Stockholm University

Department of Astronomy

Bachelor Thesis

No sign of a left-handedness in GeV photon arrival directions

Author: Julia Asplund

Supervisors: Professor Axel Brandenburg,
Assistant Professor Gudlaugur Jóhannesson

June 2, 2019

Abstract

A non-zero helicity of cosmological magnetic fields could, if detected, have important implications for models of the electroweak phase transition in the early Universe. It has been suggested that the helicity of such a field could be related to the handedness of photon arrival directions in the diffuse gamma-ray sky observed by the *Fermi* Large Area Telescope (LAT). Tashiro et al. (2014) found a left-handedness when applying this method, implying the existence of an extragalactic magnetic field with negative helicity. In this work the same method is applied, using twice as much data from the LAT due to the longer exposure time, providing higher statistical certainty. Additionally, the potential effects of the non-uniform LAT exposure and contamination of galactic emission are studied using simulated data sets. The results obtained indicate no significant signal, as the simulations suggest that the uncertainty was highly underestimated in Tashiro et al. (2014), and any observed handedness using the updated LAT data is found to be compatible with zero within this new error estimate.

Contents

1 Introduction **2**

2 Motivating theory **4**
2.1 Physical model 4
2.2 Concretized statistic 5
2.3 Generalized motivation 7

3 Data and calculations **8**
3.1 Simulated data 8
3.2 LAT data 10

4 Results and discussion **11**
4.1 Results from simulated data 11
4.1.1 Uniform data set 11
4.1.2 Isotropic data set including LAT exposure 13
4.1.3 IEM data set including LAT exposure 15
4.1.4 Combined simulated data 16
4.2 Results from LAT data 19
4.3 LAT data results combined with improved error estimate 19

5 Conclusions and further discussion **23**

Bibliography **26**

Introduction

Magnetic fields of astronomical scales are detected all over the Universe, in galaxies and galaxy clusters. They are hypothesized to originate from weaker fields, seeds, created in the early Universe, amplified by dynamo action and developed over time during large structure formation and evolution (Kronberg 1994; Brandenburg et al. 1996; Durrer & Neronov 2013). However, the process of formation of these seeds is still unknown, and a long-standing question within cosmology. One possible explanation is that the seeds are accounted for by the weak magnetic fields expected to form during the electroweak phase transition in the early Universe (Vachaspati 1991), prior to which the weak and electromagnetic forces where indistinguishable. Even though the fields may have represented a significant fraction of the energy density of the Universe at the time, because of the cosmological expansion since then, their size and strength were by many orders of magnitude different from those of galactic fields observed today. The process that would allow weak magnetic fields on such small scales to expand to the fields on megaparsec scales that are currently observed, requires them to be helical (Field & Carroll 2000).

Magnetic helicity can be described as the extent to which the field lines of a magnetic field wrap and coil around each other; one can picture it as the twisting of a rope or a cork screw. Field lines wound towards the right or left correspond to right- or left- handed helicity respectively. Formally magnetic helicity within a large valume V is defined as

$$h = \int_V dV \, \mathbf{A} \cdot \mathbf{B}, \tag{1.1}$$

where $\mathbf{A}$ is the vector potential of the magnetic field so that $\mathbf{B} = \nabla \times \mathbf{A}$. A left-handed field will correspond to a negative value of h, and vice versa, while $h = 0$ implies no helicity at all or the cancellation of equally many left- and right-handed contributions (Tashiro & Vachaspati 2013).

Because of the long evolution that currently observed astronomical magnetic fields have gone through, they do not provide enough information about the properties of the seeds to determine their origin. Therefore the approach is to try to detect these weak magnetic fields where they have not been altered too much by anything but the Universe expansion, thus in the intergalactic medium (IGM) (Durrer & Neronov 2013). Fields on megaparsec scales in the IGM have indeed been implied, and lower limits on their strength have been inferred (Neronov & Vovk 2010). However, in order to confirm the seed origin hypothesis involving electroweak phase transitions, a non-zero helicity of the field would need to be determined. The only way to measure the helicity directly is to take advantage of the magnetic field's

impact on charged particles propagating through it, and there are few ideas on how to do so. A method has been proposed by Tashiro & Vachaspati (2013), involving diffuse gamma-ray emission. TeV photons originating from blazars can interact with extragalactic background light (EBL) creating an electron-positron pair, which will then propagate through the magnetic field and change direction. The pair might eventually up-scatter photons from the cosmic microwave background (CMB) resulting in GeV gamma rays which, if detected, can probe the change of direction relative to the original TeV photon. These GeV photon cascades have previously been used to detect the presence and strength of a magnetic field in the IGM (Elyiv et al. 2009; Neronov & Vovk 2010), but not the helicity which, if observed, could have important implications. Tashiro & Vachaspati (2013) relate the position of arrival of the GeV photons to the helical part of the correlation tensor of the magnetic field and argue that diffuse extragalactic high-energy GeV emission (as opposed to that from known TeV sources) could be used in a similar way.

This approach was refined and executed in 2014 in the form of evaluating gamma-ray data collected by the LAT, with results indicating a left-handed magnetic helicity on $\sim$10 Mpc scales with a field strength of $\sim 10^{-14}$ Gauss, if the handedness (indicator of helicity) of photon arrival directions is indeed caused by a magnetic field (Tashiro et al. 2014). However, the processed data sets used were not large enough to provide satisfying statistical certainty. Since then the LAT has gathered twice as much data, and updates have been made in the reconstruction of the photon events (Atwood et al. 2013b). Additionally, possible contamination by Galactic interstellar emission (IEM) or a non-uniform exposure of the LAT was not considered quantitatively by Tashiro et al. (2014).

The aim with this work is to determine what effect the IEM and the non-uniform exposure of the LAT might have on the resulting measurements of helicity, and to reproduce the calculations done by Tashiro et al. (2014) to see if their results can be confirmed using a larger data set. There is an underlying comparison with the study by Tashiro et al. (2014) throughout this entire report, and therefore their notation and definitions are used when suitable.

Motivating theory

The following subsections provide a physical motivation for the development of the proposed method, and a derivation of the formulae used. Further details on the mathematics of magnetic fields and their helicity are not necessary when measuring just the handedness of the magnetic field, as discussed in section 2.3.

2.1 Physical model

The approach taken in measuring any helicity of a cosmological magnetic field follows that of Tashiro et al. (2014) and is motivated by a physical model. Assume there is an extragalactic blazar and the Earth is located within the jet opening angle. Photons emitted in the jet with TeV energies have a mean free path of $D_{\mathrm{TeV}} \sim 100$ Mpc before they interact with an EBL photon resulting in pair production. The electron-positron pair's trajectory will be bent by a magnetic field if present and have a mean free path of $D_e \sim 30$ kpc before eventually up-scattering a CMB photon via inverse Compton (IC) scattering (Elyiv et al. 2009; Tashiro & Vachaspati 2013). These secondary photons are referred to as cascade photons because of the multiplying effect, as they in many cases still are energetic enough to stimulate pair production themselves and the process is repeated.

The distance D_{TeV} is energy dependent, where photons with higher energy have a shorter mean free path before interaction with EBL. Consider two primary photons with different energy emitted by the blazar (see Figure 2.1), generating two secondary up-scattered photons denoted by their energies E_1 and E_2 where $E_1 < E_2$. Note that every primary photon produces a pair, but since the electron and positron will travel in opposite directions in the magnetic field, only one of the IC scattered photons from each primary photon can be assumed to reach the observer. The secondary photons E_1 and E_2 reach the observation plane at vectorial position $\mathbf{\Theta_1}$ and $\mathbf{\Theta_2}$ respectively, the origin of the vector plane defined as the point where the line of sight from the observer to the source meets the observational plane. As $E_1 < E_2$, $D_{\mathrm{TeV1}} > D_{\mathrm{TeV2}}$, so the observed photons will probe the magnetic field in two different regions which is what makes it possible to retrieve the helicity.

It is now suitable to define a correlation function

$$G(E_1, E_2) = \langle \mathbf{\Theta_1} \times \mathbf{\Theta_2} \cdot \hat{\mathbf{x}} \rangle, \tag{2.1}$$

with $\hat{\mathbf{x}}$ pointing towards the source, perpendicular to the observational plane, and the average is to be taken over all photons originating from the blazar. This correlator is related to the

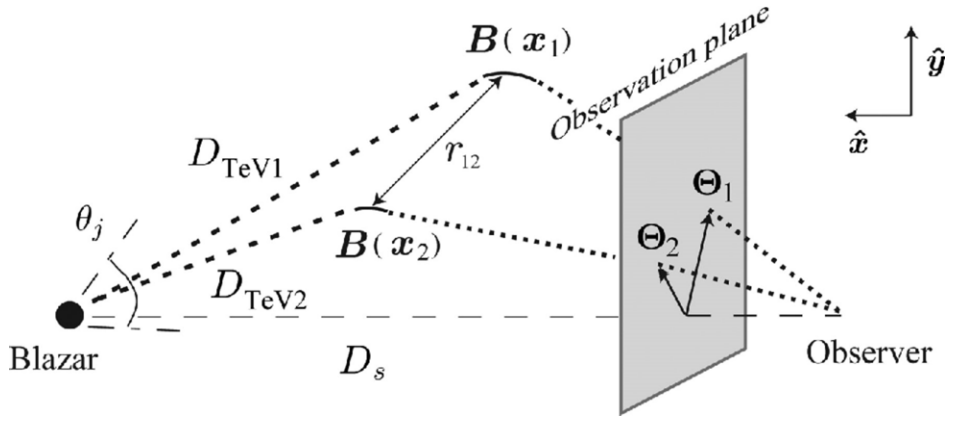


Figure 2.1: Schematic of how two TeV photons originating from a blazar at a distance D_s can probe the direction of a magnetic field. The solid curved lines are the distance $D_e \sim 30$ kpc which the charged particles travel and hence the region size the field can be sampled. The regions are separated by a distance r_{12} which depends on the energy difference between the two initial photons, but can be several orders of magnitude larger than D_e. $\mathbf{B}(x_i)$ denotes the magnetic field in the two regions. Figure taken from Tashiro & Vachaspati (2013)

helical correlation function M_H of the magnetic field as $G(E_1, E_2) \propto \frac{1}{2} M_H(|r_{12}|) r_{12}$. M_H is the helical part of the two-point correlator of the magnetic field, also containing a normal part. This is the outline of how observing gamma rays via the correlator $G(E_1, E_2)$ can reveal the helicity of a cosmological magnetic field (Tashiro & Vachaspati 2013).

2.2 Concretized statistic

When putting this method into practice, there are a few issues with the correlator defined in equation (2.1). For one, it cannot be guaranteed that the line of sight to the Earth lies within the jet opening angle, meaning the blazar might not be visible, excluding the use of the unit vector $\hat{\mathbf{x}}$. But, suppose a third photon E_3 is observed, with a higher energy than the other two so that $E_1 < E_2 < E_3$. The higher the energy of the photon, the smaller is the change in trajectory, so E_3 can be used to approximate the position of the blazar. This results in a new correlator

$$G(E_1, E_2, E_3) = \langle (\mathbf{\Theta_1} - \mathbf{\Theta_3}) \times (\mathbf{\Theta_2} - \mathbf{\Theta_3} \cdot \hat{\mathbf{x}}_\mathbf{3} \rangle \tag{2.2}$$

where $\hat{\mathbf{x}}$ points in the direction of E_3. If the blazar is not visible it is of course inconvenient to use the definition of the origin involving the line of sight, but now the choice of origin is arbitrary.

As diffuse gamma rays are observed in the sky, the correlator needs to be adjusted to a sphere rather than an observational plane. Since the relevant information about each photon is only the their position, the correlator can be defined in terms of the unit vectors (n_a) pointing to their location on the sphere (Tashiro et al. 2014),

$$Q'(E_1, E_2, E_3) = \langle (\mathbf{n_1} - \mathbf{n_3}) \times (\mathbf{n_2} - \mathbf{n_3}) \cdot \mathbf{n_3} \rangle = \langle \mathbf{n_1} \times \mathbf{n_2} \cdot \mathbf{n_3} \rangle. \tag{2.3}$$

This equation could have been enough, but recall that the average is to be taken over all observed photons originating from one source. When observing the diffuse gamma-ray sky, there is no telling which photons come from which source nor if they are the result of IC scattering or not. This is the key in the method developed by Tashiro et al. (2014). The idea is that presumably some of the photons within a certain neighborhood of the E_3 photon can be assumed to originate from the same source, and be the result of a cascade. So, if the average is taken over all photons within a reasonable region around the photon probing the source, the statistic in equation (2.3) should still work. Many photons included will of course have completely different origin, but their handedness would then be completely random and should add up to zero. To strengthen the statistical certainty, the ensemble average is also taken over all E_3 photons as there is also no way of guaranteeing their origin being from a TeV source. To mask out a region around each E_3 photon, a straightforward approach is to define a top-hat (top-hat refers to the rectangular shape of the function) window function that will eliminate any E_1 and E_2 photons outside a chosen radius R of the region. The resulting statistic is then

$$Q(E_1, E_2, E_3, R) = \frac{1}{N_1 N_2 N_3} \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \sum_{k=1}^{N_3} W_R(\mathbf{n_1}_{,i} \cdot \mathbf{n_3}_{,k}) W_R(\mathbf{n_2}_{,j} \cdot \mathbf{n_3}_{,k}) (\mathbf{n_1}_{,i} \times \mathbf{n_2}_{,j} \cdot \mathbf{n_3}_{,k}). \quad (2.4)$$

The subscripts i, j, k correspond to different photons, and N_a to the number of photons with energy E_a. W_R is the top-hat window function defined as

$$W_R(\cos\alpha) = \begin{cases} 1, & \text{for } \alpha \leq R \\ 0, & \text{otherwise.} \end{cases} \quad (2.5)$$

$\cos\alpha$ will be the result of the dot product between the unit vectors of an E_3 and E_1 or E_2 photon, so W_R will be zero if their separation angle is larger than R. Using equation (2.5), equation (2.4) can be simplified. We define

$$\eta_{\mathbf{a}} = \frac{1}{N_a} \sum_{i \in D(\mathbf{n_k}, R)} \mathbf{n_a}_{,i}, \quad (2.6)$$

where $a = 1, 2$ and $D(\mathbf{n_k}, R)$ is the list of photons in the masked out region in the sky centred on $\mathbf{n_3}_{,k}$. Basically $\eta_{\mathbf{a}}$ is the unit vector to the average location of all E_a photons within the chosen patch radius. This definition simplifies equation (2.4) to

$$Q(E_1, E_2, E_3, R) = \frac{1}{N_3} \sum_{k=1}^{N_3} \eta_{\mathbf{1}} \times \eta_{\mathbf{2}} \cdot \mathbf{n_3}_{,k} \quad (2.7)$$

so that the correlator is actually just calculated once for every E_3 photon, but using the averaged normal vectors for E_1 and E_2, which saves computational time.

Figure 2.2 shows the essential function of the correlator. The sky is shown, with the Galactic plane cut out to reduce the Galactic emission, and photons are shown as dots in three different colours corresponding to different energies. A patch radius is chosen, within which the E_1 and E_2 photons are assumed to come from the same source as the E_3 photon

in the patch centre. Then each combination of E_1 and E_2 photons is checked to see if they relative to the E_3 photon, with declining energy, are positioned in a left or right turning spiral, corresponding respectively to left- and right-handed helicity thus negative and positive Q value.

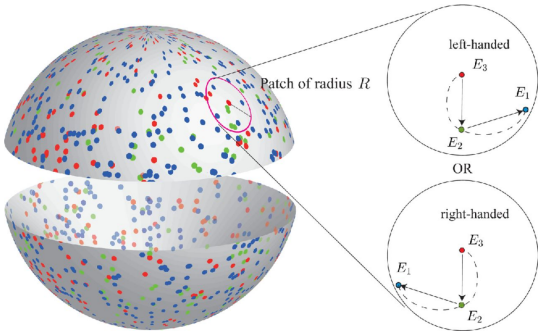


Figure 2.2: Illustration of the gamma-ray sky featuring only three different energies, with the galactic plane cut away that demonstrates how patches of radius R are centred on the highest energy photon E_3. The handedness of all triplets inside the patch is determined by whether the photons in order of declining energy are ordered along a spiral twisting to the right or left. Figure taken from Tashiro et al. (2014).

2.3 Generalized motivation

This subsection provides a more general motivation for equation (2.7). When considering the sky, as in Figure 2.2, one can always construct left- or right-handed triplets of ordered energy. But note that the energy only enters in the ordering of the photons. In telling the handedness of the triplet, the only thing that matters is their relative positioning. A mirrored image of a triplet will give opposite handedness, so to quantify it, one needs a mathematical construct that will also be odd under parity inversion. Such a construct is a pseudoscalar, an example of which is

$$q = \mathbf{n_1} \times \mathbf{n_2} \cdot \mathbf{n_3}. \tag{2.8}$$

Now if these were completely randomly distributed, their combined handedness should of course add up to zero. If one then sums equation (2.8) over all possible ordered triplets, a non-zero result would imply the presence of an underlying handedness, which could be provided by a helical magnetic field. Summing equation (2.8) over all triplets almost results in equation (2.7). However, the sum is not taken over all possible ordered triplets in the sky but over all triplets with the energies E_1, E_2 and E_3. The reason for this is the physical motivation, that the E_3 photon is expected to be of a certain energy to actually probe the direction of the blazar. This is the choice made by Tashiro et al. (2014), although this more general approach should in principle work and has been applied when inferring the helicity of solar magnetic fields (Bourdin & Brandenburg 2018). An argument for using the more physically motivated approach as in this and the referenced work, is that summing over all possible triplets may introduce too much noise, the extragalactic signal could then drown within any intergalactic sign of helicity.

Data and calculations

The calculations of Q were performed in the same way for all data sets described in the following subsections. For each calculation, one combination of energy bins was used, a latitude cut applied and the patch radius R varied in integer steps from 1° to 25°, applying equation (2.7) at each R. For the uncertainty estimate of each Q, the standard deviation of the mean was used. It is given by the standard deviation σ of the distribution of Q values from all patches (i.e. based on all E_3 photons) divided by the square root of the number of patches (the same as N_3 the number of E_3 photons):

$$\delta_Q = \frac{\sigma}{\sqrt{N_3}}. \tag{3.1}$$

To provide a fair comparison of the results, the same parameters were used as by Tashiro et al. (2014), apart from R where they used a maximum value of 20°. The larger maximum value of R used in this work is only to further map the R dependence. The photons are binned in five different energy bins of width 10 GeV, ranging from 10 GeV to 60 GeV. The energy bins will be denoted by their lower boundary, meaning 10 GeV photons refers to events in the range 10–20 GeV. As $E_3 = 50$ GeV for all calculations, these choices lead to six possible combinations of energy bins fulfilling the condition $E_1 < E_2 < E_3$ for each latitude cut. The choice of latitude cuts applied to avoid galactic emission were ±60°, ±70° and ±80° although only the latter two were presented by Tashiro et al. (2014). The stated latitude cut is only applied to the E_3 photons; E_1 and E_2 have a lower cut so that the patches form full circles around all E_3 photons. Additional calculations were performed on the simulated datasets with other choices of parameters for verification, see 4.1.1.

3.1 Simulated data

In the reference study (Tashiro et al. 2014), the mean value and standard deviation of Q was calculated for a set of synthetic data with a uniform distribution for comparison with their results and error calculation. As they mention, a uniform distribution should be a reasonable simulation of the diffuse gamma-ray background. However, the exposure of the LAT is not uniformly distributed over the sky and could therefore affect the Q statistic.

The exposure of a position in the sky is given by the live time, the time spent observing a given position at a given inclination angle, multiplied by the effective area. The effective area is the efficiency of the instrument, which depends on the photon energy and incoming

inclination angle. This means that the exposure is actually, if perhaps only slightly, energy dependent. It is possible that the relative enhancements in exposure leaves a pattern in the data that could have a handedness, or perhaps give rise to excessive outliers of Q values. Therefore a set of synthetic photon data, mimicking an isotropic signal but including the varying exposure was generated for this project. The set consisted of 200 simulations of photon data over the entire sky, each simulation generated with a different seed using `gtobssim` from the Fermitools[1]. Then Q was calculated for each simulation using the same code as for the real data.

The useful result from these calculations is the statistics compiled over all seeds. First of all, comparing the mean of all δ_Q and the standard deviation of the distribution of 200 Q values is a way of determining the validity of the error estimate. Additionally, the mean Q value can reveal if the exposure introduces a systematic deviation of Q from zero, which would have to be corrected for in the data.

Aside from the varying exposure times, one hypothesis tested in this project was that the IEM also could have a significant effect on Q as measured. In Tashiro et al. (2014), it is argued that the contribution to Q from photon triples that happen to not originate from the same source, which is liable to be the majority of the triples, will add up to zero. However, this cannot be guaranteed to be the case as the Galactic emission does not have a uniform distribution, it is dominant around the galactic equator and decreases towards the poles. In addition, there are extended regions of the sky where excessive IEM is observed, especially in the higher energy bins used in this work, that not only enhances the latitude dependence but also the differences in emission between the northern and southern hemisphere (Acero et al. 2016). Galactic photons would therefore ideally be completely excluded from the LAT data in the analysis, but in reality that is impossible to achieve. Although a broad band around the Galactic plane is cut out of the LAT data set (see section 3.2, there is still undoubtedly a significant amount of Galactic emission left. To make sure that it doesn't have a significant impact on the calculated Q values, a set of `gtobssim` simulated synthetic data generated to mimic galactic emission, also including varying exposure, was generated and used to perform the same calculations and statistical analysis as described for the isotropic data set. The model used for the simulations is the `gll_iem_v06.fit` available from the Fermi Science Support Center [2].

Finally, to create simulated diffuse emission data as realistic as possible, the isotropic and IEM simulated data sets including the exposure are combined and adjusted, forming a data set with approximately the same number of photons in each energy bin as the real LAT data. The statistical results from this simulated data is used to determine the accurate uncertainties of the calculations on real data.

For comparison, and to verify the code, 500 simulations of random uniformly distributed events were also generated. Since it is completely uniform there is no energy dependence in the photon positions, and it was therefore generated to the same size as only one of the energy bin combinations in the LAT data, namely $E_1, E_2, E_3 = 30, 40, 50$ GeV. Calculations were then performed in the same way as for the non-uniform data, so their statistics could

[1]Fermitools version 1.0.1, available through `https://fermi.gsfc.nasa.gov/ssc/data/analysis/software/`

[2]`https://fermi.gsfc.nasa.gov/ssc/data/access/lat/BackgroundModels.html`

be easily compared. The reasoning behind the method of determining Q is based upon the idea that it must sum up to zero for a uniform distribution, so evaluating such a data set is appropriate to test the calculations.

3.2 LAT data

For the evaluation of Q, close to ten years of P8R3[3] photon data (rougly 10^6 events in the selected energy range, from 1/9/2008 to 1/4/2019) from within a 60° radius from the Galactic poles was retreived from the Fermi Science Support Center [4]. P8R3 is the most recent version of the data with the most accurate reprocessing since the launch of the telescope (Bruel et al. 2018). As the only data relevant in the chosen way of calculating Q is diffuse extra galactic emission, several cuts are made to the data set, resulting in in a total number of events of the order 10^5 between 10 GeV and 60 GeV.

Much of the Galactic emission is excluded by only selecting events from within 60° around the poles. Next, two steps are taken to avoid contamination from photons originating from the Earth limb. First, using the `gtselect` script in the Fermitools [5], the maximum accepted zenith angle for the data is set to 105 degrees, based on recommendations given by the *Fermi*-LAT collaboration (The-Fermi-LAT-collaboration 2019). Second, using the `gtmktime` script, the spacecraft's rocking angle from zenith is selected to be less than 52°.

The final cut made to the data is to remove events assumed to originate from known point sources. Many of them are extragalactic, but because the statistical approach to calculate Q requires diffuse emission, a point source would be too dominant and could reduce the signal to noise (Tashiro et al. 2014). Therefore, a 2° angular diameter region is masked out around every known point source given in the *Fermi*-LAT Fourth Source Catalog (4FGL) (The-Fermi-LAT-collaboration 2019).

Regarding the "purity" of the photons, i.e. the probability of the observed event actually being a photon, the default event class SOURCE is used, as it is stated by the *Fermi*-LAT collaboration to be sufficient.

After all cuts, the number of photons left to use in the analysis in energy bins E_1–E_5 above a latitude cut of 60° were 63985, 16089, 7389, 3870 and 643, respectively. This is a factor 9–10 more in all energy bins except the highest, which is about a factor 3 larger compared with in Tashiro et al. (2014), at the same latitude cut.

[3]P8R3 denotes the pass 8 release of the data, revised for the third time.

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

[5]Fermitools version 1.0.1, available through `https://fermi.gsfc.nasa.gov/ssc/data/analysis/software/`

Results and discussion

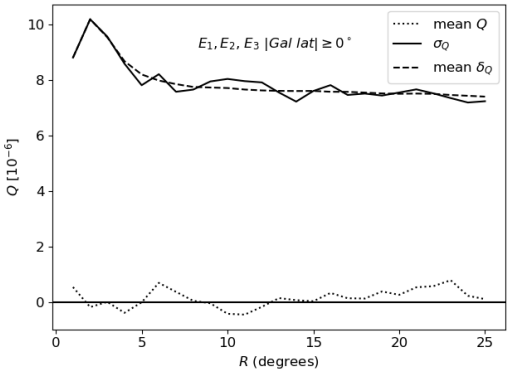
4.1 Results from simulated data

The statistical results of the calculations using the different sets of simulated data are all presented in the same way; Figure 4.1(a) shows an example of such a plot. The dotted line shows how the mean value of the distribution of Q values from all of the simulations (with different seeds) varies with R. The standard deviation of that distribution, denoted by σ_Q, is represented by the solid line, and the mean of δ_Q of the individually seeded simulations by the dashed line. So if the uncertainty estimation is valid, then the dashed line should match up with the solid line as the the spread in the distribution of Q values should be the same as the average of the errors for the individual Q values. In most of the plots presented in the following subsections, all three quantities are plotted for data from the northern and southern galactic hemisphere both separately and combined, denoted by different colours. The reason for this is that both the LAT exposure and the IEM distribution vary between the hemispheres. Unless explicitly stated otherwise, the latitude cut stated in each plot is only applied to the highest energy bin, while the other have a 25° lower cut; see Figures 4.1(a) and (b).

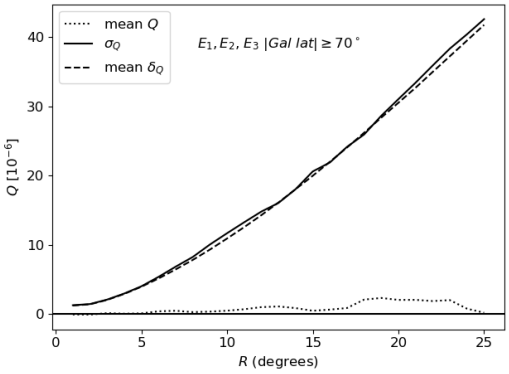
4.1.1 Uniform data set

The first calculations are performed with the uniform data set, primarily to test the calculations, but the results revealed an unexpected effect of the latitude cut. Figure 4.1 displays the result from three calculations on the uniform data set. The first plot, (a), has no latitude cut applied at all, the second, (b) has a cut at ±70° for all energy bins, and the third, (c), a cut at ±70° for E_3, and ±45° for E_1 and E_2. Note that all three show an average Q that is consistent with zero within errors for all R, verifying the code used for the calculations. The error estimates and standard deviation of Q are of further interest.

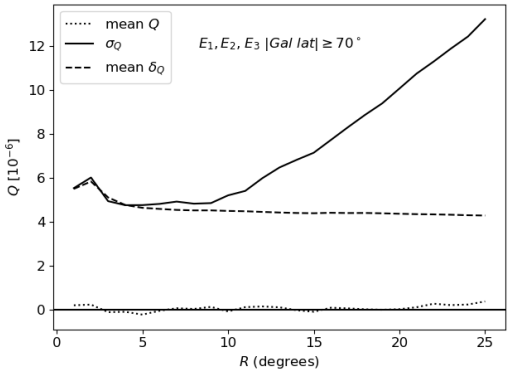
Both Figures 4.1(a) and (b) show expected results. With no latitude cut at all, as in Figure 4.1(a), a larger patch radius will always include more photons, uniformly distributed around E_3, leading to more Q values calculated within each patch. Both δ_Q and σ_Q are fairly constant past $R = 5°$, but declining slightly as the statistic improves with R. In Figure 4.1(b) it is clear that applying a latitude cut causes both the spread and the value of δ_Q to increase with R. This is expected for geometrical reasons; when the same latitude cut is applied for all energy bins most patches will not be circular but have a sharp cut off. This means the photons will not be uniformly distributed around E_3, but concentrated in one half. So the



(a) No latitude cut applied.



(b) A latitude cut of ±70° is applied to all energy bins.



(c) A latitude cut of ±70° is applied to the highest energy bin, and ±45° to the two lower.

Figure 4.1: The statistical results of calculating Q using 500 subsets of uniformly distributed simulated data. Plotted is the mean Q value and error estimate δ_Q, and the standard deviation σ_Q of the obtained distribution. Each subplot represents a different variation of latitude cut.

contributions to the averaged Q value in each patch will not be uniformly distributed. This increases δ_Q and the absolute values of Q, which in turn increases the standard deviation of the whole distribution of Q values.

The plot in Figure 4.1(c) however shows an unexpected result. Here the curve of the error estimate has the same shape as in Figure 4.1(a), while the standard deviation of the Q distribution starts to deviate from δ_Q around $R = 5°$ and takes off in a shape much like in Figure 4.1(b), although not to as high values. Running the same calculations for other latitude cuts showed that the effect becomes stronger and begins at lower R for cuts at larger latitudes. This probably has to do with the fact that, even though the lower latitude cut for E_1 and E_2 allows for a full, circular patch around each E_3, the photons below the E_3 cut are not accessed by as many E_3 photons as those above, and only from one direction. Within the patches there is no difference to the case of no latitude cut, which explains why δ_Q has the same shape. When considering the entire distribution of Q values however, the contributions from photons below the cut will not cancel out to the same extent, which is likely to be what

causes the increase the in standard deviation with R. Although we have not been able to provide a full mathematical explanation for this behavior, it is sufficient motivation for concluding that δ_Q is a strong underestimate of the uncertainty of the calculations, due to the latitude cut applied. It is not yet enough to determine a new error estimate, as the effect of the exposure and IEM might should also be accounted for.

4.1.2 Isotropic data set including LAT exposure

The discussion in this section is on the analysis of the data set with an isotropic signal simulated based on LAT exposure data. Six examples of the results, for two different latitude cuts and three different energy bin combinations, are displayed in Figure 4.2. The left column has a cut at 60° and the right at 80°. In these plots the results from the northern and southern Galactic hemisphere are displayed both separately and combined, as they have different exposure. Several conclusions can be drawn from these results.

First of all, the varying exposure seems to change the signal to noise ratio, as δ_Q increases slightly with R past $R = 10°$ in all plots while it decreased with the uniform data. Again the standard deviation of the distribution deviates from the error estimate, but with a steeper dependence than for the uniform data. In the right column, the curves are closer to the linear shape in Figure 4.1(c), most likely because the effect of the latitude cut is more dominant when the cut is larger. The values of the uncertainties increase with the cut and with the values of the two lower energy bins. This is mainly because the number of photons decreases. Overall, the relative increase in σ_Q compared to δ_Q is larger than in the uniform data, suggesting that the non-uniformity of the LAT exposure implies an even larger realistic error estimate.

Additionally, the mean Q values deviate more from zero when considering the exposure, and apparently 200 realizations of data is not enough for all Q to cancel out to zero. This effect is strongest for $E_1, E_2 = 10, 20$ GeV; in Figures 4.2(a) and (b), which display a negative bias. Individual plots of Q for the different realizations show that the bias is due to outliers rather than an overall shift of the values. The mean Q starts to deviate from zero first at large R, but not prominently for all energy bin combinations and consistently within uncertainties. Thus the conclusion is that the varying exposure does not significantly bias the value of Q as measured.

A potential difference between the signal in the data from the northern versus the southern hemisphere was mentioned in Tashiro et al. (2014) as something to be evaluated further. They presented an example where the northern and southern results were plotted separately, showing a clear discrepancy between them becoming significant at $R \sim 12°$. It is mentioned that this may be due to the larger amount of photons in the northern data. In Figure 4.2, a very similar behavior is observed in panels (b) and (f). A discrepancy is in fact found in all plots except in Figure 4.2(d), but it becomes prominent at varying R. Note that there is no particular pattern in whether the northern or southern values are negative or positive or deviate more from zero. It is hard to determine exactly what causes the discrepancy. It could still be because of the difference in the number of photons, or the exposure, or perhaps a combination as the two are closely related. Either way it appears quite random in the results, and indicates that a certain difference between northern and southern Q values can be expected in the real data without it being a sign of an actual difference in the underlying

signal.

A surprising result in all of the plots in Figure 4.2 is that the standard deviation of the Q distribution is larger in the northern data than the southern for large R, even though the number of photons is slightly larger in the northern data. This is most likely caused by the exposure introducing a larger spread.

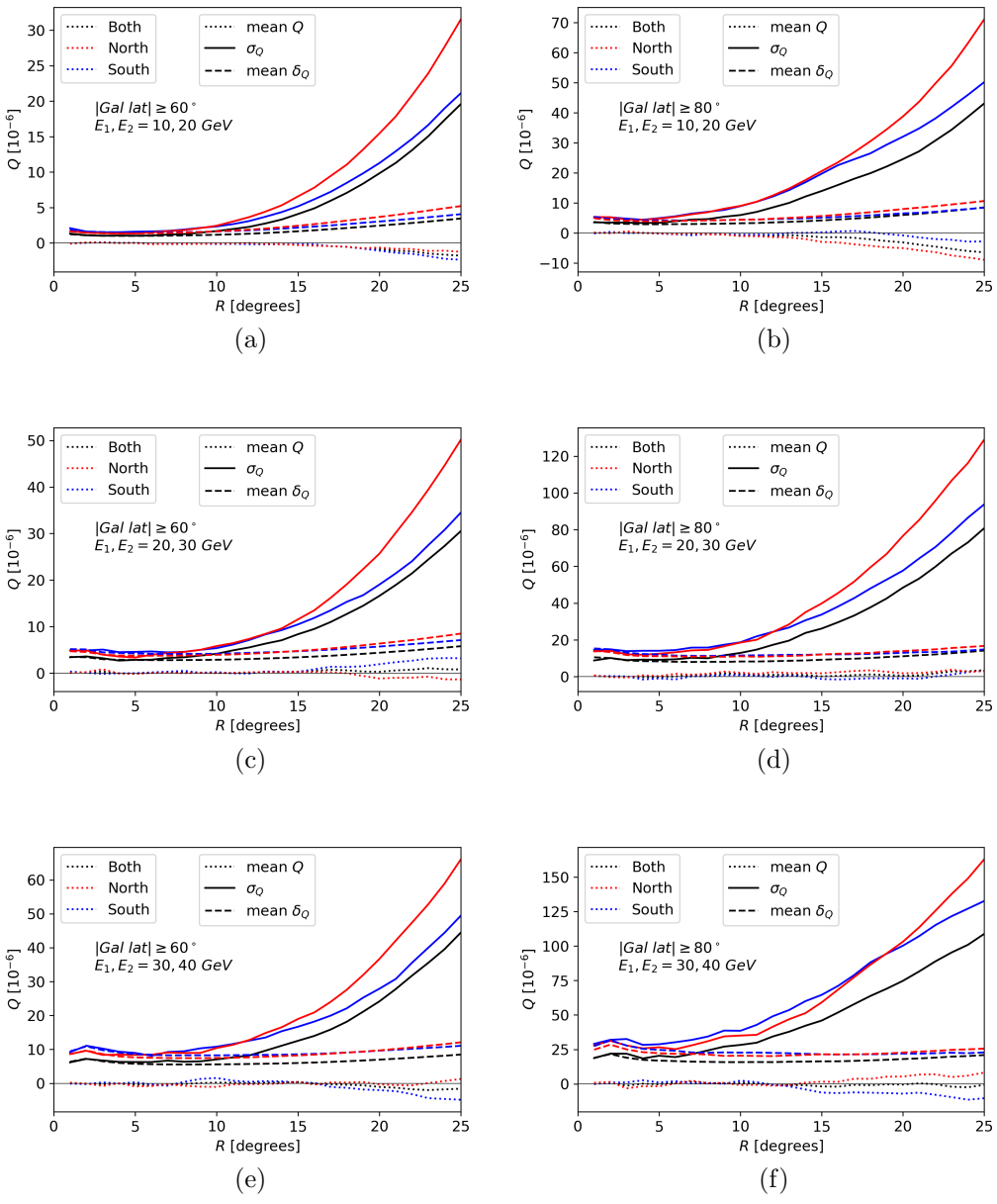


Figure 4.2: The statistical results of calculations on 200 simulations of data mimicking an isotropic signal, but distributed according to the LAT exposure. The two columns represent different latitude cuts, and each row a different energy bin combination.

4.1.3 IEM data set including LAT exposure

Figure 4.3 shows six examples of the results from the calculations on the data set simulated using IEM with the LAT exposure, including the same choices of latitude cut and energy bin combinations as presented for the isotropic set.

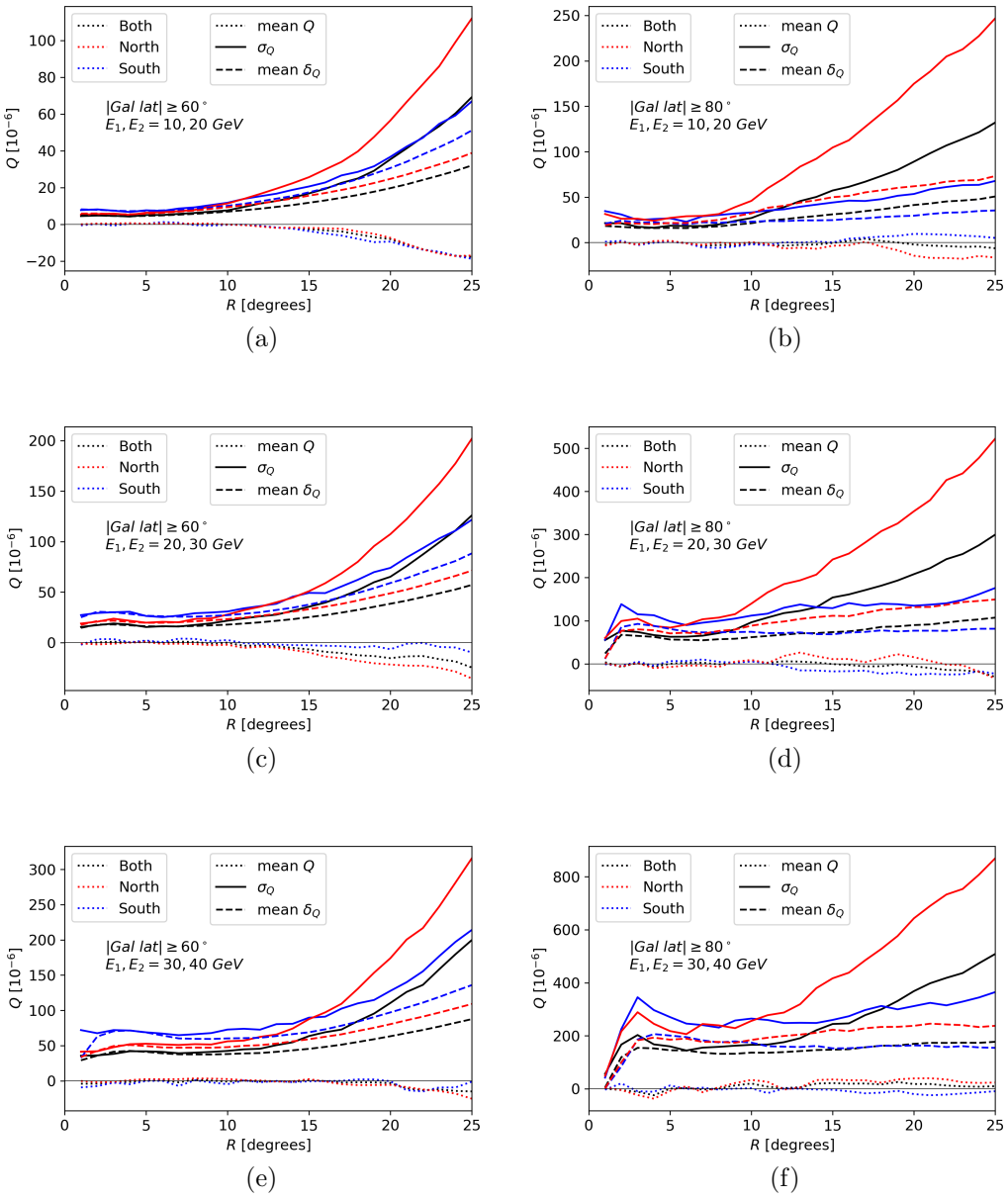


Figure 4.3: The statistical results of calculations on 200 simulations of data mimicking only galactic emission, distributed according to the LAT exposure. The two columns represent different latitude cuts, and each row a different energy bin combination.

Overall these results show larger spread, error estimate, and deviation of Q from zero than in the isotropic data set. It is important to note however, that, because the IEM is

concentrated towards the Galactic plane, the number of photons decreases more with latitude cut. So in the left column of Figure 4.3, where the cut is at $\pm 60°$, the number of $E_1 = 10$ GeV photons is roughly the same as in the isotropic set, but in the right column, where the cut is at $\pm 80°$, there are half as many. The number of E_3 photons relative to lower energy photons is small overall in the IEM data compared to the isotropic. This is because the spectral curve of the IEM declines steeper with energy than in the isotropic signal (Ackermann et al. 2015). That being said, even considering the difference in the amount of data used, the results in the left column of Figure 4.3 strongly indicate that the IEM introduces additional noise to the calculated Q values.

In all of the displayed plots, the standard deviation of the Q distribution is larger for the northern than the southern data at large R, just like in the isotropic data, but the difference is more prominent, especially at the highest latitude cut. Unfortunately, it is not suitable to conclude too much from this as there are only seven E_3 photons in the north above the $\pm 80°$ cut in the IEM data and ten in the south. The photons in the south could just happen to be positioned so that the effect of the latitude cut is less significant. It could also be that the discrepancy in exposure is accentuated and therefore causes a much larger spread in the northern Q values, or that it is caused by the position of the extended regions of excessive IEM. However, no solid conclusions can be drawn based on so little data.

All plots except 4.3(f) show a negative bias towards large R in the mean Q values of the combined data, and it is in fact rather strong in panels (a), (c), (d) and (e). When considering all results from the calculations on the IEM data, there is a negative bias in all but three, while in the isotropic data there is a more random spread. The value of R where it becomes visible however, varies from $5°$ to over $20°$. It is clear that both the IEM and the varying exposure widen the spread in the Q values, which results in a non-zero mean value of the 200 realizations. To determine whether they might cause an underlying bias that should be corrected for in the real data, it is more suitable to consider the results of the combined IEM and isotropic data, that form a realistic simulation of the sky.

4.1.4 Combined simulated data

Six examples of the results of the combined isotropic and IEM data simulated based on the LAT exposure are displayed in Figure 4.4. Before the calculations, the data set was resized so that the number of photons in the energy bins 10–40 GeV is no more than 10% more or less than the corresponding number in the LAT data, and no more than 5% in the highest energy bin. This is because the results presented in the previous subsections have all indicated that the original error estimate δ_Q is insufficient except for small R, so the combined data set is needed to provide a valid uncertainty. These results confirm that δ_Q highly underestimates the uncertainty of the measurement, starting already at $R = 10°$ at the highest latitude cut. Therefore, when analysing the LAT data results to conclude whether there is an underlying non-zero Q signal, the standard deviation (of north and south combined) of Q in this simulated data set will be used as the uncertainty.

Since the contribution of IEM photons is very small at high latitudes, the shape of the curves of δ_Q and σ_Q are very similar to those of the isotropic data results. There seems to be a consistency in that the spread in the north data is smaller at first and then exceeds the southern spread somewhere between $R = 10°$ and $15°$. A satisfying explanation for this has

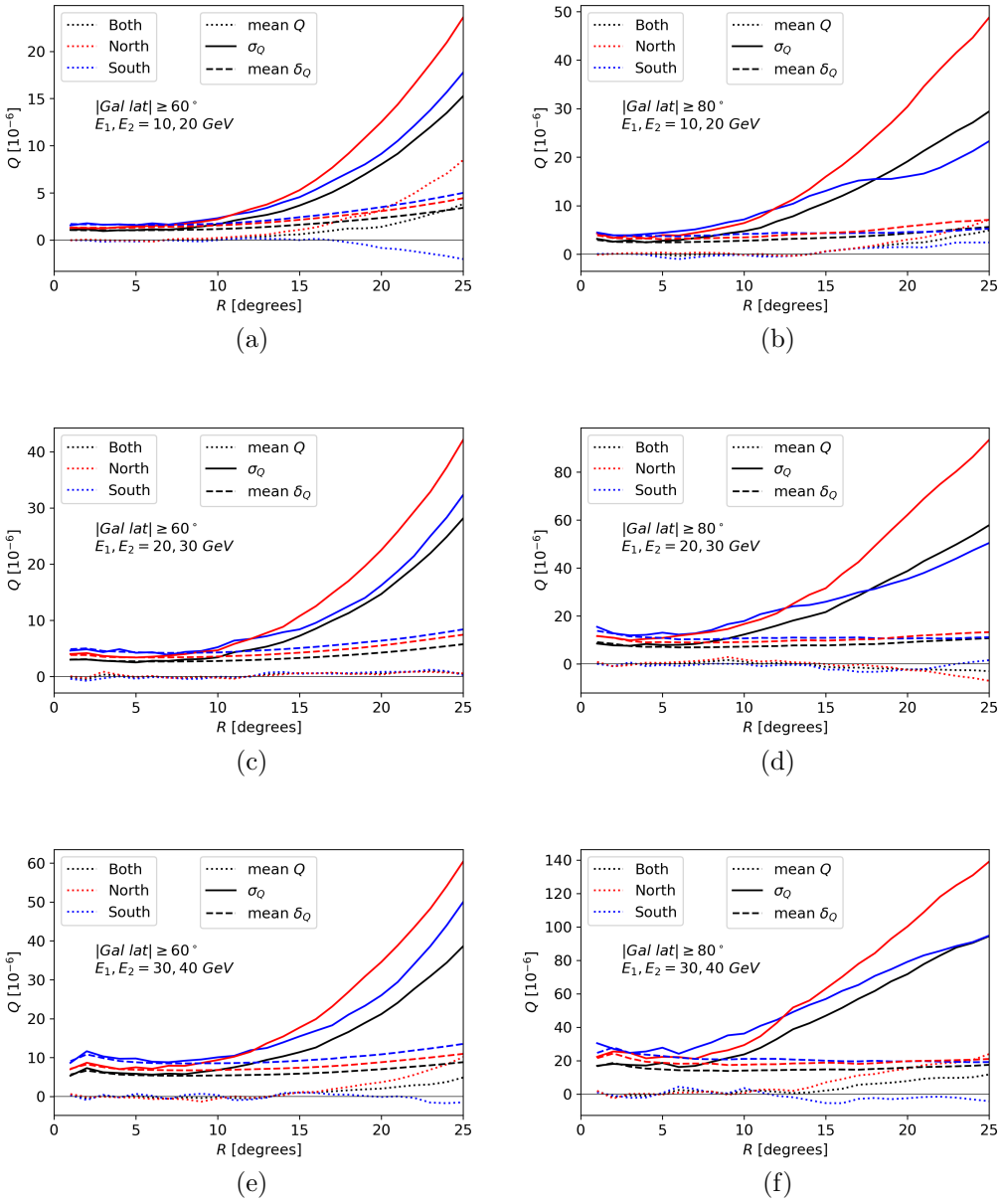


Figure 4.4: The statistical results of calculations on 200 simulations of data mimicking only Galactic emission, combined with 200 mimicking an isotropic signal, distributed according to the LAT exposure. The two columns represent different latitude cuts, and each row a different energy bin combination.

not been found, except that the results suggest this to be due to the exposure.

Four of these six plots show a significant bias at large R. It is in fact larger relative to the errors than it was in the separate results, and all positive. An attempt was made to correct for this by weighting the unit vectors of each photon with the effective area. To obtain the effective area, one needs the instrument response functions of the LAT, which were accessed using the `pyIrfloader` python script. Interestingly, the weighting did not change the results

significantly, so to correct for the exposure ,the live time of the LAT probably also needs to be considered. However, the bias in Q only appears at $R \geq 15°$ (recall Tashiro et al. (2014) did not include $R > 20°$) and is contained well within the standard deviation. Therefore this should be kept in mind when analysing the LAT data, but it is not crucial to correct for to obtain useful results.

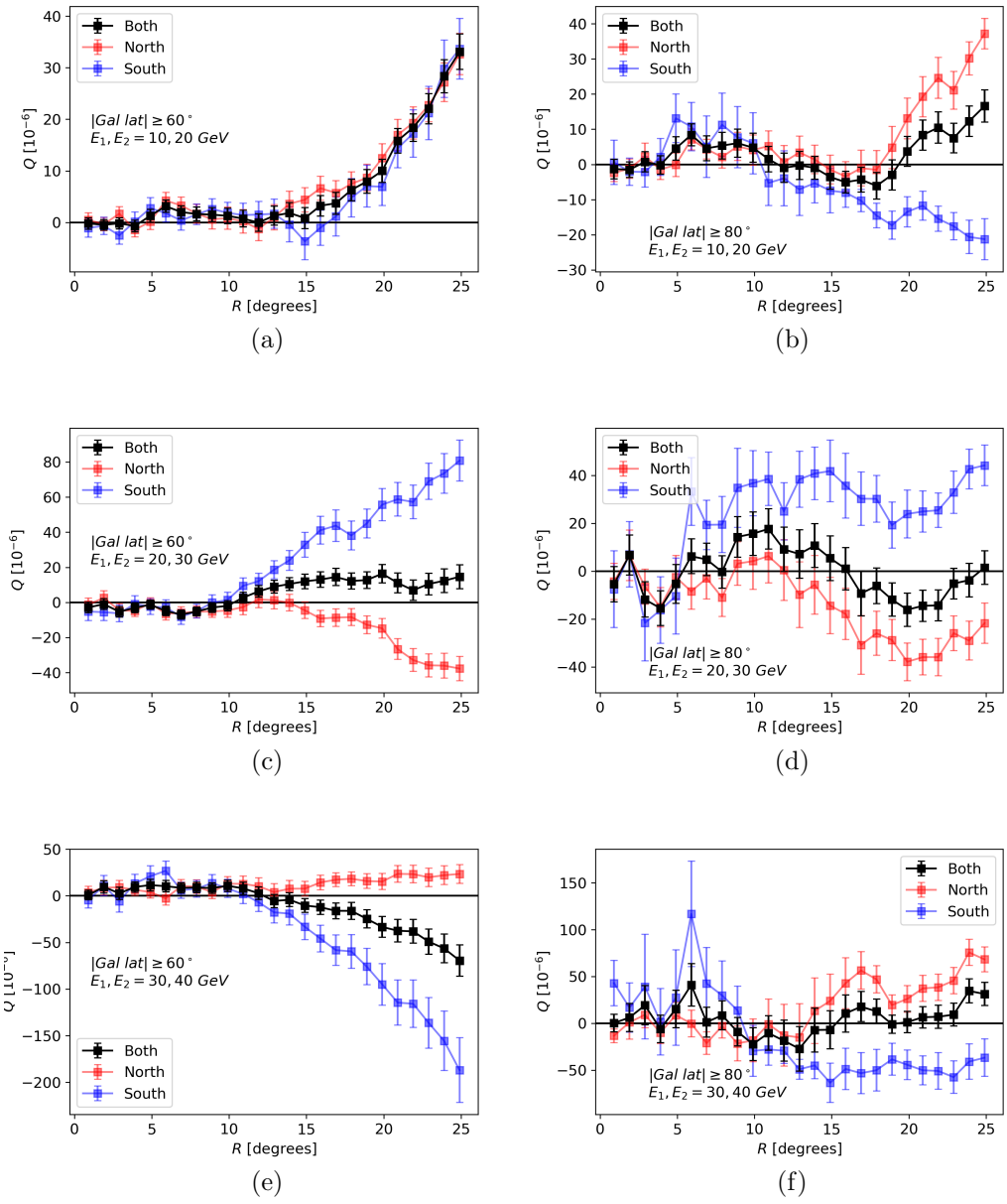


Figure 4.5: Examples of the results from calculating Q from the LAT data. The error bars are δ_Q. The two columns represent different latitude cuts, and each row a different energy bin combination.

4.2 Results from LAT data

Figure 4.5 shows six examples of results from calculations using LAT data. In these plots the error bars are still δ_Q, so they do not give a valid statistical certainty of the measurements and are only included as an example and to study the hemispheric dependence. Since the sky can only provide one realization of data, these results are no statistics but the actual calculated Q value at each R.

The hemispheric dependence does not appear to have an obvious pattern, although a majority of the plots show the previously observed behavior of a discrepancy between the northern and southern values becoming significant between $R = 10°$ and $20°$. However when comparing each individual plot with the corresponding plot in Figure 4.4, there is little to no correlation. For example Figure 4.4(a) shows a very strong discrepancy between north and south, while Figure 4.5(a) lacks one completely. On the other hand, the overall bias they display agrees well between these two plots, while not so well between Figures 4.4(e) and 4.5(e). Given the ambiguous results from both LAT and simulated data, the conclusion is that, independent of R, there does not seem to be a systematic north-south discrepancy in the gamma-ray sky that needs to be considered when inferring Q. Hemispheric differences will appear, but are most likely random effects, possibly caused by the varying LAT exposure. Therefore the combined data will be used for the final analysis of Q, as in Tashiro et al. (2014).

When it comes to the bias, the results are inconclusive. Nevertheless, there is an agreement between the combined simulated data and LAT data in a majority of the resulting plots, and it will therefore be included in further discussion.

4.3 LAT data results combined with improved error estimate

All the results of the Q calculations using the LAT data are displayed in Figure 4.6. The simulated data is used to estimate the statistical uncertainty. There are three columns, one for each latitude cut, each holding six plots corresponding to the different combinations of energy bins. Q for the LAT is plotted in blue with δ_Q error bars over the mean of the Q distribution from the combined simulated data in green with the σ_Q as error bar. As in Tashiro et al. (2014), any Q values that lie more than two σ away from the simulated data values is marked in red. The simulated data values are used as reference instead of zero to take the bias into account, although examining the plots suggests it wouldn't have made a difference. It is obvious from these results however that using σ_Q as the error estimate makes a crucial difference in determining whether a non-zero Q value is statistically significant or not. As seen in the plots a majority of the deviations lie within $1\sigma_Q$ of the simulated results. In Tashiro et al. (2014), simulated data was also used for comparison, but with a uniform distribution only. They found δ_Q to be comparable with the standard deviation of the distribution of simulated data results, so the strong amplification of the uncertainty displayed here is due to the LAT exposure and the IEM.

The three columns show good agreement, where plots corresponding to the same selec-

tions of energy bins show similar shape and structure between latitude cuts. There is also a clear trend of the deviation from zero becoming smaller with increasing latitude cut, mostly at large R. This is interesting because at lower latitudes there is more contamination from Galactic emission, so this trend implies that the strong deviation from zero is a Galactic effect.

For comparison, the results of Tashiro et al. (2014) are shown in Figure 4.7. Comparing columns 4.6(b) and 4.6(c) to the two columns in Figure 4.7 results in good agreement in the structure of the curves, apart from the second row, i.e. the energy bin combination $E_1, E_2 = 10, 30$ GeV. An important difference however is that the peak-like structure observed by Tashiro et al. (2014) around $R = 12^\circ$ is not prominent in the results of this work. In the cases that a peak is visible, for example in the bottom plot in 4.6(c), its maximum absolute value lies well within the valid uncertainty. These results also do not show at all as many outliers over 2σ as were observed in Figure 4.7. In fact, in Figure 4.6 no plot has more than two values marked in red. There are five plots that have a positive outlier appearing at $R \approx 7^\circ$, and three of them are at the highest latitude cut, suggesting it is not a Galactic signal. However, since it is a single value in all cases, it is not a significant result. Within $2\sigma_Q$ is the 95% confidence level, so a 2σ outlier has a probability of one in 20. As there are 25 calculations in each plot (one for each R) and 18 plots there is a total of 450 calculations which are also not all independent as photons within all previous R values are included within the next. Around 20 outliers at 2σ can therefore be expected using a single realization of data.

Overall the curves in these results have a more subdued structure than in Tashiro et al. (2014), which is what would be expected if the underlying signal was zero. Few data points would then give a non-zero signal that comes closer to zero as more data is added.

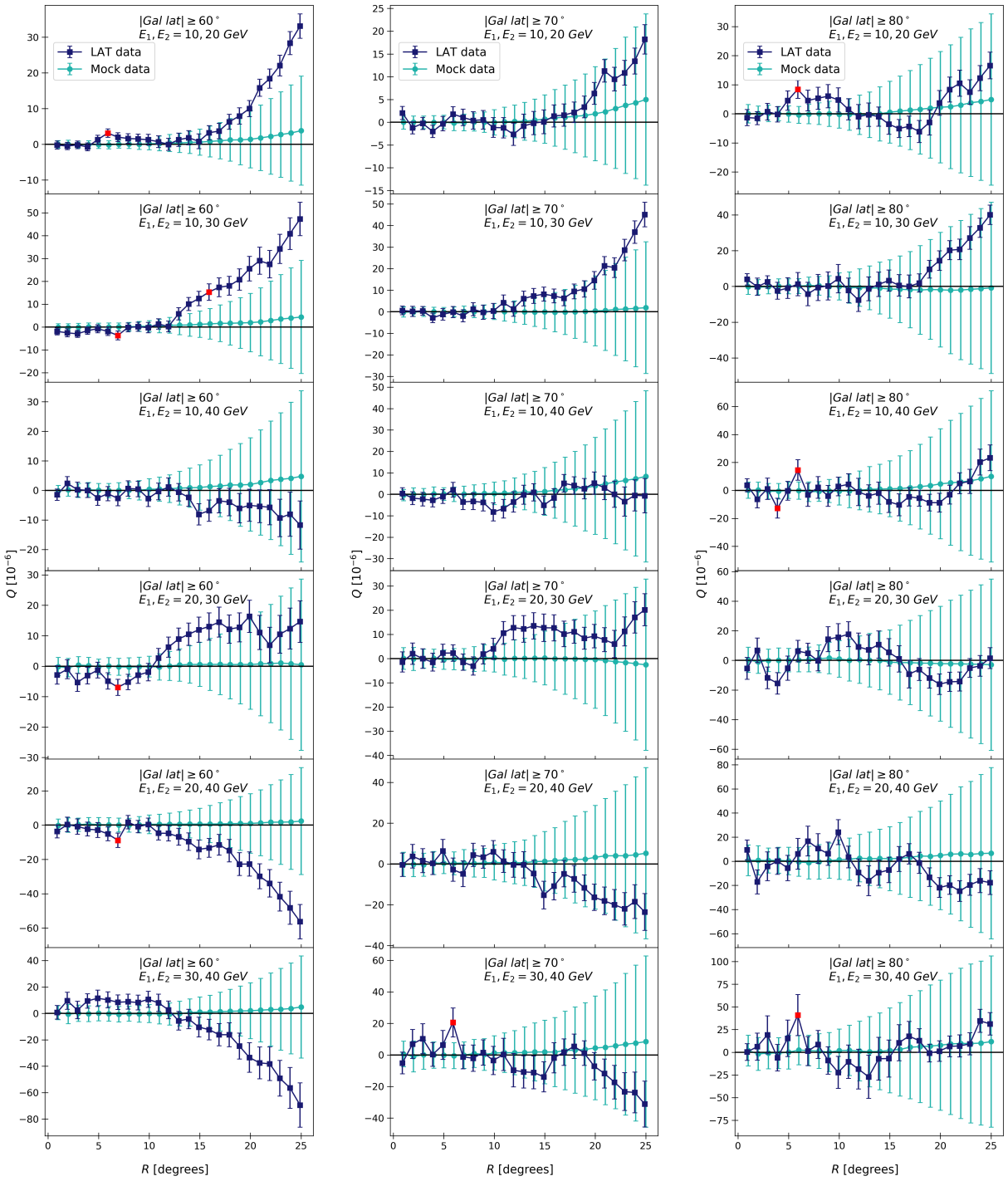


Figure 4.6: All results from calculating Q using the LAT data. The Q values calculated with LAT data with σ_Q as error bar are plotted over the mean Q and standard deviation of a distribution of results from 200 simulations of simulated data including LAT exposure IEM effects. Each column corresponds to a different latitude cut and each row a different energy bin combination.

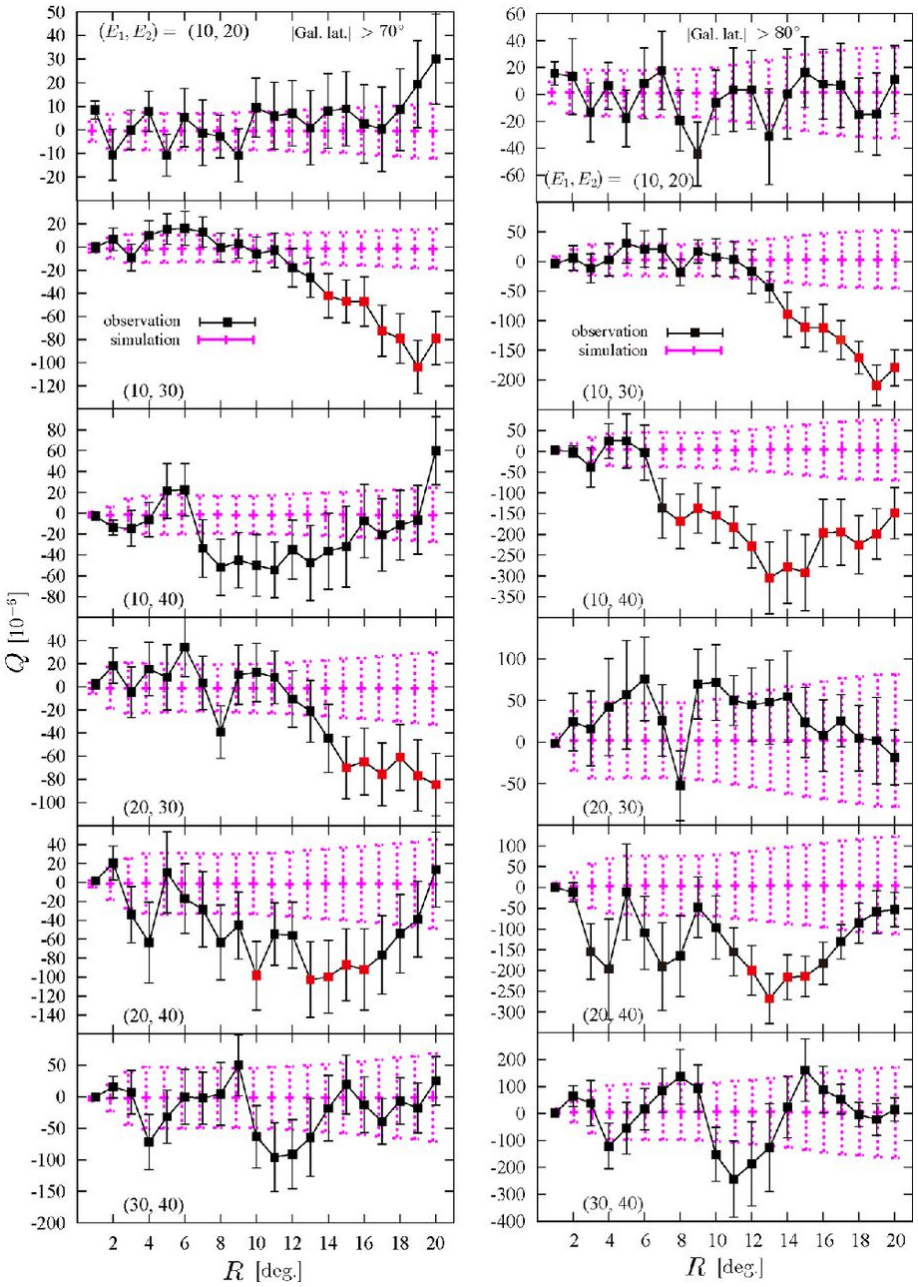


Figure 4.7: Results obtained by Tashiro et al. (2014). The Q values calculated with LAT data with σ_Q as error bar is plotted over the mean Q and standard deviation of a distribution of results from 10^4 simulations of uniformly distributed data. Each column corresponds to a different latitude cut and each row a different energy bin combination. Figure taken from Tashiro et al. (2014).

Conclusions and further discussion

In conclusion, the combined results of this work presented in section 4.3, indicate that the Q values as measured are compatible with zero. This suggests no sign of a helical cosmological magnetic field, in contradiction with Tashiro et al. (2014). There are two main factors behind this contradictory result. For one, including the additional data that has been gathered by the LAT lead to more subdued curves lacking the peak structure that was a main argument for the conclusions in Tashiro et al. (2014). Furthermore, and perhaps even more importantly, the error estimate is highly improved in this work, as the LAT exposure and IEM effect were not considered by Tashiro et al. (2014) and have been shown in the present work to significantly increase the uncertainties.

This conclusion does not rule out the existence of a helical cosmological magnetic field. Several assumptions are made in the physical motivation described in sections 2.1 and 2.2. For instance, there is no way of knowing how many of the photon triples used actually do originate from the same source. It may very well be so that there are so few that any signal they might carry is completely drowned by all of the random data. In other words, the constructed statistic, Q, in equation 2.7 may in practice not be as closely related to the helical part of the correlator of the magnetic field as assumed. One way to improve the current method to bring the practical execution more in line with the physical model, would be to include a larger energy range. The reason only photons above 10 GeV are used in the analysis is that the LAT performance is energy dependent below this value, but it is basically constant from 10 GeV up to 1 TeV (Atwood et al. 2013a) [1]. Therefore, since more E_3 photons improves the statistic, and the higher the energy of the photon the better it is thought to probe the position of the source, there is no reason not to include a much wider energy range for the E_3 photons. Doing so might allow a non-zero signal, if present, to stand out over the noise. There is a higher risk of a look-elsewhere effect in this approach, so significant results would have to be properly confirmed by comparison to simulations.

Furthermore, it should be confirmed that applying the updated uncertainty obtained in this work to the results in Tashiro et al. (2014) indeed makes the strong negative signal found in some of them compatible with zero. A method to test this in the future would be to sample a simulated data set from the combined simulated data set used in this work, with the same number of photons in each energy bin as the LAT data used in Tashiro et al. (2014), and overlay the statistical results of Q calculations to see if their signal can also be contained within the increased uncertainty. If not, a comparison should be made between

[1] Plots demonstrating the energy dependence of the performance can be found on the Fermi LAT Performance page, `http://www.slac.stanford.edu/exp/glast/groups/canda/lat_Performance.htm`

their results and calculations on $P8R3$ LAT data from the same time to see if the update made a significant difference. The results of this work could also be confirmed by repeating the analysis using a more strict event class for the LAT data, like ULTRACLEAN, as data contamination due to Galactic cosmic rays is lower in this event class.

Considering that the sought magnetic field may exist despite the undetected signal, a different approach may be needed. A method discussed during this project would be to select out the point sources rather than excluding them. The quantity Q as defined by Tashiro et al. (2014) requires diffuse emission, but the statistical argument should be transferable to summing over all point sources. Instead of, as currently, summing over all E_3 photons and assuming that a significant fraction are cascade photons resulting from TeV blazar emission and that the patch selected around them will contain a few photons from the same source, one could sum over all point sources, being more certain of the surrounding photons origin. The current point source catalogue holds over 3000 identified or associated sources of the blazar class that could be used in such an analysis (The-Fermi-LAT-collaboration 2019). Even if only the ones fully identified would be used, that would still be an amount of the same order as the number of E_3 photons at the highest latitude cut, so no statistical certainty is lost.

Another approach to be tested would be to select photons within intervals of R for each calculation as opposed to circular patches with different radii. So by defining shells of a certain width, preferably so that roughly the same amount of photons lie within each shell, Q can be studied as a function of distance from the E_3 photon (or point source). That way, every calculation would only contain new photons in the two lower energy bins, instead of always including those already used in the previous patches. A signal at a particular radius wouldn't be diluted by data not carrying a handedness and could be more prominent.

It should be pointed out that the results of Tashiro et al. (2014) were remarkable regardless of any connection to magnetic fields, as any handedness in the arrival directions of gamma rays would be unexpected and need an explanation. This work provides such an explanation, that the left-handed result they obtained was a statistically insignificant variation misinterpreted because of the underestimated uncertainty. But since only a relatively small energy range has been considered in the analysis, it has also not been proven that there is no overall handedness in the arrival directions of extragalactic gamma rays in general. In order to do so, one should use the generalized approach described in section 2.3, and sum equation 2.8 over all possible combinations fulfilling $E_1 < E_2 < E_3$, but without a fixed value for E_3 unlike what has been done before. Since this approach is not constrained by the physical model, all energy ranges from 10 GeV up to 1 TeV could be included. At high energies the photons are more likely to be extragalactic and the additional data allows using only a high latitude cut of $\pm 80^\circ$ to exclude galactic emission and still have enough data to achieve reasonable statistical certainty.

To summarize, the results of this work have two possible interpretations. Either there is in fact no helical cosmological magnetic field present in the IGM, which would have consequences for different models of the early Universe (Vachaspati 2001). Or, the method developed by Tashiro et al. (2014) and used in this work needs to be altered further in order to detect an existing non-zero helicity of intergalactic magnetic fields.

Acknowledgements

I would like to thank my two supervisors for their much needed guidance, for all of the long discussions and for their inspiring dedication to the project. They never failed to answer any of my endless questions, but above all they made me believe in my own ideas, for which I am so grateful. This hands-on experience in performing research was much more than I could hope for.

Of course I also want to thank everybody else at Nordita West for making me feel most welcome. A special thanks to Sveva Castello for being a solid sounding board and good friend.

I am also grateful to my parents for proof reading this report and for their endless support.

Last but not least, I want to thank Rafat Yousif for being my rock, my sounding board, for making me laugh and for always believing in me more than I do.

Bibliography

Acero, F., Ackermann, M., Ajello, M., et al. 2016, The Astrophysical Journal Supplement Series, 223, 26

Ackermann, M., Ajello, M., Albert, A., et al. 2015, The Astrophysical Journal, 799, 86

Atwood, W., Albert, A., Baldini, L., et al. 2013a, arXiv:1303.3514 [astro-ph]

Atwood, W. B., Baldini, L., Bregeon, J., et al. 2013b, The Astrophysical Journal, 774, 76

Bourdin, P.-A. & Brandenburg, A. 2018, The Astrophysical Journal, 869, 3

Brandenburg, A., Enqvist, K., & Olesen, P. 1996, Physical Review D, 54, 1291

Bruel, P., Burnett, T. H., Digel, S. W., et al. 2018, arXiv:1810.11394 [astro-ph]

Durrer, R. & Neronov, A. 2013, The Astronomy and Astrophysics Review, 21, 62

Elyiv, A., Neronov, A., & Semikoz, D. V. 2009, Physical Review D, 80, 023010

Field, G. B. & Carroll, S. M. 2000, Physical Review D, 62, 103008

Kronberg, P. P. 1994, Reports on Progress in Physics, 57, 325

Neronov, A. & Vovk, I. 2010, Science, 328, 73

Tashiro, H., Chen, W., Ferrer, F., & Vachaspati, T. 2014, Monthly Notices of the Royal Astronomical Society: Letters, 445, L41

Tashiro, H. & Vachaspati, T. 2013, Physical Review D, 87, 123527

The-Fermi-LAT-collaboration. 2019, arXiv:1902.10045 [astro-ph]

Vachaspati, T. 1991, Physics Letters B, 265, 258

Vachaspati, T. 2001, arXiv:astro-ph/0111124