University of California at Santa Cruz

Search for CP Violating Gamma-Ray Signals with Fermi-LAT Pass 8 Data

Kevin P. Rodriguez

August 21, 2018

A thesis presented in the partial satisfaction
of the requirements for the degree of

Bachelor of Science

in

Physics

The thesis of Kevin P. Rodriguez has been approved by

Professor William B. Atwood	Professor David Belanger
Faculty Advisor	Senior Theses Coordinator

Professor Robert P. Johnson
Chair, Department of Physics

Acknowledgements

It goes without saying that none of the work that I have presented here would have been possible without the combined efforts of those scientists, engineers, and dreamers who collectively compose the wonderful Fermi-LAT collaboration. Within my own limb of the Fermi group at Santa Cruz, I would especially like to thank Professors Steve Ritz, Robert Johnson, and Bill Atwood for their inexhaustible support on this project, as well as their guidance and counseling through my various steps (and missteps) in my short physics career thus far. My meetings with Professor Atwood first began long before I had even declared physics as my major, and I surely would not know half of what I do now without his perennial assistance and encouragement.

Additionally from the collaboration, I am indebted to Regina Caputo and PhD candidates Brendan Wells and Christian Johnson for taking the time to help me overcome various roadblocks in my research. And finally, within my own cohort of physics undergraduates, I am particularly grateful to Kfir Dolev, Lia Gianfortone, Brutus Gruey, and Thea Dumont for always pushing me toward a better self.

Search for CP Violating Gamma-Ray Signals with Fermi-LAT Pass 8 Data

Abstract

A parity-odd correlator (Q) capable of probing the handedness of primordial extragalactic magnetic fields is constructed and computed with Fermi-LAT Pass-8 gamma-ray data (10-60 GeV). Two phases of analysis are considered: (I) diffuse and (II) source-anchored. Known AGNs are masked with a radius of 1.5° for the diffuse analysis, while the same AGNs' coordinates are used for anchoring the helical patterns used in calculating Q in the source-anchored phase. The differences in signal strength in the northern and southern hemispheres are discussed for the diffuse analysis. Although we observe a weak right-handed signal in some modes of analysis, this signal is absent from the northern hemisphere and is therefore deemed an artifact unrelated to CP violating mechanisms. We report neither a coherent right- nor left-handed signal, with essentially all values of Q lying within 95% of a null signal ($Q = 0$).

Contents

Introduction

It is clear from our surroundings that our universe is dominated by matter rather than antimatter. What is not clear is *how* our universe came into such an asymmetry. Modern cosmological models predict that matter and antimatter were created in equal amounts in the early universe, yet no antimatter-dominated regions have ever been detected.[1] Reconciling these two ideas requires a mechanism by which baryonic matter could grow to its current excess following inflation. The three necessary conditions for this so-called "baryogenisis" were first described by Sakharov [9] in 1967:

1. Non-conservation of baryon-number (B)
2. Charge-Parity (CP) violation
3. Interactions out of thermal equilibrium

The first process has not been observed in any decay modes.[2] CP violation, however, has been observed in a number of channels, all of which involve the weak interaction. The third condition can be attained during an electroweak phase transition as the universe cools down, breaking the $SU(2) \times U(1)$ symmetry. In what follows, we will focus exclusively on the second Sakharov Condition. In Section 0.1 we establish the effects of CP violation on the matter-antimatter asymmetry as observed in known particle decays. In Section 0.2 we describe the relation between electroweak baryogenesis and the present-day helicity of extragalactic magnetic fields. Finally, in Section 0.3, we propose a method of indirectly probing the helicity of these same magnetic fields with Fermi data.

0.1 CP Violation and the Matter-Antimatter Asymmetry

Historically, CP was proposed by Lev Landau as the "true" symmetry of nature following the failure of P-symmetry observed by the Wu Experiment [14] in 1957. On its own, a parity transformation amounts to the inversion of the spatial coordinates of the object at hand. For a particle with wavefunction $\psi(\mathbf{r})$,

$$\hat{P}\psi(\mathbf{r}) = \psi(-\mathbf{r}). \tag{1}$$

[1] An antimatter-dominated region would be easily visible in gamma-rays near its interface with nearby matter-dominated regions due to pair annihilation.

[2] The Standard Model does leave room for B violation, but these interactions are suppressed by a factor of $\exp\{-4\pi \sin^2 \theta_W/\alpha\}$, where $\alpha \approx 1/137$ is the fine structure constant and $\theta_W \approx 28°$ is Weinberg's weak-mixing angle [8]. With these values, the present-day exponential suppression is $\sim 10^{-165}$. [10]

On the other hand, charge conjugation converts a particle into its corresponding antiparticle

$$\hat{C}|p\rangle = |\bar{p}\rangle. \tag{2}$$

The canonical probe of CP symmetry is the neutral kaon – the bound state of a down quark and a strange antiquark $|K^0\rangle = |d\rangle \otimes |\bar{s}\rangle$. A parity transformation reveals that the kaon is a *pseudoscalar* particle

$$\hat{P}|K^0\rangle = \hat{P}\{|d\rangle|\bar{s}\rangle\} = (+1)(-1)|d\rangle|\bar{s}\rangle = -|K^0\rangle,$$

where I have used the fact that parity is a multiplicative quantum number, as well as the antisymmetry of fermionic fields. The pseudoscalar[3] nature of the neutral kaon allows for the construction of CP eigenstates

$$|K_1\rangle = \frac{|K^0\rangle - |\bar{K}^0\rangle}{\sqrt{2}} \qquad\qquad |K_2\rangle = \frac{|K^0\rangle + |\bar{K}^0\rangle}{\sqrt{2}}, \tag{3}$$

with CP eigenvalues $+1$ and -1, as can be checked by applying the operations in Eqs. (1) and (2) to the states in Eq. (3). Now, if CP is indeed a perfect symmetry of nature, any decay products of $|K_1\rangle$ or $|K_2\rangle$ must retain the original value of CP after the interactions have taken place. The neutral pion, while lighter than the kaon, is also a pseudoscalar, and so two valid decay channels via the weak interaction include

$$K_1 \to 2\pi^0 \quad (CP = +1)$$
$$K_2 \to 3\pi^0 \quad (CP = -1).$$

So, if one were to construct a beam of neutral kaons (K^0), the K_1 component would decay nearly immediately into two pions, with the remaining K_2's decaying into three pions much further down the line. If, however, the K_2's were ever observed decaying into only two pions, this would be a stark violation of Charge-Parity ($CP = -1$ before and $CP = +1$ afterwards). This was exactly the setup for the experiment devised by Fitch and Cronin [4] in 1964, in which they observed a CP violating decay about once in every 500 decays.[4]

Further investigation of CP violation in the weak sector in the case of semi-leptonic (electron/positron) decays has shown that this asymmetry persists in other weak interactions as well. The process $K_L \to \pi^- + e^+ + \nu_e$, for example, occurs slightly more often than its CP conjugate $K_L \to \pi^+ + e^- + \bar{\nu_e}$ [8]. Interactions such as these indicate fundamentally dissimilar treatments of matter and antimatter by nature. It is this fact that posits the connection between CP violation and a matter-antimatter asymmetry.[5]

[3] "Scalar" for its lack of spin, and "pseudo" for its vector-like behavior under parity.

[4] As the K_1 and K_2 were revealed by Fitch and Cronin to not be perfect CP eigenstates after all, another set of states are defined to reflect this, denoted $|K_L\rangle = (|K_2\rangle + \epsilon|K_1\rangle)/\sqrt{1+|\epsilon|^2}$ and $|K_S\rangle = (|K_1\rangle + \epsilon|K_2\rangle)/\sqrt{1+|\epsilon|^2}$, where $\epsilon \approx 2.3 \times 10^{-3}$ reflects the small mixing due to CP violation. The subscripts L and S refer to the long and short lifetimes of each state, respectively. [5]

[5] This is not the end of the story, however. The amount by which the CP symmetry is broken in observed channels is extremely small ($\sim 10^{-3}$) [5] and it is believed that another mechanism must be involved to account for the present-day abundance of baryonic matter.

0.2 Primordial Magnetic Fields and Baryogenesis

An alternative probe of CP violation involves looking directly to the cosmos themselves in a search for a primordial relic left behind from inflation, similar to how the Cosmic Microwave Background (CMB) informs us of the temperature, composition, and expansion rate of the early universe. The idea is that CP violating mechanisms prior to inflation may have laid the seeds for chiral magnetic fields, whose helicity imposes a positive feedback loop during inflation, allowing those same magnetic fields to propagate throughout the entire universe at present-day. The theoretical relation between the magnetic helicity h[6] and the baryon density n_B was derived by Vachaspati [13] to be

$$h = \frac{1}{V}\int \mathbf{A}\cdot\mathbf{B}\,\mathrm{d}^3 r \approx -\frac{n_B}{\alpha}, \tag{4}$$

where $\mathbf{A}$ is the magnetic vector potential, $\mathbf{B} = \nabla \times \mathbf{A}$ is the magnetic field, and $\alpha \approx 1/137$ is the fine-structure constant. The baryon density at present day is positive [8], and so Eq. (4) tells us that the helicity of the extragalactic magnetic fields should be negative ("left-handed"). Note that the helicity as defined above is a pseudoscalar – this is what enables us to probe the magnetic fields for parity-odd effects. The problem of determining the parity-odd nature of these primordial magnetic fields seeded by CP violation then reduces to the problem of observing the helicity of the fields themselves in some way. A proposed method of observation has been suggested by Tashiro and Vachaspati [12] using Fermi-LAT photon data at GeV energy scales.

0.3 A Parity-Odd Signal from Active Galactic Nuclei

Active Galactic Nuclei (AGNs) are the most consistently energetic objects in the observable universe[7] and the TeV energy gamma-rays ejected from an AGN are ripe for pair-production with Extragalactic Background Light (EBL) nearby.[8] In the presence of a magnetic field, the electron (positron) will follow a spiraled trajectory, lending some of its enormous kinetic energy to any CMB photons who intercept its path via inverse-Compton scattering. The result is a cascade of ∼GeV energy photons emanating from points along the spiraling path of the electron (positron), each with less energy than the last. As the handedness of the spiral corresponds to the handedness of the magnetic field, such a cascade of photons gives an indirect probe of these fields' helicity. A schematic depiction of these chiral gamma-ray spirals is shown in Fig. (1).

In practice, a parity-odd signal will have to be teased out of the densely populated sea of gamma-rays corresponding to other background phenomena. The signal-to-noise ratio is then expected to be very small, and an algorithm for systematically scanning the sky for all possible parity-odd gamma-ray spirals is needed. Such an algorithm has also been suggested by Tashiro and Vachaspati [12], yielding a parity-odd correlator proportional to h. The development and implementation of this algorithm is discussed in Chapter 1, as well as in the Methods (Chapter 2).

[6] As defined here, h is actually a magnetic helicity *density*, but the last word is usually omitted and understood to refer to helicity per unit volume.

[7] "Galactic Nucleus" refers to a supermassive blackhole at the center of a galaxy. "Active" just means that it is currently (in our own frame of reference) spewing matter and light out into many light-year-long jets along its axis of rotation.

[8] We consider only gamma-rays with energies in the TeV range here, as AGN-sourced photons with energies greater than ~ 0.8 PeV will be capable of pair producing with CMB photons while still in the radiation field of the AGN. See Appendix A.

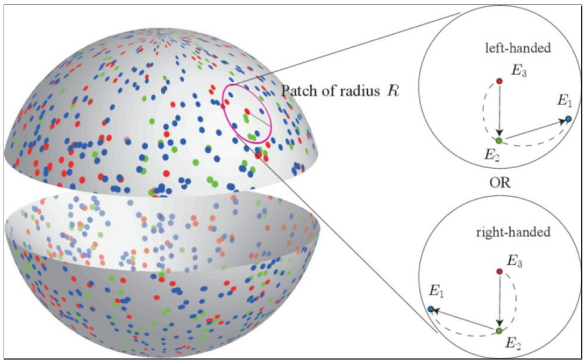


Figure 1: Possible left- and right-handed gamma-ray cascade signals. Figure from Tashiro, Chen, Ferrer, and Vachaspati, 2014 [11].

As Tashiro and Vachaspati have already claimed that a left-handed signal exists with $\sim 2.7\sigma$ significance [11], the purpose of my own analysis is to search for the same signal but with newer, more sophisticatedly processed data. The Fermi-LAT has been taking data for an additional 4 years since the release of Tashiro and Vachaspati's results, and the entire catalog of Fermi data has also since been reprocessed with Pass 8 (replacing Pass 7) algorithms and data categorization. Regardless of whether or not the signal claimed by Tashiro and Vachaspati is boosted or washed away by the addition of new data following my own analysis, the substantial amount of new data used will tighten the uncertainty on the resulting signal.

Details on the data used and techniques of computation are discussed in the Methods (Chapter 2). The analysis itself is dissected into both diffuse and AGN-sourced phases, both presented in Chapter 3. The diffuse analysis involved the masking of known gamma-ray sources, while second phase of analysis investigates the possibility that a CP-violating signal may actually be boosted by instead using the coordinates of these gamma-ray sources directly. The diffuse and AGN-sourced analyses are discussed in Sections 3.1 and 3.2, respectively. Finally, the results, along with current limitations, are discussed holistically with thoughts on how they might be addressed in future work in Chapter 4.

Chapter 1

Constructing the Parity-Odd Correlator Q

As mentioned in the Introduction, our detection of any CP-violating signal will rely on a careful analysis of Fermi-LAT photon data in the GeV range. Before discussing the specifics of the data used, however, we must first establish the correlator that defines the algorithm used to search for such a signal.

1.1 Tracing the Signal to Its Origins

The signal that we seek is thought to be encoded in the helicity of the primordial extragalactic magnetic fields. As probing the nature of the extragalactic magnetic fields directly is clearly not an option, more creative routes have been sought. Fermi allows one of these routes to be explored indirectly via electromagnetic phenomena that relate the primordial magnetic fields to AGN-sourced photons, and eminently to Fermi's own detectors.[1]

The origins of our photons of interest may be traced back to those emitted in the jets of AGNs scattered across our observable universe. These TeV energy gamma-rays are not, however, the same photons that give rise to our signal. While the detection of AGN-sourced photons is useful in informing us of the location of the AGNs, they alone would carry no information regarding the extragalactic magnetic fields they have traversed to reach the Fermi-LAT. Additionally, the mean free path for TeV energy gamma-rays traversing the EBL is only about 10 Mpc, meaning that essentially no photons with energies greater than $\sim$ 80 TeV will be able to reach the LAT at all, which itself lies gigaparsecs away from even the nearest AGNs.[2] The short-lived life of these ultra-high energy gamma-rays is enforced by a non-zero cross-section for the process $\gamma + \gamma \to e^+ + e^-$ past threshold ($\sqrt{s} \geq 2m_e c^2$), which is easily attained collectively for a TeV energy gamma-ray and an optical EBL photon ($\varepsilon \approx 1$ eV) at a typical scattering angle of $\pm 90°$.[3]

Fortunately, these same positron-electron pairs resulting from the aforementioned threshold process are indeed sensitive to any magnetic fields they may encounter on their path. While traversing an extragalactic magnetic field, these charged particles will trace out curved trajectories in accordance with Lorentz' force law, which when com-

[1] The probability for an individual photon emmitted isotropically from the nearest quasar (which itself is a naive assumption) to then make its way in a straight line to the Fermi-LAT is $\sim 10^{-52}$.

[2] The mean-free path is inversely proportional to the cross-section of pair-production and the density of EBL photons (about 3×10^{-3} cm^{-3}). Additionally, $\gamma + \gamma \to \gamma + \gamma$ processes may further suppress the upper limit on photon energy that Fermi can "see", but those processes are additionally attenuated by a factor of α^{-4}. See Appendix A.

[3] The kinematic constraints for two-photon pair-production are discussed in Appendix A.

bined with Newton's 2nd Law yields:

$$\mathbf{x}(t) = \frac{q}{m} \int_0^t \int_0^{t'} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \, dt'' dt'. \tag{1.1}$$

Without explicit knowledge of the **E** and **B** fields, this integral is certainly not tractable. Furthermore, it is not entirely clear why we should be interested in the paths of these charged particles, as the Fermi-LAT detects the resulting photons, not the originating electrons themselves.

The missing link is an additional process of the form $e + \gamma_{CMB} \rightarrow e' + \gamma'_{CMB}$, where e represents the electron or positron pair-produced earlier from the interaction of AGN photons with the EBL.[4] Following creation, our electron of interest will take off carrying half the energy of the original TeV gamma ray, on average.[5] As the electron traverses further away from the AGN, it may subsequently scatter off of the CMB, imparting a significant fraction of its newly gained kinetic energy to a CMB photon, kicking the photon's energy up to the GeV range. This process is referred to as inverse-Compton scattering, or up-scattering. The same electron may actually up-scatter several CMB photons before losing the bulk of its kinetic energy, with the curve of its trajectory tightening with each impact before traversing a cooling length of several kpc [7]. It is these up-scattered, "cascade" GeV energy gamma-rays that we are interested in analyzing for our signal.

We know from electrodynamics that a charged particle in a magnetic field follows a trajectory along the field lines with a radius of curvature inversely proportional to the strength of the field. One may imagine the transverse motion of the charged particles as spiraling tightly around a given field line due to Lorentz' force law while the average net velocity remains tangent to the field line at any point along the trajectory.[6] On the largest scales, then, our pair produced electrons' trajectories will mimic the curvature of the primordial fields. Therefore, the helicity h of the field is proportional to the curl of the trajectory $\mathbf{x}(t)$:

$$\begin{aligned}
\nabla \times \mathbf{x}(t) &= \frac{q}{m} \int_0^t \int_0^{t'} (\nabla \times \mathbf{E} + \nabla \times (\mathbf{v} \times \mathbf{B})) \, dt'' dt' \\
&= \frac{q}{m} \int_0^t \int_0^{t'} \left(-\frac{\partial \mathbf{B}}{\partial t} + \nabla \times (\mathbf{v} \times \mathbf{B}) \right) dt'' dt' \\
&\approx \frac{q}{m} \int_0^t \int_0^{t'} \nabla \times (\mathbf{v} \times \mathbf{B}) \, dt'' dt',
\end{aligned}$$

where I have made use of the fact that the extragalactic magnetic fields we are interested are primordial and will not

[4] The electron or positron could theoretically also scatter off of EBL photons, but the relevant cross-section here actually *decreases* with increasing photon energy, so the CMB ($\varepsilon_{CMB} \approx 1$ meV) is energetically preferred to the EBL ($\varepsilon_{EBL} \approx 1$ eV). Additionally, the number density of the CMB is greater than that of the EBL by about a factor of 10^5, making an up-scattering event much more likely to occur for the former.

[5] The splitting function for the outgoing positron-electron pair does allow for an uneven division of the excess center-of-mass energy between the two, so long as momentum is still conserved. However, because the probability distribution is symmetric, each of the two particles will generally take off with half of the excess energy available, on average.

[6] Although the more granular spiraling about the field line will be reversed in motion for a positron compared to an electron, the net motion is still along the field line itself. It is the larger arc of motion that we are interested in and that is able to be resolved by up-scattered photons that intercept this path.

change significantly over the time of propagation, so the term $\partial \mathbf{B}/\partial t$ is expected to be extremely small. We then write that

$$h \sim \frac{q}{m} \int_0^t \int_0^{t'} \nabla \times (\mathbf{v} \times \mathbf{B}) \, dt'' dt'. \tag{1.2}$$

1.2 A Tangible Signal From the Data at Hand

Evaluating the integral (1.2) to obtain the expected electron trajectories is still not possible without additional knowledge of $\mathbf{B}$, but doing so is not actually necessary for our purposes. Even if we could calculate the exact electron trajectories a priori, we still would not be able to predict the locations at which scattering with the CMB would occur, or which of these photons would make it to the LAT at all.[7] Then there is the problem of determining which of the photons that are detected by the LAT are actually the result of these AGN-related happenings and not some completely unrelated process. For these reasons, we restrict our interest to a single quantity that is – like the right hand side of Eq. (1.2) – proportional to the helicity of the fields, but is scaled to fall only between -1 (totally left-handed) and $+1$ (totally right-handed).

This quantity, which we shall refer to hereafter as Q, must be able to probe the helical nature of the fields only through the detection of gamma-rays' locations and energies, as these are the only physical observables that the Fermi-LAT can measure.[8] We know that if we were somehow able to meticulously track the cascade of photons resulting from the scatterings of only a single electron while ignoring all other sources on the sky, these gamma-rays would inarguably be decreasing in energy, as the electron responsible for the up-scattering is able to impart less momentum with each successive impact. Furthermore, we have already established that the general shape of the electron's trajectory will be some sort of helical curve due to the Lorentz force. The minimum number of points to define a simple interpolatable curve is three, so establishing our prototype signal as one derived from the locations of three photons with decreasing energies $E_3 > E_2 > E_1$ is a good place to start.[9]

The locations of the photons of interest can be assigned unit vectors $\hat{\mathbf{n}}_3, \hat{\mathbf{n}}_2, \hat{\mathbf{n}}_1$ on the sky, with $\hat{\mathbf{n}}_3$ corresponding to the location of the photon with energy E_3, etc. The handedness of the spiral between the three points is then determined by constructing a triple product with the vectors that connect the locations of the photons in space (the subscript p stands for "prototype"):

$$\begin{aligned} Q_p &= [(\hat{\mathbf{n}}_1 - \hat{\mathbf{n}}_3) \times (\hat{\mathbf{n}}_2 - \hat{\mathbf{n}}_3)] \cdot \hat{\mathbf{n}}_3 \\ &= (\hat{\mathbf{n}}_1 \times \hat{\mathbf{n}}_2) \cdot \hat{\mathbf{n}}_3, \end{aligned} \tag{1.3}$$

using the fact that $\hat{\mathbf{n}}_i \times \hat{\mathbf{n}}_i = 0$. This quantity does indeed take on values only in the range $[-1, 1]$ while characterizing the handedness of the spiral from which is it composed. Additionally, the quantity is sensitive to parity transforma-

[7] See footnote 1.

[8] There has been some research into how to investigate the polarization of detected gamma-rays, but the results are preliminary and beyond the scope of this analysis.

[9] We are setting up a prototype signal based on only one hypothetical electron's trajectory, so the energies E_a represent particular values of that electron's up-scattered photon counterparts. In practice, we will not know the energies beforehand and the three energies inputted for analysis must come instead from three ranges of energies instead. The ranges of energy values allowed in our analysis will be discussed shortly.

tions, as it is a pseudoscalar.[10] The geometry of the unit vectors used in computing a value of Q_p is shown in Figure 1.1.

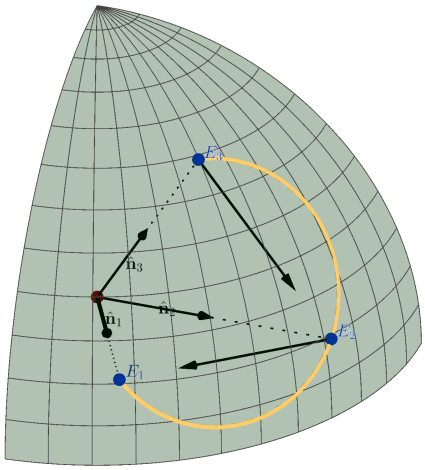


Figure 1.1: Geometry of the triple product used in computing a prototype signal Q_p. Three detected gamma-ray locations (blue) with energies E_3, E_2, and E_1 are chosen, with unit vectors pointing from the LAT to their locations on the sky drawn accordingly. Ordered by decreasing energies $E_3 > E_2 > E_1$, the differences between the corresponding unit vectors are then used as in Eq. (1.3) to compute the scalar triple product. The orientation of the gold spiral corresponds to the helicity of the signal. In this example, Q_p is positive (right-handed). The angular distances between the photons shown here have been exaggerated for ease of viewing.

To account for our lack of knowledge regarding whether the photons actually originated from an AGN-sourced up-scattering process, we simply average over all possible triplets of photons that meet our criteria. Miscellaneous photons included from other processes that still happen to fall within our criteria for data collection will contribute nothing to the signal, on average. Taking all of this into account, the parity-odd correlator can be written as

$$Q_{E_1E_2E_3}(r) = \frac{1}{N_1N_2N_3}\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\sum_{k=1}^{N_3}\left\{\hat{\mathbf{n}}_i(E_1)\times\hat{\mathbf{n}}_j(E_2)\right\}\cdot\hat{\mathbf{n}}_k(E_3)H(r-\alpha_{ik})H(r-\beta_{jk}) \tag{1.4}$$

where N_3, N_2, N_1 are the numbers of photons with energies E_3, E_2, E_1, respectively, r is size of the radial patch surrounding each k^{th} "anchor" photon with energy E_3, and α_{ik} and β_{jk} are the angles between the unit vector pairs $\{\hat{\mathbf{n}}_i(E_1), \hat{\mathbf{n}}_k(E_3)\}$ and $\{\hat{\mathbf{n}}_j(E_2), \hat{\mathbf{n}}_k(E_3)\}$, respectively. $H(x)$ is the Heaviside step function, which is defined in the following way:

$$H(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{cases}$$

In context, Heaviside's function is inserted as a filter to ensure that only photons within an angular distance r of the anchor photons are included in the analysis.

[10] Evidently, pseudoscalars – whether they be in the form of mathematical abstractions like Q or actual particles like the K_0 – are the vessels for which CP violating signals are best extracted.

1.3 Characteristics of the Correlator Q

In calculating the signal, not all high energy photons labeled E_3 will have the same energy, nor will the lowest energies E_1 be the same. For this reason, we adopt the convention that E_1 in Eq. (1.4) corresponds to any energies lying in the range $[E_1, E_1 + \Delta E]$, and similarly for the E_2 and E_3 energy photons. The actual anticipated energies of photons up-scattered in the processes mentioned before have been calculated by Vachaspati and Tarishimo to lie between 10 and 60 GeV [11]. If we choose our bin width to be $\Delta E = 10$ GeV, this provides us 5 bins with baseline energies of $\{10, 20, 30, 40, 50\}$ GeV to choose from. Then if we always choose 50 GeV as the E_3 bin baseline, we are left with $\binom{4}{2} = 6$ possible combinations of energy bin triplets on which we can perform our analysis.

Another ambiguity in our definition of Q is the radius r of the patch enclosing each anchor photon. As the magnitude of the magnetic fields of interest are not known, it is not possible to predict the size of the helical trajectories carved out by the electrons as they up-scatter nearby photons. For this reason, we perform each analysis of Q 20 times, for each integer value $r \in [0°, 20°]$. After performing the analysis for each radial patch size, we will be able to plot our 6 Q values (one for each energy bin combination) as functions of r. We use the notation $Q_{E_1 E_2}(r)$ henceforth to represent these functions, omitting the third energy bin label since we have always chosen $E_3 = 50$ GeV.

It is important to note that "choosing a patch size" does not amount to drawing all E_1 and E_2 photons from a single region of interest (ROI) on the sky. Rather, for *every* anchor photon with energy E_3, a new circular region of interest with radius r is constructed around this point on the sky. Then, all eligible E_1 and E_2 photons within range of this particular anchor photon are used to calculate numerous values of Q. For example, if there are 1000 photons with $E_1 = 10$ GeV and 500 photons with $E_2 = 30$ GeV within 5° of a particular E_3 photon, then this single E_3 photon alone will contribute $500 \times 1000 = 5 \times 10^5$ terms to the ensemble average that makes up our signal for $Q_{10,30}(r = 5°)$. Conversely, just as each E_3 photon may contain many lower energy photons within its ROI, any particular E_1 or E_2 photon may be included in the ROI of many different anchor photons and thus can contribute to many different terms of the sum (1.4) as well. Techniques used to actualize such extensive calculations as those implied by Eq. (1.4) are discussed more thoroughly in the Methods (Chapter 2).

Finally, there is the issue of signal contamination due to the galactic plane. We noted earlier that miscellaneous photons that make their way into the analysis will contribute nothing to the signal on average, but this is under the assumption that the spatial distributions of the photons are homogeneous in solid angle (i.e. $\mathrm{d}N/\mathrm{d}\Omega \approx$ const.) As is well known amongst those familiar with Fermi data, our positioning within the galactic plane inherently gives rise to a large artificial excess[11] of gamma rays, particularly for low galactic latitudes, $|b| < 50°$. The simplest way to avoid contamination and non-homogeneity near the plane is to avoid it entirely in our analysis. We then restrict our attention to the galactic poles, where $|b| \geq 50°$. The largest patch radius that we will be using is $r = 20°$, so our anchor photons must then be further restricted to $|b| \geq 70°$ corresponding to a galactic zenith angle cutoff of $\theta_{max} = 20°$.

[11] The excess is "artificial" only for our purpose of trying to look far beyond our own galaxy, but being unable to verify which photons in this region are created in our own galaxy and which have AGN origins.

Chapter 2

Methods

The correlator Q has been calculated with data drawn entirely from photons detected by the Fermi-LAT between weeks 9 and 521 of operation. Selecting the appropriate data for analysis is done via the Fermi Science Tools, which are available to the public for use with the LAT Data Server hosted by NASA. With the selected data in hand, there is then the issue of maximizing efficiency and accuracy in calculating Q itself. Each of these issues will be discussed in this chapter in turn.

2.1 The Fermi Large Area Telescope

Fermi was launched June 11, 2008 and soon after the LAT began accumulating gamma-ray data, collecting photons with energies ranging from 20 MeV to greater than 300 GeV [2]. The LAT is the primary instrument carried on the Fermi Gamma-Ray Space Telescope, which is supplemented by the Gamma-Ray Burst Monitor (GBM) in detecting and tracking potential torrents of photons of lower energy (8 keV-30 MeV) from a single location over a very short period of time. The data we are interested in for this analysis is primarily from the LAT, as we are interested in full-sky, long-exposure data.

The detection of a gamma-ray by the LAT relies on an induced pair-production into a positron-electron pair, both of which can be tracked to extrapolate the original energy and direction of the incident photon. The LAT has therefore been constructed to make this process as efficient, accurate, and precise as possible. This is accomplished with the coordinated efforts of the Tracker, Calorimeter, Anticoincidence Detector, and Data Acquisition System, the four of which are the essential components of the LAT.

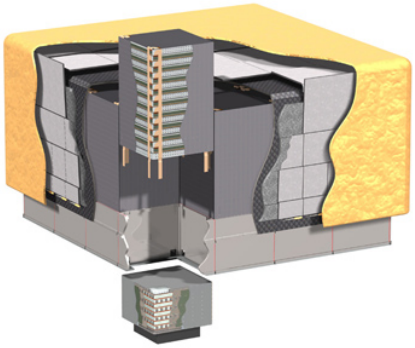


Figure 2.1: LAT cutaway showing one of the 16 modular towers. The top of each tower holds the Tracker module, while the Calorimeter and Data Acquisition System lie in the lower module. Figure from Atwood et al., 2007. [2]

The Anticoincidence Detector (ACD) consists of several plastic tiles that essentially make note of any charged particles (mostly cosmic rays) that enter the LAT to prevent them from giving rise to false signals. When a gamma-ray approaches the LAT, however, it will pass undetected through the ACD, subsequently interacting with 1 of 16 thin tungsten sheets which col-

lectively serve as a catalyst for pair production inside the LAT.

The Tracker and Calorimeter modules are segmented into 16 adjacent towers in a 4x4 array, one of which depicted in the LAT cutaway diagram in Figure 2.1. Each modular tower is constructed with 19 stacked trays at the front end, the first 16 of which hold tungsten conversion sheets. Each tungsten sheet is then immediately followed by two silicon strip detectors (SSDs): one for x coordinate tracking and another for y, for a grand total of 36 SSDs per tower. A standard event trigger requires the registration of 3 contiguous (x, y) coordinate pairs within the same module after conversion. For this reason, the last two trays contain only the x and y SSDs, as a pair conversion via tungsten sheet this deep into the tower would not allow for sufficient positioning data to be collected. A schematic of the pair conversion and tracking process throughout the Tracker trays is depicted in Figure 2.2.

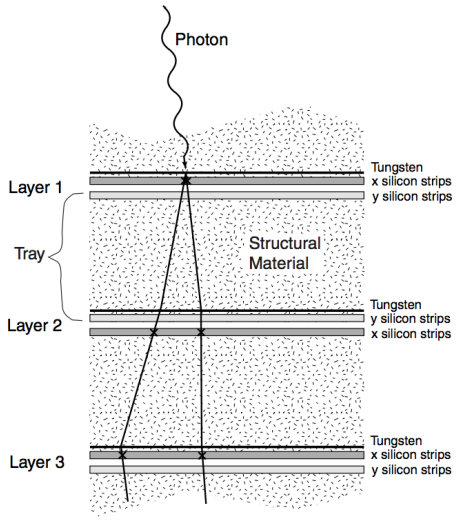


Figure 2.2: Heuristic schematic of the front end of a Tracker tower. A gamma-ray incident upon the first tungsten sheet converts to a positron-electron pair whose x and y coordinates are recorded by the subsequent interwoven SSDs. Figure from Atwood et al., 2007. [2]

Finally, after traversing through the Tracker trays, the electron and positron will deposit their energy into the Calorimeter, which itself is composed primarily of cesium-iodide crystals. The scintillation light produced by the calorimeter upon impact of the electron and positron is proportional to the energy delivered by the particles, which then inform us of the initial energy of the parent gamma-ray.

The Data Acquisition System (DAQ) lies at the back end of the LAT while gathering all of the collected information into a cohesive picture of the event corresponding to each detected photon while rejecting false signals from charged cosmic rays and gamma-rays produced in the Earth's own atmosphere. Once the photon data has been processed as a legitimate signal, the information can be uploaded to NASA's servers for post-processing to be released to the public for use in individual analyses.

2.2 Data Used for Analysis

2.2.1 The Pass 8 Event Post-Processing Package

The post-processing package used to analyze and categorize photon data between collection by the LAT and uploading to the Data Server is called Pass 8. Fermi was initially launched with Pass 6 in 2008, with Pass 7 being released in August of 2011. Pass 8 was then released in June 2015, with which the Fermi team has processed all data taken since. Additionally, all data recorded prior to the release of Pass 8 has been reprocessed with the newer algorithms as well.

The main improvements in Pass 8 over Pass 7 include a reduced background contamination, greater effective area, better point-spread function (PSF), and an extension of the range of reliably detectable energies into the sub-100

MeV range and those with energy above a few hundred GeV [1]. The specific re-workings of the algorithms from Pass 7 to Pass 8 is mostly beyond the scope of this paper but were primarily motivated by an improved understanding of the systematic errors accumulated over the first 5 years of data acquisition. A switch from a combinatoric tracking algorithm to a global tree-based one was crucial in refining the new PSF, while implementing a clustering stage in the calorimeter (as opposed to single-particle) led to an increase in effective area.

The improvements in Pass 8 also laid the foundation for a new event class called ULTRACLEANVETO (evclass=1024). The 1024 class is described as having a background rate 2-4 times lower than that of the SOURCE class (evclass=128) for photons with energies between 10 MeV and 100 GeV – the range in which our photons of interest lie. As mentioned before, the original analysis by Tashiro and Vachaspati used Pass 7 data [11], whose cleanest class at the time was equivalent to Pass 8's ULTRACLEAN (evclass=512). The photons processed into the newer ULTRACLEANVETO class should give rise to a more reliable signal, although reflecting this in our analysis will require the convolution of the current PSF with the data simulated to calculate the 68% and 95% containment bands of the signal Q. This will be discussed further in the Analysis.

Regardless of event class, however, we are now armed with 10 full years of data which will unequivocally put a tighter constraint on the statistical significance of Q, as the uncertainties and containment band widths are proportional to $1/\sqrt{N_3}$, where N_3 is the number of anchor photons with energies between 50 and 60 GeV.

2.2.2 Further Cuts With the Fermi Science Tools

The Fermi Science Tools enable one to make additional cuts on data retrieved from the Data Server. Our starting point was the Pass 8 Weekly Files – which are recommended for all-sky analyses – utilizing weeks 9 through 521. We used the `gtselect` tool to select only ULTRACLEANVETO (evclass=1024) photons with front and back processing (evtype=3) with energies between 10 and 60 GeV.

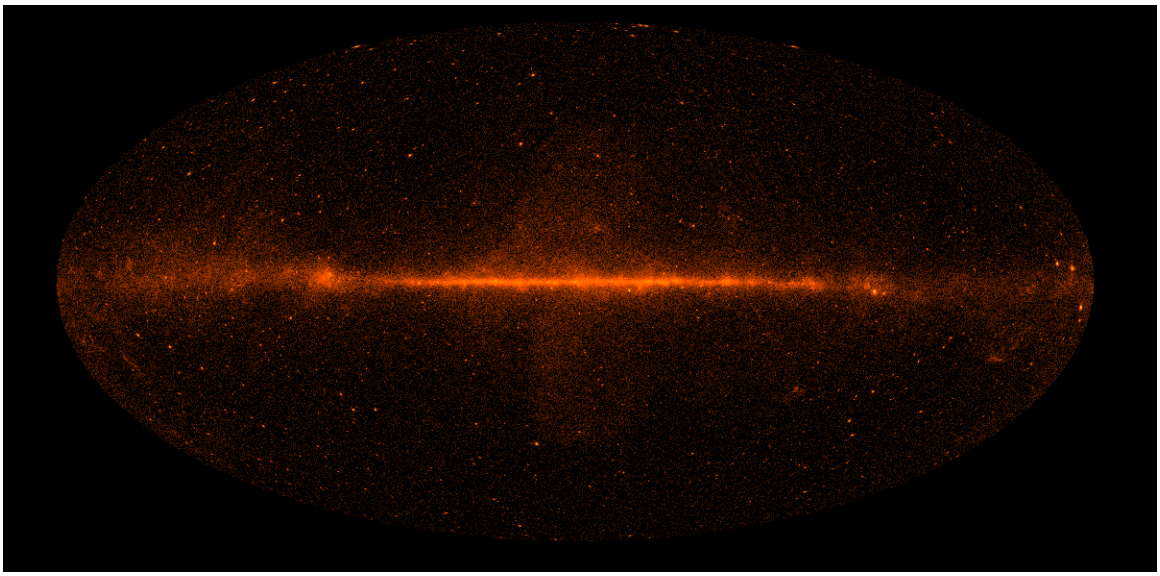


Figure 2.3: $2D$ map of ULTRACLEANVETO photon density (log-scaled) after final cuts were made with `gtselect` and `gtmktime`. The usual Aitoff projection is used to transform the galactic coordinates $(l, b) \mapsto (x, y)$. The galactic plane clearly contributes the bulk of the flux, especially near the galactic center where the "Fermi bubbles" balloon outwards from the plane up to $|b| \approx 50°$. Point sources near the galactic plane are primarily pulsars, while the overwhelming majority of points off of the plane are AGNs.

The corresponding Spacecraft Data file was used in conjunction with the all-sky photon files to filter out photons that were not taken during "good time intervals" (GTIs). The Fermi satellite orbits the Earth every ~ 96 minutes, alternating rocking directions with each orbit. Only $\sim 1/5$ of the sky is visible at any given time, and even if a particular source is in the field of view, there are phenomena such as the South Atlantic Anomaly that can interfere with data collection. To filter out photons that do not have a GTI, the `gtmktime` tool was used with the parameters `(DATA_QUAL>0)&&(LAT_CONFIG==1)&&(ABS(ROCK_ANGLE)<52)`. These are the standard parameter choices recommended by the Fermi team for which the satellite is in normal scanning mode. After the `gtselect` and `gtmktime` tools have been used, the data is ready for use in calculating the signal Q. A count-density map of the photon data used for analysis following these cuts is shown in Figure 2.3.

2.3 Calculation of the Correlator Q

If we were to perform our analysis with a set of totally random data (scrambled energies and positions) one would expect to measure a null signal ($Q \approx 0$), but our simulations showed that this was almost never the case. This begs the question: How do we know that our results are statistically significant in implying a CP violating signal, and are not any more meaningless than those of a totally random data set? To investigate the statistical significance of our signal, then, we generated 500 sets of random data to be analyzed by the same algorithm implied by Eq. (1.4), which we place here for convenience:

$$Q_{E_1E_2E_3}(r) = \frac{1}{N_1N_2N_3}\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\sum_{k=1}^{N_3}\left\{\hat{\mathbf{n}}_i(E_1)\times\hat{\mathbf{n}}_j(E_2)\right\}\cdot\hat{\mathbf{n}}_k(E_3)H(r-\alpha_{ik})H(r-\beta_{jk}) \tag{2.1}$$

Recall that we aim to calculate $Q_{E_1E_2}(r)$ for 6 different combinations of energies E_1, E_2, E_3, each for patch radii in the range $r \in [0°, 20°]$. The analysis of a single (real or simulated) data set then requires the evaluation of $20 \times 6 \times (N_{10}N_{20}N_{50} + N_{10}N_{30}N_{50} + N_{10}N_{40}N_{50} + N_{20}N_{30}N_{50} + N_{20}N_{40}N_{50} + N_{30}N_{40}N_{50})$ individual triple products of the form $\hat{\mathbf{n}}_i \times \hat{\mathbf{n}}_j \cdot \hat{\mathbf{n}}_k$, where N_{10} is the total number of photons with 10 GeV in the data set, etc.

While performing this computation once is manageable, performing it several thousands of times is not. This, however, is exactly what is required in order to obtain the error bars for various values of $Q_{E_1E_2}(r)$. As the number of simulations increases, the average of the simulated signals $\langle Q \rangle$ tends towards zero, and thus the variance in Q approaches a constant value.[1] To address the massive amount of computation that must be carried out in order to produce the containment band widths $\sigma_{\langle Q \rangle}$, we can exploit the linearity of the cross product. Instead of iteratively

[1] An investigation of this limiting process of $\sigma_{\langle Q \rangle}$ is discussed in Appendix B.

running through each of the three sums over E_i, E_j, and E_k, we may first expand the triple product:

$$\begin{aligned}
\sum_{ijk}\left\{\hat{\mathbf{n}}_i(E_1)\times\hat{\mathbf{n}}_j(E_2)\right\}\cdot\hat{\mathbf{n}}_k(E_3) &= \sum_{ij}\left\{\hat{\mathbf{n}}_i(E_1)\times\hat{\mathbf{n}}_j(E_2)\right\}\cdot\left(\sum_k\hat{\mathbf{n}}_k(E_3)\right)\\
&= \sum_i\left\{\hat{\mathbf{n}}_i(E_1)\times\hat{\mathbf{n}}_1(E_2)+\hat{\mathbf{n}}_i(E_1)\times\hat{\mathbf{n}}_2(E_2)+\cdots+\hat{\mathbf{n}}_i(E_1)\times\hat{\mathbf{n}}_{N_2}(E_2)\right\}\cdot\left(\sum_k\hat{\mathbf{n}}_k(E_3)\right)\\
&= \sum_i\left\{\hat{\mathbf{n}}_i(E_1)\times\left(\sum_j\hat{\mathbf{n}}_j(E_2)\right)\right\}\cdot\left(\sum_k\hat{\mathbf{n}}_k(E_3)\right)\\
&= \left\{\left(\sum_i\hat{\mathbf{n}}_i(E_1)\right)\times\left(\sum_j\hat{\mathbf{n}}_j(E_2)\right)\right\}\cdot\left(\sum_k\hat{\mathbf{n}}_k(E_3)\right).
\end{aligned}$$

We have reduced the triple sum to a product of sums resembling averages. We have however neglected the presence of the two Heaviside functions which depend on i, j, and k. We can compromise by writing

$$Q_{E_1E_2E_3}(r) = \frac{1}{N_3}\sum_{k=1}^{N_3}\left\{\langle\hat{\mathbf{n}}(E_1)\rangle_{\in R_k}\times\langle\hat{\mathbf{n}}(E_2)\rangle_{\in R_k}\right\}\cdot\hat{\mathbf{n}}_k(E_3). \tag{2.2}$$

Here, $\langle\hat{\mathbf{n}}(E_1)\rangle_{\in R_k}$ represents the average location of all unit vectors pointing to E_1 photons within the radial patch R_k surrounding the k^{th} anchor photon, and likewise for $\langle\hat{\mathbf{n}}(E_2)\rangle_{\in R_k}$.

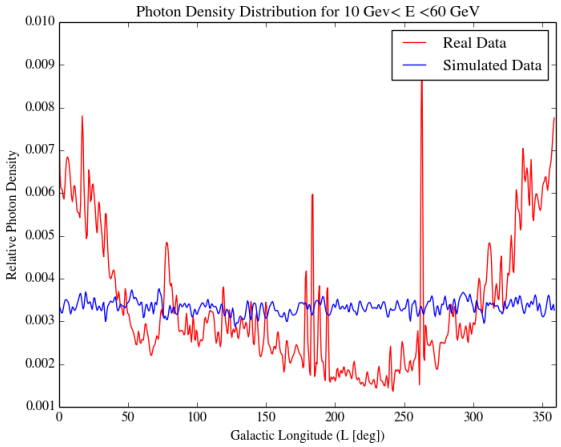
Mathematically, Eq. (2.2) is equivalent to Eq. (1.4), but the two expressions are computationally very different. By implementing an algorithm resembling Eq. (2.2), we only need to check the positions of each E_1 and E_2 photon once for each value of k. Thus we need only to calculate $120\times N_3$ triple products, as opposed to the more formidable number mentioned in the preceding paragraph.

2.4 Generating Simulated Data

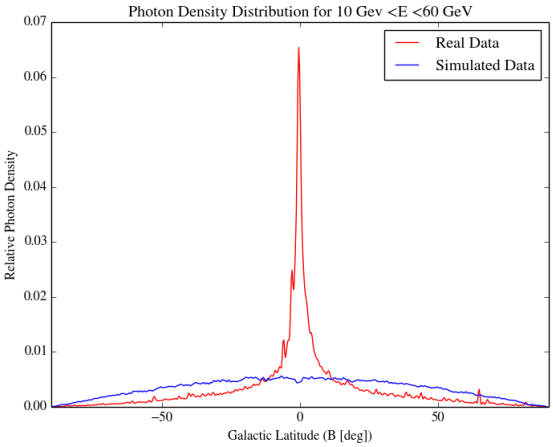
As mentioned previously, the distribution of gamma-rays in the universe is expected to be homogeneous and isotropic such that $\mathrm{d}N/\mathrm{d}\Omega$ is constant on average. The radiation due to the galactic plane of our own galaxy complicates things considerably. Comparisons of the actual and ideal distributions of photons observed by the LAT with energies between 10-60 GeV are shown in Figures 2.4a and 2.4b. We noted earlier that we would restrict our field of view entirely to photons that lie outside the galactic plane, i.e. those with $|b|> 50°$. Clearly this would take care of the massive central spike in Figure 2.4b, and doing so would simultaneously flatten the distribution in Figure 2.4a as well, as the two distributions are inexplicably correlated. There still remain sharp irregularities throughout the remaining sections of the distributions, though.

Addressing the issue of these irregularities is not a trivial task. Some of the sharper points are due to known gamma-ray sources (mostly AGNs), while other areas of deviation from the expected, flat distributions cannot be explained away as easily. The problem remains then: How can we generate our containment bands for Q using simulated data that does not reflect the data we are given in reality? Removing photons from known gamma-ray sources will re-

Figure 2.4: Actual photon distributions (red) and simulated data (blue) generated from the expected flat distribution (+ Gaussian noise) with $d\Omega = d(\cos b)\,dl$. All distributions have been normalized such that $\int (\partial N/\partial \cos b)\, d(\cos b) = \int (\partial N/\partial l)\, dl = 1$.



(a) Comparison of real and ideal distributions over l.

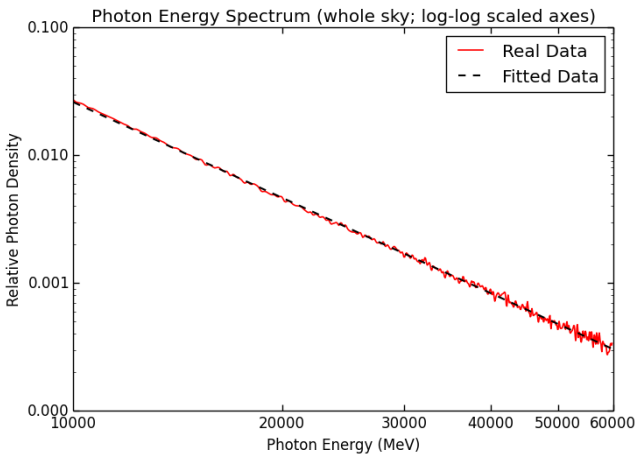


(b) Comparison of real and ideal distributions over b.

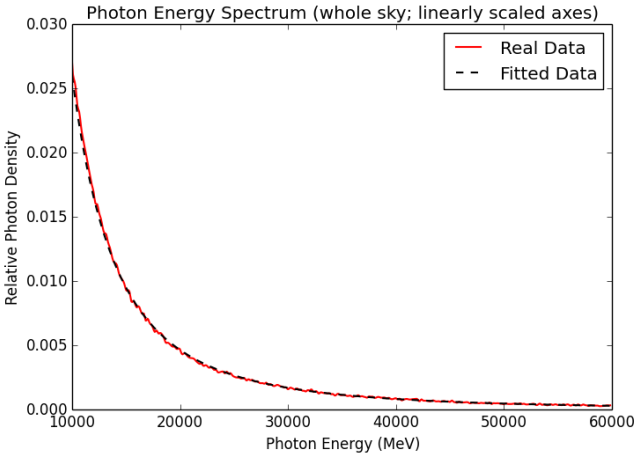
move some of the most offensive irregularities – this is what is required of a so-called "diffuse" analysis, typically with a masking radius of 1.5° for each source.[2] We found that, even with point sources and galactic plane removed, the ideal homogeneous distributions still did not faithfully represent the real ones.

For these reasons, we opted to generate our simulated data from the observed distributions themselves (red in Figures 2.4a and 2.4b) via inverse transform sampling. This ensures that our simulations will not endure any bias due to an artificially flat distribution of photons, and this will in turn guarantee that our resulting error bands are meaningful with regards characterizing the strength of the true observed signal.

Figure 2.5: Actual photon energy spectrum (red) and fitted distribution via linear regression (dashed black) over log-log (left) and linear (right) axes scaling. The distributions have been normalized to unity.



(a) Log-log scaled energy distribution (red) and regression line (dashed black), with slope $m = -2.50 \pm 4 \times 10^{-3}$.



(b) Linearly scaled energy distribution (red) and power law fit (dashed black). For energies $E \in [10, 60]$ GeV, $\partial N/\partial E \sim E^{-2.5}$.

We have not yet discussed the energy distribution of detected photons. Guessing the distribution of photon

[2] In our Analysis, we will examine the signal for both diffuse (Phase I) and source-anchored (Phase II) scenarios.

energies a priori would be a colossal undertaking and is beyond the scope of this report. The more predictable and well understood components of the spectrum (blackbody distributions of CMB and EBL, light from stars) are many orders of magnitude below the energy range that we are probing, and a deep understanding of the most energetic phenomena in our universe (blazars, quasars, nebulae, etc.) as well as all of the processes in our own galaxy would be necessary to predict the spectrum observed by Fermi. However, because the bandwidth of interest is relatively narrow ($10 - 60$ GeV), we might guess that a power law would suffice in describing the distribution empirically.

In Figure 2.5a, the observed energy spectrum is plotted across log-log scaled axes. A linear regression routine reveals that the slope of the transformed distribution (red) is $m \approx -2.5$. The resulting regression line is plotted in dashed black alongside the data. This same value of m is then used in Figure 2.5b to re-express the extrapolated distribution (dashed black again) on set of linearly-scaled axes against the true energy spectrum (red) as a simple power law.

Figure 2.5b reveals that, to a decent approximation,[3] $\partial N/\partial E \sim E^{-2.5}$. However, it is not necessary to formally analyze the residuals to see that that the fit in Figure 2.5a is not exact. Instead of attempting to model the true distribution's fluctuations from the regression line by adding Gaussian noise or convolving with a model PSF, we appeal once more to the method of inverse transform sampling. All energy values used for simulations are then drawn directly from the true energy spectrum.

[3] The coefficient of determination for the linear regression is $R^2 = 0.998$.

Chapter 3

Analysis & Results

The first phase of the analysis involved a calculation of the signal $Q_{E_1E_2}(r)$ using only diffuse data with a masking radius of 1.5°, inspired by Tashiro and Vachaspati's original paper [11] but with newer and more plentiful Pass 8 data. The second phase, however, involves just the opposite: Using known gamma-ray sources as the anchoring points for each helix, instead of anchoring with E_3 photons. The reasoning for this is that, if the primordial magnetic fields are helical in nature, plugging in the locations of the expected centers of each helical cascade of photons should certainly boost the signal. If the signal is the same or even weaker than that derived from the diffuse phase, then this would be a strong indicator that our method of extrapolating the helicity of the extragalactic fields is flawed or insufficient, or that the fields are not helical at all.

For the diffuse analysis, we examined the northern and southern hemispheres separately, and then again as a whole. Additionally, for both the diffuse and source-anchored phases, we used galactic zenith cutoff angles of $\theta_{max} = 10°$ and $\theta_{max} = 20°$ for E_3 photons, corresponding to $b \geq 60°$ or $b \geq 50°$ respectively for the remaining lower energy photons.[1]

3.1 Phase I: Diffuse Analysis

3.1.1 E_3 Zenith Cutoff: $\theta_{max} = 10°$

A zenith cutoff of 10° on the coordinates of each E_3 photon corresponds to a global cut of $b \geq 60°$ for the remaining E_1 and E_2 photons. This is to ensure that cascade spirals can be constructed for triple products around each E_3 photon up to a radius of $r = 20°$. Giving each E_3 photon the same amount of room to sow its roots ensures that there is no artificial gradient in photon density that could bias the signal.

As mentioned in Section 2.3, the containment bands were constructed from the averaged results of $N = 500$ Monte Carlo simulations, each with photon energies and coordinates generated according to the real, observed distributions (red lines in Figures 2.4a, 2.4b, and 2.5b). This was achieved using inverse transform sampling, as described in 2.4. The results for $b \geq 60°$ are shown in Figure 3.1.

[1] For the southern hemisphere, the inequalities and signs of the angles would be reversed.

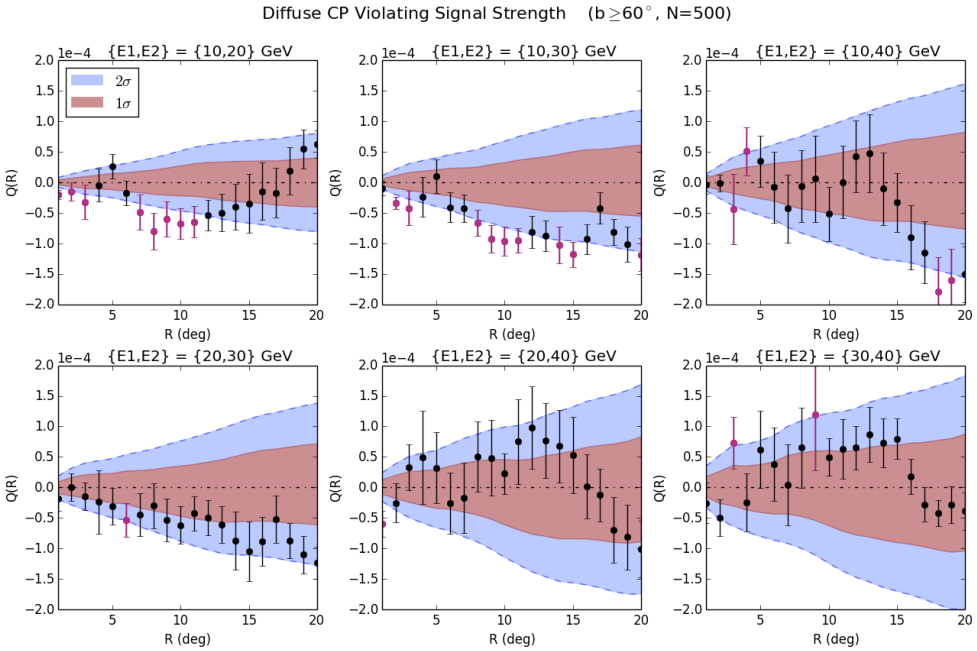


Figure 3.1: $Q_{E_1E_2}(r)$ for diffuse, northern hemisphere photons ($b \geq 60°$). The red and blue bands represent 68% and 95% containment, respectively, computed from $N = 500$ Monte Carlo simulations. Principal measurements of Q that lie outside the 2σ bands are plotted in magenta.

The first two subplots depicting $Q_{10,20}(r)$ and $Q_{10,30}(r)$ in Figure 3.1 above show a left-handed signal with strength $\approx (-1.0 \pm 0.3) \times 10^{-4}$ for $r \in (5°, 15°)$, while the signal stretches down as far as $(-1.8 \pm 0.5) \times 10^{-4}$ for $Q_{10,40}(18°)$ in the third subplot. These results are still consistent with a null signal, as 2σ and occasional 3σ values are expected. The plots for $Q_{20,40}(r)$ and $Q_{30,40}(r)$ seem to indicate a right-handed signal, but the errors σ_Q are equal to or even greater than the values of Q themselves. The wider 1σ and 2σ bands and error bars for the $Q_{10,40}(r)$, $Q_{20,40}(r)$, $Q_{30,40}(r)$ plots can be attributed simply to the fact that there are far less photons in the 40 GeV bin than there are in the 10, 20, and 30 GeV bins, as can be seen by glancing at Figure 2.5*b*. For this plot in particular, numbers of photons in each bin were $N_{10} = 6591$, $N_{20} = 1671$, $N_{30} = 686$, $N_{40} = 386$, and $N_{50} = 217$.

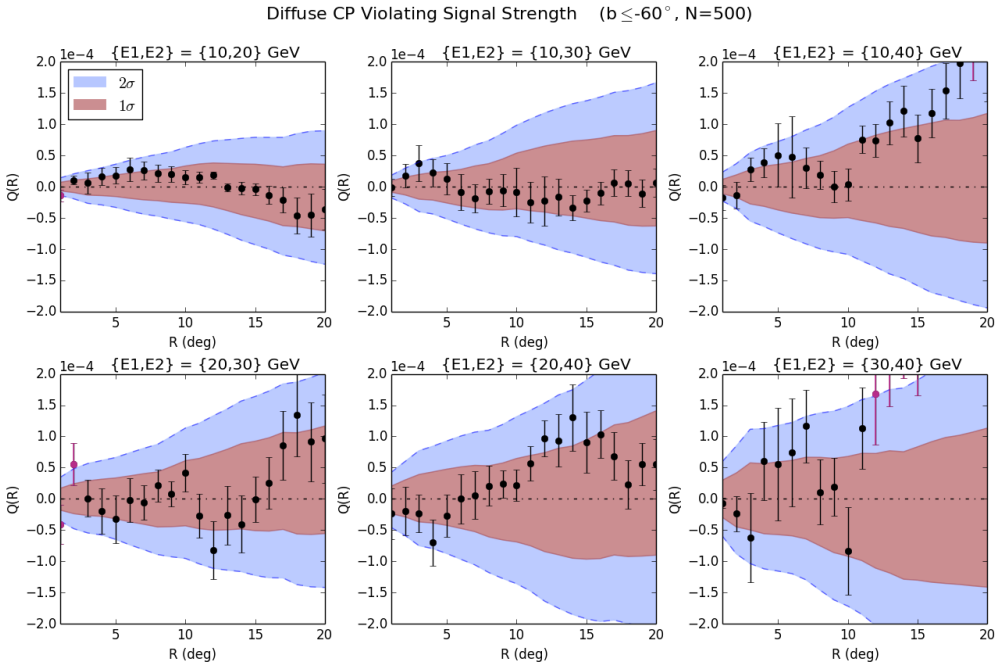


Figure 3.2: $Q_{E_1E_2}(r)$ for diffuse, southern hemisphere photons ($b \leq -60°$). The red and blue bands represent 68% and 95% containment, respectively, computed from $N = 500$ Monte Carlo simulations. Principal measurements of Q that lie outside the 2σ bands are plotted in magenta.

Figure 3.2 above shows the results of the same diffuse analysis, but for the southern hemisphere ($b \leq 60°$). The first two subplots for $Q_{10,20}(r)$ and $Q_{10,30}(r)$ no longer show a negative signal, but the remaining 4 subplots, especially the third ($Q_{10,40}(r)$) and sixth ($Q_{30,40}(r)$) show a positive signal of magnitude $\approx (2.0 \pm 0.9) \times 10^{-4}$ for $r > 12°$. In the $Q_{30,40}(r)$ subplot, the signal jumps to a maximum of $(5.0 \pm 1.0) \times 10^{-4}$ at $r = 17°$. The containment bands and error bars, however, are larger than that of the northern hemisphere in the previous plot because of Fermi's uneven exposure and sensitivity in the northern and southern hemispheres. For the southern hemisphere, the bin totals were $N_{10} = 5724, N_{20} = 1398, N_{30} = 625, N_{40} = 280$, and $N_{50} = 189$. Each bin contains less than its northern hemisphere counterpart.

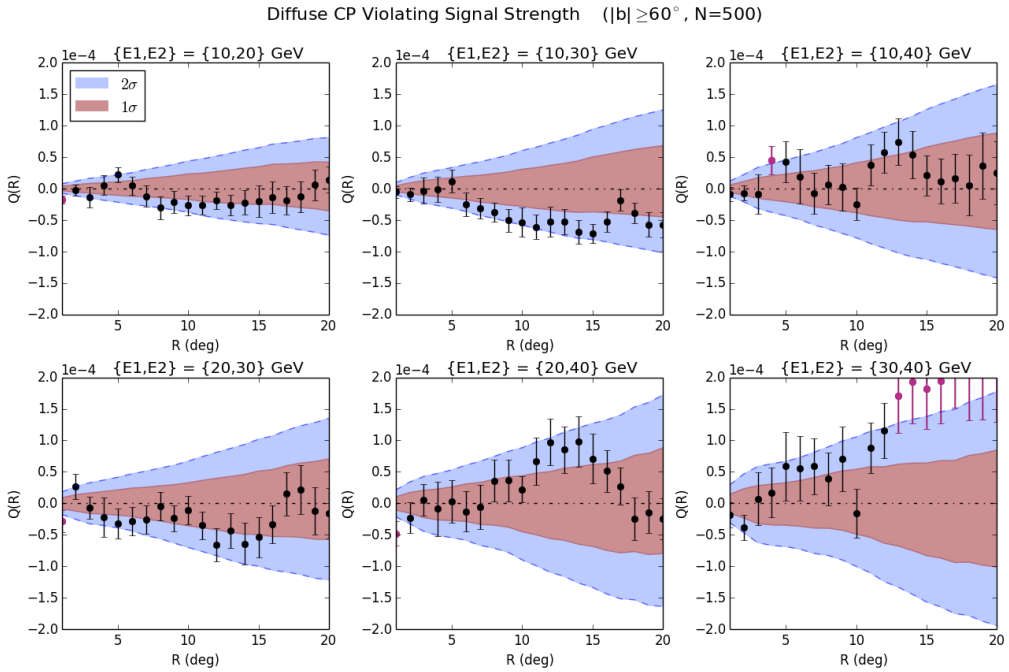


Figure 3.3: $Q_{E_1E_2}(r)$ for diffuse northern and southern hemisphere photons ($|b| \geq 60°$). The red and blue bands represent 68% and 95% containment, respectively, computed from $N = 500$ Monte Carlo simulations. Principal measurements of Q that lie outside the 2σ bands are plotted in magenta.

It is interesting to analyze the northern and southern hemispheres separately, but ultimately the signal that we are looking for should be extragalactic and isotropic: A relatively constant positive (right-handed) or negative (left-handed) helicity measured from any direction that we choose to look towards in the sky. Figure 3.3 shows the combined results for the northern and southern hemispheres for $|b| \geq 60°$. The bin counts for the combined northern and southern hemispheres were $N_{10} = 12315, N_{20} = 3069, N_{30} = 1311, N_{40} = 648$, and $N_{50} = 406$.

Most of the previous values of Q from the isolated northern and southern hemisphere cuts are flattened by averaging these results, with most measurements lying within 1σ of a null signal, and very few with significance greater than 2σ. An exception is the sixth subplot for which $Q_{30,40}(r)$ hovers at $(2.0 \pm 0.7) \times 10^{-4}$ for $r \geq 17°$. Were one to look only at the results for the combined northern and southern hemispheres, one might indeed claim a weak $(2\text{—}3\sigma)$ right-handed signal, but upon closer examination we have seen that same signal is absent in the northern hemisphere for $Q_{30,40}(r \geq 17°)$ (6th subplot of Fig. 3.1).

Figures 3.1, 3.2, and 3.3 have shown that the observed signal is unexpectedly dependent on where we look in the sky. There are several possible conclusions that could be drawn from this observation. The simplest explanation is that the signal we are observing does not have its origins in helical magnetic fields on an extragalactic scale, but rather

we are observing some other parity-odd process that is locally a function of the galactic latitude b. This does not imply that these helical primordial fields do not exist, but only that the observed signal is an artifact of some other process.

It could also be argued, however, that the signal component due to the primordial fields is indeed present in our results, but that it is being washed out by other processes, helical in nature or not. In other words, the signal-to-noise ratio is too low, meaning that there is either not enough data available representing CP-violating processes compared to other miscellaneous ones, or that Fermi is not sensitive enough to discern them (or both). The latter case is unlikely, as Fermi's angular resolution for GeV gamma-rays is $< 0.15°$ [1], and our radial patches of observation are $1-2$ orders of magnitude above this. Short of boosting the signal at the source itself[2], increasing the signal-to-noise ratio can then only be accomplished by either methodically filtering out unwanted noise (making additional cuts) or by adding more samples to the mean, as the uncertainties will decrease as the square-root of the number of sample measurements.

The most obvious additional cut to make would be to constrain our ROI even closer to the galactic poles, as the density of photons is more nearly homogeneous in those regions, as discussed in Section 2.4. Doing so, however, would consequently increase our uncertainties by providing less E_3 anchor photons with which to calculate Q. Unless we are able to conjure a routine that only filters out seemingly helical processes *not* due to primordial fields (which would require knowing the parameters of the fields a priori), our only other option is to increase the number of anchor photons to sample from, even if it requires dipping our ROI further towards into the galactic plane. We limit ourselves to increasing our ROI only from $|b| \geq 60°$ to $|b| \geq 50°$ to avoid oversaturating the photon density distributions with galactic plane gamma-rays.

3.1.2 E_3 Zenith Cutoff: $\theta_{max} = 20°$

The combined northern and southern hemisphere results for our extended ROI of $|b| \geq 50°$ are shown below in Figure 3.4. The numbers of photons in each bin are now $N_{10} = 22775$, $N_{20} = 5680$, $N_{30} = 2431$, $N_{40} = 1237$, and $N_{50} = 737$. Compared to the earlier $|b| \geq 60°$ analysis above, it is clear that including over double the number of photons in the analysis has only flattened the signal for all $Q_{E_1E_2}(r)$. There are remnants of the right-handed signal for $Q_{30,40}(r \geq 17°)$ in the sixth subplot, but the strength has diminished by a factor of ~ 4 to $(5.0 \pm 0.3) \times 10^{-5}$. Despite the contraction of the 1σ and 2σ bands, nearly every value of Q remains within 95% containment of a null signal.

An interesting common feature of the $Q_{10,40}(r)$, $Q_{20,40}(r)$, and $Q_{30,40}(r)$ subplots is a jump from a negative signal to a positive one near $r \approx 11°$. Although the significance of the values of Q themselves are low (less than 2σ and consistent with a null signal), the common appearance of this local minimum of the $Q_{E_1E_2}(r)$ for the three $E_2 = 40$ GeV combinations does not seem to be random.

[2] This is, as we will see in Section 3.2, the idea behind the source-anchored analysis in Phase II. For our diffuse analysis, this is not an option, however.

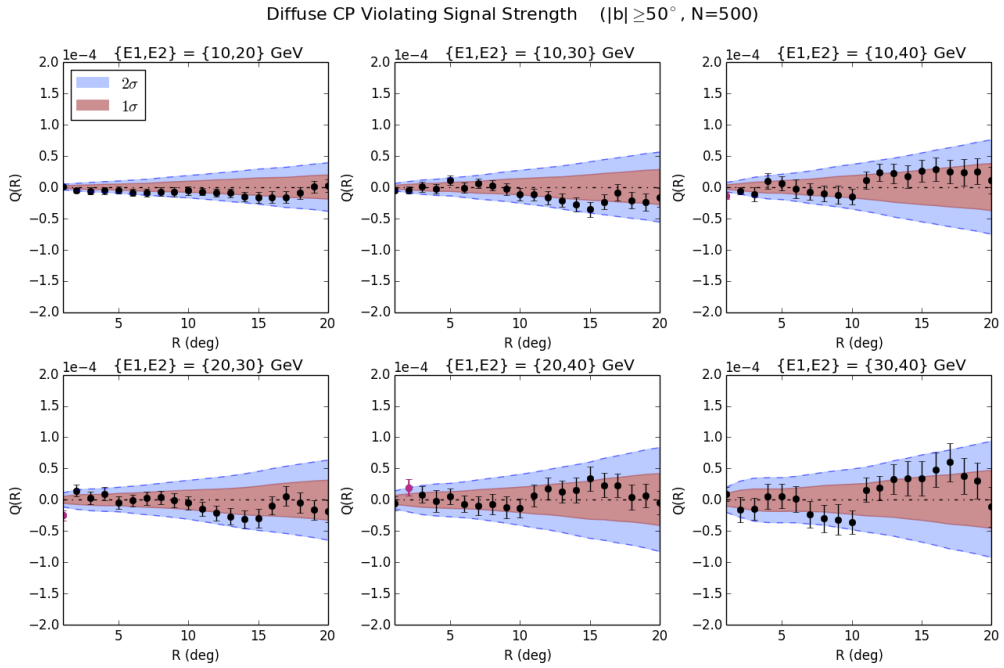


Figure 3.4: $Q_{E_1E_2}(r)$ for diffuse northern and southern hemisphere photons ($|b| \geq 50°$). The red and blue bands represent 68% and 95% containment, respectively, computed from $N = 500$ Monte Carlo simulations. Principal measurements of Q that lie outside the 2σ bands are plotted in magenta.

Because the deviations of the real photon distributions from the ideal, homogeneous ones grow faster than the size of the zenith angle cut, increasing our ROI will only flatten the signal-to-noise ratio even further. We then turn towards a non-diffuse analysis, now with the intention of indirectly boosting the signal source itself.

3.2 Phase II: Anchoring to Known Sources

The concept for this phase of the analysis is to replace all E_3 anchor photons with the actual locations of known AGNs. Specifically, we imported the locations of the 521 verified BL LAC-type objects with non-zero Fermi-LAT flux from the BZCAT multiwavelength AGN catalog [6]. Theoretically, this should boost the strength of any CP violating signal that is due to helical extragalactic magnetic fields.

Recall from Section 1.1 that the triple product of the locations of the E_3, E_2, and E_1 photons were used to characterize the helical nature of up-scattered EBL photons as electrons traverse the primordial magnetic fields after being sourced from pair-production near AGNs. If the E_3 photons are to be the closest to a given AGN after up-scattering, then using the actual locations of gamma-ray loud sources should be a more accurate characterization of the helical paths taken by the pair-produced electrons and positrons. Then, if all preceding assumptions about the primordial fields are correct, this method of analysis amounts to boosting the signal at its source, rather than attempting to increase the signal-to-noise ratio via additional cuts and filters.

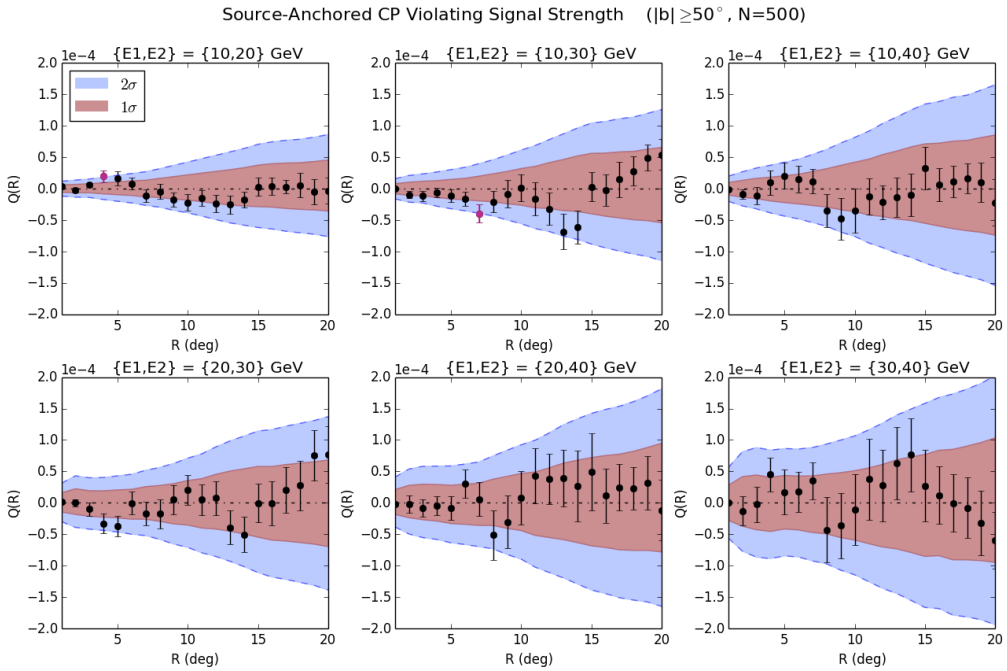


Figure 3.5: $Q_{E_1E_2}(r)$ for source-anchored northern and southern hemisphere photons ($|b| \geq 50°$). The red and blue bands represent 68% and 95% containment, respectively, computed from $N = 500$ Monte Carlo simulations. Principal measurements of Q that lie outside the 2σ bands are plotted in magenta.

In Figure 3.5 we have plotted the combined northern and southern hemisphere results for $|b| \geq 50°$ for our source-anchored analysis. There are values of Q that have greater magnitude than their diffuse counterpart (Fig. 3.4) such as the previous area of interest for $Q_{30,40}(r \geq 17°)$, but the increase in signal is marginal ($\sim 25\%$ increase) while there is still no discernibly right- or left-handed signal to be gleaned collectively from the 6 subplots. Each $Q_{E_1E_2}(r)$ meanders about the line $Q = 0$, with essentially all values lying within of the 95% containment of a null signal.

It is interesting to note that there is a definite similarity of the functions $Q_{E_1E_2}(r)$ for all but the first subplot in Figure 3.5: Each of these subplots features a local minimum near $r \approx 7°$. However, only the second subplot has a value that lies outside the 2σ band, with a value of $Q_{10,30}(7°) = (-3.9 \pm 1.4) \times 10^{-4}$. The small size of the other local minimums relative to the uncertainties does not say much in claiming a left-handedness of the signal at $r = 7°$, but it is interesting nonetheless, and we do not yet have an explanation for its appearance.

The main drawback of using known sources as anchor points for calculating triple products is that the Fermi source catalog, while impressive, provides much fewer points to anchor to in the first place. The total numbers of photons in each bin for this phase were $N_{10} = 29730$, $N_{20} = 7734$, $N_{30} = 3411$, $N_{40} = 1840$, and $N_{AGN} = 121$. Comparing the 121 AGNs within our ROI to the 737 E_3 photons available for the diffuse analysis, it is no wonder that the statistical significance of the source-anchored results would be much less, even if the principal values of Q are larger.

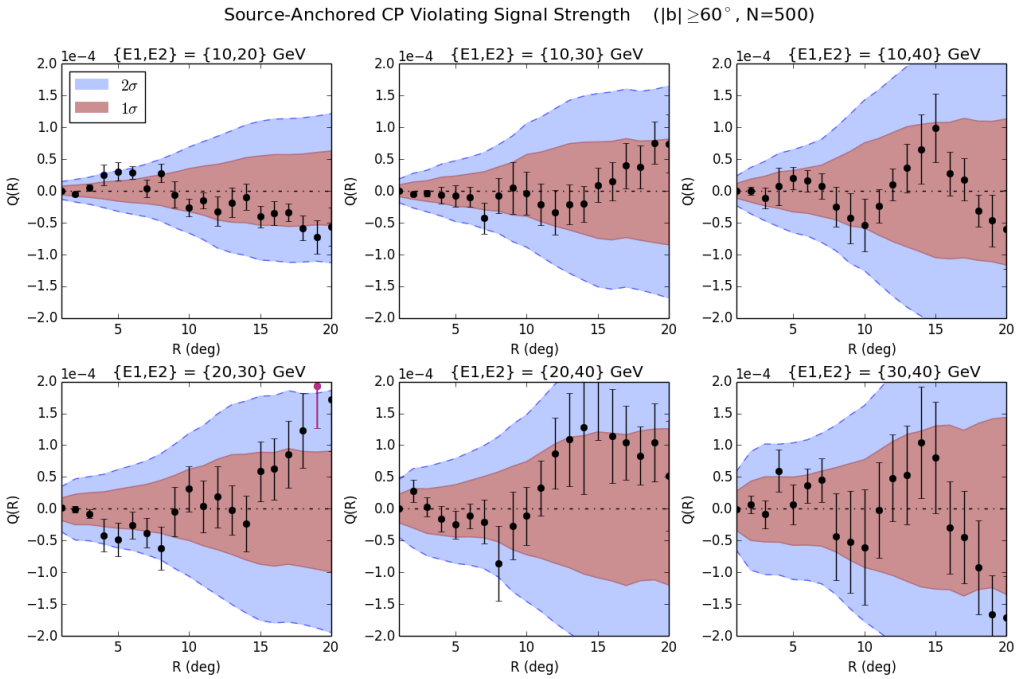


Figure 3.6: $Q_{E_1E_2}(r)$ for source-anchored northern and southern hemisphere photons ($|b| \geq 60°$). The red and blue bands represent 68% and 95% containment, respectively, computed from $N = 500$ Monte Carlo simulations. Principal measurements of Q that lie outside the 2σ bands are plotted in magenta.

Despite the even smaller sample size, we also include the source-anchored results for our smaller ROI of $|b| \geq 60°$ in Figure 3.6 above for completeness. The numbers of photons within the ROI were $N_{10} = 16401$, $N_{20} = 4295$, $N_{30} = 1889$, $N_{40} = 1040$, and $N_{AGN} = 71$. As expected, the magnitude of Q is greater as we close our ROI towards the poles, reflecting the results that we found above in the diffuse analysis of Phase I, but all values remain consistent with a null signal. The general structure of each subplot is also retained in going from $|b| \geq 50°$ to $|b| \geq 60°$ for the source-anchored analysis.

Chapter 4

Discussion & Further Work

Summarizing the results stated in the previous section into a single number or statement would not be possible. We have scanned over several parameters in search of a CP violating signal corresponding to the handedness of theorized primordial magnetic fields. These open parameters included the radius r encompassing each anchor photon as well as the choice of E_1 and E_2 bins, the zenith cut θ_{max} for anchor photons, as well as the type of anchor photons used (diffuse vs. AGN-sourced).

Looking at the diffuse and sourced-anchored results (Fig. 3.4 and 3.5) for our most inclusive ROI of $|b| \geq 50°$, we can say with confidence that there is no consistent right- or left- handed signal with greater than 2σ statistical significance. In other words, we have no reason to suspect that our results are due to anything more then the usual anticipated random deviations from $Q = 0$ that are expected to arise in such an analysis. Tashiro and Vachaspati originally claimed a significance in their left-handed signal with peak structure at $r \approx 12°$ of $\sim 2.7\sigma$. It is certainly possible that we had slight differences in our analyses of the data, but there still remains the fact that four additional years of Fermi data have been made available for my own analysis at the time of writing.

Going forward, there are actually more stringent criteria that I would need to subject my simulated data to, all of which would increase the size of the 1σ and 2σ containment bands. Each parameter ($r, E_1, E_2, \theta_{max}$, etc.) scanned over should take up an additional degree of freedom in our analysis which we would need to account for in our quoted error bars for each Q value. Additionally, I have neglected to account for the sensitivity of Fermi's PSF in my Monte Carlo simulations.

Then of course there is the fact that I used only $N = 500$ simulations due to time constraints. With more time (or access to more computing cores) I could certainly run more simulations, which may slightly alter the widths of the containment bands, but the effect would be negligible compared to that which would be done to increase uncertainty when accounting for the aforementioned effects. It should be noted, however, that the majority of the averaged simulation sample variances from which the containment bands are computed converge almost completely after only 300 simulations or so, as shown in Appendix B.

Appendix A

Electron-Positron Pair Production: Threshold and Cross-Section

As mentioned in the main text, the signal photons with energies $E_a \in [10, 60]$ GeV from which the correlator $Q_{E_1 E_2 E_3}(r)$ is calculated can have their origins traced back to those ultra high energy ($\sim$TeV) photons emitted from the jets of Active Galactic Nuclei. Once outside the strong radiation field of the AGN, these TeV energy photons can interact with the EBL to create positron-electron pairs. These newly created electrons and positrons then up-scatter CMB photons to GeV energies as they traverse a cooling length of several kpc, which Fermi can then detect.

The derivation for the total cross section for the process $\gamma + \gamma \to e^+ + e^-$ requires the full arsenal of quantumelectrodynamics and is beyond the scope of this report. We simply quote the result, first derived by Breit and Wheeler [3] in 1934:

$$\sigma_{BW}(\beta) = \frac{\pi r_e^2}{2}(1-\beta^2)\left[2\beta(\beta^2-2) + (3-\beta^4)\ln\left(\frac{1+\beta}{1-\beta}\right)\right]. \tag{A.1}$$

In the above formula, $r_e = 2.817 \times 10^{-15}$ m^2 is the classical electron radius, $\beta = \sqrt{1-(m_e c^2)^2/\varepsilon_1 \varepsilon_2}$ is the interaction parameter, and ε_1 and ε_2 are the energies of the two photons. The cross section is only nonzero[1] above threshold ($\beta > 0$), i.e. when the energy available in the center of momentum frame is just sufficient enough to produce an electron-positron pair at rest, as we shall soon see.

The kinematic constraints of the pair-production threshold can be derived semiclassically. First we will consider the more general case where the energy available in the center of momentum frame is appreciable enough to lend some kinetic energy to the electron-positron pair following production. Assume that the two photons with energies ε_1 and ε_2 are traveling in directions $\hat{\mathbf{k}}_1$ and $\hat{\mathbf{k}}_2$ before colliding at an angle θ in an external ("lab") frame of reference. Then the four-momenta[2] of the two photons are simply

$$\mathbb{K}_1 = \varepsilon_1/c \begin{pmatrix} 1 \\ \hat{\mathbf{k}}_1 \end{pmatrix} \qquad\qquad \mathbb{K}_2 = \varepsilon_2/c \begin{pmatrix} 1 \\ \hat{\mathbf{k}}_2 \end{pmatrix},$$

[1] Eq. (A.1) is well behaved and real for $-1 < \beta < 1$, but the output for negative β is also negative, corresponding to unphysical behavior (a negative area). For $|\beta| \geq 1$, the function quickly diverges towards $\widetilde{\infty}$.

[2] I omit the usual contra/co-variant index notation here for simplicity and instead denote four-vectors with a chalkboard bold script: $\mathbb{V}$.

while in the center of momentum frame, the electron and positron emerge afterwards back-to-back:

$$\mathbb{P}_{e^-} = m_e\gamma \begin{pmatrix} c \\ +\mathbf{v} \end{pmatrix} \qquad\qquad \mathbb{P}_{e^+} = m_e\gamma \begin{pmatrix} c \\ -\mathbf{v} \end{pmatrix}.$$

Here, $\gamma = (1-\beta^2)^{-1/2}$ is the usual relativistic factor and m_e is of course the mass of the electron. The invariance of the four-vector inner product allows us to write

$$\begin{aligned} \Big(\mathbb{K}_1 + \mathbb{K}_2\Big)^2_{\text{lab}} &= \Big(\mathbb{P}_{e^-} + \mathbb{P}_{e^+}\Big)^2_{\text{c.o.m.}} \\ \frac{2\varepsilon_1\varepsilon_1}{c^2}\Big(1 - \hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_2\Big) &= 2(m_e c)^2 + 2(m_e\gamma)^2\Big(c^2 + \mathbf{v}\cdot\mathbf{v}\Big). \end{aligned}$$

Using the fact that $\hat{\mathbf{k}}_1\cdot\hat{\mathbf{k}}_2 = \cos\theta$ and the definition of $\beta = |\mathbf{v}|/c$, we can rearrange the previous line as

$$\frac{2\varepsilon_1\varepsilon_1}{c^2}(1-\cos\theta) = m_e^2c^2\left(1 + \frac{1+\beta^2}{1-\beta^2}\right).$$

Finally, after solving for β we obtain

$$\beta = \sqrt{1 - \frac{2(m_e c^2)^2}{\varepsilon_1\varepsilon_2(1-\cos\theta)}}. \tag{A.2}$$

This expression gives the magnitude of the (normalized) velocity of either the positron or electron in the center of momentum frame following pair-production. Notice that if we consider a head-on collision of the photons such that $\theta = \pi$, we recover the same β that appears as the interaction parameter in Eq. (A.1). Setting $\beta = 0$ in Eq. (A.2) and solving for ε_2 will then give us the minimum threshold energy required for the positron-electron pair to be produced at rest, given the other photon's energy:

$$\varepsilon_2 = \frac{2(m_e c^2)^2}{\varepsilon_1(1-\cos\theta)}.$$

For a typical collision angle[3], we set $\theta = \pi/2$:

$$\varepsilon_2 = \frac{2(m_e c^2)^2}{\varepsilon_1}. \tag{A.3}$$

Plugging in some numbers, we then see that the energy required of a photon to engage in pair-production with the CMB ($\varepsilon_1 \approx 1$ meV) is about 0.8 PeV, while the energy required to pair-produce with the EBL ($\varepsilon_1 \approx 1$ eV) is about 0.5 TeV.

This gives weight to the claim that ∼TeV energy gamma-rays emitted by AGNs can provide the charged particles necessary for up-scattering, but it is still not entirely clear why the PeV energy photons would not suffice. The reason is that, because of the high density of CMB photons relative to the EBL, PeV gamma-rays will actually pair-produce *too quickly* while still in the extremely strong radiation field of the AGN, making interstellar space effectively

[3] The bare minimum energy would again be given by a head-on collision, but such events are unlikly. Averaging over all angles gives $\langle\cos\theta\rangle = 0$, or $\theta = \pm\pi/2$.

opaque to such photons beyond a certain length scale. The mean free path for an interaction to occur is defined as

$$\lambda_{\gamma\gamma} = \frac{1}{n\sigma_{BW}}, \tag{A.4}$$

where σ_{BW} is the pair-production cross-section given earlier by Eq. (A.1) and n is the number density of the target medium. For the CMB, $n \approx 400$ cm^{-3}, while for the EBL, $n \approx 3 \times 10^{-3}$ cm^{-3}. Using Eq. (A.4) along with the threshold energies of the incoming gamma-rays given by Eq. (A.3) gives a mean free path of about 600 Mpc for the TeV-EBL process, but only about 5 kpc for the PeV-CMB interaction. Because most galaxies have sizes on the order of 10—100 kpc, virtually none of the PeV gamma-rays or their daughter electron-positron pairs will even make it outside of their host galaxies.

Appendix B

Convergence of the Containment Bands

I remarked in Chapter 4 that each set of 1σ and 2σ containment bands for the plots of $Q_{E_1E_2E_3}(r)$ in Chapter 3 were generated from only $N = 500$ Monte Carlo simulations, due to computing and time constraints. While this number of iterations does not adhere strictly to the standard $N = 10^4$ used by Tashiro and Vachaspati in their original analysis [11], it can be shown that the decrease in maximum error of the Monte Carlo estimator is marginal past a certain point.

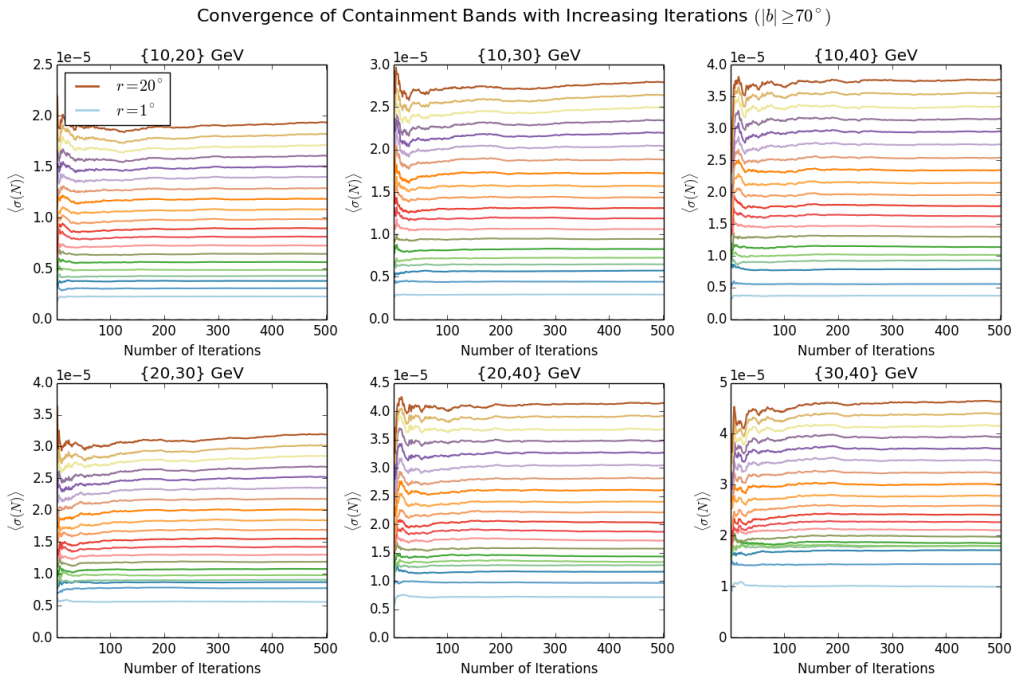


Figure B.1: Running ensemble averages of the 68% containment band widths σ plotted as functions of the number of iterations $N \in [0, 500]$ for all values $r \in [0°, 20°]$ in each of the six energy bin pairs. The size of the randomized data used in each simulation is pulled directly from the number of real diffuse photons with $|b| \geq 70°$, and with simulated Q and σ_Q then calculated from the inverse-transformed l, b, and E distributions shown in red in Figures 2.4a, 2.4b, and 2.5b, respectively.

Shown in Figure B.1 are the 1σ containment band widths for each radius $r \in [0°, 20°]$ in each of the six energy bin combinations $\{E_1, E_2\} = \{10, 20, 30, 40\}$ GeV. For the lowest radial patch sizes ($r < 10°$), the convergence is very quick, and the width of the band is proportionally small relative to those of the larger patch radii. Only a hundred or so iterations are necessary for convergence for these values of r. As r increases towards 20°, however, the corresponding $\langle\sigma(N)\rangle$ take a few hundred or so iterations for the function to smoothen out. For $r > 17°$, however, the band widths appear to be increasing still up to $N = 500$, albeit slowly ($\partial\langle\sigma\rangle/\partial N \ll 1$) and increasing at an ever slower rate ($\partial^2\langle\sigma\rangle/\partial N^2 \leq 0$). Additional iterations would be necessary for complete convergence of the containment band widths for these higher r values.

References

[1] W. Atwood, A. Albert, L. Baldini, M. Tinivella, J. Bregeon, M. Pesce-Rollins, C. Sgrò, P. Bruel, E. Charles, A. Drlica-Wagner, A. Franckowiak, T. Jogler, L. Rochester, T. Usher, M. Wood, J. Cohen-Tanugi, and S. Zimmer for the Fermi-LAT Collaboration. Pass 8: Toward the full realization of the fermi-lat scientific potential, 2013. arXiv:1303.3514.

[2] W.B. Atwood, R. Bagagli, L. Baldini, R. Bellazzini, G. Barbiellini, F. Belli, T. Borden, A. Brez, M. Brigida, G.A. Caliandro, C. Cecchi, J. Cohen-Tanugi, A. De Angelis, P. Drell, C. Favuzzi, Y. Fukazawa, P. Fusco, F. Gargano, S. Germania, R. Giannitrapani, and N. Giglietto et al. Design and initial tests of the tracker-converter of the gamma-ray large area space telescope. *Astroparticle Physics*, 2007.

[3] G. Breit and John A. Wheeler. Collision of two light quanta. *Phys. Rev.*, 46:1087–1091, Dec 1934.

[4] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2π decay of the k_2^0 meson. *Phys. Rev. Lett.*, 13:138–140, 1964.

[5] David J. Griffiths. *Introduction to Elementary Particles*. Wiley, 2014.

[6] Enrico Massaro, Alessandro Maselli, Cristina Leto, Paolo Marchegiani, Matteo Perri, Paolo Giommi, and Silvia Piranomonte. The 5th edition of the roma-bzcat. a short presentation. 2015. arXiv:1502.07755.

[7] Andrii Neronov. High-energy astrophysics, November 2017. Astronomy Department of the University of Geneva.

[8] K. A. Olive et al. Review of Particle Physics. *Chin. Phys.*, C38, 2014.

[9] A. D. Sakharov. Violation of CP Invariance, C Asymmetry, and Baryon Asymmetry of the Universe. *Soviet Journal of Experimental and Theoretical Physics Letters*, 5:24, 1967.

[10] G. 't Hooft. Symmetry breaking through bell-jackiw anomalies. *Phys. Rev. Lett.*, 37:8–11, 1976.

[11] Hiroyuki Tashiro, Wenlei Chen, Francesc Ferrer, and Tanmay Vachaspati. Search for cp violating signature of intergalactic magnetic helicity in the gamma ray sky. Monthly Notices of the Royal Astronomical Society: Letters 2014 445 (1): L41-L45, 2013. arXiv:1310.4826.

[12] Hiroyuki Tashiro and Tanmay Vachaspati. Cosmological magnetic field correlators from blazar induced cascade, 2013. arXiv:1305.0181.

[13] Tanmay Vachaspati. Primordial magnetic fields and cp violation in the sky, 2001. arXiv:astro-ph/0111124.

[14] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Phys. Rev.*, 105:1413–1415, 1957.