Title

Detection of the 511 keV positron annihilation line with the Compton Spectrometer and Imager

Author

Kierans, Carolyn

Publication Date

2018

Detection of the 511 keV positron annihilation line with the Compton Spectrometer and Imager

by

Carolyn Kierans

A dissertation submitted in partial satisfaction of the
requirements for the degree of

Doctor of Philosophy
in
Physics
in the
Graduate Division
of the
University of California, Berkeley

Committee in charge:

Professor Steven Boggs, Co-Chair
Professor Stuart Bale, Co-Chair
Professor Daniel McKinsey
Professor Karl van Bibber

Summer 2018

ABSTRACT

Detection of the 511 keV positron annihilation line with the Compton Spectrometer and Imager

by

Carolyn Kierans

Doctor of Philosophy

University of California, Berkeley

Professor Steven Boggs, Co-Chair
Professor Stuart Bale, Co-Chair

The signature of positron annihilation, namely the 511 keV γ-ray line, was first detected coming from the direction of the Galactic center in the 1970's, but the source of Galactic positrons still remains a puzzle. The measured flux of the annihilation corresponds to an intense steady source of positron production, with an annihilation rate of ~10^{43} e^+/s. Spatially, the 511 keV intensity is strongest in the Galactic center region, with an additional component that is consistent with the Galactic disk; however, the unique morphology is not well constrained.

The Compton Spectrometer and Imager (COSI) is a balloon-borne soft γ-ray (0.2–5 MeV) telescope designed to perform wide-field imaging and high-resolution spectroscopy, with a goal of furthering our understanding of Galactic positrons. COSI employs a compact Compton telescope design, using 12 cross-strip germanium detectors to track the trajectory of incident photons, where position and energy deposits from Compton interactions allow for a reconstruction of the source sky position and significant background reduction.

COSI had a record-breaking 46-day balloon flight in May–July 2016 from Wanaka, New Zealand, and here we report on the detection and analyses of the 511 keV emission from those observations. To isolate the Galactic positron annihilation emission, we have developed a background subtraction technique utilizing the COMPTEL Data Space. With this new method, we find a $7.2\,\sigma$ detection of the 511 keV line and a broader spatial distribution of the emission than has been previously reported.

ACKNOWLEDGMENTS

First thank you to Professor Steve Boggs, my advisor, for your constant positivity and support. Thank you for advocating for me and giving me opportunities to excel.

A big thank you to the whole COSI team! Thanks to Dr. Andreas Zoglauer for being so encouraging and never getting tired of my countless programming questions. Thanks to Dr. John Tomsick for the science discussions and for stepping up as a leader in the group. Thanks to Brent Mochizuki and Steve McBride for the engineering expertise, grilling skills, and, above all else, friendship. Thanks to Dr. Alex Lowell for having an answer to everything, from science, to electronics, to random internet facts. And thanks to Clio Sleator with whom I have shared more than just an office, a room while abroad, and a love (verging on obsession) of tea. The COSI group, with everyone's individual contribution, was the reason I had such an amazing time over the past six years.

Thanks to my friends in Berkeley and those back home in Vancouver. Thank you to my family for the constant support.

The largest thanks of all goes out to my mother, Catherine. Thank you for sacrificing everything for me, thank you for encouraging me to be who I am, and thank you for reading every page of this thesis; you always go above and beyond.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

COSI	Compton Spectrometer and Imager
NCT	Nuclear Compton Telescope
HEAO-3	High Energy Astronomy Observatory 3
HRGRS	High Resolution Gamma-Ray Spectrometer (on HEAO-3)
SMM	Solar Maximum Mission
GRS	Gamma Ray Spectrometer (on SMM)
CGRO	Compton Gamma Ray Observatory
COMPTEL	Imaging Compton Telescope (on CGRO)
OSSE	Oriented Scintillation Spectrometer Experiment (on CGRO)
BATSE	Burst And Transient Source Experiment (on CGRO)
EGRET	Energetic Gamma Ray Experiment Telescope (on CGRO)
TGRS	Transient Gamma-Ray Spectrometer (on *Wind*)
INTEGRAL	International Gamma Ray Astrophysics Laboratory
SPI	Spectrometer on INTEGRAL
IBIS	Imager on Board the INTEGRAL Satellite
PICsIT	PIxelated CsI Telescope (on INTEGRAL)
BAT	Burst Alert Telescope (on *Swift*)
BeppoSAX	Beppo Satellite per Astronomia a raggi X
WFC	Wide Field Camera (on BeppoSAX)
CZTI	Cadmium Zinc Telluride Imager (on *AstroSat*)

SGD	Soft Gamma-ray Detector (on *Hitomi*)
NuSTAR	Nuclear Spectroscopic Telescope Array
AMEGO	All-sky Medium Energy Gamma-ray Observatory
e-ASTROGAM	enhanced ASTROMeV/GAMMA-LIGHT
VLA	Very Large Array
ESA	European Space Agency
NASA	National Aeronautics and Space Administration
JAXA	Japan Aerospace Exploration Agency
LBNL	Lawrence Berkeley National Laboratory
IRAP	Institut de Recherche en Astrophysique et Planétologie
CERN	European Organization for Nuclear Research
SSL	Space Sciences Laboratory
MEGAlib	Medium Energy Gamma-ray Astronomy library
DEE	detectors effects engine
MEM	Maximum Entropy Method
MLM	Maximum Likelihood Method
ML-EM	maximum-likelihood expectation-maximization
LM-ML-EM	list-mode maximum-likelihood expectation-maximization
CKD	Compton Kinematic Discrimination
ARM	angular resolution measure
CTD	collection time difference
GTI	good time interval
EHC	earth horizon cut
GSE	ground support electronics

GCN	Gamma-ray Coordination Network
CDS	COMPTEL Data Space
GC	Galactic Center
Sgr A*	Sagittarius A*
Cas A	Cassiopeia A
Cyg X-1	Cygnus X-1
Cen A	Centaurus A
ISM	interstellar medium
CMZ	central molecular zone
AGN	active galactic nuclei
WR	Wolf-Rayet
AGB	asymptotic giant branch
XRB	X-ray binary
HMXB	high-mass X-ray binary
LMXB	low-mass X-ray binary
TDE	tidal disruption event
GRB	gamma-ray burst
SN	supernova
CCSN	core-collapse supernova
SNIa	Type Ia supernova
SNR	supernova remnant
BH	black hole
NS	neutron star
WD	white dwarf

REP	relativistic electron precipitation
DREP	duskside relativistic electron precipitation
DM	dark matter
NFW	Navarro-Frenk-White
WIMP	weakly-interacting massive particle
MHD	magnetohydrodynamic
Ps	positronium
o-Ps	ortho-positronium
p-Ps	para-positronium
EC	electron capture
BR	branching ratio
FOV	field of view
ACS	anti-coincidence shield
PMT	photomultiplier tube
CCT	compact Compton telescope
TRL	technology readiness levels
ADC	analog-to-digital converter
HV	high voltage
GeD	germanium detector
HPGe	high-purity germanium
LN2	liquid nitrogen
BGO	bismuth germanium oxide $(Bi_4Ge_3O_12)$
CsI	cesium iodide
NaI	sodium iodide

SPB	Super Pressure Balloon
ULDB	ultra-long duration balloon
CSBF	Columbia Scientific Ballooning Facility
SIP	Support Instrument Package
LOS	line-of-sight
MCF	million ft^3
GPS	Global Positioning System
UTC	Coordinated Universal Time
FWHM	full width half max

1

INTRODUCTION

The initial astrophysical studies of the sky began with Galileo and optical telescopes, but soon astronomers were building instruments to observe the cosmos using other bands in the electromagnetic spectrum. Each energy range, e.g., radio waves, infrared, or X-rays, reveals unique information about the cosmic source in question. As technology advanced over time and telescopes became more sensitive across the whole spectrum, there was one range that fell behind: γ-rays in the MeV regime. This gap in sensitivity, often referred to as the *MeV Gap*, is not due to a lack of compelling science but instead to a limitation in technology, and there remain many open questions.

The 511 keV γ-ray line produced in the Galaxy through the annihilation of an electron and its anti-particle, the positron, was first detected in the 1970's. Telescopes measure an extended source with a relatively high flux ~1×10^{-3} $\gamma/\mathrm{cm}^2/\mathrm{s}$ and a strong concentration in the center of the Galaxy, an emission that is unlike any other wavelength. While ^{26}Al, a positron emitting isotope, could explain the positron population in the disk of the Galaxy, there is no convincing explanation for the excess of positrons in the Galactic Center region. This remains one of the major puzzles in γ-ray astrophysics over the last half-century.

This dissertation presents the detection and analyses of the Galactic positron annihilation signal with the Compton Spectrometer and Imager (COSI). COSI is a balloon-borne Compton γ-ray telescope that is designed to further our understanding of Galactic positrons, perform novel polarization measurements of compact sources, and map the emission of nucleosynthesis in our Galaxy. The COSI instrument had a successful 46-day balloon flight from Wanaka, New Zealand, in May–July 2016. The flight path provided over 1.5 Ms of exposure of the Galactic Center (GC).

One of the difficulties of measurements in this energy range is the dominance of the γ-ray background radiation. Through the work leading up to this dissertation, it was determined that traditional background subtraction techniques were not suitable for Compton telescopes or for the strongly varying background emissions seen throughout the balloon flight. Therefore, a method for background subtraction of line emission has been formulated based on the CDS. With the new routine, a 7 σ detection of the 511 keV line of the Galactic center region from the COSI 2016 balloon flight was achieved. Furthermore, a spatial distribution of the positron annihilation signal was detected that is broader than the previously reported models of the emission. In addition, the CDS leads to a binned-imaging algorithm that gives the most direct image of the Galactic emission to date.

Part I of this dissertation provides the historical context. Chapter 2 begins with the initial detections of the 511 keV annihilation line and summarizes the current state-of-the-art measurements by INTEGRAL/SPI. Chapter 2 continues with a review of the present understanding of Galactic positrons and concludes with an overview of the possible cosmic sources.

Part II covers the COSI mission. Chapter 3 gives a brief introduction to γ-ray astronomy and focuses on Compton telescope technology. The COSI instrument is described in detail. Chapter 3 finishes with an overview of the software programs used by the COSI collaboration. Chapter 4 details the calibration of COSI prior to the 2016 balloon flight and concludes with the measured higher-level performance of the telescope. Chapter 5 summarizes the 2016 balloon campaign from Wanaka, New Zealand.

Part III focuses on the signature of Galactic positrons detected during the 2016 flight. Chapter 6 overviews the GC observations and motivates the need for a new background subtraction technique. Chapter 7 introduces the CDS and details the routine developed for background subtraction for the positron annihilation analyses. Chapter 8 validates the routine with simulations of the Galactic emission. Chapter 9 presents the final results of the detected positron annihilation signature from the Galaxy. Chapter 10 concludes with a summary of the work presented and a look to the future of γ-ray astrophysics and the Galactic positron puzzle.

PART I

BACKGROUND MATERIAL

2

POSITRON ANNIHILATION IN THE MILKY WAY

2.1 INTRODUCTION TO GALACTIC POSITRONS

Positrons are the anti-particles of electrons, which means they have the same mass but opposite charge. Their existence was predicted by Dirac (1931), and since the first measurements of positron tracks in cloud chambers were disregarded (Skobeltzyn, 1985; Chao, 1930), the credit of the particle's discovery was given to Anderson (1932), for which he was given the Nobel prize in 1936. The signature of electron-positron direct annihilation is two γ-rays emitted at the rest energy of the electron (and positron) at 511 keV, abiding by the conservation of momentum. This signature was first observed by Klemperer & Chadwick (1934), which confirmed the anti-particle nature of the positron. Mohorovicic (1934) predicted the existence of a hydrogen-like bound state between a positron and electron that can be formed before annihilation at low energies. This was finally measured by Deutsch (1951) and that bound state became known as positronium (Ps).

Formed by two spin-$\frac{1}{2}$ particles, positronium has four basis states, analogous to the hydrogen atom: a singlet-state with a total spin angular momentum of $s = 0$, called para-positronium (p-Ps)

$$|0,0\rangle = \frac{1}{\sqrt{2}}(\uparrow\downarrow - \downarrow\uparrow), \tag{2.1}$$

and a triplet state with a total spin angular momentum of $s = 1$, called ortho-positronium (o-Ps)

$$|1,-1\rangle = \upuparrows \qquad |1,0\rangle = \frac{1}{\sqrt{2}}(\uparrow\downarrow + \downarrow\uparrow) \qquad |1,1\rangle = \downdownarrows\,. \tag{2.2}$$

The singlet-state will form 25% of the time. With an even charge parity, it decays via the emission of two photons (lifetime = 0.125 ns; Adkins et al. 2003), and results in a spectrum with a narrow line at 511 keV, like that seen from

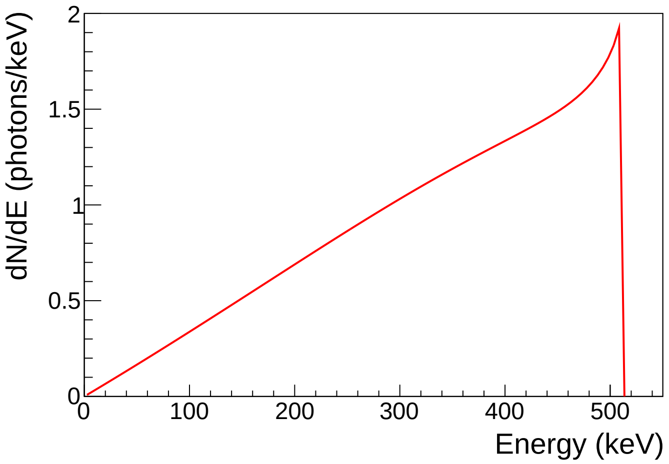


Figure 2.1: The predicted spectrum from Ore & Powell (1949) of the annihilation of the triplet-state of positronium, o-Ps, described in Equation 2.3.

direct-annihilation. The triplet-state will form 75% of the time, and with odd charge parity it decays via the emission of three photons (lifetime = 140 ns; Vallery et al. 2003) with energies totaling 1.022 MeV. Ore & Powell (1949) calculated the lifetime and spectrum resulting from o-Ps annihilation

$$F(k_1) = \int_{m-k_1}^{m} 2\left[\frac{k_1(m-k_1)}{(2m-k_1)^2} - \frac{2m(m-k_1)^2}{(2m-k_1)^3}\ln\frac{m-k_1}{m} + \frac{2m-k_1}{k_1} + \frac{2m(m-k_1)}{k_1^2}\ln\frac{m-k_1}{m}\right] \tag{2.3}$$

where m is the electron mass, and k_1 is the momentum quanta of one of the photons ($k_1 + k_2 + k_3 = 2m$). The spectrum as defined in Equation 2.3 is plotted in Figure 2.1.

Stecker (1969) was the first to determine that low-energy positrons in the Galaxy will predominantly form Ps before annihilating. In their work on solar flares, Crannell et al. (1976) found that a positronium fraction of $f_{P_s} = 1$ is theoretically expected for most astrophysical plasmas, meaning 100% of positrons would first form Ps before annihilation.

The 511 keV signature of electron-positron annihilation and the low-energy continuum from o-Ps were observed coming from the Galactic Center (GC) region in the 1970's, marking it as the first γ-ray line to be detected originating from outside of the solar system. After almost 50 years of studies, it is now known that the positron annihilation emission from the Galaxy is extended in size with a strong concentration in the center region of the Galaxy, called the

Galactic bulge, in addition to a low surface-brightness emission consistent with the Galactic disk. This observed spatial distribution is unlike that seen in any other wavelength. The steady source has a flux of ~1×10^{-3} $\gamma/\mathrm{cm}^2/\mathrm{s}$, and (depending on the spatial model) this corresponds to an annihilation rate of ~10^{43} e^+/s in the Galaxy.

The origin of Galactic positrons remains a major puzzle in γ-ray astrophysics. Although there are many theorized sources of positrons in the Galaxy, no source distribution matches the observed annihilation emission. The only confirmed source of positrons to date is from the β^+ decay of ^{26}Al. The Galactic distribution of ^{26}Al has been mapped by its 1.8 MeV γ-ray emissions and the observed distribution can possibly account for the 511 keV emission seen along the Galactic disk, but the 511 keV emission in the Galactic bulge is still unexplained.

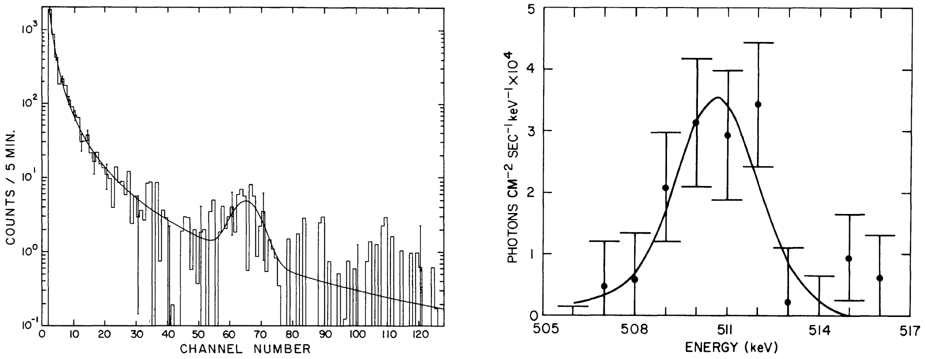
A comprehensive review of our understanding of Galactic positrons can be found in Prantzos et al. (2011), and some of the discussion presented here follows from their work.

2.2 HISTORICAL OVERVIEW

2.2.1 EARLY BALLOON MEASUREMENTS

The 511 keV signature of positron annihilation was first observed from the direction of the GC when γ-ray spectrometers were launched onboard high altitude balloons in the 1970's. In 1972, Johnson III et al. (1972) reported a 4 σ spectral feature at 473±30 keV with a flux of $1.8\pm0.5\times10^{-3}$ $\gamma/\mathrm{cm}^2/\mathrm{s}$, as shown in Figure 2.2 (a). This result was from a 17 hour balloon flight with a collimated sodium iodide (NaI) scintillator detector using on-off pointings of the GC region. Though they discussed positron annihilation as a possible source, they thought it unlikely unless the emission was red-shifted by ~0.07 in energy. Leventhal (1973) and Clayton (1973) were quick to point out that the expected spectrum for Galactic positron annihilation would be a line at 511 keV from p-Ps and direct annihilation, and a low-energy continuum from o-Ps annihilation, which, if observed by a spectrometer with poor energy resolution, would result in an apparent red shift.

Leventhal et al. (1978) were the first to fly a high-resolution γ-ray detector. The instrument consisted of a large ~130 cm^2 germanium detector with a 3.2 keV FWHM resolution. With 17.3 hours of data and alternating on-off pointings every 20 mins, they found a 5.6 σ excess at 511 keV when the instru-



(a) First detection of Galactic annihilation.

(b) 511 keV line.

Figure 2.2: (a) The 4 σ first detection of the 511 keV line from the Galaxy, and when converted from channel number to energy, this corresponds to a line at 473±30 keV. Due to the poor resolution of the NaI detectors, the 511 keV line and the o-Ps continuum appeared as a broad line at a lower energy. Figure 4 from Johnson III et al. (1972). (b) The first high-spectral resolution measurement of the 511 keV emission from the Galactic center region. The line is centered at 510.7±0.5 keV and the line-width is limited by the spectral resolution at $\lesssim$ 3.2 keV. Figure 2b from Leventhal et al. (1978).

ment was directed at the GC region; see Figure 2.2 (b). This was the first measurement that showed evidence of Galactic o-Ps annihilation below 511 keV; with the addition of the three-photon continuum to the spectral fit, an improvement in χ^2 was seen for a positronium fraction f_{Ps} close to 1. Such a high f_{Ps} confirmed that most of the positrons must be at low energies before annihilation. With the contribution of the o-Ps annihilation, the authors concluded that the measured flux corresponds to a substantial Galactic annihilation rate of ~3.3×10^{43} e^+/s (assuming a distance to the GC of 10 kpc).

After a number of successive launches with balloon-borne telescopes (Johnson III et al., 1972; Johnson III & Haymes, 1973; Haymes et al., 1975; Ling et al., 1977; Albernhe et al., 1981), it was found that the measured 511 keV flux varied with each measurement. The conclusion was that the source of positrons was itself variable, and for decades this was the prevailing assumption (Leventhal, 1991). Figure 2.3 shows a compilation of these early balloon-borne measurements and additional satellite measurements from 1970-1990. In 1981, Albernhe et al. (1981) recognized that the flux measured with each detector was correlated with the telescope field of view (FOV), and the variability could be explained as a steady extended emission. However, this view was not accepted by the majority of researchers until years of satellite observations in the 1990's provided no detections of variable activity.

Along with these initial measurements of the 511 keV signature of Galactic positron annihilation, a number of astrophysical sources were proposed to be the birth-site of these positrons. In 1973, Clayton (1973) discussed how explosive nucleosynthesis is a plausible source for positrons, with the primary contributors being the ^{56}Ni and ^{44}Ti decay chains. He therefore claimed that the measurement by Johnson III et al. (1972) was the first confirmation of explosive nucleosynthesis (which he helped pioneer). Nucleosynthesis, through the decay of ^{26}Al, has been the only confirmed source of Galactic positrons to date, and this is further discussed in Section 2.4.1. Leventhal et al. (1978) proposed a cosmic-ray (Section 2.4.3.1) or radio pulsar (Section 2.4.2.2) origin, but added "undoubtedly, the positrons giving rise to the observed feature come from a variety of processes." Lingenfelter et al. (1981) suggested the possibility of positron production and annihilation around a massive black hole in the Galactic center (Section 2.4.2.3). After decades of measurements of the Galactic 511 keV signal and attempts at sophisticated simulations, not many of these possible sources have been ruled out and the true source of Galactic positrons is still a mystery.

2.2.2 EARLY SATELLITE OBSERVATIONS

2.2.2.1 CONFIRMATION OF A STEADY SOURCE

The first satellite to detect the signature of Galactic positrons was the High Energy Astronomy Observatory 3 (HEAO-3), launched in September 1979. The High Resolution Gamma-Ray Spectrometer (HRGRS) instrument (Mahoney et al., 1980) consisted of 4 large high-purity germanium (HPGe) crystals surrounded by a cesium iodide (CsI) anti-coincidence shield (ACS), and it operated until June 1980 when the onboard liquid nitrogen was exhausted. HEAO-3 reports of significant time variability of the 511 keV flux (Riegler et al., 1981) further raised the (false) evidence for a variable source near the Galactic center. A measured decrease in flux by almost a factor of three in a six month period led the authors to conclude that the source was singular and compact $\lesssim 10^{18}$ cm (Lingenfelter & Ramaty, 1982). Figure 2.3 shows the "variability" of the 511 keV flux from the GC including the HEAO-3 results, as reported by Leventhal (1991).

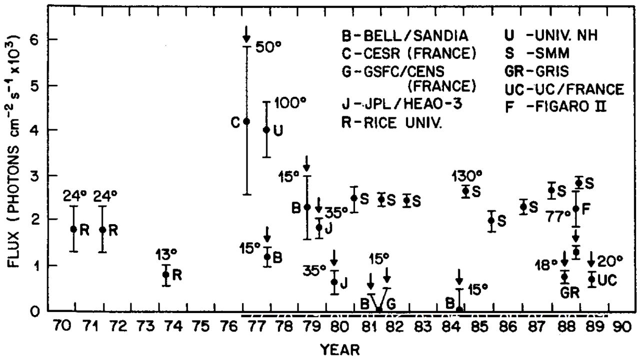


Figure 2.3: The measured Galactic center 511 keV flux from balloon-borne and satellite instruments assuming a point source emission. Published in 1991, it was still not entirely accepted that the positron annihilation in the Galaxy was a steady source. Figure 1 from Leventhal (1991).

The Solar Maximum Mission (SMM), launched in 1980 and operated till 1989, provided the first long-term monitoring of the inner Galaxy with its Gamma Ray Spectrometer (GRS) instrument (Forrest et al., 1980). The GRS instrument consisted of seven 7.6 cm diameter × 7.6 cm thick NaI detectors surround by a thick annulus of CsI acting as an active shield. With a FOV of 130° in the solar direction, the GC transited the instrument aperture in December each year of its operation and resulted in a measurement of the 511 keV line with significance of $> 30\sigma$. With these measurements, Share et al. (1988) constrained the year-to-year variability of the 511 keV flux to $< 30\%$. Furthermore, it was discussed in (Share et al., 1988) and (Share et al., 1990) that the detected flux during the GC transits is consistent with an extended steady source with intensity of $\sim 2 \times 10^{-3}$ $\gamma/\mathrm{cm}^2/\mathrm{s}$, weakening the argument for a variable source at the Galactic center. Although SMM observations seemed to indicate a steady diffuse source, a compact variable source at the Galactic center could not yet be ruled out.

The launch of the first γ-ray coded mask imager SIGMA on *GRANAT* in 1989 further confused the scenario of 511 keV variability. In 1991, an outburst of hard X-ray source 1E 1740.7-2942, later classified as the first microquasar (Mirabel et al., 1992), resulted in an apparent strong and broadened line at 440 keV with a FWHM ~200 keV (Sunyaev et al., 1991). Due to its close proximity to the GC, it was thought to be the long-sought variable source of annihilating positrons, where the spectrum could be explained with a broadened, red-shifted annihilation line. The source 1E 1740.7-2942 was dubbed the "Great Annihilator." It was proposed that the binary system would produce variable jets of e^--e^+ pairs; some of the positrons would annihilate in the accretion disk creating a signature like SIGMA observed, and some would escape and lose energy to give rise to the time-variable 511 keV line emission seen by previous instruments (Sunyaev et al., 1991; Gilfanov et al., 1994). Years later, however, it was reported that no 511 keV feature was seen in the concurrent observations with OSSE (Jung et al., 1995) and the Burst And Transient Source Experiment (BATSE) (Smith et al., 1996a) aboard the Compton Gamma Ray Observatory (CGRO).

Finally, with the combination of years of SMM data (Harris et al., 1994), a lack of 511 keV features detected by BATSE (Smith et al., 1996b), and a reanalysis of HEAO-3 data which refuted the previously reported transient activity (Mahoney et al., 1994), the steady-source nature of the 511 keV Galactic emission was firmly established. And with it came the necessity for the source to be spatially extended. The "variability" of the flux, as shown in Figure 2.3, can be largely explained by instruments with different FOVs measuring an extended source distribution, where the viewing angle of the instrument is smaller than the solid angle of the source. However, the transient emissions reported by SIGMA and HEAO-3 are limited to the analysis of a single instrument's data, and the lesson learned here is that γ-ray analysis, especially in the MeV range, is difficult and plagued with potential systematics. This will be further emphasized in the imaging attempts discussed below. Section 3.1 discusses these difficulties and ways to reduce systematics in further detail.

2.2.2.2 THE FIRST IMAGES

It was not until the launch of CGRO in 1991 with the Oriented Scintillation Spectrometer Experiment (OSSE) (Johnson III et al., 1993) onboard that the 511 keV emission was imaged, and it became fully accepted that the emission was extended. OSSE consists of 4 identical NaI-CsI phoswich detectors with a 3.8°×11.4° aperture and energy resolution ~9% at 511 keV. The OSSE observing strategy typically included 2 min pointings alternating between source and off-pointing background measurements, described in Purcell et al. (1992). With 67 weeks of observations of the GC and Galactic plane after 9 years of flight, Figure 2.4 is the first image of the 511 keV emission from the Galaxy reported by Purcell et al. (1997). This analysis combined the OSSE data with data from SMM and TGRS on *Wind* to provide a flux normalization. The reconstructed image, which shows an extended source with a symmetric bulge centered on the GC as well as emission along the Galactic plane, excluded a

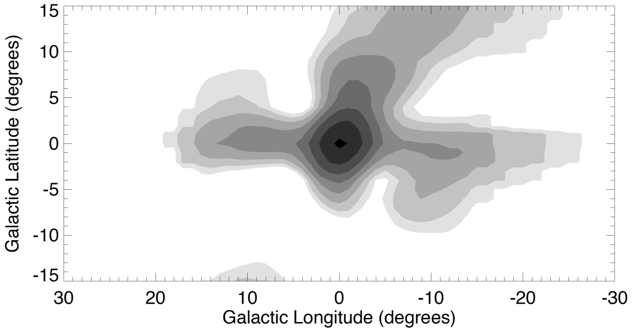


Figure 2.4: The first map of the Galactic 511 keV positron annihilation line from OSSE measurements, combined with TGRS and SMM data. A strong Galactic bulge component is seen along with emission consistent with the Galactic plane. The emission at higher latitudes was later confirmed to be an imaging artifact. The contours are exponentially scaled to show the weaker Galactic plane emission. Figure 5 from Purcell et al. (1997).

single point source as the origin of the 511 keV emission once and for all. The emission at higher latitudes b ~12°, referred to as the Positive Latitude Enhancement, or colloquially the "positron fountain," was suggested to be jet activity in the Galactic nucleus or from a more local origin. However, this enhancement was not seen in the image reconstruction of the o-Ps continuum (Milne et al., 2001) and would later prove to be an imaging artifact, again highlighting the difficulty of MeV γ-ray analysis.

2.3 RECENT RESULTS

2.3.1 IMAGING AND SPECTROSCOPY WITH INTEGRAL/SPI

The International Gamma Ray Astrophysics Laboratory (INTEGRAL) launched by ESA in 2002 is still in operation. One of the main science goals of the Spectrometer on INTEGRAL (SPI) (Vedrenne et al., 2003) is to study the emission from Galactic positrons. With high-spectral resolution germanium detectors (GeDs) and modest imaging abilities with a tungsten coded-aperture mask, SPI significantly advanced our understanding of Galactic positrons with thorough studies of the annihilation spectrum and the spatial distribution of the emission.

The SPI instrument consists of 19 high-purity GeDs with total geometrical area of 508 cm^2, giving an energy resolution of 2.1 keV FWHM at 511 keV.

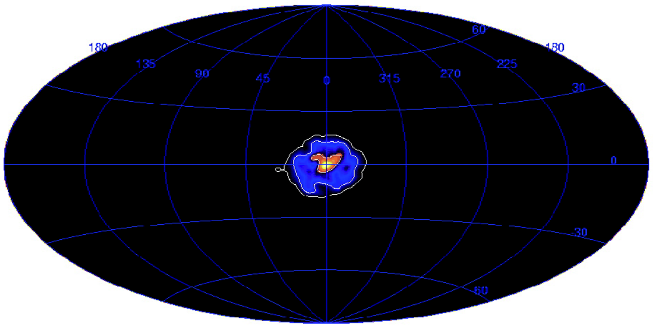


Figure 2.5: The image of the Galactic 511 keV emission with 1 year of SPI data, shown in Galactic coordinates.The emission from the Galactic bulge dominates the image and the low-surface-brightness emission from the disk is not yet detected. The contour levels indicate intensity of 10^{-2}, 10^{-3}, and 10^{-4} $\gamma/\text{cm}^2/\text{s}/\text{sr}$ from the center outwards. Figure 4 from Knödlseder et al. (2005).

Imaging is achieved through a coded-aperture mask made of tungsten blocks with a thickness of 3 cm set at a distance above the GeDs of 1.71 m, which gives an angular resolution of ~2.5°. The GeDs are surrounded by an ACS of bismuth germanium oxide ($Bi_4Ge_3O_12$) (BGO) crystals limiting the FOV to 16°. Additionally, a thin plastic scintillator is located just below the tungsten mask to allow for a further reduction of the 511 keV background from prompt interactions in the mask.

Using one year of SPI data, Knödlseder et al. (2005) showed the 511 keV positron annihilation emission to be predominantly in the Galactic bulge region, which was highly symmetric and centered on the GC with an extension of ~8° FWHM (see Figure 2.5). The Galactic disk emission, as seen by OSSE in Figure 2.4, is absent in the first year of data. Through model fitting, this study found that the 511 keV line flux ratio between the Galactic bulge and disk component was in the range of B/D ~1–3, which is more extreme than for any other wavelength. The authors concluded from the observed spatial distribution that the positrons annihilating in Galactic bulge must be from a different source than those annihilating in the Galactic disk.

With a total of 4 years of SPI data analyzed by Weidenspointner et al. (2008), the low surface brightness of the Galactic plane emission is finally apparent in the image; however, the new image shows an asymmetric distribution in the Galactic disk, (see Figure 2.6). The authors explained this asymmetry as resembling the asymmetry in the distribution of spectrally-hard low-mass X-ray binaries (LMXBs). As it is expected that the annihilation spatial distribu-

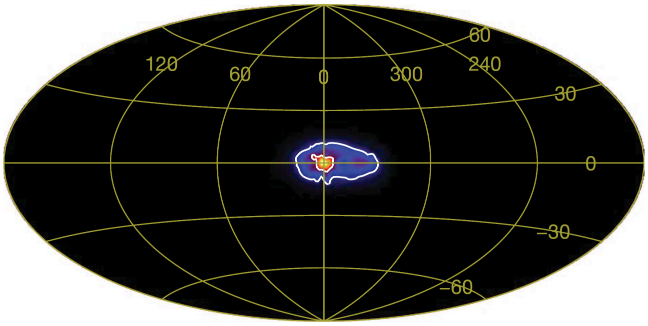


Figure 2.6: The measured Galactic 511 keV emission with 4 years of SPI data. The asymmetry in the emission along the Galactic plane was thought to trace the spatial distribution of LMXBs. The contours correspond to intensity levels of 10^{-2} and 10^{-3} $\gamma/\mathrm{cm}^2/\mathrm{s}/\mathrm{sr}$ from the center outwards. Figure 1 from Weidenspointner et al. (2008).

tion should match that of the positron source, it was suggested that LMXBs may be the dominant source of Galactic positrons. The authors estimate a positron production rate of $10^{41}\mathrm{e}^+/\mathrm{s}$ per hard LMXB which, in addition to known positron contribution from β^+ decay of ^{26}Al (see Section 2.3.3), can explain the Galactic annihilation emission. However, subsequent analysis by Bouchet et al. (2010) determines that the disk asymmetry can be better described as an offset in the Galactic bulge emission.

Another attempt at constraining the morphology of the Galactic positron annihilation was performed by Skinner et al. (2014). The authors analyzed more than 10 years of SPI data and considered a wide range of methods to model the background. Skinner et al. introduce a simple phenomenological model, referred to by the authors as the 'Baseline' model but referenced here as the 'Skinner Model,' that fits the observations as well as or better than astrophysically inspired models. The Skinner Model is illustrated in Galactic coordinates in Figure 2.7, and the different model component parameters are detailed in Table 2.1. Not only do they conclude that the observed asymmetry is better represented as an offset in the position of the Gaussian representing the Galactic bulge, agreeing with previous studies (Bouchet et al., 2010), they are the first to report that the best fit includes a point source component at the center of the Galaxy. The position of the point source in their model is chosen to be the location of Sagittarius A* (Sgr A*); with the resolution of 3°, a point source at the position of Sgr A* at (-0.06, -0.05) is indistinguishable from the GC. A point source contribution from Sgr A* implies that a fraction of the

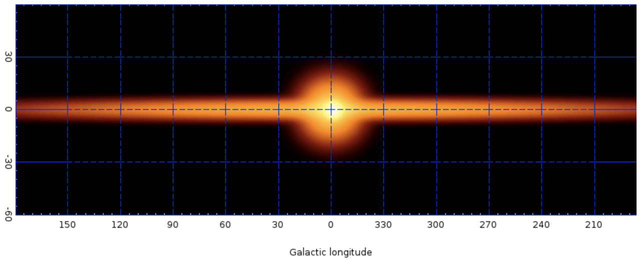


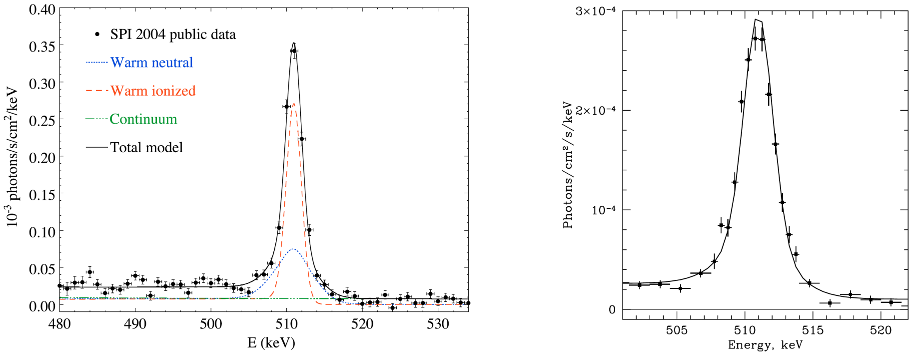
Figure 2.7: All-sky distribution of the Skinner 'Baseline' model in Galactic coordinates plotted on a logarithmic scale to show the low surface brightness of the disk emission. The small offset of the Galactic bugle Gaussian component is visible here. Figure 1 of Skinner et al. (2014).

Model Component	Position (l, b) [deg]	FWHM (σ_l, σ_b) [deg]	Flux [$\times 10^{-4}$ $\gamma/\mathrm{cm}^2/\mathrm{s}$]
Disk	(0, 0)	(212, 7)	14
Broad bulge	(0, 0)	(20.5, 20.5)	7.3
Narrow bulge	(-1.15, -0.25)	(5.9 5.9)	2.8
Central point source	(-0.06, -0.05)	(0, 0)	1.2

Table 2.1: Skinner Model components representing the Galactic 511 keV emission (Skinner et al., 2014).

positrons are being produced in the region around the massive Galactic black hole (see Section 2.4.2.3); however, this contribution can also be interpreted as extended emission from the nuclear bulge (Crocker et al., 2017).

Spectral analysis of the 511 keV line emission and the o-Ps continuum can reveal information about the physical conditions of the interstellar medium (ISM) in which the positrons annihilate. The two early examples of this analysis are the independent spectral studies by Jean et al. (2006) and Churazov et al. (2005) using one year of public SPI data (see Figure 2.8). The spectrum measured by SPI is extracted by model fitting of the sky distribution (using the Knödlseder et al. (2005) model available at the time, shown in Figure 2.5) and then fit with physical spectral models of annihilation in different phases of the ISM (see Section 2.5.6). Jean et al. show that the emission is best fit with narrow and broad Gaussian components, both centered at 511 keV, in addition to the o-Ps continuum with $f_{Ps} = 0.97 \pm 0.02$. When fit with a physical model of the ISM, the line widths and intensities of these components match best with positron annihilation in 50% warm ionized and 50% warm neutral



(a) Spectral fit from Jean et al. (2006).

(b) Spectral fit from Churazov et al. (2005)

Figure 2.8: Analysis of the INTEGRAL/SPI positron annihilation spectrum after 1 year of public data. (a) Jean et al. (2006) find the 511 keV line and o-Ps continuum are best fit with models of positron annihilation in 50% of warm ionized and 50% warm neutral phases of the ISM. (b) A complementary analysis by Churazov et al. (2005) show the data is best fit with positron annihilation in an ISM with ionization of 0.1 and temperature of 8000 K. Figure 5 from Jean et al. (2006) and Figure 10 from Churazov et al. (2005).

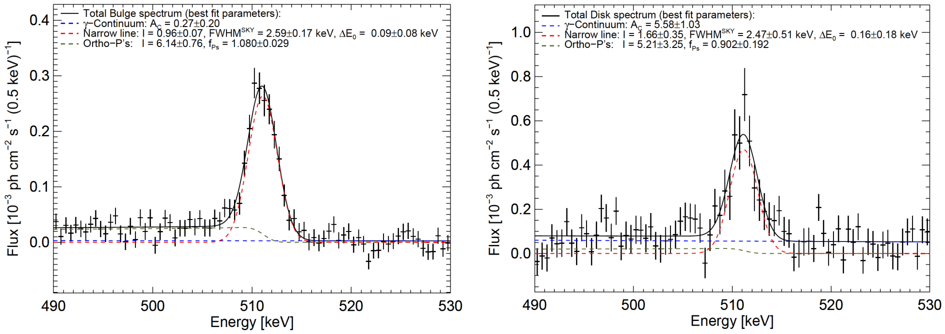
phases of the ISM. Churazov et al. performed a complimentary analysis by leaving the temperature and ionization fraction as the fit parameters. The authors find the positron annihilation spectrum is best fit with an ISM temperature of 8000 K with a degree of ionization of ~0.1, in agreement with Jean et al. (2006). The conclusion from these studies is that positrons annihilate in warm phases of the ISM, which implies that they may propagate away from the hotter sites of production, i.e. star forming regions. With a high Ps fraction of ~95%, this is another confirmation that positrons must have low energies (~10 eV) before annihilation.

A more detailed study of the SPI positron annihilation spectrum was performed by Siegert et al. (2016a). Using 11 years of SPI data and an enhanced approach to background modeling, Siegert et al. refined the spatial model of the annihilation line. With this larger amount of data, the 511 keV line emission from the inner Galaxy is detected at 58 σ, and the o-Ps continuum is detected at 29σ. For the first time, the authors were able to spatially discriminate between the spectra of the Galactic bulge and disk regions (see Figure 2.9). One may expect different spectral signatures from positrons annihilating in the Galactic disk and bulge due to the different gas phases which dominate these regions (Guessoum et al., 2005); however, there is no statistically significant difference observed. Although the main purpose of this study was the spectral analysis, the most significant detail may be the spatial modeling. Siegert et al. find the best fit to the SPI data is very similar to the Skinner Model; see Figure 2.10 and Table 2.2 for the details of this model. With the longer exposure of the GC, Siegert et al. (2016a) report a 5σ detection of the point-source component consistent with Sgr A*. The modeled Galactic disk emission in the Siegert Model is much thicker than the Skinner Model and the B/D ratio of the 511 keV line flux is reported to be 0.58±0.13, a ratio that is much easier to explain with known astrophysical sources. The positron production rate is concluded to be $2{\times}10^{43}$ e^+/s and $3{\times}10^{43}$ e^+/s for the Galactic bulge and disk, respectively. The estimate of the positron production rate in the disk from this study is more than previous estimates of the rate in the entire Galaxy.

Even with SPI observing the skies for the past 16 years (and counting), there remain many open questions about Galactic positrons. The imaging, which was expected to reveal the source, has been far from conclusive. With a coded-aperture mask instrument, a model-fitting approach is taken since there is no way to directly image a source; while straightforward for point source emission, this becomes much more difficult for extended sources. The spatial models of Skinner et al. (2014) and Siegert et al. (2016a) describing the Galactic 511 keV emission show a broad bulge emission. It is not yet known if this distribution is truly diffuse or composed of a population of point sources within the bulge. Furthermore, the extent of the disk emission is still poorly constrained. Through the coded-aperture mask model fitting routines, each of the spatial components are correlated (Siegert et al., 2016a) and the chosen model for each component influences the others; a direct imaging telescope would be needed to measure the true emission to fully determine the latitudinal extent of the disk. Further areas of advancement are discussed in Section 2.7.

2.3.2 MEV CONTINUUM CONSTRAINTS

A very constraining measurement on positron production models is from observations of the MeV continuum emission from the inner Galaxy. If positrons are injected into the ISM at relativistic energies, one would expect a fraction to directly annihilate with electrons in-flight while undergoing the energy loss necessary to result in the 511 keV and o-Ps continuum emission that is observed; therefore, some spectral signature of these energetic positrons would



(a) Galactic bulge emission.

(b) Galactic disk emission.

Figure 2.9: Extracted through model fitting, separate positron annihilation spectra are obtained for the Galactic bulge emission (a) and the Galactic disk emission (b). The different ISM gas phases which dominate these two regions should give different spectral signatures; however, no statistical difference is seen here. Figure 4 from Siegert et al. (2016a).

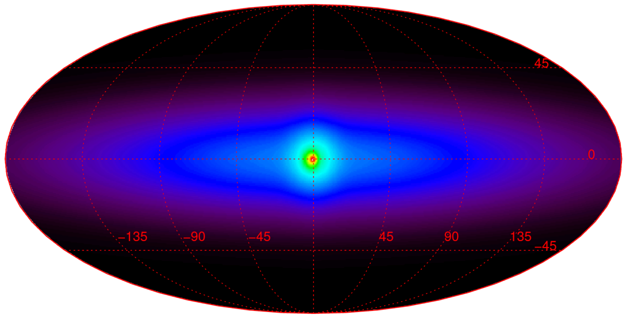


Figure 2.10: The Siegert Model positron annihilation phenomenological model has a much thicker Galactic disk contribution compared to the Skinner Model. Figure 2 from Siegert et al. (2016a). The plot is on a logarithmic scale to show the low surface brightness of the disk emission.

Model Component	Position (l, b) [deg]	FWHM (σ_l, σ_b) [deg]	Flux [$\times 10^{-4}\ \gamma/\mathrm{cm}^2/\mathrm{s}$]
Disk	(0, 0)	(141.29, 24.73)	16.6
Broad bulge	(0, 0)	(20.55, 20.55)	6.9
Narrow bulge	(-1.25, -0.25)	(5.75, 5.75)	2.7
Central point source	(-0.06, -0.05)	(0, 0)	0.8

Table 2.2: Siegert Model components representing the Galactic 511 keV emission (Siegert et al., 2016a).

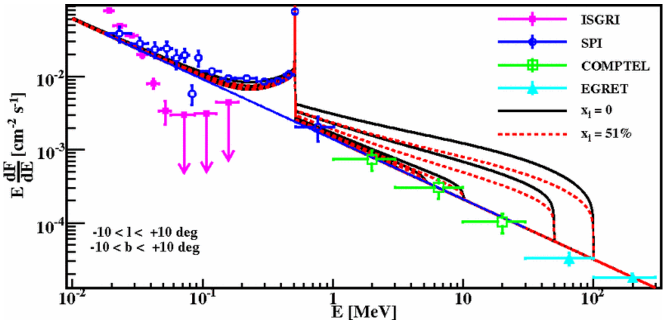


Figure 2.11: The measured flux from a 20° region around the Galactic center. The 511 keV line is clearly visible in the SPI spectrum, and at higher energies a power-law background is consistent with both SPI, COMPTEL, and EGRET data. Overlaid on the measured flux is the theoretical annihilation spectrum in neutral (solid black curves) and 51% ionized (dashed red curves) gas phases assuming initial positron kinetic energies of 1, 3, 5, 10, 50 and 100 MeV (top curve). These measurements place a stringent constraint on the initial positron energy to be $\lesssim$3 MeV. Figure 7 from Sizun et al. (2006).

be detected above 511 keV. Beacom & Yüksel (2006) and Sizun et al. (2006) performed independent studies of the Galactic γ-ray continuum emission using Imaging Compton Telescope (COMPTEL), SPI, and the Energetic Gamma Ray Experiment Telescope (EGRET) data, and developed models of in-flight annihilation for a number of positron initial kinetic energies (Figure 2.11). These studies place a stringent upper limit on the initial kinetic energy of the injected Galactic positrons to be $\lesssim$ 3 MeV since there has been no detected high-energy emission. A limitation on the initial kinetic energy of the positrons injected into the ISM is in turn a strong constraint on the models for Galactic positron production. For example, positrons produced from proton-proton collisions in cosmic-ray interactions would result in a kinetic energy of ~ 30 MeV (Murphy et al., 1987), which would exclude them as a major source of Galactic positrons. This constraint will be discussed in further detail when the possible positron sources are reviewed in Section 2.4.

2.3.3 MAPPING OF ^{26}AL EMISSION

The next Galactic γ-ray line to be detected after the 511 keV emission was the 1.8 MeV line from the radioactive decay of unstable aluminum-26. Measurements of the ^{26}Al characteristic γ-ray line was the first confirmation of

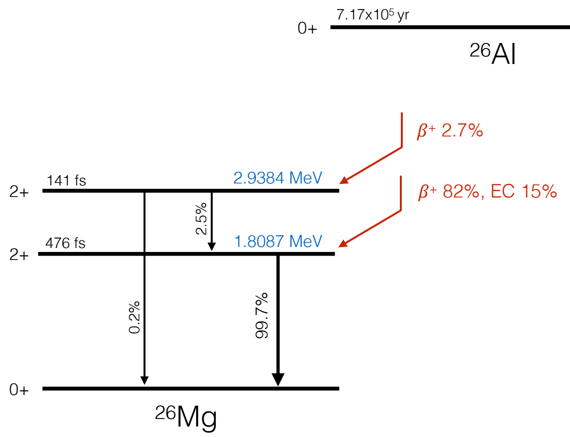


Figure 2.12: ^{26}Al decay scheme. The branching ratio for the 1.8 MeV γ-ray is 0.997, and 85% of decays result in positron production.

active nucleosynthesis in our Galaxy (Clayton, 1984). ^{26}Al has a lifetime of 7.17×10^5 years and in the Galaxy we expect its emission to be diffuse. The main mode of decay for ^{26}Al is the emission of a positron through β^+ decay (see the ^{26}Al decay scheme in Figure 2.12). The 1.8 MeV line was first detected from the direction of the Galactic center by the HEAO-3 satellite at a flux of 4×10^{-4} γ/cm^2/s, which corresponds to ~2 M$_\odot$ of ^{26}Al in the Galaxy (Mahoney et al., 1984). The measured Galactic content of ^{26}Al was larger than stellar models predicted (Ramaty & Lingenfelter, 1977) and corresponds to a positron production rate of ~4×10^{42} e$^+$/s in the Galactic disk. This is the only confirmed source of Galactic positrons to date, though the contributing fraction of ^{26}Al to the total Galactic positron population is still unknown.

^{26}Al is produced in large quantities in stellar nucleosynthesis by proton capture on ^{25}Mg (Hoffman et al., 1995), and therefore is produced in regions rich with protons or magnesium. ^{26}Al can be injected into the ISM in four ways: 1) through stellar winds of massive stars, called Wolf-Rayet (WR) stars, with $\gtrsim$30 M$_\odot$, 2) through core-collapse supernova (CCSN) of less massive stars ~10–30 M$_\odot$, 3) from stars of intermediate or low mass $\lesssim$9 M$_\odot$ during the asymptotic giant branch (AGB) evolution stage, and 4) from O-Ne-Mg rich novae. The importance of each stellar source to the overall Galactic abundance of ^{26}Al is still not known, though measurements of the 1.8 MeV line have excluded novae and AGB stars as major contributors. In massive stars, ^{26}Al is produced through hydrostatic H-burning through the NeNa-MgAl

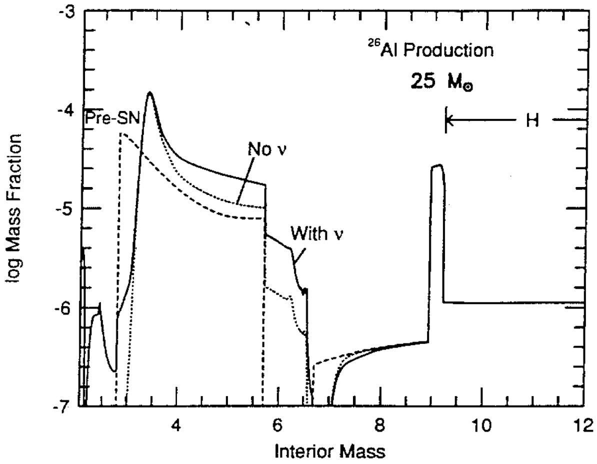


Figure 2.13: The predicted ^{26}Al abundance profile inside a 25 M$_\odot$ star, in terms of the mass fraction. The dashed line is the pre-supernova production from stellar nucleosynthesis, the dotted line is the production from explosive nucleosynthesis, and the solid curve is the production from explosive nucleosynthesis with an enhancement from neutrino processes. This shows that ^{26}Al is produced in two regions: the H-burning shell (from 7–12 M$_\odot$), where ^{26}Al can then be expelled through stellar winds, and the C-Ne-O shell (from 3–7 M$_\odot$), where ^{26}Al can be ejected through the supernova explosion. Figure 1 from Hoffman et al. (1995).

sequence (Cox & Giuli, 1968) and is then ejected into the ISM by the radiation-driven mass-loss which forms a WR star (Abbott & Conti, 1987). In CCSN, ^{26}Al is produced in the O-Ne cycle during the explosion (Woosley & Weaver, 1980). Figure 2.13 shows the predicted production of ^{26}Al in a 25 M$_\odot$ star from Plüschke et al. (2001), where the ^{26}Al is created in significant amounts in the inner and outer regions within a massive star. Another way to help distinguish the major source of ^{26}Al would be to compare its γ-ray line flux to the flux produced by another isotope, ^{60}Fe, which is also produced in CCSN but not in stellar winds (Hoffman et al., 1995). However, these measurements are not yet precise enough for a clear conclusion (Smith, 2004; Wang et al., 2007).

The 1.8 MeV emission was mapped by CGRO/COMPTEL after its launch in 1991 and provided the first all-sky map of a single γ-ray line. Figure 2.14 shows an analysis of the spatial distribution of the ^{26}Al emission after 9 years of data using the Maximum Entropy Method (MEM) imaging algorithm (Plüschke et al., 2001). The emission is predominantly along the Galactic plane and

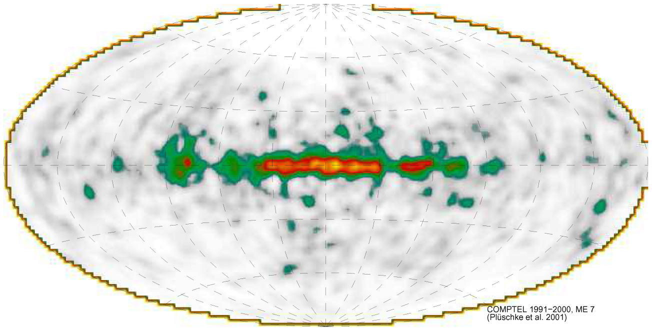


Figure 2.14: All sky map of the ^{26}Al 1.809 MeV γ-ray line measured by COMPTEL with 9 years of data, confirming that ^{26}Al traces the regions of massive stars. The image is obtained through the MEM algorithm. Figure 2 from Plüschke et al. (2001)

shows a few prominent hot spots of stronger intensity, which are approximately in the direction tangent to the spiral arms of the Galaxy. From this emission map, it can be determined that massive stars, either CCSN or WR stars, are the origins of ^{26}Al in the Galaxy. Since we know that 82% of the ^{26}Al decays that produce a 1.8 MeV γ-ray also produce a positron (see Figure 2.12), and since we also expect the positron source distribution to match the annihilation distribution (see Section 2.6), it is puzzling that the 1.8 MeV map and the 511 keV map are so different. Positrons from the β^+ decay of ^{26}Al and other unstable isotopes are further discussed in Section 2.4.1.

The characteristic γ-ray of ^{26}Al has also been detected with INTEGRAL/SPI (Diehl et al., 2006). With 10 years of observation, reconstructed images have not revealed more detail than that of COMPTEL (Bouchet et al., 2015) due to the difficulty in imaging extended sources with coded-aperture mask instruments. SPI has performed spectral studies of various hot spots seen in the Galactic emission (Martin et al., 2008; Voss et al., 2012; Siegert & Diehl, 2016) and Doppler shift measurements of the line centroid due to Galactic rotation (Kretschmer et al., 2013); however, none of these results have progressed the story of positron annihilation.

2.4 POSITRON PRODUCTION MECHANISMS AND GALACTIC SOURCES

We have learned from observations of the 511 keV emission from positron annihilation in the Galaxy that:

1. the high f_{Ps} fraction requires positrons to have low energies of $\lesssim$10 eV before annihilation,
2. the line shape analysis indicates positrons annihilate predominantly in warm phases of the ISM,
3. the lack of emission in the MeV continuum above 511 keV constrains the initial kinetic energies of positrons to $\lesssim$3 MeV,
4. imaging analysis shows a strong bulge component ($\sim 2\times10^{43}$ e^+/s) and emission along the Galactic disk (0.8–3×10^{43} e^+/s, depending on the spatial model),
5. and some fraction of the positrons annihilating in the Galactic disk can be explained by β^+ decay of ^{26}Al ($\sim 4\times10^{42}$ e^+/s).

It is not known whether positrons are produced from a truly diffuse source or a population of point sources densely packed in the Galactic bulge region. Additionally, it seems likely that more than one astrophysical source is contributing to the population of positrons in the Galaxy. In this section, we will review the mechanisms of positron production and relevant astrophysical sources. A range of exotic phenomena have been proposed: Q balls (Kasuya & Takahashi, 2005), relic particles (Picciotto & Pospelov, 2005), axions (Hooper & Wang, 2004), primordial black holes (Frampton & Kephart, 2005; Bambi et al., 2008), superconducting cosmic strings (Ferrer & Vachaspati, 2005), sterile neutrinos (Khalil & Seto, 2008), etc., which will not be discussed here.

A major limitation in our knowledge of Galactic positrons is the unknown distance that they travel from their point of production to their point of annihilation. If positrons are born in less dense hot regions of the Galaxy, it is understandable that they could travel until reaching a more dense warm phase of the ISM in which they annihilate. Some authors have stressed that the source distribution must be weighted with the distribution of different gas phases/densities in the ISM when determining the expected annihilation emission (Guessoum et al., 1991; Siegert et al., 2016a). To further understand this effect, simulations of positron propagation have been performed and will be further discussed in Section 2.6.

2.4.1 β^+ DECAY OF RADIOACTIVE NUCLEI

The most familiar production mechanism of positrons is through radioactive decay in which a proton is converted to a neutron, positron, and electron neutrino:

$$^{A}_{Z}\mathrm{X}_{\mathrm{N}} \rightarrow {}^{\;\;A}_{Z-1}\mathrm{Y}_{\mathrm{N+1}} + \mathrm{e}^{+} + \nu_e. \tag{2.4}$$

β^+ decay is mediated by the relatively slow weak interaction and can occur in unstable "proton-rich" nuclei. With the conversion of one of the parent isotope's Z protons into a neutron in the daughter isotope, the mass number $A = Z + N$ remains constant. The daughter isotope has one fewer proton than the parent isotope, and therefore, in its atomic form it must also shed one electron to return to a neutral atom. The decay is energetically possible if the mass difference between the parent and daughter nucleus is at least $2m_ec^2$, where m_e is the mass of the electron or positron (one m_e for the positron produced and one m_e for the electron lost). The energy release in the form of kinetic energy, or Q-value, of β^+ decay is

$$Q = (m_i - m_f)c^2 = \left[m({}^{A}_{Z}\mathrm{X}) - m({}^{\;\;A}_{Z-1}\mathrm{Y}) - 2m_e\right]c^2, \tag{2.5}$$

ignoring the small mass of ν_e. For reactions where the difference in mass is $< 2m_ec^2$, electron capture (EC), sometimes referred to as inverse β decay, is the only decay mode available. In EC the nucleus absorbs an inner-shell electron to mediate the conversion of a proton to a neutron:

$$^{A}_{Z}\mathrm{X}_{\mathrm{N}} + \mathrm{e}^{-} \rightarrow {}^{\;\;A}_{Z-1}\mathrm{Y}_{\mathrm{N+1}} + \nu_e. \tag{2.6}$$

In the three-body process of β^+ decay, the energy released is divided between the positron, the neutrino, and the recoiling daughter nucleus. This results in a distribution of energies for the positron where the maximum possible is the Q-value, usually between 0.5-3 MeV, which satisfies the MeV constraints on the initial energy of Galactic positrons (Section 2.3.2). The daughter nucleus is often produced in an excited state, and thus will undergo a subsequent decay via γ-ray emission. The life-time of an unstable nucleus is related to the Q-value of the reaction, and for β-decay this can be millions of years.

Decay chain	Half life	γ-ray [keV] (BR)	Production site
${}^{26}_{13}\text{Al} \rightarrow {}^{26}_{12}\text{Mg}$	7.4×10^{5} yr	1809(1)	WR stars, CCSN
${}^{44}_{22}\text{Ti} \rightarrow {}^{44}_{21}\text{Sc} \rightarrow {}^{44}_{20}\text{Ca}$	59 yr	68(0.94), 78(0.96), 1157(1)	CCSN
${}^{22}_{11}\text{Na} \rightarrow {}^{22}_{10}\text{Ne}$	2.6 yr	1275(1)	Novae
${}^{56}_{28}\text{Ni} \rightarrow {}^{56}_{27}\text{Co} \rightarrow {}^{56}_{26}\text{Fe}$	77.2 d	847(1), 1238(0.68), 1771(0.15), 2598(0.17)	SNIa

Table 2.3: Nucleosynthesis products that decay via β^{+} decay with half-lives that are relevant to Galactic positron production. When more than one transition is listed in the decay chain, the longest half life is listed. The major γ-ray emission lines are listed with the branching ratio (BR) included in brackets.

Proton-rich environments that are conducive to producing β^{+}-unstable nuclei are found in hydrostatic and explosive nucleosynthesis, i.e., in the cores of massive stars and supernova (SN). At the end of hydrostatic burning, ^{28}Si ($Z = N = 14$) is the most abundant isotope in massive stars, whereas ^{12}C ($Z = N = 6$) and ^{16}O ($Z = N = 8$) are the most abundant isotopes in the cores of the white dwarf (WD) progenitors to Type Ia supernova (SNIa). After hydrostatic burning, an explosion, either through a SNIa or CCSN, ends the life of the star. A dominance of $Z = N$ isotopes is maintained throughout explosive nucleosynthesis and large quantities of ^{56}Ni ($Z = N = 28$) are produced. Heavy stable isotopes tend to have $N > Z$, therefore the heavy $Z = N$ isotopes created in stellar nucleosynthesis are proton-rich and susceptible to β^{+} decay. Table 2.3 lists relevant isotopes that are created in abundance in hydrostatic and explosive nucleosynthesis, which are all $Z = N$ nuclei, that decay via β^{+} emission (and have a long enough half-life to escape their sites of production). The characteristic γ-rays that are emitted from the excited nuclei after β^{+} decay are also listed. These isotopes are potential sources of Galactic positrons and their yields can be detected via their γ-ray emission.

There remain many unknowns when it comes to nucleosynthesis positron production: the yields of isotopes in hydrostatic and explosive nucleosynthesis, the SN rate and frequency of other nucleosynthesis sites, and the fraction of the isotopes or positrons that escape from a collapsing star. The next

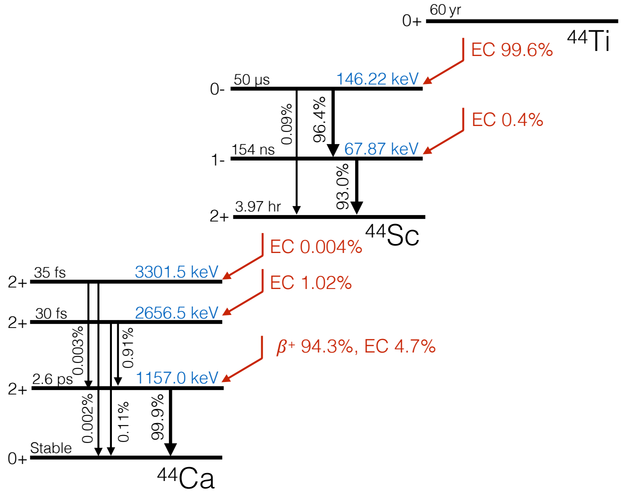


Figure 2.15: ^{44}Ti and ^{44}Sc decay scheme (Bé et al., 2016, modified from). ^{44}Ti decays via EC to ^{44}Sc, which then predominantly decays via β^+ decay to stable ^{44}Ca.

few sections overview the main β^+ decay isotopes that may contribute to the Galactic positron population.

2.4.1.1 ^{26}AL AND ^{44}TI FROM MASSIVE STARS

As discussed in Section 2.3.3, the characteristic γ-ray line from ^{26}Al decay has been studied in detail by COMPTEL and SPI. From these observations, it has been found that 2.8±0.8 $M_\odot$ of ^{26}Al is currently in the Galaxy (Diehl et al., 2006), which corresponds to a positron production of ~4×10^{42} e^+/s in the Galactic disk, a considerable fraction of the measured annihilation rate. Stellar models predict a yield of ~10^{-4} $M_\odot$ of ^{26}Al produced in each massive star, and with a CCSN rate of 2/100 yrs, this corresponds to a production/decay rate (assuming steady state) of 2.7 $M_\odot$/Myr, consistent with the flux observations (Diehl et al., 2006). ^{26}Al in the Galaxy is fairly well understood and is accepted in the scientific community as one of the major sources of Galactic positrons.

There remains much uncertainty about the yields of ^{44}Ti and therefore its contribution to Galactic positrons. ^{44}Ti is produced within the inner regions of a CCSN near the dividing line that separates the material that is ejected from that which collapses back on the compact object (Thielemann et al., 1996). The unknowns of the SN explosion details, such as the the explosion energy and possible asymmetries, vary the predicted ^{44}Ti yields significantly (Woosley & Weaver, 1995; Magkotsios et al., 2010).

With a half life of 60 years (Wietfeldt et al., 1999), ^{44}Ti can be observed from a young supernova remnant (SNR). The characteristic X-rays and γ-rays from the ^{44}Ti decay chain, namely 68 keV and 78 keV from ^{44}Ti, and 1.157 MeV from ^{44}Sc (Figure 2.15), have been detected. The first observation of ^{44}Ti in the Galaxy was through the ^{44}Sc characteristic 1.157 MeV γ-ray measured by COMPTEL from Cassiopeia A (Cas A) SNR (Iyudin et al., 1994). More recently, the ^{44}Ti 67.9 keV and 78.4 keV lines were detected from Cas A with the Imager on Board the INTEGRAL Satellite (IBIS) (Renaud et al., 2006) and the Nuclear Spectroscopic Telescope Array (NuSTAR) (Grefenstette et al., 2014). The CCSN explosion creating Cas A is dated to 1681 (Minkowski, 1959), making Cas A one of the youngest and closest SNR. It is particularly famous for being aspherical and its jet-like features that have been imaged in X-ray (Hwang et al., 2004) and optical wavelengths (Fesen et al., 2006). Interestingly, the ^{44}Ti γ-ray flux measurements result in a large ^{44}Ti/^{56}Ni ratio, factors greater than predicted by spherically symmetric explosions. Nagataki et al. (1998) show that an axisymmetric explosion model can explain the large ratio of ^{44}Ti/^{56}Ni as detected in the γ-ray fluxes from Cas A, further confirming the asymmetric nature of the explosion.

The only other source for which ^{44}Ti X-ray and γ-rays have been detected is SN1987A. At 51.4 kpc it is the closest CCSN in recent history. As with Cas A, SN1987A also shows an asymmetric explosion and large ^{44}Ti/^{56}Ni ratio (Boggs et al., 2015). Both SNRs have a derived ^{44}Ti yield of ~2×10^{-4} $M_\odot$. Efforts have been made to detect ^{44}Ti from other young SNRs and it is significant that none have been found despite the expected rate of 2–3 CCSN per century (The et al., 2006). ^{44}Ti is also theorized to be produced by a subset of SNIa, specifically sub-Chandrasekhar mass models (Woosley & Weaver, 1994); however, the occurrence rate of these particular sources is totally unknown. The only confirmed sources of ^{44}Ti in the Galaxy are asymmetric CCSN which produce higher than expected yields, but these sources seem rare compared to standard CCSN.

Prantzos et al. (2011) estimates the production rate of positrons from ^{44}Ti by considering: 1) the solar ^{44}Ca/^{56}Fe ratio, which should be equivalent to the parent nuclei ratio ^{44}Ti/^{56}Ni, and 2) the typical yields of ^{56}Fe in SNIa and CCSN. The authors project a rate of ~3×10^{42} e$^+$/s from ^{44}Ti, which is comparable to the rate expected from ^{26}Al. The measured positron annihilation rate from the Galactic disk is 3×10^{43} e$^+$/s in Siegert's Model (Siegert et al., 2016a) and 0.8×10^{43} e^+/s from earlier spatial models (Bouchet et al., 2010); therefore, the positrons produced in the β^+ decay of ^{26}Al and ^{44}Ti could explain a large fraction, if not all, of the disk emission. However, it is important to stress that the ^{44}Ti yields in SN are still not well understood and improved modeling and more significant detections are needed to constrain these values.

2.4.1.2 ^{56}NI FROM TYPE IA SUPERNOVA

Nickel-56 is created in vast quantities through silicon burning in SN. It has a half-life of only 6.1 days and decays by electron capture into unstable cobalt-56. ^{56}Co decays with a half-life of 77.2 days into ^{56}Fe through β^+-emission with a branching ratio of 19%. Explosion models predict on order of 1 $M_\odot$ of ^{56}Ni is synthesized in SNIa and 0.1 $M_\odot$ in CCSN, and with a SN rate of 0.5 per century for SNIa and 2 per century for CCSN (Mannucci et al., 2005), the ^{56}Ni decay chain alone would account for an overabundance of positron production. However, at a short half life of 77 days, there is a question about what fraction of the positrons produced from the decay chain of ^{56}Ni can escape the SN explosion region to annihilate in the ISM. There still are many unknowns with regards to SNIa explosion models (Hillebrandt & Niemeyer, 2000), and the uncertain distribution of isotopes within the SN can lead to significantly different yields of ^{56}Ni. Additionally, the magnetic field, the ionization fraction, the temperature and the density of the SN explosion, all which are poorly constrained or entirely unknown, can affect the positron escape fraction (Chan & Lingenfelter, 1993).

SNIa optical light-curves are powered by the decay of ^{56}Ni, and a measured difference in peak brightness of an explosion is often attributed to a difference of ^{56}Ni produced. The inferred mass of ^{56}Ni for different SNIa range from ~0.07 $M_\odot$ to 0.92 $M_\odot$ (Mazzali et al., 1997; Khoklov et al., 1993). The hot, dense ejecta is opaque to γ-rays until ~100 days after the explosion, and after this time, the luminosity results from the deposited energy from positrons. Chan & Lingenfelter (1993) calculated the escape fraction of positrons from SNIa for a range of models and found fractions of 0.1–27%; however, with the knowledge of the observed Galactic annihilation rate, they placed a bound on the escape fraction of 0.5±0.25% to 2±1%. Comparing models to observations of SNIa light curves, Milne et al. (1999) found that ~3% of positrons from ^{56}Ni escape the ejecta, which gives an average yield of ~8×10^{52} e^+ per SNIa. Assuming a SNIa rate of 0.5 per century, this corresponds to a positron production rate of 1.3×10^{43} e^+/s, which is a good fraction of the observed

Galactic rate. However, it is important to emphasize that the positron escape fraction has not been directly measured and there is much uncertainty in the e^+ yield.

The recent and close SN2014J provided the first clear detection of γ-ray lines associated with the ^{56}Ni decay chain from a SNIa, reported by Churazov et al. (2014). From the measured line fluxes, the authors determine 0.6±0.1 $M_\odot$ of ^{56}Ni was synthesized in the explosion. Direct searches for 511 keV emission from SN, which would allow for a measure of the e^+ escape fraction, have also been attempted. Kalemci et al. (2006) were only able to set an upper limit on the 511 keV flux from SN1006 of 0.59×10^{-4} γ/cm^2/s for a 1 Ms observation. This limit corresponds to a escape fraction of 7.5%, which, the authors conclude, rules out the possibility that ^{56}Ni from SNIa are the only source of Galactic positrons.

As SNIa models are not yet complete, and especially since the positron escape fraction is not well constrained, it is not known if positrons from ^{56}Ni decay could actually contribute 1.3×10^{43} e^+/s, as suggested in Milne et al. (1999) or if the number is substantially lower, as some studies imply. The old stellar progenitors of SNIa are more abundant in the inner Galaxy than the younger stars responsible for the production of ^{26}Al and ^{44}Ti. Therefore, it is possible that ^{56}Ni from SNIa could contribute to a higher fraction of the positrons in the Galactic bulge.

2.4.1.3 ^{22}NA FROM NOVAE

Novae are non-destructive explosions that occur when accreted hydrogen from a companion star explosively burns on the surface of a WD. The different classes of WDs, namely carbon oxygen (CO) and oxygen neon (ONe), will produce different novae signatures since the accreted hydrogen will mix with the WD core material (Hernanz, 2005). Explosive hydrogen burning synthesizes some β^+-unstable nuclei, namely ^{13}N and ^{18}F from the CNO cycle in all WDs, and ^{22}Na from ONe WDs. The half-life of ^{13}N and ^{18}F are 110 min and 10 min, respectively, and therefore the positron escape fraction from these isotopes is expected to be negligible (Prantzos et al., 2011). ^{22}Na has a half-life of 2.6 yr and calculations suggest that 10^{-9} to 10^{-8} $M_\odot$ of ^{22}Na could be produced in ONe novae (Hernanz & José, 2006; Gómez-Gomar et al., 1998), which corresponds to $\sim10^{48}$ e^+ per nova. With a Galactic nova rate of ~30 per year (Shafter, 2002), one-third being ONe nova (Gil-Pons et al., 2003), the positron production rate would be $\sim1.5\times10^{41}$ e^+/s, which is $\lesssim$1% of the measured rate. Therefore, novae are not expected to be major contributors to the population of Galactic positrons.

2.4.2 PAIR PRODUCTION

Through the symmetric laws of physics, an electron and positron can annihilate into two photons, and likewise two photons can convert to an electron and positron. The latter process is called pair production. The inverse of electron-positron direct annihilation is pair production through $\gamma - \gamma$ interactions (Breit & Wheeler, 1934; Gould & Schréder, 1967)

$$\gamma + \gamma \rightarrow \mathrm{e}^- + \mathrm{e}^+. \tag{2.7}$$

From conservation of energy and momentum, this occurs if the energies of the two photons E_{γ_1} and E_{γ_2} satisfy

$$E_{\gamma_1} E_{\gamma_2} > \frac{2m_e^2c^4}{(1-\cos\theta)}, \tag{2.8}$$

where θ is the angle between the two photon directions. The minimum energy required is for a head-on collision, $\theta = \pi$, and the threshold is $E_{\gamma_1} E_{\gamma_2} > m_e^2c^4$. The interaction cross section has a maximum for a photon energy of $E_\gamma \sim 700$ keV, which leaves an energy of (1.4 MeV - 1.022 MeV) ≈ 0.4 MeV to be distributed to the e^- and e^+ (Patrignani & Particle Data Group, 2016). $\gamma - \gamma$ pair production is only efficient at very high photon densities, for example in accretion disks.

As a photon is the force mediator for electromagnetic force, pair production can occur with a single photon in a strong enough electric or magnetic field, provided that the interacting photon satisfies the energy threshold (Motz et al., 1969). A photon interacting with the Coulomb field around a nucleus is the dominant pair-production process of light-matter interactions for MeV γ-rays:

$$\gamma + {}^{A}_{Z}\mathrm{X}_\mathrm{N} \rightarrow {}^{A}_{Z}\mathrm{X}_\mathrm{N} + \mathrm{e}^- + \mathrm{e}^+. \tag{2.9}$$

This interaction has a threshold of $E_\gamma > 2m_ec^2 = 1.022$ MeV from conservation of energy, taking the recoil of the nucleus to be neglible. Pair creation in the Coulomb field of a free or bound electron is called triplet production (Perrin, 1933)

$$\gamma + \mathrm{e}^- \rightarrow \mathrm{e}^- + \mathrm{e}^- + \mathrm{e}^+, \tag{2.10}$$

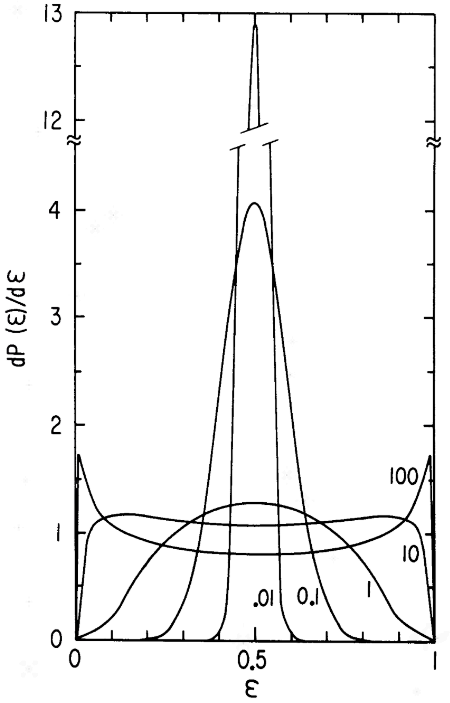


Figure 2.16: The energy distribution of the positron (or electron), where ε is the fraction of total energy contained in one particle, from B pair production for different levels of B field and γ energy, parameterized by $\chi = \frac{\hbar\omega}{2mc^2}\frac{B}{B_{cr}}$ where $B_{cr} = \frac{m^2c^3}{e\hbar} = 4.414{\times}10^{13}G$. Figure 7 from Daugherty & Harding (1983).

and the energy threshold becomes $E_\gamma > 4m_ec^2$. Pair production in a strong magnetic field

$$\gamma + B \rightarrow \mathrm{e}^- + \mathrm{e}^+, \tag{2.11}$$

has a threshold of $E_\gamma > 2m_ec^2/\sin\theta_{kB}$, where θ_{kB} is the angle between the photon direction and the magnetic field vector. For γ's near this threshold energy, the interaction does not occur until $B \gtrsim 10^{12}$ G (Daugherty & Harding, 1983).

The energies of the positrons created in pair-production depend on the initial γ energy. In $\gamma - \gamma$ interactions and pair production in the Coulomb field around a nucleus and free electron, the available kinetic energy, which is just $(E_\gamma - 2m_ec^2)$, is distributed on average equally among the generated particles. Interestingly, for γ interactions in a strong B field, it is more likely that either the electron or positron is imparted with the majority of available energy, as opposed to an equal distribution. Figure 2.16 shows this energy distribution of the e^--e^+ pairs as a function of the B field and E_γ.

For pair production in astrophysical sources, one needs high-energy photons, which are known to occur around luminous compact objects, such as black holes (BHs), microquasars, and active galactic nuclei (AGN), or strong magnetic fields, which can be found in neutron stars, pulsars, and magnetars.

2.4.2.1 X-RAY BINARIES AND MICROQUASARS

An X-ray binary (XRB) is a system with a compact object, either a black hole or neutron star (NS), accreting mass from a companion star (Verbunt, 1993). The more massive component of the binary is called the primary and the companion star is called the secondary. There is a further sub-classification from optical observations that depend on the mass of the companion star: high-mass X-ray binaries (HMXBs) involve massive O or B stars, while low-mass X-ray binaries (LMXBs) generally have $\lesssim \mathrm{M}_\odot$ and are main sequence stars, red giants, or WDs. The X-ray emission results from the in-fall of the accreting matter from Roche lobe overflow. For LMXBs, the spectra are fairly soft ($kT \lesssim$5 keV) and they are some of the brightest X-ray objects in the Galaxy, as they emit almost all of their energy in X-rays (Verbunt, 1993). In HMXBs, the optical emission from the secondary star is comparable to the X-ray emission from the accreting compact object, which is about an order of magnitude less than LMXBs. HMXBs often contain X-ray pulsars and have a harder spectrum ($kT \sim$15 keV) than LMXBs (Verbunt, 1993). Some of these systems will periodically have relativistic jets of matter observable in the radio wavebands that are attributed to synchrotron radiation of electrons in the jets (Mirabel & Rodríguez, 1999). These systems are called microquasars due to their resemblance to quasars, but they are on a much smaller physical scale.

There are approximately 200 LMXBs (Liu et al., 2007), about 100 HMXB (Liu et al., 2006), and only about 20 known microquasars detected in the Galaxy. Grimm et al. (2002) find that LMXBs are clustered in the Galactic bulge region while HMXBs are distributed along the plane of the Galaxy, since they are generally associated with young stellar populations (see Figure 2.17 for an illustration of the Galactic distribution of XRBs). Due to the spatial distribution of HMXBs, they can be excluded as the source of positrons in the Galactic bulge, so we will focus on LMXBs and microquasars as possible positron sources.

XRBs have a long history as the potential source of Galactic positrons. The "Great Annihilator" (1E 1740.7-2942) observed by SIMGA (see Section 2.2.2)

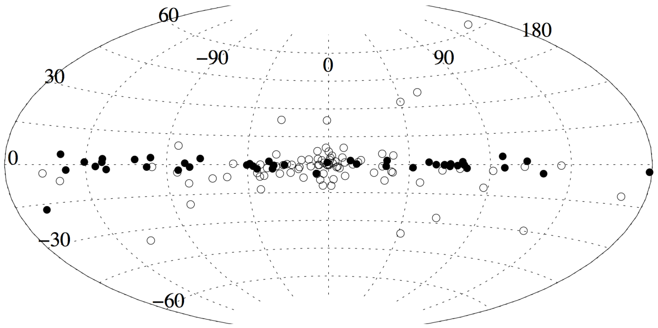


Figure 2.17: The Galactic distribution of XRBs; the open circles represent the known LMXBs and the black circles represent the known HMXBs. Figure 1 from Grimm et al. (2002).

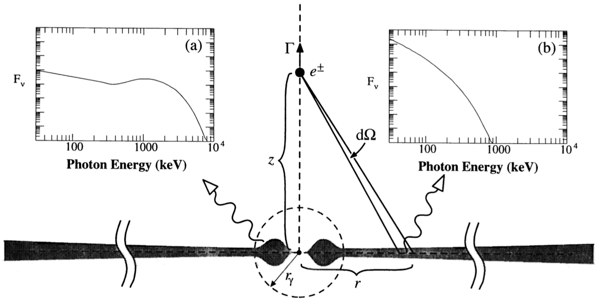


Figure 2.18: Pair-production in the accretion disk, the corona, or at the base of the jet, creates electron-positron pairs that can be funneled into the ISM. Figure 1 from Li & Liang (1996).

was confirmed as the first detected microquasar in 1992 when the Very Large Array (VLA) imaged well-collimated jets that were consistent with the X-ray source (Mirabel et al., 1992). The reported 511 keV emission was later refuted (Smith et al., 1996a). In 2008, the asymmetric map of 511 keV emission produced by Weidenspointner et al. (2008) was explained as resembling the asymmetry observed in hard LMXBs (Figure 2.17). However, this disk asymmetry is now thought to be explained by an offset bulge component (Bouchet et al., 2010).

Positrons can be created through $\gamma-\gamma$ pair-production from the X-rays/γ-rays in the hot inner accretion disk near the compact object, in the X-ray corona, or at the base of the jet (Li & Liang, 1996; Paredes, 2005), and some fraction of positrons produced could then be funneled into the ISM by the relativistic jets, as shown in Figure 2.18. There still are major unknowns about XRB jets. For example, it is not known if the jets are leptonic (e$^-$-e$^+$ pairs) or hadronic (protons and pions); therefore, the possible number of positrons produced and their energies are uncertain. Although many jet acceleration scenarios result in ultra-relativistic positrons, Li & Liang (1996) show that electron positron pairs could be born with a kinetic energy of several hundred keV and accelerated in the jet to only ~MeV. Guessoum et al. (2006) give an estimate of the pair ejection rate into the ISM of 10^{41} to 10^{42} e^+/s per source, and the authors predict rates of 511 keV emission from the brightest LMXBs and microquasars at $\sim 1\times 10^{-4}$ $\gamma/cm^2/s$. These fluxes could be detected as point sources with the next-generation of γ-ray telescopes.

Recently, Siegert et al. (2016c) reported a possible detection of a broadened 511 keV signature from the microquasar V404 Cygni during a period of strong flare activity in 2015 (Siegert et al., 2016c). Although the results are somewhat controversial in the community (Roques & Jourdain, 2016), the authors report a flux that corresponds to a positron production rate of 10^{42} e^+/s, and therefore this could be evidence that microquasars are a significant contributor to the Galactic positron population.

Positron production from pair creation in XRBs might be large enough to significantly contribute to the population of Galactic positrons (if the leptonic jet model is correct), but the morphology of these sources is not concentrated enough in the GC to explain the Galactic bulge emission. This is discussed further in Section 2.6.

2.4.2.2 PULSARS AND MAGNETARS

Pulsars and magnetars are rapidly rotating magnetized NSs formed during a CCSN explosion. Pulsars (Seiradakis & Wielebinski, 2004) have stronger than average magnetic fields (~10^{12} G) with jets of relativistic charged particles ejected along its two poles. The emission, when observed from an angle, results in radio pulsation analogous to the pulses of light seen from a lighthouse. The frequency of pulsations is generally ~Hz, which means these objects are spinning extremely quickly. Pulsars that are born in binary systems generally have higher rotation rates and lower magnetic fields, and are called millisecond pulsars (Lorimer, 2008). Magnetars are rare isolated neutron stars that have stronger magnetic fields than the average pulsar at $B \sim 10^{14}$ G (Harding & Lai, 2006), but display no persistent radio pulsations. Magnetars are responsible for the soft gamma-ray repeater and anomalous X-ray pulsar source classes. There are almost 2000 pulsars (Manchester et al., 2005), about 150 millisecond pulsars (Lorimer, 2008), and about 10 magnetars (Harding & Lai, 2006) that have been detected in our Galaxy.

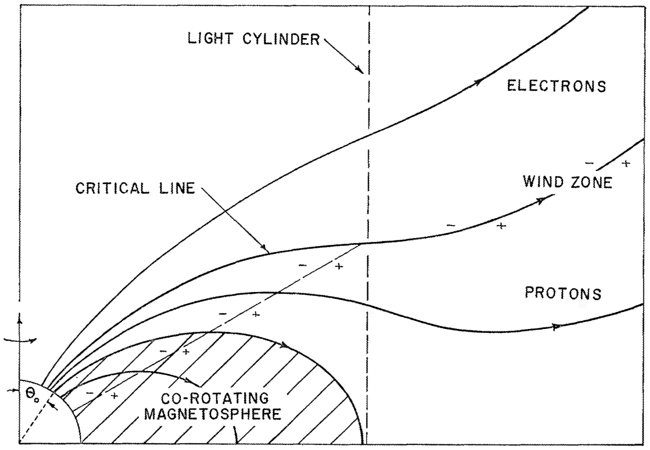


Figure 2.19: Diagram of the pulsar model developed by Goldreich & Julian (1969). The star is in the lower left corner, and the electron positron pairs are produced in pair cascades above the polar cap or in the outer gap region. The *light cylinder* is the radius at which the rotation velocity is equal to the speed of light $R_L = \frac{cT}{2\pi}$, where T is the period of the star, in seconds, and it separates which field lines are opened or closed. Figure 1 from Goldreich & Julian (1969).

The standard model for pulsars includes a dipolar magnetic field which creates a dense magnetosphere co-rotating with the star (Goldreich & Julian, 1969), as illustrated in Figure 2.19. Charged particles, namely electrons and protons, flowing along the open field lines will emit curvature radiation in γ-rays (Sturrock, 1971). Positrons can then be produced in γ-γ collisions and through γ-B pair production in the strong magnetic field of the NS. Pulsars are perfect environments for pair production, either in regions above the polar caps, which serve as the origin for the open field lines (Daugherty & Harding, 1982), or in the so called *outer gap*, which is near the light cylinder (Cheng et al., 1986). Electron-positron pairs created through γ-B interactions would be predominantly produced in directions perpendicular to the field lines, and are therefore in a position to radiate via synchrotron emission; the emitted synchrotron γ-rays then could produce another generation of pairs (Daugherty & Harding, 1982). The pair creation cascade will continue as long as the photons that are produced are above the pair-creation threshold, see Section 2.4.2. It is thought that the electron-positron pairs are further accelerated up to relativistic energies at the shock front of the pulsar wind nebula (Chi et al., 1996; de Jager & Djannati-Ataï, 2009).

The positron production rate from pulsars and magnetars can be estimated by

$$\dot{n}_{e^+} \approx 2.8 \times 10^{37} B_{12}^{10/7} P^{-8/21} \ \mathrm{e^+/s}, \tag{2.12}$$

where B_{12} is the dipolar magnetic field strength in units of 10^{12} G, and P is the rotation period of the pulsar in seconds (Zhang & Cheng, 1997; Prantzos et al., 2011). For pulsars and millisecond pulsars the production rate is $\sim 4 \times 10^{37}$ $\mathrm{e^+/s}$, and for magnetars this can be up to 4×10^{40} $\mathrm{e^+/s}$. However, pair production in magnetars is not yet well understood. Magnetars do not exhibit radio pulsations, which have been linked to the pair cascades of pulsars, and therefore it is believed that pair-creation is suppressed when the magnetic fields become as large as they are in magnetars (Harding & Lai, 2006). Additionally, the fraction of positrons that escape from the pulsar into the ISM is not well known, but is thought to be close to one (Wang et al., 2006). Most pulsar models predict that positrons escape from the pulsar with relativistic energies $\gtrsim$30 MeV (de Jager & Djannati-Ataï, 2009), which eliminates these sources as the major contribution to Galactic positrons, see Section 2.3.2.

Pulsars and magnetars are young and are expected to trace star forming regions of the Galaxy, so they would not be considered as significant contributors to the Galactic bulge. The distribution of millisecond pulsars should match that of SNIa, which allows for a potentially significant bulge contribution; however, the fact that positrons produced in pulsars have energies $\gtrsim$ 30 MeV rule them out as a major contributor of Galactic positrons.

2.4.2.3 SAGITTARIUS A*

Positrons can be created through γ-γ interactions in the inner hot accretion disk of a BH, similar to the case of XRBs. The X-ray emissivity of the massive black hole at the center of the Galaxy, Sagittarius A* (Sgr A*), is orders of magnitude weaker in comparison to the emission from the population of Galactic XRBs. It is difficult then to see how Sgr A* could be a major Galactic positron contributor unless there were periods in the past with higher activity. For the Galactic positron annihilation emission, we have been assuming a steady state: a balance between the $\mathrm{e^+}$ production and annihilation rates. The steady-state production assumption can be dismissed if the diffusion time for positrons before annihilation is long. Two scenarios are proposed: 1) accretion of the gas surrounding Sgr A* was 10^4 times higher in the past, but dropped suddenly to its current rate about 300 yr ago (Totani, 2006)

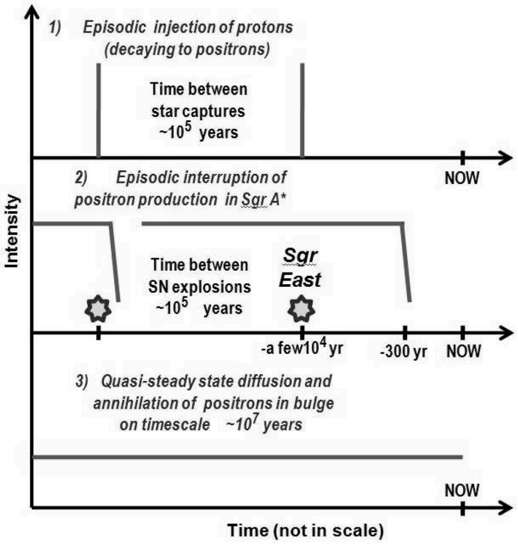


Figure 2.20: Figure 17 from Prantzos et al. (2011) showing the two proposed scenarios for past activity around Sgr A*; see text. Both scenarios require a long diffusion timescale to result in quasi-steady state annihilation emission.

and 2) a tidal disruption event (TDE) could increase the accretion activity for 10-100 yr (Rees, 1988) every 10^4-10^5 years (Cheng et al., 2007). These two scenarios are illustrated in Figure 2.20.

Beloborodov (1999) describes how an e^--e^+ atmosphere, created through γ-γ interactions, is inevitably created around a BH accretion disk, similar to the case of XRBs. Totani (2006) considers a past time with a higher Sgr A* accretion rate, such as described in Scenario 1 above, which may be supported by observations (Su et al., 2010). With an accretion rate of $\dot{M} \sim 10^{-4}$ $M_\odot$/yr, the author shows that a positron production rate of $\sim 10^{43}$ e^+/s can be obtained with $E_{e^+} \sim 1$ MeV, satisfying the high-energy γ-ray continuum constraints. Totani argues the past high accretion rate could have been reduced to its current rate ~300 yr ago by the shell passage of the SNR Sagittarius A East.

Scenario 2 listed above would involve a TDE-induced relativistic jet of protons ejected into the ISM. These protons could collide and positrons would be produced through the decay of the resulting π^+, and this will be discussed in more detail in Section 2.4.3 after the introduction of positron production through meson decay.

A difficulty with the emission models around Sgr A* concerns the propagation of positrons. Though the most recent SPI analyses show evidence for a point source consistent with Sgr A* (Siegert & Diehl, 2016; Skinner et al., 2014), the majority of positrons must diffuse away from Sgr A* to fill the bulge before annihilation. This scenario was explored in Alexis et al. (2014). The authors found that the Galactic bulge 511 keV distribution could be explained by a transient source injecting a large number of positrons in the GC, as discussed. However, they determined through propagation simulations that an outburst as recent as $3{\times}10^5$ yr can be ruled out since there is not enough time to fill the Galactic bulge, and therefore they dismiss the theory of Totani (2006). Positron propagation will be further discussed in Section 2.6.

2.4.3 MESON AND LEPTON DECAY

With a mass of $m_{\pi\pm} \approx$139.6 MeV, charged (and neutral) π mesons have long been considered a keystone to γ-ray astrophysics (Pollack & Fazio, 1963). For the consideration of positron production, we will limit this discussion to the positively charged pion. π^+ dominantly decays into a muon and muon-neutrino with a mean lifetime of $2.6{\times}10^{-8}$ s (Greenberg et al., 1969):

$$\pi^+ \rightarrow \mu^+ + \nu_\mu. \tag{2.13}$$

The muon, with a mass of $m_\mu \approx$105.7 MeV and lifetime 2.20 μs (Bardin et al., 1984), then decays into a positron:

$$\mu^+ \rightarrow \mathrm{e}^+ + \nu_e + \bar{\nu}_\mu. \tag{2.14}$$

Pions are produced in the Galaxy predominantly through the interactions of cosmic-ray protons with neutral or ionized hydrogen in the ISM (Blattnig et al., 2000). The dominant interactions are

$$\begin{aligned} p + {}^1_1\mathrm{H}^+ &\rightarrow p + n + \pi^+ \\ p + {}^1_1\mathrm{H}^+ &\rightarrow {}^2_1\mathrm{H} + \pi^+. \end{aligned} \tag{2.15}$$

The threshold kinetic energy for the incident proton in these reactions is 290 MeV (Pollack & Fazio, 1963). At higher proton energies ($\gtrsim$ GeV), interactions which create multiple π^+-π^- pairs and neutral pions dominate. With a mass of $m_K \approx$493.7 MeV, charged kaon mesons can be produced in more energetic $p-p$ collisions and there are a number of decay routes that lead to the production of positrons (Moskalenko & Strong, 1998):

$$\begin{aligned} K^+ &\rightarrow \pi^0 + \pi^+ \\ K^+ &\rightarrow \mu^+ + \nu_\mu, \end{aligned} \tag{2.16}$$

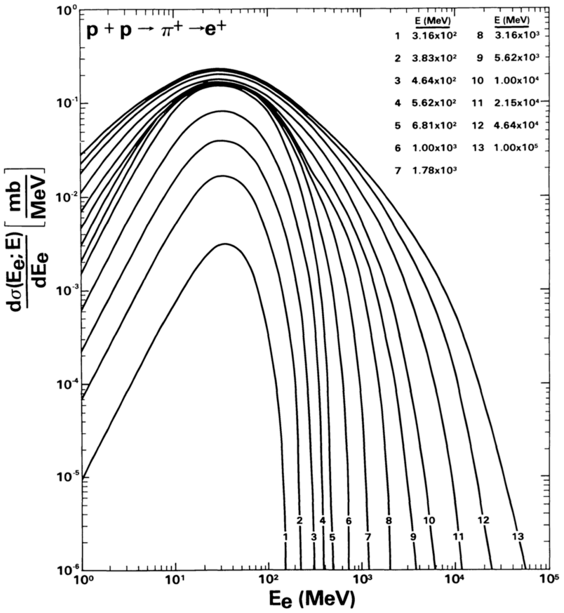


Figure 2.21: Energy spectra of positrons resulting from the decay of π^+ produced in $p - p$ interactions. Initial kinetic energies of the proton between $3.16{\times}10^2$ and $1.0{\times}10^5$ MeV are considered. Independent of the initial proton energy, the positron energies peaks around 30 MeV. Figure 7 from Murphy et al. (1987).

where the decay of the pion and muon follow Equation 2.13 and 2.14, respectively.

Pion and kaon production from $p - p$ interactions occur throughout the Galaxy as cosmic rays are prevalent and the dominant components of the ISM are neutral and ionized hydrogen (Moskalenko & Strong, 1998). Murphy et al. (1987) explored the expected energy distribution of positrons from $p - p$ interactions (Figure 2.21). From calculations with various initial proton kinetic energies, the authors find that the positrons produced have a spectrum that peaks between 30-40 MeV. Therefore, interactions of this type have difficulty satisfying the observed MeV continuum discussed in Section 2.3.2

2.4.3.1 COSMIC RAYS

Cosmic rays are high-energy particles of a cosmic origin discovered by Victor Hess in 1912 in balloon-borne experiments. About 90% of cosmic rays are protons, 9% are α-particles, and the remaining percentage consists of heavy ions, electrons, and a few anti-particles (Perkins, 2003). Cosmic-rays span an impressive range of energies from ~1 GeV to energies above 10^{20} eV. Cosmic-rays of medium energy (10^{10}–10^{18} eV) are of Galactic origin and are thought to be produced in SN, SNR, and with a contribution from pulsars and XRBs (Strong et al., 2007).

The high-energy particles produced in the accelerating source are considered primary cosmic rays, whereas particles produced by primary interactions in the ISM are considered secondary cosmic rays. The majority of cosmic-ray positrons are believed to be secondary cosmic rays produced in $p-p$ interactions within the ISM, and thus models of production rely heavily on propagation models (Strong et al., 2007). Porter et al. (2008) perform a detailed analysis of the X-ray and soft γ-ray emission from secondary cosmic rays in the Galactic plane. The authors find the production rate of positrons as secondary cosmic rays in the Galaxy is ~2×10^{42} e^+/s. Although cosmic rays can provide a large source of positrons in the Galaxy, they must be excluded as a major contributor to the Galactic positron signal since they are created at energies in excess of ~30 MeV, as discussed in Section 2.4.3.

2.4.3.2 SAGITTARIUS A*

As discussed in Section 2.4.2.3, Sgr A* has been proposed as a source of Galactic positrons from γ-γ pair creation. A similar non-steady-state scenario can be invoked for positrons produced through $p-p$ collisions in a TDE, as discussed in Cheng et al. (2007). Due to the high energy of positrons produced in the subsequence π^+ decay, the issue of in-flight annihilation is a problem for these models; however, Chernyshov et al. (2010) argue that if the magnetic field in the Galactic bulge is large enough (> 0.4 mG), then the positrons will lose their energy rapidly enough for in-flight annihilation to be negligible. Current observations do not favor such a strong magnetic field (Ferrière, 2009).

2.4.4 DARK MATTER

After the release of the early 511 keV spatial maps from SPI, shown in Figures 2.5 and 2.6, dark matter (DM) was proposed as a possible source as the distribution of DM is expected to be strongly concentrated in the center of the Galaxy (Bœhm et al., 2004). As positrons and electrons are the lightest leptons, they are predicted to be formed in pairs through the annihilation, de-excitation, or decay of DM particles. There are various means of DM positron production proposed (Hooper & Wang, 2004; Picciotto & Pospelov, 2005; Oaknin & Zhitnitsky, 2005), but here we will focus on the more "natural" scenarios: direct annihilation of low-mass DM particles and decay of weakly-interacting massive particles (WIMPs).

Sizun et al. (2006) propose that self-annihilation of low-mass DM particles could be responsible for the Galactic positron excess. DM particles with mass $m_\chi < 200$ MeV are expected to annihlate into electron-positron pairs, and the total kinetic energy of the pair will be $2m_\chi$. The author's work and that of Beacom & Yüksel (2006) limit the mass for annihilating DM as a source of Galactic positrons to be $\lesssim 7$ MeV and $\lesssim 3$ MeV, respectively. However, with updated cosmological data, Wilkinson et al. (2016) claim that annihilation of light DM particles is strongly disfavored.

Pospelov & Ritz (2007) propose that the decay of WIMPs could be the primary source of MeV positrons. The decay can occur after collisional excitation or long-lived states surviving from the Big Bang, where the WIMP has MeV-scale energy-level differences. The authors find that the cross section required to produce a large fraction of Galactic positrons through collisional excitation is too large and must be ruled out. Additionally, the authors note that the decay of a metastable WIMP state could explain the positron production rate if it had a lifetime of 10^9–10^{13} yr; however, they conclude that such a suggestion requires *ad hoc* tuning of the WIMP interaction parameters. Cembranos & Strigari (2008), however, claim that they can explain both the 511 keV emission and the diffuse higher energy MeV emission seen by COMPTEL and SMM (Weidenspointner et al., 2000; Watanabe et al., 2000) with decaying WIMPs.

Two of the leading spatial models that are thought to describe the DM halo of the Galaxy are the Navarro-Frenk-White (NFW) profile (Navarro et al., 1997) and the Einasto profile (Einasto, 1965). Vincent et al. (2012) find the 511 keV morphology fits well with DM described by the Einasto profile, Figure 2.22, while Bœhm & Ascasibar (2004) find a better fit to the NFW profile, though this was with an earlier release of the SPI data (Knödlseder et al., 2005). On the other hand, Lingenfelter et al. (2009) argue that DM sources cannot account for the observed spectral o-Ps fraction without significant propagation, and with propagation, the spatial motivation for DM to explain the unaccounted for 511 keV flux in the budge is no longer valid.

If DM is indeed a source of Galactic positrons, then one should detect emission from dwarf galaxies as they are expected to be DM dominated. Recent attempts have been made to observe 511 keV emission from Milky Way satellite

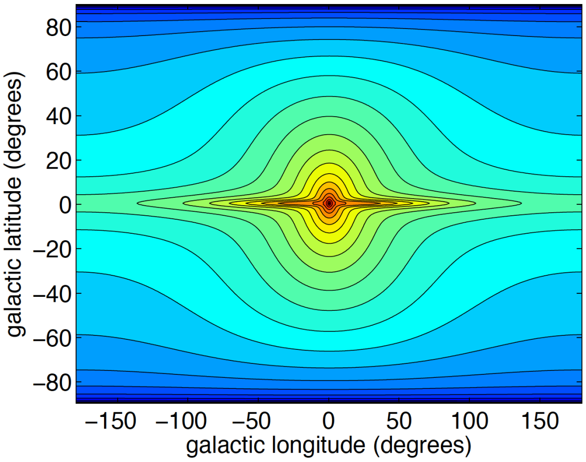


Figure 2.22: Predicted 511 keV intensity sky map from the Einasto and disk model after a DM halo model fit with the observed SPI data. The bulge component is from scattering or annihilation in the DM halo, and the disk component is attributed to β^+ decay of ^{26}Al and ^{44}Ti. The authors assume positron transport is small. Figure 1 from Vincent et al. (2012).

galaxies with SPI (Siegert et al., 2016b). In a survey of 39 galaxies, the authors report a 3.1 σ detection of 511 keV emission from the dwarf galaxy Reticulum II and a 2 σ detection in 5 other dwarf galaxies, though stacking the signals of all 39 galaxies does not result in a positive signal. The authors have placed a firm upper limit on the 511 keV emission from DM in such systems, but higher sensitivity will be needed to constrain DM and positron theories.

So little is known about DM and it is trivial to conceive of processes that could produce positrons; as a result, finding an explanation for the Galactic positron source with a few theory tweaks is not difficult. Until the community knows more about DM or the nature of Galactic positrons, it will be difficult to rule out DM as a possible source.

2.5 POSITRON INTERACTIONS AND ANNIHILATION WITHIN THE ISM

In the production processes discussed in Section 2.4, positrons are created with an initial kinetic energy of ~ MeV, which is higher than the ambient energy in the ISM. The positrons must therefore undergo energy loss, thermal-

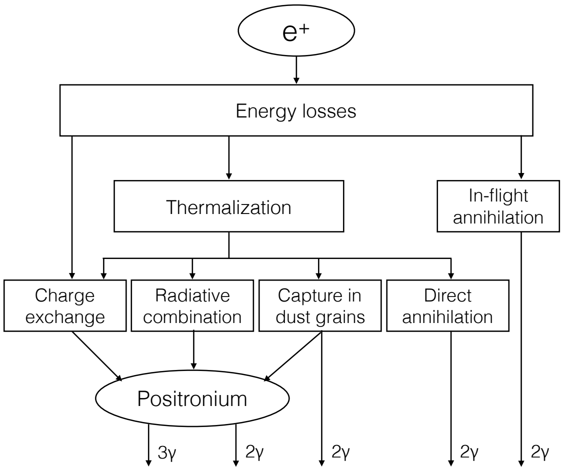


Figure 2.23: Summary of positronium formation and positron annihilation processes. Modification of Figure 1 from Guessoum et al. (1991).

ization, and finally annihilation, often after forming positronium, to produce the Galactic signal observed. This requires an accurate description of the ISM components to understand these interactions. First, an overview of the ISM will be given in Section 2.5.1. A summary of the interaction processes for positrons is presented in Figure 2.23, and will be discussed in further detail in Sections 2.5.2-2.5.5. Finally, the spectral signature of positron interactions in different phases of the ISM will be discussed in Section 2.5.6.

2.5.1 MODEL OF THE INTERSTELLAR MEDIUM

An accurate model of the ISM, including temperature, ionization fraction, density, etc., of the gas and dust is necessary in order to understand the interactions within the Galaxy and to estimate the overall positron annihilation spectrum. The most detailed description of the components of the ISM and their spatial distributions have been compiled by Ferrière in three separate papers that describe: the innermost 10 pc around Sgr A* (Ferrière, 2012), the inner 3 kpc of the Galaxy (Ferrière et al., 2007), and the Galactic disk (Ferrière, 1998). The general properties of the ISM as described in these papers are summarized here.

The ISM contains ordinary matter, cosmic rays, and magnetic fields, all which interact with each other and have comparable pressures. The interstellar matter accounts for 10–15% of the total mass of the Galactic disk and

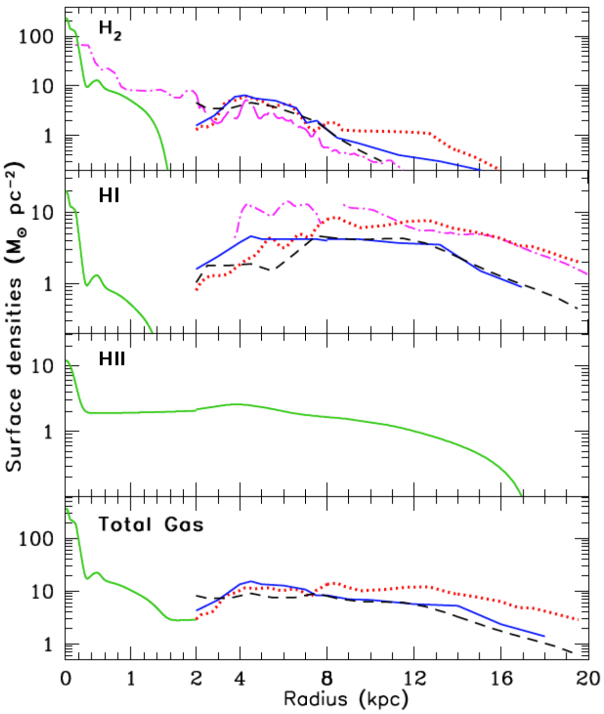
Component	T [K]	f_{ion}	n [cm^{-3}]	$\mathcal{M}$ [10^9 M$_\odot$]
Molecular	10–20	$\lesssim 10^{-4}$	10^2 – 10^6	1.3 – 2.5
Cold atomic	50–100	4×10^{-4}–10^{-3}	20–50	$\}\gtrsim 6.0$
Warm atomic	6000–10000	0.007–0.05	0.2–0.5	
Warm ionized	~8000	0.6–0.9	0.2–0.5	$\}\gtrsim 1.6$
Hot ionized	~10^6	1	~0.0065	

Table 2.4: Temperature (T), ionization fraction (f_{ion}), density (n), and mass contained in the Galaxy ($\mathcal{M}$) of the five components of the ISM. The neutral component, in the cold and warm atomic phase, accounts for $\gtrsim 6.0\times10^9$ M$_\odot$, and the warm and hot ionized phases account for $\gtrsim 1.6\times10^9$ M$_\odot$. Modified from Table 1 from Ferrière (2001) and Table 1 from Jean et al. (2009).

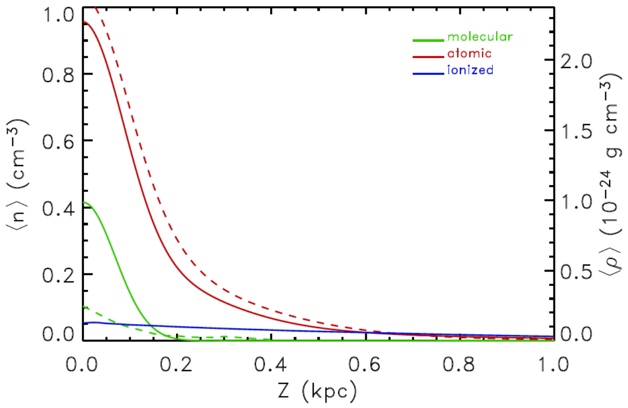
can be in the form of gas or dust particles. By mass, 70.4% of the interstellar matter is hydrogen, 28.1% is helium, and 1.5% consists of heavier elements commonly called "metals." The gas can be found in five phases: molecular, cold atomic, warm atomic, warm ionized, and hot ionized; see Table 2.4 for a summary of these components.

About half of the interstellar matter by mass is confined in cold molecular and atomic clouds. The molecular clouds are gravitationally bound and can be large structures up to 20–80 pc with a mass of 10^6 M$_\odot$, or small with mass $\lesssim 10^3$ M$_\odot$ (Goldsmith, 1987). The most abundant molecules in the molecular clouds are H_2 and CO. The atomic clouds are filamentary and are likely created by stellar winds and SN explosions. Warm ionized regions are found surrounding hot O and B stars, in addition to a diffuse complex structure with a larger scale height that is thought to be potentially from cosmic rays (Valinia & Marshall, 1998) or dust particles (Reynolds & Cox, 1992), among other theories. Hot ionized interstellar gas is accepted to be produced around SN and stellar winds. About 0.5–1% of the interstellar matter is in the form of dust and it is spatially correlated with atomic and molecular regions. The radial and vertical distribution of these ISM components are shown in Figure 2.24.

The stellar Galactic bulge is defined as the region of the Galaxy inside R$\lesssim$3 kpc and is completely distinct from the 511 keV Galactic bulge emission. The stellar Galactic bulge is generally hotter and denser than the Galactic disk, and there are more defined spatial features. Between 1.5–3 kpc, the Galactic bulge is relatively devoid of gas. With a radius ~1.5 kpc, a tilted elliptical disk contains dense atomic (10–20%) and molecular (80-90%) hydrogen, and is thought to have an elliptical hole in the center. The molecular hydrogen is



(a) Radial density profile of ISM hydrogen.



(b) Volume density versus height of ISM hydrogen.

Figure 2.24: (a) Azimuthally-averaged surface densities of molecular (H_2), atomic (HI), and ionized (HII) hydrogen as a function of Galactic radius where each curve is from a different study, see caption for Figure 8 from Prantzos et al. (2011) for a list of references. (b) Average volume densities of molecular, atomic, and ionized hydrogen as a function of distance from the Galactic plane, averaged along the solar circle ($R = R_\odot$). The curves are from the same references as in (a).

mostly confined to the thin central molecular zone (CMZ), which appears as a smaller ellipse aligned with the Galactic plane with radius ~200 pc and scale height ~30 pc, sitting within the atomic tilted disk. The hot and warm ionized gases are not constrained to the tilted disk or the CMZ and appear to fill the entire Galactic bulge. The total mass of the interstellar matter in the Galactic bulge is ~1.3×10^8 $M_\odot$, with 41% in molecular form, 4% in atomic form, and 55% in ionized form.

The interstellar magnetic field is thought to be nearly azimuthal with a pitch angle of ~8° and a field strength of a few μG (Ferrière, 2009), though there still remains much uncertainty here. The strength of the magnetic field increases with density toward the Galactic center, reaching perhaps 6-7 μG. Close to the GC ($\lesssim$300 pc), the field is thought to be poloidal with strengths of 10 μG up to 1 mG in the molecular cloud regions (Ferrière, 2009).

2.5.2 ENERGY LOSS MECHANISMS

As a charged particle, the positron loses energy via collisional interactions with particles and scattering off magnetic fields in the ISM. The interaction type depends on the density and ionization fraction of the ISM and the energy of the positron. Above 1 GeV, the dominant energy-loss processes for positrons are inverse Compton scattering with cosmic microwave background photons, synchrotron radiation near magnetic fields, and bremsstrahlung radiation from interactions with ions, electrons and atoms (Blumenthal & Gould, 1970). At lower energies, the relevant range for the Galactic positrons considered here, Coulomb scattering is the dominant process in ionized media, and the excitation and ionization of atoms, in addition to Ps formation by charge exchange, dominate in a neutral media (Guessoum et al., 1991).

Coulomb scattering is elastic scatting of a charged particle in the Coulomb field of a target particle. Coulomb scattering with free electrons is the dominant energy-loss process for MeV positrons in ionized media (Jean et al., 2009). The target electrons for scatters at these energies can be considered at rest, and the energy loss rate of the positron can be determined by the collision rate in a cold plasma (Prantzos et al., 2011). The energy loss rate from Coulomb scattering and those from higher-energy processes are shown in Figure 2.25.

In neutral phases, positrons lose energy through inelastic collisions with ISM atoms and molecules, resulting in excitation or ionization, namely of H, H_2, and He (Guessoum et al., 2005). Additionally, charge exchange, which

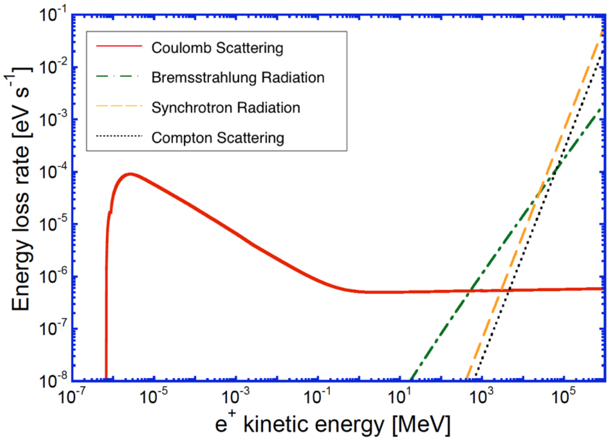


Figure 2.25: Energy loss rate in eV/s for the dominant processes as a function of the positron energy from eV to TeV, assuming a fully ionized ISM with $T = 8000$ K. The energy loss rate for synchrotron radiation assumes $B = 5\mu$G, and the energy loss rate for inverse Compton scattering assumes $U_{ph} = 0.26$ eV/cm^3. For positrons with energy $\lesssim$1 MeV, the dominant process is Coulomb scattering. Figure 1 of Jean et al. (2009).

occurs when a positron takes the electron from an atom forming positronium and an ion, dominates at energies below ~30 eV (Bussard et al., 1979). In these interactions, the positrons lose a substantial amount of their energy in one interaction, unlike the continuous process of Coulomb interactions. Table 2.5 lists the ionization, excitation, and charge exchange interactions relevant for Galactic positrons and their respective energy thresholds. Figure 2.26 shows the cross section for the ionization, excitation, and positronium formation through charge exchange on atomic hydrogen.

2.5.3 MAGNETOHYDRODYNAMIC WAVE SCATTERING

Under certain conditions, the plasma nature of the ISM allows for positrons to interact through wave-particle resonance with magnetohydrodynamic (MHD) waves (Lithwick & Goldreich, 2001). These collisionless processes are thought to be an efficient energy-loss process for positrons (Petrosian & Bykov, 2008; Higdon et al., 2009). However, Jean et al. (2009) conclude that scattering off MHD waves is only relevant for the warm ionized and hot phases, but anisotropies in the magnetic field can significantly reduce the effects. The

Process	Threshold [eV]
$e^+ + H \rightarrow Ps + H^+$	6.8
$e^+ + H \rightarrow e^+ + e^- + H^+$	13.6
$e^+ + H \rightarrow e^+ + H^*$	10.2
$e^+ + H \rightarrow e^+ + H^{**}$	12.1
$e^+ + H_2 \rightarrow Ps + H_2^+$	8.6
$e^+ + H_2 \rightarrow e^+ + e^- + H_2^+$	15.4
$e^+ + H_2 \rightarrow e^+ + H_2^*$	12.0
$e^+ + He \rightarrow Ps + He^+$	17.8
$e^+ + He \rightarrow e^+ + e^- + He^+$	24.6
$e^+ + He \rightarrow e^+ + He^*$	21.2

Table 2.5: The energy loss of positrons in neutral media is dominated by charge exchange, ionization, and excitation. The thresholds for each interaction in eV are listed. Table 1 from Guessoum et al. (2005).

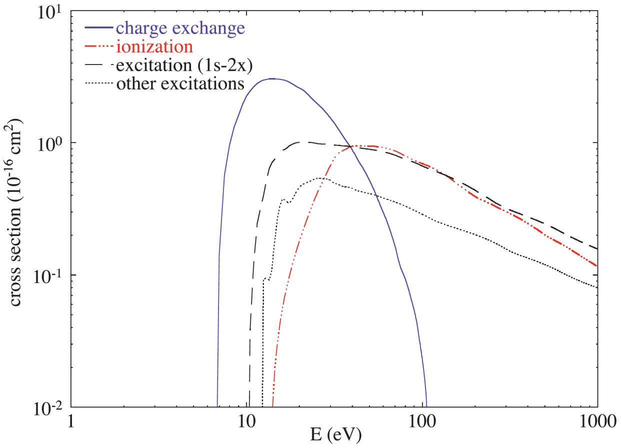


Figure 2.26: The relative cross section for positron interactions with atomic hydrogen, Figure 1 from Guessoum et al. (2005). Ionization and excitation of hydrogen dominate for high energies, where positronium formation through charge exchange dominates below ~30 eV.

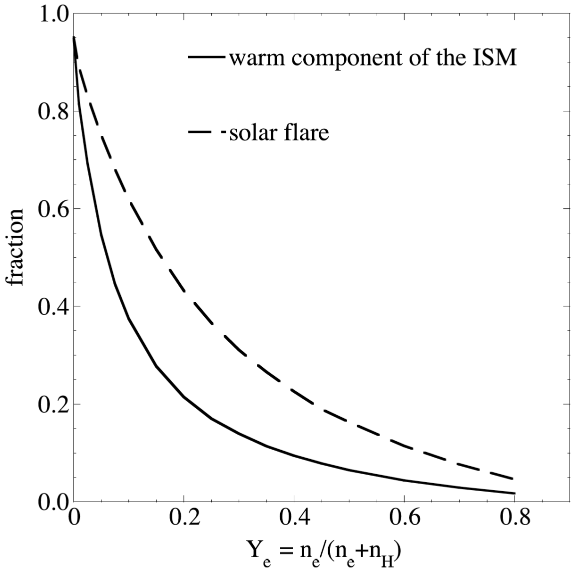


Figure 2.27: Fraction of positrons which form positronium in flight through charge exchange with atomic hydrogen as a function of the ionization fraction. The ionization fraction in the warm phases of the ISM is compared with that in solar flares, which are hotter and denser. Figure 26 from Prantzos et al. (2011).

scattering and transport along MHD waves will be further discussed in the consideration of positron propagation in Section 2.6.

2.5.4 ANNIHILATION IN FLIGHT

Positrons can directly annihilate with free or bound electrons in flight. These are continuous processes where the emitted photons will have an energy of $mc^2/2 < E_\gamma < E + mc^2/2$; however, the fraction of positrons that annihilate directly in flight is negligible for energies lower than 1 MeV (Prantzos et al., 2011).

Positronium can form in flight through charge exchange with the electron of H, H_2, and He atoms, as discussed in Section 2.5.2. Figure 2.27 shows the fraction of positrons which form positronium through charge exchange with H as a function of the ionization fraction of the medium. Once positronium is formed, the non-zero kinetic energy of the bound state will result in a Doppler-broadened 511 keV line from p-Ps annihilation.

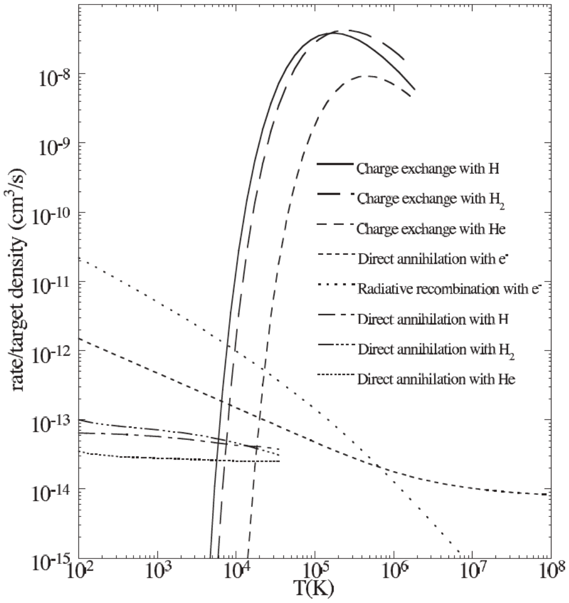


Figure 2.28: Positron reaction rates after thermalization as a function of ISM temperature. Charge exchange is the dominant process for the warm phase; at higher temperatures, few atoms remain neutral, and at lower temperatures the positrons do not satisfy the threshold for charge exchange. Below 10^4 K, radiative combination with free electrons is the main interaction. Direct annihilation with electrons is only relevant for the highest temperatures. Figure 3 from Guessoum et al. (2005).

2.5.5 ANNIHILATION AFTER THERMALIZATION

After the electrons have lost the majority of their energy, the positrons thermalize with the ISM and can be assumed to have a Maxwellian energy distribution before annihilation. Figure 2.28 shows the relative reaction rates for the possible annihilation channels as a function of ISM temperature. The main interaction for positron annihilation after thermalization in warm ISM phases ($T \gtrsim 10^4$ K) is charge exchange with H. In cold phases, radiative combination with free electrons ($e^+ + e^- \rightarrow Ps + \gamma$) dominates. Direct annihilation with free electrons has a cross section an order of magnitude less than radiative combination and is only relevant in hot ISM phases. Direction annihilation with bound electrons has the lowest cross section of all positron processes, and is only relevant in the coldest phases.

2.5.5.1 DUST GRAINS

Dust in the ISM is thought to play a role in the annihilation process of thermalized positrons, especially in the warm ionized phases (Zurek, 1985). About 0.5–1% of interstellar matter is in the form of dust (Ferrière, 2001), and it is spatially correlated with the ISM gas regions (Boulanger et al., 1996). These dust particles range in size from 0.001–1 μm and are mostly made up of silicate particles (Mg_2SiO_4 or Fe_2SiO_4) (Kruegel, 2002).

Guessoum et al. (2005) have done the most thorough study of positron annihilation in dust to date. The authors find that positron capture on dust particles is negligible in all phases of the ISM except the warm ionized and hot phases (Guessoum et al., 2005). In these phases, the positron with either 1) back-scattered or be re-emitted as a positron (15–20% of the time), 2) form positronium within the grain and annihilate inside (65-80%), 3) form positronium in the grain and be ejected (5–15%). The o-Ps that forms and annihilates within the grain will result in a 2-γ decay since the briefly-bound positron will most likely annihilate with a nearby electron in the dust grain, this process is called “pick-off” annihilation (Guessoum et al., 2005).

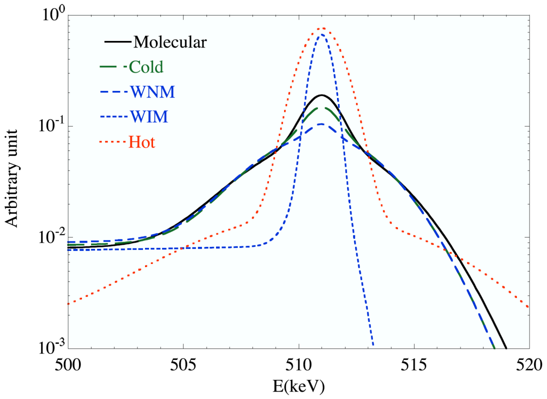
2.5.6 DISCUSSION OF SPECTRAL SIGNATURES

Each of the processes discussed in the previous sections will result in different annihilation signatures. A careful description of the ISM physical conditions, the interaction cross sections, and estimates of the line-width and positronium fractions from these interactions are needed to predict the resulting spectrum.

Using the most recent cross section and annihilation information, Guessoum et al. (2005) used Monte-Carlo simulations to find the positron annihilation spectra for each phase of the ISM and for the ISM as a whole, using the ISM models of McKee & Ostriker (1977). See Figure 2.29 for a table listing the calculated FWHM of the 511 keV line for each phase of the ISM and the expected annihilation spectral shapes. By combining the spectra from each phase of the ISM with their expected filling fractions, one can use the total spectra to fit the observed annihilation spectra from the GC, such as measured by INTEGRAL/SPI, and infer which phases have the largest contribution. Analysis of this type has been performed by Jean et al. (2006) and Churazov et al. (2005) and is discussed in Section 2.3.1. From their studies both authors

Phase	FWHM
Molecular	2.39
Cold	3.00
Warm neutral	4.78
Warm ionized	1.02
Hot	1.99
Total ISM	2.26

(a) FWHM of 511 keV line in each ISM phase.



(b) Spectrum of positron annihilation line in each phase.

Figure 2.29: (a) Predictions of the FWHM for the five different phases of the ISM, and the total combined FWHM for the modeled ISM. Table 5 from Guessoum et al. (2005). (b) Total annihilation spectra from each phase of the ISM, with arbitrary scaling. Figure 28 from Prantzos et al. (2011).

found that the measured SPI spectra is best fit with positrons annihilating in the warm neutral and warm ionized phases of the ISM.

2.6 POSITRON PROPAGATION IN THE ISM

Positrons are born with kinetic energies ~MeV, but from the observed o-Ps fraction we know they annihilate at low energies $\lesssim$ 10 eV. Therefore, the positrons must decelerate and most likely travel some distance before annihilation. There are many proposed sources of Galactic positrons that can produce a large enough rate but do not satisfy the observed morphology of the 511 keV emission, and some authors have proposed that positrons created in these sources could propagate to fill the bulge (Higdon et al., 2009; Prantzos, 2006). The big question is: How far do positrons propagate from their point of production until annihilation? And equivalently, do we expect the positron source distribution to match the annihilation emission?

To understand how far positrons can propagate, an accurate description of the Galaxy is necessary. As a charged particle, the propagation of a positron is affected by the Galactic magnetic field, and as collisions with particles in the ISM contribute to positron energy-loss, as discussed in Section 2.5, the

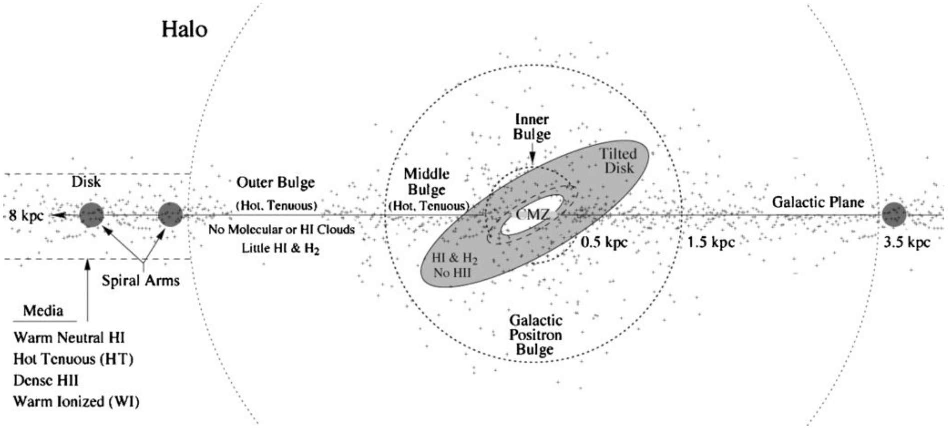


Figure 2.30: Schematic of the ISM phases and stellar distributions assumed in Higdon et al. (2009) following Ferrière et al. (2007). Figure 1 from Higdon et al. (2009).

density, temperature, and ionization fraction of the gas in different phases of the ISM will affect the propagation distance. All of the propagation simulations and studies discussed here follow the Ferrière description, reviewed in Section 2.5.1.

Higdon et al. (2009) were the first to do a thorough study of MeV positron propagation in an attempt to understand the 511 keV observations. They assumed a stellar population, with β^+ decay from SNe products ^{26}Al, ^{44}Ti, and ^{56}Ni as the source of positrons, superimposed on a gas and plasma distribution in which the deceleration and annihilation occur. Figure 2.30 illustrates the ISM model used in these simulations. The authors assume that the propagation is dominated by scattering off MHD waves, discussed in Section 2.5.3. Through propagation distance approximations, the authors conclude that positrons from β^+ decay of isotopes created in massive stars and SNe can fully account for the spatial distribution, the 511 keV spectral line width, and o-Ps fraction. Other papers, however, find issue with their analysis (Prantzos et al., 2011; Martin et al., 2012; Alexis et al., 2014). Jean et al. (2009) argue that the positron mean free-path Higdon et al. (2009) derived is only valid for interplanetary plasmas. Furthermore, Jean et al. find that collisional interactions, discussed in Section 2.5.2, are more constraining to the propagation distance than MHD scattering.

Jean et al. (2009) performed Monte-Carlo simulations to estimate the distances of positron propagation with different interaction processes: scattering off of magnetic turbulence (MHD waves) in ISM, scattering off of individual

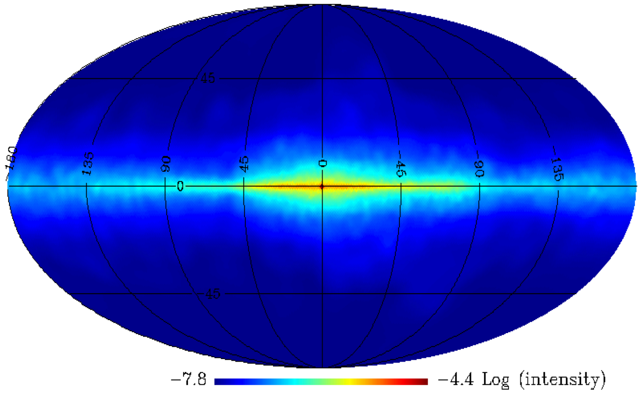


Figure 2.31: The simulated all-sky distribution of 511 keV emission from nucleosynthesis positrons, assuming an escape fraction of 5% for ^{56}Ni and a dipolar Galactic magnetic field. The emission along the disk can account for the measured 511 keV Galactic disk emission, but cannot explain the bugle emission. Figure 3 from Alexis et al. (2014).

particles in the ISM, and advection with large scale fluid motions of the ISM. The authors find that propagation for MeV energy positrons is dominated by collisions with gas particles, and the positrons can travel very large distances up to ~30 kpc along magnetic field lines.

Alexis et al. (2014) have performed the most detailed analysis of positron propagation to date. The authors developed a Monte-Carlo propagation code of ~MeV positrons that takes into account the transport through collisional interactions with the ISM. Their simulation, like the studies of Higdon et al. (2009), includes only the β^+ decay of ^{26}Al, ^{44}Ti, and ^{56}Ni (with several positron escape fractions considered) as the source of positrons. They account for the spatial distribution of the sources, the ISM phases, and three different models for the Galactic magnetic field. The authors derived 511 keV emission sky maps, light curves, and spectra and they conclude that SN β-decay products alone cannot explain the Galactic bulge emission observed by SPI. Figure 2.31 shows one of the derived sky maps with an assumed positron escape fraction of ^{56}Ni of 5% and a dipole halo Galactic magnetic field. They find that ~MeV positrons do not propagate far from their birth sites, traveling on average $\lesssim$1 kpc, and conclude that the 511 keV emission should represent the positron source distribution. Interestingly, the simulations show a sharp intensity peak at the GC due to annihilation in the dense CMZ, which can explain the point source emission seen in the Skinner and Siegert spatial models, discussed in Section 2.3.1.

2.7 FURTHERING OUR UNDERSTANDING OF GALACTIC POSITRONS

After almost 5 decades of scientific investigation, there are still big questions about Galactic positrons. The 511 keV emission from the Galactic disk is understood to be explained by nucleosynthesis β^+ decay, however, there is no conclusion as to the source of positrons in the Galactic bulge region. The spatial morphology of the emission has not been well constrained due to the fact that the most detailed studies have been performed with SPI, an indirect imaging telescope. The reported extent of the disk emission, which has a large impact on the estimated number of annihilating positrons in the Galaxy, is drastically different in the two most recent studies. Therefore, not even the annihilation rate is well known.

It is not clear if the 511 keV emission should trace the distribution of positron sources or if there is significant positron propagation. Both theoretical advancements in the understanding of positron interactions, the constituents of the ISM, and the Galactic magnetic field, and a more accurate image of the 511 keV emission are needed to further advance this discussion. Another avenue of approach that would be available with a direct image of the 511 keV emission is a comparison with the ^{26}Al 1.8 MeV map. Through a detailed comparison study of the two spatial maps, not only can the contribution of positrons from ^{26}Al be measured, the propagation of positrons could be constrained.

Most of the current observational issues arise from the use of an indirect imaging telescope. When observing an extended source with a coded-aperture mask instrument, a model fitting approach must be taken and the true source distribution remains unknown. A direct imaging telescope would be able to determine the spatial morphology, conclusively determine the extent of the disk emission, and measure the true annihilation rate from different regions of the Galaxy. Additionally, spectral studies of various spatial components, attempted with SPI data, would not have to rely on assuming an underlying model. Furthermore, an improved sensitivity would allow for a possible detection of individual sources of positron propagation, such as XRBs within the Galaxy or emission from DM dominated dwarf galaxies.

The Compton Spectrometer and Imager (COSI) is a telescope that has been developed with the goal of furthering our understanding of Galactic positrons. With its direct imaging capabilities inherent to a Compton telescope, COSI can answer some of the questions left open by SPI. The sensitivity of COSI is similar to that of SPI, even from a balloon platform; however a COSI-like instrument onboard a satellite could be at least an order of magnitude more sensitive, e.g., GRX. This is the ultimate goal of the COSI collaboration.

PART II

THE COMPTON SPECTROMETER AND IMAGER: INSTRUMENT OVERVIEW, CALIBRATIONS, AND 2016 FLIGHT

3

THE COMPTON SPECTROMETER AND IMAGER

The Compton Spectrometer and Imager (COSI) is a balloon-borne soft γ-ray telescope that is designed to perform novel polarization measurements of compact sources, study stellar nucleosynthesis, and further our understanding of Galactic positrons. The COSI mission has been in development for almost two decades; in collaboration with Taiwanese and French researchers, a small team at University of California, Berkeley, has designed, built, tested, and successfully flown the instrument.

In this chapter[1], we first discuss γ-ray astrophysics and the associated challenges in this energy range and then present an overview of the current technologies for MeV γ-ray telescopes in Section 3.1. The basics of Compton telescopes, of which COSI is an example, are introduced in Section 3.2. Details about the COSI instrument and the flight gondola will be given in Section 3.3, and an overview of the history of the COSI project will be given in Section 3.4. The MEGAlib analysis tool used by the COSI collaboration will be discussed in Section 3.5.

3.1 GAMMA-RAY ASTROPHYSICS

Gamma-rays are fantastic tools to explore astrophysical objects: they are emitted by the most extreme objects in the universe; they travel long distances without being absorbed or deviated; they probe further into objects than other wavelengths; and γ-ray line measurements act as fingerprint-like probes into sites of nucleosynthes is. However, the soft to medium γ-ray regime, from about 100 keV to 100 MeV, is one of the least astrophysically explored energy ranges, often referred to as the *MeV Gap*. The sensitivity of current instru-

---

[1] Some of this chapter is based on a previous publication ”Calibration of the Compton Spectrometer and Imager in preparation for the 2014 balloon campaign” by Kierans et al. (2014). ©2014, Society of Photo-Optical Instrumentation Engineers (SPIE). Reprinted with permission.

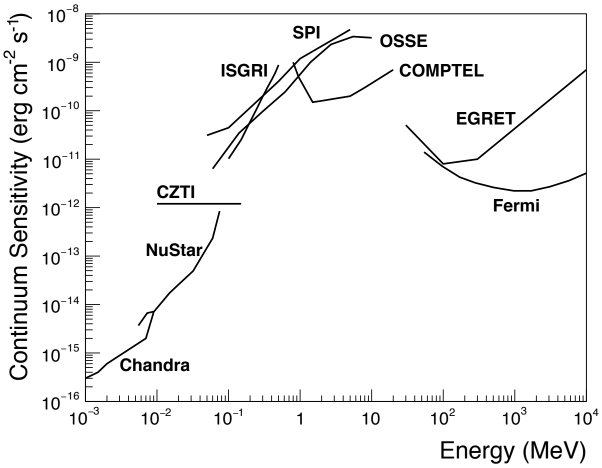


Figure 3.1: The continuum sensitivity for present and past X-ray and γ-ray astrophysics instruments. The continuum sensitivity is a measure of how faint a source can be detected; a lower sensitivity is better. The segment with the worst sensitivity from 100 keV to 100 MeV is referred to as the *MeV gap*.

ments in this energy range is orders of magnitude worse than the sensitivity in neighboring bands, as illustrated in Figure 3.1. This is due to a few major factors: low inherent interaction cross sections, high-instrumental backgrounds, and technological constraints. There are many unanswered questions, including the true source of Galactic positrons, that arise from this lack of exposure.

3.1.1 CHALLENGES IN γ-RAY ASTROPHYSICS

The aspects that make γ-rays fantastic tools for astrophysics also lead to telescope instrumentation challenges. For example, the high-penetrating power of γ-rays make them difficult to focus with lenses and mirrors common in telescopes apt for lower energies. Not only does the total photon-matter interaction cross section reach minimum at a few MeV, as shown in Figure 3.2, but in this energy range it is dominated by Compton scattering: an inelastic collision of a photon and an electron. Instead of a single interaction in which the photon will impart all of its energy, such as in photoelectric absorption, Compton scattering results in a γ-ray that “bounces” around in a large detec-

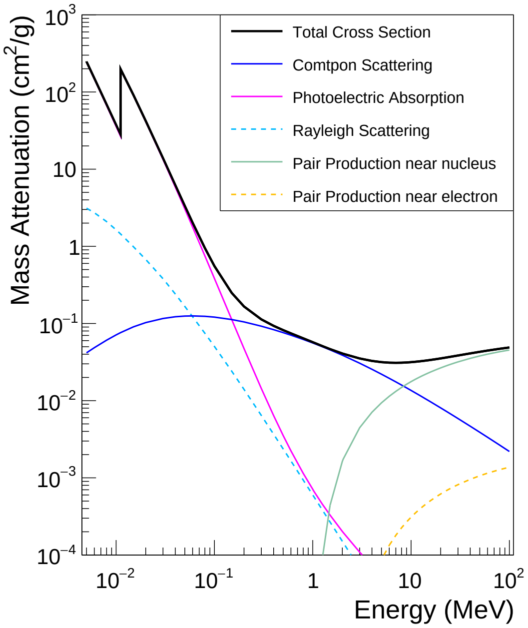


Figure 3.2: Total mass attenuation as a function of energy for germanium. At 1 MeV, the cross section is dominated by Compton scattering. The sharp edge around 200 keV is a K-edge from the germanium electron shell structure. Cross section information is from NIST XCOM (Berger et al., 2017).

tor volume. With an interaction depth of ~10 g/cm^2, a few centimeters of dense germanium are needed to fully stop MeV γ-rays.

The Earth's atmosphere is opaque to γ-rays, and therefore an instrument must be above the atmosphere to perform astrophysical studies. This can be done through three platforms: 1) satellites provide the best access to space in low radiation environments for multi-year missions but come with a large price tag; 2) rockets can achieve altitudes of 100's of km, but the exposure time of less than 10 minutes is much too short for cosmic sources of γ-ray emission ($\lesssim 10^{-3}\ \gamma/\mathrm{cm}^2/\mathrm{s}$); and 3) sub-orbital balloons can attain altitudes of 30–40 km for weeks of observations and allow for multiple flights of the same payload. Balloon-platforms are often used as a stage in satellite development, as they provide a convenient means of testing and modifying new technologies. The eventual goal of the COSI mission is to fly on a satellite platform to achieve years of exposure in a low-background environment; in the mean time, we can still perform some interesting science from the balloon-platform.

Background radiation dominates at these energies. Since there is no focusing element, the telescope aperture is the same as the size of the detector, which means that the signal-to-noise ratio is small. Cosmic-ray interactions with particles in the atmosphere will produce a broadband spectrum of γ-rays that is orders of magnitude more intense than Galactic sources. This atmospheric emission, referred to as albedo radiation, is strongest near the horizon and below a balloon-borne instrument. Furthermore, the nuclear activation lines that are produced in interesting cosmic sources can also plague the instrument as background lines. After high-energy cosmic-ray collisions occur within the instrument mass, which is inevitable at balloon-altitudes and in satellite orbits, the material around the detector will become activated and the radiation from these instrumental nuclear decays will become an additional background source.

A number of lessons about the reduction of background radiation, which drives the sensitivity of most instruments, have been learned from previous MeV telescope designs (Schönfelder, 2004): anti-coincidence shields (ACSs) surrounding the detector can significantly reduce the background at balloon altitudes; low-activation materials near the detector element will reduce instrumental background; and discrimination of neutrons and cosmic rays, either in electronics (pulse-shape identifiers), through vetoing, or in the software analysis chain can lower the dead-time and false identification of γ-ray signals. Additionally, detailed simulations for a proper benchmarking of the instrument is essential to understand the source and background response of the telescope.

The studies of the 511 keV emission over the past 50 years, reviewed in Section 2.3, show the power of γ-ray line astronomy, but also warn of the difficulties in the analysis. Systematic uncertainties in the background can swamp potential results or give a false signal of variability. Imaging is also particularly difficult in this range, requiring complex iterative deconvolution techniques with non-unique solutions to extract the source structure (Section 3.2.4).

3.1.2 SOFT γ-RAY IMAGING TELESCOPE TECHNOLOGIES

There are two telescope classes with high technology readiness levels (TRL) that can be used to image soft γ-rays: Compton telescopes and coded-aperture mask telescopes. These two technologies utilize the two interaction types that

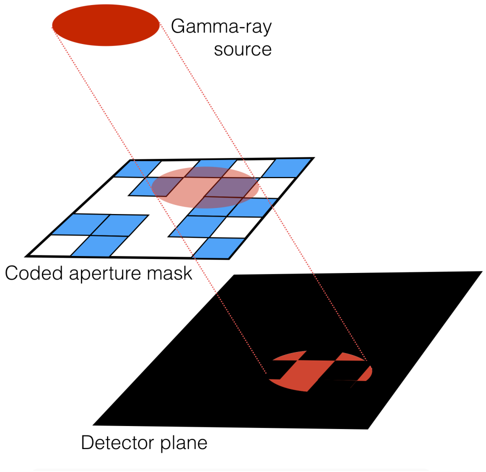


Figure 3.3: Schematic diagram of coded-mask imaging. The coded-aperture mask is made with a high-Z material with a unique set of holes so that a different shadow is created for every source sky position. (Figure adapted from B. J. Mattson, L3/NASA/Goddard Space Flight Center.)

dominate in these energy ranges, namely Compton scattering and photoelectric absorption (see Figure 3.2).

Compton telescopes use the energy and position of one or more Compton scatters in the detector volume to determine the path of the γ-ray and reconstruct the original direction of the photon. A detailed overview of Compton telescopes will be presented in Section 3.2.

Coded-aperture mask telescopes use a dense material with a unique set of holes to cast shadows on a position-sensitive detector plane, and the shadows can be used to determine the source direction. See Figure 3.3 for a schematic diagram of the coded-aperture mask imaging concept. Coded mask telescopes have been used widely in space environments. A related technology is Rotated Modulation Collimators, where the front mask spins to create a temporal variation on the detector plane (Smith et al., 2004). These technologies work well for point-source detection; however, the indirect imaging technique is difficult for diffuse emission and often requires model fitting. Coded-aperture mask telescopes for γ-ray astrophysics are reviewed in Caroli et al. (1987).

An additional technology that has been in development over the past few decades is a γ-ray condenser called a Laue lens (Smither, 1982), named after the physicist who discovered X-ray diffraction in crystalline structure. Laue lenses use Bragg diffraction to focus γ-rays onto a small detector plane; see

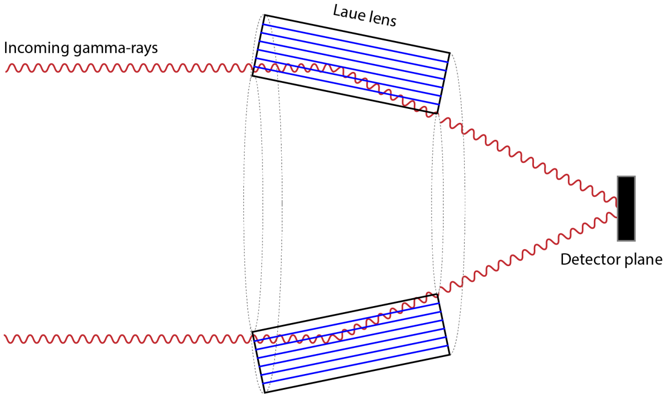


Figure 3.4: A schematic of the γ-ray condenser called a Laue lens. Gamma-rays of a specific energy will Bragg diffract off of the crystalline structure within the lens and be focused onto the small detector plane.

Figure 3.4 for the diffraction and focusing concept of Laue lenses. The main advantages of Laue lenses are the low background due to the small detector area and the achievement of arc-second angular resolutions (Frontera & von Ballmoos, 2010); however, the technology is still very much in early developmental stages.

The advantages and disadvantages of the three technologies for soft γ-ray imaging are summarized here (satellite-borne instruments which use the technology are listed):

- *Compton telescopes* have large FOVs, have good background reduction (with the help of software post-processing), are sensitive to polarization, and are a form of single-photon detection; however, they suffer from an inherently poor angular resolution.
  - CGRO/COMPTEL, *Hitomi*/SGD[2]
- *Coded-aperture mask telescopes* also have large FOVs and can have better angular resolutions than Compton telescopes; however, they have high instrumental backgrounds, limited efficiency, and only work as indirect imagers.
  - INTEGRAL/SPI, *Swift*/BAT, BeppoSAX/WFC, *AstroSat*/CZTI, etc.

---

[2] *Hitomi* was launched by JAXA in 17 February 2016; however, the satellite was lost on 26 March 2016 after an issue with the attitude control system. The Soft Gamma-ray Detector (SGD) was a Compton telescope with an energy range of 40-600 keV and would have significantly progressed the field of soft γ-ray astrophysics (Watanabe et al., 2012).

- *Laue lenses* have low backgrounds, due to small detector planes, and impressive angular resolutions; however, they have small FOVs, small bandpasses, and the current focal lengths are $\gtrsim$10 m.
    - TRL 4-5

3.2 COMPTON TELESCOPE BASICS

Compton telescopes use the interaction position and energy deposited in sequential Compton scatters to determine the initial photon's energy and source sky position (von Ballmoos et al., 1989; Boggs & Jean, 2000). The first Compton telescope to be launched on a satellite platform was COMPTEL aboard CGRO in 1991. COMPTEL is an example of a classic Compton telescope, which uses two separated planes, one to scatter the γ-ray and one to absorb it. The time-of-flight information between the two planes determines the direction of motion of the γ-ray and helps to suppress the dominating atmospheric background. After decades of technological advancements, compact Compton telescopes (CCTs) allow for higher efficiency in a small instrument volume. In a CCT there is only one active volume that acts as the scattering and absorbing material together.

The Compton scatter of a γ-ray off an electron is described by the Klein-Nishina differential cross section (Klein & Nishina, 1929):

$$\left(\frac{d\sigma}{d\Omega}\right)_C = \frac{r_e^2}{2}\left(\frac{E_f}{E_i}\right)^2\left(\frac{E_f}{E_i} + \frac{E_i}{E_f} - 2\sin^2\phi\cos^2\eta\right), \tag{3.1}$$

where the initial and final energy of the γ-ray is E_i and E_f, respectively, r_e is the classical electron radius, ϕ is the Compton scatter angle, and η is the azimuthal scatter angle. From this equation it can be seen that higher energy photons will, in general, have smaller Compton scatter angles, and lower energy photons will result in larger scatter angles. Additionally, if the incident γ-rays have an initial linear polarization, the photons will predominantly scatter at 90° relative to the the initial photon's electric field vector, defined as $\eta = 0$. Therefore, Compton telescopes can inherently detect polarization. In Equation 3.1, the electron is assumed to be unbound and at rest, and consequently this is only an approximation for interactions in a Compton telescope.

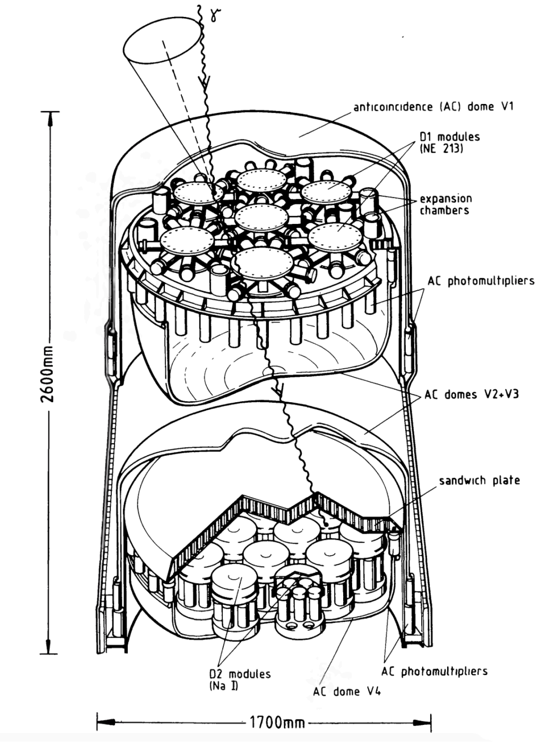


Figure 3.5: A schematic view of COMPTEL. A γ-ray will Compton scatter in the low-Z top detector plane and photoabsorb in the high-Z bottom detector plane. The direction of scatter is determined by the measured positions within the detector planes, and the Compton angle of the scatter is determined by the measured energies. The direction of the incoming γ-ray can be reduced to a cone whose axis is defined by the scatter direction and the opening angle is the Compton angle. Figure 2 of Schönfelder et al. (1993).

3.2.1 CLASSIC COMPTON TELESCOPES

The COMPTEL instrument consisted of two detector arrays: a plane of low-Z liquid scintillators (NE 213A) placed 1.5 meters above a plane of high-Z NaI scintillators (see Figure 3.5). All classic Compton telescopes use two separate detector planes, the first as the scattering material and the second as the absorbing material, where they are spaced far enough away for time-of-flight information to confirm downward-going events.

If a γ-ray Compton scatters in the first layer and deposits its remaining energy in a photoabsorption interaction in the second layer, then the incoming direction of the γ-ray can be constrained to a circle on the sky. The measured position of the Compton scatter and the photoabsorption defines the direction of the scattered photon, as shown in Figure 3.5. With an accurate measure of the energy deposited in the scattering plane (E_s) and the absorbing plane (E_a), the total energy of the γ-ray can be obtained ($E_\gamma = E_a + E_b$), and the angle of the Compton scatter ϕ can be determined with the Compton equation:

$$\cos\phi = 1 - \frac{m_e c^2}{E_a} + \frac{m_e c^2}{E_s + E_a}. \tag{3.2}$$

The Compton scatter angle then defines the cone on the sky whose axis is defined by the direction of the scattered γ-ray. The cone is known as a *Compton cone* and its projection on the celestial sphere as an *event circle*. When multiple photons from the same source interact in the detector, the resulting event circles will overlap at the source sky position and iterative deconvolution techniques can be used to create an image (Wilderman et al., 1998); this will be further discussed in Section 3.2.4.

Although COMPTEL was a hugely successful mission opening the MeV skies for the first time, it suffered from many limitations. The segmented detector modules had poor position resolution (~40 cm^3) and the scintillator detectors had poor energy resolution (5-10% FWHM). The separated planes restricted the FOV and limited range of Compton scatter angles, which in turn limited the efficiency to only ~1% (Schönfelder et al., 1993).

3.2.2 COMPACT COMPTON TELESCOPES

In a compact Compton telescope (CCT), the scattering and absorbing detector planes are reduced to a single active volume. A γ-ray will typically scatter a few times, the exact number depends on the photon energy and atomic number of the detector material, and will lose energy with each scatter before finally stopping with a photoabsorption interaction. Figure 3.6 shows a schematic of a typical event sequence in a CCT, where the blue rectangular area depicts the active volume of a 3D position-sensitive detector.

The original energy of the photon is given by the sum of the deposited energies, e.g., $E_\gamma = E_1 + E_2 + E_3$ for a 3-site event as shown in Figure 3.6. Unlike the classic Compton telescopes, the CCTs can track multiple scatters, so the Compton scatter angle ϕ of the first scatter can be determined from the kinematics of the interaction; specifically, for a 3-site event:

$$\cos\phi = 1 - \frac{m_e c^2}{E_2 + E_3} + \frac{m_e c^2}{E_1 + E_2 + E_3}. \tag{3.3}$$

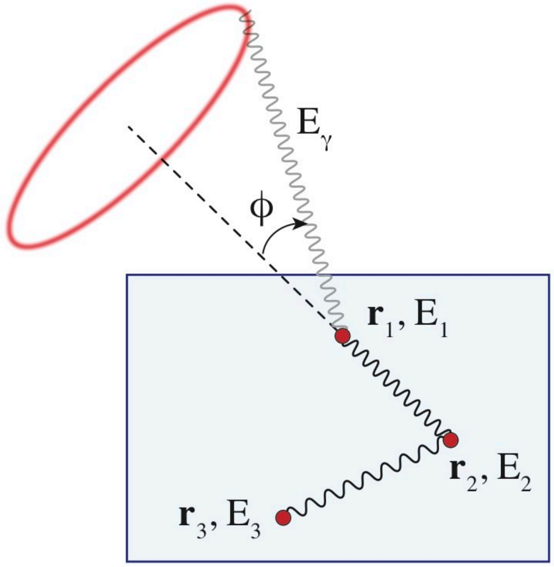


Figure 3.6: A typical event in a CCT. The incoming γ-ray Compton scatters a number of times before ending with a photoabsorption event. If the event is completely absorbed, the energy of each scatter can be summed to find E_γ and the Compton scatter angle ϕ of the first interaction can be found. The incoming γ-ray direction can then be constrained to a cone whose axis is defined by the first scatter direction between $\mathbf{r_1}$ and $\mathbf{r_2}$, and the opening angle is the Compton scatter of the first interaction. The width of the event circle is the angular resolution of the telescope.

The initial photon scatter direction, between sites 1 and 2, defines the axis from which ϕ is taken, as described in Section 3.2.1 and depicted in Figure 3.6.

The major benefits of CCTs are that one can track 2+ site events and there are no geometric limitations on the Compton scatter angle; both features increase the efficiency of the instrument and its polarization capabilities. CCTs can be solid-state detectors, like those used in COSI, or time projection chambers, as long as a precise measure of the interaction position and energy can be determined. There is a sub-class of Compton telescopes that are referred to as *tracking* Compton telescopes when the recoil electron can be traced. If the recoil direction from the Compton scatter is known, then the event circle can be reduced to an arc, and the origin of the γ-ray can be more accurately determined which improves background rejection.

There is an added complication when dealing with compact Compton telescopes: the time-of-flight between successive scatters in the instrument is a couple orders of magnitude less than the timing resolution achievable (in GeDs, Δt_{res} ~10 ns); therefore, the relative order of Compton scatters within an event is unknown. Compton Kinematic Discrimination (CKD) uses redundant information in the energy deposits and angles of 3+ site events to determine the most likely interaction sequence (Boggs & Jean, 2000). The process of determining the most probable sequence of interactions is also referred to as *Compton reconstruction* or *event reconstruction* and will be discussed further in Section 3.2.3.

There are now several groups around the world developing compact Compton telescopes with the aim to fill the MeV sensitivity gap (Kurfess et al., 2000; Takeda et al., 2007; O'Neill et al., 2003; Ueno et al., 2012). COSI remains one of the furthest developed and most rigorously tested compact Compton telescope to date.

3.2.3 COMPTON EVENT RECONSTRUCTION

In a CCT, all of the scatters from a single γ-ray event occur "simultaneously" since the time resolution of the detector is greater than the time its takes light to cross the detector volume. Therefore, the true sequence of interactions within the detector must be determined from physics principles and statistics in a process called event reconstruction.

For events that consist of 3+ interactions, each of the $N!$ possible combinations of the interaction order is analyzed to determine the most probable sequence. Since there is redundant information in the energies and positions, the Compton scatter angle of the central interaction(s), denoted by l, can be determined in two ways: kinematically with the Compton equation

$$\cos\phi^{kin} = 1 - \frac{m_e c^2}{E_{l_+}} + \frac{m_e c^2}{E_l + E_{l_+}}, \tag{3.4}$$

and geometrically considering the scatter angles

$$\cos\phi^{geo} = \frac{\vec{g}_k \cdot \vec{g}_l}{|\vec{g}_k||\vec{g}_l|}. \tag{3.5}$$

For the central interaction l, E_{l_+} is the total energy of all the interactions after l, $\vec{g}_k$ is the incoming γ-ray direction and $\vec{g}_l$ is the outgoing direction. See Figure 3.7 for a schematic description of the variables used in these equations. The two measures of $\cos\phi$, from Equation 3.4 and 3.5, should be identical for the correct order of interactions in an ideal instrument. In the classic

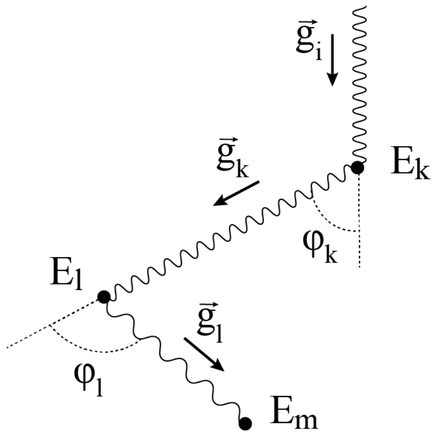


Figure 3.7: An example 3-site interaction with labeled variables used in event reconstruction. The Compton scatter angle ϕ_l of the central interaction can be determined with redundant information, either kinematically with the energies E or geometrically with the scatter directions $\vec{g}$. Minimizing the difference between these two measures for each Compton scatter leads to the most probable sequence of interactions.

CKD approach, a χ^2 quality factor Q is assigned to each permutation of scatters (Zoglauer, 2005)

$$Q = \sum_{i=2}^{N-1} \frac{(\cos\phi_i^{kin} - \cos\phi_i^{geo})^2}{\Delta\cos^2\phi_i^{kin} + \Delta\cos^2\phi_i^{geo}}, \tag{3.6}$$

where $\Delta\cos\phi_i$ are the measurement errors and i is the interaction index. The sequence with the lowest quality factor is the best estimate of the correct kinematic ordering of the event.

For 2-site events, with one Compton scatter and a photoabsorption, the order of only a small fraction of events can be determine unambiguously. For all other events, one can utilize the Klein-Nishina equation (Equation 3.1) to determine which hit order is more probable. Additionally, the probability of photoabsorption for a certain distance between interactions can be taken into account.

One of the strengths of Compton reconstruction is that it allows for the high background suppression necessary to attain reasonable sensitivities in the MeV regime. Background events or incompletely absorbed events will often result in poor quality factors; therefore, only selecting events with a low quality factor will reject most of the background signal. Currently, the COSI collaboration is developing new methods of event reconstruction involving machine learning algorithms that can more accurately determine the correct sequence of interactions and better reject background events.

Once the true sequence has been determined, the position and energies of the second and higher interactions are used to find the initial Compton scatter angle and the total energy of the event, as described in Section 3.2.2. These parameters can then be used as additional indicators of the quality of the event, which will be further discussed in Section 3.5.5.

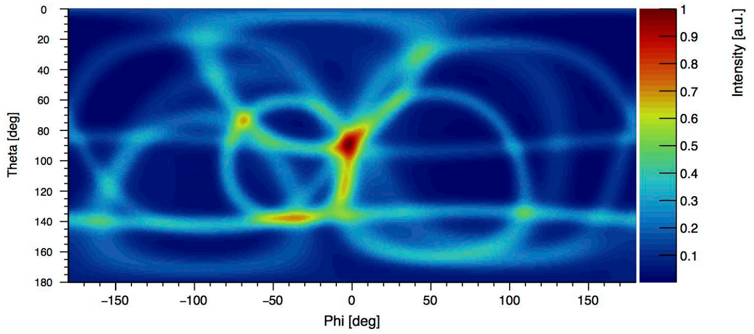
3.2.4 IMAGE RECONSTRUCTION

The Compton angle of the first interaction describes the circle on the sky from which the photon is known to originate. If multiple γ-rays are detected from the same source, then the event circles will overlap at the source sky position, if they are properly reconstructed. The source position can be seen with a simple back projection of these event circles on the sky. See Figure 3.8 for an example of the back projection of event circles from a simulation of 40 γ-ray events (a), and 200 events (b). The red hot spot in the center of the image where the circles overlap reveals the simulated source location.

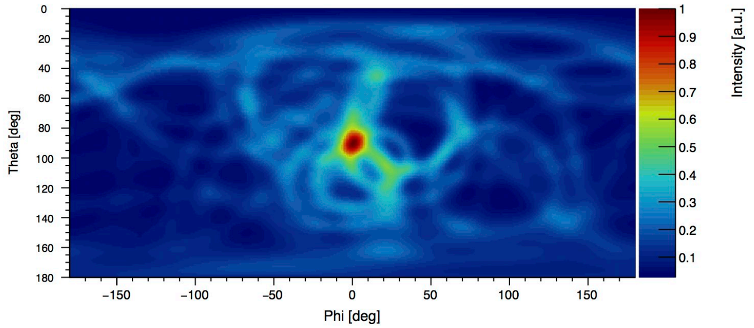
The data measured by a Compton telescope is generated by convolving the sky distribution with the instrument response and the addition of background. Image reconstruction is the process of inverting the measurement operation to recover the source distribution. Since there is no unique solution for these images, iterative deconvolution techniques are used.

The imaging data space for CCTs can contain a huge number of bins and is sparsely populated, which make computation difficult. Therefore, most imaging reconstruction for CCTs has been done with list-mode algorithms. Instead of keeping the event information in a binned data-space, a list of events and all of their parameters is stored and the size of the data space is proportional to the number of events.

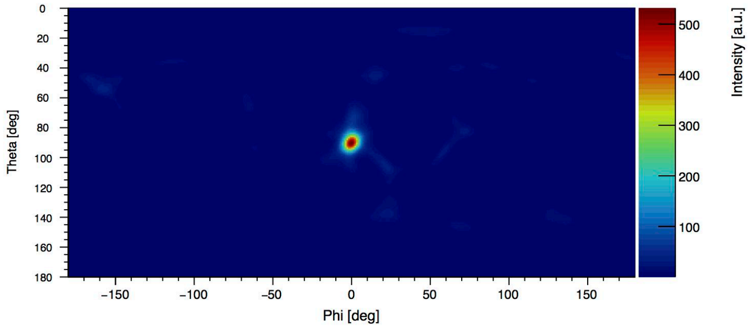
The major advantage of list-mode image reconstruction is that all of the event information is kept with full precision. Additionally, information that is not traditionally in the image data space, like the position of the first interaction, or the distance between two scatters, can be saved and used to further select the quality of the events. The imaging algorithm used for COSI is the list-mode maximum-likelihood expectation-maximization (LM-ML-EM) (Wilderman et al., 1998; Zoglauer et al., 2011). See Figure 3.8 (c) for an example of an image of a simulated point source after 5 iterations of the LM-ML-EM algorithm where the point source is clearly recovered. List-mode imaging, how-



(a) Back projection of a point source simulation with 40 events.

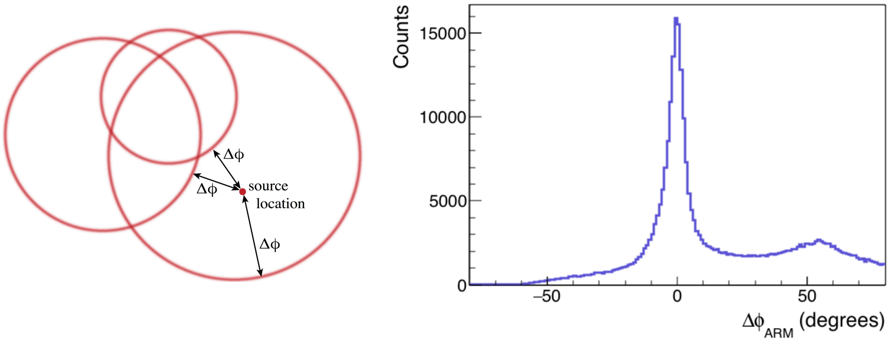


(b) Back projection of point source simulation with 200 events.



(c) Image obtained after 5 iterations of the LM-ML-EM algorithm.

Figure 3.8: (a) The back projection of a 500 keV point source simulation with COSI with only 40 events. The event circles overlap at the location of the source in the center of the image shown with the red hot spot. (b) Same as in (a) except with 200 events; the location of the source becomes much more prominent. (c) After 5 iterations of the LM-ML-EM imaging algorithm on the image in (b) the point source is reconstructed.



(a) The ARM $(\Delta\phi)$ defined for three event circles.

(b) Histogram of ARM values.

Figure 3.9: (a) The smallest angular distance between the known source location and each Compton event circle is the ARM $\Delta\phi$. The ARM is defined as negative when the source is inside the event circle. (b) An example ARM histogram from a ^{22}Na calibration measurement with COSI. The total ARM histogram from a sample of Compton events is the effective point spread function of telescope. The FWHM of the ARM distribution, which is 6° here, defines the angular resolution of a Compton telescope and is a measure of the width of the event circle.

ever, has a few drawbacks: it is difficult to recover the true flux of the source, and the ML-EM iterative deconvolution is optimized for point sources. Motivated by the 511 keV studies presented here, COSI is currently implementing a binned-mode image reconstruction algorithm which can account for exposure, give an accurate measure of flux, estimate background, and image extended sources (Zoglauer et al., in prep).

3.2.5 ANGULAR RESOLUTION MEASURE

The angular resolution of a Compton telescope is described by the angular resolution measure (ARM). The ARM $(\Delta\phi)$ is the smallest angular distance between the *known* source location and the event circle for each event, as shown in Figure 3.9 (a). The distribution of all ARM values from a sample of Compton events represents the effective width of the point spread function of a Compton telescope; the FWHM of the ARM distribution defines the achievable angular resolution after event reconstruction. For the calibration measurement of the 511 keV line from ^{22}Na shown in Figure 3.9 (b), for example, the ARM FWHM gives an angular resolution of 6°. A positive $\Delta\phi$ occurs when the known source location is outside of the event circle, and this is generally due to an incompletely absorbed recoil electron or incorrect path due to an escaped photon. A negative $\Delta\phi$ occurs when the source location is inside the event circle and results mostly from an incompletely absorbed γ-ray.

More precise measurements of the energy and position will give a better ARM; however, Compton telescopes have a fundamental limit on angular resolution. The use of the Compton scatter equation (Equation 3.3) assumes the electron with which the γ-ray interacts is free and at rest. If one takes into account that the electron is bound and has some momentum, then the energy distribution between the γ-ray and recoil electron will change. This means there will be a broadening in the measured Compton scattered γ-ray energy and in turn this limits the angular resolution. This limitation is referred to as "Doppler broadening" and is dependent on the atomic number of the detector material and the energy of the γ-ray. For germanium, the angular resolution is limited to 1.25° at 500 keV (Zoglauer & Kanbach, 2003).

3.3 THE COSI INSTRUMENT

The Compton Spectrometer and Imager mission has been in development for the past decade through a collaboration between the Space Sciences Laboratory (SSL) and the Lawrence Berkeley National Laboratory (LBNL) at the University of California, Berkeley; IRAP in Toulouse; and several Taiwanese universities led by the National Tsing Hua University. The instrument has been designed with high sensitivity, excellent spectral resolution, and moderate angular resolution, in an effort to perform novel studies of astrophysical sources while testing the technologies and analysis techniques necessary for future missions, e.g., COSI-X.

3.3.1 3-D POSITION-SENSITIVE GERMANIUM DETECTORS

The heart of COSI consists of twelve HPGe cross-strip detectors developed using LBNL's amorphous germanium contact technology (Luke et al., 1992; Amman et al., 2007). Each GeD is fabricated by first depositing a thin layer of amorphous germanium over the entire surface of an $8 \times 8 \times 1.5$ cm HPGe crystal. This layer has been shown to serve as an effective bipolar blocking contact. The aluminum strip electrodes are then evaporated onto the sur-

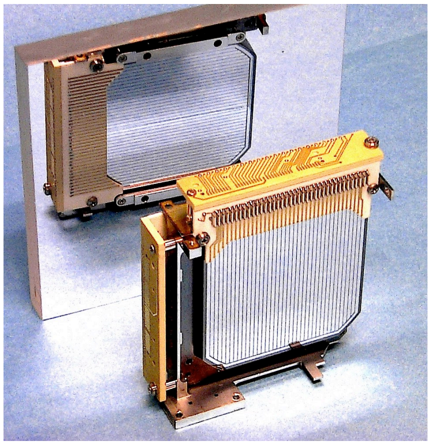


Figure 3.10: Single COSI GeD with detector boards and aluminum mounts. The 37 strips on the front face of the detector visible in the vertical direction are orthogonal to the strips on the back face of the detector, which are visible through the mirror.

face, as these give the detector its position sensitivity (Amman & Luke, 2000; Coburn et al., 2003). Each of COSI's twelve GeDs has 37 strips with 2 mm strip pitch deposited on the anode and cathode in an orthogonal orientation. Figure 3.10 is a photograph of a COSI GeD which shows the 37 vertical aluminum electrodes on the front face, and the horizontal electrodes on the back face are seen through a mirror.

Good angular resolution through Compton reconstruction requires high energy resolution and precise 3D position measurements. Germanium inherently provides an excellent energy resolution; however, this can be significantly degraded when using a detector with segmented electrodes. The strip pitch and gap between neighboring strips have been optimized during the development of these detectors to reduce the charge loss, one of the main spectrum-degrading effects (Coburn et al., 2003). The position sensing occurs in two ways: the X and Y positions are given by the intersection of the activated anode and cathode electrodes, and the Z-positioning is obtained from the difference in the arrival times of the electrons and the holes at their respective electrodes (for details, see Section 4.5). A 2 mm wide guard ring surrounds the 37 strips on each side and vetoes interactions close to the edge of the detector, where fringes in the electric field and high leakage currents can degrade the detector response. The COSI GeDs have demonstrated good position resolution (2 mm^3) while maintaining the excellent spectral resolution of germanium (~0.2-1%).

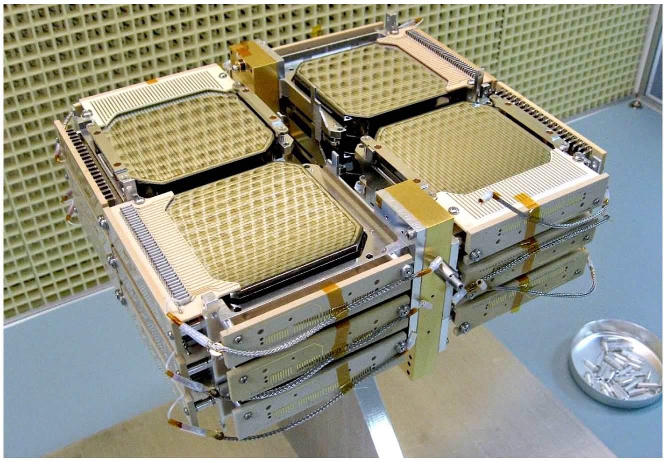


Figure 3.11: Photograph of the COSI detector array before integration. All detectors are individually mounted to an anodized U-shaped copper coldfinger which is visibile between the detectors. The detectors together have a total active volume of 972 cm^3. Care has been taken in the design to minimize the passive material between the individual GeDs.

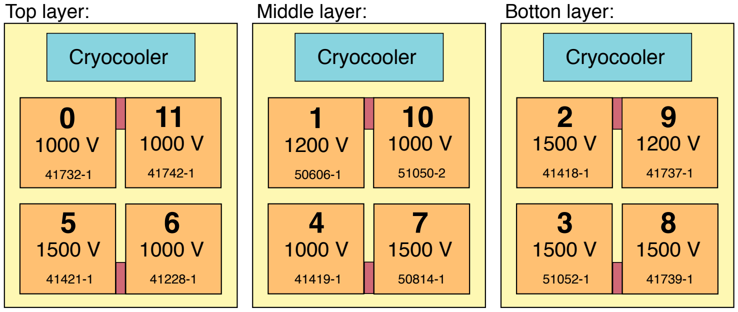


Figure 3.12: The position of each detector within the GeD array with the detector operating voltage and crystal ID listed. The location of the coldfinger, to which each of the GeDs is individually mounted, is shown in red.

The twelve GeDs are stacked in a $2 \times 2 \times 3$ configuration and have a total active volume of 972 cm^3; see Figure 3.11 for a photograph of the twelve GeDs. The detectors are fully depleted and operate at voltages ranging from 1000-1500 V. The readout electronics on the high-voltage side of the detectors are AC-coupled, and as such, the cathode is often referred to as the *AC side* and the anode is referred to as the *DC side*. A list of the 14 detectors (2 spares) developed for the COSI mission and some of their characteristics are listed in Table 3.1. The relative detector positions are shown in Figure 3.12.

Detector Number	Crystal ID	Type	Crystal Thickness [mm]	Depletion Voltage [V]	Operating Voltage [V]	Defects
0	HP41732-1	n	14.9	600	1000	Slow strip leakage increase with time
1	HP50606-1	p	14.5	800	1200	Thin groove region (cosmetic)
2	HP41418-1	p	15.0	700	1500	Poor electron trapping
4	HP41419-1	p?	15.1	600	1000	
5	HP41421-1	p	15.0	1000	1500	Third wing chips (cosmetic)
6	HP41228-1	p	14.8	600	1000	Third wing chips and circular etch feature (cosmetic)
7	HP50814-1	p	14.7	400	1500	
8	HP41739-1	p	14.9	1200	1500	
9	HP41737-1	n	14.7	700	1200	Top edge strip indicates surface channel
10	HP51050-2	p	14.2	150	1000	
11	HP41742-1	n	14.5	250	1000	Slow strip leakage increase with time
Spare	HP41228-2	p	14.7	450	1000	Damaged during reprocessing after COSI'14
Spare	HP41418-3	p?	15.1	400	800	High GR leakage above 800 V

Table 3.1: Physical and electronic details for the 14 COSI GeDs provided by Mark Amman, who fabricated the detectors. The Detector Number is based on the crystal location within the cryostat, where the Crystal ID is the name given by the manufacturer. The type of impurity in the crystal will change the classification of the semiconductor, n- or p-type; a question mark indicates the type could not be definitively measured.

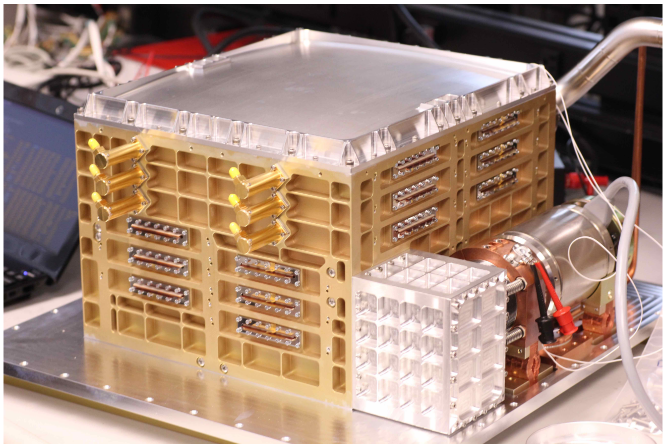


Figure 3.13: Photograph of the COSI cryostat. The mechanical cryocooler is seen on the right-hand side of the cryostat. No read-out electronics are attached to the cryostat, but three rows of Kapton-manganin flex circuit feedthroughs are visible on each side. Six of the high-voltage feedthroughs are seen on the front-left side of the cryostat.

3.3.2 CRYOSTAT AND CSI SHIELDS

The array of GeDs is housed in an anodized aluminum cryostat, as shown without any of the readout electronics in Figure 3.13. Within the cryostat, the twelve detectors are surrounded by a thin infrared shield and multi-layer insulation. The cryostat is evacuated to ~ 10^{-6} Torr and is maintained at low pressure through the use of activated charcoal.

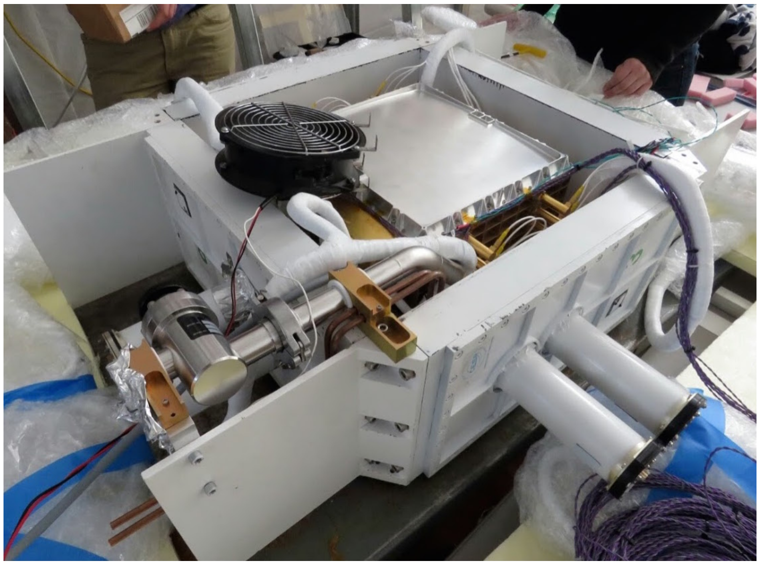
One disadvantage of using germanium detectors is the low operation temperature of < 100 K. Conventionally, liquid nitrogen (LN2) is used to cool GeDs; however, the finite amount of LN2 that can realistically be launched on a balloon payload greatly constrains the flight duration. To enable COSI to fly on ultra-long duration balloon (ULDB) flights, the system was designed to use a Sunpower CryoTel CT mechanical cryocooler, completely eliminating the flight duration constraints: as long as the cryocooler has power, the detectors can remain cold. The coldtip of the cryocooler is maintained at a constant temperature of 77 K through a temperature-controlled feedback loop. The detectors are individually bolted to a copper coldfinger, which is then thermally coupled to the tip of the cryocooler. We observe a seven degree temperature gradient between the tip of the cryocooler and the coldfinger, and as such, the operation temperature of the GeDs is maintained at ~84 K.

The cryostat is surrounded on five sides by CsI scintillator blocks instrumented with photomultiplier tubes (PMTs), as shown in Figure 3.14. The CsI blocks, which are 30 lb each, act as a passive shield by absorbing a large fraction of the incident γ-rays originating from non-astrophysical sources, i.e., the earth's albedo, effectively constraining the field-of-view to 25% of the sky. The scintillators also function as an ACS: GeD events with a CsI coincidence signal are vetoed to reject incompletely absorbed events. The CsI shields play an important role in suppressing the background detected in COSI, which directly affects the sensitivity of the instrument.

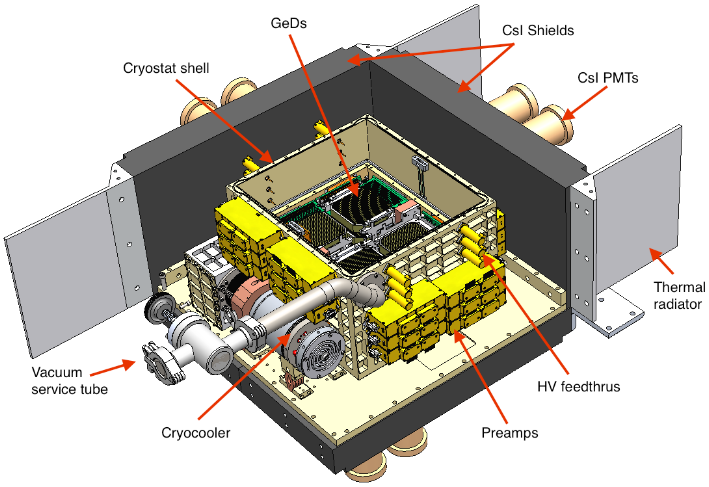
The background radiation is further reduced through the use of a graded-Z shield placed above the cryostat. Graded-Z shields, which are thin layers of material with decreasing atomic number, have been shown to reduce the effects of ionizing radiation in space environments (Rojdev et al., 2009). The top high-Z layer is efficient at blocking low-energy γ-rays, but produces X-ray fluorescence. Each lower-Z layer absorbs the X-rays emitted in the previous layer and produces X-rays of lower energy until an acceptable level is reached. For COSI, the graded-Z shield consists of a 0.22 mm thick layer of tin and a 0.4 mm layer of copper. In general, aluminum is included as a third layer, but the aluminum top of the cryostat, with an average thickness of 0.75 mm, serves as the final COSI graded-Z layer.

3.3.3 ELECTRONIC READOUT

Each strip of COSI (37 strips × 2 sides × 12 detectors = 888 strips) has an individual readout channel consisting of amplifying and shaping analog electronics. The strip electrode signals (AC-coupled at the cathode and DC-coupled at the anode) are fed through the cryostat walls on Kapton-manganin flex circuits and are coupled to low-power, low-noise preamplifiers (Fabris et al., 1999), which are mounted on the side walls of the cryostat; see Figures 3.14 (b) and 3.15 (a). The output from each preamplifier is fed into custom-made pulse-shaping amplifier circuits that consist of two channels: a fast channel for timing and a slow channel for energy. The fast-timing channel, with a 40 keV threshold, uses a bipolar shaper with a 170 ns rise time to accurately define the interaction time. The slow-energy channel, with a 20 keV threshold, uses a unipolar shaper with a slow 6 μs shaping time for noise reduction and accurate pulse height determination. There is one "card cage," shown in Figure 3.15 (b), assigned to each detector that provides the high-voltage power to the GeD, delivers the power for two preamplifier boxes, and con-

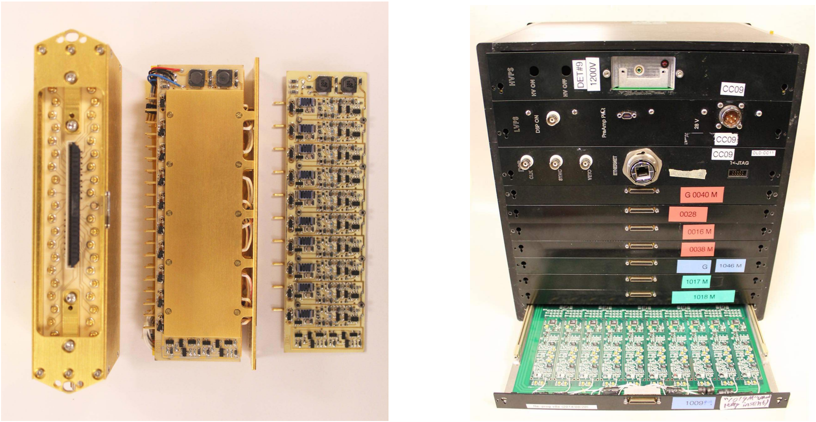


(a) Photograph of COSI cryostat with shields attached.



(b) `SolidWorks` drawing of COSI cryostat.

Figure 3.14: (a) Photograph of the cryostat with four surrounding shields. The PMTs from the closest CsI shield are visible on the right side. The fan placed on top of the cryostat is used for cooling the cryocooler during ground operations before the liquid cooling system is installed. The graded-Z shield has been removed in the photo. (b) Labeled `SolidWorks` model of the cryostat in the same orientation as (a) with two of the side CsI shields removed. The top of the cryostat is made transparent so the GeDs are visible inside.



(a) (b)

Figure 3.15: (a) The low-noise, low-power preamplifier boxes, which are 1” thick, mount directly to the COSI cryostat. The full preamp box is shown on the left. The preamp boards (*right*) have 10 channels per board, and there are four boards per box (*middle*). (b) The COSI card cages contain the analogue pulse-shaping amplifier circuits, in addition to the high voltage supply and DSP board. One of the analogue boards in this photograph is pulled out of the box and shows the 10 channels per board. There is one card cage per GeD.

tains the analog shaping electronics for each strip on that detector. The card cages only trigger on events in which both an X and Y strip are activated.

A low-power, dual-core, single-board flight computer controls the basic operations of the gondola. It communicates with the card cages via ethernet, through which it receives science and housekeeping data, and sends out a 10 MHz clock signal for the event timing in the card cages. Raw data is stored on three redundant 1 TB flash drives and the flight computer does a rudimentary on-board analysis to parse out interesting science data, i.e. possible Compton events or triggered gamma-ray bursts (GRBs), for real time telemetry. The flight computer communicates with the cryocooler controller, CsI shield controller, GPS system, twelve card cages, and the rest of the gondola sub-systems.

3.3.4 GONDOLA AND OTHER SYSTEMS

The three-tiered aluminum frame structure of the COSI gondola is designed to be lightweight and simple for ULDB flights; see Figure 3.16 for a labeled picture of the entire gondola with the foam side panels removed. The cryostat and CsI shields sit in the top layer, where the large FOV of COSI circumvents any pointing requirements. All of the flight electronics, including the card cage electronics, flight computer, and power distribution, are mounted in the thermally-insulated middle tier. Three aspect systems enable us to accurately reconstruct our pointing throughout the flight. A Magellan ADU5 differential GPS is our main aspect system and a Trimble BX982 GPS and Applied Physics Systems Model 544 magnetometer are flown as backup systems for redundancy.

The Columbia Scientific Ballooning Facility (CSBF) provides the power system for the COSI flights. A peak power of 1500 W is provided by 15 SunCat Solar panels, placed at 34° orientation for the Wanaka, New Zealand, latitude. At a weight of ~600 lb, 24 Odyssey PC1100 batteries provide 480 Ah at 24V, which, with our average power consumption of ~450 W during flight, allows for 15+ hour night cycles. The power system is complete with a Charge Controller from MPPT by Morningstar Corporation.

Three types of telemetry through the CSBF Support Instrument Package (SIP) are used for in-flight commanding of the flight computer and data downlink. Two line-of-sight (LOS) L-band transmitters with a rate of 512 kbps each allow for high-throughput in the first few hours of flight. Two Iridium Pilot antennae using the Iridium Openport Satellite relay network are the main telemetry route for COSI with 130 kbps each. As a backup, the Iridium dial-up network can be used at a rate of up to 2 kbps. See Figure 3.17 for the placement of the flight antennae, highlighted in red, and the COSI GPS antennae, as well as the placement of the CSBF cameras. To ensure a uninterrupted data set, a program was developed to search the down-linked data and fetch any dropped packets. The average COSI science data rate was ~ 60 kbps in the 2016 flight, while count rates in the GeDs were on average ~50 cts/s.

3.3.4.1 THERMAL DESIGN

Like any space environment, the thermal environment for a balloon flights needs to be designed carefully. Through detailed thermal modeling of all gondola subsystems, one can predict the temperatures in flight and thus design insulation and thermal-links to keep all systems within their specified tem-

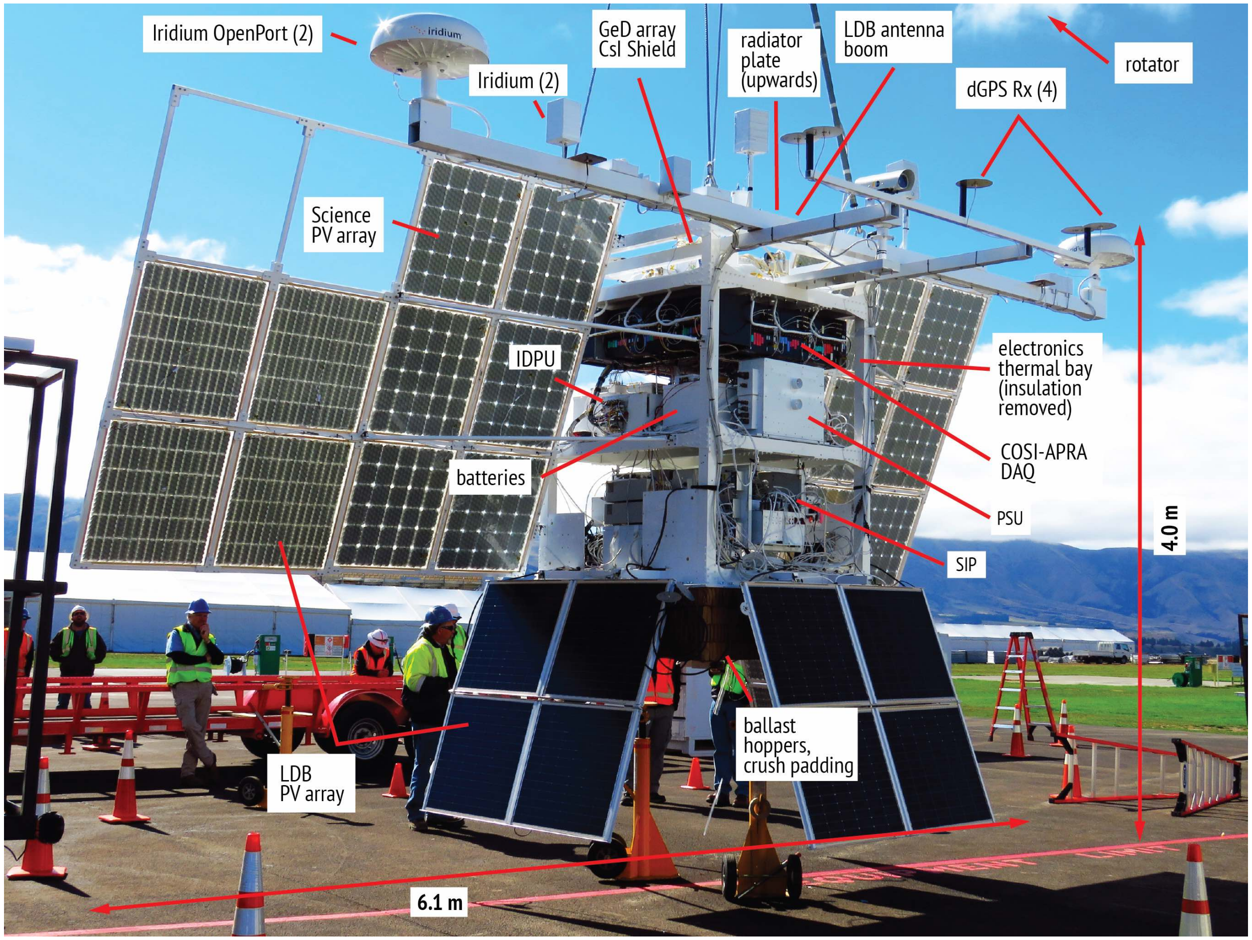


Figure 3.16: Labeled photograph of the COSI gondola before the 2016 flight from Wanaka, New Zealand. The insulating side panels are removed for this photograph to show the interior of the gondola. The total suspended weight of the 2016 COSI gondola without ballast was 4590 lb.

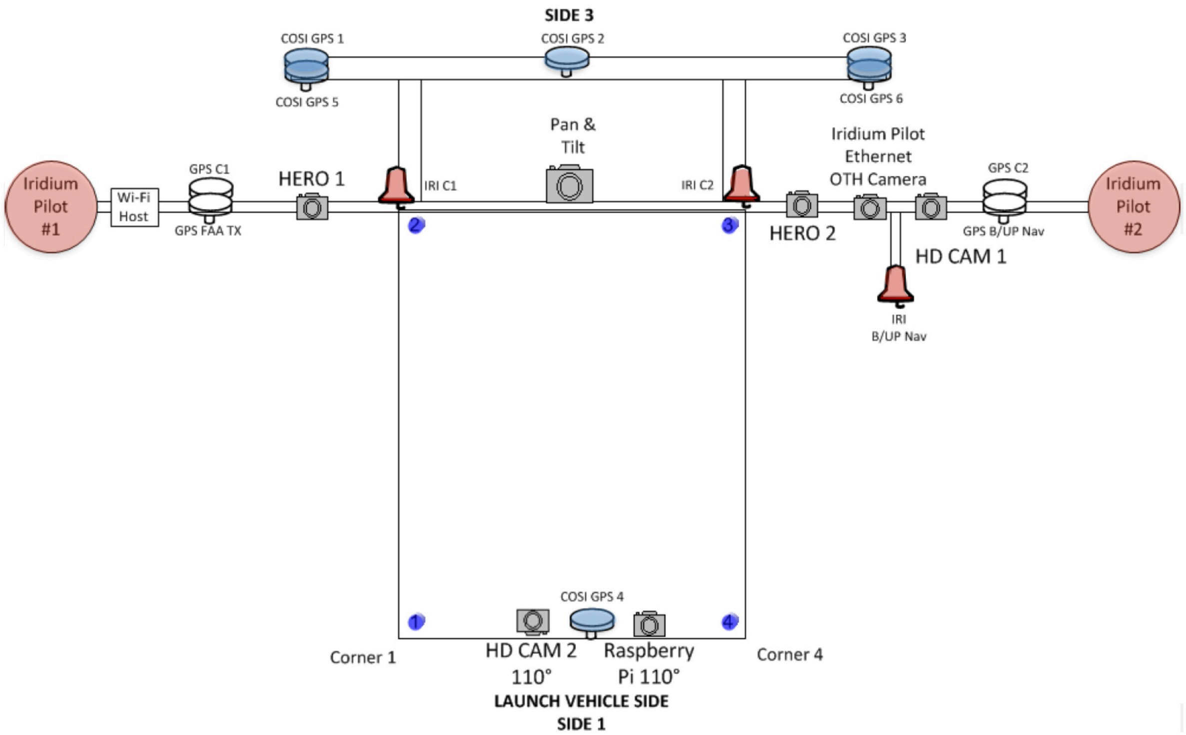


Figure 3.17: Layout of GPS antennas for the COSI 2016 flight, highlighted in blue (the magnetometer, which is mounted in Corner 4 is not shown). The redundant telemetry systems are highlighted in red. The Iridium Pilot antennae are the main telemetry route for the COSI flight, with a total data rate of 260 kbps. They are separated by 20' to reduce interference. The Iridium dial-up network uses the three cone-shaped antennae. The LOS antennae, which hang below the gondola, are not shown here.

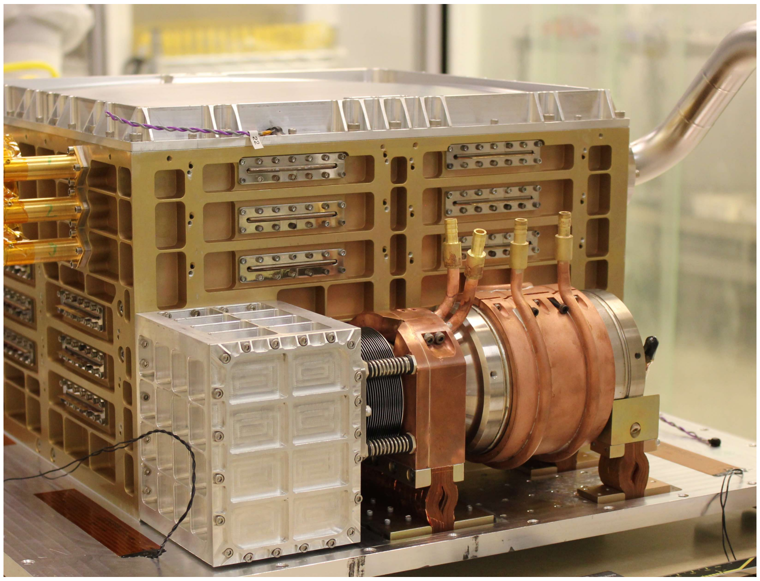
perature range. The range of ambient temperature for a mid-latitude flight can be between -100 to 80 °C, but most electronics operate at a minimum of -40 °C and a maximum of 70 °C. See Table 3.2 for an overview of the temperature limits for each of the COSI sub-systems. The COSI thermal design consists of passive foam insulation, CHO-THERM pads and thermal grease for better thermal conductivity, and heaters for critical components. There are 68 LM335 temperature sensors positioned around the COSI gondola and instrument to carefully monitor the temperature during operations.

The cryocooler dissipates a large amount of power and requires a more sophisticated heat rejection scheme than the radiation that all other components rely on. A liquid cooling system using 3M Novec HFE-7200 fluid is utilized to actively cool the cryocooler. The fluid is pumped through a custom-built crycooler collar and sleeve, shown in Figure 3.18 (a), and into a 1'×4' copper radiator where the heat dissipates. The radiator is mounted at the back of the gondola frame and is shown in Figure 3.18 (b). Mounted on the radiator is a 1.8 L liquid reservoir, and the liquid is transported to and from the cryocooler in flexible Tygon tubing. Two redundant Fluid-o-Tech FG200 pumps are mounted on the underside of the radiator and separated with a check-valve. The pumps have a tunable pumping speed that allows for a tuning of the heat-rejection efficiency. The performance of the active cooling system and the thermal design will be discussed in Chapter 5, which gives an overview of the 2016 COSI flight.

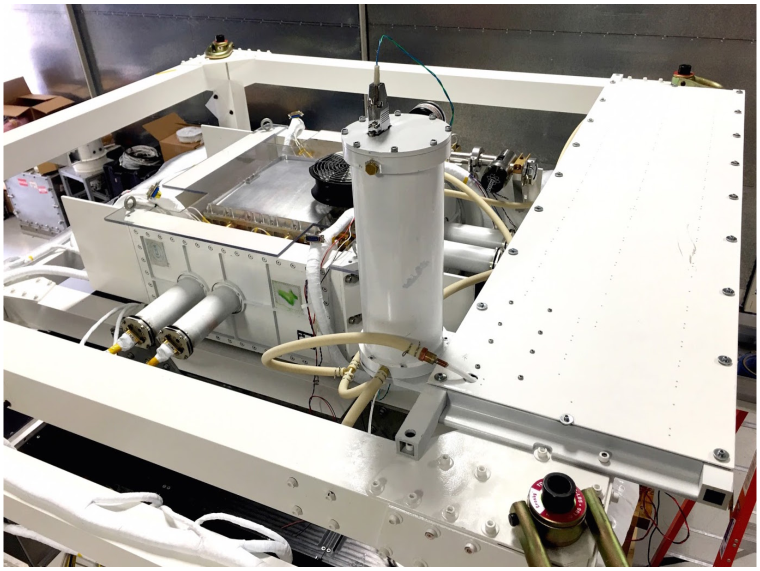
The top two tiers of the gondola are surrounded in Polystyrene insulating foam wrapped in aluminized mylar. The foam, which is 2" thick on the front and back (where front is defined as the sun-pointing side) and 1" thick on the sides and top of the gondola, entirely encloses the gondola frame. Beneath this foam, the gondola has powder-coated aluminum side panels that provide further thermal insulation and radio-frequency isolation to the electronics bay. The GPS antennae, the radiator fluid reservoir, and pumps are positioned outside of the gondola-enclosing foam box and have individual 1" thick foam boxes surrounding them. To better insulate a few temperature-sensitive components in the electronics bay, namely the flight computer and the ADU5 GPS control box, a 1" piece of foam was placed on top of the individual boxes, and additionally for the flight computer, a 1/4" layer of foam was placed underneath.

Component	Nominal	Temp. Limit		Heater	Thermostat
	Power [W]	Cold	Hot [°C]	[W]	Temp. [°C]
Flight Computer	36	0	55	18	10/15
PDU	10	-40	60	–	–
Ethernet Switch	6.4	-40	80	–	–
GPS - Power Box	2	-40	90	–	–
GPS - ADU5	6	-20	55	18	-30/-20
GPS - Trimble	3.8	-40	70	–	–
Magnetometer	0.3	0	70	–	–
Pump Box	9	-40	70	–	–
Pump	0.02	-40	70	36	-40/-30
Heater Box	0	-40	100	–	–
Cryocooler Plate	47	-40	70	–	–
Cryocooler	100	-20	70	–	–
Shield Box	11	-40	85	–	–
Shields	(6×0.35)	-50	45	–	–
Card Cages	(12×18)	-55	60	72	-35/-25
Preamps	(24×1)	-40	70	–	–
Cryostat	126	-20		70	N/A

Table 3.2: Power estimates and temperature limits for the COSI systems. The shields, card cages, and preamps have the number of units and individual power consumption written in parentheses. The 126 W listed for the cryostat is the sum of the power from the cryocooler, all of the preamps, and the shields. Kapton heaters with the available listed power were put on components that had high minimum temperature or were vital to the instrument operations. These heaters were turned on/off at the temperatures listed, with a thermostat-like software control.



(a) Copper cryocooler collar and sleeve for the active cooling system.



(b) Top of the COSI gondola showing the radiator and reservoir.

Figure 3.18: (a) Photograph of the cryocooler and cryostat immediately after integration. The copper sleeve and collar around the cryocooler have fluid lines for the active cooling system. (b) The radiator and cryostat integrated into the gondola before flight. The radiator mounts along the back of the gondola with fluid lines running to the crycooler collar and sleeve shown in (a). The reservoir is front and center, while the pumps controlling the liquid cooling system are mounted on the bottom of the radiator and not visible here.

3.4 PREVIOUS FLIGHTS WITH COSI AND NCT

COSI's precursor instrument, the Nuclear Compton Telescope (NCT), had two successful flights from Fort Sumners, New Mexico, in 2005 and 2009. During the 5.5 hour NCT prototype flight, in which only 2 GeDs were flown, the instrumental γ-ray background was measured (Bowen et al., 2006; Bowen, 2009). NCT was flown again in 2009 with 9 operational GeDs. The flight lasted 39 hours and resulted in a 4 σ detection of the Crab Nebula, the first detection of an astrophysical source by a compact Compton telescope (Bandstra et al., 2011; Zoglauer & Boggs, 2013).

A third flight of NCT was attempted in 2010, but the payload suffered a launch mishap; while the detectors and most of the readout electronics were unharmed, the gondola was completely destroyed. The NCT collaboration used this opportunity to redesign the instrument and optimize it for ULDB flights and polarization sensitivity. The upgraded design has twelve detectors and a modified configuration to increase the effective area. A mechanical cryocooler is now used to maintain low GeD temperatures instead of the conventional consumable LN2. The old BGO scintillator shields were traded in for CsI shields that were a better mechanical fit to the new cryostat design. The gondola was simplified with a fixed-zenith design. And finally, with these major upgrades came the new name: the Compton Spectrometer and Imager.

The newly designed COSI gondola was first flown from McMurdo Station, Antarctica, in December 2014. NASA's new Super Pressure Balloon (SPB) (Section 5.1.1) has been in development for decades, and this was the first flight that allowed a science instrument onboard as a Mission of Opportunity. It was expected that the flight would last for weeks and that the instrument would leave the continent and float north; however, a leak in the balloon caused the flight to be terminated after just 43 hours. The quick termination allowed for a recovery of the entire instrument before the end of the Antarctic summer season, which enabled the COSI team to have a quick turn-around before setting out on campaign again in early 2016.

3.5 MEGALIB ANALYSIS TOOLS

The analysis software used for COSI, its precursor NCT, and a number of other γ-ray instruments, is the Medium Energy Gamma-ray Astronomy library (MEGAlib) (Zoglauer et al., 2006). MEGAlib contains a suite of tools: `geomega` is used to create a realistic mass model of the instrument; `nuclearizer` con-

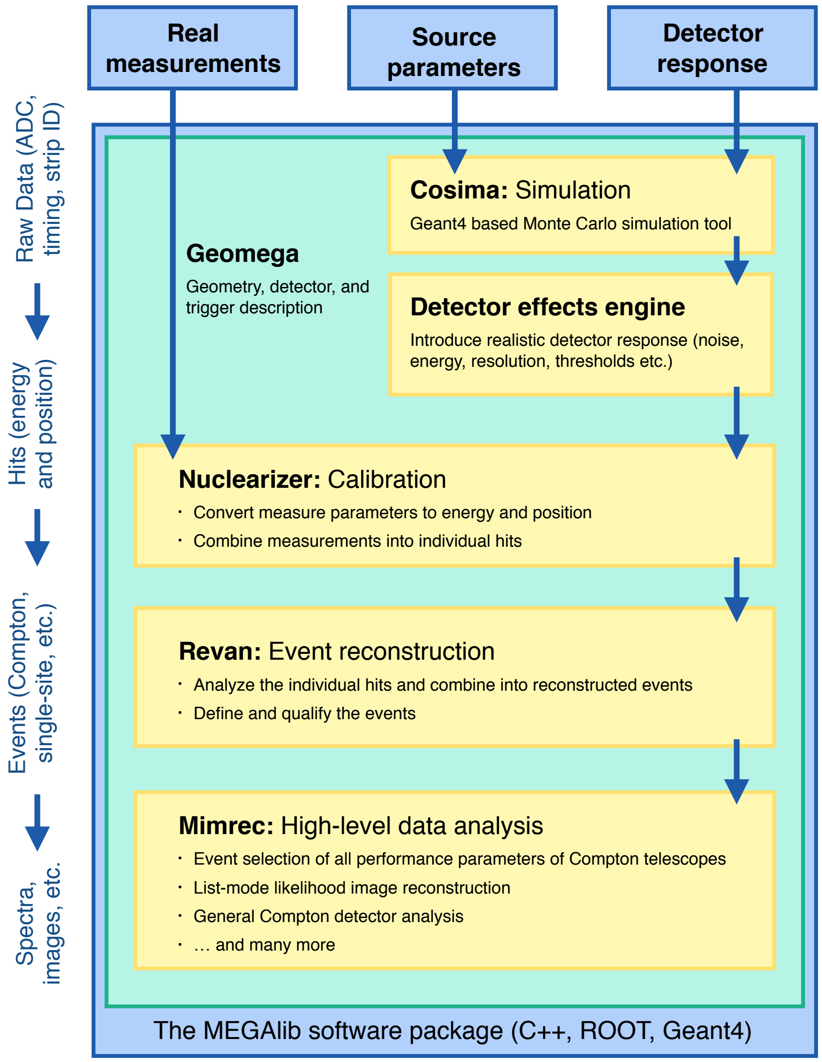


Figure 3.19: Flowchart for the MEGAlib software package (Zoglauer et al., 2006). The simulations, based on `Geant4`, take the same path through the analysis pipeline as the measured raw data. The programs are written in `C++` using `ROOT` (Brun & Rademakers, 1997).

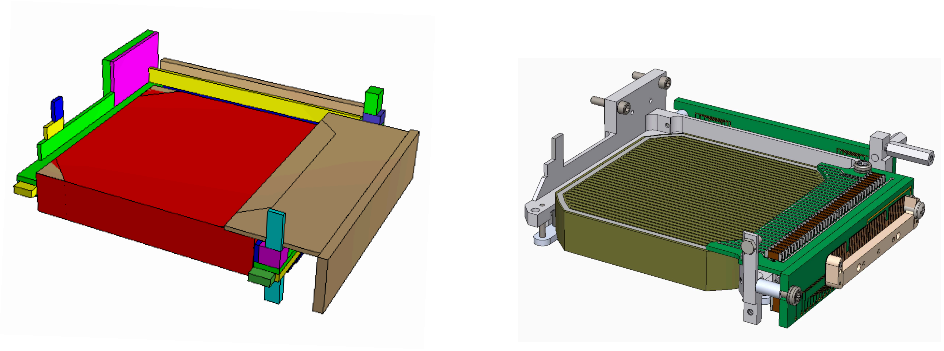
tains a detectors effects engine (DEE) that allows one to convert simulated data into realistic events; `cosima` is used for simulations of the detector response to arbitrarily complex γ-ray sources; `revan` performs the Compton event reconstruction of data and simulations alike; and `mimrec` allows for image reconstruction and higher level analysis. In addition, an automated energy calibration program `melinator` has been recently developed and is used for the energy calibrations discussed in Section 4.2. The analysis path is the same for real and simulated data; see Figure 3.19 for a flowchart describing how these programs fit together. These programs will be discussed in further detail below.

3.5.1 GEOMEGA

The COSI mass model is built in the `geomega` program. The mass model is a geometrical representation of the instrument to be used with MEGAlib simulations and analysis. In `geomega`, one defines the volume of each component, the material composition, and whether the part is an active detector, with given sensitivity and trigger criteria, or passive material. An accurate representation of the material surrounding the detectors is necessary to for accurate simulations and detector response matrices and the mass model is based on the COSI `SolidWorks` model to confirm that the dimensions and masses are correct. Figure 3.20 (a) shows the mass model of a single COSI GeD created in `geomega`, and Figure 3.20 (c) shows the full cryostat, which includes the twelve GeDs, and the shields, PMTs, preamps, cryocooler, etc. There is another version of the COSI mass model that includes the entire gondola structure, but to save on simulation time, this is not often used.

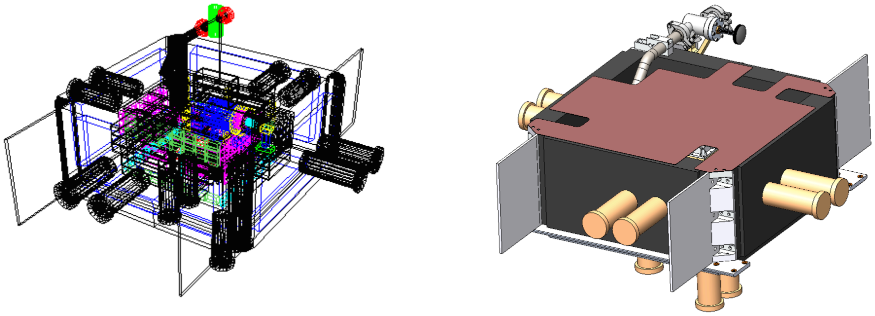
3.5.2 COSIMA

The cosmic simulator for MEGAlib, `cosima`, is a simulation tool based on `Geant4` (Agostinelli et al., 2003), a Monte-Carlo tool developed by CERN that tracks the passage of particles through matter. The input to `cosima` is a spectral and spatial description of a source (photons or particles) and the program preforms a simulation of how the particles will interact in the `geomega` mass model. The simulation output contains the energy and positions of photon or particle interactions within the detector. A wide range of spectral and spatial models are included as predefined functions; however, the user can include a file describing an arbitrary source.



(a) `geomega` model of single GeD.

(b) `SolidWorks` model of single GeD.



(c) `geomega` model of cryostat and shields.

(d) `SolidWorks` model of cryostat and shields.

Figure 3.20: (a) A single COSI GeD built in `geomega` based on `Solidworks` model is shown in (b). In the `geomega` drawing, the germanium crystal is red and the detector board and mechanical mounts are a variety of colors. (c) The `geomega` model of the COSI cryostat and surrounding shields with each component displayed as wire frames. All measurements are based on the `SolidWorks` model, shown in (d).

A recent addition to the `cosima` program is the inclusion of *orientation*. A file containing the aspect information for an instrument can be loaded and the source position can be defined in Galactic coordinates. The aspect information can be used to determine where the source is in the instrument's FOV at all times, and the Galactic coordinates are modeled by a rotating celestial sphere around the instrument mass model. `cosima` can also take into account the atmospheric attenuation, where the effective column density is based on the NRLMSISE00 model (Picone et al., 2002) and a simple Monte-Carlo simulation is used to "attenuate" the incident photons. With these two relatively new additions, `cosima` is capable of performing simulations that are accurate representations of flight conditions.

Instrument activation from the bombardment of cosmic-rays in space-like environments can also be simulated in `cosima`. This will be discussed further in Section 7.3.1 when we give a detailed description of the background simulation performed to reproduce the emission detected in the 2016 COSI flight.

3.5.3 NUCLEARIZER

`nuclearizer` is the detector calibration program. The raw data from COSI is in the form of analog-to-digital converter (ADC) values for the peak height, timing between the strips on the AC and DC side of the detector, in addition to the Strip and Detector ID in which the pulses are measured. In order to perform Compton event reconstruction, the energy and position of each interaction must be determined. This is done through instrument calibrations. The raw data from the flight computer is loaded in `nuclearizer`, and the calibrations are applied to convert the ADC into energy and the timing into depth of interaction. A flow chart showing the instrument calibrations in `nuclearizer` is shown in Figure 3.21 and this will be discussed further in Chapter 4.

To make the `cosima` simulations more realistic, a detectors effects engine (DEE) has been included in `nuclearizer`. The DEE applies the inverse of the instrument calibrations, converting the simulated energy into an ADC and the depth of interaction into timing, so that after the DEE is applied, the simulated data will look almost exactly like the raw data. The electronic thresholds and any dead strips are also included in the DEE. When reading in a simulation file in `nuclearizer`, first the DEE is applied and then the simulated data can

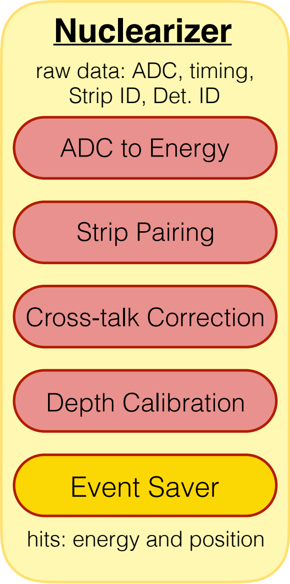


Figure 3.21: Flowchart for the `nuclearizer` calibration program implemented for COSI. Each red block represents a seperate instrument calibration module and the full process converts the measured raw data into hits, with energy and position.

go through the same calibrations as shown in Figure 3.21. The rest of the analysis pipeline is then identical for simulations and real data.

3.5.4 REVAN

`revan` is the real event analyzer for MEGAlib, which performs the event reconstruction. After `nuclearizer`, all interactions have an energy and interaction position, and `revan` matches coincident events and reconstructs the Compton path, as discussed in Section 3.2.3. In general, `revan` also can handle the event reconstruction of electron tracking Compton detectors and pair conversion telescopes that are commonly used for detection of higher energies γ-rays.

3.5.5 MIMREC

All higher-level analysis is done through the image reconstruction program `mimrec`. `mimrec` is used to make images, spectra, ARM distributions, and light curves of observations. Additionally, there are a variety of useful analysis tools contained in `mimrec`: histograms of the Compton scatter angle, distributions of the distance between interactions, exposure maps, plots of the azimuthal scatter angle distribution for polarization analysis, and many more. MEGAlib also allows for user-defined stand-alone programs that can be extended past the routines contained in `mimrec`.

3.5.5.1 EVENT SELECTIONS

In Compton telescopes, a careful choice of event selections can significantly improve detection sensitivity. The main event selections available in `mimrec` are:

- *total energy of γ-ray* - Selecting the range of allowable energies can confine one's analysis to a single γ-ray line, or exclude specific background lines, e.g., 511 keV.
- *Compton scatter angle* (ϕ) - The Compton scatter angle of the first interaction has a large effect on the quality of the event. Backscatters, with $\phi > 90°$, are generally harder to reconstruct, so limiting ϕ to less than 90° can reject these lower-quality events.
- *minimum distance between the first interactions* - The distance, measured in cm, between the first two interactions is the lever arm of the Compton cone; with a longer baseline, the direction of the γ-ray can be more accurately determined. For most analyses, this is chosen to be 0.5 cm.
- *minimum distance between any two interactions* - When the distance between two sequential interactions is small, it can be more difficult to properly reconstruct the event. By limiting this distance to a minimum of 0.3 cm, considering the width of the strips in COSI is 0.2 cm, poorer quality events can be excluded.
- *length of Compton sequence* - The number of allowable interactions can be selected, where a length of 2 corresponds to a 2-site event, i.e., one Compton scatter followed by a photoabsorption interaction. For COSI, this is generally chosen to be 2-7, since very few events have more interactions.
- *earth horizon cut (EHC)* - The EHC rejects any event whose Compton circle overlaps with the horizon. This is the most rigorous method to reduce the background from albedo radiation during flight.

- *origin cut selection* - A selection can be made on a location in image space, within a given radius, which can be used when the source location is known. Only the Compton circles which overlap with this origin selection will be kept. This selection can be used in spherical, Cartesian, or Galactic coordinates.
- *pointing selection* - A pointing selection can be made for a certain Galactic latitude and longitude and a certain radius, and only the events detected when the instrument's zenith is within this pointing cut are saved. This is used for flight data (when the aspect information is known) to exclude times when a source is outside of the FOV.
- *good time interval (GTI)* - A GTI can be used to restrict the observation, either set as a time range in UTC or by uploading a file including a list of "good" time intervals. These intervals are used to reject times of low altitude, high-background, or when the instrument is not in nominal operation.

3.5.6 MELINATOR

`melinator` is MEGAlib's line calibrator, which performs an automated energy calibration. This will be further discussed in Section 4.2.

4

CALIBRATIONS OF THE COMPTON SPECTROMETER AND IMAGER

Through instrument calibrations, we convert the measured detector parameters into the physical parameters necessary for Compton event reconstruction, namely the energies and positions of interaction. The resolution to which we can determine the positions and energies affects the accuracy of the event reconstruction and, in turn, the sensitivity and angular resolution of the telescope. Therefore, it is extremely important to properly calibrate a Compton telescope to reach optimal performance.

In this chapter[1], we will give an overview the COSI calibration pipeline in Section 4.1 and discuss how the calibration models for each step have been determined through laboratory measurements in Section 4.2–4.5. Once the instrument has been calibrated, we can establish its higher-level performance and we will conclude this chapter with an overview of angular resolution, effective area, and polarization response in Section 4.6.

4.1 OVERVIEW OF INSTRUMENT CALIBRATIONS

The parameters available from the detector readout electronics are the peak height, which is in ADC channel value, the pulse timing from the anode and cathode strips, the Detector ID, and the Strip ID. By performing detector calibrations, we can determine the conversion between these parameters and meaningful physical values: the interaction energy and position. The COSI instrument calibration pipeline is a six step procedure in `nuclearizer` (Figure 3.21):

---

[1] Some of this chapter is based on a previous publication "Calibration of the Compton Spectrometer and Imager in preparation for the 2014 balloon campaign" by Kierans et al. (2014). ©2014, Society of Photo-Optical Instrumentation Engineers (SPIE). Reprinted with permission.

1. *Load Data* - the measured data for one event can contain multiple active strips in multiple detectors; for each activated strip there is the Strip ID, Detector ID, ADC, and timing information.
2. *Energy Calibration* - the ADC peak height measured from each activated strip is converted to a measure of deposited energy in keV.
3. *Strip Pairing* - the activated strips on opposite sides of the detector are "paired" after the correct combination is determined.
4. *Cross-Talk Correction* - the signal-enhancing effect for neighboring activated strips is corrected.
5. *Depth Calibration* - the 3D position of the interaction is determined in reference to the mass model from the Strip ID, Detector ID, and signal timing.
6. *Save Data* - each event is saved. Single-site events will be described by one energy deposit and one 3D position and multi-site events will contain multiple energy deposits in multiple positions and even in multiple detectors.

The calibration of the instrument predominately involves defining the correct conversion for energy-to-ADC and timing-to-depth. These conversions are determined through thorough laboratory measurements. Additionally, the results of the calibrations are used in the DEE of our simulation software, which enables us to mimic the attributes of our detectors in simulations, thus creating a more accurate instrument model.

In order for COSI to perform diffuse spectroscopic imaging in wide field surveys, we need to fully understand the sensitivity and angular resolution over the entire FOV of the instrument. Extensive measurements and simulations of the integrated instrument are needed to properly benchmark the simulations and understand the detector response. These higher-level instrument tests are discussed in Section 4.6.

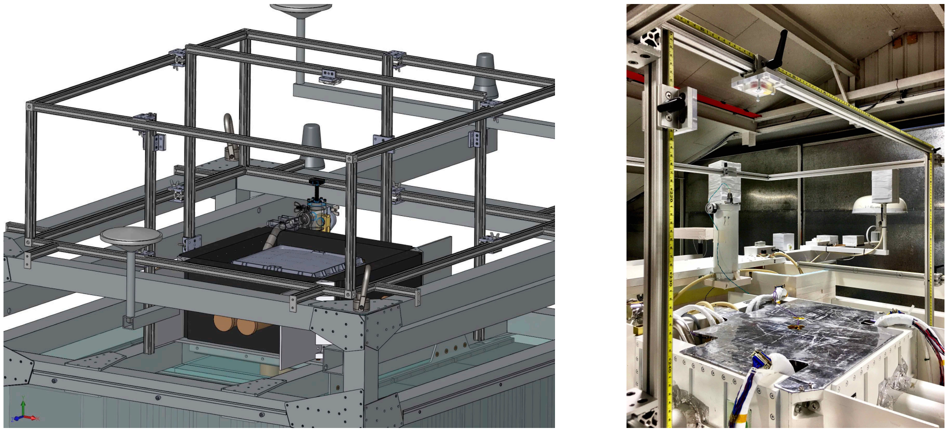
All laboratory calibration measurements are done with sealed Type D disk γ-ray sources from Eckert & Ziegler. A list of the COSI γ-ray calibration sources and relevant information is listed in Table 4.1.

Source	Half-life	Activity* [μCi]	Energy [keV] (BR)
^{241}Am	432 yr	99.4	59.54 (0.36)
^{57}Co	271 d	2.18	122.06 (0.86)
			136.47 (0.12)
^{133}Ba	10.7 yr	8.64	81.00 (0.33)
			276.40 (0.07)
			302.85 (0.18)
			356.01 (0.62)
			383.85 (0.09)
^{22}Na	2.6 yr	6.34	511.0 (0.90)
			1274.54 (1.0)
^{137}Cs	30.2 yr	91.0	661.66 (0.85)
^{88}Y	107 d	18.7	898.04 (0.94)
			1836.05 (0.99)
^{60}Co	5.3 yr	58.3	1173.23 (0.99)
			1332.49 (0.99)

Table 4.1: The γ-ray sources used in the COSI instrument calibrations with the most prominent γ-ray lines and branching ratios listed. The listed activity is valid for 7 March 2016, which was the start of full instrument calibration measurements during the COSI Wanaka campaign.

4.1.1 LABORATORY CALIBRATION STRUCTURE

In order to have accurate and reproducible placements of the laboratory calibration sources, an 80-20 frame structure was designed with translation in the horizontal and vertical direction to cover the COSI FOV; see Figure 4.1. The range of translations allows for virtually all positions in the FOV to be sampled, except for a ~3 cm wide region in the corners of the structure. Clear plastic holders for the γ-ray sources, with an open window for the active region, were designed to mount to the support structure. All calibrations were performed with a fully integrated instrument with the γ-ray calibration sources mounted in the 80-20 structure, unless noted otherwise.



(a) `SolidWorks` model of 80-20 calibration structure.

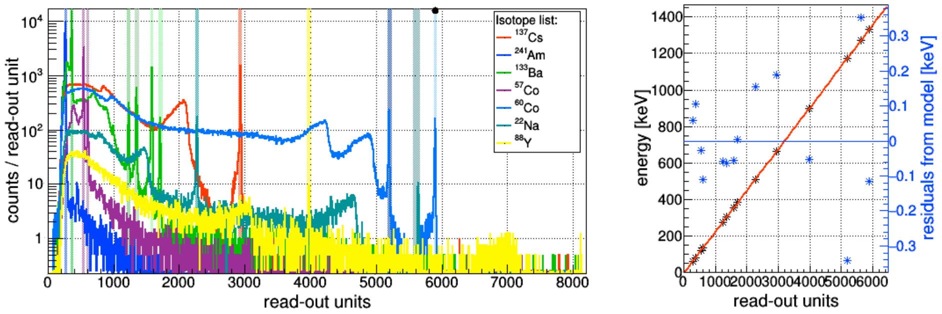
(b) Photograph of calibration structure.

Figure 4.1: (a) `SolidWorks` model of the 80-20 calibration structure mounted on top of the gondola. (b) Photograph of the calibration structure during calibrations. Rulers with mm markings attached to each translation bar are used for precise and reproducible source placement. The structure allows for three sources to be accurately placed in the FOV simultaneously.

4.2 ENERGY CALIBRATION

When a γ-ray scatters in a GeD, the recoil electron will create an equal number of electrons and holes, where the number of electron-hole pairs is directly proportional to the initial energy of the recoil electron (2.96 eV per *e*-*h* pair). Under the presence of an electric field, the electrons and holes will drift in opposite directions along the field lines, and it is this motion that induces a signal on the anode and the cathode segmented electrodes. Therefore, the output of the slow shaping channel, an ADC channel value which corresponds to the peak height of the induced signal, is approximately proportional to the energy deposited. Small deviations from linearity occur due to the response of the germanium and the shaping electronics. To achieve a good understanding of the conversion from ADC channel to energy, we sample the energy range of the detectors with 10–15 nuclear lines of known energy (Table 4.1).

For the energy calibration measurements, radioactive sources are placed in the 80-20 calibration structure ~50 cm above the cryostat to evenly illuminate all detector strips. For the lower energy sources with γ-ray lines $\lesssim 400$ keV, i.e., ^{241}Am, ^{57}Co, and ^{133}Ba, the mean free path of the radiation through ger-



(a) Total combined spectra in `melinator`.

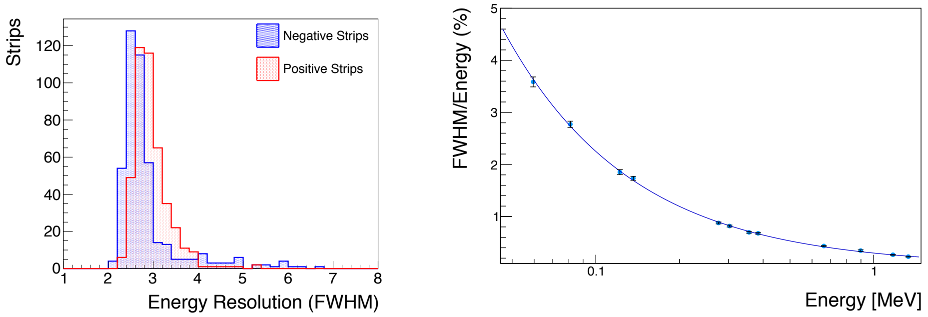
(b) Energy calibration polynomial fit.

Figure 4.2: (a) Energy calibration data for DC strip 13 on Detector #0 from `melinator`. The spectrum from each calibration measurement is superimposed and the photopeak lines are automatically found and fit. (b) With the photopeak centroid of each γ-ray line found in (a), `melinator` fits the energy-ADC relation to determine the calibration model. The fit residuals, which are less than 0.4 keV, are shown in the right.

manium is much smaller than the depth of a single HPGe crystal. We achieve illumination of the bottom and middle detectors by positioning the low-energy source at the edge of the cryostat aligned with the gap between the detector planes. All measurements are taken for each source independently.

By measuring the ADC value for γ-ray lines of known energy, we can find the appropriate conversion between the two parameters for each detector strip. Due to the large number of required line fits for this analysis (888 strips with $\gtrsim$ 10 lines each), the COSI team developed an automated line calibration program within MEGAlib. `melinator` parses the calibration measurement files to find the spectrum of each source measured by each strip, automatically finds and fits the lines, and determines the calibration model describing the conversion from ADC to energy for each strip. To precisely find the centroid and FWHM of the measured energy lines, the spectra are fit with a phenomenological model: a delta function plus an exponential tail convolved with Gaussian noise (see Bandstra, 2010, Eq. 6.6).

Figure 4.2 (a) shows the raw ADC spectrum from the seven γ-ray calibration sources superimposed in `melinator`. With at least 10 nuclear lines within the COSI single-strip energy range, `melinator` automatically fits the resulting ADC-energy relation, shown in Figure 4.2 (b), with an empirical model to account for non-linear deviations; in the current implementation of the cali-



(a) Energy resolution at 662 keV.

(b) Energy resolution as a function of energy.

Figure 4.3: (a) The single-strip energy resolution measured as the FWHM of the ^{137}Cs line at 662 keV for all of the 888 strips in the twelve COSI GeDs. The AC and DC strips have been plotted separately. The average single-site energy resolution of the COSI GeDs is 2.9 keV, or 0.4% at 662 keV. (b) The averaged single-strip energy resolution as a function of energy as measured by COSI. The error bars are $\sigma/\sqrt{N}$ and the relation is fit with $y = Ax^k$, where k is found to be -0.9.

bration we use a 3rd or 4th order polynomial. Although the program is automated, the calibration result for each strip is manually checked for poor line fits which could degrade the calibration model. The output of `melinator` is a file that contains the best-fit polynomial description of the energy calibration for each strip, which, used as an input into the energy calibration module in `nuclearizer`, allows for the energy to be evaluated at any measured ADC value.

An important measure of strip and detector performance, and validation of our energy calibration, is the measured energy resolution of each strip. A histogram containing the energy resolutions of all 888 strips in the twelve COSI GeDs is shown in Figure 4.3 (a), where the resolution is defined as the FWHM of the `melinator` line fit at 662 keV. The AC and DC strips have been plotted separately and a slight difference is seen in the two populations. The average single-strip energy resolution of the DC strips is 2.86 keV, compared to 2.95 keV for the AC strips. It is thought that the small degradation of the AC strip resolution is due to electron trapping within the HPGe crystals. Combining the AC and DC strips, the average single-strip energy resolution of COSI is 2.9 keV FWHM, or 0.4%, at 662 keV, and the standard deviation is 0.6 keV. Figure 4.3 (b) shows the averaged single-strip energy resolution as a function of the energy.

Sources	UTC Run Time	Average Preamp Temp [°C]
^{137}Cs & ^{241}Am	April 23rd, 2016, 16:36-17:06	12.1
^{137}Cs & ^{241}Am	April 3rd, 2016, 17:43-18:23	14.6
^{137}Cs & ^{241}Am	April 1st, 2016, 18:40-19:40	26.2
^{137}Cs & ^{241}Am	April 3rd, 2016, 12:50-13:30	18.7
^{137}Cs & ^{241}Am	April 23rd, 2016, 08:06-08:36	34.1

Table 4.2: Date and times for the preamp temperature dependence measurements.

4.2.1 TEMPERATURE DEPENDENCE

A surprising realization made shortly before the Wanaka campaign was that the COSI preamplifier boards, which have been flown in three previous balloon campaigns, are temperature sensitive. It was confirmed with bench-top measurements that the preamp circuit could induced a peak shift of up to 0.5 keV/°C at 662 keV. Considering the temperature variations expected in flight, it is vital that the temperature effect is well characterized to maintain the excellent energy resolution necessary for the scientific success of COSI. To properly record the temperatures so an accurate correction can be applied, the COSI team added 24 new temperature sensors to the system so that each preamp could be monitored individually.

As with the energy calibration, the temperature correction is determined for each strip. To perform these calibration measurements in Wanaka, the ambient temperature in the hangar was significantly reduced on three separate occasions once the instrument was fully integrated so all strips could be calibrated simultaneously. We preformed the measurements with ^{137}Cs since the count rate and penetrating power of the source allow for a determination of the 662 keV peak position for each strip within ~30 min. We also included the ^{241}Am source for these measurements to check the energy dependence of the gain shift. Table 4.2 lists the five measurement times used for this calibration and the average preamp temperature for each.

Preforming the line fits in `melinator`, we find a linear relation between the preamp temperature and the ADC peak location. Figure 4.4 shows the measured ADC peak centroid of the 662 keV line as a function of the preamp temperature for the randomly chosen Strip 21 of Detector #0. Because of this linearity, we can define our temperature correction to be a simple offset, as measured as a percent of the energy, which depends on the preamp temperature. Within the energy calibration module in `nuclearizer`, we have added

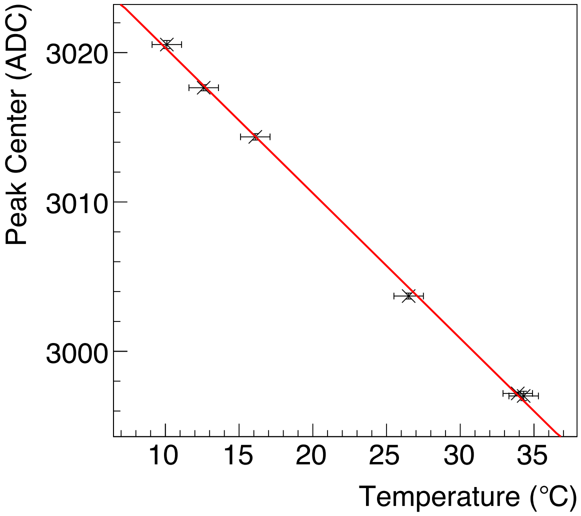
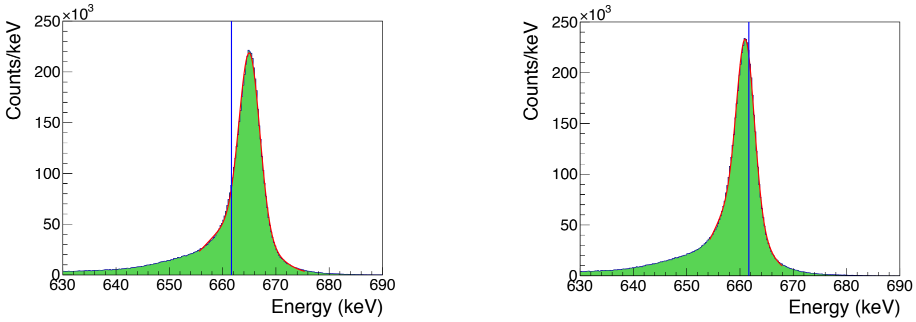


Figure 4.4: Measurements of the ADC shift for the 662 keV line of ^{137}Cs as a function of temperature for each of the runs listed in Table 4.2. The linear fit to the data gives the temperature correction for this one strip: DC strip 21 of Detector #0.

a correction that shifts the ADC to what the preamps would have measured without any temperature sensitivity. One difficulty with this correction is that it requires that all of the standard ADC-energy calibration measurements be preformed at the same temperature. Since the temperature in the hangar would naturally vary during day-night cycles, we used the hangar's climate control to keep the ambient temperature within a degree or two through long night measurements for the energy calibrations.

Figure 4.5 shows the spectra for the ^{137}Cs 662 keV line from a calibration run when the temperature varied from 11–20° within one hour. The full spectrum is shown with and without the temperature correction in (a) and (b), respectively. Without the temperature correction, a Gaussian fit to the line gives a centroid at 665.160±0.002 keV and line width of σ = 1.913±0.002 keV, where the errors are from the fit function. This line is broadened and shifted to a higher energy due to the lower temperatures of this calibration run. When the correction is applied, as shown in Figure 4.5 (b), we now obtain a line centroid at 661.046±0.002 keV with a width of σ = 1.634±0.003 keV. Though our corrected line is slightly lower than the expected 661.7 keV, it is a signif-



(a) Without temperature correction.

(b) With temperature correction.

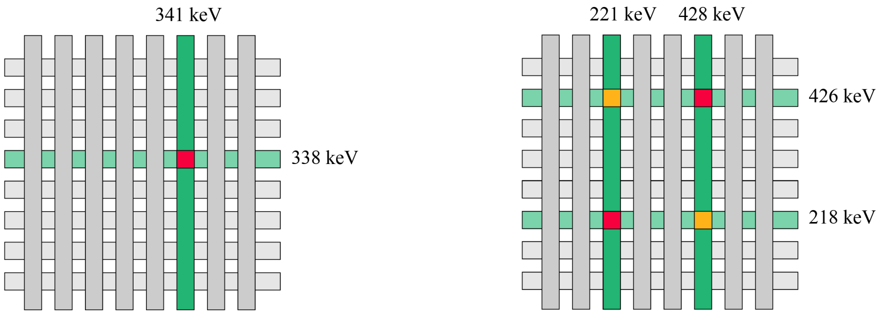
Figure 4.5: (a) For an hour long calibration measurement where the ambient temperature ranges between 11–20°, the line centroid is measured at 665.2 keV due to the temperature-dependent preamplifiers. (b) When the temperature correction is applied, the line is narrower in width and the center is shifted to 661.05 keV.

icant improvement upon the uncorrected value, reducing the relative difference from 0.5% to 0.1% of the correct line centroid.

4.3 STRIP PAIRING

The interaction location in cross-strip GeDs is determined by the activated strips. The orthogonal strips allow for the position of the interaction to be determined by the cross section of the strips, as Figure 4.6 (a) illustrates for a single scatter. If two interactions occur simultaneously in the same detector, there are two possible combinations of strip-pairs that could define the positions of the two interactions; see Figure 4.6 (b). The strip pairing calibration module in `nuclearizer` determines which combination is more probable to describe the interaction positions, which are then called “hits.” The most straightforward way to do this is to compare the energy deposited on all of the strips. An ideal interaction will result in the two corresponding strips measuring the same amount of deposited energy. However, in a real detector it is not this simple. Dead strips, poor energy resolution, high thresholds, cross-talk and charge sharing events can all affect the outcome of the pairing algorithm.

Figure 4.7 shows the results after the strip pairing algorithm has been applied to a calibration measurement with a ^{22}Na source. For a perfect pairing,



(a) One interaction with a single GeD.

(b) Two interactions within a single GeD.

Figure 4.6: (a) With one interaction in the detector volume, the cross section of the two active strips, colored green, gives the interaction position which is denoted by the red square. (b) If two interactions occur in one detector, then both the red and orange squares are possible combinations of interaction positions that result in the same four activated strips. The true pairing is determined by matching the energies of the activated strips which identifies the red squares as the correct hit positions.

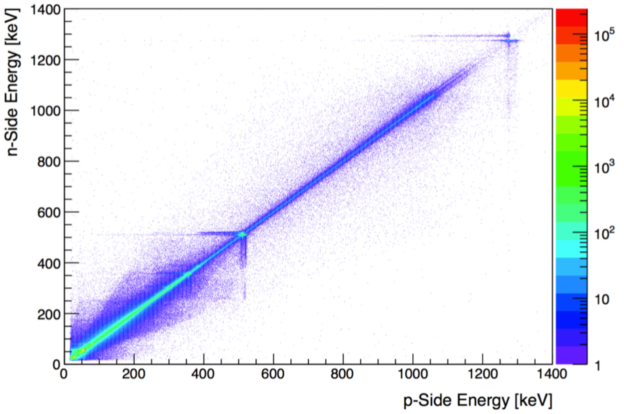


Figure 4.7: The result of COSI's strip pairing algorithm for a ^{22}Na calibration source. The n-side and p-side energy, corresponding to the anode and cathode of the detector, should be equal for a correct pairing. This source has two lines: at 511 keV and 1.27 MeV, which can be seen here as hot spots. The horizontal and vertical lines originating from these energies are from charge-loss. The offsets to the right and above the measured 1.27 MeV point are due to cross-talk.

the measured energy on both sides of the detector, denoted here by n-side and p-side, would be equal. In addition to showing that the strip pairing algorithm performs well, this figure reveals much about the response of our instrument. The ^{22}Na source has two γ-ray lines, at 511 keV and 1.27 MeV, and the hot spots at these energies result from single photoabsorption events; everything else along the n = p line is from a multi-site event when the full γ-ray energy in not deposited on a single interaction. The horizontal and vertical lines originating from the photoabsorption hot spots at 511 keV and 1.27 MeV are due to charge-loss, whereas the offset points to the right and above the measured single-site energies, especially obvious in the higher-energy hot spot, are due to cross-talk. These two effects will be further discussed in the next section.

4.4 CROSS-TALK AND CHARGE-LOSS

More difficulties arise when dealing with the spectra of multi-strip events. The energy spectrum is degraded primarily through two phenomena: cross-talk between adjacent electronic channels and charge-loss in the 0.5 mm gap between strips. Figure 4.8 shows the full instrument spectrum of the 662 keV line of a ^{137}Cs calibration source after the strip pairing algorithm. We expect a single narrow Gaussian centered at 661.7 keV, which is indicated with the vertical blue line; however, there is structure from cross-talk seen at higher energies and a low-energy tail from charge-loss. These two effects are fairly well understood and we have implemented the corrections developed for past NCT calibrations (Bandstra, 2010).

We measure significant cross-talk when interactions occur on neighboring strips of our GeDs. Due to the proximity of an activated strip, a neighboring strip will measure an artificially enhanced energy. Often, the resulting enhancement in energy is not seen since it is below the threshold of the neighboring strip, but when two neighboring strips have significant charge collection, then the effect becomes apparent. This is seen in the fully-reconstructed Compton spectrum of a ^{137}Cs source shown in Figure 4.8, where we expect the emission peak to be centered at 661.7 keV. The bump at ~670 keV is due to cross-talk. Cross-talk occurs for single-site and multi-site events regardless of the underlying interaction.

To determine the cross-talk correction factor we isolate only the multi-site events that have two activated neighboring strips on one side of the detector and two non-adjacent strips on the other side. Looking only at these events, we find the enhanced measured energy is very well determined by the total

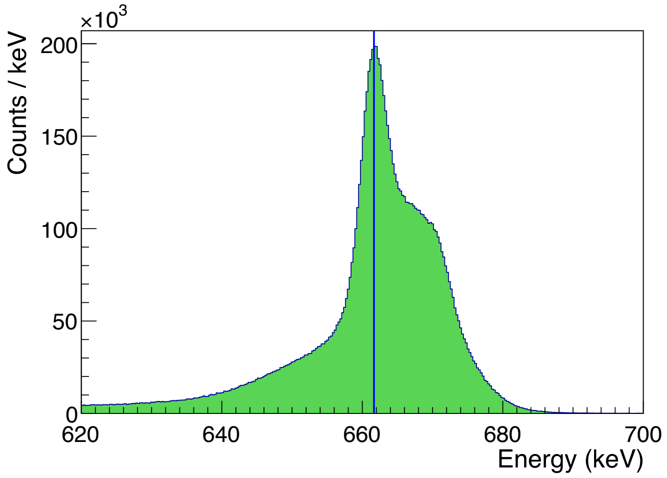


Figure 4.8: Full multi-site energy spectrum after strip pairing algorithm.

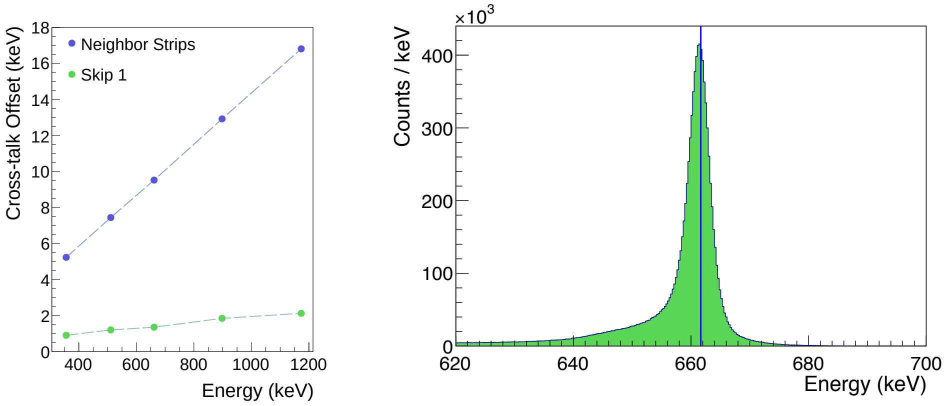
energy deposited. By using the seven γ-ray calibration sources, we can find the energy enhancement as a function of line energy, and since it is fortunately a linear effect, we can easily correct for this. The results of these calibration measurements are a linear correction for each side of each detector, for both neighboring and next-nearest-neighbor strips.

Figure 4.9 (a) shows the measured enhancement for the DC-side of Detector #6. Fitting a line to these two relations gives the expected cross-talk enhancement for any deposited energy. To recover the true energy of the interaction, we simply subtract off this known offset for any interaction measured in the GeDs that results in two adjacent or near-adjacent activated strips. Figure 4.9 (b) shows the same spectrum as Figure 4.8 but now with the cross-talk correction applied. A narrow Gaussian line at 662 keV is recovered.

The COSI GeDs do not show significant charge-loss on the DC side of the detector. Although the reason for this is not entirely understood, to mitigate the degradation of measured energy from charge-loss, the energy measured on the DC side of the detector is used for every interaction. Therefore, even though previous NCT calibrations have included a correction for charge-loss, we find it unnecessary.

4.5 DEPTH CALIBRATION

The X and Y position of an interaction in the COSI GeDs is determined by the intersection of the activated anode and cathode strips, as shown in Fig-



(a) Cross-talk correction for the DC side of Detector #6.

(b) Full spectrum with cross-talk correction.

Figure 4.9: (a) The measured cross-talk enhancement and linear correction for DC-side of Detector #6. The cross-talk effect is largest for neighboring strips, but still is detectable for Skip 1 neighbors. (b) The total multi-site energy spectrum as in Figure 4.8 but now with the cross-talk correction applied.

ure 4.6 (a). From the strip width of 2 mm, this defines an X-Y pixel size of 4 mm^2. The depth of the interaction, however, cannot be directly measured and thus needs to be inferred from the timing of the signals deposited on the anode and cathode. The signal collection time difference (CTD) is the difference between the time when the electrons arrive at the anode and the time when holes arrive at the cathode. For example, if an interaction occurs near the anode, the electrons have a much shorter distance to travel and will arrive before the holes drift through the depth of the detector to be collected on the cathode. An interaction that occurs within the center of the detector would lead to an equal arrival time of electrons and holes, resulting in a CTD ~ 0. This is illustrated in Figure 4.10.

The depth calibration consists of determining the relation between CTD and the true depth of interaction for each pixel of each detector ($37^2 \times 12 = 16428$ pixels). As each pixel will have a slightly different response due to the timing electronics (resulting in a CTD offset) and electric field differences in the detector (resulting in a scaling of measured CTD), each pixel of each detector must be individually calibrated. In general, an ideal depth-to-CTD conversion is found for each detector through charge transport simulations; see Figure 4.11 (a). The detector mass model is used to simulate the expected depth distribution, which is converted to an expected CTD distribution, or CTD tem-

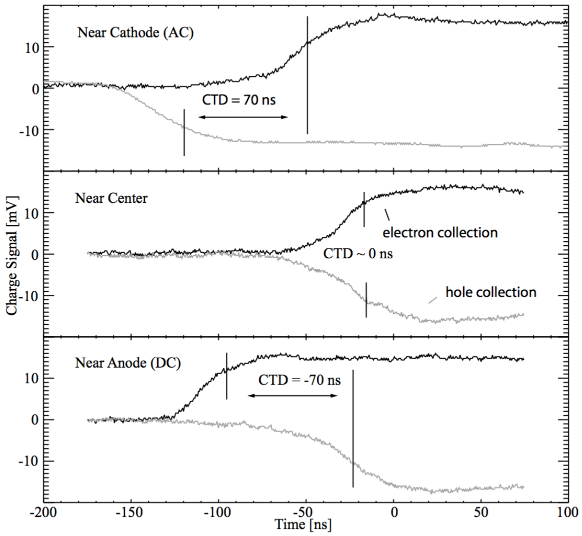
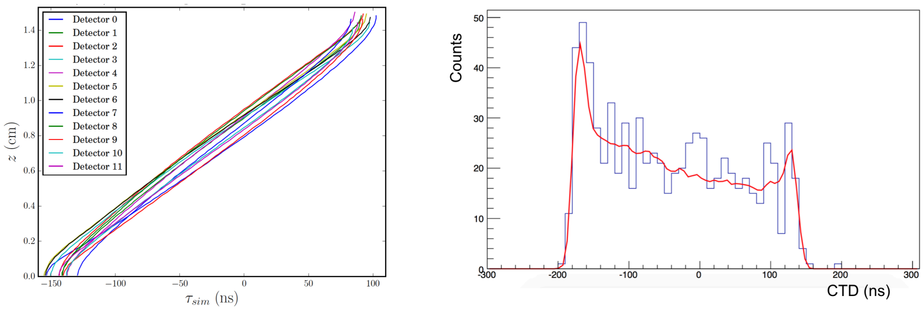


Figure 4.10: The measured signal from three interactions at different depths in a COSI GeD illustrating the definition of CTD. In the top panel, the interaction occurs close to the cathode of the detector, and the signal from the holes arrives before the signal from the electrons, resulting in a positive CTD. In the middle panel, the interaction occurs near the center of the detector and CTD ~ 0. With the interaction occurring closest to the anode, shown in the bottom panel, the electron collection signal arrives first and a negative CTD is measured. Figure 2 from Coburn et al. (2003).



(a) Depth-to-CTD curves.

(b) Single pixel CTD distribution.

Figure 4.11: (a) CTD to depth conversion from charge transport simulations for all 12 of COSIs detectors. Each detector has a slightly different curve due to the measured thickness of the detector and the high-voltage applied. (b) The measured CTD from a ^{137}Cs source in a single pixel. The red curve is the fit of the CTD template, which shows a good match to the measured data. Figures 4 and 5 from Lowell et al. (2016).

plate, through the ideal depth-to-CTD conversion. The conversion is then fine-tuned for each pixel by fitting the CTD template to measured CTD data to obtain the offset and scaling parameters for each pixel; see Figure 4.11 (b). See Lowell et al. (2016) for a more detailed discussion of the depth calibration for COSI.

4.6 HIGHER-LEVEL BENCHMARKING

After the energy and depth calibrations are complete, we can perform reliable Compton reconstruction and are able to explore the higher-level performance of the telescope. An extensive characterization of the angular resolution and effective area over the entire FOV of the instrument is critical for the success of the telescope. The COSI team performed extensive laboratory calibrations and simulations to characterize the instrument response and benchmark the DEE. The preliminary higher-level performance will be briefly discussed in Section 4.6.1 and 4.6.2. Additionally, the polarization response of the instrument needs to be validated in the laboratory prior to launch, and this will be briefly introduced in Section 4.6.3. See Sleator et al. (2016) for a comparison of the efficiency and angular resolution from calibration data to the simulations with the DEE, and see Lowell (2017) for a detailed overview of the polarization response validation.

4.6.1 ANGULAR RESOLUTION

In Compton telescopes, the angular resolution is described by the FWHM of the ARM distribution, introduced in Section 3.2.5. Understanding the ARM response over the entire FOV of the instrument is crucial for the wide-field survey operating mode of COSI. Furthermore, as the angular resolution is limited by the accuracy of the energy and position determination within the instrument, the ARM distribution offers a sensitive comparison between data and simulations. After precise instrument calibrations, any differences seen between the data and simulations should be accounted for in the DEE.

To preform the angular resolution calibration measurements, the COSI team took over 20 observations of each γ-ray calibration source in the COSI FOV, with the precise location of the source known from the 80-20 structure. By analyzing the ARM distribution from the γ-ray line sources, we can find the energy dependence of the telescope angular resolution. We will only present

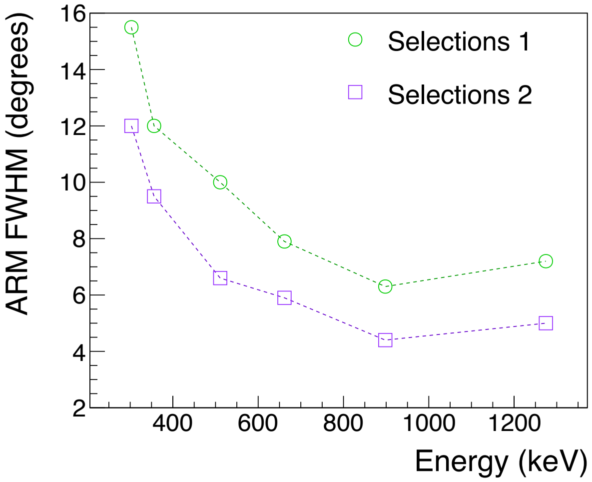


Figure 4.12: The FWHM of the ARM distribution gives the angular resolution of the telescope, which is plotted here as a function of the energy as measured in calibrations. Two different levels of event selections are shown. The green open circles show the angular resolution with open event selections. The angular resolution is dramatically improved wit a 1 cm distance cut and with the Compton scatter angle ϕ restricted to less than 90°, shown with the purple open squares. The COSI's angular resolution is a minimum of 4.5° at ~900 keV.

the preliminary on-axis data here, as the analysis of the full calibration set is still a work in progress.

We follow a 3-step process to determine the ARM FWHM for each source: 1) fit the measured spectrum to find the Gaussian width of the γ-ray line, 2) apply an energy cut of $\pm 1.5\ \sigma$ to select only the photopeak events. 3) make a histogram of all ARM values and fit with a phenomenological model to find the FWHM of the distribution. The angular resolution as a function of energy from on-axis calibration measurements is shown in Figure 4.12. There are two curves corresponding to two different levels of event selections. The green curve with open circles has no restrictions on the event selections other than the $\pm 1.5\ \sigma$ energy cut. The purple curve with open squares includes a more stringent 1 cm first interaction distance cut, and a Compton scatter angle restriction of < 90° (see Section 3.5.5.1 for an overview of the event selections in `mimrec`). With these more restrictive sections, the angular resolution of the telescope is improved dramatically.

With these event selections, the COSI angular resolution reaches a minimum of 4.5° around 900 keV. At 511 keV, we find an optimized angular resolution of 6.6° for this calibration run. More restrictive event selections could reduce this further; however, this will in turn reduce the efficiency.

4.6.2 EFFECTIVE AREA

The effective area A_{eff} is a measure of the efficiency of a telescope. It is defined as the area that an *ideal* absorber needs to detect an equivalent number of incident photons. The effective area is defined by:

$$A_{eff} = A_{start}\frac{N_{det}}{N_{start}}. \tag{4.1}$$

For simulations, A_{start} is the simulated area surrounding the mass model, N_{start} is the number of photons that are initiated, and N_{det} is the number of photons that pass given event selections. In calibration measurements, N_{start} is determined from the source activity and exposure time, and $A_{start} = 4\pi R^2$, where R is the nominal distance from the source to the detector. Not only is the effective area an important qualifier for a telescope, but, just like the angular resolution, it is a useful benchmarking parameter that can be used to determine any issues with the analysis pipeline or DEE (Sleator et al., 2016).

The calibration measurements used for the ARM determination are also used to calculated the A_{eff}, and the event selections are the same except an additional origin cut of ±1 FWHM of the ARM distribution is used for the effective area measurement so as to select on events that are properly reconstructed. Therefore, the ARM and A_{eff} measurements are closely related and easily calculated together.

Figure 4.13 shows the preliminary calculated A_{eff} effective area as a function of energy for on-axis calibrations sources. Again, two separate event selections are shown: the blue triangles include all event within the photopeak with no origin selection. The green circles show the measured effective area with the ±1 FWHM ARM cut applied to each measurement. No error bars are included in these measurements since the proper error propagation has not yet been implemented, though we believe these measurements are dominated by the error of the source activity. One would expect the effective area to smoothly vary with the γ-ray energy, and we are working to understand the structure seen here. Currently, our measured effective area drops drastically at higher energies, and the COSI team is still investigating the reason for

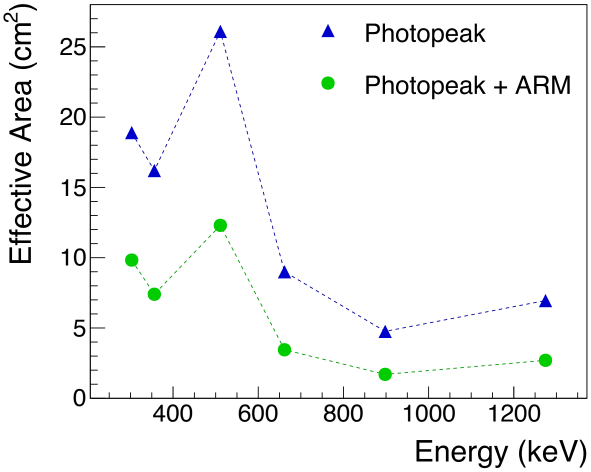


Figure 4.13: The preliminary effective area measurements as a function of energy for calibration sources. The maximum effective area is observed at 511 keV.

this. As per design, the effective area is largest for the 511 keV calibration line which is key to the science analysis considered here.

As with the angular resolution discussed in the previous section, a careful choice of event selections can optimize the effective area; however, there is a constant trade off between efficiency and quality of the event reconstruction.

4.6.3 POLARIZATION RESPONSE

Compton telescopes are inherently sensitive to polarization. According to the Klein-Nishina scattering cross section (Equation 3.1), if the incident photons are linearly polarized, we expect to see a sinusoidal modulation of counts as a function of the azimuthal angle in a Compton telescope (Lei et al., 1997). However, the polarization response is fairly difficult to validate in the laboratory setting, as we cannot easily modify the source polarization. Ideally, we would use a 100% polarized source; however, for these calibrations we are restricted to using Compton scattering to produce a partially polarized beam. Figure 4.14 shows the resulting azimuthal scatter distribution from one of these laboratory validation measurements. See Lowell (2017) for a discussion of the polarization response of COSI and measurements prior to the 2016 flight.

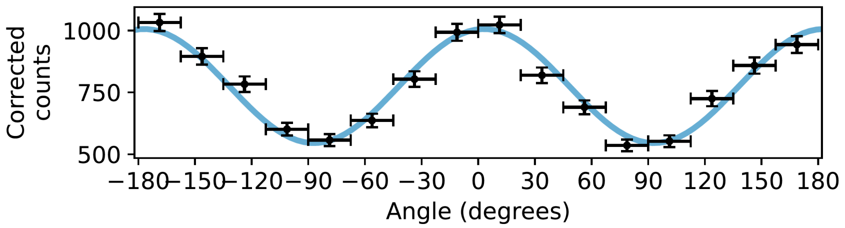


Figure 4.14: Azimuthal scatter distribution measured in COSI due to a ~50% polarized source. The polarized signal is normalized by a corresponding unpolarized source to minimize false modulations due to the geometry of the instrument. Figure 5.7 from Lowell (2017).

5

THE 2016 COSI BALLOON CAMPAIGN

The COSI 2016 flight was the first science flight from NASA's new launch location in Wanaka, New Zealand, and the first mid-latitude Mission of Opportunity with NASA's new Super Pressure Balloon (SPB) technology. The COSI team arrived in Wanaka on February 7th, 2016, and the instrument was flight ready on April 1st, 2016. The team then waited six weeks for the low surface winds that are crucial for a safe SPB launch. In this chapter[1], we will give a summary of the 2016 flight and an overview of the preliminary results.

5.1 FLIGHT OVERVIEW

COSI was launched from Wanaka (45° S, 169° E), New Zealand, on May 17, 2016 (23:35 06/16/16 UTC). See Figure 5.1 for a photograph of the COSI and SPB flight train after inflation of the balloon. Initially after the launch, the balloon floated westwards towards Australia, flying almost directly over Sydney, before shifting to lower latitudes and heading with the prevailing winds to the east. COSI underwent a full circumnavigation within 14 days, and then spent much of the remaining flight fairly stagnant above the South Pacific Ocean. The altitude drops during cold nights, the concern for the health of the balloon, and the aim to recover COSI and the balloon cut the flight short, and COSI was terminated above land 200 km north-west of Arequipa, Peru, on July 2 (19:54 07/02/16 UTC; 16° S, 72° W). See Figure 5.2 for the full path of the 2016 fight during its 46-day journey. The COSI instrument was successfully recovered on July 14th and showed no major damage. Figure 5.3 shows a photograph of the gondola at its landing site in Peru.

---

[1] Some of this chapter is based on a previous publication "The 2016 Super Pressure Balloon flight of the Compton Spectrometer and Imager" by Kierans et al. (2016). Copyright owned by the authors under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

Figure 5.1: Photograph after the Super Pressure balloon has been inflated and immediately before launch on May 17th, 2016, in Wanaka, New Zealand. COSI, hanging from the launch vehicle, is visible on the left side of the photograph, the orange parachute is in the middle, and the balloon is on the right. Only partially inflated on the ground, once the balloon reaches float altitudes it becomes fully expanded with a diameter of over 100 m. The Wanaka Airport can be seen in the background, which includes the hangar within which all integration occurred.

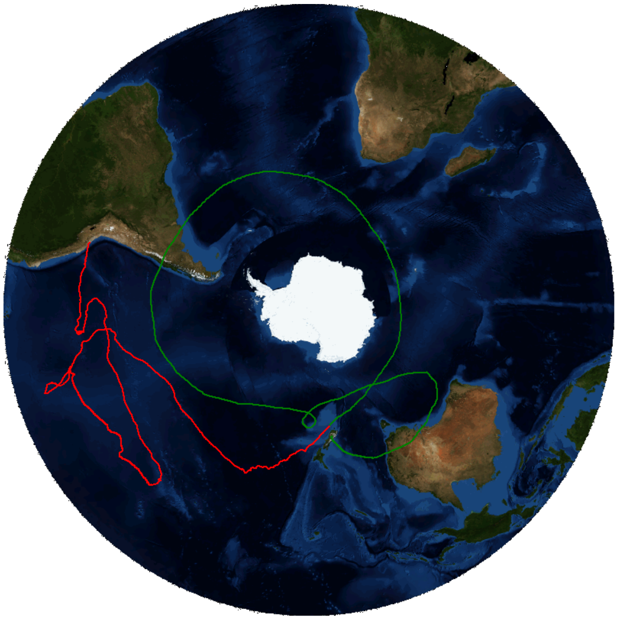


Figure 5.2: COSI was launched from Wanaka, NZ, on May 17th, 2016, circumnavigated the globe, and landed in Peru 46 days later. The first circumnavigation, completed in 14 days, is depicted in green. There is no steering available for high-altitude balloons, so the prevailing winds carry the balloon and the only trajectory choice the ground crew has is when to terminate the flight.

Figure 5.3: The instrument was located in the desert north of Arequipa, Peru, on July 7, 2016, just 5 days after termination. Although the gondola frame was bent and the antenna booms and some solar panels were destroyed, all of the COSI instrument systems have been tested after flight and are operational.

Within the initial weeks of the flight, COSI was at high southern latitudes, which increased the instrument background but introduced some interesting electron precipitation events (see Section 5.2.1). During one such event on May 30th, COSI detected a bright GRB. The flight path gave excellent exposure of the Galactic Center and a few bright compact γ-ray sources. The sources detected in the 2016 flight will be discussed in Section 5.3.

Two of the systems operated by CSBF had issues during the flight. One of the Openport antennas failed on June 11th; however, the bandwidth of the other alone was enough for the real-time telemetry. The rotator, which is required to keep the solar panels pointed at the sun, was not tracking properly for two days at the end of May, which led to an incomplete charge on the batteries; however, we mitigated any problems by going into a lower power mode for one night. CSBF was able to fix the issue and there were no further concerns.

The minimum success requirement for the COSI/SPB 2016 campaign was 14 days afloat. With 9 out of 12 of COSI's GeDs operating (see Section 5.1.3) for 46 days, the flight was declared a mission success.

Figure 5.4: A day time photograph of the SPB from the on-board camera in the 2016 COSI flight.

5.1.1 SUPER PRESSURE BALLOON

NASA has recently developed the new SPB platform designed to fly at near constant altitude through day-night cycles by maintaining a positive differential pressure. Conventional scientific helium balloons are referred to as “zero-pressure” balloons since they maintain the same pressure as the ambient atmosphere through a small opening in the bottom of the balloon. This means that during the cold nights when the temperature drops, the balloon also drops in altitude, and it never quite recovers the same altitude the following day. In contrast, the SPB is a closed system that is slightly pressurized compared to the ambient pressure. With a maximum differential pressure of 180 Pa, this small difference is enough to keep the balloon at the same altitude through the temperature range of day-night cycles, and thus the balloon can maintain altitude for a much longer duration.

The SPB is impressive; the current design weighs 5240 lb and has a price tag of over one million dollars. Once it is fully inflated with helium, the diameter of the balloon is 100 m. The thickness of the polyethylene film that makes up the balloon is only 0.04 mm. There are 280 gores, which are long, light-weight but high-strength tendons that run from the top to bottom of the balloon to support the large stresses on the material; each tendon can maintain a force of ~7000 N. The capabilities of the SPB are groundbreaking for scientific ballooning.

The volume of the current SPB design is 18 million ft^3 (MCF) balloonand it can support a maximum suspended weight of 5000 lb. COSI is the first instrument specifically designed for 18 MCF SPB flights: it is light-weight with no consumables, has autonomous observing, and has the ability to telemeter all science data in real time. Without the realization of the SPB giving balloon payloads long-duration access to the mid-latitude skies, COSI would be unable to achieve many of its science goals.

5.1.2 ALTITUDE VARIATIONS

The SPB promises constant altitudes with a goal of 100-day flights; however, the technology is still in development. The 2016 COSI flight was only the second SPB flight at mid-latitudes. COSI was accepted as the Mission of Opportunity on-board the SPB, but with this acceptance came the agreement that everything might not go smoothly. The 2016 flight was viewed, first and foremost, as a test flight of the SPB.

The balloon started showing signs of a problem in early June. It was expected that the balloon would lose a little altitude during very cold nights over ocean storms, which were experienced a few times within the first week of flight, but in June the altitude was dropping farther than the meteorologists had predicted. It was quickly apparent that the balloon had somehow developed a small leak. After mid-June, the balloon lost a few km of altitude every night, even at moderate temperatures. See Figure 5.5 for the altitude profile of the 2016 flight. Once the instrument flew over solid land, the flight was terminated to ensure recovery of COSI and the balloon for further testing.

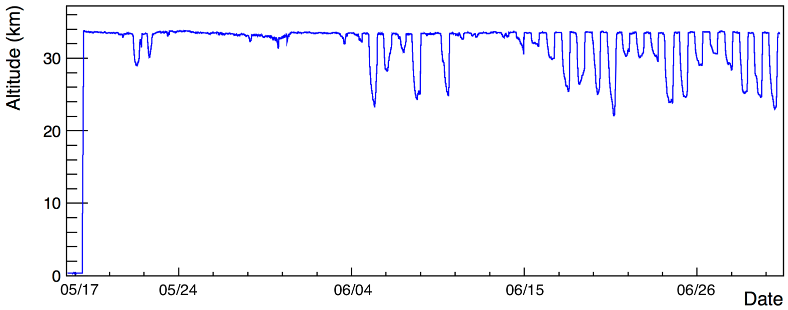


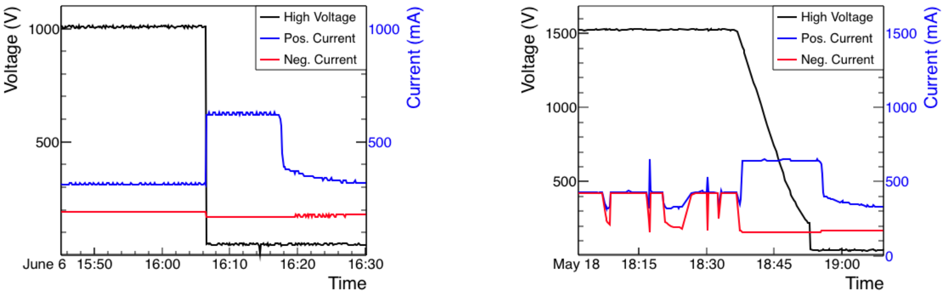
Figure 5.5: The altitude profile during the COSI 2016 flight. The nominal altitude was 34 km. After June 4th, 2016, the balloon started to drop in altitude more than the meteorologists had predicted and it was apparent that the balloon had developed a leak.

Unfortunately, for the science discussed here, the GC was in the COSI FOV only during the night hours. A decrease in altitude increases the atmospheric depth, which lowers the transmission of Galactic γ-rays through the atmosphere. The effect of this will be discussed more in Chapter 6.

5.1.3 HIGH VOLTAGE ISSUES DURING FLIGHT

While nine of the GeD detectors worked flawlessly, three failed at different times during flight. Detector #8, in the bottom layer of the array, failed within the first 12 hours, and Detector #5 in the top layer failed 24 hours later. Both had similar modes of failure. Detector #0 failed on June 6th, but exhibited a different behavior. We believe these three failures are related to the high voltage (HV), but the cause of the first two failures is not yet clear.

Detector #0 had its high voltage reading suddenly drop to zero on June 6th, 2016, and the preamps responded as if the high voltage had been ramped down, as shown in Figure 5.6 (a). Even with our limited monitoring in flight, it was clear that the HV supply had stopped working. After flight, the failure of the Detector #0 HV supply was reproduced in the laboratory vacuum chamber at SSL. The issue was traced to a flaw in the design that resulted in too much current in a critical resistor, but now the circuit has been modified and the board has been vacuum tested with no further issues.



(a) Detector #0 failure.

(b) Detector #5 failure.

Figure 5.6: (a) The HV on Detector #0 dropped suddenly to zero on June 6th, 2016. Since the preamp currents usually saturate when the HV ramps down, it was confirmed that this was a true change in voltage.. (b) The preamps on Detector #5 started to go in and out of saturation during the second day of flight; however, the HV monitor showed no issues. Detector #8 displayed the same behavior.

For Detector #5 and #8, the card cage HV monitor showed no signs of a problem, but the current monitor for the preamps displayed spontaneous high values. Figure 5.6 (b) shows the time-series of the HV monitor and preamp currents measured during the failure of Detector #5. The fast, recurrent impulses in the preamp currents had not been observed before. The detector was behaving as if the HV was being turned off and on very quickly. After the issue was detected, the card cages were turned off since the detectors were inoperable. Laboratory vacuum tests of the card cages and HV cables were done after the failures in flight, but the observed behavior could not be reproduced. The current theory as to the cause of the issue is HV breakdown in the cryostat HV feedthrough.

5.1.4 PHOSPHORESCENCE IN CSI

Once COSI reached float altitude, significant and frequent spikes in the CsI shield count-rate were observed. Figure 5.7 shows the CsI count-rate for a 1-hour segment of flight. These spikes, which are believed to be phosphorescent states induced by cosmic rays (Hurley, 1978), were also seen in the CsI detector PICsIT (Segreto et al., 2003) on INTEGRAL. For COSI, these rate spikes were initially problematic since the CsI shields were being used as the primary GRB trigger mechanism. However, this was averted quickly after launch with a slight change to the trigger algorithm by implementing a three-bin median filter that smoothed out the short duration spikes.

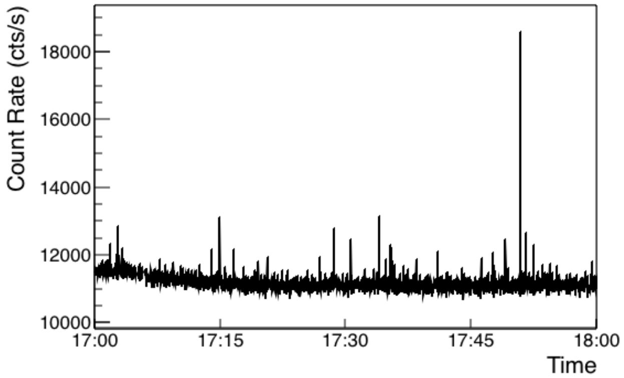


Figure 5.7: The count rate in the CsI shield detectors during one hour of flight on June 6th, 2016. The high-amplitude short-duration spikes are phosphorescent states in the CsI detectors induced by cosmic rays.

5.1.5 THERMAL PERFORMANCE

The COSI thermal environment was well within operating range for all systems. Most notably, the mechanical cryocooler operated efficiently with the thermal stabilization of the cooling system, with no detectable loss in cooling fluid. Once the instrument reached float altitude, it was quickly apparent that the liquid cooling system was actually *too* efficient. If the pump was kept on, then the cryocooler would get colder than its performance specification. This was resolved by defining an ideal temperature range for the cryocooler collar, between 30° and 40 °C, which automatically turned on and off the pump so that the cryocooler collar would stay within this range. Figure 5.8 shows the temperature of the crycooler collar for one hour in flight; the fast drops in temperature correspond to the times that the pump is on and the slow rise is when the pump is off. Figure 5.9 shows the temperature of the cryocooler collar and the fluid input and output temperature of the copper radiator for a 24-hour day-night cycle.

The passive insulation around the gondola kept the electronics bay at a comfortable temperature for the COSI subsystems. See Figure 5.10 for the temperatures of some of the subsystems during a 24-hour day-night cycle. For the elements that had more stringent temperature constraints (see Table 3.2), Kapton heaters were toggled on and off in the flight computer code, mimicking a thermostat. The flight computer CPU is one of these temperature sensitive parts, and the profile in Figure 5.10 shows the up-tick in temperature when the associated heater was turned on around 19:00 hr.

5.1.6 MONITORING

With the expectation of a 100-day flight, the COSI team developed a sophisticated monitoring system. An interactive real-time monitor, written in `python`, acted as the ground support electronics (GSE), both in flight and during calibrations. The real-time monitor allows one to send commands to the instrument and receive the detector count rates and the housekeeping data; however, one needs to be connected to the local data distributor. For a higher-level of monitoring also accessible to all COSI team members, time profiles of the housekeeping data were posted on a Web GSE, which allowed for a quick overview of the instrument's health. To mitigate the necessity of staring at a computer for 100 days, text message alerts were sent to members of the flight operations team if a parameter went out of range, if there were unexpected

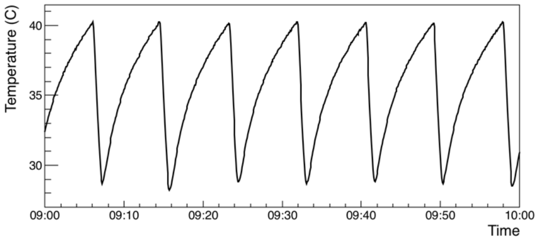


Figure 5.8: Cryocooler collar temperature over one hour of flight. The liquid cooling system was toggled on and off to keep the collar temperature between 30° and 40 °C. The pump would turn on for ~2 min to allow the flow of cold liquid to cool the collar. Then with the pump off, it would take the system ~8 min to warm up.

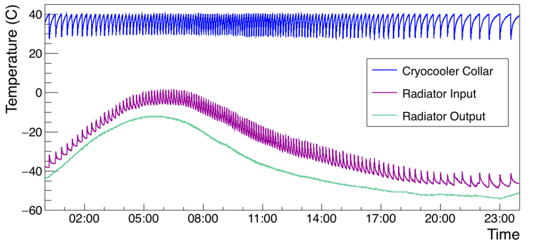


Figure 5.9: A 24-hour day-night cycle showing the temperature of the cryocooler collar, the radiator fluid input, and radiator fluid output. During the warmer parts of the day, the frequency at which the pump was turned on and off was higher than during the colder nights.

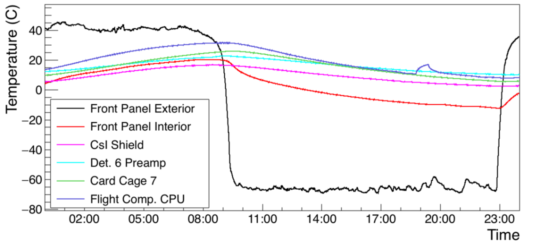
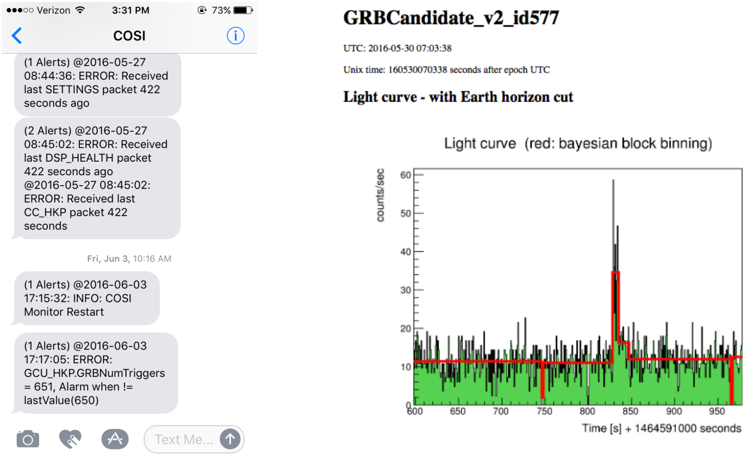


Figure 5.10: The temperature of a few of the COSI subsystems and gondola parts during a 24-hour day-night cycle. The most extreme temperatures seen were on the sun-facing front panel, shown in black. The flight computer had a 18 W heater that was turned on when the temperature of the CPU reached 10°C, which occurred around 19:00.



(1 Alerts) @2016-05-27 08:44:36: ERROR: Received last SETTINGS packet 422 seconds ago

(2 Alerts) @2016-05-27 08:45:02: ERROR: Received last DSP_HEALTH packet 422 seconds ago @2016-05-27 08:45:02: ERROR: Received last CC_HKP packet 422 seconds

Fri, Jun 3, 10:16 AM

(1 Alerts) @2016-06-03 17:15:32: INFO: COSI Monitor Restart

(1 Alerts) @2016-06-03 17:17:05: ERROR: GCU_HKP.GRBNumTriggers = 651, Alarm when != lastValue(650)

GRBCandidate_v2_id577

UTC: 2016-05-30 07:03:38

Unix time: 160530070338 seconds after epoch UTC

Light curve - with Earth horizon cut

(a) (b)

Figure 5.11: Two of the monitoring programs developed for the COSI 2016 flight. (a) Text message alerts were sent to the team if any monitored value exceeded its limit, if there were gaps in the data, or if a possible GRB was detected. (b) Real-time analysis for GRB detections allowed for the report of GRB 160530A, the first GRB reported to the GCN from a balloon.

drops in data, or if there was a potential GRB trigger. See Figure 5.11 (a) for examples of a few text messages that were sent from COSI. A real-time automated GRB analysis program was also developed that produced light curves and images of a possible triggered GRB; see Figure 5.11 (b). It was through this automated monitoring program that COSI was able to be the first balloon telescope to submit a Gamma-ray Coordination Network (GCN) Notice (Section 5.3).

5.2 COSI OBSERVATIONS OVERVIEW

The COSI 2016 flight was record-breaking. Not only was this the first mid-latitude science flight with NASA's new SPB, but also it was the longest mid-latitude balloon flight for a large scientific balloon. Although the altitude drops and the three lost detectors affected the quality of observing, the data from the 2016 flight hold the promise of interesting scientific results. COSI had excellent exposure of the GC and the Galactic plane, and when the instrument floated close to the equator in mid-June, COSI had a few days of good exposure of Cygnus X-1 (Cyg X-1) and the Crab Nebula (see Section 5.3). Fig-

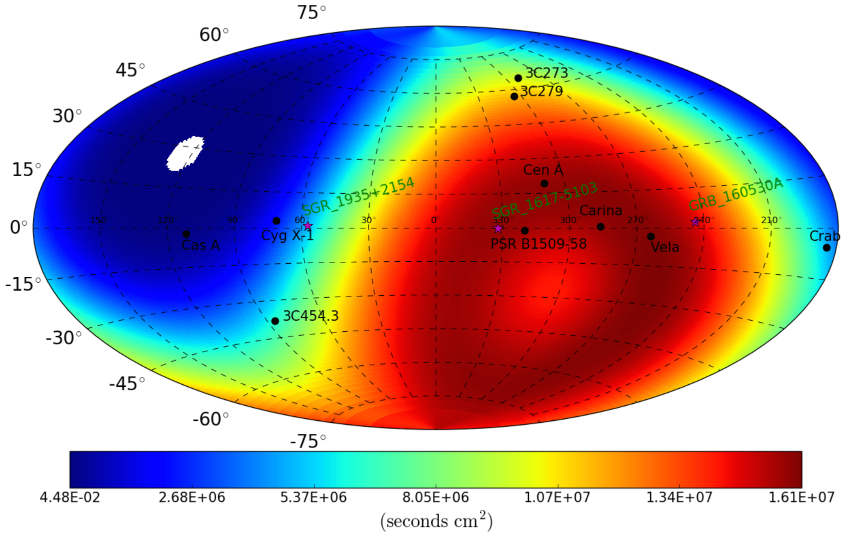


Figure 5.12: Exposure map from the 2016 COSI flight assuming an energy-independent effective area of 20 cm^2. COSI had excellent exposure of the GC and Galactic plane, as well as some of the brighter γ-ray compact objects in the sky, labeled here.

ure 5.12 shows the exposure map from the entire flight, where the effective area is assumed to be 20 cm^2. The count rate for a top GeD, which gets the best exposure, is shown in Figure 5.13. There are a number of different features that can be seen in the count-rate profile. In the first 15 days of flight, the slow modulation shows the change in background radiation levels due to changes in latitude, and the largest count-rate spikes are from electron precipitation events when the instrument was closest to the South Magnetic Pole. In the latter half of the flight, the count rate increased when the altitude dropped at night. The effect of the latitude and altitude changes and the impact on background radiation will be further discussed in Section 6.1.1.

The total integrated spectrum from the entire flight is shown in Figure 5.14. Instrument activation lines and the prominent (mostly) background line at 511 keV from β^+-decay in the atmosphere are labeled.

5.2.1 RELATIVISTIC ELECTRON PRECIPITATION EVENTS

COSI observed several relativistic electron precipitation (REP) events, which are caused by high-energy electrons from the Earth's radiation belts that precipitate into the atmosphere after a geomagnetic storm. The electrons collide

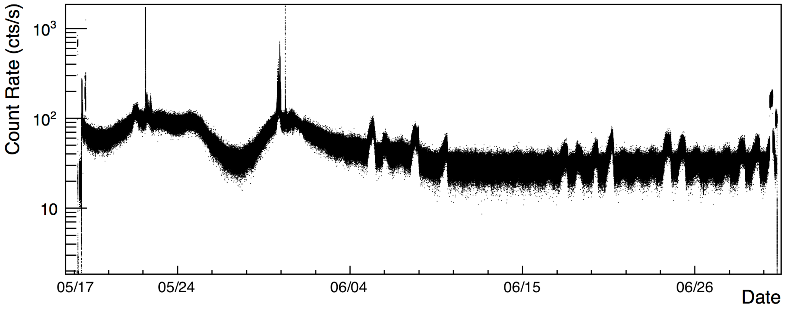


Figure 5.13: Detector rates of one of the top GeDs over the duration of the flight. The initial slow variations are due to changes in latitude, and the sharper variations in the latter half of the flight are from altitude drops at night. Multiple intense REP events are seen at the highest latitudes when background was largest; GRB 160530A was observed during the second of these events.

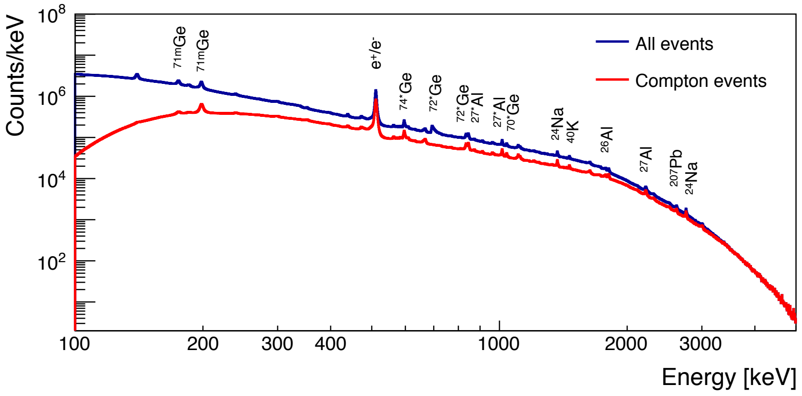


Figure 5.14: Total integrated energy spectrum, including single-site and multi-site events, from the duration of the flight. The intense 511 keV atmospheric background line and known activation lines have been identified.

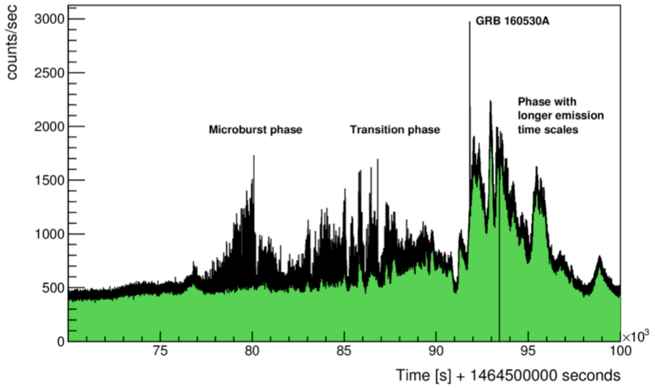


Figure 5.15: The light curve of one of the microburst events, characterized by short bursts, detected with COSI. This event shows a transition to a phase with longer timescales. GRB 160530A was detected during this transition phase.

with atoms in the atmosphere and emit γ-rays and X-rays through bremsstrahlung radiation (Millan et al., 2002). This marks the first time these events have been detected with a wide-FOV imaging γ-ray detector. REP events are divided into two main categories: duskside relativistic electron precipitation (DREP) and microburst precipitation. DREP events, as obvious from the name, occur mostly during dusk hours and are characterized by ultra-low frequency flux changes. Microbursts are rapid, intense bursts lasting ~250 ms; see Figure 5.15 for one such event observed by COSI. There are still many open questions regarding REP events. The mechanism responsible is not well understood, and the scale size and relative importance of these events to radiation belt depletion are still unknown. COSI's time resolution, spectral resolution, imaging and polarization capabilities, unprecedented in previous REP detections, are promising for further understanding these events. See Millan & Thorne (2007) for a review of REP events.

5.3 COSI 2016 DETECTED SOURCES

Long-duration GRB 160530A was detected by COSI on May 30th, 2016, coinciding with an REP event. See Figure 5.16 for the COSI image and the light curve of this GRB. The real-time automated GRB analysis program was triggered (Figure 5.11) and minutes later the COSI team recognized the event as a GRB candidate. Within 4 hours, the COSI team sent out a *Swift* Target of

Opportunity for X-ray and optical followup. Unfortunately, no afterglow was found. The COSI team sent out a GCN circular for the GRB 15 hours after the initial trigger (Tomsick & the COSI team, 2016), marking COSI as the first balloon payload to do so. The GRB was later reported by Konus-Wind, INTEGRAL/ACS, and *Astrosat*/CZTI, which allowed for an absolute timing calibration of the COSI data and a comparison of the measured spectrum. Throughout the flight, a number of other GRBs were detected in the COSI CsI shields, but none other than GRB 160530A were bright enough for detatiled analysis in the GeDs.

One of the main science goals for COSI is to perform γ-ray polarization measurements of GRBs. Polarization measurements offer unique information about the emission mechanism, the magnetic field, and the outflow geometry, and a sample of measurements of γ-ray polarization could help distinguish between competing GRB models (Toma et al., 2009). Only a limited number of GRB γ-ray polarization measurements exist to date (McConnell, 2017). Lowell et al. (2017a) developed a Maximum Likelihood Method (MLM) analysis to increase the sensitivity of polarization measurements with COSI. Applying this new technique to GRB 160530A, Lowell et al. (2017a) report a 90% upper limit on the polarization level of 46%, as shown in Figure 5.17. COSI is the only balloon-borne instrument to have reported a GRB polarization result, and this strong upper-limit is a significant addition to the small number of measurements currently available.

Three compact γ-ray sources were detected during the 2016 flight: the Crab Nebula, a SNR that is the brightest X-ray and soft γ-ray compact source in the sky; Cyg X-1, an HMXB classified as a microquasar and known to be persistently bright; and Centaurus A (Cen A), also known as NGC 5128, one of the closest and brightest AGN in the sky. All three of these objects have been extensively studied by other telescopes, so they are good benchmarking tools for COSI to prove its detection abilities. Even though the Crab was only high in COSI's FOV for a few days, as it is the brightest of these sources, it has been the test case for developing the COSI analysis tools. Work is ongoing to build the tools to perform an accurate background subtraction and to enable polarization measurements for compact objects (Sleator et al., in prep). Once the pipeline has been developed for the Crab data, it will be applied to the other detected sources. The main scientific goal of these studies will be to produce polarization measurements of these three compact objects.

The Galactic positron annihilation emission detected during the 2016 COSI flight will be discussed in Part III.

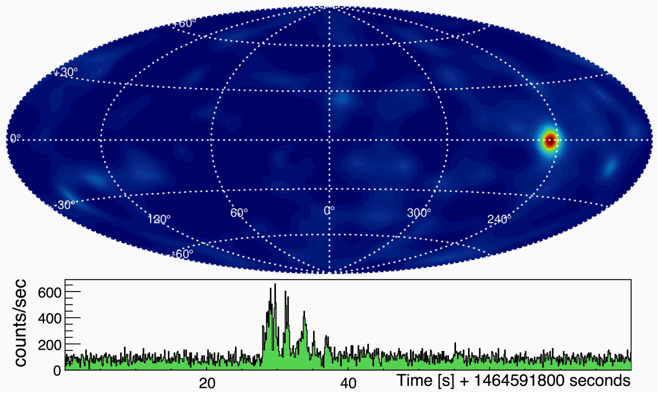


Figure 5.16: The image and light curve for GRB 160530A as detected by COSI during the 2016 flight. The image is obtained after 10 iterations of the LM-ML-EM algorithm. The GRB had a duration of ~16 s, and it was 43.5° off-axis in COSI's FOV. The detection was coincident with an REP event.

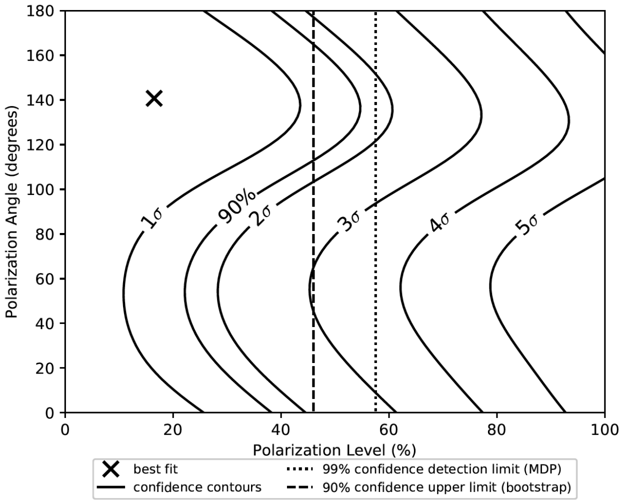


Figure 5.17: Confidence contours for the polarization of GRB 160530A derived from the new MLM technique developed for COSI polarization analysis (Lowell et al., 2017a). The cross signifies the best fit polarization level, but only an upper limit of 46% can be placed on this measurement. Figure 6 from Lowell et al. (2017b).

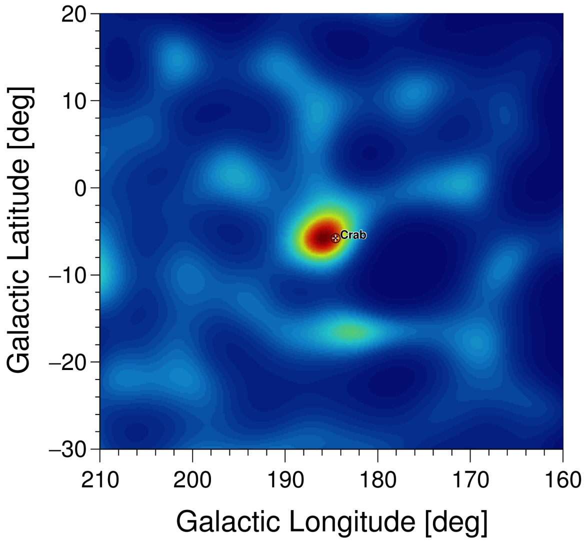
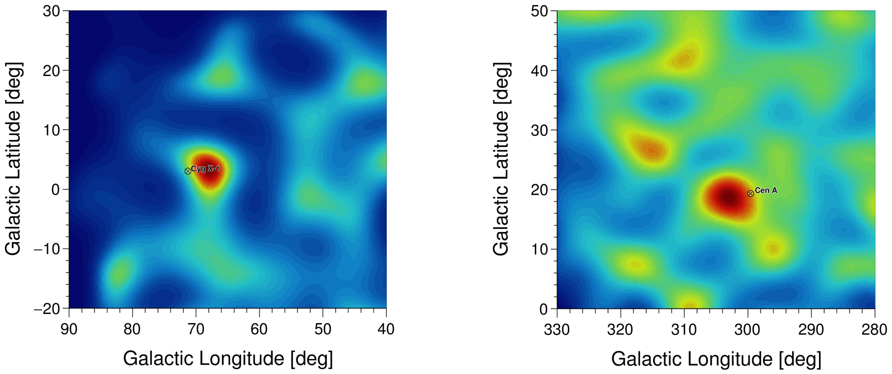


Figure 5.18: Image of the Crab Nebula from the 2016 COSI flight after 40 iterations of the LM-ML-EM algorithm. Only the two days when the Crab was highest in COSI's FOV, June 12th and 13th, were used. This image was produced with γ-ray energies from 0–480 keV and 530–1500 keV to cut out the large 511 keV background line. All of the point sources detected by COSI are offset from the true position by a few degrees, perhaps due to the lack of exposure correction in the LM-ML-EM imaging.



(a) Cyg X-1 image.

(b) Cen A image.

Figure 5.19: Images of the two other compact objects detected in the COSI 2016 flight: Cyg X-1 and Cen A. The event selections used for these images are the same for the Crab image above; however, both of these images were made after 60 iterations of the LM-ML-EM algorithm.

PART III

ANALYSIS OF THE GALACTIC POSITRON ANNIHILATION SIGNAL BY COSI

6

OBSERVATION OF THE GALACTIC CENTER REGION

The southern latitudes provide the excellent exposure of the GC region that is necessary for Galactic 511 keV studies. From the 2016 flight, COSI had a total of 1.6 Ms of exposure of the GC region, considering times when the GC was within 40° of COSI's zenith. Figure 6.1 shows the elevation of the GC for the duration of the flight, where 90° corresponds to COSI's zenith, i.e., directly overhead, and 0° represents the horizon. Also indicated on this plot are the times that the altitude of the payload descended below 32 km, depicted in red. The GC was in COSI's FOV only during the night hours when the altitude would drop due to the lower temperatures. At lower altitudes, there is more attenuation of the γ-rays in the atmosphere and this degrades the observations; at the expected float altitude of 33 km, the nominal transmission probability at 500 keV is 49%, but at 27 km the transmission probability is only 18%. The total GC exposure time when the altitude was above 33 km was only 610 ks, and we estimate that the altitude drops correspond to an additional loss of 15% of 511 keV γ-rays from the GC compared to the nominal altitude.

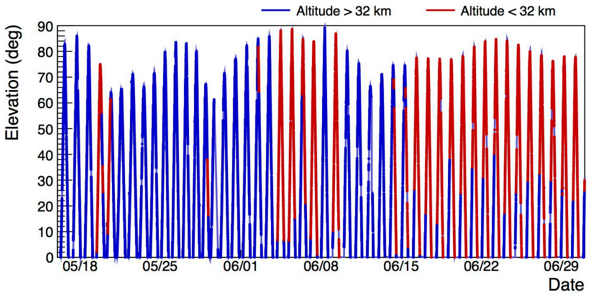


Figure 6.1: Elevation of the GC in COSI's FOV during the 2016 flight. An elevation of 90° degrees occurs when the GC is directly overhead and consistent with COSI's zenith. The times in which the altitude dropped below 32 km are indicated in red.

An LM-ML-EM image of 511 keV emission from the Galaxy can be easily obtained from `mimrec`. Figure 6.2 shows the COSI image of the 511 keV emission after 40 iterations of the ML-EM algorithm. Specific event selections were chosen to make the central emission more prominent; see Table 6.1 for a list of the event selections used to create this image.

There are a number of caveats regarding this image. As discussed in Section 3.2.4, the LM-ML-EM algorithm can not be corrected for exposure. Therefore, the bright, ring-like emission seen to the right and bottom of the GC is potentially due to a difference of exposure over this region. This could also be the explanation for the offset of the emission peak from the GC. Furthermore, the iteration algorithm is optimized for point source distributions, so with enough iterations, an extended source will be reduced into a point-source-like object. Therefore, one can not determine the true spatial distribution of the emission from this image. Recent work by Zoglauer et al. (in prep) has been to develop a binned-mode imaging algorithm for COSI analysis to address these issues. This will be further discussed in Chapter 9. Regardless, it is clear from this image that the 2016 COSI flight resulted in a strong detection of the 511 keV emission from the GC region.

The studies presented in Part III will focus on obtaining the Galactic positron annihilation spectrum and measuring the morphology of the emission through the ARM distribution. To extract the spectrum of the Galactic emission, significant background subtraction must be done.

6.1 SUMMARY OF THE BACKGROUND RADIATION ENVIRONMENT

One of the difficulties of MeV γ-ray astrophysics is the large background radiation involved, especially at balloon altitudes. Figure 5.14 shows the total accumulated spectrum from the COSI 2016 flight. The majority of photons in this spectrum are from atmospheric emissions, i.e., γ-rays from cosmic-ray interactions in the atmosphere. The intensity of this background emission is dependent on atmospheric depth, geomagnetic cutoff rigidity, and zenith angle.

In Figure 6.3 (a), the total count rate in the CsI shields is plotted for the entire flight; this serves as a measure of the background radiation intensity. Figure 6.3 (b) shows the altitude, and relatedly, the atmospheric depth, during the flight. The atmospheric depth x, or interaction depth, is a measure of how

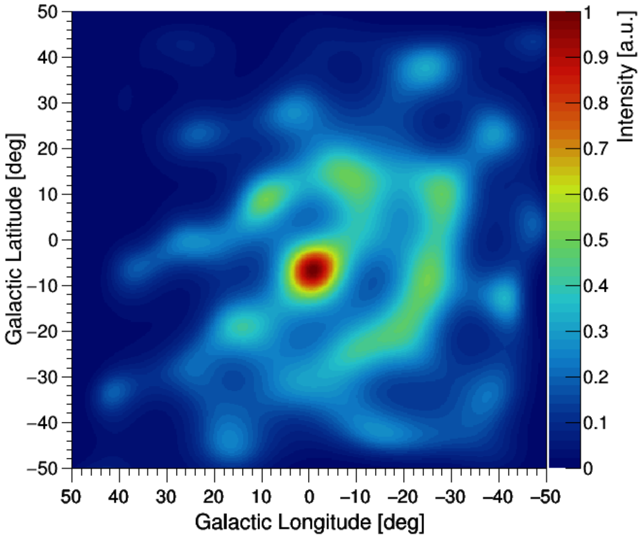
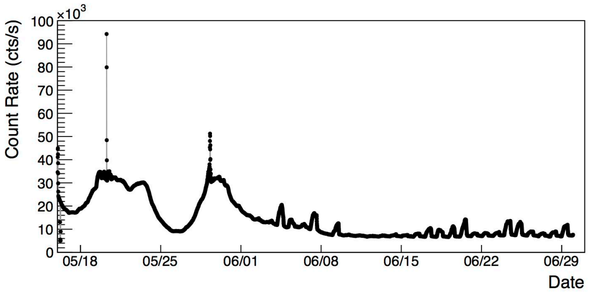


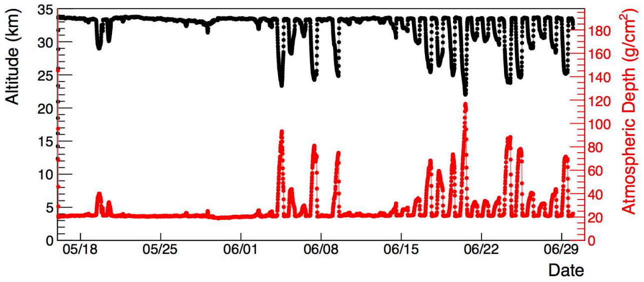
Figure 6.2: LM-ML-EM of the 511 keV emission from the GC during the 2016 COSI flight, after 40 iterations. Although the emission is very prominent, this image is not exposure-corrected, and due to the iteration algorithm used, the spatial distribution here is not truly representative of the emission.

Parameter	Allowed Range
Pointing selection	< 60 degrees of GC
Altitude	> 31 km
Photon energy	507–514 keV
Number of interactions	2 - 7
Compton scatter angle	10 - 35
Distance between first 2 interactions	> 0.5 cm
Distance between any interaction	> 0.0 cm
Earth horizon cut	Reject if 99% of Compton circle is below horizon

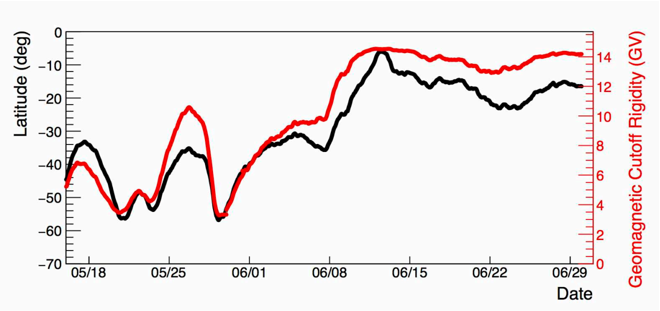
Table 6.1: Event selections for the LM-ML-EM 511 keV image shown in Figure 6.2. Refer to Section 3.5.5.1 for an overview of the `mimrec` event selections.



(a) CsI shield count rate.



(b) Altitude and atmospheric depth.



(c) Latitude and geomagnetic cutoff rigidity.

Figure 6.3: (a) CsI shield count rate for the entire flight, in one-hour averaged increments. The slow modulations at the beginning of flight are due to changes in the latitude, and the more frequent increases in flux during the latter half of flight are due to altitude drops at night. (b) The altitude and atmospheric depth for the duration of the flight, averaged over one-hour increments. CSBF provided the pressure measurements included in the atmospheric depth calculation. (c) Geographic latitude and geomagnetic cutoff rigidity for the duration of the flight. When the latitude is close to the South Magnetic Pole, the cutoff rigidity is the low, and the background radiation from cosmic-ray interactions in the atmosphere is large.

much material the γ-rays must pass through. It can be calculated from the altitude, but can be found more simply from the pressure P in the atmosphere $x = P/g$, where g is the acceleration due to gravity. A larger atmospheric depth will cause more absorption and also gives more material for the cosmic rays to interact with. Figure 6.3 (c) shows the geographical latitude of COSI throughout the flight, and the related geomagnetic cutoff rigidity. The cutoff rigidity is a measure of the magnetic shielding of the Earth from cosmic rays. The geomagnetic cutoff rigidity, estimated here as only the vertical arrival direction, is given by

$$R_{cv} = \frac{14.27 \cos^4 \lambda}{r^2}, \tag{6.1}$$

where λ is the geomagnetic latitude and r is the radius in units of the Earth's radii. This is a measure of the minimum rigidity, or momentum, that a particle must have to reach that given position. At higher latitudes, i.e., closer to the South Magnetic Pole, the cutoff rigidity is lower and therefore the background from cosmic rays is higher. Both the atmospheric depth and the cutoff rigidity affect the background environment, and the consequence of both can be seen in the CsI shield rates shown in Figure 6.3 (a).

Another component that affects the background environment is activation. When the instrument is bombarded with protons, neutrons and other cosmic-ray particles in the upper atmosphere, nuclear reactions will be induced within the instrument material and the radioactive isotopes that are produced will then decay. Decay products from isotopes with a half-life that is less than the timing resolution of the detector will most likely be vetoed by the CsI since the high-energy cosmic-rays are likely to pass through the GeDs and interact in the shields. Isotopes that have a half-life longer than the timing resolution of the detector, but less than the flight duration, will eventually decay, and if the emitted γ-ray is detected in the GeD, the decay could appear as a Compton event. Some of these activation lines, which are mostly from germanium, are labeled in Figure 5.14. For β^+-unstable isotopes, the emitted positron will quickly annihilate within the instrument material and produce two 511 keV γ-rays. If one of these 511 keV photons escapes the detector without interacting in the CsI shields, and the other is absorbed in the GeDs, then the signal will be indistinguishable from a cosmic 511 keV photon.

6.1.1 OVERVIEW OF BACKGROUND SUBTRACTION

The goal of background subtraction is to obtain the true spectrum of the source. Since MeV observations are dominated by background radiation, one must first determine the spectral signature of the background underlying the observation so that it can be accurately subtracted off to reveal the spectrum of the source. One of the simplest ways to achieve this would be to determine the total spectrum when the source is within the FOV and subtract off the accumulated spectrum when the source is not within the FOV, after a normalization to account for different levels of exposure. However, for the dynamic background radiation observed in COSI's flight, as discussed in the previous section, this method proved to have unquenchable systematics. It became apparent that the background and source spectrum needed to be determined simultaneously.

With COSI's imaging abilities, the next most straightforward method is to select two regions in the sky concurrently, obtain the spectrum from each one, and perform an image-space background subtraction. Since the background emission depends on the zenith angle, it is more appropriate to choose a background region that has the same zenith angle as the source in COSI's FOV, as opposed to a background region that is separated from the source region by a certain degree in Galactic coordinates. Once an appropriate background region is selected, the spectra from the source region and the background region can be subtracted to obtain the emission from the source. To reduce the statistical uncertainty in the subtracted spectrum, a large background region is chosen.

Figure 6.4 is a schematic illustrating the choice of background regions used for the COSI image-space background subtraction. The red region represents the position of the GC with a origin selection of 16°, meaning any event whose Compton circle is within 16° of the GC location will be included in the spectrum for this region. The blue regions represent the 9 positions that have the same zenith angle as the source region. The spectra data from these 9 regions are averaged to get an accurate representation of the background radiation. The sky position that these background regions survey is constantly changing as the source region moves through the FOV of the telescope. The separation of each background region is azimuthally 30°, and the two regions separated by ±30° from the source position are not included to reduce the overlap between the source and background regions. As will be shown in the next few sections, this has not been the best option for Compton telescopes. Chapter 7

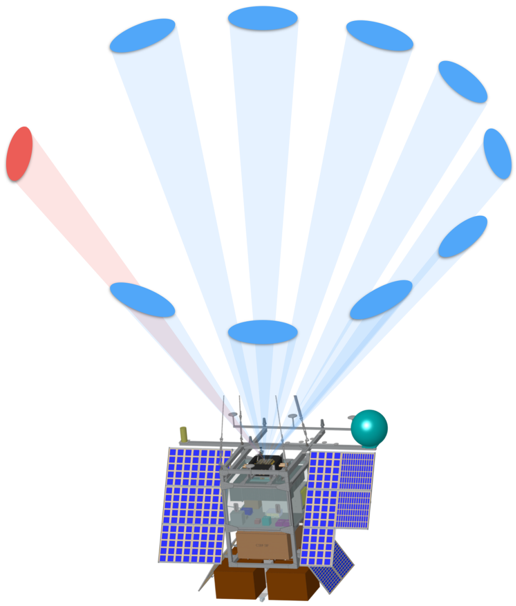


Figure 6.4: Schematic of the image-space background subtraction method. The red region in the sky represents the position of the GC, and the blue regions are the chosen background regions. All selected regions have the same zenith angle in COSI detector coordinates and are each separated azimuthally by 30°, except for the region on either side of the GC where there is a 60° gap. The regions, which have a radius of 16°, have not been drawn to scale.

will introduce the COMPTEL Data Space, which we will show is a much more appropriate method for Compton telescope background subtraction, but to motivate that choice, the image-space subtraction will be discussed below.

6.2 IMAGE-SPACE BACKGROUND-SUBTRACTED GC 511 KEV SPECTRUM

For every Compton event detected by COSI when the GC was within 40° of COSI's zenith, the reconstructed Compton circle was checked for consistency with each of the 10 regions shown in Figure 6.4. If consistent, the energy of the event would be added to the spectrum of that region. As the GC location changed in COSI's FOV, the background regions would change appropriately so as to always have the same zenith angle as the source region. After the spectra for each region were accumulated from the full flight dataset, the background spectra were averaged to get a low-variance single background spectrum.

Figure 6.5 is the image-space background-subtracted spectrum of the Galactic positron annihilation measured in flight for a 16° radius around the GC. The 16° radius of the source and background regions were chosen to optimize the significance of the signal, though it is larger than the modeled FWHM from SPI analyses. Additional event selections were used to make the 511 keV line feature more prominent; see Table 6.2 for a list of these selections. The gray histogram in Figure 6.5 is the total spectrum within the 16° origin selection around the GC, before the spectrum subtraction. The number of photons between 506–516 keV in the background spectrum shown in gray is 20.8×10^3 cts, and after subtraction we recover 458±138 photons from the same region around the GC. This corresponds to a ~3.3 σ detection of the positron annihilation signal, where the signal-to-noise is ~2%.

We can estimate the number of 511 keV γ-rays we expect to detect in the 2016 flight by:

$$N_{511} = \text{Flux} \times \text{transmission probability} \times A_{eff} \times \text{Exposure time}. \quad (6.2)$$

Assuming a 511 keV flux of 1×10^{-3} from the GC region, an effective area of 1.5 cm^2 (Section 4.6.2), an exposure time of 1.6 Ms, and an average transmission probability of 50%, we would expect to detect ~1200 photons. Although there is a clear line signature at 511 keV in the subtracted spectrum, it seems as if we are not seeing the entire signal.

The subtracted spectrum is fit with a single Gaussian, and the fit gives a line centroid at 511±1 keV, with a width of 3.1±0.8 keV σ. The reduced χ^2 of this fit is 0.7. There is no significant detection of the o-Ps flux below 511 keV. The method for estimating the error bars is discussed below.

The Compton circle from a single event can be consistent with the 16° region around the GC, in addition to one or more of the background regions, and therefore the energy from a single event can be counted in multiple spectra. This "double-counting" needs to be taken into account in the error projection. Following the discussion in Bandstra et al. (2011), we calculate the error on each bin of the subtracted spectrum considering that the events in the source and background spectra are not independent Poisson processes. We define the label $GC + k$ for an event which intersects the GC region as well as k background regions, and $\overline{GC} + k$ for an event which intersects k background regions and not the GC region, where k, ranging from 0 to 9, is

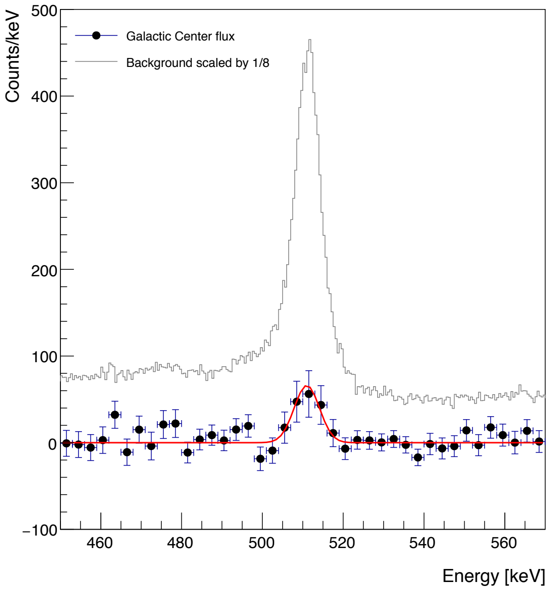


Figure 6.5: Image-space background-subtracted 511 keV spectrum from a 16° region around the GC. The gray histogram is the spectrum of the 16° region around the GC before background subtraction, which has been scaled by $\frac{1}{8}$. The error bars are determined through the N_{GC} method as discussed in the text. The significance of the 511 keV line is 3.3 σ.

Parameter	Allowed Range
Pointing selection	< 40 degrees of GC
Altitude	> 27 km
Origin selection	16°
Photon energy	N/A
Number of interactions	3 - 7
Compton scatter angle	10 - 45
Distance between first 2 interactions	> 0.4 cm
Distance between any interaction	> 0.0 cm
Earth horizon cut	Reject if 99% of Compton circle is below horizon

Table 6.2: Event selections for the image-space background-subtracted spectrum shown in Figure 6.5. Refer to Section 3.5.5.1 for an overview of the `mimrec` event selections.

the number of background regions considered. The number of excess counts can be written:

$$\begin{aligned} N_{excess} &= N_{GC} - N_{bkg} \\ &= N_{GC} - \frac{1}{9}\sum_{i=1}^{9} N_i \\ &= N_{GC+0} + \sum_{k=1}^{9}\left(1 - \frac{k}{9}\right) N_{GC+k} - \frac{1}{9}\sum_{k=1}^{9} k N_{\overline{GC}+k} \end{aligned} \tag{6.3}$$

for each of the i background regions and considering an event is counted in k regions. The statistical uncertainty for the excess counts is then

$$\sigma^2_{excess} = N_{GC+0} + \sum_{k=1}^{9}\left(1 - \frac{k}{9}\right)^2 N_{GC+k} + \sum_{k=1}^{9}\left(\frac{k}{9}\right)^2 N_{\overline{GC}+k}. \tag{6.4}$$

See Bandstra et al. (2011) for a more detailed derivation of these equations.

For each bin in the subtracted spectrum, N_{excess} is determined from Equation 6.3 and the error is calculated with Equation 6.4, where N_{GC+k} and $N_{\overline{GC}+k}$ values are determined for each energy bin. An example of the counts in N_{GC+k} and $N_{\overline{GC}+k}$ for one error bar calculation is shown in Figure 6.6. These histograms include all events which satisfy the selections shown in Table 6.2, but only for the selected energy range of 510–511 keV. Similar histograms were obtained for all bins in the subtracted spectrum. For this particular example, there are only 114 events which are consistent with the only GC, while there are 474 events which are consistent with all 10 regions. This figure conveys some of the drawbacks of the image-space background subtraction, as will be discussed in the next section.

6.3 DRAWBACKS OF IMAGE-SPACE SUBTRACTION

After looking at Figure 6.4 and the numbers in Figure 6.6, one major drawback of the image-space background subtraction may already be clear; that is, there are times when the GC is close to COSI's zenith and the background regions start to overlap with the chosen source region. Since the regions are 16° in radius, this starts to occurs when the elevation of the GC is above ~60°, and there is a complete overlap when the source is at COSI's zenith. This makes the differentiability of the source and background events effectively impossible. An attempt was made to use fixed background regions that are separated in Galactic coordinates during the times when the GC had the highest elevation,

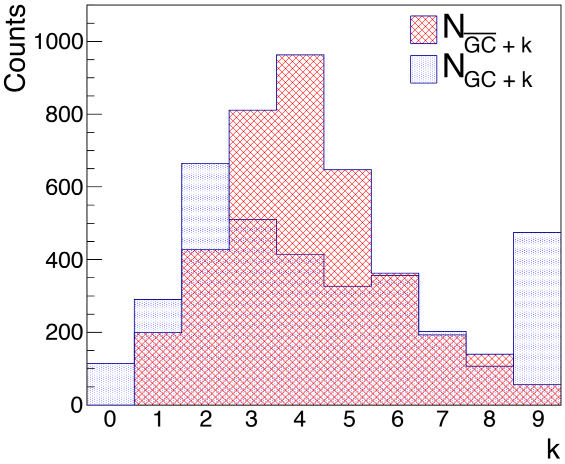


Figure 6.6: The number of events consistent with k background regions, but not the GC region, is shown in red, while the number consistent with the GC and k background regions is displayed in blue. These are all of the events from flight with E_γ between 510–511 keV that satisfy the selections listed in Table 6.2. Similar histograms were obtained for every 1 keV bin to calculate the error in the subtracted spectrum shown in Figure 6.5.

but the zenith dependence of the background made this difficult. Ultimately, many of the events were counted in both the source region and background region; therefore, a significant number of good events were lost.

In general, since each event is represented by its projected event circle in image-space, the same event will often be consistent with multiple background regions, even when the regions do not overlap with the source location. Choosing 9 regions for the background had the added advantage of reducing the variance in the background spectrum, but it increased the likelihood that the event circle from a good source photon would overlap with a background region, and thus decreased the significance of the detection.

To prove that this choice of background subtraction method was limiting the number of detected 511 keV photons from the GC region, a simple GC source simulation was performed. The source spatial distribution was based on the Skinner Model, which included the two Gaussian bulge components and the point source but not the disk component (see Table 2.1). The source spectrum was chosen to be a Gaussian line at 511 keV with a width of σ = 2 keV. The simulation was performed using the aspect information from the flight and no background radiation was included. The output of

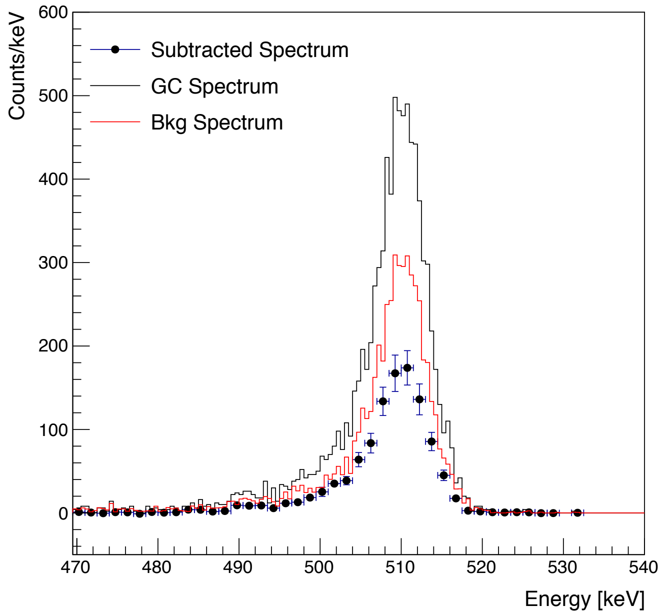


Figure 6.7: Image-space background-subtraction of a simulated Gaussian source at the GC with FWHM of 16° and a mono-energetic spectrum at 511 keV. The GC spectrum includes all events consistent with a 16° origin cut around the GC source location. Since there is no simulation of the background radiation, any event in the background spectrum, shown in red, is actually from the simulated GC source.

the simulation was then analyzed with the same program that produced Figure 6.5. Figure 6.7 shows the background-subtraction result from this GC source simulation. The spectrum from a 16° origin cut around the GC is plotted in black, along with the averaged spectrum from the 9 background regions in red, which are defined in the same way as discussed above. As there is only a source at the location of the GC and no background radiation in this simulation, any counts in the background spectra are source events whose event circles are consistent with one or more of the background regions. Only 37% of events from the simulated source are recovered in the subtracted spectrum. Therefore, we can conclude that the image-space subtracted spectrum from the 2016 flight shown in Figure 6.5 is only recovering approximately one-third of the available 511 keV γ-rays from the Galactic positron annihilation.

7

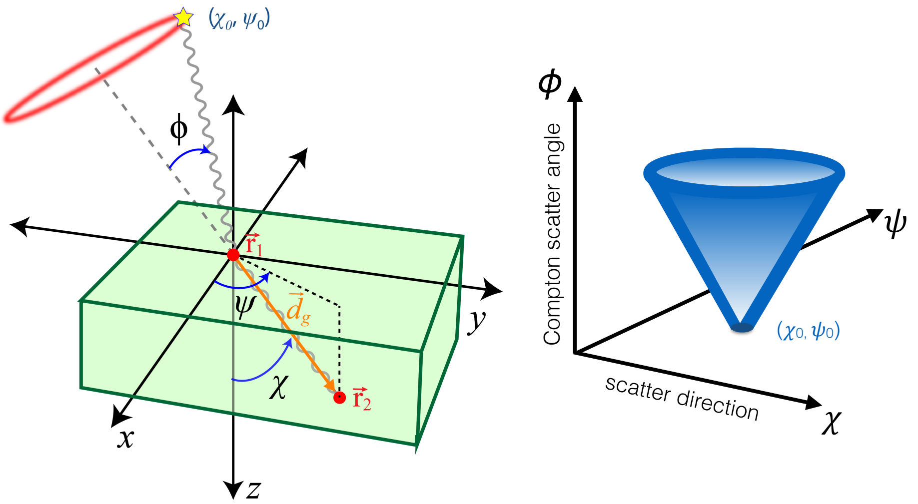
COMPTEL DATA SPACE

The COMPTEL collaboration pioneered the analysis tools for Compton telescopes. In particular, they performed the majority of their analysis in a three-dimensional data space, referred to here as the COMPTEL Data Space (CDS). The CDS is introduced in Section 7.1. First, we will show how events are defined in the CDS relative to the familiar image-space. We will then walk through the simplified case of an on-axis source where the CDS can be reduced to 2D. And finally, we will demonstrate that the analysis of a source at any location can be simplified by transforming each event into the 2D CDS with a known rotation matrix.

The motivation for using the CDS is that it allows for accurate background subtraction since each event is represented as a point instead of the event circle in image-space. The background subtraction routine developed here for the positron annihilation spectral analysis is outlined in 7.2. The method is validated with detailed background simulations and the recovery of the correct flux of a simulated Galactic point source, described in Section 7.3. Through this background subtraction, we can also get a measure of the spatial distribution of the source, detailed in Section 7.4. Furthermore, the CDS leads naturally to a binned-mode imaging algorithm, as will be discussed in Section 7.5.

7.1 INTRODUCTION TO THE COMPTEL DATA SPACE

The three orthogonal axes of the CDS are defined by the polar and azimuthal angles, χ and ψ, respectively, of the first Compton scatter direction $\vec{d}_g$ in detector coordinates, and the Compton scatter angle ϕ of the first interaction. The total energy of the γ-ray can be considered the fourth dimension of this data space, though this will not be explicitly written here. In Figure 7.1 (a), a schematic illustrates the definition of the three CDS angles for a single event. In contrast to the projected event circle in image-space, each Compton event



(a) Compton event with source at (χ_0, ϕ_0) in image space.

(b) Accumulation of events with source at (χ_0, ϕ_0) in CDS.

Figure 7.1: (a) Schematic diagram of the first two interactions of a Compton event showing the CDS angles. The source is at (χ_0, ϕ_0), and in the far field, the radius of the Compton circle in image-space is equal to the Compton scatter angle ϕ, which defines one of the axes of the CDS. The polar and azimuthal angles, χ and ψ, of the Compton-scattered γ-ray direction $\vec{d}_g$ define the two other axes of the CDS. (b) Once many photons from the same source are accumulated, with each event represented as a point at (χ, ψ, ϕ), the source is mapped as a cone in the CDS with its apex at the source position (χ_0, ψ_0).

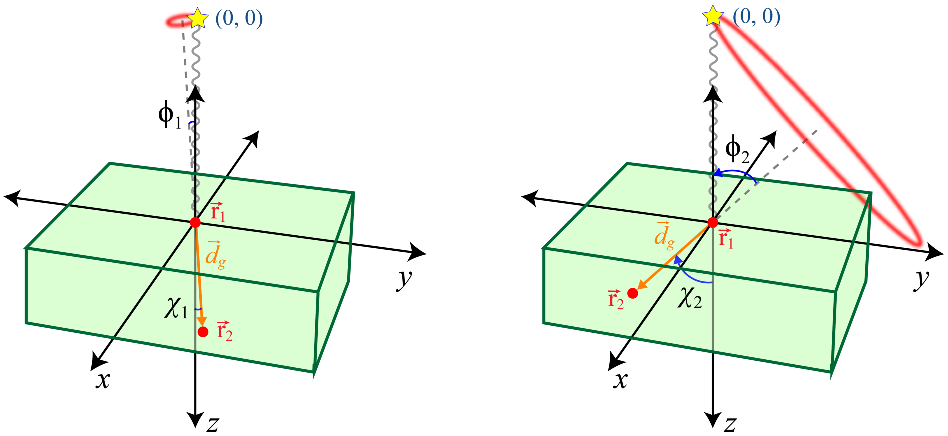
is a point in the CDS at (χ, ψ, ϕ). The accumulation of Compton events from point-source emission in the instrument's FOV, depicted by the yellow star at (χ_0, ψ_0) in Figure 7.1 (a), fills the CDS as a 3D cone. The CDS cone has its apex at the source position (χ_0, ψ_0) in detector polar coordinates because in the limit that $\phi \rightarrow 0$, $\vec{d}_g$ will point towards the source location. The opening angle of the cone is 90° since the Compton scatter angle and polar scatter direction increase at the same rate; see Figure 7.1 (b) for a drawing of this cone shape in the CDS. It is important to note that although only the scatter direction and Compton scatter angle are needed to describe the CDS, all other parameters of each Compton event are stored for event selections, such as the total γ-ray energy, the distance between interactions, and the aspect information.

7.1.1 ON-AXIS SOURCE IN THE CDS

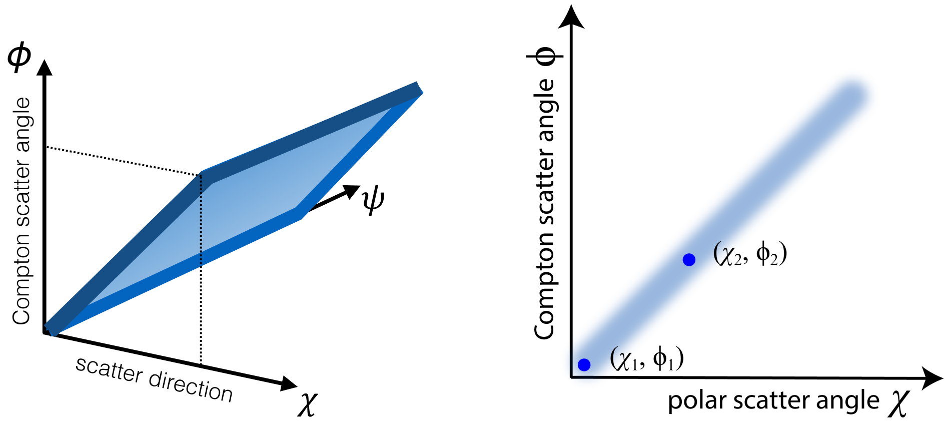
When a far-field source is at the instrument's zenith (0, 0), the apex of the CDS cone is at the origin of the data space and the cone shape is transformed into a 2D plane at 45°, as illustrated in Figure 7.2 (c). The azimuthal scatter direction ψ in detector coordinates is now equal to the Klein-Nishina azimuthal scatter angle, η from Equation 3.1, and therefore ψ encodes the polarization. For a polarized source, the sinusoidal modulation would be visible along this axis, as discussed in Section 4.6.3, which makes the CDS a powerful tool for polarization analysis; however, the Galactic positron annihilation emission is not polarized and this will not be further explored. For an unpolarized source, by ignoring the azimuthal scatter angle ψ, the CDS can be projected into 2D defined by χ and ϕ, as shown in the drawing Figure 7.2 (d). Any on-axis source in this 2D space will be mapped to a line at $\chi = \phi$. Since this 2D space is geometrically simpler than the 3D cone and there is no loss of information for unpolarized sources, in this analysis we apply a rotation matrix so we can always work in the χ-ϕ 2D CDS; this will be detailed in Section 7.1.2.

We will walk through a few simple illustrations of how the CDS is populated for an on-axis source at (0, 0), with a comparison to the familiar image-space event circles. Figure 7.2 (a) shows a γ-ray interaction that results in a small Compton scatter angle ϕ_1. The direction of the scattered photon $\vec{d}_g$ defines the axis of the image-space event circle, and the polar angle of the scattered photon χ_1 will be equal to ϕ_1 since they are both measured relative to the z-axis. This event is mapped to (χ_1, ϕ_1) near the origin of the 2D CDS shown in Figure 7.2 (d). For a slightly larger Compton scatter angle ϕ_2, shown in Figure 7.2 (b), again, the polar scatter angle χ_2 and the Compton scatter angle are equivalent. This event is also added at (χ_2, ϕ_2) to the CDS in Figure 7.2 (d). Each event from the zenith source will ideally land on the $\chi = \phi$ line and the CDS source distribution will be filled in, drawn as the blurred blue line in Figure 7.2 (d).

A simulation of a point source at COSI's zenith is shown in Figure 7.3 (a). As in Figure 7.2 (d), we have drawn the CDS in 2D, defined by χ and ϕ, by integrating along ψ. Due to the finite energy and position resolution in our detectors and Doppler broadening, there is a spread to this distribution. The deviation from the ideal $\chi = \phi$ line is equivalent to the ARM distribution, i.e., the effective width of the point spread function. Now that each event is represented by a point in the CDS space instead of a circle in image-space, the ARM definition in Section 3.2.5 is no longer the correct parameter of each event to

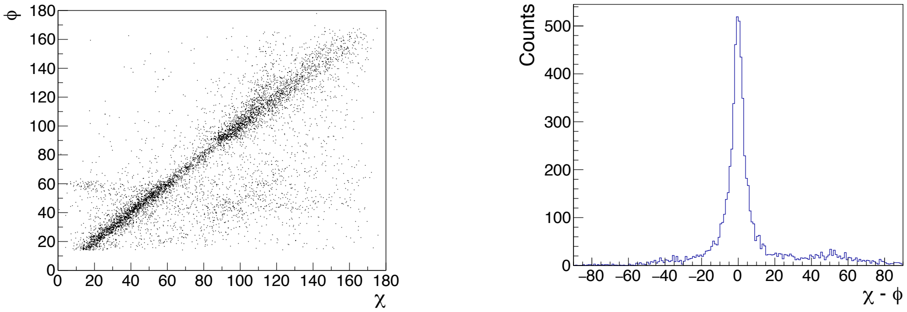


(a) Event with (χ_1, ϕ_1) in image-space. (b) Event with (χ_2, ϕ_2) in image-space.



(c) CDS for an on-axis source. (d) On-axis source projected in the 2D CDS.

Figure 7.2: (a) An example Compton event from an on-axis source shown in detector coordinates. The Compton angle ϕ_1 is equal to the polar scatter angle χ_1 (see text) and the event is represented as a point at (χ_1, ϕ_1) in the 2D CDS shown in (d). (b) Another event from the same on-axis source position with a larger Compton scatter angle ϕ_2 and polar angle χ_2. (c) When a source is on-axis, the CDS cone is transformed into a plane along $\chi = \phi$. (d) For an unpolarized, on-axis source, the CDS is projected down to 2D. The blue shaded line at $\chi = \phi$ will be the distribution of events as the CDS is filled, where the width represents the measurement uncertainty.



(a) On-axis point source simulation.

(b) CDS-ARM of point source simulation.

Figure 7.3: (a) Simulation of a 511 keV point source at (0, 0) mapped in the CDS. Only total γ-ray energies between 506–516 keV are included so as to select on the fully absorbed and properly reconstructed events to best represent the angular resolution. (b) The deviation of events from the true source location line at $\chi = \phi$, given by the distribution of $(\chi - \phi)$ shown here for the point-source simulation, defines the angular resolution of the telescope. This is referred to as the CDS-ARM and is equivalent to the ARM distribution defined previously in image-space.

define the angular resolution. In the CDS, a histogram of the deviation from the $\chi = \phi$ line is equivalent to the ARM distribution. We define the distance of each event from the $\chi = \phi$ line as $(\chi - \phi)$, as opposed to the closest distance to the line, given by $(\chi - \phi)/\sqrt{2}$ for reasons that will become apparent. We will refer to this angular distance as the CDS-ARM. The CDS-ARM histogram of the on-axis point source simulation in Figure 7.3 (a) is shown in Figure 7.3 (b). As the equivalent to the standard ARM distribution, the FWHM of the $(\chi - \phi)$ distribution is the achievable angular resolution.

7.1.2 ROTATION OF THE CDS FOR OFF-AXIS SOURCES

To simplify the analyses by working in 2D instead of 3D, we rotate any off-axis source into the origin of the CDS by using Rodrigues' rotation formula (Rodriguez, 1840). Consider a source with position defined by vector $\vec{v}$ and polar angles (Θ, Φ). One can rotate $\vec{v}$ to the origin $(0, 0)$ in detector coordinates with a rotation matrix $\mathbf{R}_{\mathrm{To0}}$ defined by:

$$\mathbf{R}_{\mathrm{To0}} = \mathbf{I} + (\sin\alpha)\mathbf{K} + (1 - \cos\alpha)\mathbf{K}^2, \tag{7.1}$$

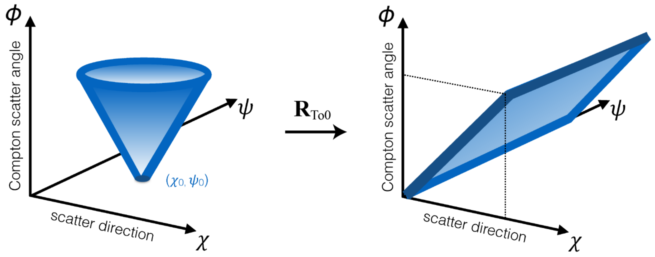


Figure 7.4: Rotation of an off-axis source position at (χ_0, ϕ_0) to the origin of the CDS, transforming the cone into a plane along $\chi = \phi$. With the rotation $\mathbf{R}_{\mathrm{To0}}$, any source location can be reduced to the 2D χ-ϕ CDS to simplify the analysis.

where

$$\mathbf{K} = \begin{vmatrix} 0 & \sin\Theta\sin\Phi & \cos\Theta \\ -\sin\Theta\sin\Phi & 0 & 0 \\ -\cos\Theta & 0 & 0 \end{vmatrix}, \tag{7.2}$$

and $\alpha = \arccos(\sin\Theta\cos\Phi)$ is the angle between $\vec{v}$ and $\hat{z}$. This rotation matrix is multiplied by the scatter direction

$$\vec{d}_{g0} = \mathbf{R}_{\mathrm{To0}} * \vec{d}_g \tag{7.3}$$

for each event in the observation. The polar and azimuthal angle of the rotated scatter direction $\vec{d}_{g0}$ fill the CDS for any source position. An illustration of the rotation with $\mathbf{R}_{\mathrm{To0}}$ that transforms the CDS cone into a plane is shown in Figure 7.4. Once the planar distribution is achieved, then the 2D CDS can be used by integrating along the ψ axis.

7.1.3 CDS IN GALACTIC COORDINATES

For flight observations, we convert the CDS to Galactic coordinates. The relative rotation between COSI's detector coordinates and Galactic coordinates is given by the Galactic Pointing Rotation Matrix $\mathbf{R}_{\mathrm{Gal}}$, calculated from COSI's aspect information and already assigned to every event in MEGAlib. The γ-ray scatter direction $\vec{d}_g$ is rotated into Galactic coordinates by $\vec{d}'_g = \mathbf{R}_{\mathrm{Gal}} * \vec{d}_g$. Again, we prefer to work in the 2D CDS, so the same $\mathbf{R}_{\mathrm{To0}}$ rotation from Equation 7.2 can be defined in Galactic coordinates, where $(\Theta = \frac{\pi}{2} - l, \Phi = b)$ for

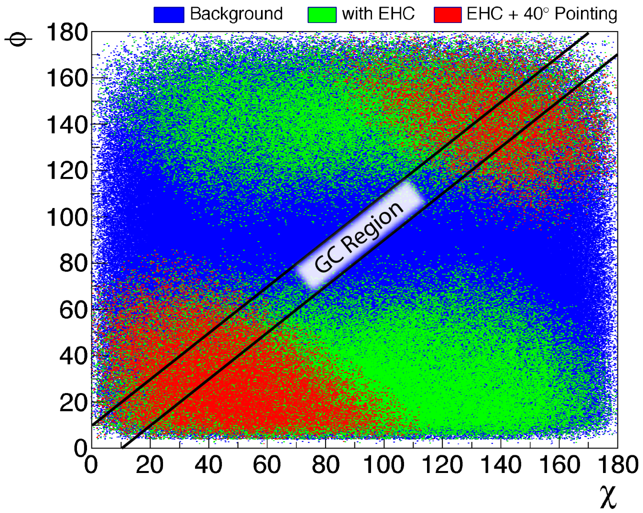


Figure 7.5: The CDS populated after a single day of flight, June 16th, 2016. The blue points show all of the Compton events from this single day. With the EHC applied to this data set, which then reduces to the green points, any event with a Compton circle that intersects the horizon is rejected. When a 40° pointing selection is applied, only events that occur when COSI's zenith is within 40° of the GC are included; these events are shown in red. Only 2.4% of the total events from this day pass the EHC and pointing selection. The black diagonal lines on this plot represent the region around the GC with a 10° origin cut, meaning any event that has $|\chi - \phi| \leq 10°$.

the Galactic latitude and longitude coordinates (l, b) of the Galactic source position. In Galactic coordinates, the rotated scatter direction is

$$\vec{d}'_{g_0} = \mathbf{R}_{\text{To0}} * (\mathbf{R}_{\text{Gal}} * \vec{d}_g). \tag{7.4}$$

For a source at the GC, which is the origin of the Galactic CDS at $(l = 0, b = 0)$, $\alpha = 0$. and from Equation 7.1, $\mathbf{R}_{\text{To0}}$ is reduced to the identity matrix.

7.1.4 BACKGROUND IN THE CDS

Since COSI observations are dominated by background radiation, it is important to understand what the background distribution $n_B(\chi, \psi, \phi)$ in the CDS looks like. Figure 7.5 shows the accumulated background events from one day of flight (June 16th, 2016), after the rotations in Equation 7.4 have been applied with the GC chosen as the source location. The photons from the GC region will be within a spread of the $\chi = \phi$ line; however, they are not visible here due to the dominance of the background radiation. The equivalent of the origin cut is now a selection on $(\chi - \phi)$, and the boundaries for events that are consistent within a 10° region around the GC are indicated with the diagonal black lines.

There are three different data sets overlaid in Figure 7.5. All of the Compton events from the full day, with no restriction on the event selections (refer to Section 3.5.5.1 for an overview of `mimrec` event selections), are shown in blue. This is without the EHC and background emissions come from every direction through atmospheric γ-rays or activation within the instrument. The total number of Compton events in this one day of flight is 1.7×10^6 cts. When the EHC is applied any event with a Compton circle that intersects the horizon is rejected, the events that remain are shown with the green points. With the EHC, the background radiation is significantly reduced and, as a by-product, the middle Compton scatter angles from ~60–120° are suppressed due to the limited effective area when the elevation is $\lesssim$60°. The total number of events is now reduced to 10% of the original. When a 40° pointing selection is applied, only events that occur when the GC is within 40° of COSI's zenith are included; these events are depicted in red. The number of events which have passed these two selections is 4.2×10^4, which corresponds to only 2.4% of the original events. As can be seen in Figure 7.5, these two event selections significantly reduce the background emission; the EHC is used in all COSI flight analyses since it is the best reduction of background emission, and the pointing selection limits observations to only when the source is within the FOV. Even with this reduction, the background still dominates, and the γ-rays from the GC, which will be within the boundary region shown in Figure 7.5, will not be isolated except when background subtraction is employed.

7.2 BACKGROUND SUBTRACTION IN THE CDS

The number of counts in each bin $n(\chi,\psi,\phi)$ of the CDS for a Galactic source with intensity distribution $I(l,b)$ is

$$\begin{aligned} n(\chi,\psi,\phi) &= n_B(\chi,\psi,\phi) + \int_l \int_b A(l,b) I(l,b) f(\chi,\psi,\phi|l,b) \cos b dl db \\ &= n_B(\chi,\psi,\phi) + n_S(\chi,\psi,\phi), \end{aligned} \tag{7.5}$$

where the integral is taken over the Galactic longitude and latitude, l and b, respectively. The average integrated effective area of the telescope is $A(l,b)$, $f(\chi,\psi,\phi|l,b)$ is the point spread function, and $n_S(\chi,\psi,\phi)$ is the source contribution in each bin obtained from the integral of the source intensity distribution. The goal of background subtraction is to determine $n_S(\chi,\psi,\phi)$ by first finding $n_B(\chi,\psi,\phi)$.

For the positron annihilation emission from the GC, limiting our analysis to the 2D CDS, we expect the source photons to be within a spread of the $\chi = \phi$ line, depending on the spatial extent of the annihilation emission. We expect approximately a 16° origin cut will be an appropriate first estimate for the Galactic 511 keV distribution from the analysis presented in Section 6.2. We will refer to the region defined by $|\chi - \phi| < \Delta$ as the source region **SR**, where Δ is the origin cut. We can find the spectrum of all events that are consistent with this region $N^{\mathrm{SR}}(E)$ in the CDS:

$$N^{\mathrm{SR}}(E) = S^{\mathrm{SR}}(E) + B^{\mathrm{SR}}(E) = \sum_{\phi}\sum_{\chi}\Bigg|_{|\chi-\phi|<\Delta} n_S(\chi,\phi,E) + n_B(\chi,\phi,E). \tag{7.6}$$

We are now explicitly including the total γ-ray energy E as a dimension in the CDS, and the superscript SR defines the region of the CDS from which the spectrum is taken. $N(E)$ is the measured spectrum with both the source and background contributions, $S(E)$ and $B(E)$, respectively. The source spectrum $S(E)$ can be isolated through an inversion of Equation 7.6:

$$\begin{aligned} S^{\mathrm{SR}}(E) &= N^{\mathrm{SR}}(E) - B^{\mathrm{SR}}(E) \\ &= \sum_{\phi}\sum_{\chi}\Bigg|_{|\chi-\phi|<\Delta} \left[n(\chi,\phi,E) - n_B(\chi,\phi,E)\right]. \end{aligned} \tag{7.7}$$

To find $S^{\mathrm{SR}}(E)$, we need to determine the spectrum of the underlying background emission $B^{\mathrm{SR}}(E)$ within the source region.

To reduce potential systematics, a background region **BR** from which we can extract a background spectrum is chosen to be two concentric cones: one that closely surrounds and one that lies within the source cone in the CDS. Or, in our reduced 2D CDS, this would be represented by two line regions that are adjacent to the source region, as shown in Figure 7.6. We will refer to these two separate background regions as $\mathbf{BR_{out}}$ and $\mathbf{BR_{in}}$, where the subscript refers to the position of the cone in the 3D CDS. It is from these background regions that the background spectrum $B^{\mathrm{SR}}(E) \approx N^{\mathrm{BR}}(E)$ can be determined.

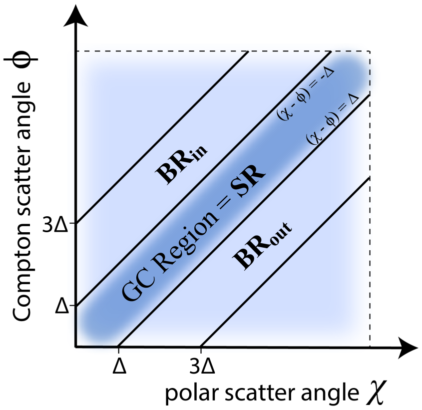


Figure 7.6: Schematic of the 2D χ-ϕ CDS showing the source and background regions, which are defined by an origin cut of Δ. In the 3D CDS, the background regions are chosen as concentric cones that surround and sit within the source cone. In 2D, these cones are transformed into lines that lie adjacent to the source line at $\chi = \phi$. Only $\mathbf{BR_{out}}$ is used in the positron annihilation emission analysis, as discussed in the text.

If there is no difference in exposure between the **SR** and **BR** regions, then spectral background subtraction can be done without further complications:

$$S^{\mathrm{SR}}(E) = N^{\mathrm{SR}}(E) - \frac{1}{2}N^{\mathrm{BR}}(E), \tag{7.8}$$

where the $\frac{1}{2}$ is included since the sum of the two background regions should be approximately twice the area of **SR**. However, as can be seen from Figure 7.5, the source region and background region will not be uniformly populated, especially if a large origin cut is used with the EHC, so we will need to scale the spectra from the background regions $N^{\mathrm{BR}}(E)$ to get an accurate representation of the total spectrum of background photons within **SR**. As we know the GC positron annihilation source will have emission at 511 keV and below, we can scale the two background region spectra by normalizing the total counts in a higher energy interval (e.g., 520–720 keV) where there will be no positron annihilation contribution in the spectrum from the GC region. The details of this, considering the energy dependencies of the CDS, will be discussed in the next section.

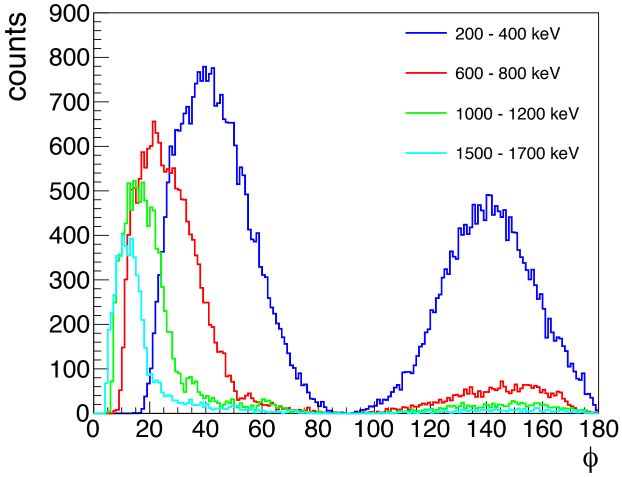
The population of the background regions depends on ϕ. As can been seen in Figure 7.6, the $\mathbf{BR_{in}}$ region does not begin to be populated until the Compton scatter angle $\phi = \Delta$, where Δ is the origin cut, i.e., the boundary line between **SR** and the background regions. For an origin cut of $\Delta = 16°$, this means that no events will fill the $\mathbf{BR_{in}}$ region with $\phi < 16°$. For the analysis routine presented in Section 7.2.2, only the $\mathbf{BR_{out}}$ region is used to determine the background spectrum.

7.2.1 ENERGY DEPENDENCIES

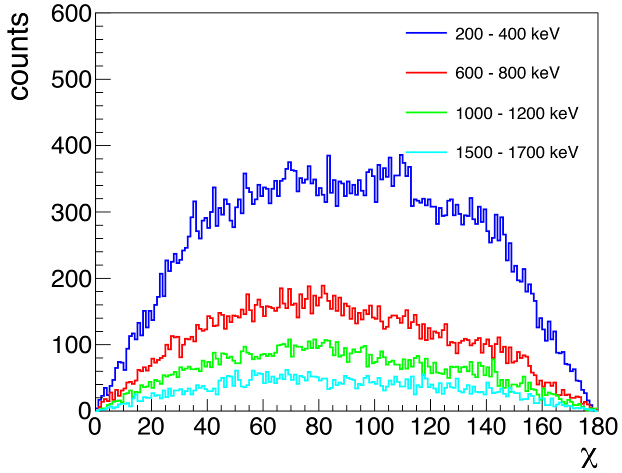
The Compton scatter angle ϕ is energy dependent, as one can see in Equation 3.3. Higher energy γ-rays will, on average, result in smaller ϕ interactions, and lower energy γ-rays will have larger scatter angles, on average. The ϕ dependence on energy is illustrated in Figure 7.7 (a). For four different energy ranges, the distribution of Compton scatter angles is plotted using only the green events from the data shown in Figure 7.5, after the EHC has been applied. Figure 7.7 (a) is an integration of Figure 7.5 along χ for the different energy ranges. Again, events with ϕ between 60–120° are suppressed with the EHC. These separate energy ranges have not been scaled, so the relative amplitude is representative of the energy-dependent effective area. This ϕ-energy dependence will need to be taken into account when attempting the spectral background subtraction.

Smaller energy γ-rays are more likely to have a backscatter event, with $\phi > 90°$. The 200–400 keV range histogram in Figure 7.7 (a) shows a second peak at large ϕ angles, but there is very little contribution of large ϕ at the higher energies. In general, these backscatter events are more difficult to reconstruct, and some ambiguous two-site events will be improperly labeled as backscatter interactions. Due to the difficulties reconstructing these events, they will on average be of lesser quality, so we will exclude $\phi > 90°$ from our analysis.

The other CDS parameter, χ, is independent of energy. Figure 7.7 (b) shows the distribution of χ for the same four energy ranges as displayed in Figure 7.7 (a). Analogous to the ϕ histograms, these distributions are obtained by integrating Figure 7.5 along the ϕ dimension. The overall shape, with fewer counts in the highest and lowest χ bins, comes from the integration along the azimuthal scatter angle ψ direction and our choice of binning. The difference in amplitude of the four histograms can be attributed to the energy-dependence of the effective area, but their overall shape is relatively the same, and this will be confirmed with a statistical test in the next section. For background subtraction, we can therefore use an energy range that contains no source contribution (e.g., 520–720 keV for the positron annihilation emis-



(a) Energy dependence of Compton scatter angle.



(b) Energy dependence of polar scatter angle.

Figure 7.7: (a) Compton scatter angle distribution for four different energy ranges using 1 day of flight data and the EHC selection. The histograms have not been scaled, so their relative amplitudes reflect the effective area for the different energies. (b) The polar scatter angle distribution is energy independent, as can be seen with the same 1 day of flight data. The difference in the relative amplitude of each scatter angle distribution is a result of the energy-dependent effective area, but the overall shape is the same.

sion), as a proxy for the χ shape of the underlying background distribution in the CDS.

7.2.2 CDS BACKGROUND SUBTRACTION ROUTINE FOR GC POSITRON ANNIHILATION

The energy dependencies presented in Section 7.2.1 led to the background subtraction routine developed by COMPTEL and modified here for COSI and the 511 keV analysis. To mitigate the energy dependence of ϕ, we slice up the data space for each ϕ bin and find the ϕ-dependent spectrum from the source and background regions, $N^{\mathrm{SR}}(\phi, E)$ and $N^{\mathrm{BR}}(\phi, E)$, respectively. We can then scale $N^{\mathrm{BR}}(\phi, E)$ such that the total number of counts within a high-energy range is equal to the number of counts in $N^{\mathrm{SR}}(\phi, E)$ for that same energy range. This is equivalent to finding the relative number of counts consistent with **SR** and **BR** in the χ distribution for each ϕ slice. The COMPTEL collaboration describe this routine another way in Knödlseder et al. (1996): "normalize the data space distribution of events taken from adjacent energy intervals so that their ϕ-profile equals that of the line interval."

The spectral background subtraction routine we developed for the positron annihilation emission analysis with the CDS is a four-step process:

1. Fill the CDS with all events using no energy selection. Split up the data into 1° slices in ϕ and find the total spectrum for **SR** and **BR** for each Compton scatter angle. For an origin cut of Δ, the ϕ-dependent spectrum for the region consistent with **SR** for Compton scatter angle ϕ is

$$N^{\mathrm{SR}}(\phi, E) = \sum_{\chi=\phi-\Delta}^{\phi+\Delta} n(\chi, \phi, E). \tag{7.9}$$

   The ϕ-dependent spectrum for the outer background region $\mathbf{BR_{out}}$ is

$$N^{\mathrm{BR}}(\phi, E) = \sum_{\chi=\phi+\Delta}^{\phi+3\Delta} n(\chi, \phi, E). \tag{7.10}$$

2. Find the scaling factor for the background spectra $N^{\mathrm{BR}}(\phi, E)$ so that the number of counts in a higher-energy range, 520–720 keV for these

studies, equals that in $N^{\mathrm{SR}}(\phi, E)$ within the same range. The scaling factor for each Compton scatter angle is:

$$F_{\phi} = \frac{N^{\mathrm{SR}}(\phi, E|520 < E < 720)}{N^{\mathrm{BR}}(\phi, E|520 < E < 720)}. \tag{7.11}$$

3. For each ϕ slice, scale the background region spectrum $N^{\mathrm{BR}}(\phi, E)$ by F_{ϕ},

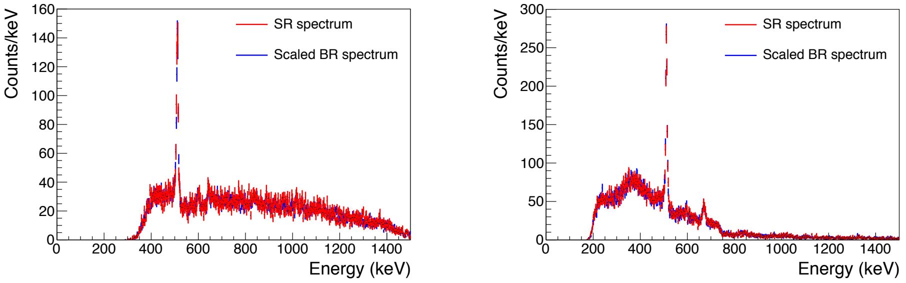
$$B^{\mathrm{SR}}(\phi, E) = F_{\phi} N^{\mathrm{BR}}(\phi, E). \tag{7.12}$$

4. Subtract the scaled background spectrum from $N^{\mathrm{SR}}(\phi, E)$ to find the total source spectrum:

$$S^{\mathrm{SR}}(E) = \sum_{\phi} \left[N^{\mathrm{SR}}(\phi, E) - B^{\mathrm{SR}}(\phi, E) \right]. \tag{7.13}$$

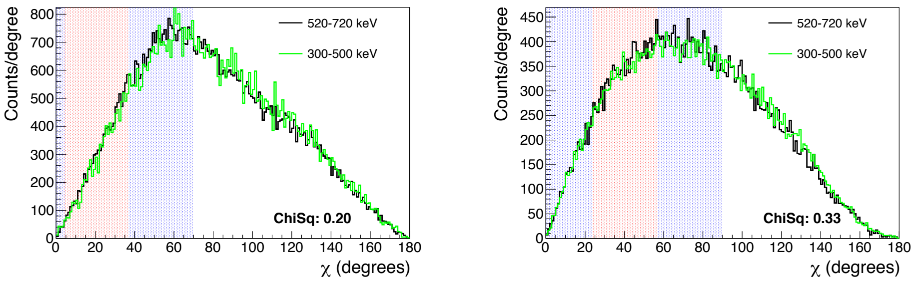
The CDS background subtraction routine is illustrated in Figure 7.8. Figure 7.8 (a) and Figure 7.8 (b) show the flight **SR** spectrum in red for two different Compton scatter ranges, ϕ = 20–21° and ϕ = 40–41°. The difference seen in these two spectral shapes results from the ϕ energy-dependence and clearly demonstrates the need to perform a spectral subtraction for each ϕ angle. The background spectrum from **BR** is plotted in blue after it has been scaled so that the number of counts within 520–720 keV is equal to that in the **SR** spectrum. For each Compton scatter range, there is a very good match between the source region spectrum and the scaled background region spectrum.

By using the energy range from 520–720 keV to scale the background ϕ-dependent spectra, we are relying on the energy-independence of χ. Figure 7.8 (c) and (d) show the χ distribution from simulations for two different energy ranges above and below the 511 keV line emission: 300–500 keV is shown in green, and 520–720 keV is shown in black. Simulation data is used since the positron annihilation spectrum is known to ave contribution below 511 keV. The χ histograms have been normalized, and though the shape visually appears consistent, a chi-square statistical test was used to determine if the distributions from the two energy ranges are statistically different. Based on the P-values from these tests, ranging from 0.01-0.99 with an average of P = 0.53 for ϕ between 15° and 60°, it can be concluded that there is no statistical difference between the χ distributions for the two energy ranges in the simulation data. Therefore, our method of scaling the ϕ-dependent spectra is



(a) **SR** and **BR** spectra for ϕ = 20–21°.

(b) **SR** and **BR** spectra for ϕ = 40–41°.



(c) χ distribution for ϕ = 20–21°.

(d) χ distribution for ϕ = 40–41°

Figure 7.8: (a) Spectrum of the measured flight data in the source region **SR** with Compton scatter angle ϕ = 20–21°. The spectrum from the background region **BR**, shown in blue, has been scaled so that the number of counts between 520–720 keV is equal to that in the **SR** spectrum. (b) Same as (a) except for Compton scatter angle ϕ = 40–41°. (c) Using simulation data, the χ distribution is compared for two energy ranges: 520–720 keV and 300–500 keV. We performed a chi-square test to confirm there are no statistical differences in the two χ distributions and the resulting P-value is listed on the figure. The red and blue regions signify the location of the **SR** and **BR**, respectively. (d) Same as (c) except for Compton scatter angle ϕ = 40–41°.

valid. The results from the statistical tests for the two Compton scatter angles shown in Figure 7.8 (c) and (d) are listed on the plot as **ChiSq**.

Overlaid on the χ histograms in Figure 7.8 (c) and (d) are the locations of the source region and the background regions, shaded in red and blue, respectively. For each ϕ value, we know the polar scatter angles that are consistent with the source region satisfy $|\chi - \phi| < \Delta$, where we define our origin cut to be $\pm\Delta$. The χ values that are consistent with the GC with a $\Delta = 16°$ origin cut are shaded in red. Alternatively, the χ values that are consistent with the two background regions are shaded in blue. The total number of counts within the red **SR** of the 520–720 keV χ distribution is the same as what is obtained by integrating the **SR** spectrum within 520–720 keV. The is also true for the **BR**. In essence, we are taking advantage of the χ energy independence for our background subtraction routine.

The $\mathbf{BR_{in}}$ region is to the left of the source region in the χ histograms, and for $\phi = 20–21°$ it is hardly populated. Even for $\phi = 40–41°$, the number of counts in the inner background region are significantly fewer than in the source region and outer background region. This lack of statistics is the main reason we only use $\mathbf{BR_{out}}$ as the background region for this analysis. If the origin cut was smaller, this would be less of an issue.

The background subtraction procedure will be further illustrated in Section 7.3, where we first validate this method by using a background simulation, then we recover the correct flux of a simulated point source.

7.3 SPECTRAL BACKGROUND SUBTRACTION VALIDATION

The CDS background-subtraction routine described in Section 7.2.2 has been developed for the analysis presented here, so it is important to validate the method with simulations before applying it to flight data. In Section 7.3.1, we will describe a detailed background simulation that closely matches the emission measured in flight. With this simulation, we can perform the CDS background subtraction described above, and with no source present we would expect to recover a flat spectrum. This will allow us to check for potential systematics induced in the subtraction procedure. Similarly, it is important to show that we can recover the correct flux of a simulated source. A simulation of a GC point source that takes into account the atmospheric absorption and aspect information from flight will be discussed in Section 7.3.2.

7.3.1 BACKGROUND SIMULATION

The verification of the CDS background subtraction will be shown with a simulation of background radiation, and with that aim, we will describe the detailed background simulation with a comparison to the measured flight data. Two background components need to be simulated: 1) the atmospheric γ-ray emission that comes predominantly from bremsstrahlung radiation of secondary electrons produced in cosmic-ray interactions in the atmosphere, and 2) the internal activation of the telescope when the instrument material is bombarded by cosmic-rays.

7.3.1.1 BACKGROUND MODEL

We define the atmospheric γ-ray emission with the model from Ling (1975), an empirical model derived from the measured flux from a balloon-borne germanium detector. The Ling Model describes the angular and depth dependence of γ-rays within 0.3-10 MeV. From Ling (1975), the differential γ-ray flux at an angle θ from the zenith, for a detector at an atmospheric depth of h (g/cm^2), is given by

$$\begin{aligned}\frac{dF(E',h)}{d\Omega} &= \int_r S(E',x)\rho(x)\exp\left[-\int_0^r \mu(E')\rho(r)dr\right]\frac{dr}{4\pi} \\ &+ \frac{dF_c(E')}{d\Omega}\exp\left[-\int_0^\infty \mu(E')\rho(r)dr\right]\gamma/\mathrm{cm}^2/s/\mathrm{sr}/\mathrm{MeV}. \qquad (7.14)\end{aligned}$$

The production of γ-rays in a unit of air mass is given by $S(E',x)$, the isotropic source function, and is defined in Ling (1975). Also given in Ling (1975) is the intensity of cosmic γ-rays incident at the top of the atmosphere, $F_c(E')$. The other parameters, the density of air $\rho(x)$ at depth x and the mass absorption coefficient $\mu(E')$, are determined from a physical model of the atmosphere. The Ling Model also defines a separate source function for line emissions, and we can therefore include the contribution of the atmospheric γ-rays at 511 keV from β^+ decay in the atmosphere.

To describe the interaction depth x, density ρ, and mass absorption coefficient μ within the atmosphere at balloon altitudes, which are inputs to the Ling Model, we use the NRLMSISE-00 model (Picone et al., 2002). This model allows for the computation of the density of molecular oxygen (O_2) and nitrogen (N_2), and atomic argon (Ar), which are the three most abundant components of the atmosphere, as a function of altitude, geographic location, and time of the year. For our input we used June 13th, 2016, at 19:00 hr, with a geographic location of -6° latitude and 254° longitude. These values were chosen since they corresponded to the best times of the Crab Nebula observations, which is a separate analysis being performed for COSI, but the averaged parameters need not be precise since the full background spectrum can normalized. From the NRLMSISE model, we obtain the density of the three molecules for a range of altitudes relevant to COSI, 26-34 km, and the mass absorption coefficients are obtained from the NIST XCOM library (Berger et al., 2017). To confirm our implementation of the Ling Model, we have calculated the γ-ray flux as a function of the zenith angle for an atmospheric depth of 2.2 g/cm^2 to compare to Figure 6 from Ling (1975); our results are shown in Figure 7.9 and are correct to within 2%. The strength of the 511 keV atmospheric line is included in this plot.

Once the differential atmospheric γ-ray flux has been determined, we simulate the energy- and zenith-dependent emission in detector coordinates. A `cosima` orientation file was included in the simulation so that the flight aspect information is taken into account and the simulation can be analyzed in Galactic coordinates. The atmospheric γ-ray simulation assumes an altitude of 33.5 km. Altitude drops and changes in geographical location were not taken into account; instead, the simulated flux is scaled so that the background spectrum from the 46 day simulation matches the total accumulated spectrum from flight data.

In addition to the atmospheric photon emission, activation within the instrument is simulated. When high-energy cosmic rays bombard the telescope, radioactive isotopes are created within the passive and active material. The main cosmic ray contributors to activation are protons, neutrons, positrons, and α particles. With a proper description of the flux and energy of the cosmic rays, the resulting prompt emission and long-lived isotopes can be simulated. Activation simulations in `cosima` are a three-step process: 1) simulate the incoming cosmic ray particles to save the prompt component and a list of isotopes produced, 2) from the isotope list generated in Step 1, calculate the activation per isotope per volume after a certain time of irradiation, and 3) using the output of Step 2, simulate the decay products and save the delayed emission. For these simulations, we use an irradiation time of 40 days. The detector timing resolution of 5 μs separates the prompt emission from the delayed activation component.

The differential fluxes of the cosmic ray particles for the activation simulation are obtained from the EXPACS model (Sato, 2015), where we have used a date of June 1st, 2016, a geographic position of -30° latitude and -100°

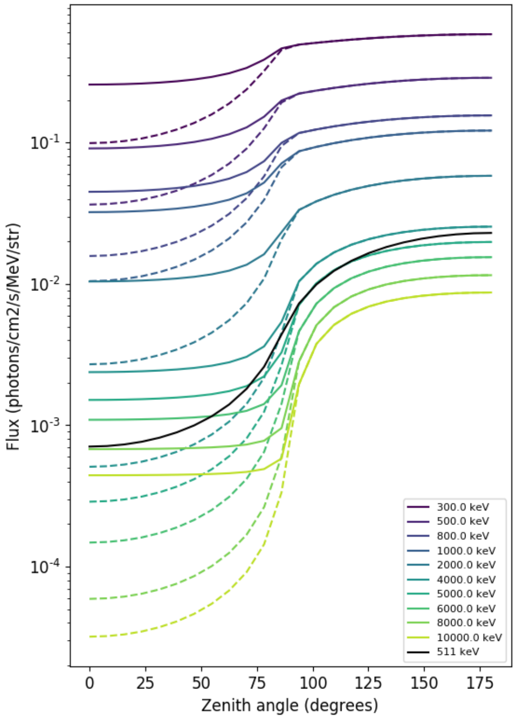


Figure 7.9: The Ling Model for γ-ray flux as a function of zenith angle at balloon altitudes. For this plot, an atmospheric depth of 2.2 g/cm^2 is used so as to compare to the values published in Figure 6 from Ling (1975). Both the atmospheric and cosmic-ray γ-rays are included, where the dashed line shows the atmospheric contribution, and the solid line includes the cosmic γ-rays. The 511 keV atmospheric background component is shown as the solid black curve. A zenith angle of 0° corresponds to downward going radiation. Figure from A. Lowell (private communication, 2017).

longitude, and an atmospheric depth of 6 g/cm^2 as input parameters. The detector response for each incident particle is split into two separate components, the prompt emission and the delayed decay emission, which are the outputs of Step 1 and 3 from the `cosmia` simulation description above. The prompt emission occurs within the timing resolution of the detector and can be from inelastic scatters of the incident particle or a decay of an activated isotope with a half-life shorter than the timing resolution. The delayed emission is any detected interactions from the byproducts of isotopes that have a long half-life, e.g., positrons or γ-rays from β^+ decay.

The vast majority of prompt activation events are vetoed by the CsI shields, since most of these high-energy particles will not be fully stopped in the GeDs and will therefore interact with the scintillators. In addition to producing radioactive isotopes, these cosmic-ray particles will often result in a large number inelastic scatters within the GeDs, which will be rejected in the `revan` event reconstruction. In the prompt neutron simulation, for example, 8.5×10^7 of the simulated particles interact in the COSI GeDs and ~98% of these are vetoed by the shields or rejected in the event reconstruction. The remaining 2% of events are then improperly tagged as Compton events. When the EHC selection is included, this contribution is further reduced and only 0.2% of the original interactions are falsely registered as valid events in `mimrec`.

The delayed decay events are rejected in much the same way as the prompt emission. Most radioactive decays result in the emission of both a particle and γ-ray, and the subsequent interactions of these byproducts in the detector will produce positions and energy deposits that will not conform to the Compton event reconstruction algorithm and thus will be rejected. In additional, a fraction of the decay byproducts will interact within the CsI shields, which will veto the event. A relevant example would be a delayed β^+ decay, which would produce a positron that would immediately decay into two 511 keV photons. If both 511 keV photons interact within the GeD, then the event is unlikely to pass the `revan` event reconstruction, since the interactions will not be consistent with a single photon event. If one of the 511 keV photons is detected in the GeDs and the other interacts in the CsI shields, which is likely since they surround the majority of the cryostat, then the event will be vetoed by the shields. However, if one 511 keV photon interacts with the GeDs and the other does not interact in COSI at all, then this event will be indistinguishable from a cosmic 511 keV γ-ray. However, the EHC can potentially reject the event if the incoming direction of the γ-ray was found to be consistent with below the horizon.

Even with all of COSI's background reducing techniques, some activation emission can still be detected. Figure 7.10 shows the resulting activation spectrum, split into the prompt and delayed components, from the four simulated cosmic ray particles. The contribution from comic ray muons and electrons was also simulated, but the emission was negligible. Activation has not been previously simulated for the COSI instrument, but since there is a strong activation line at 511 keV, we need to confirm that this emission does not affect the CDS background subtraction routine.

A comparison of the total background spectra from flight and simulations is shown in Figure 7.11. The activation and γ-ray atmospheric components, combined and plotted in red, have been separately scaled before combination

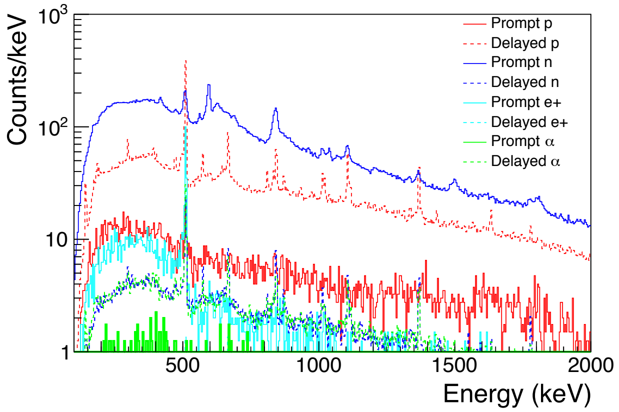


Figure 7.10: Simulated COSI activation spectra of multi-site events resulting from proton, neutron, positron, and α-particle cosmic rays. The prompt emission is any deposited energy that appears to be simultaneous with the original particle interaction, either from a subsequent scatter or a radioactive isotope with a half-life shorter than the detector timing resolution. The delayed emission is the measured byproducts from any isotope that has a half-life longer than the detector timing resolution. The simulation for each component was 100,000 s, and the majority of interactions are vetoed by the CsI shields.

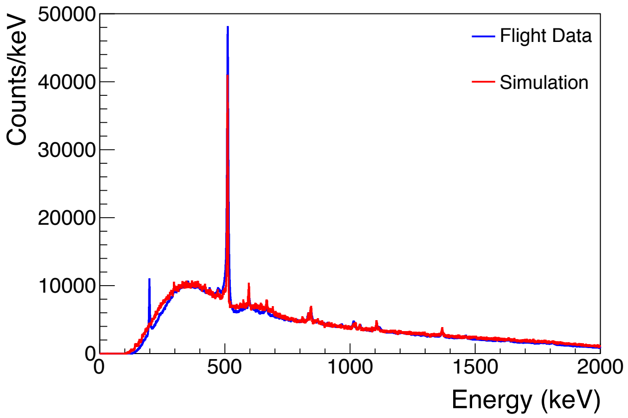


Figure 7.11: The total simulated multi-site background spectrum in comparison with the measured total integrated flight data, where both include the EHC. The simulated atmospheric γ-ray emission and activation spectra, combined and shown in red, are separately normalized such that the total spectrum matches the measured flight spectrum, shown in blue.

so that the overall shape of the spectrum and size of the activation lines match the total measured spectrum from flight, shown in blue. Although the match is not perfect, for our purposes of confirming the CDS background subtraction routine, the simulated background spectrum is adequate.

7.3.1.2 CDS BACKGROUND SUBTRACTION WITH BACKGROUND SIMULATION

Now that we have a fairly accurate representation of the background emissions during the flight, we can use our CDS background subtraction routine, described in Section 7.2.2, to confirm we can obtain a flat spectrum free of systematics around 511 keV. The analysis performed here will be the same as for the flight data, which is presented in Chapter 9. The DEE has been applied to the simulations to better mimic the true detector response.

Since the aspect information from flight was added to the background simulation, we can analyze the data in Galactic coordinates. The source region is chosen to be the GC to most closely match the flight data analysis; however, as further validation we will test other Galactic positions. For this analysis, we use a 16° origin cut, since this is our first guess as to the extent of the spatial distribution based on preliminary analysis. We are particularly interested to see if there is any detectable excess between 506–516 keV in the subtracted spectrum, which would imply that the atmospheric or activation component of the background is not properly subtracted and could potentially influence our measured flight spectrum.

Figure 7.12 shows the resulting CDS background-subtracted spectrum for the background simulation, assuming the GC as the source location. The spectrum is visually consistent with a flat spectrum and the number of counts between 506–516 keV is 38±286 cts. However, on closer inspection, a slight negative excess can be seen between ~250–350 keV. The reduced χ^2 of this distribution compared to a flat line at 0 counts/keV is 0.92 for the range 400–1000 keV, but 1.3 between 250–350 keV, for 50 and 300 degrees of freedom, respectively. The significance of this negative excess at low energies depends on the event selections, most notably the Compton scatter angle range, and seems to be a function of the exposure of the GC region, as will be illustrated when we choose different Galactic positions for our analysis.

The event selections for this analysis are listed in Table 7.1. The accepted Compton scatter angle range is 16–60°, which may seem arbitrary at first; however, this corresponds to most of the full range of angles allowable for energies close to 511 keV. Figure 7.13 shows the distribution of Compton

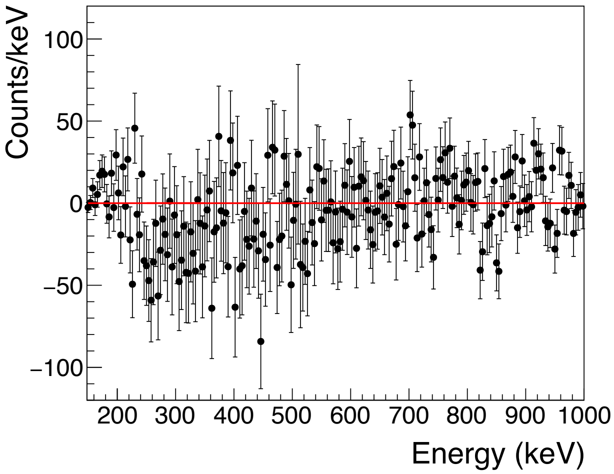


Figure 7.12: CDS background-subtracted spectrum of the background simulation, where a red line is draw along Counts/keV = 0 to represent the flat distribution we expect. The number of counts within 506–516 keV is 38±286 cts, which confirms a proper background subtraction of the atmospheric background and instrumental activation at 511 keV.

Parameter	Allowed Range
Pointing selection	None
Altitude	N/A
Origin selection	16°
Photon energy	N/A
Number of interactions	2 – 7
Compton scatter angle	16 – 60 °
Distance between first 2 interactions	> 0.3 cm
Distance between any interaction	> 0.0 cm
Earth horizon cut	Reject if 99% of Compton circle is below horizon

Table 7.1: Event selections for the CDS background-subtracted spectrum for the background simulation shown in Figure 7.12. Refer to Section 3.5.5.1 for an overview of the `mimrec` event selections.

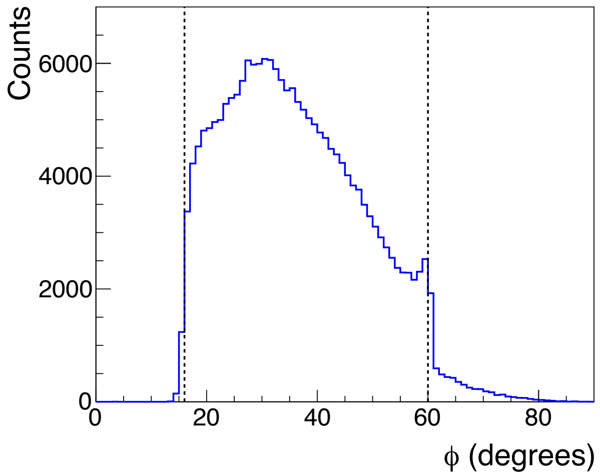
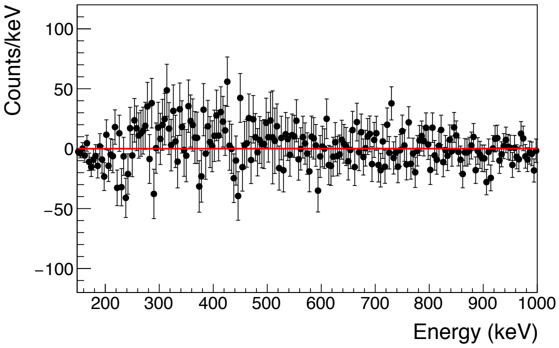


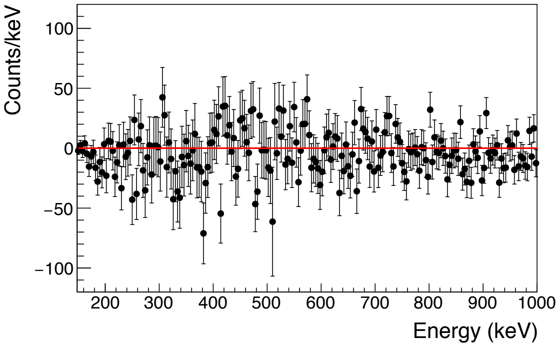
Figure 7.13: The distribution of Compton scatter angles for event with E_γ = 506–516 keV from the background simulation. The dotted vertical lines show the event selections used in this analysis, which is the majority of events.

scatter angle for the full background simulation of events with E_γ = 506–516 keV. The dotted vertical lines indicate the range used for this background analysis. For now, our goal is to show that we can achieve the expected distribution with selections that are fairly representative of the final analysis, but it is important to note that further limiting the event selections could allow for a more sensitive background subtraction.

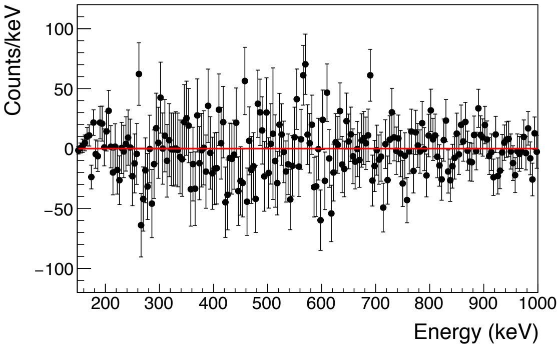
To further test the background subtraction procedure, we have chosen three random positions in Galactic coordinates and repeated the CDS background subtraction with our simulated data. We have ensured these positions had exposure during the COSI flight. For each of these locations, we want to confirm that there is no detected significance, particularly around the 511 keV line. Figure 7.14 shows the resulting spectrum for the chosen source locations of (-150, 30), (120, -6), and (180, -90), in (a)–(c), respectively. All event selections are the same as was used in the GC subtraction, as listed in Table 7.1. The number of counts between 506–516 keV for these three spectra are 49±200, -244±237, and 40±287 cts, respectively. All are consistent with zero. The negative excess seen at low energies in Figure 7.12 is not apparent here, though a slight positive excess is seen in the spectrum from (30, -150). This effect can be attributed to how the CDS is populated, since the data being used for these tests are identical.



(a) Source location: (-150, 30).



(b) Source location: (120, -60).



(c) Source location: (180, -90).

Figure 7.14: Background-subtracted spectra obtained for three different source locations in the background simulation. The Galactic coordinates were selected randomly from the flight exposure map: (-150, 30), (120, -60), and (180, -90) for (a)–(c). The number of counts between 506–516 keV for these spectra are all consistent with no source.

7.3.2 POINT SOURCE SIMULATION

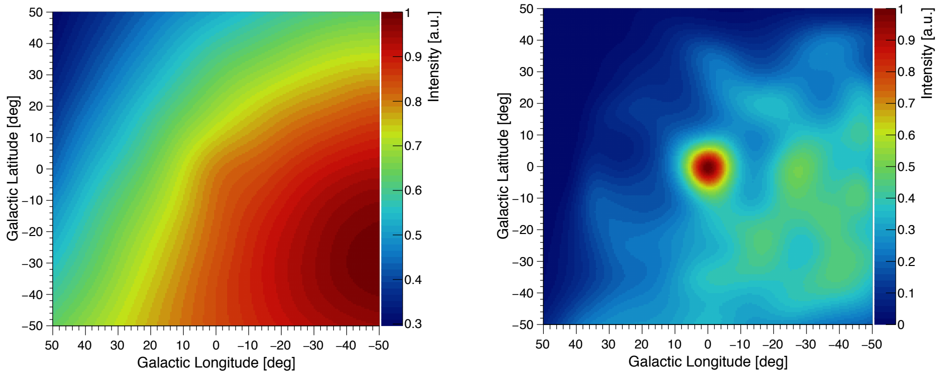
The background-minus-background spectra are fairly free of systematics, but a more definite proof of concept of the CDS background subtraction algorithm would be an accurate extraction of a simulated source spectrum. To this end, we have simulated a point source at the GC to be combined with the background simulation described in the previous section. Again, we have used all aspect information from the flight, but now will take into account the absorption of the source γ-rays in the atmosphere, which will be more severe for the altitude drops at night; our method will be explained in Section 7.3.2.1.

We have performed the point source and background simulations separately, and we can analyze them by combining them together, which would closely represent flight-like conditions, or we can analyze the point source simulation alone to determine the expected number of counts and spectral shape without the background contribution. The simulated source is a single Gaussian line at 511 keV with a width of 2 keV and a flux of 0.0023 $\gamma/\text{cm}^2/\text{s}$.

Figures 7.15 (a) and (b) show an image of the combined background and point source simulation in Galactic coordinates. The image in (a) is a simple back projection of all events that have E_γ = 506–516 keV. Since the emission is dominated by the background radiation, this back projection closely represents the exposure of the region around the GC. Considering that the pointing information from flight was used for these simulations, this should be identical to the exposure map from flight for this region. Figure 7.15 (b) shows the same image after 5 iterations of the LM-ML-EM algorithm, and the point source emission at the GC becomes clearly visible. In Figure 7.15 (c), we have plotted the full spectrum with a 6° origin cut around the GC, which includes both the atmospheric γ-ray and activation background, as well as the point source contribution at 511 keV. This is the **SR** spectrum. We can analyze the source simulation separately, and Figure 7.15 (d) shows the ARM distribution obtained in `mimrec` for events with E_γ = 506–516 keV. When fit with the ARM distribution fit function, a double Lorentzian and an asymmetric Gaussian, shown in red, the FWHM is found to be 6.5°. This is close to what we have achieved from calibration sources (Chapter 4) and affirms the choice for an origin cut of 6° for this point source simulation.

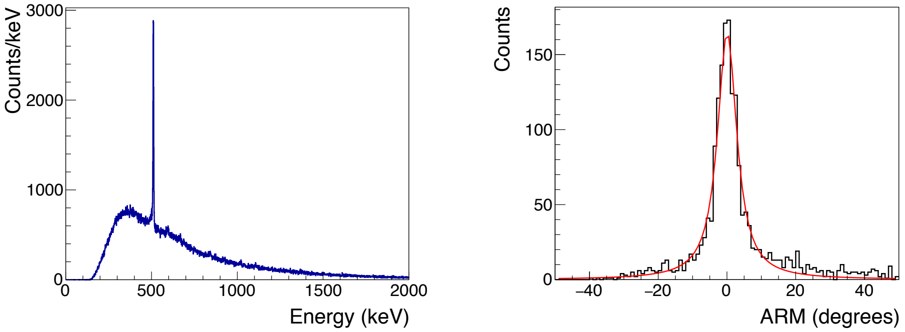
7.3.2.1 SIMULATING ATMOSPHERIC TRANSMISSION

To produce a realistic simulation of a GC point source to better compare with the flight data, we simulate the affects of γ-ray absorption in the atmosphere,



(a) Image backprojection.

(b) Image after 5 iterations of LM-ML-EM.



(c) Spectrum of background and point source.

(d) ARM distribution of GC point source.

Figure 7.15: (a) Back-projection of the GC point source and background simulation. The hot spot in the lower right of the image shows the peak of the exposure. (b) The point source emission is visible in the image after 5 iterations of the LM-ML-EM algorithm. (c) The total spectrum of the point source and background simulation with a 6° pointing cut around the GC. This is **SR**. (d) The ARM distribution of the GC point source, without the background simulation, to show the expected angular distribution.

Altitude	Exposure time [s]	
range [km]	2 dead dets	3 dead dets
27–28	5306	54401
28–29	23334	41172
29–30	35545	92187
30–31	29347	144653
31–32	34878	106156
32–33	164006	136563
> 33	1312948	901394
	Total time = 3081890 s	

Table 7.2: The total COSI flight time has been divided into different altitude ranges. For each altitude listed, the time has been further split into the number of dead detectors, i.e., before or after June 6th. Each altitude and detector combination corresponds to a different simulation. The total simulated time of 3.08×10^6 s is less than the entire flight duration of 46 days = 3.98×10^6 s, since there are a number of bad-time intervals have already been removed from the flight data.

taking into account the drops in altitude. This was done by creating 14 different `cosmia` orientation files with seven different altitude ranges and two mass models, including: the two dead detectors (Det #5 and Det #8) before June 6th, and the three dead detectors (Dets #5, #8, and #0) after June 6th. Table 7.2 lists the different simulation times for each of these orientation files, where these values are obtained from the flight data. A separate simulation was performed for each of these 14 different orientation files; the number of dead detectors determines the mass model, and the altitude range selects the different absorption file to use in the simulation.

The absorption files are created by using the NRLMSISE-00 model (Picone et al., 2002), which was also used to determine the density in the atmosphere for the Ling background model. From the densities of O_2, N_2, and Ar, we use the NIST XCOM library (Berger et al., 2017) to find the cross-section of interaction as a function of γ-ray energy. Then, by integrating along the line-of-site through the atmosphere, we calculate the transmission probability for 18 different off-axis angles between 0–90° and 25 different energies between 50 keV–10 MeV. The altitude is kept as a independent parameter in the model. The transmission probabilities for two different altitudes and two different off-axis angles calculated from this model are shown in Figure 7.16.

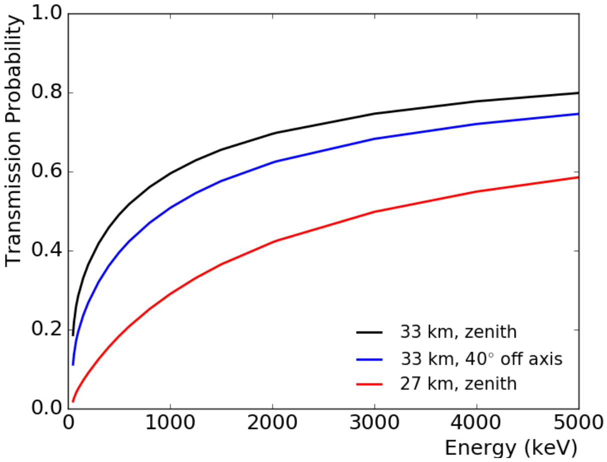


Figure 7.16: The calculated atmospheric transmission probability as a function of γ-ray energy. The probability of transmission is highest for downward-going, high-energy γ-rays. The black curve shows the transmission probability at an altitude of 33 km for downward-going γ-rays. When a Galactic source is off-axis or the instrument has a lower altitude, there is more atmosphere for the γ-rays to pass through and the transmission probability is reduced. This is shown for a 40° off-axis angle at an altitude of 33 km, and downward-going events at an altitude of 27 km.

The results illustrate the detrimental effects of the nightly altitude drops; at 500 keV the transmission probability decreases from 0.49 to 0.18 at zenith, for an altitude of 33 km compared to 27 km. The transmission probability is lower for off-axis sources because there is more atmosphere to affect transmission between the source and the instrument.

In `cosima`, the absorption is simulated with a simple Monte-Carlo using the probabilities from the transmission model. This is only done for Galactic sources, since the absorption of the background radiation within the atmosphere is already taken into account in the Ling Model.

7.3.2.2 CDS BACKGROUND SUBTRACTION WITH GC POINT SOURCE

To perform the CDS background subtraction of the GC point source we use an origin cut of 6°. Figure 7.17 shows the results of the background subtraction, where the event selections are listed in Table 7.3. The total number of counts within 506–516 keV is 1167±194 cts, where the expected number from the point source simulation is 1218 cts. For comparison, the point source spectrum obtained through `mimrec` with the same event selections is plotted in blue in Figure 7.17. We are able to extract the correct number of simulated photons from this point source. The total number of counts between 506–516 keV in the unsubtracted spectrum shown in Figure 7.15 (c) is 1.9×10^4, and the signal-to-noise ratio is therefore 6%, which is similar to what we expect to see for the Galactic positron annihilation emission. Though the error bars are large due to the strong background, this proves the CDS background-subtraction routine can recover the spectrum of a simulated source.

It is also important to confirm that we can find the correct spectral parameters in the CDS background-subtracted spectrum. When the spectrum shown in Figure 7.17 is fit with a single Gaussian, we obtain a line centroid at 511.5±0.5 keV, with a width of σ = 3.0±0.6 keV. The fit of the source-only spectrum in `mimrec` gives a mean energy of 510.54±0.05 keV, and a width of σ = 2.9±0.4 keV. Though the line centroid is slightly shifted, the Gaussian line width of the subtracted spectrum is consistent with the source-only spectrum.

We can retrieve the initial simulated flux from the point source simulation by first calculating the effective area, making sure to use the same event selections. The simulated flux was 0.0023 $\gamma/\text{cm}^2/\text{s}$, and the total exposure time was 3.08×10^6 s; therefore, with 1218 cts measured between 506–516 keV in the point source spectrum, this gives a *flight-averaged* effective area of

$$\begin{aligned} \overline{A}_{Eff} &= A_{start}\frac{N_{det}}{N_{start}} = \frac{N_{det}}{\text{Flux}\times\text{time}} \\ &= \frac{1218\ \text{cts}}{0.0023\ \gamma/\text{cm}^2/\text{s}\times3.08\times10^6\ \text{s}} = 0.175\ \text{cm}^2. \end{aligned} \tag{7.15}$$

This value includes the atmospheric absorption, an accurate representation of the dead detectors in flight, and the event selections listed in Table 7.3. In addition, this flight-averaged effective area calculation also includes the time when the source is not within the FOV, which is why the value is so much lower than the effective area reported in Section 4.6.2. The GC is within 40° of COSI's zenith for only 20% of the total flight time, however the full flight path was simulated without a pointing cut since no selection is used in the CDS analysis. We can work from the subtracted spectrum to convert the measured number of counts back into a source flux:

$$\begin{aligned} \text{Flux} &= \frac{N_{det}}{\overline{A}_{Eff}\times\text{time}} \\ &= \frac{1167\pm194\ \text{cts}}{0.175\ \text{cm}^2\times3.08\times10^6\ \text{s}} = 0.0022\pm0.0004\ \gamma/\text{cm}^2/\text{s}, \end{aligned} \tag{7.16}$$

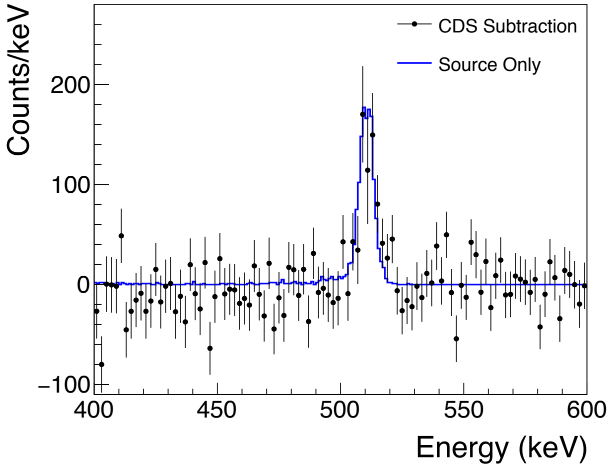


Figure 7.17: CDS background subtraction for the GC point source and background simulation. The number of counts within 506–516 keV is 1167±194, which corresponds to a ~6 σ detection. The spectrum of the GC point source with out background is shown in blue, and the number of counts in the same energy range is 1218 cts.

Parameter	Allowed Range
Pointing selection	None
Altitude	27000 m
Origin selection	6° (if applicable)
Photon energy	506 – 516 keV (if applicable)
Number of interactions	2 – 7
Compton scatter angle	16 – 60 °
Distance between first 2 interactions	> 0.3 cm
Distance between any interaction	> 0.0 cm
Earth horizon cut	Reject if 99% of Compton circle is below horizon

Table 7.3: Event selections for the CDS background-subtracted spectrum for the point source and background simulation shown in Figure 7.17. The origin selection is only applied for the spectral subtraction and the photon energy restriction is only applied for the CDS-ARM subtraction; refer to Section 3.5.5.1 for an overview of the `mimrec` event selections.

which is in agreement with the initial simulated flux of 0.0023 $\gamma/\mathrm{cm}^2/\mathrm{s}$. To perform a flux conversion for the flight data, the effective area is obtained through simulations, as demonstrated here.

7.4 BACKGROUND-SUBTRACTED CDS-ARM DISTRIBUTION

After we find an excess in the subtracted spectrum, we can determine the spatial distribution of the emission by performing a CDS-ARM background subtraction. Analogous to the spectral subtraction, we want to subtract off the underlying CDS-ARM distribution from the background emission to recover the angular distribution of the source. To do this, we need to find an appropriate estimate of the background distribution.

Before getting into the details, it is prudent to emphasis the importance of this measure; obtaining a background-subtracted CDS-ARM distribution of the Galactic positron annihilation emission would be a direct measure of the spatial extent of the source. For example, if the emission is from a point source, then we would expect to recover a CDS-ARM distribution with a FWHM ~6°, as shown in Figure 7.3 (b). If the emission has an inherent width, then the measured CDS-ARM will be a convolution of the instrument point spread function and the spatial distribution of the source.

7.4.1 ENERGY DEPENDENCES OF CDS-ARM

The CDS-ARM is defined for each event as $(\chi - \phi)$. We have previously shown there is no statistically significant energy dependence of the polar scatter angle χ (Figure 7.8); therefore, if we split up the CDS-ARM distribution into separate ϕ bins, the CDS-ARM distribution for each ϕ bin will be energy independent. We can find the CDS-ARM for each ϕ bin for $E_\gamma = 520\text{–}720$ keV and the shape should be the same as the underlying distribution at 506–516 keV.

Figure 7.18 shows the total CDS-ARM distribution from the full COSI flight, where the majority of events in this histogram is from the background radiation. For this distribution, the GC is chosen as the source location; therefore, a CDS-ARM of 0 corresponds to the GC, and photons from the region around the GC would be within a few degrees of the origin of the CDS-ARM plot. The asymmetric distribution seen in Figure 7.18, with predominantly positive CDS-ARM values, results from the varying off-axis exposure of the GC. Events with $(\chi - \phi) < 0$ occur when the source location is inside the

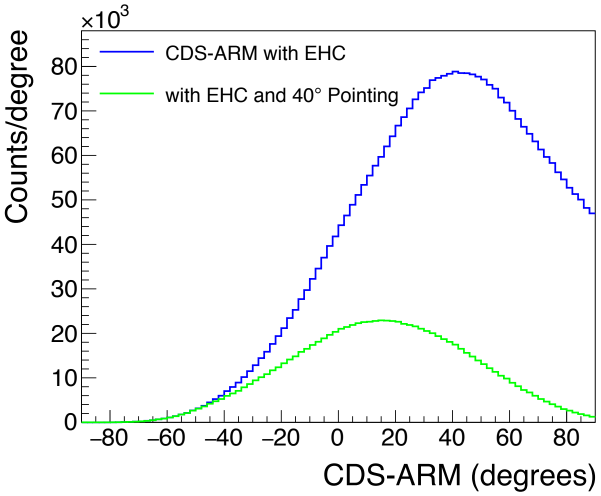
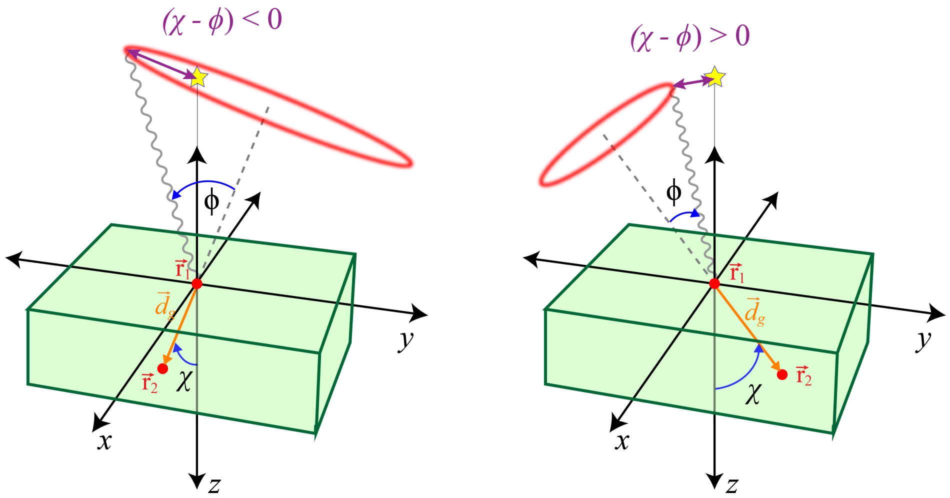


Figure 7.18: The total CDS-ARM histogram from the full COSI flight, with the GC chosen as the source location. For each event, the CDS-ARM is given by $(\chi - \phi)$, and the prevalence of the positive CDS-ARM values is due to the majority of times when the GC is outside of the FOV. This is illustrated when a 40° pointing selection on the GC is used on the data, as shown in green; a large fraction of the positive CDS-ARM values are suppressed.

projected event circle in image-space; therefore, when the EHC selection is applied, only negative ARM values are possible when the GC is within COSI's FOV. Figure 7.19 (a) illustrates an example of an event for an on-axis source which has a negative CDS-ARM. When the GC is outside of the FOV, only positive ARM values are possible, since the source location will always be outside the projected event circles. See Figure 7.19 (b) for an example of an event with a positive CDS-ARM value. The green histogram in Figure 7.18 shows the CDS-ARM distribution when a 40° pointing selection on the GC is applied; the majority of the large positive CDS-ARM events are suppressed.

Figure 7.20 shows the CDS-ARM distribution for two different ranges of Compton scatter angle: 20–25° and 50–55°. This is the (mostly) background distribution from one day of flight data with the EHC applied. The shape for these two distributions is very different. The larger Compton scatter angle range, shown with the blue-dotted line, is much flatter and extends to larger negative CDS-ARM values. This is because with a larger Compton scatter angle there is more probability that the source location will be within the large event circle as opposed to outside of it, as shown in Figure 7.19 (a). This is the opposite for smaller Compton scatter angles. The largest negative CDS-ARM



(a) Event with negative CDS-ARM.

(b) Event with positive CDS-ARM.

Figure 7.19: (a) Schematic showing an event with a negative CDS-ARM. The image-space event circle does not intersect the source position at (0, 0) since the Compton scatter is larger than the polar scatter angle. (b) Event with a positive CDS-ARM value since the Compton scatter angle is smaller than the polar scatter angle.

value for the 20–25° range of angles is -25°, because this is the maximum radius of the event circle and therefore the largest possible distance the source location can be from the circle while also being within the circle. The difference in distributions for the two Compton scatter ranges confirms the need to split up the CDS-ARM distribution into separate ϕ bins.

To determine the morphology at 511 keV, we can find the CDS-ARM distribution for each ϕ in the interval 506–516 keV. The CDS-ARM distribution from our standard high-energy range, 520–720 keV, can be used to describe the shape of the underlying background, but we will need to find an appropriate scaling factor. From the CDS spectral background subtraction, we have an accurate spectrum defining the background emission for each Compton scatter angle, $B^{SR}(\phi, E)$. We can therefore use the relative number of counts in the 506–516 keV and 520–720 keV range in this scaled background spectrum to give the correct normalization for the background CDS-ARM distribution.

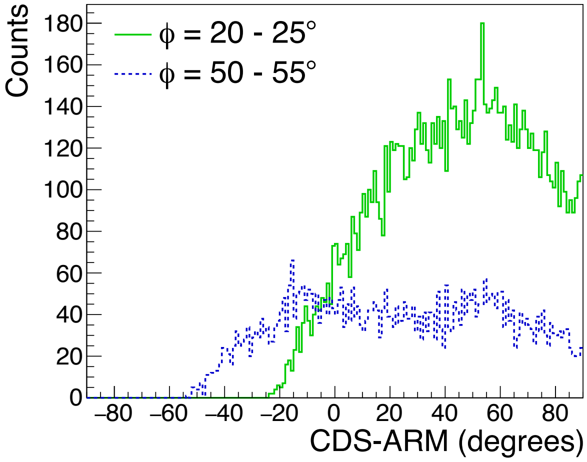


Figure 7.20: CDS-ARM distribution for two Compton scatter ranges, 20–25° and 50–55°, from one day of flight data with the GC as the source location.

7.4.2 CDS-ARM BACKGROUND SUBTRACTION ROUTINE

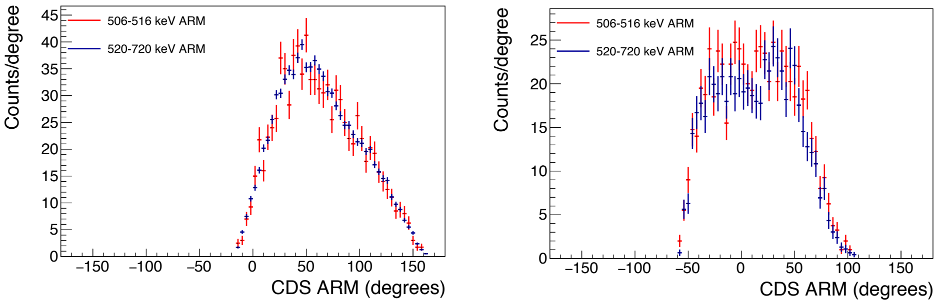
The CDS-ARM background subtraction procedure relies on the results from the spectral subtraction and is a three-step process:

1. Find a separate CDS-ARM distribution for each ϕ for two different energy ranges: the line interval $n_{\text{ARM}}(\phi, E|506 < E < 516)$ and our higher-energy range $n_{\text{ARM}}(\phi, E|520 < E < 720)$.
2. Use the scaled background spectrum $B^{\text{SR}}(\phi, E)$ from the third step in the spectral subtraction to determine the normalize factor for the higher-energy CDS-ARM:

$$A_{\phi} = \frac{\sum_{E=506}^{516} B^{\text{SR}}(\phi, E)}{\sum_{E=520}^{720} B^{\text{SR}}(\phi, E)}. \tag{7.17}$$

3. Scale and subtract off the higher-energy CDS-ARM distribution to find the source CDS-ARM distribution S_{ARM}:

$$S_{\text{ARM}} = \sum_{\phi} \left[n_{\text{ARM}}(\phi, E|506 < E < 516) - A_{\phi} n_{\text{ARM}}(\phi, E|520 < E < 720) \right]. \tag{7.18}$$



(a) CDS-ARM for ϕ = 16–17°.

(b) CDS-ARM for ϕ = 59–60°.

Figure 7.21: The CDS-ARM distribution from the background simulation for events with Compton scatter angles 16–17° and 59–60°, in (a) and (b), respectively. Events with E_γ = 506–516 keV are plotted in red. The events from the 520–720 keV interval have been scaled using $B^{\mathrm{SR}}(\phi, E)$.

Figure 7.21 shows the CDS-ARM distribution from the full background simulation (Section 7.3.1) for the line interval 506–516 keV in red. The events with Compton scatter angle 16–17° are shown in (a), and events with ϕ = 59–60° are shown in (b). The distribution from the higher-energy range, 520–720 keV, has been scaled using the background spectra $B^{\mathrm{SR}}(\phi, E)$, and is shown in blue. This scaled higher-energy CDS-ARM serves as our background distribution and closely matches the line interval distribution.

7.4.3 CDS-ARM BACKGROUND SUBTRACTION VALIDATION

Again, we will use the full background and GC point source simulations to validate the subtraction routine. Figure 7.22 shows the results from the background subtraction for the background emission, shown in (a), and the GC point source with background, in (b). For the background-only simulation, we retrieve the expected flat CDS-ARM distribution. There is no excess detected at the origin. When fit with a flat line along the axis, the reduced χ^2 is 0.91. As in the spectral subtraction, we can compare the subtracted point source distribution to the ARM histogram from `mimrec`, shown in Figure 7.15 (d) and again as the blue curve in Figure 7.22 (b). There is a close match between the subtracted CDS-ARM and the `mimrec` ARM distribution. Fitting the subtracted CDS-ARM distribution with a single Gaussian, with its centroid fixed at 0, gives a FWHM of 6±1°. This can be compared to the FWHM of the ARM distribution of the point source only, which is 6.5°. The CDS-ARM background-subtraction routine can therefore extract the correct spatial distribution of a simulated source at the GC.

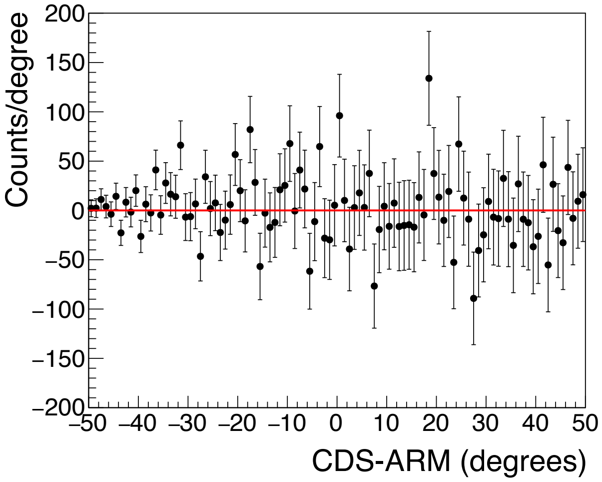
7.5 MAXIMUM-ENTROPY BINNED-MODE IMAGING

The three dimensions of the CDS (χ, ψ, ϕ), in addition to the two image-space dimensions (l, b) in Galactic coordinates, define the 5D imaging response of COSI that can be used for binned-mode imaging. The COMPTEL collaboration used the Maximum Entropy Method (MEM) for image deconvolution with the CDS for their principal imaging technique (Strong et al., 1990). The COSI collaboration is working on implementing a similar technique for a CCT for the first time (Zoglauer et al., in prep).

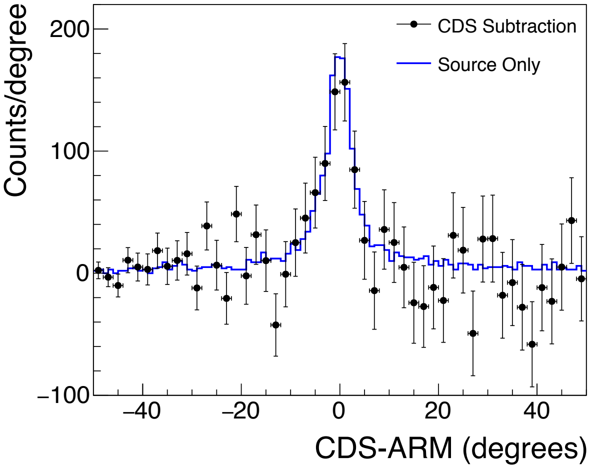
The 5D imaging response is created from a large simulation data set which is noised using the fully benchmarked COSI DEE. The advantages of this binned-mode imaging technique for the positron annihilation studies is that it will allow for an accurate measure of the flux and extended spatial distribution of the emission, and allow for background estimation and exposure correction; however, the success of the algorithm depends critically on an accurate description of the detector response and background radiation (Strong & Diehl, 1989). The preliminary results of the COSI implementation of the MEM with the CDS for the Galactic 511 keV imaging will be presented in Section 9.3.

7.6 SUMMARY

The CDS spectral and spatial background subtraction relies on the energy independence of the polar scatter direction and uses a high-energy range to determine the underlying background distribution in the CDS. The high-energy range that is used to find the background CDS distribution cannot have a source contribution; therefore, the routine developed here for the Galactic positron annihilation emission is not valid for broadband sources, though the COSI collaboration is currently developing such a CDS background-subtraction method (Sleator et al., in prep). We have shown the background-subtracted spectrum is flat when analyzing a background simulation, and we were able to recover the simulated flux of a point source, with the correct line width and spatial distribution. This is the first time this analysis has been performed for a CCT, and the routine employed here is a variant of the methods the COMPTEL collaboration published in Knödlseder et al. (1996).



(a) CDS-ARM distribution of background simulation.



(b) CDS-ARM distribution of point source simulation.

Figure 7.22: (a) Background-subtracted CDS-ARM distribution of the background simulation. A red line is drawn along 0 Counts/deg, since this is the expected distribution. (b) Background-subtracted CDS-ARM distribution of the point source simulation with background. When fit with a single Gaussian, the point source distribution gives a width of $\sigma = 6\pm1^\circ$. The expected ARM distribution, as obtained for the point source in `mimrec`, is shown in blue.

8

GALACTIC POSITRON ANNIHILATION SIMULATIONS

We have developed a background subtraction routine using the CDS to detect the Galactic positron annihilation emission with COSI. The method was introduced in Chapter 7 and validated for the simple case of point source emission. In this chapter, we will described detailed simulations of the GC positron annihilation emission that were preformed to allow for a better comparison with the flight data. It is important to confirm that the o-Ps continuum can be detected in the spectral subtraction, and that the spatial distribution of the emission can be determined from the CDS-ARM subtraction. After an overview of the preformed simulations in Section 8.1, we will do a thorough analysis of the spectral and spatial parameters of these simulations given the COSI detector response. In Section 8.2 and 8.3, we will analyze the CDS subtracted spectra and ARM distributions from these simulations.

8.1 OVERVIEW OF SIMULATIONS

The positron annihilation emission from the Galaxy is characterized by two spectral signatures: the annihilation line at 511 keV and the o-Ps continuum below 511 keV. Spatially, the emission is measured to be strongest in the center of the Galaxy, in what is referred to as the bulge emission, and there is a low surface-brightness emission associated with the Galactic disk, though the spatial extent of this component is not well constrained. The total flux of the annihilation line is ~1×10^{-3} $\gamma/\mathrm{cm}^2/\mathrm{s}$, where this value is dependent on the assumed spatial distribution. The leading descriptions for the morphology of the emission are phenomenological models derived from SPI data, namely the Skinner et al. (2014) and Siegert et al. (2016a) models.

To better understand the 2016 COSI flight measurements, we have performed detailed simulations of the Skinner and Siegert emission models. Both of these models have the same three components to describe the central GC emission: a broad Gaussian bulge with FWHM = 20.5°, an offset narrow Gaussian bulge with FWHM = 5.9°, and a point source at the location of Sgr A*. The major difference between the two models is the spatial extent of the fourth model component: the disk emission. The Skinner model has a thin disk component with latitudinal FWHM of 7°, whereas Siegert et al. (2016a) have found a width of 25° best fits the SPI data. The 511 keV flux and spatial details of the two models can be found in Tables 2.1 and 2.2, and Figure 8.1 shows the Galactic images of the models in `mimrec`. When these distributions are run through our analysis pipeline, where we include the aspect information from flight, we will only be sensitive to measurements from areas within the COSI exposure map (Figure 5.12), i.e., positive Galactic longitudes.

The simulations presented in this chapter take into account COSI's detector response and source attenuation in the atmosphere so that our results represent what COSI would have measured in the 2016 flight if either of the models correctly describe the emission. The three components within the bulge region, the narrow and broad Gaussian bulge and point source, are simulated together; however, the disk emission is separated to better understand the effects of this faint extended emission within the CDS background subtraction.

The spectral parameters in these models are assumed to be the same. For each spatial component, a 511 keV line is simulated with a Gaussian width of $\sigma = 2$ keV, and separately, the o-Ps continuum is simulated. Combining the Gaussian and o-Ps component gives a 4-parameter spectral fit function:

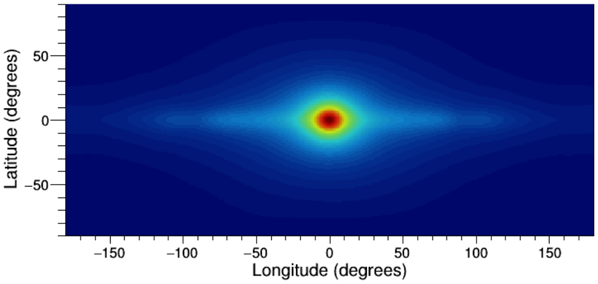
$$F(E) = A\exp\left(-\frac{(E-\mu)^2}{2\sigma^2}\right) + BF_{oPs}(E), \tag{8.1}$$

where A and B are amplitude scaling factors for each spectral component, and μ and σ are the Gaussian mean and width, respectively. The o-Ps spectrum $F_{oPs}(E)$ is given by

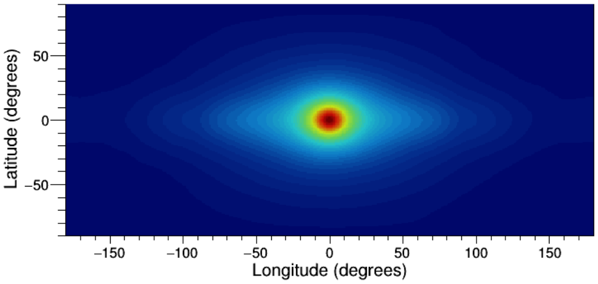
$$\begin{aligned} F_{oPs}(E) = \int_{m-E}^{m} 2\Bigg[\frac{E(m-E)}{(2m-E)^2} - \frac{2m(m-E)^2}{(2m-E)^3}\ln\frac{m-E}{m} \\ + \frac{2m-E}{E} + \frac{2m(m-E)}{E^2}\ln\frac{m-E}{m}\Bigg], \end{aligned} \tag{8.2}$$

with $m = 511$ keV. From the relative flux of the o-Ps continuum and the 511 keV line, denoted by $I_{3\gamma}/I_{2\gamma}$, one can calculate the positronium fraction:

$$f_{Ps} = \frac{8I_{3\gamma}/I_{2\gamma}}{9+6I_{3\gamma}/I_{2\gamma}}. \tag{8.3}$$



(a) Galactic distribution of Skinner model.

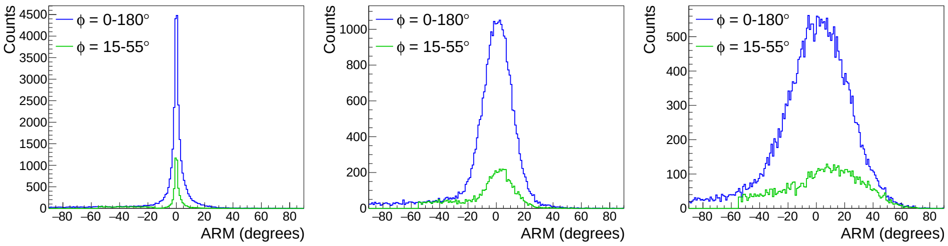


(b) Galactic distribution of Siegert model.

Figure 8.1: The spatial distribution of the two simulated positron annihilation emission models in Galactic coordinates. The only difference between the models is the size of the Galactic disk emission, which has a latitudinal FWHM of 7° for the Skinner model shown in (a) and 25° for the Siegert model in (b).

It will be important to confirm that the positronium fraction is preserved in the background subtraction. The modeled positronium fraction is $f_{Ps} = 1$ and the flux values are the same as those listed in Table 2.1. For the spectral discussions presented in the following section, we will only consider the Skinner distribution since the bulge description is the same for both models and we restrict our analysis to the 16° region around the GC.

To probe the emission morphology, we look at the CDS-ARM distribution. For a point source, the CDS-ARM gives the effective point spread function of the telescope, as seen previously. For a source that has extended emission, the ARM distribution, which is a measure of the angular separation of each event from the center of the source location, will be extended as well. The ARM distribution is effectively a measure of the average radial distribution of the source.



(a) Point source. (b) FWHM = 20.5° Gaussian. (c) FWHM = 47° Gaussian.

Figure 8.2: The ARM distributions of three simulations after ideal reconstruction with a 4π FOV detector.

The ARM distribution is fit with a phenomenological model consisting of the combination of a double Lorentzian and an asymmetric Gaussian, which gives an 8-parameter fit function:

$$C(x) = I_{L1}\frac{\Gamma_1^2}{\Gamma_1^2+(x-\mu)^2} + I_{L2}\frac{\Gamma_2^2}{\Gamma_2^2+(x-\mu)^2} + \begin{cases} A\exp\left(\frac{-(x-\mu)^2}{2\sigma_>}\right), & \text{for } x \geq \mu \\ A\exp\left(\frac{-(x-\mu)^2}{2\sigma_<}\right), & \text{for } x < \mu, \end{cases} \tag{8.4}$$

where Γ_1, Γ_2, I_{L1}, and I_{L2} are the widths and amplitudes of the two Lorentzian distributions, A is the Gaussian scaling factor, μ is the mean, and $\sigma_>$ and $\sigma_<$ are the Gaussian widths for the distribution above and below the mean, respectively. It is the FWHM of this fit function that gives the width of the ARM distribution. If the fit function does not describe the ARM distribution well, then the FWHM can be approximated graphically from the binned data.

To better understand how the ARM represents the radial extension of a source, we will look at a few quick examples. Figure 8.2 shows the ARM distributions for a point source in (a) and two extended Gaussian distributions with a FWHM of 20.5° and 47° in (b) and (c), respectively. These distributions are made from simulations with an idealized high resolution imaging detector with a 4π FOV to get a true measure of the ARM distribution without a complicated detector response. With open event selections, the FWHM of the point source simulation is 2.4°, shown with the blue curve in Figure 8.2 (a). When the Compton scatter angle range is restricted to 15–55°, which is done to match our selections in the flight analysis (Chapter 9), the ARM FWHM is reduced to 1.2° and the number of counts is much lower, as shown in green. Although this is a much smaller resolution than what we can achieve with COSI, it is useful to compare the point source distribution with that from extended emission. Figure 8.2 (b) shows the ARM distribution for a simulated Gaussian with a FWHM of 20.5°. When this distribution is fit with Equation 8.4, the FWHM is found to be 23.4°. With a restriction of the Compton scatter angle to $\phi = 15$–$55°$, the ARM FWHM is reduced to 20.3°; however, the distribution is now skewed to the right by 2.4°. Figure 8.2 (c) shows the measured ARM distribution from a 47° FWHM Gaussian simulation. The FWHM of the fit with open event selections gives 49°. When the Compton scatter angle is restricted, the ARM distribution is more skewed with a peak at 13° and the FWHM is 56°.

The restriction in the Compton scatter angle allows for a better reduction of background, which will be important for the flight data, but it skews the measured radial distribution of extended sources. The measured ARM distribution no longer matches the simulated emission. Therefore, it will be easier to make a statistical comparison of the measured CDS-ARM distribution to the simulated Skinner and Siegert ARM distributions with the same event selections, as opposed to attempting a deconvolution to extract the original emission.

As described for the point source simulation in Section 7.3.2.1, the aspect information from flight is included in the Skinner and Siegert simulations, as well as the atmospheric absorption. The background simulation described in Section 7.3.1 is used to reproduce flight-like conditions for the CDS subtraction, which will be discussed in Section 8.2 and 8.3. For the discussions that follow, the event selections listed in Table 8.1 are used unless otherwise noted.

8.1.1 SPECTRAL ANALYSIS OF GC POSITRON ANNIHILATION SIMULATIONS

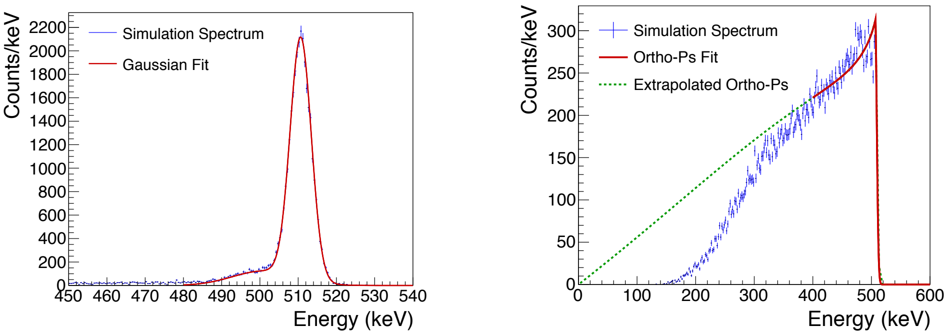
First, we will analyze the two simulated spectral components, the 511 keV line and o-Ps continuum, separately to determine the expected measured parameters. We will confirm that we obtain the same fit parameters when the two components are combined and fit simultaneously with Equation 8.1. In Section 8.2, we will include the γ-ray background emission to make sure we find consistent parameters with the CDS background subtraction. The `nuclearizer` DEE has been applied to all of these simulations, so the detector response should be fairly accurate.

Parameter	Allowed Range
Pointing selection	None
Altitude	27000 m
Origin selection	16° (if applicable)
Photon energy	506 – 516 keV (if applicable)
Number of interactions	2 – 7
Compton scatter angle	15 – 55°
Distance between first 2 interactions	> 0.5 cm
Distance between any interaction	> 0.3 cm
Earth horizon cut	Reject if 99% of Compton circle is below horizon

Table 8.1: Event selections for the CDS background subtraction of Galactic positron annihilation simulations. The origin selection is used only for the spectral subtraction, and the photon energy cut is only used for the CDS-ARM subtraction. Refer to Section 3.5.5.1 for an overview of the `mimrec` event selections, and see Appendix A for how these values were chosen for the analysis of the COSI flight data.

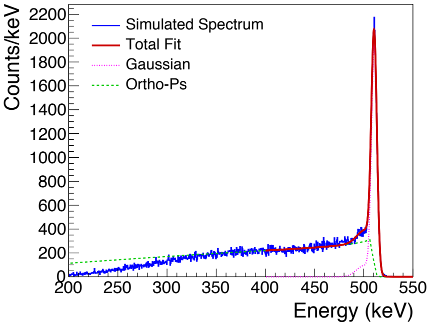
We can analyze the separate spectral components in `mimrec`, without background, and we do this for a strong simulated source to get enough statistics for a proper description of the spectral shape. We restrict our multi-site event data with a 16° origin cut on the GC so we can later compare our results with the CDS background subtraction with the same cut. For a strong simulation of the Skinner model, which is set to be 10 times the Galactic strength, the total 511 keV flux from a 16° region around the GC is 0.0133 γ/s (including the disk contribution).

The 511 keV line component, shown in Figure 8.3 (a), is fit with a Gaussian to describe the line shape. The spectral fit also includes a second Gaussian component to account for the small bump below 511 keV, which is due to charge loss in our DEE; this component is fixed to have a relative height of 6% compared to the 511 keV line, a centroid at 501 keV, and $\sigma = 8$ keV. The spectral fit of the 511 keV line shown in Figure 8.3 (a) gives a mean at 510.68±0.02 keV, and a width of $\sigma = 2.76\pm0.02$ keV. By subtracting the fit line width in quadrature with the measured COSI energy resolution of 1.85 keV at 511 keV, we find a line width of 2.1±0.3 keV, which is consistent



(a) 511 keV line component.

(b) o-Ps continuum.



(c) Combined positron annihlation spectrum with f_{Ps} = 1.

Parameter		Value
Gaussian Fit	μ	510.75±0.02 keV
	σ	2.74±0.02 keV
	A	1906±16 cts/keV
o-Ps Fit	B	164.8±0.7 cts/keV
511 keV Integral		15440±100 cts
o-Ps Integral		72860±310 cts
f_{Ps}		1.012±0.002

(d) Parameters from fit to the combined spectrum.

Figure 8.3: Spectral components of the simulated Galactic positron emission measured in the COSI detector. The 511 keV line component is shown in (a), the o-Ps continuum is shown in (b), and the combined spectrum is shown in (c). The table in (d) lists the fit parameters and the calculated flux from the separate components for the combined spectrum.

with the simulated 2 keV. The integrated flux of the line component for this simulation is 1.54×10^4 cts.

The o-Ps spectrum extending down to 0 keV is more difficult to analyze due to the energy-dependent efficiency of COSI, which drops to zero for Compton events at low energies. When fitting the spectrum with Equation 8.2 with only the amplitude as a free parameter, we find we can recover the simulated o-Ps flux, and more importantly, the relative ratio between the 511 keV line and o-Ps components, by restricting the fit range to 400–520 keV. The o-Ps continuum component is shown in Figure 8.3 (b), and the dotted green line shows the full o-Ps spectrum that is consistent with the fit from 400-520 keV, shown in red. The deviation of the measured events and the dotted-green line is due to the reduced efficiency at lower energies, where our threshold energy for Compton events is ~150 keV. It is the integral of this extrapolated spectrum that gives the o-Ps flux, which is 7.29×10^4 cts for this fit, which includes only multi-site events.

The combined 511 keV line and o-Ps spectrum is shown in Figure 8.3 (c). Fitting the combined spectrum with Equation 8.1, we are able to obtain consistent 511 keV line and o-Ps flux values to within a few percent of the values from the individual fits, and the $I_{3\gamma}/I_{2\gamma}$ ratio gives the correct simulated positronium fraction of $f_{Ps} = 1$. The parameters of the fit, as well as the calculated integrals of each component and the f_{Ps}, are included in the table in Figure 8.3 (d). This confirms that we can perform a fit of the combined spectrum and get the correct flux for each component with the correct f_{Ps}.

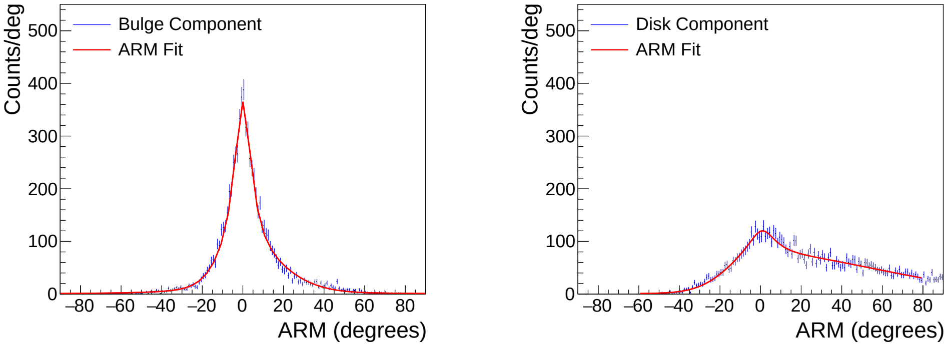
8.1.2 SPATIAL ANALYSIS OF GC POSITRON ANNIHILATION SIMULATIONS

As with the spectral analysis, we will first look at the spatial distribution of the separate components of the simulation, i.e., the bulge and disk components, in `mimrec` before analyzing the combined emission. Each ARM distribution shown here only includes events with E_γ = 506–516 keV, in addition to the event selections listed in Table 8.1, unless otherwise specified.

Figure 8.4 (a) shows the ARM distribution of the simulated bulge component for the Skinner model from `mimrec`, which includes the narrow and broad bulge and GC point source components. The measured FWHM of this distribution is 12.2° when fit with the ARM fit function in Equation 8.4. Since this is the combined distribution from the point source and narrow bulge, in addition to the broader bulge, the distribution is more strongly peaked than the single Gaussian distributions presented in Figure 8.2.

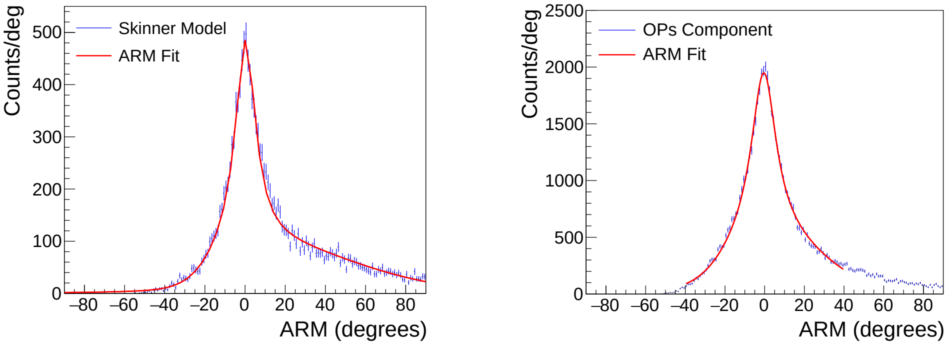
The measured ARM distribution for the disk component of the Skinner model is shown in Figure 8.4 (b). The distribution is not symmetric since the negative ARM values are suppressed with the EHC, as discussed in Section 7.4.1, and with the restriction on Compton scatter angle the distribution is skewed. The extent of the disk emission is therefore mostly apparent in the positive ARM values. We cannot separate out the latitudinal or longitudinal extent since the ARM gives only the average radial distribution; therefore, it is not easy to extract the true shape or width of the emission for large extended sources in this way.

Figure 8.4 (c) shows the ARM distribution for the total Skinner model, which combines the bulge and the disk components. The ARM fit gives a FWHM of



(a) Skinner model bulge component.

(b) Skinner model disk component.



(c) Full Skinner model.

(d) Total ARM distribution of o-Ps component

Figure 8.4: (a) The ARM distribution of the bulge component in the Skinner model, which includes the narrow and broad bulge, with a model FWHM of 5.9° and 20.5°, respectively, and the point source component at the GC. (b) The Skinner model disk component, which is modeled as a 2D Gaussian with a latitude FWHM of 7° and a longitude FWHM of 212°, shows a much broader ARM distribution that is skewed due to the EHC and restriction on Compton scatter angle. (c) The total ARM distribution of the Skinner model, which combines the distribution in (a) and (b), is dominated by the bulge component. (d) The ARM distribution of the o-Ps spectral component from the full Skinner model with E_γ = 300–500 keV.

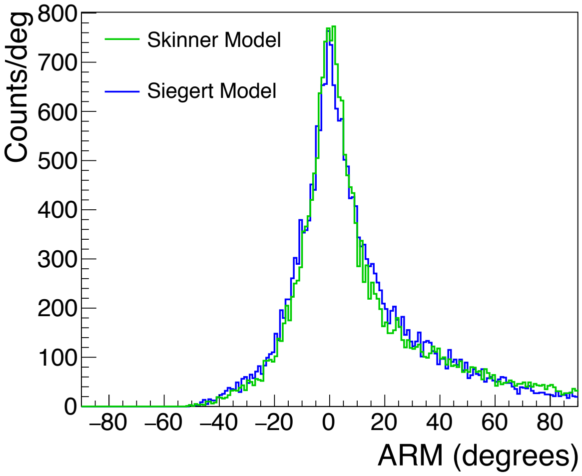


Figure 8.5: The comparison of the Skinner and Siegert model ARM distributions. With the same event selections as in Table 8.1, the thicker disk of the Siegert model, shown in green, is not significant with this analysis measure.

15.0°, which is only a few degrees larger than the bulge component alone. As we will show later in this chapter, the CDS-ARM background subtraction technique as described in Chapter 7 is not sensitive to the weak extended disk emission; therefore, we will not spend time attempting to extract the signal.

Since we have separated the 511 keV and o-Ps simulations, we can also look at the ARM distribution of only the o-Ps spectral component. Figure 8.5 (d) shows the measured ARM distribution for E_γ = 300–500 keV from the o-Ps simulation. The FWHM of the fit gives a width of 21°. The distribution is slightly wider than the 511 keV component since at these lower energies COSI's angular resolution is worse.

Finally, in Figure 8.5, we compare the ARM distribution from the Siegert and Skinner models. The distributions have been normalized so as to more easily compare the shapes. The disk emission, which is the only difference between these two models, changes the width of the distribution slightly. The ARM fit of the Siegert model gives a FWHM of 17.5°, compared to 15.0° for the Skinner model. The difference here is too small to be able to constrain the disk emission, and therefore we will focus our analysis in this chapter on only the Skinner model distribution. To better differentiate between the two disk models, COSI will need to directly image the emission; this will be discussed in Section 9.3.

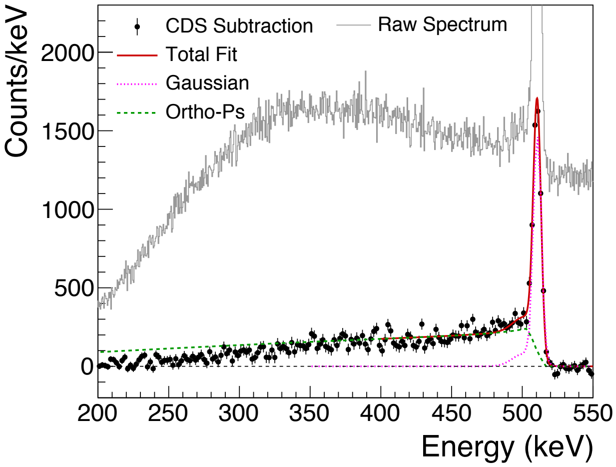
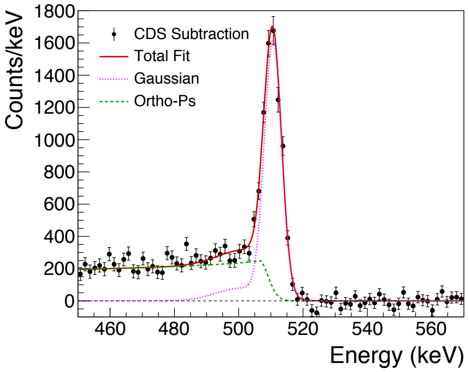


Figure 8.6: CDS background-subtracted spectrum of the Skinner model simulation. The separate components of the spectral fit defined in Equation 8.1 are shown, as is the raw unsubtracted spectrum. The event selections for this analysis are listed in Table 8.1.

8.2 CDS SPECTRAL SUBTRACTION OF GC POSITRON ANNIHILATION SIMULATIONS

The background simulation described in Section 7.3.1 is added to the Skinner and Siegert models to reproduce flight-like observations. Figure 8.6 shows the CDS background-subtracted spectrum for the strong Skinner model source, where an origin cut of 16° is used. The unsubtracted raw spectrum for the same origin cut is shown in gray for comparison. The total number of counts in the unsubtracted spectrum between 506–516 keV, which includes contribution from the 511 keV line and o-Ps continuum, is $2.9{\times}10^4$ cts. With $1.2{\times}10^4$ cts in the source-only spectrum, shown in Figure 8.3 (c), this corresponds to a signal-to-noise ratio of almost 50%, which is much larger than what is expected from the GC emission; however, this strong source will be useful to confirm our ability to recover the correct spectral parameters. In addition, we will analyze a simulation with the correct Galactic flux to better understand what we expect to measure with COSI.

Taking a closer look at the strong Skinner model simulation, Figure 8.7 (a) shows the same CDS subtracted spectrum as in Figure 8.6 but with a smaller energy range around 511 keV and finer binning for a better view of the spec-



(a) CDS subtracted spectrum.

Parameter		Value
Gaussian Fit	μ	510.67±0.08 keV
	σ	2.68±0.07 keV
	A	1580±50 cts/keV
o-Ps Fit	B	131±3 cts/keV
511 keV Integral		12190±320 cts
o-Ps Integral		58100±1300 cts
f_{Ps}		1.01±0.01

(b) Spectral parameters from fit.

Figure 8.7: (a) Spectral fit of the simulated strong Skinner model source for a 16° region around the GC. The fit parameters and calculated positronium fraction are listed in the table in (b).

tral fit function. The fit parameters for this spectrum are shown in the table in Figure 8.7 (b). When fit with the two spectral components (Equation 8.1), we find an integrated o-Ps flux of 5.8±0.1×10^4 cts, and an integrated Gaussian line component of 1.22±0.03×10^4 cts. The flux ratio gives a f_{Ps} of 1.01±0.01, which is consistent with the simulated positronium fraction. The integrated fluxes of the two components and the calculated f_{Ps} are listed in the table in Figure 8.7 (b). The 511 keV line centroid and width are consistent with the unsubtracted spectral parameters in Figure 8.3.

Though the $I_{3\gamma}/I_{2\gamma}$ ratio measured in the subtracted spectrum is correct to within a percent, the flux values are about 80% lower than those measured in the source-only spectrum shown in Figure 8.3 (c). This is due to the disk component of the emission. Since the disk extends further out than our 16° origin cut, some of the source photon from the disk will be within our defined background region $\mathbf{BR_{out}}$ from which the background spectrum is determined.

Figure 8.8 shows the Skinner source projected into the 2D CDS. Each event is a black point in the data space, the boundaries of the source region are indicated with the red dashed lines, and the boundaries of the background region are blue. Since this simulation includes no background component, any event within $\mathbf{BR_{out}}$ is from the disk emission. For the 511 keV line, we find 1.4×10^4 cts are in the source region and 2.5×10^3 cts are in $\mathbf{BR_{out}}$. Since the background spectrum is obtained from $\mathbf{BR_{out}}$, any source counts within this region will subtract off the same contribution in the source region. Therefore,

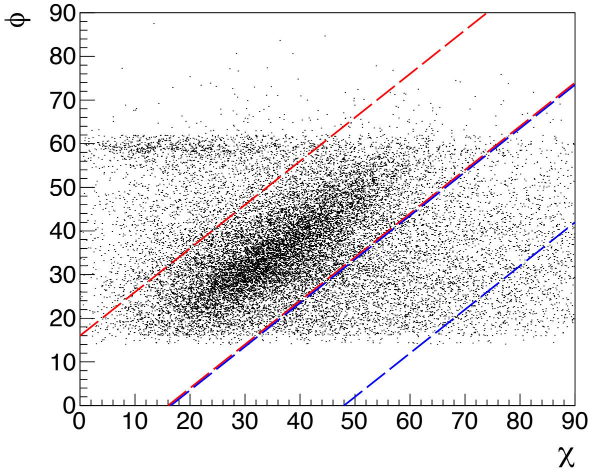
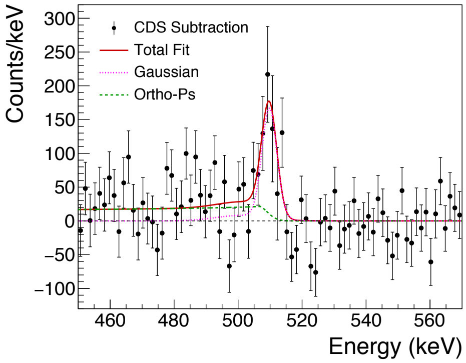


Figure 8.8: 2D CDS distribution of the Skinner model simulation. The red-dashed boundaries define the source region **SR** and the blue-dashed boundaries define $\mathbf{BR_{out}}$. The higher density of points near the $\phi = \chi$ line are from the bulge emission, while the data points spread away from the **SR** are from the disk emission.

the source signal would be reduced by 2.5×10^3 cts, and we expect to only measure 1.15×10^4 cts between 506–516 keV in the subtracted spectrum. This is exactly what we find. The number of counts with 506–516 keV in the subtracted spectrum is $1.12\pm0.03\times10^4$ cts. This reduction is unavoidable with the very extended nature of the positron annihilation signal and our choice of background region.

From the strong Skinner model simulation, we have shown that the CDS subtraction routine can recover the correct spectral parameters, including f_{Ps}. However, it may be more interesting to analyze a simulation with the expected Galactic flux to more closely compare to the background subtraction from the COSI flight. Figure 8.9 (a) shows the resulting CDS background-subtracted spectrum for a 16° region around the GC, when the simulated flux in the central region is 1.33×10^{-3} $\gamma/\mathrm{cm}^2/\mathrm{s}$. This is one-tenth the strength of the previous simulation and accurately describes the Galactic positron annihilation flux from the SPI models.

The total number of counts between 506–516 keV in the subtracted spectrum shown in Figure 8.9 is 960±240 cts. From the source-only spectrum we expect to detect 1401 cts, and therefore we are only recovering 70% of the source photons within this 16° region; as explained above, we expect this



(a) CDS subtracted spectrum.

Parameter		Value
Gaussian Fit	μ	510.2±0.7 keV
	σ	2.3±0.5 keV
	A	158±48 cts/keV
o-Ps Fit	B	9.6±3 cts/keV
511 keV Integral		1060±190 cts
o-Ps Integral		4265±1300 cts
f_{Ps}		0.97±0.09

(b) Spectral parameters from fit.

Figure 8.9: Spectral fit of the simulated Skinner model with the correct Galactic flux. The fit parameters are listed in table in (b).

to be due to the disk emission in the background spectrum. The number of counts in the background spectrum for these selections is 3.0×10^4 cts, and therefore the signal-to-noise ratio of the combined simulation is 5%.

The fit parameters for this background-subtracted spectrum are shown in the table in Figure 8.9 (b). The Gaussian component describing the 511 keV emission has a centroid of 510.2±0.8 keV, and the line width was found to be 2.3±0.5 keV; both are in agreement with the measured parameters from the stronger simulations shown in Figure 8.7. The calculated f_{Ps} is also consistent with the simulated positronium fraction of 1.

To test the statistical significance of a spectral feature, we use the F-test. The F-test, which is named in honor of the statistician Sir Ronald A. Fisher, compares the sum of squares of the null hypothesis (without the spectral feature) and an alternative model, taking into account the degrees of freedom of each fit. The significance of the spectral feature in the alternative model is then based on the P-value of the F-test. With the F-test, the significance of the spectral detection in Figure 8.9 is 3.5 σ.

We have shown that we can recover the correct spectral parameters with the CDS subtraction for a source with a signal-to-noise ratio of 5%. This give us confidence that the spectral background subtraction routine will recover the true parameters of the Galactic annihilation emission. With the inclusion of the COSI aspect information in these simulations, we expect our measured flight results to closely match the spectrum in Figure 8.9.

8.2.1 FLUX OF POSITRON ANNIHILATION EMISSION

To convert the measured counts into a source flux, we can use the Skinner model simulation to calculate the effective area. Since we do not use a pointing selection in our analysis routine, even the times when the GC is not in the FOV need to be simulated and included in a *flight-averaged* effective area $\overline{A}_{Eff}$. For the Skinner model simulation with 10 times the Galactic flux, we use the flight aspect information and take into account the drops in altitude to calculate the correct attenuation in the atmosphere. as discussed in Section 7.3.2.1. With a 16° pointing cut, we find 8775 counts in the source-only spectrum in `mimrec` between 506–516 keV. The flight-averaged effective area from this simulation is then

$$\begin{aligned} \overline{A}_{Eff} &= \frac{N_{det}}{\text{Flux} \times \text{time}} \\ &= \frac{8775 \text{ cts}}{0.0133\ \gamma/\text{cm}^2/\text{s} \times 3.08 \times 10^6 \text{ s}} = 0.214 \text{ cm}^2. \end{aligned} \tag{8.5}$$

The exposure time from the full flight simulation is 3.08×10^6 s, and the flux from the 16° region around the GC is 0.0133 $\gamma/\text{cm}^2/\text{s}$. As a test, we can use this calculated effective area to convert the number of counts detected in the CDS background-subtracted spectrum of the weaker Skinner model in Figure 8.9 into a flux:

$$\begin{aligned} \text{Flux} &= \frac{N_{det}}{\overline{A}_{Eff} \times \text{time}} \\ &= \frac{960 \pm 240 \text{ cts}}{0.214 \text{ cm}^2 \times 3.08 \times 10^6 \text{ s}} = 1.5 \pm 0.4 \times 10^{-3}\ \gamma/\text{cm}^2/\text{s}. \end{aligned} \tag{8.6}$$

This agrees with our simulated flux of 1.3×10^{-3} $\gamma/\text{cm}^2/\text{s}$.

This flight-averaged effective area is different than what was found for the point source simulation in Section 7.3.2.2: 0.175 cm^2. This is due to the extended nature of the Skinner model. When finding the effective area, we use an origin cut and all photons that are consistent with the source region will be counted in N_{det}. In the point source simulation, all of the photons simulated were from the same location; therefore, the origin cut is limiting the count of detected photons to only those that are properly reconstructed and consistent with the source location. However, in the Skinner model simulation, there will be a significant fraction of photons from the disk component that will be consistent with the source region, in addition to the events originating from within the origin cut. This gives a larger N_{det} and, in turn, a larger flight-averaged effective area. Therefore, the effective area calculation is dependent on the model of the background. There will always be a bleed-in from the disk photons into our 16° region. We must therefore make sure to include the disk component in simulation for the effective area calculation, as we do above. If the disk component of the Galactic emission is stronger than the description in the Skinner model, either spatially or in intensity, then the calculated flight-averaged effective area shown in Equation 8.5 would underestimate this extra disk contribution, and the resulting flux would be falsely enhanced.

8.3 BACKGROUND-SUBTRACTED CDS-ARM DISTRIBUTION OF GC POSITRON ANNIHILATION SIMULATIONS

With the simulated background radiation added to the Skinner model simulation, we can determine the background-subtracted CDS-ARM distribution, shown in Figure 8.10 for the strong Skinner model simulation. A 16° origin cut is used in the spectral subtraction routine to find the ϕ-dependent background spectra, and the relative number of counts in the two energy ranges (506–516 keV and 520–720 keV) in the scaled background spectra determines the scaling of each ϕ-dependent background CDS-ARM distribution (Section 7.4.2). So, although there is no origin cut in the ARM distributions, the background ARM will be scaled in a different way depending on the origin cut. This is explored in more detail in Appendix A, where the event selections for the CDS subtraction analysis are discussed.

The fit to the subtracted CDS-ARM distribution in Figure 8.10 gives a FWHM of 13.6°, which is smaller than the width of the full Skinner model distribution. The number of counts in the peak of the background-subtracted distribution is 7439±300 cts. This is only ~85% of the expected number of counts within the inner 16° region of the simulated Skinner model, but is consistent with the counts from only the bulge component: 7225 cts. The distribution at larger ARM values (> 40°) drops below 0 cts/deg; however, this is where the disk emission is seen in the unsubtracted distribution (Figure 8.4). This undershoot occurs because the $\mathbf{BR_{out}}$ region covers much of the disk region (Figure 8.8) and we are essentially subtracting this component off. In this way, we are only sensitive to the bulge emission in the subtracted CDS-ARM distribution. In addition, the measured width does not perfectly match the budge-only ARM distribution shown in Figure 8.4. Therefore,

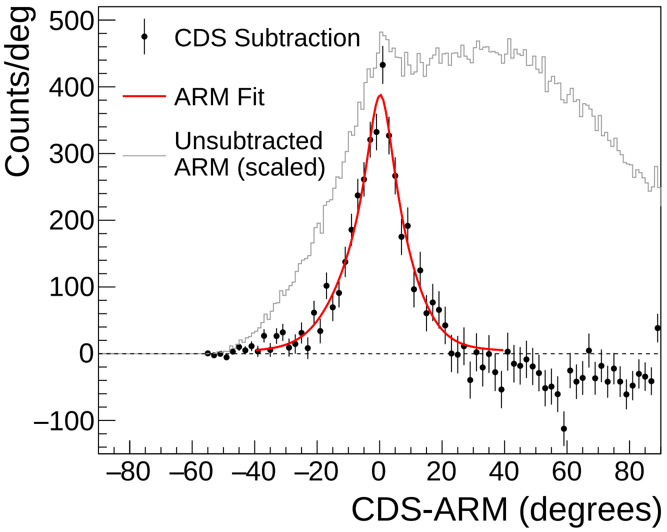
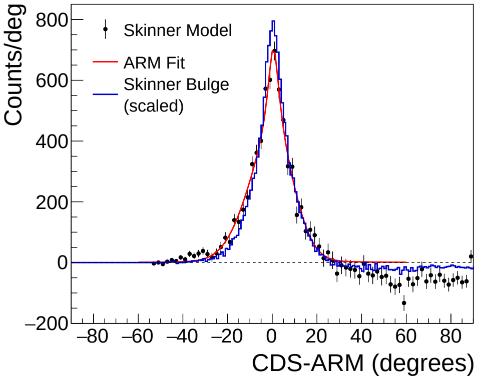


Figure 8.10: Background-subtracted CDS-ARM distribution for the 506–516 keV line from the Skinner model simulation. The fit to this distribution give 13.6° FWHM and there are 7439±300 cts within the peak. The raw unsubtracted distribution for the same energy range and event selections is scaled by $\frac{1}{3}$ and shown in gray.

we will analyze the emission shape by comparing the CDS-ARM results to the background-subtracted distribution of a very strong simulation that includes only the Skinner bulge component.

Taking a closer look at the CDS-ARM background-subtracted distribution from the strong Skinner model simulation, Figure 8.11 (a) compares the distribution with the expected distribution from only the Skinner bulge component, shown in blue. There is a good overlap between the two distributions for the total distribution except at the higher ARM values where the disk contribution results in a negative excess. The distribution is fit with Equation 8.4 with only one Lorentzian component, and the parameters are listed in the table in Figure 8.11 (b). The FWHM is determined from the fit function and the error was estimated after preforming the fit multiple times with different bin sizes. From this, we conclude that our CDS-ARM subtraction routine allows us to directly compare the measured distribution of the bulge emission with that of the Skinner model.

Figure 8.12 shows the subtracted CDS-ARM distribution from the weaker Skinner model simulation, which is derived from the spectral subtraction in Figure 8.9. With the poor statistics in this distribution, we preform a fit with a single Gaussian shown in red, instead of Equation 8.4, and the fit parameters

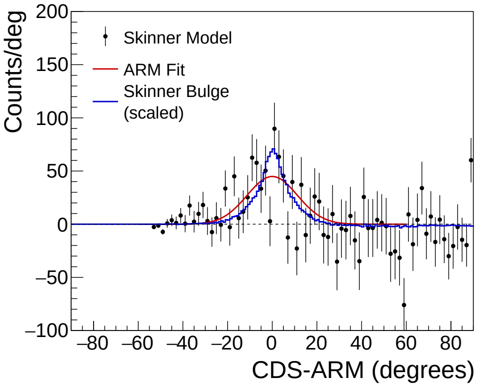


(a) Background-subtracted CDS-ARM distribution of strong Skinner source.

Parameter		Value
ARM Fit	μ	0.5±0.3°
	I_{L1}	32±54 cts/deg
	Γ_1	3.3±0.6°
	A	37±34 cts/deg
	$\sigma_>$	9.1±0.6°
	$\sigma_<$	11.3±0.6°
FWHM		14±1°

(b) CDS-ARM fit parameters.

Figure 8.11: The background-subtracted CDS-ARM distribution for the strong Skinner model with E_γ = 506–516 keV and the event selections listed in Table 8.1. This result is derived from the subtracted spectrum shown in Figure 8.7. The fit parameters from the fit function in Equation 8.4 are listed in (b); however, only one Lorentzian component was needed to give a satisfactory fit.



(a) Background-subtracted CDS-ARM distribution of weak Skinner source.

Parameter		Value
Gaussian Fit	μ	fixed at 0
	σ	11±3°
	A	45±10 cts/deg
FWHM		23±7°

(b) CDS-ARM fit parameters.

Figure 8.12: The background-subtracted CDS-ARM distribution for the weak Skinner model with E_γ = 506–516 keV and the event selections listed in Table 8.1. This result is derived from the subtracted spectrum shown in Figure 8.9. The CDS-ARM distribution is fit with single Gaussian function because the statistics does not warrant more fit parameters. The resulting fit parameters are listed in (b).

are listed in (b). We have included the expected Skinner bulge distribution, which is scaled to match the area under the curve in the subtracted distribution and is shown in blue. There is still a good visible match between the background-subtracted distribution with this weak source and the Skinner bulge distribution. A single fitted Gaussian will always be wider than the ARM fit function, but with the large error bars here, the widths are in agreement.

Unfortunately, we are not sensitive to the disk morphology with the background-subtracted CDS-ARM method presented here; however, we can recover the correct bulge emission. Directly comparing the measured CDS-ARM to that of the Skinner bulge component allows us to determine if the spatial distribution is consistent without having to perform a complicated deconvolution. We can therefore compare our background-subtracted CDS-ARM distribution from the COSI flight data to the simulations presented in this chapter to confirm whether or not the Skinner model description is an accurate representation of the Galactic bulge emission.

9

DETECTION OF THE GALACTIC POSITRON ANNIHLATION EMISSION

We have developed a background subtraction routine using the CDS to detect the Galactic positron annihilation emission. In this chapter we will give an overview of the final results from the 2016 COSI flight and make a comparison to the simulations of the Galactic emission discussed in Chapter 8. In Section 9.1, we will present the measured background-subtracted spectrum of the Galactic emission, which gives a 7 σ detection of the 511 keV line. In Section 9.2, we will present the results for the background-subtracted CDS-ARM distribution, which reveals a broader distribution than what has been previously reported. Finally, in Section 9.3, we will present some preliminary results regarding the binned-mode image of the 511 keV emission from the Galaxy.

9.1 BACKGROUND-SUBTRACTED POSITRON ANNIHILATION SPECTRUM

Figure 9.1 shows our final measured spectrum for a 16° origin cut around the GC after applying the CDS background subtraction routine described in Chapter 7. The significance of the Galactic 511 keV line is 7.2 σ (calculated with the F-test; see Section 8.2). The event selections for this spectral subtraction are the same as were used for the analysis in Chapter 8 and are listed again in Table 9.1. We measure 2500±280 cts between 506–516 keV, and have fit the spectrum with a Gaussian line component and the o-Ps continuum, where the spectral fit parameters are listed in Table 9.2.

The spectral fit of the background-subtracted signal of Galactic positron annihilation gives a Gaussian centroid at 511.8±0.3 keV with a line width of σ = 2.5±0.3 keV. The centroid, though slightly higher than the expected 511.0 keV, is not significantly shifted considering a Gaussian fit to the 511 keV

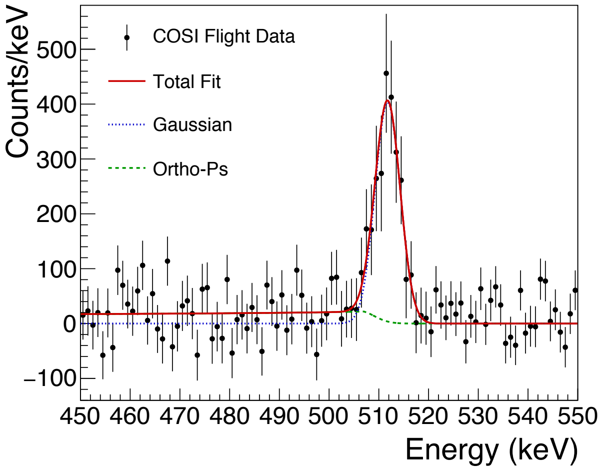


Figure 9.1: CDS background-subtracted spectrum of the positron annihilation emission detected within a 16° region around the GC from the COSI 2016 flight. The total number of counts detected between 506–516 keV is 2500±280 cts. The spectrum is fit with a single Gaussian component to describe the 511 keV line and the theoretical o-Ps continuum spectrum.

Parameter	Allowed Range
Pointing selection	None
Altitude	27000 m
Origin selection	16° (if applicable)
Photon energy	506 – 516 keV (if applicable)
Number of interactions	2 – 7
Compton scatter angle	15 – 55°
Distance between first 2 interactions	> 0.5 cm
Distance between any interaction	> 0.3 cm
Earth horizon cut	Reject if 99% of Compton circle is below horizon

Table 9.1: Event selections for the CDS background-subtracted spectrum from the COSI flight shown in Figure 9.1. The origin selection is used only for the spectral subtraction, and the photon energy cut is only used for the CDS-ARM subtraction. Refer to Section 3.5.5.1 for an overview of the `mimrec` event selections.

Parameter		Value
Gaussian Fit	μ	511.8±0.3 keV
	σ	2.5±0.3 keV
	A	403±57 cts/keV
o-Ps Fit	B	12±4 cts/keV
χ^2/d.o.f.		193.0/196
511 keV Integral		2560±300 cts
o-Ps Integral		5110±1700 cts
f_{Ps}		0.76±0.12

Table 9.2: Fit parameters for the COSI flight background-subtracted spectrum shown in Figure 9.1. The fit is made over the energy range 450–550 keV. The χ^2 of the fit, the calculated integrals of both spectral components, and the resulting f_{Ps} are listed.

line in the full-flight integrated spectrum gives a centroid of 511.54±0.01 keV, which reveals that our energy calibration is not yet perfect. We can determine the measured line width of the Galactic signal by subtracting in quadrature the inherent detector line width at 511 keV: 1.85±0.1 keV. From the fit line width σ = 2.5±0.3 keV, we determine the width of the Galactic 511 keV line to be σ = 1.7±0.4 keV. In comparison, Siegert et al. (2016a) report an average Galactic line width of 1.11±0.01 keV and centroid at 511.15±0.09 keV.

In the spectral study by Jean et al. (2006), discussed in Chapter 2, the authors reported a better fit to the Galactic 511 keV line with two Gaussian components, a narrow line with σ = 0.6 keV and a broad line with σ = 2.3 keV; however, the fit to the COSI data which has a χ^2/d.o.f. of 193.0/196, is satisfactory with only a single Gaussian. The low-energy bump from charge loss that is included in simulated spectra fits in Chapter 8 is not used here.

The measured intensity of the 511 keV line is higher than predicted from our simulations with the Skinner and Siegert models. From our simulation results (Figure 8.9), we expect approximately 1060±190 cts within 506–516 keV in the subtracted spectrum from the Skinner model with all of the same event selections, but we find over twice this number. Using the flight-averaged effective area from Section 8.2.1, the flux is calculated to be

$$\text{Flux} = \frac{2500 \pm 280 \text{ cts}}{0.21 \text{ cm}^2 \times 3.08 \times 10^6 \text{ s}} = 3.9 \pm 0.4 \times 10^{-3}\ \gamma/\text{cm}^2/\text{s}, \qquad (9.1)$$

using the flight-averaged effective are calculated in Section 8.2.1. For comparison, Siegert et al. (2016a) report a total Galactic 511 keV line flux of $2.74{\pm}0.03{\times}10^{-3}$ $\gamma/\text{cm}^2/\text{s}$.

The relative intensity of the o-Ps continuum results in a surprisingly low f_{Ps} of 0.76±0.12. This is a smaller fraction than other reported measurements, which all mostly agree with $f_{Ps} \sim 1$. The number of counts in the integrated o-Ps continuum is 5110±1700 cts, which is consistent with the Skinner model simulation shown in Figure 8.9, where an excess of 4265±1300 cts was found.

To calculate the detection significance of the measured f_{Ps}, we perform an F-test. First, with the o-Ps amplitude B fixed at 0, the reduced chi-squared of the fit is χ^2/d.o.f. = 202/197. When this is compared to the fit parameters with the o-Ps component shown in Table 9.2 the F-test value is 9.14, which corresponds to a 3.0 σ significance.

For an estimate of the statistical significance of our low o-Ps continuum measurement compared to the expected intensity, we repeat another F-test, but instead of fixing the o-Ps amplitude B to zero we force it to be equivalent to $f_{Ps} = 1$. From our simulations, we know the expected relative ratio of the 511 keV line flux and the o-Ps amplitude; therefore, assuming the measured 511 keV amplitude is correct, we calculate that a measured o-Ps amplitude of 25 cts/keV would give $f_{Ps} = 1$. Fixing the o-Ps amplitude to the expected 25 cts/keV gives a fit with χ^2/d.o.f. = 205.6/197. When we compare the fit result to the parameters in Table 9.2 and perform an F-test, we find a 3.5 σ significance of our o-Ps amplitude of 12±4 cts/keV compared to the expected amplitude of 25 cts/keV at 511 keV.

The spectral results from the COSI flight show a measured 511 keV flux that is larger than what the SPI collaboration has reported, while the f_{Ps} is lower by ~3 σ. One possible explanation for the high flux value is that the measured spatial distribution does not agree with the Skinner or the Siegert model. As discussed in Section 8.2.1, a larger disk contribution would result in a larger number of events that are consistent with the inner region of the Galaxy, and therefore our calculation of the flight-averaged effective area using the Skinner model would result in a falsely high flux. However, this scenario does not offer an explanation for the low f_{Ps}. The discrepancy, though only a 3 σ significance, is a potential signature of a previously-undetected emission component with a smaller f_{Ps} and therefore higher positron energy. Further investigation is necessary before any conclusions can be reached.

Another explanation is that our calculated 511 keV flux could be overestimated, particularly since the backgrounds at these energies are known to

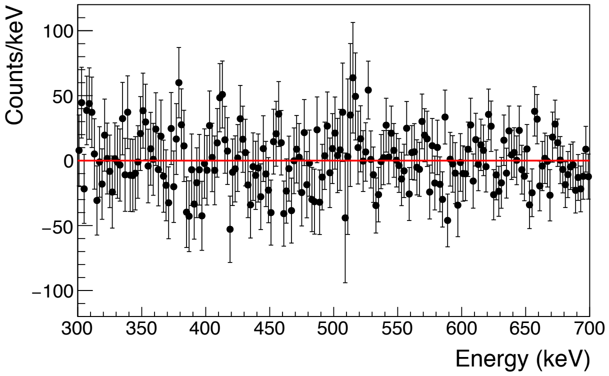


Figure 9.2: CDS background-subtracted spectrum of flight data with Galactic coordinates (120, -60) chosen to be the source location. The flat spectrum further confirms that the systematics in the CDS subtraction are minor.

be plagued with systematics. In an attempt to rule out this possibility, we performed detailed background simulations (Chapter 7) and a thorough validation of our subtraction method for Galactic models (Chapter 8). The results have confirmed that we can determine the correct spectral shapes and that the f_{Ps} is preserved through the background subtraction with simulated data. In addition, we can perform the spectral subtraction with flight data for a source location far away from the GC. Figure 9.2 shows the resulting background-subtracted spectrum from the flight when the source location is chosen to be at (120, -60) in Galactic coordinates. The moderately flat spectrum further confirms the legitimacy of our routine. Although our analysis is self-consistent in the numerous checks we have preformed, it is possible that some detail may have been missed or overlooked. This will be discussed further in the conclusion of this chapter.

9.2 BACKGROUND-SUBTRACTED ARM DISTRIBUTION OF GALACTIC POSITRON ANNIHILATION

With only 46 days of data (and poor observations due to drops in altitude and the loss of 3 detectors), we are not expecting to compete with INTEGRAL/SPI spectral measurements that include over 15 years of data. Where COSI can make a difference at this point is with measurements of the spatial distribution of the 511 keV line. As explained in Chapter 2, SPI relies on a model fitting approach to determine the morphology of the Galactic positron anni-

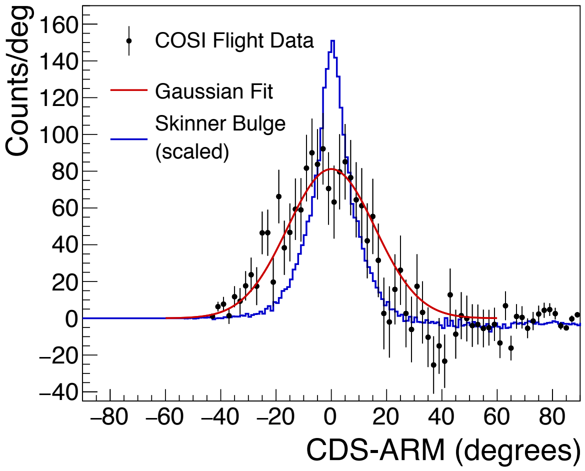


Figure 9.3: Background-subtracted CDS-ARM distribution as measured in the 2016 COSI flight for E_γ = 506–516 keV. The distribution is fit with a single Gaussian, which gives a FWHM of 33±2°. The event selections for this analysis are listed in Table 9.1; in addition, a 40° pointing cut is used.

hilation emission, and this has many limitations. A background-subtracted measurement of the CDS-ARM distribution is a much more direct way to measure the spatial extent of the Galactic emission. However, the analysis we presented in Chapter 8 concludes that we are only sensitive to the bulge emission with the current dataset and techniques.

Figure 9.3 shows the background-subtracted CDS-ARM distribution for the 511 keV line as measured in the 2016 COSI flight with a 40° pointing selection on the GC. The distribution is fit with a single Gaussian and the measured FWHM of 33±2° is significantly larger than the simulated CDS-ARM distributions of the Skinner model bulge emission shown in blue. The parameters from the Gaussian fit of the CDS-ARM distribution are shown in Table 9.3. A 40° pointing selection was used for this analysis because it was found to decrease the error of the measured distribution; when no pointing selection is used, the measured FWHM is 32±4°.

The Skinner model bulge distribution that is shown in blue in Figure 9.3 is the same as was shown in Figure 8.11 and 8.12, and it has been scaled so that the area under the curve is the same as the flight data CDS-ARM distribution. This visually shows the difference in widths of the Skinner model bulge distribution and the detected spatial distribution at 511 keV. To test the difference between these two histograms, we perform a chi-squared test, which gives a

Parameter		Value
Gaussian Fit	μ	fixed at 0
	σ	14.0±0.7°
	A	89±0.6 cts
χ^2/d.o.f.		52.1/52
FWHM		33±2

Table 9.3: Fit parameters for the flight data CDS-ARM distribution shown in Figure 9.3. The distribution is fit with a single Gaussian since the statistics do not warrant the 8-parameter fit of Equation 8.4.

P-value = 0.001, and therefore there is a 3 σ statistical significance between the COSI distribution and the Skinner bulge distribution.

A cautionary point to make is that all of the COSI detected point sources from the 2016 flight seem to be off by a few degrees in the list-mode images (Section 5.3). This could be due to a number of reasons, but the team's efforts to track down the cause have been unsuccessful. One avenue explored was the possibility that the COSI aspect reconstruction was off by a few degrees, although no issues with the reconstruction could be found. Regardless, this would result in a smearing of the COSI observations, which could explain why our list-mode images of point sources are displaced by a few degrees. As a test of this idea, we have calculated the significance of the measured CDS-ARM distribution compared to the Skinner model bulge after a convolution with a $\sigma = 3°$ Gaussian, as a liberal estimate for this effect. A chi-squared test comparing our measured distribution and a blurred Skinner bulge component shows no statistically significant difference. Therefore, this could be an explanation for the measured broad distribution. This is a somewhat unsatisfactory conclusion, but needs to be considered as a possibility. This hypothesis can be checked by performing a background-subtracted ARM measurement or binned-mode image of one of the point sources COSI detected during the flight. However, these techniques have not yet been implemented for COSI's broadband sources.

The connection with the potentially overestimated flux in the spectral subtraction must also be considered. These measurements are indeed related, since the background spectral subtraction determines the background CDS-ARM distribution scaling factors. If the excess counts in the detected 511 keV line are from a poorly modeled background component, then it is most likely that the background component would have a spatial distribution similar the

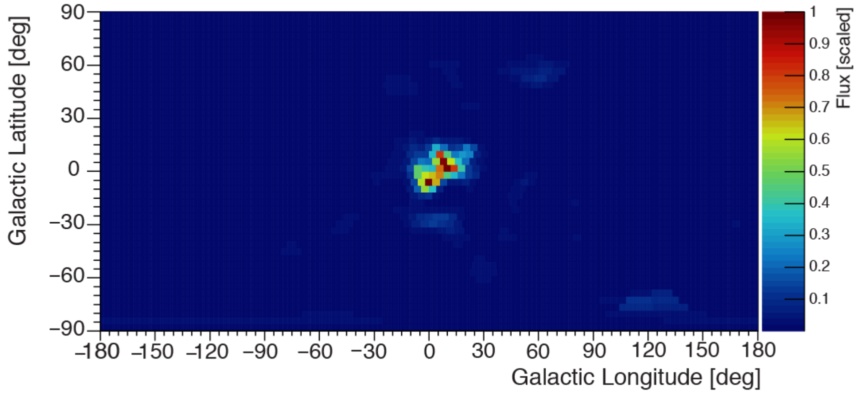


Figure 9.4: Preliminary binned-mode maximum-entropy image of the 511 keV emission from the 2016 COSI flight (Zoglauer et al., in prep).

FOV of the instrument; however, this would give a FWHM > 50°, which is much larger than the distribution that we measure and therefore seems improbable.

The large extended shape of the COSI detected CDS-ARM distribution is intriguing. In the analysis of SPI data, Skinner et al. (2014) report a bulge distribution composed of a narrow (FWHM = 5.8°) and broad (FWHM = 20.5°) Gaussian and point source component. Siegert et al. (2016a) also use this bulge description for their spectral studies, but report an alternate, yet equally significant, model in the paper's appendix. This alternate model has the morphology of the bulge emission described by two elongated components, with a longitudinal extent of the broad bulge up to FWHM ~ 55°. The COSI results seem to agree more with this alternative bulge morphology; however, more data is needed to make a strong conclusion about the spatial distribution.

9.3 BINNED-MODE IMAGE OF GC EMISSION

Figure 9.4 shows the binned-mode image of the 511 keV emission from the 2016 COSI flight using MEM image deconvolution (Zoglauer et al., in prep). This is the first time binned-mode imaging has been implemented for a CCT that is not operated in pointing mode. Figure 9.4 is the first direct image of the positron annihilation emission. However, these results are still preliminary.

Figure 9.4 uses 5° image bins, where this is not limited by the COSI spatial resolution so much as by the low statistics of the observation. The event selections for this image are the same as listed in Table 9.1. The 511 keV emission is seen strongly in the GC. Neither the disk emission nor a possible halo structure is observed. The central distribution is slightly shifted toward positive latitudes, although SPI reports an offset in the Galactic bulge emission, the shift seen here is more likely a product of our asymmetric exposure of the region; see Figure 5.12 for the full-flight exposure map.

The substructure within the central emission component is not necessarily real but most likely dominated by the statistics of the observation, as it is seen to change with event selections and the number of iterations. Aside from confirming the strong central emission, it is difficult to make any new conclusions about the spatial distribution of the positron annihilation emission from this image, but it exemplifies the power of Compton imaging and the need for more observations.

9.4 CONCLUSIONS FROM COSI GALACTIC POSITRON ANNIHILATION DETECTIONS

We have reported the first detection of the Galactic positron annihilation signal from the Galaxy with a compact Compton telescope, and to perform this analysis, we have developed an accurate background subtraction technique that is valid for sources of line-emission. In comparison with our first attempts using more traditional techniques (Section 6.2), our results are a significant improvement.

We have found a 7.2 σ detection of the 511 keV line from the Galactic center region, and a 3.0 σ detection of the o-Ps continuum after 46 days of flight. However, the relative ratio of the 511 keV line and o-Ps continuum results in a low measurement of $f_{Ps} = 0.76 \pm 0.12$.

By performing detailed γ-ray background simulations and validating the subtraction routine with models of the positron annihilation emission, we have shown we can recover the simulated spectral parameters, including f_{Ps}. However, we have also used these simulations to identify a number of limitations with our current background subtraction method. With background regions chosen adjacent to the source region in the CDS, we subtract off some of the photons due to the extended emission, which corresponds to a loss of about 70% from the analysis of the Skinner distribution. This process also results in a loss of spatial info about the disk emission that is contained within our chosen background regions. Therefore, the method as it currently stands is only sensitive to the central bulge emission in the subtracted CDS-ARM distribution.

Given the high measured flux of the 511 keV line component and the low o-Ps fraction, it is conceivable that we are detecting an excess of 511 keV photons from an improper background subtraction (though the flat spectrum above 511 keV and the lack of other systematics seem to disagree). Since the simulations in Chapter 8 showed that we can recover the correct f_{Ps} ratio with this method, there must be something about the flight data that we are not properly modeling, even though attempts were made to match the simulations to the flight emission. To attempt to find the cause of this excess would require a much more detailed background simulation of both the γ-ray radiation and the instrument activation components. For a proper treatment of the activation, the simulation would require different photon and particle differential fluxes for the full range of altitudes and geomagnetic locations to give the correct activation for each day. A full and accurate background model would require a significant effort and would be a dissertation in itself.

Nonetheless, the results from this study of Galactic positron annihilation are intriguing. The flux measurements hint at a morphology that is different than the Skinner model, and the detected angular distribution is broader than the emission models presented by the SPI collaboration. The results presented here show the power of Compton telescopes and the CDS analysis, and the need for more data is clear.

10

SUMMARY AND OUTLOOK

The origin of the positrons in the Milky Way is one of the longest standing questions in γ-ray astrophysics. The signature 511 keV emission was first detected coming from the GC region in the 1970's. Now, with over 15 years of INTEGRAL/SPI observations, significant progress in our knowledge of Galactic positrons has been made. Spectral measurements conclude that positrons almost exclusively form positronium at low energies before annihilating in warm phases of the ISM, giving rise to the observed low-energy o-Ps continuum below 511 keV. From early observations, we know there is a strong concentration of the annihilation emission in the GC region. Recent results from SPI confirm emission associated with the disk of the Galaxy, though the spatial extent is poorly constrained. The positron population in the Galactic disk is fairly easy to explain with β^+ decay of stellar nucleosynthesis products, namely ^{26}Al, ^{44}Ti, and ^{56}Ni. However, the source of positrons in the Galactic bulge region is still not clear.

One major uncertainty is the distance that positrons can propagate from their birth sites before annihilation. Until this is better understood, either through more sophisticated simulations of interactions within the ISM or from direct observations, it will be difficult to understand the relative contribution of the multitude of proposed positron sources to the annihilation signal. More constraining measurements, specifically of the morphology of the 511 keV emission, are needed.

The Compton Spectrometer and Imager is a balloon-borne γ-ray telescope that is designed to further our understanding of Galactic positrons, in addition to its goals of preforming novel γ-ray polarization measurements and mapping nucleosynthesis in the Galaxy. With the direct imaging capabilities of the compact Compton telescope design, COSI does not have to rely on a model-fitting method to determine the spectral signatures or spatial distribution of the emission. Therefore, even with the relatively short, background-dominated observations from a balloon-platform, COSI can contribute to the scientific questions that remain.

COSI had a record-breaking flight from Wanaka, New Zealand, in 2016 onboard NASA's new SPB. The flight from the southern hemisphere resulted in great exposure of the GC region necessary for positron annihilation studies. To extract the Galactic signal from the flight data, a background subtraction routine was developed based on methods from the COMPTEL collaboration. The formulated CDS routine for background subtraction of sources with line emission was presented in Chapter 7 and thoroughly validated through Galactic simulations in Chapter 8. With this routine, we obtain a 7 σ significance of the 511 keV line and a 3 σ detection of the o-Ps continuum from the COSI 2016 flight. The spatial distribution measured is significantly broader than the recently reported models of the morphology; however, the new binned-mode imaging algorithm does not yet show an indication of the disk emission. The new techniques developed here are in want of more data for definitive conclusions about the measured distributions.

To further our understanding of Galactic positrons, we require a next generation γ-ray telescope with direct imaging capabilities, an angular resolution on the order of 1°, and sensitivities ~10^{-6} $\gamma/\mathrm{cm}^2/\mathrm{s}$, all while still maintaining a high spectral resolution of $\lesssim$1%. Looking forward, there are two large mission concepts with aims to fill the MeV gap that will realize many of these science requirements: the U.S.-led All-sky Medium Energy Gamma-ray Observatory (AMEGO) and the European e-ASTROGAM[1] projects. Both missions have almost identical detector designs with cross-strip silicon plane that will function as a Compton telescope in the soft γ-ray regime and alternately as a pair-tracking telescope at energies up to ~10 GeV. On a smaller scale, the COSI mission has been recently selected for a Phase-A concept study exploring the gains of finer strip-pitch GeDs, which will improve the angular resolution by a factor of 2 at 511 keV. COSI remains at the forefront of Compton telescope design and development, and there is no doubt the COSI mission will provide significant contributions to the future success of the proposed MeV missions in the endeavor to shed light on Galactic positrons and their true cosmic source.

---

[1] Unfortunately, e-ASTROGAM recently missed the selection within the ESA M5 call; however, it is now expected that the majority of the e-ASTROGAM collaboration will join the AMEGO effort.

BIBLIOGRAPHY

Abbott, D. C., & Conti, P. S. 1987, Annual Review of Astronomy and Astrophysics, 25, 113

Adkins, G. S., McGovern, N. M., Fell, R. N., & Sapirstein, J. 2003, Physical Review A, 68, 032512

Agostinelli, S., Allison, J., Amako, K., et al. 2003, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 506, 250

Albernhe, F., Le Borgne, J. F., Vedrenne, G., et al. 1981, Astronomy and Astrophysics, 94, 214

Alexis, A., Jean, P., Martin, P., & Ferrière, K. 2014, Astronomy and Astrophysics, 564, A108

Amman, M., & Luke, P. N. 2000, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 452, 155

Amman, M., Luke, P. N., & Boggs, S. E. 2007, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 579, 886

Anderson, C. D. 1932, Science, 76, 238

Bambi, C., Dolgov, A. D., & Petrov, A. A. 2008, Physical Letters B, 670, 174

Bandstra, M. S. 2010, PhD thesis, University of California, Berkeley

Bandstra, M. S., Bellm, E. C., Boggs, S. E., et al. 2011, The Astrophysical Journal, 738, 8

Bardin, G., Duclos, J., Magnon, A., et al. 1984, Physics Letters B, 137, 135

Bé, M.-M., Chisté, V., Dulieu, C., et al. 2016, Monographie BIPM-5, Vol. 1, Table of Radionuclides (Bureau International des Poids et Mesures)

Beacom, J. F., & Yüksel, H. 2006, Physical Review Letters, 97, 071102

Beloborodov, A. M. 1999, Monthly Notices of the Royal Astronomical Society, 305, 181

Berger, M. J., Hubbell, J. H., Seltzer, S. M., et al. 2017, XCOM: Photon Cross Sections Database. `http://physics.nist.gov/xcom`

Blattnig, S. R., Swaminathan, S. R., Kruger, A. T., Ngom, M., & Norbury, J. W. 2000, Physical Review D, 62

Blumenthal, G. R., & Gould, R. J. 1970, Reviews of Modern Physics, 42, 237

Bœhm, C., & Ascasibar, Y. 2004, Physical Review D, 70, 115013

Bœhm, C., Hooper, D., Silk, J., Cassé, M., & Paul, J. 2004, Physical Review Letters, 92, 101301

Boggs, S. E., & Jean, P. 2000, Astronomy and Astrophysics Supplement Series, 145, 311

Boggs, S. E., Harrison, F. A., Miyasaka, H., et al. 2015, Science, 348, 670

Bouchet, L., Jourdain, E., & Roques, J.-P. 2015, The Astrophysical Journal, 801, 142

Bouchet, L., Roques, J. P., & Jourdain, E. 2010, The Astrophysical Journal, 720, 1772

Boulanger, F., Abergel, A., Bernard, J.-P., et al. 1996, Astronomy and Astrophysics, 312, 256

Bowen, J. D. 2009, PhD thesis, University of California, Berkeley

Bowen, J. D., Zoglauer, A., Bandstra, M. E., et al. 2006, in Bulletin of the American Astronomical Society, Vol. 38, High Energy Astrohysics Division Meeting #9, 385

Breit, G., & Wheeler, J. A. 1934, Physical Review, 46, 1087

Brun, R., & Rademakers, F. 1997, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 389, 81

Bussard, R. W., Ramaty, R., & Drachman, R. J. 1979, The Astrophysical Journal, 228, 928

Caroli, E., Stephen, J. B., Cocco, G. D., Natalucci, L., & Spizzichino, A. 1987, Space Science Reviews, 45, 349

Cembranos, J. A. R., & Strigari, L. E. 2008, Physical Review D, 77, 123519

Chan, K.-W., & Lingenfelter, R. E. 1993, The Astrophysical Journal, 405, 614

Chao, C. Y. 1930, Physical Review, 36, 1519

Cheng, K. S., Chernyshov, D. O., & Dogiel, V. A. 2007, Astronomy and Astrophysics, 473, 351

Cheng, K. S., Ho, C., & Ruderman, M. 1986, The Astrophysical Journal, 300, 500

Chernyshov, D. O., Cheng, K.-S., Dogiel, V. A., Ko, C.-M., & Ip, W.-H. 2010, Monthly Notices of the Royal Astronomical Society, 403, 817

Chi, X., Cheng, K. S., & Young, E. C. M. 1996, The Astrophysical Journal Letters, 459, L83

Churazov, E., Sunyaev, R., Sazonov, S., Revnitsev, M., & Varshalovich, D. 2005, Monthly Notices of the Royal Astronomical Society, 357, 1377

Churazov, E., Sunyaev, R., Isern, J., et al. 2014, Nature, 512, 406

Clayton, D. D. 1973, Nature Physical Science, 244, 137

—. 1984, The Astrophysical Journal, 280, 144

Coburn, W., Amrose, S., Boggs, S. E., et al. 2003, in Proceedings of SPIE, Vol. 4784

Cox, J. P., & Giuli, R. T. 1968, Principles of stellar structure (New York: Gordon and Breach)

Crannell, C. J., Joyce, G., Ramaty, R., & Werntz, C. 1976, The Astrophysical Journal, 210, 582

Crocker, R. M., Ruiter, A. J., Seitenzahl, I. R., et al. 2017, Nature Astronomy, 1, 0135

Daugherty, J. K., & Harding, A. K. 1982, The Astrophysical Journal, 252, 337

—. 1983, The Astrophysical Journal, 273, 761

de Jager, O. C., & Djannati-Ataï, A. 2009, Neutron Stars and Pulsars, ed. W. Becker (Springer), 451–479

Deutsch, M. 1951, Physical Review, 82, 455

Diehl, R., Halloin, H., Kretschmer, K., et al. 2006, Nature, 439, 45

Dirac, P. A. M. 1931, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 133, 60

Einasto, J. 1965, Trudy Astrofizicheskogo Instituta Alma-Ata, 5, 87=

Fabris, L., Madden, N. W., & Yaver, H. 1999, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 424, 545

Ferrer, F., & Vachaspati, T. 2005, Physical Review Letters, 95, 261302

Ferrière, K. 1998, The Astrophysical Journal, 497, 759

—. 2001, Reviews of Modern Physics, 73, 1031

—. 2009, Astronomy and Astrophysics, 505, 1183

—. 2012, Astronomy and Astrophysics, 540

Ferrière, K., Gillard, W., & Jean, P. 2007, Astronomy and Astrophysics, 467, 611

Fesen, R. A., Hammell, M. C., Morse, J., et al. 2006, The Astrophysical Journal, 645, 283

Forrest, D. J., Chupp, E. L., Ryan, J. M., et al. 1980, Solar Physics, 65, 15

Frampton, P. H., & Kephart, T. W. 2005, Modern Physics Letters A, 20, 1573

Frontera, F., & von Ballmoos, P. 2010, X-Ray Optics and Instrumentation Special Issue on X-Ray Focusing: Techniques and Applications, 2010

Gil-Pons, P., García-Berro, E., José, J., Hernanz, M., & Truran, J. W. 2003, Astronomy and Astrophysics, 407, 1021

Gilfanov, M., Churazov, E., Sunyaev, R., et al. 1994, The Astrophysical Journal Supplement Series, 92, 411

Goldreich, P., & Julian, W. H. 1969, The Astrophysical Journal, 157, 869

Goldsmith, P. F. 1987, in Interstellar Processes, ed. D. J. Hollenback & H. A. Thronson, 51–70

Gómez-Gomar, J., Hernanz, M., José, J., & Isern, J. 1998, Monthly Notices of the Royal Astronomical Society, 296, 913

Gould, R. J., & Schréder, G. P. 1967, Physical Review, 155, 1404

Greenberg, A. J., Ayres, D. S., Cormack, A. M., et al. 1969, Physical Review Letters, 23, 1267

Grefenstette, B. W., Harrison, F. A., Boggs, S. E., et al. 2014, Nature, 506, 339

Grimm, H.-J., Gilfanov, M., & Sunyaev, R. 2002, Astronomy and Astrophysics, 391, 923

Guessoum, N., Jean, P., & Gillard, W. 2005, Astronomy and Astrophysics, 436, 171

Guessoum, N., Jean, P., & Prantzos, N. 2006, Astronomy and Astrophysics, 457, 753

Guessoum, N., Ramaty, R., & Lingenfelter, R. E. 1991, The Astrophysical Journal, 378, 170

Harding, A. K., & Lai, D. 2006, Reports on Progress in Physics, 69, 2631

Harris, M. J., Share, G. H., & Leising, M. D. 1994, The Astronomical Journal, 433, 87

Haymes, R. C., Walraven, G. D., Meegan, C. A., et al. 1975, The Astrophysical Journal, 201, 593

Hernanz, M. 2005, in Astronomical Society of the Pacific Conference Series, Vol. 330, The Astrophysics of Cataclysmic Variables and Related Objects, ed. J. M. Hameury & J.-P. Lasota

Hernanz, M., & José, J. 2006, New Astronomy Reviews, 50, 504

Higdon, J. C., Lingenfelter, R. E., & Rothschild, R. E. 2009, The Astrophysical Journal, 698, 350

Hillebrandt, W., & Niemeyer, J. C. 2000, Annual Review of Astronomy and Astrophysics, 38, 191

Hoffman, R. D., Woosley, S. E., Weaver, T. A., et al. 1995, in Gamma Ray Sky with Compton GRO and SIGMA, ed. M. Signore, P. Salati, & G. Vedrenne, 267–278

Hooper, D., & Wang, L.-T. 2004, Physical Review D, 70, 063506

Hurley, K. 1978, Astronomy and Astrophysics, 69, 313

Hwang, U., Laming, J. M., Badenes, C., et al. 2004, The Astrophysical Journal Letters, 615, L117

Iyudin, A. F., Diehl, R., Bloemen, H., et al. 1994, Astronomy and Astrophysics, 284, L1

Jean, P., Gi, W., Marcowith, A., & Ferrière, K. 2009, Astronomy and Astrophysics, 508, 1099

Jean, P., Knödlseder, J., Gillard, W., et al. 2006, Astronomy and Astrophysics, 445, 579

Johnson III, W. N., Harnden Jr., F. R., & Haymes, R. C. 1972, The Astrophysical Journal, 172, L1

Johnson III, W. N., & Haymes, R. C. 1973, The Astrophysical Journal, 184, 103

Johnson III, W. N., Kinzer, R. L., Kurfess, J. D., et al. 1993, The Astrophysical Journal Supplement Series, 86, 693

Jung, G. V., Kurfess, D. J., Johnson III, W. N., et al. 1995, Astronomy and Astrophysics, 295, L23

Kalemci, E., Boggs, S. E., Milne, P. A., & Reynolds, S. P. 2006, The Astrophysical Journal Letters, 640, L55

Kasuya, S., & Takahashi, F. 2005, Physical Review D, 72, 085015

Khalil, S., & Seto, O. 2008, Journal of Cosmology and Astroparticle Physics, 2008, 024

Khoklov, A., Mueller, E., & Hoeflich, P. 1993, Astronomy and Astrophysics, 270, 223

Kierans, C. A., Boggs, S. E., Lowell, A., et al. 2014, in Proceedings of SPIE, Vol. 9144, Space Telescopes and Instrumentation 2014: Ultraviolet to Gamma Ray, 91443M

Kierans, C. A., Boggs, S. E., Chiu, J.-L., et al. 2016, in Gamma-Ray Astrophysics in Multi-Wavelength Perspective, Proceedings in 11th INTEGRAL Conference

Klein, O., & Nishina, Y. 1929, Zeitschrift für Physik, 52, 853

Klemperer, O., & Chadwick, J. 1934, Mathematical Proceedings of the Cambridge Philosophical Society, 30, 347

Knödlseder, J., von Ballmoos, P., Diehl, R., et al. 1996, in Proceedings of the SPIE, Vol. 2806, Gamma-Ray and Cosmic-Ray Detectors, Techniques, and Missions

Knödlseder, J., Jean, P., Lonjou, V., et al. 2005, Astronomy and Astrophysics, 441, 513

Kretschmer, K., Diehl, R., Krause, M., et al. 2013, Astronomy and Astrophysics, 559, A99

Kruegel, E. 2002, The physics of interstellar dust, ed. M. Birkinshaw, M. Elvis, & J. Silk, Series in Astronomy and Astrophysics (CRC Press)

Kurfess, J. D., Johnson III, W. N., Kroeger, R. A., & Philips, B. F. 2000, AIP Conference Proceedings, 510, 789

Lei, F., Dean, A. J., & Hills, G. L. 1997, Space Science Reviews, 82, 309

Leventhal, M. 1973, The Astrophysical Journal Letters, 183, L147

Leventhal, M. 1991, Advances in Space Research, 11, 157

Leventhal, M., MacCallum, C. J., & Stang, P. D. 1978, The Astrophysical Journal, 225, L11

Li, H., & Liang, E. P. 1996, The Astrophysical Journal, 458, 514

Ling, J. C. 1975, Journal of Geophysical Research, 80, 3241

Ling, J. C., Mahoney, W. A., Willett, J. B., & Jacobson, A. S. 1977, Journal of Geophysical Research, 82, 1463

Lingenfelter, R. E., Higdon, J. C., & Rothschild, R. E. 2009, Physical Review Letters, 103

Lingenfelter, R. E., & Ramaty, R. 1982, in AIP Conference Proceedings, Vol. 83, 148–159

Lingenfelter, R. E., Ramaty, R., & Leiter, D. 1981, in International Cosmic Ray Conference, Vol. 1, 112–115

Lithwick, Y., & Goldreich, P. 2001, The Astrophysical Journal, 562, 279

Liu, Q. Z., van Paradijs, J., & van den Heuvel, E. P. J. 2006, Astronomy and Astrophysics, 455, 1165

—. 2007, Astronomy and Astrophysics, 469, 807

Lorimer, D. R. 2008, Living Reviews in Relativity, 11, 8

Lowell, A. W. 2017, PhD thesis, University of California, Berkeley

Lowell, A. W., Boggs, S. E., Chiu, J. L., et al. 2016, Proceedings of SPIE, 9915

Lowell, A. W., Boggs, S. E., Chiu, C. L., et al. 2017a, The Astrophysical Journal, 848, 120

—. 2017b, The Astrophysical Journal, 848, 119

Luke, P. N., Cork, C. P., Mdden, N. W., Rossington, C. S., & Wesela, M. F. 1992, IEEE Transactions on Nuclear Science, 39, 590

Magkotsios, G., Timmes, F. X., Hungerford, A. L., et al. 2010, The Astrophysical Journal Supplement Series, 191, 66

Mahoney, W. A., Ling, J. C., Jacobson, A. S., & Tapphorn, R. M. 1980, Nuclear Instruments and Methods, 178, 363

Mahoney, W. A., Ling, J. C., & Wheaton, W. A. 1994, The Astrophysical Journal Supplement Series, 92, 387

Mahoney, W. A., Ling, J. C., Wheaton, W. A., & Jacobson, A. S. 1984, The Astrophysical Journal, 286, 578

Manchester, R. N., Hobbs, G. B., Teoh, A., & Hobbs, M. 2005, The Astronomical Journal, 129, 1993

Mannucci, F., Valle, M. D., Panagia, N., et al. 2005, Astronomy and Astrophysics, 433, 807

Martin, P., Knödlseder, J., & Meynet, G. 2008, New Astronomy Reviews, 52, 445

Martin, P., Strong, A. W., Jean, P., Alexis, A., & Diehl, R. 2012, Astronomy and Astrophysics, 543, A3

Mazzali, P. A., Chugai, N., Turatto, M., et al. 1997, Monthly Notices of the Royal Astronomical Society, 284, 151

McConnell, M. L. 2017, New Astronomy Reviews, 76, 1

McKee, C. F., & Ostriker, J. P. 1977, The Astrophysical Journal, 218, 148

Millan, R. M., Lin, R. P., Smith, D. M., Lorentzen, K. R., & McCarthy, M. P. 2002, Geophysical Research Letters, 29, 47

Millan, R. M., & Thorne, R. M. 2007, Journal of Atmospheric and Solar-Terrestrial Physics, 69, 362

Milne, P. A., Kurfess, J. D., Kinzer, R. L., & Leising, M. D. 2001, AIP Conference Proceedings, 587, 11

Milne, P. A., The, L.-S., & Leising, M. D. 1999, The Astrophysical Journal Supplement Series, 124, 503

Minkowski, R. 1959, in IAU Symposium, Vol. 9, Symposium - International Astronomical Union, 315–322

Mirabel, I. F., & Rodríguez, L. F. 1999, Annual Review of Astronomy and Astrophysics, 37, 409

Mirabel, I. F., Rodríguez, L. F., Cordier, B., Paul, J., & Lebrun, F. 1992, Nature, 358, 215

Mohorovicic, S. 1934, Astronomische Nachrichten, 253, 93

Moskalenko, I. V., & Strong, A. W. 1998, The Astrophysical Journal, 493, 694

Motz, J. W., Olsen, H. A., & Koch, H. W. 1969, Reviews of Modern Physics, 41, 581

Murphy, R. J., Dermer, C. D., & Ramaty, R. 1987, Astrophysical Journal Supplement Series, 63, 721

Nagataki, S., M.Hasimoto, Sato, K., Yamada, S., & Mochizuki, Y. S. 1998, The Astrophysical Journal, 492, L45

Navarro, J. F., Frank, C. S., & White, S. D. M. 1997, The Astrophysical Journal, 490, 493

Oaknin, D. H., & Zhitnitsky, A. R. 2005, Physical Review Letters, 94, 101301

O'Neill, T. J., Bhattacharya, D., Polsen, M., et al. 2003, IEEE Transactions on Nuclear Science, 50, 251

Ore, A., & Powell, J. L. 1949, Physical Review, 75, 1696

Paredes, J. M. 2005, Chinese Journal of Astronomy and Astrophysics, 5, 121

Patrignani, C., & Particle Data Group. 2016, Chinese Physics C, 40, 100001

Perkins, D. H. 2003, Particle Astrophysics (Oxford University Press)

Perrin, F. 1933, Comptes rendus de l'Académie des Sciences, 197, 1100

Petrosian, V., & Bykov, A. M. 2008, Clusters of Galaxies, ed. J. Kaastra (Springer)

Picciotto, C., & Pospelov, M. 2005, Physical Letters B, 605, 15

Picone, J. M., Hedin, A. E., Drob, D. P., & Aikin, A. C. 2002, Journal of Geophysical Research: Space Physics, 107, SIA 15

Plüschke, S., Diehl, R., Schönfelder, V., et al. 2001, Exploring the gamma-ray universe. Proceedings of the Fourth INTEGRAL Workshop

Pollack, J. B., & Fazio, G. G. 1963, Physical Review, 131, 2684

Porter, T. A., Moskalenko, I. V., Strong, A. W., & Orlando, E. 2008, The Astrophysical Journal, 682, 400

Pospelov, M., & Ritz, A. 2007, Physics Letters B, 651, 208

Prantzos, N. 2006, Astronomy and Astrophysics, 449, 869

Prantzos, N., Boehm, C., Bykov, A. M., et al. 2011, Reviews of Modern Physics, 83, 1001

Purcell, W. R., Grabelsky, D. A., Johnson III, W. N., et al. 1992, in NASA Goddard Space Flight Center, The Compton Observatory Science Workshop, ed. C. R. Shrader, N. Gehrels, & B. Dennis, Vol. 3137, 431–437

Purcell, W. R., Cheng, L.-X., Dixon, D. D., et al. 1997, The Astrophysical Journal, 491, 725

Ramaty, R., & Lingenfelter, R. E. 1977, The Astrophysical Journal Letters, 213, L5

Rees, M. J. 1988, Nature, 333, 523

Renaud, M., Vink, J., Decourchelle, A., et al. 2006, The Astrophysical Journal Letters, 647, L41

Reynolds, R. J., & Cox, D. P. 1992, The Astrophysical Journal Letters, 400, L33

Riegler, G. R., Ling, J. C., Mahoney, W. A., et al. 1981, The Astrophysical Journal Letters, 248, L13

Rodriguez, O. 1840, J Mathematiques Pures Appliquees, 5, 380

Rojdev, K., Atwell, W., Wikins, R., Gersey, B., & Badavi, F. F. 2009, in National Space and Missile Materials Symposium

Roques, J. P., & Jourdain, E. 2016, ArXiv e-prints. `https://arxiv.org/abs/1601.05289`

Sato, T. 2015, Public Library of Science One, 10, e0144679

Schönfelder, V. 2004, New Astronomy Reviews, 48, 193

Schönfelder, V., Aarts, H., Bennett, K., et al. 1993, The Astrophysical Journal Supplement Series, 86, 657

Segreto, A., Labanti, C., Bazzano, A., et al. 2003, Astronomy and Astrophysics, 411, L215

Seiradakis, J. H., & Wielebinski, R. 2004, The Astronomy and Astrophysics Review, 12, 239

Shafter, A. W. 2002, AIP Conference Proceedings, 637, 462

Share, G. H., Kinzer, R. L., Kurfess, J. D., et al. 1988, The Astrophysical Journal, 326, 717

Share, G. H., Leising, M. D., Messina, D. C., & Purcell, W. R. 1990, The Astrophysical Journal Letters, 358, L45

Siegert, T., & Diehl, R. 2016, in 14th International Symposium on Nuclei in the Cosmos, ed. S. Kubono, T. Kajino, S. Nishimura, T. Isobe, S. Nagataki, T. Shima, & Y. Takeda

Siegert, T., Diehl, R., Khachatryan, G., et al. 2016a, Astronomy and Astrophysics, 586, A84

Siegert, T., Diehl, R., Vincent, A. C., et al. 2016b, Astronomy and Astrophysics, 595, A25

Siegert, T., Diehl, R., Greiner, J., et al. 2016c, Nature, 531, 341

Sizun, P., Cassé, M., & Schanne, S. 2006, Physical Review D, 74, 063514

Skinner, G., Diehl, R., Zhang, X., Bouchet, L., & Jean, P. 2014, in 10th INTEGRAL Workshop: A Synergistic View of the High-Energy Sky

Skobeltzyn, D. V. 1985, The Early Stage of Cosmic Ray Particle Research, ed. Y. Sekido & H. Elliot (Springer Netherlands)

Sleator, C. C., Boggs, S. E., Chiu, J.-L., et al. 2016, in Gamma-Ray Astrophysics in Multi-Wavelength Perspective, Proceedings in 11th INTEGRAL Conference

Smith, D. M. 2004, New Astronomy Reviews, 48, 87

Smith, D. M., Hurford, G. J., & Boggs, S. E. 2004, New Astronomy Reviews, 48, 209

Smith, D. M., Leventhal, M., Cavallo, R., et al. 1996a, The Astrophysical Journal, 458, 576

—. 1996b, The Astrophysical Journal, 471, 783

Smither, R. K. 1982, Review of Scientific Instruments, 53, 131

Stecker, F. W. 1969, Astrophysics and Space Science, 3, 579

Strong, A. W., & Diehl, R. 1989, Maximum Entropy Image Processing in Gamma-Ray Astronomy, ed. V. Di Gesù, L. Scarsi, P. Crane, J. H. Friedman, S. Levialdi, & M. C. Maccarone (Springer US), 55–65

Strong, A. W., Moskalenko, I. V., & Ptuskin, V. S. 2007, Annual Review of Nuclear and Particle Science, 57, 285

Strong, A. W., Bennett, K., Cabeza-Orcel, P., et al. 1990, International Cosmic Ray Conference, 4, 154

Sturrock, P. A. 1971, The Astrophysical Journal, 164, 529

Su, M., Slatyer, T. R., & Finkbeiner, D. P. 2010, The Astrophysical Journal, 724, 1044

Sunyaev, R., Churazov, E., Gilfanov, M., et al. 1991, The Astrophysical Journal Letters, 383, L49

Takeda, S., Watanabe, S., Tanaka, T., et al. 2007, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 579, 859

The, L.-S., Clayton, D. D., Diehl, R., et al. 2006, Astronomy and Astrophysics, 450, 1037

Thielemann, F.-K., Nomoto, K., & Hashimoto, M. 1996, The Astrophysical Journal, 460, 408

Toma, K., Sakamoto, T., Zhang, B., et al. 2009, The Astrophysical Journal, 698, 1042

Tomsick, J. A., & the COSI team. 2016, GRB Coordinates Network, Circular Service, 19473

Totani, T. 2006, Publications of the Astronomical Society of Japan, 58, 965

Ueno, K., Mizumoto, T., Hattori, K., et al. 2012, Journal of Instrumentation, 7, 1088

Valinia, A., & Marshall, F. E. 1998, The Astrophysical Journal, 505, 134

Vallery, R. S., Zitzewitz, P. W., & Gidley, D. W. 2003, Physical Review Letters, 90, 203402

Vedrenne, G., Roques, J. P., Schönfelder, V., et al. 2003, Astronomy and Astrophysics, 411, L63

Verbunt, F. 1993, Annual Review of Astronomy and Astrophysics, 31, 93

Vincent, A. C., Martin, P., & Cline, J. M. 2012, Journal of Cosmology and Astroparticle Physics, 2012, 022

von Ballmoos, P., Diehl, R., & Schönfelder, V. 1989, Astronomy and Astrophysics Supplement Series, 221, 396

Voss, R., Martin, P., Diehl, R., et al. 2012, Astronomy and Astrophysics, 539, A66

Wang, W., Pun, C. S. J., & Cheng, K. S. 2006, Astronomy and Astrophysics, 446, 943

Wang, W., Harris, M. J., Diehl, R., et al. 2007, Astronomy and Astrophysics, 469, 1005

Watanabe, K., Leising, M. D., Share, G. H., & Kinzer, R. L. 2000, AIP Conference Proceedings, 510, 471

Watanabe, S., Tajima, H., Fukazawa, Y., et al. 2012, Proceedings of the SPIE, 8443

Weidenspointner, G., Varendorff, M., Kappadath, S. C., et al. 2000, AIP Conference Proceedings, 510, 467

Weidenspointner, G., Skinner, G., Jean, P., et al. 2008, Nature, 451, 159

Wietfeldt, F. E., Schima, F. J., Coursey, B. M., , & Hoppes, D. D. 1999, Physical Review C, 59, 528

Wilderman, S., Clinthorne, N. H., Fessler, J. A., & Rogers, W. L. 1998, in IEEE Nuclear Science Symposium and Medical Imaging Conference, Vol. 3, 1716–1720

Wilkinson, R. J., Vincent, A. C., Boehm, C., & McCabe, C. 2016, Physical Review D, 94, 103525

Woosley, S. E., & Weaver, T. A. 1980, The Astrophysical Journal, 238, 1017

—. 1994, The Astrophysical Journal, 423, 371

—. 1995, The Astrophysical Journal Supplement Series, 101, 181

Zhang, L., & Cheng, K. S. 1997, The Astrophysical Journal, 487, 370

Zoglauer, A. 2005, PhD thesis, Max-Planck-Institut für extraterrestrische Physik

Zoglauer, A., Andritschke, R., & Schopper, F. 2006, New Astronomy Reviews, 50, 629

Zoglauer, A., & Boggs, S. E. 2013, in American Astronomical Society, High Energy Astrohysics Division Meeting #13, 117.06

Zoglauer, A., Boggs, S. E., Galloway, M., et al. 2011, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 652, 568

Zoglauer, A., & Kanbach, G. 2003, Proceedings of SPIE, 4851

Zurek, W. H. 1985, The Astrophysical Journal, 289, 603

PART IV

APPENDIX

A

DISCUSSION OF EVENT SELECTIONS FOR POSITRON ANNIHILATION ANALYSIS

The first selection to be chosen is the size of the origin cut, which defines the boundaries of **SR** and **BR** in the CDS. We look at the significance of the 511 keV line from flight data as a function of the size of the origin cut. We restrict the Compton scatter angle to ϕ = 10–30° to reduce the dependence ϕ , but have kept fairly open event selections. Figure A.1 shows the number of counts within 506–516 keV in the CDS subtracted spectrum as a function of the origin cut radius. The significance, which is measured by $\sqrt{N}/N$, is low since we have the constraint in ϕ. From this analysis, we chose an origin cut of 16° around the GC since it gives a combination of a high significance and total number of counts. Increasing the origin cut will continue to increase the number of counts due to the extended disk emission. Choosing an origin cut of 16° does not mean that COSI has measured a source with that radius, instead, it is just the cut that gives the best statistics.

The Compton scatter angle range is the event selection that has the next largest effect. The minimum angle is chosen to be smallest angle for which there are events at 511 keV, since the small angle scatters are generally better reconstructed. Figure A.2 shows the Compton scatter angle range for 506–516 keV events from the flight data. For comparison, the scatter angle distribution for E_γ = 400–500 keV is normalized and shown in green. The minimum ϕ for this analysis is chosen to be 15°. Since lower energy γ-rays generally have larger Compton scatter angles, to be sensitive to the o-Ps continuum below 511 keV, it is best to keep the ϕ range large. Figure A.3 shows the number of counts within 506–516 keV for the CDS background subtracted spectrum as a function of the maximum Compton scatter angle. Again, the significance of the line is measured as $\sqrt{N}/N$. Since there is not a strong significance dependence on ϕ, we chose 15–55° as the Compton scatter angle range for this analysis.

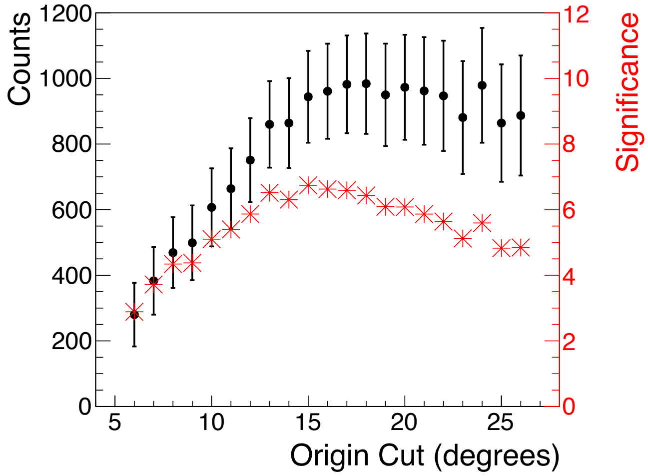


Figure A.1: Number of counts within 506–516 keV in the CDS background subtracted spectrum versus the origin cut radius. On the left axis is the significance of the line measured by $\sqrt{N}/N$, which reaches a max at ~15°.

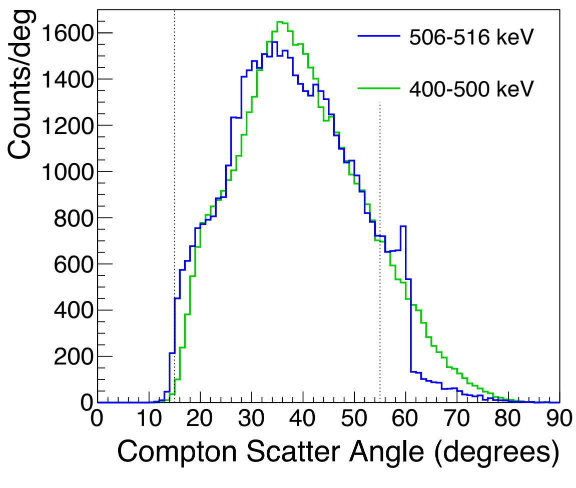


Figure A.2: The Compton scatter angle distribution for two different energy ranges, 506–516 keV in blue and 400–500 keV in green, from flight data. The chosen range of Compton scatter angles for this analysis is 15–55° and shown with the dashed vertical lines.

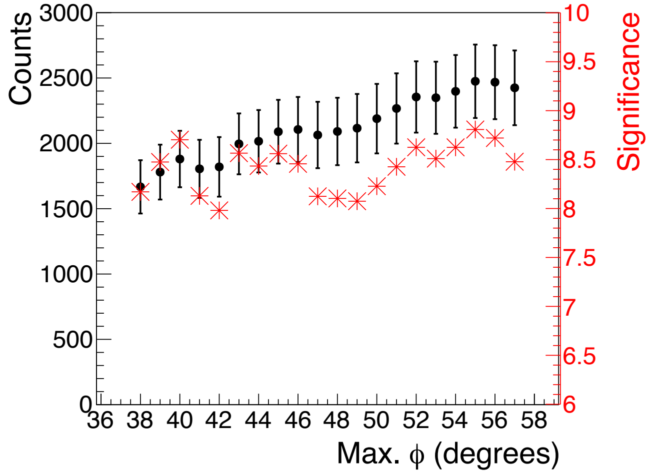
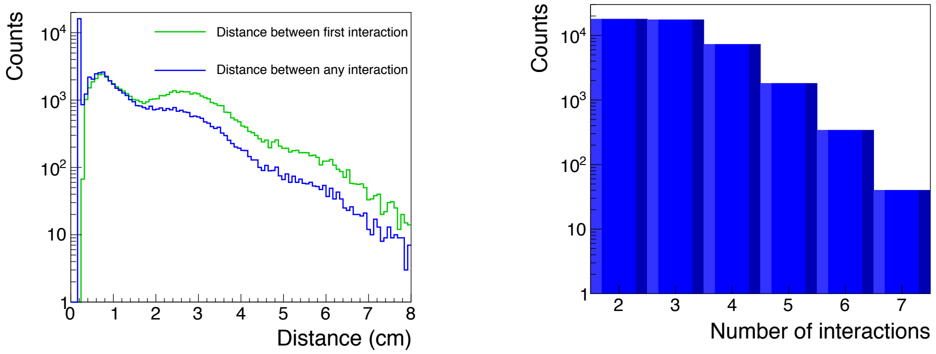


Figure A.3: The number of counts within 506–516 keV in the CDS subtracted background spectrum versus the maximum Compton scatter angle. The minimum ϕ is 15° for each of these measurements. The significance listed is $\sqrt{N}/N$.

The other event selections, namely the distance cuts and number of interactions, are less finely tuned. The first level of data-processing removes events with a distance cut of < 0.3 cm between the first two interactions and < 0.2 cm for any other interaction. This is done to reduce the number of events that are difficult to reconstruct. Figure A.4 (a) shows the distribution of the two distance measures: the distance in centimeters between the first and second interaction, and the distance between any other interaction. Cuts on the distance cut of 0.5 cm, and 0.3 cm, for the first and any interaction, respectively, were found to give a better significance, compared to the combinations 0.5 and 0.0 cm, 0.3 and 0.3 cm, and 0.3 and 0 cm, and 1.0 and 0.3 cm. The mean distance between first and second hit was 2.23 cm.

The sequence length of 2–7 interactions was chosen for this analysis; see Figure A.4 (b). The only other selection that was tested was 3–7 interactions since 2-site interactions are inherently more difficult to reconstruct. However, with an average sequence length of 2.9 scatters, the number of photons detected with just 2-site interactions was significant and therefore they were included.



(a) Distance between scatters.

(b) Compton sequence length.

Figure A.4: (a) The distance distributions for only E_γ = 506–516 keV. The distance cut for these selections is chosen to be 0.5 cm for the first two interactions and 0.3 cm for any interaction. The mean distance between the first two interactions is 2.2 cm. (b) The Compton sequence length, i.e., the number of interactions per event, is kept open for this analysis. The average sequence length is 2.9 interactions.