Studying the gamma-ray emission of the blazar Mrk 421, an 11 year data set from the Fermi-LAT space telescope.

William J Ryan

October 2019

Abstract

The Blazar Mrk 421 is the brightest HBL (High-frequency peaked BL Lac) in the sky and the only HBL with enough gamma-ray emission in its average state to build a daily (or bi-daily) light-curve from the Fermi-LAT telescope. With 11 years worth of data, 3 more years than the most recent Fermi catalog 4FGL, we can define with high accuracy the gamma-ray spectral shape of Mrk 421. I used the 'Binned Likelihood Analysis' technique to statistically compare four models, Log-parabola and Power-law, with and without an exponential cutoff over a spectrum spanning from 100 MeV to 513 GeV. I will be presenting the model comparisons and the one statistically favored. The final goal of this analysis is to create a day binned light-curve to study variability patterns of the source which may be linked to successive shocks accelerating particles.

1 An Introduction To Blazars

At the center of most galaxy lies a super massive black hole (SMBH) whose mass ranges from 10^6 to 10^{10} times the mass of our sun. Some of these objects have accretion disks that heat up as matter falls into the SMBH. This reaction often produces luminosities greater than the combined emission of the host galaxy. We call those SMBH's that actively absorb accretion disk material Active Galactic Nuclei (AGN).

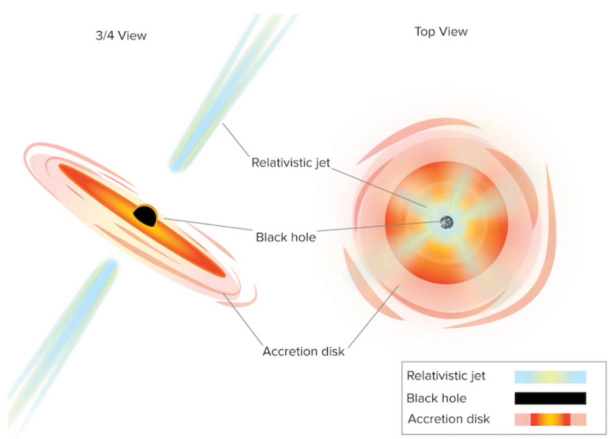


Figure 1: Left: Jetted AGN. Right: Blazar, with jet pointed at observer. (Credit: Sophia Dagnello)

About twenty percent of AGNs have jets. Jets are formed either by a SMBH's Lorentz forces or other forces within the accretion disk. The ultimate cause is still up for debate within the astrophysics community however, once a charged particle enters a jet it will produce synchrotron radiation as it interacts with the local magnetic field. A portion of this photon radiation escapes the emission zone while remaining photons interact with high energy charged particles through inverse Compton scattering which produce higher energy emission mostly in gamma ray. So a jetted AGN typically has a double bump feature in its SED as seen in fig. 2. The first bump is called the Synchrotron peak which reaches anywhere from infrared to x-ray. And the second bump is called the Inverse-Compton peak which in the case of mrk 421 reaches gamma ray. When a jet is pointed towards Earth within a few degrees[1] we call the AGN a Blazar. Should the Blazar have a Synchrotron peak in x-ray then it is called a High-frequency-peaked BL Lac[2] (HBL).

The aforementioned two bump model works under the assumption that bundles of particles are moving with a high Lorentz factor (close to the speed of light). However, most HBLs show stationary or low-speed very long baseline interferometry (VLBI[3]) radio features (radio knots). Our model requires fast moving particles but our telescopes see slow or stationary radio knots so we have an apparent contradiction.

Recollimation shocks provide a possible answer to this conundrum. When material in a gas travels faster than a wave can propagate a shock wave is formed. In a Blazar the pressure difference between the fast plasma in the jet and the stationary plasma outside the jet form a shock wave at their boundary. This shock will rebound inwards towards a single point, called a node, and rebound back towards the plasma boundary. This process repeats until the particles no longer have enough energy to propagate another shock. These Recollimation shocks can be seen on earth from any rocket or jet plane exhaust or, in figure 3, in air pumped out of a water bottle.

A particle passing through a shock will rebound a number of times before escaping and traveling to the next node. For every rebound a particle must accelerate through a magnetic field so radiation is produced

[1] Approximately less than 10 degrees

[2] BL lac: BL lac stands for BL Lacertae. This is a star categorizing scheme made by Cuno Hoffmeister in 1929. Lacertae refers to a constellation and BL is a labeling scheme for all the stars in that constellation. It was later discovered that BL Lacertae was a Blazar so the name was used to categorize Blazars similar to BL Lacertae.

[3] A radio signal from an astronomical source is collected from multiple telescopes on Earth. The time difference between detections can be used to calculate the distance between telescopes. With proper timing the data can be combined forming a final telescope with effective size as big as the max distance between them.

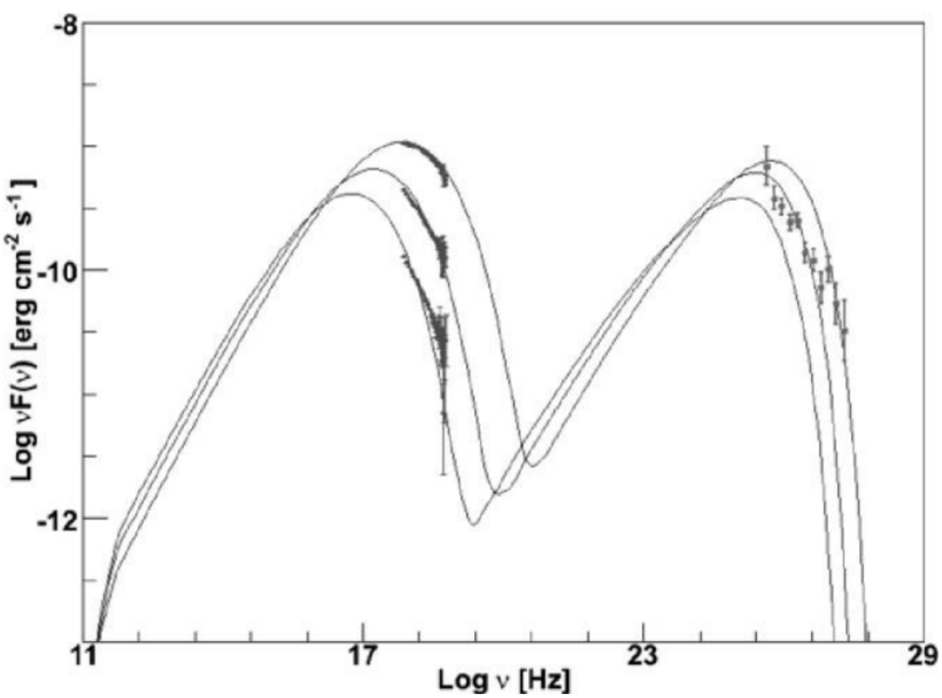


Figure 2: An emission model for MRK 421 taken from SOURCE

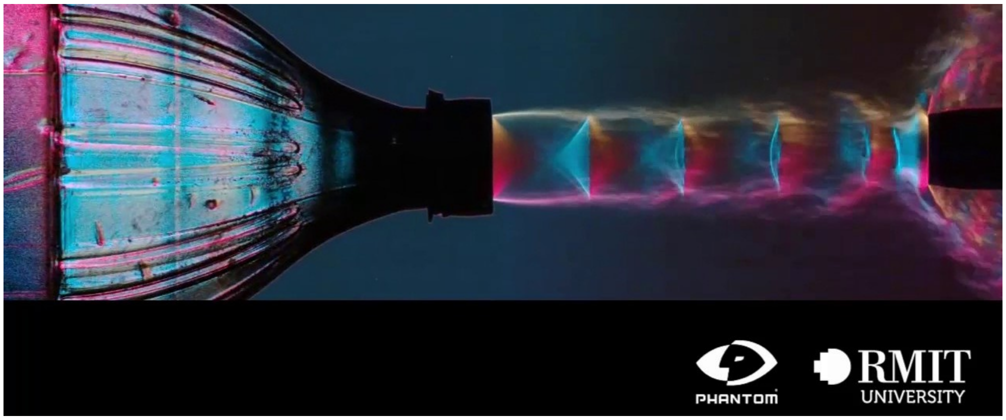


Figure 3: Recollimation shocks as seen on earth

but only at the nodes. If these nodes correspond to the radio knots we see from HBLs then we have a model that solves the previous conundrum.

As Markarian 421 (Mrk 421) is the only HBL with stationary radiation nodes and enough gamma-ray emission in its average state to produce a light-curve, it is the prime target for a full spectrum analysis. A previous analysis of Mrk 421 in the x-ray regime by Olivier Hervet pointed towards the validity of Recollimation shocks, so my gamma ray analysis will help evolve the analysis.

2 Fermi Analysis

Producing a lightcurve for MRK 421 will allow us to probe for possible patterns. In the subsequent sections I discuss Fermi's suite of python tools, Fermipy, for data collection and analysis.

2.1 Fermi Large Area Telescope

The Fermi Large Area telescope (Fermi-LAT) is a satellite telescope that detects and records gamma rays. Given gamma rays do not refract off optical mirrors, the Fermi-LAT requires a different detection setup. Fermi-LAT uses sandwiches of tungsten sheets and silicon detectors. A gamma ray will interact with a tungsten atom and create a positron electron pair whose trajectory can be mapped by the silicon detectors below. At the bottom of the detector is a calorimeter measuring the remaining energy of the particles. Adding the calorimeter energy and the detectors currents gives us the gamma ray's original energy. And, an anti-coincidence shield is placed at the opening of the detector in order to prevent any charged particles giving false data. The satellite itself orbits the Earth once every 96 minutes and can survey the entire sky every 2 orbits. When a strong flare in activity is detected, the telescope will point towards the source to collect extra data for a few hours.

2.2 Pre-Processing

All of Fermi-LAT's data is available on the official Fermi website. After inputting a source name, position, energy range, time range, and Region Of Interest (ROI) the Fermi website will provide a downloadable link containing gamma ray events and telescope positional data within the ROI. After acquiring the data we can start the analysis by running a fermipy script. The analysis begins by compiling a counts map as a sanity check and potential source identifier. This sums the photon energies per spacial bin creating a visual representation of the sky where our main source and secondary sources lie. The counts map for MRK 421 can be seen in fig. 5 below.

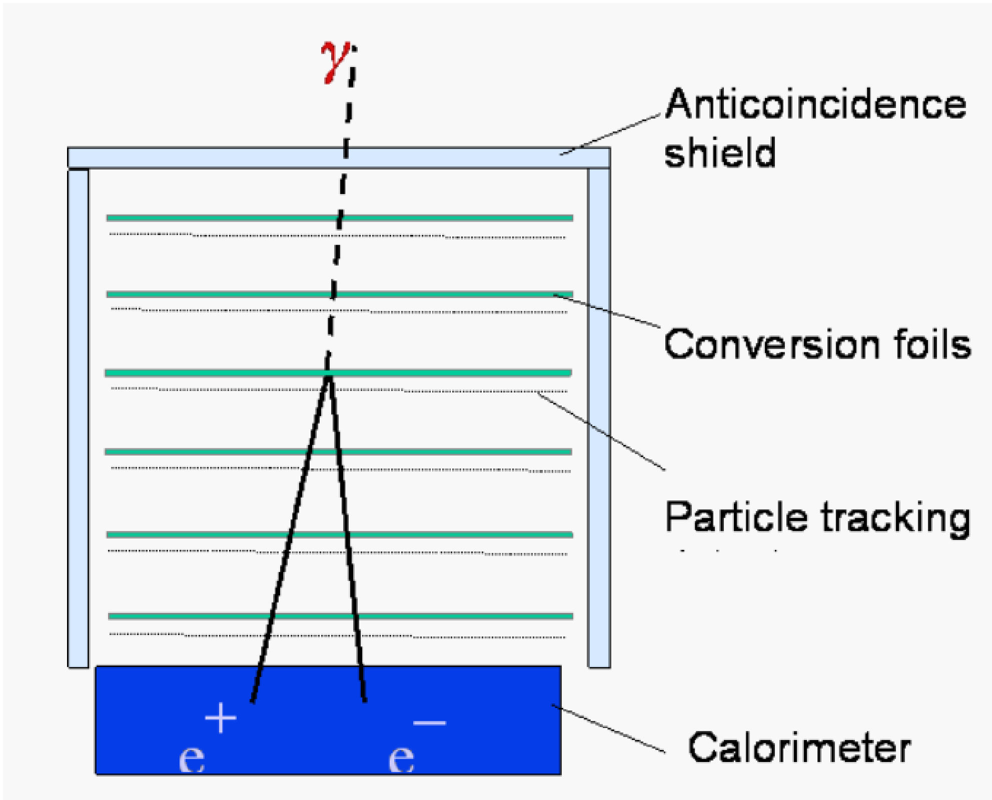


Figure 4: Fermi lat concept design

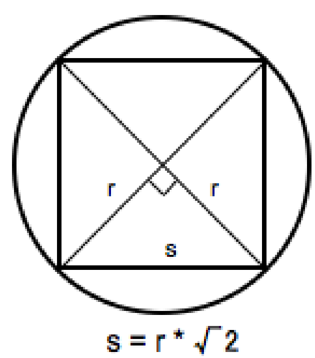


Figure 6: Counts Cube And ROI Relationship

In order to simplify calculations in the future the analysis now compiles a 3-D Counts Cube. This is a data file whose three axis are right ascension, declination, and energy. The edges of the cube are defined by the Radius of the ROI as seen in fig. 6 to the right. In my MRK 421 analysis I used eight logarithmiclly spaced bins per decade in our 100-513000 Mev energy range (NEED VALIDATION HERE).

At this point I have both diffuse models, isotropic and galactic, and the current Fermi catalog downloaded and in my configuration file. The isotropic diffuse model... The galactic diffuse model takes into account all gamma rays in the field of view that originate from the Cosmic Radiation Background (CMB) rather than the ROI sources. Subtracting these two models from the counts map gives us only those emissions originating from ROI sources. Lastly the fermi Catalog contains all models and parameters as found in older analyses. Fermipy compiles all model parameters from these three files into a single xml file for easy access.

Now fermipy computes a second data cube. The livetime cube contains sky position, right ascension and declination, and livetime as its three axis where livetime is the time that a source spent at a given inclination angle during an observation. These values are used to compute an exposure map where each spatial bin contains the total exposure for the ROI.

Taking the spectrum from the aforementioned source xml file and multiplying it with the exposure map gives us a model source map with an expected number of counts based on previous analyses. By using the

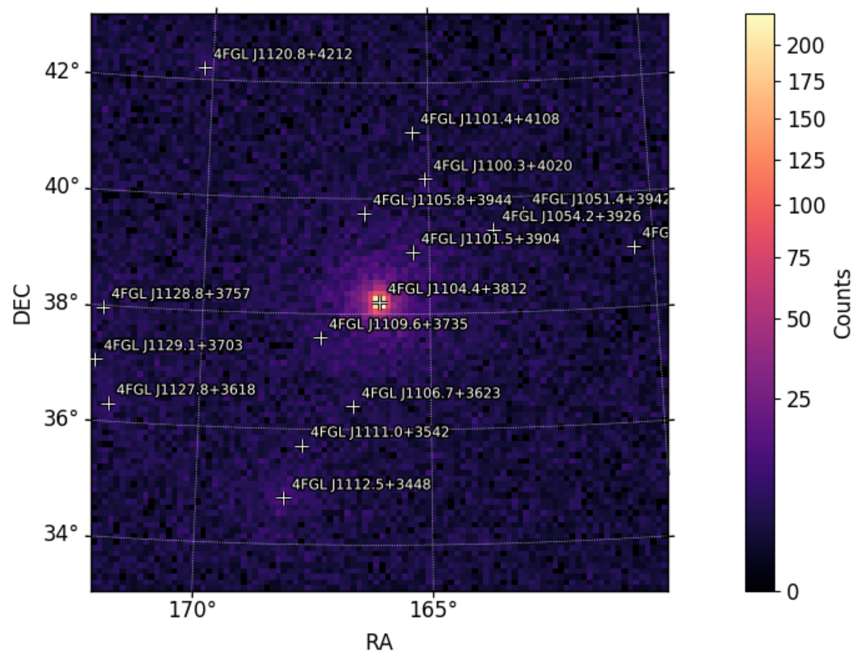


Figure 5: Counts map

expected values and true values for each energy bin fermipy performs a likelihood fit.

2.3 Binned Likelihood Analysis

Call m_i the number of expected counts in the i-th energy bin as computed in the model source map. And call n_i the number of counts in the i-th energy bin of the counts map. Then the probability of detecting n_i counts in the i-th bin is given by a Poisson distribution with average value m_i.

$$P_i = \frac{m_i^{n_i} e^{-m_i}}{n_i!}$$

Then the probability of attaining our data given our preliminary model is the product of all bin's probabilities. Hence,

$$L = e^{N_{exp}} \prod_i \frac{m_i^{n_i}}{n_i!}$$

Where N_{exp} is the sum of all expected counts m_i. The likelihood value L is related to the chi-squared statistical hypothesis test in the limit of bins with many counts. The relation is $\chi^2 = -2 \log L$. Hence maximizing L is equivalent to minimizing χ^2. This is the preferred method of fitting a model as minimizing functions are easier to compute.

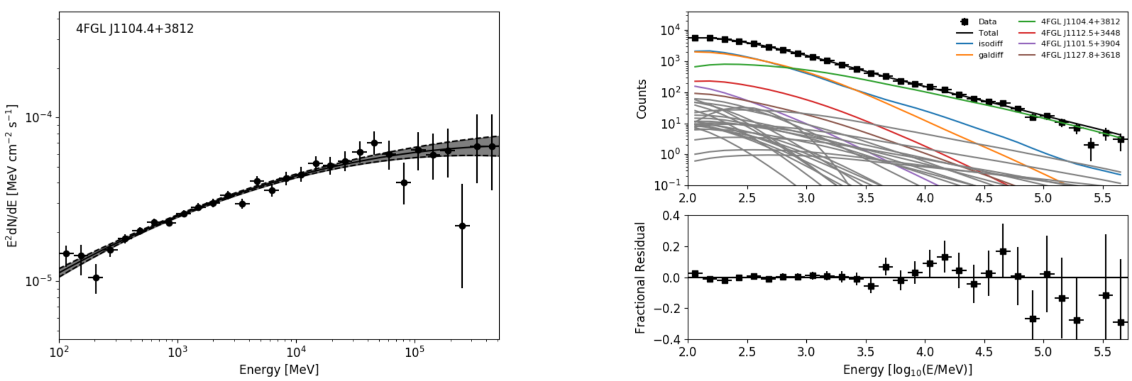
If you perform two different fits with two different models with an equivalent number of free parameters then the model with least χ^2 is the better model. If one model has more free variables than another then the first model is called a complex hypothesis while the latter model is called a simple hypothesis. If L_A is the likelihood of a simple hypotheses, L_B is the likelihood of a complex hypotheses, l_A is the log-likelihood of a simple hypotheses,and l_B is the log-likelihood of a complex hypotheses then the ratio

$$\Lambda = -2(\frac{L_A}{L_B}) = -2(l_A - l_B)$$

is related to the chi-squared values of the two models. If Λ is greater than or equal to 3.85 then you have 95 percent certainty that the complex model fits better than the simple model.

2.4 Spectral Energy Density

Given a photon has a continuous distribution of energies it could possess we can sort a source's photon data set into groups (bins) of similar energies. Each bin can be translated into a total energy, graphed, and fit with a model. This model's fancy name is a Spectral Energy Distribution or SED for short. With an SED we know how much energy is emitted from a source at a specific energy range. SEDs are useful for creating a lightcurve and understanding the basic properties of the source in question.



(a) SED

(b) Counts Spectrum

Figure 7: 2 Figures side by side

Creating an SED entails fitting a model to a source's binned counts using the likelihood method. The end result is a counts spectrum as seen in fig. 7b. Then a simple transformation of the left axis from counts to energy produces the SED seen in fig. 7a.

2.5 Lightcurve

Knowing how many photons a source emits is not as useful as knowing at which energies and with what intensity a source emits at. We can place a source's photon data into time bins (day/week/month/etc bins) and use the SED to transform a bin's total photons into a total energy value. This entails performing another likelihood fit for each time bin in order to get a more accurate description of each bin's varied emissions. Performing this operation over a data set produces a flux versus time graph called a lightcurve. A variety of analyses can be performed on a lightcurve to probe patterns within a sources emissions.

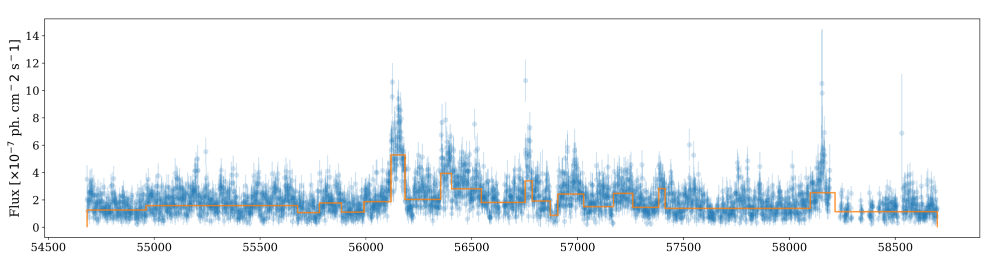


Figure 8: Fermi lat concept design

vvvvvvvvvv

3 Markarian 421 Analysis

3.1 Model Selection

3.2 Preliminary lightcurve and Bayesian Analysis

3.3 Final lightcurves and Correlation analysis