10^{-4}
	5	AR(1) + lin + sin	136.23	748 ± 61	651 - 858	9.96	1.2×10^{-3}
J2158.8-3013	1	AR(1) + lin + sin	125.16	665 ± 34	596 - 732	12.78	3.0×10^{-4}
	2	AR(1) + lin + sin	142.25	617 ± 51	553 - 684	8.11	2.8×10^{-3}
	3	AR(1) + lin + sin	134.96	564 ± 25	522 - 609	13.07	2.6×10^{-4}
	4	AR(1) + sin	123.82	606 ± 57	535 - 669	7.51	3.7×10^{-3}
	5	AR(1) + lin + harm	130.46	678 ± 24	633 - 723	8.97	1.9×10^{-3}

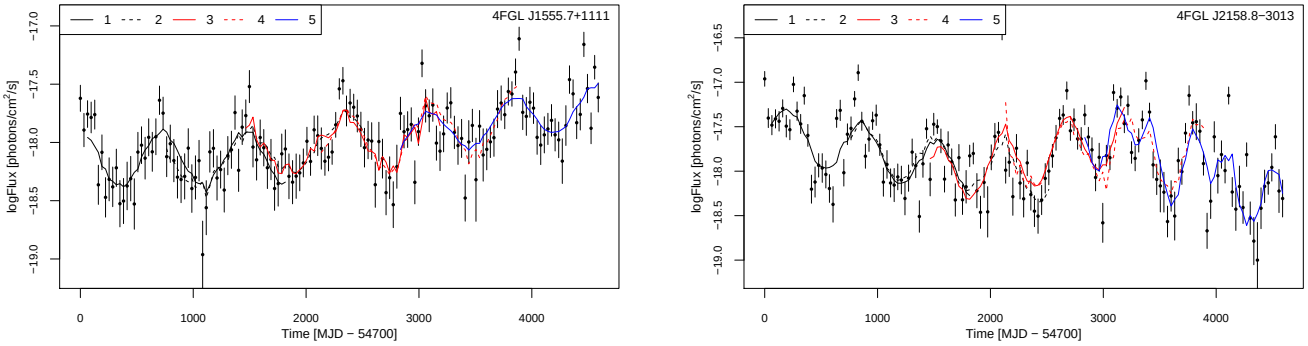

Figure 7.9. MCMC global fit of the five time windows analysis for the Fermi-LAT sample sources [*Left Panel*] J1555.7+1111 [*Right Panel*] J2158.8-3013

Table 7.6. Agatha cross-check comparison of the AGN time windows. For each source and window, the list indicates: the MCMC period mean and standard deviation in days; the Agatha period mean and standard deviation: the Agatha lnBF.

4FGL Name		$\text{Period}_{\text{MCMC}}$	$\text{Period}_{\text{AGATHA}}$	lnBF
J1555.7+1111	1	735 ± 32	737 ± 74	4.7
	2	838 ± 56	851 ± 91	4.4
	3	835 ± 74	855 ± 118	3.8
	4	758 ± 31	772 ± 63	4.7
	5	748 ± 61	746 ± 98	1.5
J2158.8-3013	1	665 ± 34	705 ± 88	3.3
	2	617 ± 51	623 ± 44	2
	3	564 ± 25	575 ± 51	3.8
	4	606 ± 57	596 ± 43	3.9
	5	678 ± 24	688 ± 64	-0.2

For PG 1553+113, a periodic component is found in every time window with a p-value below 10^{-2}. Windows 1, 4 and 5 present a period below the entire LC period of 774 days. For windows 2 and 3, the period is $\sim$ 840 days with larger error bars. A total difference of 105 days between the minimum and maximum periods was found. To quantify this dispersion, a constant fit is performed with a χ^2 test, considering the standard deviations,

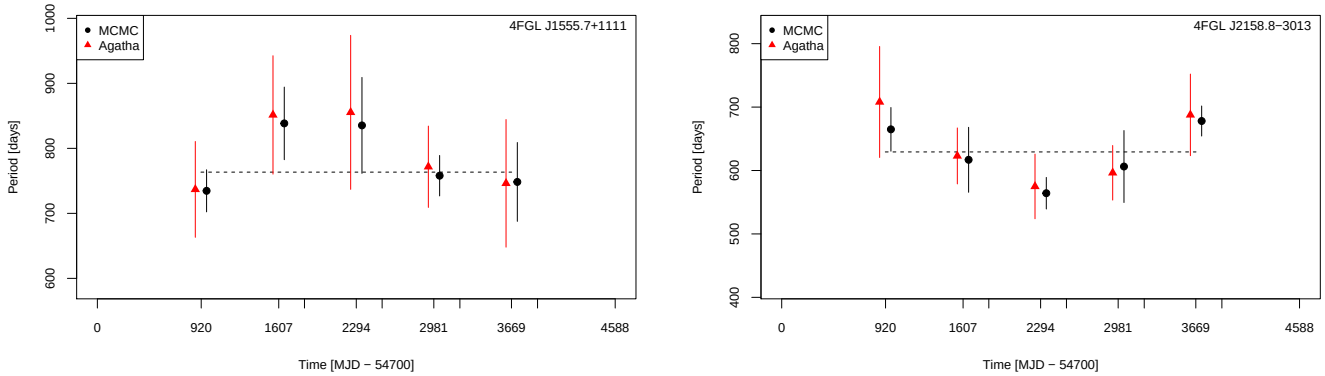

Figure 7.10. MCMC and Agatha period mean and standard deviation for each time window. Horizontal dashed line indicates the result of the constant fit value.

that indicates the probability that the different time-window periods come from a constant distribution. The fit estimate of the constant value is 756 ± 16 days, with a p-value equal to 0.64. Agatha values agree with the MCMC results with a significance above 3.8 for every time window except the last one with a value of 1.5.

For PKS 2155-304, the periodic components are also found with p-values below 10^{-2}. Windows 1 and 2 present a period higher than the entire LC period of 613 days. Windows 3 and 4 present a period below this value. For window 5, the best-fitted model corresponds to an AR(1) stochastic component, a linear trend and a sinusoidal with a second harmonic component. Again, the χ^2 test is computed, resulting in a constant value of 618 ± 25 days with a p-value of 0.02.

As explained in Appendix A, the transformed amplitude Z' of the periodic modulations is obtained from the MCMC posterior values of parameters A and B. Then, the physical amplitude Z is derived from Z' through U. Physical amplitudes as a function of the period are shown in figure 7.11. The χ^2 test is computed to quantify the amplitude variation with the period. The test gives a p-value of 0.87 and 0.95 for PG 1553+113 and PKS 2155-304, respectively, concluding that there is no variation in the oscillation amplitudes.

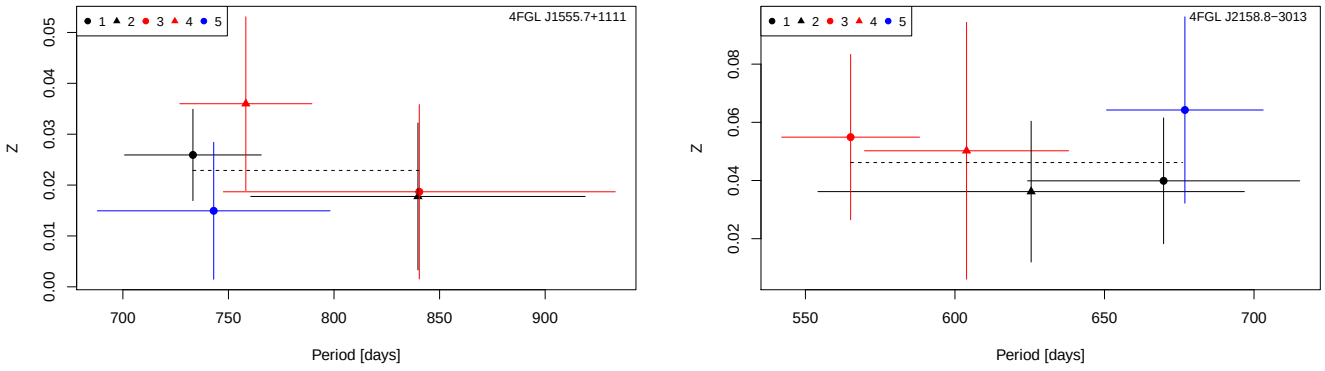

Figure 7.11. MCMC physical amplitude Z mean and standard deviation for each time-window period. Horizontal dashed line indicates the result of the constant fit value.

7.4 Discussion

In this chapter, a novel method has been employed to investigate the presence of periodicity in several HE source candidates from the Fermi-LAT 4FGL catalog. The method effectively separates the stochastic and periodic components during the light-curve fitting. Only 14 out of 27 sources in this analysis were found to be periodic, while the presence of AR features in the LCs of AGNs could explain the remaining undetected oscillations. Alternatively, the different energy ranges employed in retrieving the LCs could explain this discrepancy between this work and previous ones.

The inclusion of a periodic component in the models for 4FGL J1555.7+1111, J2158.8-3013, and J1903.2+5540 significantly enhances the data fit compared to models that consider only AR noise. In Section 7.3 we examined the period and amplitude variations over time by dividing the LC into various time windows. We focused on LCs with periods shorter than 900 days, which only includes J1555.7+1111 and J2158.8-3013 for the time-window study.

The measured period of the best candidate 4FGL J1555.7+1111 shows no significant variation over time. Additionally, the amplitude of the periodic term demonstrates only weak dependence on the period. These findings align well with a BSMBH model, specifically the 2-jet model proposed by Tavani et al. (2018). In contrast, models based on the lighthouse effect, such as the model presented by Camenzind & Krockenberger (1992) and extended for longer timescales by Ait Benkhali et al. (2020), would predict an evolution of both the period and amplitude. While these models primarily address short-term variability (≤ 1 year), their applicability to longer timescales could be possible.

In the case of J2158.8-3013, there is a marginally significant drift of the period with time. The amplitude of the oscillating component does not significantly change with time. The lack of correlation between amplitude and period disfavors again models based on the lighthouse effect. An harmonics of the period is detected in one of the time windows. This harmonics is not easily explained by pure stochastic noise (for instance by linear CARMA models). The harmonics could be however the signature of oscillations in the disk of J2158.8-3013. These oscillations would trigger quasi-periodic accretion flows with harmonic frequencies and, by coupling between the disk and the jet, quasi-periodic variations in the observed flux. The period of oscillation $P_{\rm true}$ at the source is related to the observed period $P_{\rm obs}$ and to the source redshift z by

$$P_{\rm true} = \frac{P_{\rm obs}}{(1+z)}. \tag{7.1}$$

For J2158.8-3013, one has $P_{\rm true} = 1.5$ yr. If M_{bl} is the black hole mass, $r_g = \frac{GM_{bl}}{c^2}$ is the gravitational radius, the transition from a disk or a torus to an ADAF occurs at radius

$$r_{tr} = K r_g \left((\frac{10^8 M_\odot}{M_{bl}})(\frac{P_{\rm true}}{1\ {\rm yr}}) \right)^{2/3} \tag{7.2}$$

with $K \simeq 524$ for the transition radius to a disk (Gracia et al., 2003) and $K \simeq 2100$ for the transition to a torus (Liu et al., 2006).

The estimates of the J2158.8-3013 black hole mass have a very large spread (Rieger & Volpe, 2010; Dermer et al., 2008; Aharonian et al., 2007). Taking $M_{bl} = 10^8 M_\odot$ as a typical value, one finds $r_{tr} \simeq 690 r_g$ in the disk model and $r_{tr} \simeq 2800 r_g$ in the torus model.

The value of r_{tr} for the disk model is too large according to Ait Benkhali et al. (2020). The value of r_{tr} for the torus model is similar to the value obtained by Liu et al. (2006) for BL Lac AO 0235+164.

This chapter dealt exclusively with regularly sampled LCs. In the next, we will extend the study to the whole Fermi-LAT data-set, including LCs with holes in the observations and irregular sampling.

Chapter 8

Applying the algorithm to the Fermi-LAT Light Curve Repository

This chapter is related to Sections TBA and TBA of the publication resulting from this work TBA. Here, we add Fermi-LAT irregularly sampled LCs to the dataset as we included the analysis with CAR models for the stochastic component. The LCs are obtained from the Fermi-LAT Light Curve Repository (LCR) (Abdollahi et al., 2023), and the ULs are treated as missing data points for this analysis. A deeper analysis using the UL information is foreseen. The LCR catalogue increases ~ 50 times our sample with respect to Rueda et al. (2022) (Chapter 7). Section 8.1 explains the selection criteria in the LCR, the type of sources included, and the production of LCs from this sample. Results from the MCMC search are presented in Section 8.2, and its temporal stability is examined in Section 8.3. Finally, we discuss all the findings from these analysis in Section 8.4.

8.1 Source selection and Data Analysis

For this study, we use the flux information provided by the Fermi-LAT LCR, a database with calibrated LCs of more than 1500 AGNs from the 10-year 4FGL Data Release 2 (DR2) catalog (Ballet et al., 2020). This catalog uses the same analysis methods as the original 8-year 4FGL. Even though the catalog has information for objects detected up to 10 years after the launch, the LCs are updated regularly.

The LCR is fundamental as it not only provides LCs for thousands of sources but also focuses on variable AGNs, a necessary condition to find periodicities. Only objects with a variability index of at least 21.67 are included in the LCR analysis, a value at which the probability of being a steady source is lower than 1%. The TS for the variability index is computed by comparing the log-likelihood between the flux fitted in time bins and the overall flux level of the source, as in Nolan et al. (2003):

$$TS_{\rm var} = 2\sum_i \frac{\Delta F_i^2}{\Delta F_i^2 + f^2 F_{\rm const}^2} \ln\left(\frac{L_i(F_i)}{L_i(F_{\rm const})}\right)$$

where F_i and ΔF_i are the flux and error in flux in the ith time bin, respectively, and $F_{\rm const}$ is the constant flux. $L_i(F_{\rm const})$ is the value of the likelihood in the ith bin under the null hypothesis where the source flux is constant across the full period for this hypothesis, and $L_i(F_i)$ is the value of the likelihood in the ith bin under the alternate hypothesis where

the flux in the ith bin is optimized.

In Figure 8.1, we show a skymap of the positions of every source included in the LCR. These representations show the objects as point sources with different sizes corresponding with the variability index, and the Milky Way diffusion emission. We see that the distribution of variable objects is homogeneous in the sky. The majority of these sources belong to the AGN blazar class, which can be further classified as flat spectrum radio quasars (fsrq), BL Lacs (bll), and blazar candidates of unknown type (bcu).

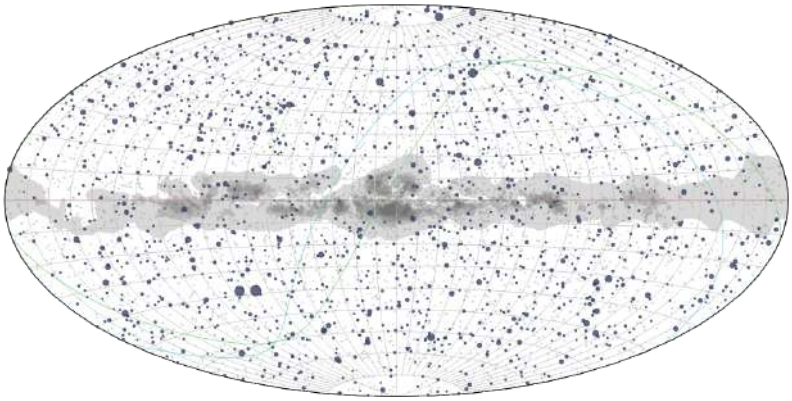

Figure 8.1. Skymap for the 1525 variable sources included in the LCR. The marker size is related to the variability index. The light gray points represent the non-variable sources in the 4FGL-DR2. The gray areas represent the Galactic Plane emission.

The LCs in the LCR are obtained by performing an unbinned likelihood analysis. The procedure is similar to Section 7.1. While in a binned likelihood analysis, the observed data are divided into several energy and spatial bins, in the unbinned analysis, the maximum likelihood optimization incorporates the complete spatial and spectral information of each photon for analysis. This method is more computationally intensive, but it retains all of the information in the data and can be more sensitive to subtle features. Thus, this method is preferred for time series analysis. See Abdollahi et al. (2023) for a detailed description of the data analysis procedure and the configuration used.

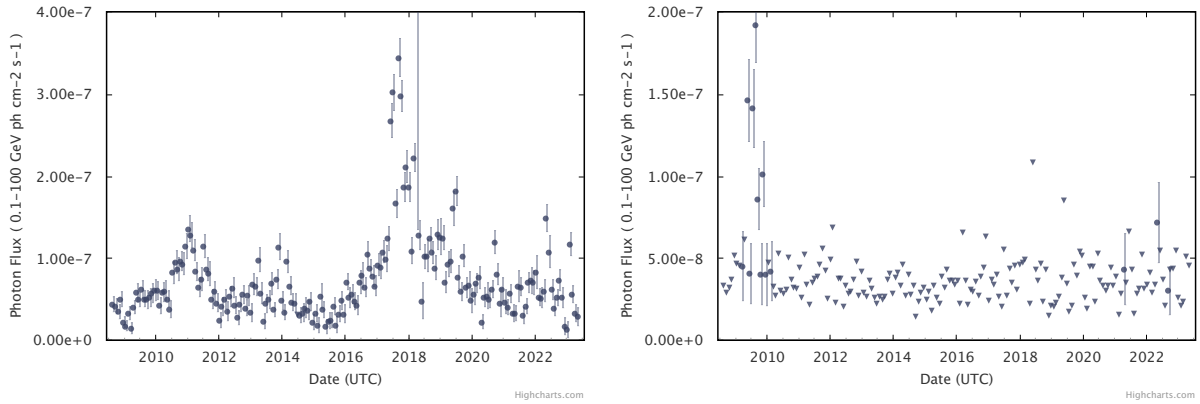

Figure 8.2. Example of LCs from the Light Curve Repository. [*Left Panel*] 4FGL J0509.4+0542 well sampled LC. [*Right Panel*] 4FGL J0358.9+6004 poorly sampled LC. Triangle marker indicates UL.

An observed LAT flux point is generated when the TS is equal to or greater than 4. Otherwise, an UL is derived. For this work, we use a monthly cadence of the data over more than 14 years of data ($\sim$ 5250 days) from the start of the mission (54682 MJD)

until February 2023 (59982 MJD). As justified in Section 6.1.1, the logarithm of the flux is applied before the MCMC analysis. Figure 8.2 shows LCs examples from the LCR. The right panel shows a regularly sampled LC for 4FGL J0509.4+0542. On the other hand, the left panel shows an irregularly sampled LC (4FGL J1303.0+2434) where most points are computed as ULs.

8.2 Detected periodic sources

Here, we use the MCMC models using CAR(1) presented in Section 6.2.1 to find periods in both regularly and irregularly sampled data from Section 8.1. The MCMC configuration is presented as follows.

The data is standardized as in the previous chapter, so we use similar vague prior distributions for the offset ($\bar{\phi}$), the amplitude terms (A, B, A', B'), and the slope (C). They are represented as a normal distribution around 0 with a standard deviation 2 on the standardized scale. In the original scale, the correlation time τ prior distribution is centered at 50 days with a standard deviation of 50 days, in the order of magnitude of the data's sampling binning. D is uniformly distributed between 10^{-5} and 10^{2}. As in Chapter 7, we account for observational uncertainties in the Fermi-LAT data by including a normal distribution $N(0, \epsilon_{obs})$ with standard deviation ϵ_{obs} corresponding to the measurement systematic 1σ errors in flux. The data points for this analysis are selected from the LCR for a TS above 4 (or 2σ, the maximum lower threshold in the LCR points), leaving some low significant detected points. Thus to avoid outliers, we remove the data points with observational errors larger than three times the mean error value.

This work considers the flux data up to $\sim$ 5250 days from the start. Periodicities above half of this time span ($\sim 2650 days, \sim 7 years$) are not considered significant. Then, the posterior distribution for the period has an UL at this value plus an interval to let the standard deviation be drawn. The lower limit is 400 days to avoid possible $\sim$ 1 year artificial periods. A normal distribution is used for the prior, centered in the space between the posterior limits ($\sim$ 1700 days) and with a standard deviation of $\sim$ 900 days. With this, we intend to use a prior as general and non-informative as possible, only adapting the values to work with all sources in the dataset properly.

In comparison to the work in the previous chapter, here we list only the results obtained from well-sampled MCMC chains. Thus, there is no prior dependence as all the results are taken using the general prior distribution, avoiding the problems discussed in Section 7.2.2.

8.2.1 High significance sources

From the $\sim$ 1500 sources analysed, 21 show explicit periodic behaviour with a CAR(1) model and well-computed MCMC parameters for the deterministic and the periodic components. From this, 13 are newly detected periodic candidates, and for 8, a periodicity was found in previous studies (see Section 5.2). These results are shown in Table 8.1 and Table 8.2, respectively. In Figure 8.3 we exhibit an example of the MCMC posterior distributions retrieved for the parameters represented in the tables.
Figures 8.4 and 8.5 show the MCMC fit for the two most significant sources in each table.

For all these sources, the p_{value} are below 0.005, which assesses the significance of the periodicity, computed from the ΔAIC (Eq. 6.53) between the periodic and the noise

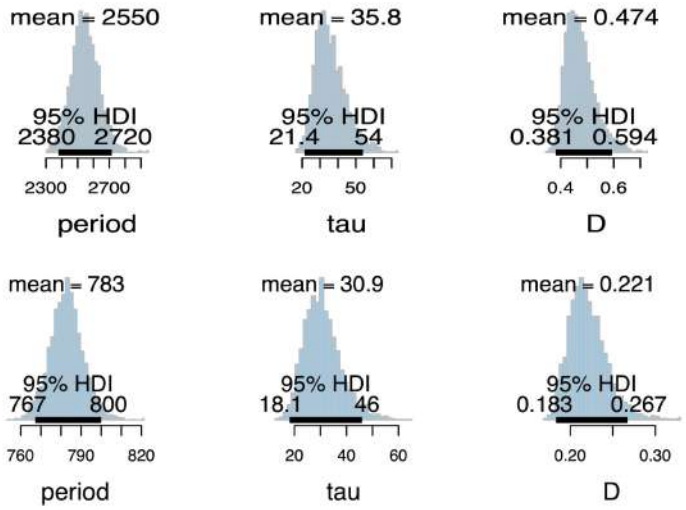

Figure 8.3. MCMC posterior distributions for the sources with the most significant periodic signals in Tables 8.1 and 8.2. [*Top panel*] J0509.4+0542. [*Bottom panel*] J1555.7+1111.

Table 8.1. MCMC fit results of the new periodic AGNs from the Fermi-LAT LCR. For each source, the list indicates: the AGN class; the number of points in the LC; the best model in terms of AIC (l, linear; s, sinusoidal; h, harmonic); the period mean and standard deviation in days; the CAR(1) τ mean in days; the CAR(1) D mean; the p-value computed from ΔAIC; the only noise model τ and D; the test for normality in residuals p-value from the AD test; the residuals Gaussian distribution fit standard deviation N_σ. The results are sorted by MCMC period detection significance.

			CAR(1) Periodic Model					CAR(1) model		Residuals	
4FGL Name	Class	np	Model	Period	τ	D	pv	τ_{noise}	D_{noise}	pv_{AD}	N_σ
J0509.4+0542	bll	175	l, h	$2567_{\pm 78}$	35	0.49	6.3×10^{-8}	96	0.95	0.11	0.39
J0739.2+0137	fsrq	170	l, s	$1743_{\pm 87}$	40	0.74	4.3×10^{-5}	64	0.99	0.75	0.56
J1640.4+3945	fsrq	120	l, s	$2058_{\pm 88}$	43	0.64	5.4×10^{-5}	68	0.85	0.46	0.48
J1740.5+5211	fsrq	145	l, s	$2100_{\pm 99}$	50	0.73	7.4×10^{-5}	79	0.99	0.06	0.52
J1913.0-8009	fsrq	120	h	$2538_{\pm 119}$	39	0.50	9.2×10^{-5}	61	0.66	0.35	0.39
J0141.4-0928	bll	167	l, h	$2279_{\pm 119}$	33	0.49	1.4×10^{-4}	50	0.62	0.54	0.40
J2311.0+3425	fsrq	152	l, s	$1393_{\pm 93}$	68	0.90	1.8×10^{-4}	97	1.2	0.22	0.56
J2243.9+2021	bll	168	s	$1965_{\pm 122}$	30	0.47	6.7×10^{-4}	40	0.55	0.24	0.4
J0921.6+6216	fsrq	160	l, h	$2607_{\pm 188}$	60	0.88	9.3×10^{-4}	82	1.11	0.13	0.55
J1754.2+3212	bll	137	l, s	$2061_{\pm 176}$	31	0.59	1.2×10^{-3}	43	0.70	0.05	0.53
J0850.0+4855	bll	123	l, s	$2482_{\pm 243}$	26	0.66	1.3×10^{-3}	37	0.76	0.12	0.57
J0030.3-4224	fsrq	130	l, s	$1791_{\pm 197}$	28	0.59	6.3×10^{-3}	33	0.64	0.12	0.50
J0217.8+0144	fsrq	150	s	$1189_{\pm 50}$	89	0.82	7.5×10^{-3}	108	0.96	0.25	0.41

model. To assess the goodness of fit following Section 6.3.2, we remark that the p_{value} of the AD test on the residuals is above 0.05 for all the sources listed, meaning that the fit's residuals have no significant deviation from normality. By looking at the number of data points for each source, we can see that the already known periodic sources have almost a regular sampling over time. On the other hand, most of the newly detected sources present an irregular sampling, with a number of points below 160 for a total of 176 time intervals. Regarding the period standard deviations, for most sources, there is a correlation between the significance and the period magnitude, meaning that the standard deviation grows for lower significances and higher periodicity values. When the period value approaches half of the time sampling of the data, fewer oscillations can be obtained,

Table 8.2. MCMC fit results of the known in literature periodic AGNs from the Fermi-LAT LCR.

			CAR(1) Periodic Model					CAR(1) model		Residuals	
4FGL Name	Class	np	Model	Period	τ	D	pv	τ_{noise}	D_{noise}	pv_{AD}	N_σ
J1555.7+1111	bll	176	l, s	$783_{\pm 8}$	31	0.22	9.8×10^{-9}	67	0.35	0.22	0.18
J0521.7+2112	bll	176	h	$2266_{\pm 54}$	42	0.47	1.5×10^{-6}	87	0.77	0.09	0.39
J1903.2+5540	bll	175	l, s	$1131_{\pm 33}$	27	0.30	1.1×10^{-4}	39	0.36	0.64	0.26
J1048.4+7143	fsrq	167	s	$1103_{\pm 37}$	105	1.1	1.5×10^{-4}	140	1.46	0.23	0.52
J0457.0-2324	fsrq	174	l, h	$2561_{\pm 110}$	93	0.77	3.4×10^{-4}	139	1.09	0.05	0.35
J0407.0-3826	fsrq	143	l, s	$1066_{\pm 163}$	51	0.78	2.6×10^{-3}	64	0.90	0.40	0.54
J2158.8-3013	bll	173	l, s	$601_{\pm 33}$	69	0.50	2.7×10^{-3}	84	0.58	0.17	0.28
J0303.4-2407	bll	171	l, h	$835_{\pm 29}$	35	0.56	2.9×10^{-3}	42	0.63	0.07	0.45

increasing the period uncertainty.

All the AGNs in these tables present a stochastic behavior characterized by a CAR model. This is similar to the results in Chapter 7. The AIC is more favorable to correlated noise models, but including such AR components reduces the significance of the periodic signal compared to WN models. Then, within the fit, including a CAR/AR component implies that the amplitude of the periodicity compared to that for a WN model is reduced. All values for τ are close or above the monthly data sampling interval. In Section 8.2.2, we list some examples where the correlated noise is incompatible with the data, then the WN model is preferred, and the amplitude and significance of the periodicities over the noise-only model is much higher.

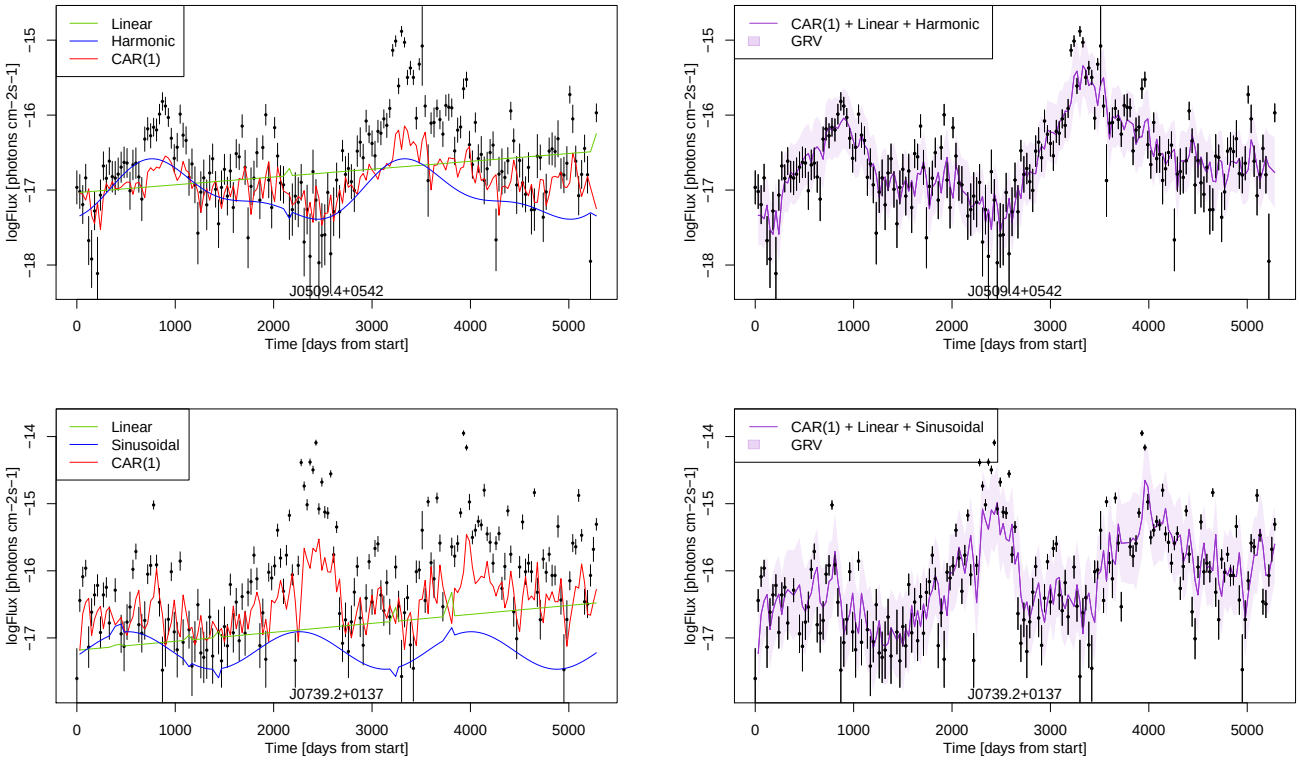

Figure 8.4. MCMC fit for the two sources with the most significant periodic signals in Table 8.1. [*Left Panel*] Separated fitted components of the model. Green line indicates the linear trend. Blue line indicates the sinusoidal term. Red line indicates the stochastic term [*Right Panel*] General fit and GRV component.

Within the new candidates, the most robust evidence of periodic behavior is observed for J0509.4+0542 (Fig. 8.4, Top), with a p-value of 6.3×10^{-8}. This indicates a highly significant periodic signal exhibited by the source. It has a fundamental oscillation of 2567 days, in the limit of the detection range, with a harmonic (1284 days). It also has a clear linear trend of $C = 1.8 \times 10^{-4}$.

The next source in significance is J0739.2+0137 (Fig. 8.4, Bottom), exhibiting a 1743 days period with a p-value of 4.3×10^{-5}, considerably higher than the previous one. For both sources the correlation time is close to the monthly data cadence, implying a lower impact of the stochastic component in the global fit.

Most of the remaining sources in this group have larger correlation times. Regarding the models, the harmonic behavior is found in almost half of the sources analyzed, and the linear trend is in most of them.

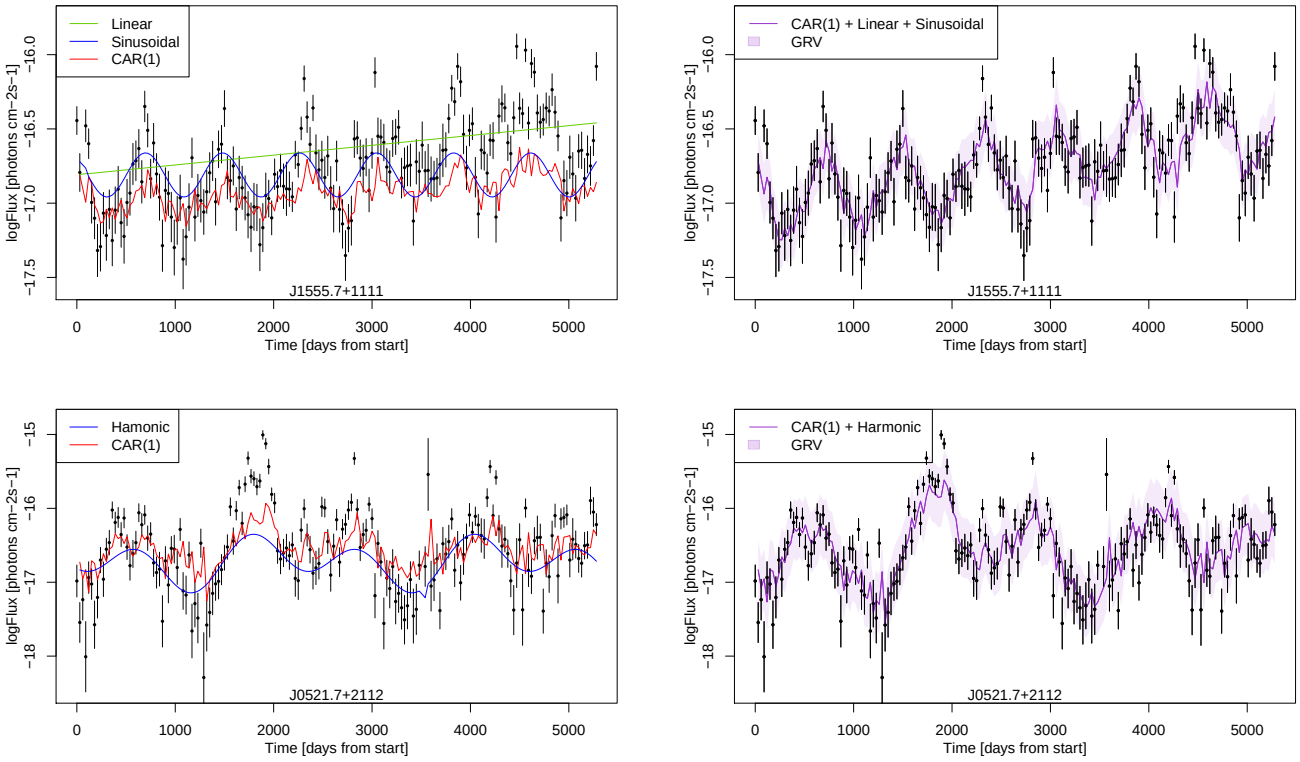

Figure 8.5. MCMC fit for the two sources with the most significant periodic signals in Table 8.2.

Within the literature candidates, 6 of them are also detected in the Chapter 7 sample, including the 5 most significant sources in Table 7.2. The strongest evidence of periodic behavior is observed for J1555.7+1111 / PG 1553+113 (Fig. 8.5, Top), with a p-value of 9.8×10^{-9}. The period is 783 days, and the slope is 1.0×10^{-4}. Again, this value is in agreement with the results in literature. These results are also similar to those in the previous chapter (Tab. 7.2), where the analysis was limited to ~ 4300 days from the start, indicating a considerable time stability of the retrieved parameters. As the number of data points has increased, and the fit to the data is still good, the significance of this new analysis is even higher. The PKS 2155-304 results are discussed in the periodicity stability section.

With a p-value of 1.5×10^{-6}, the next source is J0521.7+2112 (Fig. 8.5, Bottom),

showing a harmonic behavior of 2266 days as a primary oscillation. If we compare it with the previous chapter, this source was not found to be harmonic, but the period found of 1136 days is compatible with the harmonic component of 1133 days. The literature value of 1022 days (Peñil et al., 2020) is also compatible with the harmonic component. This comparison is displayed in Figure 8.6. We see that the first model has a higher impact of correlated noise. It consist in a AR(2) model, i.e., it has memory from the past two values, with parameters $\beta_1 = 0.375$ and $\beta_2 = 0.235$. On the other hand, the CAR(1) model has a reduced noise term with a correlation time $\tau = 32$. In the harmonic oscillation, the amplitude for the secondary term is slightly dumped, making it not far from the sinusoidal model. Nonetheless, including the secondary oscillation improves the significance of the periodic model by between 3 and 4 orders of magnitude in the p-value.

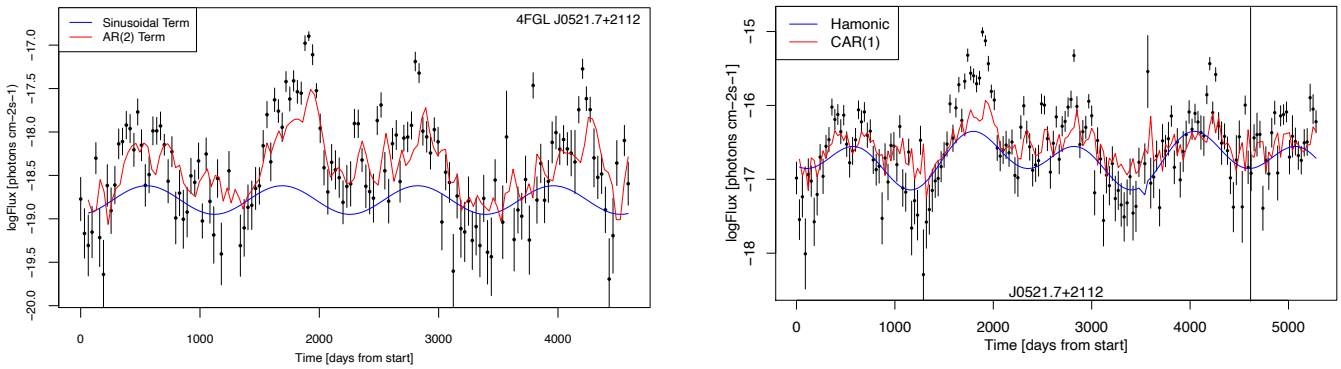

Figure 8.6. Comparison of the Chapter 7 and 8 fits for J0521.7+2112. The vertical line in the LCR plot (right) indicates the last time value in the Fermi-LAT sample data (left).

Of the remaining sources, J1048.4+7143 shows a periodic model compatible with that found in Wang et al. (2022). As in the previous chapter, J1903.2+5540 has a periodicity of 1131 days, which is not in agreement with Peñil et al. (2020) where a low significance ($> 2.5\sigma$) 1387 days is found. For J0457.0-2324, we detect a harmonic oscillation of 2561 days. The secondary harmonic component is compatible with the low significance sinusoid of 1300 days in the previous chapter. None of these results agree with the results in Peñil et al. (2020) of 949 days. For J0407.0-3826, we find the same periodicity of $\sim$ 1066 days as in Gong et al. (2022). Finally, in both chapters, we find a harmonic oscillation of $\sim$ 830 days for J0303.4-2407. In Peñil et al. (2020), a lower periodicity value of 730 days is found, whilst in Zhang et al. (2017c), a compatible value of 766 ± 109 days is found.

Spectral cross-check using Agatha

We perform a validation of our periodicity results through a spectral analysis using the Agatha software. For all sources in Tables 8.1 and 8.2, the best noise model retrieved from Agatha is parallel to our results, i.e., correlated noise of order 1. Thus, we use a MA(1) model to compute the BFPs. The results of the Agatha analysis can be found in Table 8.3 and 8.4. As an example, in Figures 8.7 and 8.8, we show the BFPs of the sources represented in Figures 8.4 and 8.5, respectively.

We see an excellent agreement between the periods retrieved with both different methods. The Agatha periods are compatible with the MCMC results for all thirteen sources with a simple sinusoidal oscillation. For the nine sources with a harmonic oscillation,

Table 8.3. Agatha cross-check comparison of the new periodic AGNs from the Fermi-LAT LCR. For each source, the list indicates: the AGN class; the number of points in the LC; the best model in terms of AIC; the MCMC period mean and standard deviation in days; the p-value computed from ΔAIC; the Agatha period mean and standard deviation; the Agatha lnBF. The results are sorted by MCMC period detection significance.

			CAR(1) Periodic Model			Agatha	
4FGL Name	Class	np	Best Model	Period	*pv*	Period	lnBF
J0509.4+0542	bll	175	h	$2567_{\pm 78}$	6.3×10^{-8}	$2639_{\pm 257}$	6.8
J0739.2+0137	fsrq	170	l,s	$1743_{\pm 87}$	4.3×10^{-5}	$1753_{\pm 169}$	7.4
J1640.4+3945	fsrq	120	l, s	$2058_{\pm 88}$	5.4×10^{-5}	$2071_{\pm 148}$	3.3
J1740.5+5211	fsrq	145	l, s	$2100_{\pm 99}$	7.4×10^{-5}	$2152_{\pm 176}$	4.8
J1913.0-8009	fsrq	120	h	$2538_{\pm 119}$	9.2×10^{-5}	$1379_{\pm 31}$	-1.1
J0141.4-0928	bll	167	l, h	$2279_{\pm 119}$	1.4×10^{-4}	$2132_{\pm 196}$	2.6
J2311.0+3425	fsrq	152	l, s	$1393_{\pm 93}$	1.8×10^{-4}	$1387_{\pm 75}$	5.3
J2243.9+2021	bll	168	s	$1965_{\pm 122}$	6.7×10^{-4}	$1981_{\pm 151}$	4
J0921.6+6216	fsrq	160	l,h	$2607_{\pm 188}$	9.3×10^{-4}	$2671_{\pm 239}$	1.2
J1754.2+3212	bll	137	l, s	$2061_{\pm 176}$	1.2×10^{-3}	$2057_{\pm 202}$	5.4
J0228.3-5547	fsrq	125	h	$1841_{\pm 114}$	1.2×10^{-3}	$914_{\pm 24}$	0.2
J0850.0+4855	bll	123	l, s	$2482_{\pm 243}$	1.3×10^{-3}	$2530_{\pm 304}$	2
J0030.3-4224	fsrq	130	l, s	$1791_{\pm 197}$	6.3×10^{-3}	$1835_{\pm 91}$	0.1
J0217.8+0144	fsrq	150	s	$1189_{\pm 50}$	7.5×10^{-3}	$1224_{\pm 30}$	1.7

Table 8.4. Agatha cross-check comparison of the known in literature periodic AGNs from the Fermi-LAT LCR.

			CAR(1) Periodic Model			Agatha	
4FGL Name	Class	np	Best Model	Period	*pv*	Period	lnBF
J1555.7+1111	bll	176	l, s	$783_{\pm 8}$	9.8×10^{-9}	$780_{\pm 30}$	17.2
J0521.7+2112	bll	176	h	$2266_{\pm 54}$	1.5×10^{-6}	$1143_{\pm 64}$	5.5
J1903.2+5540	bll	175	l, s	$1131_{\pm 33}$	1.1×10^{-4}	$1156_{\pm 52}$	10.9
J1048.4+7143	fsrq	167	s	$1103_{\pm 37}$	1.5×10^{-4}	$1144_{\pm 60}$	7.3
J0457.0-2324	fsrq	174	l, h	$2561_{\pm 110}$	3.4×10^{-4}	$1327_{\pm 38}$	-1.3
J0407.0-3826	fsrq	143	l, s	$1066_{\pm 163}$	2.6×10^{-3}	$954_{\pm 29}$	0.7
J2158.8-3013	bll	173	l, s	$601_{\pm 33}$	2.7×10^{-3}	$619_{\pm 22}$	-1.9
J0303.4-2407	bll	171	l, h	$835_{\pm 29}$	2.9×10^{-3}	$789_{\pm 48}$	1.2

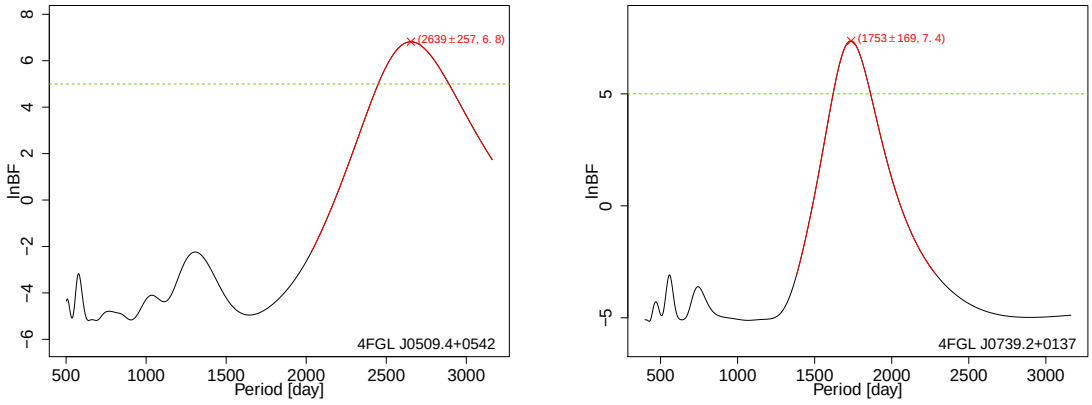

Figure 8.7. Agatha results for the two sources with the most significant periodic signals in Table 8.3. The green dashed line indicates the 5 lnBF significance as prescribed by Feng et al. (2017).

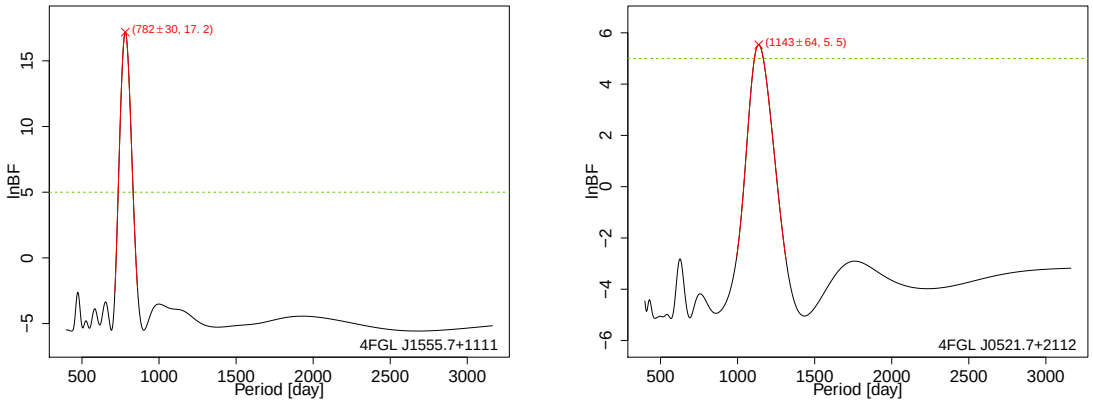

Figure 8.8. Agatha results for the two sources with the most significant periodic signals in Table 8.4

the Agatha period is compatible with the MCMC results for five of them. Another four sources show as a primary peak in Agatha the period corresponding with the harmonic component, i.e., half of the period retrieved in the MCMC sampling.

All sources but two show a positive lnBF, where a periodic model is favored over the noise model. Only J1913.0-8009 and J0457.0-2324 have a negative value. Both sources have a harmonic behavior where the power of the periodogram peaks may be shared between the primary and the secondary period. Nevertheless, it is a valuable tool to check if the primary peak in the periodogram corresponds to the period found with our MCMC.

8.2.2 Sources without correlated noise

For 13 sources, we find a newly detected periodicity corresponding to a model with well-converged deterministic components but without a CAR(1) correlation time τ parameter. In this case, the posterior MCMC sampling of the parameters draws an approximately symmetric Gaussian but is stuck at zero. Also, all the τ values are below the monthly data binning. An example of τ bad posterior distributions is presented in Figure 8.9. This means that the best stochastic description is with a WN model. In Figure 8.3, we exhibit an example of the MCMC posterior distributions retrieved for the WN and period parameters. The results are shown in Table 8.5.

We also remark that the number of data points of the WN sources is lower in mean than the CAR(1) results, meaning that these results are likely linked to fainter sources. A possible explanation for this correlation could be the following: As the sources are fainter than most CAR(1) results, several flux points are below the threshold of $TS \simeq 4$, treated as ULs. The remaining data flux points in the LC have larger error bars that could erase the correlation between flux points. Thus, as retrieving the correlated stochastic behavior is not possible, the periodicity results may not be reliable.

From Figures 8.9 and 8.10, we see that the stochastic model used does not considerably affect the period parameter's sampling. As mentioned, we also notice that the lower limit on the τ parameter is stuck at zero. Regarding model selection, for most sources, the AIC value is lower for the WN model than for the CAR(1) model, meaning that the simplest one is preferred.

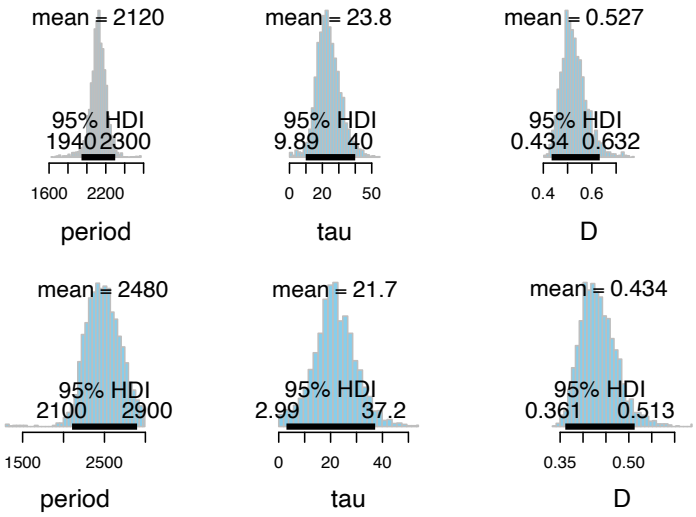

Figure 8.9. MCMC bad τ posterior distributions for the sources with the most significant periodic signals in Table 8.5. [*Top panel*] J1131.0+3815. [*Bottom panel*] J0807.1-0541.

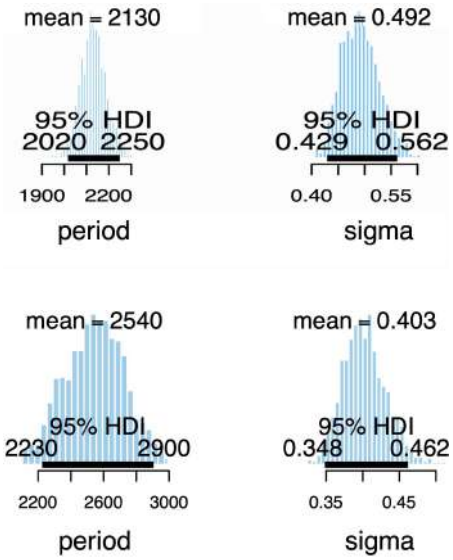

Figure 8.10. MCMC WN posterior distributions for the sources with the most significant periodic signals in Table 8.5. [*Top panel*] J1131.0+3815. [*Bottom panel*] J0807.1-0541.

We see high significances of the periodic model, with p-values below = 0.0005, denoting solid evidence of the source's periodic behaviour in comparison with only WN model. The p_{value} of the normality AD test on the residuals is above 0.05 for all the sources listed but for J0009.3+5030, which is discussed in Section 8.2.3.

By analyzing the residuals Gaussian distributions, we check that the residuals are centered close to 0, expected from well-fitted data. Also, there is a good agreement between the MCMC WN parameter σ, and the residuals standard deviation values N_σ.

Regarding the models, more than half of the sources present a linear term, and only the most significant candidate is described with a harmonic oscillation, with a $p_{value} = 9.2 \times 10^{-9}$ for a 2131 days periodicity. Figure 8.11 shows the MCMC fit for the four most

Table 8.5. MCMC fit results of the periodic AGNs without correlated noise from the Fermi-LAT LCR.

			WN Periodic Model				WN model	Residuals	
4FGL Name	Class	np	Best Model	Period	σ	pv	σ_{noise}	pv_{AD}	N_σ
J1131.0+3815	fsrq	119	l, h	$2131_{\pm 54}$	0.49	9.2×10^{-9}	0.58	0.91	0.47
J0807.1-0541	bll	108	l,s	$2544_{\pm 178}$	0.40	2.0×10^{-8}	0.47	0.13	0.39
J1107.0-4449	fsrq	112	s	$1353_{\pm 43}$	0.4	2.1×10^{-8}	0.47	0.95	0.39
J2201.5-8339	fsrq	104	s	$1782_{\pm 62}$	0.57	7.5×10^{-8}	0.65	0.25	0.55
J0902.4+2051	bll	120	s	$1648_{\pm 46}$	0.42	2.6×10^{-7}	0.47	0.06	0.41
J0143.1-3622	bcu	67	s	$2153_{\pm 111}$	0.58	3.4×10^{-7}	0.70	0.25	0.55
J1259.1-2311	bll	86	l, s	$2105_{\pm 109}$	0.38	9.4×10^{-7}	0.44	0.08	0.36
J0222.0-1616	fsrq	70	l, s	$1706_{\pm 69}$	0.41	1.5×10^{-6}	0.50	0.52	0.4
J0009.3+5030	bll	155	l, s	$1528_{\pm 64}$	0.46	1.6×10^{-6}	0.50	0.01	0.45
J1917.7-1921	bll	154	s	$1039_{\pm 26}$	0.42	9.0×10^{-6}	0.44	0.78	0.41
J0230.8+4032	fsrq	91	s	$1528_{\pm 62}$	0.63	7.5×10^{-5}	0.68	0.27	0.61
J1844.4+1547	bll	115	l, s	$1432_{\pm 66}$	0.51	8.9×10^{-5}	0.55	0.20	0.5
J1234.0-5735	bcu	115	l, s	$1026_{\pm 31}$	0.36	2.2×10^{-4}	0.38	0.14	0.35

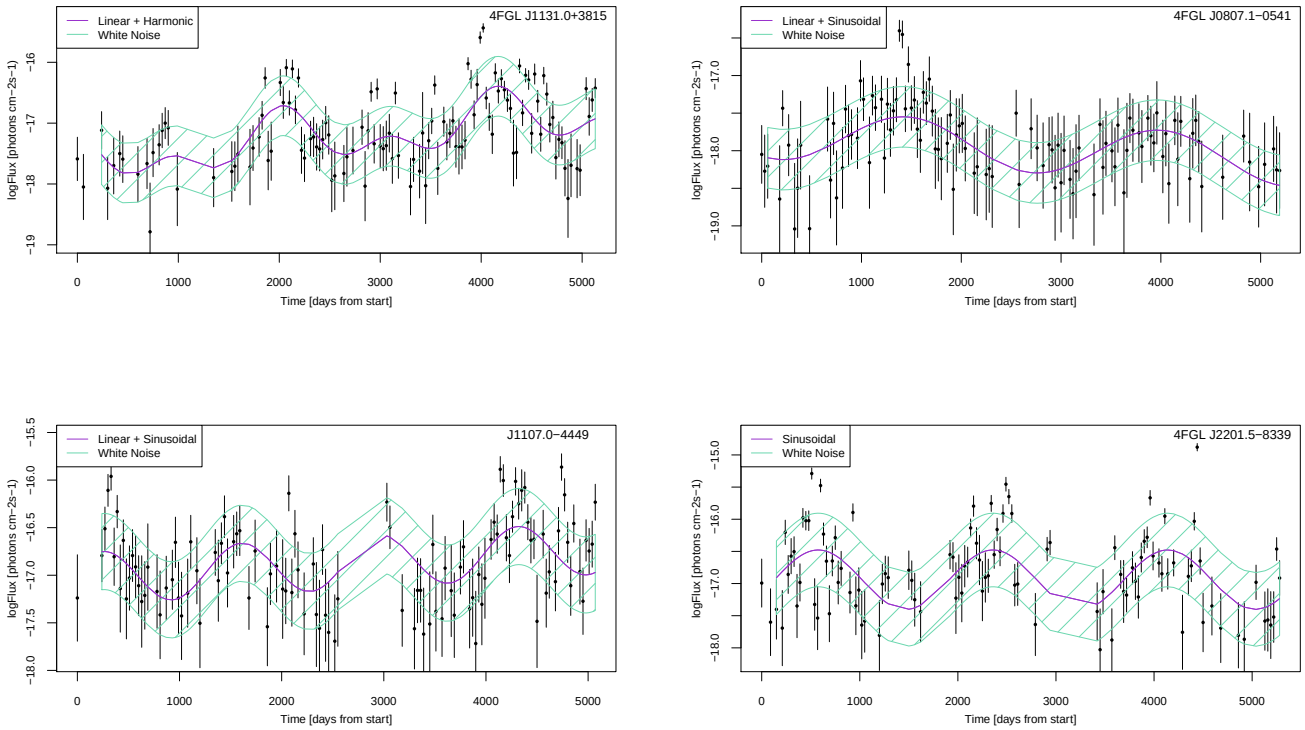

Figure 8.11. MCMC fit for the four sources with the most significant periodic signals in Table 8.5. General fit and WN σ component.

significant sources in the results table.

Spectral cross-check using Agatha

We repeat the Agatha software cross-validation of the periodicities and noise results. Again, the best noise model retrieved from Agatha for the sources in this section is the same as in our MCMC fit, i.e., WN. Thus, we use a WN model to compute the BFPs.

The results of the Agatha analysis can be found in Table 8.6. In Figure 8.12, we show the BFPs of the four most significant sources in the Table 8.5.

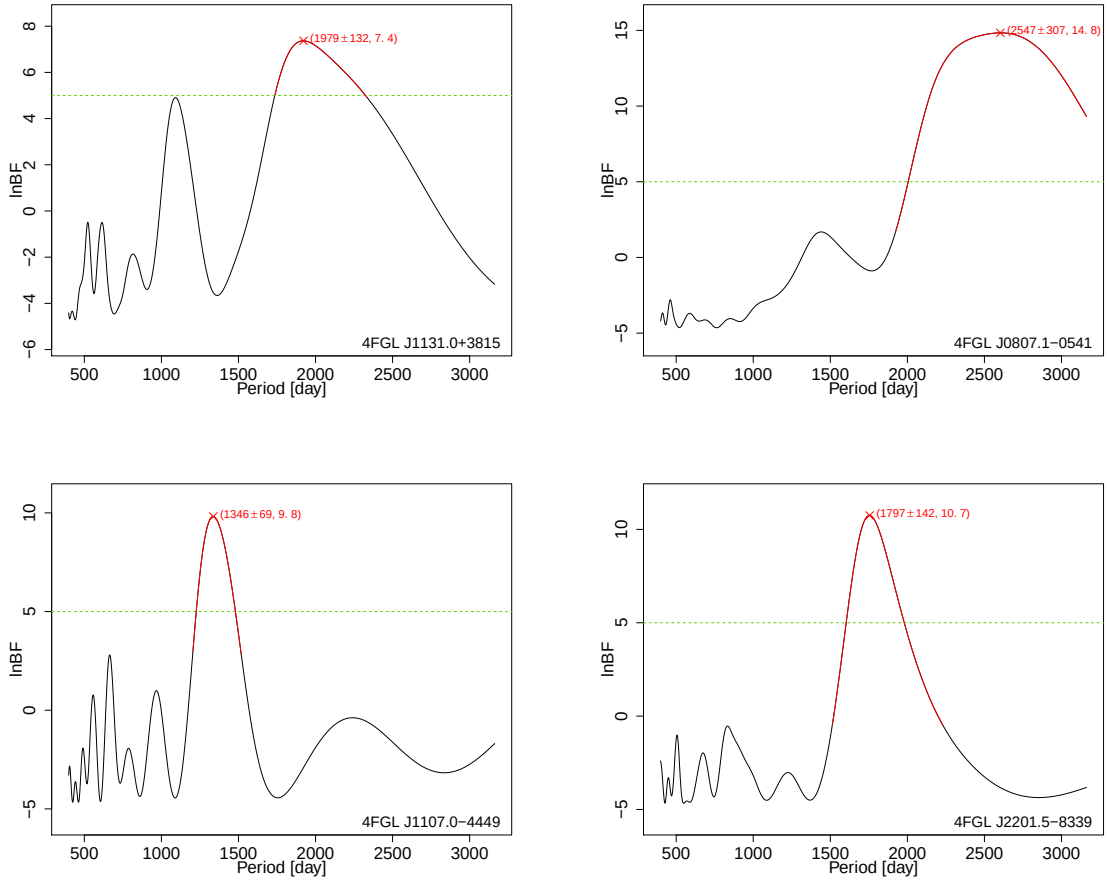

Figure 8.12. Agatha results for the four sources with the most significant periodic signals in Table 8.6.

Table 8.6. Agatha cross-check comparison of the periodic AGNs without correlated noise from the Fermi-LAT LCR.

			WN Periodic Model			Agatha	
4FGL Name	Class	np	Best Model	Period	*pv*	Period	lnBF
J1131.0+3815	fsrq	119	l, h	$2131_{\pm 54}$	9.2×10^{-9}	$1979_{\pm 132}$	7.4
J0807.1-0541	bll	108	l, s	$2544_{\pm 178}$	2.0×10^{-8}	$2547_{\pm 307}$	14.8
J1107.0-4449	fsrq	112	s	$1353_{\pm 43}$	2.1×10^{-8}	$1346_{\pm 69}$	9.8
J2201.5-8339	fsrq	104	s	$1782_{\pm 62}$	7.5×10^{-8}	$1797_{\pm 142}$	10.7
J0902.4+2051	bll	120	s	$1648_{\pm 46}$	2.6×10^{-7}	$1659_{\pm 110}$	9.1
J0143.1-3622	bcu	67	s	$2153_{\pm 111}$	3.4×10^{-7}	$2043_{\pm 116}$	4.1
J1259.1-2311	bll	86	l, s	$2105_{\pm 109}$	9.4×10^{-7}	$2134_{\pm 256}$	9.3
J0222.0-1616	fsrq	70	l, s	$1706_{\pm 69}$	1.5×10^{-6}	$1899_{\pm 370}$	8.1
J0009.3+5030	bll	155	l, s	$1528_{\pm 64}$	1.6×10^{-6}	$1513_{\pm 62}$	10.9
J1917.7-1921	bll	154	s	$1039_{\pm 26}$	9.0×10^{-6}	$1066_{\pm 48}$	4
J0230.8+4032	fsrq	91	s	$1528_{\pm 62}$	7.5×10^{-5}	$1523_{\pm 116}$	2.9
J1844.4+1547	bll	115	l, s	$1432_{\pm 66}$	8.9×10^{-5}	$1424_{\pm 79}$	5.6
J1234.0-5735	bcu	115	l, s	$1026_{\pm 31}$	2.2×10^{-4}	$1026_{\pm 56}$	2.7

As in the previous spectral cross-check analysis, there is a good agreement between the periods retrieved from the MCMC and Agatha. In contrast with the Agatha significance results of the CAR(1) periodic sources in Tables 8.3 and 8.4, here the significance of the periodicities are higher, as we are using only WN for the stochastic component. For 11 out of 13 sources the significance (lnBF) of the periodic model over the pure WN model is above 4, and above 2.5 for J0230.8+4032 and J1234.0-5735.

In the representation of the periodogram for J1131.0+3815 (Fig. 8.12, Top-Left Panel), we also notice the primary oscillation peak along with a secondary peak compatible with the presence of an harmonic component of $\sim$ 1050 days. In the Top-Right Panel, we see the high uncertainty of the Agatha period, also important for the MCMC result.

8.2.3 Systematics of the MCMC search

Tests for normality

As suggested in Section 6.3.2, we have tested the results of the LC fits for normality by performing an AD test on the residuals. We also fitted a Gaussian distribution $N(\text{mean}, \sigma)$. The results are presented and discussed in the previous sections. Here, we just present an example of the MCMC residuals in Figure 8.13. We note the presence of some high and low values in the residuals corresponding to some outliers in the MCMC fit.

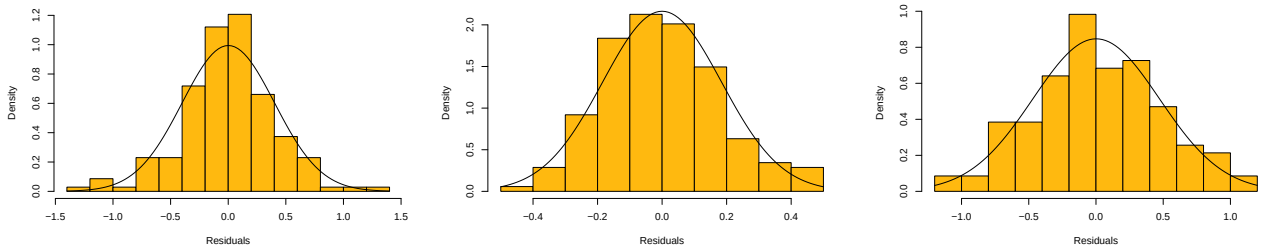

Figure 8.13. Residuals histogram for the sources with the most significant periodic signals in Tables 8.1, 8.2 and 8.5. Black line shows a Gaussian distribution fit with N_{mean} and N_σ parameters. [*Left panel*] J0509.4+0542 [*Center panel*] J1555.7+1111 [*Right panel*] J1131.0+3815.

Correlation between parameters

For every source analysed, the correlation between MCMC parameters is not as important as to affect the efficiency of the sampling chains. There is an expected correlation between the CAR(1) model parameters τ and D (see Equation 6.32). In the Figure 8.14 the corner plots for different results are shown, useful to visualize the pairwise correlations between model parameters.

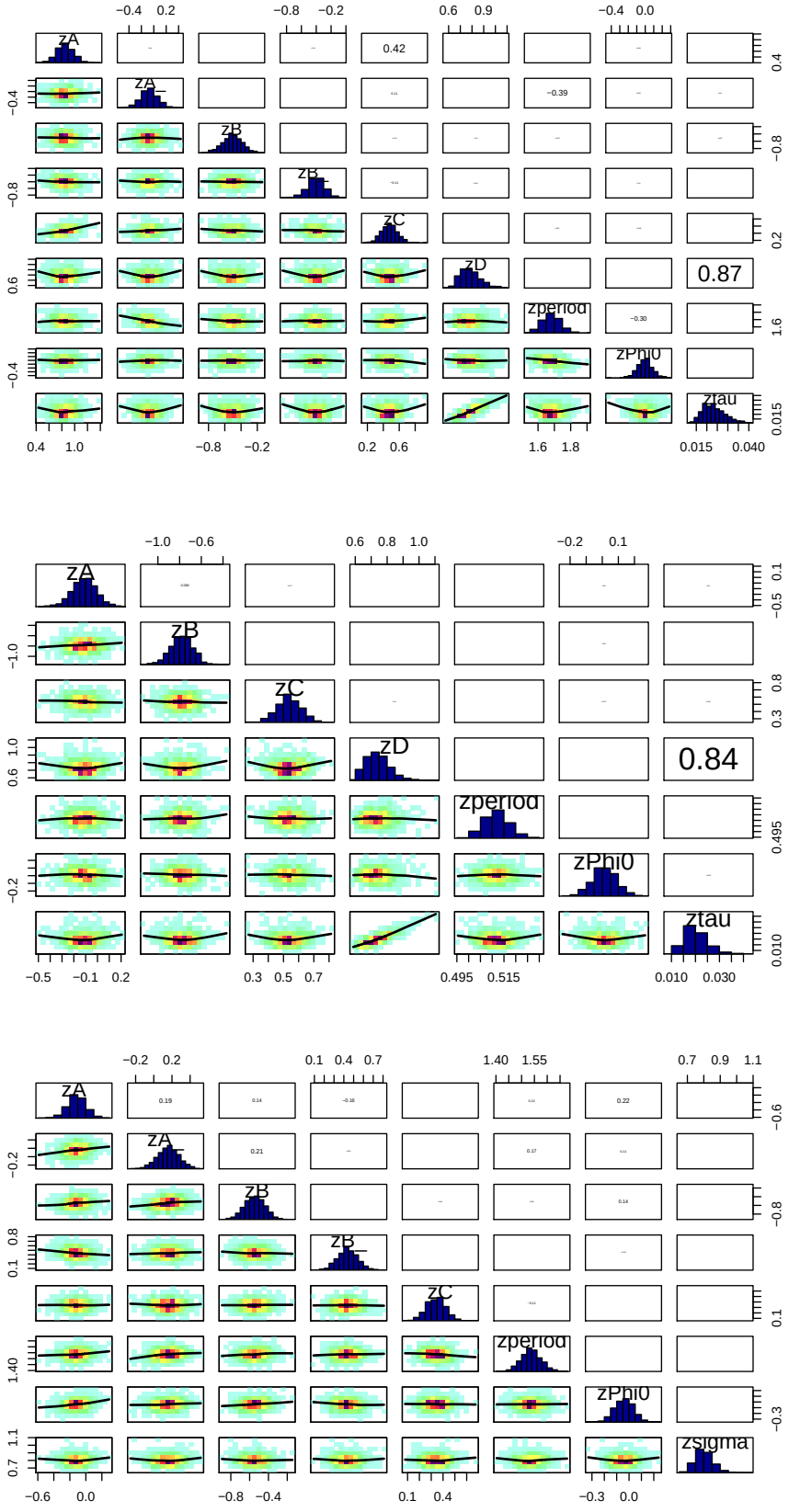

Figure 8.14. Corner plot for the sources with the most significant periodic signals in Tables 8.1, 8.2 and 8.5. [*Top panel*] J0509.4+0542. [*Center panel*] J1555.7+1111. [*Bottom panel*] J1131.0+3815. The diagonal plots show the posterior distributions for every standardize parameter (marked as z). The 2D maps below the diagonal show the density of the posterior distribution with one parameter in each axis. The values above the diagonal show the Pearson correlation coefficients between parameters. See Section 6.2.1 for a parameters' description.

8.3 Periodicity temporal stability

We follow the procedure used in the temporal stability analysis of the Fermi-LAT regularly sampled sources in the previous chapter. Using time windows with a width of 0.4 the full LC time span, we have slices of 0.4×5250 days $\sim$ 2100 days. We analyze the same sources as in Section 7.3: PG 1553+113 and PKS 2155-304. Both sources' periods are below 1050 days, allowing us to fit at least two oscillations in the windows time span. Five time windows are computed, each centered in 1050, 1838, 2625, 3412, and 4200 days from the start (MJD -54687), respectively. The MCMC fits are applied at each time window for every source included. The results are shown in Table 8.7.

Table 8.7. MCMC fit results of the Fermi-LAT LCR time windows analysis. For each source and window, the list indicates: the best model in terms of AIC; the period mean and standard deviation in days; the CAR(1) τ mean in days; the CAR(1) D mean; the p-value computed from ΔAIC; the only noise model τ and D; the test for normality in residuals p-value from the AD test; the residuals Gaussian distribution fit standard deviation N_σ.

		CAR(1) Periodic Model					CAR(1) model		Residuals	
4FGL Name		Model	Period	τ	D	pv	τ_{noise}	D_{noise}	pv_{AD}	N_σ
J1555.7+1111	1	s	$762_{\pm 28}$	19	0.17	3.2×10^{-6}	70	0.34	0.43	0.15
	2	l, s	$836_{\pm 85}$	43	0.25	1.4×10^{-4}	83	0.39	0.49	0.17
	3	l, s	$809_{\pm 89}$	51	0.26	5.9×10^{-4}	88	0.38	0.05	0.17
	4	l, s	$737_{\pm 105}$	40	0.26	2.6×10^{-3}	65	0.35	0.77	0.19
	5	l, s	$799_{\pm 112}$	56	0.34	2.1×10^{-3}	84	0.45	0.90	0.21
J2158.8-3013	1	l, s	$648_{\pm 44}$	47	0.35	2.0×10^{-4}	87	0.55	0.57	0.23
	2	l, s	$580_{\pm 61}$	50	0.44	9.5×10^{-4}	75	0.58	0.03	0.28
	3	l, s	$580_{\pm 60}$	58	0.55	4.5×10^{-3}	82	0.70	0.10	0.33
	4	l, s	$355_{\pm 20}$	83	0.66	7.0×10^{-2}	89	0.72	0.29	0.33
	5	l, s	$384_{\pm 38}$	59	0.46	2.3×10^{-3}	78	0.59	0.45	0.28

For PG 1553+113, a periodic component with a p-value below 0.005 is found in every time window, represented in Figure 8.15. Windows 1 and 4 present a period below the entire LC period of 783 days. For Windows 3 and 5, the period is $\sim$ 800 days, above the 783 days period. It is in Window 2, where the oscillation has a higher value, with a difference of $\sim$ 100 days from the lowest value in Window 4. In the first Window, the best model lacks the linear term, which appears in all remaining slices. Also, the CAR(1) noise model for this window is closer to a WN behavior, as the τ parameter has the lowest value. It is also remarkable that the later windows have higher period standard deviation and lower significances than the others. As in Section 7.3, we perform a constant fit through a χ^2 test. For this set of periods, the p-value of the test is 0.9, indicating that the values are compatible with a constant distribution with a period of 772 ± 12.

On the other hand, PKS 2155-304 shows a particular behavior, as seen in Figure 8.16. The global fit period of 601 ± 33 is compatible with the three first windows that present p-values below 0.005, with the highest periodic value in Window 1 of 648 days. However, from the date 57050 MJD, the periodic oscillation is reduced by almost half. In Window 4, we find a low-significance periodicity of 355 days, where the emission is dominated by the correlated noise ($\tau = 83$). Then, in the last window, we detect a more robust oscillation of 384 days, suggesting a change in the physical emission mechanisms in the latest years. Consequently, the constant fit gives a p-value close to 0, indicating a variable periodic distribution in time.

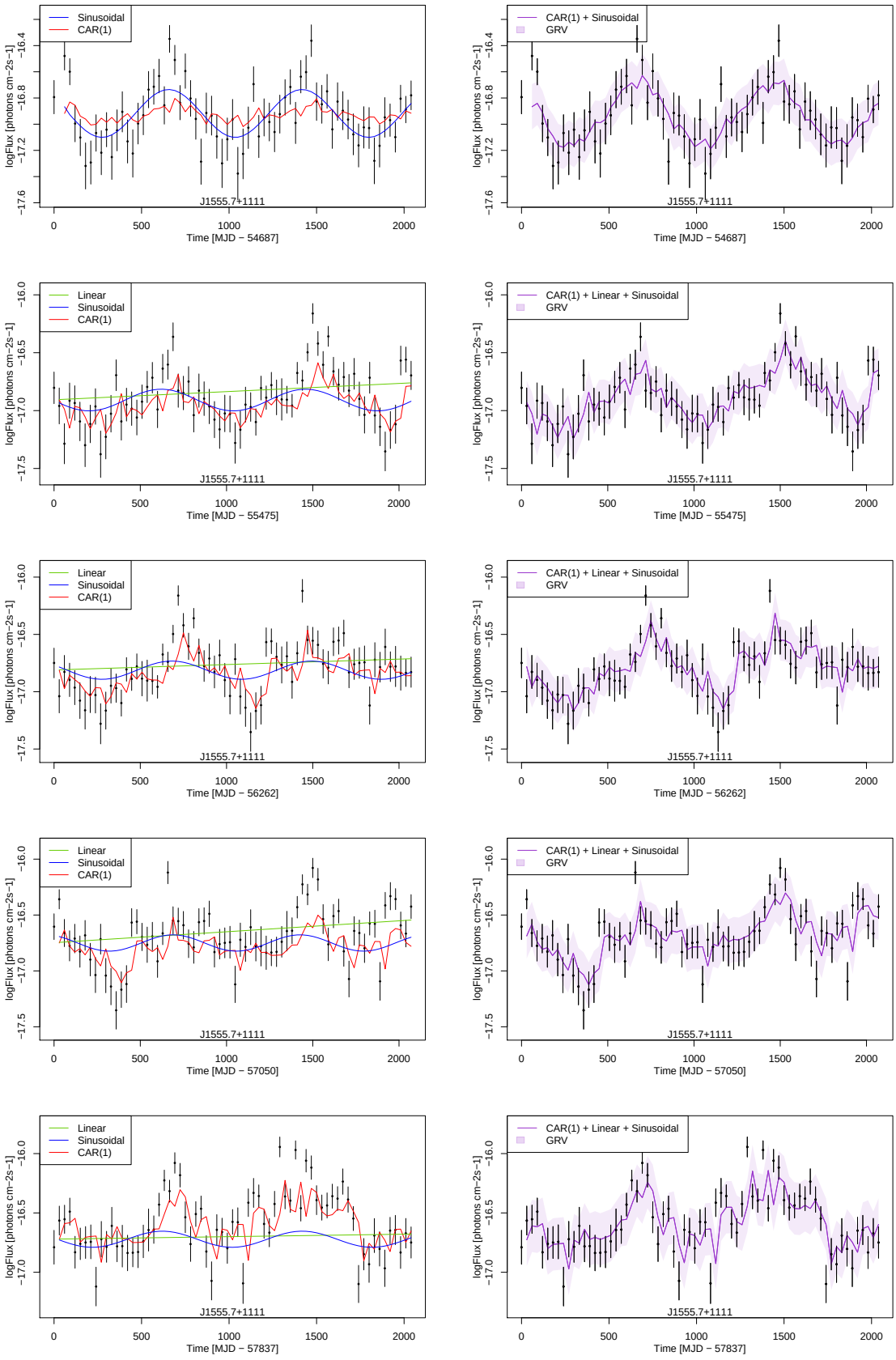

Figure 8.15. MCMC components and global fit for the five time windows analysis of J1555.7+1111.

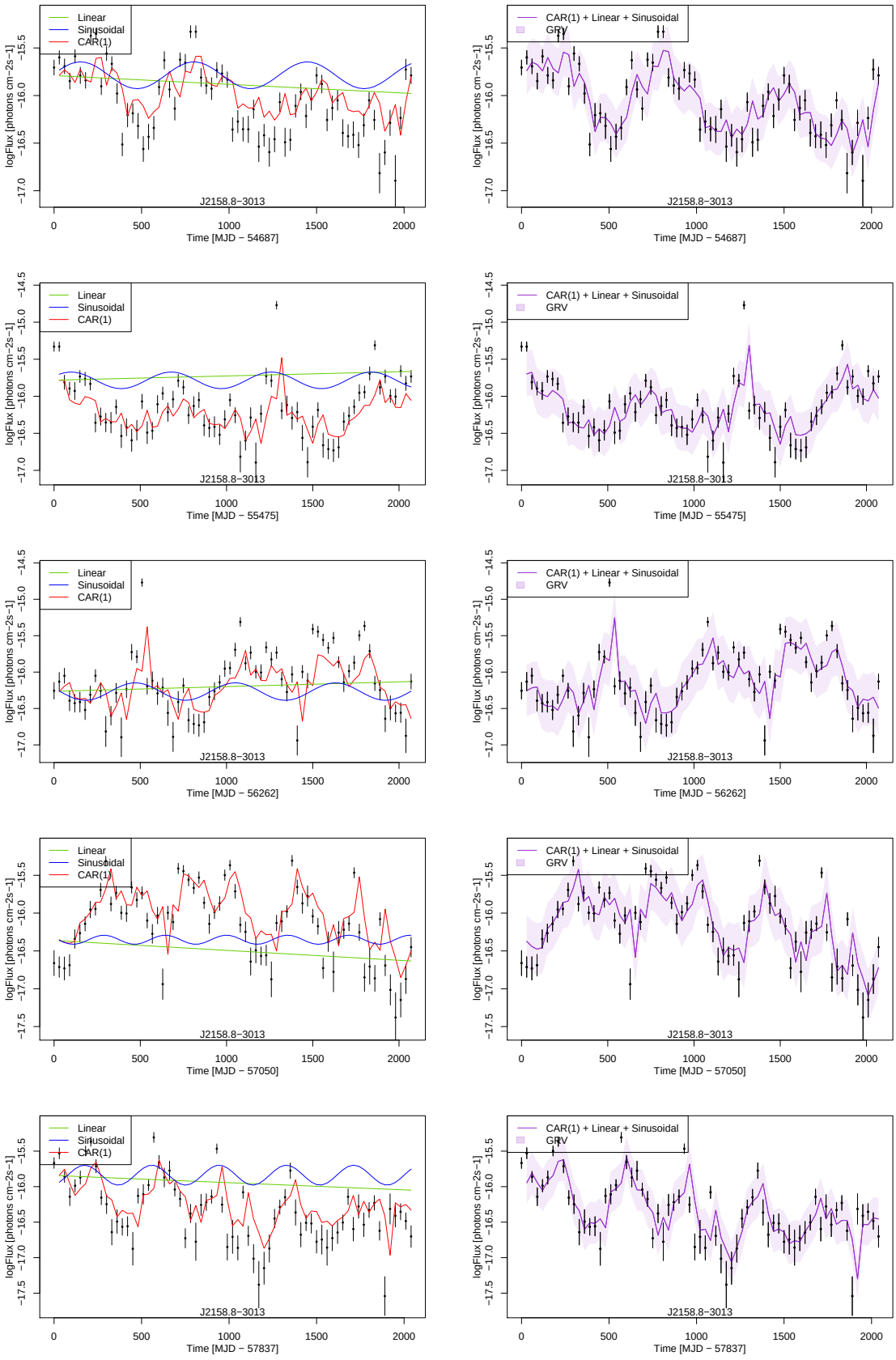

Figure 8.16. MCMC components and global fit for the five time windows analysis of J2158.8+3013.

In Table 8.8 we present the results of the spectral cross-check using the Agatha moving periodogram, as explained in Section 6.4. Figure 8.17 shows the temporal variation of the sources period for MCMC and Agatha analysis. When comparing Tables 7.10, 8.17 and Figures 7.10 and 8.17 we remark that the time span is different for previous and current analysis (4620 and 5250 days, respectively). Thus, the time value of the last window in Figure 7.10, is closer to the 4th window in Figure 8.17. It is important to note that the result for the last window in the previous chapter shows an harmonic component which can be compatible with the results of the drastic decrease of the period in the last two windows for the current chapter.

Table 8.8. Agatha cross-check comparison of the Fermi-LAT LCR time windows analysis. For each source and window, the list indicates: the best model in terms of AIC; the MCMC period mean and standard deviation in days; the p-value computed from ΔAIC; the Agatha period mean and standard deviation; the Agatha lnBF.

4FGL Name		CAR(1) Periodic Model			Agatha	
		Model	Period	*pv*	Period	lnBF
J1555.7+1111	1	s	$762_{\pm 28}$	3.2×10^{-6}	$753_{\pm 72}$	2.8
	2	l, s	$836_{\pm 85}$	1.4×10^{-4}	$863_{\pm 108}$	2.4
	3	l, s	$809_{\pm 89}$	5.9×10^{-4}	$831_{\pm 78}$	1.8
	4	l, s	$737_{\pm 105}$	2.6×10^{-3}	$736_{\pm 61}$	1.7
	5	l, s	$799_{\pm 112}$	2.1×10^{-3}	$799_{\pm 82}$	3.6
J2158.8-3013	1	l, s	$648_{\pm 44}$	2.0×10^{-4}	$689_{\pm 88}$	6.4
	2	l, s	$580_{\pm 61}$	9.5×10^{-4}	$617_{\pm 25}$	3.4
	3	l, s	$580_{\pm 60}$	4.5×10^{-3}	$585_{\pm 28}$	-1.2
	4	l, s	$355_{\pm 20}$	7.0×10^{-2}	$368_{\pm 10}$	-1.7
	5	l, s	$384_{\pm 38}$	2.3×10^{-3}	$390_{\pm 19}$	0.7

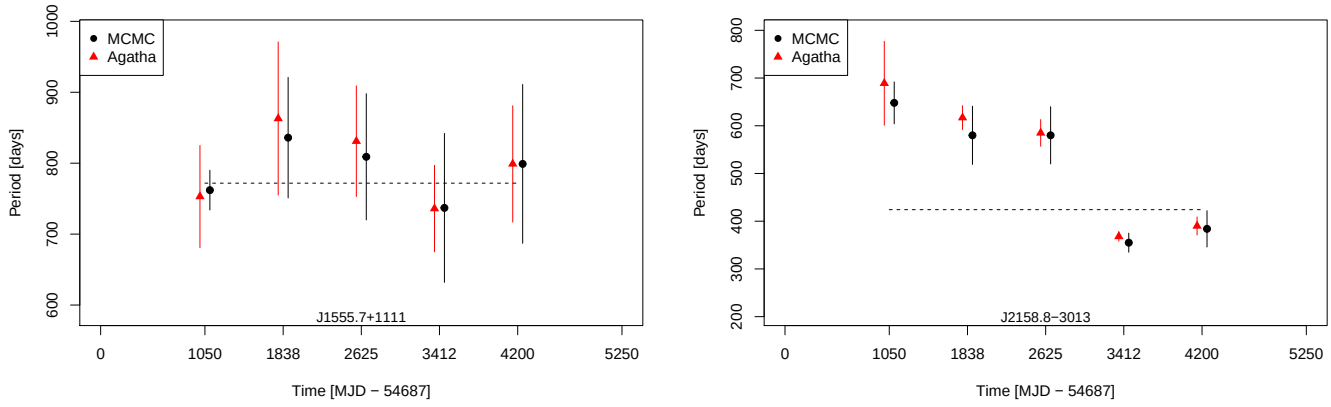

Figure 8.17. MCMC and Agatha period mean and standard deviation for each time window. Horizontal dashed line indicates the result of the constant fit value.

For PG 1553+113, we see a good agreement between Agatha values and the MCMC results with lnBF values between 1.7 to 3.6. For PKS 2155-304, the first two windows are well computed with higher significances, while Windows 3 and 4 present similar periodicities negative lnBF values. The spectral method also obtains the drop in periodicity for the two last windows. From Figure 8.17 we see that the period variations are compatible between the two methods employed.

Finally, following the procedure in the previous chapter, in Figure 8.18 we show the variation of the oscillation amplitude as a function of the period. The constant χ^2 test

is computed to quantify the amplitude variation with the period, giving a p-value of 0.92 for both PG 1553+113 and PKS 2155-304. Again, the amplitude of the oscillation is independent of its width.

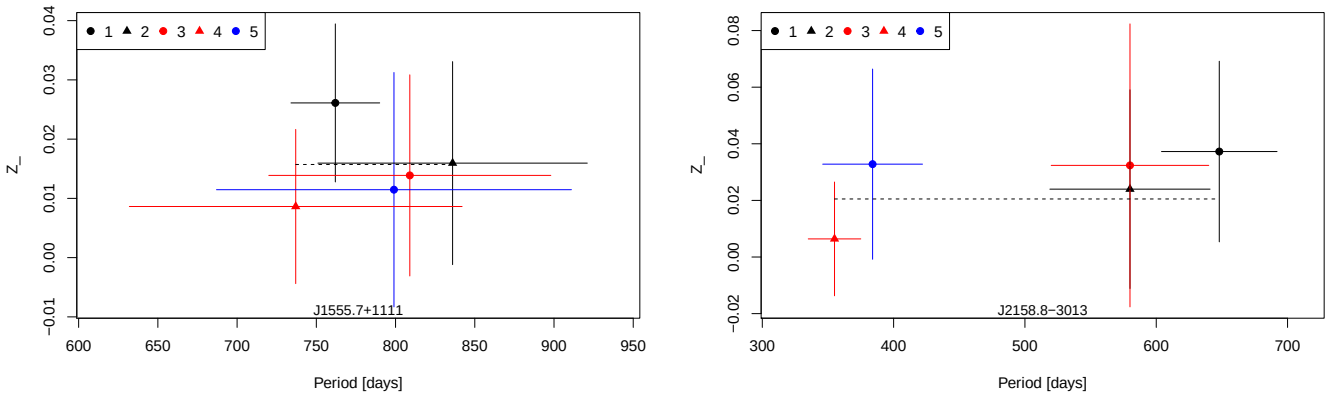

Figure 8.18. MCMC physical amplitude Z mean and standard deviation for each time-window period. Horizontal dashed line indicates the result of the constant fit value.

8.4 Discussion

In this chapter we have extended the previous formalism for period searches in γ-ray AGN LCs (Chapter 7). The new extension involves the inclusion of a novel time-series approach using a CAR-based method capable of analyzing irregularly sampled data. As demonstrated, this method effectively separates the stochastic and periodic components during the LC fitting process.

We applied the new algorithm to the entire AGN population of the Fermi-LAT LCR, including both regularly and irregularly sampled data, resulting in a significant increase in the number of available sources. By using a CAR(1) stochastic model for the noise component, we were able to detect several new periodic sources and confirm the presence of oscillations in others. We also present a list of new periodic sources that show no sign of noise correlation and are modeled only as WN. To validate our results, we cross-checked them with the Agatha software spectral method, which yielded similar stochastic noise models and periodic signals for the majority of the sources analyzed.

We can compare the periodicity results obtained in this chapter using CAR models with the results of the previous chapter using AR models for regularly sampled data (Tables 7.2 and 8.2). For AGNs that appear in both the AR and CAR results, the periodic signals found are consistent. However, it should be noted that for most of the sources marked with * or ** in the table 7.2 of the previous chapter (indicating some systematic problem in the MCMC chains), the periodic model is not recovered using the CAR formalism presented in this chapter. This can be attributed to the imposition of stricter constraints on the quality checks for MCMC convergence in this chapter, leading to the exclusion of sources where convergence was not achieved with a uniform prior.

By analyzing the number of data points for each source, we found that the already known periodic sources had almost regular sampling over time. Conversely, most of the newly detected sources showed irregular sampling with less than 160 data points out of a total of 176 time intervals. This suggests that the inclusion of models such as CAR or

CARMA, which are able to interpret the variability of irregularly sampled data, can lead to the detection of new sources.

In both AR/CAR analyses we have fitted the LCs with different deterministic models, including harmonic oscillations or linear trends. Gao et al. (2023) provide an explanation for the linear trend in PG 1553+113 that can be applied to all the linear models found in the AGN fluxes in this and the previous chapter. First, they explain the 2.18 ± 0.08 year QPO with a helical jet motion model (see section 5.1.1, example 3). They then suggest that the upward trend is due to long-term variability with a timescale of ~ 42 years (19 times the medium-term periodicity of ~ 2.2 years), reproducible with a small modification of the proton injection into the source. Both values for the periodic and linear terms are compatible with our results for PG 1553+113.

Again, as in the previous chapter, we examined the period and amplitude variations over time of PG 1553+113 and PKS 2155-304 by dividing the LC into different time windows. Similarly, J1555.7+1111 shows no significant variation in either period or amplitude over time. This result again disfavors the lighthouse effect models (Tavani et al., 2018; Ait Benkhali et al., 2020) and is consistent with the BSMBH hypothesis. In the case of J2158.8-3013, the period shows a drastic decrease for the second half of the observations. First the period decreases by ~ 70 days for windows 2 and 3, and then it decreases to ~ 350 days for the last two windows. This could be related to the ~ 315 days periodic oscillation in the optical (Zhang et al., 2014), which is almost half of the general fit for the source and the literature results in γ-rays. The increase in the amplitude of the harmonic component in the periodic emission could explain this behavior in the LC. A more detailed study of physical models compatible with this result needs to be carried out.

To continue this line of research, we propose some next steps to improve the understanding of periodic SMBH. First, we intend to use weekly and daily sampling in the LCR dataset. With this finer sampling resolution, we expect to obtain more detailed information about the patterns of AGN variability, and possibly reveal shorter-period oscillations that may not be apparent in coarser-sampled data, as in Ren et al. (2023). This may help to uncover faster oscillations and provide insights into the physical mechanisms underlying AGN variability. Another aspect to consider is the use of automatic AR/CAR models up to higher orders to capture more complex temporal dependencies and improve the accuracy of our periodicity analysis. To improve the analysis, we will also consider the information provided by the ULs in the LCs. By incorporating these bounds as constraints in the MCMC fitting process, we can refine the determination of periodic signals and improve the accuracy of our results.

Furthermore, we plan to investigate the correlation between the MCMC output parameters, such as the stochastic correlation times, the linearity, or the period values, with known physical variables of AGNs. Pei et al. (2022) present an effective method to estimate four important parameters in the physical framework of blazars, namely the UL of the central BH mass, the Doppler factor, the distance along the axis to the site of γ-ray production, and the propagation angle with respect to the axis of the accretion disk. These parameters alone are essential to profile the different physical models of periodic AGNs, but by studying these correlations we aim to gain a deeper understanding of the underlying physical processes driving the observed periodic behavior.

In this chapter we have applied our developed time-domain algorithm to detect periodicities in the Fermi-LAT LCs, including irregularly sampled data. This is used as a starting point for the next chapter, where we intend to apply similar techniques to understand

AGN variability in the VHE domain using H.E.S.S. data. CAR models are fundamental to this next step, as H.E.S.S. and other IACTs have LCs with different features due to their observational capabilities and performance.

Chapter 9

Applying the algorithm to H.E.S.S. data

This chapter focuses on the analysis of various sources observed by the H.E.S.S. telescope array. H.E.S.S. observations differ significantly from those conducted by the Fermi-LAT. Instead of continuous monitoring, H.E.S.S. observations are performed based on approved proposals or as part of Target of Opportunity (ToO) observations. Consequently, the data obtained from H.E.S.S. observations are irregularly sampled, introducing challenges in analyzing the LCs. Additionally, since H.E.S.S. is a ground-based observatory located in the southern hemisphere, the visibility of certain objects may vary throughout the year. As a result, the LCs data points are often grouped within a few months each year, resulting in significant flux gaps. These factors necessitate specific considerations and techniques for analyzing and interpreting the LC data obtained from H.E.S.S. observations.

Accurately characterizing the noise emission can be challenging when dealing with stochastic noise models for irregularly sampled LCs. However, employing a CAR model provides valuable insights into understanding and modeling the correlated noise component. By utilizing the CAR model, we can better understand the source's power spectrum, allowing us to compare our findings with existing literature that employs alternative methods. The irregular sampling of most sources also makes detecting periods particularly difficult. Furthermore, these sources often have a limited number of data points, making it challenging to analyze and assess the significance of the periodic behavior.

In Section 9.1, we start by modeling the noise behavior of a sample of sources from the H.E.S.S. Extragalactic Survey (HEGS), a catalog of 23 sources from the H.E.S.S. phase I observations. We present the MCMC fit results using the period-wise and night-wise LCs, and the PSD derived from the MCMC noise parameters. After, in Section 9.2, we make a joint periodicity search analysis using data from Fermi-LAT and H.E.S.S. for PKS 2155-304. First, we analyze the HEGS LC, and then, we use the information provided by the H.E.S.S. phase II observations to analyze a longer LC. Finally, we discuss the results and the physical implications in Section 9.4.

9.1 Modeling the stochastic behavior for the HEGS

9.1.1 HEGS sources

In 2018, the H.E.S.S. Galactic plane survey (HGPS, Abdalla, H. et al.) was presented, a survey of the Galactic plane observations in VHE γ-rays. The HGPS carries information on a decade-long data from the H.E.S.S. I array from 2004 to 2013 and contains 78 VHE sources, including PWNs, SNR, and γ-ray binaries.

During H.E.S.S. Phase I, an extragalactic observing programme was also carried out to detect and follow different types of sources, such as galaxy clusters and AGNs in the VHE range. The HEGS[1] is a project to re-analyze this extragalactic archival data coherently, as for the galactic survey. The goal is to detect within this framework the extragalactic sources which are already published by H.E.S.S. and to analyze in more depth the sources of particular interest, such as M87, CenA, or PKS 2155-304. It also aims to analyze a possible source of bias caused by the fact that the observations were conducted based on multi-wavelength flaring information, as bright X-ray or Fermi-LAT sources.

The HEGS catalogue is composed of 23 sources, listed in the Table 9.1, of which 22 are AGNs and NGC 253 is a starburst galaxy. In this chapter, we make use of the LCs provided from the analysis of these extragalactic sources.

Table 9.1

H.E.S.S. Name	Assoc	Type	redshift
J0013-189	SHBL J001355d9-185406	bll	0.095
J0152+017	RGB J0152+017	bll	0.08
J0232+202	1ES 0229+200	bll	0.1396
J0303-2411	PKS 0301-243	bll	0.2657
J0349-119	1ES 0347-121	bll	0.188
J0416+010	1ES 0414+009	bll	0.287
J0449-438	PKS 0447-439	bll	0.343
J0550-322	PKS 0548-322	bll	0.069
J1010-313	1RXS J101015d9-311909	bll	0.142639
J1103-234	1ES 1101-232	bll	0.186
J1315-426	1ES 1312-423	bll	0.105
J1444-391	PKS 1440-389	bll	0.1385
J1517-243	AP Librae	bll	0.049
J1555+111	PG 1553+113	bll	0.433
J2009-488	PKS 2005-489	bll	0.071
J2158-302	PKS 2155-304	bll	0.117
J2324-406	1ES 2322-40d9	bll	0.17359
J2359-306	H 2356-309	bll	0.165
J0627-354	PKS 0625-354	rg	0.05486
J1230+123	M 87	rg	0.0044
J1325-430	Centaurus A	rg	0.00183
J0047-253	PKS 1510-089	fsrq	0.361
J1512-090	NGC 253	Starburst	0.000864

[1] http://vietnam.in2p3.fr/2018/vhepu/transparencies/03_wednesday/01_morning/4_PS3/05_Bonnefoy.pdf

9.1.2 CAR(1) Modeling

Here use the MCMC algorithm developed to fit the noise components of the LCs. A linear term can also be obtained from the fit. Only the periodic term cannot be retrieved due to the reduced flux information for long-term variability. Given the irregularly sampled and grouped data, we use a CAR(1) model to obtain the correlated noise relaxation time τ, along with the GRV component. Most of the HEGS sources present LCs with insufficient data points to account for auto-regressive behavior. In that case, we present the σ value for the only WN model, with the linear term, if any.

Night-wise

First, in Table 9.2, we present the results on the night-wise LCs, where the sampling time between the data points at each observation group is on the order of one day, as it is integrated on a nightly basis. For PKS 2155-304, a known variable source, an exceptional flaring state was detected in July 2006, where the source reached a flux much above the average with minute time-scale variability (Aharonian et al., 2007). As we are also interested in analyzing the long-term behavior of the source, we remove the flare from the LC to better understand the correlation between all the other groups of data. Here, PKS 2155-304* indicates the analysis without the flare points.

Table 9.2. NIght-wise. MCMC fit results of the HEGS sources. For the CAR(1) model, the list indicates: the correlation time τ in days; the GRV parameter D; the linear term slope c in [logFlux day^{-1}] and the p-value computed from ΔAIC between the CAR(1) and the WN models. And for the WN model: the sigma value of the Gaussian distribution and the linear term slope c.

	Model $_{\mathrm{CAR(1)}}$				Model $_{\mathrm{WN}}$	
Name	τ	D	c	pv	σ	c
SHBL J001355d9-185406	-	-	-	-	1.3	-
RGB J0152+017	-	-	-	-	1.0	-1.0×10^{-3}
1ES 0229+200	-	-	-	-	1.1	-
PKS 0301-243	-	-	-	-	1.3	-
1ES 0347-121	1.1	0.79	-5.1×10^{-4}	8.3×10^{-3}	0.71	-6.0×10^{-4}
1ES 0414+009	-	-	-	-	1.4	-
PKS 0447-439	-	-	-	-	1.0	-
PKS 0548-322	-	-	-	-	0.94	-1.2×10^{-3}
1RXS J101015d9-311909	-	-	-	-	0.98	-
1ES 1101-232	4.2	1.5	-	8.0×10^{-3}	0.82	-
1ES 1312-423	-	-	-	-	1.2	-8.1×10^{-4}
PKS 1440-389	-	-	-	-	0.68	-
AP Librae	-	-	-	-	0.86	-
PG 1553+113	-	-	-	-	0.51	4.9×10^{-4}
PKS 2005-489	2.7	0.89	3.4×10^{-4}	6.3×10^{-13}	0.70	2.8×10^{-4}
PKS 2155-304	1.5	0.69	-3.0×10^{-4}	5.7×10^{-15}	0.66	-3.0×10^{-4}
PKS 2155-304*	1.7	0.56	-2.4×10^{-4}	7.9×10^{-12}	0.51	-2.4×10^{-4}
1ES 2322-40d9	-	-	-	-	1.2	-
H 2356-309	1.1	0.76	-	5.0×10^{-5}	0.71	-
PKS 0625-354	-	-	-	-	1.08	-
M 87	2.5	0.93	-	1.9×10^{-5}	0.82	-
Centaurus A	-	-	-	-	0.92	-
PKS 1510-089	-	-	-	-	0.88	-
NGC 253	-	-	-	-	1.1	-

By fitting the long-term data with a CAR(1) model, we produce information on the noise behavior for every group in the LCs. We see that for most sources (17 out of 24), we cannot retrieve the CAR(1) fitting. This does not necessarily mean that these sources' noise is uncorrelated but that the flux is not sampled enough. As expected, the correlation time τ is close but higher than the minimum data separation time (> 1 day). For two sources, the p-value significance of the CAR(1) model over the pure WN model is greater than 0.005 and higher than 0.00001 for another two. For PKS 2005-489 and PKS 2155-304, the CAR(1) models are quite significant with p-values below $\sim 10^{-12}$. A linear component is obtained for 8 of the HEGS sources with slopes of the order of $\sim 10^{-3}$ and $\sim 10^{-4}$, where 6 of them have negative trends, indicating a lowering emission with time.

In Figure 9.1 we show the fit example for the two most significant sources in Table 9.2, PKS 2005-489 and PKS 2155-304*.

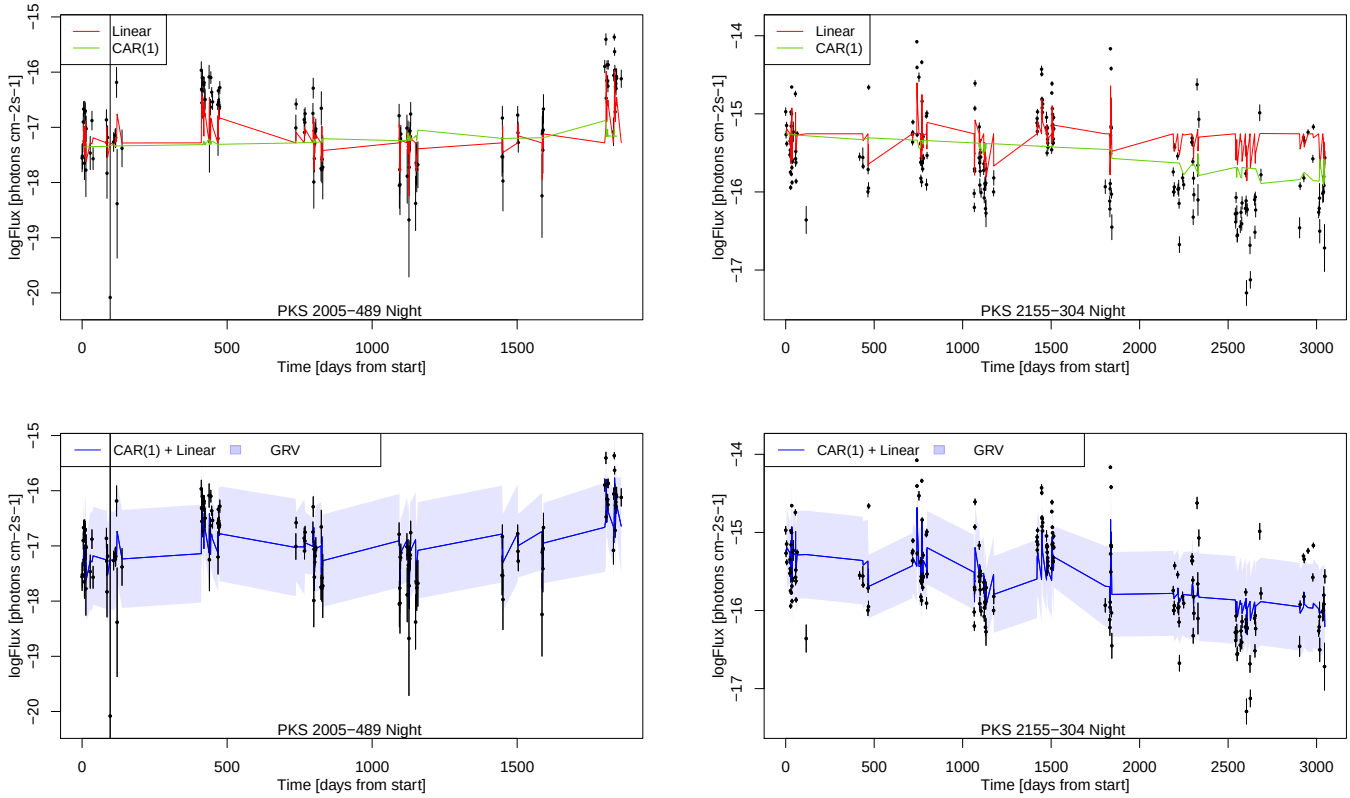

Figure 9.1. Night-wise. MCMC CAR(1) fit for the two sources with the most significant periodic signals in Table 9.2: [*Left Panels*] PKS 2005-489. [*Right Panels*]. PKS 2155-304*. [*Top Panels*] Separated fitted components of the model. Green line indicates the linear trend. Red line indicates the stochastic term [*Bottom Panels*] General fit and GRV component.

From the figures, we see that the τ values 2.7 and 1.7 for each source make that the auto-regressive behavior with previous data points has only memory between nights range. For comparison, the Figure 9.2 presents the fit example of the WN models for both sources. We see that the linear terms retrieved by both models are similar, and the GRV (depending on D and τ, eq. 6.7) and σ WN terms are comparable.

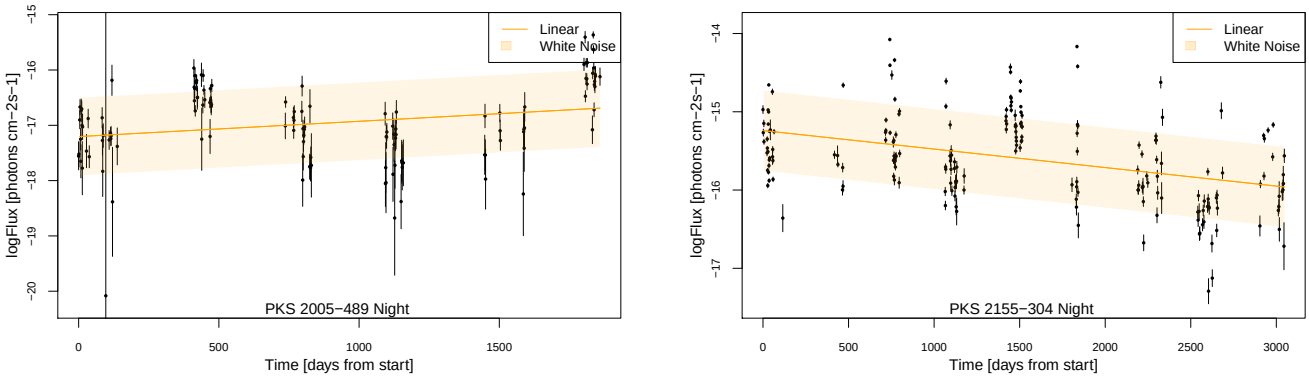

Figure 9.2. Night-wise. MCMC fit for [*Left Panel*] PKS 2005-489. [*Right Panel*]. PKS 2155-304*.

Period-wise

Now, in Table 9.3 and Figure 9.3, we present the results on the period-wise LCs, where the sampling time between the data points at each observation group is on the order of $\sim$ 30 day, as all the observations are integrated on a monthly basis. Only the two most significant sources in Table 9.2 PKS 2005-489 and PKS 2155-304*, represented in Figures 9.1 and 9.2, show a suitable CAR(1) model in the period-wise data, where we can account for a longer correlation range.

Table 9.3. Period-wise. MCMC fit results of the HEGS sources. For the CAR(1) model, the list indicates: the the correlation time τ in days; the GRV parameter D and the p-value computed from ΔAIC between the CAR(1) and the WN models. And for the WN model: the sigma value of the Gaussian distribution.

	Model $_{\mathrm{CAR(1)}}$			Model $_{\mathrm{WN}}$
Name	τ	D	pv	σ
PKS 2005-489	90.3	0.85	1.3×10^{-3}	0.63
PKS 2155-304*	52	0.59	3.2×10^{-3}	0.47

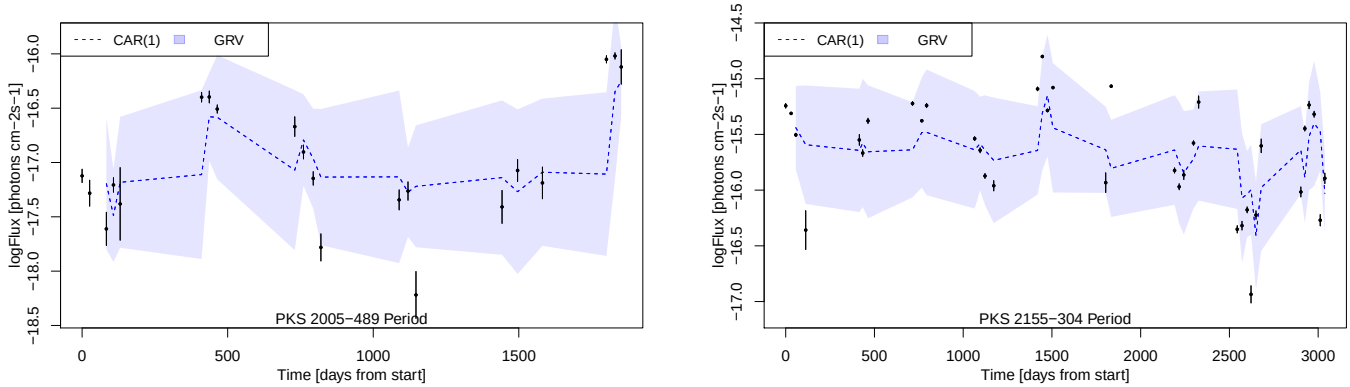

Figure 9.3. Period-wise. MCMC CAR(1) fit for [*Left Panel*] PKS 2005-489. [*Right Panel*]. PKS 2155-304*.

From these results, we see the correlation in a monthly range. The PKS 2005-489 relaxation time is considerably higher ($\sim$ 90 days) than the one for PKS 2155-304 ($\sim$ 50

days). As expected, both values are above $\sim$ 30 days, the median time binning of the flux data points, by a factor of $\sim$ 3 and $\sim$ 1.6, respectively. It is to remark that, for the period-wise data, the linear term does not increase the significance of the fit. For PKS 2005-489, we can see that the last group of points in Figure 9.3 (Left Panel) is accounted for by the high value of τ. Again, as a comparison, we include the WN model fit in Figure 9.4.

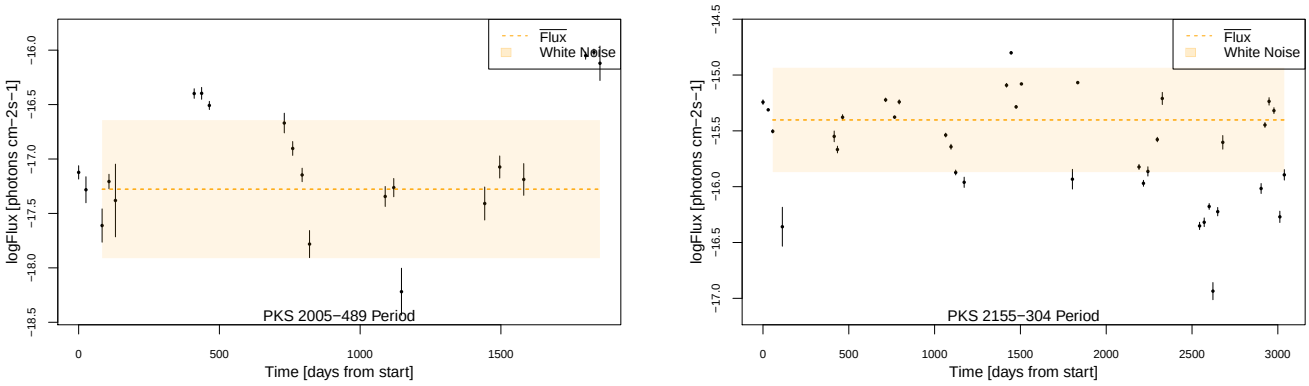

Figure 9.4. Period-wise. MCMC fit for [*Left Panel*] PKS 2005-489. [*Right Panel*]. PKS 2155-304*.

9.1.3 PSD results

With the information on the stochastic components provided by the MCMC fits on Tables 9.2 and 9.3, we can use the Equation 6.61 to produce the PSD in the frequency domain for PKS 2005-489 and PKS 2155-304. In Figure 9.5, we represent together the CAR(1) models PSD of the night and period-wise fluxes, with the WN PSD for comparison.

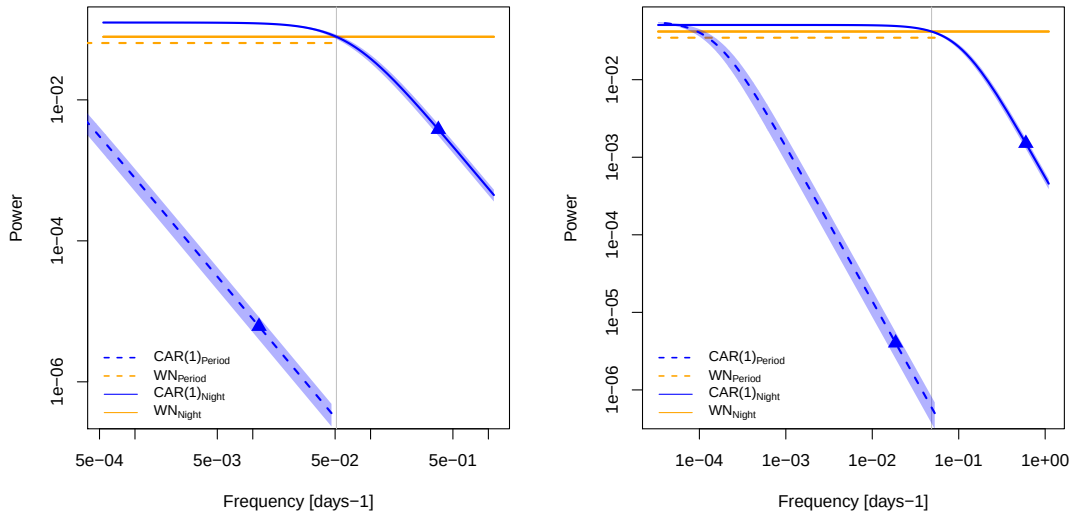

Figure 9.5. PSD from the night and period-wise data CAR(1) and MCMC models for [*Left Panel*] PKS 2005-489. [*Right Panel*] PKS 2155-304*. The triangle symbols indicate the frequency corresponding to the τ values for each set of data. The vertical grey line indicates the cutoff frequency between the CAR(1) and the WN models.

For the CAR(1) model in the shorter time-scale (night-wise), the PSD behaves as a power law ($\propto \nu^{-\beta}$), where $\beta \sim 2$, with a cutoff for a relaxation time scale where the uncorrelated WN is dominant. This is expected for the use of the Equation 6.61 as we are constraining the analysis to CAR(1) models. The cutoff is located at $\sim$ 20 days for both. This means that, for the τ values retrieved, the CAR(1) noise model can account for auto-regressive behavior between the data points if they are separated by 20 days or less for a night-wise LC. For a time-scale greater than the cutoff, the noise in the flux is purely WN.

For the long-term data (period-wise), we are just able to see the auto-regressive behavior part of the spectral density, described as a power-law. For PKS 2005-489, the WN flattering is expected to occur at $\sim$ 13000 days and at $\sim$ 8000 days for PKS 2155-304. Both values are beyond the total time interval of the observations, which are spanned over $\sim$ 1800 and $\sim$ 3000 days, respectively.

9.2 Multiwavelenght analysis of PKS 2155-304

Here, we make a preliminary joint HE and VHE periodicity analysis for PKS 2155-304 using Fermi-LAT and H.E.S.S. data. For the H.E.S.S. data, we use the period-wise LCs, since the data points monthly sampling is similar to the Fermi-LAT LC obtained from the LCR. The starting point of the joint LC is selected to be the first Fermi-LAT flux point (from August 2008), in order to avoid the flux gaps of the H.E.S.S. LC.

To perform the analysis, we apply the following procedure:

- We use the logarithmic values of the flux for both Fermi-LAT and H.E.S.S. datasets.
- As we are interested in the relative behavior of the LCs (linear or periodic terms), we normalize the flux for both datasets in order to make them comparable.
- For every H.E.S.S. flux point, we find the nearest Fermi-LAT flux point in time. We match both values and make a correlation analysis.
- Then, we substitute the Fermi-LAT flux point with its corresponding match H.E.S.S. flux point.
- Finally, we apply the periodicity search algorithm described in Chapter 6 to produce a joint result.
- To compare the fit accuracy on both datasets, we produce residual distributions and theirs normality test for the Fermi-LAT and H.E.S.S. points, separately.

This analysis is first performed using the HEGS data, where only H.E.S.S. Phase I data is included (up to 2012). Then, we produce the source LC for Phase I and II, using the observations available on the source up to 2019.

9.2.1 Phase I

In Figure 9.6, we show the results of the correlation study by plotting the H.E.S.S. in front of the Fermi-LAT data points. By drawing the median flux value for each set

of data (horizontal and vertical lines), we divide the plot into four quadrants. A first hint of correlation is given by counting the number of points at each quadrant, with 7 points in the bottom-left and upper-right panels and only 3 points in the upper-left and bottom-right ones. Then, we apply the Pearson correlation test that quantifies the linear correlation between two sets of data. It is obtained by computing the ratio between the covariance and the standard deviations of both variables ($P_{cor} = cov(H, F)/\sigma_H \sigma_F$). A Pearson's coefficient of 0.47 is retrieved, indicating a non-negligible correlation.

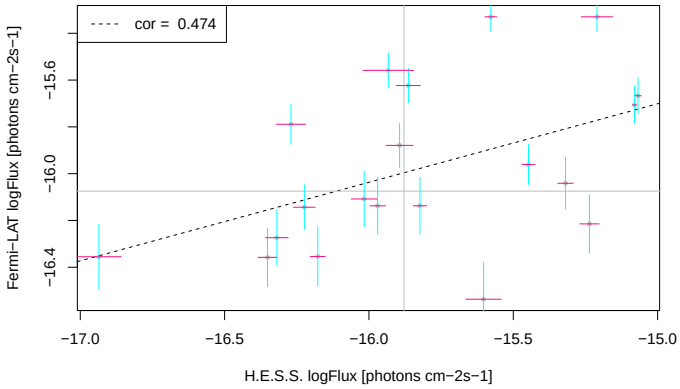

Figure 9.6. PKS 2155-304. Phase I. Fermi-LAT vs H.E.S.S. joint flux data points. Blue bars indicate the Fermi-LAT 1σ error. Pink bars indicate the H.E.S.S. 1σ error. Horizontal and vertical lines indicate the median flux values. Dashed line indicates the linear correlation.

Now, we apply the MCMC algorithm using a CAR(1) model for the stochastic component. The MCMC configuration is similar to the one describe in Chapter 8. The results of this analysis are represented in Table 9.4 and Figure 9.7.

Table 9.4. Phase I. MCMC fit results for the joint H.E.S.S. and Fermi-LAT LC of PKS 2155-304. The list indicates: the best model in terms of AIC; the period mean and standard deviation in days; the CAR(1) τ mean in days; the CAR(1) D mean; the p-value computed from ΔAIC; the only noise model τ and D; the test for normality p-value from the AD test for the global, H.E.S.S. and Fermi-LAT residuals.

	CAR(1) Periodic Model					CAR(1) model		Residuals		
4FGL Name	Best Model	Period	τ	D	pv	τ_{noise}	D_{noise}	pv_{AD}	$pv_{\mathrm{AD,\,H}}$	$pv_{\mathrm{AD,\,F}}$
J2158.8-3013	l, s	$597_{\pm 22}$	54	0.47	7.6×10^{-4}	74	0.59	0.31	0.59	0.78

As in Chapters 7 and 8, the best MCMC model is a linear plus sinusoidal. The ~ 600 days periodicity found, with a p_{value} far below 0.005, is also compatible with previous results and literature on the source oscillation. Including the H.E.S.S. data creates an irregular sampling different from the 30 days binning in the LCR. This can explain the reduction of the correlation time τ parameters up to 54 days with respect to the results in Table 8.2.

By visually inspecting the Figure 9.7, we can see the proximity of the H.E.S.S. (cyan) and Fermi-LAT (pink) data points. The H.E.S.S. values follow well the periodic and linear trend of the general fit. The highest oscillation states at ~ 900 and ~ 1500 days and the lowest state at ~ 1200 days account also well for the H.E.S.S. data points.

It may be that the inclusion of the H.E.S.S. data in the fit may not significantly affect the output of the MCMC fit, since the number of Fermi-LAT data points (158) is

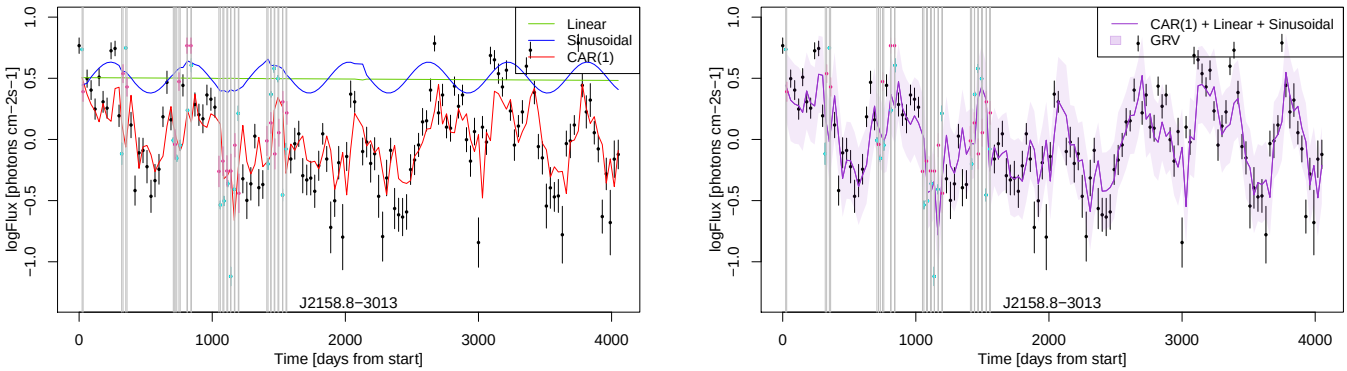

Figure 9.7. PKS 2155-304. Phase I. MCMC fit for the joint H.E.S.S. and Fermi-LAT LC. Blue points indicate the H.E.S.S. values included in the fit. Pink points indicate the Fermi-LAT values removed from the fit. Black points indicate the Fermi-LAT values included in the fit. Vertical grey lines indicate the position of the matched flux points. [*Left Panel*] Separated fitted components of the model. Green line indicates the linear trend. Blue line indicates the sinusoidal term. Red line indicates the stochastic term [*Right Panel*] General fit and GRV component.

considerably higher than the number of H.E.S.S. data points (20). Nevertheless, as we have mentioned, both fluxes have some correlation and follow the deterministic trends in the LCs fit. To measure how accurately both datasets are described by the output parameters, we separate their residuals to test the goodness of fit, as represented in Figure 9.8. The AD test p-value for the global fit is 0.31. The fit for the Fermi-LAT data is better than for H.E.S.S., with values of 0.78 and 0.60, respectively. It is also noticeable that the separate residual analysis is better than the global one. Nevertheless, all of them are above 0.05, meaning that the fit's residuals have no significant deviation from normality.

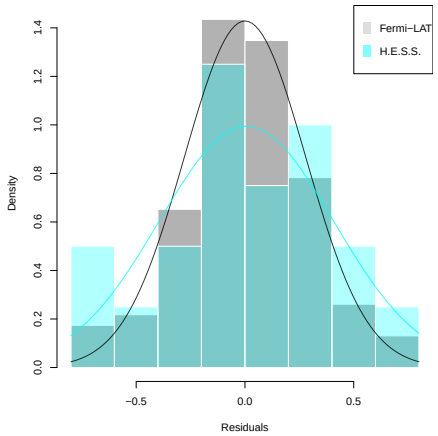

Figure 9.8. Phase I. Residuals histogram for the H.E.S.S. and Fermi-LAT datasets. Cyan and Black lines indicate Gaussian distribution fits with N_{mean} and N_{σ} parameters.

9.2.2 Phase I & II

In this section, we perform the same analysis for the HEGS PKS 2155-304 data but adding observations from the H.E.S.S. Phase II (up to October 2019). Before doing so, we briefly

describe the H.E.S.S. analysis configuration and output on the sources. The data is analyzed using the *standard cuts* of the semi-analytical *Model ++* analysis framework in order to reconstruct the events and perform the gamma-hadron separation. We select a *HESS I Stereo* profile, in which the analysis is performed using the four smaller telescopes (CT1-4) in Stereo mode, where each event is recorded by at least two telescopes.

In Figure 9.9, we present the analysis results in terms of significance. The θ^2 distribution shows the number of counts for every squared angular distance to the source, from which the computed significance is $\sigma = 251$ for 430 hours. The sky map shows the spatial distribution of the signal significance, where the point source PKS 2155-304 is retrieved at a high significance surrounded by the background signal distribution. Also, a significance distribution is displayed, where the black lines correspond to the source signal.

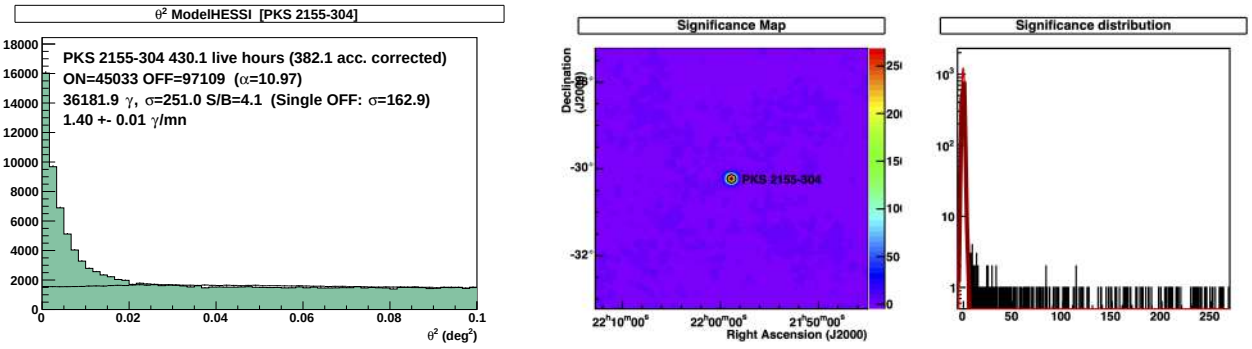

Figure 9.9. PKS 2155-304. Phase I & II. Results on the Model++ HESSI Stereo analysis. [*Left Panel*] θ^2 distribution. [*Center Panel*] Significance Map. [*Right Panel*] Significance distribution.

From this point-like source result, we can perform an analysis to obtain the spectra and LC. In Figure 9.10, we present the spectral analysis results.

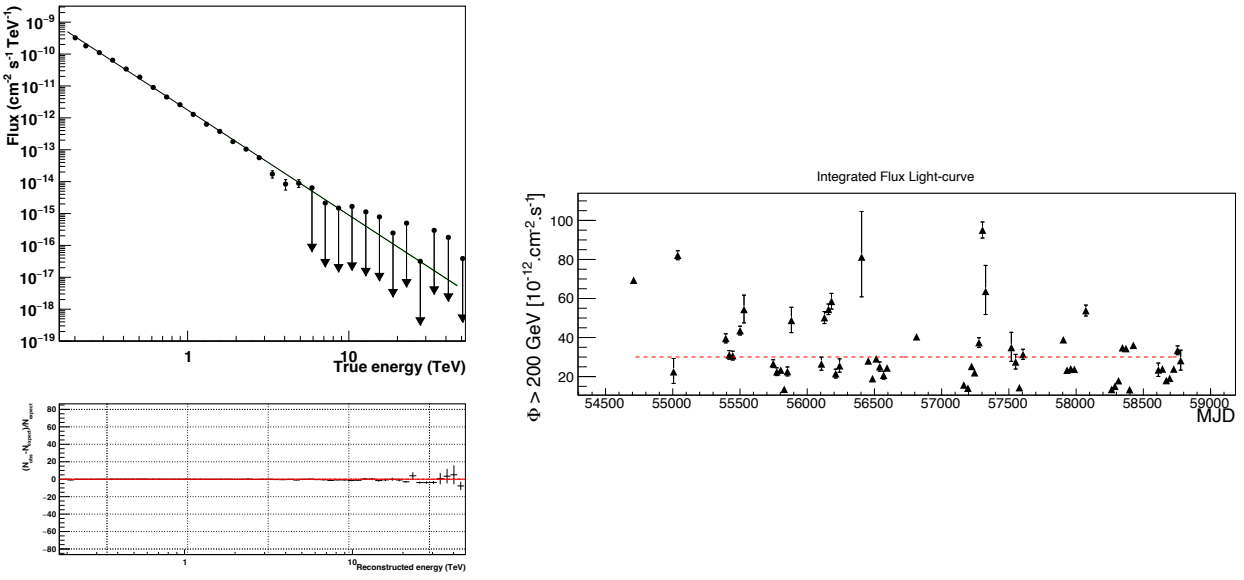

Figure 9.10. PKS 2155-304. Phase I & II. Results on the Model++ HESSI Stereo spectral analysis. [*Left Panel*] Spectrum and residuals. [*Right Panel*] Period-wise LC.

The spectrum is performed using a power-law fit giving the following results: a flux normalization of $\Phi_0 = 38.38_{\pm 0.20} \times 10^{-8}$ m^{-2} s^{-1} TeV^{-1} with an spectral index of $\Gamma =$

$3.30_{\pm 0.01}$ at a reference energy of $E_0 = 0.39$ TeV and for an energy range from 0.18 to 47.06 TeV. Now, from the spectrum results, we compute a period-wise LC using an energy range between 0.02 and 125 TeV for a reference energy of 0.2 TeV.

Now we can repeat the procedure applied to the Phase I data. In Figure 9.11, we show the results of the correlation study by plotting the H.E.S.S. in front of the Fermi-LAT data points. Again, the difference in data point numbers in the diagonal quadrants is a first hint of flux correlation, with 21 points in the bottom-left, 22 upper-right panels and only 7 points in the upper-left and bottom-right ones. The Pearson's coefficient for this set of data is 0.47, identical than the Phase I result, which implies a temporal stability of the emission correlation.

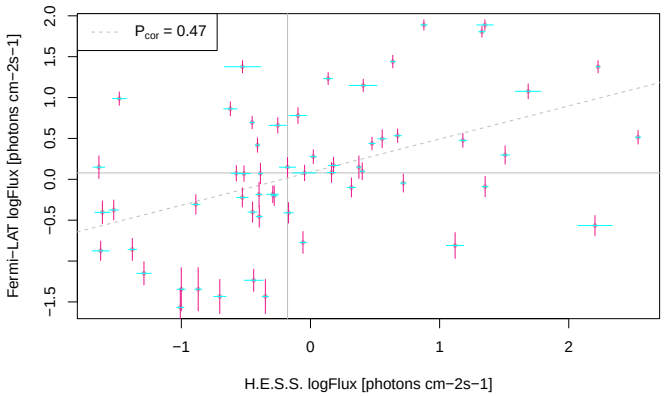

Figure 9.11. PKS 2155-304. Phase I & II. Fermi-LAT vs H.E.S.S. joint flux data points. Blue bars indicate the Fermi-LAT 1σ error. Pink bars indicate the H.E.S.S. 1σ error. Horizontal and vertical lines indicate the median flux values. Dashed line indicates the linear correlation.

The same MCMC configuration is used to obtain the results presented in Table 9.5 and Figure 9.12.

Table 9.5. Phase I & II. MCMC fit results for the joint H.E.S.S. and Fermi-LAT LC. The list indicates: the best model in terms of AIC; the period mean and standard deviation in days; the CAR(1) τ mean in days; the CAR(1) D mean; the p-value computed from ΔAIC; the only noise model τ and D; the test for normality p-value from the AD test for the global, H.E.S.S. and Fermi-LAT residuals.

	CAR(1) Periodic Model					CAR(1) model		Residuals		
4FGL Name	Best Model	Period	τ	D	pv	τ_{noise}	D_{noise}	pv_{AD}	$pv_{\mathrm{AD,\,H}}$	$pv_{\mathrm{AD,\,F}}$
J2158.8-3013	l, s	$599_{\pm 14}$	30	0.97	3.1×10^{-4}	40	1.15	0.12	0.22	0.65

No significant difference is found in the periodicity result. Again, the linear plus sinusoidal model is fitted with a period of ~ 600. However, adding more H.E.S.S. to the analysis improves the fit, as the standard deviation in the period parameter is reduced to 14 days. Also, we can notice the influence of the H.E.S.S. data in the stochastic CAR(1) components, as the noise τ parameter has a value of 30 days, a reduction by almost half if compared with the Phase I result.

Now, we can visually inspect the Figure 9.12. Again, the H.E.S.S. points follow well the fit for the oscillation states at ~ 900, ~ 1200 and ~ 1500 days. At ~ 2500 days from start we see the upward trend of the cyan points following the sinusoidal behavior, and the coinciding high emission state of HE and VHE data. In the last three group of

points ($\sim$ 3500, 3600 and 4000 days from start), we also see the LC points following the low/high emission state as expected by the fit.

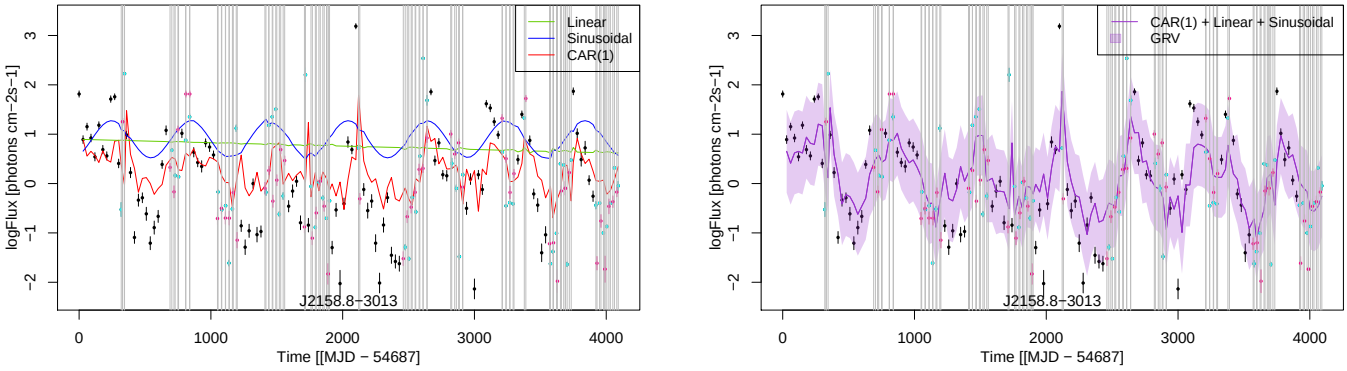

Figure 9.12. PKS 2155-304. Phase I & II. MCMC fit for the joint H.E.S.S. and Fermi-LAT LC. Blue points indicate the H.E.S.S. values included in the fit. Pink points indicate the Fermi-LAT values removed from the fit. Black points indicate the Fermi-LAT values included in the fit. Vertical grey lines indicate the position of the matched flux points. [*Left Panel*] Separated fitted components of the model. Green line indicates the linear trend. Blue line indicates the sinusoidal term. Red line indicates the stochastic term [*Right Panel*] General fit and GRV component.

From this Phase I & II analysis, the AD test p-value is 0.22 for the global fit, and 0.81 for both H.E.S.S. and Fermi-LAT datasets. In comparison with the Phase I, the number of H.E.S.S. data points has increased by almost three, with a total of 57 points for the total of 80 Fermi-LAT points.

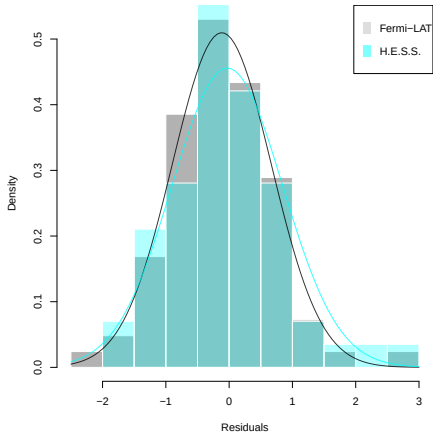

Figure 9.13. PKS 2155-304. Phase I & II. Residuals histogram for the H.E.S.S. and Fermi-LAT datasets. Cyan and Pink lines indicate Gaussian distribution fits with N_{mean} and N_{σ} parameters.

9.3 Multiwavelenght correlation of PG 1553+113

PG 1553+113 is a periodic source well studied in literature and in this thesis. The number of observations by H.E.S.S. is not enough to perform a periodicity search analysis on the source. As done for PKS 2155-304 in previous section, we obtain the period-wise LC in order to compare it with the Fermi-LAT photons. In Figure 9.14

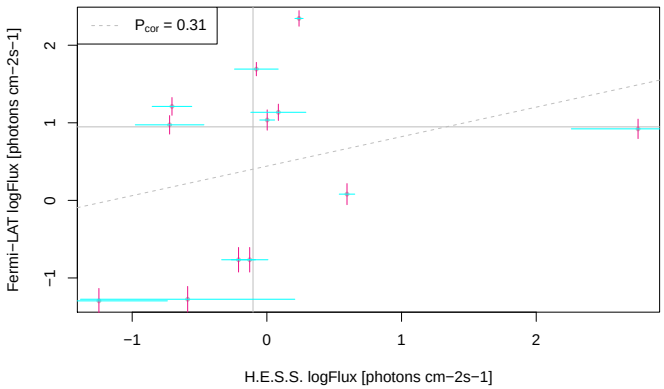

Figure 9.14. PG 1553+113. Phase I & II. Fermi-LAT vs H.E.S.S. joint flux data points. Blue bars indicate the Fermi-LAT 1σ error. Pink bars indicate the H.E.S.S. 1σ error. Horizontal and vertical lines indicate the median flux values. Dashed line indicates the linear correlation.

For this source, we also see a difference in data point numbers between the diagonal quadrants divided by the mean flux values. The positive linear correlation is indicated with 4 points in the bottom-left and upper-right panels and only 2 points in the upper-left and bottom-right ones. Pearson's coefficient for this case is 0.31, suggesting a relatively low but not negligible positive correlation between the data sets.

9.4 Discussion

The PSD represents the variability of the blazar emission as shown in the CAR(1) fitted components in Figures 9.1 and 9.3. It provides information about the contribution of variability power at the different scales of a time-series, represented in the frequency domain ($\nu \propto 1/t$). Here we use the PSD to model the stochastic noise behavior of the blazar emission after subtracting the deterministic trends, e.g. linear. We show the results for different data sampling and the effect on the noise power for different time-scales, valuable information for understanding the variability characteristics of AGN LCs.

In Thiersen et al. (2019) the long-term variability is studied in the framework of leptonic blazar emission. They simulate the variability associated with the accretion flows with different time-dependent variations of the LCs. They show that these variations are compatible with multiwavelength power-law PSDs with an index of $\beta = 2$, as in our results for the PSDs derived from the stochastic CAR(1) components.

Goyal (2019) present a PSD study for PKS 2155-304 considering intra-night (run-wise) and long-term (period-wise) time-scales, using a nightly and a minute-binned H.E.S.S. LC for a decade of observations. The LCs are modeled by CARMA processes, where a CARMA(2,1) best describes the long-term variability, while the intra-night variability is described by a CARMA(1,0) or CAR(1) model. Unlike our method, they do not fit the deterministic components of the LC as they are modeling the auto-correlation function. They interpret that the presence of different slopes $\beta \sim 1-2$ in the PSD characterisation could be explained if the emission is generated in an extended, highly turbulent jet. Using a similar PSD analysis, multiwavelength (HE, X-ray, optical, radio) variability studies are carried out on Fermi-LAT blazars such as PKS 0735+178 (Goyal et al., 2017) and OJ 287 (Goyal et al., 2018).

In (Moreno et al., 2019) they develop stochastic modeling techniques to describe AGN variability using CARMA components and their corresponding PSDs. They propose the use of a Damped Harmonic Oscillation (DHO) modeled as a CARMA(2,1) (similar to the CAR(2) components described in Appendix B.1) to measure AGN periodicities or QPOs. In contrast, in Chapters 7 and 8 we describe the periodic AGN emission by a CAR(1) component and an independent sinusoidal oscillation of the LC. A detailed PSD analysis on a selected sample of Fermi-LAT blazars using ten-year LCs is performed in Tarnopolski et al. (2020). Among other methods, they use ARMA and CARMA models and find power-law PSDs in the $\beta \sim 1-2$ range. They also confirm the QPO of PKS 2155-304, but find no other significant periodicity in the Fermi-LAT blazar sample. A similar analysis of simulated LCs, optical and X-ray sources is presented by Kelly et al. (2014).

Regarding the HE/VHE periodic fit of PKS 2155-304, it should be noted that with a more balanced number of points between the two wavelengths, the fit still gives the same periodicity and noise model output, even with a higher significance. The p-value for the H.E.S.S. Phase I & II analysis ($p_{value} = 4.5 \times 10^{-5}$) is better than that for the HEGS data set ($p_{value} = 7.6 \times 10^{-4}$). It is also better than the results for the Fermi-LAT data only, where p_{values} are of the order of $\sim 10^{-3}$. A correct fit to the data is also expected from the ~ 0.5 correlation between the two data sets. A previous work by (Abdalla et al., 2017) investigated the HE/VHE PKS 2155-304 emission correlation and compared the PSDs of both energy bands.

From here, some work can be done to improve the joint analysis on PKS 2155-304 and other AGNs. First, adding more data points as they are available from H.E.S.S. observations. Also, perform the analysis on other H.E.S.S. sources for which the number of period-wise data points is greater than ~ 20 in order to be able to perform meaningful correlation analysis, depending on the Fermi-LAT data. The step beyond is to perform the search using γ-ray photons and well-sampled data in different spectrum frequencies. In Ackermann et al. (2015), they provided the first multiwavelength report of a QPO for PG 1553+113 in the radio, optical, and γ-ray LCs where the oscillations were correlated, as in the results for this chapter. This supposes more robust evidence of the oscillation detected. In Otero-Santos et al. (2020), they report QPO behavior for the blazars 3C 66A and B2 1633+38 in the optical and γ-ray bands.

We could also have taken into account the possible time lag between HE and VHE emissions, if any between these bands. Several papers have provided multiwavelength results for various γ-ray extragalactic sources on the time correlation between HE/VHE photons with other lower energy bands (Bonnoli et al., 2011; Lindfors et al., 2016; Liodakis et al., 2018, 2019; Acharyya & Sadun, 2023). Understanding the multi-wavelength variability correlations and their corresponding time lags can provide information and constraints on physical emission jet models. In Max-Moerbeck et al. (2014), the expected lagged observation can be described by different moving emission regions in the jet, producing photons at each energy for γ-ray and other wavelengths. The relative distances of the jet regions can be calculated by estimating the time difference between the two detected emissions. In Zhang et al. (2019), the leptonic model predicts correlated variability between the low (radio and optical) and high (γ-rays) energy bands on both short and long time scales, which is not expected for a hadronic scenario.

In this chapter, we have analyzed the limitations of ground-based γ-ray telescopes to produce long-term time-domain analysis. Nevertheless, with sufficient data sampling, we

have produced valuable information on the stochastic behavior for the extragalactic VHE γ-ray sky of H.E.S.S.. We have also taken advantage of the HE regularly sampled LCs from Fermi-LAT to produce a time-correlation analysis between the HE and VHE band, resulting in a joint periodicity search for PKS 2155-304.

Chapter 10

Forecasting using the retrieved parameters

Flux forecasting in astrophysics is valuable for various reasons. It helps optimize observation plans by prioritizing objects with expected flux changes and capturing important events, facilitating observations across different wavelengths. It also helps validate theoretical models by comparing forecasted fluxes with actual observations, improving our understanding of astrophysical processes, and helping to constrain the system parameters.

Specifically, in the context of SMBBH, it is particularly relevant and beneficial. By forecasting the SMBBHs fluxes, we can predict the timing and amplitude of the periodic flux variations, allowing for optimized observation strategies. Flux forecasting enables the identification of specific observational windows when the binary system is expected to exhibit enhanced activity, such as accretion events. By understanding and predicting flux variations in SMBBHs, we gain insights into their dynamics, evolution, and the underlying astrophysical processes in these intriguing systems.

After the application and validation of the MCMC time-domain periodicity search algorithm presented in Chapter 6, we use the emission information provided by the results on the LCR sources in Chapter 8 to predict the future HE flux behavior as accurately as possible. First, in Section 10.1, we describe the process followed to determine the expected values and their uncertainty. In Section 10.3 we evaluate the forecasting algorithm performance by fitting a portion of the available data for each source and then comparing the forecasted values with the remaining data. Then, in Section 10.4, we use the parameters retrieved from the LCR analysis to predict the future values and the next high emission state. Finally, the results are discussed and some remarks are done in Section 10.5.

10.1 Simulating the future values and the prediction intervals

To do this analysis, we use the results obtained for the Fermi-LAT LCR periodic sources represented in Table 8.1 and 8.2. These results are obtained for a CAR(1) stochastic model, so in order to sample the following values, we consider the Equations 6.31 and 6.32. For the deterministic part, we use the fitted best model, using Equations 6.33, 6.34 and 6.35. Nevertheless, the method we present below can be applied for any model equation if the parameters that describe the model are available.

The idea is to sample the next values using the posterior distribution for every parameter in the models. It is essential to take into account the correlation between every parameter. For example, we can consider the correlation results of J0509.4+0542 and J1555.7+1111 from Figure 8.14. After the de-standardization, the parameters τ and D are highly correlated, with a value of 0.87 and 0.84, respectively. Then, if we sample random values from these parameters using each mean and standard deviation, the results would have a much larger dispersion than expected. Thus, to address for correlation, we directly use all the samples from the MCMC chains for each considered source (see Figure 8.3).

With this, for every source, the procedure to determine the forecast values with prediction intervals is:

- To compute the first forecast $\hat{\phi}(t_{n+1})$, the last time value t_n, the last lagged time value $s_n = t_{n-1}$, and the last flux value $\phi(t_n)$ are taken from the last observed point.
- We obtain the N samples from the parameters MCMC posterior distribution. In this case, $N = 3000$.
- Using the N samples, we compute N realizations of the next forecast value $\phi_N(t_{n+1})$ using the equations for the specific model.
- To compute the GRV, we draw N random normally distributed samples using the Equation 6.31 as the standard deviation.
- By taking the mean from the N realizations, we determine the next forecast value $\hat{\phi}(t_{n+1})$. By taking the standard deviation from the Nrealizations, we determine the Prediction Intervals (PIs) $\hat{\sigma}_{n+1}$. Then, the first forecast value is $\hat{\phi}(t_{n+1}) \pm \hat{\sigma}_{n+1}$
- To compute the following h forecast time values, we add the desired step to the previous time value, i.e., $t_{n+h} = t_{n+h-1} + \Delta t_{n+h}$. The forecast lagged value is $s_{n+h} = t_{n+h-1}$ Using a CAR(1) model for the noise, we can use any Δt at every step.
- To compute the following h forecasts flux values, we use the previous value as the last flux point $\hat{\phi}(t_{n+h-1})$. Then, at each step h, we repeat the sampling of N realizations and get the average and the standard deviation, giving $\hat{\phi}(t_{n+h}) \pm \hat{\sigma}_{n+h}$.

This sampling approach is analogous to randomly drawing parameters from the MCMC parameters distribution characterized by their mean and standard deviation while considering the correlation between parameters through a covariance matrix. Using both approaches, we can derive the PI accurately, which provide a measure of uncertainty. When forecasting the first values, the standard deviation of the forecast closely resembles the standard deviation of the fitted values residuals. As we look further into the future, the PIs associated with the forecasts progressively widen due to the propagation of parameter uncertainties.

10.2 Evaluating the forecast accuracy

In this section, the aim is to present the validation tools for the forecasting method described above.

10.2.1 Evaluation data sets

The accuracy of forecasts can be evaluated by assessing the performance of a model on new data that were not included during the model fitting process.

To do so, we split the data into two sets:

- **Training set:** Observation values used to fit the data as we have done in previous chapters using the MCMC algorithm, retrieving the model parameters we use to forecast.
- **Validation set:** Observation values coming after the fitted data points, used to evaluate the forecast accuracy.

This analysis allows us to evaluate the accuracy and reliability of the forecasting method by quantifying the agreement between the predicted values and the unseen observations.

10.2.2 Evaluation metrics

To assess the goodness of forecasting we compute the Forecast Errors (FE), that express the difference between the prediction mean values and the observation data points from the validation set:

$$e_{n+h} = \phi(t_{n+h}) - \hat{\phi}(t_{n+h}) \tag{10.1}$$

As proposed by Hyndman & Athanasopoulos (2021), from the FE, the following accuracy metrics are computed:

- **Scale-dependent errors**

  The errors scale is the same as the underlying data scale, and they are expressed in the same units. With this feature, they are easy to interpret In this case, the scale and units are given by the Fermi-LAT logarithmic flux in [photons cm^{-2} s^{-1}]

  For this type of error, we compute the following metric:

  - **Mean Error (ME)**

$$\mathrm{ME} = \frac{1}{h}\sum_{i=1}^{h}(\phi(t_{n+i}) - \hat{\phi}(t_{n+i})) = \frac{1}{h}\sum_{i=1}^{h} e_{n+i} \tag{10.2}$$

  - **Root Mean Squared Error (RMSE)**

$$\mathrm{RMSE} = \sqrt{\frac{1}{h}\sum_{i=1}^{h}(\phi(t_{n+i}) - \hat{\phi}(t_{n+i}))^2} = \sqrt{\frac{1}{h}\sum_{i=1}^{h} e_{n+i}^2} \tag{10.3}$$

- **Percentage errors**

  The errors scale is relative (scaled to 1) and they are unit-free. This allows to interpret easily and compare the forecast results between different time series data sets. Some of the percentage error metrics face the problem of having extreme values when the time series values are close to 0. Also, they tend to penalize more negative than positive errors. The logarithmic nature of the Fermi-LAT dataset avoids the aforementioned problem.

  For this type of error, we compute the following metric:

– **Mean Absolute Percentage Error (MAPE)**

$$\text{MAPE} = \frac{1}{h}\sum_{i=1}^{h}\left|\frac{\phi(t_{n+i}) - \hat{\phi}(t_{n+i})}{\phi(t_{n+i})}\right| = \frac{1}{h}\sum_{i=1}^{h}\left|\frac{e_{n+i}}{\phi(t_{n+i})}\right| \tag{10.4}$$

These error metrics only consider the information provided by the mean future values. Additionally, in order to use all the information available, we propose the following metrics:

- **Prediction intervals accuracy**

  In the previous section, we described the computation of the forecast PIs. To assess their validity, we use two metrics: the percentage of data points that lie within the intervals ($\text{PI}_{\%\text{ values}}$): and the percentage of data points, considering their 1σ observational uncertainty, that lie within the intervals ($\text{PI}_{\%\text{ errors}}$).

- **High periodic state determination**

  One of the most valuable tools about forecasting periodic sources is to predict the next higher flux values in order to prepare specific observations, if necessary. To do so, we use the Bayesian change point (BCP) analysis given by Wang & Emerson (2015) based on the work by Barry & Hartigan (1993). BCP is a statistical method that detects changes in data sequences by modeling them with different segments and evaluating posterior probabilities of change points. It estimates the likelihood of change points occurring at different positions, identifying significant shifts in the underlying process. With this, we detect the time of the highest state in the unseen data ($t_{\text{MAX, BCP}}$). Then, we compare the highest state given by the BCP with the forecasted one ($t_{\text{MAX, F}}$). With this, the new metric we propose is the percentage error given by the division of the distance between the two high states and the period value of each specific source, namely:

  – **Weighted Prediction Distance of the High State (wPD_{HS})**

$$\text{wPD}_{\text{HS}} = \frac{|t_{\text{MAX, BCP}} - t_{\text{MAX, F}}|}{\text{period}} \tag{10.5}$$

10.3 Forecasting the past

We apply the method and evaluation described to the detected periodic sources in Section 8.2.1. In this case, the training set is the data up to 4500 days from the observation's start, corresponding to $\sim 85\%$ of the total observation time considered. Consequently, the validation set is the data from the last training value ($\sim$ 4500 days) up to the last observation value ($\sim$ 5250 days). In this case, to coincide with the Fermi-LAT LCR time binning and be able to compare, we use a fixed monthly step for the forecast points, i.e., $\Delta t_{n+h} = 30$ days.

From the 22 sources analyzed (Table 8.1 and 8.2), 15 show explicit periodic behavior in the training data set (up to $\sim$ 4500 days) with a CAR(1) model and well-computed MCMC parameters for the deterministic and the periodic components. All the remaining non-periodic sources in this training analysis have periodicities above half of the time

Table 10.1. MCMC fit results for the training data set of the periodic AGNs from the Fermi-LAT LCR. For each source, the list indicates for the MCMC fit: the best model in terms of AIC; the period mean and standard deviation in days; the CAR(1) τ mean in days; the CAR(1) D mean; the p-value computed from ΔAIC. The results are sorted by MCMC period detection significance.

	CAR(1) Periodic Model				
4FGL Name	Model	Period	τ	D	pv
J1555.7+1111	l, s	$776_{\pm 13}$	35	0.23	1.2×10^{-6}
J0407.0-3826	l, s	$1093_{\pm 50}$	38	0.68	3.3×10^{-5}
J1048.4+7143	l, s	$1128_{\pm 40}$	97	1.2	4.4×10^{-5}
J1903.2+5540	l, s	$1180_{\pm 43}$	26	0.31	1.1×10^{-4}
J0739.2+0137	l, h	$1652_{\pm 78}$	26	0.55	2.5×10^{-4}
J0030.3-4224	s	$1741_{\pm 78}$	26	0.55	2.5×10^{-4}
J0521.7+2112	s	$1130_{\pm 52}$	55	0.63	4.8×10^{-4}
J1640.4+3945	l, h	$1935_{\pm 146}$	39	0.62	1.1×10^{-3}
J1754.2+3212	s	$2115_{\pm 171}$	40	0.70	1.3×10^{-3}
J1740.5+5211	s	$2041_{\pm 242}$	56	0.81	2.0×10^{-3}
J0217.8+0144	s	$1210_{\pm 52}$	74	0.80	2.5×10^{-3}
J2311.0+3425	s	$1371_{\pm 138}$	82	1.1	3.2×10^{-3}
J2158.8-3013	l, s	$618_{\pm 34}$	72	0.53	4.9×10^{-3}
J0303.4-2407	l, h	$818_{\pm 35}$	36	0.65	6.3×10^{-3}
J2243.9+2021	s	$1827_{\pm 358}$	31	0.50	9.8×10^{-3}

span ($\sim$ 2250) in Chapter 8, so they are not retrieved in the MCMC sampling. In Table 10.1, we show the results on the MCMC fit parameters.

From the MCMC fits, we see that the periodic models are also significant, with all p-values below 0.01. For most of the sources, the deterministic model is different from the one for the $\sim$ 5300 days analysis, implying some variations in the linear or the sinusoidal/harmonic components with time. Regarding the period values, for every source, they are compatible and very close to the values in Chapter 8. For J0521.7+2112, the period corresponds to half the harmonic in the previous analysis, which is equal to the results in the Fermi-LAT sample chapter, as described in Figure 8.6.

10.3.1 Expected metrics from the simulated forecast

Before computing the forecast metrics, it is valuable to get an expectation on the metrics normal values for each source, based on the MCMC fit results and uncertainties. The PIs are taken on a step-wise range, for the N realizations alltogether. In contrast, we can simulate one realization at a time for the entire forecast time range, from $\sim$ 4500 to 5250 days from start. Then, by producing $N = 300$ trajectories which we can evaluate in terms of ME, RMSE and MAPE with respect to the forecast values, which is the mean trajectory of the N realizations.

To extract the metrics from the simulations, for Equation 10.1, we consider $\hat{\phi}(t_{n+h})$ as each of the N realizations of the h steps and $\phi(t_{n+h})$ as the mean value of the N realizations. By extracting the metrics for each of the N trajectories, we obtain a distribution of ME, RMSE, and MAPE valuable to compare with the metrics from the observed data assessment in the next section.

In Figure 10.1, we show the example of the N = 1, 10, 50, 100, 1000 and 3000 simulations for PG 1553+113. In the N = 3000 (the MCMC full samples), we also show the computed forecast mean with PIs. We see that by increasing N, the simulation

converges to the deterministic trends, as expected.

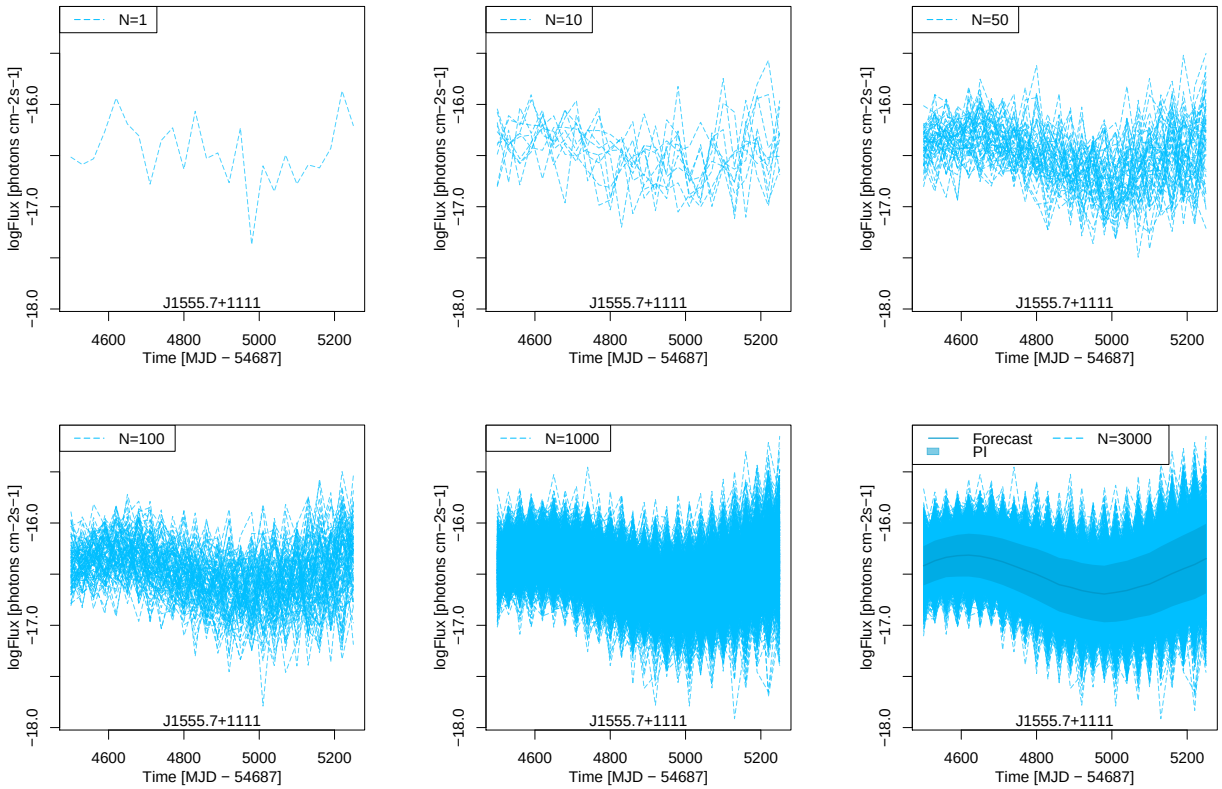

Figure 10.1. Simulated forecast trajectories for J1555.7+1111 based on the MCMC results of Table 10.1. The case of $N = 1,\ 10,\ 50,\ 100,\ 1000$ and 3000 realizations are shown. In the $N = 3000$, the forecast values with PIs are shown.

Then, in Figure 10.2, we show the simulated metrics distribution for the N realizations of the PG 1553+113 forecast. The mean values with standard deviation are drawn (for ME the distribution are centered at 0), which gives a estimation of the expected values for the metrics computed with the observed training data in next section.

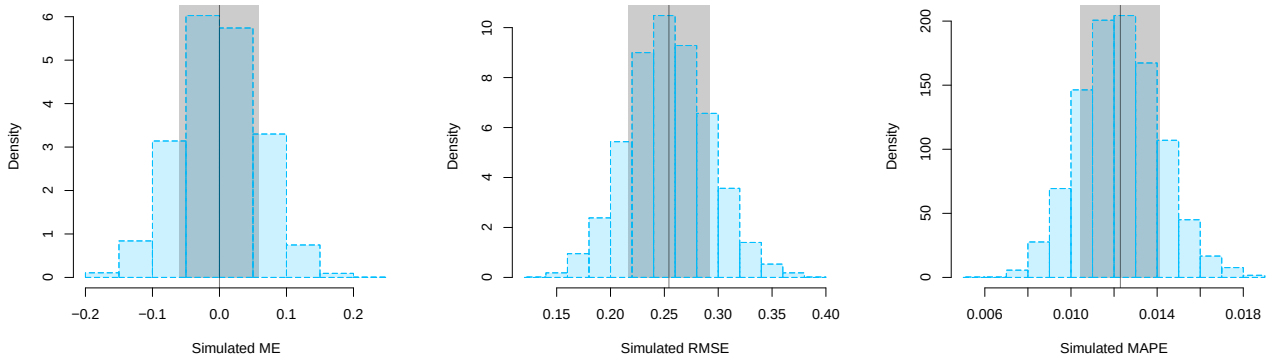

Figure 10.2. Simulated metrics distribution for the J1555.7+1111 $N = 3000$ forecast. [*Left Panel*] ME. [*Center Panel*] RMSE. [*Right Panel*] MAPE. For every metric, the vertical black line indicated the mean value and the shaded area indicates the standard deviation.

By following this procedure for every source analyzed, in Table 10.3 we present the

expected metrics along with the true metrics obtained in the next section, which can be compared to assess the accuracy of the forecast. If any of the metrics is beyond the mean value with standard deviation, we can consider that the forecast or the MCMC fit could be improved, or that the retrieved emission model does not describe precisely the AGN behavior.

10.3.2 Evaluation results on the Fermi-LAT validation set

Now, from the forecast mean and PIs obtained from the $N = 3000$ realizations through the MCMC results, we obtain the metrics by comparing with the observed Fermi-LAT data from the validation data set. In this case, as explained in Section 10.2, we consider $\hat{\phi}(t_{n+h})$ as the mean value of the N realizations and $\phi(t_{n+h})$ as the Fermi-LAT data points. In this case, we compute just one value of each metric for every source, presented in Table 10.2. Figures 10.3, 10.4, 10.5 and 10.6 show, for the four most significant sources in Table 10.1, the MCMC fit, the forecast, and the error distributions. Figure 10.7 shows the example for PKS 2155-304.

Table 10.2. Expected and Forecast metrics from the MCMC fit results of Table 10.1 of the periodic AGNs from the Fermi-LAT LCR. For each source, the list indicates: the expected ME; the expected RMSE; and the expected MAPE (in a scale of 1) the forecast ME; the forecast RMSE; and the forecast MAPE (in a scale of 1); the percentage of data points with and without error bars that lie within the PIs (in a scale of 1); the wPD_{HS}. The results are sorted by MCMC period detection significance.

	Expected Metrics			Forecast Metrics					
4FGL Name	$_{\text{E}}$ME	$_{\text{E}}$RMSE	$_{\text{E}}$MAPE	ME	RMSE	MAPE	$\text{PI}_{\%\text{values}}$	$\text{PI}_{\%\text{errors}}$	wPD_{HS}
J1555.7+1111	± 0.06	$0.26_{\pm 0.04}$	$0.012_{\pm 0.002}$	-0.04	0.22	0.011	0.74	0.92	0.08
J0407.0-3826	± 0.19	$0.78_{\pm 0.11}$	$0.038_{\pm 0.006}$	-0.22	0.83	0.038	0.68	0.76	0.22
J1048.4+7143	± 0.19	$0.75_{\pm 0.11}$	$0.038_{\pm 0.006}$	-0.45	0.78	0.045	0.50	0.83	0
J1903.2+5540	± 0.10	$0.39_{\pm 0.06}$	$0.018_{\pm 0.003}$	-0.06	0.28	0.013	0.73	1	0.48
J0739.2+0137	± 0.22	$0.84_{\pm 0.13}$	$0.042_{\pm 0.007}$	-0.06	0.64	0.031	0.85	0.88	0.09
J0030.3-4224	± 0.16	$0.64_{\pm 0.09}$	$0.030_{\pm 0.005}$	0.23	0.75	0.034	0.65	0.91	0.03
J0521.7+2112	± 0.12	$0.55_{\pm 0.08}$	$0.027_{\pm 0.004}$	0.07	0.43	0.023	0.81	0.92	0.08
J1640.4+3945	± 0.26	$1.05_{\pm 0.15}$	$0.047_{\pm 0.007}$	0.94	1.05	0.057	0.73	0.73	0.25
J1754.2+3212	± 0.18	$0.78_{\pm 0.11}$	$0.035_{\pm 0.005}$	-0.59	0.83	0.033	0.70	0.95	0.26
J1740.5+5211	± 0.18	$0.67_{\pm 0.10}$	$0.031_{\pm 0.005}$	0.45	0.71	0.036	0.65	0.73	0.04
J0217.8+0144	± 0.13	$0.59_{\pm 0.08}$	$0.028_{\pm 0.004}$	0.70	0.92	0.050	0.39	0.43	0.10
J2311.0+3425	± 0.15	$0.76_{\pm 0.11}$	$0.036_{\pm 0.005}$	-0.26	0.81	0.041	0.57	0.78	0.07
J2158.8-3013	± 0.10	$0.41_{\pm 0.06}$	$0.020_{\pm 0.003}$	-0.22	0.40	0.019	0.73	0.92	0.05
J0303.4-2407	± 0.19	$0.84_{\pm 0.13}$	$0.037_{\pm 0.006}$	0.22	0.39	0.018	0.92	0.96	0.04
J2243.9+2021	± 0.14	$0.60_{\pm 0.09}$	$0.027_{\pm 0.004}$	0.26	0.41	0.020	0.88	0.92	0.05

If we analyze all the forecast metrics in Table 10.1, we can obtain different knowledge on the method and on the forecast power. First, given the small number of validation Fermi-LAT points, we cannot expect that the error distribution for the observed metrics is normally distributed. For the simulated forecast we have 3000 realizations and the metrics distributions are normally distributed. Instead, a good forecasting method yields errors with zero mean, or close to zero for a small number of validation points.

Some of the sources have ME metrics in agreement with the limits given by the $_{\text{E}}$ME. Other sources are slightly beyond the expected deviation and a few have larger dispersion. A high difference in the ME metric indicates that the prediction has underestimated the flux in the case of a positive value and overestimated in the case of a negative value. For

most of the sources, the actual values for RMSE and MAPE are below or close to the expected dispersion, indicating a good sign on the power of the forecast to retrieve the general trend of the emission.

We see some clearly biased examples as J1640.4+3045 and J0217.8+0144, where the ME is close to one order of magnitude in flux, and much higher than the $_{\mathrm{E}}$ME. The RMSE values are fairly correlated with the ME. The MAPE is quiet stable over all the sources, which means that the FE are higher for higher flux values. By checking the PI metrics, we see that we can predict more than the 65% of the values for 12 out of 15 sources (more than 75% considering errors). Finally, from the $\mathrm{wPD_{HS}}$, we see that the relative distance between the predicted and the actual emission high state is lower than 15% for 11 out of 15 sources.

The figures left panels show the MCMC fit, the following predicted values, and their PIs, along with vertical lines indicating the predicted next high state and the actual high state calculated using the BCP in the real flux values. As expected, the PIs for the first forecast value is close to the GRV of the last value and grows as the forecast steps in the future.

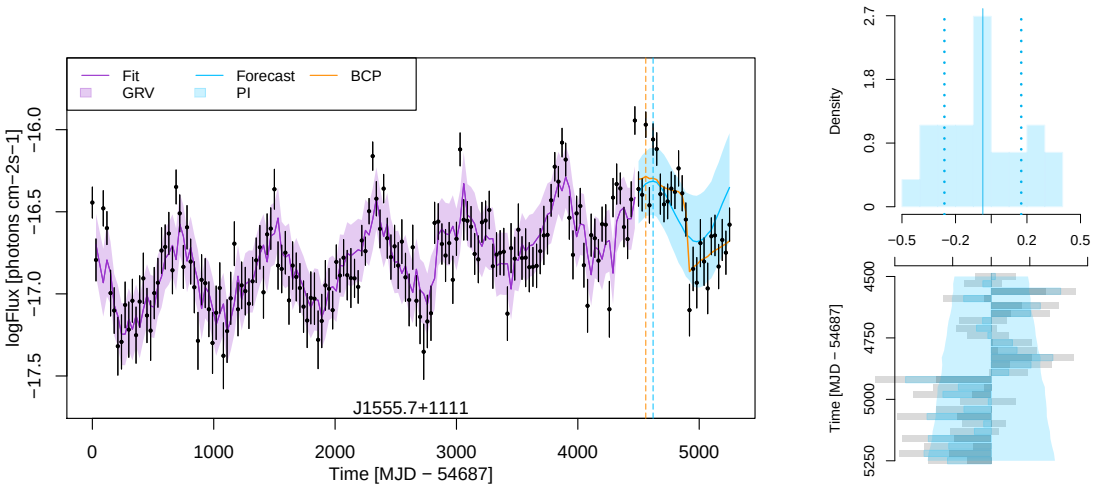

Figure 10.3. J1555.7+1111 [*Left Panel*] MCMC fit with GRV for the training data set and forecast values with PIs for the validation data set. The vertical blue line indicates the predicted next high state of the source. The vertical orange line indicates the BCP high state of the validation data set. [Right Top Panel] FE distribution. The blue vertical solid and dashed lines indicates the ME and the RMSE, respectively. [Right Bottom Panel] FE as a function of forecast time in the future. The light grey bars indicates the FE considering the data error bars. The light blue area shows the PIs.

By a visual assessment, we can say that the forecast may not predict the following flux values with very high precision but instead can retrieve the trend of the source emission. This feature might be more important to understand the source's physical nature and prepare future observations. Analyzing Figures 10.3 and 10.4, we see that the predictions follow the data well and the PIs account for some of the unexpected features in the LCs. Also, the expected future high state of the emission is well retrieved for both sources, with $\mathrm{wPD_{HS}}$ below 0.10.

Analyzing Figure 10.5, we see that, even though the first future values are overestimated, we can predict the high state of the subsequent oscillation. This is a particular case of periodicity, where the oscillation started after $\sim$ 1000 days of the first Fermi-LAT observations. That is why the MCMC fit has a strong CAR(1) correlation time

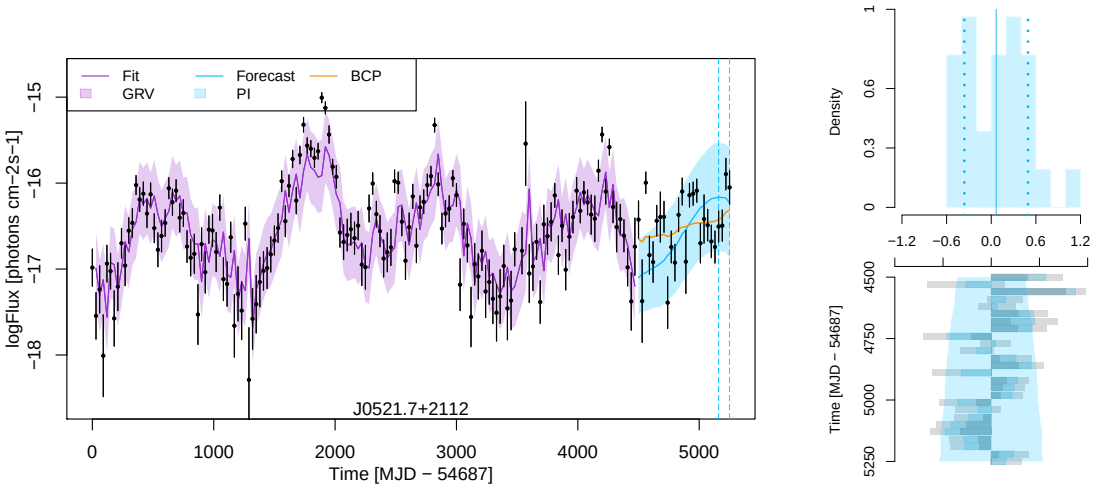

Figure 10.4. J0521.7+2112. Same description as for Figure 10.3

of 105 days (Table 8.2), and then, the forecasting may under/overestimate the emission. The amplitude of this oscillation is considerably high compared to what we see for other sources. Thus, being able to catch most of the values with the PIs and to predict the high states is a promising sign for the forecast method and a strong indication of true periodic behavior and periodic temporal stability.

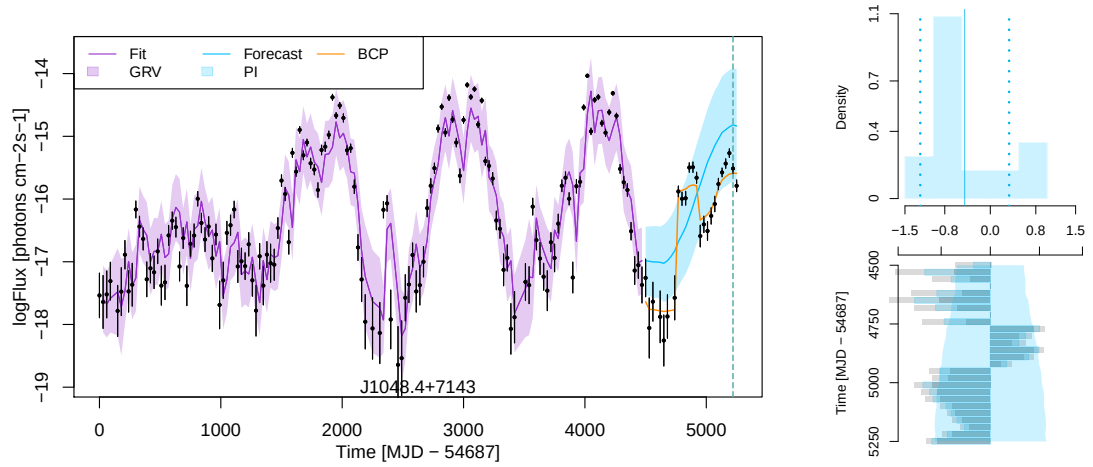

Figure 10.5. J1048.4+7143. Same description as for Figure 10.3

On the other hand, the periodic trend is retrieved for J1903.2+5540 in Figure 10.6, but the forecast first overestimates and last underestimates the flux. Also, we see a slight flare at around 5000 days, which makes the BCP high state determination significantly far from the predicted high state.

Finally, Figure 10.7 shows the example of PKS 2155-304. In Section 8.3, we show an evident period variation in the last slice of the source LC, from 57050 MJD, where the periodicity decreases almost by half, from $\sim$ 600 to $\sim$ 350 days. With the forecasting, we can obtain similar conclusions to those of the time window analysis. As we can see, the forecast line has a period wider than the observed data from the validation dataset. With this, as expected, the predicted high state, even very close to the last training dataset

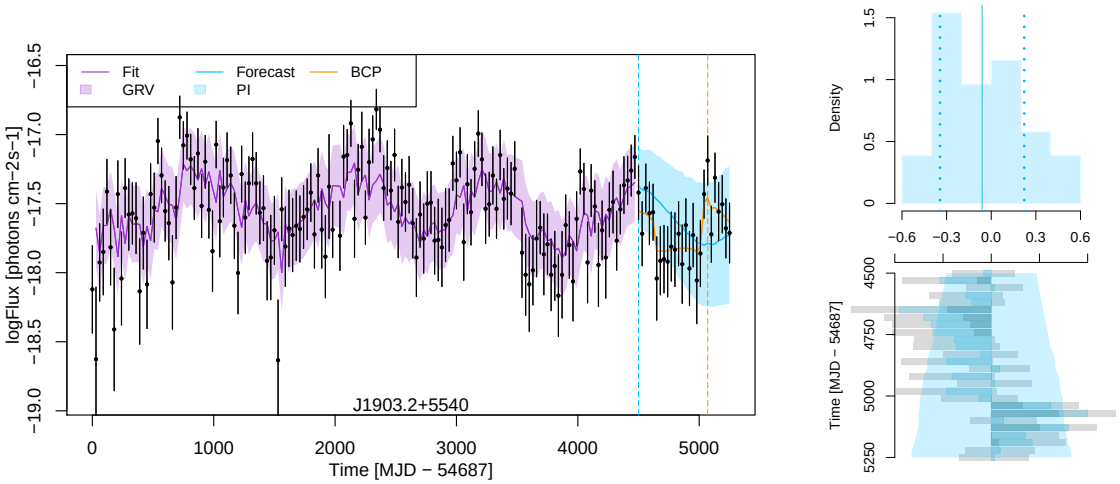

Figure 10.6. J1903.2+5540. Same description as for Figure 10.3

value, is estimated after the actual high state of the oscillation.

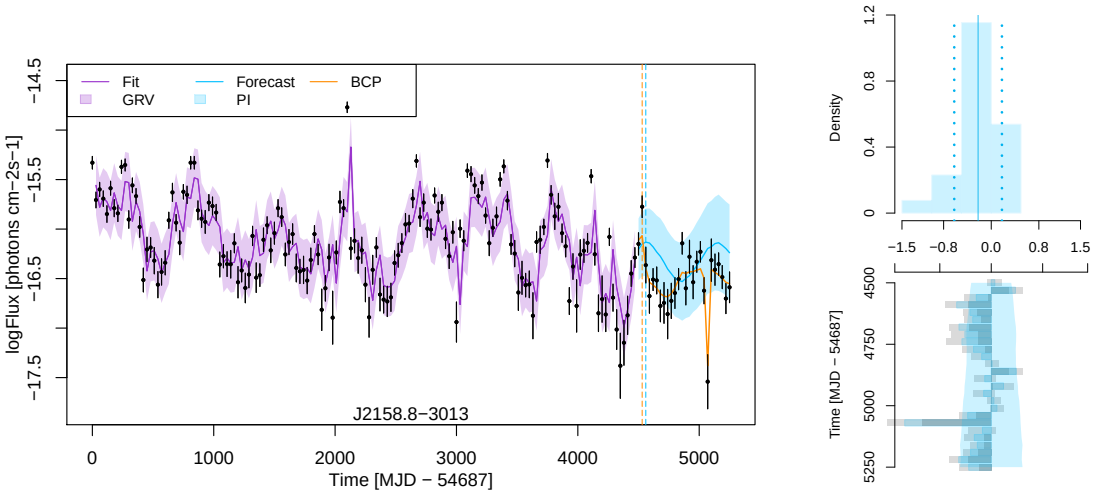

Figure 10.7. J2158.8-3013. Same description as for Figure 10.3

From all of the above, we understand that the specific features, as unexpected diminished flux or flares, may be explained with the use of correlated noise AR/CAR models. In the forecasting, as we are computing the mean expectation of the next values, this auto-regressive behavior is erased. Nevertheless, the uncertainty of this stochastic noise behaviour is refelcted in the growing PIs.

10.4 Forecasting the future

Finally, we follow the same procedure to predict the next unseen flux values and emission high states. To do this, we use the MCMC results on Section 8.2.1, on which all the Fermi-LAT data points are used for the fit, and we follow the forecasting procedure described in Section 10.1.

The last point used for the MCMC fit in the LCR dataset is at ~ 5250 days from start, which means that the last date value is $5250 + 54686 = 59936$ MJD (December 2022), where 54686 MJD is the starting day of the Fermi-LAT observations (August 2008) from the LCR. The future forecast is extended in time to around one period oscillation beyond the first estimated high estate of the emission, which is different for every source.

Table 10.3. Forecast results of the periodic AGNs from the Fermi-LAT LCR presented in Section 8.2.1. For each source, the list indicates: the next 1st ($t_{n+1} = 5280$ days from the start), 5th ($t_{n+5} = 5400$) and 10th ($t_{n+10} = 5550$) forecast logFlux value; the next forecast emission low state time in days from the start and the corresponding logFlux value; the next forecast emission high state time in days from the start and the corresponding logFlux value. For all logFLux values the corresponding PIs is indicated. The logFlux values and PIsunits are [photons cm $^{-2}$s^{-1}]. The results are sorted by MCMC period detection significance (see Tables 8.1 and 8.2).

	Next Forecast Values			Next Low/High State			
4FGL Name	$\hat{\phi}(t_{n+1})_{\pm\text{PI}}$	$\hat{\phi}(t_{n+5})_{\pm\text{PI}}$	$\hat{\phi}(t_{n+10})_{\pm\text{PI}}$	t_{ls}	$\hat{\phi}(t_{ls})_{\pm\text{PI}}$	t_{hs}	$\hat{\phi}(t_{hs})_{\pm\text{PI}}$
J1555.7+1111	$-16.39_{\pm 0.19}$	$-16.22_{\pm 0.21}$	$-16.35_{\pm 0.24}$	5820	$-16.64_{\pm 0.26}$	5400	$-16.22_{\pm 0.21}$
J0509.4+0542	$-16.50_{\pm 0.46}$	$-16.21_{\pm 0.56}$	$-15.85_{\pm 0.71}$	7590	$-16.27_{\pm 1.32}$	5910	$-15.24_{\pm 0.83}$
J0521.7+2112	$-16.43_{\pm 0.43}$	$-16.80_{\pm 0.47}$	$-17.13_{\pm 0.55}$	5640	$-17.23_{\pm 0.58}$	6360	$-15.97_{\pm 0.78}$
J0739.2+0137	$-15.65_{\pm 0.57}$	$-15.29_{\pm 0.65}$	$-15.03_{\pm 0.78}$	6660	$-15.67_{\pm 1.43}$	5760	$-14.82_{\pm 0.91}$
J1640.4+3945	$-17.08_{\pm 0.50}$	$-17.21_{\pm 0.56}$	$-17.46_{\pm 0.65}$	5850	$-17.77_{\pm 0.84}$	6900	$-16.82_{\pm 1.19}$
J1740.5+5211	$-17.30_{\pm 0.55}$	$-17.36_{\pm 0.62}$	$-17.36_{\pm 0.67}$	5490	$-17.39_{\pm 0.63}$	6600	$-16.07_{\pm 1.24}$
J1913.0-8009	$-17.516_{\pm 0.42}$	$-17.25_{\pm 0.45}$	$-17.29_{\pm 0.49}$	5640	$-17.33_{\pm 0.52}$	6330	$-16.58_{\pm 0.92}$
J1903.2+5540	$-17.57_{\pm 0.28}$	$-17.47_{\pm 0.30}$	$-17.46_{\pm 0.33}$	6030	$-17.83_{\pm 0.43}$	5520	$-17.45_{\pm 0.32}$
J0141.4-0928	$-16.27_{\pm 0.43}$	$-16.16_{\pm 0.47}$	$-15.93_{\pm 0.54}$	5310	$-16.27_{\pm 0.44}$	5820	$-15.64_{\pm 0.64}$
J1048.4+7143	$-15.50_{\pm 0.54}$	$-15.90_{\pm 0.58}$	$-16.71_{\pm 0.62}$	5730	$-17.31_{\pm 0.64}$	6300	$-15.34_{\pm 0.80}$
J2311.0+3425	$-17.66_{\pm 0.59}$	$-18.01_{\pm 0.65}$	$-18.20_{\pm 0.72}$	5550	$-18.20_{\pm 0.72}$	6270	$-16.93_{\pm 1.03}$
J0457.0-2324	$-15.38_{\pm 0.39}$	$-15.51_{\pm 0.47}$	$-15.82_{\pm 0.57}$	5910	$-16.40_{\pm 0.68}$	6690	$-15.14_{\pm 0.99}$
J2243.9+2021	$-17.91_{\pm 0.42}$	$-18.07_{\pm 0.45}$	$-18.14_{\pm 0.48}$	5610	$-18.17_{\pm 0.49}$	6690	$-17.68_{\pm 0.74}$
J0921.6+6216	$-16.53_{\pm 0.58}$	$-16.22_{\pm 0.64}$	$-15.87_{\pm 0.72}$	7440	$-16.21_{\pm 1.76}$	5940	$-15.41_{\pm 0.96}$
J1754.2+3212	$-18.99_{\pm 0.57}$	$-19.09_{\pm 0.69}$	$-19.29_{\pm 0.81}$	5940	$-19.60_{\pm 1.05}$	6960	$-19.04_{\pm 1.50}$
J0228.3-5547	$-16.98_{\pm 0.50}$	$-16.87_{\pm 0.55}$	$-16.82_{\pm 0.60}$	6000	$-17.91_{\pm 0.91}$	6660	$-16.29_{\pm 1.27}$
J0850.0+4855	$-18.27_{\pm 0.80}$	$-18.31_{\pm 0.92}$	$-18.27_{\pm 1.07}$	5370	$-18.32_{\pm 0.87}$	6690	$-17.82_{\pm 2.14}$
J0407.0-3826	$-17.17_{\pm 0.56}$	$-17.06_{\pm 0.63}$	$-16.79_{\pm 0.74}$	5280	$-17.17_{\pm 0.56}$	5820	$-16.47_{\pm 0.88}$
J2158.8-3013	$-16.65_{\pm 0.30}$	$-16.66_{\pm 0.33}$	$-16.46_{\pm 0.35}$	5310	$-16.67_{\pm 0.31}$	5670	$-16.30_{\pm 0.40}$
J0303.4-2407	$-17.71_{\pm 0.48}$	$-17.57_{\pm 0.53}$	$-17.79_{\pm 0.59}$	5910	$-18.17_{\pm 0.76}$	5370	$-17.54_{\pm 0.51}$
J0030.3-4224	$-17.18_{\pm 0.53}$	$-17.22_{\pm 0.58}$	$-17.20_{\pm 0.66}$	5430	$-17.24_{\pm 0.60}$	6360	$-16.55_{\pm 1.10}$
J0217.8+0144	$-17.21_{\pm 0.40}$	$-17.33_{\pm 0.42}$	$-17.48_{\pm 0.45}$	5580	$-17.49_{\pm 0.46}$	6180	$-16.57_{\pm 0.55}$

In Table 10.3, we provide the future forecast results of the LCR periodic sources. First, we present the 1st, 5th and 10th forecast flux values with PI, which correspond to the dates $5280 + 54686 = 59966$ MJD (January 2023), $54000 + 54686 = 60086$ MJD (May 2023) and $5550 + 54686 = 60236$ MJD (October 2023), respectively. Then, we list the estimated low/high state day from start with the forecast flux value with PI. The high state of the emission is of interest as can be helpful to schedule observations and to detect possible flares. However, observing the lower flux values is also important for time-domain analysis in order to record all possible states of the emission and to avoid biasing the LCs.

Some examples of the future forecasting estimation with low/high states are shown in Figure 10.8. We see sources with different linear and period values, sinusoidal and harmonic models, and various uncertainties in future predictions based on the standard deviation of the MCMC parameters.

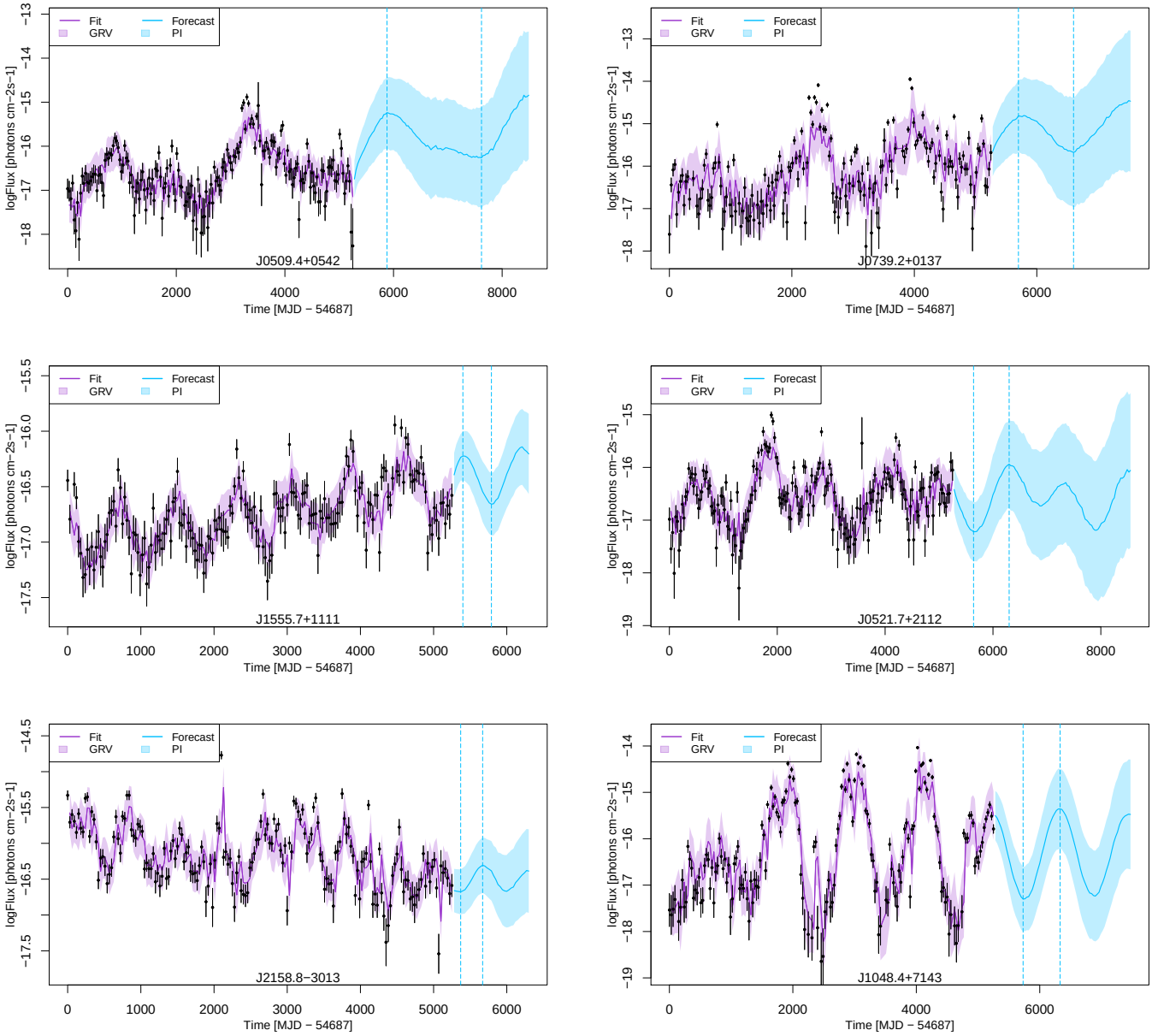

Figure 10.8. MCMC fit with GRV for some of the the LCR periodic sources from Table 10.3 and forecast values with PIsfor the future. The vertical blue line indicates the predicted next low/high state of the source.

10.5 Discussion

In this chapter, we have provided a forecasting method based on the MCMC parametric description of the time-series under analysis to produce and evaluate the predicted future values. First, we have tested the prediction power for the Fermi-LAT AGNs fluxes, including information on their stochastic behavior through a CAR(1) noise model and their deterministic linear and periodic trends, fundamental in understanding the prediction. We have used the simulated N forecast realizations to generate information about the expected error metrics based on the knowledge gain from the MCMC fit of the sources' LCs.

There are not many literature references on LCs forecasting in astrophysics. Some works have implemented artificial neural network to predict LCs for optical data (Zhang & Zou, 2018; Kumar et al., 2023). Goumiri et al. (2022) propose a Gaussian process based method to interpolate and forecast photometric LCs. Also, Kelly et al. (2014) analyzed the PSDs related to CARMA processes and used the stochastic models' information to interpolate and forecast different sets of simulated and observations time series.

Forecasting evaluation, using training and validation data sets, is another valuable tool for looking at periodicity temporal stability. A wrong metrics output in the validation set could imply that the periodic model used to fit the data is not accurate or stable in time. In this case, it may be interesting to check whether the parametric description or the periodic behavior has changed, as for PKS 2155-304. As seen in Table 10.1, the deviation in the next predicted high state is 10%. Now, if we consider the last high state in Figure 10.7 at around 5100 days, the deviation between the prediction and the data rises to 20 %. Such deviation is expected from the LCR temporal stability results from Section 8.3, where we find a significative decrease in the source periodicity from ~ 600 to ~ 355 days.

Time series cross-validation (Hyndman & Athanasopoulos (2021), Section 3.4) is a different evaluation approach that can improve forecasting accuracy. It consists in creating multiple validation sets using only one data point from a specific date. Then, we add more data to the training set, and the validation data point moves one step forward, as illustrated in Figure 10.9.

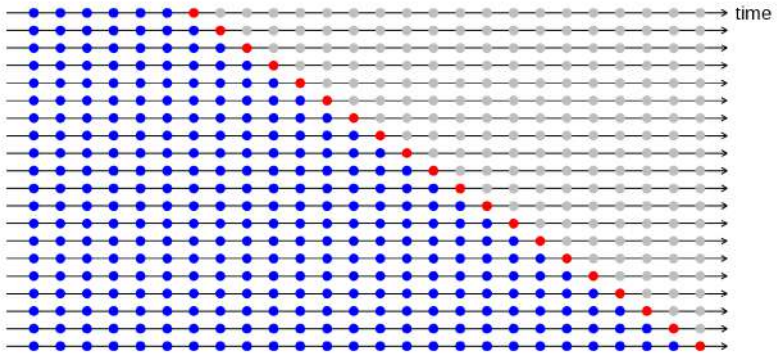

Figure 10.9. (Hyndman & Athanasopoulos, 2021) Time series cross-validation training (blue) and validation (red) data sets.

In order to use this method in our context, we could fix an MCMC model to fit the N training sets and compute the RMSE on the corresponding N validation sets. Then, we can evaluate distinct models looking to minimize the RMSE. This technique is helpful if we are only interested in the next forecast value. In the context of periodic SMBBH, this may not be as useful as determining the following low/high emission state.

Indeed, from our predictions, one of the most valuable knowledge we can get is about the next lowest and highest emission states. Both low and high states, and even intermediate emission states, can help to optimize observation strategies. This is useful to properly sample different astrophysical sources using IACTs, where the observation are planned in advance when the object is expected to be detectable. Also, this information on the low-energy γ-ray emission from Fermi-LAT can be used to make joint multiwavelenght or multi-messenger observations.

The 26th of February 2022, the NASA's time-somain and multi-messenger alert system

General Coordinates Network, emitted a circular (Garrappa et al., 2022) on a Fermi-LAT analysis of observations of the vicinity of the HE neutrino event IC220225A. They found a coincidence enhanced γ-ray emission for the FSRQ PKS 0215+015 (J0217.8+0144), with a high state flux 7 times greater than the average flux for the source. Interestingly, this is a newly detected periodic source from the LCR analysis in Chapter 8, with a value of $\sim$ 1200 days for both MCMC and Agatha results (Tables 8.1 and 8.3). Also, this source is found periodic from the forecasting training set in Section 10.3, which uses data from the Fermi-LAT launch on 2008 to December 2020, and its forecasting is evaluated from January 2021 to January 2023, as illustrated in Figure 10.10.

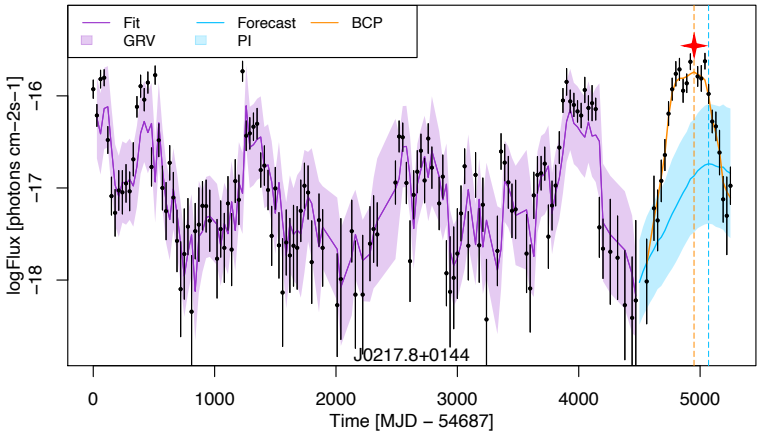

Figure 10.10. MCMC fit with GRV for the training data set and forecast values with PIsfor the validation data set for J0217.8+0144. The vertical blue line indicates the predicted next high state of the source. The vertical orange line indicates the BCP high state of the validation data set. The red star indicates the position of the neutrino event IC220225A.

The next high state for the validation data set was predicted on June 2022, while the actual flaring activity occurred in February with a considerably higher flux than estimated. This is $\sim$ 120 days before the prediction and, compared to the periodicity value, the estimation error is only a 10% (Table 10.1). This means that we could have predicted this neutrino coincidence γ-ray flaring state, considering that the estimation is only for the highest point in the flaring activity, while the enhanced emission is estimated for months, depending on the period value. Now, from the results in the future forecast (Section 10.4), we predict the next flaring state to happen in June 2025.

Similarly, on May 2021, Yusafzai & Garrappa (2021) emitted an Astronomer's Telegram (ATel) regarding the Fermi-LAT detection of a flaring activity from the FSRQ PKS B0027-426 (J0030.3-4224). As in the previous case, J0030.3-4224 is a periodic source from the LCR analysis, with a value of $\sim$ 1800 days for both the MCMC and the spectral cross-check. In Figure 10.11 we present the forecasting evaluation results on the source.

For this source, the predicted high state was for December 2020, $\sim$ 150 days before the flaring notification. Thus, 11% is the difference between the two dates, weighted by the periodicity value. From Figure 10.11, we see that other lower flares occurred before and after the predicted line, coinciding with the upward trend of the detected oscillation. J0030.3-4224 is one of the 137 blazars for which Abbasi et al. (2022) made a stacking analysis searching for astrophysical neutrino production in the first Fermi-LAT low energy catalog (1FLE, Principe et al. (2018)) using 10 yr of IceCube muon–neutrino

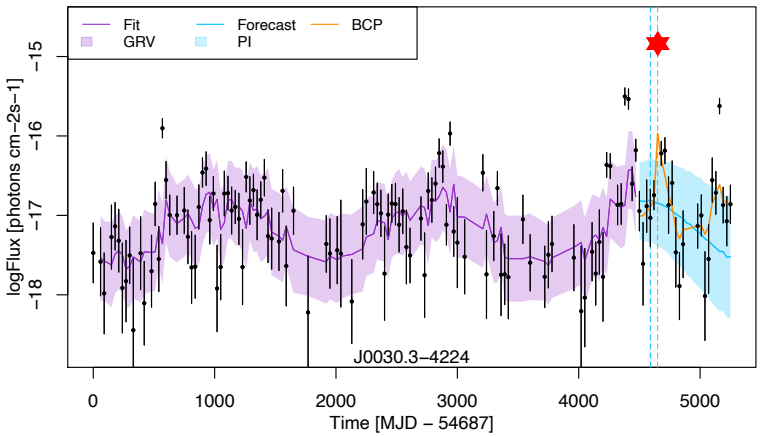

Figure 10.11. MCMC fit with GRV for the training data set and forecast values with PIsfor the validation data set for J0030.3-4224. The vertical blue line indicates the predicted next high state of the source. The vertical orange line indicates the BCP high state of the validation data set. The red star indicates the position of the flaring state from the ATel (Yusafzai & Garrappa, 2021).

data, where they were able to set ULs on the neutrino flux. It may be interesting to prepare observations during the enhanced γ-ray activity of the source. From the future LC values forecast in Section 10.4, we estimate that the next flaring state will happen in January 2026.

For the Fermi-LAT LCR sources analyzed in this work, the forecast values can be assessed on a monthly basis as new data points are added to the source observations. Also, this allows for updating the MCMC analysis by incorporating the new data point and generating forecasts for subsequent observations along with an expectation on the error metrics. In view of the above, for these confirmed periodic sources, a dedicated public catalog in a web page form could be implemented with the results of the stochastic and deterministic parameters describing the LCs and the future forecast values along with low and high states.

Conclusions

Conclusions

The main achievement of this thesis has been the development, testing and application of an MCMC periodicity search algorithm using a time-domain approach for γ-ray regularly and irregularly sampled data. The use of a time-domain approach offers several key advantages. One is the ability to separate the different components within a LC, such as the periodic signal and the stochastic noise. We have shown that adding a periodic component can significantly improve the LC fit to the data compared to AR/CAR noise-only models. Similarly, adding an AR/CAR component usually improves the fit to the data compared to WN models. We have also proven the ability of our method to detect harmonic oscillations and to analyze the evolution of periods and amplitudes over time.

With this, we have contributed to the detection and analysis of new periodic AGN candidates in the Fermi-LAT observations, and we have performed variability and correlation analysis for the H.E.S.S. extragalactic sky. This thesis emphasizes the importance of multi-wavelength analysis and the potential for joint analysis with multiple data origins. While the advantages of the time-domain approach are evident, Chapter 9 also highlights the limitations when analyzing ground-based telescope data. The irregular and limited observations obtained by telescopes like H.E.S.S. pose challenges compared to the continuous scanning capabilities of Fermi-LAT. Nonetheless, despite these limitations, the time-domain analysis can still provide valuable insights and contribute to our understanding of the physical mechanisms behind the HE and VHE emission from astrophysical sources.

Finally, we have demonstrated the ability of the algorithm to predict future behavior. The prediction method is presented with a procedure to evaluate the predictions based on past data already seen. Future emission values are given together with the expected FE based on the MCMC uncertainties for each source. The results for this technique can contribute to our understanding of the long-term behavior of AGN fluxes and provide a valuable tool for observational planning and the study of periodic phenomena.

In Chapter 3 the importance of VLZA observations for the detection of γ-ray emission beyond 100 TeV with IACTs is highlighted. The feasibility of this technique is demonstrated by the detection and analysis of the MGRO J2019+37 region with H.E.S.S. The results obtained with current detector responses indicate that VLZA observations can extend the observing time available for transient phenomena. However, there is still room for improvement in terms of angular resolution, atmospheric correction, and gamma-hadron separation, which can enhance the analysis of the observations. VLZA observations with the current generation of IACTs also serve as a pathfinder for future telescopes aimed at scientific goals beyond γ-ray detection, such as UHE neutrino detection and cosmic-ray mass composition studies.

Chapter 4 focuses on the evaluation and testing of the timing system of the NectarCAM

camera for the CTA. The results show that the camera fulfills the timing requirements of CTA, which is crucial for accurately analyzing the collected data. The experiments demonstrate that the camera achieves a time resolution of less than 1 ns for sufficient light intensities and that the timing offsets between pixels can be effectively reduced through PMT transit time calibration. Additionally, the camera trigger timing accuracy surpasses the CTA requirements. These findings ensure the reliability and precision of the NectarCAM timing system for future CTA observations.

Chapter 6 introduces a novel method for investigating the presence of periodicity in HE source candidates, which is applied in Chapter 7 to a subsample of AGNs from the Fermi-LAT 4FGL catalog. The method effectively separates the stochastic and periodic components in the LC fitting. Of the 27 sources analyzed, 14 were found to exhibit periodic behavior, while the remaining undetected oscillations could potentially be explained by the presence of AR features in the LCs. The study also examines the period and amplitude variations with time of PG 1553+113 and PKS 2155-304, providing insight into the behavior of these sources. The results support the applicability of various physical models and shed light on the nature of periodic flux variations in AGNs.

In chapter 8 we extended the formalism used for period searches by introducing a CAR-based method capable of analyzing irregularly sampled data. Applying this algorithm to the entire AGN population of the Fermi-LAT LCR, we discovered several new periodic sources and confirmed oscillations in others. Cross-validation with the Agatha software spectral method yielded similar results. Comparing the periodicity results obtained with CAR models and previous AR models, consistent periodic signals were found for AGNs appearing in both analyses. However, tighter constraints on MCMC convergence affected the recovery of periodic models for some sources. The time stability of the periodicity is also studied within this formalism, finding a drastic decrease in the period of PKS 2155-304. This result can be valuable to link the source to a theoretical model and to constrain the physical parameters describing the emission mechanisms.

Chapter 9 focuses on the PSD analysis of blazar emission, which provides information about their variability in different time scales. The analysis utilizes the CAR models to describe the stochastic noise behavior and quantify the power of variability at different frequencies. The results demonstrate power-law PSDs with slopes of $\beta \sim 2$ are suitable and consistent with previous studies. We also performed a joint Fermi-LAT and H.E.S.S. periodicity search analysis for PKS 2155-304. With a non-negligible correlation between both HE and VHE fluxes, the algorithm is able to retrieve significant and efficiently the oscillation value resulted in previous chapters and in literature.

Finally, Chapter 10 presents a forecasting method based on MCMC modeling of time-series data. The method allows the prediction of future values by incorporating the stochastic and deterministic trends observed in the data. The forecasting approach is tested on Fermi-LAT AGN fluxes, considering both the stochastic behavior described by a CAR noise model and the deterministic linear and periodic trends. The results demonstrate the predictive power of the method and its potential application in optimizing observing strategies and performing multi-wavelength or multi-messenger observations. The prediction evaluation provides insight into the temporal stability of the periodic behavior and allows the estimation of future high and low emission states.

The above is in turn a starting point for further work in this area. Possible next steps include the analysis of finer sampling resolutions, the use of higher-order AR/CAR models, the inclusion of ULs as constraints, and the investigation of correlations between MCMC

output parameters and known physical variables of AGNs. The algorithm developed in this thesis has versatile applications beyond the study of γ-ray AGNs. It can be used to study the emission from galactic sources, such as binaries or the Galactic Center, in order to discover or improve our knowledge of their periodic behavior, taking into account correlated noise and other trends. Furthermore, the analysis can be extended to include data from other wavelengths, such as optical or radio observations, in order to have a broader understanding of the different astrophysical mechanisms. In addition, the thesis recognizes the importance of future facilities such as the CTA, which will provide more comprehensive and detailed observations for time-domain analysis.

Appendices

Appendix A

Derivation of the models physical variables

Subtracting k times equation 6.4, one gets:

$$\begin{aligned} z &= \phi(t_n) - \textstyle\sum_{j=1}^{k} \beta_j \phi(t_{n-j}) \\ &= (1 - \textstyle\sum_{j=1}^{k} \beta_j + \sum_{j=1}^{k} \beta_j j\delta t)\bar{\phi} + Ct_n(1 - \textstyle\sum_{j=1}^{k} \beta_j) + w(\delta t) + S_n \\ S_n &= \textstyle\sum_j (A_j(\cos(\omega_j t_n) - \sum_{l=1}^{k} \beta_l \cos(\omega_j t_{n-l})) + B_j(\sin(\omega_j t_n) - \sum_{l=1}^{k} \beta_l \sin(\omega_j t_{n-l}))) \end{aligned}$$

with $\delta t = t_1 - t_0$. S_n can be simplified by defining:

$$U_j \exp\{i\psi_j\} = 1 - \sum_{l=1}^{k} \beta_l \exp\{(-il\omega_j \delta t)\} \tag{A.1}$$

with U_j real. Then:

$$S_n = \sum_j (A_j U_j \cos(\omega_j t_n + \psi_j) + B_j U_j \sin(\omega_j t_n + \psi_j)) \tag{A.2}$$

Finally, z can be written as

$$z = \bar{\phi}' + \sum_j (A'_j \cos(\omega_j t_n) + B'_j \sin(\omega_j t_n)) + C' t_n + w(\delta t) \tag{A.3}$$

with

$$\bar{\phi}' = (1 - \textstyle\sum_{j=1}^{k} \beta_j + \sum_{j=1}^{k} \beta_j j\delta t)\bar{\phi} \tag{A.4}$$

$$C' = C(1 - \textstyle\sum_{j=1}^{k} \beta_j) \tag{A.5}$$

$$A'_j = U_j(A_j \cos(\psi_j) + B_j \sin(\psi_j)) \tag{A.6}$$

$$B'_j = U_j(-A_j \sin(\psi_j) + B_j \cos(\psi_j)) \tag{A.7}$$

The time-averaged square amplitude of each oscillating term is

$$Z'^2_j = 1/2(A'^2_j + B'^2_j) = 1/2(U_j)^2(A_j^2 + B_j^2) = (U_j)^2 Z_j^2 \tag{A.8}$$

The transformed amplitude Z' picks up an additional period dependence compared to the physical amplitude Z. Specializing to the case of an AR(1) model with a single period $T = \frac{2\pi}{\omega}$, the ratio of the transformed amplitude to the physical amplitude is

$$\frac{Z'}{Z} = U = \sqrt{(1 + \beta^2 - 2\beta \cos(\frac{2\pi\delta t}{T}))}. \tag{A.9}$$

If $\beta > 0$ as in the OU model (equation 6.11), $\frac{Z'}{Z}$ is a decreasing function of T in the limit of small sampling times ($\delta t/T \ll 1$).

Appendix B

Properties of the gaussian vector N for CAR models

B.1 Components of N in the CAR(2) model

The N_1, N_2 GRVs can be written as Wiener integrals:

$$N_1(t-t_0,\xi_1,\xi_2) = \frac{\sqrt{D}}{\tau^{3/2}(\xi_2-\xi_1)}\int_{t_0}^{t} ds\zeta(\exp\{(\xi_2(t-s))\}-\exp\{(\xi_1(t-s))\}) \quad \text{(B.1)}$$

$$N_2(t-t_0,\xi_1,\xi_2) = \frac{\sqrt{D}}{\tau^{3/2}(\xi_2-\xi_1)}\int_{t_0}^{t} ds\zeta(\xi_2\exp\{(\xi_2(t-s))\}-\xi_1\exp\{(\xi_1(t-s))\}) \quad \text{(B.2)}$$

The expectations of N_1 and N_2 are :

$$E(N_1) = E(N_2) = 0 \quad \text{(B.3)}$$

The components of the covariance matrix of vector of (N_1, N_2) can be computed by applying theorem 2.3.4 of Kuo (2006).

B.1.1 Quasi-periodic oscillations

In this section, ξ_1 and ξ_2 have imaginary parts:

$$\xi_1 = -\lambda - i\omega \quad \text{(B.4)}$$

$$\xi_2 = -\lambda + i\omega. \quad \text{(B.5)}$$

Adapting from equations 6.17 and 6.18, the evolution of ϵ and $\dot{\epsilon}$ becomes is given by:

$$\epsilon(t) = \exp\{-\lambda(t-t_0)\}(\epsilon(t_0)\cos(\omega(t-t_0))+ (\tfrac{\dot{\epsilon}(t_0)}{\omega}+\tfrac{\lambda}{\omega}\epsilon(t_0))\sin(\omega(t-t_0))) + N_1(t-t_0) \quad \text{(B.6)}$$

$$\dot{\epsilon}(t) = \exp\{-\lambda(t-t_0)\}(\dot{\epsilon}(t_0)\cos(\omega(t-t_0))- (\tfrac{\lambda\dot{\epsilon}(t_0)}{\omega}+\tfrac{(\lambda^2+\omega^2)}{\omega}\epsilon(t_0))\sin(\omega(t-t_0))) + N_2(t-t_0) \quad \text{(B.7)}$$

The covariance matrix of the (N_1, N_2) gaussian vector has the following elements:

$$E((N_1)^2) = \tfrac{D}{4\lambda(\lambda^2+\omega^2)\tau^3}(1-\exp\{(-2\lambda(t-t_0))\}(1+ 2\tfrac{\lambda^2}{\omega^2}\sin^2(\omega(t-t_0)) + 2\tfrac{\lambda}{\omega}\sin(\omega(t-t_0))\cos(\omega(t-t_0)))) \quad \text{(B.8)}$$

$$E(N_1N_2) = \frac{D}{2\omega^2\tau^3}\exp\{(-2\lambda(t-t_0))\}\sin^2(\omega(t-t_0)) \tag{B.9}$$

$$\begin{aligned} &E((N_2)^2) = \tfrac{D}{4\lambda\tau^3}(1-\exp\{(-2\lambda(t-t_0))\}(1+ \\ &2\tfrac{\lambda^2}{\omega^2}\sin^2(\omega(t-t_0)) - 2\tfrac{\lambda}{\omega}\sin(\omega(t-t_0))\cos(\omega(t-t_0)))) \end{aligned} \tag{B.10}$$

Equations B.8, B.9, B.10 were first derived by Wang & Uhlenbeck (1945).

B.1.2 Coupled exponentials

If ξ_1 and ξ_2 are real, then, defining λ and $\delta\lambda$ by:

$$-\lambda = \tfrac{1}{2}(\xi_1+\xi_2) \tag{B.11}$$

$$\delta\lambda = \tfrac{1}{2}(\xi_1-\xi_2), \tag{B.12}$$

the components of the covariance matrix of (N_1, N_2) are:

$$\begin{aligned} &E((N_1)^2) = \tfrac{D}{4\lambda(\lambda^2+(\delta\lambda)^2)\tau^3}(1-\exp\{(-2\lambda(t-t_0))\}(1+ \\ &2\tfrac{\lambda^2}{(\delta\lambda)^2}\sinh^2((\delta\lambda)(t-t_0)) + 2\tfrac{\lambda}{(\delta\lambda)}\sinh(\delta\lambda(t-t_0))\cosh(\delta\lambda(t-t_0)))), \end{aligned} \tag{B.13}$$

$$E(N_1N_2) = \frac{D}{2\delta\lambda^2\tau^3}\exp\{(-2\lambda(t-t_0))\}\sinh^2(\delta\lambda(t-t_0)), \tag{B.14}$$

$$\begin{aligned} &E((N_2)^2) = \tfrac{D}{4\lambda\tau^3}(1-\exp\{(-2\lambda(t-t_0))\}(1+ \\ &2\tfrac{\lambda^2}{\delta\lambda^2}\sinh^2(\delta\lambda(t-t_0)) - 2\tfrac{\lambda}{\delta\lambda}\sinh(\delta\lambda(t-t_0))\cosh(\delta\lambda(t-t_0)))) \end{aligned} \tag{B.15}$$

B.2 Components of N in the CAR(N) model

The components of the gaussian vector **N** are given by a generalization of equations B.1 and B.1. Define the p row vectors $y_k = \left(\xi_1^k\exp\{\xi_1(t-s)\}\xi_2^k\exp\{\xi_2(t-s)\}\ldots\xi_p^k\exp\{\xi_p(t-s)\}\right)$ for k from 0 to $(p-1)$. Then the k-th component of vector **N** is:

$$N_k(t-t_0,\xi_1,\xi_2,\ldots,\xi_p) = \int_{t_0}^{t} ds\zeta y_k V^{-1}\mathbf{r} \tag{B.16}$$

and the generic element of the covariance matrix of **N** is:

$$E(N_kN_k') = 1/2\int_{t_0}^{t} ds\left(((y_k)^*V^{-1}\mathbf{r})\left((y_k')V^{-1}\mathbf{r}\right) + \left((y_k)V^{-1}\mathbf{r}\right)\left((y_k'^*)V^{-1}\mathbf{r}\right)\right) \tag{B.17}$$

Bibliography

Abbasi, R., Ackermann, M., Adams, J., et al. 2022, ApJ, 938, 38, doi: 10.3847/1538-4357/ac8de4

Abdalla, H., Abramowski, A., Aharonian, F., et al. 2017, A&A, 598, A39, doi: 10.1051/0004-6361/201629419

Abdalla, H., et al. 2019, Astrophys. J., 870, 93, doi: 10.3847/1538-4357/aaf1c4

Abdalla, H., Aharonian, F., Ait Benkhali, F., et al. 2021, A&A, 653, A152, doi: 10.1051/0004-6361/202140962

Abdalla, H., Abramowski, A., Aharonian, F., et al. 2018, A&A, 612, A1, doi: 10.1051/0004-6361/201732098

Abdo, A. A., Allen, B., Berley, D., et al. 2007a, ApJ, 658, L33, doi: 10.1086/513696

—. 2007b, ApJ, 664, L91, doi: 10.1086/520717

Abdo, A. A., Ackermann, M., Ajello, M., et al. 2010a, A&A, 524, A75, doi: 10.1051/0004-6361/201014458

Abdo, A. A., Ackermann, M., Agudo, I., et al. 2010b, ApJ, 716, 30, doi: 10.1088/0004-637X/716/1/30

Abdollahi, S., Acero, F., Ackermann, M., et al. 2020, ApJS, 247, 33, doi: 10.3847/1538-4365/ab6bcb

Abdollahi, S., Acero, F., Baldini, L., et al. 2022, ApJS, 260, 53, doi: 10.3847/1538-4365/ac6751

Abdollahi, S., Ajello, M., Baldini, L., et al. 2023, ApJS, 265, 31, doi: 10.3847/1538-4365/acbb6a

Abeysekara, A. U., Albert, A., Alfaro, R., et al. 2017, ApJ, 843, 40, doi: 10.3847/1538-4357/aa7556

Abeysekara, A. U., Archer, A., Aune, T., et al. 2018, ApJ, 861, 134, doi: 10.3847/1538-4357/aac4a2

Abeysekara, A. U., Albert, A., Alfaro, R., et al. 2020, Phys. Rev. Lett., 124, 021102, doi: 10.1103/PhysRevLett.124.021102

Abramowski, A., Aharonian, F., Benkhali, F. A., et al. 2016, Nature, 531, 476, doi: 10.1038/nature17147

Acciari, V. A., Ansoldi, S., Antonelli, L. A., et al. 2020a, A&A, 635, A158, doi: 10.1051/0004-6361/201936899

—. 2020b, A&A, 642, A190, doi: 10.1051/0004-6361/201936896

Acharya, B. S., Actis, M., Aghajani, T., et al. 2013, Astroparticle Physics, 43, 3, doi: 10.1016/j.astropartphys.2013.01.007

Acharya, B. S., et al. 2018, Science with the Cherenkov Telescope Array (WSP), doi: 10.1142/10986

Acharyya, A., & Sadun, A. C. 2023, Galaxies, 11, doi: 10.3390/galaxies11040081

Ackermann, M., Ajello, M., Albert, A., et al. 2014, The Astrophysical Journal, 787, 15, doi: 10.1088/0004-637X/787/1/15

Ackermann, M., Ajello, M., Albert, A., et al. 2014, Science, 345, 554, doi: 10.1126/science.1253947

—. 2015, ApJ, 813, L41, doi: 10.1088/2041-8205/813/2/L41

Adams, C. B., Benbow, W., Brill, A., et al. 2021, ApJ, 913, 115, doi: 10.3847/1538-4357/abf926

Aharonian, F., Akhperjanian, A. G., Aye, K. M., et al. 2004, Astroparticle Physics, 22, 109, doi: 10.1016/j.astropartphys.2004.06.006

—. 2005, A&A, 437, 95, doi: 10.1051/0004-6361:20053050

Aharonian, F., Akhperjanian, A. G., Bazer-Bachi, A. R., et al. 2006, A&A, 457, 899, doi: 10.1051/0004-6361:20065351

—. 2007, ApJ, 664, L71, doi: 10.1086/520635

Ahnen, M. L., et al. 2018, Astropart. Phys., 102, 77, doi: 10.1016/j.astropartphys.2018.05.002

Ait Benkhali, F., Hofmann, W., Rieger, F. M., & Chakraborty, N. 2020, A&A, 634, A120, doi: 10.1051/0004-6361/201935117

Ajello, M., Arimoto, M., Axelsson, M., et al. 2019, ApJ, 878, 52, doi: 10.3847/1538-4357/ab1d4e

Akaike, H. 1974, IEEE Transactions on Automatic Control, 19, 716, doi: 10.1109/TAC.1974.1100705

Albert, A., et al. 2021, Astrophys. J., 911, 143, doi: 10.3847/1538-4357/abecda

Aliu, E., Aune, T., Behera, B., et al. 2014, ApJ, 788, 78, doi: 10.1088/0004-637X/788/1/78

Anderson, T. W., & Darling, D. A. 1954, Journal of the American Statistical Association, 49, 765, doi: 10.1080/01621459.1954.10501232

Ashton, T., Backes, M., Balzer, A., et al. 2020, Astroparticle Physics, 118, 102425, doi: 10.1016/j.astropartphys.2019.102425

Atwood, W., Albert, A., Baldini, L., et al. 2013, arXiv e-prints, arXiv:1303.3514. https://arxiv.org/abs/1303.3514

Atwood, W. B., Abdo, A. A., Ackermann, M., et al. 2009, ApJ, 697, 1071, doi: 10.1088/0004-637X/697/2/1071

Ballet, J., Burnett, T. H., Digel, S. W., & Lott, B. 2020, arXiv e-prints, arXiv:2005.11208, doi: 10.48550/arXiv.2005.11208

Balzer, A., Füßling, M., Gajdus, M., et al. 2014, Astroparticle Physics, 54, 67, doi: 10.1016/j.astropartphys.2013.11.007

Barry, D., & Hartigan, J. A. 1993, Journal of the American Statistical Association, 88, 309. http://www.jstor.org/stable/2290726

Bi, B., Barcelo, M., Bauer, C., et al. 2022, in 37th International Cosmic Ray Conference, 743, doi: 10.22323/1.395.0743

Bonnoli, G., Ghisellini, G., Foschini, L., Tavecchio, F., & Ghirlanda, G. 2011, MNRAS, 410, 368, doi: 10.1111/j.1365-2966.2010.17450.x

Bradascio, F., Rueda, H., Brun, F., et al. 2022, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 12182, Ground-based and Airborne Telescopes IX, ed. H. K. Marshall, J. Spyromilio, & T. Usuda, 121823B, doi: 10.1117/12.2629386

Bradascio, F., Rueda, H., Barrio, J., et al. 2023, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 1054, 168398, doi: https://doi.org/10.1016/j.nima.2023.168398

Brockwell, P. J., & Davis, R. A. 1991, Time Series: Theory and Methods (New York: Springer-Verlag)

—. 2016, Introduction to Time Series and Forecasting, Springer Texts in Statistics (Springer International Publishing), doi: 10.1007/978-3-319-29854-2

Brockwell, P. J., & Marquardt, T. 2005, Statistica Sinica, 15, 477. http://www.jstor.org/stable/24307365

Bruel, P., Burnett, T. H., Digel, S. W., et al. 2018, arXiv e-prints, arXiv:1810.11394, doi: 10.48550/arXiv.1810.11394

Camenzind, M., & Krockenberger, M. 1992, A&A, 255, 59

Cao, Z., Aharonian, F. A., An, Q., et al. 2021, Nature, 594, 33, doi: 10.1038/s41586-021-03498-z

Caproni, A., Abraham, Z., & Monteiro, H. 2013, MNRAS, 428, 280, doi: 10.1093/mnras/sts014

Casella, G., & George, E. I. 1992, The American Statistician, 46, 167. http://www.jstor.org/stable/2685208

Cash, W. 1979, ApJ, 228, 939, doi: 10.1086/156922

Cavaliere, A., Tavani, M., & Vittorini, V. 2017, ApJ, 836, 220, doi: 10.3847/1538-4357/836/2/220

Champion, C., et al. 2018, in 16th International Conference on Accelerator and Large Experimental Physics Control Systems, TUPHA090, doi: 10.18429/JACoW-ICALEPCS2017-TUPHA090

Chevalier, J., Sanchez, D. A., Serpico, P. D., Lenain, J. P., & Maurin, G. 2019, MNRAS, 484, 749, doi: 10.1093/mnras/stz027

Covino, S., Landoni, M., Sandrinelli, A., & Treves, A. 2020, ApJ, 895, 122, doi: 10.3847/1538-4357/ab8bd4

Das, A. K., Prince, R., Gupta, A. C., & Kushwaha, P. 2023, ApJ, 950, 173, doi: 10.3847/1538-4357/acd17f

de Naurois, M., & Mazin, D. 2015, Comptes Rendus Physique, 16, 610, doi: 10.1016/j.crhy.2015.08.011

de Naurois, M., & Rolland, L. 2009, Astroparticle Physics, 32, 231, doi: 10.1016/j.astropartphys.2009.09.001

Degrange, B., & Fontaine, G. 2015, Comptes Rendus Physique, 16, 587, doi: 10.1016/j.crhy.2015.07.003

Delagnes, E., et al. 2011, in 2011 IEEE Nuclear Science Symposium Conference Record, 1457–1462, doi: 10.1109/NSSMIC.2011.6154348

Dermer, C. D., Finke, J. D., & Menon, G. 2008, arXiv e-prints, arXiv:0810.1055. https://arxiv.org/abs/0810.1055

di Sciascio, G., & Lhaaso Collaboration. 2016, Nuclear and Particle Physics Proceedings, 279-281, 166, doi: 10.1016/j.nuclphysbps.2016.10.024

Diehl, R. 2001, Gamma-Ray Production and Absorption Processes, ed. V. Schönfelder (Berlin, Heidelberg: Springer Berlin Heidelberg), 9–25, doi: 10.1007/978-3-662-04593-0_2

Durbin, J., & Koopman, S. 2001, Time series analysis by state space methods (Oxford; New York: Oxford University Press). http://www.worldcat.org/search?qt=worldcat_org_all&q=0198523548

Efron, B., & Hastie, T. 2016, Computer Age Statistical Inference: Algorithms, Evidence, and Data Science, 1st edn. (USA: Cambridge University Press)

Feigelson, E. D., Babu, G. J., & Caceres, G. A. 2018, Frontiers in Physics, 6, 80, doi: 10.3389/fphy.2018.00080

Feng, F., Tuomi, M., & Jones, H. R. A. 2017, MNRAS, 470, 4794, doi: 10.1093/mnras/stx1126

Feng, F., Tuomi, M., Jones, H. R. A., Butler, R. P., & Vogt, S. 2016, Monthly Notices of the Royal Astronomical Society, 461, 2440, doi: 10.1093/mnras/stw1478

Foster, G. 1996, AJ, 112, 1709, doi: 10.1086/118137

Frigessi, A., & Heidergott, B. 2011, Markov Chains (Berlin, Heidelberg: Springer Berlin Heidelberg), 772–775, doi: 10.1007/978-3-642-04898-2_347

F.R.S., S. E. R., & da C. Andrade B.Sc. Ph.D., E. N. 1914, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 28, 263, doi: 10.1080/14786440808635207

Gao, Q.-G., Lu, F.-W., Qin, L.-h., et al. 2023, ApJ, 945, 146, doi: 10.3847/1538-4357/acbe3e

Garrappa, S., Buson, S., Cheung, C. C., Sinapius, J., & Fermi-LAT Collaboration. 2022, GRB Coordinates Network, 31653, 1

Gelman, A., & Rubin, D. B. 1992, Statistical Science, 7, 457 , doi: 10.1214/ss/1177011136

Gerward, L., & Rassat, A. 2000, Comptes Rendus de l'Académie des Sciences - Series IV - Physics, 1, 965, doi: https://doi.org/10.1016/S1296-2147(00)01094-5

Ginsburg, A., Sipőcz, B. M., Brasseur, C. E., et al. 2019, AJ, 157, 98, doi: 10.3847/1538-3881/aafc33

Glicenstein, J.-F. 2016, PoS, ICRC2015, 937, doi: 10.22323/1.236.0937

Gong, Y., Zhou, L., Yuan, M., et al. 2022, ApJ, 931, 168, doi: 10.3847/1538-4357/ac6c8c

Gordon, Y. A., O'Dea, C. P., Baum, S. A., et al. 2023, ApJ, 948, L9, doi: 10.3847/2041-8213/accf0a

Goumiri, I. R., Dunton, A. M., Muyskens, A. L., Priest, B. W., & Armstrong, R. E. 2022, Light curve completion and forecasting using fast and scalable Gaussian processes (MuyGPs). https://arxiv.org/abs/2208.14592

Goyal, A. 2019, Galaxies, 7, 73, doi: 10.3390/galaxies7030073

Goyal, A., Stawarz, Ł., Ostrowski, M., et al. 2017, ApJ, 837, 127, doi: 10.3847/1538-4357/aa6000

Goyal, A., Stawarz, Ł., Zola, S., et al. 2018, ApJ, 863, 175, doi: 10.3847/1538-4357/aad2de

Gracia, J., Peitz, J., Keller, C., & Camenzind, M. 2003, MNRAS, 344, 468, doi: 10.1046/j.1365-8711.2003.06832.x

Gu, M. F., & Ai, Y. L. 2011, A&A, 534, A59, doi: 10.1051/0004-6361/201117467

Heck, D., Pierog, T., & Knapp, J. 2012, CORSIKA: An Air Shower Simulation Program, Astrophysics Source Code Library, record ascl:1202.006. http://ascl.net/1202.006

Hénault, F., et al. 2013, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8834, Nonimaging Optics: Efficient Design for Illumination and Solar Concentration X, ed. R. Winston & J. Gordon, 883405, doi: 10.1117/12.2024049

Hess, V. 2018, arXiv e-prints, arXiv:1808.02927, doi: 10.48550/arXiv.1808.02927

Hillas, A. M. 1985, in International Cosmic Ray Conference, Vol. 3, 19th International Cosmic Ray Conference (ICRC19), Volume 3, 445

Hinton, J. A., & HESS Collaboration. 2004, New A Rev., 48, 331, doi: 10.1016/j.newar.2003.12.004

Hoffmann, D., et al. 2017, Journal of Physics: Conference Series, 898, 032015, doi: 10.1088/1742-6596/898/3/032015

Holch, T. L., Leuschner, F., Schäfer, J., & Steinmassl, S. 2022, J. Phys. Conf. Ser., 2398, 012017, doi: 10.1088/1742-6596/2398/1/012017

Holder, J. 2015, arXiv e-prints, arXiv:1510.05675, doi: 10.48550/arXiv.1510.05675

Hyndman, R., & Athanasopoulos, G. 2021, Forecasting: Principles and Practice, 3rd edn. (Australia: OTexts)

James, F., & Roos, M. 1975, Computer Physics Communications, 10, 343, doi: 10.1016/0010-4655(75)90039-9

Kelly, B. C., Becker, A. C., Sobolewska, M., Siemiginowska, A., & Uttley, P. 2014, ApJ, 788, 33, doi: 10.1088/0004-637X/788/1/33

Kruschke, J. K. 2015, in Doing Bayesian Data Analysis (Second Edition) (Boston: Academic Press), 143–191, doi: https://doi.org/10.1016/B978-0-12-405888-0.00007-6

Kumar, N., Bhardwaj, A., Singh, H. P., et al. 2023, MNRAS, 522, 1504, doi: 10.1093/mnras/stad937

Kuo, H. 2006, Introduction to Stochastic Integration, Universitext (Springer New York). https://books.google.fr/books?id=hrRpqLYOsyAC

L'Annunziata, M. F. 2007, in Radioactivity, ed. M. F. L'Annunziata (Amsterdam: Elsevier Science B.V.), 187–215, doi: https://doi.org/10.1016/B978-044452715-8.50006-2

Leo, W. R. 1994, Techniques for nuclear and particle physics experiments: a how-to approach; 2nd ed. (Berlin: Springer), doi: 10.1007/978-3-642-57920-2

Li, H. Z., Jiang, Y. G., Yi, T. F., et al. 2018, Ap&SS, 363, 45, doi: 10.1007/s10509-018-3256-7

Li, T. P., & Ma, Y. Q. 1983, Astrophys. J., 272, 317, doi: 10.1086/161295

Lindfors, E. J., Hovatta, T., Nilsson, K., et al. 2016, A&A, 593, A98, doi: 10.1051/0004-6361/201628420

Liodakis, I., Romani, R. W., Filippenko, A. V., et al. 2018, MNRAS, 480, 5517, doi: 10.1093/mnras/sty2264

Liodakis, I., Romani, R. W., Filippenko, A. V., Kocevski, D., & Zheng, W. 2019, ApJ, 880, 32, doi: 10.3847/1538-4357/ab26b7

Liu, F. K., Zhao, G., & Wu, X.-B. 2006, ApJ, 650, 749, doi: 10.1086/507267

Liu, H., & Lhaaso Collaboration. 2022, in 37th International Cosmic Ray Conference, 245, doi: 10.22323/1.395.0245

Lomb, N. R. 1976, Ap&SS, 39, 447, doi: 10.1007/BF00648343

Longair, M. S. 2011, High Energy Astrophysics

Lypova, I., et al. 2021, PoS, ICRC2021, 704, doi: 10.22323/1.395.0704

Mattox, J. R., Bertsch, D. L., Chiang, J., et al. 1996, ApJ, 461, 396, doi: 10.1086/177068

Max-Moerbeck, W., Hovatta, T., Richards, J. L., et al. 2014, MNRAS, 445, 428, doi: 10.1093/mnras/stu1749

Meegan, C., Lichti, G., Bhat, P. N., et al. 2009, ApJ, 702, 791, doi: 10.1088/0004-637X/702/1/791

Mirzoyan, R. 2004, European Physical Journal C, 33, 938, doi: 10.1140/epjcd/s2004-03-1630-2

Mohan, P., & Mangalam, A. 2015, The Astrophysical Journal, 805, 91, doi: 10.1088/0004-637X/805/2/91

Moreno, J., Vogeley, M. S., Richards, G. T., & Yu, W. 2019, PASP, 131, 063001, doi: 10.1088/1538-3873/ab1597

Murtaugh, P. A. 2014, Ecology, 95, 611, doi: https://doi.org/10.1890/13-0590.1

Neronov, A., Semikoz, D. V., Vovk, I., & Mirzoyan, R. 2016, Phys. Rev. D, 94, 123018, doi: 10.1103/PhysRevD.94.123018

Netzer, H. 2015, ARA&A, 53, 365, doi: 10.1146/annurev-astro-082214-122302

Nolan, P. L., Tompkins, W. F., Grenier, I. A., & Michelson, P. F. 2003, ApJ, 597, 615, doi: 10.1086/378353

Olivera-Nieto, L., et al. 2021, PoS, ICRC2021, 754, doi: 10.22323/1.395.0754

Otero-Santos, J., Acosta-Pulido, J. A., Becerra González, J., et al. 2020, MNRAS, 492, 5524, doi: 10.1093/mnras/staa134

Parsons, R. D., & Hinton, J. A. 2014, Astropart. Phys., 56, 26, doi: 10.1016/j.astropartphys.2014.03.002

Peñil, P., Domínguez, A., Buson, S., et al. 2020, ApJ, 896, 134, doi: 10.3847/1538-4357/ab910d

Pei, Z., Fan, J., Yang, J., Huang, D., & Li, Z. 2022, ApJ, 925, 97, doi: 10.3847/1538-4357/ac3aeb

Plummer, M. 2012, JAGS: Just Another Gibbs Sampler. http://ascl.net/1209.002

Principe, G., Malyshev, D., Ballet, J., & Funk, S. 2018, A&A, 618, A22, doi: 10.1051/0004-6361/201833116

Prokhorov, D. A., & Moraghan, A. 2017, MNRAS, 471, 3036, doi: 10.1093/mnras/stx1742

Ren, H. X., Cerruti, M., & Sahakyan, N. 2023, A&A, 672, A86, doi: 10.1051/0004-6361/202244754

Rieger, F. M. 2004, ApJ, 615, L5, doi: 10.1086/426018

Rieger, F. M., & Volpe, F. 2010, A&A, 520, A23, doi: 10.1051/0004-6361/201014273

Robert, C. 2007, The Bayesian Choice From Decision-Theoretic Foundations to Computational Implementation, 2nd edn., Springer Texts in Statistics (New York, NY: Springer New York)

Romero, G. E., Fan, J.-H., & Nuza, S. E. 2003, Chinese J. Astron. Astrophys., 3, 513, doi: 10.1088/1009-9271/3/6/513

Rueda, H., Glicenstein, J.-F., & Brun, F. 2022, ApJ, 934, 6, doi: 10.3847/1538-4357/ac771c

Sanchez, D. A., & Deil, C. 2013, in International Cosmic Ray Conference, Vol. 33, International Cosmic Ray Conference, 2784. https://arxiv.org/abs/1307.4534

Sandrinelli, A., Covino, S., Dotti, M., & Treves, A. 2016a, AJ, 151, 54, doi: 10.3847/0004-6256/151/3/54

Sandrinelli, A., Covino, S., & Treves, A. 2014, ApJ, 793, L1, doi: 10.1088/2041-8205/793/1/L1

—. 2016b, ApJ, 820, 20, doi: 10.3847/0004-637X/820/1/20

Sandrinelli, A., Covino, S., Treves, A., et al. 2018, A&A, 615, A118, doi: 10.1051/0004-6361/201732550

—. 2017, A&A, 600, A132, doi: 10.1051/0004-6361/201630288

Scargle, J. D. 1982, ApJ, 263, 835, doi: 10.1086/160554

Schwanke, U., et al. 2015, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 782, 92, doi: https://doi.org/10.1016/j.nima.2015.01.096

Shannon, C. E. 1948, Bell System Technical Journal, 27, 379, doi: https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Shapiro, S. L., & Teukolsky, S. A. 1983, Black holes, white dwarfs and neutron stars. The physics of compact objects, doi: 10.1002/9783527617661

Sobacchi, E., Sormani, M. C., & Stamerra, A. 2017, MNRAS, 465, 161, doi: 10.1093/mnras/stw2684

Sommers, P., & Elbert, J. W. 1987, Journal of Physics G Nuclear Physics, 13, 553, doi: 10.1088/0305-4616/13/4/019

Tarnopolski, M., Żywucka, N., Marchenko, V., & Pascual-Granado, J. 2020, ApJS, 250, 1, doi: 10.3847/1538-4365/aba2c7

Tatischeff, V., & Hernanz, M. 2007, ApJ, 663, L101, doi: 10.1086/520049

Tavani, M., Cavaliere, A., Munar-Adrover, P., & Argan, A. 2018, ApJ, 854, 11, doi: 10.3847/1538-4357/aaa3f4

Tavernier, T., et al. 2020, PoS, ICRC2019, 805, doi: 10.22323/1.358.0805

Tejedor, L. A., Barrio, J. A., Peñil, P., et al. 2022, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 1027, 166058, doi: https://doi.org/10.1016/j.nima.2021.166058

Tejedor, L. A., et al. 2013, IEEE Transactions on Nuclear Science, 60, 2905, doi: 10.1109/TNS.2013.2271300

Thiersen, H., Zacharias, M., & Böttcher, M. 2019, Galaxies, 7, 35, doi: 10.3390/galaxies7010035

Torrence, C., & Compo, G. P. 1998, Bulletin of the American Meteorological Society, 79, 61, doi: 10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2

Tsiahina, A., et al. 2021, Nuclear Instruments and Methods in Physics Research A, 1007, 165413, doi: 10.1016/j.nima.2021.165413

Uhlenbeck, G. E., & Ornstein, L. S. 1930, Phys. Rev., 36, 823, doi: 10.1103/PhysRev.36.823

Urry, C. M., & Padovani, P. 1995, PASP, 107, 803, doi: 10.1086/133630

VanderPlas, J. T. 2018, ApJS, 236, 16, doi: 10.3847/1538-4365/aab766

Wang, G. G., Cai, J. T., & Fan, J. H. 2022, ApJ, 929, 130, doi: 10.3847/1538-4357/ac5b08

Wang, M. C., & Uhlenbeck, G. E. 1945, Reviews of Modern Physics, 17, 323, doi: 10.1103/RevModPhys.17.323

Wang, X., & Emerson, J. W. 2015, Bayesian Change Point Analysis of Linear Models on Graphs. https://arxiv.org/abs/1509.00817

Weekes, T. C., Badran, H., Biller, S. D., et al. 2002, Astroparticle Physics, 17, 221, doi: 10.1016/S0927-6505(01)00152-9

Westerhoff, S. 2014, Advances in Space Research, 53, 1492, doi: 10.1016/j.asr.2013.03.016

Yang, S., Yan, D., Zhang, P., Dai, B., & Zhang, L. 2021, ApJ, 907, 105, doi: 10.3847/1538-4357/abcbff

Yodh, G. B. 1996, Space Sci. Rev., 75, 199, doi: 10.1007/BF00195035

Yusafzai, A., & Garrappa, S. 2021, The Astronomer's Telegram, 14648, 1

Zanotti, O., Rezzolla, L., & Font, J. A. 2003, MNRAS, 341, 832, doi: 10.1046/j.1365-8711.2003.06474.x

Zhang, B.-K., Zhao, X.-Y., Wang, C.-X., & Dai, B.-Z. 2014, Research in Astronomy and Astrophysics, 14, 933, doi: 10.1088/1674-4527/14/8/004

Zhang, H., Fang, K., Li, H., et al. 2019, ApJ, 876, 109, doi: 10.3847/1538-4357/ab158d

Zhang, H., Yan, D., Zhang, P., Yang, S., & Zhang, L. 2021, ApJ, 919, 58, doi: 10.3847/1538-4357/ac0cf0

Zhang, P.-f., Yan, D.-h., Liao, N.-h., & Wang, J.-c. 2017a, ApJ, 835, 260, doi: 10.3847/1538-4357/835/2/260

Zhang, P.-f., Yan, D.-h., Liao, N.-h., et al. 2017b, ApJ, 842, 10, doi: 10.3847/1538-4357/aa7465

Zhang, P.-F., Yan, D.-H., Zhou, J.-N., et al. 2017c, ApJ, 845, 82, doi: 10.3847/1538-4357/aa7ecd

Zhang, R., & Zou, Q. 2018, in Journal of Physics Conference Series, Vol. 1061, Journal of Physics Conference Series, 012012, doi: 10.1088/1742-6596/1061/1/012012

Zhou, J., Wang, Z., Chen, L., et al. 2018, Nature Communications, 9, 4599, doi: 10.1038/s41467-018-07103-2