Summer Student Programme 2010
Final Report

Estimation of the probability of observing a gamma-ray flare based on the analysis of Fermi data

Matteo Palermo
Universitá degli Studi di Perugia

Supervisors:
Elisa Bernardini **Jose Luis Bazo Alba**

20 July 2010 - 09 September 2010, DESY-Zeuthen

1 Introduction

The Earth is continuously hit by energetic particles which either interact with the atmosphere or stop in the Earth or pass through; such energetic particles are called cosmic rays. By studying them (e.g. measuring properties as the energy) it is possible to explore the Universe, for example to understand what the stars are made of or which kind of objects are in the Universe. It's possible to classify cosmic rays in three categories:

- **charged particles:** ($\sim 79\%$), mainly protons, electrons and ionized nuclei. Their energy spectrum extends several orders of magnitude, from few 10^8 eV, reaching up to energies of $\sim 10^{20}$. Since these particles are charged their main feature is that their path is changed by the magnetic field (few microgauss) which is present in the interstellar medium, thus they don't carry any information about the direction of their source.
- **electromagnetic radiations:** these radiations cover the whole electromagnetic spectrum; in this report we are particularly interested in gamma rays (for reasons that will be discussed later). The gamma rays energy band extends from 0.5 MeV up to 100 TeV.
- **neutrinos:** these particles are unique messengers to explore the Universe because they have no charge and can only interact via the weak force, thus their path is not affected by the magnetic field and their flux practically remains unchanged along the path from the source to the Earth.

There are basically two kinds of models which can explain the presence of such energetic gamma rays and neutrinos:

- **leptonic models**
- **hadronic models**

2 Multi-messenger Astronomy and the NToO

In the hadronic models take place reaction like:

$$p\gamma \rightarrow \Delta^+ \rightarrow n\pi^+ \tag{1}$$

$$\text{or } \rightarrow p\pi^0 \tag{2}$$

It is well known that the charged pions decay in a muon and a neutrino

$$\pi^+ \rightarrow \mu^+ \nu_\mu$$

while the neutral pion decays in two gammas

$$\pi^0 \rightarrow \gamma\gamma \ .$$

So neutrinos can be associated with high energy photons and protons, therefore correlation studies between these particles are as natural as reasonable. This kind of analysis is also called the multi-messenger astronomy. This concept is the main reason for the work that is going to discuss in this report therefore we need to spend few words on this topic in order to portray it better. In particular we will focus on the correlation between photons and neutrinos.
Multi-messenger data can increase the discovery potential and constrain the phenomenological interpretation of high energy emission of certain sources. This kind of studies in principle can be accomplished off-line, searching for correlations between the measured intensity curves in the electromagnetic spectrum and the time of detected neutrinos. Such a time dependent analysis enhances the discovery probability by profiting from the photon-neutrino correlation. However the main problem for these off-line studies is that for many sources there might be missing photon data because many gamma-ray telescopes have a small field of view (e.g. Imaging Atmospheric Cherenkov Telescopes, IACTs, like MAGIC) so they cannot look at a wide region of sources at the same time as in the case of neutrino detectors (e.g. IceCube) which can look at the entire sky. Moreover these telescopes are not taking data continuously (overall duty cycle of around $\sim$15%).
In order to avoid these problems an ONLINE alerts program, called Neutrino triggered Target of Opportunity (NToO) [2], was developed first in 2006 for AMANDA-II and MAGIC and now is planned for IceCube and MAGIC but in principle can be used for any other IACTs. The idea is that a significant neutrino observation from IceCube can trigger the follow-up observations from IACTs: when a neutrino (or a multiplet with significance above a given threshold) is detected a signal is sent to MAGIC in order to let it focus on the incoming neutrino's direction.
Now, to define a *significant* neutrino observation an online filter was developed, consisting of a significance calculation. Once combined observations have been performed, it is necessary to estimate the overall probability of random particle detection, namely the probability that the observed neutrino and gamma ray emission are not casually connected. This can be done by calculating the following quantity:

$$\alpha = \sum_{m=N_{obs}}^{+\infty} \frac{(N_{bck})^m}{m!} e^{-N_{bck}} \cdot \sum_{j=N_{coinc}}^{m} \frac{m!}{j!(m-j)!} (P_{gam})^j (1-P_{gam})^{m-j} \qquad (3)$$

which represents the probability of observing at least N_{obs} neutrino alerts above the significance threshold and detecting at least N_{coinc} coincident gamma-ray flares. In the first term (clearly poissonian) N_{bck} represents the number of background alerts expected (note that the null hypothesis is that all the neutrinos are of atmospheric origin), while in the second term the P_{gam} is the probability of observing a gamma-ray flare of a particular source.
This P_{gam} is the quantity of interest for this report. In order to avoid statistical biases it is mandatory that this statistical test is defined a-priori. To estimate the quantity P_{gam} we selected certain sources which respect certain criteria and

then we studied their light curves in a way that is going to be presented later on. The main requirement for such a study is that we need that the photon data to analyze must be taken continuously and so, due to the reasons given above, is not possible to use the IACTs data. Therefore to estimate P_{gam} we used Fermi data because Fermi is a satellite telescope and as satellite has a large field of view (2.4 steradians, thus about 20% of the sky) and is taking data continuosly; however it covers a slightly lower energy range. Later on this report will be discussed the issue about the energy ranges.

3 Fermi experiment

The Fermi Gamma-ray Space Telescope ([4]) is an observatory for the study of gamma-ray emission from astrophysical sources. Fermi scientific objectives span from the detailed study of pulsar, AGNs and diffuse emission, to the search for new classes of gamma-ray emitters and the possible signals of new physics, such as from the dark matter. Fermi has two main instruments:

- **the Large Area Telescope (LAT)**: a gamma-ray imager operating in the energy band between 30 MeV and 300 GeV;
- **the Gamma Ray Burst Monitor (GBM)**: a detector covering the 8 keV-20 MeV energy range, devoted to the study of the Gamma Ray Bursts.

Since for this work we used LAT data we will describe its performances briefly. The LAT instrument has a large field of view of 2.4 sr and an effective area of $\approx 8000\ cm^2$ for normal incidence at 1 GeV. Furthermore the silicon detectors provide high angular resolution (0.15, 0.9 and 3.5° of single photon angular resolution, respectively at 10, 1 and 0.1 GeV).

Fermi operates primarily in an all-sky scanning survey mode that maximize observing time while maintaining excellent uniformity; this means that the LAT rarely stares at a single point in the sky (this peculiarity is actually very important for the purpose of this report). Rather, the boresight of the LAT continuously scans the sky, alternating between the northern and southern hemispheres each orbit; this provides 30 minutes of lifetime on each point in the sky every two orbits (approximately three hours).

4 Analysis

In this chapter we present first the basic idea used to estimate the probability of observing a gamma-ray flare from a certain source, P_{gam}, then the actual procedure and all the details, and last but not least we will justify this procedure. In the next chapter we will discuss the results for each source that has been analyzed, the improvements of this analysis and we will compare, whenever it's

possible, our results with the ones of other experiments (e.g. MAGIC).
The basic idea to estimate P_{gam} is very simple: we make a light curve of the LAT data for a certain source, then we need to set a threshold in order to define what is a flare and what is not, so the estimation of the P_{gam} will be simply the ratio between the amount of time that the light curve is over the threshold and the total period of data taking, which in our case is approximately two years. So first of all let see what a light curve is and how to obtain it.

4.1 Light curve

A light curve is a graph which shows the light intensity (flux, measured in photons/cm^2s) of an object over a period of time. Such a graph gives information about the variability of the source that we are looking at and it's very useful to understand processes at work within the astrophysical objects and identify specific categories (or classes) of astrophysical events. Here, in this work, we use the light curve to estimate P_{gam}. To make a light curve we collected all the data available from the LAT detector so far (thus from August 2008) and then we used the following tools provided by the Fermi collaboration [5]:

- **gtselect:** this tool provides a selection of the collected data in energy and direction. In particular we set the energy range from 100 MeV to 300 GeV (almost all the range available);
- **gtmktime:** this tool creates the Good Time Intervals (GTIs) which represent the periods that the LAT detector was working;
- **gtbin:** this tool was used to bin the data in time. We set the bin size equal to 1 day (thus 86400 s);
- **gtexposure:** this tool calculates the exposure[1] associated with each time bin. To do that it requires the photon spectral index to be used for weighting the exposure as a function of energy. In first approximation we used the same spectral index (=-2.1) for all the selected sources.

4.2 Threshold

Once we have the light curve for a certain source we have to define the threshold to discriminate between the flaring state and the steady state. In order to do that in a statistically meaningful way we analyzed the flux distribution, which is simply the marginal probability density function for the flux (projecting the light curve's points onto the flux axis) in three different ways:

- **log-normal fit**
- **gaussian fit of log(flux) distribution**

[1] it represents the effective sensitive detector area times the time that the detector can take data. Usually is energy dependent

- **gaussian fit with restrictions**

The reasons for fitting with three different functions are many. First we want to check if everything is alright by comparing the first and the second modes, second we check which one is the best fit from the statistic point of view (for example the third one should be the worst), then to study how the threshold depends on the particular mode (and on the spectral index chosen to compute the exposure) and if the behavior of the threshold with respect the fit mode is always the same.
The threshold, for each fit mode, is defined as the mean value of the fit plus 5 (or 3 in some cases, see next section) sigma (square root of the variance, as given by Root).

4.3 The cumulative approach

Once we defined the threshold we computed the P_{gam} by evaluating the cumulative of the flux distribution; to be more precise the probability ends up to be the following:

$$P_{gam} = 1 - \frac{\int_0^{Thr} flux\, d(flux)}{N} \,, \tag{4}$$

$$\text{where N is the normalization factor } N = \int_0^{+\infty} flux\, d(flux) \,. \tag{5}$$

In practice we integrated the resulting function from the lognormal fit. In fig. 1 there is (left) an example of the resulting cumulative distribution where the histogram represents the cumulative of the data and the line band on top of it represents the cumulative of the lognormal fit function (with its errors). It is shown on the right an example of P_{gam} as a function of the threshold (Thr). In the latter the red vertical band represents the threshold as obtained from the gaussian fit, the blue one is from the gaussian log(flux) fit and the black one from the lognormal fit. For each mode the lower limit is computed as follows:

$$Thr_{lower} = (mean - \delta(mean)) + (n_{sigma}) \cdot (\sigma - \delta\sigma)$$

where $\delta(mean)$ and $\delta(\sigma)$ are the errors of the fitting parameters; similarly the upper limit is:

$$Thr_{upper} = (mean + \delta(mean)) + (n_{sigma}) \cdot (\sigma + \delta\sigma) \,.$$

The upper limit of the curve is evaluated by using the lower limit lognormal fit function whereas the lower limit of the curve by using the upper limit lognormal fit function. Furthermore the bounds for P_{gam}, for each fit mode, are calculated as follow:

$$\text{the upper limit } P_{gam-UP} = 1 - \int_0^{Thr_{lower}} (lognormal - fit - upper - bound)$$

and

the lower limit $P_{gam-LOW} = 1 - \int_0^{Thr_{upper}} (lognormal - fit - lower - bound)$.

For the fit modes which the errors on the mean and the sigma are not available yet (see improvements in the conclusion's section), namely the lognormal and log(flux) fit modes, in order to evaluate the P_{gam}'s bounds we simply integrate until the threshold (Thr).

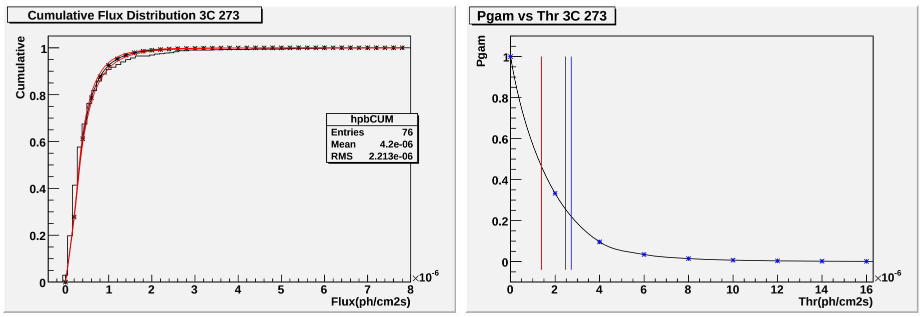

Figure 1: Left: cumulative distribution, right: P_{gam} as a function of the threshold

4.4 Energy range issue

As anticipated in the previous chapter the LAT detector covers a different energy range with respect to IACTs like MAGIC so the question is: is the P_{gam} still the same in the TeV scale?
In order to answer this question we made the same analysis above but for a restricted energy range, namely from 1 GeV to 300 GeV and look for for P_{gam} values; if P_{gam} doesn't change it's likely that it won't change even in the TeV scale.
Since the first energy range includes the second one an improvement of this analysis can be done by studying the energy ranges 100 MeV-1 GeV and 1 GeV-300 GeV separately. Nonetheless, since the energy spectrum should follows a power law we expect more flux for smaller energy, this means that in the first energy range (the bigger) most of the contribution should come from the low part of the range so our comparison can still be reliable.

5 Sources

In this section we will introduce the sources used for the analysis and the criteria used. Then the results for each source will be presented and commented.

The criteria used to select them are the following:

- they have been classified as *variable*[8] in the Fermi/LAT bright[2] source catalog[3];
- they should have been observed in TeV scale[4];
- they are monitored by MAGIC.

As you can see all of them are AGNs[5] (basically Blazars), which are among the brightest extragalactic objects in the Universe. Moreover these kind of sources are known to be variable.

6 Results

Here are presented the results just for the first energy range; in the next section we will discuss all the results.

6.1 Energy range (100 MeV-300 GeV)

- **3C 273**

Fit mode	χ^2/ndf	**Mean**	$\sqrt{Variance}$	**Thr**	$P_{gam}(5\sigma)$
Lognormal	1.7	$4.4 \cdot 10^{-7}$	$4.1 \cdot 10^{-7}$	$2.5 \cdot 10^{-6}$	$0.0049^{+0.002}_{-0.0015}$
Log(flux)	1.4	$3.7 \cdot 10^{-7}$	$4.7 \cdot 10^{-7}$	$2.7 \cdot 10^{-6}$	$0.0035^{+0.0015}_{-0.002}$
Gaussian	14.3	$3.16 \cdot 10^{-7}$	$2.14 \cdot 10^{-7}$	$1.39 \cdot 10^{-6}$	$0.032^{+0.014}_{-0.011}$

Table 1: Fit results for 3C 273

- **3C 66A/B**

[2] Flux $> 1.1 \cdot 10^{-7} ph\, cm^{-2}\, s^{-1}$ in the energy range 100MeV-1GeV

[3] Most of them have been observed previously by the Energetic Gamma Ray Experiment Telescope (EGRET)

[4] Just the 3C 273 has not been observed yet in the TeV scale

[5] In this type of galaxies the central emission out-shines the stars. They are composed by a supermassive black hole ($10^6 - 10^{10}$ solar masses) with an accretion disk rotating around it; they often show two ultrarelativistic jets moving in opposite directions, perpendicular to the accretion disk. AGNs are classified based on their orientation with respect to the Earth

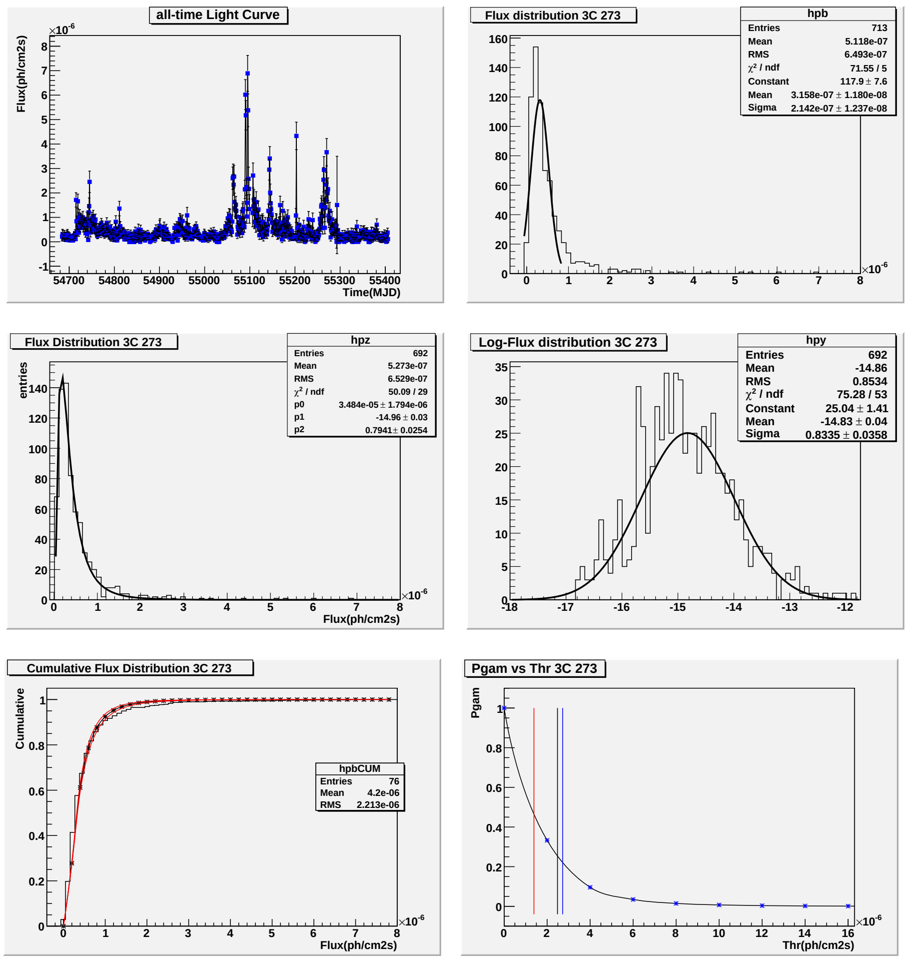

Figure 2: top-left: light curve, top-right:gaussian fit, middle-left:lognormal fit, middle-right: gaussian log(flux) fit, bottom-left: cumulative distribution, bottom-right: P_{gam} vs Thr for 3C 273

Fit mode	χ^2/ndf	**Mean**	$\sqrt{Variance}$	**Thr**	$P_{gam}(5\sigma)$
Lognormal	2.3	$3.7 \cdot 10^{-7}$	$1.9 \cdot 10^{-7}$	$1.3 \cdot 10^{-6}$	$0.0019^{+0.0011}_{-0.0008}$
Log(flux)	1.8	$3.4 \cdot 10^{-7}$	$2 \cdot 10^{-7}$	$1.4 \cdot 10^{-6}$	$0.0014^{+0.0009}_{-0.0006}$
Gaussian	1.3	$3.32 \cdot 10^{-7}$	$1.57 \cdot 10^{-7}$	$1.12 \cdot 10^{-6}$	$0.005^{+0.005}_{-0.003}$

Table 2: Fit results for 3C 66A/B

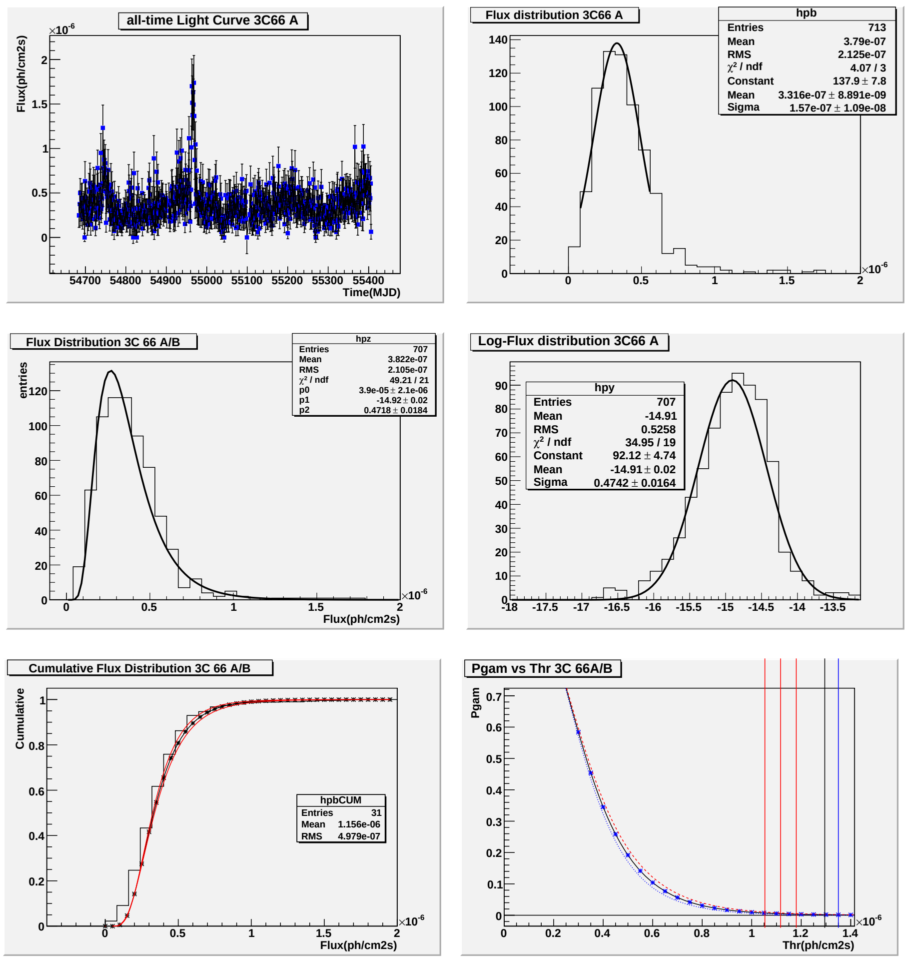

Figure 3: top-left: light curve, top-right:gaussian fit, middle-left:lognormal fit, middle-right: gaussian log(flux) fit, bottom-left: cumulative distribution, bottom-right: P_{gam} vs Thr for 3C 66A/B

7 Comments and Conclusions

By comparing the results between the two different energy ranges for each source we saw that:

- by looking at the P_{gam} values relative to the lognormal fit, they basically remain constant for each source so this can be a good hint to extrapolate these values also in the TeV scale, that is actually the aim of this work.
- as expected the lognormal fit and the log(flux) fit give basically the same results in each source.

- the errors of P_{gam} are asymmetric just because of their evaluation method (explained in section 4.3).

Moreover a clarification about the source 3C 66A/B is mandatory: these are actually two sources which have an angular distance $\sim 0.09°$ while the angular resolution of the LAT detector is, as anticipated, at most $0.15°$ (at 10 GeV) so it's impossible to resolve them.
So far, with this analysis, we can conclude that the possibility to extrapolate the results for the P_{gam} at the TeV scale is actually realistic, nevertheless it's clear that this work needs many improvements and corrections as:

- evaluate the errors for the mean and the variance of the lognormal and log(flux) fit mode (by error propagation);
- re-do the same analysis but with the energy ranges completely separated;
- define the confidence level for each results (so far it is undefined);
- evaluate the actual spectral index for each source and re-do the analysis.

8 Acknowledgments

I would like to thank the DESY summer student programme for giving me such a great opportunity, then the AT group and in particular all the members of the Young Investigators Group. A special thanks to my supervisors, Elisa Bernardini for her support and Jose Luis Bazo Alba for his infinite patience although his huge commitments. Lastly, I thank all the summer students for the time spent together at DESY-Zeuthen, 2010.

References

[1] Bazo-Alba, J. L.: *Search for steady and flaring astrophysical neutrino point sources with the IceCube Detector*, PhD thesis, Humboldt-Universität zu Berlin, 2010

[2] Neutrino triggered high-energy gamma-ray follow-up with IceCube., *Franke R., Bernardini E. et al. for the IceCube Collab.* Proceedings of the 31^{st} ICRC, Lodz, Poland, Jul 2009

[3] Vitale, V. (for the LAT Collaboration), *FERMI/GLAST and the Future of High Energy Gamma Ray Astronomy*

[4] FSSC web-site, *http://fermi.gsfc.nasa.gov/ssc/observations/types/allsky/*

[5] Robin Corbet, FSSC *LAT light curve analysis* http://users.monash.edu.au/dgallow/fdaw/doc/Robin-LightCurveAnalysis.pdf

[6] B. Giebels, B. Degrange, *Lognormal variability in BL Lacertae* A&A 503, 797-799 (2009)

[7] Degrange et al. 2008, *The emission of blazars in VHE gamma-rays viewed as a random stationary process: the case of PKS 2155-304* in Proceedings of Workshop on Blazar Variability across the Electromagnetic Spectrum, PoS(BLAZARS2008)016

[8] Abdo, A. A.; et al (Fermi Collaboration), *Fermi Large Area Telescope Bright Gamma-ray Source List,* Astrophys. J. Suppl. 183 (2009) 46

[9] The Fermi-LAT Collaboration, *Spectral Properties of Bright Fermi-detected Blazars in the Gamma-ray Band,* Jan 2010, arXiv:1001.4097