The Gamma-Ray Large Area Space Telescope Mission

Understanding Nature's Highest Energy Accelerators
and Using Them to Probe the Universe.

Edited by
Elliott D. Bloom and Linda Lee Evans

SLAC-R-497
March 1997

Stanford Linear Accelerator Center, Stanford University, Stanford, CA 94309
Prepared for the Department of Energy under contract number DE-AC03-76SF00515

Printed in the United States of America. Available from the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Road, Springfield, VA 22161.

Foreword

It has been some time between the time of the first Gamma-ray Large Area Space Telescope (GLAST) workshop, Towards a Next Generation High-Energy Gamma-Ray Telescope, in late August 1994, and the publication of a partial proceedings of that meeting. Since then there has been considerable progress in both the technical and project development of GLAST. From its origins at SLAC/Stanford in early 1992, the collaboration has currently grown to more than 20 institutions from France, Germany, Italy, Japan, and the United States, and is still growing. About half of these are astrophysics/astronomy institutions; the other half are high-energy physics institutions. About 100 astronomers, astrophysicists, and particle physicists are currently spending some fraction of their time on the GLAST R&D program.

The late publication date of this proceedings has resulted in some additions to the original content of the meeting. The first paper is actually a brochure prepared for NASA by Peter Michelson in early 1996. We have decided to present it in its entirety and in its original color format as one demonstration of how the NASA style is being integrated with the particle physics style in the GLAST project. Peter actually did give a talk on this subject at the meeting, but we deemed the brochure more up-to-date. Except for the appendix, the other papers in the proceedings were presented at the conference, and written up over the following two years. Some presentations were never written up. These speakers shall go nameless in these proceedings, though we greatly appreciate their contributions to the workshop. The appendix originates from the 1995 NASA Supporting Research and Technology Program (SR&T) and DOE detector R&D proposals, both of which were approved after a very competitive peer review in 1996: SR&T by NASA headquarters, and DOE-HEP after consideration by the Scientific Assessment Group for Experiments in Non-Accelerator Physics (SAGENAP) advisory panel. These approvals have led to a greatly expanded GLAST R&D in FY 1997 as compared to previous years.

At this writing, the DOE-HEP has yet to make a decision as to whether or not to fund the GLAST experiment. The SLAC management is enthusiastic about the GLAST program, and has been from the very beginning of the GLAST development. However, the HEPAP Subpanel on Planning for the Future of U.S. High Energy Physics, popularly known as the Gilman Committee, after chairman Fred Gilman, will have much influence in the decision of the DOE to fund GLAST. Indeed, the Gilman Committee is considering the entire effort of nonaccelerator physics (NAP) in the context of the future DOE-HEP program, and GLAST would be a major element of the NAP program. The results of the deliberation of the Gilman Committee will be publicly available in early 1998.

In contrast to the ongoing decision process within the DOE, the GLAST mission is well established as a high priority mission in the NASA program. Notes from NASA Headquarters tell

the story. From Alan Bunner, Science Program Director for Structure and Evolution of the Universe, June 1997:

"The May 1997 NASA Strategic Planning Retreat is now behind us. This exercise, organized by NASA's Office of Space Science (OSS), had as its main goal reaching a community consensus on a plan for budget initiatives in the next few years (augmentations and 'new starts' in the FY 2000–2004 period), as well as developing a long-term 'Roadmap' of needed space science missions over a period stretching out to about 2020. The meeting took place at Breckenridge, Colorado, and involved some 75 scientists, science communication experts, and NASA Administration officials.

"The new OSS strategic plan is designed around the theme of 'Origins, Evolution, and Destiny,' a theme that encompasses all of space science and serves astrophysics quite well. The basis for the plan is a set of broad fundamental questions and goals that were critiqued and edited by the retreat attendees. Wes Huntress, Associate Administrator for Space Science, put forward a straw man mission set that contained missions in four categories: (1) the existing (funded) program, (2) initiatives in the President's FY 1998 budget (and understood to be continued in the 'out years'), (3) additional, proposed new starts and augmentations up through 2004, (4) proposed missions beyond 2004. Category 2 initiatives are considered almost 'sold' to Congress. Category 3 includes new starts for the Gamma Ray Large Area Space Telescope (GLAST) in FY 2001, the High Throughput X-ray Spectroscopy mission (HTXS) in FY 2004, a Solar Terrestrial Probe series, an OSS-wide international mission of opportunity line beginning with funding for a US role in ESA's Far Infrared and Submillimetre Telescope (FIRST) beginning in FY 2000, and a Mission Operation and Data Acquisition(MO&DA) augmentation for an extended mission (beyond the nominal 5-year life) for the Advanced X-ray Astronomy Facility (AXAF).

"In short, we have an OSS strategic plan that is inclusive of most of the priorities set by the community and documented by the Structure and Evolution of the Universe Subcommittee (SEUS) and the Search for Origins Subcommittee and their respective Roadmap documents. Cosmology and black holes are included specifically in the strategic plan's fundamental questions and implicitly in the OSS mission statement. The Roadmap also emphasizes the past success and future promise of the Explorer program.

"Next steps in the realization of these plans for FY 2000–2004 include solidifying the technology funding that has been identified as necessary for the advent of FIRST, GLAST, and HTXS. Pre-Project Offices have been formed (at JPL for FIRST; at GSFC for GLAST and HTXS) to guide the current phase of these studies. All of us are hopeful that the space science community will rally behind this plan and press our friends at OSTP and OMB and in Congress for its implementation."

At the grass roots it is clear that GLAST science has great appeal in both the particle physics and astronomy/astrophysics communities. The challenge is to convince the HEP community that exciting particle physics can be done not only with accelerators, but with other tools as well.

Elliott Bloom
June 1997

Table of Contents

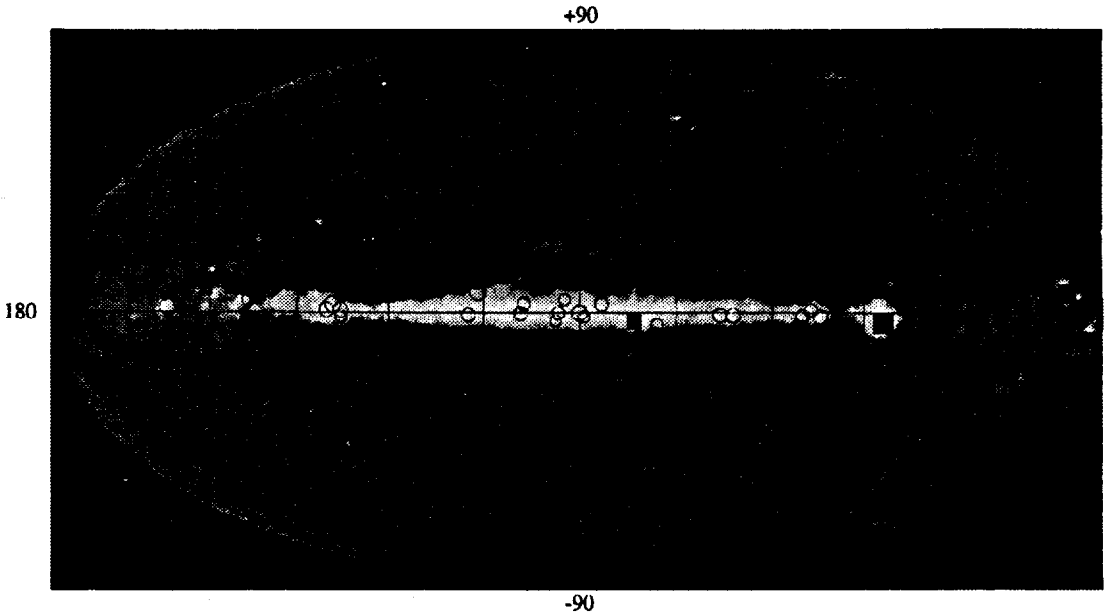
The Gamma-Ray Large Area Space Telescope Mission

Exploring the Astrophysics of Extremes with Nature's Highest Energy Accelerators

Peter Michaelson
Stanford University

The Gamma-ray Large Area Space Telescope (GLAST) Mission is under study for flight in the first decade of the next century. The GLAST mission is a next generation high-energy gamma-ray observatory designed for making observations of celestial gamma-ray sources in the energy band extending from 10 MeV to more than 100 GeV. This mission will (i) identify and study nature's high-energy particle accelerators through observations of active galactic nuclei, pulsars, stellar-mass black holes, supernova remnants, gamma-ray bursts, and the diffuse galactic and extragalactic high-energy radiation and (ii) use these sources to probe important physical parameters of the galaxy and the universe that are not readily measured with other observations. The mission's scientific objectives require an instrument with a large collecting area and imaging capability over a wide field of view. New detector technologies that offer significant improvements over existing hardware (a factor of 100 improvement in source sensitivity) are near flight readiness and will allow these requirements to be met within the cost constraints of an intermediate class mission.

EGRET All-Sky Survey above 100 MeV



◆ Active Galactic Nuclei
○ Unidentified EGRET sources
■ Pulsars
▲ LMC

2-97
8270A1

1. Introduction

High-energy gamma-ray astronomy is currently in a period of discovery and vigor unparalleled in its history. In particular, the Energetic Gamma-Ray Experiment Telescope (EGRET) on the Compton Gamma Ray Observatory (CGRO) has moved the field from a few detections of a small number of sources to multifaceted studies of several classes of galactic and extragalactic objects. The CGRO-EGRET discoveries of gamma-ray blazars, pulsars, high-energy gamma-ray bursts, and a large class of unidentified high-energy sources have given us a new view of the high-energy sky, while raising fundamental new questions about the origin, evolution, and destiny of high-energy sources.

High-energy gamma rays probe the most energetic phenomena occurring in nature. These phenomena typically involve dynamical nonthermal processes and include interactions of high energy particles (electrons, positrons, protons, pions, etc.) with matter, photons, and magnetic fields; high energy nuclear interactions; matter-antimatter annihilation; and possibly other fundamental elementary particle interactions. High-energy gamma rays are emitted over a wide range of angular scales from a diverse population of astrophysical sources—stellar mass objects, in particular neutron stars and black holes; the nuclei of active galaxies that likely contain massive black holes; interstellar gas in the galaxy that interacts with high-energy cosmic rays; the diffuse extragalactic background; supernovae that may be galactic sites of cosmic ray acceleration; and gamma-ray bursts. Even the Sun has been found to produce high-energy gamma rays during active periods. Many of the sources exhibit transient phenomena, ranging from the few-second timescale of gamma-ray bursts to active galactic nuclei (AGN) flares lasting days or more, and they often radiate the bulk of their power at gamma-ray energies.

3C 279: A cosmic high-energy accelerator laboratory

It is likely that most galaxies have nuclei and that at some stage in the galaxy's evolution, the nucleus is active in the sense that there is a substantial energy source in addition to the thermonuclear sources inside the constituent stars of the galaxy.

The nuclear activity of AGN is observed from radio frequencies ($\leq 10^8$ Hz) to high-energy γ-rays ($> 10^{25}$ Hz) and is directly manifest on scales as large as the lobes of giant radio sources (≤ 6 Mpc $\approx 2 \times 10^{25}$ cm) and as small as the distance travelled by light in the shortest observed variability timescales ($\approx 2 \times 10^{12}$ cm). That AGN emit significant power over such a broad range of frequencies and spatial scales shows that they are far from thermal equilibrium. The high efficiency with which they convert rest mass into radiant energy cannot be explained by nuclear and atomic processes. The currently most-favored theoretical model of the AGN central engine involves a massive, rotating black hole.

The quasar 3C 279 exhibits essentially all of the different manifestations of nuclear activity seen in AGN. (See figure at right, counterclockwise from top left.)

i) VLBI radio images of the nucleus at different epochs reveal an apparent transverse velocity of expansion of features in the radio brightness distribution that exceeds the velocity of light.

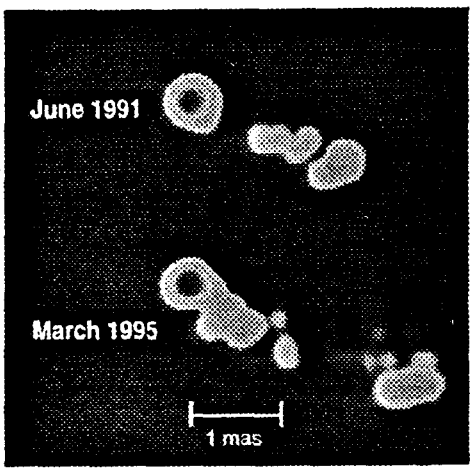
ii) This "super-luminal motion" can be explained if the radio emitting material is ejected from the central engine in a jet moving close to the speed of light with the axis of the jet nearly aligned to the line-of-sight.

iii) The first observation of 3C 279 by EGRET detected a dramatic high-energy γ-ray flare with a variability timescale of a few days. This result indicates that the source of the γ-ray emission is also relativistically beamed because, if the radiation were not beamed, the energy density of the γ-rays would be so great that γ-γ collisions would degrade the emission by producing electron-positron pairs before the high-energy radiation could escape from the source region.

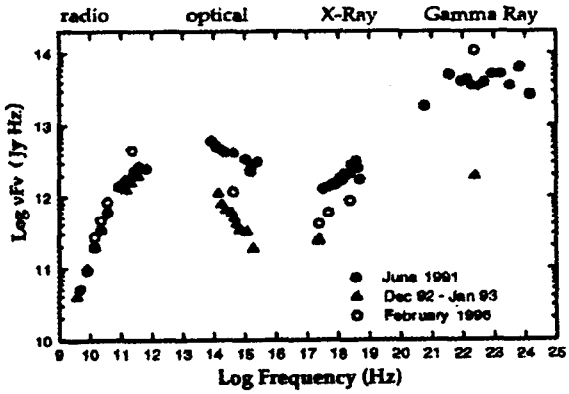
iv) In June 1991, 3C 279 was one of the brightest high-energy sources in the sky, more luminous than its neighbor 3C 273.

v) Multiwavelength observations reveal that the high-energy emission can dominate the emission from 3C 279.

In the future, simultaneous multiwavelength observations (radio to γ-rays) of flares will be key to understanding the physics of these cosmic accelerators.

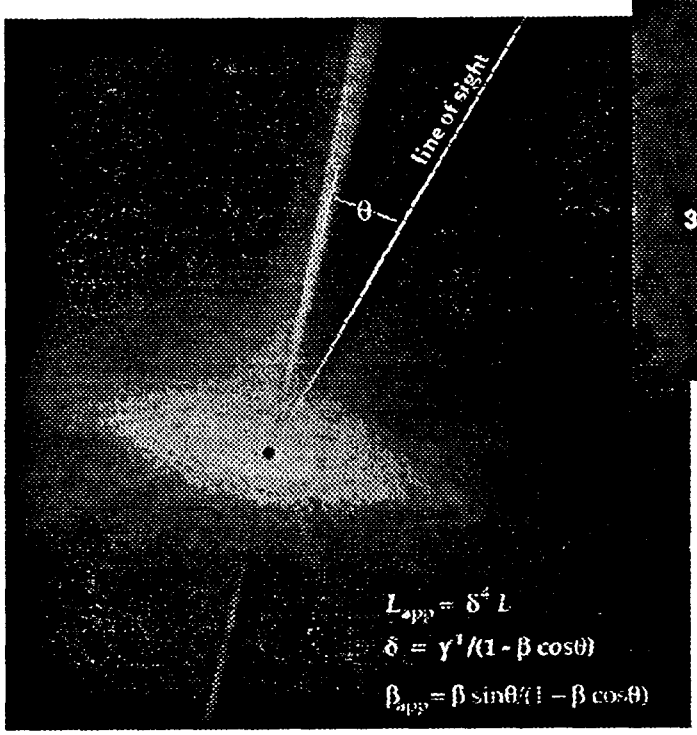


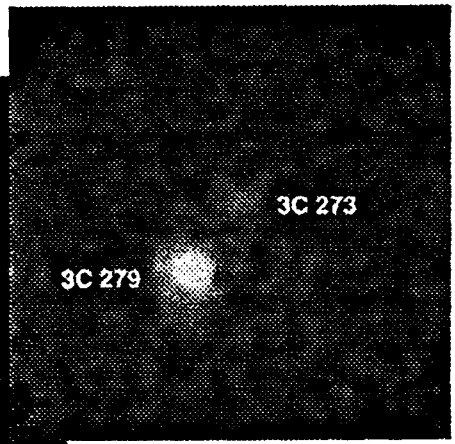
VLBI image of 3C 279 at 22.217 GHz

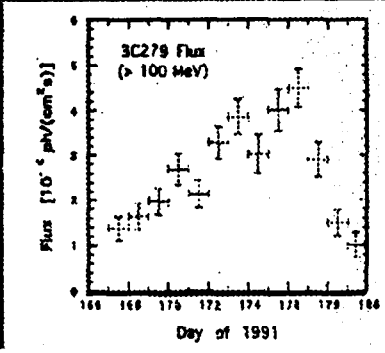


Multiwavelength Spectrum of 3C 279

EGRET Observation, June 1991







2-97
8270A7

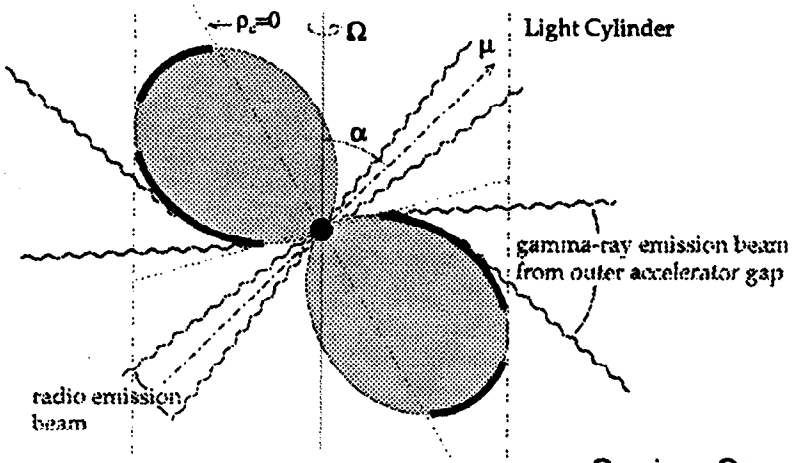
Geminga: it really is there

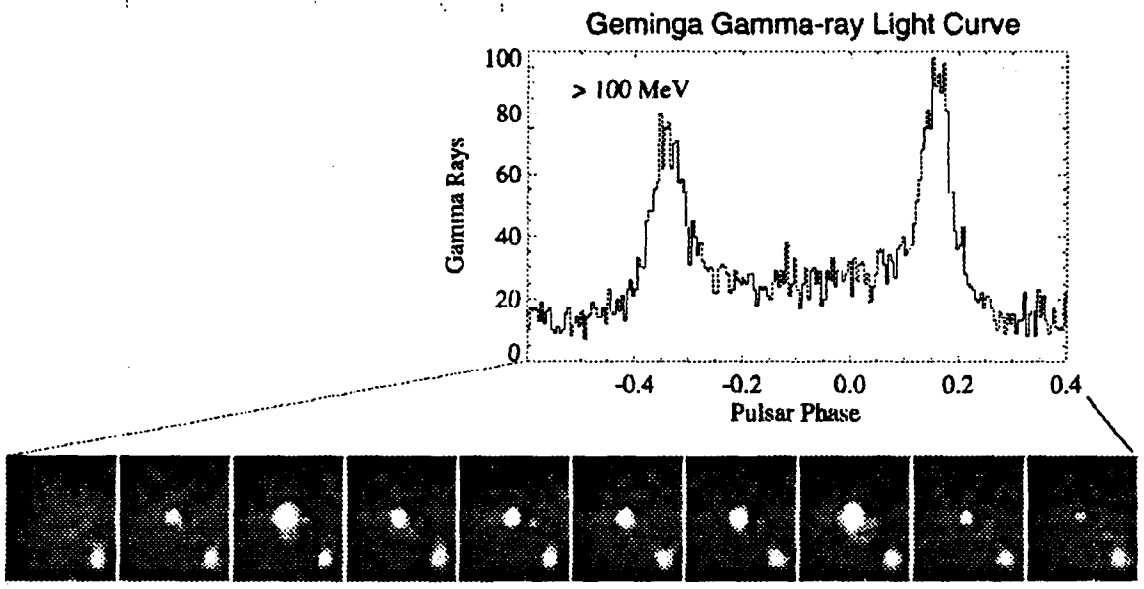
This bright γ-ray source was first discovered by the SAS-II satellite in 1975 and later observed in more detail by COS-B. Attempts to find a unique counterpart at lower energies were hindered by the relatively large γ-ray source error box, although it was long suspected that the X-ray source 1E0630+178, discovered by the Einstein Observatory, was also the source of the γ-rays. Early attempts to search for an unknown periodicity in the γ-ray emission, expected if the source is a pulsar, were also unsuccessful because of the paucity of γ-ray photons detected by SAS-II and COS-B. So enigmatic was the nature of this object that it was named, by the Italian astronomer Giovanni Bignami, "Geminga," a Northern Italian dialect word that means "it's not there."

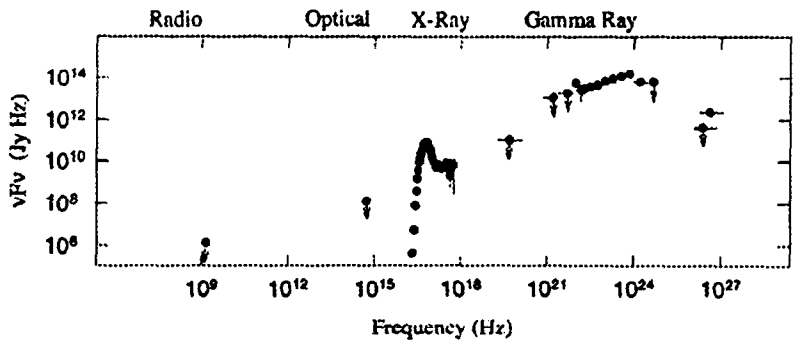
With the detection in 1992, by ROSAT, of 237 ms coherent X-ray pulsations from the Einstein source, and observations of high-energy γ-rays with EGRET, a sensitive search for coherent γ-ray pulsations from Geminga was made. A pulsed γ-ray signal was found at the same period, thus finally solving the mystery of the nature of Geminga: it is an isolated rotating neutron star. Interestingly, no radio (or optical) pulsations have ever been detected from this source. In this regard, Geminga is, so far, unique among γ-ray pulsars.

Because of the quality of the EGRET data and because observations of the source were made over several months, the rate at which the spin period of Geminga changes was also determined. This number and the spin period, combined with the simplest model of a rotating, magnetized neutron star, indicate that Geminga is about 300,000 years old, has a magnetic field (at its surface) of 1.6×10^{12} gauss, and loses energy at a rate of 3.5×10^{34} erg s^{-1}.

Theoretical models of γ-ray emission from pulsars have the γ-rays arising from particles accelerated in the pulsar magnetosphere, either near the magnetic polar cap surfaces or in "accelerator gaps" in the outer magnetosphere. A schematic diagram of one version of the outer gap model is shown at the top of the figure at left. In this model the lack of observed radio pulsations from Geminga is explained because the radio beam is narrower than the γ-ray beam and points in a different direction. This model predicts that there should be many more spin-powered pulsars that will only be seen because of their high-energy emission.







Spectrum of Pulsed Emission

2-97
8270A5

Observations of high-energy gamma rays provide unique astrophysical information:

- They identify sites where extreme particle acceleration processes occur and provide probes of the physical processes operating at those sites. The gamma-ray outburst of 3C 279 observed in 1991 by EGRET was unanticipated, showing that remarkably high luminosities can be attained along with rapid turn-on and turn-off of the source. The emission processes likely involve shock acceleration and particle cascades that produce the observed gamma-ray spectrum. In some systems, the acceleration may be electromagnetic: the Blandford-Znajek mechanism for producing a large voltage drop near a rotating black hole and the acceleration of charged particles in the magnetosphere of a rotating neutron star are two examples.

- High-energy gamma-ray observations have established what might be called "GeV source classes," that is, classes of sources preferentially visible at high energies. Geminga is an example of a pulsar that is far more conspicuous in gamma rays than at lower energies and there should be many others like it. All of the 'surprising' EGRET detections—blazars, gamma-ray bursts, solar flares, and the Geminga pulsar—are discoveries of GeV source classes, in the sense that the energetics of the GeV contribution to the total spectrum was unexpectedly high. In addition to the point sources, high-energy observations at high galactic latitudes have revealed an isotropic "diffuses" background.

- Characterization of these GeV source classes has only begun. For example, many of the sources near the Galactic plane are correlated with supernova remnants and OB associations. In both cases, it is possible that the majority of the unidentified EGRET point sources in these regions along the Galactic plane are gamma-ray pulsars. If this proves correct, these pulsars will provide an important new window into the history of recent massive star deaths in the solar neighborhood. Eventual source identifications may also reveal distinctly new source classes.

- The study of the galactic component of the diffuse emission and of emission from molecular clouds and supernovae remnants are important uses of high-energy gamma-ray observations. This provides information about the distribution of cosmic rays and matter in the galaxy which in turn have implications about the origin and propagation of cosmic rays.

- Space-based GeV observations, combined with ground-based TeV observations that are sensitive above 100 GeV, can provide measurements of absorption cutoffs in the spectra of extragalactic sources (for example, Mkn 421). These cutoffs may be useful as indicators of the intergalactic IR radiation field (high-energy photons incident on IR photons are attenuated through pair production). Detection of gamma rays and γ-γ spectral cutoffs in the 10–200 GeV range from high-redshift quasars can constrain the intensity and era of formation of the integrated extragalactic background light (EBL) in the near-UV, optical, and near-IR portions of the spectrum. The tremendous cosmological importance of measuring the EBL stems from the prospects of using it as a probe of galaxy formation and evolution. In particular, deep observations of GeV sources at redshifts $Z > 2$ may provide, for the first time, direct constraints on starlight at early cosmological epochs and thus probe the era during which galaxies and the EBL formed.

- X-ray and gamma-ray observations of predicted scattered "halos" of keV to tens of MeV radiation surrounding distant beamed high-energy gamma-ray sources can probe the intergalactic magnetic field.

Investigation of the rich variety of astronomical phenomena in the high-energy gamma-ray sky place strong demands on the performance of the next-generation high-energy gamma-ray telescope. The telescope must have angular resolution that will meet the requirement to identify point sources with objects at other wavelengths, yet have a wide field-of-view that will permit the study of sources that exhibit extreme intensity variations on timescales from seconds to months, or longer. The telescope must also have sufficient sensitivity to detect a large sample of sources and determine their energy spectra.

2. Key Scientific Questions

The Gamma-Ray Large Area Space Telescope (GLAST) Mission study team reviewed the current state of knowledge in high-energy gamma-ray astrophysics and formulated a key set of questions that remain unanswered or that have been raised by recent observations. Emphasis was placed on the connections of high-energy gamma-ray astronomy to all of astrophysics.The key questions include the following:

- How do active galactic nuclei (AGNs) form and evolve?
- What is the nature of jets emanating from AGNs and Galactic black holes, and how are the particles in the jets accelerated? How are these structures connected with similar structures seen at smaller scales?
- At what energies are the gamma-ray spectra of AGNs cutoff? Are high energy spectral cutoffs due to source-intrinsic absorption effects or to absorption by extragalactic background light? What is the redshift dependence of these effects? Is there a class of AGNs that can be used as high-energy "standard candles" in the sense that their intrinsic spectral shapes are similar?
- What is the origin of the isotropic "diffuse" gamma-ray background?
- What are the sites and mechanisms of cosmic-ray acceleration?
- How do rotation-powered pulsars generate high-energy gamma rays, and what is the relation of this radiation to emission in lower-energy bands?
- What is the rate of supernovae in the Galaxy, and where are the unobserved supernovae of the past several hundred years?
- What are gamma-ray bursts, and how do they generate high-energy radiation?
- What are the unidentified high-energy gamma-ray sources?

3. Scientific Objectives

The objectives of a next-generation high-energy gamma-ray mission follow from consideration of the measurements necessary to answer, or at least strongly impact, the key questions posed above. The scientific objectives listed below, organized by source class, can be met with a realizable, intermediate-class, high-energy gamma-ray mission such as GLAST. The important characteristics of the GLAST mission are summarized in the next section.

Active Galactic Nuclei

- Determine the mechanisms of AGN jet formation, particle acceleration, radiation, and flaring by studying gamma-ray emission from all known blazars (and possibly other AGN classes) and correlating these observations with those at other wavelengths.
- Increase the sample of high-energy sources detected, including objects at large distance, by more than an order of magnitude.
- Measure high-energy spectral turnovers in a large sample of sources distributed over a large range of redshifts.
- Determine if the average spectrum is consistent with that of the isotropic component of the high-energy gamma-ray background.

Unidentified EGRET Sources

Determine the type of object(s) and the mechanisms for gamma-ray emission from the unidentified EGRET sources by measuring precise positions of these sources, searching the gamma-ray data directly for periodicities, and determining the time variations of the emission in coordination with observations in other wavelength bands.

Isotropic Background Radiation

Determine if the high-energy background is resolvable into point sources or if there is a truly diffuse component, by a deep survey of high latitude fields.

Gamma-Ray Bursts

Provide constraints on physical mechanisms for gamma-ray bursts by detecting high-energy radiation from 50–150 bursts per year and study the GeV : keV-MeV emission ratio as a function of time; image burst positions to a few arcminutes or better, allowing deep "real-time" multiwave-length observations.

Endpoints of Stellar Evolution (Supernovae, Neutron Stars, and Black Holes)

- Provide direct evidence of proton cosmic-ray acceleration in supernova remnants by gamma-ray mapping and energy spectral measurements.
- Distinguish between models for high-energy gamma-ray emission from pulsars by measuring detailed phase-resolved spectra, detecting many more radio pulsars at gamma-ray energies, and identifying other radio-quiet pulsars like Geminga.
- Search for emission from millisecond pulsars; search for stellar-size black holes that produce high-energy particle beams analogous to those seen in AGN.

Molecular Clouds, Normal Galaxies and Clusters

- Probe cosmic-ray distribution in dense molecular cloud and in nearby galaxies (LMC, SMC, M31) by gamma-ray mapping and measuring the spectra of diffuse emission from these objects.
- Search for extended emission from possible cold, dark, matter clouds in the Galaxy and from galaxy clusters as a signature of unusual concentrations of unseen gas or cosmic rays.

4. Mission Description

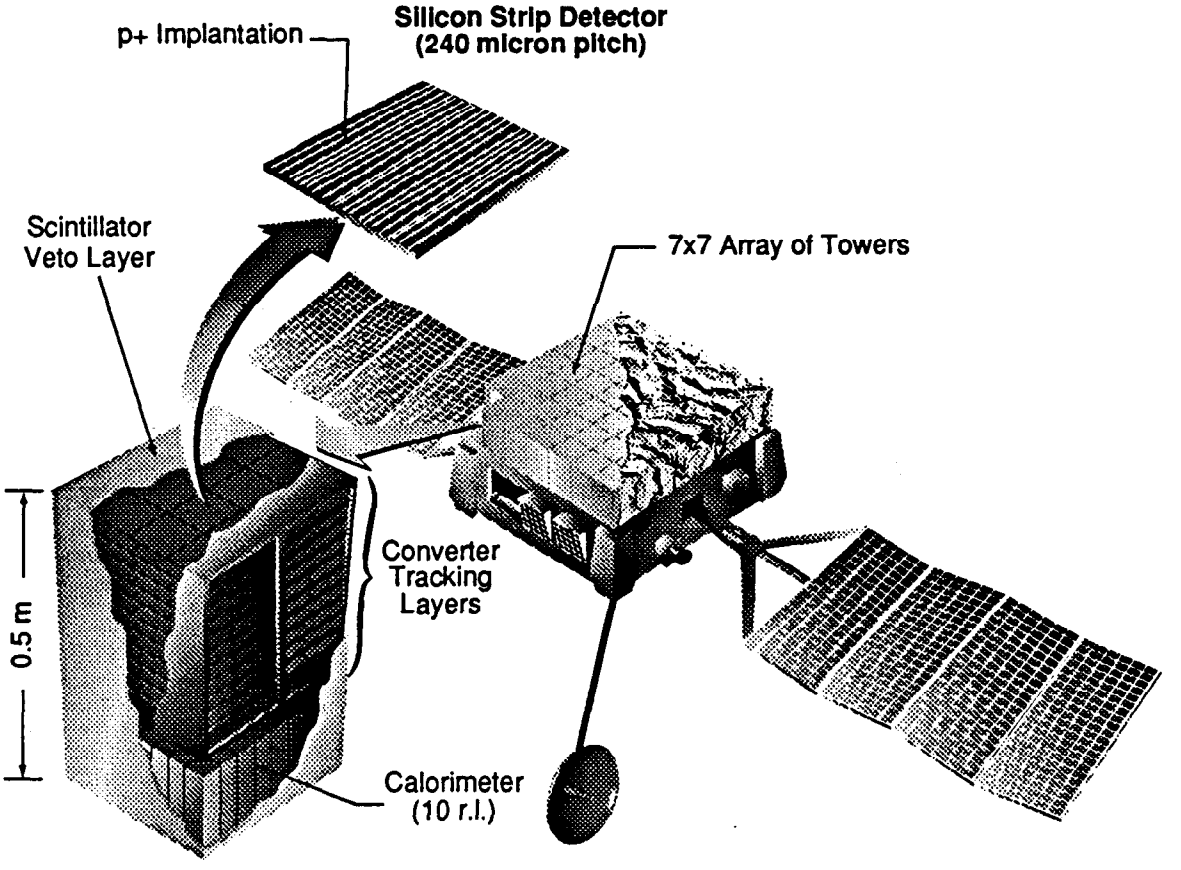
The primary instrument required on a high-energy gamma-ray mission is an imaging wide field-of-view telescope that covers the energy range from approximately 10 MeV to more than 100 GeV. This telescope identifies incident gamma rays by recording the characteristic track signature that results from pair conversion in the presence of a nucleus. The telescope consists of many thin layers of high-Z pair production material (metal foils) interleaved with position sensitive, charged-particle detectors. This gamma-ray converter and track imaging system is followed by an energy-measuring calorimeter system. Measurement of the energy and direction of the resulting electron and positron provide information about the energy and direction of the incident photon. Finally, the telescope requires a very efficient anticoincidence system for rejecting the much higher flux of background particles, and an on-board trigger and data-acquisition system. Modern particle tracking detectors and sophisticated on-board processing will allow the required major advance in observational capability needed to achieve the science goals outlined above, within the cost constraints of an intermediate class (Delta II launch) astrophysics mission.

Table 1: Characteristics of High-Energy Gamma-Ray Mission

Primary Instrument: Imaging Pair Conversion Telescope	
Energy Range:	10 MeV to >100 GeV
Energy Resolution:	10%
Effective Area:	> 8,000 cm^2 (above 1 GeV)
Single Photon Angular Resolution:	< 2.5° x (100 MeV/E) (10 MeV - 3 GeV)
(68% containment angle)	< 0.10° (E > 10 GeV)
Field of View:	> 1.5 sr
Point Source Sensitivity:	2 x 10^{-9} ph cm^{-2} s^{-1}
Source Location Determination:	30 arcsec - 5 arcmin
Mass:	3,000 kg
Power:	600 W
Telemetry:	100 kbps
Mission Requirements	
Mission life:	> 2 years
Orbit:	low inclination
Spacecraft pointing:	10 arcsec knowledge; 2° accuracy
Operating modes:	all-sky survey mode; pointed observation mode: any direction at any time

A mission concept that the study team has examined intensively is the Gamma-ray Large Area Space Telescope (GLAST) Mission, selected as a NASA New Mission Concept in Astrophysics. Table 1 (page 13) summarizes the important characteristics of the baseline GLAST Mission.

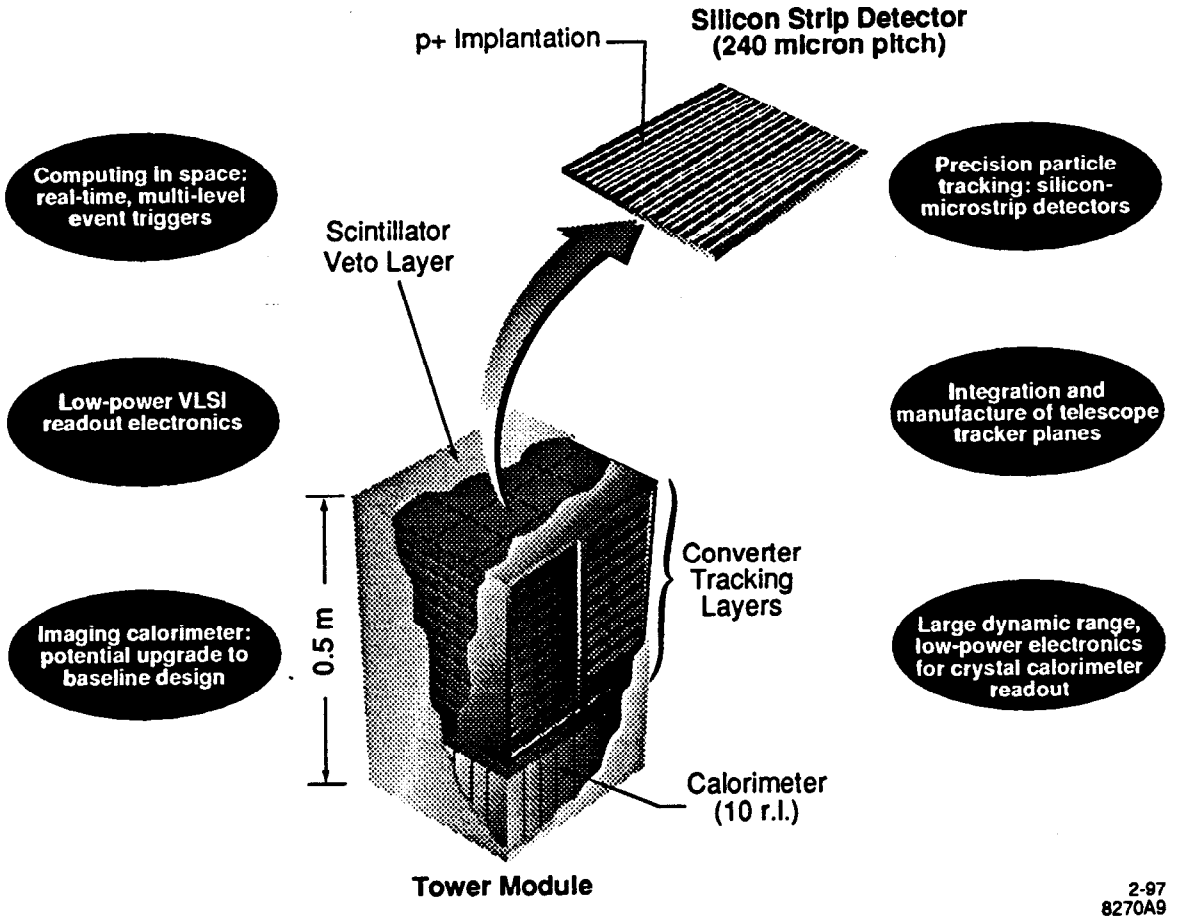
The GLAST instrument design is modular. The principal elements of the telescope are a segmented charged particle anticoincidence shield, a gamma-ray tracker /converter, a calorimeter, and an on-board trigger and data acquisition system. Elements of all of these are present in each GLAST tower module as shown in the figure below. The modular design of GLAST has the advantages associated with redundancy and avoids many of the dead-time and data-rate problems associated with more monolithic designs.



The Gamma-Ray Large Area Space Telescope. The total area of the telescope, made of 49 tower modules, is 2.8 sq. m. Each tower, with a frontal area of 24 cm x 24 cm, has a charged particle veto layer, followed by 12 tracker/converter layers, and a 10 radiation length calorimeter. Each of the first 10 tracker/converter layers has 0.05 r.l. of high Z converter, followed by 2 planes of (x,y) silicon strip detectors. These are followed by 2 tracking layers. The baseline calorimeter is an array of 3 cm x 3 cm x 10 r.l. CsI (T1) crystals.

Developments in using semiconductors for particle detection over the past decade are the main technical stimulus for GLAST. In particular, the development of large area silicon strip detectors for use in particle tracking has resulted in working devices now in widespread use in major high-energy particle accelerator experiments. Much recent innovation was brought about because of the detector challenges that were presented by the Superconducting Super Collider (SSC) background. Detailed Monte Carlo simulations show that GLAST can reach the goal of having residual background rates less than 1% of the isotropic diffuse gamma-ray flux.

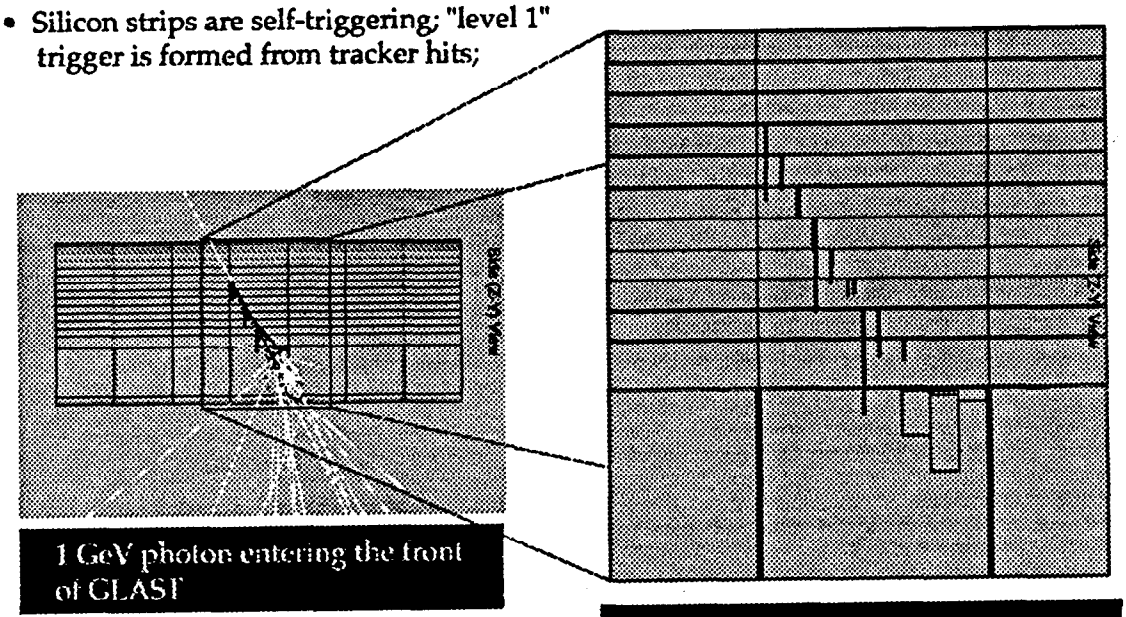
In the GLAST design, each plane in the tracker/converter has two sets of strip detectors that can accurately measure the charged particle tracks in two orthogonal dimensions. These tracks can be used to identify gamma-ray interactions, because the showers resulting from gamma-ray conversions have different properties than those caused by high-energy hadrons (protons, neutrons, etc.). By observing the pattern of charged particle "hits" in the silicon strip tracker and the energy-deposition pattern in the calorimeter, events caused by gamma rays that enter through the front of the instrument can be distinguished from the much higher fluxes of cosmic rays and trapped radiation incident on the instrument in orbit. Also, earth-albedo gamma rays are easily identified and eliminated by their directional signature.



High-energy γ-ray telescope technology requirements

Real-time, multilevel event triggers

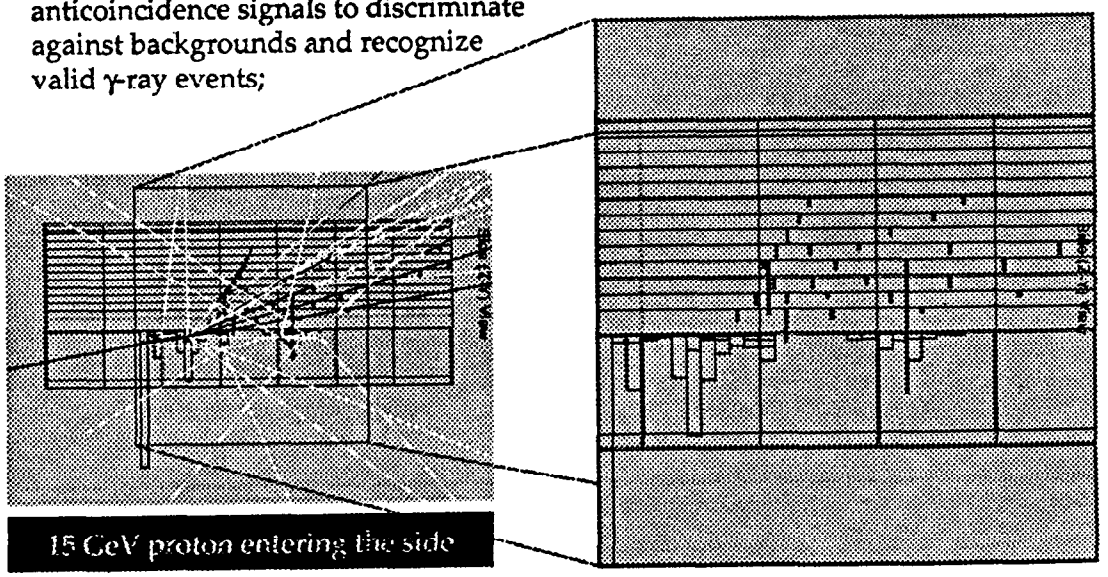
- Silicon strips are self-triggering; "level 1" trigger is formed from tracker hits;



1 GeV photon entering the front of GLAST

- "level 2" software trigger uses hit pattern, calorimeter signals, and anticoincidence signals to discriminate against backgrounds and recognize valid γ-ray events;

Simulated hit patterns and calorimeter energy deposition



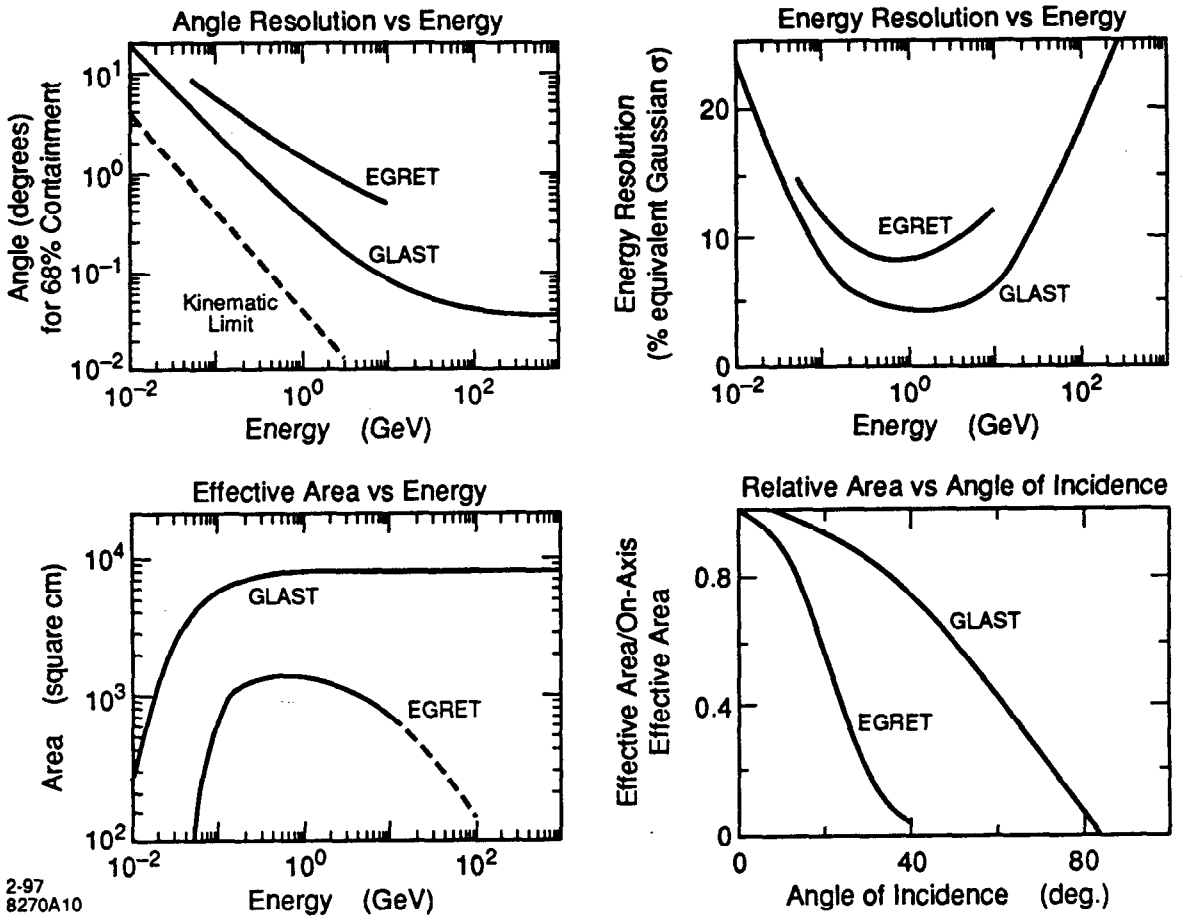
15 GeV proton entering the side

- $\geq$ 15 MIPS of on-board processor power is required

2-97
8270A4

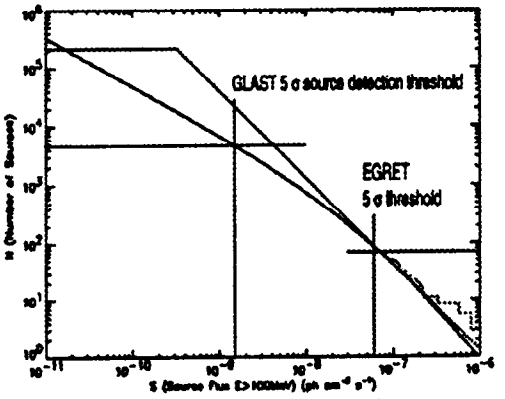
5. Comparison of Capabilities with Egret

Shown below is a comparison of the GLAST and EGRET effective areas (gamma-ray detection efficiency times area), angular resolutions, and energy resolutions versus energy. Also shown are the relative effective areas (@ 1 GeV) versus the angle of incidence. Inefficiencies due to analysis cuts for background rejection have been included.

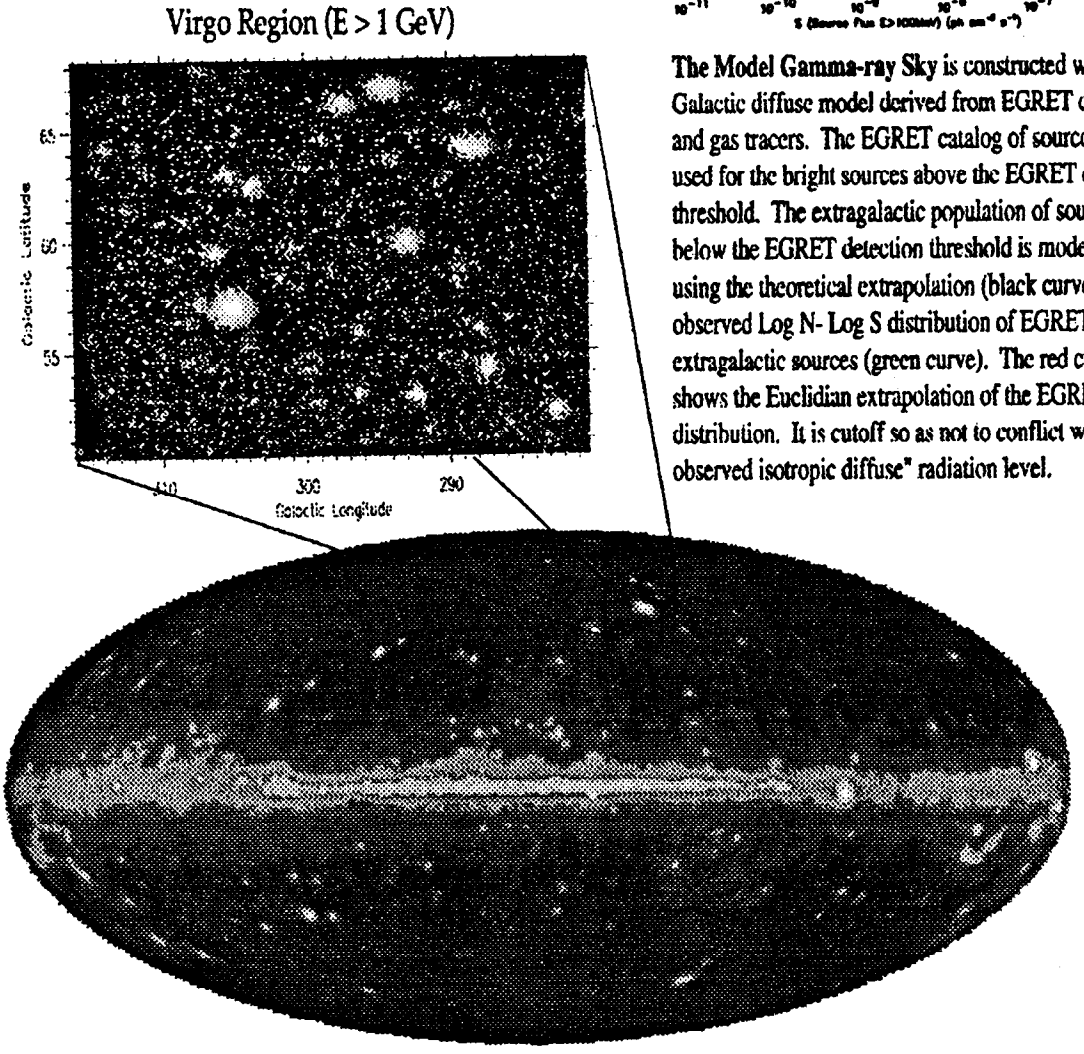


The figures on this and following page show a comparison of simulations of all-sky survey data obtained with the GLAST and EGRET instruments.

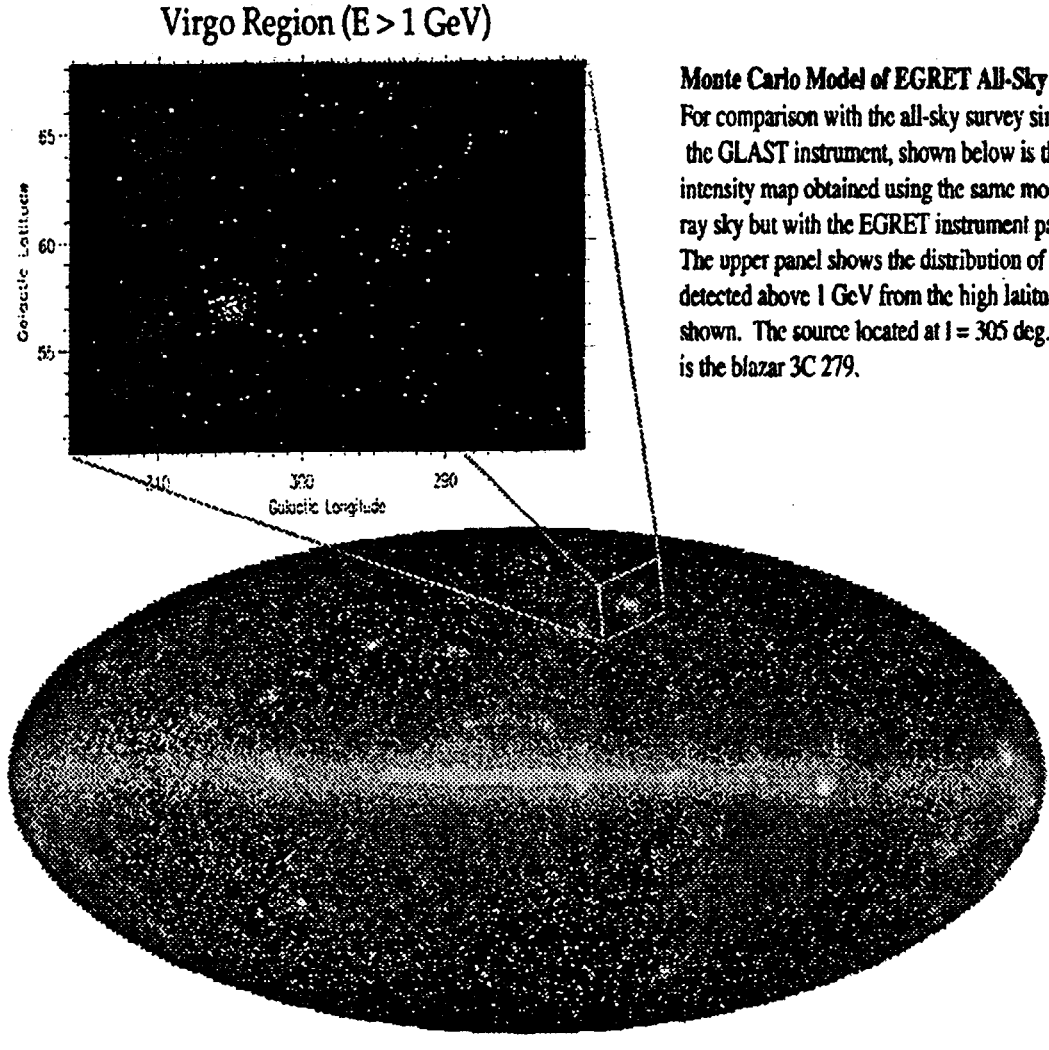
One-year All-Sky Survey Simulation. The lower panel shows an all-sky intensity map above 100 MeV obtained from a Monte Carlo simulation of a one-year all-sky survey with a next-generation high-energy gamma-ray telescope with the GLAST instrument characteristics. The panel immediately below shows a simulated intensity map above 1 GeV for a high latitude region.



The Model Gamma-ray Sky is constructed with a Galactic diffuse model derived from EGRET data and gas tracers. The EGRET catalog of sources is used for the bright sources above the EGRET detection threshold. The extragalactic population of sources below the EGRET detection threshold is modelled using the theoretical extrapolation (black curve) of the observed Log N- Log S distribution of EGRET extragalactic sources (green curve). The red curve shows the Euclidian extrapolation of the EGRET distribution. It is cutoff so as not to conflict with the observed isotropic diffuse" radiation level.



2-97
8270A3



Monte Carlo Model of EGRET All-Sky Survey. For comparison with the all-sky survey simulation of the GLAST instrument, shown below is the simulated intensity map obtained using the same model gamma-ray sky but with the EGRET instrument parameters. The upper panel shows the distribution of photons detected above 1 GeV from the high latitude region shown. The source located at l = 305 deg., b = 57 deg. is the blazar 3C 279.

2-97
8270A2

The Study Team

Y. C. Lin, P. F. Michelson, P. L. Nolan	Hansen Experimental Physics Laboratory, Stanford University
W. B. Atwood, E. D. Bloom, G. L. Godfrey	Stanford Linear Accelerator Center, Stanford University
D. Bertsch, N. Gehrels, R. C. Hartman, S. D. Hunter, J. Norris, J. Ormes, R. Streitmatter, D. Thompson, J. Tueller	NASA/Goddard Space Flight Center
E. Grove, P. Hertz, W. N. Johnson, M. Lovellette, G. Share, M. Wolff, K. Wood, D. Yentis	Naval Research Laboratory
R. Johnson	University of California, Santa Cruz
C. Covault, R. Ong, M. Oreglia	University of Chicago
J. Mattox	University of Maryland
T. H. Burnett	University of Washington
D. L. Chenette, G. Nakano	Lockheed Research Laboratory, Palo Alto, CA
L. Cominsky	Sonoma State University
H. Mayer-Hasselwander	Max Planck Institut für Extraterrestriche Physik, Garching, Germany
G. Barbiellini	Instituto Nazionale di Fisica Nucleare, Trieste, Italy
A. Colavita	International Center for Theoretical Physics, Trieste, Italy
A. Morselli	University of Rome, Rome, Italy
T. Kamae	University of Tokyo, Tokyo, Japan
K. Kasahara	Kanagawa University, Tokyo, Japan

Main Presentations

Included in this Section:

Gamma Large Area Space Telescope (GLAST)*

W. B. Atwood
Stanford Linear Accelerator Center, Stanford University

Abstract

The recent discoveries and excitement generated by the space satellite experiment EGRET have prompted an investigation into modern detector technologies for a next generation, space based, gamma ray observatory. The GLAST proposal is based on silicon strip detectors as the "technology of choice" for space application: no consumables, no gas volume, robust (versus fragile), long lived, and self triggerable. The GLAST detector has two main components: a tracking module preceding a calorimeter. The tracking module has planes of crossed (x,y) strip silicon detectors in close proximity to a thin radiator to measure the coordinates of converted γ-rays. The gap between the layers (~3 cm) provides a lever arm for track fitting resulting in an angular resolution per γ of $<0.1^{o}$ at high energy. The status of this R&D effort is discussed including details on triggering the instrument, the organization of the detector electronics and readout, and work on computer simulations of this instrument.

1.0 Introduction

The history of γ-ray astronomy has been covered in Peter Michelson's contribution to this workshop[1]. It suffices to say that the tantalizing glimpses of the universe in the >10 MeV region provide more than ample motivation to push for a major new instrument in this wavelength band. Many of the early devices were based on gas detector technologies and presented numerous problems in their adaptation to space operation. Our approach from the outset was to abandon these technologies because of the inherent difficulties.

Over the past decade alternatives have evolved for tracking charged particles . The advances were to a certain extent spin-offs from the now discontinued SSC project. Radiation hardness, cost, and power per channel were all central issues for the large detectors envisioned and our approach is to take these now well developed technologies, in particular silicon strip detectors, and adapt them to the design of a γ-ray pair conversion telescope.

Figure 1 is a schematic of the GLAST design. At the smallest level are silicon strip detectors (SSD's) measuring 6 cm on a side and 300-500 μm thick. The dimensions of the SSDs were chosen to match what is currently available from industry. The readout pitch was chosen to be ~240 μm, mainly to minimize the number of channels while not compromising the performance. Four

detectors are connected end-to-end, thus forming 24-cm-long readout channels. (The length needs to be limited to control the detector capacitance as seen by the front end electronics.) Two square planes of the 4-detector-long "ladders" together with a thin radiator mounted directly in front form the basic unit (a "tray") of the pair telescope.

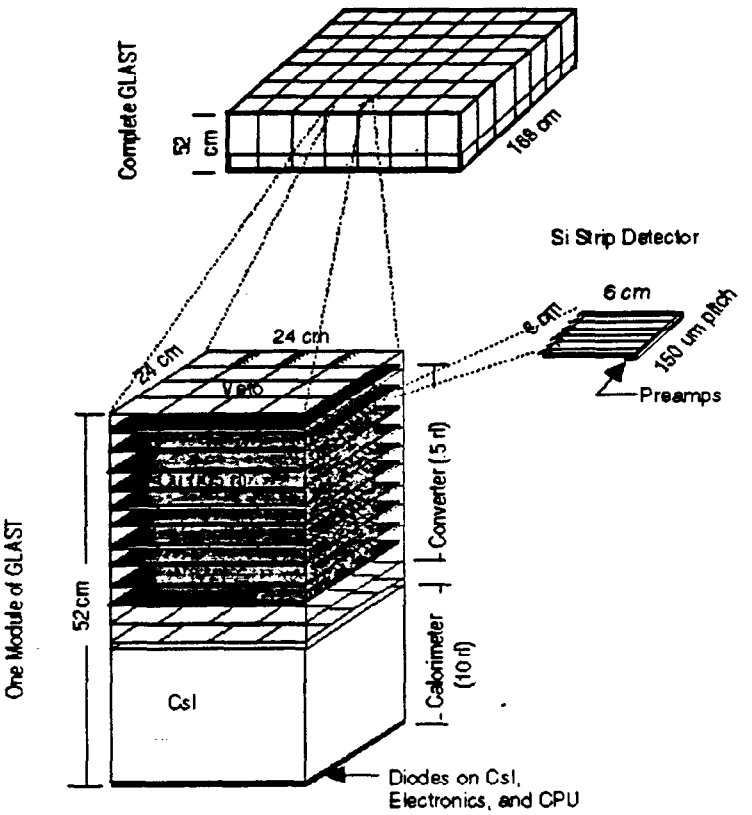


Figure 1. A schematic drawing of the instrument concept. Thin radiators in the converter portion induce pair conversion of high energy γ-rays. The energy of resulting electromagnetic shower is measured in the calorimeter section.

A series of 13 trays form the tracking portion of the one tower module of the telescope. The first and last two trays in the sequence have no extra converter material. In the case of the first tray, which serves as a tracking veto "shield" to discriminate against the entry of charged particles, conversion here would diminish the flux seen by the following layers. The last two trays also have no extra material since the minimum requirement to recognize a track is at least 3 distinct (x,y) hits. Radiator material in these last two trays would contribute to multiple scattering errors for already converted γ's, while γ's converting that late in the tower would not be reconstructible.

To measure the energy of the converted γ-rays, the tracking portion of each tower is followed by an array of 10-radiation-length long CsI crystals. This calorimeter is broken up into course "pixels" to improve the pattern matching between reconstructed tracks in the upper portion of the tower with the recorded energy depositions. The pixel size that has been studied has ranged from 2-6 cm on a side. Collectively the tracking section using the silicon strip detectors and the calorimeter form a tower module that is the basic unit from which GLAST is constructed.

GLAST is formed by making an array of tower modules. The number of modules has been chosen to fit within the launch capabilities of a Delta II rocket. Serendipitously, this scopes the instrument to just fill the niche between what is best done in space due to atmospheric obscuration and the emerging 100+ GeV ground based techniques[2]. But as shown in Figure 2, GLAST is not much larger than EGRET. It weighs approximately 50% more, and has >6 times the effective area at all energies. Notice also that the EGRET time-of-flight system(TOF) is not a component of the GLAST design. Because of the use of a triggered spark chamber to track converted γ-rays in EGRET, the trigger rate needs to be severely limited to conserve the gas supply; the TOF reduces the hardware trigger rate by a factor of 3-5 [3].

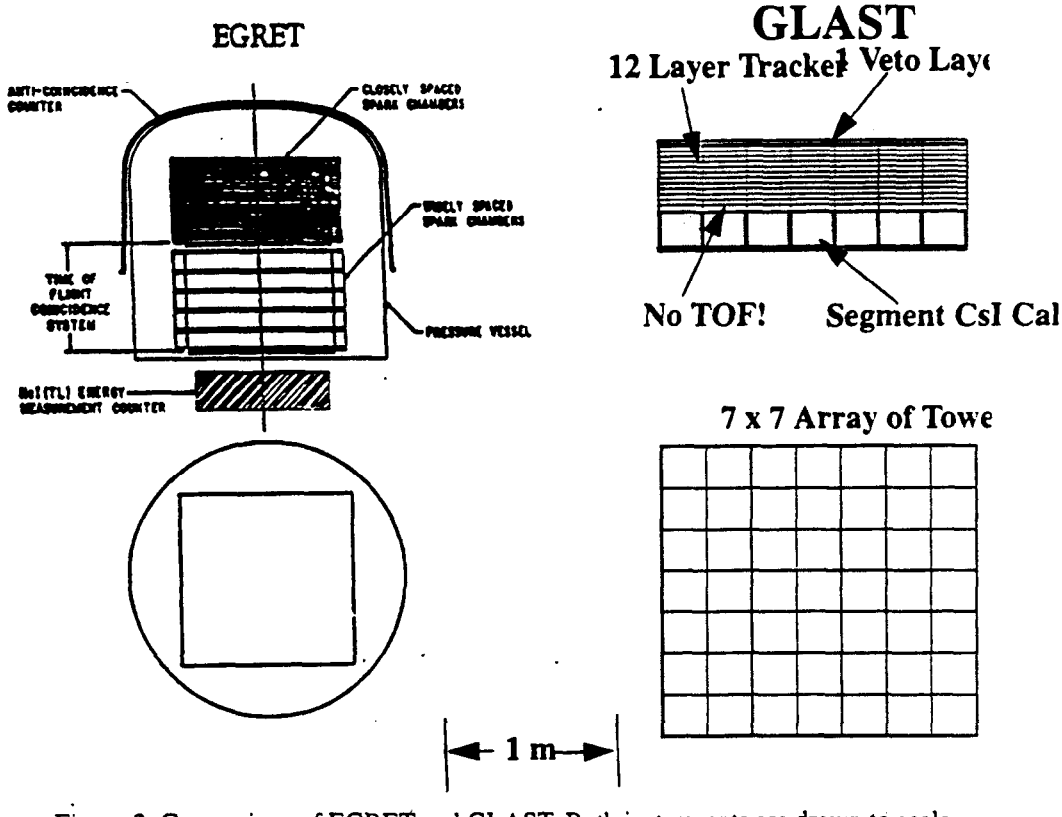


Figure 2. Comparison of EGRET and GLAST. Both instruments are drawn to scale.

The resulting GLAST instrument dramatically increases the observational parameter space available with EGRET in all aspects. In Figure 3a, the now familiar (and famous) EGRET All-Sky-Survey is shown in galactic coordinates. The data for this picture was gathered over the first two years of the EGRET mission and due to the limited field of view required nearly 40 individual settings ("pointings") to cover all 4π sr. Superimposed on the EGRET All-Sky Survey are two circles roughly centered on the Galactic Center showing the GLAST field of view (in white) and that of EGRET (in gray). This is perhaps the most striking advance over EGRET. It is seen from this figure that GLAST covers almost half the sky at any given instance. Usually "all sky monitors" (such as GLAST) are not imaging but as shown in Figure 3b the GLAST single photon reconstruction accuracy exceeds that of EGRET by a factor of 2 - 5 depending on energy. Figure 3c shows the expected effective area of GLAST and EGRET as a function of energy. Also shown on this figure is the increase in effective area that might be realizable if the calorimeter can be sufficiently finely segmented (in depth as well as transversely... see Section 5.0). The combined impact of all of these improvements over the present is dramatic!

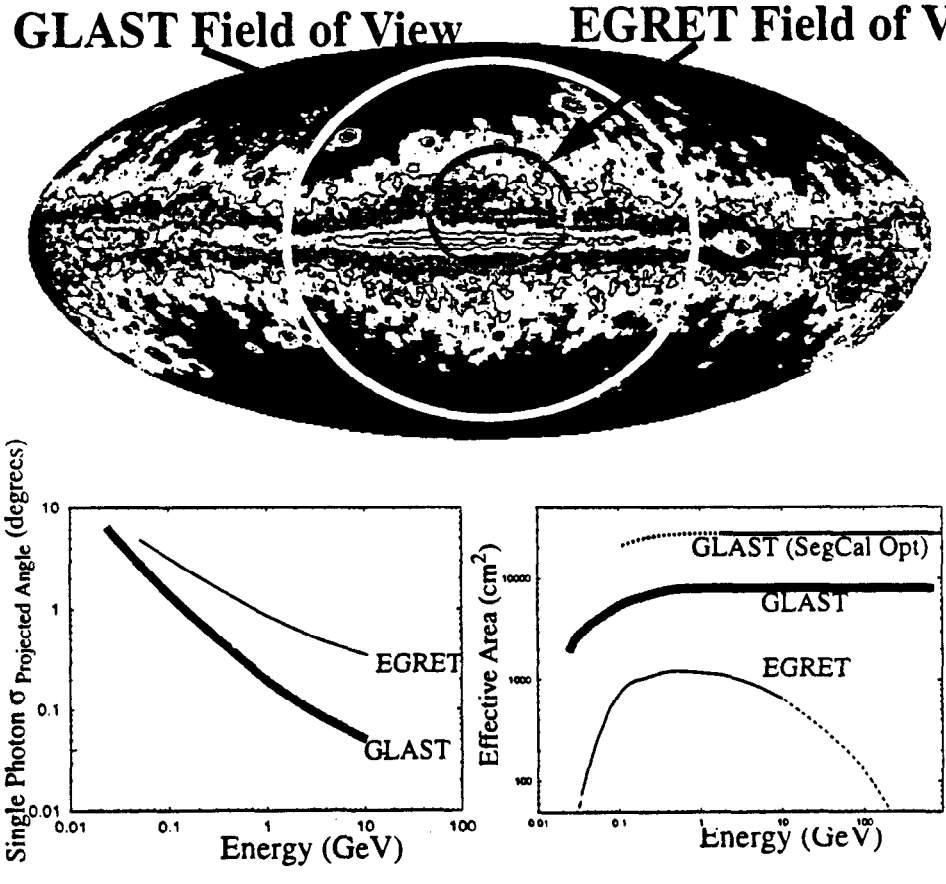


Figure 3. Comparison of Egret's acceptance, angular resolution, and effective area to GLAST's. The relative acceptances are overlayed on EGRET's all-sky survey map. A typical setting made by Egret to preform the "all sky survey" are shown by the small circle. One setting of GLAST is shown by the large circle centered on $(0^o,0^o)$.

We close the introduction with a short discussion of why we have chosen SSDs. The technical benefits are obvious: NO consumables, low voltage operation, and a basic unit that lends itself to a highly modular design. This technology is robust and mature and the detectors are long lived. Furthermore, there is extensive space experience with silicon wafers arranged in large arrays: the solar panels supplying electrical power for many spacecraft!

There are also scientific advantages. By mounting the silicon strip measuring planes directly behind the radiators, multiple scattering in the first radiator is essentially eliminated as there is practically no lever arm for the angular error to act through. A simple estimation of the improvement factor this leads to over, for example, a detector having a continuous radiator shows that the discrete radiator design can be made about 4.2 times thicker per measuring plane and still achieve the same angular resolution.

The compact, flat designs that SSD's lead to also effects the low energy cut-off of the instrument. This improvement is realized because of two factors: first without a TOF system low energy particles are less likely to range-out and second they are less like to scatter out of the telescope before hitting the calorimeter. EGRET has a low energy limit of ~ 35 MeV, while GLAST still has an appreciable area of 300 cm^2 at 10 MeV. Lastly SSD's provide excellent "two track" resolution. The result is fewer "merged" hits and hence a much more detailed "picture" of the event. This becomes important when distinguishing γ-ray interactions from cosmic rays.

2.0 Computer Simulations of GLAST

A detailed computer model of GLAST was built using the C++ radiation transport class library, Gismo[4]. The physics of particles decaying and interaction with matter are taken from EGS4 and Gheisha for electromagnetic and hadronic particles respectively. The ionization energy losses in the active regions of the detectors are used to produced a facsimile event data stream. This data stream is subsequently processed first using an event reconstruction package followed by analysis using the interactive histogramming and "cuts" package called "HippoDraw"[5]. Only at this last stage is the Monte Carlo event "truth" information used to measure resolutions, efficiencies, and biases.

The computer model of GLAST is an array of 7 x 7 identical "tower" modules with 1 mm gaps between each. Each tower includes a 13 layer tracking section followed by an 8 x 8 array of 10 radiation-length-long CsI crystals and is shown in Figure 4. The towers are square and have an inside transverse width of 24 cm. The support structure is simulated by 1) a 2mm thick aluminum wall around all sides except the front and 2) plastic foam between each of the tracking layers, which

are spaced 3 cm apart. The front entrance window is 100 μm of aluminum foil. Each of the layers in the tracking section is modelled as two 300 μm thick sheets of silicon (providing an x and a y hit readout) and a 1 mm thick layer of silicon 15 mm wide along two edges of the layer. Gaps of 300 μm are places around edges of each 6 cm x 6 cm SSD. For the 2nd through the 11th layer 270, μm of lead (~5% of a rad. len) are placed directly in front of the silicon layers.

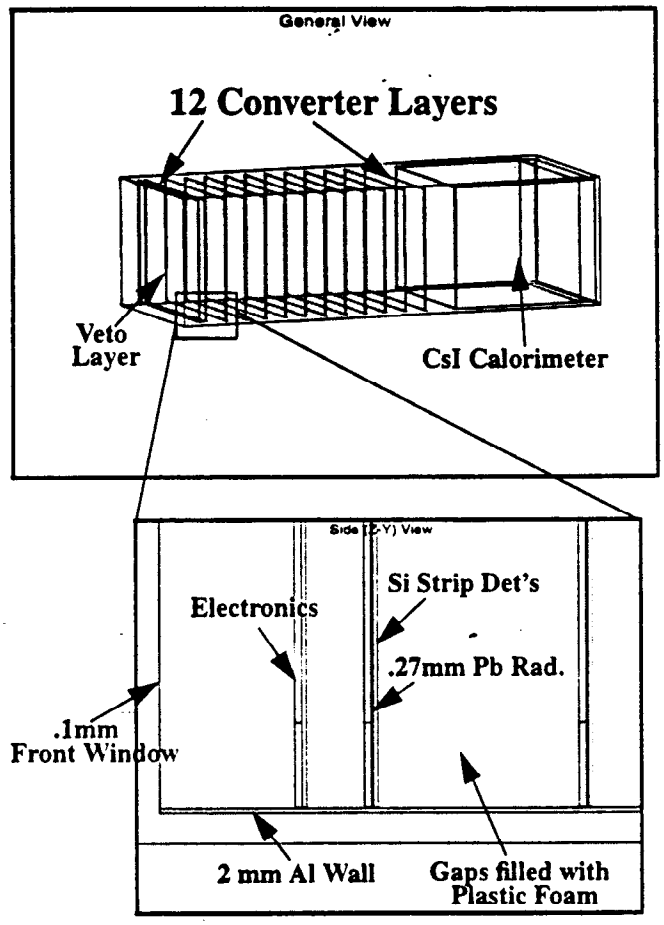


Figure 4. A "wire frame" drawing of the GLAST elements present in the computer model

Particles traversing the silicon layers have the energy loss apportioned three dimensionally among the SSD strip volumes intersected. Additionally charged particles crossing the photodiode readouts at the rear of the calorimeter crystals register an additional 2.5 MeV / $\cos(\theta_{normal})$.

A 1 GeV γ-ray incident from the front of GLAST is shown in Figure 5. Charge tracks are shown as black lines while neutrals are shown in white. The hit silicon strips are shown by the dark bars trailing layers crossed by charge tracks. The energy deposited in the CsI crystals is reflected by the rectangles drawn at the entrance to the calorimeter. The blowup of the point at which the incoming

γ-ray interacts, illustrating the fine details of the hit information available for event reconstruction.

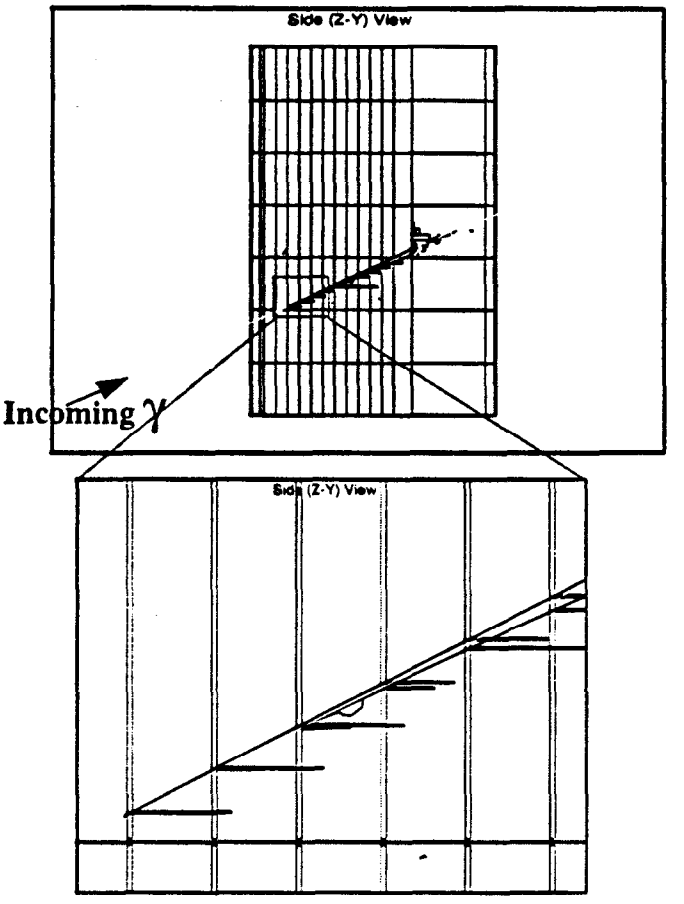


Figure 5. A 1 GeV γ-ray event in GLAST. The blow-up shows the details of the initial interaction and the hits in the silicon strip detectors, illustrated by the horizontal bars. The energies deposited in the CsI crystals of the calorimeter are depicted by the heights of the towers shown in the calorimeter.

The event reconstruction program assumes that the event is a γ-ray. First, the energy centroid in the calorimeter is found and used as a crude space point lying on the γ-ray trajectory. Combinations of x and y hits in each track layer in each tower are tried as candidate, second hits. The line formed by these hits allows for other hits in the planes between them to be added. A straight line least-squares fit is made with individual hits weighted by the predicted multiple scattering plus measurement error expected under the γ-ray hypothesis. This search continues until a suitable candidate is found or all possibilities have been exhausted. Other tracks are then looked for, first as a potential pair associated with the first track and then as stand-alone trajectories. The γ-ray is assumed to be best approximated by the first track found. This approach is admittedly simple, but to date has been adequate for the initial studies.

2.1 Acceptance and Projected Event Rates

The angular acceptance of GLAST resulting from the simulations for 1 GeV γ-rays is shown in Figure 6 as a function of $\cos(\theta_{normal})$. The approximately linear fall-off results not only from the projected area, but also from the compensating effects of increasing effective radiator thickness and decreasing reconstruction efficiency. This figure shows that GLAST has a finite acceptance to see essentially down to the instruments horizon. The wide field of view results primarily because of the relatively flat layout of the instrument afforded by both the silicon technology and the elimination of a TOF system.

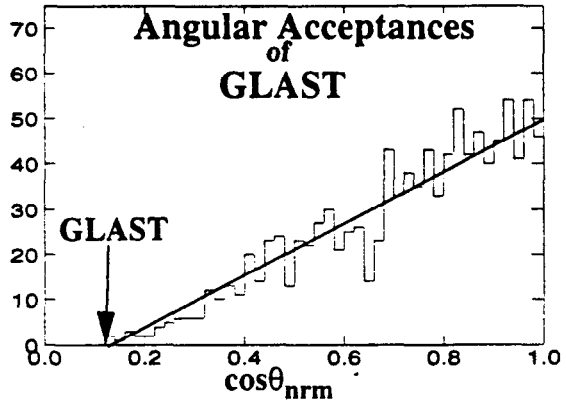


Figure 6. The acceptance of GLAST as a function of the cosine of the angle off the instrument's axis.

Given that GLAST has sensitivity to ~1.6π sr., it is possible to use the instrument in a zenith pointed mode, each orbit covering ~80% of the entire sky. The resulting inefficiency relative to pointing at a single source is on average only about a factor of two. This may be a small price to pay to effect an imaging "all sky" monitor.

Before presenting a comparison of the projected event rates in GLAST with those measured in EGRET we need to compensate for the energy range accessible to the two instruments. EGRET reports on sources for $E_\gamma > 100$ MeV. GLAST should be able to provide information well below this (down to a cutoff of ~10 MeV) with a subsequent boost in rate. In addition GLAST's effective area will remain essentially constant above a GeV as the self veto problem experienced by EGRET, is essentially eliminated by the finely segmented veto system. The net result for GLAST is an increase in rate by a factor of ~4. EGRET also records data to lower energies, but the effective area is decreasing rapidly. If this data were used, it would only increase the EGRET rate by ~ 20%

Table 1 gives the events rates for EGRET and GLAST for the weakest source reported on by EGRET (~ 10^{-7}/cm^2-sec above 100 MeV). It has been assumed that GLAST is in a scanning mode while EGRET's duty cycle on source has been derated by the ratio of its solid angle acceptance to 4π. The GLAST numbers are for the average source, that is 50% of the sky is covered by sensitivities greater than those shown and half with less sensitivity. The numbers given in parentheses include the boosts expected if the energy range is extended.

Table 1:

Rate	Egret	GLAST
γ's/year	104 (208)	4530 (16800)
γ's/week	2 (4)	87 (322)
γ's/day	29 (.57)	12.5 (46)
γ's/Orbit	.02 (.04)	.7 (2.5)

2.2 Angular and Energy Resolutions

The projected angular error GLAST makes in the reconstruction of the direction of the incident γ-ray has been studied as a function of energy. The projected angular error distributions for on axis γ-rays at four different energies ranging from 50 MeV to 10 GeV are shown in Figure 7. The overlaid fits are Gaussians which, although not being poor representations of the data, do undershoot the peak and the wings. The resolutions below ~5 GeV are dominated by multiple scattering. Above this the finite silicon strip pitch plays an increasing role and finally limits the resolution to a little less then 1 mrad. A more sophisticated event reconstruction could result in some improvement on the overall resolution.

The energy measurement afford by the 10 rad. len. CsI calorimeter for the same set of runs is given in Figure 8. A correction for the shower development occurring in the tracker portion has been applied and is particularly important below about 100 MeV. The systematically low mean observed energy results from shower leakage and no attempt has been made to correct for this . Also note that there are not anomalously large high energy tails due to charge particles crossing

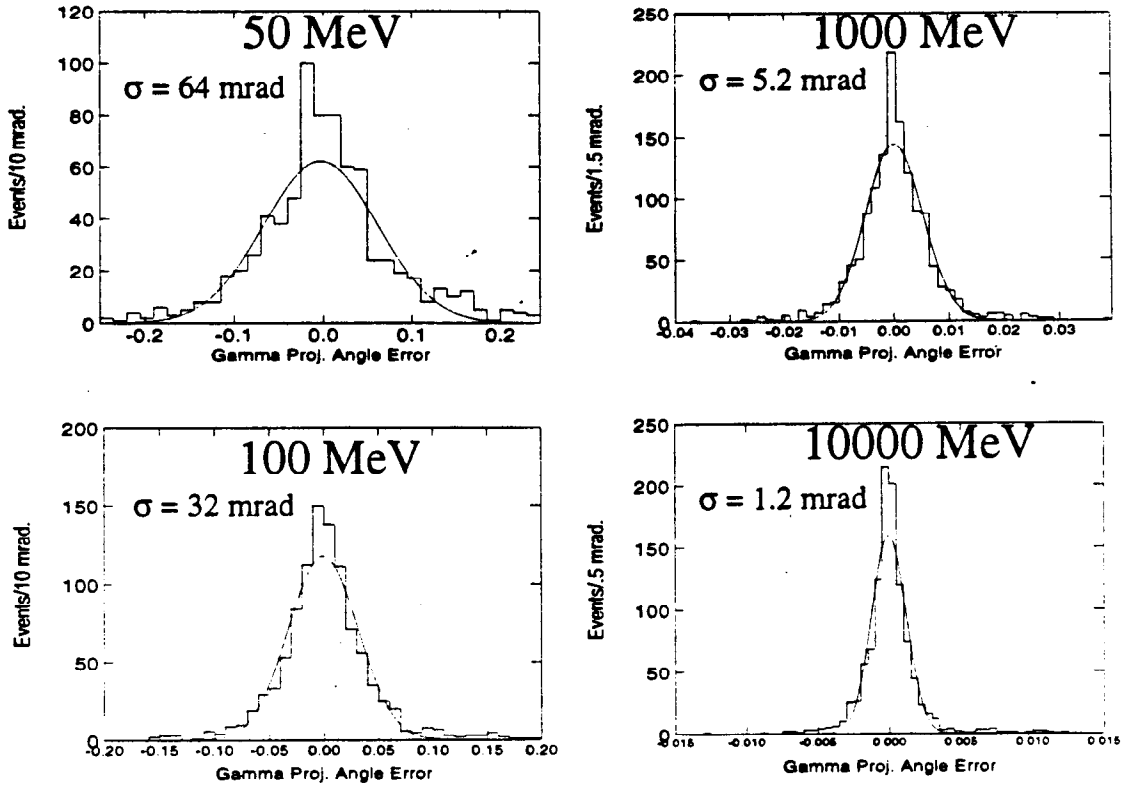


Figure 7. Projected angular error distribution from the GLAST simulations at 50, 100, 10^3, and 10^4 MeV.

the photodiodes (as already mentioned, this effect is included in the simulations). Again, overlaid on the plots are again Gaussian fits as reference guides. While a Gaussian is a fair representation of the data at low energy, at high energy the large, low energy tail from leakage skews the distributions considerably. A calorimeter with more radiation lengths would postpone this effect to higher energies, but the depth of the calorimetry must be balanced against the aperture and fit within the weight limits of the intended mission. Though the energy measurement becomes increasingly inaccurate with increasing energy, a 10 rad. len. calorimeter will provide useful information well beyond 100 GeV. Finally for an instrument with an effective area of ~1 m^2, the number of expected events above 100 GeV are to few to merit better than an ~25% energy measurement.

2.3Background Rejection

One of the most challenging requirements for astrophysical γ-ray instruments is the ability to distinguish the genuine γ-ray signal in a flood of backgrounds. In low Earth orbit the

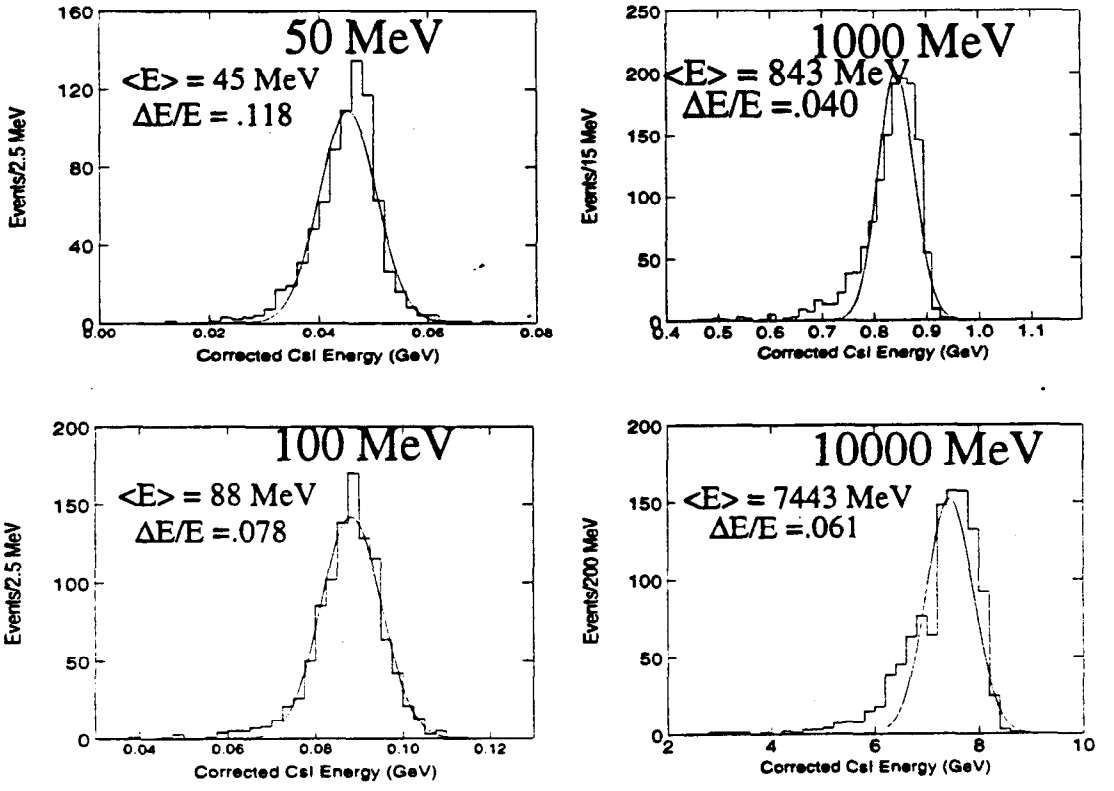


Figure 8. Reconstructed energy distributions from the GLAST simulations at 50, 100, 10^3, and 10^4 MeV.

direct extraterrestial cosmic rays are cut off below a few GeV by the Earth's magnetic field but extend to higher energies falling approximately as $1/E^2$. These high energy cosmic rays are mostly protons, the remainder being heavier nuclei, mostly α particles. Interactions of these particles with the spacecraft, the calorimeter, etc. can occasionally mimic γ-ray interactions. Furthermore these cosmic rays interact with the earth's upper atmosphere and create a plethora of secondaries, usually referred to as albedo, which in the direction of the earth's limb are exceedingly intense. Among these sources we find that the most challenging to eliminate are the direct high energy cosmic rays.

We have used the CREME program[6] to mimic the energy spectrum of the direct cosmic ray. The GLAST instrument was subjected to an isotropically distributed sample of these protons from on axis to -20° below the instrument's horizon. This approximately covers the direct exposure for GLAST operated in a zenith pointed mode. These events were processed by the event reconstruction and then techniques to distinguish them from γ-rays were developed. While the material model for the GLAST instrument is fairly complete, left out of the computer modelling

are the effects of the material in the spacecraft, but this is thought to represent only a small overall increase as by far the most massive "target" around is the calorimeter.

The first requirement placed on events is that they appear to "materialize" within the active volume of the detector. In practice this means that the reconstructed charged particle trajectories in an event cannot be traceable all the way to the edge of the detector. This is a tracking version of the charged particle anti-coincidence shield used in EGRET and previous γ-ray instruments. The advantage is that effective veto area used in each event is only a small fraction of the full aperture of GLAST. Few γ-ray events are hence eliminated by having a small energy deposition caused by shower back-splash from the calorimeter. EGRET's efficiency fall-off above a few GeV is attributed to this effect. In practice the veto function is accomplished by tracing all reconstructed charged particle trajectories backwards from the first hit (the one furthest from the calorimeter) and requiring that it past through at least one layer of live silicon detectors (i.e., not a gap). In each such layer crossed, the nearest hit to the projected track is found and the minimum distance of closest approach (DOCA) is recorded for all tracks in the layers crossed. Plots of this DOCA for good γ-ray events and for the cosmic proton sample are shown in Figure 9. Events in the overflow bin are ones in which no hit was found in the effective veto layer(s). Requiring that there be no hits within a radius of 10 cm about the track(s) in the veto layer(s) results in only about 4% of the γ-rays being eliminated, while 99.4% of the protons are killed. Most of these are cosmic rays that enter the instrument through the front; however, events entering the calorimeter from the sides and back, which produce interactions with secondaries that leave through the tracker are also eliminated.

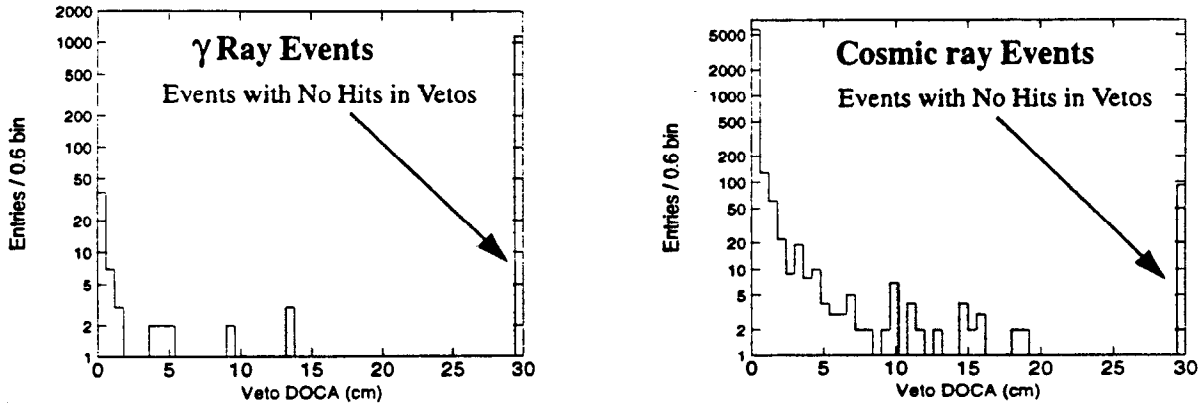


Figure 9. Distributions of the distance from the reconstructed γ-ray trajectory to the nearest hit in a veto layer. On the left the results for 1 GeV γ-rays is shown while on the right are the results for cosmic ray proton event.. Events with no veto hits are shown in the overflow bin. The shaded area indicates the events that are retained.

Next the goodness of the least squares fit (χ^2/deg. of freedom) is used. The error component

attributable to multiple scattering is computed from the observed energy deposition in the calorimeter. Interacting protons usually only leave a fraction of their energy in the calorimeter and this energy is not well correlated with the secondary producing the track observed in the silicon. Figure 10 shows the distribution of χ^2/deg. of freedom for both γ-rays and the remaining protons after the veto cut. Requiring χ^2/deg. of freedom to be less then 10 eliminates only 1% of the γ-rays while almost 3 out of 4 of the remaining protons are killed.

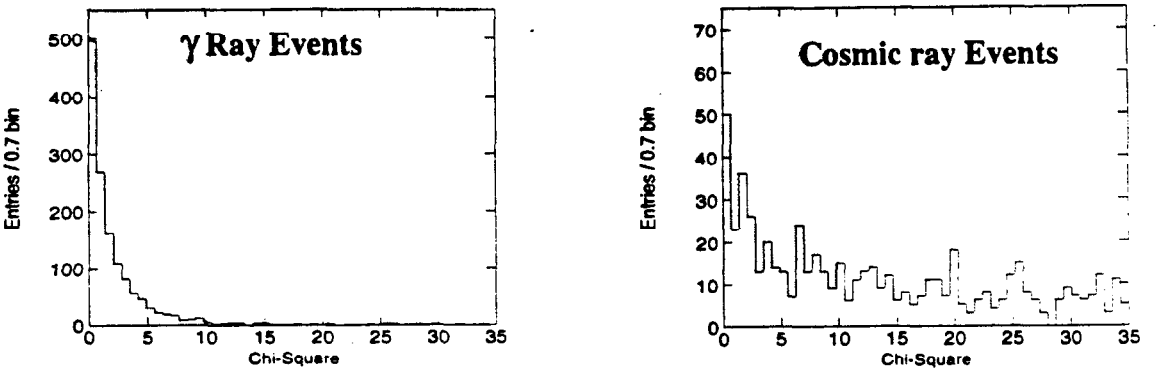


Figure 10. χ^2/deg-of-freedom for γ-ray events and cosmic ray protons. Events in the shaded regions are retained.

The spacial information in the tracker is compared with the energy centroid in the calorimeter to further reduce the hadronic background. To do this, the tracker trajectory is projected into the calorimeter to a depth approximately corresponding to shower maximum, and the transverse distance between this point and the centroid is recorded. For γ-ray events this is strongly correlated while for protons the correlation is poor, as shown in Figure 11 (The events shown in this figure have only had the veto cut applied to maintain a reasonable statistical level.) After this cut, 3×10^{-4} of the original proton sample is left , but only another ~2% of the γ-rays are lost.

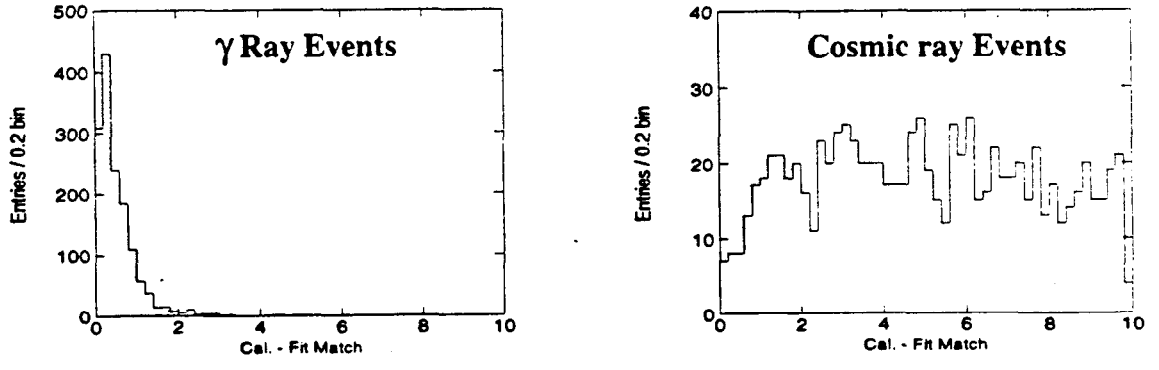


Figure 11 Distance between the energy centroid in the CsI crystals of the calorimeter and the projected γ-ray trajectory for γ-ray events and cosmic ray protons. Events in the shaded regions are retained.

An examination of the proton events that remain reveal three categories: 1) calorimeter interactions where a soft secondary prong ranged out in the tracker, 2) events producing many soft secondary particles (and hence lots of hits), usually leaving very little energy in the calorimeter in which random uncorrelated sets of hits were grouped together to form reconstructed trajectories, and 3) events where the proton enters the side between tracker layers and interacts in the converter material, producing tracks headed for the calorimeter. All of these categories tend to share the common feature that the reconstructed trajectory has few, if any, hits close to it other then those used in its own fit. This is not true for γ-ray events since the developing shower almost always has nearby hits (until the energy gets so high that the hits become merged). To quantify this observation, a virtual cylinder is formed around the supposed converted γ-ray trajectory and the hits within the cylinder are counted. The ratio of this number to the number of silicon planes is formed. Figure 12 gives this "surplus hit"ratio for γ-ray events and for protons: note that the γ-ray events tend to have this ratio >1 while the proton events are predominantly at exactly 1. This cut improves the rejection of protons over γ-rays by another order of magnitude, but at a cost of about an additional 10% inefficiency for γ-rays. The effective areas plotted in the first section of this talk include all of the inefficiencies incurred by the cuts necessary to eliminate the cosmic ray backgrounds.

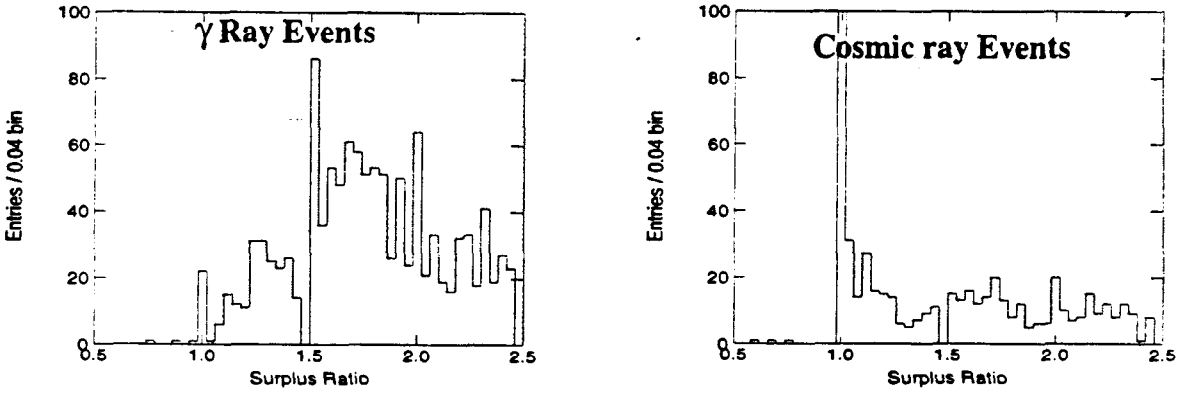


Figure 12. The surplus hit ratio (see text) distributions for a)γ-ray events and b)cosmic ray protons. Events in the shade regions are retained.

The remaining efficiency for accepting protons is 3×10^{-5}, which would seem to leave an unacceptably high level of background events in the data sample. But this is not the case. Figure 13 shows the spectrum of generated proton energies (for events after the veto cut) and the spectrum of observed energies after all cuts. The generated spectrum shows the geomagnetic low energy cutoff and a peak at ~15 GeV, but the observed energies of the events after filtering are very small.

Indeed only one event deposited more then 100 MeV! It is now also apparent where these events come from. Low energy in the calorimeter has the effect of relaxing the multiple scattering errors applied in the pattern recognition/fitting process. Events with many random hits and low energy can satisfy the cut criteria, but would be easily recognized by a more sophisticated analysis (note: much of the EGRET data is "hand scanned" to remove this sort of thing). Even so, if we compare the high latitude diffuse γ-ray event rate above 100 MeV (~2 x 10^{-5} /cm^2-sec-sr) to the integrated cosmic ray rate(~.2/cm^2-sec integrated over all solid angle), we compute a signal-to-noise of ~20: 1. Given the lack of sophistication, and limited time spent to date on developing this filtering algorithm, we are confident that backgrounds will not be a problem for GLAST.

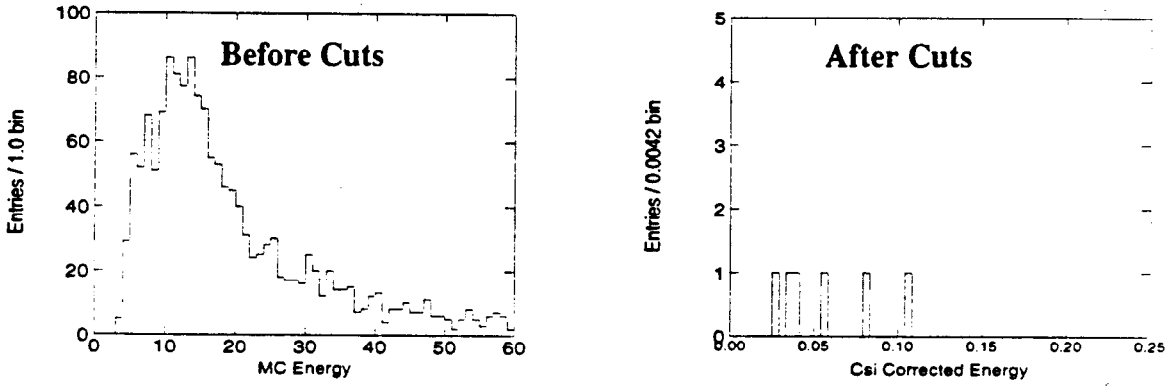


Figure 13. Energy distributions of cosmic ray proton events a) after the veto cut but before the other cuts and b) as reconstructed in the CsI for the remaining events after all analysis cuts. Only one event remains with a recorded energy greater than 100 MeV

3.0 Preliminary Instrument Design

3.1 Tray Electrical and Structural Design

The GLAST instrument will have more SSDs than have been previously used in experiments. Furthermore, the detectors must be mounted in a manner suitable for a rocket launch. The number of channels in our present design is approximately 1.3 x 10^6. As each channel is comprises of four SSD strips connected end-to-end, the number of interconnections is really four times larger than the channel count.

The newer technologies for mounting silicon integrated circuits offers some promising solutions to this problem: specifically the TAB (tape automated bonding) methods. Connections are made in blocks in TAB processes. The silicon circuit presents its electrical connections as bonding pads around the periphery of the chip. The mount is commonly copper foils on a Kapton backing arrayed

in matching pattern. Conductive epoxy or gold bumps on the pads make the contact between the matching pads. These connections can be made in one set for each circuit rather than one at a time as is the case for wire bonding. Furthermore, the connections are more robust, which is potentially very important in our application.

The current GLAST design is to mount the detectors for a single coordinate readout of a tray on a sheet of kapton, which carries the electrical interconnects. Two such lay-ups would then be glued back-to-back, forming an x,y pair. On top of this, the radiator material (5% rad. len. of lead) is mounted (again using a sheet of kapton as a backing). Lastly, the whole assembly is glued to a carbon fiber panel for mounting in a GLAST tower module. This scheme is pictured in Figure 14. Prototypes of these trays are now under construction to unearth problems and further refine the design.

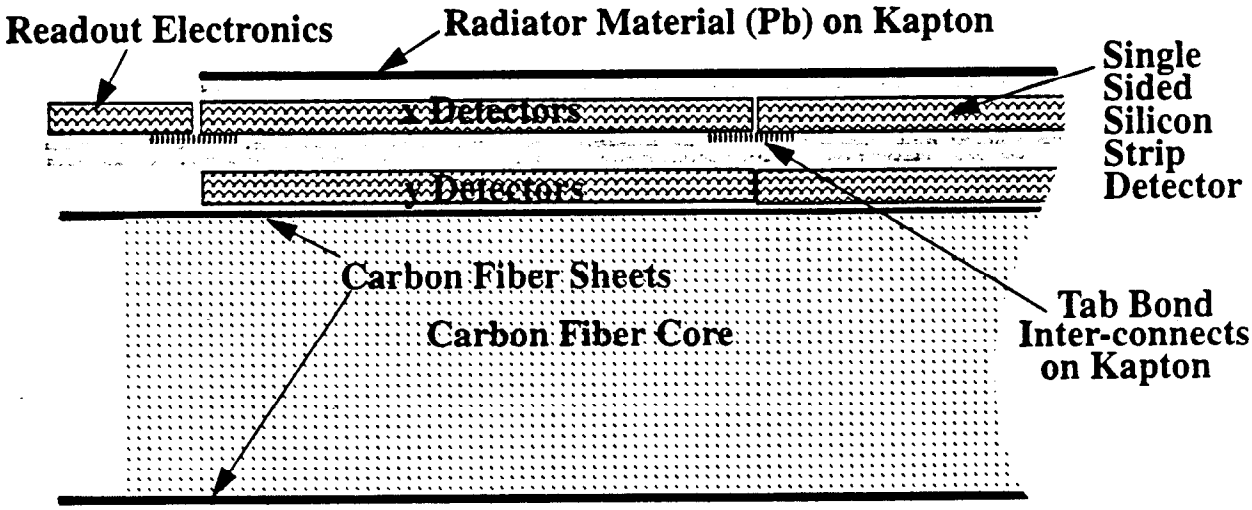


Figure 14. Schematic of a GLAST detector tray.

3.2 Low Power Preamps

Critical to GLAST is the development of low power front-end electronics for the silicon strip detectors. In High Energy Particle Physics experiments the major penalty to using "high power" pre-amps on SSD vertex detectors is heating. For the SSC detectors the scale of the trackers became large enough to warrant a serious effort to minimize power consumption. The results were pre-amps using less then 1 mW per channel. For SSC applications there was also a requirement for high speed which in general pushes the power consumption up. However, for a GLAST application the speed requirement can be relaxed by over an order of magnitude and allows the usage of CMOS instead of the more power hungry bi-polar technology. From early on we were confident that front

ends could be developed using less then .5 mW per channel. Currently three options are being evaluated.

At Los Alamos National Lab our collaborators have designed a new, grounds up, pre-amp front end based on CMOS and our shaping time requirement of ~ 1 μsec. This design is projected to use ~100 - 200 μW per channel and the first round prototypes are at Los Alamos being evaluated.

Other collaborators at UCSC, with strong connections into the work done for SSC detectors, undertook to redesign the SVX II front end pre-amp again with our requirements in mind. The starting point for this design was the fully developed and debugged SVX II chip which used ~ 1 mW/channel. Slowing down the shaping time along with optimizing for our projected strip capacitance resulted in a design that should be ~.5 mW/channel. Again prototype chips have been ordered and will be evaluated this year.

The third option comes from our Italian collaborators in Triest. They have been studying and prototyping SSD pair telescopes since the late '80s. Their work has evolved a front end chip, complete with a serial readout, which is now in its second round of prototypes. They expect the power in this design to be ~ 600 μW/channel. They have shared some of these chips with us for evaluation which is currently under way.

3.3 Level I Trigger

The basic hardware trigger in GLAST is derived from the SSDs directly.The challenges are to design a system requiring no external inputs, that is totally asynchronous, and that minimizes dead time. To this end we have proposed the following design. In each tray of each tower in GLAST all the signals from the discriminated front end pre-amps are OR'd together for each coordinate (1024 channels each). The singles rate on each silicon strip is calculated to be ~50 Hz based on the assumption that the discrimination threshold will be set at 4σ above the noise (approx ~1/4 min. ionizing). Hence these OR'd signals will have a rate of ~50 kHz. A coincidence is formed from the x and y signals with a ~ 1 μsec width, reducing the rate to ~2.5 kHz.

The tray coincidences are fed to logic associated with each tower, which then OR's each layer with all the trays from the same layer in the neighboring towers (9 x-y coin. signals total). This is required to avoid edge effects and to allow steeply inclined tracks to trigger. The layer rate runs at ~23 kHz. These signals are subsequently fed to a programmable gate array that produces an output if any three adjacent layers are in coincidence. Again, using a 1 μsec pulse width, the resulting effective threefold coincidence should be ~0.2 Hz per tower (or 10 Hz total

over all towers). This is just the stochastic noise rate, and the real charge particle rate is much larger (~6 kHz).

The charged particle rate can be reduced considerably at Level I by feeding the gate arrays with the veto signals from the top layer and demanding that they be in anti-coincidence. The bulk of the remaining rate will be due to charged particles entering through the sides. In either case the trigger rate (much less than 10 kHz) is easily read out by simply feeding back to the planes a Level I trigger signal, causing the hit strips to be latched into a readout buffer. By making this buffer just two deep, the dead time is reduce to essentially zero. Only trays having an x,y coincidence will be latched in, by simply putting the level I trigger in coincidence with the existing tray x,y coincidence signal.

3.4 Level II Trigger

As in many particle physics experiments today, the final event trigger is realized in the data acquisition computer. Given that limited processing power will be available on the GLAST spacecraft, a sufficiently simple approach must be found to limit the rate. To this end the following scheme is proposed .

The nature of a silicon strip detector is digital. The data stream from the Level I trigger is a series of addresses of the silicon strips that have fired their discriminators. The addresses are most naturally arranged in ascending (or descending) order in each layer. As such, the low order bits of the difference between hit strip addresses in adjacent layers is proportional to the slope of a line connecting them. A straight through trajectory will give the same address difference in each pair of layers it passes through. By binning these address differences, charged particle trajectories will show up as peaks in the distribution.

In addition to calculating the address difference for each pair of hits, for each such pair if the projected address of the strip in the layer preceding them is computed, an effective "veto" can be realized. This is done by asking if the projected address is valid. If it isn't, then the projection is outside of the detector. By flagging these pairs, the peaks in the address difference distribution to which they contribute can be identified as charged particles entering the sides. Similarly if the front veto layers are not included in the Level I trigger, flagging of pairs pointing to veto hits closes the aperture. The event is passed if a peak exceeds a progammed threshold and contains no flagged pairs. The veto cut reduces the charged particle rate by a factor of ~250. The resulting software trigger rate will then be less then 30 Hz. As our event size is estimated to be ~0.5 kBytes, this rate would result in less than 100 kBit/sec down link rate.

Because the above address difference peaks are really just the trajectories of charged particles in

instrument coordinates, on can eliminate events coming from known and unwanted sources such as albedo from the earth's limb. This crude pattern recognition can also be used to quickly identify transient phenomena in the sky as well as serve as the front end data reduction for a more extensive analysis on the spacecraft to further limit the downlink rate.

The required computing capacity to perform this type of software trigger is estimated at less than 20 MIP's. Such computers are now in the process of being prepared for space qualification and flight testing as part of the USA experiment onboard the ARGOS satellite by our collaborators at NRL[7] due for launch in 1996.

4.0 Instrument Extensions and Options

In this section we discuss an option for GLAST that could increase the effective area by about a factor of three. In particle physics experiments the calorimeter is often divided up longitudinally as well as transversely. As such, information of the longitudinal shower development is obtained, and can be used to aid in the identification of electromagnetic showers as opposed to hadronic showers. In addition, it can be used to reconstruct the direction of the particle initiating the shower. The angular resolution is poor compared with that obtained with a layered pair spectrometer (as in the tracker section of GLAST), but at high energy it may be adequate for the study of objects at high galactic latitude.

To study this possibility the 10 rad. len. GLAST calorimeter was divided into 4 layers, each of 2.5 rad. len. To improve the angular resolution for γ-rays converting in the calorimeter, the "pixilation" of the first layer was increased by reducing the crystals to1 cm x 1 cm. This results in far too many individual parts, and is justified only as a fast way to perform a simple first study. A typical 2 GeV γ-ray shower in this layered calorimeter is shown in Figure 15. Event reconstruction was done by simply performing a least squares fit to the energy centroids in each layer. The resulting angular resolution follows the expected $1/(\sqrt{E\gamma})$ behavior and is shown in Figure 16. While this angular resolution is considerably worse than that obtained by γ-rays converting in the tracker, it could result in useful data for isolated sources, particularly if they are transient in nature. Background rejection has yet to be investigated for this trigger mode in GLAST.

More possibilities for GLAST exist. For example, adding BATSE like counters at the corners would greatly improve the γ-ray burst capabilities while not adding substantially to the load on the spacecraft and its systems. Such corner detectors would see the 10 keV - 10 MeV γ-rays, while GLAST would cover the energies above this. Together the location of the sources for burst as well as their high energy tails could be studied in detail.

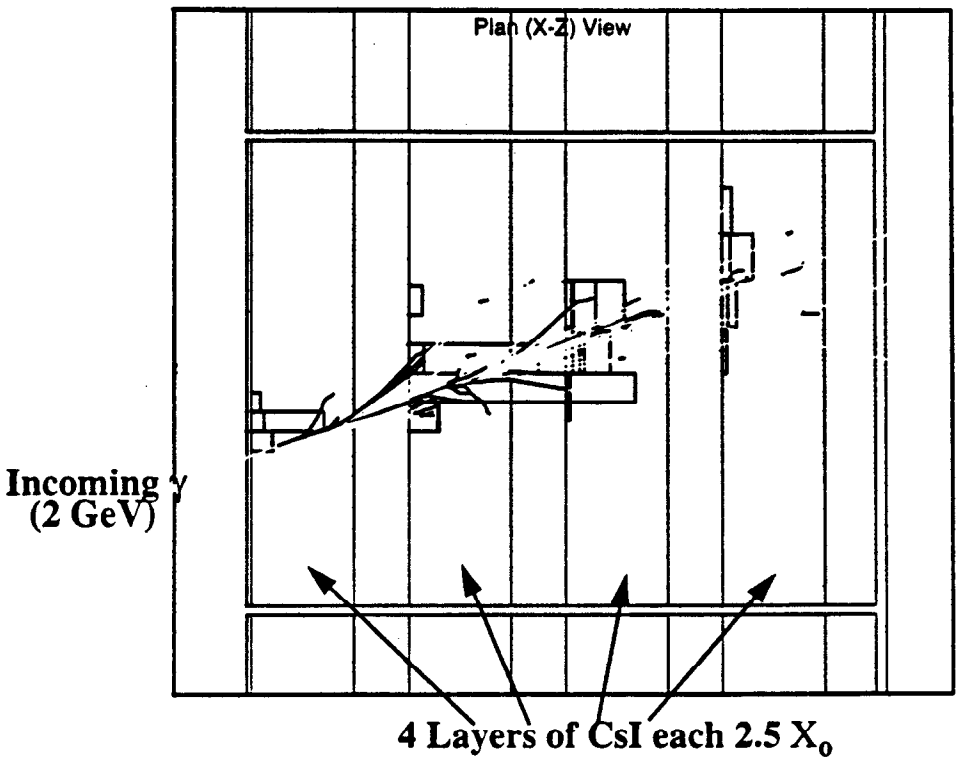


Figure 15. A 2 GeV γ-ray event in a longitudinally segmented calorimeter (see text).

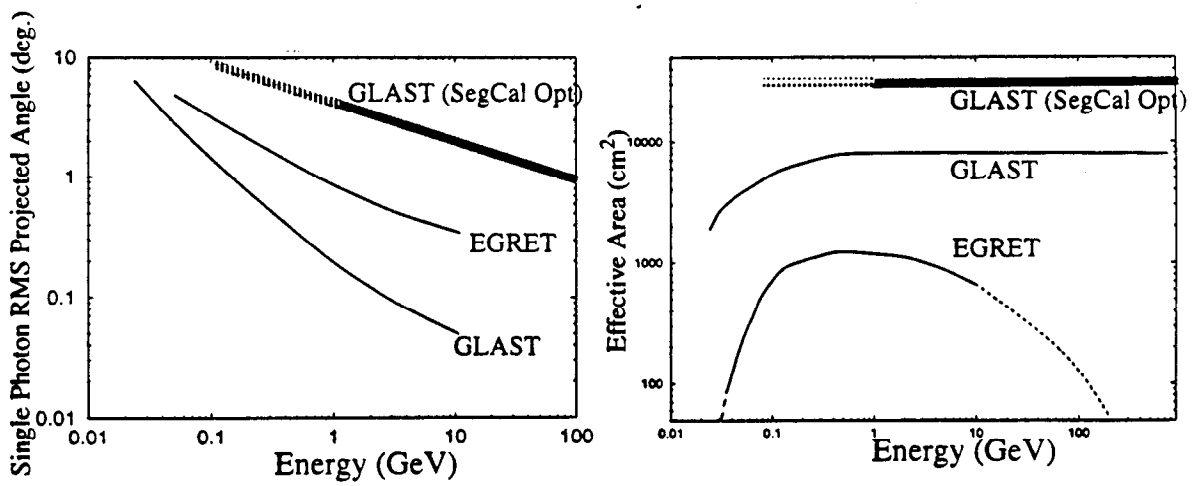


Figure 16. Angular error and effective area as a function of γ-ray energy for the segmented calorimeter option. The curves for EGRET and GLAST (without this option) are shown for comparison.

It has been posited that GLAST may have some sensitivity to γ-ray polarizations below ~ 100 MeV. This comes about for γ-rays which convert in the support material after the SSDs. There is ~.5% of a rad. len. of such material and this may be thin enough to see a polarization signal. Furthermore,

a portion of the events converting in the support material after the SSDs in any of GLAST's tracker planes will be identifiable by having multiple hits in the first tracking plane with hits. The azimuthal orientation of these hits about the shower axis reflects on the γ-ray's polarization.

The list goes on with items such as cosmic ray calorimetry, e+e- spectra, the (remote) possibility of coincident events (recall GLAST's timing is at the ~ 1 μsec level), etc.

5.0 Conclusions

In this section, I won't recapitulate the many aspects of the GLAST Instrument dealt with in this talk. I will simply give my conclusions drawn from having worked on this project for the past two years.

1) GLAST would extend EGRET's discovery reach by a factor of ~100. It is rare in any field of science when less than highly speculative technology can so advance a field.

2) GLAST would provide an excellent monitor for transient sources. The emerging ground based γ-ray instruments (CAT in particular) are limited field of view instruments and would benefit tremendously by having a somewhat lower energy, all sky monitor (i.e. $E_\gamma < 100$ GeV) to alert them of transient "targets of opportunity." GLAST's effective area is too small to effectively study such phenomena above 100 GeV, furthermore, why make a super space based instrument for this ultrahigh energy domain when it can be done from the ground?

3) The GLAST instrument concept is viable. The simulations done to date have increased both our enthusiasm for GLAST and confidence that it will work at least as well as EGRET and it predecessors. Note that these previous experiments were not nearly as thoroughly investigated prior to construction and launch as GLAST.

4) The technology that GLAST is based on is "IN HAND." There is nothing new to invent, but there are engineering problems to be solved (i.e., low power preamps, segmenting the calorimeter, etc.). GLAST is simply an adaptation and optimization of technology now 10 years old in particle physics. SSDs have been shown in many experiments to be robust, highly efficient, and reliable.

GLAST is a tremendous Scientific Opportunity!

GAMMA-RAYS: TEV ENERGIES AND BEYOND

Rene A. Ong
University of Chicago

1. Introduction

In discussing the field of high energy gamma-ray astronomy, one must use as a starting point recent exciting discoveries made by the EGRET detector aboard the Compton Gamma Ray Observatory [1], discussed by Gottfried Kanbach at this workshop. EGRET, operating in the energy regime between 30 MeV and 20 GeV, has detected gamma-ray emission from approximately 130 point sources. Of these point sources, six are identified as spin-down pulsars, approximately forty are identified as active galactic nuclei (AGN), and the rest are unidentified.The AGN were one of the big surprises of the EGRET observations. Only one source (3C273) had been detected by the earlier generation of gamma-ray instruments (SAS-2, COS-B). Since AGN are distant objects, the copious amounts of high energy gamma-ray emission detected at Earth implies tremendous source luminosities, even when possible beaming effects are taken into account.

More than 90% of the gamma-radiation seen by EGRET is in the form of diffuse emission from the galactic plane. This emission comes from the collisions of cosmic rays with gas and dust producing energetic electrons and pions. Electrons radiate and neutral pions decay to give gamma-rays. EGRET has also detected diffuse emission from the Large Magellenic Cloud (LMC) in a 7 degree by 7 degree box. Studies of diffuse emission are important because of the way in which the gamma-radiation traces the cosmic ray flux. For example, recent observations by EGRET of the Small Magellenic Cloud (SMC) [2] provide evidence that cosmic rays are galactic in origin. If cosmic rays were universal and extended to the SMC, one would expect a diffuse gamma-ray flux of (2.4-4.2) x 10^{-7} photons/cm^2/sec from the SMC. The EGRET flux limit of $< 0.5 \times 10^{-7}$ photons/cm^2/sec argues against such an universal picture. This result illustrates the close interplay between cosmic ray and gamma-ray research.

Another exciting result from the Compton Gamma Ray Observatory is the detection of gamma-ray burst sources, discussed by Dieter Hartmann at this workshop. Over 1000 gamma-ray bursts (GRB) have been detected by the BATSE instrument [3] and ten have been seen in coincidence by EGRET. The bursts are observed to be isotropically distributed and to have a wide variety of spectral and temporal features that defy classification. Since the bursts have not been demonstrated to repeat and have not been positively identified with objects seen a lower energy, their origin remains a mystery. The general view is that GRB are cosmological or from a galactic halo population with a scale of 70 kpc for more. Very surprising was the detection by EGRET of a burst on February 17, 1994, in which high energy emission was seen roughly 80 minutes after the initial burst.

Above 20 GeV the gamma-ray fluxes become so low as to be undetectable by EGRET. Therefore, only ground-based experiments have sufficient collection area to make gamma-ray detection feasible. These experiments and their results are the subject of this talk.

2. Experimental Technique

Ground based gamma-ray detectors make use of the method of extensive air showers. Air shower detectors function simultaneously as cosmic ray and gamma-ray instruments; the cosmic rays are the principle background to doing astronomy. There are two major types of ground-based gamma-ray detectors: those that use the atmospheric Cherenkov technique (ACT) and those that use the air shower technique (AST).

The major challenge facing all ground-based experiments lies in their ability to reject the cosmic ray background. This challenge arises because cosmic rays outnumber gamma-rays by three to four orders of magnitude. This statement can be easily understood by considering our conventional picture of cosmic rays. We believe that cosmic rays below 10^{14} eV are galactic in origin. Since they are charged, they bend and spiral in the galactic magnetic field and are loosely confined by it. Cosmic rays have lifetimes in our Galaxy of approximately ten million years. Gamma-rays are not confined, but are created at a source and propagate outward with a $1/r^2$ fall-off. We have only one chance to intercept a gamma-ray before it continues on its journey out of the Galaxy.

2.1. Atmospheric Cherenkov Experiments:

When a high energy gamma-ray strikes the upper atmosphere, it creates an air shower of electromagnetic particles that propagate towards the Earth's surface. These relativistic particles are moving faster than the speed of light in the atmospheric medium and therefore emit Cherenkov radiation. Even when the charged particles do not reach the Earth's surface, there is substantial Cherenkov radiation that is beamed to the ground in a narrow cone of half angle of approximately 1 degree (Figure 1). The radiation is spread out over a circle of roughly 200 m in diameter. It arrives in the space of a few nanoseconds, and is copious: a 1 TeV gamma-ray will produce a shower with roughly 50 photons/m^2 at the ground.

Atmospheric Cherenkov detectors consist of a large mirror (or several mirrors) to reflect the optical-UV Cherenkov photons onto a camera. The Cherenkov image of the shower is converted into an estimate of the incident gamma-ray arrival direction. The angular resolution achievable by Cherenkov telescopes is excellent: 0.2 degrees or better.

As discussed earlier, the key feature of ground-based gamma-ray astronomy is the necessity to reduce the background of hadronic cosmic rays. Background rejection is difficult because experiments on the ground do not see the first interaction of the primary particle. They only infer something about the nature of the primary from the detection of the secondary shower particles. Atmospheric Cherenkov telescopes achieve background

rejection by either of two methods: imaging or sampling. The imaging method takes advantage of the fact that showers initiated by hadronic cosmic rays are much more irregular in development than those initiated by gamma-rays. This irregularity comes from the fluctuations in the development of the hadronic core of cosmic ray showers and in the fraction of energy carried away by hadrons and muons. With a multiple-pixel camera, typically an array of photomultiplier tubes (PMTs), Cherenkov telescopes record an image of the profile of the shower in its lateral and longitudinal extent. Images for gamma-ray showers are elliptical and point towards the center of the field-of-view. Images for showers created by hadronic cosmic rays are typically more irregular in the focal plane (Figure 2). Current state-of-the-art imaging detectors achieve a background rejection of greater than 99%, while keeping the majority of the gamma-ray events.

The sampling method takes advantage of the fact that most of the Cherenkov radiation in a shower reaches the ground in a sharp conical wavefront. By recording the time of arrival of this wavefront in several detectors separated by some distance, the shower direction can be reconstructed with accuracies approaching a few milliradians. In this way, background rejection is achieved through excellent angular resolution. Additional rejection can be obtained from the lateral distribution of the Cherenkov light on the ground.

The atmospheric Cherenkov technique has been used by ground-based telescopes for more than thirty years. It is only in the last ten years that significant advances have been made in the power of the instruments to reject background. These advances are directly responsible for the detections of three gamma-ray sources by Cherenkov telescopes. State-of-the-art experiments exist on several continents. In the United States, the premier instrument is the Whipple Observatory, located on Mt. Hopkins near Tucson, AZ. Whipple consists of two mirrors (diameters 10 and 11 m), separated by approximately 120 m. Each mirror has an imaging camera of roughly 100 PMTs. In Australia, a Japanese-Australian collaboration has built an imaging detector known as CANGAROO. The main element is a 3.5-m-diameter mirror with a 95 pixel camera. In France, there are two detectors operating at the site of the Themis solar facility in the Pyrenees mountains. Each detector has multiple mirrors to make use of the sampling method. The ASGAT experiment consists of an array of seven mirrors (each 4 m diameter) and THEMISTOCLE is an array of 18 smaller mirrors (each 0.8 m diameter). The mirrors in ASGAT are equipped with an array of PMTs while those in THEMISTOCLE have a single photomultiplier tube.

Current atmospheric Cherenkov telescopes have typical threshold energies for gamma-ray detection between 0.25 and 2 TeV. The threshold energy is set by the minimal density of Cherenkov radiation needed to trigger the telescope. A relatively straightforward analysis [4] shows that this threshold energy is proportional to the following:

$$E(th) = \mathrm{sqrt}(B\ \Omega\ t / A\ e)$$

where B is the background light level, Omega is the solid angle viewed by each photomultiplier tube, t is the integration time, A is the mirror area, and e is the efficiency of light collection. In the quest for lower energy thresholds, the parameter that is most straightforward to control is the mirror area. Increasing the size of the mirror collection area lowers the energy threshold of the detector.

2.2. Air Shower Arrays:

Air shower arrays are sampling detectors consisting of a number of elements (e.g., scintillation counters) distributed on the ground. The elements typically comprise an active collection area of one percent of the surface area. Relative timing measurements between elements are used to determine the shower direction. The angular resolution is typically 0.5 degrees, depending on the shower size. Pulse-height measurements are used to estimate the number of charged particles in the shower and hence the primary particle energy.

Air shower arrays have several advantages over atmospheric Cherenkov detectors. Cherenkov telescopes can only operate on clear moonless nights; arrays have close to 100% live-time. The field-of-view of a Cherenkov detector is small, typically a few degrees, while an air shower array views roughly 1 steradian of the overhead sky. On the other hand, air showers arrays require the shower particles to reach the earth's surface. A Cherenkov detector is sensitive to photons produced high in the atmosphere, near shower maximum. This difference leads to significantly higher energy thresholds for air shower arrays. Typical air shower arrays at mountain altitudes have gamma-ray energy thresholds of 50-100 TeV.

There is another important consequence to the fact that surface arrays detect the charged particles that are mostly created in the radiation length right above them (as opposed to the Cherenkov photons created near shower maximum). The lateral and angular distributions of the charged particles contain less information in regards to the nature of the primary particle compared to the same distributions of the Cherenkov light. This effect makes it difficult for air shower arrays to achieve the level of background rejection obtainable with Cherenkov detectors. In essence air shower arrays measure the leakage energy out the back of a thick calorimeter of 30 radiation lengths.

We do expect showers initiated by gamma-rays to have significantly fewer muons (by about a factor of 30) than those initiated by hadrons. This expectation comes from the fact that gamma-ray showers are almost completely electromagnetic in nature. Recent measurements at the HERA electron-proton collider at DESY in Hamburg, Germany, have confirmed this property of high energy photons, where the cross section for photo-pion production is found to be much smaller than that for electron-positron pair production (Bethe-Heitler). Several recent experiments (CASA-MIA, HEGRA) have built extensive muon detectors to aid in background rejection. The experiments typically reject 90% of the hadronic cosmic rays.

Motivated in part by reports in the mid 1980s of bright gamma-ray sources at 1000 TeV, a number of second generation air shower arrays were recently constructed for the purposes of gamma-ray astronomy. These arrays are significantly larger, have lower energy thresholds, and have better angular resolution than experiments of the earlier generation. The foremost of the recent arrays are CYGNUS, Tibet, HEGRA, and CASA-MIA. CYGNUS, located at Los Alamos National Laboratory, NM, consists of roughly 200 scintillation detectors covering an area of 70,000 m^2. The experiment has been taking data and will be decomissioned and moved to surround the future MILAGRO detector. The Tibet array consists of 49 detector stations on a 15 m and is located in Yangbajing, Tibet, at an altitude of 4300 m. The high altitude of the experiment leads to an energy threshold of 10 TeV which is the lowest of any existing air shower array. HEGRA, located on the Canary Island of La Palma, consists of a variety of detector elements. The charged particle array consists of 219 scintillation stations and encloses an area 32,400 m^2. Within the same area is an array of wide angle Cherenkov detectors. There are also muon detectors with a combined area of 300 m^2. Located in Dugway, UT, CASA-MIA is the largest air shower array operating in this energy region. CASA is a surface detector consisting of 1056 scintillation stations and covering an area of 220,000 m^2. MIA consists of 2400 m^2 of buried scintillator spread out over the same area as CASA and sensitive to muons with energies greater than 0.7 GeV. In addition to measuring the muon and electron components of air showers, images of the longitudinal Cherenkov profile are recorded by a detector (DICE) consisting of two stations, each containing a bank of 256 PMTs viewing a single mirror.

3. Recent Results

3.1. Very High Energy:

Very High Energy (VHE) gamma-rays are defined as those near 1 TeV in energy. Historically, this energy range has been accessible only to atmospheric Cherenkov telescopes, but air shower arrays (e.g., Tibet, MILAGRO) are beginning to encroach. There are now three steady sources of VHE gamma-rays. The three sources have also been detected at lower energies by EGRET. The Crab and PSR 1706-44 are spin-down pulsars located in the galactic plane. The third source (Markarian 421) is an extragalactic source of the AGN type.

The Crab, first unambiguously detected by the Whipple Observatory [5], has now been seen by at least five other Cherenkov telescopes. The TeV Crab emission is observed to be unpulsed, which is somewhat of a surprise because at EGRET energies the dominant emission is pulsed. This fact supports the contention that there must be two components that contribute to the gamma-ray emission at high energies. A model has been proposed in which the pulsed component is associated with synchrotron radiation near the pulsar itself, while the TeV emission comes from inverse-Compton scattering in the nebula [6]. This Self-Synchrotron-Compton (SSC) model is quite successful at describing the high energy emission from the Crab; however it is not overly constrained in the region

between 20 GeV and 250 GeV, where no observations have been made. Figure 3 shows some of the latest gamma-ray observations of the Crab spectrum.

There are problems of consistency in the experimental data on the Crab. The Whipple group has recently lowered their flux estimate from their earlier paper. Their new flux [7] disagrees with measurements made by THEMISTOCLE [8]. On the other hand, the Tibet air shower array does not detect the Crab and sets an upper limit [9] which may be in contradiction with the detection level of THEMISTOCLE. One concludes that although that Crab has been convincingly detected at TeV energies, its energy spectrum is not completely understood.

Markarian 421 is an AGN of the Bl-Lac variety. Located at a redshift of 0.031, it is the closest AGN in the EGRET catalog and also the weakest AGN source detected by EGRET. Nevertheless, the discovery by Whipple in 1992 [10] of TeV photons from this object was an important one. Markarian 421 gave proof that there was a source other than the Crab at these energies. The fact that an extra-galactic object some 400 million light-years away could emit TeV gamma-rays detectable on Earth is remarkable. Figure 4 shows the differential energy spectrum of Markarian 421 from MeV to TeV energies. The spectrum is remarkably well fit by a single power law of form E^{-2.06}. More recent data from Whipple [11] may not support this simple fit quite as well.

AGN are assumed to consist of a massive, compact central source, which is presumed to be a super-massive black hole. The central source is surrounded by an accretion disk and jets of material are often seen along the axis of the disk. In most models, gamma-rays are produced by shock acceleration in the relativistic jets. The acceleration is powered by the infall of matter from the accretion disk into the black hole. In the simplest picture, the AGN observed by EGRET are ones having associated jets that are beamed towards the observer.

A very exciting recent result is the detection by Whipple of a flare from Markarian 421 in which the flux increased by about an order of magnitude in a couple of days [12]. The flare was also seen by an x-ray satellite (ASCA) at approximately the same time [13]. The flare detection indicates that the AGN is the source of violent activity on short time scales. Similar short duration outbursts from other AGN were seen by EGRET, which argues for compact acceleration regions.

3.2. Ultra High Energy (> 10 TeV):

Gamma-rays above 10 TeV are in the Ultra High Energy (UHE) region. This region is the domain of air shower arrays that historically operated at energies near 10^{15} eV (1 PeV). The reports of bright UHE emission from x-ray binary systems like Cygnus X-3 and Hercules X-1 stimulated the construction of a new generation of air shower arrays in the late 1980s. The apparent brightness of Cygnus X-3 motivated acceleration models in which the origin of cosmic rays could be explained by a few powerful sources. Unfortunately, the earlier results have not been confirmed; the new experiments have been unable to see steady emission from any source, including x-ray binaries.

Figure 5 shows some current limits on UHE gamma-ray emission from Cygnus X-3. Also shown are claimed detections of this object by earlier experiments. The current limits are up to two orders of magnitude lower than the earlier detections. In principle, the source (in this case Cygnus X-3) could have shut off its gamma-ray emission in the intervening few years. However, since the earlier detections were marginal in significance (typically four to five standard deviations) and seemed to come and go, it is much more likely that these earlier ``detections" were in fact statistical fluctuations at the limit of the experimental sensitivity. Cygnus X-3 has been also ruled out by Whipple as a gamma-ray source at TeV energies.

Air shower experiments have also searched for emission from other sources, including the Crab and Markarian 421 [14-17]. So far, no steady state emission has been detected and the upper limits to the integral photon flux are $10^{-14}/cm^2/sec$ or lower. These limits represent sensitivities comparable to those achieved by Cherenkov telescopes. In Figure 3, for example, upper limits from air shower experiments approach, and in some cases go below, the flux extrapolated from Cherenkov detectors. The UHE limits indicate that either the source spectra roll over at high energies or the photons above 10 TeV are aborbed on their way to Earth. In the case of a source like the Crab, current models indicate that the first possibility is more likely. For extragalactic objects like Markarian 421, there is good reason to expect absorption to be a factor, as we will see in the following section.

One very likely astrophysical source of ultra high energy photons is diffuse gamma-ray emission from pi-zero decay in the galactic plane. Detecting diffuse galactic emission is important because it could signal locations of enhanced cosmic ray production. Measurements of diffuse emission can only realistically be made by experiments having some sort of gamma-ray identification. Unfortunately, Cherenkov telescopes, with narrow fields-of-view, are poorly suited for such measurements. Air shower arrays withmuon detectors identify gamma-rays by the the paucity of muons in their showers. The current limit on the ratio of the diffuse gamma-ray flux to the cosmic ray flux in the direction of the galactic plane is 1.0×10^{-4} [18]. There may also be an isotropic gamma-ray flux that comes from cascading processes resulting from the interaction of extremely high energy cosmic rays with the 3K microwave background radiation. This flux, being omnidirectional, is more difficult to detect than the galactic radiation, but we should see sensitive limits from a couple of experiments (CASA-MIA, HEGRA) in the near future.

3.3. Intergalactic Absorption:

The EGRET experiment has detected high energy gamma-ray emission from over 100 point sources. At very high energies, Cherenkov telescopes have detected three steady sources. At ultra high energies, air shower experiments have failed to detect any steady sources. The three types of experiments have approximately the same sensitivity for sources with a differential power law spectral index of 2.0. This paucity of sources seen by experiments at high energies could be explained by a number of possibilities:

(1) The source production mechanism leads to a cut-off in gamma-ray emission above some energy, or gamma-rays are produced at the source but are absorbed there.

(2) The sources are episodic. This could explain why Cherenkov telescopes with narrow fields-of-view fail to see a large number of sources, but it would not explain why full-aperture air shower arrays do not see many sources.

(3) The gamma-rays are absorbed on their way to Earth.

We know that the first possibility is very likely to be occurring for some sources. There are many sources detected by EGRET that have steep spectra that would make detection at higher energies difficult. The self-synchrotron Compton model for the Crab emission predicts a rapid fall-off in the spectrum above 10 TeV. The second possibility is not likely to be an important factor for most sources.

The possibility that gamma-rays are absorbed as they cross intergalactic space is motivated by the AGN data. EGRET has detected emission from forty AGN at energies up to 10 GeV. Cherenkov telescopes have searched for emission from more than half of these AGN and, except in the case of Markarian 421, have failed to detect any emission. For example, the quasar 3C279 was seen by EGRET to flare to a level thirty times that of Markarian 421, but 3C279 is not seen at TeV energies. Since 3C279 is much further away than Markarian 421 (redshift of 0.54 as opposed to 0.031), it is natural to suggest that photons from 3C279 are being absorbed on their journey to Earth [19].

The dominant absorption process for TeV photons is pair production of soft photons in the infrared region. Diffuse infrared radiation is created by normal galaxy formation. At the present time, the infrared radiation field has not been well measured, but recent estimations place it at approximately the right level to be significant in the absorption of TeV photons from distant sources like 3C279.

It has been pointed out that high energy gamma-ray astronomy could make use of this absorption process to probe the nature of the intergalactic radiation field. Ideally, one wants to measure the spectral features of several sources at different redshift values. Since the amount of absorption depends on the infrared photon density and the distance to the source, the observation of spectral breaks that are correlated with source distance could be used to estimate the infrared photon density. Conversely, if the infrared density is measured directly by other experiments, the gamma-ray spectra can be used to infer the distance scale and possibly measure the Hubble constant [20]. Recent calculations have sought to model the formation of galaxies in order to estimate the diffuse infrared photon density from first principles [21]. These calculations assume that the uncertainty in the photon density is dominated by the epoch at which galaxies begin to form. Therefore, high energy gamma-ray observations could be used to probe different models

of galaxy formation that depend, for example, on the relative fraction of cold and hot dark matter in the Universe.

4. Future Directions

With the successes of atmospheric Cherenkov telescopes like Whipple, ground-based gamma-ray astronomy has come into its own. The exciting results from EGRET motivate better instrumentation on the ground to understand what happens to the source spectra at high energies. It is important to realize that detectors on the ground are complementary to those in space. Only in space using direct detection methods will one be able tocontinuously monitor the entire gamma-ray sky. The great sensitivity of ground-based instruments, resulting from enormous collection areas, will be key in making directed observations of weaker sources. In addition, ground-based experiments will be superior for observations at higher energies where the source fluxes will invariably be low. Since the gamma-ray sky above 20 GeV is largely uncharted territory, there is great motivation for the development of new generations of both space and ground-based instruments.

At the 1994 Snowmass Workshop, the future of high energy gamma-ray astronomy was discussed by the community. In the report of the working group on ground-based gamma-ray astronomy [22], three major design goals for new detectors were identified:

1. High sensitivity and good energy resolution.
2. All-sky coverage with high duty-factor.
3. Lowest possible energy threshold.

Experiments with greater sensitivity will hopefully permit the detection of more sources. For strong sources, increased sensitivity will permit flux monitoring on shorter time scales which is important for studies of variability. An experiment combining the imaging and sampling techniques should reach a sensitivity of 0.01 Crab. There are a number of efforts worldwide to develop finer pixel imaging telescopes and detectors using multiple mirrors with imaging capability. The Whipple group is proposing to reduce the pixel size of their imaging camera by increasing the number of PMTs. The Durham group has a significant upgrade underway in Australia. The most ambitious idea is the Telescope Array, shown in Figure 6. This experiment [23] is proposed to consist of 200 telescopes on a 50-m grid. Each telescope would be composed of a 3-m-diameter mirror viewed by 16 multianode photomultiplier tubes (64 channels per tube). The experiment would have a flux sensitivity of 5 mCrab and an energy threshold of 100 GeV. Prototype telescopes have already been built for the dual purpose of detecting extremely high energy cosmic rays via the nitrogen fluorescence technique.

EGRET has shown that many gamma-ray sources are highly variable. In addition, a sensitive all-sky survey has not been done at TeV energies. These facts argue for experiments that have wide field-of-view and high-duty cycle. One of the most promising ideas for experiments of this type is that of the extended water Cherenkov

detector. Such detectors have lower energy thresholds than conventional sampling air shower arrays. The MILAGRO experiment [24], currently under construction in the mountains of New Mexico, is the first water Cherenkov detector. The experiment should be the first air shower experiment to operate near energies of 1 TeV. Although the sensitivity of MILAGRO will not approach that of atmospheric Cherenkov detectors, its high duty factor will make it ideal to search for strong transient phenomena, like gamma-ray bursts.

The energy range between 20 and 250 GeV represents one of the last remaining unexplored sections of the electromagnetic spectrum. We know that there must be interesting observations to be made in this region simply because of the disparity in the number of sources seen by EGRET and the number seen by ground-based instruments. In this ``unopened window", the Crab emission changes from being dominantly pulsed to dominantly unpulsed in nature (Figure 3). AGN spectra may show features of self-absorption above 10 GeV. We may be able to probe the nature of intergalactic radiation fields by examining AGN absorption features at different redshift values. Finally, EGRET has detected emission from gamma-ray bursts up to an energy of 20 GeV. If sensitive observations of bursts could be made up to 200 GeV, we would be in a position to answer the question of distance scale for gamma-ray bursts. These compelling scientific issues argue for the development of instruments in the energy region between 20 and 250 GeV.

One way to open this window is with a new satellite instrument with a much greater energy reach than EGRET. There are design concepts for several such experiments worldwide. The focus of this workshop is the GLAST design concept. A new gamma-ray satellite experiment will be expensive and will take considerable time to fund and construct. Therefore, it is natural to consider what can be done on the ground. As discussed earlier, the energy threshold of atmospheric Cherenkov telescopes can most effectively be lowered by increasing the mirror area for Cherenkov light collection. A Cherenkov detector using a very large mirror could reach energies as low as 10 GeV. The design for such an experiment (Big Bowl) has been recently formulated [25]. It would consist of an enormous spherical mirror (up to 500 min diameter) to reflect Cherenkov light onto a movable camera box. The main advantage of the design is that almost all the information contained in the Cherenkov radiation is preserved so that good low energy response and excellent background rejection can be obtained at the same time. The disadvantage is the long development time and high cost.

Very large mirror arrays have been constructed around the world for solar energy research. In principle, a gamma-ray telescope could be built which would use existing solar heliostats (mirrors) at a fraction of the cost of an experiment that would build mirrors from scratch. Several groups are pursuing the possibility of using solar heliostats to reflect Cherenkov light to detecting elements on a central tower. A French group is conducting tests at the site of the Themis power plant in the Pyrenees [26]. The Themis plant, now decomissioned, consists of 200 heliostats, each having a mirror surface of 50 m^2.

An American-Canadian group has been evaluating the Solar One plant near Barstow, California, for the purposes of gamma-ray astronomy. This facility consists of 1818 heliostats (mirror area 40 m^2 each) and 109 newly installed heliostats (mirror area 75 m^2 each) (Figure 7). The group has conducted tests in 1994 demonstrating that the night sky background conditions are acceptable and that the optical and mechanical characteristics of the heliostats are excellent [27]. In November 1994, the group detected Cherenkov radiation using four heliostats and a large 1 m^2 Fresnel lens as a secondary collector. In 1995, there are plans to image Cherenkov light from as many as ten heliostats using a large 1.7-m parabolic reflector on the central tower. The next step would be to build a prototype detector that would use 45-50 heliostats and that should achieve an energy threshold below 100 GeV.

5. Summary

Space-based gamma-ray instruments run out of sensitivity at high energies because of rapidly falling fluxes. Therefore, we must turn to ground-based experiments for information at the highest energies. These experiments have enormous collection area, but until recent years were unable to detect firm gamma-ray signals because of the charged particle background. In the last five years, ground-based detectors using the atmospheric Cherenkov technique have succeeded in detecting TeV point sources. The frontier now is the energy region between 20-250 GeV where no observations have been made. This region, one of the last unexplored segments of the electromagnetic spectrum, should be finally explored in the next decade by more sensitive space-based experiments (like GLAST) and by ground-based instruments.

References:

[1] C.A. Fichtel et al., ApJ Suppl. 94:551, 1994.

[2] P. Sreekumar et al., Phys. Rev. Lett. 70:127, 1993.

[3] G.J. Fishman et al., ApJ Suppl. 92:229, 1994.

[4] M.F. Cawley and T.C. Weekes, submitted to Ex. Astr, 1994.

[5] G. Vacanti et al., ApJ 377:467, 1991.

[6] O.C. De Jager and A.K. Harding, ApJ 396:161, 1992.

[7] D.A. Lewis et al., Proc. 23rd Int. Cosmic Ray Conf. (Calgary), 1:279, 1993.

[8] P. Baillon et al., Astroparticle Phys., 1:341, 1993.

[9] M. Amenomori et al., Proc. 23rd Int. Cosmic Ray Conf. (Calgary), 1:342, 1993.

[10] M. Punch et al., Nature 358:477, 1992.

[11] G. Mohanty et al., Proc. 23rd Int. Cosmic Ray Conf. (Calgary), 1:440, 1993.

[12] A.D. Kerrick et al., ApJ 438:L59, 1995.

[13] T. Takahashi et al., IAU Circular 6063, 1994.

[14] D.E. Alexandreas et al., ApJ 383:L53, 1991.

[15] M. Amenomori et al., Phys. Rev. Lett. 69:2468, 1992.

[16] T.A. Mckay et al., ApJ 417:742, 1993.

[17] A. Karle et al, MPI-PhE/95-01, submitted to Astroparticle Phys., 1995.

[18] C.E. Covault et al., Proc. 2nd Compton Symposium (College Park), AIP Conf. Proc. 304:499, 1993.

[19] F.W. Stecker at al., ApJ 415:L71, 1992.

[20] M.H. Salamon et al., ApJ 423:L1, 1994.

[21] D. MacMinn and J. Primack, TeV Gamma Ray Astrophysics, ed. H. Volk, Space Science Reviews, in press, 1995.

[22] Richard C. Lamb, Rene A. Ong, Corbin E. Covault, and David A. Smith, Ground Based Gamma-Ray Astronomy, Proc. Snowmass '94, in press.

[23] M. Teshima et al., Proc. Tokyo Workshop on Techniques for the Study of EHE Cosmic Rays, ed. M. Nagano, 109, 1993.

[24] S. Barwick et al., Proc. 23rd Int. Cosmic Ray Conf. (Calgary), 1:346, 1993.

[25] S. Ahlen et al., Nucl. Inst. and Meth., 351A:493, 1994.

[26] E. Pare, Proc. Towards a Major Atmospheric Cherenkov Detector-II (Calgary), ed. R.C. Lamb, 250, 993.

[27] R.A. Ong et al., Proc. Towards a Major Atmospheric Cherenkov Detector-III (Tokyo), ed. T. Kifune, Univeral Academy Press, 295, 1994.

Figure Captions:

Figure 1: Schematic representation of an air shower initiated by a gamma-ray of energy near 1 TeV. The gamma-ray interacts in the upper atmosphere and creates an electromagnetic cascade that reaches its maximal extent at an altitude of 10 km. Charged particles in the cascade radiate a cone of Cherenkov light that can be detected by optical instruments on the ground.

Figure 2: A comparison of images recorded in the focal plane of an atmospheric Cherenkov detector for showers initiated by gamma-rays (left) and by hadronic cosmic rays (right).

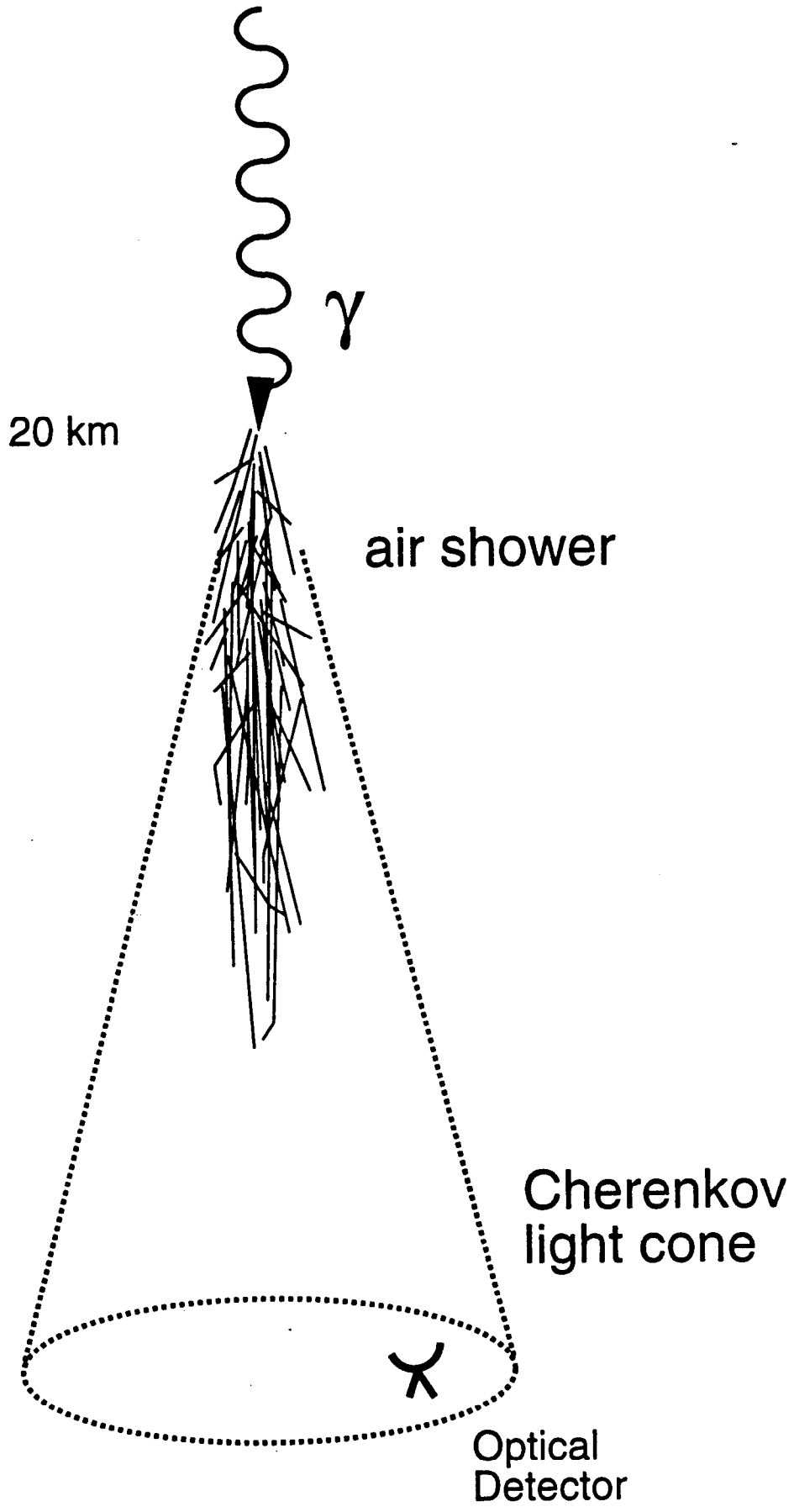
Figure 3: Gamma-ray observations of the Crab at high energies (EGRET), very high energies (Cherenkov telescopes), and at ultra high energies (air shower arrays). The integral gamma-ray flux from the Crab is shown as a function of energy. The Cherenkov telescope results are detections; the air shower array results are upper limits. Shown is the most recent Whipple data [7]. The dashed line is the extrapolated flux from the previous Whipple data [5].

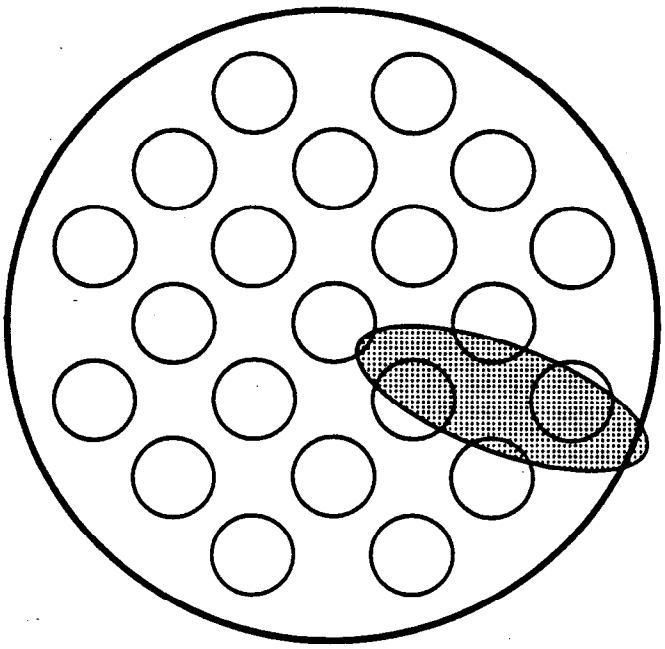
Figure 4: Gamma-ray observations of Markarian 421 at high energies (EGRET) and very high energies (Whipple). The differential gamma-ray flux from Markarian is shown as a function of energy. The dashed line is a single power law fit to the EGRET data.

Figure 5: Gamma-ray observations of Cygnus X-3 at ultra high energies. Upper limits from four air shower arrays are shown. Also shown are claimed detections from earlier experiments in the 1980s (Kiel, Haverah Park). The solid curve is a model proposed by Hillas to explain the bright emission from Cygnus X-3 and the origin of cosmic rays below $10^{(16)}$ eV.

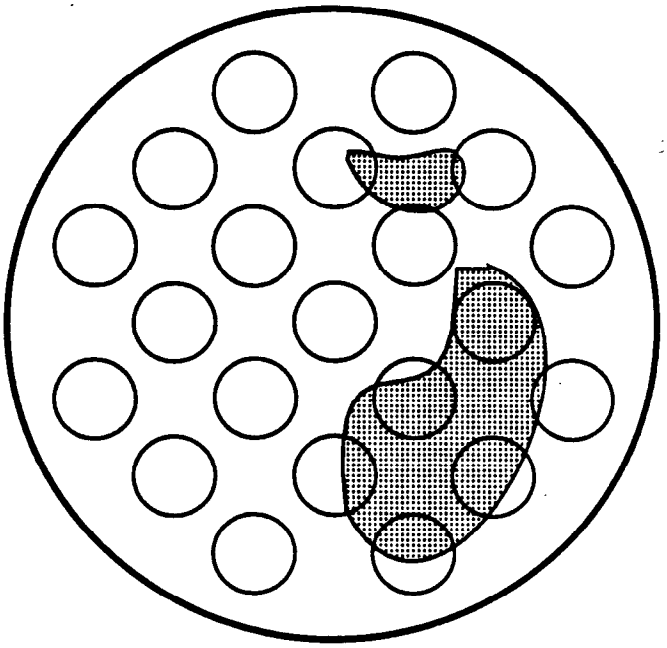
Figure 6: Proposed Telescope Array detector for very high energy gamma-ray astronomy (Japan). The array would consist of 200 telescopes, each having 1024 PMT channels.

Figure 7: Solar-One heliostat array near Barstow, California. Each of the 1818 heliostats has an independent altitude-azimuth and a mirror surface of 40 m^2. Heliostats represented by diamond symbols were used for spot size measurements. Heliostats indicated by rectangles were used to detect Cherenkov radiation in November 1994.

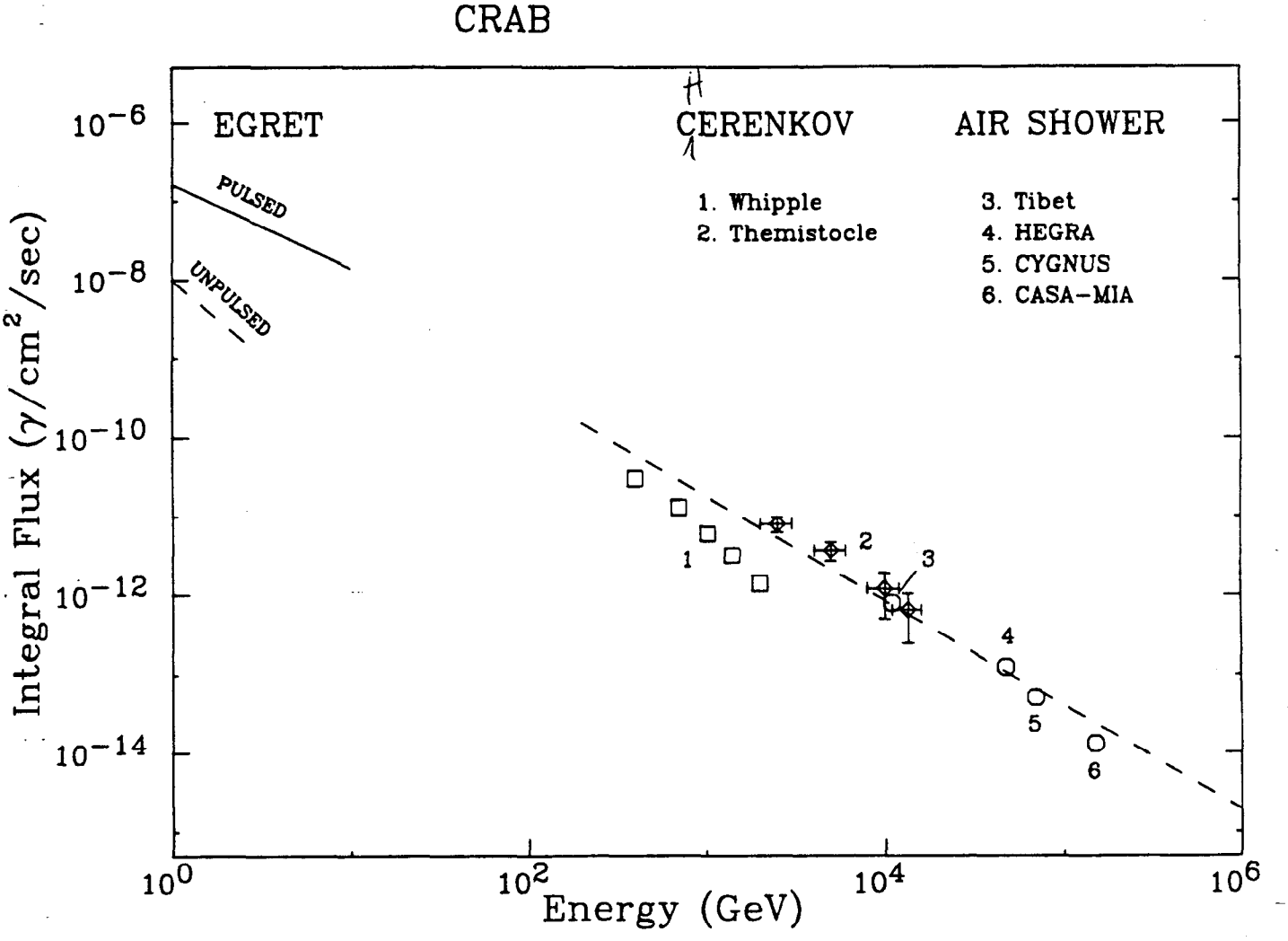

γ
20 km
air shower
Cherenkov
light cone
Optical
Detector

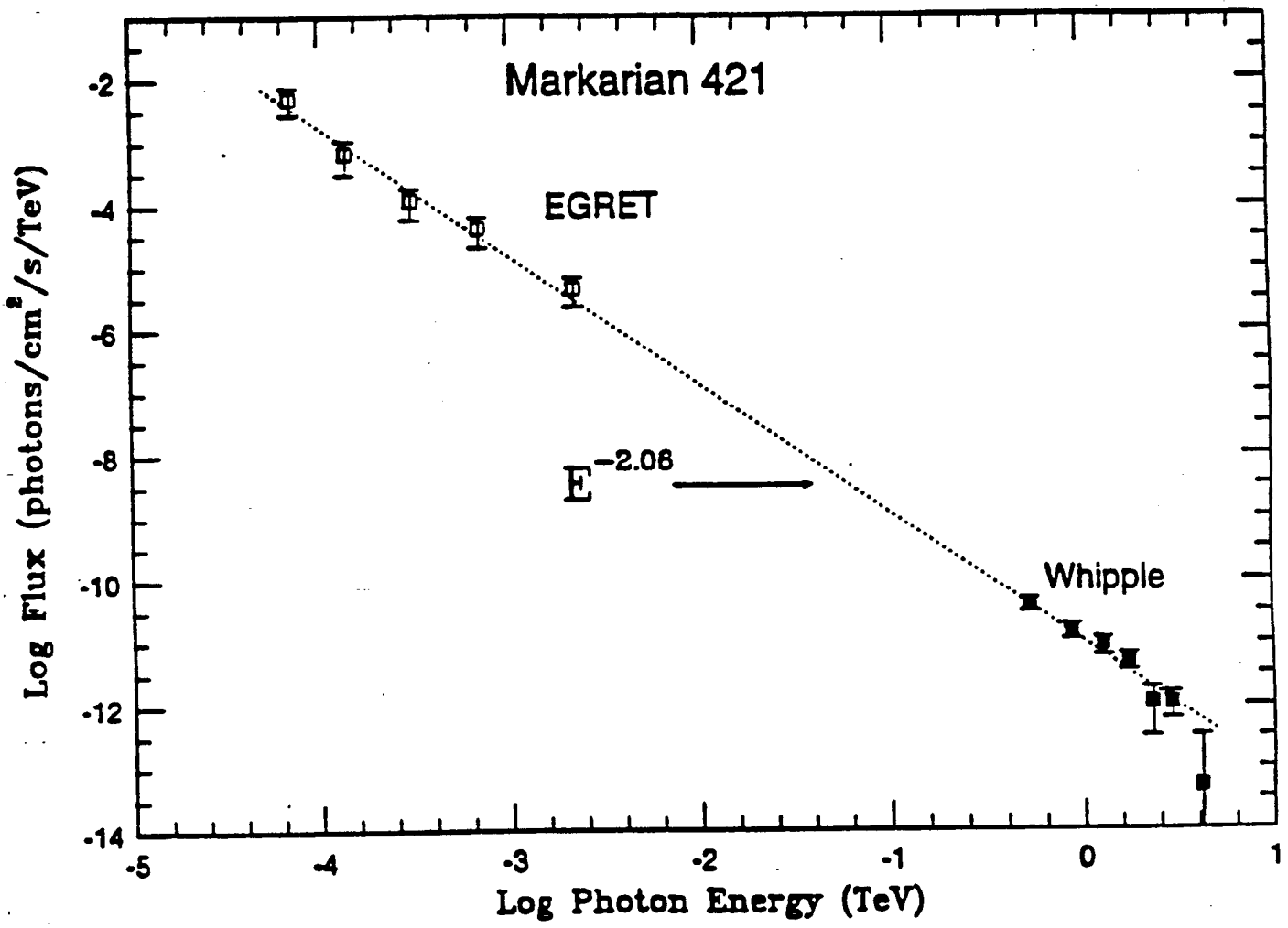


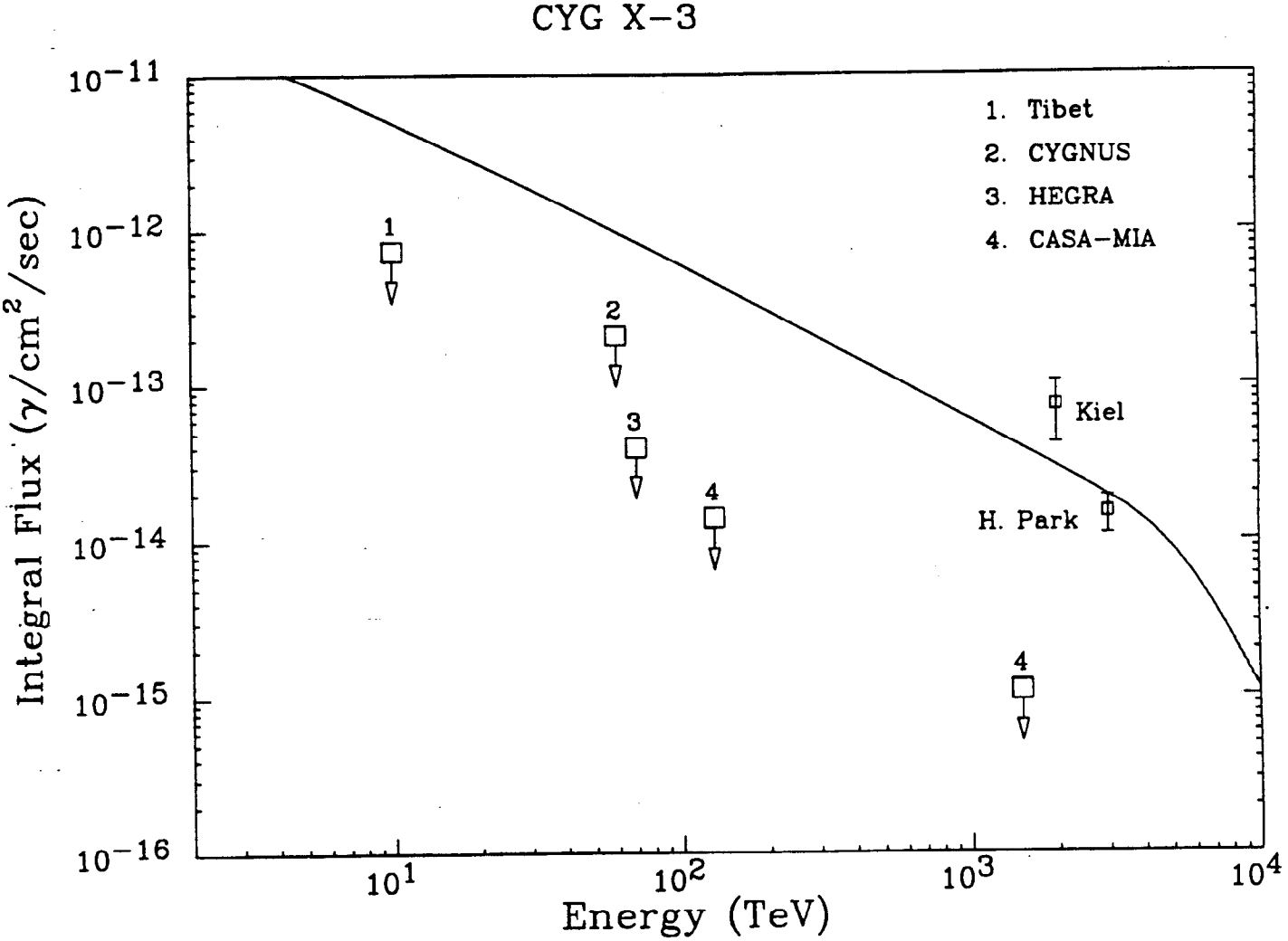
gamma

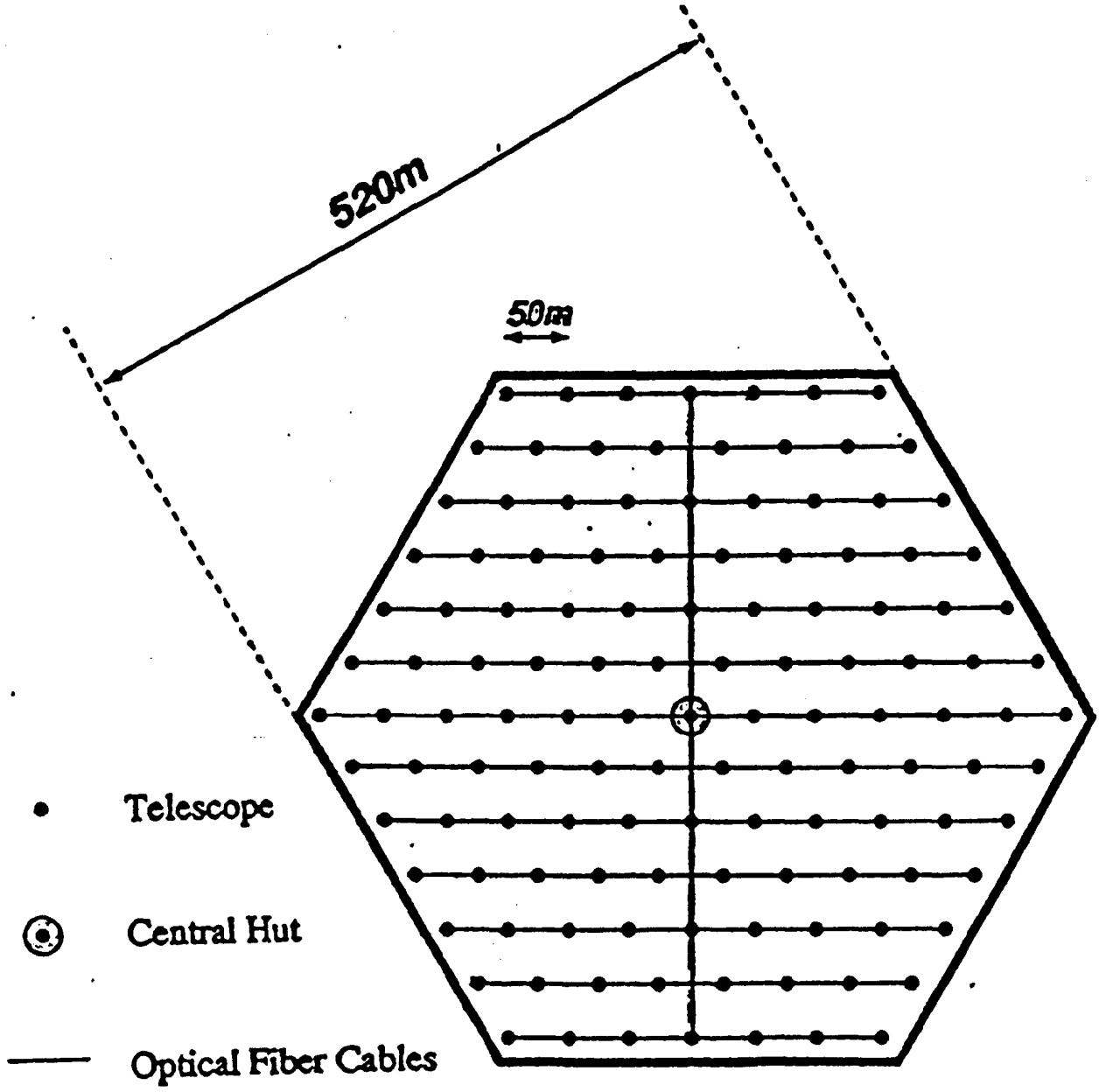


proton

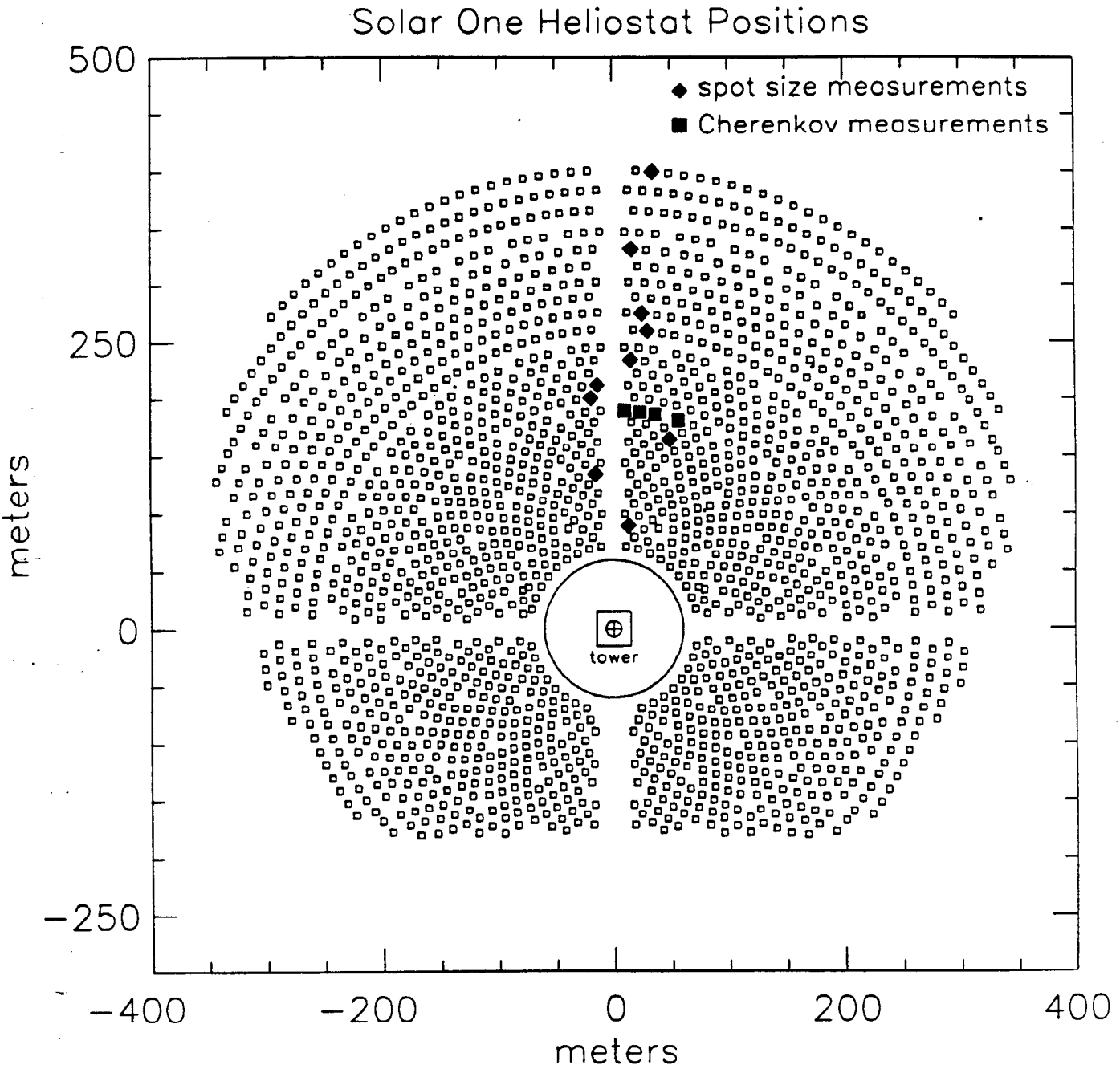

CRAB
EGRET
CERENKOV
AIR SHOWER
PULSED
UNPULSED
1. Whipple
2. Themistocle
3. Tibet
4. HEGRA
5. CYGNUS
6. CASA-MIA
Integral Flux (γ/cm²/sec)
Energy (GeV)
10⁻⁶
10⁻⁸
10⁻¹⁰
10⁻¹²
10⁻¹⁴
10⁰
10²
10⁴
10⁶


Markarian 421
EGRET
$E^{-2.06}$
Whipple
Log Flux (photons/cm^2/s/TeV)
Log Photon Energy (TeV)
-2
-4
-6
-8
-10
-12
-14
-5
-4
-3
-2
-1
0
1


CYG X-3
Integral Flux (γ/cm²/sec)
10⁻¹¹
10⁻¹²
10⁻¹³
10⁻¹⁴
10⁻¹⁵
10⁻¹⁶
10¹
10²
10³
10⁴
Energy (TeV)
1. Tibet
2. CYGNUS
3. HEGRA
4. CASA-MIA
Kiel
H. Park
1
2
3
4
4



TELESCOPE ARRANGEMENT


Solar One Heliostat Positions
spot size measurements
Cherenkov measurements
tower
500
250
0
−250
meters
−400
−200
0
200
400
meters

The Energy Emission Mechanisms of γ-ray Blazars

Ann E. Wehrle
Infrared Processing and Analysis Center
Jet Propulsion Laboratory and California Institute of Technology,
Pasadena, CA 91109

I. Introduction

Our ultimate goals in studying the class of extremely variable quasars known as "blazars" are to understand the nature of the "central engine", how nonthermal jets emerge from this region, and how they propagate and evolve. The low-frequency part of the nonthermal emission (radio through infrared wavelengths) originates in synchrotron radiation from relativistic particles, probably electrons and positrons, moving in the magnetic field near a massive black hole. The nonthermal emission is tightly collimated in two oppositely-directed beams with opening angles of order 10°. If an observer's line of sight falls near or within the collimated beam of emission, the observer sees polarized, highly variable emission: an object observed in this way is called variously, a BL Lac object, an Optically Violent Variable, a Highly Polarized Quasar, or simply, a "blazar". Blazars vary dramatically at all wavelengths; a flare may increase the luminosity at a given wavelength by factors of 2-10 on timescales of days to months. Variability is generally stronger at high frequencies (x-ray, γ-ray) than low frequencies (radio) (eg., Maraschi et al. 1994, Von Montigny et al. 1995). Relativistic beaming both amplifies the radiation along the direction of the jet and compresses the observed timescale of variability. Correlations between variability at high and low energies can reveal much about the formation, propagation and evolution of jets in quasars. The key data are *simultaneous* spectral energy distributions from radio to gamma-rays, ideally at more than one epoch. The availability of infrared, ultraviolet, x-ray and γ-ray data from several active spacecraft allows us the unprecedented opportunity to relate changes in the parsec-scale radio structures, for example, the ejection of a new plasma blob or the brightening of a stationary shock feature, to variability across the spectrum.

γ-ray blazars (see First and Second EGRET catalogs; Fichtel et al. 1994 and Thompson et al. 1995) are particularly interesting observational targets for two reasons: first, their γ-ray luminosity and fast variablility are most challenging to accomodate under jet theories; second, starlight and related thermal processes contribute relatively little to the strongly Doppler-boosted spectral energy distribution, i.e., the *observed* spectral energy distribution of a blazar is predominantly nonthermal continuum emission. Multiwavelength approaches are needed because the most conclusive tests of the jet models involve correlation of time variability in various wavebands, from gamma rays to radio, especially in response to a gamma ray flare.

The majority of models, discussed below, proposed for gamma-ray emission in blazars use inverse-Compton scattering of low-energy seed photons from relativistic electrons to produce gamma rays. A matter of much debate is the origin of the seed photons: either synchrotron photons within the jet, or other photons external to the jet, e.g., from an accretion disk or ambient photon field. Alternative models include electron-proton jets which produce high energy emission through photomeson production (Mannheim and Biermann, 1989, 1992), and decaying ultrarelativistic neutrons emitted near the central

engine which in turn produce radiation by secondary electrons (Kazanas 1993). Specific questions include:

- How far from the supermassive black hole does the gamma-ray emission originate? Electron-positron pair production limits optical depth to gamma-ray emission at the base of the relativistic jet.
- Which blazar characteristics are particularly conducive to high gamma ray emission, e.g., rapidity of variability, superluminal expansion, frequency at which the flat radio spectrum turns over, submillimeter or infrared spectral index, strong X-ray emission, optical polarization, etc.?
- Do gamma ray flares correspond to outbursts at lower frequencies, for example, the 22 and 37 GHz outbursts observed by Valtaoja and Terasranta (1995)? If so, are the flares simultaneous or are there time delays between the different wavebands?

As an example of what may be learned about blazar emission mechanisms, multiwavelength monitoring of 3C279 (at 2 epochs separated by 1 year) showed a dramatic decrease in flux for $\nu \geq 10^{14}$ Hz (Maraschi et al. 1994), with the largest change in gamma-rays (Figure 1). When faint, the gamma-ray luminosity was comparable to the infrared-optical luminosity, while in the bright state it was $\sim$ 30 times greater, indicating that both the scattering electrons and seed photons varied together as in a synchrotron self-Compton model (Maraschi et al. 1994). But this was only one object, with limited time sampling. A critical observing gap in the far infrared (to be filled by the Infrared Space Observatory, scheduled launch in November 1995) occurred where we believe the peak in the seed photon distribution must be located. Compton scattering takes the peak of the low-frequency (infrared) seed photon distribution to the peak of the high-frequency (γ-ray) scattered photons.

II. Background

The power output in the 100 – 1000 MeV range of blazars detected by the Compton Gamma Ray Observatory is comparable to, and in several cases exceeds, that in the rest of the electromagnetic spectrum. Thus, explaining the γ-rays is of key importance for a theoretical understanding of these objects. Various models have been proposed, the majority favoring the inverse Compton mechanism for γ-ray production (for reviews, see Sikora 1993, Maraschi, Ghisellini and Celotti, 1994, Marscher and Bloom 1994). This is natural since the spectrum from radio to ultraviolet frequencies is commonly interpreted as synchrotron radiation from high energy electrons in a jet with relativistic bulk flow (e.g. Königl 1989). The same electrons can produce gamma-rays by upscattering either the synchrotron photons (Maraschi, Ghisellini and Celotti 1992, Marscher and Bloom 1992) or other photons external to the jet (Dermer, Schlickeiser and Mastichiadis 1992, Blandford 1993, Sikora, Begelman and Rees 1994, Coppi, Kartje and Konigl 1993, Dermer and Schlikeiser 1994). In both cases a strong correlation between the synchrotron and inverse-Compton emission is predicted, since they derive from the same relativistic electrons. The exact quantitative relation depends on the specific model, and the study of the simultaneous variations of the synchrotron (millimeter – ultraviolet – soft X-rays) and inverse-Compton (X-rays – γ-rays) components offers a unique opportunity to constrain the models.

Fast variability, of order a few days (Kniffen et al. 1993, von Montigny et al. 1995) implies that the emitting regions are highly compact, and suggest that the gamma rays originate near the base of the relativistic jet (within the "core") or in shocks located within the jet, possibly the superluminal "knots" observed with Very Long Baseline Interferometry (VLBI). Explaining the origin of the gamma-ray continuum is therefore critical to understanding blazars, and by extension the central engines of all AGN.

III. Monitoring Programs on γ-ray Blazars

Two well-known blazars, 3C273 and 3C279, happen to lie within a few degrees of each other such that they can be conveniently viewed by EGRET in both "wide field of view" and "narrow field of view" observing modes. This fortuitous proximity enables spacecraft and groundbased observatories to efficiently concentrate observing time during the few weeks of the year (December/January and May/June) when the pair are observable at night from the ground and when the sources are accessible by spacecraft that must avoid pointing too near or too far from the sun's direction (e.g., to keep the sun away from the science instrumentation or to keep the solar power panels illuminated). In 1992/1993, two extensive campaigns were undertaken to observe 3C273 and 3C279 (respectively, Lichti et al. 1995 and Maraschi et al. 1994). The participants obtained ultraviolet observations with IUE, X-Ray observations with ROSAT, and γ-ray observations with CGRO, as well as flux measurements in other wavebands. Three space-based and several dozen groundbased observatories cooperated in arranging the 3C279 and 3C273 campaigns. These data formed almost-simultaneous spectral energy distributions which we have compared with predictions of relativistic jet models for 3C279 (Maraschi *et al.* 1994), and which various other authors have done for 3C273.

A multi-object multifrequency campaign led by Hartman (of the EGRET Instrument Team) with 63 co-investigators began in the 1993-1994 CGRO observing cycle; it has targets continued over to be observed in 1995. Related programs led by Michelson and Fichtel of the EGRET Instrument Team form a deep exposure (8 weeks) of the Virgo Region, including 3C279 and 3C273, both of which have been strongly detected by EGRET at various times during the CGRO mission. A very bright blazar can be detected in a day or two, most solid 5σ detections involve EGRET integration times of one or two weeks. COMPTEL and OSSE, lower energy instruments on CGRO, have less sensitivity to blazars; most of their AGN detections are Seyfert galaxies.

The EGRET-detected blazars are ranked among the strongest flat-spectrum radio sources in the sky. Many have been monitored for years with optical and radio telescopes (e.g., Aller et al. 1985, Steppe et al. 1994; Robson et al. 1993; Stevens et al. 1994); they are also secondary calibrators at the millimeter observatories IRAM (Spain), Nobeyama (Japan), SEST (Chile), and Owens Valley Radio Observatory (California) as well as the Very Large Array (New Mexico). The radio spectra appear flat because the radio sources are composed of several compact "blobs" which have different turnover frequencies in the range 1-40 GHz; the superposition of emission from various blobs adds up to a fairly flat spectrum. Below the turnover frequency, the spectrum of the radio-emitting blobs is "self-absorbed", namely, the blobs become optically thick with spectral indices of 2.5. Typical sizes, as measured with VLBI are roughly a milliarcsecond, which translates to about 4 parsecs (13 light years) at redshifts of 0.5-1.0. A useful summary of relationships between

VLBI-observed and inferred physical quantities can be found in Pearson and Zensus (1987). VLBI images for a good fraction have been made or are in progress (eg., Wehrle et al. 1992 and references therein, Pearson and Readhead 1988; Marscher, Moore, Xu and Wehrle in progress).

Radio emission from compact jet and core components becomes optically thin in the millimeter/submillimeter part of the spectrum, such that infrared and submm fluxes vary simultaneously (Robson et al. 1993; Maraschi et al. 1994 and references therein). Millimeter and submillimeter monitoring is especially difficult because of water absorption in the atmosphere; clear, dry weather on Mauna Kea in Hawaii (or in the Andes in Chile) is required to observe in the few bands where the atmosphere is relatively transparent.

IV. Results from the 3C279 Campaign in 1992-1993

A. Summary of Results

In establishing goals for long-term, sensitive monitoring of AGN by a future γ-ray mission, it is useful to examine what has been learned in current multiwavelength work on the bright γ-ray blazar, 3C 279. Simultaneous observations of 3C 279 at radio, millimeter, near-infrared, optical, ultraviolet (with IUE) and X-ray (with ROSAT) wavelengths were obtained in December 1992-January 1993, during a three-week pointing at the source by the Compton Gamma Ray Observatory. Preliminary results, briefly described here, are presented in Maraschi *et al.* (1994). The blazar was in a quiescent or "low" state during this period. Comparing the multiwavelength energy distribution to that from June 1991, when 3C279 was in its brightest recorded γ-ray state, we found that:

1) 3C 279 faded dramatically at all frequencies above 10^{14} Hz, while the flux variations at low frequencies (radio-mm) were minor.

2) The near-infrared-optical-ultraviolet spectral shape was softer (steeper) in the quiescent state, and the X-ray spectra also appear softer, although the spectral index measured by ROSAT refers to a lower energy band than that measured earlier with Ginga.

3) The ratio of the γ-ray luminosity to that across all other frequencies decreased from a value of $\simeq 10$ in the flaring state to a value $\simeq 1$ in the quiescent state.

These findings imply that the production of γ-rays is closely related to the optical-ultraviolet continuum, in agreement with models where γ-rays are produced through inverse Compton scattering by relativistic electrons emitting the synchrotron continuum. To explain the observed non-linear relation between the synchrotron and inverse Compton fluxes requires both a change in the electron spectrum *and* an associated change in the seed photons.

B. Discussion of Implications and Constraints on Theoretical Models

The seed photons to be upscattered could be the synchrotron photons themselves, as in the synchrotron self-Compton model (Maraschi, Ghisellini, & Celotti, 1992; Bloom & Marscher, 1993), or photons external to the jet. These may derive from an accretion disk (Dermer, Schlickeiser & Mastichiadis,1992, Dermer & Schlickeiser 1993) and could possibly be isotropized by scattering in a hot intercloud medium and/or may be emitted by the broad-line region itself (Blandford 1993; Sikora, Begelman, & Rees 1994). The relative importance of the two types of seed photons (synchrotron or external), estimated on the basis of observed quantities, depends on a high power of the beaming factor ($\sim \delta^5$; Sikora,

Begelman, & Rees 1994). However, in neither case is it easy to reproduce accurately the observed energy distributions with a simple homogeneous model; multiple components or an inhomogeneous jet are required.

A detailed fit to the data with multi-parameter models is being made by Maraschi et al., but based on the extraordinary observed variability, some general remarks can be made. In the case of the simplest synchrotron self-Compton model, the most important photons for up-scattering are those at the peak of the synchrotron emission (i.e., 10^{13-14} Hz). The similarity of the spectral shape of the synchrotron and inverse Compton components suggests that these photons are upscattered to the γ-ray band (10^{22-23} Hz), which requires a relatively high electron Lorentz factor, $\gamma \sim 10^4$, and a rather low magnetic field $B = 0.3(\frac{\gamma}{10^4})^{-2}\delta^{-1}$ Gauss.

External seed photons, even if isotropic in the rest frame of the active galactic nucleus, would appear blue-shifted and therefore enhanced in the jet frame. The typical frequency of accretion disk or broad-line photons as seen from the jet would be in the EUV/soft-X-ray band, thus providing an inverse-Compton γ-ray component peaking at higher energies than in the case of the synchrotron self-Compton mechanism (for the same mean electron Lorentz factor) and/or allowing for higher values of the magnetic field.

A general property of the synchrotron self-Compton model is that when the electron distribution decreases in number and/or average energy, the synchrotron luminosity decreases proportionately while the inverse-Compton (gamma-ray) luminosity decreases quadratically (eg., Maraschi, Ghisellini and Cellotti 1992). If instead the photon energy density as seen in the jet frame is dominated by external photons, when the electron spectrum varies, the synchrotron and inverse-Compton emission components will both scale proportionately, unless the external photon field varies at the same time and in the same sense as the electron spectrum.

The energy distributions of 3C279 in high and low state are qualitatively consistent with the synchrotron self-Compton model, since the γ-ray luminosity, assumed to be representative of the inverse Compton luminosity, varies more than the synchrotron luminosity, peaking presumably in the far infrared range. However, one can still salvage the external photon models fairly naturally by postulating that the synchrotron decrease is caused by a change in the *bulk* Lorentz factor of the radiating electrons. In that case, the external photon field *as seen by the relativistic electrons* also decreases, and if at the same time the spectrum of relativistic electrons in the jet decreases, the resulting variation of the γ-ray flux is larger than that of the synchrotron one. A decrease of δ by a factor of 2 would account for the observed variation in the broad-band energy distribution of 3C279.

It is interesting to recall that a steepening of the electron distribution is a natural consequence of a diminishing strength of a shock, while a flattening of the electron spectrum is expected if a shock gets stronger or a new shock forms (Kirk and Schneider 1987, Schneider and Kirk 1987). Moreover, the evolution of a relativistic shock may offer a physical framework, whereby the relativistic amplification factor for the observed flux (more complicated than a simple Doppler correction, Lind & Blandford 1985, Celotti, Maraschi & Treves, 1991) could be related to the shock strength. Thus shocks in a jet can qualitatively explain the observed variation in L_γ/L_{bol}. If external photons dominate the inverse Compton mechanism the shock front should accelerate or slow down substantially

when getting respectively stronger or weaker.

The "observables" are related to physical jet characteristics as follows. If the seed photons are synchrotron photons from the jet, then if they increase by a factor X due to an increase in the number of energetic electrons, the γ-ray flux will increase as X^2 (Maraschi et al. 1994). If the seed photons are external to the jet, they may or may not be directly observable; in our frame, they could be swamped by the beamed synchroton component yet because of Doppler beaming still appear dominant in the jet frame (Sikora, Begelman and Rees 1993). An increase in synchrotron electrons would then cause equal increases in γ-ray and synchrotron components, i.e., gamma-ray flux is proportional to X. Finally, if the external seed photons change, which will likely be correlated with change in isotropic emission (e.g., emission lines), the γ-rays should change while the observed synchrotron flux remains constant. There are no predictions of which blazar will flare. Ideally, the future γ-ray telescope would be able to observe a large number of potentially flaring blazars *simultaneously* with high sensitivity, such that *quiescent* levels of emission as well as rising and decaying flares could be observed.

Conclusions

The multiwavelength γ-ray blazar studies outlined above require a large number of objects in order to sample the range of magnetic fields, intrinsic luminosities, orientation angles, electron energy spectra, Doppler factors, source environments (narrow line clouds, accretion disks), and the possibility of observing various stages of activity (flaring, quiescent, incipient and fading). A tremendous amount of spacecraft and ground-based observing time will be required to understand the influence of all the factors above. Multiplexing the γ-ray targets with a very wide field of view on a new γ-ray telescope maximizes the opportunities to identify and track the evolution of flaring blazars.

Acknowledgements

This paper is based on collaborative work, principally with R. Hartman, L. Maraschi, A. Marscher, P. Michelson, and C. M. Urry. The author is grateful to C. Fichtel and the EGRET team for cooperation. She acknowledges support from the NASA Compton Observatory Guest Investigator Program, NAS7-1260.

References

Aller, H., Aller, M., Latimer, G., and Hodge, P. 1985, ApJS 59, 513

Blandford, R.D. 1993, in Compton Gamma-Ray Science Workshop, ed. Friedlander et al., p. 533

Bloom, S. and Marscher, A.P. 1993 in Proc. of Second Compton Symposium (Maryland), in press.

Bregman,J. 1990, A&AR 2, 125

Coppi, P., Kartje, and Konigl, A. 1993 in Proc. of Second Compton Symposium (Maryland), in press.

Dermer, C., Schlickeiser, R. and Mastiachiadis, A. 1992, A&A 256, L27

Dermer, C. and Schlickeiser, R. 1994 ApJS 90, 945

Fichtel, C., Bertsch, D., Chiang, J. et al. 1994 ApJS 94, 551

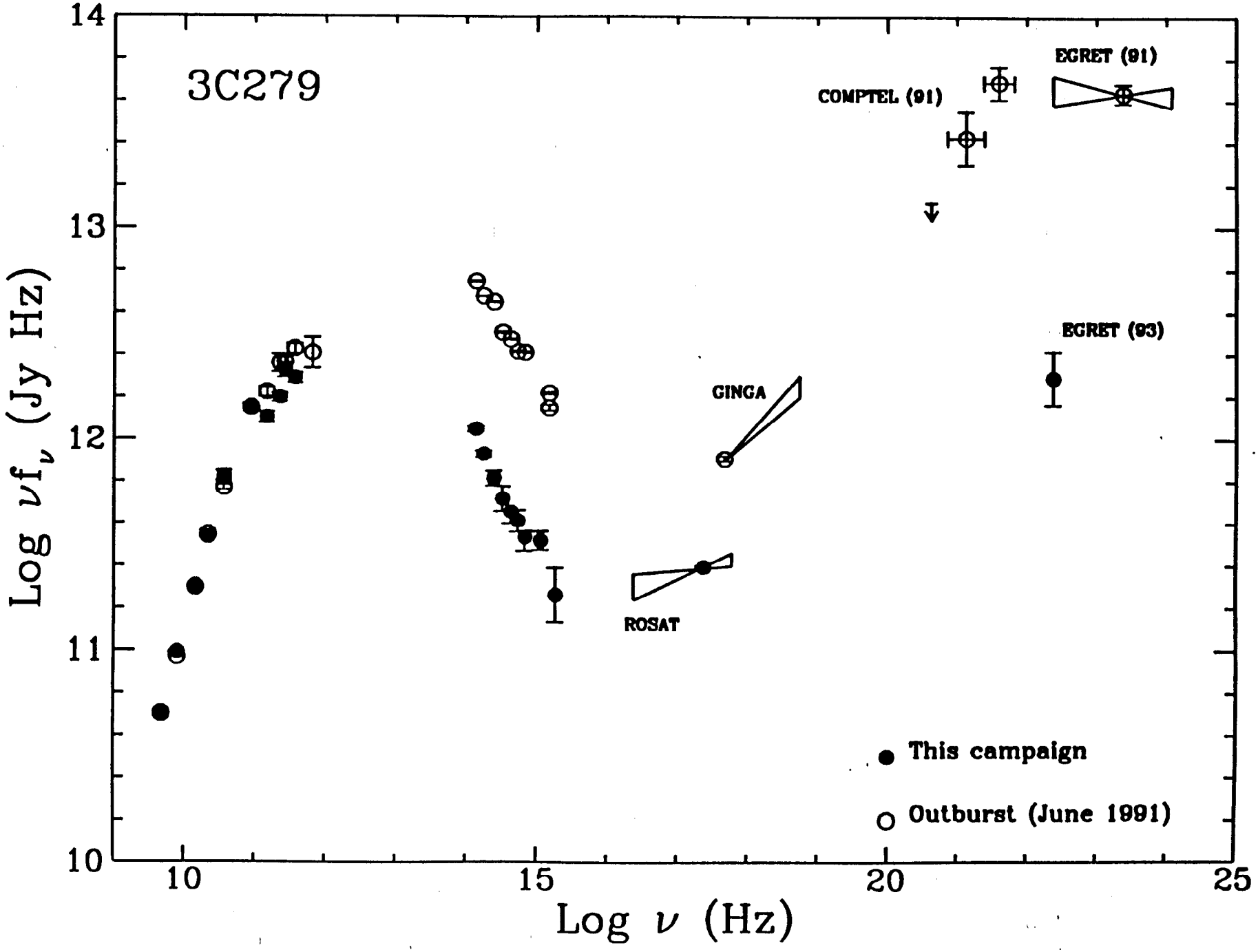
Hartman, R.C., Bertsch, D.L., Fichtel, C.E. et al. 1992 ApJ 385, 1

Hartman, R.C. 1993, in Proc. of Second Compton Symposium (Maryland), in press.

Impey, C. and Neugebauer, G. 1988 AJ 95, 307
Kazanas, D. 1993 in Proc. of Second Compton Symposium (Maryland), in press.
Proc. of Second Compton Symposium (Maryland), in press.
Kirk, J. and Schneider, J. 1987 ApJ 322, 256
Konigl, A. 1989, in "BL Lac Objects", eds. Maraschi et al., p. 321
Kniffen, D.A., Bertsch, D.L., Fichtel, C.E. et al. 1993 ApJ 411, 133.
Lichti, G., Balonek, T., Courvoisier, T.J.-L. et al. 1995 A&A 298, 711.
Mannheim, K. and Biermann, P.L. 1989 A &A 221, 211.
Mannheim, K. and Biermann, P.L. 1992, A&A 253, 21.
Maraschi, L., Ghisellini, G. and Celotti, L. 1992, ApJ 397, L5
Maraschi, L., Grandi, P., Urry, M., et al. 1994, ApJ 435, L91
Marscher, A. and Bloom, S. 1992 in The Compton Observatory Science Workshop, ed. C. Shrader, N. Gehrels and B. Dennis (NASA Conf. Publ. 3137), 346.
Marscher, A. and Bloom, S. 1994 in Proc. of Second Compton Symposium (Maryland), in press.
Pearson, T. and Readhead, A. 1988 ApJ 328, 114
Pearson, T. J. and Zensus, J.A. 1987, in "Superluminal Radio Sources", ed. J.A. Zensus and T.J. Pearson, (Cambridge University Press: Cambridge, p. 1.
Robson, I. et al. 1993, MNRAS, 262, 249
Schneider, P. and Kirk, J. 1987 ApJ 323, L87
Sikora, M. in Proc. of Second Compton Symposium (Maryland 9/93), in press.
Sikora, M., Begelman, M., and Rees, M. 1994, ApJ 421, 153
Steppe, H. et al. 1993 A&AS 102, 611
Stevens, J.A., Litchfield, S.J., Robson, E.I, et al. 1994, ApJ 437, 91.
Thompson, D. J., Bertsch, D.L., Dingus, B.L. et al. 1995 ApJS, in press.
Urry, C.M. and Padovani, P. 1991 ApJ 371, 60.
Von Montigny, C., Bertsch, D.L., Chiang, J. et al. 1995 ApJ 440, 525.
Wehrle, A.E., Cohen, M.H., Unwin, S.C., Aller, H.D., Aller, M.F., and Nicolson, G. 1992 ApJ 391, 589

Figure 1

Spectral energy distribution of 3C279 at two epochs, June 1991 and January 1993 ("This campaign") from Maraschi et al. 1994.


3C279
EGRET (91)
COMPTEL (91)
EGRET (93)
GINGA
ROSAT
This campaign
Outburst (June 1991)
Log νf_ν (Jy Hz)
Log ν (Hz)
14
13
12
11
10
10
15
20
25

GAMMA-RAY BURSTS IN THE BATSE AND POST-BATSE ERA

D. H. Hartmann[1] and S. E. Woosley[2,3]

[1]*Department of Physics & Astronomy*

Clemson University, Clemson, SC 29634

[2]*Astronomy Board, UCSC, Santa Cruz, CA 95064*

[3]*General Studies Division, LLNL, Livermore, CA 94550*

ABSTRACT

The origin of gamma-ray bursts (GRBs) continues to be a great mystery. We review some relevant observations and a number of recent models. While no clear solution exists at the present time, the parameter space for Galactic halo models is becoming very constrained. Cosmological models on the other hand require both enormous total energy and the concentration of that energy into a small mass. This implies compact objects, probably accreting black holes of stellar size. We review some of the physics of such accreting black holes, emphasizing the possibility that a rapidly accreting black hole of stellar mass might precess. This precession, coupled to beaming, could impose additional time structure on the burst and its spectrum. In the event of a "failed" supernova model, the wind of the Wolf-Rayet star prior to the event could provide the beam dump where the jet generates gamma-rays. Enduring emission that grows harder with time might be expected for several hours as the density in the vicinity of the black hole declines. How will future observations settle some, or all, of the mysteries GRBs provide? Counterpart searches, very high energy γ-ray observations, and X-ray spectroscopy appear to be promising avenues

for future studies of GRBs in the post BATSE era.

INTRODUCTION

"But the pendulum does swing in science,..."

J. Silk, 1994

With the launch of the Compton Gamma-Ray Observatory (CGRO) it was hoped that the origin of γ-ray bursts (GRB's) would be conclusively resolved. Much to the delight and consternation of all involved, data gathered with BATSE instead deepened the mystery. Whereas most theorists had favored a burst origin associated with Galactic neutron stars, this association appears to be inconsistent with the observed inhomogeneous spatial distribution and near perfect isotropy on the sky. A new paradigm has emerged that involves sources at cosmological distances, although local solutions are still being considered. Suggestions for Galactic models range from the Oort cloud to a very extended Galactic halo, and many of the new models still retain neutron stars as building blocks. A recent compilation, already incomplete, listed 135 separate models for gamma-ray bursts, two-thirds of which involved neutron stars (Nemiroff 1994). A guide to the literature is provided in several recent conference proceedings (Paciesas & Fishman 1992; Ho, Epstein, & Fenimore 1992; Fishman, Brainerd, & Hurley 1994) and review articles (Hartmann 1994abc). Figure 1 summarizes some of the basic observational facts.

Figure 1. Clockwise from upper left the figure shows a typical lightcurve, the isotropic distribution in Galactic coordinates, the non-uniform spatial distribution implied by the peak flux statistic, and the bimodal duration distribution (from Hartmann, 1994a).

There is, at present, no single GRB model, local or cosmological, that provides an acceptable explanation, consistent with all major observational constraints, for the bulk of classical bursts. There are pieces of models, and numerous scenarios for achieving the requisite bolometric energy at any distance. However, no model starts with a credible astronomical event and follows that event through to the production of γ-radiation having the requisite spectrum, time history, spatial distribution, and event rate. This lack of a credible model reflects both continued uncertainty about the distance scale and a general difficulty, both in Galactic and cosmological models, in producing either well ordered, highly relativistic jets or matter characterized by exceptionally high entropy per baryon. It is also difficult to reach definitive conclusions when the bursts almost certainly involve magnetohydrodynamics and radiation transport in a situation that is far from thermal equilibrium. Highly relativistic motion, perhaps coupled with tight beaming seems to be required in cosmologically distant models to allow photons with $\sim$ 10 GeV to escape, but that same beaming increases the number of sources or their specific burst rate. Estimates of the rate at which neutron stars merge together with the implied sampling depth of BATSE ($z \sim 1$) are consistent only if strong beaming, say to $\lesssim$1% of the sky, does not occur. We recently presented a review of theoretical models of classical GRBs (Hartmann & Woosley 1994), where additional details can be found.

SOME OUTSTANDING ISSUES

There remain intriguing suggestions that some classical GRB's are associated with Galactic neutron stars: the possible existence of repeating classical burst sources; the existence of three soft gamma-repeaters, which have been linked to supernova

remnants; controversial cyclotron line features; small residual anisotropies in the BATSE source distribution; and observations of allegedly thermal x-rays which imply D(kpc)/R(km) $\sim$ 1, where R is the size of the emitting region and D the source distance. The BATSE data on isotropy and C/C_{max} remain the strongest constraints on such local models, but there are loopholes. If both the kick velocity of the progenitor neutron stars and the uncertain potential of the Galactic halo can be finely (but perhaps reasonably) tuned (much larger than 700 km s^{-1} for the former), an extreme Galactic halo model consistent with BATSE data can be geometrically constructed (Podsiadlowski, Rees, & Ruderman 1994; Hartmann *et al.* 1994), although it is clear that the challenge of providing truly convincing physical arguments in favor of such models has not yet been met.

The detection by EGRET of continuing high energy emission, for as much as 90 minutes, of photons with energies up to $\sim$ 10 GeV (Sommer *et al.* 1994; Hurley *et al.* 1994) provides the the most recent challenge to model builders.

Against this background, the resolution to the gamma-ray burst problem is likely to come only with more observations. We must keep BATSE operating for an extended time and, if possible, have another detector launched with at least 4 times the area of BATSE to see, at twice the distance, whether: 1) a burst excess from M31 might emerge 2) the color and time dilations attributed to cosmology (Norris *et al.* 1994; Band 1994) are real; 3) repetition is indeed present (Quashnock & Lamb 1993; Wang & Lingenfelter 1993; Hartmann *et al.* 1994) in some fraction of classical GRB's; and 4) whether the angular correlation function for gamma-ray bursts agrees with that of distant galaxies (Hartmann & Blumenthal 1989; Lamb & Quashnock 1993). More statistics will allow a more meaningful result to be obtained when the sample is subdivided, and we might also eventually detect a gravitational lensing event (Nemiroff *et al.* 1994).

THEORETICAL BASICS

"There are theories,and then there are theories"

L. Lederman

For isotropic emission, the total energy released in a burst is

$$E_r \sim 10^{33} \left(\frac{F}{10^{-6}\ \mathrm{erg/cm^2\ s}}\right) \left(\frac{\tau}{10\ \mathrm{s}}\right) D_{\mathrm{pc}}^2\ \mathrm{erg} \qquad (1)$$

For an average recurrence time τ_r during a total lifetime T_r, the total energy required is $E_{\mathrm{tot}} \sim E_r\ T_r/\tau_r$. Consider distance scales of 10^4 AU (solar system), 100 pc (local stars), 1 kpc (Galactic disk), 100 kpc (extended Galactic halo), and 1 Gpc (cosmological). For non-repeating sources, the corresponding energies range from 10^{30} ergs to 10^{51} ergs. Simply providing this energy without consideration of efficiency for conversion to gamma-rays poses no great difficulty, even at cosmological distances. The formation of a neutron star of $\sim 1\ M_\odot$ and radius of ~ 10 km releases a gravitational binding energy or about 3×10^{53} ergs. Accretion from the last stable orbit of a black hole provides similar specific energy. However, for the requisite large accretion rates into black holes or for neutron star formation, one expects most of this energy to be carried away by neutrinos.

Prior to the launch of CGRO most researchers thought that the solution lay relatively close by. Calculations of neutron star distributions (Hartmann, Epstein, & Woosley 1990; Paczynski 1990) and the resulting statistical estimates of brightness distributions and sky maps suggested that the Milky Way association should be revealed to a sensitive BATSE. Alternatively, the idea that bursts originate at cosmological distances was suggested soon after the discovery of the phenomenon (Prilutski & Usov 1975; Usov & Chibisov 1975), but cosmological scenarios did not receive much attention until Paczynski (1986) drew attention to two coincidences: (1) for redshifts of order unity the required energy is comparable to typical

energies released in supernova explosions; and (2) the resulting thermal spectra would have temperatures of several MeV and moderate redshifts would roughly account for the peak in the power spectra. Paczynski (1986, 1987) also suggested that repetition could be caused by gravitational lensing of high redshift sources by foreground galaxies, and raised two issues - the predicted thermal spectra would have to be altered to agree with the observations, and the angular distribution of faint bursts should be isotropic. Subsequently Goodman (1986) showed that even optically thick models (expected for cosmological distances) can produce very hard spectra (see also Goodman *et al.* 1987).

The energy required for cosmological models is about $10^{51}\delta\Omega$ ergs, with $\delta\Omega$ the beaming fraction of the burst. The observed variability of γ-ray burst light curves suggests a small volume, comparable to the size of neutron stars, for the delivery of this energy. Even if this energy is delivered as pure radiation, two-photon pair production ($\gamma\gamma \rightarrow e^- e^+$) would create a pair plasma that would be thick to gamma-rays. If no baryonic matter is present a pair fireball will be created at temperatures of $\sim$ 100 MeV with a Compton optical depth exceeding 10^{10}. The fireball will thermalize regardless of the energy injection mechanism. The resulting bulk motion of these expanding pair fluids leads to Lorentz factors $\Gamma \sim 10^3$. While these factors are sufficient to Doppler boost the spectrum into the γ band, the spectrum would still be approximately thermal, in conflict with the observations.

What bulk Lorentz factors are needed to produce a GRB and avoid the pair catastrophe? The most severe constraints follow from demanding that the opacity is below unity at all photon energies. The maximal photon energy observed from GRBs exceeds $\sim$ 10 GeV. The cross section is approximately given by the Thomson cross section, but is angle and energy dependent. It scales with the square of the center-of-momentum energy of the interacting photons. The higher the energy of photons that apparently escape from the source region without suffering extinction the more severe will be the limits resulting from the opacity constraint. We ignore the angle dependence and assume an average photon energy to provide an order of magnitude estimate. Using the observed time variability, $\delta t = 10^{-3}\ \tau_{ms}\ s$,

to estimate the size of the source region. Allowing for relativistic expansion with bulk Lorentz factor Γ the scale is given by $R = c\ \delta t\ \Gamma \sim 3 \times 10^7 \tau_{ms}\ \Gamma$ cm. The mean observed photon energy is ϵ so that, in the rest frame of the emitting source, photons have energies of order $\epsilon_0 = \epsilon\Gamma^{-1}$. If we assume that the expansion leads to beaming into a solid angle $\delta\Omega \sim \Gamma^{-2}$, the observed GRB flux, f, yields the source luminosity, L, for given distance D, from $f = L\ D^{-2}\ \Gamma^2$. Relativistic beaming thus reduces the luminosity requirements, but consequently increases the required number of sources by the same factor.[27] The reduced luminosity implies smaller photon number densities at the source which reduces the optical depth to photon-photon pair production

$$\tau_{\gamma\gamma} \sim \sigma_{\rm T}\ n_\gamma\ R \sim 10^5\ D^2_{\rm kpc}\ \tau_{ms}^{-2}\ \Gamma^{-4}, \tag{2}$$

where we assumed a typical photon energy of $\sim 10^{-6}$ ergs ($\sim$ 1 MeV) in the rest frame and $F_6 \sim 1$. At cosmological distances Lorentz factors would have to exceed $\Gamma \sim 10^4$ to significantly reduce the opacity. An additional effect of relativistic beaming is the suppression of pair creation by reducing the threshold energy. A test photon with energy ϵ_0 interacting with a field of photons with energy ϵ' will create pairs only if

$$\epsilon\ \epsilon'\ (1-\mu) \geq 2\ \left(m_e c^2\right)^2 \tag{3}$$

is satisfied, where μ is the interaction angle in the center-of-momentum frame. Large Lorentz factors reduce observed photon energies of $\sim$ 100 MeV to rest frame energies that can fall below this threshold. In addition, the spectral shape of the emitting source and possible deviations from isotropy as well as the geometric extent of the expanding shell all affect the opacity calculation. Also, pair creation further increases the opacity via Klein-Nishina scattering off the additional leptons. Detailed calculations (Harding & Baring 1994) suggest that $\Gamma \sim 10^3$ for cosmological models. Meszaros & Rees (1994) discussed how this might be reduced to $\Gamma \sim 100$ in a non-steady jet.

Early attempts (Fenimore *et al.* 1992) to combine the count statistics of PVO with

that of BATSE, suggested that source evolution was necessary to explain the deficit of faint BATSE bursts relative to the extrapolated PVO rate. A redetermination of the effective lifetime of PVO reduced the discrepancy significantly (Fenimore *et al.* 1993), so that the combined PVO/BATSE statistics appears to be consistent with non-evolving standard candles observed to a maximum redshift of $z_{max} \sim 1$ (Fenimore *et al.* 1993; Wickramashinghe *et al.* 1993). However, if the observed time dilation (Norris *et al.* 1994) implies that the maximum redshift is in fact $z_{max} \sim 2$, models with zero cosmological constant ($\Lambda = 0$) actually require some form of source evolution (Horack, Emslie, & Hartmann 1994). Assuming an $\Omega = 1$ universe with Hubble constant $H = 75$ km s^{-1} Mpc^{-1} the brightness distribution implies a source luminosity of $L \sim 6 \times 10^{50}$ ergs s^{-1} and an event density of $\rho = 22$ Bursts yr^{-1} Gpc^{-3} h_{75}^3. The detector thresholds correspond to maximum sampling redshifts of z_{max}(PVO) ~ 0.2, and z_{max}(BATSE) ~ 1. An upper limit on redshift can be obtained from the requirement that the photons of highest energy not be attenuated by pair production off the intergalactic IR field. The optical depth is approximately given by $\tau \sim 10\ z\ E^{3/2}$, where the observed photon energy, E, is measured in TeV. EGRET recorded several bursts with GeV emission. The highest photon energy recorded from GB940217 is roughly 10 GeV (Hurley *et al.* 1994), thus this particular burst could not have originated at a redshift larger than ~ 100, ruling out superconducting string models. If future observations extend burst detections into the TeV range, powerful constraints on the burster distance scale would result. Some of the GeV emission from GB940217 was delayed by over one hour, which is hard to explain in any "promt" scenario, but models in which the GRB occurs via the interaction of an expanding fireball with surrounding matter show that delayed GeV emission may in fact be a characteristic feature of cosmological fireball models (Meszaros & Rees 1994). The sampling redshift relates to an average event rate per host galaxy. The total number of galaxies within redshift z is given by the following set of equations

$$N(z) = N_0 \int_0^z dz\, y^2\, \mathrm{E}^{-1} \quad , \tag{4}$$

$$y(z) = \left(\frac{\Omega}{2}\right)^{-2} (1+z)^{-1} \left\{\Omega z/2 + (\Omega/2 - 1)\left[(1+\Omega z)^{1/2} - 1\right]\right\} \quad , \tag{5}$$

$$N_0 = 4\pi\, \mathrm{n_0}\, {\mathrm{L_H}}^3 \sim 8 \times 10^9\, \mathrm{H_{75}^{-3}}\, \mathrm{n_7}\ . \tag{6}$$

If host sites are spiral galaxies, the B-band emissivity of the universe ($L_u \sim 2 \times 10^8\ \mathrm{L_{B\odot}\ Mpc^{-3}}$) and the B-band luminosity of our Galaxy ($L_g \sim 2 \times 10^{10} L_{B\odot}$) imply a host density of $n \sim 10^7\ \mathrm{Gpc^{-3}}$, i.e., $\mathrm{n_7} \sim 1$. If emission is beamed into the fraction $\delta\Omega \ll 1$, the specific event rate per host is

$$R = 2.5 \times 10^{-6}\, R_3\, n_7^{-1}\, \delta\Omega^{-1}\ \ \mathrm{yr^{-1}}. \tag{7}$$

The corresponding time between bursts in any given galaxy is thus of order $R^{-1} \sim 4 \times 10^5$ yrs. If strong beaming (say, $\delta\Omega \sim 10^{-2}$) were an essential part of cosmological burst models, specific rates as high as $10^{-4}\ \mathrm{yr^{-1}}$ may be needed. Many of the recent cosmological burst models invoke stellar mergers in neutron star – neutron star or neutron star – black hole systems. Although it seems established that these events must occur in galaxies, the expected specific rates are of order $10^{-5}\ \mathrm{yr^{-1}}$ and could therefore fall short by a few orders of magnitude.

During the history of the universe there have been about $10^{13}/\delta\Omega$ gamma-ray bursts within the distance sampled by CGRO (with beaming factor $\delta\Omega \sim 0.001$ - 0.1). If these have come from $\sim 10^9$ galaxies, each galaxy has produced about $10^4/\delta\Omega$ bursts, or one event every $10^6\ \delta\Omega$ years. If each event corresponds to the accretion of a sun-like mass, then over $10^4\ \mathrm{M_\odot}/\delta\Omega$ of material has fallen into black holes in each galaxy. The number could be larger because not all events involving the same sources must produce gamma-ray bursts. This suggests either a large number, $\sim 10^6$ of stellar mass black holes, or else the repeated accretion of stars

by a very massive black hole, e.g., an active galactic nucleus (Carter 1992). The latter may pose a problem in terms of its long characteristic time scale. The orbital period for a 10^6 $M_\odot$ black hole at 3 R_s is

$$P \approx 400\ (10^6\ M_\odot/M)^{1/2} \quad \text{sec}\ , \tag{8}$$

much longer than the characteristic time scale for variation, or even duration of a typical GRB. Many of the mechanisms alleged to produce hard emission in AGN's rely on the disk making many orbits, for example, to build up the equipartition magnetic field required for efficient extraction of rotational energy (Blandford & Znajek 1977; Blandford 1989), or for centrifugal acceleration. The tidal radius for a 10^6 $M_\odot$ black hole is also uncomfortably close to the event horizon, even for main sequence stars.

While it is possible to envision intermediate cases - e.g. 10^4 black holes of 100 $M_\odot$ - it is more interesting to consider the lower limit where the black holes have mass of roughly 5 - 10 $M_\odot$, i.e., what might be produced in many supernova explosions, especially "failed" ones (Woosley 1993). Alternatively, but probably with much less frequency, such black holes may merge with neutron stars (Paczynski 1991; Narayan, Paczynski, & Piran 1992), or white dwarfs that might be tidally detonated. In these situations one expects after some delay, perhaps many seconds in the case of failed supernovae, to have a rapidly rotating black hole accreting material from a disk at a rate of about 0.1 $M_\odot$ s^{-1}. The total mass of the disk is $\sim$ 1 $M_\odot$ with a center of mass around $\lesssim 10^8$ cm (Bodenheimer & Woosley 1983). Interestingly this situation could lead to gravitomagnetic precession of the black hole (Hartle *et al.* 1986)

$$\Omega_{GM} \sim 0.1 \left(\frac{M_{disk}}{M_\odot}\right) \left(\frac{M_{BH}}{5\ M_\odot}\right)^{1/2} \left(\frac{10^8 \text{cm}}{\text{b}}\right)^{2.5} \text{s}^{-1}\ . \tag{9}$$

This precession of the black hole would not necessarily be strictly periodic since both the distance, b, and the mass of the disk are time variable. The inner part of

the accretion disk would be warped into the equatorial plane of the rotating black hole by the Bardeen-Petterson effect and would precess with the black hole. The precession frequency is quite sensitive to the Bardeen-Petterson radius, R_{BP}, which is frequently assumed (Hartle *et al.* 1986) to be in excess of $\sim 30\ R_s$. However, at the very large accretion rates we are considering here, R_{BP} may be significantly reduced and Ω_{GM} would be greater. Given that the black hole has just been formed in a system which presumably has the angular momentum of the accretion disk and black hole very nearly aligned, one would not expect a large precession angle, still pulsars are somehow born in similar situations with magnetic and rotational axes that are not aligned. In any case, the precession angle is probably not more than a few degrees and any observational effect would depend on beaming to a smaller angle than that.

The possibility of precessionally induced modulation is especially interesting in light of numerous GRB's of the long complex variety, e.g., GRB920221, whose light curves exhibits quasiperiodic behavior. Perhaps we only see the burst, or at least see the greatest intensity when the jet passes across our line of sight, somewhat analogous to a pulsar. This would have several consequences in addition to obvious effects on the time profile. A given burst of instantaneous opening angle θ would illuminate a larger portion of the sky as it passed over many different locations. There could thus be fewer events. The total duration of a burst might also be significantly longer than what is seen from Earth by BATSE and one might not be surprised to see burst activity at various wavelengths, both substantially before and after the BATSE trigger. Because it emits a non-trivial part of its rest mass as neutrinos at the speed of light, the black hole and its disk may also wander somewhat in orientation owing to the back reaction from any anisotropy in this emission. A burst might even appear to turn off and reappear at a later time (e.g., GRB910503).

Though our discussion has centered so far on black holes that either existed previously or formed in failed supernovae, similar circumstances may arise in the case of merging neutron stars. While the neutron stars are merging, they are heated to an

unknown extent by tidal interaction (Pringle 1992) and emit neutrinos. However the degree of this heating is uncertain and preliminary hydrodynamical calculations of neutron star mergers suggest a relatively gentle event (Wilson, private communication). Even if the neutron stars become hot and emit neutrinos, it will be difficult to produce matter moving at sufficiently relativistic energies because the neutrinos will tend to drive winds (Pringle 1992; Meszaros & Rees 1992, 1993; Woosley & Baron 1992) as well as deposit energy along the rotational axis (see however Duncan, Shapiro, & Wasserman 1986 and Usov 1992 for more optimistic conclusions). Given that the critical mass for black hole formation may be significantly less than 2.5 $M_{\odot}$, a likely outcome of this merger process, a second or so after the merger, is a black hole surrounded by a residual accretion disk (Woosley 1993; Narayan, Paczynski, & Piran 1992) of $\sim$0.1 $M_{\odot}$. This is less mass, and therefore possibly less energy available for making a GRB than in failed supernovae, but there is less overlying material and possibly less need for focusing of the emission into a jet. The burst can also commence quickly and have a shorter duration. Perhaps merging neutron stars are more applicable to short gamma-ray bursts (Woosley 1993).

Once an accreting black hole has been produced, new characteristic time scales are introduced. One is the viscous time scale for the accretion disk. How long does it take for material in the disk to move into the black hole? Very crude estimates (Woosley 1993) suggest that a time scale of seconds to minutes might be reasonable. Additionally, in the case of the failed supernova, there is the time it takes for outlying material to free fall to the disk. Starting at a density of $\sim 10^2$ g cm $^{-3}$ at the outer edge of the helium core, this time is also a minute or so. Indeed it might take hours for all the lower density matter to accrete. This suggests that continuing post burst emission should be sought from bursts at all wavelengths and suggests a possible explanation for the hard radiation seen at $\sim$ 90 minutes for burst GRB940217 by EGRET (Hurley *et al.* 1994). The spectrum produced by a rapidly rotating accreting black hole might harden as the density of matter near the black hole declines. Raw energy is delivered by accretion at a rate

$$\dot{E} \sim 5 \times 10^{52} \left(\dot{M} / 0.1 \, M_\odot \, s^{-1} \right) \, erg \, s^{-1}. \qquad (10)$$

The total mass reservoir in the failed supernova case is about 10 $M_\odot$ for a total energy of almost 10^{54} erg. Of this we wish to convert only about 0.01% into gamma-radiation (we presume here that such a burst will be beamed to about 1 - 10% of the sky) so there is some room for inefficiency, e.g., for neutrino losses. Some of the neutrinos may contribute to the gamma-ray burst, but this is a relatively inefficient process that may not lead to sufficient energy density to make a highly relativistic beam.

One expects, however, that the magnetic field of the disk will be greatly amplified by differential rotation. Unlike the very massive black holes of AGN's, the accretion disk for stellar mass black holes experiences many orbits within a GRB time scale. At a radius of 100 km and for an accretion rate of 0.1 $M_\odot$ s^{-1}, the equipartition field, $B^2/8\pi \sim \rho v^2$ would be about 10^{14} Gauss. Dissipation of this magnetic energy would release considerable energy that might be collimated by the thick disk into jets[39]. One also expects that black holes form with a large amount of rotational energy. A potential of 10^{21} volts could be developed along the magnetic field coaligned with the black hole rotation axis by the Blandford-Znajek (1977) process. The presence of a large amount of plasma in the vicinity of the black hole could pose problems by short circuiting this potential and keeping it from accelerating particles (the Goldreich-Julian density). But both the magnetic field and the plasma density are about 10 orders of magnitude greater here than in AGN's, so one might think that what AGN's can do, this model for γ-ray bursts can do as well. AGN's emit both collimated relativistic beams and high energy γ rays. The Goldreich-Julian density has historically been a problem for AGN's (Katz 1987).

The greatest shortcoming of this class of models is the lack of a detailed physical description of how accretion and rotational energy gets converted into (beamed ?) γ-ray emission and how this emission is affected by the concentration of matter around the source. For example, in the failed supernova model one expects there

to be a lot of matter ($\rho \sim 100$ g cm^{-3} extending to $\sim 10^{10}$ cm) that gets in the way of the beam and could degrade its initial energy. The configuration produced by merging neutron stars may not be all that much better, unless energy is converted into relativistic outflow with near 100% efficiency. A subrelativistic wind[42,43] of only 0.01 $M_\odot$ s^{-1} would provide a surrounding density of $\sim 10^6$ r_7^{-2} g cm^{-3}. Depending upon its density and distribution in space, external matter could be either a problem or an asset. Some low density matter in the vicinity of the burst might aid in converting the kinetic energy of the beam into radiation. Radiation diffusing into the beam might experience inverse Compton, and the pressure of the external medium might even serve to collimate the beam and keep its opening angle small (Mochkovitch *et al.* 1994; Meszaros & Rees 1992b). Details are lacking here and calculations are needed. Ultimately if Γ is to be 10^{2-3} one needs to have the energy in the burst, about 10^{49} ergs for beaming into 1% of the sky, concentrated into only about $10^{-8} - 10^{-7}$ $M_\odot$. This may be tough. It may be easier to put more energy in a larger amount of matter, say 10^{53} erg into $10^{-4} - 10^{-3}$ $M_\odot$, and then invoke a lower efficiency of kinetic energy conversion to gamma-rays and a larger beaming fraction.

Pair neutrino losses are $\sim 10^{25}\ T_{MeV}^9$ erg cm^{-3} s^{-1} and the internal energy of a pair plasma, $\sim 10^{26} T_{MeV}^4$ erg cm^{-3}, implying a neutrino cooling time of about 10 T_{MeV}^{-5} sec. Unless the temperature in the energy deposition region along the rotational axis exceeds about 5 MeV, the expansion time scale, $\sim 10^8 \mathrm{cm}/c$, would be shorter than the neutrino cooling time. This is unlikely because, even without expansion, pair neutrino losses in a volume of 10^{23} cm^3 would dissipate 5×10^{52} erg s^{-1} at a temperature of only 3 MeV. Thus it is reasonable to assume that the energy deposited, by whatever means, in the matter just above the rotating hole, will be converted into expansion kinetic energy with reasonable efficiency. This of course neglects neutrino dissipation in the disk itself. All in all, a total (time integrated) energy deposition in the outgoing matter of $10^{51} - 10^{52}$ erg, does not seem unreasonable.

As previously noted, the subsequent propagation of this expanding matter, now

a "jet", through the overlying material is an uncertain aspect of the model which can and should be studied with multidimensional hydrodynamical calculations. For now we assume that the jet reaches the surface of the star relatively intact with an opening angle of about 15 degrees (similar to quasars) and a relativistic Γ of about 100. This jet then produces the GRB by interaction with surrounding material, one possibility being synchrotron emission from the reverse shock (Woosley & Baron 1992; Begelman, Blandford, & Rees 1984).

Interestingly, the wind of a WR star prior to core collapse may play an important role here. Taking 10^{-5} $M_\odot$ y^{-1} as a representative value implies a baryonic concentration of 10^{-9} g cm^{-3} r_{10}^{-2} and a Thomson depth of about 10 from the surface of the star to infinity. In fact, along polar angles and after most of the star has collapsed, there is considerably more mass in this wind per unit solid angle than in the star. Thus the wind of the precollapse star may provide the "beam dump" in which the energy of the jet is converted to gamma-rays (Woosley & Baron 1992). The duration of this wind is typically 10^5 to 10^6 years and thus, depending on circumstellar interaction, most of the mass would probably be located at a parsec or so. The beam would encounter its own mass at a considerably shorter distance.

SOURCE IDENTIFICATION

We believe that continuing efforts at source identification are essential. The HETE mission, to be launched in 1995, will give precise localizations, 10 arc seconds if detectable (12th magnitude) UV emission accompanies the burst. Rapid localization with the BACODINE system should be coupled to a dedicated radio observatory to detect the possible delayed radio afterglow from GRBs. Just as in the case of radio pulsars, the measurement of a GRB dispersion measure could finally fix the unknown burster distance scale. Promising ground based searches for simultaneous optical transients are being carried out. The BACODINE–GROCSE system at LLNL (Akerlof *et al.* 1994) is able to slew to an ongoing burst (in the proper location, at night, in good weather) within about 10 seconds and observe with a limiting optical magnitude of 8. The Explosive Transient System (ETC) has already succeeded in placing optical magnitude limits ranging from 6 to 8 on four gamma-ray bursts (Krimm, Vanderspek, & Ricker 1994) whose error boxes had at least partial overlap with its field of view and were seen by BATSE during April, 1991 - May, 1993. However, the inferred ratios of gamma-ray to optical luminosity were all less than about 100, about 10 times less than inferred for historical optical transients in modern GRB error boxes (Schaefer 1981). In the future ETC will receive BATSE coordinates by way of BACODINE and slew to bursts within 10 s. This will greatly increase both their event rate and the likelihood of placing limits of up to $10^5 = L_\gamma/L_{opt}$ on the brightest bursts, or else seeing a counterpart.

Missions, such as HETE, that have significantly greater sensitivity than BATSE below 50 keV, will be better able to address the search for cyclotron lines and additional "soft gamma-ray repeaters" (SGR). Are these repeaters truly a separate family (as generally assumed in cosmological models) or is there a relation between SGRs and GRB's? At some energy one also expects the spectrum to become self absorbed and display a low energy roll over. HETE will be able to push the search for such roll overs down to 1.5 keV. A rollover to a Rayleigh-Jeans spectrum could indicate the emission measure of the emitting material and, coupled with estimates

of the size and temperature, imply a distance (Fenimore, Laros, & Ulmer 1994). Finally gravity wave detectors, such as LIGO, might observe events correlated with GRB's if the merger of compact objects is indeed responsible for GRBs. Paczynski & Xu (1994) recently proposed that neutrino bursts accessible to future detectors may accompany GRB's. In the remainder of this paper, we review a number of physical models and constraints on GRB's. Though we concentrate on cosmological models, the issue is far from resolved.

To push the localization accuracy into the sub-arcsecond regime, novel concepts are required. One of the most promising approaches appears to be the Energetic Transient Array (ETA), essentially a network of "space buoys" in solar orbit (Ricker 1990). In this concept a network of microsatellites with a 2 AU circular baseline constitute a timing array that would improve the localization accuracy of the IPN networks significantly.

This work has been supported by NASA (NAGW 2525, NAG 5-1578 and SC-A-292701)

REFERENCES

Akerlof, K., *et al.* 1994, in *Gamma-Ray Bursts* eds. G. J. Fishman, K. Hurley, & J. Brainerd, (AIP: New York), 307, 633

Band, D. 1994, ApJL, 432, L23

Begelman, M., Blandford, R., & Rees, M. J. 1984, Rev M. Phys., 56, 255

Blandford, R. D., & Znajek, R. L. 1977, MNRAS 179, 433

Blandford, R. D. 1989, in *Theory of Accretion Disks* eds. F. Meyer *et al.*, (Kluwer Academic Publishers)

Bodenheimer, P., & Woosley, S. E. 1983, ApJ 269, 281

Carter, B. 1992, ApJL 391, L67

Duncan, R. C., Shapiro, S. L., & Wasserman, I. 1986, ApJ 309, 141

Fenimore, E. E., Laros, J. G., & Ulmer, A. 1994, ApJ, in press

Fenimore, E.E. *et al.* 1992, Nature 357, 140

Fenimore, E.E. *et al.* 1993, Nature 366, 40

Fishman, G. J., Brainerd, J. J., & Hurley, K. 1994, *Gamma-Ray Bursts*, AIP 307.

Goodman, J. 1986, ApJL 308, L47

Goodman, J. *et al.* 1987, ApJL 314, L7.

Harding, A. K., & Baring, M. G. 1994, in *Gamma-Ray Bursts* eds. G. Fishman, J. Brainerd, & K. Hurley, (AIP: New York), 307, 520

Hartle, J., *et al.* 1986, in *Black Holes: The Membrane Paradigm*, eds. K. Thorne, R. Price, & D. MacDonald, (Yale Univ. Press), p. 173

Hartmann, D. H., & Blumenthal, G. R. 1989, ApJ, 342, 521

Hartmann, D. Epstein, R.I. & Woosley, S.E. 1990, ApJ 348, 625

Hartmann, D. H. 1994a, Science, 263, 47

Hartmann, D. H. 1994b, in *High Energy Astrophysics*, ed. J. M. Mathews, (World Scientific, pp. 69-106

Hartmann, D. H. 1994c, in *The gamma-ray sky with COMPTEL and SIGMA*, eds. M. Signore, P. Salati, and G. Vedrenne, Les Houches 1994 Proceedings

Hartmann, D. H., & Woosley, S. E. 1994, Proceedings of the 30th COSPAR Scientific Assembly, ed. K. Pollock & N. Gehrels, (Pergamon Press), in press

Hartmann, D. H. *et al.* 1994, ApJS, 90, 893

Hartmann, D. H., *et al.* 1994, ApJ, in press

Horack, J. M., Emslie, A. G., & Hartmann, D. H. 1994, ApJ, in press

Hurley, K., *et al.* 1994, Nature, in press

Katz, J. 1987, *High Energy Astrophysics*, (Addison Wesley), p. 306

Krimm, H. A., Vanderspek, R. K., & Ricker, G. R. 1994, in *Gamma Ray Bursts*, ed. G. Fishman, K. Hurley, & J. Brainerd, (AIP: NY) 307, 423

Lamb, D. Q., & Quashnock, J. M. 1993, ApJ, 415, L1

Mészáros, P. & Rees, M. J. 1992a, ApJ 397, 570

Mészáros, P. & Rees, M.J. 1992b, MNRAS 257, 29P

Mészáros, P. & Rees, M. J. 1993, ApJ 405, 278

Mészáros, P., & Rees, M. J. 1994, MNRAS 269, L41

Mochkovitch, R., Loiseau, S., Hernanz, M., & Isern, J. 1994, in *Gamma Ray Bursts* eds. G. Fishman, J. Brainerd, & K. Hurley, (AIP: NY) 307, 537

Narayan, R., Paczynski, B., & Piran, T. 1992, ApJL 395, L83

Nemiroff, R. J., *et al.* 1994, ApJ, 432, 478

Nemiroff, R. J. 1994, in *Gamma-Ray Bursts*, eds. G. Fishman, J. Brainerd, & K. Hurley, (AIP: New York), 307, 730

Ho, C., Epstein, R. I., & Fenimore, E. E. 1992, *Gamma-Ray Bursts*, (Cambridge University Press)

Norris, J. P., *et al.* 1994, ApJ, 424, 540

Paciesas, W. S., & Fishman, G. J. 1992, *Gamma-Ray Bursts*, AIP 265

Paczynski, B. 1986, ApJL 308, L43

Paczynski, B. 1987, ApJL 317, L51

Paczynski, B. 1990, ApJ 348, 485

Paczynski, B. 1991, Acta Astr. 41, 257

Paczynski, B., & Xu, G. 1994, ApJ, 427,708

Podsiadlowski, P., Rees. M. J., & Ruderman, M. 1994, MNRAS, submitted

Prilutski, O.F. & Usov, V.V. 1975, Ap&SS 34, 395

Pringle, J. E. 1992, MNRAS, 258, 811

Quashnok, J. M., & Lamb, D. Q. 1993, MNRAS, 265, L45

Ricker, G. R. 1990, in *High Energy Astrophysics in the 21st Century*, ed. P. C. Joss, AIP, 211, 365

Schaefer, B. E., 1981, Nature, 294, 722

Sommer, M., *et al.* 1994, ApJL, 422, L63

Usov, V.V. & Chibisov, G. V. 1975, Sov. Astron. 19, 115

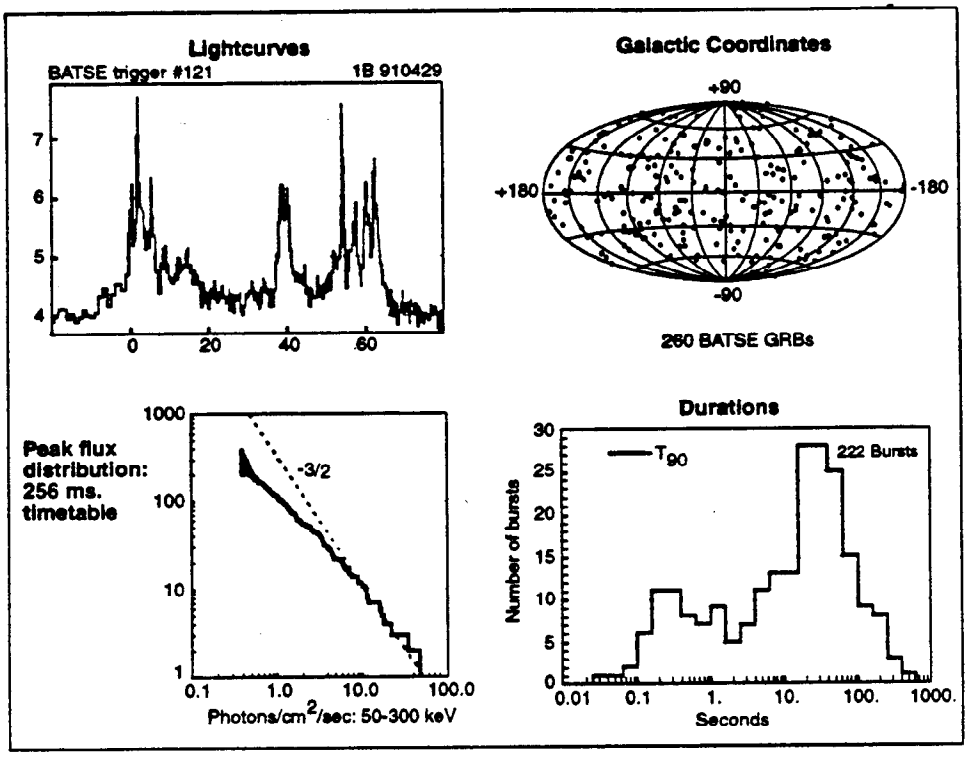
Usov, V.V. 1992, Nature 357, 472

Wang, V. C., & Lingenfelter, R. E. 1993, ApJ, 416, L13

Wickramashinghe, W.A.D.T. *et al.* 1993, ApJL 411, L55

Woosley, S. E. & Baron, E. 1992, ApJ 391, 228

Woosley, S.E. 1993, ApJ 405, 273


Lightcurves
BATSE trigger #121
1B 910429
7
6
5
4
0
20
40
60
Galactic Coordinates
+90
+180
-180
-90
260 BATSE GRBs
Peak flux distribution: 256 ms. timetable
1000
100
10
1
-3/2
0.1
1.0
10.0
100.0
Photons/cm2/sec: 50-300 keV
Durations
30
25
20
15
10
5
0
T90
222 Bursts
Number of bursts
0.01
0.1
1.
10.
100.
1000.
Seconds

Astro-E Hard X-ray Detector

T. Kamae
Department of Physics, University of Tokyo
7-3-1, Hongo, Bunkyo-ku, Tokyo 113, Japan

Abstract

Astro-E is the 5th Japanese X-ray satellite project to be launched in year 2000 by the Institute of Space and Astronautical Science (ISAS). This report primarily deals with the design and expected performance of the Hard X-ray Detector (HXD) now being developed as one of the 3 experiments aboard the Astro-E satellite. The HXD is a combination of YAP(or GSO)/BGO well-type phoswich counters and silicon PIN diodes: these two combined will cover a wide energy band of 10−700 keV. The detector is characterized by its low background of around several$\times 10^{-6}$c/s/cm^2/keV, or equivalently, a high sensitivity, at least one order of magnitude higher than any other past missions between a few 10 keV and several 100 keV. Combined with the other 2 experiments, a micro-calorimeter array (XRS) and 4 CCD arrays (XIS), all with X-ray mirrors, the Astro-E mission will cover the entire range of soft and hard X-ray at a highest sensitivity ever achieved in the past.

1 Introduction

The fifth Japanese X-ray astronomy satellite, ASTRO-E, following *Hakucho*, *Tenma*, *Ginga*, and *ASCA*, is scheduled for launch in year 2000 by ISAS's new launcher M-V-4 from Kagoshima Space Center. This satellite will carry three experiments: the Hard X-ray Detector (HXD) covering the energy band from 10 keV to 700 keV; a micro-calorimeter array with an X-ray mirror (X-ray Spectrometer – XRS); 4 CCDs with 4 X-ray mirrors (X-ray Imaging Spectrometer – XIS). The latter two experiments cover the soft X-ray energy band with highest energy resolution (micro-calorimeter: $\Delta E \sim$ 12eV) or with medium energy resolution (CCDs: $\Delta E \sim$ 150eV at 5.9keV), both with imaging capability. All 3 experiments combined, Astro-E will become a spectrometer facility covering the energy band from 0.4 keV to 700 keV with moderate spatial resolution (XRT: $\sim$ 1 arcmin., HXD: $\sim$ 20 arcmin.).

The HXD has been jointly developed by scientists at Department of Physics, University of Tokyo, Institute of Space and Astronautical Sciences (ISAS), National Laboratory for High Energy Physics (KEK), and others. It is basically an upgraded version of the well-type phoswich counters successfully flown on balloons [1] [2] [3] [4] [5] [6]. Silicon PIN diodes are the important addition introduced to lower the energy coverage as well as to improve the energy resolution in the lower energy band. The design and characteristics of the HXD as of July 1995 are described here together with its expected performances.

2 Overview of the Mission

The total weight and power available for the 3 experiments will be about 1600kg and 500W, respectively. The M-V rocket will put the satellite into a near-circular orbit of radius 550km with an inclination of 31 deg. The 3 scientific instruments, the X-Ray Spectrometer (XRS) with an XRT (focal length ~3.5m), the 4 X-ray Imaging Spectrometers (XIS) with an XRT each (focal length ~4.5m), and the HXD, have several new and inovative technologies. The micro-calorimeter array made of 36 HgTe elements will be the first such instrument to fly in a satellite and will have the ultimate energy resolution (~12eV) in the soft X-ray band. The XRTs will be made of replica foils and will allow 1 arcmin. imaging despite their light weight. The X-ray CCDs will include one back-illuminated one and their energy resolution (fwhm) will be at the theoretical limit for CCD's: below 55eV at Oxygen K-line and 145 eV at 5.9 keV. The large (~ 20mm×20mm) 2mm-thick silicon PIN diodes used in the HXD will be the first of this kind and will give us energy resolution (fwhm) better than 3keV in the hard X-ray band. The well-type phoswich counter itself is a new concept and the scintillators that are now being developed for it include GSO(Ce) and YAP(Ce), newly developed fast and high light-yield inorganic scintillators [7] [8] [9] [10]. These instruments will be prepared by ISAS, Univ. of Tokyo, Osaka Univ., Kyoto Univ., Nagoya Univ., Tokyo Metropolitan Univ., RIKEN, NASA/GSFC, Univ. of Wisconsin, and MIT.

The time schedule of the project is as follows: the R/D phase to complete in Japanese FY of 1994, the Engineering Model (EM) to be ready in mid 1996, the Flight Model (FM) to complete in mid-1998, the final assembly and tests to finish in mid 1999, and the Launch to come late in Japanese FY of 1999.

3 Hard X-ray Detector

3.1 Overview

The HXD detector assembly is schematically shown in Fig.1. The total weight of this assembly will be about 200kg. The HXD consists of 16 (=4×4) modular units and has an overall photon collecting area of about 350 cm^2. As seen in Fig.1, each unit is a phoswich counter made of a fast inorganic scintillator, YAP ($YAlO_3$:Ce) or GSO (Gd_2SiO_5:Ce) and the BGO active shield. A cosmic hard X-ray is detected as a clean-hit signal if their full energy is deposited in the fast scintillator or the detection part. The active collimation part of the BGO shield forms four deep-wells (350mm deep, ~25mm × 25mm in area), limiting the field of view to ~ 4° × 4°. The YAP (or GSO) crystals (four per unit, each measuring in area 24 mm×24 mm and in thickness 20 mm for YAP or 5-7 mm for GSO) are glued at the bottom surface of the wells. The whole assembly is viewed by a common 2-inch phototube from the exterior surface of the shielding part.

The silicon PIN diodes, each measuring 22 mm×22 mm×2 mm, are also buried in the deep BGO well just above the detection part. Two PIN diodes are placed per well, making a total of eight per unit. The HXD contains 128 Si diodes in all, achieving a photon collecting area of about 240 cm^2.

Each unit will be equipped with an independent high voltage supply for the phototube, and a common DC power supply for 8 PIN diodes. Fig. 2 shows the effective areas of the phoswich scintillators and the silicon PIN diodes, in the respective energy ranges of 40-700 keV and 10-70 keV.

The 4×4 matrix of phoswich counters are surrounded by 20 units of thick BGO anti-counters for additional shielding. Furthuremore, a fine collimator made of phosphor bronze

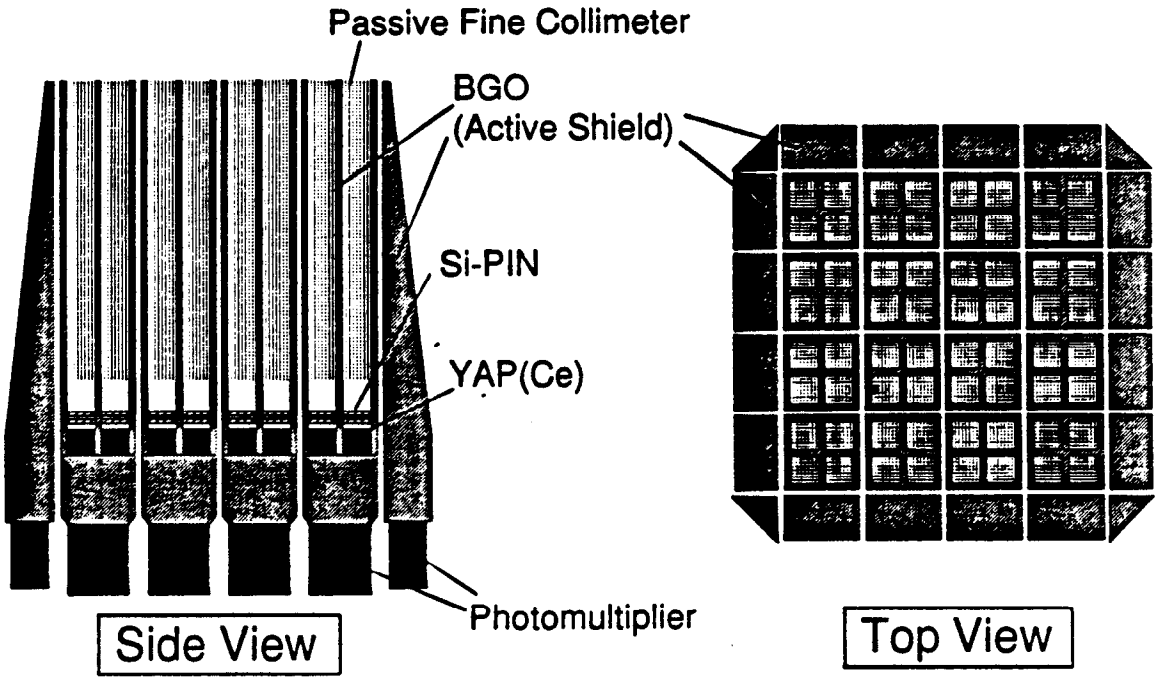


Figure 1: Schematic views of the HXD.

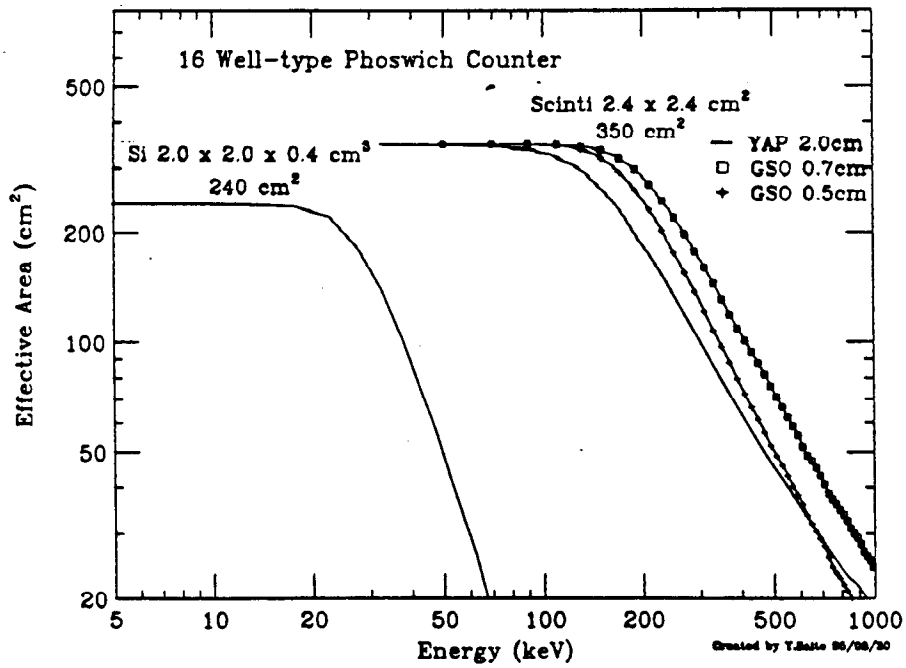


Figure 2: Effective area of the phoswich counters and PIN diodes. Three options are shown for the phoswich counters: YAP 20 mm thick, GSO 5 mm thick, and GSO 7 mm thick.

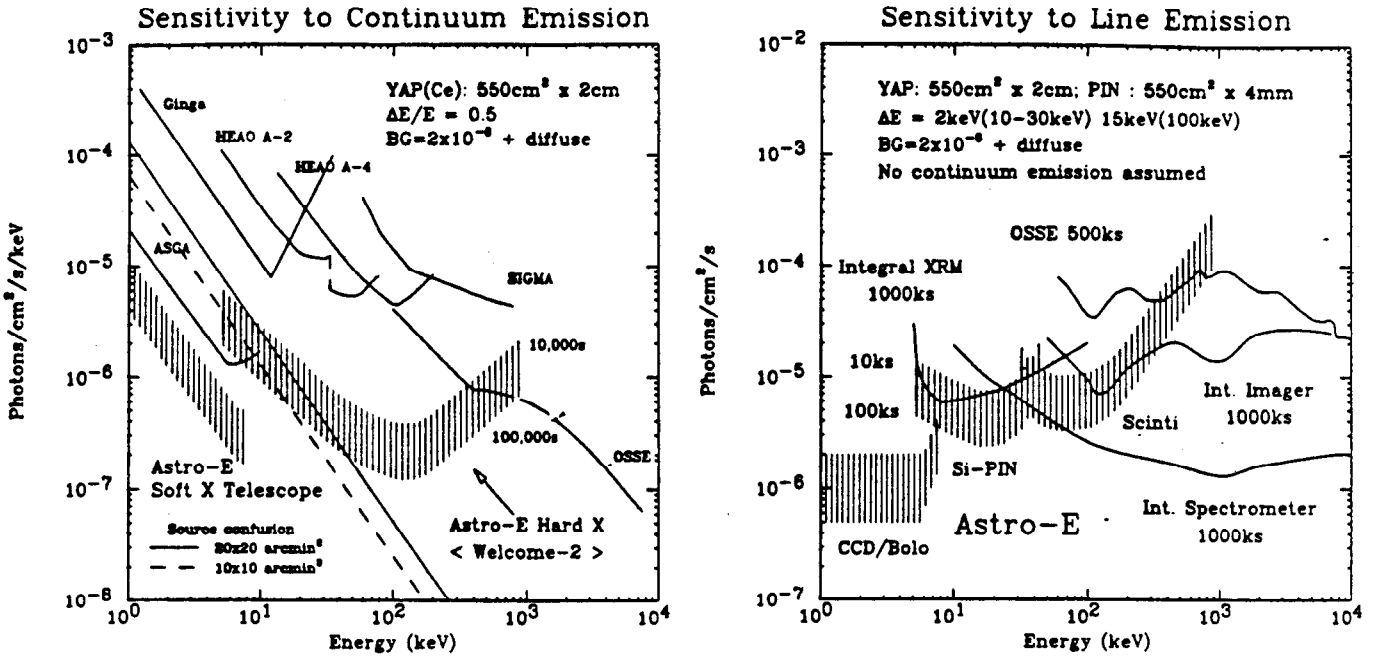


Figure 3: Sensitivities (3σ) to the continuum (left) and line (right) emissions of selected missions. The geometrical area of the HXD is assumed to be 350 cm^2 with 20 mm thick YAP for the phoswich counters and 240 cm^2 for the PIN diodes.

sheet (50μm thick) is placed inside the BGO wells to match the HXD field of view to that of the soft X-ray telescopes ($\sim 0.5°$ in diameter). This collimeter is expected to reduce the the cosmic diffuse X-ray background that may otherwise become a dominant background source for the PIN diodes. In the soft γ-ray band, background is reduced by mutual anti-coincidence among neighboring modules. We expect the detector background to be below 10^{-6} c/s/cm^2/keV for the Si PIN diodes and $\sim 5 \times 10^{-6}$ c/s/cm^2/keV for the scintillators.

In Fig. 3, we have compared the expected 3σ sensitivities to the continuum and to the line emission for the Astro-E HXD with those achieved by the past satellite missions. The sensitivity of Astro-E HXD will is be more than one order of magnitude higher than any other past missions between 10 keV to several 100 keV. We therefore expect to detect and study many new cosmic hard X-ray sources.

3.2 Well-type phoswich counter

The phoswich counter consists of two kinds of scintillators whose scintillation decay times are distinclty different. The faster scintillator is placed in the front as the detection part and the slower one in the back as the shielding part. Phototube signals generated purely by the faster scintillator are selected by using an appropiate pulse-shape discriminator (PSD). Signals with an appreciable contribution by the slower scintillator, eg. those by hard X-rays scattered by the shielding part and those by charged particles penetrated through the shielding part, are efficiently rejected (see Sec.4.3). This phoswich technique has been used for many years in cosmic γ-ray detection. The uniqueness of the "well-type" phoswich counters is that the well-shaped shielding part acts also as an active collimator and that each counter anti-coincidences out unwanted hard X-rays down to a low energy deposition (50 – 100keV) unit by unit [1] [3]. The detection part (the fast scintillator) being buried deep in the active

Table 1: Characteristics of NaI, BGO, GSO and YAP

Chemical composition	NaI(Tℓ) NaI (Tℓ)	BGO $Bi_4Ge_3O_{12}$	GSO(Ce) Gd_2SiO_5 (Ce)	YAP(Ce) $YAlO_3$ (Ce)
Eff. atomic number	50	74	59	35
Density (g/cm^3)	3.7	7.1	6.7	5.5
Rad. length (cm)	2.6	1.2	1.4	2.6
Index of refraction	1.85	2.15	1.9	1.94
At around 20°C				
Decay time (ns)	~230	~300	~60	~30
Light yield (relative)	100	~12	~28	~35
Peak emission (nm)	410	480	430	347
At around −20°C: data on GSO and YAP are preliminary.				
Decay time (ns)	~500	~600	~80	~30 and > 500
Light yield (relative)	~75	~15	~30	~20
Peak emission (nm)				
Reference	[14] [15]	[16]	[7] [8]	[9] [10]

anti-coincidence well also reduces efficiently background due to nuclear activity. The details on the development of the well-type phoswich counter are given in [1] [2] [3] [4].

Choice of the two scintillation materials becomes important in reducing the background and improving energy resolution. For its large effective atomic number and long scintillation decay time, BGO emergied as our first choice for the shielding part. As we tested a prototype well, we noted that BGO scintillators commercially available at that time were contaminated at a level barely tolerable by a radioactive isotope ^{207}Bi [11] [3]. A previous study had existed finding that the amount of contamination largely depends on where the Bi ore comes from [12]. The BGO scintillators now commercially available have substantially reduced ^{207}Bi contamination [13].

Radioactive contamination in the detection part contributes to the background more directly and should be absolutely minimized. Radioactive contamination has been measured for two high-light-yield scintillation materials with fast decay times GSO(Ce) and YAP(Ce) (see Table 1) [1] [3] and possible activation has been studied by irradiating protons with a kinetic energy (~100 MeV) typical in the satellite orbit [17] [18]. The study has shown that the number of long-life line γ-rays in the energy range of HXD is comparable for 2cm thick YAP than for 5-7 mm GSO. Note that the radiation length is quite different for the two crystals (see Table 1). We found recently that scintillation light yield decreases as temperature drops below 0°C for YAP but increases for GSO. We also found that a slowly decaying scintillation component appears below 0°C for YAP. We are continuing our study on GSO(Ce) and YAP(Ce) so that we can choose the best scintillator for the HXD by the end of 1995.

Possible in-orbit backgrounds have been estimated for the present counter design: YAP of $2.4 \times 2.4 \times 2.0$ cm^3 or GSO of $2.4 \times 2.4 \times (0.5 - 0.7)$ cm^3 shielded by BGO of ~ 5 cm thickness all around. The proton flux has been assumed to be that of the model given in [18] at solar minimum. The results are summarized in Table 2: one can see that the internal and cosmic-ray induced radioactivities will be the dominant source of background below 300 keV. We therefore anticipate that they will limit the ultimate sensitivity of the well-type

Table 2: Expected background rates

In-orbit activation	:	$\sim (0.5 \sim 1) \times 10^{-5}$ /sec/cm^2/keV at 100 keV
Leak-thru γ-rays	:	$\sim 1 \times 10^{-6}$ /sec/cm^2/keV at 300 keV
		$\sim 2 \times 10^{-5}$ /sec/cm^2/keV at 400 keV
		$\sim 2 \times 10^{-4}$ /sec/cm^2/keV at 500 keV
Off-aperture CXB	:	$\sim 7 \times 10^{-6}$ /sec/cm^2/keV at 50 keV
		$\sim 1 \times 10^{-6}$ /sec/cm^2/keV at 100 keV
Radioactive impurities	:	$0.5 \sim 1 \times 10^{-6}$ /sec/cm^2/keV

phoswich counters.

3.3 Silicon PIN diode

Li-drifted silicon diodes can be made a few mm thick and have been used as X-ray detectors in past satellite experiments. We have decided, however, to use high-purity PIN silicon diodes because the Li-drifted diode must always be kept at liquid nitrogen temperature and suffer from long-term deterioration.

The PIN diodes are imbedded at the bottom of the deep BGO wells in front of the detection part of the phoswich conters. Softer X-rays (below 40 – 70keV) will be photo-asborbed in the two layers of PIN diodes, while harder photons pass through the diode layers and reach the YAP or GSO crystal.

The PIN diodes are introduced to fill the possible gap in energy coverage between the phoswich detectors ($\geq$ 50 keV) and the CCDs ($\leq$ 10 keV). The BGO well provides the PIN detectors with very low background environment and the diodes act as anti-coincidence shield for the scintillator against low-energy charged-particles.

There are several technically critical issues in developing thick PIN diodes. One needs ultra high purity ($\sim$ ultra high resistivity) silicon wafers that give little volume leakage current. Even with such wafers, extreme care must be taken in the diode fabrication process not to increase edge leakage current. Sample diodes with thickness of 1-1.5 mm have been produced by Hamamatsu Photonics and Micron. We plan to develope 2 mm thick diodes and stack two of them to obtain 4 mm effective thickness. To reduce the leakage current to a reasonable level (< a few nA), we plan to operate the diodes at around -20°C.

Improving the energy resolution is the most demanding issue in developing our large area thick silicon PIN diode. We currently set our goal at an energy resolution of ΔE(FWHM) $\simeq$ 2.5 keV at -20°C. For the scintillator, the goal is ΔE(FWHM) $\simeq 7\%/\sqrt{\mathrm{E\ (MeV)}}$ keV in the temperature range.

For PIN diodes the energy resolution is determined by electronic noise introduced in the amplification system. Assuming external noise is absent, the energy resolution is determined primarily by the diode leakage current I_d, the input capacitance C_{in}, and the transconductance g_m of the input FET. The root mean square (rms) value of output noise voltage is expressed as

$$v_{rms}^2 \simeq \alpha \cdot I_d \cdot \tau_s + \beta \cdot \frac{C_{in}^2}{g_m} \cdot \frac{1}{\tau_s} + \gamma$$

where τ_s is the shaping time of the amplifier, α and β are constants, and γ represents contribution of the 1/f noise. In our case C_{in} is the sum of the junction capacitances of two PIN diodes ($\sim 2 \times 20$ pF) and the capacitance of the cables connecting the diodes to the

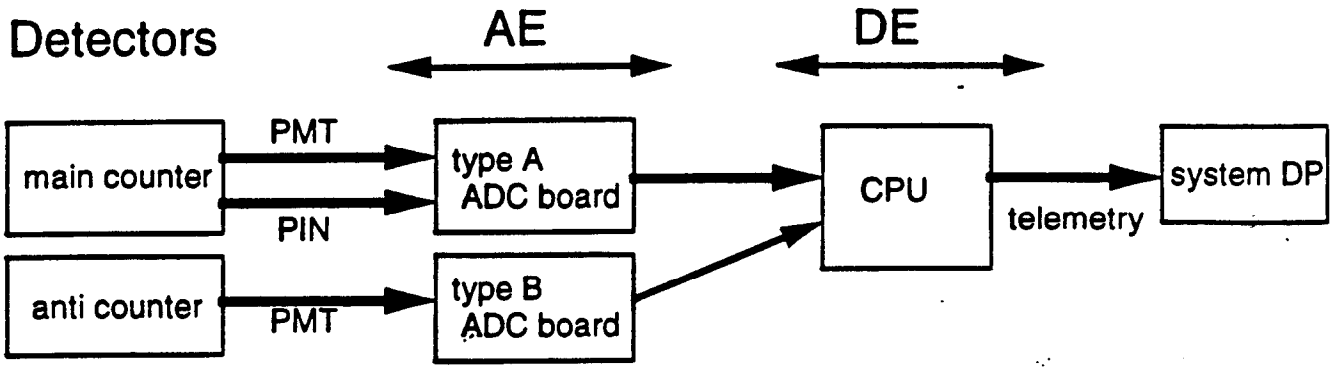


Figure 4: Flow chart of the on-board data processing for the HXD.

preamplifiers ($\sim$ 30 pF) [19] [20]. To meet our goal, we need a low capacitance cable, an FET with very large g_m and PIN diodes with very small I_d: they are important targets of our research and development efforts.

4 On-Board Signal Processing

4.1 Overview

A simplified schematical diagram of the on-board electronics system for the HXD is shown in Fig.4. It consists of the analog electronics part (AE) and the digital electronics part (DE). Sensor signals reach AE via 116 independent coaxial signal cables, 6 lines for each of the 16 phoswich units and 1 line for each of the 20 anti counters. The 6 lines from a phoswich unit are for the anode and the last dynode signals of the phototube and 4 sets of PIN diode pairs. Here 4 diode signal lines and the phototube dynode signal line are sent to their respective charge preamplifers while the phototube anode signal line is sent directly to the PSD.

In the AE part, the pulse-shape discriminator (PSD) plays the most crucial part: it filters out hard X-ray signals whose time profile is consistent with that of the scintillator used in the detection part (YAP or GSO) not contaminated with slow-decaying component of BGO. These "clean-hit" events are sent to the DE part and analyzed by on-board CPUs and transmitted to the ground via telemetry. The PIN diode siganals are digitized independent of each other if no slow-decaying scintillation light is detected by the PSD for the phoswich unit the diodes are housed.

4.2 Analog data processing

The analog electronic (AE) part consists of 8 ADC boards classified into two types, type A for the phoswich counters and type B for the anti counters.

One ADC board of type A reads signals from 4 phoswich counters. In total four type A boards are required, each with four PSDs and related circuits as shown in Fig.5. The output from the phototube dynode is first fed to a charge sensitive preamplifier and then to a PSD circuit, while the anode output is used as a fast trigger to generate a peak-hold gate signal of PSD.

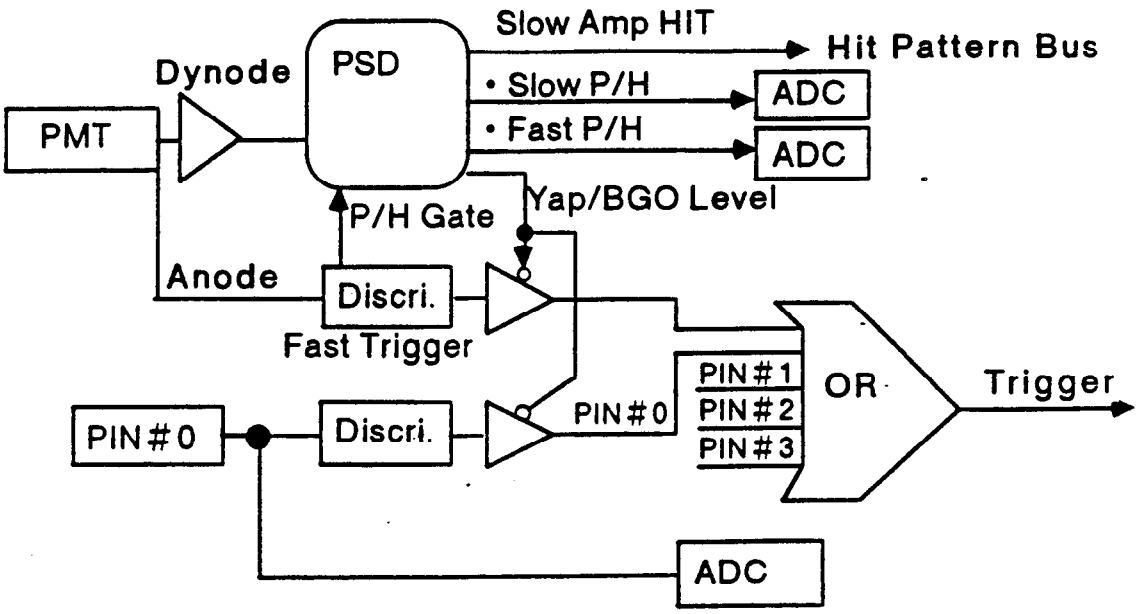


Figure 5: A schematic view of one unit on a type A board including PSD, 6 ADC chips and other related circuits.

The type B ADC board reads signals from 5 BGO anti counters, making the total type B boards four. This board also outputs hit pattern information to the hit pattern bus. Furthermore, two kinds of histograms are produced for each anti counter: one pulse-height-analyzed in 4 energy bands every 10 msec to monitor γ-ray bursts and the other in 32 energy bands at 512 msec interval to study transient sources.

4.3 Pulse Shape Discriminator

The PSD distinguishes events from the two scintillators, the fast-decaying scintillator (YAP or GSO) and the slow-decaying scintillator (BGO). There are several PSD methods developed and used in γ-ray astronomy and nuclear physics, one of which is the double shaping method adopted in the HXD. In this method, the signal from a phototube is integrated with two different shaping times and these two output pulse heights are compared. This method is less affected by external electronic noise, while its demerit is that the circuit becomes somewhat complicated. This demerit can be solved by implementing the circuit on an LSI chip [21]. In our LSI chip, the two integration times are set at τ_{fast} = 100 ns and τ_{slow} = 500 ns (see Fig.6).

The outputs of the two shapers have different pulse heights for BGO signals, while they are almost equal for pure YAP (or GSO) signals (Fig.7). Currently, the pulse-shape discimination LSI is being developed using a semi-customed LSI and the YAP/BGO phoswich combination. Shown in Fig.8 is the result of a test done by using the first prototype LSI circuit: One can see that pure YAP events are seperated clearly from BGO events and Compton scatterd events.

4.4 CPU processing

When the on-board CPU is interrupted by an ADC trigger signal, the data are collected and analyzed with a 32-bit processor 80386 (and a co-processor 80387) operated at a clock rate of 8MHz with a memory of 512 kB.

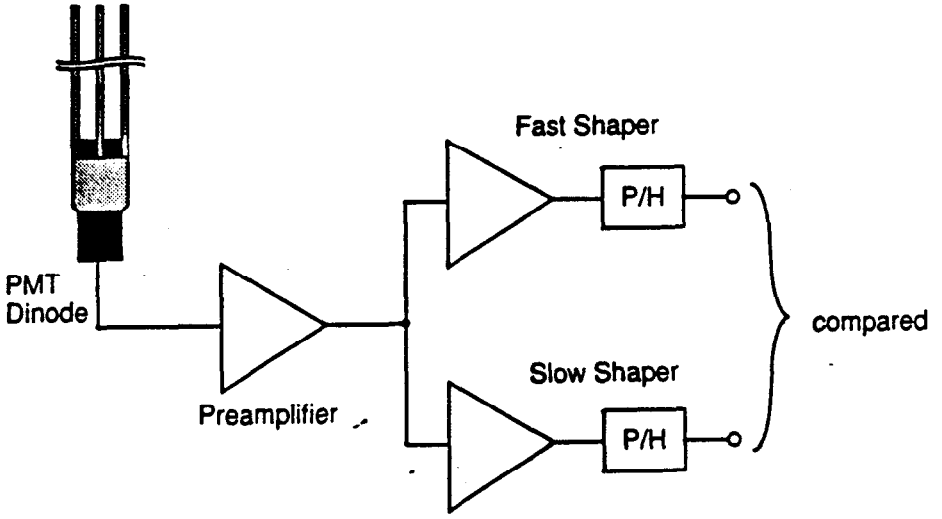


Figure 6: The double shaping method employed in the HXD.

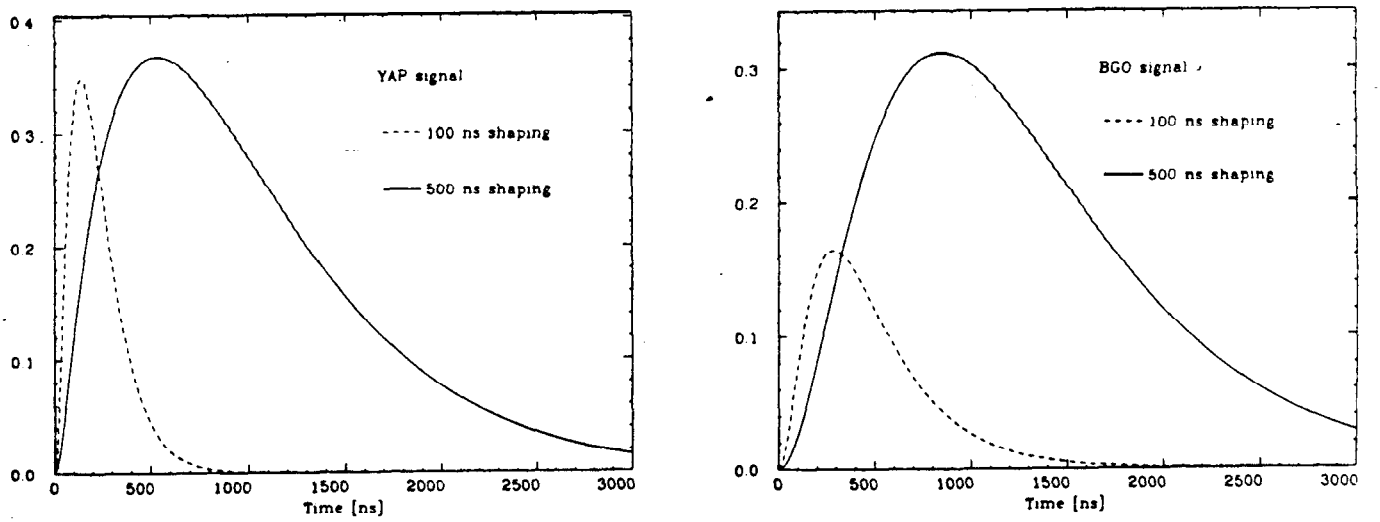


Figure 7: Output of two shaping amplifiers with different time constants (100 ns and 500 ns). The two traces in the left figure are for YAP or GSO signals and those in the right figure for BGO signals.

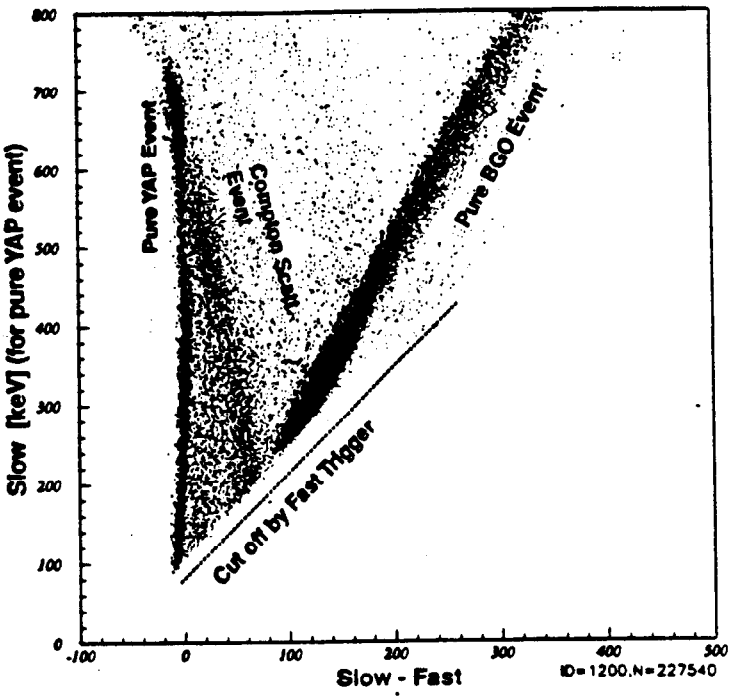


Figure 8: Test results of the first PSD LSI. The abscissa is the difference of the two pulse heights (τ = 100 ns and 500 ns) and the ordinate the pulse heights of the slower shaper.

The CPU reads hit pattern information of all the 32 counters and if a HIT signal exists in the surrounding units, rejecting Compton scattered events and particle interaction events. Thus, the "compound-eye" configuration combined with the well-type phoswich concept are expected to work together in reducing the external background caused by the off-aperture X/γ-rays and particle interaction in the counter, as well as reducing the internal background caused by the $\beta - \gamma$ decays of radioactive nuclei.

5 Future Plan

The construction of the engineering model HXD for ASTRO-E is now underway. Design, fabrication and testing of the proto-model HXD is scheduled in 1995 and 1996. The proto-model will probably consists of 4 main counters, 5 anti counters, one type A ADC board, and one type B ADC board. The important ground programs before the launch include acutal calibration of the detector and establishing in-flight calibration procedures.

References

[1] Kamae, T., et al.,*Proc. SPIE*, **1734**, 2 (1992)

[2] Takahashi, T., et al.,*Proc. SPIE*, **1734**, 44 (1992)

[3] Kamae, T., et al.,*IEEE Trans. Nucl. Sci.*, **40**, 204 (1993)

[4] Takahashi, T., et al., *IEEE Trans. Nucl. Sci.*, **40**, 890 (1993)

[5] Gunji, S. et al., *Astrophys. Jour.*, **397**, L83 (1992)

[6] Gunji, S. et al., *Astrophys. Jour.*, **428**, 284 (1994)

[7] Takagi, K. and Fukazawa, T., *Appl. Phys. Lett.*, **42**, 43 (1983)

[8] Ishibashi, H., Shimizu, K., Susa, K., and Kubota, S., *IEEE Trans. Nucl. Sci.*, **36**, 170 (1989)

[9] Baryshevsky, V. G. et al., *Nucl. Instr. Meth.* **B58**, 291 (1991)

[10] Korzhik, M. V., Miesvich, O. V., and Fyodorov, A. A., *Nucl. Instr. Meth.*, **B72**, 499 (1992)

[11] Nakao, K., Senior Thesis, Dept. of Physics, Univ. of Tokyo (1990)

[12] Lewis, T. A., *Nucl. Instr. Meth.*, **A264**, 534 (1987)

[13] For example, BGO crystals supplied by Crismatec.

[14] Werkheiser, A. H. and Miller, T. G., *Nucl. Instr. Meth.*, **75**, 167 (1969)

[15] Schweitzer, J. S. and Ziehl, W., *IEEE Trans. Nucl. Sci.*, **30(1)**, 380 (1983)

[16] Zucchiatti, A. et al., *Nucl. Instr. Meth.*, **A281**, 341 (1989)

[17] Matsuzaki, K., Master Thesis, Dept. of Physics, Univ. of Tokyo (1995)

[18] Stassinopoulos, E. G., in Rester, Jr., A. C. and Trombka, J. I. (eds.) "High-Energy Radiation Background in Space", AIP Conference Proceedings 186 (1989) p.3

[19] Kaneda, H., Master Thesis, Dept. of Physics, Univ. of Tokyo (1994)

[20] Tamura, T., Master Thesis, Dept. of Physics, Univ. of Tokyo (1995)

[21] Ezawa, K., Master Thesis, Dept. of Physics, Univ. of Tokyo (1995)

Working Group Talks

Included in this section:

GLAST SPACECRAFT GROUP REVIEW

ROBERT J. TWIGGS
AERONAUTICS AND ASTRONAUTICS DEPT.
STANFORD UNIVERSITY

1. Mission Goals

The mission goals required by the spacecraft are to (1) meet the science needs, (2) meet cost requirements to assure funding for the spacecraft, (3) assure that the spacecraft can use a variety of launch vehicles, and (4) provide the maximum access to the science of interest.

1.1 Meeting the Science Needs--Mission Baseline Requirements

The mission life requirement is 3 years minimum. This provides the opportunity to collect useful data and to gain sufficient experience with the spacecraft operations that the major operations activities are devoted to the science mission.

The orbital parameters that can be met directly from the launch vehicle and not require additional orbital adjustments by the spacecraft are for an altitude of 650 km at an inclination of 28.5 degrees. The 28.5-degree inclination will provide science observation near the equator, looking into the galactic plain, and is the most efficient launch orbit from Cape Kennedy in Florida.

The initial baseline launch vehicle is the Delta-2. This vehicle was initially selected since it can provide the launch for the GLAST satellite as a primary payload similar to the ARGOS satellite. It is also one of the more cost effective vehicles.

1.2 Meeting the Science Needs--Payload Requirements

The payload size is expected to be 1.8 meters x 1.8 meters x 0.85 meters. The initial payload mass allocation is 3000 Kg with a goal of 2500 Kg, leaving a 500-Kg margin. The field of view provided by the GLAST instrument is 7 Sr, thus requiring an unobstructed mounting to the satellite bus.

The instrument power requirement is estimated at 600 watts with the tracker requiring 300 watts, calorimeter 200 watts, and data acquisition 50 watts, leaving a margin of 50 watts. The design objective would be to work toward a total power consumption of 500 watts, leaving a 100-watt margin.

The thermal requirement of the payload is for the Si temperature to be less than or equal to a 20-30 degree centigrade range. A design objective is to not exceed 30 degrees centigrade at the mission end of life. The power dissipation in the instrument is approximately 150 microwatts/preamp and 150 microwatts/channel for other electronics. The total dissipation/channel is then expected to be 300 watts. The thermal control goal for GLAST is to use passive thermal control methods with cold biases, radiators and heaters.

1.3 Spacecraft Requirements

The selection of the Delta-2 for launch would provide the capability of attaining a 650-km orbit at 28.5 degrees in inclination with a lift-to-orbit capability of 4700 Kg.

The power budget requirement objective for the satellite is 400 watts for the spacecraft, with a goal of 300 watts to allow a 100-watt margin. The payload requirement is for 600 watts.

The attitude requirement for the spacecraft is a knowledge of attitude to within one arc minute. The control requirement is to within 10 degrees of final orientation of an antinadir direction. The amount of jitter is to be determined . ** Discussions after the conference lead to an additional consideration of inducing a north-south scanning of the satellite at a period of several orbits to gather data for an all-sky scan towards the earth. The projected data collection rate would be 60 kilobits per second at 100% duty cycle. This would require on-board data storage and data retrieval by methods of independent ground station(s), use of the NASA satellite data network, or possible use of the TDRSS geostationary NASA data transfer satellite.

1.4 Operational Requirements

The operational data requirements for downlink are: 1/4/5 Mbps for the science data. The allows for multiple transmission rates depending upon the amount of data needed to be transferred to the ground data collection centers. A slower mode of 128 Kbps would be used for real-time housekeeping operations on the satellite bus and for a quick look at science data without requiring the complete on-board stored details. Also a 4 Kbps downlink could be provided with an omni directional antenna for initial on-orbit operation and for the spacecraft operational housekeeping data.

The uplink requirement is 2 Kbps, which could be used for reprogramming upload operations as well as normal command and control operations.

1.5 Objective in Meeting the Spacecraft Cost Requirements

The design objectives for the satellite are geared to keeping the costs low so as to increase the program funding probabilities. The spacecraft design will be kept simple by allowing wide margins for control, using a passive thermal system, and allowing 2-30% margins for power and weight considerations. Using a known, proven launcher such as the Delta-2 reduces the launch risk.

2. CONCLUSION

The spacecraft should be designed in conjunction with the instrument development to provide for optimum tradeoff of science performance and the performance requirements of the spacecraft, thus ensuring increased reliability and lower costs.

TRACKER WORKING GROUP SUMMARY

Robert Johnson
UC Santa Cruz

The tracker working group concentrated on issues related to silicon strip detectors, but there were two talks on other tracking technologies. T. Kashiwagi gave a talk on lithium-drifted silicon detectors that utilize surface barriers. These are very large detectors, with wide pitch, and are fabricated in-house at low cost. The lithium-drifted silicon allows full depletion of 2-mm-thick detectors at reasonable voltages, so the signal is large. These detectors are especially well suited for construction of calorimeters with very good spatial resolution.

K. Kasahara gave a presentation during the joint meeting with the simulations group. He presented work done with scintillating fiber trackers and calorimeters, which are read out by image intensifiers. The detectors have been tested in particle beams and in balloon flights and are being used in a gamma-ray telescope and in an experiment for detection of primary cosmic ray electrons.

1. Data Acquisition and Cooling

Data acquisition was discussed during a joint meeting with the DAQ working group. A. Colavita presented an innovative proposal for a data-push architecture. There would be no trigger. Instead, each channel would time-stamp hits and pass them onto a data stream that flows into a chip containing a hardware sorting algorithm. All hits within a tower would be sorted by time into events. Data from multiple towers would be collected together in much the same fashion. Simple cuts would be made on the data along the way to keep the data volume at manageable levels. This approach seems to be well-suited for the modular GLAST concept.

Otherwise, the working group concentrated on issues connected directly with silicon strip detectors and their use in a pair-conversion telescope. The discussions were mostly restricted to the detector wafers themselves and the front-end chips required to amplify the signals. There are, of course, many other engineering issues involving the construction and integration of the modules, but there were no participants in this group who were working on such issues. Nonetheless, there was some discussion of cooling. The precise temperature of the amplifiers is not critical, since the noise only goes up as the square root of the absolute temperature. However, stability of the comparator threshold is crucial. The threshold depends on the matching of transistors in the comparator and, for most designs, the second gain stage, but it also depends on the gain of the amplifier. To the extent that these effects are sensitive to temperature, the stability of the temperature over time is important. Within the interior of a large GLAST-

scale instrument, the temporal temperature variations would be negligible, as long as the cooling system is always functioning.

The absolute temperature of the detectors is important, since the leakage current depends exponentially on temperature, and the shot-noise of the current is proportional to the square root of the current. Therefore, one wants the detectors to be as cold as possible. Although they themselves produce very little heat, they must be located very close to the amplifiers and therefore will equilibrate at the temperature of the cooling structure on which the amplifiers are mounted. The cooling would have to be done by conducting heat down the sides of towers, past the calorimeter, to radiators on the back of the instrument. The use of heat pipes that rely on vaporizing a liquid would be ideal, since then the temperature could be maintained everywhere at the same constant value. Conduction through a metal or ceramic structure running along the sides of the towers would be much simpler but would probably result in the front silicon layers being on the order of 20 degrees warmer than the back layers.

2. Radiation Damage

J. Krizmanic gave a presentation on expected levels of radiation dose as a function of orbit parameters and the amount of shielding. He found that the dose is severe for an unshielded detector but that only minimal shielding of, for example, 0.2 gr/cm2 is necessary to reduce the dose to a reasonable level of 0.3 kRad/yr for an equatorial orbit or 10 kRad/yr for a polar orbit. R. Johnson reported on some preliminary measurements of radiation damage made at UCSC on a GLAST prototype detector manufactured by Hamamatsu. The detector is unusual in that it has a large strip pitch (300 mm) and a narrow strip width (30 mm). With only a 10-kRad exposure to ionizing radiation from a Co-60 source, the leakage current measured on individual strips was found to increase by a factor of nearly 20, from about 3 nA per strip up to 50 nA per strip. A similar increase took place upon increasing the exposure to 20 kRad. Judging from earlier work, less than a nA increase could be accounted for by considering damage to the bulk silicon, so the excess current must be generated at the surface. There was some speculation that the large interstrip area of the detectors might aggravate this effect, but more work would have to be done to understand it. However, even with such a large rate of increase in leakage current, in an equatorial orbit only the veto layer of GLAST would be at risk of unacceptable noise levels from leakage current, if it were left unshielded. Regarding the CMOS front end electronics, there is no problem with radiation damage, since radiation hardened CMOS processes that can withstand greater than 1 MRad of exposure are widely available and in use within the particle physics community.

3. Front End Electronics

Geoff Mills presented work done at LANL on design of a very low power front end chip. They have designed and submitted for fabrication a prototype amplifier-discriminator CMOS chip that is expected to use only 100 mW of power per channel while achieving a noise performance with a shaping time of 1ms of 890 electrons equivalent noise charge for a detector capacitance of 12pF. They are also working on a design of a digital readout section. R. Johnson presented noise measurements made at UCSC of a preamplifier obtained from LBL (the preamp for the SVX-II chip). For a power consumption of 200uW in the front-end FET only, one has with a shaping time of 1ms ENC=83+14*C electrons, with C the detector capacitance. If the power in the front end FET is cut in half, then the noise increases by about 30%.

4. Detector Optimization

The working group held several discussions on optimization of the silicon detectors for a GLAST-type tracker. Several geometric parameters need to be determined, namely the strip width, strip pitch, the readout pitch, the thickness, and the strip length. From the standpoint of minimizing the capacitance, it is desirable to minimize the ratio of strip width to pitch. However, the strips must be sufficiently wide such that the AC coupling capacitance is large compared with the detector capacitance and such that the series resistance of the strip does not contribute to the noise. Furthermore, recent measurements made at UCSC on GLAST prototype detectors with 300-mm strip pitch and 30-mm width indicate poor charge collection in the region between strips. That is probably due to the electric field lines in the intermediate region pointing toward the surface rather than toward the strips. The motion of the holes toward the surface then does not generate very much current on the strips, and if one then assumes that the holes stick to the surface for some length of time the net signal seen from holes is very small. Therefore, to have a 300-mm readout pitch, the strips should either be made quite wide (>150 mm), or else the strip pitch should be reduced, with adjacent pairs or triplets of strips ganged together. This work has been summarized in the preprint SCIPP-94/31 at UCSC.

The readout pitch is the subject of many optimization questions. First, to find the optimum value for measurements one must consider both the angular resolution and the two-track separation. Second, one can optimize the power requirements versus noise performance. The capacitance decreases with decreasing pitch, until the interstrip capacitance begins to dominate (below a pitch of about 150 mm for very narrow strips and 300 mm thick detectors). Furthermore, the equivalent noise charge (ENC) varies linearly with the capacitance. Therefore, increasing the pitch to a large value can have the result of requiring more power in the front-end transistor per channel in order to maintain the required noise performance.

A similar issue is the strip length. In this case the capacitance is simply proportional to the strip length, so the ENC of a given amplifier varies linearly with the length. There was some disagreement at the workshop on exactly how the amplifier power requirements, for a given noise level, would scale with the capacitance. Considering only the input FET, the noise scales inversely with the square root of the transconductance, gm. From textbooks, gm should be proportional to the square root of the drain current, and hence the power. That is generally correct for a given transistor operating in the strong inversion region. From that, one would predict that the noise would scale inversely with the fourth root of the power, which would indicate that doubling the capacitance would require 16 times the power to get back to the same noise level. In fact, the scaling is not quite that simple. First, the noise is not directly proportional to the capacitance, since there is always a constant term. Second, when an amplifier is designed for a larger detector, the width of the input FET is increased proportionally to the size of the detector in order to optimize the noise performance. Because of that, the current density Id/W (drain current divided by width) tends to stay at a constant value, with the optimal value being around the transition region between week inversion (where gm is proportional to Id) and strong inversion (where gm is proportional to the square root of Id). In a plot of gm/Id versus Id/W, transistors of vastly different dimensions tend to fall on a universal curve. Therefore, if one goes from a small detector to a large detector and at the same time redesigns the front end chip appropriately, keeping Id/W the same for both designs, then the ratio of gm between the two designs would be the same as the ratio of Id.

Thus the most appropriate scaling to use may be to assume that the noise goes inversely with the square root of the power in the front end FET. To get some idea of what this means, we can take the LANL simulation of their front end design, with a noise performance given by ENC=450 + 37*C electrons, for a total of 890 electrons with a detector capacitance of C=12 pF. The power used is about 100 mW, equally divided between the first FET and the remainder of the circuit, not including any digital readout. Assuming that the digital circuitry requires only another 50 mW per channel, then the first FET uses 33% of the power. In this situation, if the detector capacitance is doubled to 24 pF, the ENC would increase by 50%. To get the ENC back down to 890 electrons would require increasing the power in the first FET by a factor of roughly 1.5**2=2.4, assuming that the FET were redesigned to be optimal for the new input capacitance. The power per channel would then increase by 50%, but the number of channels would be reduced by a factor of 2, for an overall power savings of 25%. If the power scaling were taken instead to be the 4th power, then the overall power consumption would increase by about 20% in the same scenario. More detailed studies should be made to understand the scaling and the power figures better before deciding on the optimal strip length. However, other important factors must also be considered. In particular, decreasing the strip length leads to a smaller tower size and therefore more gaps and a larger fraction of dead mass in the system, plus a larger numbers of required cooling, power, and data transmission lines.

The thickness of the silicon is easier to optimize. The signal increases linearly with increasing thickness (and the noise decreases slightly due to a reduction in

capacitance), while the reverse bias voltage needed for full depletion increases quadratically with thickness. High energy physics experiments generally have used 300 mm thick detectors. However, the optimization in that case tends to be driven by minimization of multiple scattering, which is not such an issue for GLAST, where the silicon mass is small compared with that of the radiator. The working group concluded that 500 mm thickness would be a better choice for GLAST. For a typical silicon resistivity of about 5000 ohm-cm, only about 150 volts would be necessary to overdeplete the detectors. The extra thickness would also increase slightly the leakage current, due to the extra volume of depleted silicon, but that effect would be insignificant compared with the factor of 5/3 increase in signal.

5. Conclusion

For the future, more measurements of detectors and amplifiers will be carried out, along with simulations, to get the necessary information needed to optimize the designs of the individual detectors, the front-end amplifiers, and the detector modules. Prototype chips designed by LANL and A. Colavita will be tested and evaluated, and work will continue on chip design, of both the analog and digital sections, with the goal of achieving the necessary performance with as low power as possible. Then work must proceed on a host of engineering issues, including the mechanical design, cooling, power distribution, and electronics connections.

Calorimetry for a Next Generation Gamma Ray Telescope

David Strom*

Department of Physics

University of Oregon

ABSTRACT

Various calorimetry options for use with a gamma ray telescope are explored in light of the requirements imposed on the calorimeter by the demands of energy resolution, pattern recognition and pointing accuracy. As a baseline option, a homogeneous CsI(Tl) calorimeter is considered. The direction of the gamma rays is determined in a tracker/converter. Three longitudinally segmented options which all employ CsI(Tl) for the first 4 radiation lengths are considered. The use of longitudinal segmentation allows the direction of the gamma rays that do not convert in the tracker/converter to be measured, considerably increasing the sensitivity of the telescope at high energies above a few GeV.

1 Introduction

The success of the EGRET instrument aboard the Compton Gamma Ray Observatory in identifying galactic and extragalactic gamma ray sources provides ample motivation for a next generation gamma ray telescope to search for gamma ray sources [1]. Advances in technology since the conception of the EGRET mission will allow the energy reach of a new instrument to be extended well beyond that of EGRET (30 Mev to 20 GeV) into the region currently probed only by ground based Cherenkov telescopes (200 GeV and above). This region is completely unexplored and may hold many interesting discoveries [2]. The physics issues associated with

*Members of the calorimetry working group were: G. Barbiellini, J. E. Grove, K. Kasahara, T. Kashiwagi, A. Luebke, M. Merck, R. Ong (convener), M. Oreglia,D. Strom (convener), R. Taylor

the detection and characterization of these gamma ray sources are discussed elsewhere in these proceedings. This report explores various options for calorimetry in a next generation gamma ray telescope. First, we attempt to establish the requirements which a calorimeter must meet to be considered for use in a gamma ray telescope. This is followed by a discussion of four different calorimetry options. Three of these options employ longitudinal sampling and can be used to establish both the direction and energy of the incoming gamma ray. The addition of pointing capability to the calorimeter can dramatically improve the efficiency of the telescope for high energy gamma rays.

2 Detector Requirements

An ideal calorimeter would determine accurately the energy, shape, and direction of electromagnetic showers from gamma rays. While such a calorimeter can be constructed, the weight and power limitations imposed by satellite operation require that the parameters of the calorimeter be carefully optimized. The most important job of the calorimeter is to determine the energy of the gamma ray. This requirement is discussed first, followed by a discussion of background rejection and pointing resolution.

2.1 Energy Determination

The first requirement of the calorimeter is that it be sensitive to gamma rays of energies that can be reconstructed in the tracking section of the telescope. The lowest energy photons that can be reconstructed in the converter and will penetrate to the calorimeter are of order 20 MeV. On the upper end, the converter can reconstruct tracks from converted gamma rays of arbitrarily high energies. Here the needed sensitivity of the detector is set by the effective area of the telescope and the known gamma ray flux at very high energies from ground based imaging atmospheric Cherenkov observations [2]. For a calorimeter with an effective area of approximately 1 m^2, this corresponds to a few hundred GeV.

To measure the low energy gamma rays, it is essential that the first few radiation lengths of the calorimeter be instrumented with a homogeneous calorimeter such as

CsI(Tl) crystals. In most instances the longitudinal segmentation of a sampling calorimeter would be too coarse to allow efficient detection of gamma rays below 100 MeV.

While any calorimeter with more than a few radiation lengths will show an increase of signal with energy (see figure 1), the detection of high energy gamma rays is only useful if they can be unambiguously separated from low energy ones. In most cases, it is expected that the flux of gamma rays from a given source will be a falling continuum. For example, the photon spectrum (dN/dE) from the Crab Pulsar and Nebula falls as $E^{-2.13 \pm 0.04}$ [3]. One task of the calorimeter will be to determine the power with which the photon energy spectra for a given source falls, but perhaps the most important task will be to look for breaks in the spectrum. The only AGN to have been observed by ground based Cherenkov techniques is Markarian 421 [4], which is to be compared with almost 40 such sources observed in the GeV range by EGRET.[1] This is notable as many other, brighter extragalactic sources have been observed by EGRET in the GeV range. Thus it is imperative for a next generation gamma ray telescope to determine where the gamma ray spectra of the other AGNs cut off. The accuracy with which the breaks in these spectra can be determined will largely be a function of the effective area and energy resolution of the calorimeter. These breaks are of considerable physics interest. For example, it has been proposed that the cutoff between GeV and TeV energies observed in the spectra of AGNs may be due to absorption of the high energy gamma rays in photon-photon interactions with starlight [5].

Figure 2 shows Monte Carlo simulations of 50 GeV and 200 GeV gamma rays observed in 6 X_0 and 8 X_0 CsI(Tl) calorimeters. The distribution of angles of the incident gamma rays corresponds to that which would be observed by a satellite in survey mode. Because the gamma rays are required to convert in the tracker/converter and because most gamma rays hit the calorimeter at an angle, the energy resolution is somewhat better than might be naively expected. The apparent energy resolution (rms/mean) of the 6 X_0 calorimeter is 40%, however, due to the nonlinear energy response the actual resolution is somewhat worse. For the 8 X_0 calorimeter the energy resolution is approximately 30%. Note that the 8 X_0 calorimeter separates the 50 GeV gamma rays from the 200 GeV gamma much better than 6 X_0 calorimeter. The ability of the telescope to detect a break in the energy spectrum of a source will depend on the spectral index of the source and the available statistics. From figure 2 it is clear that for a steeply falling source

spectrum and small statistics, the 6 X_0 calorimeter could only detect the location of a break in the spectrum between 50 and 200 GeV with difficulty.

2.2 Background Rejection

The potential background rejection obtainable with the calorimeter has not yet been studied in detail. The most important role of the calorimeter will be (a) to provide an additional space point of the shower which can be used as a constraint on the direction of the gamma ray and (b) to provide some information about shower shape. For many of the proposed calorimeters the position information from the calorimeter will be in tower format, which will provide a true space point, as opposed to the information from the tracker which provides separate x and y projections.

At low energies the ability of the calorimeter to help with pattern recognition will be limited by the granularity of the detector. At high energies the differences between electromagnetic and hadronic shower shapes can be used to reduce the cosmic ray background. In high energy physics experiments, differences in the transverse shower shape between electrons and pions have been used to obtain rejection factors of 100 or more [6] for particles of the same momentum. Additional information about the longitudinal shower shape can be used if a sampling calorimeter is used or if the homogeneous calorimeter is segmented longitudinally. The longitudinal segmentation will be especially helpful for identifying background that enters the calorimeter from behind.

Because the background from cosmic rays is also rapidly falling, it may be possible to reconstruct gamma rays above 1 GeV without requiring that they convert in the tracker/converter. After requiring that a high energy shower (1 GeV or more) is seen in the calorimeter, the remaining cosmic ray background can be reduced by requiring that little activity be seen in the tracker and by using the reconstructed longitudinal and transverse shower shape. Since the calorimeter is likely to comprise less than one interaction length, most high energy cosmic rays will only deposit a small fraction of their energy in the calorimeter. Experience with electron identification in high energy physics [6] indicates that requiring that a large signal is seen in the first few radiation lengths in an electromagnetic calorimeter is extremely effective in reducing hadronic background. The rejection factors that can

be obtained in a gamma ray telescope will depend primarily on the rate of high energy cosmic rays that shower early in the detector.

2.3 Pointing Resolution

A dramatic improvement in the high energy effective area of the telescope can be obtained if the direction of the gamma rays can be reconstructed in the calorimeter. Figure 3 shows a simulation of the possible improvement in the effective area when a pointing calorimeter is incorporated into a gamma ray telescope. This improvement is especially important as the flux of gamma rays from most sources is expected to fall rapidly with energy. The large gain in effective area is possible because only about 30% of all gamma rays will convert in a 1 X_0 converter. While it is possible to increase the thickness of the converter in each layer of the detector or to increase the number of layers in the tracker, both of these options would lead to an unacceptable response at low energies.

It is important to note that the angular resolution of the calorimeter need not be as good as that of the converter portion of the telescope. For many purposes an angular resolution of a few degrees will be enough to associate a gamma ray with a known source. In those instances when a better pointing accuracy is required, only the gamma rays that convert in the tracker portion of the telescope can be used.

The use of sampling calorimeters has been explored by both the GLAST (A Gamma-ray Imaging Large Detector for Astronomy) group in the US [7] and by the GILDA (a Gamma-ray Imaging Large Detector for Astronomy) group in Europe [8]. Pointing calorimetry has been demonstrated by a number of calorimeters used in high energy physics and astrophysics experiments. [9,10,11]. These calorimeters sample the shower at several different longitudinal positions and obtain the direction of the shower from the evolution of the transverse shower position as a function of depth. Figure 4 shows the radial position resolution of the OPAL silicon-tungsten luminosity monitor [6] as a function of shower depth for 45 GeV showers. This detector has radial strips with a width of 2.5 mm. Near shower maximum, radial resolutions of better than 200 microns are obtained. At shower maximum still better resolution could be obtained if narrower strips were employed or if the spacing between layers was increased to allow the transverse size of the shower to increase.

The angular resolution of the calorimeter depends critically on the spacing between the layers. Moving the layers apart broadens the showers to allow their position to be reconstructed with a coarse pitched sampling and it increases the level arm used to determine the angle. The best pointing resolution at low energies obtained thus far is that of the WiZard collaboration, which obtained a pointing resolution (in one dimension) of approximately $1^{o}/\sqrt{E}$, where E is given in GeV[10]. A silicon tungsten calorimeter with $1/2\,X_0$ sampling and a spacing between layers of 24.5 mm was used. The strip pitch was 3.6 mm and 20 layers were used. A similar resolution of $2.2^{o}/\sqrt{E}$ (E in GeV) was obtained in reference [9] using a Si(Li)/Pb calorimeter. In this test the first two layers of the calorimeter (5 mm Pb, 2 mm Si(LI)) were separated from the rest of the calorimeter by 300 mm. The resolution of the OPAL calorimeter was 8 mrad at 45 GeV when information from all 19 layers ($22\,X_0$) was used. Using only the first 10 radiation lengths, an angular resolution of 12 mrad is obtained at 45 GeV or roughly $5^{o}/\sqrt{E}$. This is worse than the other two calorimeters as the OPAL calorimeter was designed to be as compact as possible with a 6.5 mm spacing between layers.

The main limitation of the calorimeters discussed above is their large channel count and their limited response to gamma rays in the MeV region. One solution to this problem is to instrument the front of the calorimeter with a homogeneous calorimeter backed up by a sampling calorimeter. From figure 4, it is clear the sampling calorimeter could start sampling the shower at approximately $4\,X_0$ without too much loss of information. A minimal sampling calorimeter might consist of two to four layers of x-y strips interleaved with absorber.

Since the pointing resolution of the detector will scale linearly with spacing between samplings, the layers of the calorimeter should not be placed too near to each other. On the other hand, increasing the spacing too far would change the aspect ratio of the detector and reduce its useful solid angle. It seems reasonable to assume that a resolution of order $1.5\ mm/\sqrt{E}$ (E in GeV) could be obtained in each plane. If the planes were placed 50 mm apart from each other, this would lead to an angular resolution (1-dimension) of roughly $2^{o}/\sqrt{E}$ (E in GeV) . Energy resolution consideration may require that the shower be sampled longitudinally somewhat more finely. Adding additional layers of Si strips would improve the pattern recognition capabilities of the device as well as its pointing resolution, but might increase the channel count and power consumption too much to be practical.

Alternatively, the "absorber" could be an active material such as CsI(Tl) and also be readout.

3 Survey of Available Technologies

In this section we discuss several possible calorimeters that might be used as a part of a next generation gamma ray telescope. We begin with a discussion of the forward part of the calorimeter and continue with a discussion of possible technologies for a sampling calorimetry. At the end of this section the properties of the different calorimeters are compared.

In all of the options considered the first few radiation lengths of the calorimeter consist of CsI(Tl). CsI(Tl) offers good resolution at low energies and can be conveniently read out using photo diodes. The properties of CsI(Tl) calorimeters have been extensively studied in high energy physics experiments [12]. Using photo diodes, readout noise levels corresponding to 0.5 MeV have been achieved in large systems. At high energies, the resolution of the calorimeter will be dominated by leakage, as shown in section 2.1. At low energies sampling fluctuations will dominate. Ideally, the relative energy in resolution would scale as $1/\sqrt{E}$. In practice the energy resolution degrades more quickly at low energy. Nevertheless the 3.8% resolution at 100 MeV reported in reference [12] would be more than adequate for a gamma ray telescope. Readout via photo diode requires no high voltage photo tubes and is well suited to operation in space.

The choice of technologies in a sampling portion of the calorimeter is less clear. In the following, three technologies are discussed: (a) segmented CsI(Tl), (b) silicon with lead or tungsten absorber, and (c) scintillator with lead or tungsten absorber.

3.1 Segmented CsI(Tl) Calorimetry

Longitudinal segmentation of CsI(Tl) crystals offers the possibility of good pointing resolution without significantly harming the energy resolution. The main difficulty with this approach is that very fine segmentation (1 cm x 1 cm) will be needed at the front of the calorimeter to obtain good position resolution. Since it seems

impractical to segment the CsI this finely, it will probably be necessary to sample the shower with either Si or scintillating fibers after the first few radiation lengths. The issues associated with Si or scintillating fibers are discussed in more detail below. The main difficulty in a segmented CsI(Tl) calorimeter is the support structure necessary to hold the CsI(Tl) at launch time.

One innovative method suggested at the workshop [13] was the use of CsI(Tl) crystals in a transverse orientation (see figure 5). This has the potential to reduce the channel count; however, in the context of the GLAST instrument with modular towers, it would be difficult to arrange readout without significant dead space between the towers.

3.2 Silicon Sampling Calorimetry

The use of Si as sampling medium has several advantages. It is radiation resistant and it requires only modest voltages (~100 V) for operation. Furthermore, there is extensive experience with the use of silicon calorimeters in high energy and astrophysics experiments [9][10][11]. The primary disadvantage of Si as sampling medium is the power consumption of the electronics needed for low noise readout [14]. The equivalent noise charge of a Si detector is dominated by leakage current and the noise introduced by the first transistor in the amplification stage. Assuming that a FET is used, the equivalent noise charge is approximately

$$\mathrm{ENC}^2 = \frac{C_t^2}{\tau}\left[\frac{8\,kT}{3\,g_m}\right] + 2eI_L\,\tau$$

where C_t is the capacitance of the detector and input transistor, g_m the transconductance of the FET, I_L the leakage current, and τ the shaping time of the amplifier. For a given transconductance the first term of the equivalent noise depends on the capacitance of the Si strips. For wide strips this is dominated by the detector capacitance $C_d = \varepsilon A/d$ where A is the total area of the strip and d is the thickness of the Si wafers. To reduce the capacitance, the Si detectors should be made as thick as possible; however, this may have an adverse effect on the cost of the detectors. For strips 180 mm by 2 mm and 0.5 mm thick, the strip capacitance would be 70 pF.

0Using Si detectors segmented into strips 180 mm by 2 mm with the currently available Viking preamp (power consumption 1.2 mW/channel), an equivalent noise of approximately 1000 electrons could be achieved compared to the expected signal of 30,000 electrons from a minimum ionizing particle. With the 1.5 μs shaping time of the Viking chip and the expected leakage current of approximately 10 nA/cm^2, the contribution to noise from the leakage current is approximately 430 electrons. Combining this with the noise from the first stage of the preamplifier, the total noise would be 1100 electrons. If sensitivity to minimizing particles is not needed in the sampling calorimeter, the strip area could be increased considerably, although increasing the strip width might degrade the position resolution.

The power consumption of the amplifier is driven by g_m. For values of current density that are not too small, the transconductance of FET scales as

$$g_m \propto \sqrt{\frac{w}{L} I_d}$$

where w is the FET channel width, L the FET channel length and I_d is the drain current. The minimum value of L is set by the feature size of the CMOS process used. w can be as long as 5 mm. For smaller values of the current density, g_m scales linearly with I_d . In the Viking chip [15] the current in the first stage of the preamplifier has been reduced to only 200 μA and w/L = 4300 μm/ 0.8 μm. It might appear advantageous to segment the Si more finely and decrease I_d, in order to reduce the total power consumption while keeping the signal to noise constant. However, the power consumption of the subsequent stages of the amplifier would then dominate the power consumption of each channel and the power would begin to scale with the number of channels in the detector.

3.3 Pb-Scintillator Calorimetry

The use of scintillating fibers together with Pb absorbers has been widely studied as a possible calorimeter for High Energy Physics experiments [16] and has been proposed for use in astrophysics experiments [8]. The use of transverse scintillating fibers and a lead absorber offers the same fine granularity (~1 mm) that can be obtained with Si readout and it allows one to escape the scaling between area and power that is present in the case of Si calorimetry. The fibers used in the

scintillating calorimeter can be made several meters long before their performance begins to degrade. In the context of the GLAST experiment, this would require that the sampling calorimeter be assembled as a single piece rather than in modules with fibers running the entire length or width of the calorimeter.

The optimum technology for reading out the scintillating fibers is not clear. In high energy physics experiments phototubes have generally been used. Other possibilities include avalanche photodiodes and CCDs. The later possibility is especially attractive as CCD readout provides the possibility of both readout and storage. The CCD could either be operated in a mode where the readout is triggered by the CsI(Tl) calorimeter, or CCD samples could be constantly shifted to the output and digitized. The later mode of operation is dead time free; however, the information from all events that occur during one readout time are overlaid. Commonly used pixel size in CCDs are small, typically 25 μm x 25 μm, but these pixels can be binned at the readout stage to obtain effective pixel sizes on the order of the size of a fiber.

For a telescope such as GLAST with an area of 1.8 m x 1.8 m, each layer of the calorimeter would have 1800 scintillating fibers. Assuming CCDs with an area of 15 mm x 15 mm, each CCD could service 180 fibers. Thus only 10 CCDs per layer would be needed. In principle a calorimeter with 4 x-layers and 4 y-layers could then be read out with only 80 channels. Assuming 10 μs/channel are needed to read out each binned pixel, the total time needed to read the CCDs is 1.8 ms. If the sampling calorimeter is used only for GeV showers, this latency is probably acceptable. If faster readout is needed, several amplifiers could be used with each CCD. Another potential difficulty with the CCDs is that the best noise performance is obtained at low temperatures (-50° C). Operation at 0° C is possible, but might compromise performance at low energies.

3.4 Comparison of Calorimetry Options

Four different possible calorimeters for a gamma ray telescope have been considered. To compare the calorimeters with each other, we assume an instrument similar to the proposed GLAST. The calorimeter is taken to be a 10 x 10 array of modules 18 cm x 18 cm and to be 10 X_0 deep.

The four options considered are as follows. Option I consists of a homogeneous CsI(Tl) calorimeter without longitudinal segmentation. The CsI(Tl) crystals are taken to be 2 cm x 2 cm. Option II has the CsI(Tl) divided longitudinally into lengths corresponding to 4 X_0, 3 X_0 and 3 X_0. A double layer of Si strips (for x and y readout) with 2 mm pitch has been added after the first CsI(Tl) crystal. The Si strips are 18 cm long. This is more economical then trying to subdivide the first layer of CsI(Tl) crystals further. Option III consists of 4 X_0 CsI(Tl) followed by 3 layers of Silicon strips (2 mm pitch) interleaved with 2 layers of tungsten absorber (3 X_0).

Table 1 shows the three options along with the channel count for each detector. Table 2 gives rough estimates of power, mass, materials cost, and construction complexity. To calculate power 1-10 mW per channel has been assumed. Currently available electronics such as the AMPLEX [17] chip are available at the upper end range. The VIKING chip offers better signal noise and uses less power, but its dynamic range may be too restricted for use in the calorimeter. Custom electronics with an emphasis on low power could certainly reduce the power consumption below that of the AMPLEX chip. Additional power will be needed for digitization of the calorimeter signals and in some scenarios this could dominate the power consumption. Estimates of materials cost include only the cost of expensive materials such as CsI and Si. For Si a cost of \$15/cm^2 is assumed and for CsI(Tl) a cost of \$25/cm^3 is assumed.

Table 1

		Segmentation		Channels	
Option	**Material**	**Longitudinal**	**Transverse**	**Module**	**Total**
I	CsI(Tl)	$10\ X_0$	2 x 2 cm	81	
					8,100
II	CsI(Tl)	$4\ X_0\ 3\ X_0\ 3\ X_0$	2 x 2 cm	243	
	Si		2mm strips	180	
					42,300
III	CsI(Tl)	$4\ X_0$	2 x 2 cm	81	
	Si-W	$2 \times 3\ X_0$	2mm strips	540	
					62,100
IV	CsI(Tl)	$4\ X_0$	2 x 2 cm	81	
	Pb/Scint	$6 \times 1X_0$	1 mm x 1.8 m	-	
					22,500

Table 1: Comparison of segmentation and channel counts in the four options considered.

Table II:

Option	I	II	III	IV
Material	CsI(Tl)	CsI(Tl)-Si	CsI(Tl)-Si-W	CsI(Tl)-Pb/Scint
Power [W]	8 - 80	42 - 420	62 - 620	23-225
Materials Cost [M$]				
Cs	2.2	2.2	0.9	0.9
Si/Scint		1.0	3.0	>0.1
Weight [kg]	2,800	2,800	2,400	2,300

Table 2: Estimates of power consumption, materials cost and weight for the four calorimetry options.

The main difficulty with most of the solutions involving longitudinal segmentation (II-IV) is the power cost associated with the large number of channels. The total power consumption available for an entire satellite will be approximately 1000 W and most of this will be needed by the tracker portion of the telescope. At 10 mW/Channel a Si-W sampling calorimeter is probably not viable.

In the tables, we have assumed that miniature phototubes or APDs are used to read out the scintillating fiber. If CCD readout is possible, the number of channels could be dramatically reduced. However, this readout scheme has several disadvantages. The major disadvantage is the long time needed to read out the CCD. In addition, the use of the CCD as an analog memory introduces a single point of failure to each group of 180 channels, which is avoided if the channels are read out individually. Finally, in contrast to Si strips or phototube readout of scintillating fibers, use of CCD readout of scintillating fibers is still not a proven technology.

A minor advantage of a sampling calorimeter over CsI is that it can be made of high Z materials, which provide a greater number of radiation lengths for a given mass.

A more significant advantage of sampling calorimetry over a homogeneous calorimeter is that the material used in the absorber of the calorimeter could be incorporated into the space craft structure. Although all the options employ CsI(Tl) crystals, the $4\,X_0$ crystals will be considerably easier to support than $10\,X_0$ crystals. The tungsten absorber would have the best structural properties. By incorporating the tungsten into the support structure of the satellite, it might be possible to increase the portion of the satellite mass that is used as an absorber. Even if Pb is used as an absorber, it could be combined with an Al or Fe support without degrading the energy and pointing response of the calorimeter.

4 Conclusion

Several possible calorimeters for use in a gamma ray telescope have been examined. A calorimeter with a $4\,X_0$ of CsI(Tl) followed by a sampling calorimeter with either tungsten or lead absorber maintains good energy resolution at low energies, but at the same time provides enough longitudinal segmentation to allow the direction of the 70% of incident gamma rays that do not convert in the tracker/converter portion of the telescope to be reconstructed. The main difficulty with a longitudinally segmented calorimeter is the high channel count. In the case of a Pb/scintillator (or W/scintillator) calorimeter it may be possible to greatly reduce this channel count by using CCD readout of scintillating fibers. The use of sampling calorimeter offers the possibility of using the absorber as part of the structure of the satellite, which may simplify the construction of the satellite and may allow a more massive calorimeter than otherwise possible.

Acknowledgments

I wish to thank the organizers for their successful efforts to provide an interesting and useful workshop. In addition, I wish to thank Mark Oreglia for kindly providing the results of his simulations of gamma rays in CsI(Tl) of the GRSST instrument.

References

1. G. Kanbach, these proceedings.

2. J. W. Cronin, K. G. Gibbs, and T. C. Weekes, "The Search for Discrete Astrophysical Sources of Gamma Radiation," Ann. Rev. Nucl. Part. Sci. 43(1993)883.

3. R. Nolan, et al., "Observation of the Crab Pulsar and Nebula by the EGRET Telescope on the Compton Gamma-Ray Observatory," ApJ, 409 (1993) 697.

4. M. Punch, et al., "Detection of TeV Photons from the Active Galaxy Markarian 421," Nature, 358 (1992) 477.

5. F. W. Stecker, O. C. De Jasger, and M. H. Salamon, 'TeV Gamma Rays from 3C 279; A Possible Probe of Origin and Intergalatic Infrared Radiation Fields." ApJ 390 (1992) L49.

6. B. Michel, "Electron-Pion Separation in a Highly Segmented Electromagnetic Calorimeter Designed for the ALEPH Experiment," Nucl. Instr. and Meth. A263 (1988) 43.

7. B. Atwood, these proceedings.

8. "GILDA: a Gamma-ray Imaging Large Detector for Astronomy," to be submitted to Nucl. Inst. Meth.

9. T. Kashiwagi, et al., "A 5in Si(Li)/Pb Sampling Calorimeter Telescope for Observation of Cosmic Gamma Rays in the GeV Region," Nucl. Instrum. and Meth. A290 (1990) 579.

10. M. Bocciolini, et al., "A Silicon Imaging Calorimeter Prototype for Antimatter Search in Space: Experimental Results," Nucl Instr. and Meth. A333 (1993) 560.

11. D. Strom, "Measurement of Ratio of the Invisible to Leptonic Widths of the Z^0 using the OPAL detector at LEP," Proceeding of DPF94, Albuquerque, New Mexico, August 2-6, 1994.

12. See, for example, Y. Kubota et al., "The CLEO II detector," Nucl. Instr. and Meth. A320 (1990) 66.

13. P. Persson, "Electromagnetic Calorimetry for Cosmic Gamma Rays," diploma Thesis, KTH, Stockholm, 1994.

14. R. Johnson, these proceedings.

15. P. Aspell, et al., "CMOS Low Noise Monolithic Frontends for Si Strip Detector Readout," Nucl. Instr. and Meth. A315 (1992) 425.

16. R. Wigmans, "Recent Results from the Spaghetti Calorimeter Project," Nucl. Instr. and Meth. A315(1992) 299.

17. E. Beuville et al., AMPLEX, "A Low Noise, Low Poser Analog CMOS Signal Processor for Multi-Element Particle Detectors," Nucl. Inst. and Meth. A288 (1990) 157.

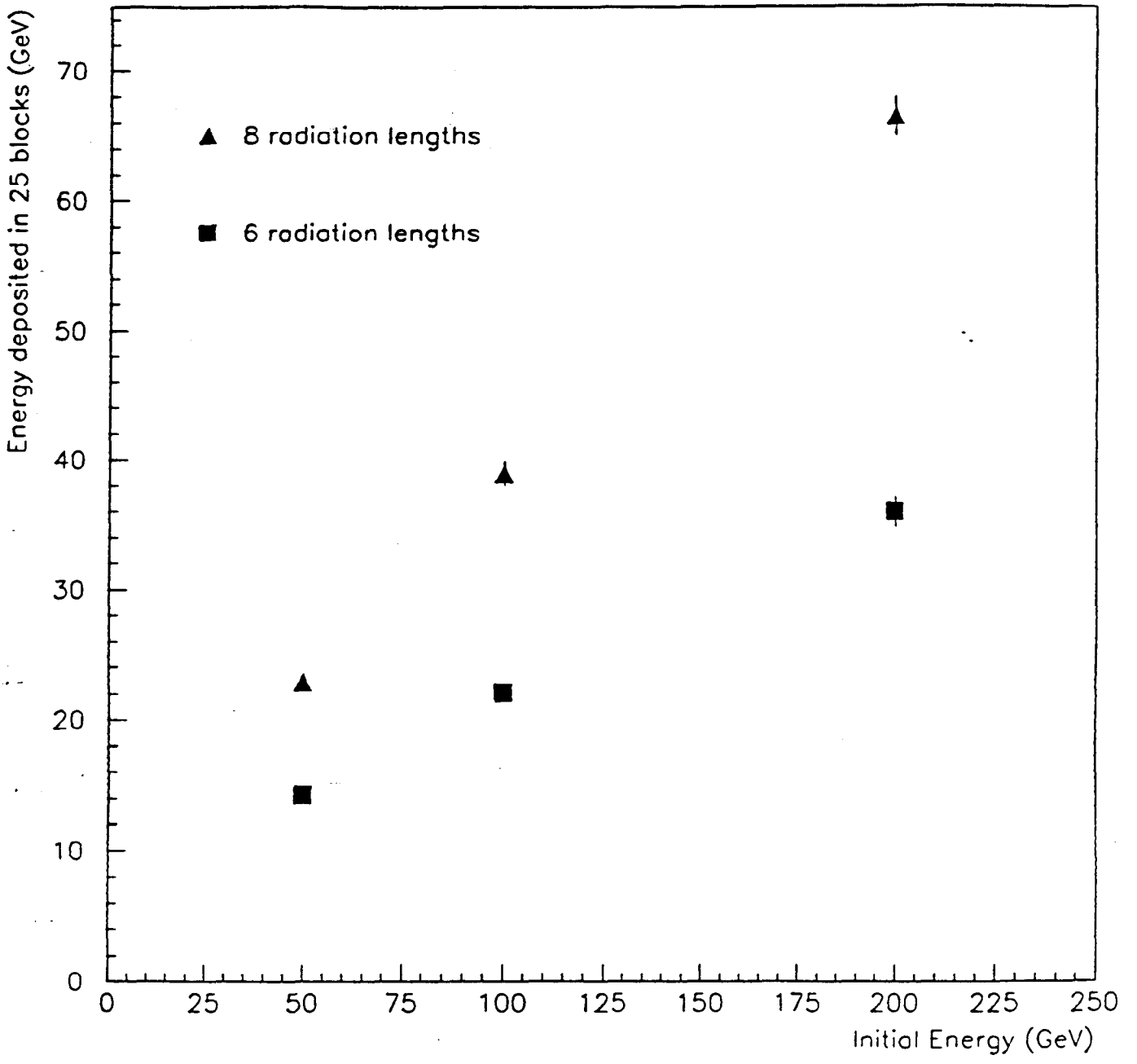


Figure 1. Simulation of the energy deposited in the 25 CsI(Tl) (2 cm x 2 cm) blocks centered on the shower as a function of incident energy. Only gamma rays that convert in the tracker/converter are plotted. Gamma rays whose showers develop near the transverse edges of the calorimeter have also been eliminated.

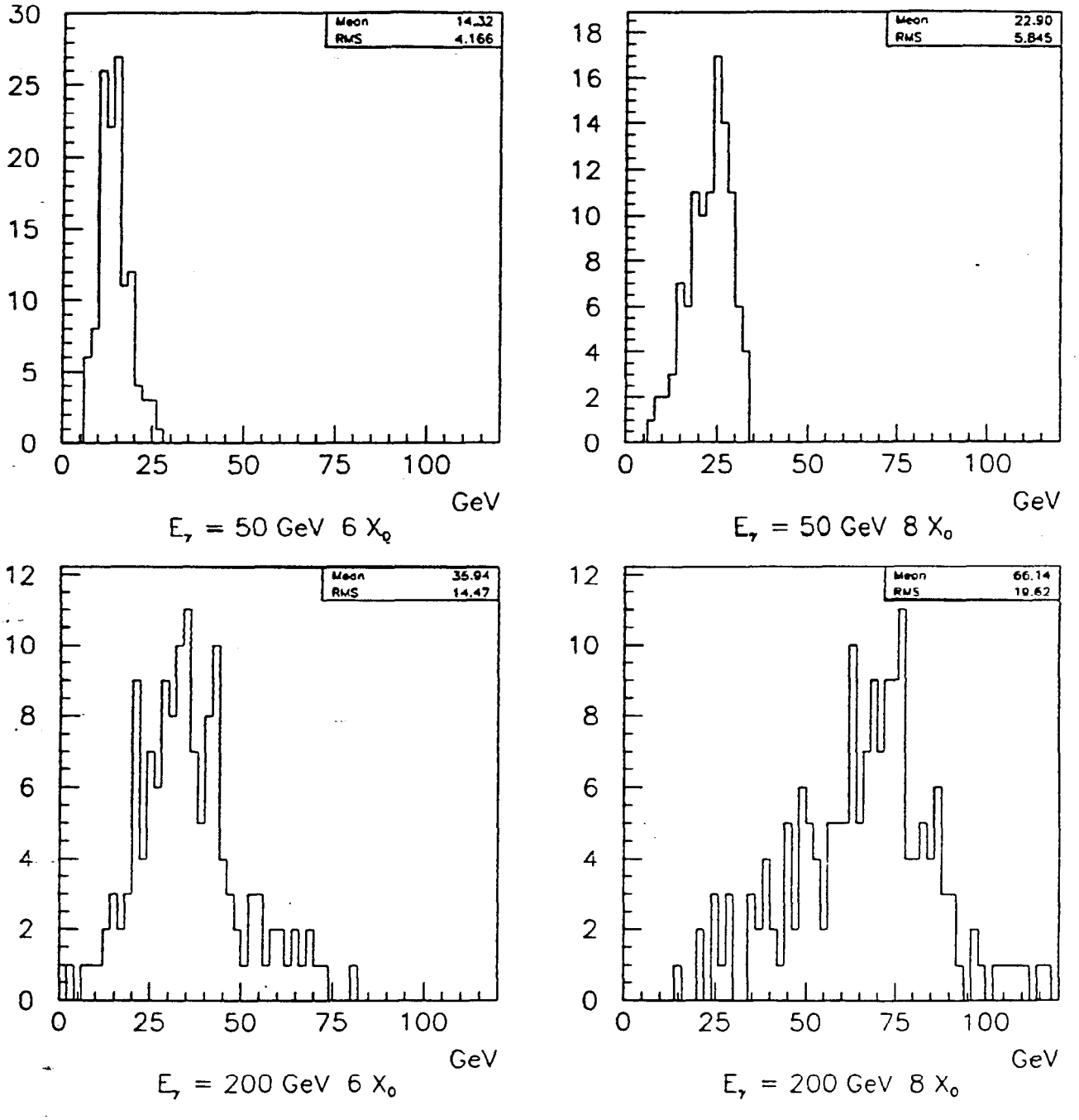
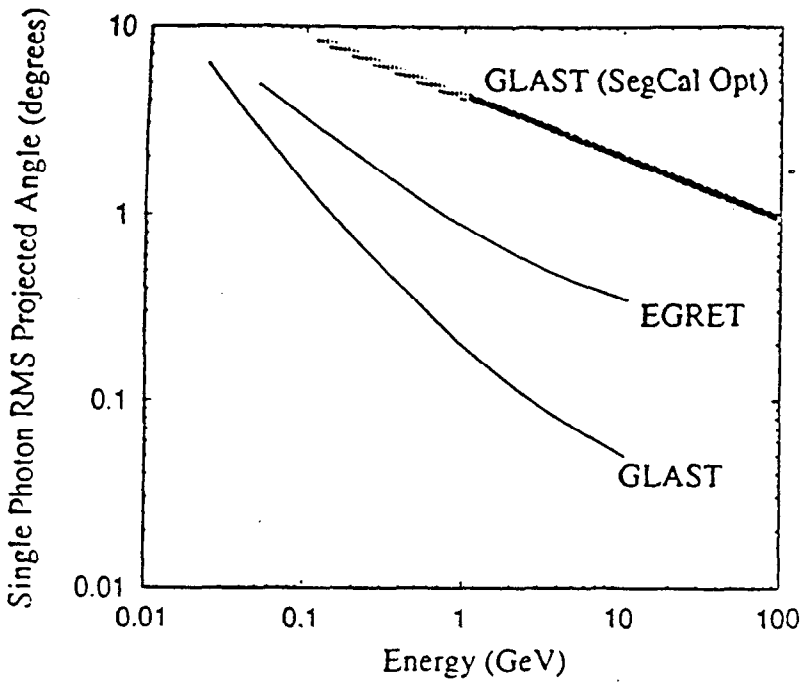


Figure 2. Comparison of 50 GeV and 200 GeV incident gamma rays in CsI calorimeters with 6 X_0 and 8 X_0 longitudinal sampling.



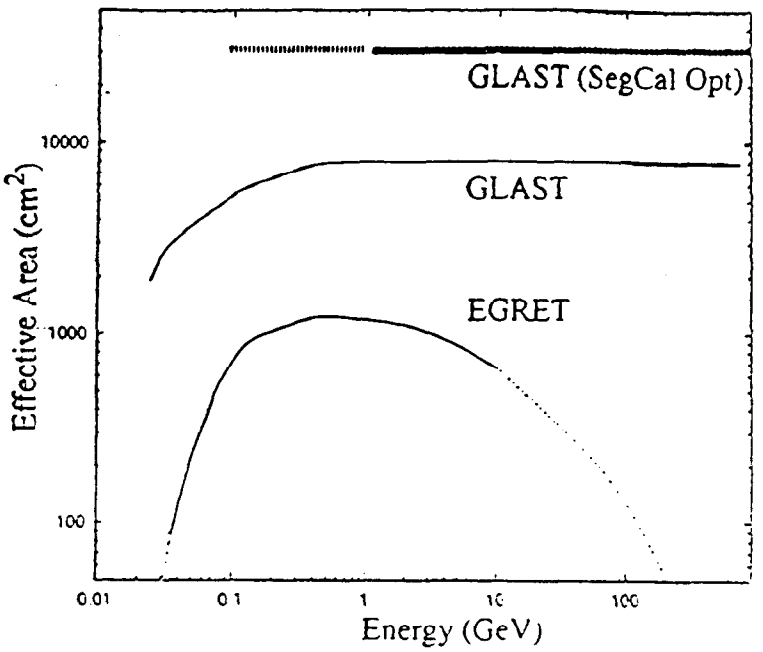


Figure 3. Angular resolution and effective area of the proposed GLAST instrument compared with EGRET. The upper curves show the gain in effective area of the satellite when a longitudinally segmented calorimeter is incorporated into GLAST.

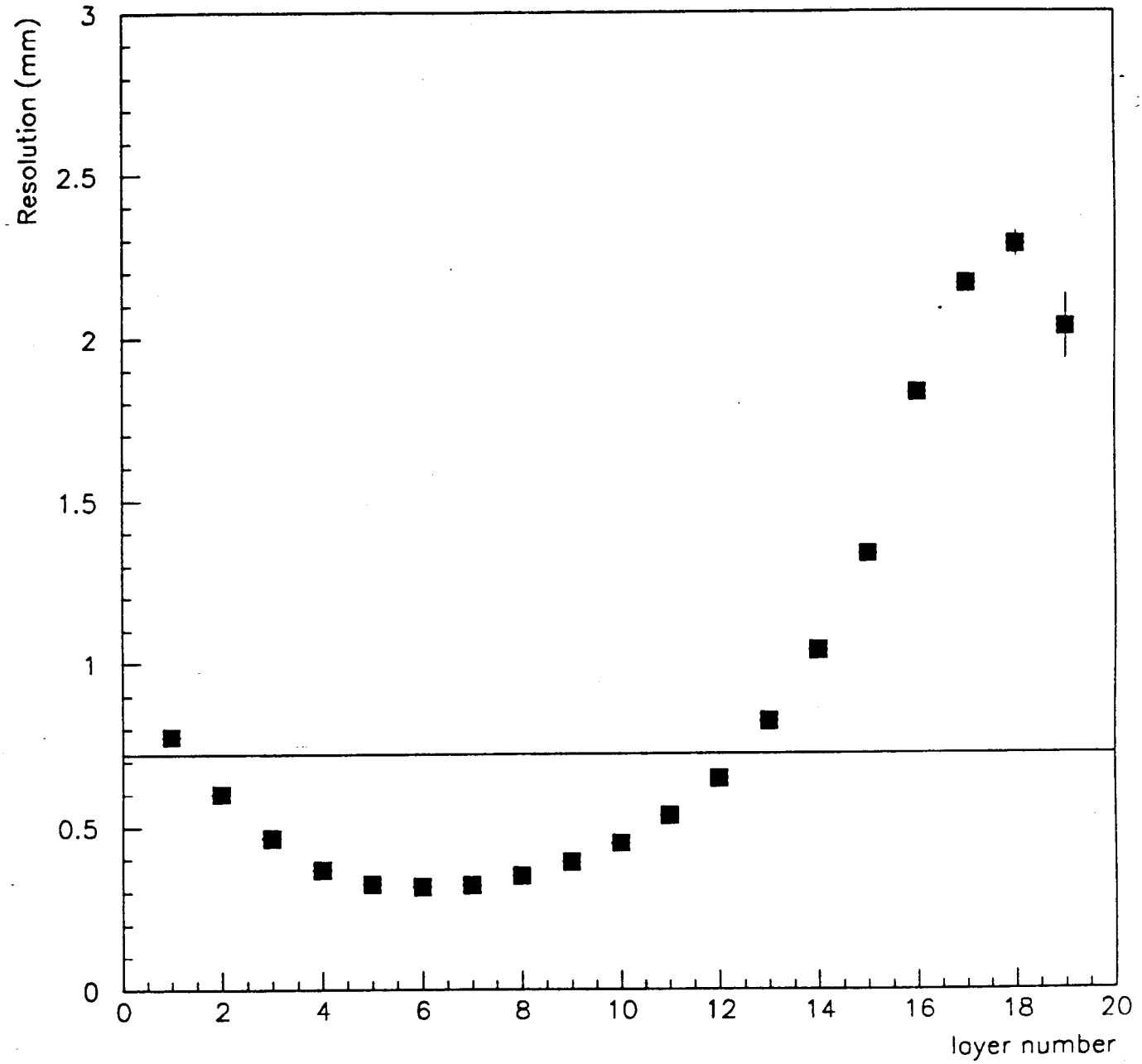


Figure 4. Position resolution of the OPAL Silicon-Tungsten luminosity monitor as a function of layer number. The first 14 layers have 1 X_0 spacing and the last 4 layers have 2 X_0 spacing. The solid line corresponds to 2.5 mm/$\sqrt{12}$.

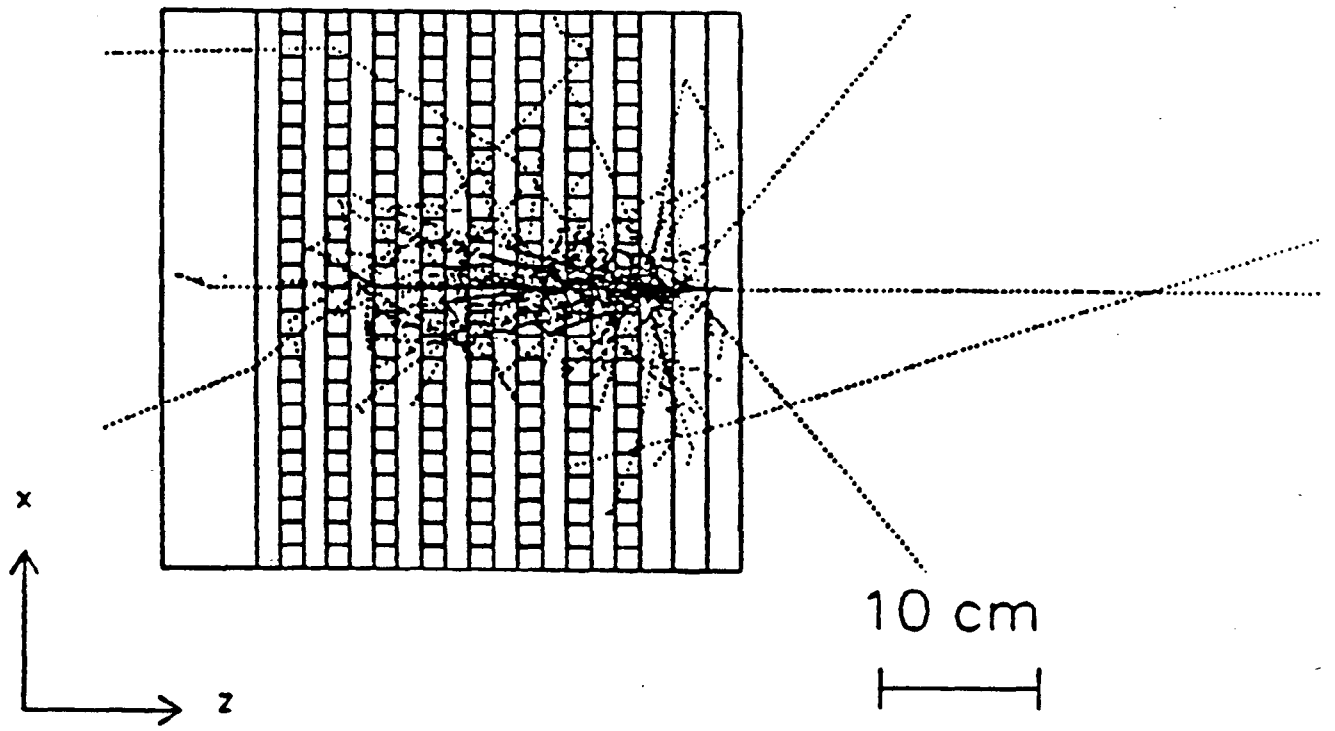


Figure 5. A simulated electromagnetic shower of a 1.0 GeV gamma ray in a calorimeter with transverse CsI(Tl) crystals [13]. Alternate layers of the calorimeter are oriented at right angles to each other.

Appendix: Physics with GLAST

High-energy gamma rays probe the most energetic phenomena in nature, typically involving nonthermal processes. High-energy gamma rays are emitted over a wide range of angular scales from a diverse population of astrophysical sources—stellar mass objects; the nuclei of active galaxies; interstellar gas in the galaxy interacting with high-energy cosmic rays; the diffuse extragalactic background; supernovae that may be sites of cosmic-ray acceleration; gamma-ray bursts; and the Sun, which during active periods can produce high-energy gamma rays. Many of the sources exhibit transient phenomena, ranging from the few second timescale of gamma-ray bursts to AGN flares lasting days or more, and often radiate the bulk of their power at gamma-ray energies.

Observations of high-energy gamma rays provide unique astrophysical information. For example:

- They identify sites where extreme particle acceleration processes occur and provide probes of the physical processes operating at those sites. The gamma-ray outburst of 3C 279 (Hartman et al., 1992; Kniffen et al., 1993) observed by EGRET was unanticipated, showing that remarkably high luminosities can be attained along with rapid turn-on and turn-off of the source. The emission processes likely involve shock acceleration and particle cascades that produce the observed gamma-ray spectrum. In some systems, the acceleration may be electromagnetic; two examples are the Blandford-Znajek mechanism for producing a large voltage drop near a rotating black hole and the acceleration of charged particles in the magnetosphere of a rotating neutron star. In addition to the point sources, high-energy observations at high galactic latitudes present a diffuse background puzzle quite distinct from the one that has been studied since the 1960s in X-rays and low-energy gamma rays.
- High-energy gamma-ray observations have identified what might be called 'GeV source classes'. Geminga is an example of a pulsar that is far more conspicuous in gamma rays than at lower energies and there should be many others like it. All of the 'surprising' EGRET detections — blazars, gamma-ray bursts, solar flares, and Geminga pulsars — amount to discoveries of GeV source classes, in the sense that the energetics of the GeV contribution to the total spectrum was unexpectedly high. Detailed characterization of these classes is only beginning and eventual identification of the unidentified EGRET sources may reveal distinctly new source classes.
- The study of the galactic component of the diffuse emission and of emission from molecular clouds and supernovae remnants are important uses of high-energy gamma-ray

observations that provide information about the distribution of cosmic rays and matter in the galaxy, which, in turn, have implications about the origin and propagation of cosmic rays.

- Space-based GeV observations combined with ground-based TeV observations sensitive above 50 to 100 GeV can provide measurements of absorption cutoffs in the spectra of extragalactic sources (e.g., Mkn 421) which can be used as indicators of the intergalactic IR radiation field (high-energy photons incident on IR photons are attenuated through pair production (Stecker, DeJager, and Salamon, 1992).

Because of its larger area, wider field of view, better angular resolution, and broader energy range compared to EGRET, the GLAST instrument will be well suited to the study of high-energy astrophysical phenomena that occur over a broad range of angular and temporal scales. GLAST is essentially a large, high-energy particle detector with a flexible trigger; its operating modes can be reconfigured and optimized for different scientific objectives. This capability is unprecedented for an orbiting high-energy telescope and presents the possibility of fundamental new discoveries in high-energy astrophysics. In the following subsections, the instrument performance parameters and the observational breakthroughs likely to be obtained with GLAST in several of the areas mentioned above are discussed.

2.1 GLAST Observational Performance Parameters

Figure 1 summarizes the instrument's effective area, angular resolution, and energy resolution versus energy for an on-axis source. These parameters were determined with Monte Carlo simulations. Inefficiencies due to background rejection analysis cuts have been included. The off-axis effective area of GLAST is essentially given by the geometric projection of the effective area at normal incidence. The energy resolution shown is the equivalent Gaussian σ determined from Monte Carlo distributions of the energy deposited in the calorimeter, corrected for energy loss in the tracker. Tracker corrections become increasingly important below 100 MeV. Fluctuations of the shower leakage out the back of the calorimeter begin to become important above 10 GeV. The energy measurement reach of GLAST would extend into the multi-TeV range with a thicker calorimeter, but the number of photons expected does not justify the extra weight. With a 10 r.l. thick calorimeter, simulations show that an energy measurement is still possible (σ ~33%) at 800 GeV.

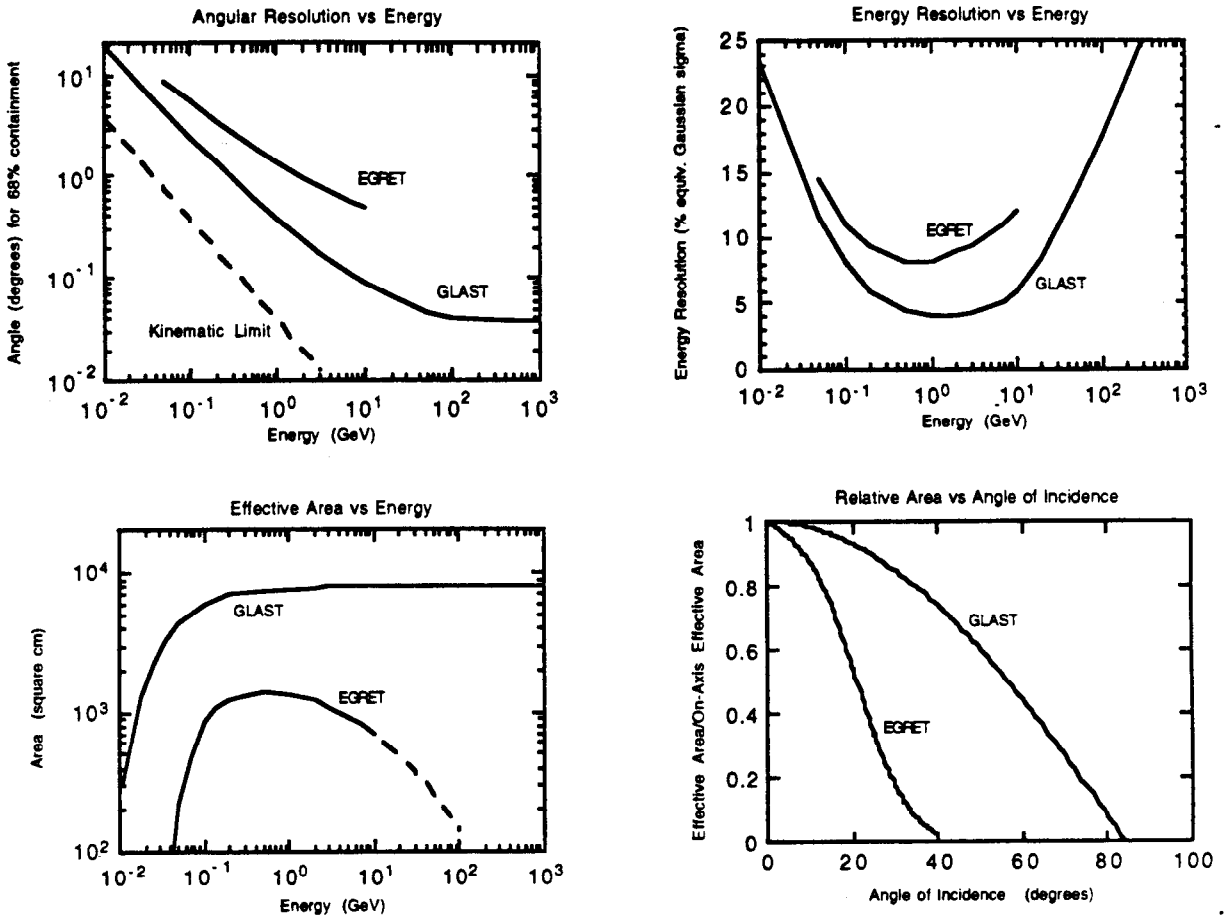


Figure 1. GLAST and EGRET angular resolution, effective area, and energy resolution versus energy for photons incident normal to the detectors. Also shown are the relative effective areas (@ 1 GeV) versus the angle of incidence. Inefficiencies due to analysis cuts for background rejection have been included.

Figure 2(a) shows the source sensitivity versus threshold energy E obtained from a 1-year all-sky survey; GLAST will be 50 to 100 times as sensitive as EGRET. The diffuse background assumed in calculating the sensitivity is 2 x 10^{-5} photons $cm^{-2}\ s^{-1}\ sr^{-1}$ $(100\ MeV/E)^{1.1}$, typical of the background seen by EGRET at high galactic latitudes. The source differential photon number spectrum is assumed to have a power law index of -2, typical of many of the sources observed by EGRET. Above 1 GeV, the sensitivity shown is not limited by background, but rather by the conservative, but somewhat arbitrary, requirement that the number of source photons detected be at least five.

The limiting continuum source flux spectral sensitivity ϕ in units of (photons $cm^{-2}\ s^{-1}\ MeV^{-1}$) is calculated from $\phi = \sigma/(2Af\Delta E)\ [\ \sigma/T + (\sigma^2/T^2 + 4\phi_b \Delta E A\Omega/T\ )^{1/2}\ ]$,

where σ is the statistical significance of the flux measurement, A is the effective area in cm^2, ΔE is the energy interval over which the flux measurement is made in MeV, T is the

observation time in seconds, ϕ_b is the spectrum of the gamma-ray diffuse background in (photons $cm^{-2}\ s^{-1}\ sr^{-1}\ MeV^{-1}$), and $\Omega(f)$ is the solid angle within which a fraction f of the reconstructed directions of photons from the point source lie. Fig. 2(b) shows S, assuming $\Delta E = E$.

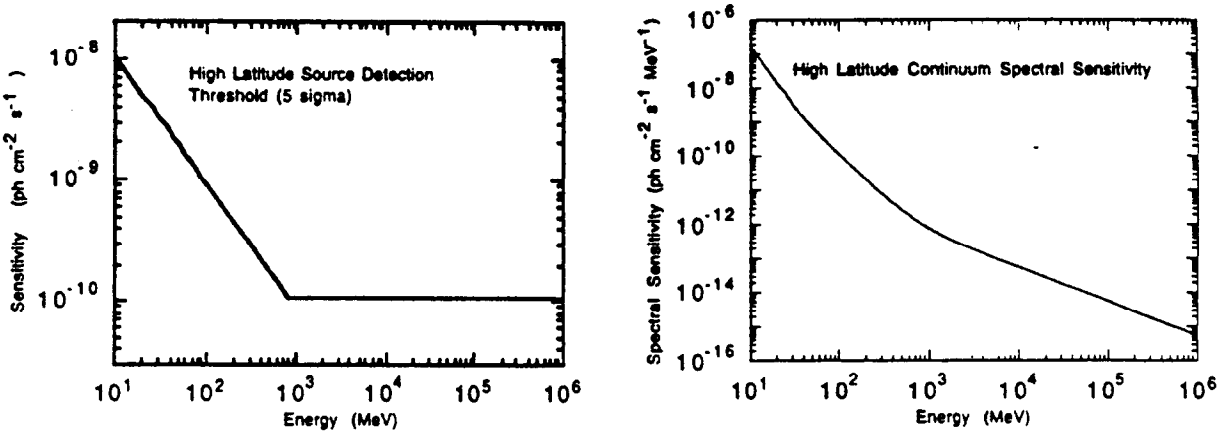


Figure 2. (a) Source detection threshold for GLAST 1-year all-sky survey. (b) Continuum source flux spectral sensitivity for GLAST 1-year all-sky survey.

2.2 Extragalactic Sources: Active Galactic Nuclei

With detections of more than 50 active galaxies, EGRET has opened gamma-ray astronomy to the study of extragalactic sources (see von Montigny et al., 1995, for a summary). Most of these sources, distributed over a wide range of redshifts (0.03 to 2.28), have characteristics of the blazar class of AGN (BL Lacs, OVV and HPQ quasars) and many of them are associated with superluminal radio sources. In many of the sources the apparent gamma-ray luminosity is dominant over the flux in lower-energy bands and exhibits extreme variations in intensity on timescales from days to months. These time variations imply a source significantly smaller than the smallest resolved radio feature.

Ground-based TeV observations of several of the EGRET gamma-ray loud AGNs have also been made. Detection of Mkn 421 above 0.5 TeV has been reported (Punch et al., 1992). While 3C 279 is a relatively strong EGRET source by comparison, it was not detected in the TeV range, indicating a very strong spectral break above the EGRET energy range. This break has been interpreted as due to attenuation of high-energy photons by g-g pair production processes occurring in the intergalactic IR radiation field. Recently, the detection of Mkn 501 at TeV energies has been reported (Schubnell et al., January 1996, AAS Meeting, San Antonio).

AGNs are widely believed to be ultimately powered by accretion onto a massive black hole (for reviews, see Rees, 1984; Begelman, Blandford, and Rees, 1984). Most models for

gamma-ray emission involve beamed emission from a jet of highly relativistic particles that originate near the black hole central engine. While there appears to be little doubt that relativistic jets are involved, the mechanisms responsible for collimating and accelerating the jets are still speculative. Indeed, there is still major uncertainty about whether the jets are primarily made of an electron-positron plasma or an electron-proton plasma. Most models of the high-energy emission involve inverse Compton scattering of low-energy photons from a population of relativistic electrons in the jet. Some models involve a cascade that begins with photomeson production by very high-energy protons ($\geq$ 107 TeV) that are accelerated by diffusive shocks in the jet (see von Montigny et al., 1995, for a recent review of these models).

Despite the importance of the EGRET blazar discoveries, understanding of the gamma-ray emission mechanisms in these sources is limited by current instrumental capabilities. With GLAST:

- An all-sky survey should produce more than 1,000 AGN detections. This will allow an accurate determination of their logN-logS distribution which, in turn, will allow an accurate calculation of their contribution to the high-energy diffuse extragalactic radiation. GLAST's sensitivity also offers the possibility of detecting high-energy emission from less luminous sources such as Seyfert galaxies and radio galaxies.
- Measurement of features, such as high-energy cutoffs, in the spectra of the brighter sources can constrain acceleration and emission models. For example, jet models involving proton initiated cascades have the general property that power injected at energies well above a TeV is cascade-reprocessed to finally emerge in the entire energy band from keV to TeV. These models predict that in some sources the differential gamma-ray photon spectrum above a few GeV should significantly depart from a power-law and actually increase (Mannheim and Biermann 1989, 1992; Mannheim 1993). Extending the measured spectra of AGNs significantly beyond 10 GeV may also provide a probe of the intergalactic IR radiation field, because of attenuation of high-energy photons due to pair production processes (Stecker, DeJager and Salamon 1992).
- The large field-of-view and effective area of GLAST will allow the monitoring of sources for variability over long periods of time and will facilitate multi-wavelength observations with ground-based telescopes and other orbiting instruments. With GLAST, the onset of an event such as the June 1991 gamma-ray flare in 3C 279 can be detected in less than 90 minutes.

2.3 Compact Stellar Objects

With CGRO, the number of gamma-ray pulsars seen above 30 MeV has risen from two (Crab and Vela) to six. Five of these are relatively young, spin-powered radio pulsars, while the sixth is Geminga, the only known radio-quiet pulsar in this class (summary by Fierro 1995 and Thompson 1996). It is well established that pulsars are rotating neutron stars with strong magnetic and electric fields. Theoretical work on gamma-ray pulsars, stimulated by the EGRET results, has not conclusively resolved the question of whether the gamma-ray emission region is close to the magnetic polar cap of the neutron star (e.g., Daugherty and Harding, 1994; Sturner and Dermer, 1994) or is in the outer magnetosphere (e.g., Romani and Yadigaroglu, 1994).

- With larger area and extended energy response, GLAST can explore the phase-resolved energy spectra of pulsars in far greater detail, well into the region where the models predict "cutoffs" in the spectrum. GLAST can also study in detail the unpulsed emission from the Crab in the transition region between the possibly time-variable synchrotron emission and the Compton component that extends to TeV energies (DeJager et al., 1996).
- The high sensitivity of GLAST will also result in a much larger sample of pulsars, both radio pulsars (model estimates indicate that up to 40 of these should be visible, including PSR 0540-69 in the LMC if its spectrum is Crab-like) and radio-quiet pulsars like Geminga. The gamma-ray rate in GLAST will be high enough to search for pulsations in many of the brightest unidentified sources. This larger sample is needed to test pulsar models and to address questions about the rate and distribution of supernovas in the Galaxy.
- GLAST will also facilitate searches for gamma-ray emission from the older, millisecond pulsars not seen by EGRET (Fierro et al., 1995; Michelson et al., 1994) but predicted by some models to be gamma-ray sources (Chiang and Romani, 1992; Chen and Ruderman, 1993). Gamma-ray emission from the interaction of the putative relativistic (TeV) e± pulsar wind with the surrounding medium or with the mass outflow from the companion star in the case of a binary (Phinney et al., 1988; Tavani, 1991; Arons and Tavani, 1993) may be detectable.

2.4 Unidentified EGRET Sources

At high latitudes, EGRET has detected more than 50 sources for which no strong candidate identifications have been made. Many have spectral and temporal properties similar to the gamma-ray emitting blazars (Hartman et al., 1993). This suggests that either they are also blazars, in which case either radio brightness need not accompany gamma-ray brightness, or the objects brighten in radio by as much as an order of magnitude around the time of a gamma-ray flare, or these sources are members of a new source class.

There are also localized gamma-ray sources observed near the Galactic plane that have no obvious counterparts at other wavelengths. Within 10o of the plane, EGRET has observed more than 30 such sources above 5s. These sources appear to fall into two categories: those with relatively constant gamma-ray flux, many of which could be radio-quiet pulsars like Geminga (Romani and Yadigaroglu, 1994), and a substantial fraction that exhibit significant time variability. Of the latter source category, several may be AGNs, but the majority are probably galactic and could represent a new source class.

- GLAST's improved angular resolution, especially at high energies, will provide source error boxes between 10 and 100 times smaller than EGRET's. With its large field-of-view, GLAST can monitor unidentified sources for temporal variability over long periods of time. Both of these factors will aid in source identification and "handoff" to other energy bands.
- The expected high rate of GeV photons detected from more unidentified sources (comparable to Geminga in EGRET) will enable the detection of 'Gamma only' pulsars without prior knowledge of periodicity from other wavelength observations.

2.5 Diffuse Emission: Galactic, Extragalactic and Molecular Clouds

The galactic diffuse radiation is the most prominent feature in the gamma-ray sky. Much of it results from cosmic ray nuclei and electrons interacting with interstellar gas. EGRET observations of the Galactic plane have been compared with detailed models of the Galactic gamma-ray emission to indirectly deduce the Galactic cosmic ray distribution and spectrum. The angular resolution of EGRET limits these comparisons because it is relatively large compared to much of the structure predicted by the models. Among the outstanding issues to still be resolved are the relative contributions from various processes as a function of energy and location, the question of cosmic-ray/matter coupling versus cosmic-ray gradients, the physics of high-energy particles in clouds, emission from the Galactic halo, and the contribution from unresolved sources such as the galactic pulsar population (see, e.g., Strong, 1993).

With EGRET, many nearby molecular clouds, including those in Orion and Ophiuchus, have been detected and spatially resolved (Digel and Hunter, 1993). However, with EGRET's sensitivity, the gamma-ray spectrum can only be determined for the entire cloud. Similarly, the LMC has been detected (Sreekumar et al., 1993) but not mapped in detail in gamma rays.

At high galactic latitudes the observed "diffuse" radiation contains both a galactic component and a component that appears isotropic on large scales. Theories of the origin of this radiation generally fall into two categories: one class assumes the radiation was produced at an early epoch in intergalactic space and is truly isotropic and diffuse, while the other class of theories assumes that the radiation is really the integrated flux from a large number of unresolved discrete sources. GLAST will

- Map the galactic emission on a finer angular scale and to higher energies than was previously possible.
- Observe nearby dense molecular cloud complexes on smaller spatial scales to look for variations in the cosmic ray spectrum, possible exclusion of cosmic rays from the dense cores of clouds, and regions of cosmic ray acceleration, and should allow detection of nearby galaxies such as M31, the SMC, and detailed study of the LMC. These data will give information about the distribution of cosmic rays in these objects.
- Observe unidentified EGRET Galactic plane sources that statistically appear to be associated with supernovae remnants. GLAST's angular resolution and sensitivity is needed to confirm these identifications and will provide improved spectra. If supernovae are the injectors and accelerators of high-energy Galactic cosmic rays, then the gamma-ray spectrum of SNRs is expected to be harder than the Galactic diffuse emission.
- Map the high-latitude diffuse radiation on finer angular scales and determine the spectrum to higher energies. Combined with new information gained about active galaxies, this will place important constraints on models of extragalactic emission.

2.6 Gamma-Ray Bursts

The GRB detection rate for CGRO/BATSE is ~one per day. The GRB spatial distribution is isotropic, with no significant quadrupole or dipole moment found with respect to any coordinate system (Briggs et al., 1995), whereas the number intensity relation implies an

effective limit to the depth of the spatial distribution — BATSE "samples the edge" (Meegan et al., 1996). However, physical models which invoke the two remaining viable distance scales, galactic halo (~ few 100 kpc) and cosmological (~ few Gpc), encounter interesting difficulties from source-population (Hartmann, 1996) and spectral considerations (Baring, 1995). The latter arise especially from EGRET detections in several bursts of > GeV emission, which can persist up to tens of minutes longer than the BATSE event (Hurley et al., 1994). Straightforward calculations indicate that a large fraction of GRBs should exhibit this very high-energy emission. But, because EGRET has a high fraction of dead time during intense GRBs, the GeV : keV-MeV emission ratio as a function of time is highly uncertain.

- With its large acceptance/area factor and essentially 100% live time during GRB events, GLAST, in conjunction with a lower-energy burst monitor, will detect ~ 50 – 100 bursts per year and will be able to study the GeV : keV-MeV emission ratio as a function of time, thus providing much improved constraints on physical mechanisms. Stand-alone GLAST will image positions of DIM bursts to ~ few arc minutes, allowing deep "real time" observations at other wavelengths.

References

Arons, J., and Tavani, M. 1993, ApJ, 403, 24

Baring, M. G., 1995, in *Currents in High-Energy Astrophysics*, (eds. M. M. Shapiro, et al.), (Kluwer Academic: Amsterdam), 21.

Begelman, M., Blandford, R., and Rees, M.J. 1984, *Rev. Mod. Phys.*, 56, 255

Berridge, S. C., et al. 1990, *IEEE Trans. Nuc. Sci.*, 37, 1191

Briggs, M. M., et al., 1995, *ApJ* 459, in press

Chen, K., and Ruderman, M. 1993, *ApJ*, 402, 264

Chiang, J., and Romani, R. W. 1992, *ApJ*, 400, 629

Daugherty, J. K., and Harding, A. K. 1994, *ApJ*, 429, 325

DeJager, O.C., et al. 1996, *ApJ*, 457, 253

Digel, S. and Hunter, S., 1993, in Proc. of *The 2nd Compton Symp.*, AIP Conf. Proc. 304, 484 (AIP, New York)

Fierro, J., et al. 1995, *ApJ*, 447, 807

Hartmann, D. 1996, in *Proc. 3rd GRB Symposium*, (eds. C. Kouveliotou, M. S. Briggs and G. J. Fishman), AIP, in press.

Hartman, R.C., et. al. 1992, *ApJ*, 385, L1

Hartman, R.C., et al. 1993, in *Proc. of The 2nd Compton Symp.*, AIP Conf. Proc. 304, 563 (AIP, New York)

Hurley, K., et al. 1994, *Nature*, 372, 652

Kniffen, D.A., et al. 1993, *ApJ*, 411, 133

Mannheim, K. 1993, *Phys. Rev. D*, 48, 2408

Mannheim, K. and Biermann, P. 1989, *A and A*, 221, 211

Mannheim, K., and Biermann, P. 1992, I, 53, L21

Meegan, C.A., et al., 1996, in Proc. I., eds. C. Kouveliotou, M. S. Briggs and G. J. Fishman), *AIP*, in press.

Michelson, P.F., et al. 1994, *ApJ*, 435, 218

Phinney, E.S., et al. 1988, *Nature*, 333, 832

Punch, M., et al. 1992, *Nature*, 358, 477

Rees, M.J. 1984, *Ann. Rev. Astron. Astrophys.*, 22, 471

Romani, R., and Yadigaroglu, I. -A. 1994, *ApJ*, 438, 314

Schubnell et al., January 1996, AAS Meeting, San Antonio

Sreekumar, et al. 1993, *ApJ*, 400, L67

Stecker, F.W., DeJager, O.C. and Salamon, M.H. 1992, *ApJ*, 390, L49

Strong, A. 1993, in *Proc. of the 2nd Compton Symp.*, AIP Conf. Proc. 304, 463 (AIP, New York)

Sturner, S.J. and Dermer, C. 1994, *ApJ.*, 420, L79

Tavani, M. 1991, *ApJ*, 379, L69

Thompson, D.J., et al. 1996, in Proc. *IAU Colloq.* 160, to be published

von Montigny, C., et al. 1995, *ApJ*, 440, 525