# NASA

## Technical Memorandum 85029

# Spatial Fluctuations in the Diffuse Cosmic X-Ray Background

**Richard Arrick Shafer**

**MAY 1983**

National Aeronautics and
Space Administration

**Goddard Space Flight Center**
Greenbelt, Maryland 20771

SPATIAL FLUCTUATIONS IN THE
DIFFUSE COSMIC X-RAY BACKGROUND*

by
Richard Arrick Shafer

Dissertation submitted to the Faculty of the Graduate School
of the University of Maryland in partial fulfillment
of the requirements for the degree of
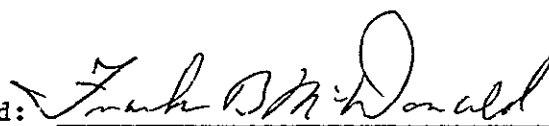Doctor of Philosophy
1983

APPROVAL SHEET


Title of Dissertation:  Spatial Fluctuations in the Diffuse Cosmic
X-ray Background


Name of Candidate:  Richard Arrick Shafer
Doctor of Philosophy, 1983


Dissertation and
      Abstract Approved:

Frank B. McDonald
Adjunct Professor
Department of Physics


Date Approved: _____ 8 April 1983

ABSTRACT

Title of Dissertation:   Spatial Fluctuations in the Diffuse Cosmic
X-Ray Background

Richard Arrick Shafer, Doctor of Philosophy, 1983

Dissertation directed by: Professor Frank B. McDonald
Adjunct Professor, Department of Physics

Studies of the bright, essentially isotropic, X-ray sky flux above
2 keV yield information on the universe at large distances. However, a
definitive understanding of the origin of the flux is lacking. Some
fraction of the total flux is contributed by active galactic nuclei and
clusters of galaxies, but less than one percent of the total is
contributed by the $\geq 3$ keV band resolved sources, which is the band where
the sky flux is directly observed. The extension of N(S) (number of
sources versus flux) to lower fluxes than can be studied from the
resolved sources—is possible by investigating the distribution of the
fluctuations in the sky intensity. Data from the A-2 experiment on HEAO
1 present the best opportunity to date to study anisotropies in the
2.5-13 keV band. The quality of the data encourages a critical
assessment of the fluctuations analysis technique. An excellent fit to
the data is given by model fluctuations distributions (P(D) curves) for
a Euclidean N(S) with an additional, pure Gaussian, component (excess
variance). The Euclidean component indicates a continuation of the
resolved source counts at least a decade below the flux limit of the A-2
resolved source flux limit. The acceptable range for the excess
variance is inconsistent with the expected behavior of known sources at
low fluxes unless there is: (1) significant evolution of the known
sources; (2) a new population of sources; (3) large scale source
clumping; or (4) clumping of a truly diffuse source of X-rays.

Parametric models of AGN (quasar) luminosity function evolution

are examined. They are most constrained by the total sky flux. The acceptability of particular models hinges on assumptions currently not directly testable. The comparison with the <u>Einstein</u> Observatory 1-3 keV low flux source counts is hampered by spectral uncertainties.

A tentative measurement of a large scale dipole anisotropy is consistent with the velocity and direction derived from the dipole in the microwave background. The impact of the X-ray anisotropy limits for other scales on studies of large-scale structure in the universe is sketched. Models of the origins of the X-ray sky flux are reviewed, and future observational programs outlined.

CURRICULUM VITAE


Name:  Richard Arrick Shafer.
Permanent address:   c/o W.R. Shafer
                     24756 Embajadores
                     Mission Viejo, CA
                     92691

Degree and date to be conferred: Doctor of Philosophy, 1983.

Date of birth:            ██████████ .

Place of birth:           █████████████ .

Secondary education:   Princess Anne High School
                       Virginia Beach, Virginia
                       Graduated June 1971.

| Collegiate institutions attended | Dates | Degree | Date of Degree |
|---|---|---|---|
| Massachusettes Institute of Technology | Aug 71–Aug 75 | S.B. | September, 1975 |
| University of Maryland | Aug 75–Jan 83 | Ph.D. | May, 1983 |

Major:        Physics.


Professional publications:

The Fluctuating and Isotropic Components of the Cosmic X-Ray Background,
      E. Boldt, A. Stottlemyer, R. Shafer, F.Marshall, R. Mushotzky, S.
      Holt, R. Rothschild, and P. Serlemitsos, IAU Symposium 1978.

The Cosmic X-ray Background: II. Fluctuations, E.Boldt, A.Stottlemyer,
      R.Shafer, S.Holt, R. Rothschild, and P.Serlemitsos, (COSPAR) X-ray
      Astronomy, edited by W.A.Baity and L.E.Peterson, (Pergamon Press:
      Oxford and New York), 1979.

A Complete X-Ray Sample of the High Latitude ($|b|>20$ degrees) Sky from
      HEAO-1 A-2: Log N-Log S and Luminosity Functions, G.Piccinotti,
      R.F.Mushotzky, E.A.Boldt, S.S.Holt, F.E.Marshall, P.J.Serlemitsos

and R.A. Shafer, $\underline{\text{Astroph. J.}}$, $\underline{253}$, 485, 1982.

A Large Scale Height Galactic Component of the Diffuse 2-60 keV Background, DeAnn Iwan, E.A.Boldt, F.E.Marshall, R.F.Mushotzky, R.A.Shafer, and A.Stottlemyer, $\underline{\text{Astroph. J.}}$, $\underline{260}$, 111, 1982.

The Anisotropy of the X-ray Sky, R.A. Shafer and A.C. Fabian, in $\underline{\text{Proceedings of IAU Symposium 104: Early Evolution of the Universe and Its Present Structure}}$, Abell and Chincarini, eds., (Reidel:Dordrecht), 1983.

The $e^+$-$e^-$ Annihilation Line and the Cosmic X-ray Background, D. Kazanas and R. Shafer, in $\underline{\text{Proceedings of IAU Symposium 104: Early Evolution of the Universe and Its Present Structure}}$, Abell and Chincarini, eds., (Reidel:Dordrecht), 1983.

$e^+$-$e^-$ Annihilation and the Cosmic X-ray Background, Demos Kazanas and Richard Arrick Shafer, in $\underline{\text{Proceedings of the GSFC Conference on }e^+\text{-}e^-\text{ Pair Processes in Astrophysics}}$, M. Burns, A. Harding, and R. Ramaty, eds., (AIP:New York), 1983.


Meeting presentations:

Fluctuations in the High Galactic Latitude Background, R.A. Shafer, E.A.Boldt, S.S.Holt and P.J.Serlemitsos, $\underline{\text{BAAS}}$, $\underline{9}$, 616, 1977.

HEAO-1 Measurements of the Fluctuations in the Cosmic X-Ray Background, R.A. Shafer, E. Boldt, L.A. Rose, S.S. Holt, R.E. Rothschild, and P.J. Serlemitsos, $\underline{\text{BAPS}}$, $\underline{23}$, 526, 1978.

HEAO-1 Measurement of the Energy Dependence of the Fluctuations in the Cosmic X-Ray Background, R.A. Shafer, E. Boldt, A.E. Szymkowiak, S.S. Holt, R. Rothschild, and P. Serlemitsos, $\underline{\text{BAAS}}$, $\underline{10}$, 390, 1978.

Fluctuations in the Diffuse X-Ray Background with HEAO-1: Simple Models and their Comparison with Resolved Sources, R.A. Shafer, E.A. Boldt, and G. Piccinotti, $\underline{\text{BAAS}}$, $\underline{12}$, 543, 1980.

Spatial Fluctuations in the Diffuse Cosmic X-ray Background and Models of Point Source Contributions, R.A. Shafer, E.A. Boldt, and R.F. Mushotzky, $\underline{\text{BAPS}}$, $\underline{26}$, 1981.


Professional positions held:

1973-1975:  Undergraduate Research Opportunites Program at M.I.T.

Research performed at the M.I.T. Center for Space Research, Cambridge, Massachusettes, 02139. Summer and part time term employment in analysis of x-ray and optical astronomical data. Supervised by Professor Hale Bradt.

1975-1976: Teaching Asistantship in Physics Department, University of Maryland, College Park, Maryland, 20742. Recitation Instructor for Introductory Mechanics for engineering students.

Summer 1976: Summer research assistantship at the Laboratory for High Energy Astrophysics, Code 661, NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771. Work with Dr. Elihu Boldt and others in the X-ray astronomy group.

1976-1977: Teaching Asistantship as a mechanics laboratory instructor and as a grader for graduate level thermodynamics course.

1977-1983: Full time research assistantship at Goddard, supervised by Dr. Boldt.

1983-present: Research Associate at the Institute of Astronomy, Madingley Road, Cambridge University, Cambridge CB3 OHA, England. Work supervised by Professor A.C. Fabian.

# Prologue

In classical theater a prologue allowed the author to set the scene and alert the audience to important themes, outside the formal strictures of the play. Practice, University regulations, and my advisors tell me that a doctoral dissertation must be a serious, formal work. Therefore, before I assume the impersonal voice of the dissertational "we", I use this prologue as my informal introduction to the reader and to give advice to aid in the reading the text.

The word "Prologue" is not used to imply that this is any work of drama. The thesis does have some dramatic elements: a problem or conflict, developement, and even a certain amount of tension ("Will he ever finish?"). However, as at the conclusion I cannot offer even the pretense of a resolution of the problem, this dissertation is at best a peculiar mixture of minimalist/turgid theater. Retitling it "Waiting for the X-ray Background" would be appropriate, given the length of its gestation. The actual title, Spatial Fluctuations in the Diffuse Cosmic X-ray Background, is more descriptive. A reader might note, though, that I make no explicit reference in the text to a "diffuse background," instead using the less familiar phrase "X-ray sky flux." This choice of words avoids prejudging the true nature of the bright essentially isotropic emission from 3 to 100 keV that current observation is unable to resolve into individual sources.

In addition to failing as drama, this work is obviously not a "light read." It is, in fact, too long to expect any but the most motivated and dedicated to actually read through from front to back. Many potential readers may be put off by the bulk and after skimming the abstract set it aside. This is unfortunate, as I believe almost any astrophysicist would find something of interest in the topics discussed. In addition to the X-ray background, I will touch on design features of the HEAO 1 A-2 experiment, a general measure of the distinguishabiltiy of two models, statistical issues, determinations of number counts of sources versus flux and how they are effected by cosmology and detector

spectral bandpass, and "much, much more!" To make all this material more accessable, I have provided a series of mini-abstracts that head each section, which outline the material covered. These are collected in a "Précis" before the standard Table of Contents. Therefore, the mildly interested astrophysicist's best strategy is to read Chapter I for an introduction to the broad topic of the X-ray background, then to skip to Chapter IX for the main conclusions, and follow with the Précis to determine which individual sections merit a detailed reading.

Those who for whatever reason feel required to wade through the material in its entirety will also benefit from the mini-abstracts. They may also benefit from the two words in large, friendly letters that I received when I had despaired of ever actually finishing (see Frontispiece following).

# Don't Panic

This work is dedicated to
MY FAMILY,
particularly my parents,
Walter Richard and Joanne ▮▮▮▮ Shafer,
as well as my sibs, co-sibs,
and divers other Friends and Relations.

Through word and deed, they have shown me that the true value of the examined life is in both the examination and in having something to examine.

## Acknowledgments

Almost anyone with whom I've had a more than casual interaction with during the last two+ years has had some contribution to this work. These contributions can be classified roughly into three catagories: contributions to the science proper of the dissertation; general scientific education and support during the writing of the dissertation; and assistence in making the past years of my life both interesting and enjoyable so that I could see this project through to its completion. Many individuals were outstanding contributors to two or all three of these catagories. To properly aknowledge everyone, and to catalogue their total role in this work would require many more pages to be added to this already long opus. (In addition, these words are written during the last hours of my graduate student career.) Therefore, to those I may miss, whose contribution may not be fully recognized, please be assured that all have my lasting and heartfelt gratitude.

Traditionally, the aknowledgements begin with the student's advisor. I have been fortunate in having had several top notch scientists to play that role, both in a formal and informal capacity, each with their own style and advantages. Elihu Boldt, as my research advisor for the last six years, provided impetus and encouragement for my studies of the diffuse X-ray background. He also showed the advantage of an open, inquiring, mind in a field where surprise is one of the great rewards. Prof. Frank McDonald, as my formal academic advisor, helped me to negotiate the shoals of my academic career, as well as making my fellow graduate students and me feel that we make a definite contribution to the science and other aspects of life at the Labratory for High Energy Astrophysics, Goddard Space Flight Center. As my unofficial cheerleader, source of information, and general pundit to the masses, Dr. Rich Mushotzky played an important role in making me see that the research of this dissertation might have something to say about a wide range of current problems in high energy astrophysics. Though this may appall him, Chapter VII owes its breadth to his "Let's try to understand what this means."

vii

For specific contributions to the science, I wish to thank:

My predecessors in the analysis of the fluctuations, Prof. Andy Fabian, and Drs. Dan Schwartz, John Pye, and Bob Warwick. All expressed continuing interest in my results, even after long delays. All also kindly shared many unpublished details of their own analysis programs.

Prof. Maarten Schmidt and Dr. Martin Elvis for communication of their results in advance of publication.

Prof. Paul Smith of the University of Maryland for pointing me to the proper proof of Boole's Inequality as a technique for including non-independent measurements.

Dr. Fred Derrick for providing some additional thoughts and helpful hints for tackling the sticky statistical problems that arose.

Guy Harris for directing me to the improved radix-8 FFT algorithm that made the model fits merely expensive rather than horrendous.

Dr. Rick Harnden for providing important information on the spectral window used in the Einstein Observatory Deep Survey.

My early appreciation of the importance of understanding the properties of the detector internal background owes much to a reading of an unpublished analysis of the fluctuations performed by Drs. France Cordova and Gordon Garmire.


The particular contributions of past and present members of the LHEA/GSFC X-ray group are themselves almost too numerous to mention. Besides Drs. Boldt and Mushotzky, particular mention should be made of:

Drs. DeAnn Iwan and Frank Marshall, who provided much of the initial work for the large scale galactic structure. Frank also did the preliminary analysis of the Compton-Getting signal, and also performed the statistical analysis of the confusion noise contribution to fit source intensities (Figure III:15).

Dr. Alan Rose and Allyn Tennant each performed an in-depth analysis of the environmental response of the detector internal background.

Dr. Jean Swank, Dr. Robert Miller, and others performed the careful work required to derive useful detector spectral and spatial

response functions. Jean and Dr. Nick White shared their expertise on the low luminosity galactic sources.

Individuals who also improved the quality of the dissertation through their particular fields of experience include Drs. Gail Reichert, Esther Hu, and Richard White.

Special recognition for aid in the understanding and debugging of the various quirks in the analysis algorithms goes to Andrew Szymkowiak. His unflagging interest and informed skepticism kept me from sweeping some issues under the rug as my predecessors had. Andy is also responsible for a large body of analysis software, one example of which is the source of the information contained in the table in Section II:4.

The sucess and value of the numerical results of the dissertation rests on the excellant performance of the A-2 experiment on HEAO 1. The experiment design, execution, integration, and callibration were carried out by a crack team of scientists, engineers, and technicians. Of particular note are the efforts of Drs. Boldt, Serlemitsos, and Rothschild. The initial data analysis software was written under the supervision of Dr. Allan Stottlemyer.

Out of the great host of software writers, documenters, and maintainers I want to give special thanks to Marilyn Newhouse and Debbie Derrick for their additional talents as moral supporters. An early analysis of the fluctuations used data taken from the OSO 8 B detector. The initial software for this project was produced by Keith Hope.

The preparation of the text of this thesis was aided by an enormous effort by the Thesis Defense Team. Early drafts of the thesis were read in their entirety by my stalwart proofreaders: Elihu Boldt, Esther Hu, Richard Mushotzky, Andy Szymkowiak, and Meg Urry[*]. Each of my proofreaders offered their own particular interpretation of the draft. Meg Urry, for instance, has methodically expunged all of my more colorful (for which read ungrammatical or archaic) phrases. Additional

_____

[*]There has been some concern among my proofreaders that they will be the only people to negotiate more than 90% of the document in its entirety. Should counterexamples of this hypothesis exist, they should contact M. Urry at GSFC Code 661.

proofing and esthetic advice has been provided by Julia L. Robinson-Saba. Joe Jones can lay claim to providing useful editorial services for both my bachelor's thesis and this dissertation.

The physical production of the thesis was made using the SCRIPT word processing software of the University of Waterloo, as implemented on the GSFC IBM 3081 computer. The physical hard copy was created using an intermediate program that allowed multiple font output to a single wheel Anderson-Jacobson AJ832 terminal. A critical piece of that software package was written by Dr. Nand Lal. The figures were produced through the talents of Adam Thompson, with additional graphics support from Frank Shaffer and Harry Trexel. The software for producing the contour and grey scale plots of Figures I.2 and I.3 was written by Allyn Tennant and Andy Szymkowiak. Additional support, far beyond the normal call of friendship, was provided by Gail Reichart, Andy Szymkowiak, and Esther Hu in the assembly of the final typescript.

The wealth of moral support I was able to draw upon made it possible for my graduate career to come to this successful conclusion. Part of this support was due to the wide range of social activities provided by friends to remind me that there is life after dissertation: from lunch with Steve early in my graduate career to pleasant dinner conversation with my Thesis Conscience at the conclusion; from formal dinner parties in Baltimore highrises to dancing in the streets in New Orleans; from the Fun Crowd activities to stalking the deer. Special ambience aknowledgements go to Mark of WHFS, who made week night music to lucubrate by.

On a more serious note: it is doubtful that I could have successfully survived all the self-inflicted wounds to my graduate transcript if I had not been able to avail myself of the knowledge, skill, and general motherly advice of Mary Dobbins, late of the Physics Department Graduate Records. No student could have asked for a better den mother.

Financial support during my thesis research was provided by a

## Précis

These fluctuations are also shown to have an impact on other measurements of the X-ray sky, as well as measurements of low intensity sources.

response to get n(I). The distribution for a Euclidean power law is studied, contrasting the full distribution with analysis based only on the variance of the distribution. We present plots of the fluctuations for various values of $\gamma$, the power law index, comparing them with the Euclidean form. The importance of sharp cutoffs in N(S), particularly for certain values of $\gamma$, is presented, along with an outline of how to include them as a simple extension of power law models. Finally we consider variants of the single power law, including models whose behavior is fixed at a certain flux, as well as general modifications to low-flux behavior represented as a Gaussian term.

77

5: Comparisons between N(S) models - We derive a general tool, the likelihood difference, to tell how similar a test distribution is to the (presumed) "true" distribution. The relationship of this tool to the likelihood statistical tests allows us, when testing a parametric family of models, to estimate which range of parameters will, on average and for a given number of observations, be outside acceptable confidence contours. We apply the tool to pure Gaussian distributions, and to the fluctuation models. The likelihood difference provides a quantifiable indicator of the sensitivity of the fluctuations to details of models for N(S). The ability to distinguish two models for the fluctuations is degraded in the presence of non-fluctuations noise. We find that the behavior of the lowest intensity sources can be well modeled as a purely Gaussian contribution.

95

6: The Evaluation of Ideal Experiments for the Measurement of the Fluctuations - We consider the competing effects of counting statistics, number of observations, and variations in the non X-ray count rate in the design and analysis of experiments to measure the fluctuations. When we wish to measure the W parameter of N(S), large sky coverage is preferred to a few deep measurements. An optimal beam size for such measurements exists, depending on the internal background contribution to the noise. To measure the behavior of the low flux sources through the invariant excess, $\sigma^2_S$, lower noise is more important

our measurements. In this section we determine the variation of the internal background as a possible source of noise. The internal background mean count rate is measured, and it is shown that for data selected for low values of McIlwain "L", the internal background is well represented by a constant rate modified by a slow secular drift. Any residual variation in the internal background is small. A much larger effect on measurements of the fluctuations is the contamination due to large scale structure in the sky. A thick disk of emission associated with the galaxy is the principal component. A residual "24-hour" signal is seen with only slight significance, ~95%. This signal is consistent in direction and magnitude with that from motion of the solar system with respect to the cosmic background as indicated by similar anisotropy measurements in the microwave.

counting statistics. The final distribution is a convolution of the $P_I$ distribution with a Gaussian representing the additional noise. A particular difficulty is encountered, though, when the size of the added variation changes from one measurement to the next. This means that each measurement is drawn from a unique distribution. A further refinement is necessary if in order to minimize the effect of bright sources on model fitting, a high intensity selection cutoff is imposed on the data. The final result of this section is a derivation of the probability distributions that properly describe the data set.

2: Statistical Techniques Used − Model fitting for the fluctuations requires care in the selection of the proper statistical tools. The integral probabilities, $P_j(r_j)$, may be used in a family of statistical tests that will tell the goodness of the fit of the model. For the determination of confidence contours the likelihood statistic is the preferred technique. We also outline a general rigorous approach to combine results from several non−independent determinations of a confidence contour based on Boole's Inequality.

1: Recapitulation of Data Characteristics − We review the features of the data chosen for fitting models. The various subsets of independent, and non−independent data are described. The fitting process is sketched.

2: The Reconciliation of the Observed and Model Means − The fluctuations distribution, $P_I'(I-\bar{I})$, is calculated from the model $N(S)$ to give a distribution about the mean count rate. If in performing the fits the model mean is identified with the mean count rate of the observations, a systematic bias is introduced. We show in this section that the value for the size of the fluctuations, $W_{eff}$, can be strongly correlated with the value chosen for $\bar{I}$. The bias and correlations are due to the assumption that the model has the correct contribution from resolved sources to

the mean count rate. A simple correction term removes both bias and correlation, allowing the separation of the uncertainty in $\overline{I}$ from determinations of $W_{eff}$ of the uncertainty.

182

3: Fits of Euclidean Models — We test a power law model for $N(S)$ with index $\gamma = 5/2$, characteristic of a uniform distribution of sources in Euclidean space. The quality of the fit is good. The various subsets are used to determine the confidence region for the width, W. The answers from all the sets are in agreement. By fitting models with sharp cutoffs at low fluxes we are able to estimate the limitations of our models in describing the low flux region of $N(S)$. The limit is set by a variance condition, rather than by the number of sources per field of view. The best-fit $\gamma=5/2$ model is used to define a fiducial $N(S)$ to which subsequent models are compared.

186

4: Fits of Gamma and W — We consider non-Euclidean power law $N(S)$ models, i.e. those with $\gamma$ not fixed at 5/2. We find that the WW (whole sky weighted) data-set provides much greater constraints on the allowed values for W and $\gamma$ than a dataset of independent measurements. Euclidean and near-Euclidean models provide a good description of the data. The acceptable models show that the $N(S)$ is most tightly constrained at fluxes of $3 \times 10^{-4}$ counts $s^{-1}$ $cm^{-2}$, an order of magnitude below the limits of resolved source counts. A low flux cutoff, imposed by the requirement of not exceeding the total X-ray sky flux, has a significant impact by allowing with the confidence region high $\gamma$ models that otherwise would have been excluded. The technique used for calculating the mean count rate also has an impact. Pegged models, where $N(S)$ is constrained at a given flux $S_p$, are examined.

191

5: Fits of the Excess Variance and W — We describe fits when $N(S)$ is parameterized as a pure Euclidean power law with an additional component whose contribution to the fluctuations is a pure Gaussian. The variance of the Gaussian, $\sigma^2_I$, and the value of W for the power law are strongly correlated. We present the 90% confidence contours for these parameters. The data

are consistent with a zero excess variance, with a 90% upper bound to $\sigma^2{}_I$ of 0.064 (counts $\exp^{-1}$)$^2$, or in collimator-size-independent units of flux, $\sigma^2{}_s$, 0.0008 (counts $s^{-1}$ $cm^{-2}$)$^2$. This limit is derived from the overlapping data WW dataset. The constraints imposed -by the sets of independent measurements are much less severe. The limits on the excess variance is used to place limits on a broken power law model for N(S). We show how the confidence range would be modified if independent information could provide an exact value for W.

at the highest fluxes. Corrections for the Malmquist bias are also estimated. The range of values for K is consistent with the Euclidean component of the fluctuations. We combine the resolved source fits with the fluctuations analysis to produce a significant reduction in the confidence regions for K (or W) and the excess variance, especially for the independent data subsets. The 90% confidence range for $\sigma^2_I$ is 0.0 to 0.05 (counts $exp^{-1}$)$^2$, (or $\sigma^2_S$ is 0 to $6x10^{-4}$ (counts $s^{-1}$ $cm^{-2}$)$^2$) using overlapping data.

<div align="right">220</div>

2: The Expected Excess Variance from Non-Evolving Populations — Analysis of the fluctuations has been reduced to applying the determination of two parameters: the strength of the fluctuations (measured by K) for a Euclidean N(S) component and an excess variance characterizing the deviation from a strictly Euclidean form. In this section we examine in detail the populations of high latitude resolved sources as studied in the complete sample of Piccinotti et al. [1982], and predict their contribution to both the K values and the excess variance, assuming no evolution. The total excess variance for all the populations is ~$-3.3x10^{-4}$ (counts $s^{-1}$ $cm^{-2}$)$^2$. This negative value is dominated by the contribution from clusters of galaxies, due to the rapid effect of the cosmological redshift on their contribution in the HED 1 Layer 1 band. The value is significantly less than that allowed by the fluctuations, except for our most conservative estimate using non-overlapping data. The fluctuations therfore require an additional source of variance, perhaps a new population of X-ray sources or evolution in one of the known populations.

<div align="right">234</div>

3: Sources of Excess Variance: New and Evolving Populations — New or evolving populations of sources may provide the amount of invariant excess required by the fluctuations analysis, compensating for the deficit calculated from the non-evolving luminosity functions of the known sources. The upper limit on the invariant excess places constraints on any new population that could provide the bulk of the sky flux. There must be at least 50 such new sources per square degree, with an average flux of no more than ~$3x10^{-13}$ ergs $s^{-1}$ $cm^{-2}$. Any evolution of the AGN

luminosity function would also affect the AGN contribution to the excess and the sky flux. The requirement that the observed sky flux not be exceeded places the strongest constraints on the amount of such evolution. For pure luminosity evolution, it is difficult to provide the missing invariant excess without overproducing the sky flux, unless moderate luminosity objects ($\gtrsim 0.1$ $L_{44}$) undergo more evolution than the lower luminosity sources. Models of luminosity-dependent density evolution (index evolution) are also considered, and are consistent with the required behavior. The possible presence and form of a high luminosity extension of the AGN luminosity function can affect the high flux N(S). While observations of the resolved sources cannot constrain a non-evolving extension of the luminosity function, they do rule out strong luminosity evolution by these high luminosity sources.

254

4: Results from the <u>Einstein</u> Observatory: Source Counts — We review the results of the low flux source counts obtained from the <u>Einstein</u> Observatory. Though the differing energy bands of <u>Einstein</u> and the A-2 experiment make comparisons strongly dependent on spectral assumptions, the "Medium Survey roughly equivalent to the range of greatest sensitivity of the fluctuations. The N(S) curves derived are in agreement, again depending on spectral assumptions. The identification of the sources which constitute the Medium Survey implies some evolution by AGN from the luminosity function observed in the XCS. If the evolution is pure luminosity evolution, where $L(\tau) = L(0) \exp(Q_L \tau)$, ($\tau$ is the fractional lookback time), then the Medium Survey source counts and the XCS derived current epoch luminosity function predict $2 \lesssim Q_L \lesssim 4$, $q_o = 0.1$. If a substantial number of the low luminosity AGN are not absorbed or are only partially covered with absorbing material, the allowed range must be significantly reduced. Similar results are obtained for the original "Deep Survey The analysis of these sources is sensitive to the behavior of N(S) in a narrow flux range, and its interpretation depends on the assumed power law index of N(S). No conclusion about the presence or absence of evolution for BL Lacertae objects can be drawn solely from their relative frequency in the XCS and Medium

5: Comparisons of AGN models with Optical Data – Studies of X-ray sources in other spectral bands, besides having an intrinsic value for elucidating the physical mechanisms of the emission, can provide indirect information about the behavior of those sources in the X-ray band. For example, with data on the connection between optical and X-ray fluxes, optical source counts can be used to predict the contribution of AGN to the total X-ray sky flux. Such a program is complicated by nonlinear correlations between optical and X-ray luminosities. Many authors have found that increases in the optical luminosity produce a smaller increase in the X-ray luminosity than strict proportionality would predict. That is, $L_x \propto (L_{opt})^\delta$, with $\delta < 1$. However, this is difficult to reconcile with the fact that the X-ray luminosity function is flatter than the optical luminosity function. The optically derived-parameters for evolution must be corrected by the deviations from simple proportionality, in order to apply them to the X-ray band. The observed correlations between X-ray and optical luminosity predict that $Q_{L;x}$ must be much less than $Q_{L;opt}$, the exact value depending on the actual form of the correlation. The use of the optical source counts to predict the X-ray sky flux involves several important extrapolations.

309

6: Excess Variance from Non-Poisson Fluctuations – Non-Poisson distributions of X-ray sources can produce significant additional contribution to the fluctuations. We present estimates of the amplification of the fluctuations due to sources clumped on supercluster scales. We also sketch the utility of studies of the X-ray anisotropy for probing the general distribution of matter. On the very largest angular scales, the comparison of the Compton-Getting (dipole) signal in the X-ray and microwave bands supports the hypothesis that the overdensity responsible for our observed velocity is "local" compared to the scale of the sources that produce the X-ray sky flux. The limits of the excess variance constrain the magnitude of lumps of X-ray

volume emissivity, the size of the limit depending on the lump ·size. For data from our HED 1 smeared detector, the strongest constraints are placed on structures larger than ~400 Mpc. These constraints can be translated to limits on matter clumping of the universe at scales otherwise not easily accessible. Using a naive Euclidean approximation, the fluctuations do not permit highly clumped structures ($\delta\rho/\rho>1$) larger than ~100 Mpc.

Table of Contents

Contents

Contents

Contents

# List of Figures

# List of Tables

CHAPTER I

INTRODUCTION AND OVERVIEW

The sky as experiment

The fundamental difference between astrophysical observations and experimental measurements in other sciences has been noted often. In astrophysics the processes being investigated are removed from the direct control of the "experimenter." In other sciences, experimental variation and controls can simplify the data analysis needed to make direct tests of theoretical models. Astrophysicists are unable to perform an experiment with different initial conditions and controls but the range of astrophysical phenomena allows the observer to treat each object of a class as an additional run of the experiment. The careful comparison of data taken over the entire range of behavior in a class can be a rich source of information about physical processes in the universe. For example, much of our understanding of stellar interiors, origins, and evolution is based on measurements taken from stars across the H-R diagram, including extreme non-main-sequence objects.

Studies of populations of objects lend themselves not only to "direct" measurements of processes in the objects, but also to indirect measurements of their environment as well as determine the evolution of the objects. For instance, the velocities of galaxies in clusters indicate the presence of a significant amount of matter not contained in normal light-emitting objects. The fraction of this matter in intracluster gas may then be related to the morphological structure of the cluster members. In another example, the population statistics of members of a stellar cluster offer tests of stellar formation and evolution. The largest scale example of this approach is an observational program that counts the contents of the entire universe in its measurement sample, such as the measurement of large scale correlation of galaxy positions in the sky [see Peebles 1980].

The extension and further abstraction of these universal population studies is an examination of the sky as a whole. The brightness of the sky, even without an explicit identification of its origins, tells us something about the sources that contribute to it. This has been recognized since Edmund Halley first advanced the argument now popularly known as Olbers' Paradox: the darkness of the night sky places a limit on the age or the distribution of stars. The study of the sky as a whole is clearly crucial to our understanding of the 2.7 Kelvin black-body radiation, which apparently does not originate in discrete sources and has been identified with relic radiation from the Big Bang. Even if this identification is not correct, the isotropy and physical transparency of the universe at these wavelengths argue that its origins are at large distances from our galaxy, and the sheer magnitude of the radiation intensity makes it important to understand.

## The cosmic background spectrum

Though the microwave background is dramatic in its intensity and possible cosmological implications, the sky at other energies is equally interesting. Figure I.1 shows the cosmic background intensity over a wide spectral band. Though the details may have changed, the essential picture has been known for a decade [Boldt 1971]. The dominant feature is the 2.7 Kelvin black-body component [reviewed by Weiss 1980] discussed above. The power-law component in the radio, from $10^6$–$10^{8.5}$ Hz, is not directly observed, but inferred after the removal of the synchrotron signal produced by cosmic rays in our galaxy [reviewed by Longair 1978]. The non-galactic component is well explained as a superposition of radio galaxies and other discrete radio sources. The intensity of the radio emission is consistent with the convergence of the number of radio sources at low flux limits [e.g. Jauncey 1977; Wall, Scheuer, Pauliny-Toth, and Witzel 1982]. The optical and ultraviolet points are upper bounds to a background probably composed of starlight and other emission from distant galaxies. At the ionization limit of hydrogen, 13.6 eV = $3.3 \times 10^{15}$ Hz, interstellar gas becomes opaque. At higher energies helium and higher Z material contribute significantly to the absorption so that until photon energies exceed
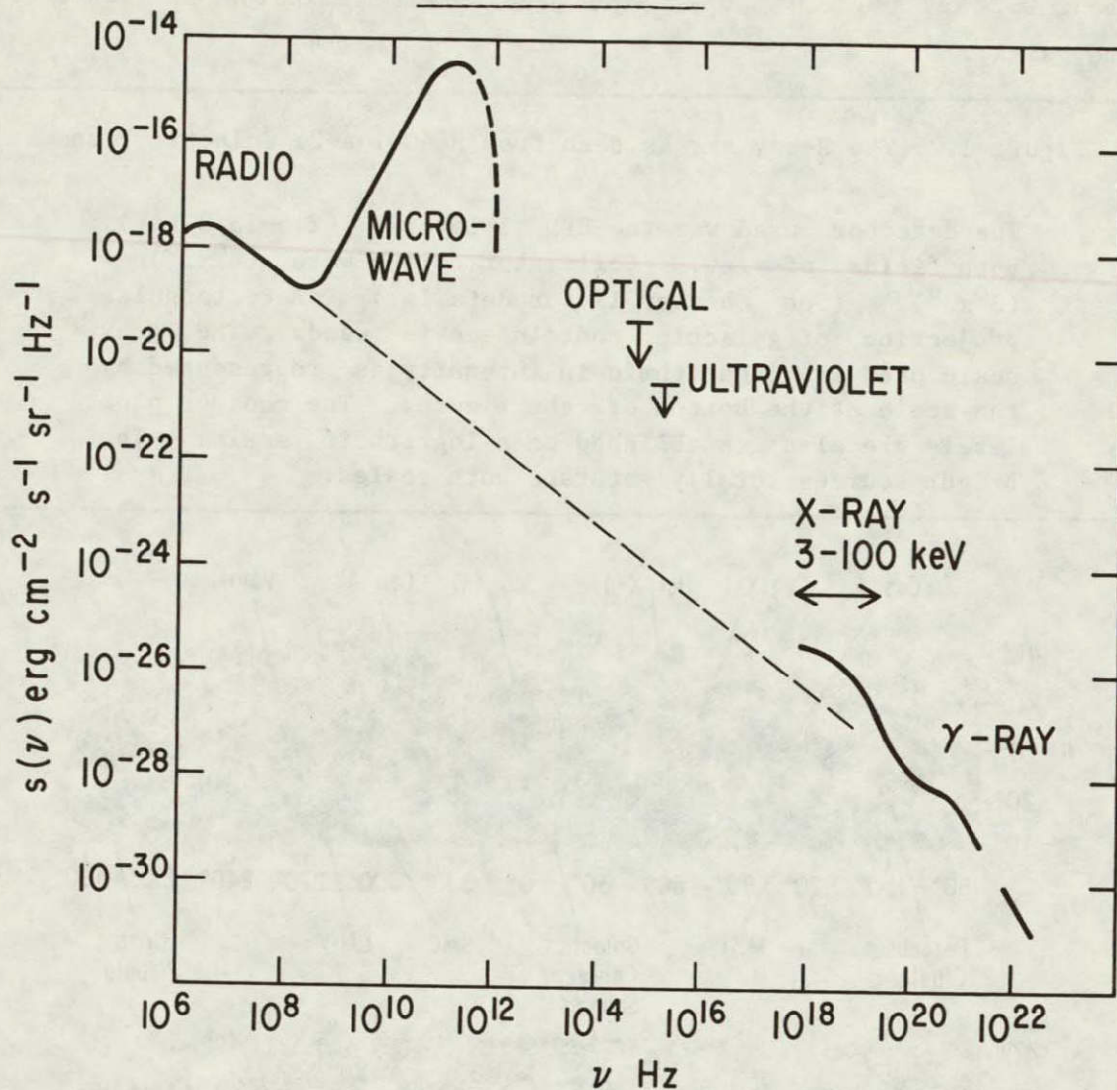
Figure I.1: The isotropic sky flux, s(ν) versus ν

The dashed line extends the emission of the radio band
discrete sources to higher energies, assuming that their
spectra continue to follow the observed power-law form
with energy index, α, ∼0.7 ($f_\nu \propto \nu^{-\alpha}$). Boldt [1971]
suggested from such a figure that the same sources may
make a significant contribution to the X- and γ-ray
backgrounds. The references: Radio and microwave curves
from Longair [1978]; optical 90% upper bound, Dube, Wickes
and Wilkinson [1979]; ultraviolet bound, Paresce, McKee
and Bowyer [1980]; X-ray and low energy γ-ray, Marshall et
al. [1980] and Rothschild et al. [1983]; and high energy
γ-ray reviewed in Fichtel and Trombka [1981].

about 0.2 keV ($4.8 \times 10^{16}$ Hz), almost no outside emission can penetrate the galaxy.

---

Figure I.2: The X-ray sky as seen from HEAO 1 A-2: Galactic Plane

The detector used was the HED 1 layer 1 (2.5-13.3 keV), both fields of view. Collimator sizes were ($6^{\circ} \times 3^{\circ}$) + ($3^{\circ} \times 3^{\circ}$). (See Chapter II for details.) A rectangular projection of galactic coordinates is used. The grey scale used is logarithmic in intensity as represented by the scale at the bottom of the figure. The contour plot levels are also established on a logrithmic scale. The bright sources totally saturate both scales.

---
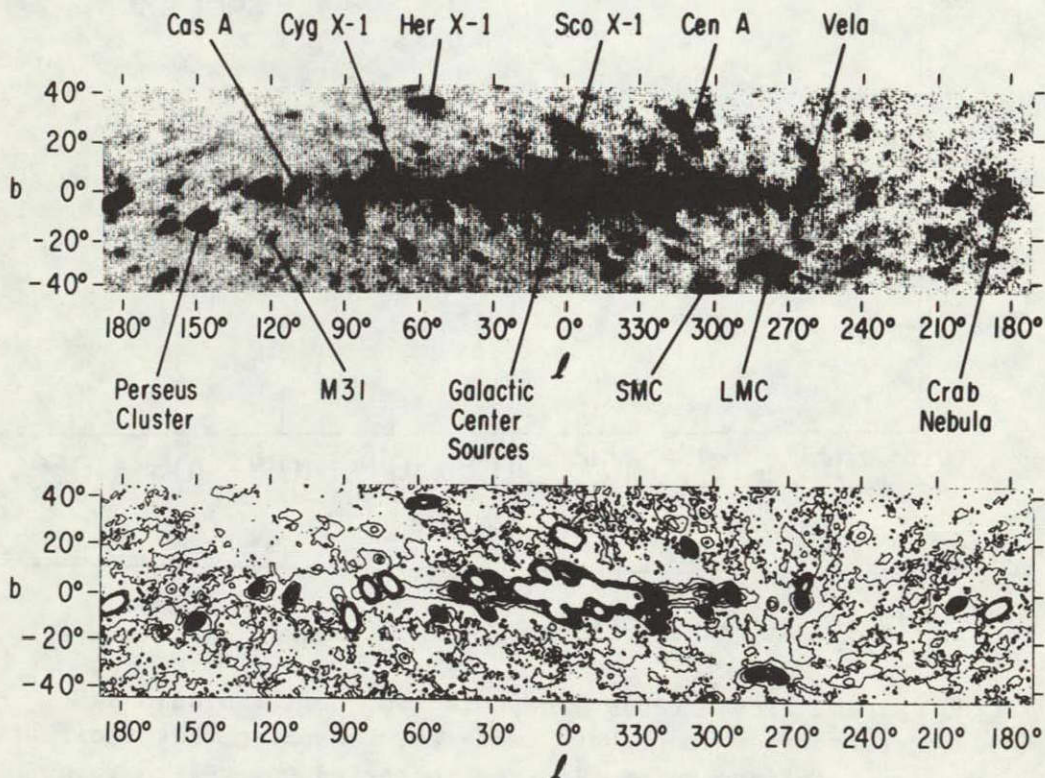


The X-ray sky

The story at still higher energies changes again. (A review of

the physics important for studies of the X-ray sky can be found in Boldt [1974]. Later reviews of recent data and their consequences are given by Boldt [1981b] and Fabian [1981].) The best current all-sky survey in the X-ray band was performed with the first High Energy Astrophysical Observatory, HEAO 1. (A general description of HEAO 1, its four experiments and a summary of their important results as of December, 1980 is given by Boldt [1981a].) A surface brightness map produced using one of the detectors of the A-2 experiment[*] onboard HEAO 1 is presented in Figures I.2 and I.3. Figure I.2 covers the region between ±45° in galactic latitude. Point sources are smudged because of the finite size of the collimators compared to the map resolution. The galactic plane is easily recognized by the line of sources, with the galactic center an unresolved (at this resolution) bright region. As well as the bright galactic sources (Sco X-1, the Cygnus region sources, and others throughout the plane) this map contains extragalactic objects: the Magellanic Clouds, the brightest cluster of galaxies (Perseus), the active galactic nucleus associated with the radio source Centaurus A, and the brightest and nearest "normal" galaxy M31. The region about the north galactic pole (Figure I.3) also contains interesting, relatively bright, sources: the Coma cluster, the Virgo cluster (center of our local supercluster), the Seyfert I galaxy NGC 4151 and the QSO 3C273. In contrast the south galactic pole (also shown in Figure I.3) is undistinguished, containing a few discernible low intensity sources associated with various clusters and active galactic nuclei.

Though the south polar region seems uninteresting by virtue of its few resolved sources, it and the other maps are dominated by a feature of great interest: a bright, isotropic, spectrally intriguing unresolved flux. The salient aspects of this unresolved X-ray sky are:

(1) It is bright. The extra-solar all-sky flux, $S_{as}$, measured by an omnidirectional detector with a bandpass of 2-10 keV is $\sim 7.4 \times 10^{-7}$ ergs s$^{-1}$ cm$^{-2}$. In comparison, the quiet sun, with a

[*]The HEAO 1 A-2 experiment was a collaborative effort led by E. Boldt of Goddard Space Flight Center (GSFC) and G. Garmire then at the California Institute of Technology (now at Pennsylvania State University), with collaborators at Goddard, Cal Tech, the Jet Propulsion Laboratory, and the University of California at Berkeley.

---

Figure I.3: The X-ray sky as seen from HEAO 1 A-2: Galactic poles

Same configuration as for I.2, restricted to the north and
south polar regions. The gray scale has been changed to
promote the resolution of detail. The concentric circles
are lines of 80° and 60° latitude. The bars to the right
of the gray scale are a histogram of the intensity
distribution in the two fields. This distribution, wider
than can be expected from counting statistics alone, is
the subject of this dissertation.

---

NORTH

Coma Cluster  0°  3C273

90°  −270°

NGC 4151  180°  Virgo

0°

90°  −270°

180°

SOUTH

SMC  0°

270°  −90°

180°

0°

270°  −90°

180°

---
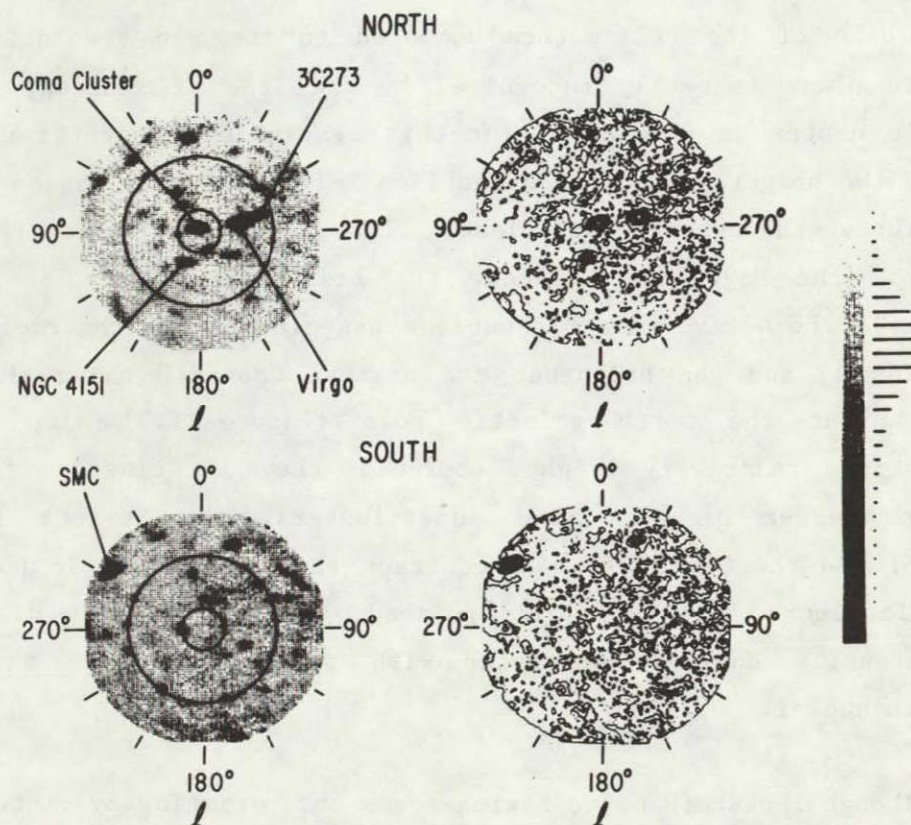
much softer spectrum, contributes only $5\times10^{-7}$, while the
brightest extra-solar object, Sco X-1, has a flux of $2\times10^{-7}$.
For studies at energies greater than ~3 keV, the limit of
sensitivity of pre-HEAO 1 experiments was about $3\times10^{-11}$ ergs
$s^{-1}$ $cm^{-2}$. An all-sky survey performed with the HEAO 1 A-2
experiment has detected 66 sources at high galactic latitudes
($|b| > 20°$) above $2\times10^{-11}$ ergs $s^{-1}$ $cm^{-2}$, which were not

Introduction

"Page missing from available version"

bremsstrahlung with a temperature of 40±5 keV. A thermal
spectrum, which is essentially exponential, ~exp(-E/kT),
provides a good representation because of a relatively sharp
break in the data at about 40 keV. No population of single
index power-law spectra provides the right shape. While a
suitably evolving population of sources with a sharp change in
the power-law index has been fit in the 3-50 keV range
[De Zotti et al. 1982], strong constraints are placed on such
a population.

All three of these properties make the study of the X-ray sky of
astrophysical interest. The production of X-rays indicates that the
originating processes involve intrinsically high energies. The
brightness shows that such phenomena are not uncommon. The isotropy
indicates that these are probably distributed out to cosmological
distances, ($z \gtrsim 0.5$). The structure in the spectra may give clues about
the astrophysical nature of the processes, acting as a general
diagnostic about phenomena presently not directly observable in any
other way.

## Models for the X-ray sky

Since resolved extragalactic sources are observed, viz. clusters
of galaxies and active galactic nuclei, their continued existence at
fluxes below the resolution limit is a natural source for at least part
of the sky's brightness. Is this extension continued to low enough
fluxes to make up the total sky brightness? Current estimates of the
density of both clusters [McKee et al. 1980; Piccinotti et al. 1982]
and Seyferts [Piccinotti et al. 1982] indicate that they could make up
roughly 25% of the total flux. This estimate assumes no evolution in
the number or luminosity of such objects. Seyferts and other active
galactic nuclei are thought to be the low luminosity counterparts of
QSOs. QSOs are a population that undergo significant evolution in
luminosity and/or number density [Green and Schmidt 1978; Cheney and
Rowan-Robinson 1981a; Schmidt and Green 1983] so it would not be
surprising if other active galaxies also evolve. In fact the lower
luminosity active galaxies can make up the 2-10 keV sky flux with less

evolution than has been inferred from optical studies for QSOs [Avni 1978].

This simple picture must satisfy the stringent constraints placed by the form of the sky and various source spectra. There is at present no coherent accepted picture to explain the spectrum and intensity of the X-ray sky from 3 to 100 keV. Clusters of galaxies and active galactic nuclei dominate the resolved sources at the limits of the HEAO 1 survey, but neither class is typified by spectra that could provide the diffuse flux spectrum. The emission of clusters of galaxies is characterized by thermal spectra, but the temperatures are much too low, $\lesssim$ 10 keV [Mushotzky et al. 1978]. Seyfert and other active galactic nucleus X-ray spectra are well modelled by a single power-law component with an index steeper than the diffuse sky from 2-20 keV [Mushotzky et al. 1980; Rothschild et al. 1983]. Although they may be expected to make a significant contribution at energies $\gtrsim$ 100 keV where the thermal fit no longer provides a good description, they are not able to make up the thermal component without significant spectral evolution.

The radio-loud QSO 3C273 is an exception. Using A-2 data Worrall et al. [1979] were able to fit the 2 to 50 keV spectrum with a power law of energy index, $\alpha$, equal to 0.41. This makes it the only extragalactic source whose spectrum appears compatible with the requirements of the unresolved sky spectrum in the 2-20 keV range. However, the single power law interpretation of the spectrum may be obscuring important structure. For example, at the higher energies accessible to the A-4 experiment (13-120 keV) Primini et al. [1979] are able to fit the spectrum with an $\alpha$ = 0.7 index. This is entirely consistent with the behavior of other active galactic nuclei and is significantly steeper than the required behavior for making up the sky flux. Primini et al. suggest that the spectrum may break ~ 20 keV, a behavior which is weakly evident in the anlysis of Worrall et al. (see their Figure 1). In addition the detection of 3C273 by the COS-B satellite in the $\gamma$-ray range in conjunction with limits on the extragalactic background at that flux implies that objects with spectra like 3C273 cannot contribute more than ~5% of the 2-10 keV flux without

overproducing the γ-ray background [Setti and Woltjer 1979]. QSOs are observed to be significant emitters of X-rays in the Einstein 0.5-3.0 keV band. Most of these objects are radio-quiet quasars, as are most of the optical quasars. The only radio-quiet QSO with good broad band X-ray spectral information, 0241+622, has a best fit spectrum that is steeper than the average Seyfert's, although it is consistent with the average spectrum at the 90% level [Worrall et al. 1980].

Current observations can provide no examples of sources whose spectrum are consistent with the unresolved X-ray sky spectrum. Thus from spectral information alone, if the X-ray sky flux is produced by unresolved point sources, they are either a new population not currently observed in the 2-50 keV regime, or a known population of sources that has undergone significant spectral evolution. Such evolution could indicate a fundamental shift in the physical processes which power these sources. Models of the former type have been advanced by Bookbinder et al. [1980] (phenomena associated with young protogalaxies) and Carr [1980] (accretion onto primordial high mass black holes). A specific example of the latter type [Leiter and Boldt 1982], presented a model for the spectral evolution of active galactic nuclei, with a unified explanation for spectral features in both the X-ray and gamma-ray sky flux. If unresolved point sources do not generate the sky flux there is a third possibility. The thermal component could be due to a significant amount of hot intergalactic gas [Field and Perrenod 1977]. If the gas were uniformly distributed throughout the universe it would require a large energy input to be heated to the required high temperature. Clumping the gas would reduce this difficult energy requirement (but raises problems of its own, as we shall see [Fabian 1981]).

## Fluctuations in the diffuse sky and discrete sources

As only 1% of the sky flux is accounted for in the high energy observations of resolved sources, it is important to try to understand their behavior at lower flux levels. Questions to be asked are: do they show convergence in their numbers; do they instead show strong evolution

at low fluxes; do their summed spectra show any change? Without actual
observations of individual sources these questions appear unanswerable.
However, following the lead of radio astronomers we can utilize the
expected properties of the sources, en masse, to place limits on their
behavior at flux levels that are inaccessible directly. The discrete
nature of the sources and their presumed random distribution across the
sky will introduce an inherent "noise" anisotropy to the sky brightness.
The statistics of this noise are grounded in the statistics of Poisson
noise.

Consider a series of ideal measurements of electric current,
performed by integrating the charge for a fixed length of time. For
suitably low currents, even with ideal apparatus the total integrated
charge would not be equal for each measurement. The distribution of the
measurements would be roughly Gaussian. A natural interpretation is
that we are seeing the effect on the measurements of counting statistics
due to charge being carried in discrete units, i.e. electrons.
Assuming the number of charges counted in each measurement is governed
by Poisson statistics, how the variance of the distribution changes as
the experimenter modifies the integration time or current allows the
inference of the charge of the electron. Thus a property of the
particle can be determined without the examination of an individual
particle.

In similar fashion, measurements taken at different positions of
the sky will differ in intensity from variations in the number of
unresolved sources in the individual measurements. It is not expected
that the sources, unlike electrons, will have the same "charge" (flux).
Even so, models can be made of the expected distribution of sky flux
measurements based on particular forms for the distribution of source
numbers versus their flux. In comparing these models to the observed
distribution of sky flux we will see that at high fluxes (just a decade
below the resolved source survey limits) the behavior of unresolved
sources is consistent with the expectations based on the study of
resolved sources. At lower fluxes, though, the expected behavior of the
source counts is insufficient to provide the observed fluctuations.

Evolution in the source luminosity function could provide the missing fluctuations. Besides requiring an additional source of fluctuations, the data also has an upper limit to the fluctuations. The upper limit when applied to the fluctuations produced by source evolution can constrain the degree of the evolution. (Reviews of analyses of the anisotropies in the X-ray sky and their impact on sources of the sky flux have been given by Schwartz [1980], Rees [1981], and Fabian [1981]. These reviews do not include results from HEAO 1.)

## Dissertation outline, organization, and notation

The next chapter, Chapter II, is a general description of the A-2 experiment on HEAO 1. It considers in detail the design properties of the detector that were used to study and understand the unresolved sky flux as well as to continuously monitor the internal, non-X-ray, background. A broad review of the spectral sensitivity of the experiment, its history of operation in orbit, and a general database are also given. Chapter III presents a detailed derivation of how unresolved sources produce a distribution of sky flux measurements. Issues (normally ignored) are developed to understand the strengths and limitations of the techniques. Chapter IV describes the creation and organization of the database subsets used to study the fluctuations and other anisotropies. It also presents the results of fitting the HEAO 1 data for variations not explicitly modelled from the unresolved sources, including emission associated with our galaxy. A possible measurement of the Compton-Getting effect is also discussed. The stability of the internal background is measured. Chapter V is a quick overview of the process of fitting models to the fluctuations. Statistical problems and other issues are addressed. In Chapter VI we fit the models described in Chapter III to our data. The results of these fits are compared to expectations from resolved sources in Chapter VII. The upper bound for other sources of variation is used to place a general restriction on any "new" component to the sky flux. Chapter VII also places the results of the fluctuations in the context of recent results in other bandpasses, including results from the Einstein X-ray Observatory. The possibility of significant limits on large scale structure in the universe based on

the fluctuations measurements are sketched. Chapter VIII investigates current models for the origin of the X-ray sky flux, and outlines how future studies of the fluctuations could yield interesting results. The results and conclusions are summarized in Chapter IX. A series of appendices present issues and information of a technical nature.

Each chapter, appendix, and section is headed by a brief synopsis of its contents, which is repeated in the Prècis. Sections, equations, figures and tables are numbered by chapter (roman numerals) or appendix ("A" followed by an arabic numeral). In the discussion of cosmological issues, Hubble's constant, $H_o$, is assumed to be 50 km s$^{-1}$ Mpc$^{-1}$. Some relationships will indicate their dependence on the value of $H_o$ through the dimensionless $h_{50} \equiv H_o/50$. In the figures and most of the following discussion, all flux measurements will be in detector dependent units. Unless specifically labeled otherwise, units of "counts s$^{-1}$ cm$^{-2}$" are measurements made with the first layer of the HED 1 counter from the A-2 experiment of HEAO 1. The label "H1L1" also refers to measurements in these units. Converting to absolute units depends on the incident spectrum, but roughly 1 H1L1 $\approx$ 1.3x10$^{-11}$ ergs s$^{-1}$ cm$^{-2}$ integrated over the range from 2 to 10 keV.

Power law energy spectra are defined so that a positive index, $\alpha > 0$, corresponds to a decreasing flux with energy, $(s(\nu) \propto \nu^{-\alpha})$. Power law models for the source counts $(N(S) \propto S^{-\gamma})$ and luminosity function $(\phi(L) \propto L^{-\beta})$ are similarly defined.

CHAPTER II

THE HEAO 1 COSMIC X-RAY EXPERIMENT (A-2)

In this chapter we briefly review the physical and organizational aspects of the A-2 experiment, particularly those relevant to studies of structure in the X-ray sky.

The HEAO 1 satellite marked a watershed in X-ray astronomy over the energy range from below 0.1 keV to several hundred keV. With its large total area, and complement of specialized detectors, it perfomed the culmination of the X-ray all-sky surveys of the seventies, work begun by the UHURU and Ariel V satellites. Except at the lower energy range ($\lesssim$ 3 keV), HEAO 1 will remain the best source for all-sky X-ray studies for the foreseeable future. The A-2 experiment, one of four on HEAO 1, had several design features for studies of the X-ray sky. We first outline the physical organization of the experiment as well as a general technique for estimating the detectors' internal background and making background-free measurements of the unresolved sky flux. This technique is implemented using a system of different solid angle collimators. The spectral response is then considered with a particular eye to characterizing the broad energy band measurements subsequently used. We review the in-orbit history of the experiment. Finally the preliminary data preparation techniques are described, including the definition of data quality conditions used in the following analysis.

---

1: The HEAO 1 A-2 Experiment

Characteristics for the study of the X-ray sky

One of the primary goals in the design of the A-2 experiment was to study and characterize the unresolved X-ray sky flux over a broad energy band, from 0.1 keV to about 60 keV. In this section we give a general description of the experiment and elaborate on those aspects that made this experiment well suited for this project. A more detailed

description is given in Rothschild et al. [1979].

The A-2 experiment consisted of three types of gas proportional counters. There were six detectors in all: two propane-filled thin window flow counters, designated LEDs for low energy detectors; a single argon-filled sealed counter, called the MED (medium energy detector); and three sealed xenon counters, the HED 1, 2, and 3 (high energy

---

Figure II.1: Configuration of the A-2 experiment on HEAO 1

The six detectors are shown along with the three spacecraft axes.

---
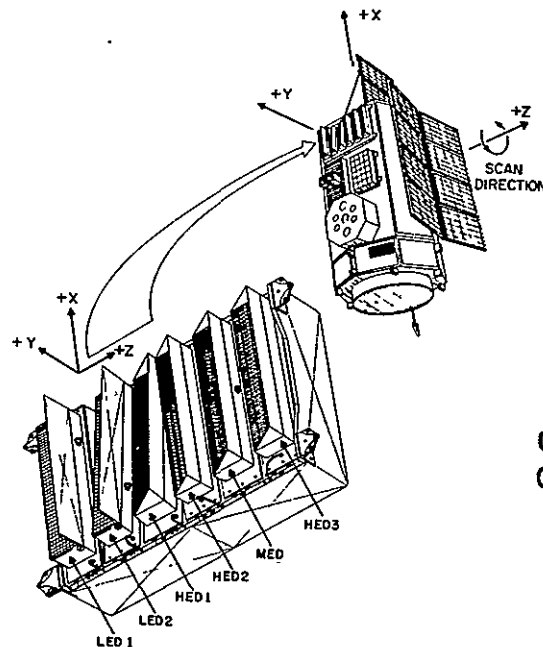


ORIGINAL PAGE IS
OF POOR QUALITY

---

detectors). The different active gases allow the three kinds of detectors to be sensitive over different energy bands. The LEDs cover the range from 0.15-3 keV. The analysis of their data was handled in

large part by the West Coast arm of the A-2 consortium: groups at the California Institute of Technology, the Jet Propulsion Laboratories, and the University of California at Berkeley. In this low energy band, the X-ray sky is dominated by an anisotropic component of a local, i.e. galactic, nature. Because of this we will not consider these data here. The MED covers the range from 1.2 to 20 keV, while the HEDs cover the 2.5-60 keV range.

The detectors were oriented generally along the spacecraft +Y axis, (see Figure II.1). HED1 was offset by 6 degrees in the +X direction. The solar cells were pointed along the +Z axis, and spacecraft operations were constrained so that this axis was oriented towards the sun. The spacecraft operated in two modes, pointing and scanning. When the spacecraft was scanning it rotated clockwise about the sun-pointed +Z axis with a 33 minute nominal scan period, i.e. at a rate of about 1 degree every 5.5 seconds. The experiment thus surveyed a great circle scan that followed a line of constant ecliptic longitude, passing through the north and south ecliptic poles with each scan. In order to maintain the proper orientation to the sun the spin axis was stepped by 0.5 degrees roughly every 12 hours. At the end of six months of scanning operations the spin axis was antiparallel to its original position. During that time the entire sky was surveyed, although the parts of the sky at the ecliptic equator had less total exposure than areas at high ecliptic latitudes.

### The counters

Proportional counters detect incident X-rays through the photoelectric effect. An incident X-ray photoionizes an atom of the active gas. The ejected electron has an energy equal to that of the incident X-ray less the ionization potential. A high voltage field accelerates the electron so that it in turn ionizes further atoms, producing additional electrons all of which are collected on an anode wire and detected as a pulse. The number of ion-electron pairs created, and thus the total charge in the pulse, is roughly proportional to the energy of the ionizing photon. This determination of photon energy is

not exact, at the very least because of Poisson statistical processes in the creation of the secondary electron-ion pairs. An additional complication is that occasionally the initial X-ray interaction event will produce a secondary fluorescence X-ray as well as the primary photo-electron. This secondary X-ray may pass out of the detector so that the collected pulse for this event will have its corresponding energy reduced by the energy lost through the fluorescent X-ray. The complicated spectral responses of the detectors are displayed with the discussion in Section II:3.

The detectors are all multi-layered, having at least two layers of active gas for counting purposes. These are surrounded on three sides by additional active gas layers wired for use in anticoincidence logic. A penetrating charged particle, unlike an X-ray, tends to deposit its energy along its entire path, being detected by several wires and thus

---

Figure II.2: Cross-sectional view of MED and HED 2

---



ORIGINAL PAGE IS
OF POOR QUALITY

---
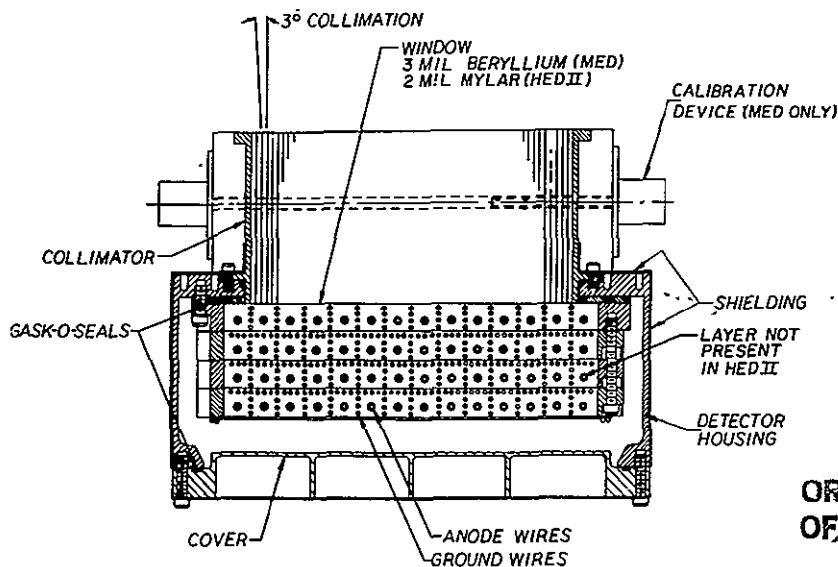
easily rejected. A major source of background not measured by this arrangement is soft non-penetrating electrons entering the detector through the window. These could deposit all their energy in a single

Figure II.3: Cross-sectional view of HED 1 and 3
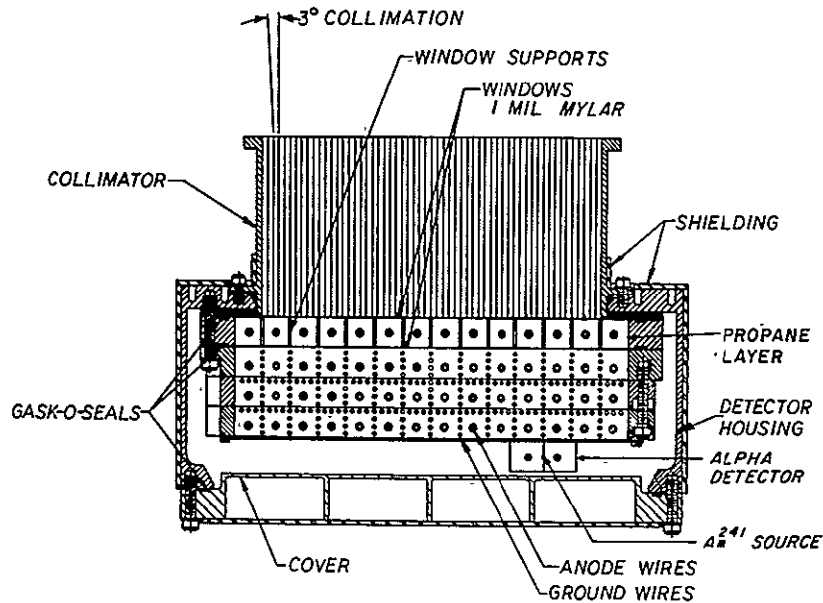


layer and would thus not be caught by the standard anticoincidence arrangement. A layer of propane gas was placed behind the collimator and ahead of the xenon active gas in the HED 1 and 3 detectors (see Figure II.3). Propane is transparent to most of the X-rays that the HED's xenon is sensitive to, but will either absorb the incident electrons or remove via anticoincidence those that pass through it to the first xenon layer below. Although neither HED 2 or the MED have such a layer, the second layer of active gas benefits from a similar anticoincidence screening effect by the first layer. As the second layer of the MED had substantial response over most of the MED energy range, this first layer-second layer scheme was used instead of a front layer propane anticoincidence, avoiding for at least one detector any possible complications of the additional propane gas system.

## Internal-background-free sky flux measurements

The multiwire aspect of the detectors was exploited in a novel fashion to allow a direct measurement of the diffuse sky flux with a

---

Figure II.4: Detector grid connections for HED 1 and 3

The L1, L2, and R1, R2 grid are the "right" and "left" rates for the first and second layer, respectively. V1 is the general outer cell anti-coincidence veto rate, while V2 is (for these detectors) the rate from the propane layer used for electron rejection. The "alpha" wire is used in calibration. A similar arangement is used in the other detectors.

---

**COLLIMATOR**
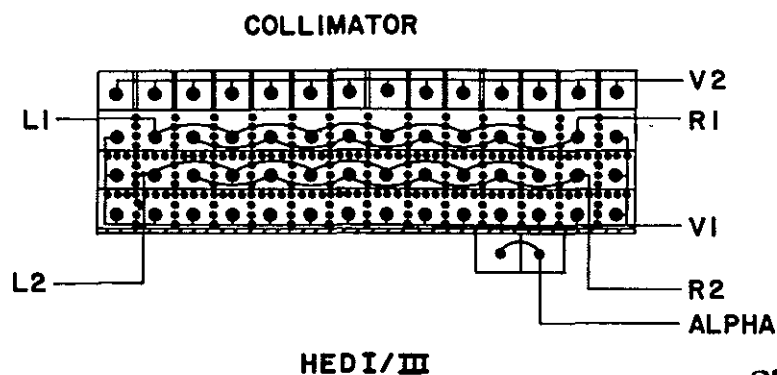


HED I/III

## DETECTOR GRID CONNECTIONS

---

simultaneous estimate of the internal background rate. As illustrated in Figure II.4, alternating cells of a layer are combined into two different rates, e.g. R1 and L1 for layer one, referred to as the "right" and "left" layer one rates. As they share the same active gas, have essentially identical geometrical relationship to the spacecraft, and pass through a common amplifier (so they have identical gain), both have a near identical internal background contribution, B. The value of this condition is utilized by a clever design so that the unresolved sky count rate is not identical in the two "sides". Let us label properties of the separate sides by "L" and "S". (The significance of such letters for the A-2 experiment will become apparent.)

For instance if the X-ray sky rate is denoted by $X$ $s^{-1}$ $cm^{-2}$ $sr^{-1}$,

we can let the two "sides" have differing efficiencies, $\alpha_L$ and $\alpha_S$. Assume that $\alpha_L > \alpha_S$. Then the two rates would be given by:

$$L = \alpha_L X + B ;$$                                                    [II.1a]

$$S = \alpha_S X + B .$$                                                    [II.1b]

This can be solved so that

$$X = \frac{L - S}{\alpha_L - \alpha_S}.$$                                  [II.2]

Note that this mean rate has no contribution from the balanced internal background rate. Therefore the background rate can vary enormously during an integration, but as long as it is the same in both sides the mean rate found in equation [II.2] will remain unaffected.

This advantage is not gained without some price, notably increased imprecision due to counting statistics. If the two rates, L and S, are statistically independent and have variances $\sigma_L^2$ and $\sigma_S^2$, respectively, then the variance of their difference is the sum of their variances, and

$$\sigma_X^2 = \frac{\sigma_L^2 + \sigma_S^2}{(\alpha_L - \alpha_S)^2} .$$    [II.3]

Thus the determination of X in [II.2] is less contaminated by uncertainties in B compared to what could be achieved with a single rate but at the cost of reducing the signal to noise ratio. If the variance in L and S is purely Poisson,

$$(X/\sigma_X)_{L-S} \propto \frac{L - S}{(L + S)^{1/2}} \propto \frac{1 - \alpha_S/\alpha_L}{(1 + \alpha_S/\alpha_L + 2B/(X\alpha_L))^{1/2}} .$$    [II.4]

For fixed values of X, B, and $\alpha_L$ the signal to noise is maximized by setting $\alpha_S$ to zero. By doing this, though, half of the detector area would be effectively dedicated to monitoring the internal background.

Simultaneous with the determination of X, we can solve [II.1] for

B :

$$B = \frac{S\alpha_L - L\alpha_S}{\alpha_L - \alpha_S} \text{, with}$$ [II.5a]

$$\sigma_B^2 = \frac{\alpha_L^2 \, \sigma_S^2 + \alpha_S^2 \, \sigma_L^2}{(\alpha_L - \alpha_S)^2} \, .$$ [II.5b]

This is useful for monitoring changes in the internal background during the lifetime of the experiment under a variety of external conditions. If with such studies it can be shown that the rate B is constant, or at least predictable from other observables, then X may be estimated directly from one of the two rates, L or S, or for even better signal to noise, from their sum: L + S. That is,

$$X_{L+S} = \frac{L+S - 2B}{\alpha_L + \alpha_S} \, ,$$ [II.6]

where the subscript L+S indicates that $X_{L+S}$ is an estimate of X based on the summed rates. The signal to noise for the summed rate, ignoring any imprecisions introduced by uncertainties in the rate B, is found by replacing the minus sign in the numerator of [II.4] with a plus sign. The ratio of the signal to noise for the summed rate versus that for the difference rate is

$$\frac{(X/\sigma_X)_{L+S}}{(X/\sigma_X)_{L-S}} = \frac{\alpha_L + \alpha_S}{\alpha_L - \alpha_S} \, .$$ [II.7]

Several techniques may be used to create unequal values for $\alpha_L$ and $\alpha_S$. One pre-HEAO 1 approach was used on an argon detector, the "B" detector, that was part of the GSFC Cosmic X-ray Spectroscopy Experiment aboard OSO 8, the 8th Orbiting Solar Observatory. The B detector had a similar arrangement of alternating cells, with different window materials over the two groups of cells. These materials had different amounts of X-ray transmission at low energies, so equation [II.2] could be used to determine unambiguously the X-ray flux in the 1.2-9.0 keV range. At higher energies the X-ray transmission of the two window

The A-2 Experiment

Experiment Characteristics

materials was essentially identical. The entire experiment is described in more detail by Pravdo [1976].

On the HEAO 1 A-2 experiment the technique was extended to higher energies by providing the two groups of cells not with different windows but with exposure to different amounts of solid angle. The two groups are refered to as the Large Field Of View, or LFOV, and Small Field Of View, SFOV. If the diffuse flux were totally uniform, it is obvious that the approach of [II.2] and following equations could be used. We will see in the next section that even in the presence of spatial non-uniformity, equation [II.2] provides an unambiguous measurement of X-ray flux without a contribution from internal background in a systematics-reduced way. Equation [II.5] is not correct on an instantaneous basis, but will provide, on average, a good measurement of the internal background.

The crux of either the OSO 8 or A-2 approach is the presumption that B, the internal background, is the same for both sides. In particular there can be only a negligible difference in the non-X-ray count rate due to the differing structures, windows or collimators, over each set of cells. Cavallo, Horstman, and Moretti-Horstman [1977] have pointed out that, without this presumption or an a priori estimate of the difference, it is formally impossible to separate measurements of the diffuse sky from the internal background. Our confidence in the applicability of this technique rests on our understanding that the impact on the internal background due to the structure of the collimator is small. This understanding is bolstered by the agreement for diffuse measurements of separate detectors (HED 1 and 3). The warnings of Cavallo, Horstman and Moretti-Horstman are more important for energy regimes where the internal background dominates the total rate, particularly the high X- and γ- ray portions of the spectrum, where the γ-rays from cosmic ray conversions in the collimator may be a significant source of the total non-X-ray counts.

———————————————

## 2: Spatial Response

The spatial response of the detectors is described, including the effects of scanning. When different "sides" have different response functions, their difference gives an internal-background-free measurement of X-ray intensity. An alternate combination allows a measurement of the internal background. We find some simple characterizations of the response for general and ideal detectors.

Start with a point source of flux $S_{ps}$. The number of counts per exposure observed in a detector due to this source, $I_{ps}$, depends on the position of the source in the detector's field of view:

$$I_{ps}(\theta,\phi) = r(\theta,\phi) \, S_{ps} \qquad\qquad [II.8]$$

where $r(\theta,\phi)$ is the response function at a position $\theta$ and $\phi$. These are angles in a detector-defined coordinate system. The angular dependence of $r$ is a measure of the increasing shadowing of the detector window by the collimator and an additional cosine term due to the inclination of the detector face. This last term is negligible in detectors with a small or moderate sized field of view. The A-2 detectors are collimated by a slat-type system consisting of hundreds of rectangular cross section tubes. In this case the response function will be proportional to

$$r(\theta,\phi) \propto \left( 1 - \frac{\tan \theta}{\tan \theta_0} \right) \left( 1 - \frac{\tan \phi}{\tan \phi_0} \right) \qquad\qquad [II.9]$$

which in the limit of small angles reduces to $r(\theta,\phi) \propto (1-\theta/\theta_0)(1-\phi/\phi_0)$. The critical angles $\theta_0$ and $\phi_0$ represent the half width where the response goes to zero, which is equivalent at small angles to the Full Width Half Maximum, FWHM, normally quoted to describe spatial response. The small-angle response function has a triangular cross section. The two dimensional form of the response is sometimes referred to as a "pyramidal" response, although an examination of a contour plot of the response shows that this is an evocative rather than a mathematically precise description.

We will assume throughout that $r(\theta,\phi)$ is independent of the energy of the incident X-rays. This is a good approximation for most photons

Figure II.5: HED 1 LFOV r($\theta$,$\phi$=0) versus $\theta$

The cross section of the LFOV collimator's instantaneous response versus angle in the scan plane, where $\phi$=0.
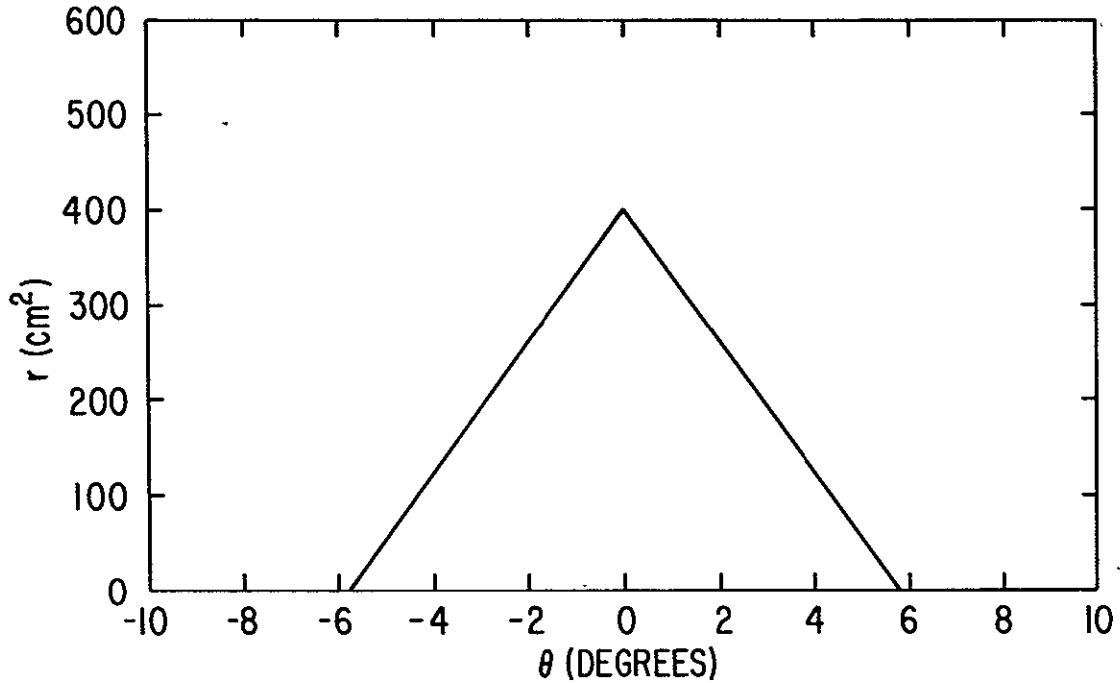


in studies of the diffuse background. However for the highest energy photons, the collimator tubes begin to become transparent. The collimator is thus less efficient at restricting incident photons. This causes bands of increased response over angles out to 2 $\theta_o$, 3 $\theta_o$ etc. Each outer band is dimmed by the absorption by additional collimator tubes. At energies above 50 keV, although the attenuation is still large, the solid angle covered by these additional bands is a significant contributor to measurements of the sky surface brightness. Fortunately the vast majority of photons analyzed in this dissertation have energies less than 30 keV.

The absolute value of r in equation [II.9] depends on the units in which S is measured. Flux is often measured in ergs $s^{-1}$ $cm^{-2}$, but

The A-2 Experiment                                    Spatial Response

Figure II.6: HED 1 LFOV two-dimensional response, $r(\theta,\phi)$

Instantaneous response. Contours show lines of constant
r, contour intervals of $50 \ cm^2$.
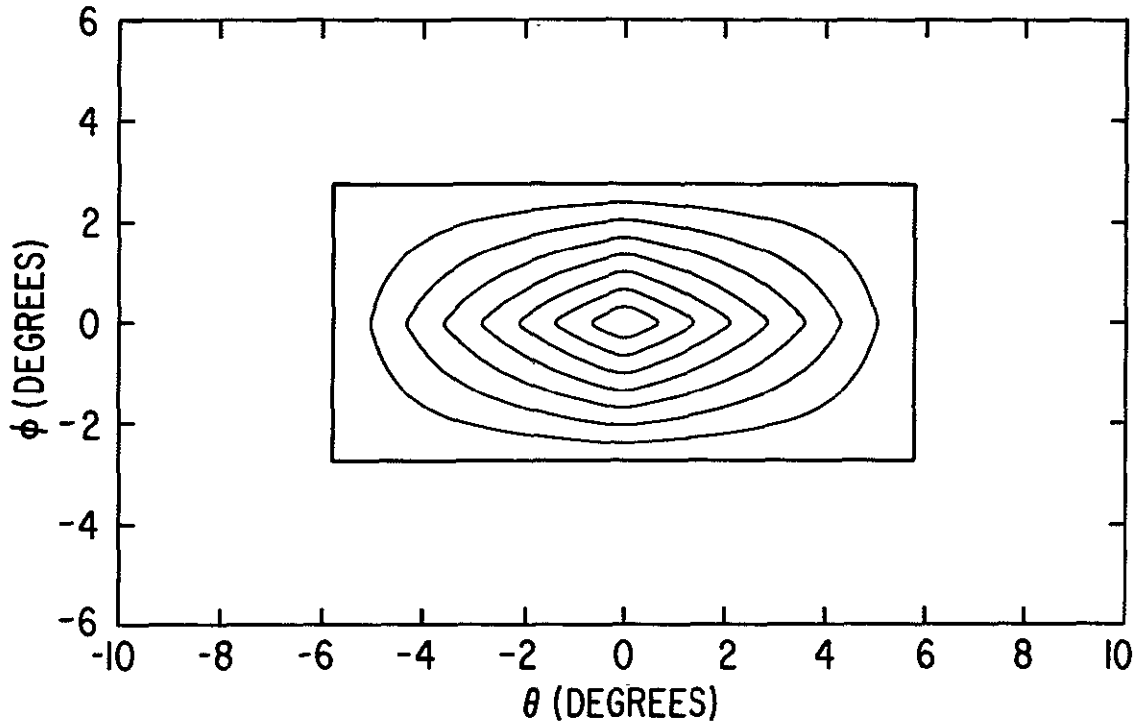


which S is measured. Flux is often measured in ergs $s^{-1} \ cm^{-2}$, but predicting how many counts in an exposure this produces in the detector depends on the source spectrum as well as the detector's spectral efficiency. To avoid this ambiguity we will measure S in terms of the number of counts $s^{-1}$ the source would produce on-axis in a detector of identical construction of area one $cm^2$; that is, the units of S are [S] = counts $s^{-1} \ cm^{-2}$. To relate such a unit to astrophysically interesting quantities (such as flux measured in ergs $s^{-1} \ cm^{-2}$) still requires the spectrally dependent conversion, as we will discuss in a later section. With this choice of units for S, the units for $r(\theta,\phi)$ are [r] = $cm^2$ s exposure$^{-1}$. The peak on-axis response for an exposure of length $\Delta t$ with a detector of area A is just $A \ \Delta t$. The "instantaneous response" of a detector is given by $\lim_{\Delta t \to 0} r/\Delta t$, the units of which are $cm^2$.

The A-2 Experiment                                                    Spatial Response

The detectors are oriented with respect to the spacecraft axes so that $\theta$ and $\phi$ are respectively parallel and perpendicular to the scan direction. In all the detectors, $\phi_O$, the size of the collimator perpendicular to the scan, is a nominal 3 degrees so that in a FWHM sense each scan samples a 3 degree wide strip centered on a great circle. The value of $\theta_O$ varies from detector to detector, and from "side" to "side", ranging from a nominal 12 to 1.5 degrees. $\theta_O$ for the LFOV side of a detector is roughly twice that of the detector's SFOV.

### Measurements with combinations of FOVs

As discussed in the previous section the data from the two FOVs may be combined in a variety of ways. In an extension to equations [II.1] et seq. the counts per exposure observed in the large field of view are

$$I_L = B + \int d\Omega \ r_L(\theta,\phi) \ \Sigma(\theta,\phi) \qquad [\text{II}.10]$$

where $r_L$ is the LFOV response, $\Sigma(\theta,\phi)$ is the X-ray surface brightness in counts $s^{-1}$ $cm^{-2}$ $sr^{-1}$ and B is the internal background contributed, in counts exposure$^{-1}$. A point source of flux $S_p$ at a position $\theta_p,\phi_p$ can be said to give a contribution to $\Sigma$ of $S_p \ \delta(\theta-\theta_p) \ \delta(\phi-\phi_p)$. A similar relationship holds for the SFOV replacing $r_L$ with $r_S$. The difference, $I_{L-S}=I_L-I_S$, is equivalent to an internal-background-free measurement using a detector with a spatial response $r_{L-S}$:

$$I_{L-S} = \int d\Omega \ r_{L-S}(\theta,\phi) \ \Sigma(\theta,\phi) \ , \qquad [\text{II}.11]$$

where $r_{L-S}(\theta,\phi) = r_L(\theta,\phi) - r_S(\theta,\phi)$.

We can define $\alpha_L = \int d\Omega \ r_L(\theta,\phi)$, and a similar $\alpha_S$. The $\alpha$s are the detectors' geometry factors times the exposure time. These are equivalent to what is often called a detector's solid angle in the sense that for a uniform diffuse surface brightness $\Sigma$, the contribution $\Sigma$ makes to the count rate, $I_{sb}$, would be $\alpha\Sigma$. For the instantaneous detector the $\alpha$s are identical to the geometry factors, and the units of $\alpha$ are just $cm^2$ sr.

Following the practice outlined in Equation [II.5] for measuring the internal background, B, the combination of $I_L$ and $I_S$ to use is

$$I_B = \frac{\alpha_L I_S - \alpha_S I_L}{\alpha_L - \alpha_S} = B + \int d\Omega \; r_B \Sigma \; . \; _ . \qquad [II.12]$$

This is similar to equation [II.5], except in that case there was no contribution to the estimate of B from the X-ray sky, while here $I_B$ includes the effect from a detector with response

$$r_B(\theta,\phi) = \frac{\alpha_L r_S - \alpha_S r_L}{\alpha_L - \alpha_S} \; . \qquad [II.13]$$

By the definition of $I_B$, $\alpha_B = \int d\Omega \; r_B$ is zero so that for uniform surface brightness the last term of [II.12] vanishes. Even when the sky is non-uniform the term will be on average zero, but the sky will contribute an additional variance to any particular measurement of B.

---

Table II.1: Detector collimators and areas

We here give the FWHM of the collimator response perpendicular to the scan direction, the FWHM in the scan direction for the two fields of view, the on-axis open areas, and the effective-area solid-angle product for the instantaneous response (geometry factor). Numbers are based on post-launch measurements of the Crab Nebula. The HED 1 data came from scans past the Crab while for the other detectors the data were from an early "point" operation at the Crab.

| Detector | $\phi_o$ | $\theta_{oL}$ | $\theta_{oS}$ | $A_L$ | $A_S$ | $\alpha_L$ | $\alpha_S$ | $\alpha_{L+S}$ | $\alpha_{L-S}$ |
|---|---|---|---|---|---|---|---|---|---|
| | | degrees | | cm$^2$ | | | cm$^2$ sr | | |
| HED 1 | 2.74 | 5.78 | 2.85 | 400 | 392 | 1.94 | 0.94 | 2.87 | 1.00 |
| HED 2 | 2.84 | 5.92 | 2.91 | 451 | 443 | 2.32 | 1.11 | 3.43 | 1.20 |
| MED | 2.90 | 2.94 | 1.40 | 440 | 402 | 1.14 | 0.50 | 1.64 | 0.65 |
| HED 3 | 2.79 | 2.91 | 1.47 | 429 | 402 | 1.06 | 0.50 | 1.57 | 0.56 |

---

The combination of different fields of view produces detector response functions of a shape unusual in X-ray astronomy experiments.

Figure II.7: HED 1: FOV combinations response cross-section

(a) $r_{L+S}(\theta,0)$     (b) $r_{L-S}(\theta,0)$
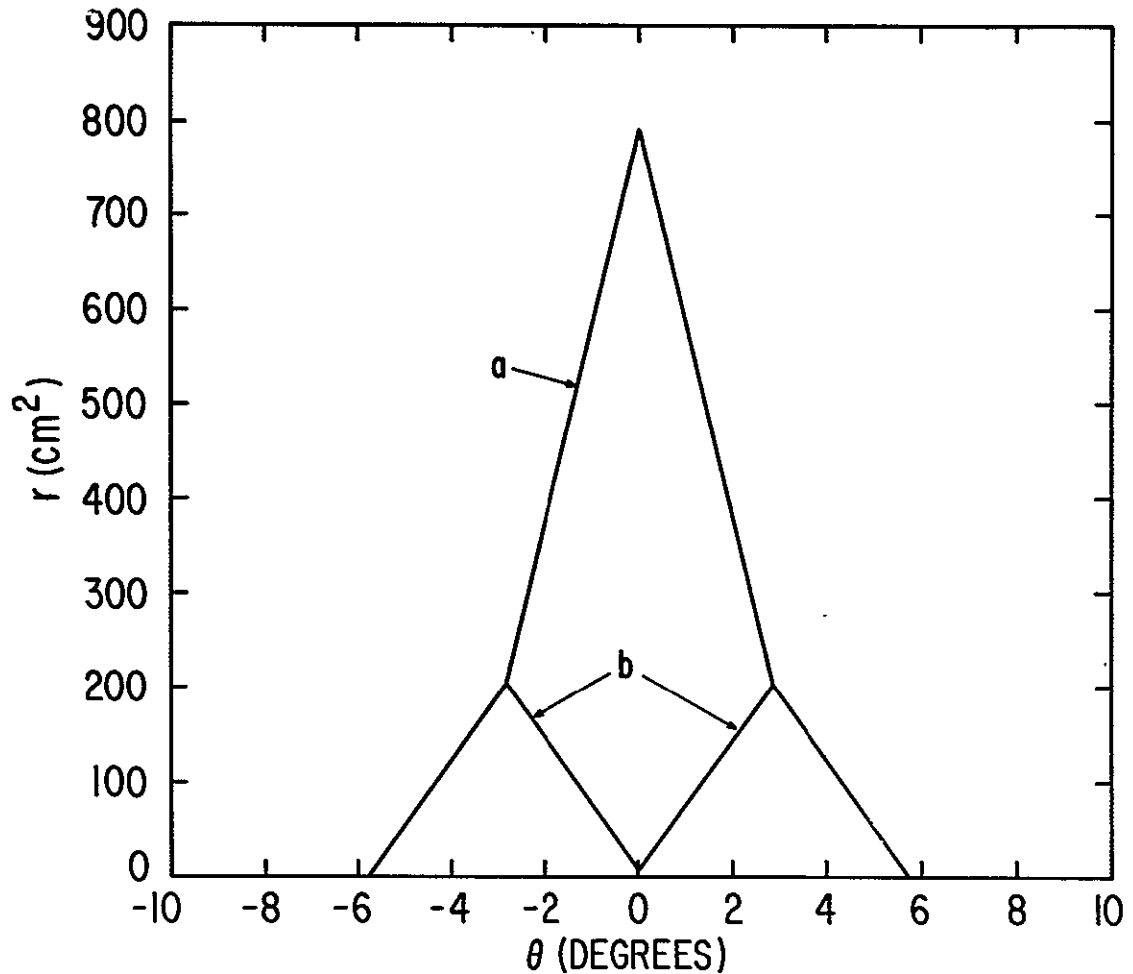


In particular for the L-S response, rather than a single "pyramidal" shape, the response is two abutting "pyramids" with a peak response roughly half of the peak response of a single side. For the study of a point source, the L-SFOV would provide at best only a quarter as many counts as would be possible with the L+SFOV with the same absolute value of the uncertainty from counting statistics. A close examination of the plot of $r_{L-S}$ in Figure II.7 shows that the on-axis response, $r(0,0)$, is not exactly zero. This is because the on-axis open areas of the two sides are not identical, in part because the SFOV required more

The A-2 Experiment                                                    Spatial Response

Figure II.8: HED 1: FOV combinations, $r(\theta,\phi)$

Instantaneous response.     Contours at equal r,     contour
interval of 50 cm$^2$.   (a) $r_{L+S}(\theta,\phi)$.  (b) $r_{L-S}(\theta,\phi)$
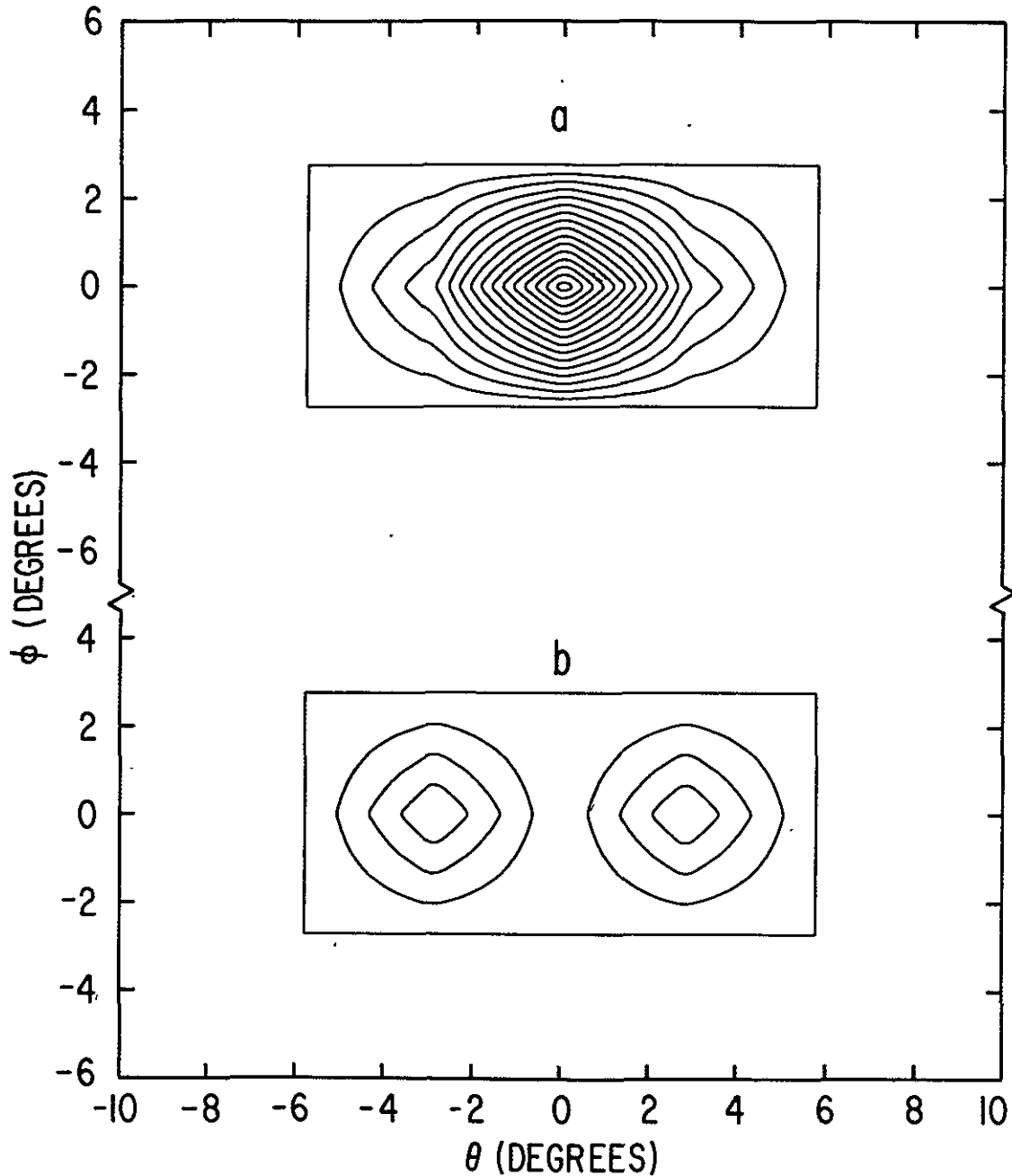


collimation tubes to produce the smaller field of view, covering more of

the detector face. Part of the motivation for measuring flux, S, in terms of counts per unit area is that, by dividing out the area, we can with greater ease compare flux measurements in the various FOV combinations. Measurements made per unit area also smooth comparisons between HED 1 and HED 3, which while essentially identical in spectral response, have different areas.

The A-2 experiment's implementation of the equal-internal-background/unequal-sky-intensity technique uses the different solid angles of the L and S rates. This has an advantage in that a source positioned on-axis can be measured using the total detector area, $r_{L+S}(0,0) \approx 2\ r_L(0,0)$. In contrast the implementation for the OSO 8 B detector, using the different X-ray absorption properties of the two windows, was achieved at the cost of a permanent decrease in the effective area by the absorbing window material.

### Smeared response functions

So far we have described the instantaneous detector response. But the exposures were of finite length, during which the detector scanned along the sky. This tended to "smear" the collimator response, since a given position in the sky sampled a range of the collimator response function as the collimator moved past it. The smeared response from an integration of length $\Delta t$ is

$$r_{\Delta t}(\theta,\phi) = \int_{-\Delta t/2}^{+\Delta t/2} dt\ r_0(\theta + \omega t, \phi),$$                [II.14]

where $r_0$ stands for the instantaneous response and $\omega$ is the scan rate. This integration over $\Delta t$ makes the units of $r_{\Delta t}$ cm$^2$ s, versus cm$^2$ for the instantaneous response. In analyzing the fluctuations we used data taken with 1.28 second exposures during which the detector scanned ~0.22 degrees. This is small compared to the scale of the HED 1 detector but is more significant for the HED 3. The principal effect is to multiply the response by $\Delta t$, but in addition there is a wing added to the edges of the response and the peak is slightly rounded. In the L-SFOV combination the trough between the peaks is somewhat filled in.

Even with the large area of the A-2 detectors, their relatively small solid angle means that a single 1.28 second exposure has a relatively high uncertainty in a sky flux measurement due to counting statistics, from 26%, one sigma, for the HED 1 first layer L+S rate to 110% for HED 3 layer 1 L-S. To make a reasonable measurement we combine many such exposures. By superposing many scans, over periods up to 2 days, and taking the exposures that contribute to a section of the superposed scan plane, $\Delta\theta$, we greatly reduce the effect of counting statistics. This combining of data for a single estimate of the X-ray flux is equivalent to a single exposure with an "average" response produced by additional smearing. The smearing in the scan direction is simply:

$$r_{\Delta\theta}(\theta,\phi) = (\Delta\theta)^{-1} \int_{-\Delta\theta/2}^{+\Delta\theta/2} d\theta' \, r_{\Delta t}(\theta+\theta',\phi) \qquad [II.15]$$

where the $(\Delta\theta)^{-1}$ represents the fact that we are averaging over $\Delta\theta$. This smearing can have a significant effect on the response function. Although the integral of the response over all solid angles is not affected, the peak response is greatly reduced. A point source at a given position in the sky therefore contributes on average fewer counts to I than it would if the detector were not smeared. If $\Delta\theta$ is large enough, $>\theta_{FWHM}$, then the trough for $r_{L-S}$ can be completely filled in, and the response function looks fairly flat in its $\theta$ dependence (compare Figure II.9 with the unsmeared response in Figure II.7).

Because the data are combined over a period of days, during which the nominal spin axis moves, there is an additional smearing in the $\phi$ direction. In principle the amount of smear varies depending on the location of the measurement in ecliptic coordinates. Those measurements near the pole suffer little additional smearing. Those at the equator are affected the most. But because of wobble in the actual spin axis position, the smear is at minimum $\Delta\phi \sim 0.5$ degrees. The difference between the two extremes is then generally not a large effect on the average response. Unlike the $\Delta t$ smearing, which is formally correct unless a source varies while the detector scans over its position, the angular smearing is an approximation that assumes that the individual

Figure II.9: HED 1 smeared response cross section (a) $r_{L+S}$ and (b) $r_{L-S}$ versus $\theta$

$\Delta t$ = 1.28 s, scan period = 35 minutes, $\Delta\theta$ = 6 degrees, $\Delta\phi$ = 1.5 degrees.
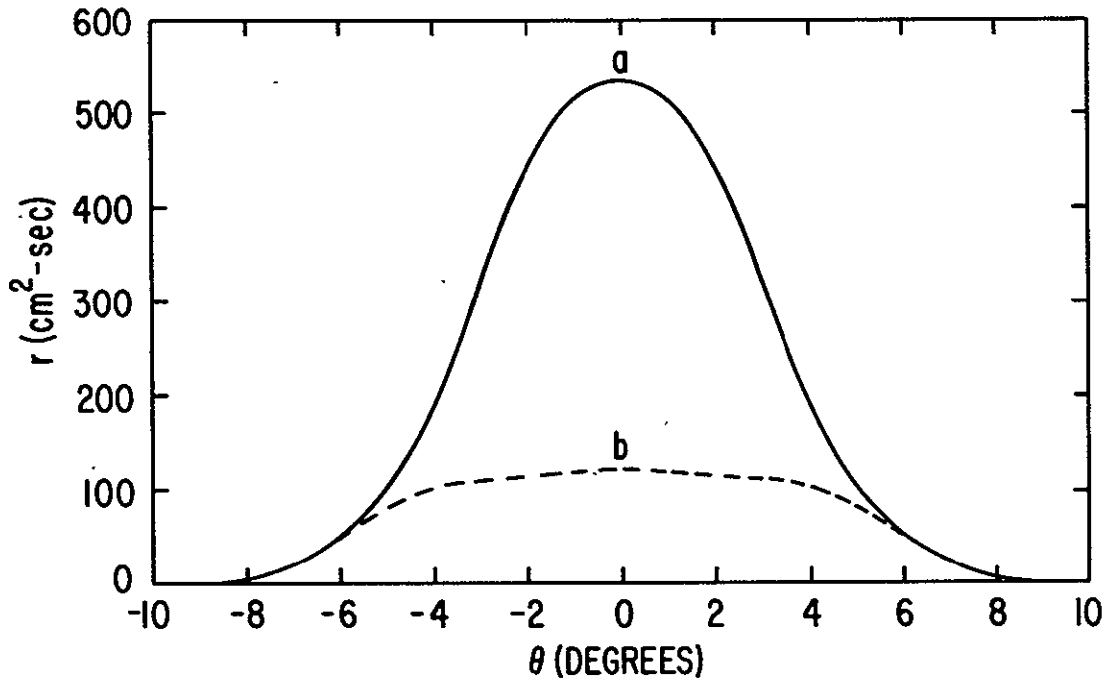


exposures are uniformly distributed throughout the smear area $\Delta\theta \times \Delta\phi$.

## Characterization of the detectors' angular size

When the detector response is smeared as much as we use in some of the fluctuations analysis, it is inappropriate to characterize the response function in terms of FWHM. The correct, but unwieldly representation of the response would be to plot the two dimensional form of $r$, as in Figure II.10. However it is difficult to integrate by eye such a picture to see from what area most of the counts come. In this sense the FWHMxFWHM of an unsmeared rectangularly collimated detector is a box which contains 56% of the diffuse sky counts received by the instantaneous detector.

Figure II.10: HED 1 smeared two dimensional response, $r(\theta,\phi)$

$\Delta t=1.28$ s,    scan period = 35 minutes,    $\Delta\theta$ = 6 degrees, $\Delta\phi$ = 1.5 degrees. Contours at 50 cm$^2$ s.    (a) L+SFOV; (b) L-SFOV.
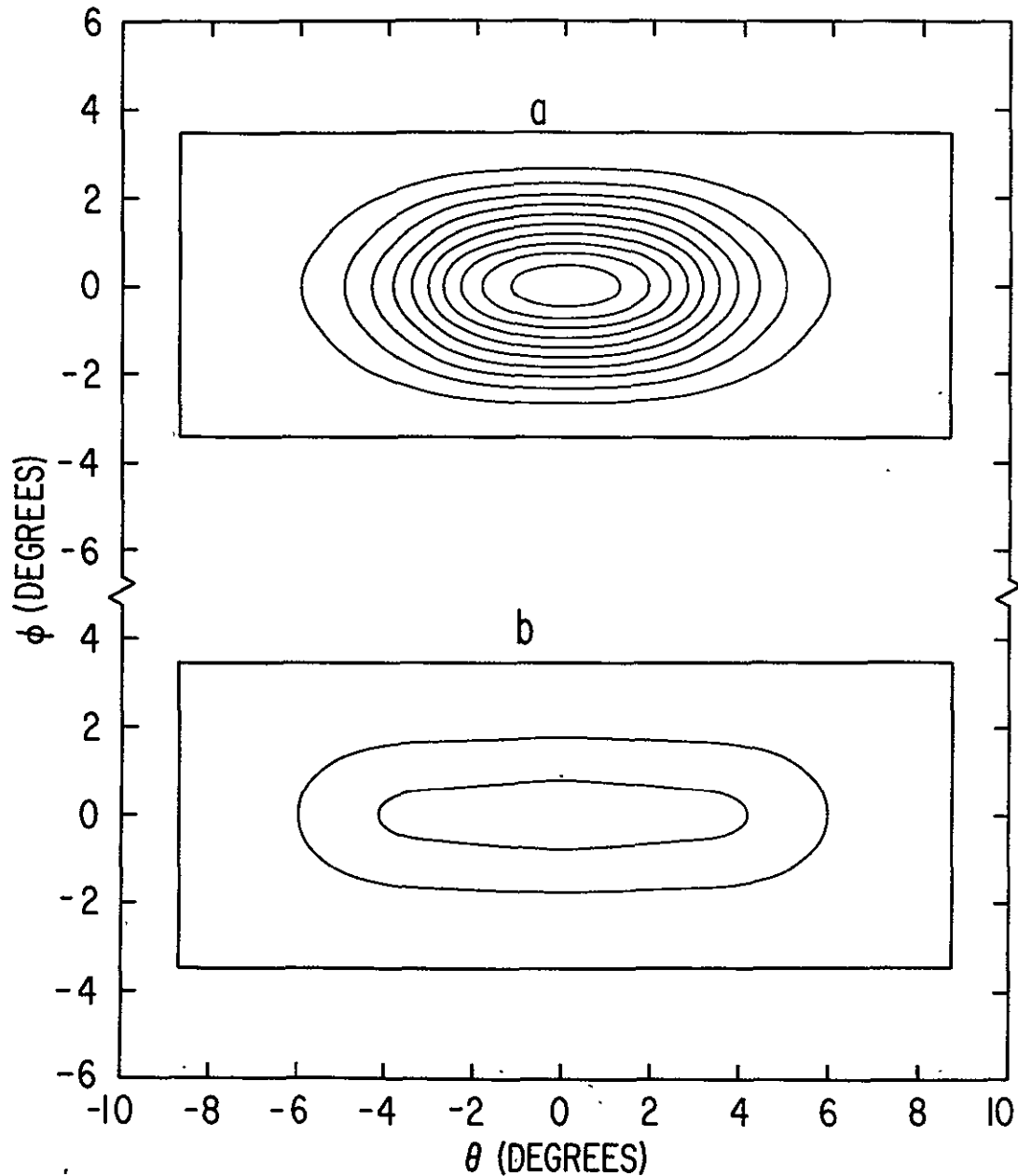


By plotting the contours that show what percentage of the total
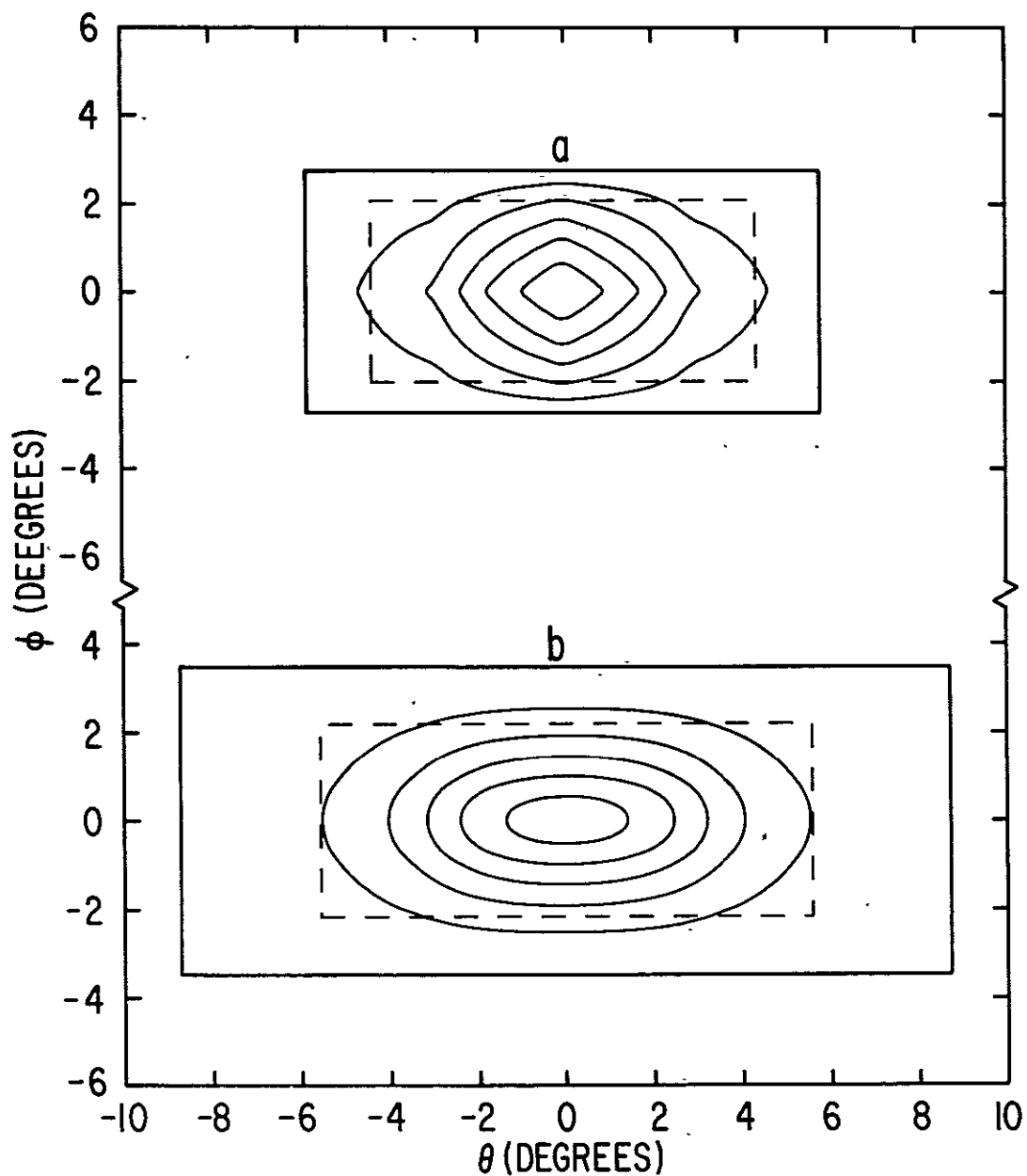
Figure II.11: HED 1 L+S $r(\theta,\phi)$, outline of fractional contribution areas

The solid contours are lines of constant r. From the outside in they contain 100, 90, 70, 50, 30, and 10% of the total count rate. The dashed rectangle includes 90% of the counts, the FW90%DCR. (a) Instantaneous response. (b) $\Delta t = 1.28$ s, scan period=35 minutes, $\theta = 6$ degrees,

The A-2 Experiment                                              Spatial Response

$\phi$ = 1.5 degrees.

count rate comes from what solid angle, we see in Figure II.11 that the size of the smeared detector is not as large as the region of non-zero response might suggest. Again, plotting a series of percentile contours is awkward, so to characterize the size of the detector in terms of a pair of numbers we calculate a single rectangle which contains 90% of the total count rate due to a diffuse surface brightness. For an unsmeared rectangular collimator this would be at $1.55 \theta_o$ x $1.55 \phi_o$ . We call this size the FW90%DCR, Full Width 90% Diffuse Count Rate. After smearing by $\Delta\theta$ = $6^o$ and $\Delta\phi$ = $1.5^o$ the HED 1 FW90%DCR is 11.2 degrees by 4.4 degrees.

## Response frequency and response weights

Quite often we will need to calculate many integrals of the form $\int d_\Omega$ f( $r(\theta,\phi)$ ). The two dimensional nature of this integral can be computationally taxing. In those cases where the integrand has no explicit dependence on $\theta$ and $\phi$ (and implicit dependence only through the response function r), the integral can be converted to a one dimensional response "frequency" integral,

$$\int dr \; \nu(r) \; f(r) = \int d_\Omega \; f( \; r(\theta,\phi) \; ), \qquad [II.16]$$

where $\nu(r)$ is the differential number of steradians with response between r and r+dr.

Later another characterization of the detector response function will be useful, the weighted response of the detector:

$$R_a = [ \; (4_\pi)^{-1} \int dr \; \nu(r) \; r^a \; ]^{1/a}, \qquad [II.17]$$

where $R_a$ is the "a"th weight response. Note that $[R_a]$ = $[r(\theta,\phi)]$ = cm2 s. The factor of $4_\pi$ removes the steradian units of $\nu(r)$. If $S_{as}$, which is defined as $4_\pi \; \Sigma$, is the total flux of the X-ray sky, not per steradian, then $I_{as}$= $R_1 \; S_{as}$ is the average number of counts per exposure the detector will see from the sky flux. $R_1$ is invariant under any angular smearing and is proportional to $\Delta t$ for temporal integrations.

Figure II.12: Response freguency, $\nu(r)$ versus r

HED 1 L+S $\Delta t$ =1.28 s.  (a) no angular smear.  (b) $\Delta\theta$ = 6 degrees,  $\Delta\phi$ = 1.5 degrees.  The peaking of the response frequency is another indication of how the smearing flattens out the detector response.
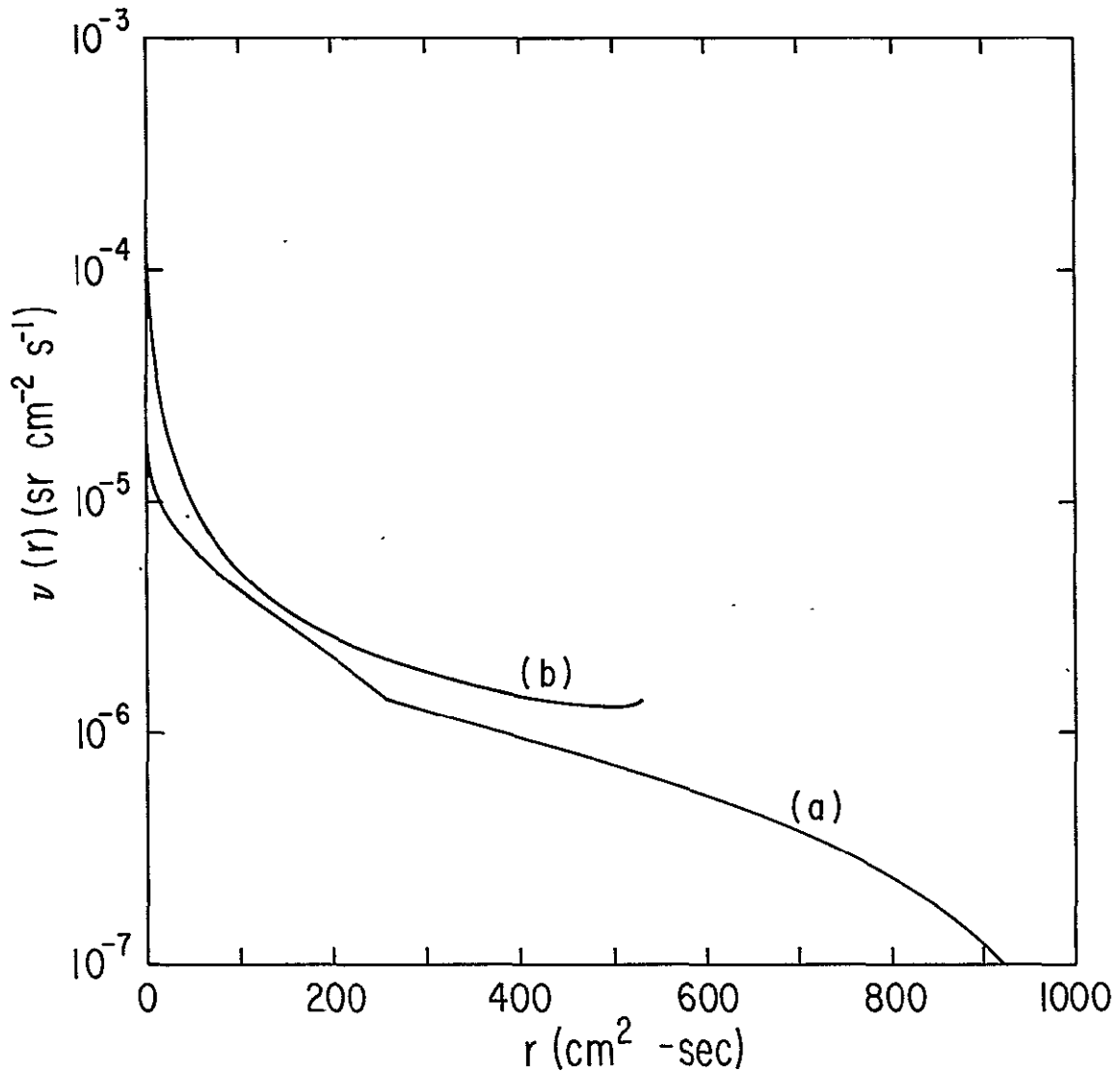


In terms of the descriptions of the instantaneous detector,

$$R_1 = \alpha \ \Delta t \ / \ (4 \ \pi).$$ [II.18]

For some applications we will need to separate out the solid angle

The A-2 Experiment                                    Spatial Response

dependence of the weights:

$$C_a = R_a / (A \, \Delta t). \hspace{4cm} [II.19]$$

Equations [II.17] and [II.19] are of course undefined for a = 0, but we can define a dimensionless covering factor:

$$C_o = (4\pi)^{-1} \int_{>0} dr \; \nu(r). \hspace{3cm} [II.20]$$

$C_o$ is just the fraction of the total sky contributing a non-zero response.

For purposes of illustrating aspects of measurements of the background we will have occasion to use an idealized detector. This detector has a flat response over a solid angle $\Delta\Omega$. The weighted responses are

$$R_a = C^{1/a} \, A \, \Delta t. \hspace{4cm} [II.21]$$

A is the ideal detector's area, and $C = \Delta\Omega/4\pi$ is identical to $C_o$, the fraction of the sky covered by the detector.

The measurement of the internal background is not without contamination from the X-ray sky, as noted in the discussion following equation [II.12], above. The contamination is due to possible anisotropies in the X-ray surface brightness. By smearing the response $r_B$ (equation [II.13]) we can reduce this contamination, but even smears of $\Delta\theta$ on the order of the FWHM leave significant signal. However, if we increase the smear greatly, we see in Figure II.13 that this contamination can be greatly reduced. $R_1$ is zero in all cases. $R_2$ is reduced by the additional smearing. As we will show, $R_2$ is proportional to the excess variance the X-ray sky contributes to a measurement of the internal background, so the impact of this sky variance on studies of the internal background may be greatly reduced by analyzing data with large amounts of smear.

For measurements involving both positive and negative response, such as the internal background measurements and for phase-switching interferometers often used in studies of the radio background, the non-integral weighted responses are undefined. However, these weights are

Figure II.13: Response for the internal background FOV combination,

$$r_B(\theta, 0)$$

$\Delta t$ = 1.28 s,  $\Delta \phi$ = 1.5 degrees.    solid:  $\Delta \theta$ = 6 degrees.
dashed:  $\Delta \theta$ = 30 degrees.
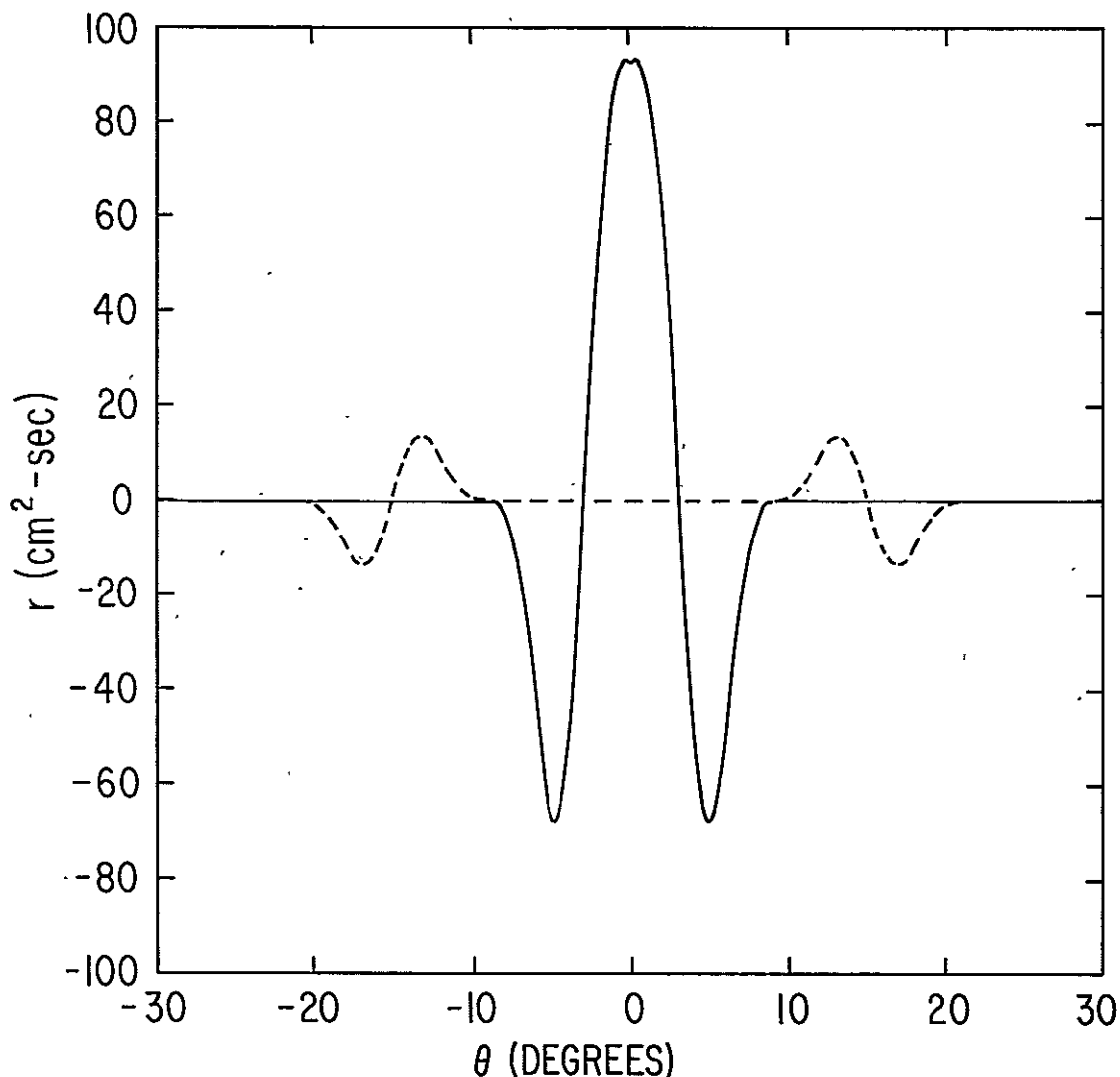


particularly useful in making models of the contribution of unresolved sources to the distribution of measured intensities. For detectors with negative response the best approach is to divide the response into two halves: $r_+(\theta, \phi)$ where $r > 0$, and $r_-(\theta, \phi)$ where $r < 0$. The response frequency of the "−" detector can be represented as the positive

Figure II.14: Response frequency, $\nu(r)$, for $r_B$

$\Delta t = 1.28$ s, $\Delta\phi = 1.5$ degrees.   $\Delta\theta = 6$ degrees.
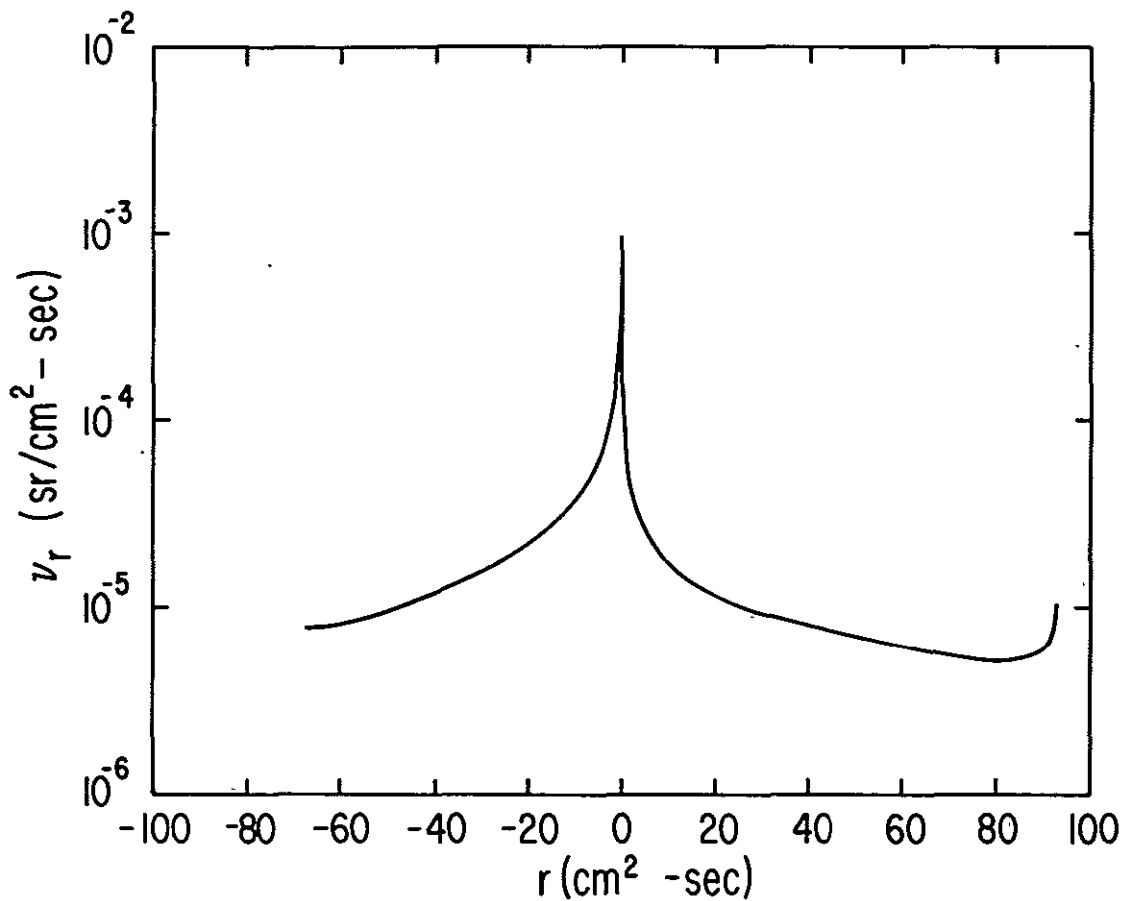


quantity:

$$\nu_-(r) = \nu(-r) \ ; \ r < 0. \tag{II.22}$$

The moments of the "-" detector will be always defined, but care must be exercised. The techniques to recombine the two halves are treated in more detail in Appendix A3.

## 3: Spectral Response and Flux Measurements

The spectral response of the detector determines how various
incident spectra affect the count rate. A brief description
of this response and of an approach to convert fluxes in
counts to ergs (2-10 keV) is given. The availablity of
several spectral windows and their effective spectral range
is discussed.

The proportional counter technology used in the A-2 detectors
provides moderate spectral resolution. At 6 keV an incident photon's
energy may be _inferred_ from pulse height logic with a statistical
imprecision of ~16%, FWHM. To correctly deduce the incident spectrum
requires a good understanding of the atomic physics underpinning the
spectral response. The detection process involves two aspects:     (1)
what is the probability that an incident photon of energy E will be
detected, and (2) what is the probability that such a detected photon
will be counted in a particular pulse height channel. These aspects may
be combined in terms of a response matrix. A given incident photon
spectrum may be folded through this matrix to predict the expected count
rate in a given channel.

Fine pulse-height binning is important for characterizing the
spectra of the brighter resolved sources and that of the total diffuse
flux. However, in studying the fluctuations and other structure in the
X-ray sky there will be insufficient counting statistics to provide
accurate information at the finest bin sizes. Instead we consider broad
band measurements over a relatively large range in pulse-height space.
The very broadest are measurements from an entire layer of the detector
above a certain minimum pulse threshold. These measurements might be
characterized by a "typical" energy range; that is, the energy range
spanned by the photons that dominate the measurement. There are a
variety of ways to determine this range, and we will develop two in this
section.

The total efficiency

Let R(E,h) be the differential response function, the probability

that incident photons of energy E produce counts with pulse height h.
That is, $R(E,h)$ $E^{-1}$ $s(E)$ dE dh counts are produced in the pulse height
interval from h to h+dh by an incident energy spectrum of $s(E)$ dE. The
factor of $E^{-1}$ converts the energy spectrum to photons. A major
underpinning of the spectral analysis programs at Goddard and elsewhere
has been the detailed determination of the structure of $R(E,h)$ for the
various detectors using experimental measurements and theoretical
modelling of the detector physics. A description of this process as
applied to the earlier but similar OSO 8 detectors is found in Pravdo
[1976]. The response, and subsequent abstractions from it, are subject
to change over the detector lifetime. A particular case of this is
drift in the detector "gain", the number of electrons or total pulse
height, produced by a typical detection process. The gain is dependent
on such possibly variable quantities as active gas density (and
therefore the temperature), anode high voltage, and the presence of
various impurities. The A-2 detectors were remarkably stable over the
short run and evidenced only modest variability of the gain in the first
year of the mission. Nonetheless, it should be kept in mind that the
results of this section are formally applicable to a limited range of
time around a particular epoch. All graphs, figures and tables depend
on calibrations based on studies of the Crab Nebula spectrum on or about
day 257 (14 September 1977).

We consider a measurement taken from a window in pulse height
space, $\Delta h$. One way of looking at the energy information in such a
measurement is the total efficiency,

$$Q(E) = \int_{\Delta h} dh \, R(E,h) \, , \qquad\qquad [\text{II}.23]$$

which is just the total probability that a photon of energy E will be
detected as a count in the window $\Delta h$. The plot of $Q(E)$ in Figure II.15
shows that the first layer of an HED has greater than 10% efficiency
over the range from 2.5 to 46 keV. We will define $\Delta E_q$ as the energy
range that contains all places where $Q(E)$ is at least 10% of its peak
value.

Figure II.15: Total efficiency, Q(E), for HED 1

Solid: the first layer efficiency. Dashed: the second
layer. The reduction of the second layer efficiency is
due to the almost total interception of the lower energy
X-rays in the first layer. At high energies, where the
first layer becomes more transparent, the two efficiencies
are nearly identical. A similar plot for HED 3 has the
same shape. Based on gain and other fits for day 257
(1977).
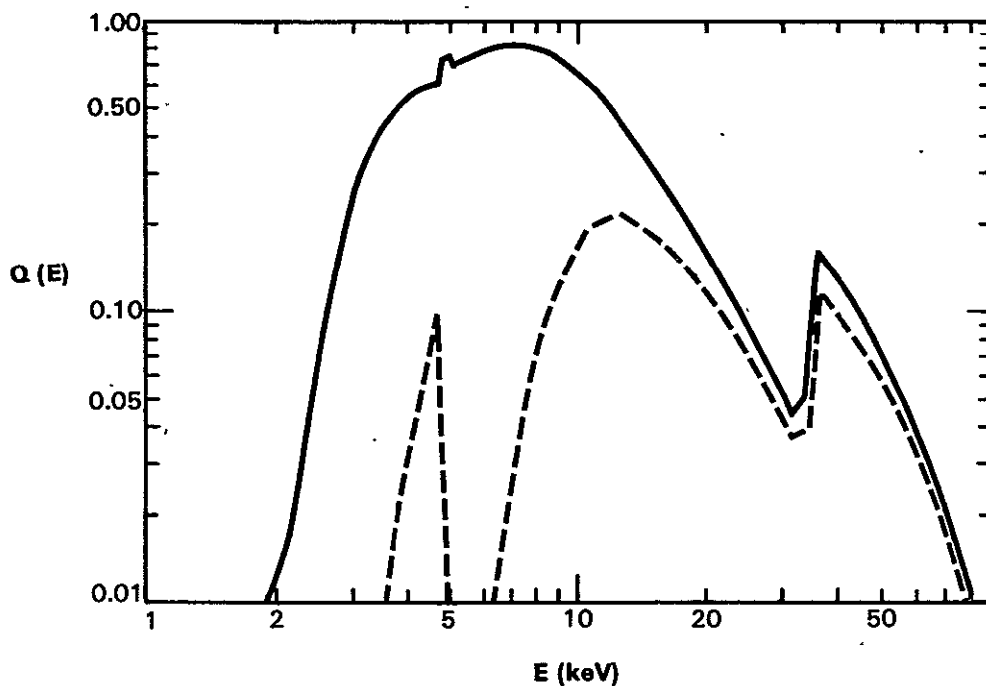


Incident spectra weighted energy ranges

Q(E) as a measurement of a window's spectral response is
independent of the incident spectrum, s(E) dE. For a flat incident
photon spectrum Q(E) will indicate "where the photons come from."
However most astrophysical spectra are at least moderately steep. Thus
there are fewer incident counts at energies of 30 keV or greater than
below 10 keV. The significant values of Q(E) at high energies might
incorrectly lead us to feel that a significant fraction of the observed

The A-2 Experiment                                    Spectral Response

ORIGINAL PAGE IS
OF POOR QUALITY

Figure II.16: Total efficiency for MED, day 257 (1977)

Solid: Layer 1. Dashed: Layer 2. The total efficiency of
the second layer exceeds that of the first layer at high
energies. This is because the second layer of the MED is
effectively twice as thick.
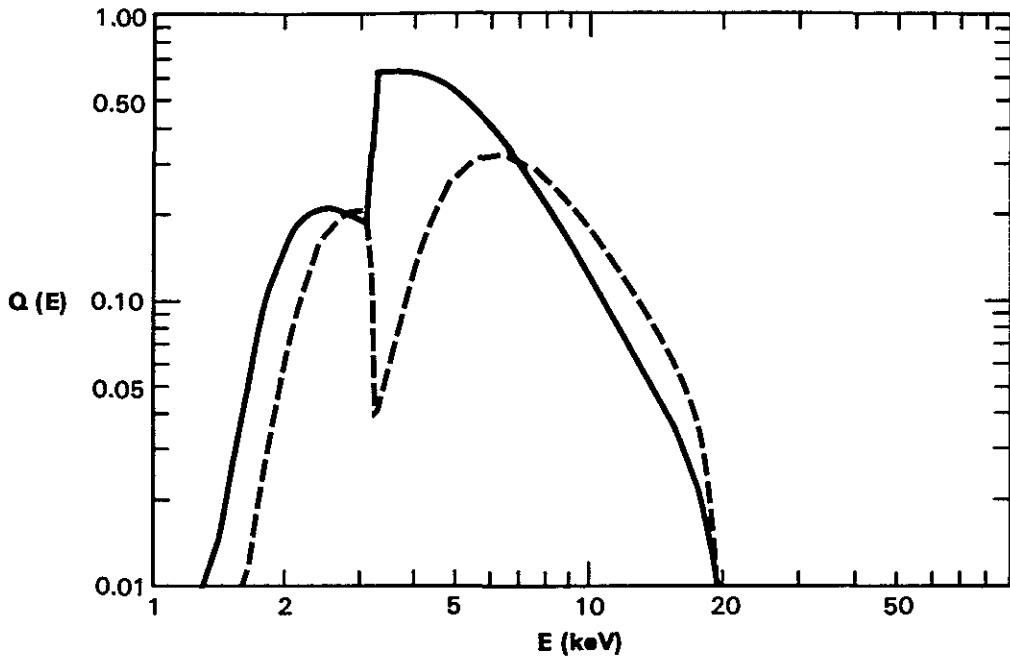


counts of a broad band measurement come from these energies. To
compensate we weight Q(E) by the incident photon flux. The total count
rate is

$$S_{counts} = \int_{\Delta h} dh \int dE\ R(E,h)\ E^{-1}\ s(E)\ .$$

[II.24]

If we reverse the order of integration the dh integral is equivalent to
the definition of Q(E) in [II.23]. We define a function

$$P(E)\ dE = dE\ Q(E)\ E^{-1}\ s(E)\ /\ S_{counts}\ .$$

[II.25]

The normalization by $S_{counts}$ allows us to interpret P(E) as a
probability density. P(E) dE is the probability that a particular
observed count started off as a photon with energy between E and E+dE.
We can use P(E) to define a pulse height window's energy range by the

The A-2 Experiment                                    Spectral Response
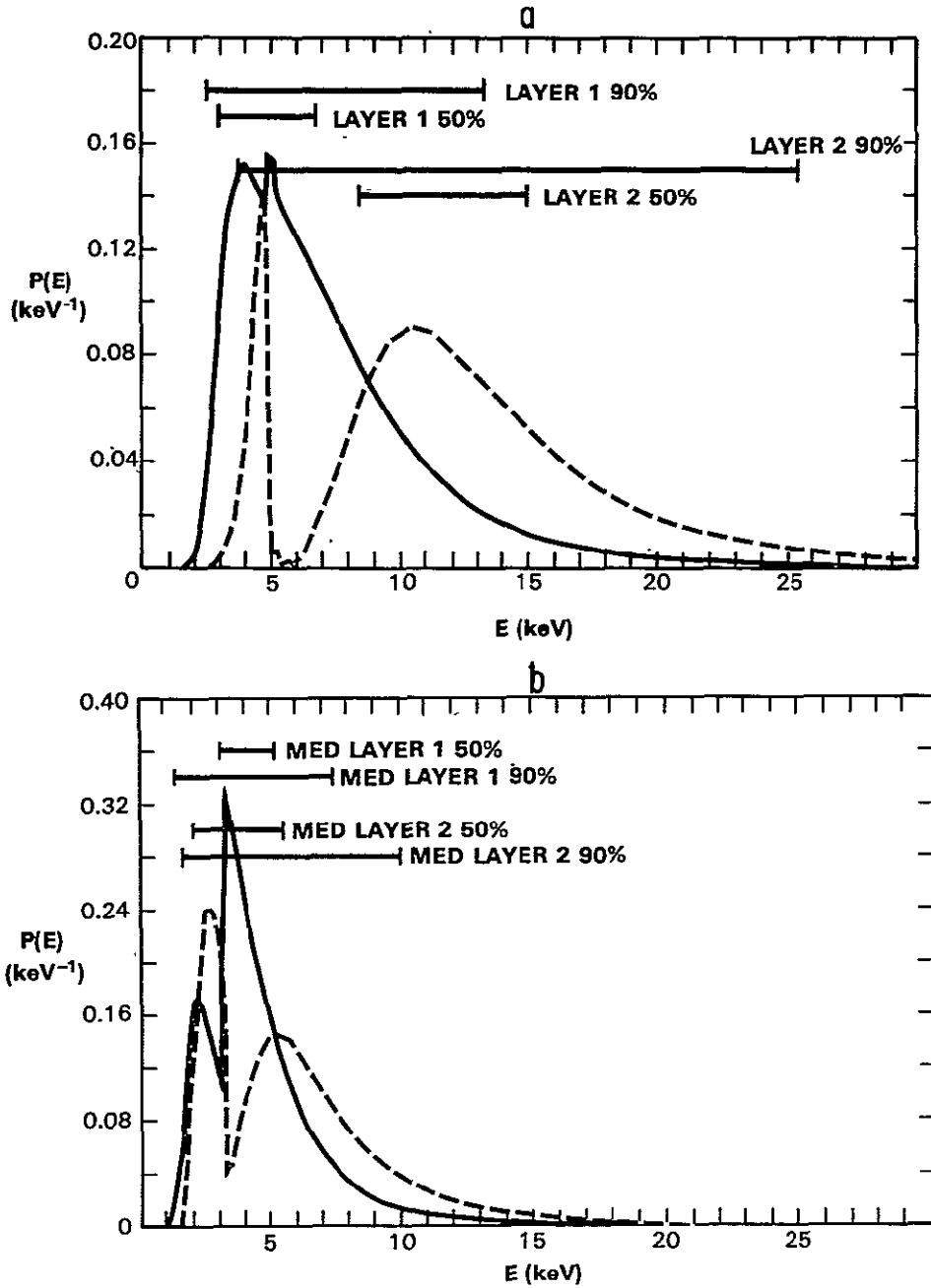
Figure II.17: HED 1 and MED incident photon distribution, P(E)

Solid: Layer 1. Dashed: Layer 2. (a) HED 1. (b) MED.
Incident spectrum was a thermal bremsstrahlung with kT =
40 keV. The bars indicate the energy intervals where 50%
and 90% of the counts originate. Note that the scale of

Spectral Response

this plot is linear, in contrast to the log-log plot of
II.15. Based on day 257 detector response data.

shortest interval that gave rise to a given fraction of the counts, in
analogy to the definition of the FW90%DCR as a characterization of the
detector's spatial size. Examining Figure II.17 we note that for the
spectrum of the diffuse background, the P(E) for the two different
layers of an HED have a markedly different distribution of incident
photons. The 50% of total count rate intervals, $\Delta E_{50}$, are disjoint:
3-6.7 keV for layer 1 and 8.4-15 keV for layer two. However the large
peak in the response of the second layer below the edge at 4.8 keV
contributes a substantial fraction of the total counts so that for $\Delta E_{90}$
there is substantial overlap: 2.5-13.3 keV for layer 1 and 3.7-25.4 keV
for layer 2. The extension of the layer 2 range to high energies shows
that it is more sensitive to high energy effects than to the low,
compared to the first layer. The strong overlap cautions against
treating the two as totally independent measurements.

Although $\Delta E$ properly weights the lower energies, insofar as it
does not imply extension of the window to high energies where there are
very few incident photons, it depends on the exact form of the incident
spectrum. No single spectral form over the energies covered by the A2
detector even roughly approximates the variety of astrophysical spectra
seen. Yet many extragalactic objects can be classified into two broad
spectral categories: clusters of galaxies having thermal spectra with
kT typically in the range of 4 to 8 keV, and active galactic nuclei,
principally Seyfert galaxies, having power law spectra with energy
indices of 0.55 to 0.80. A third, still different, spectrum is the
spectrum of the diffuse background itself. The actual energy range of a
window can have a large change with spectrum, particularly if the
incident spectrum is steep over the peak Q(E) region of the window.

### Conversion from counts to ergs

Though measuring S in detector based units, such as counts $s^{-1}$
$cm^{-2}$, may be the most honest policy when the incident spectrum is not

specifiable, it does not lend itself to making comparisons with other detectors or with astronomical and astrophysical data from other spectral bands. A more typically quoted unit for flux in X-ray astronomy is ergs $s^{-1}$ $cm^{-2}$ over a given spectral region, easily related to the incident spectrum by

$$S_{ergs} = \int_{\Delta E} dE \; s(E) \; .$$                                            [II.26]

Traditional ranges for $\Delta E$ are 2-10 keV for the higher energy proportional counters, such as the MED and HEDs of the A2 experiment, the Ariel V SSI, and UHURU; and 1-3 keV for the instruments limited by the telescope response of the Einstein Observatory. To translate from $S_{counts}$ to $S_{ergs}$ we define the conversion coefficient

$$\varepsilon = S_{ergs} \; / \; S_{counts} \; .$$                                         [II.27]

$\varepsilon$ is the number of ergs in the range $\Delta E$ produced by a source with an observed flux of 1 count $s^{-1}$ $cm^{-2}$. This conversion depends explicitly on the incident spectrum s(E). We also note that the numerator is an integral over the defining interval $\Delta E$, while in the denominator the integral is over all energies where the response or spectrum is non-negligible. Therefore the conversion factor, $\varepsilon$, may be affected by the behavior of s(E) over regions outside the interval $\Delta E$. Particular values for $\varepsilon$ for a variety of incident spectra and detector pulse height windows are presented in Section A7.1.

The term "flux" also refers to the incident differential energy flux, s(E), measured at a particular energy. This is often used in optical and radio flux measurements; for radio work the typical unit is the Jansky (= $10^{-23}$ ergs $Hz^{-1}$ $s^{-1}$ $cm^{-2}$). Such a characterization for X-ray measurements is of value in those cases analagous to measurements in other spectral bands: high spectral precision, i.e. resolution comparable to structure in the spectrum, and/or an a priori understanding of the shape of s(E). Such cases rarely obtain for the broad band measurements of the fluctuations and other structure of the diffuse background.

The pulse height windows available on the A-2 experiment cover a wide range, from PHA channels, with size as small as 1/128th of the total pulse height range, to the entire layer as limited by lower and upper discriminators. The basic experiment electronics were capable of producing a wealth of information, conceivably giving the PHA channel and time of arrival information of each detected count. The severe constraint of the total telemetry available to the experiment required a compression of this information. There was a great deal of flexibility in the experiment in selecting which information to send, a feature often used when studying individual sources in point mode. During scan operations, the telemetry was normally in one of a few standard configurations. The PHA data were available on a 10.24 second basis, corresponding to a smear of 1.75 degrees for a 35 minute scan period, or 7 degrees for 40.96 seconds.

The Discovery Scalers provided a set of rates available more often, every 1.28 seconds (0.22 degrees). Each detector had 4 pairs of Discovery Scalers, the members of each pair were spectrally identical but associated with separate fields of view. The first pair were denoted Discovery Scalers 1 and 2, the second pair 3 and 4, etc. With this nomenclature the odd member of each pair by convention corresponded to the nominal 3° by 3° field of view, i.e. the SFOV for HED 1 and 2, LFOV for MED and HED 3. The first two pairs read out the counts from the two detector layers above a low pulse height descriminator. The remaining two pairs could be used to measure the rate in smaller pulse height windows. The first layer was divided into four such pulse height windows 1A, 1B, 1C, and 1D, while the second layer had only two windows: 2A and 2B. The exact combination of windows selected for the last two pairs was commandable from the ground, and was modified at intervals throughout the mission. These window combinations allow finer spectral resolution than the layers alone can give. From Figure II.18(b) we see that the HED 1 rates defined by the sum of the 1A and 1B windows and the sum of the 1C and 1D windows are significantly more disjoint, even at the 90% level, than the first layer versus the second. In analyzing the

Figure II.18: HED 1 layer 1 pulse height windows: (a) Q(E), (b) P(E)

Solid: 1A=!B window. Dashed: 1C+1D window.
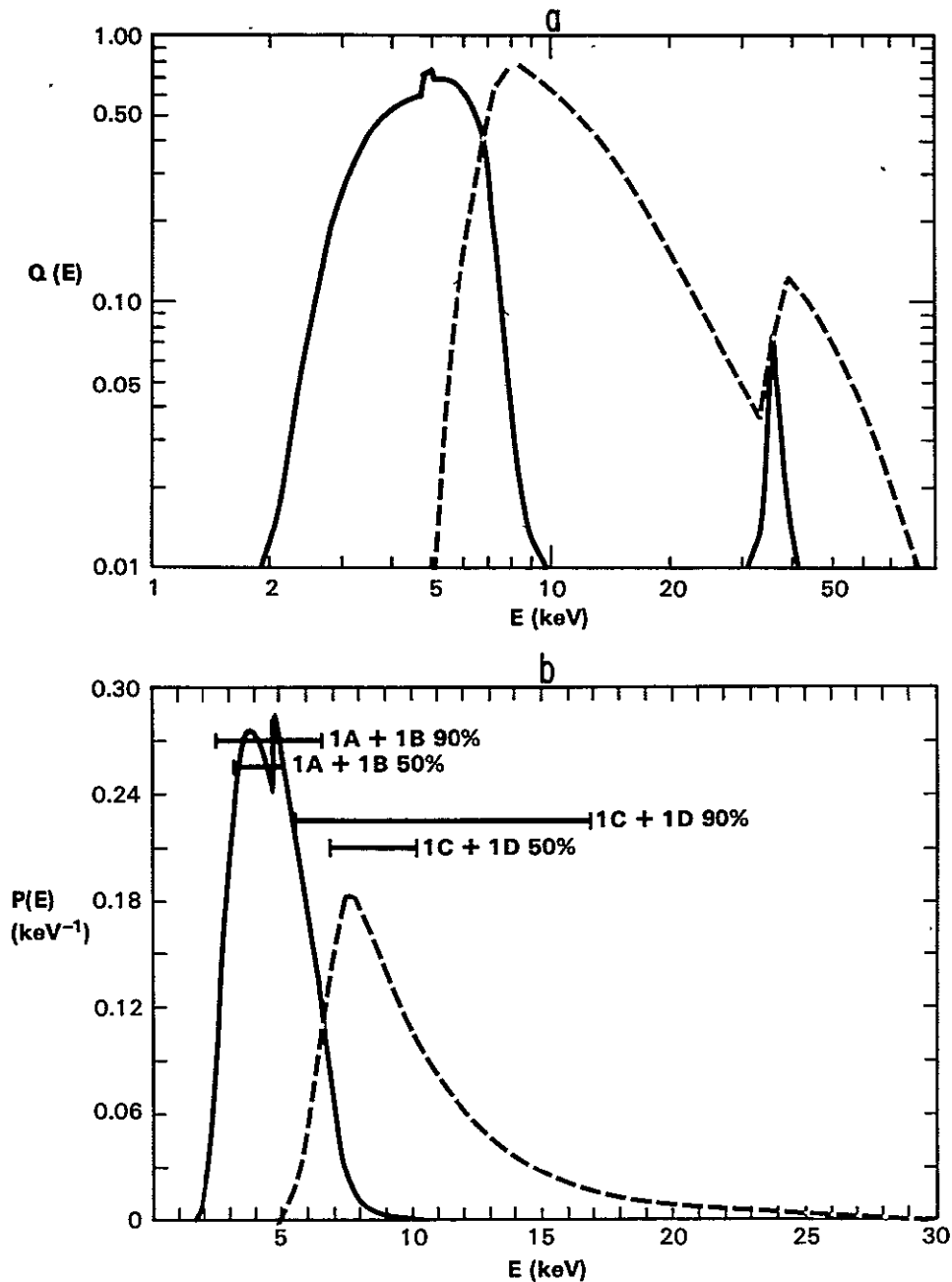


fluctuations we will confine ourselves to the information available from the Discovery Scalers, with particular attention to the first layer of

The A-2 Experiment                                        Spectral Response

the HEDs, as that provides the best signal to noise ratio.

---

### 4: Experiment History

The HEAO 1 spacecraft was launched 12 August 1977 and the A-2 detectors were operating fully within a week. The initial orbit was circular, with 445 km apogee and an inclination of 22.75 degrees. The orbital period was 93 minutes. The observatory spin period during scan observations was initially set at 33 minutes, but was soon increased to 35 minutes for most of the mission.

Time for the data was given in day of year, 1977, and milliseconds. We will here report time in decimal days, 1977. The experiment went through a variety of configuration changes during the first 6 days, but a period of stability was achieved for the HED detectors at day 234.95. The MED did not reach a long term state until day 248.15. Subsequently, the windows selected for the third and forth pairs of Discovery Scalers were modified.

By using combinations of the Discovery Scalers, we have access to window combinations not directly chosen. For instance, during the interval from day 321 to 476, we can subtract the number of counts in the third pair (window 1B) and the fourth pair (windows 1C+1D) from the number of counts in the first pair (layer 1=1A+1B+1C+1D) to get the number of counts that a Discovery Scaler for window 1A would have gotten. The emphasis on the word counts is to indicate that the subtraction gives the true number of observed photons in that window. The uncertainty in the rate due to counting statistics is just the square root of this number of counts, and is not the square root of the sum of the variances of the Discovery Scalers used to generate the counts. This contrasts with estimates of the diffuse sky rate derived from the large minus the small FOV and the associated counting statistics as discussed in equations [II.2] and [II.3].

| | HED 1 | | HED 2 | | MED | | HED 3 | |
|---|---|---|---|---|---|---|---|---|
| DOY (1977) | Third pair | Fourth pair | Third pair | Fourth pair | Third pair | Fourth pair | Third pair | Fourth pair |
| 234.94 | 1B | 1D+2B | 1B | 1D+2B | 1A | XX | 1B | 1D+2B |
| 238.89 | " | " | " | " | 1B | XX | " | " |
| 242.20 | " | " | " | " | " | 1D | " | " |
| 248.15 | " | " | " | " | " | 2B | " | " |
| 304.93 | " | 1C | " | 1C | " | " | " | 1C |
| 321.79 | " | 1C+1D | " | 1C+1D | " | " | " | 1C+1D |
| 476.20 | " | " | " | " | " | " | 1A | " |
| 488.82 | " | " | " | " | " | " | 1B | " |
| 615.27 | " | " | " | " | " | 1C+1D | " | " |

Table II.2: History of discovery scaler definitions

NOTE: During the interval from 235.63 to 245.69 the MED processed no events in the 2B window, so that the second layer rate as measured in the second Discovery Scaler pair measured only the 2A window.

The detectors' performance in orbit was monitored, and their response and internal background were stable, within limits to be discussed later. The effects of long term gain drift on the definition of the Discovery Scalers' energy ranges amounted to a few tenths of a keV over the duration of the experiment. On or about day 544 the experiment suffered an "event" of unknown origin, although it is suspicious that this was coincident with the loss of propane gas by the LED flow counters. At this event the mean count rate for the detectors increased by a few percent, ascribed to a change in gain. The HED 3 seemed most affected, while HED 1 appeared to change least, less than one percent. After the event, detector gain stability, particularly for HED 3, was degraded.

The observatory was almost entirely in scan mode during the first 6 months of the mission, completing an all sky survey. After this major

mission goal was accomplished, the observatory was with greater
frequency placed in point mode for the study of individual objects and
small areas of the sky. This had the obvious effect of decreasing the
total coverage, in seconds, available for studies of the unresolved sky
flux. Near the end of the mission, the spacecraft was typically in
point mode over 50% of the time.

The quality of the aspect solution degraded during January 1979,
with the attitude gas running out shortly thereafter. HEAO 1 re-entered
the Earth's atmosphere on 14 March, 1979. Almost one and a half years
of data were accumulated, providing two and a half complete scans of the
sky.

————————————————

## 5: The XRATES Database

The Discovery Scaler data we will use to analyze the fluctuations
and other large scale structure in the X-ray sky required a fair amount
of detailed and specialized processing. Details of the resulting data
base, called the SCANER tapes, will be covered in a later chapter. The
initial data used by the programs that eventually produced the SCANER
tapes were from the XRATES database. We here describe the production of
XRATES, emphasizing the data used in our subsequent processing to
produce SCANER.

The HEAO 1 A-2 experiment telemetry, in data tapes provided by the
Goddard Information Processing Division (IPD), was initially processed
by a series of programs written and maintained by the Computer Sciences
Corporation (CSC). The three programs used to produce information for
the fluctuations data sets were: FLAP, the First Look Analysis Program,
which provided microfilm plots of the principal detector rates; FRAPPE,
the Frame Re-blocking and Production Processing Executive, which
reformatted the tapes provided by IPD and calculated the principal data
quality condition flags, output on the MAX database; and finally SKYMAP,
which produced an extract of the MAX tapes (principally restricted to

the Discovery Scaler data), and calculated a few further data quality flags, creating the XRATES tapes.

Figure II.19: Production system for XRATES tapes
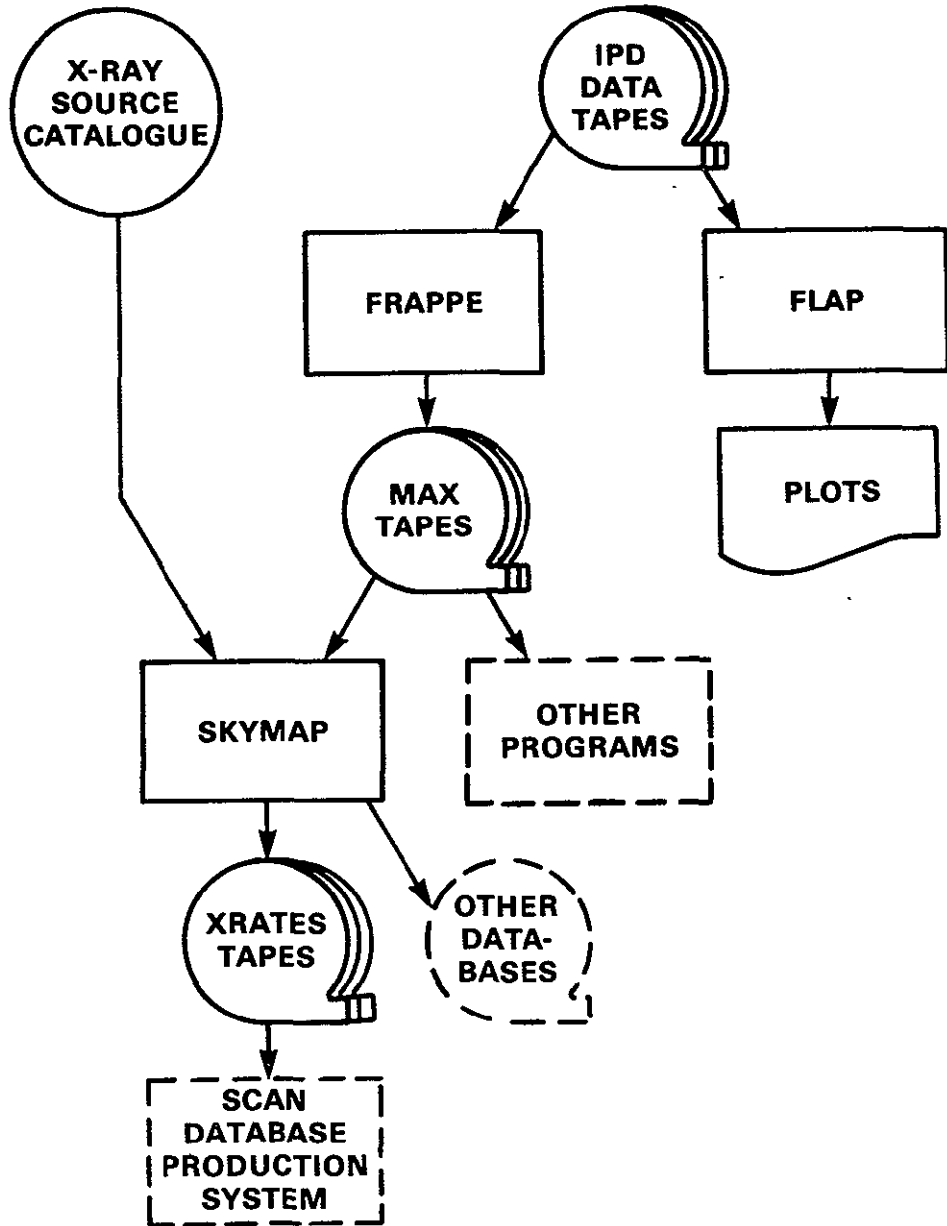
In Figure II.19 a schematic diagram is given showing how the XRATES tapes are produced. A detailed description of the programs and tape formats may be found in the two CSC documents: TM 79/6231 and TM 81/6123, describing FRAPPE and SKYMAP respectively.

Many of the data quality flags used in the selection of data for the SCANER database are produced by FRAPPE and SKYMAP. They are produced on a major frame basis, that is, every 40.96 seconds. There are 32 1.28-second Discovery Scaler integrations every major frame. In scan mode, the detectors cover 7 degrees in one major frame. Two flags applying to the spacecraft as a whole are a point flag, indicating whether the experiment is pointing or scanning, and the SAA/NPA flag, showing when the spacecraft is within the geomagnetic anamoly regions over the South Atlantic and North Pacific. (These are regions that correlate with a high flux of trapped electrons which produce high internal backgrounds [Shulz and Lanzerotti 1974].)

Other data flags are individually calculated for the various detectors, e.g. HED 1, 2, and 3, and the MED. The CLEAN flag indicates whether the detector satisfied several conditions: the detector field of view excluded the Earth, with an additional 100 km allowance for effects of the atmosphere; the detector high voltage was stable; there were no data transmission errors including fill data and bit errors; and, in the case of the MED, the calibration source was not in the field of view. An additional Earth occultation flag, HEOCC1, indicates if the detector was occulted by the Earth plus a 200 km additional distance.

The electron contamination flag, HECONF, indicates if the "electron rate" for each detector was within a certain range, one which presents an acceptably low rate of contamination by charged particles of the detector's X-ray count rate. The electron rate is formulated from "housekeeping rate" information about coincident events and events in the "veto" layers. Both the formulation and a reasonable electron rate range were based on an analysis done by Dr. Alan Rose from data taken the first few weeks in orbit. (A reasonable range for the electron rates means a range which rejects most of the bursts of electron

contamination while not having too high a rate of accidental rejection of good data.) It was later decided that the initially chosen threshold for the MED was too strict, rejecting too much good data, particularly when one wished to use only data from the second layer of the detector. The second layer benefits from the electron-trapping efficiency of the first layer, which acts as an electron veto layer in much the same way as the propane layer does on the HED.1 and 3 detectors. For this reason the MED has two electron rate thresholds, allowing accumulation of two sets of MED data: "electron clean" data using the relaxed criterion and "electron super-clean" data satisfying the original, more stringent, electron rate criterion.

Other information available on a 40.96 second basis includes the position of the Moon, in spacecraft coordinates, and the value of the McIlwain L parameter, which characterizes the magnetic shell in which trapped electrons drift in terms of the equatorial radius of a field line in a dipole field [McIlwain 1961]. The L parameter strongly correlates with residual variations in the level of the internal background, as we shall see, because a major source of the internal background is activation of the spacecraft by cosmic rays. L is a measure of the cut-off rigidity and thus the amount of shielding from these particles that was provided by the Earth's magnetic field.

The XRATES tapes provide Discovery Scaler information for each of the detectors on a 1.28 second basis. Aspect information is also provided, giving the orientation of the spacecraft on a 1.28 second basis with an accuracy generally well within 0.1 degrees.

CHAPTER III

## THE ORIGIN AND DESCRIPTION OF THE FLUCTUATIONS

A detailed study of how unresolved discrete sources will
produce fluctuations in an otherwise isotropic background is
presented. The effect depends on models of the source
number versus intensity relationship. A family of
reasonable models is developed and we discuss the ultimate
sensitivity of the fluctuations to the source behavior.
These fluctuations are also shown to have an impact on other
measurements of the X-ray sky, as well as measurements of
low intensity sources.

That deviations from isotropy contain useful information about the
origins of the X-ray background has long been recognized. Early
treatments of the problem, when only upper bounds to the strength of the
fluctuations were available, were at first somewhat unsophisticated. As
the quality of X-ray surface brightness measurements improved and with
the launching of the first all-sky survey instruments on board UHURU and
Ariel 5, the models and mathematical tools used to analyze the
fluctuations became more complex, following the example of analysis of
fluctuations in the radio background. Though there are several
derivations and explications of the techniques extant in the literature,
with advent of the HEAO 1 A-2 database, a careful accounting of the
origins and limitations of these tools should be made. Many issues that
previously could be ignored with little harm must now be directly faced.

In this chapter we review the techniques for modelling the
fluctuations and present an analysis of their possible pitfalls based on
rigorous assessment of the limitations of any finite experiment.

––––––––––––––––––––

### 1: Discrete Source Luminosity Functions and N(S)

A general description of luminosity functions is made. It
is shown that objects uniformly distributed in an infinite

Euclidean space will produce a power law form for the
differential source counts function N(S), of index -5/2, for
any form of the luminosity function. Deviations from this
simple form of N(S) are calculated for the elementary cases
of a finite size Euclidean space and for a nonuniform
distribution of sources such as an exponential scale height
disk population. The effects of different luminosity
functions are then apparent.

Distribution functions are common tools used in the studies of
ensemble properties of a class of objects. If x is an intrinsic
property of the objects then we are often interested in the distribution
$\phi(x;z,...)$, the density of objects with a value of x between x and x+dx.
Examples of intrinsic properties might be bolometric luminosity,
temperature, or morphological type. The subsidiary variables of $\phi$,
represented by the z,... above, show that the distribution might change
as a function of extrinsic variables such as position and epoch. It is
often assumed that this density varies only slowly with the spatial
coordinates. To be more explicit, we consider volumes large enough that
effects of clumping or other correlations between the objects are small.
With current observational and theoretical work indicating structures as
large as 100 Mpc [e.g. Chincarini and Rood 1979; Gregory, Thompson, and
Tifft 1981; Kirshner, Oemler, Schechter and Shectman 1981; and Davis et
al. 1982] it is important to recognize this assumption and estimate the
impact of these clumpings and voids. But in this chapter we will start
by examining non-clumped distributions of objects. In Section VII:6 we
present a simple analysis of the impact of large-scale spatial
inhomogeneities.

The distribution $\phi$ is often parameterized in terms of the object's
luminosity, and in such a form is called a luminosity function. As we
remarked in Section II:3 the value of a source flux based on broad band
spectral measurements is dependent on the assumed incident spectral
form. If all members of the class are well described by a known form
then the transformation from counts to ergs, and thus to luminosity is
well defined. One can use the bolometric luminosity, or the
differential luminosity at a particular energy, but it is more common in
X-ray astronomy to use the integrated luminosity over a particular band.

Such a parameterization is less sensitive to imprecision in the spectral shape than is the differential luminosity, particularly if the spectral band is well matched to the detector band pass and the source spectrum does not vary too widely over this range.

In order to have a good understanding of the spectral properties of a set of objects, we use a luminosity function of the form $\phi(\ell(E),E;z,...)$, where $\ell(E)$ is the differential luminosity. This makes it possible to relate measurements made at widely different spectral bands; e.g., the HEAO 1 A-2 medium- and high-energy data versus data derived from observations carried out with the Einstein Observatory (HEAO 2). When observations are made in a cosmological setting, the effect of spectral form on interpreting the luminosity function, because of the redshift, becomes important long before other cosmological effects due to geometry, evolution, or high redshift cutoffs.

If one has a well defined and complete sample of objects, then techniques exist for deriving the luminosity function directly. (For examples of this in the X-ray regime see Piccinotti et al. [1982], McKee et al. [ 1980], Pye and Warwick [1979], or Schwartz [1978].) Such a determination is often inadequate because it is based on observations dominated by nearby objects, making it difficult to measure any variation of $\phi$ with distance. Broad limitations on the behavior of the luminosity function can be placed by examining its impact on other observables: for example, the number versus flux relationship, represented by the differential relationship $N(S)$ $dS$, and the total sky flux.

### Uniform Euclidean space

To make the connection between $\phi(L;...)$ and $N(S)$, the number of objects in the entire sky with flux between S and S+dS (not per steradian as is often given in other presentations of this topic), we need a relationship giving the flux observed for an object of luminosity L, as a function of position in the sky and distance (or redshift). In cases where the source spectrum does not depend on luminosity, such a

relationship is linear in the luminosity, i.e. $S = L\ A(\vec{r})$. In a Euclidean space with no position-dependent absorption and where both L and S are measured in the same energy units (e.g. ergs, 2-10 keV), the form of A is given by the inverse square law:

$$S = L\ A(\vec{r}) = L\ /\ 4_\pi r^2\ ,$$                                    [III.1]

where r is the magnitude of $\vec{r}$.

Consider objects with luminosity L and density $\phi(L)dL$, uniformly distributed through the Euclidean space, so that $\phi$ is independent of r. The number of objects in an infinitesimal shell at a distance r is

$$(\phi(L)dL)\ dV = (\phi(L)dL)\ 4_\pi\ r^2\ dr\ .$$                        [III.2]

Solving [III.1] for r and determining dr in terms of dS (holding L fixed) we get

$$(\phi(L)\ dL)\ dV = (\phi(L)dL)\ 2_\pi\ (L/4_\pi)^{3/2}\ S^{-5/2}\ dS\ .$$   [III.3]

Integrating over the entire luminosity function we get the standard Euclidean N(S) law

$$N(S)\ dS = 4_\pi\ K\ S^{-5/2}\ dS$$                                         [III.4]

where

$$K = \int\ dL\ \frac{L^{3/2}\ \phi(L)}{2\ (4_\pi)^{3/2}}\ .$$                 [III.5]

So we see that for any shape of luminosity function the form of N(S) is just a -5/2 power law; the only impact the luminosity function has is in terms of the integral magnitude represented through K.

To find the contribution of a class of objects to the all-sky flux, $S_{as}$, we calculate their luminosity density:

$$\Lambda(\vec{r},\ldots) = \overline{L}(\vec{r},\ldots)\ \rho(\vec{r},\ldots) = \int dL\ L\ \phi(L;\vec{r},\ldots)$$   [III.6]

where $\Lambda(\vec{r},\ldots)$ is the luminosity per unit volume, and $\overline{L}$ and $\rho$ are just the mean luminosity and the total density, respectively, of the objects. $\Lambda$ is also referred to as the volume emissivity. The contribution to

the all-sky flux of this population is simply

$$S_{as} = \int dV \ A(\vec{r},...) \ A(\vec{r},...)$$                      [III.7]

where $A(\vec{r},...)$ is the relationship between luminosity and flux in equation [III.1]. For a Euclidean space with uniform volume emissivity, $S_{as}$ equals $A$ R, where R is the radius of the volume. In order to avoid an infinite flux, R must be finite. (This is just a simple formulation of Olber's paradox.)

One can calculate $S_{as}$ just from the form of N(S) of course,

$$S_{as} = \int dS \ S \ N(S) = 4\pi \ K \int dS \ S^{-3/2} \ ,$$          [III.8]

indicating that the 5/2 power law for N(S) can not be extended down to arbitrarily low fluxes. If infinite values for $S_{as}$ are to be avoided, N(S) must break to a form flatter than $S^{-2}$ so that the integrand for $S_{as}$ will be flatter than $S^{-1}$.

### Model luminosity function forms

The form of $\phi(L)$ derived from a complete sample of resolved and identified objects is often well described by a power law over a range of luminosities from $L_{min}$ to $L_{max}$:

$$\phi(L) = \kappa \ L^{-\beta} \ ; \ L_{max} > L > L_{min} \ .$$           [III.9]

The limits on the luminosity range are required to avoid divergent values for K or $A$. The impact of $L_{min}$ or $L_{max}$ on these quantities depends on the value of the power law index, $\beta$ in the following way:

$\beta < 2$    :  $A$ and $S_{as}$ dominated by $L_{max}$

$\beta > 2$    :  $A$ and $S_{as}$ dominated by $L_{min}$

$\beta < 5/2$  :  K (and source counts) dominated by $L_{max}$

$\beta > 5/2$  :  K dominated by $L_{min}$.

For $2 < \beta < 2.5$ the study of any flux-limited sample will be dominated by high luminosity objects. This may produce a significant imprecision in an estimate of $L_{min}$, with a consequent imprecision in estimates of the contribution of low luminosity sources to the all-sky flux.

Other functional forms for the luminosity function have been

considered for theoretical reasons [e.g. Bahcall 1979], while an exponential luminosity function has the property that it is naturally convergent independent of $L_{min}$ and $L_{max}$. However, Piccinotti et al. [1982] report that the power law form is always a good fit for the HEAO 1 A-2 survey-derived luminosity functions for extragalactic objects, and other forms are either no better or specifically excluded by the observations.

### Effects of a finite sized universe

To show how the edge of a finite Euclidean universe, of radius R, produces the needed break in the source counts, we examine the case for power-law-type luminosity functions. There is no deviation of N(S) from a 5/2 power law above a flux level

$$S' = L_{max} / 4\pi R^2,$$                                         [III.10]

at which point the very highest luminosity objects have all been counted at the fluxes S>S'. This causes an effective upper truncation of the luminosity function as S decreases further. For a given flux S there can be no sources with luminosities greater than $L'(S) = 4\pi R^2 S = L_{max} S/S'$. Above $S_{min} = L_{min} /(4\pi R^2)$ all the sources have been counted.

We can generalize the derivation of the N(S) law from a uniform distribution of sources to one where $\phi$ has some dependence on position or distance. We replace the constant K in [III.4] with a K(S) function given by

$$K(S) = \int dL \frac{L^{3/2} \phi(L; r=(L/4\pi S)^{1/2})}{2 (4\pi)^{3/2}}.$$    [III.11]

If $\phi$ has no r dependence, then $\phi$ is independent of S, and K(S) reduces to equation [III.5]. In the case of a sharp cutoff in $\phi$ at a particular r the effect is to modify the limits of the L integration as a function of S. For power law forms of $\phi(L)$ (equation [III.9]), the deviation

from Euclidean form becomes

$$K(S) = \frac{\kappa}{(4\pi)^{3/2}(5-2\beta)} L^{5/2-\beta} \Big|_{L_{min}}^{L'(S)} \qquad [III.12]$$

For $S > S'$, $K(S)$ is a constant ($K_0$) found by evaluating the above equation with $L'(S)$ fixed by the luminosity function's intrinsic upper limit $L_{max}$.

---

Figure III.1: K(S) for finite Euclidean space

$L_{max} = 1000\ L_{min}$. Two values of $\beta$, the power law luminosity function index, illustrate the dependence of the behavior of $K(S)$ below $S'$ on the form of the luminosity function.

---
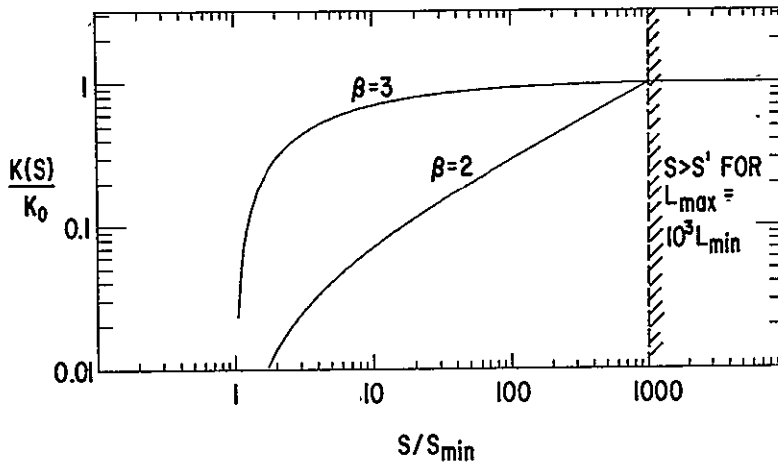


By plotting $K(S)$ rather than $N(S) = 4\pi\ K(S)\ S^{-5/2}$, we are able to compress the information into a more compact figure, but at the cost of needing to retrain our intuition. $N(S)$ has not rolled over until $K(S)$ steepens faster than $S^{+5/2}$. However for the $S_{as}$ to be convergent, $K(S)$

need only steepen faster than $S^{1/2}$ at low S.

In cases where $\beta < 2.5$ and $L_{max} \gg L_{min}$, there is an immediate break in K(S) at S′ and a consequent break in N(S). Indeed for S≲S′, N(S) ∝ $S^{-\beta}$. Thus the size of the break depends on the flatness of $\phi$(L). For cases where $\beta > 2.5$ the break is quite broad and does not become significant until S ~ $S_{min}$. For steep luminosity functions it is not possible to draw an accurate conclusion about the shape of the luminosity function from the shape of the source number counts alone, until S ~ $S_{min}$.

### Source counts for a disk population

A more complicated situation still involving the simplicity of Euclidean geometry is encountered when $\phi$(L;$\vec{r}$) is a function of position. One example is the estimation of the contribution from a population of galactic sources to the high latitude source counts, considered in detail by Rosner et al. [1981]. We assume that the sources are distributed in an infinite disk with an exponential scale height, h:

$$\phi(L;\vec{r}) = \phi(L) \, e^{-|z|/h} = \phi(L) \, \exp( -(r/h) \sin |b| ) \qquad [III.13]$$

where $\phi$(L) is the luminosity function in the plane of the galaxy and b is the galactic latitude. Because of the dependence on angular position we can be interested in more than just the flux distribution of sources over the entire sky; that is, interested in the distribution in a particular patch of solid angle, N(S,$\hat{r}$)dS dΩ. At latitudes and energies where absorption may be neglected so that the relationship between S, L and r is given by [III.1], a derivation similar to the case of the uniform luminosity function gives

$$N(S,\hat{r}) \, dS \, d\Omega = K(S,\hat{r}) \, S^{-5/2} \, dS \, d\Omega \qquad [III.14]$$

where

$$K(S,\hat{r}) = 0.5 \int dL \, \left(\frac{L}{4\pi}\right)^{3/2} \phi(L) \, \exp\left[ \left(\frac{L}{4\pi S}\right)^{1/2} \frac{\sin |b|}{h} \right] . \qquad [III.15]$$

The unit vector, $\hat{r}$, indicates the functional dependence on the direction

in the sky. By considering a collection of objects with the same luminosity L it is easy to see that for fluxes

$$S \gg S'(L,\hat{r}) = \frac{L \sin^2 |b|}{4\pi \ h^2} \ ,$$

[III.16]

the exponential term does not contribute and the source counts are well fit by the uniform $\phi$ Euclidean model. Essentially, for $S \gg S'$ we haven't traveled far enough out of the plane of the galaxy for the variation in $\phi$ to be significant. However, for $S \ll S'$ the value of $K(S,\hat{r})$ quickly

---

Figure III.2: $K(S,\hat{r})$ for exponential scale height distribution

The solid line is the case of a collection of single luminosity objects with luminosity $L_{min}$. The dashed lines are for power law luminosity functions, $L_{max} = 1000 \ L_{min}$, with $\beta$ the power law index. The scale height is assumed to be independent of L.

---



$S/\dot{S}' \ (L_{min},\hat{r})$

---
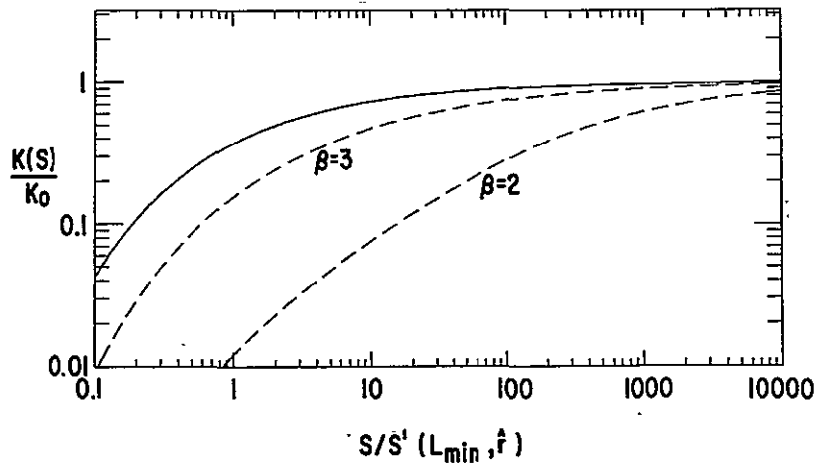
rolls off like $\exp(-S^{-1/2})$. Applying equation [III.15] to the power law

form of $\phi$, we see in Figure III.2 qualitativly similar behaviour to the finite uniform Euclidean space of Figure III.1.

Integrating $N(S,\hat{r})$ over all solid angles gives

$$N(S) \ dS = \int d\Omega \ N(S,\hat{r}) \ dS \propto S^{-2} \ dS; \quad S \ll L/4\pi h^2 \ . \qquad [III.17]$$

Note however that the break to a $-2$ power law for $S \ll S'(L,\hat{r}=$galactic pole) is a property of the all-sky counts, $N(S)$, and the maintenance of the new slope over a wide range is due to the continued contribution of the low-latitude sources. For a particular latitude range of high latitudes, the counts eventually flatten and roll over as $K(S,\hat{r})$ enters the exponential-dominated phase. The roll over occurs at lower and lower fluxes, as the latitude range approaches the galactic equator, at which point the source counts (in this simple model) never depart from the Euclidean form. These observations of the behavior of exponential-scale-height sources will be useful in estimating the effect of galactic source contamination to our analysis where source identifications cannot always be made. This is particularly important for the fluctuations.

The all-sky flux from this infinite disk population also requires an integration over all solid angles:

$$S_{as} = \int d\Omega \ (h/sin \ |b|) \ \Lambda(z=0) \qquad [III.18]$$

where $\Lambda(z=0)$ is the volume emissivity in the plane. The angular dependence of the surface brightness, the integrand, has the familiar cosecant $|b|$ dependence of any infinite plane distribution of sources. Note also that the surface brightness is finite in any direction out of the plane, while the integrated all-sky flux is divergent, in accordance with the all-sky $N(S)$ asymptotic behavior of $S^{-2}$.

Another interesting modification to the Euclidean assumption comes from the effects on $N(S)$ of cosmological corrections with and without evolutionary changes in $\phi$. A simple closed form solution that includes detector bandpass, complicated source spectra, and the cosmological volume and luminosity distance corrections is not possible. The integral equations are derived in Appendix A1 and their application will be outlined in subsequent chapters when we compare information about

N(S) derived from the fluctuations with particular models for the objects that dominate the X-ray sky. A principal complication is that the non-Euclidean form of equation [III.1] is not easily inverted to give z in terms of S and L as we were able to do for r in deriving [III.3]. Two general conclusions will be noted here: for large enough values of S, N(S) approaches a 5/2 Euclidean power law; and any population of objects with a non-evolving luminosity function and a non-inverted (monotonically decreasing with energy) photon spectrum will never have a N(S) with a steeper rise than the Euclidean 5/2 (i.e. K(S) is always a decreasing function of S as S → 0) [Weinberg 1972, equation 14.7.35].

Applying the analysis of this section to the optical number counts versus flux literature is complicated by the fact that luminosities and fluxes are usually represented in terms of magnitudes, which are logarithmic quantities. Because of this, plots of the $\log_{10}(N(S))$ or $\log_{10}(N(>S))$, the integral number of sources with flux greater than S, versus magnitude, will be linear wherever N(S) is well described by a power law. The relationship between the power law index, $\gamma$, and the log(N)-versus-magnitude slope, $\alpha$, is

$$\alpha = (\gamma-1) / 2.5 .$$                                       [III.19]

The slope is the same for differential or integral counts (N(S) or N(>S)), but the intercept changes. It is the intercept that contains information about the value of K, the N(S) coefficient.

-------------------

## 2: Fluctuations in Diffuse Sky Measurements Due to Discrete Sources

Surface brightness measurements cannot be isotropic when unresolved discrete sources are in the field of view. The general Poisson nature of the variation is derived for constant flux sources, and then generalized for power law distributions of source fluxes. The effect of these variations, the fluctuations, can be characterized by their distribution cumulants. The impact of the collimator solid angle response on these cumulants depends on the response weights. However for power law N(S) models, these cumulants

are formally infinite unless N(S) has both high- and low-flux limits.

The N(S) curve allows us to make some conclusions about the discrete source populations, even when we are unable to make accurate identifications and distance measurements of individual sources. Yet the use of N(S) as an analytical tool is limited when we are unable to make accurate estimates of the source intensities, or even decide if one or more than one source is being measured. This confusion is intimately related to the extension of the N(S) curve to low flux values. As we will show here for ideal cases, analysis of the form of this confusion can yield as much information about the source populations as could the resolved source counts, or more. Even in the non-ideal cases we consider when analyzing real data, interesting constraints can be set.

### Sources with equal flux

We begin with a simple characterization of the effects that discrete sources have on measurements of the sky intensity. Consider dividing all sources into classes of equal flux. Let $N_i$ be the number of objects having flux $S_i$. These objects are assumed to be distributed completely at random on the celestial sphere. A single observation of the sky intensity is made, using a detector with response characteristics $r(\theta,\phi)$. We divide the response into infinitesimal cells of equal solid angle $\delta\Omega$. The response in the jth cell is $r_j$. $n_{ij}$ is the number of objects from the ith class contributing to the intensity measurement in the jth cell. $n_{ij}$ is not a fixed number, but is a random variable. $n_{ij}$ is distributed as a binomial distribution of $N_i$ objects with a mean

$$\overline{n_{ij}} = N_i \; \delta\Omega \; / \; 4 \; \pi. \qquad\qquad [III.20]$$

If the solid angle is small enough and $N_i$ large enough the binomial distribution is well modeled by a Poisson distribution with the same mean.

A probability distribution is often described in terms of its cumulants (see Section A2:1). The first cumulant is the distribution's mean, the second its variance. The third and fourth cumulants are related to the skewness and kurtosis. For a Gaussian distribution all cumulants higher than the second are zero. The units for the nth cumulant of a distribution of a variable x are the units of x raised to the nth power. A common construct is the reduced cumulant,

$$\gamma_n \equiv \kappa_n / (\kappa_2)^{n/2} , \qquad\qquad [\text{III.21}]$$

where $\kappa_n$ is the nth cumulant and $\kappa_2$ is the second cumulant, the variance. The reduced cumulant is dimensionless. The size of the nth reduced cumulant, $\gamma_n$, when n is greater than two, measures the deviation of a distribution from a Gaussian. A complete definition of a distribution's cumulants and a discussion of how they can completely describe a distribution is given in Section A2:1.

The intensity of these $n_{ij}$ objects is

$$I_{ij} = S_i \ r_j \ n_{ij}. \qquad\qquad [\text{III.22}]$$

$I_{ij}$ is also a random variable. Its distribution can be given in terms of its cumulants; with the mth cumulant being:

$$\kappa_{I_{ij};m} = (S_i \ r_j)^m \ \kappa_{n_{ij};m} = (S_i \ r_j)^m \ \overline{n_{ij}} , \qquad\qquad [\text{III.23}]$$

where $\kappa_{n_{ij};m}$ is the mth cumulant of the distribution of $n_{ij}$. For the $n_{ij}$, which follow a Poisson distribution, all the higher cumulants are equal to the mean.

If we integrate over the response of the entire detector we can define $I_i$ to be the measured intensity for all sources from the ith class. If the distribution of these sources were completely random, i.e. there was no clumping or other correlation, and the total number of the sources $N_i$ is large enough that the binomial distribution was well modelled by the Poisson, then the contributions from the different

cells are statistically independent. The cumulants for $I_i$ are

$$\kappa_{I_i;m} = (R_m \, S_i)^m \, N_i \, , \tag{III.24}$$

since the cumulants for the sum of independent variables are just the sum of the cumulants. $R_m$ is the mth weighted response as defined in equation [II.17]. The mean intensity, $\bar{I}$, is proportional to $R_1$, while the variance is proportional to $R_2$. We can characterize the fluctuation in the intensity $I_i$ to first order by its fractional RMS size:

$$\sigma_{I_i}/\bar{I}_i = (R_2/R_1) \, N_i^{-1/2}. \tag{III.25}$$

Thus the larger the number of objects, the smaller the RMS noise, and the more the graininess of the signal is reduced.

For an example we consider the case of the ideal detector described for equation [II.21]. In the case of this ideal detector the RMS fractional noise is

$$\sigma_{I_i}/\bar{I}_i = (C \, N_i)^{-1/2}. \tag{III.26}$$

For detectors with smaller and smaller solid angles, represented by the sky coverage factor C, the fractional fluctuation is increased. However if we examine the absolute size of the fluctuations, we see that their magnitude increases as the detector solid angle increases. If the measurements of I are limited by an irreducible internal noise then a more accurate measurement of the fluctuations would require a larger value of C. In the case where the uncertainty in I is due to counting statistics, the ratio of the fluctuations variance to the photon variance is, for a single integration,

$$\sigma^2_{I_i} / \sigma^2_{photon} = S_i \, R_2^2 / R_1 \, . \tag{III.27}$$

For an ideal detector this ratio is independent of solid angle and proportional to $(A \, \Delta t)^2$, where A and $\Delta t$ are the detector area and integration time, respectively. So, although the fractional size of the fluctuations may be increased for smaller solid angle detectors, this by itself does not increase the statistical significance of a single measurement of the fluctuations. It is also worth noting that the signal-to-noise ratio of [III.27] is independent of the total number of

sources, but does depend on the class flux strength, $S_i$.

To find the total intensity we integrate over the various intensity classes by replacing $N_i$ with $N(S)$ $dS$. Again, as long as there is no correlation between objects of the different classes, the cumulants of the total intensity, I, follow directly from the defining integral:

$$\kappa_{I;m} = R_m{}^m \int dS \ S^m \ N(S). \qquad [III.28]$$

In principle, as the cumulants of a distribution contain all its information, we can test models for $N(S)$ by calculating their cumulants using [III.28] and comparing them with the estimated cumulants from a set of observations. Such an approach is somewhat restrictive as there is no established prescription for estimating the goodness of fit of a particular model $N(S)$, or for setting any confidence ranges on any parameter of such models. Worse, we see in what follows that for the very models that we expect to best describe the true behavior, the cumulants show awkward infinities.

### The cumulants for power law $N(S)$

We have seen that an expected form for $N(S)$ is a power law with an index of $-5/2$. We generalize this to models that have the form

$$N(S) \ dS = 4\pi \ K \ S^{-\gamma} \ dS, \qquad [III.29]$$

with the index, $\gamma$, taking any value. If such models are to hold over the full range of S from zero to infinity, the cumulants we calculate with equation [III.28] contain divergent integrals. The divergences can be removed by modifying the models to include cutoffs at both the high- and low-flux levels. Let $S_0$ be the lower sharp cutoff and $S_{max}$ be the high-flux limit. When the power law behavior covers many decades, the value of each cumulant will depend essentially on only one flux limit:

$$-\gamma + m + 1 < 0 \quad : \quad \kappa_{I;m} \propto K \ S_0{}^{-\gamma+m+1}; \qquad [III.30a]$$

$$-\gamma + m + 1 > 0 \quad : \quad \kappa_{I;m} \propto K \ S_{max}{}^{-\gamma+m+1}. \qquad [III.30b]$$

For all models with indices steeper than 2.0, the first cumulant diverges as $S_0$ approaches zero — Olber's familiar paradox again. But what is not so familiar, for models with indices flatter than three (e.g. the Euclidean model), is that the variance and all higher cumulants diverge with increasing $S_{max}$. The nearby high-flux sources dominate any description of the distribution, based on its cumulants, of the fluctuations from these models. Since these sources are relatively few in number, we would expect the shape of the distribution to be non-Gaussian; a Poisson distribution is poorly described by a Gaussian when the mean number of objects is small. The non-Gaussian nature of all these distributions is assured by the non-negligible values of the skewness and other higher reduced cumulants.

Now we know that there must be an $S_{max}$ supplied by nature so that the true cumulants of the observed distribution do not diverge. That limit is flux of the brightest observed object. A danger is that at these very highest limits the models for $N(S)$ might be ill-described by the power law that may suit the bulk of the sources. Indeed, the resolved high-latitude source counts as observed by Piccinotti et al. [1982] are somewhat steeper than the expected Euclidean model, in all probability due to local inhomogeneities in the source distribution or the inevitable uncertainties due to measurements based on a low number of objects. If we could exclude data that contain any high-flux sources we might arbitrarily set $S_{max}$ at a low enough value so that such inhomogeneities and uncertainties are not a problem, but in that case we must insure that no observations have any contribution from sources with fluxes greater than $S_{max}$.

Early characterizations of the fluctuations of the cosmic X-ray background [e.g. Schwartz et al. 1971; Fabian and Sanford 1971; Holt et al. 1974] followed the lead of the early characterizations of the isotropy of the microwave background. (A detailed example of this technique, applied to the microwave sky fluctuations, is found in Boynton and Partridge [1974].) In the microwave case, where the small-scale anisotropies were not expected to be due to a power law $N(S)$ distribution of microwave sources, the upper limit to a Gaussian

variance characterized the fluctuations. But in the X-ray case, a non-Gaussian distribution of the fluctuations was expected as soon as extragalactic sources were identified, when a possible model for the sky flux was unresolved point source emission. The first attempts to characterize the X-ray distribution consisted of estimating the excess variance in addition to the intrinsic variance due to counting statistics and variations in the non-X-ray counts. Both the preliminary upper bound published by Holt et al. [1974] and the actual detected excess variance measured by Schwartz, Murray, and Gursky [1976] from UHURU data showed that the fluctuations were smaller than expected for then-current estimates of the value of K in the N(S) law based on resolved sources in the UHURU catalogue. (However, see the discussion of these results in Section VI:6.) The resolved sources K values were subsequently revised downward in agreement with the fluctuations. However to rely on the variance to characterize the fluctuations is fundamentally limited in that it depends mostly on the behavior of N(S) very near the limit of resolved sources. In addition, the statistical tools ($\chi^2$) used by these authors for setting confidence limits on the variance (derived for studies of the microwave background) are only accurate to first order for non-Gaussian distributions. Using estimators of the higher cumulants is not much help. They are even more strongly determined by the behavior of the high-flux sources than is the variance.

If we are to extend our understanding much below $S_{max}$ we need an alternative approach. Such an approach was pioneered by radio astronomers, who first used it to characterize the "confusion noise" that unresolved sources produced on measurements of the intensity of low-flux objects [Scheuer 1957, Condon 1974]. The confusion noise is described by a probability distribution called a P(D) curve, ("P of D") for the deflection distribution. P(D) dD is the differential probability that an observation of sky intensity has a deflection, D, from the mean intensity. Even though the cumulants are not well-behaved, we will show that the general shape derived for the P(D) curve is well behaved, with the added benefit that much of the shape is generated by the behavior of sources at lower intensities than those

that dominate the variance. The suggested application of this technique
for the analysis of the X-ray sky was a natural one [Fabian 1972;
Scheuer 1974; Rowan-Robinson and Fabian 1974] and its first application
to UHURU data by Fabian [1975] not only repeated the message that the
resolved source counts were too high, but reported that the fluctuations
were roughly what was expected from a Euclidean power law. This
additional information is unattainable with studies restricted to the
determination of a single parameter such as the excess variance.

The remainder of this chapter is devoted to a derivation and
analysis of the P(D) technique. The ground work of this derivation was
laid by the previous authors, but because of the quality of the HEAO 1
data we will play close attention to aspects of the process of model
making that were ignored in earlier X-ray work.

A form of analysis midway in sophistication between the P(D)
analysis and the simple variance technique was used on data obtained
with the OSO 8 B detector. The crude results (reported in Shafer et al.
[1977]) for the magnitude of the fluctuations were slightly lower than
the earlier reported UHURU results, but still within the range of
consistency. However, the statistical aspects of the technique used
then were very primative, possibly incorrect. By the time that greater
understanding of the statistics and other portions of the the
fluctuations problem was acheived, the HEAO 1 all-sky survey data was
available. The OSO 8 B detector data was very clean, the imaginative
use of different window material on the two halves of the detector
allowed us to subtract out both the internal background and the possible
electron contamination. (The identical electron transmission of the two
window material allowed the electron removal, there was no propane veto
layer as was used in the HEAO HEDs.) However, the total area of the
detector was small, effectively ~10 cm$^2$ (the exact value depends on the
incident spectrum). In addition, though the detector was able to obtain
some very long exposures, the total sky coverage was limited. Large
portions of the satellite's lifetime were spent so that instruments
alligned parallel to the B detector could study sources in the galactic
plane. From the standpoint of area and sky coverage, the HEAO database

provides a better opportunity for the study of the fluctuations.

------------------------

### 3: Derivation of the Fluctuations Distribution Function

The calculation of the fluctuations distribution is a two-step process. The first step involves folding N(S) through the collimator spatial response to give the number versus intensity relationship. The second is to take this n(I) law and, using properties of the Fourier transform, derive the distribution of intensities.

As noted in the last section, the technique for deriving the distribution function of the fluctuations was pioneered by Scheuer [1957]. The use of Fourier transforms in the derivation was detailed by Scheuer [1974] and Condon [1974]. Though the technique described here is no different than what was used before, the notation has been changed somewhat. Originally, this was to make clear the distinction between a source flux S and its contribution to an intensity measurement I in the context of the fluctuations. The change in notation will also cleanly demonstrate the scaling relationships of the probability distributions for simple power law models for N(S). In addition, rather than presenting the "closed form" solutions of Scheuer and Condon we leave them in terms of Fourier transforms, which is sufficient for our data fitting program.

### Going from N(S) to n(I)

Our ultimate goal is to obtain a probability distribution $P_I(I)$ dI produced by a collection of point sources randomly distributed in the detector field of view. The first step is to calculate the mean number of sources with intensity I observed in the solid angle of the observation. In a derivation similar to equations [III.20] and [III.22] we consider a small patch of solid angle $d\Omega$ with response r. In this patch a source with flux S has an intensity I = rS. Then the mean

number of sources with intensity between I and I+dI is

$$n_{d\Omega}(I) \; dI \; = \frac{d\Omega}{4\pi} \; N(S) \; dS \; = \frac{d\Omega}{4\pi} \; N(I/r) \; d(I/r). \qquad [III.31]$$

To get the mean number for the entire exposure we integrate over the detector's solid angle:

$$n(I) \; dI \; = \int \; dr \; \frac{\nu(r)}{4\pi} \; N(I/r) \; r^{-1} \; dI \qquad [III.32]$$

(see equation [II.16]). In general such an integral would have no closed-form solution, and for every model N(S) would require a separate integration. One way to accelerate such a process is to note the similarity of equation [III.32] to a convolution, although instead of involving sums and differences (see Section A2:1), the integral involves multiplications and divisions of variables. Just as Fourier transforms provide an equivalent and computationally faster representation of a convolution, equation [III.32] may be recast in terms of Mellin transforms (see Section A2:3). These can in turn be rewritten as Fourier transforms in order to use the Cooley-Tukey algorithm and its variants, for fast Fourier transforms (e.g. Bergland [1969].

It is important to note the change in going from N(S) to n(I). There is no simple proportional way to completely identify objects of intensity I with objects of a single flux S, save in the simple case of the constant response ideal detector. The inverse problem, going from S to I, is also not simple but it is more straightforward. For a given source of flux S that is observed somewhere in the detector, the probability that it will contribute an intensity I is

$$P_{I(S)}(I;S) \; dI \; = \frac{\nu(I/S)}{4\pi \, C_0} \; \frac{dI}{S}. \qquad [III.33]$$

The coefficient $C_0$ is just the total fraction of the sky covered by the detector as defined in equation [II.20]. The mean value of I(S) is

clearly

$$\langle I(S) \rangle = S \; R_1 \; / \; C_o \; , \qquad\qquad\qquad\qquad [III.34]$$

i.e., S times the mean response of the detector, $R_1/C_o$ (see [II.17] and the discussion of the weighted detector responses). The variance around this mean intensity is

$$\sigma^2_{I(S)} = S^2 \; (R_2^2/C_o - R_1^2/C_o^2). \qquad\qquad\qquad [III.35]$$

For a source observed to have an intensity I, the probability that it has a flux S is proportional to the probability in equation [III.33] weighted by N(S). A priori knowledge that there are more low-flux sources than high is thereby reflected in $P_{S(I)}(S;I)$. It can be shown that when N(S) is a power law with index $\gamma$, the expected mean value of the flux of a source with intensity I is

$$\langle S(I) \rangle = I \; (R_{\gamma-1})^{\gamma-1} \; / \; (R_\gamma)^\gamma \; . \qquad\qquad\qquad [III.36]$$

Only for an ideal flat-spatial-response detector is this $\langle S(I) \rangle = I/A\Delta t$ $= I \; C/R_1$, the simple inverse of the relationship for the mean I(S). In other cases it is an error to assume that the flux S is given by I divided by the peak detector response (in our units, A $\Delta t$). This is true even for the unsmeared detector. The error is compounded if the detector is substantially smeared (see the discussion in Section II:2 on the effects of smearing).

### From n(I) to $P_I(I)$

From the distribution of source intensities we can find the probability distribution of the intensity of one of them drawn at random:

$$P_1(I_1) \; dI_1 = \mu^{-1} \; n(I_1) \; dI_1 \; , \qquad\qquad\qquad [III.37]$$

where $\mu$, the total number of sources in the distribution n(I), acts as a normalization constant.

If we wanted the probability distribution for the intensity of two

sources chosen at random and assuming no correlation:

$$P_2(I_2) \; dI_2 = \int_{-\infty}^{+\infty} dI_1 \; P_1(I_2 - I_1) \; P_1(I_1) \; , \qquad\qquad [III.38]$$

This convolution may be rewritten in the Fourier domain as

$$\mathbb{F}(P_2) = \mathbb{F}(P_1) \; \mathbb{F}(P_1) = \mathbb{F}(P_1)^2 \; . \qquad\qquad [III.39]$$

The Fourier transform of a probability distribution is often referred to as the characteristic function of the distribution. Thus this technique for deriving the fluctuations distribution is sometimes called the method of characteristic functions.

By extending the process with multiple convolutions we can find the Fourier transform of the intensity distribution of three, four, or an arbitrary number of sources; i.e., $\mathbb{F}(P_n) = \mathbb{F}(P_1)^n$. Of course each of these distributions applies to a case with a known (and fixed) number of sources drawn from $P_1$. Though this number is not known a priori, we know by reasoning parallel to the derivation of [III.28] that the number of sources is a random variable drawn from a Poisson distribution with mean $\mu$. In this case the distribution of the intensity I from all sources is

$$P_I(I) = \sum_n \frac{\exp(-\mu)}{n!} \mu^n \; P_n(I). \qquad\qquad [III.40]$$

By the linearity of the Fourier transform:

$$P_I(I) = \exp(-\mu) \; \mathbb{F}^{-1}\{ \; \sum_n \frac{\mu^n \; \mathbb{F}(P_1)^n}{n!} \; \} \qquad\qquad [III.41]$$

$$= \mathbb{F}^{-1}\{ \; \exp[\mu \; \mathbb{F}(P_1) - \mu \; ]\} \qquad\qquad [III.42]$$

$$= \mathbb{F}^{-1}\{ \; \exp[\mathbb{F}(n(I)) - \mu]\}. \qquad\qquad [III.43]$$

With the speed and convenience of Fast Fourier Transforms, equation [III.43] provides a total description of the fluctuations induced by the sources described by n(i). An alternative technique [Fabian 1972; Wall et al. 1982] is to provide a stochastic estimate of $P_I$ by using a random number generator first to determine the number of objects in an

observation and second to determine the intensities of these n objects from $P_1$. The details of this and the issues of an implementation of the Fourier transform technique are discussed in detail in Appendix A3.

------------------------

## 4: The P(D) curves for Simple Models of N(S)

> For better studies of the fluctuations we re-parameterize the power law models for N(S). This simplifies the problem of folding through the detector spatial response to get n(I). The distribution for a Euclidean power law is studied, contrasting the full distribution with analysis based only on the variance of the distribution. We present plots of the fluctuations for various values of $\gamma$, the power law index, comparing them with the Euclidean form. The importance of sharp cutoffs in N(S), particularly for certain values of $\gamma$, is presented, along with an outline of how to include them as a simple extension of power law models. Finally we consider variants of the single power law, including models whose behavior is fixed at a certain flux, as well as general modifications to low-flux behavior represented as a Gaussian term.

### P(D) for single power law N(S)

Armed with the general prescription for deriving a P(D) curve, we will apply it to a specific class of simple models, the power law models:

$$N(S)\ dS = 4\pi\ K\ S^{-\gamma}\ dS \equiv (S/W)^{-\gamma}\ d(S/W). \qquad [III.44]$$

The second half of the equation defines a reparameterization of N(S) in terms of the parameter W. One of the virtues of W is that it has simpler units than K, W and S having the same units, so that in comparing two different experiments that use different units of flux, W transforms the same as S. It will also turn out that the width of the $P_I$ curve depends linearly on W. The virtue of K, besides its almost universal use in other published studies of N(S), is that N(S) is proportional to K, allowing for linear combinations of two populations with the same value for $\gamma$. That is, given two populations with K values

of $K_1$ and $K_2$, their net $N(S) = (K_1 + K_2)\ S^{-\gamma}\ dS$. The conversion between $K$ and $W$ is

$$4\pi\ K = W^{\gamma-1}\ . \qquad\qquad [III.45]$$

Inspection of this relationship demonstrates that the re-parameterization in terms of $W$ is inappropriate when $\gamma$ equals one.

Equation [III.32], which calculates $n(I)$, has a particularly simple form for pure power law $N(S)$ models, i.e. ones with no cutoffs at high or low-flux:

$$n(I)\ dI = \int\ dr\ (4\pi)^{-1}\ \nu(r)\ (I/Wr)^{-\gamma}\ (Wr)^{-1}\ dI\ . \qquad [III.46]$$

The $I$ dependence can be moved outside the integral, indicating that $n(I)$ is a power law with the same index as $N(S)$. Only the coefficient has changed. We can simplify this equation by making the following definitions:

$$W_{eff} \equiv R_{\gamma-1}\ W\ , \qquad\qquad [III.47]$$

where $R_{\gamma-1}$ is the '$\gamma-1$'th weight of the detector response defined in [II.17]. As $R_{\gamma-1}$ has the same units as the detector response, the units for $W_{eff}$ are the same as the units used to measure intensity, i.e. counts per exposure. This allows us to define the dimensionless parameter $\xi$:

$$\xi \equiv I\ /\ W_{eff}\ . \qquad\qquad [III.48]$$

We can recast the $n(I)$ distribution into a $n_\xi(\xi)\ d\xi$ distribution. Equation [III.46] simplifies to

$$n_\xi(I\ /\ W_{eff})\ d(I/W_{eff}) = (\ I\ /\ W_{eff}\ )^{-\gamma}\ d(I/W_{eff})\ , \qquad [III.49]$$

or

$$n_\xi(\xi)\ d\xi = \xi^{-\gamma}\ d\xi\ . \qquad\qquad [III.50]$$

$n_\xi$ can be used in equation [III.43] to derive $P_\xi(\xi)$, the probability that an observation will be of intensity between $\xi$ ($= I/W_{eff}$) and $\xi+d\xi$. The shape of $P_\xi$ is dependent only on the value of $\gamma$, as there is no $W$, or $W_{eff}$, dependence. To convert $P_\xi$ to the intensity-probability distribution involves the simple scaling law:

$$P_I(I)\ dI = P_\xi(I/W_{eff})\ d(I/W_{eff})\ . \qquad\qquad [III.51]$$

Thus we need perform the involved calculation of Fourier transforms only

once for each value of $\gamma$.

The literature on the fluctuations uses a different, though equivalent, notation:

$$n(I) \ dI = \Omega_{eff} \ K \ I^{-\gamma} \ dI \ . \hspace{4cm} [III.52]$$

The factor $\Omega_{eff}$ serves to convert the coefficient K from S units to I units, as $R_{\gamma-1}$ was used to convert W to $W_{eff}$. One can be expressed in terms of the other by

$$\Omega_{eff} = 4\pi \ (R_{\gamma-1})^{\gamma-1} \ . \hspace{4cm} [III.53]$$

This shows that $\Omega_{eff}$ has, through the factor of $4\pi$, units of solid angle and is often referred to as the detector's "effective solid angle" for fluctuations from power law N(S) models. But the equation shows that $\Omega_{eff}$ also has units of $(cm^2 \ s)^{\gamma-1}$ from the response units of $R_{\gamma-1}$. Thus to refer to the conversion factor from N(S) to n(I) as just an effective solid angle obscures the difference between S and I which, in turn, encourages the erroneous practice of attributing fluctuation effects caused by sources with observed intensity I to those caused by sources with flux $S = I \ / \ r_{max}$. That is, it assumes that all sources are contributing as if they were at the peak of the response, rather than scattered throughout the detector field of view. (Recall the discussion of equation [III.33].)

With our model $n_\xi$ and the algorithm of the last section we now seem prepared to calculate and plot the probability distributions. Before we do this, however, we should note that we have not yet put any lower or upper limits on the range of the power law. Without such limits we are led to infinite values of the cumulants (see equation [III.30]) and since the cumulants are directly related to the probability distribution (see Section A2:1), we appear to be ignoring considerable difficulties. The change of representation from S to $\xi$ is no help:

$$\kappa_{\xi;m} = \int d\xi \ \xi^{-\gamma+m} \hspace{4cm} [III.54]$$

is equally dependent on the limits of integration for finite values. Although the cumulants are formally infinite without these limits, the

shape of the distribution will be fairly insensitive to the cutoff values for $\gamma$ near 5/2 and not near two or three.

The infinity of the first cumulant, the mean, is handled simply by removing the cumulant from consideration. If we are not interested in the absolute distribution of $P_\xi(\xi)$, we can instead investigate the distribution of measurements about the mean by defining

$$\xi' \equiv \xi - \bar{\xi} \; . \qquad\qquad\qquad\qquad\qquad\qquad\qquad [III.55]$$

Even as $\bar{\xi}$ tends to infinity due to more and more low $\xi$ sources, the distribution of $P_{\xi'}(\xi - \bar{\xi})$ is well behaved. In moving from $\xi$ to $\xi'$ we have lost any chance of making an absolute correspondence between our expected distribution of intensities and the observations. For the intensities we have $P_{I'}(I - \bar{I})$, where by analogy $I' = I - \bar{I}$. $\bar{I}$ is now a free parameter, independent of the particulars of the model for $n_I$. We have only a distribution for the fluctuations about some mean. This is actually an advantage over an absolute distribution as there may be sources and processes contributing to the count rate I but not included in the $n_I$ model, e.g. internal background or a truly diffuse component (not associated with the unresolved sources). It is tempting to identify $\bar{I}$ with the mean observed sky count rate, but this must be done carefully or unintentional bias will be introduced when comparing different models.

The higher cumulants, for $\gamma < 3$, diverge due to the high-flux sources. A plot of $P_{\xi'}(\xi')$ is shown in Figure III.3 for a Euclidean model, $\gamma = 5/2$. The non-Gaussian nature of the distribution is evident. The peak probability is offset substantially from the mean and there is a long tail trailing off to positive infinity. This tail asymptotically approaches the form $(\xi')^{-\gamma}$. Measurements far out on the tail are dominated by the presence of a single, relatively intense source. The probability of such occurrences is roughly proportional to $P_1(\xi)$, which in turn is proportional to $n(\xi)$.

It is this long tail due to the most intense sources that leads to infinite values for the variance and the higher cumulants. The bulk of

---

Figure III.3: Euclidean model fluctuations distribution

$P_{\xi'}(\xi')$ versus $\xi'$ for $\gamma = 5/2$. Because the model for N(S) is a simple power law, the only detector dependence is in the scaling parameter $W_{eff}$.

---



$$\xi' = \frac{I - \bar{I}}{W_{eff}}$$

---
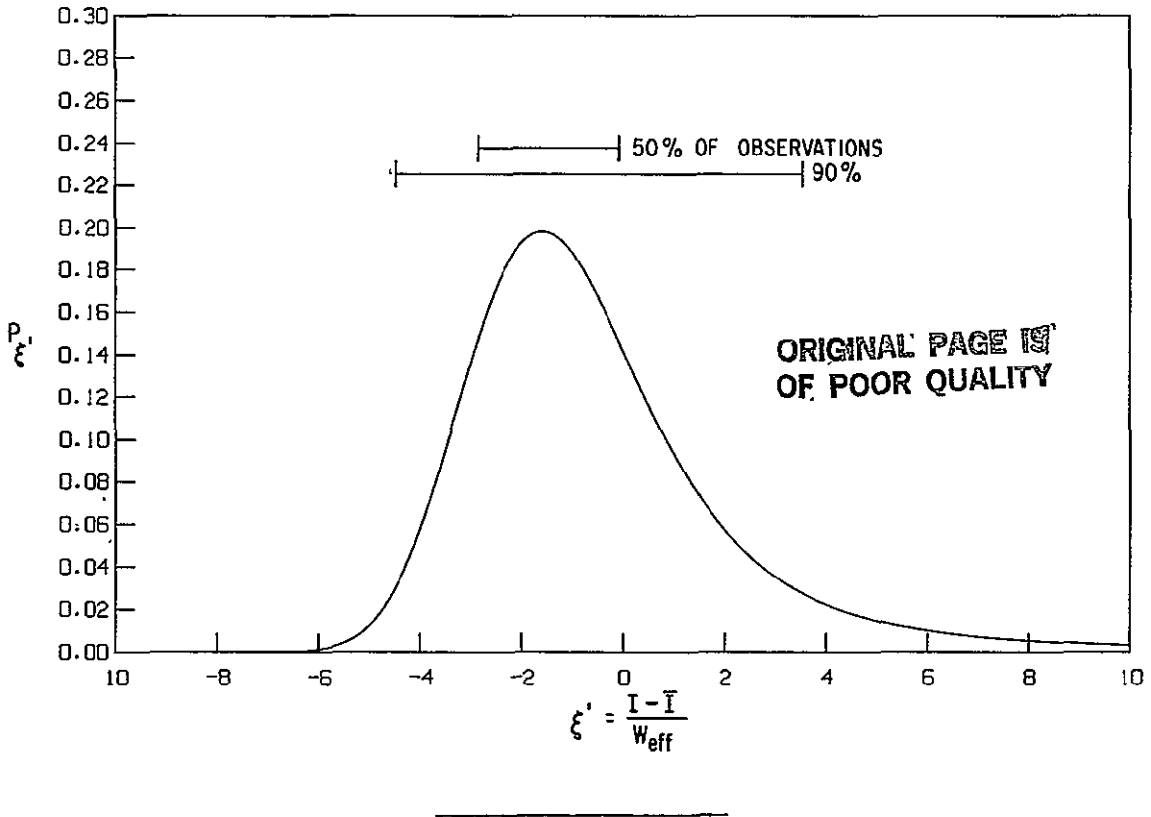
the observations are far removed from that region, and appear to be approximately described by a skewed Gaussian. We can describe the size of the fluctuations in terms of the smallest range in $\xi'$ containing a given fraction of the observations. Fifty per cent of the observations are in the range $-2.78 < \xi' < -0.01$. Thus, as well as the occasional high fluctuation, there are a large number of observations lying below the mean. The 90% range is more symmetric about the mean since the long tail begins to have an effect: $-4.46 < \xi' < +3.51$. To make a comparison with better-behaved distributions, which do not suffer from a formally infinite variance, we note that for a Gaussian the equivalent 90% range is from $-1.645$ $\sigma$ to $+1.645$ $\sigma$. We can characterize the $\gamma$ equals $5/2$

Fluctuations Theory                                                    P(D) for Simple Models

$P_{\xi'}(\xi')$ curve as having an "effective sigma" of ~2.4. For the intensity distribution, $P_{I'}(I')$, this is a size of ~2.4 $W_{eff}$ counts per exposure. $W_{eff}$ is therefore of the magnitude to characterize the size of the fluctuations and it is appropriate to refer to $W_{eff}$ as the "effective width" of the $P_{I'}$ distribution.

## Observations of the fluctuations variance

To restore finiteness to the cumulants of $P_{I'}$, the sample can be restricted to observations below a certain limit, $I'_u$. With this upper-intensity limit the variance becomes finite and has been used as the measurement of the fluctuations [Condon 1974; Schwartz, Murray, and Gursky 1976]. Obviously the exact value of the variance will depend on the chosen $I'_u$. It is frequently set at the point above which a high positive fluctuation would be identified with high confidence as the detection of a single source with intensity $>$ ~$I'_u$. The standard practice has been to connect measurement of this truncated variance, $\sigma^2_{I'}(I'_u)$, with models for N(S) via equation [III.54] (using m=2 for the variance):

$$\sigma^2_{I'}(I'_u) \approx (3-\gamma)^{-1} W_{eff}^2 (I'_u/W_{eff})^{-\gamma+3} . \qquad [III.56]$$

This measurement is dominated by the behavior of $n_I$ in the region very close to the cutoff $I'_u$. One might constrain $n_I$ at lower intensities using the fluctuations by choosing a suitably low value for $I'_u$.

However, equation [III.56] is only an approximation, which is usually valid for $I'_u \gg W_{eff}$. This approximation equates the cutoff in the distribution of sky intensity observations at $I'_u$ to the truncation of the integral in the source distribution integral in [III.54] at a source intensity $I = I'_u$. The approximation is valid when most of $I'$ is due to a single bright source, i.e. far up the tail of the $P_{I'}$ curve. At lower values of $I'$ the measurement may be increasingly affected by the behavior of sources other than the single brightest one. Clearly at the very lowest values of $I'$, where $I' \ll W_{eff}$, equation [III.56] breaks down, predicting vanishing variance even though the distribution of observations with $I'<0$ still has a significant extent.

Figure III.4: $\sigma^2_{\xi}{'}(\xi'_u)$ versus $\xi'_u$, $\gamma = 5/2$

The dashed line is the approximation of equation [III.56].
The solid lines are the actual values for cases with
varying amounts of non-fluctuations noise: $\sigma_{noise}/W_{eff} = $
0, 1, 2, and 4. The y axis is the variance $\sigma^2_{\xi}{'}$, which is
$\sigma^2_{I}{'}$ divided by $W_{eff}^2$.
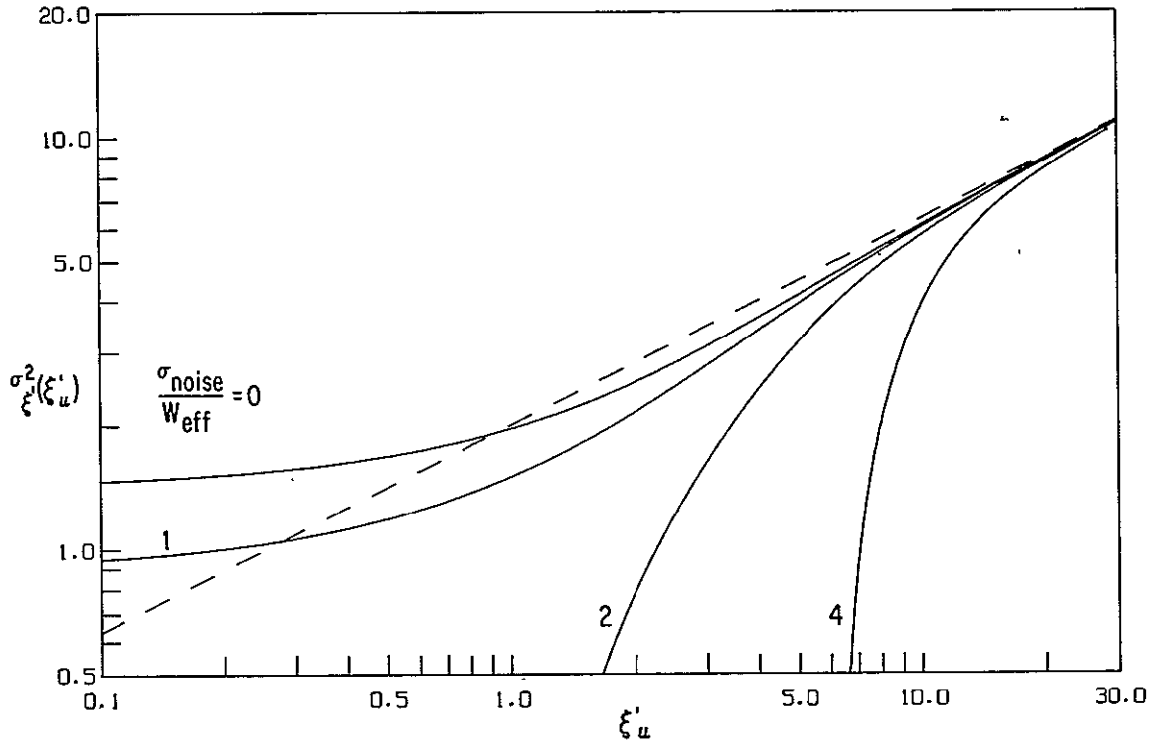


The characterization of $P_{I}{'}$ simply by its variance is further
complicated because observations of $I'$ are not exact, but are instead
measured in the presence of noise such as photon counting statistics or
an inherent variance in the non-X-ray background. Let the observed
count rate be $I_{obs}$, with a probability distribution, $P_{obs}(I_{obs})$, found
by convolving $P_{I}{'}$ with the noise distribution. It is correct to first
order, and certainly convenient, to represent the noise distribution as
a Gaussian with variance $\sigma^2_{noise}$. If $\sigma^2_{obs}(I'_u)$ is the measured
variance of observations with $I_{obs} < I'_u$, then for $I'_u \gg \sigma_{noise}$, an

estimate of $\sigma^2_{I'}$ is

$$\sigma^2_{I'} \approx \sigma^2_{obs} - \sigma^2_{noise} .$$                    [III.57]

However, as seen in Figure III.4, reducing $I'_u$ to $\sim$ a few $\sigma_{noise}$ causes this approximation to break down. When the upper limit is placed this low, the variance of the noise distribution can no longer simply be added to the modeled variance of the fluctuations to get the variance of the truncated set of observations.

Of course with some suitably sophisticated programs one could explicitly calculate $\sigma^2_{obs}$ as a function of $W_{eff}$ and $\gamma$, thus eliminating the need for both this approximation and the one in [III.56]. $I'_u$ could then be arbitrarily low but having gone to this much computational trouble we could just as well fit the model parameters explicitly using the entire shape of the $P_{obs}(I_{obs})$ curve, and not just from its variance. Doing the former is based more evenly on the behavior of $P_{I'}$ throughout the distribution, and not just at the high $I'$ tail where most of the variance is generated.

### Non-Euclidean power law models

Power law models where $\gamma$ is not equal to 5/2 can be easily generated for $2 < \gamma < 3$. As seen in Figure III.5, as $n(I)$ becomes flatter and flatter, the distribution becomes more and more skewed, and the fluctuations are increasingly dominated by the bright sources in the tail. For steeper $\gamma$, the distribution becomes more symmetrical and Gaussian as the numerous low intensity sources dominate. Note the percentage of total probability bars in Figure III.5. Not only does their skewness change with $\gamma$, but their magnitude is subject to slight changes. Thus we should remember that the identification of $W_{eff}$ as a characteristic width is good but not exact, or at least the identification depends on the value of $\gamma$.

Because of the analytic properties of these power law models for $n(I)$, direct solutions of $P_{I'}$ can be found in terms of an integral equation [Condon 1974] or a somewhat slowly convergent power series expansion [Scheuer 1974]. However, both of these solutions fail when $\gamma$

Figure III.5: $P_{\xi'}(\xi')$ versus $\xi'$, non-Euclidean models

(a) $\gamma = 2.1$. (b) $\gamma = 2.9$. The bars indicate the range for 50% and 90% of the observations. The dashed curve is the $\gamma = 5/2$ Euclidean case for comparison.
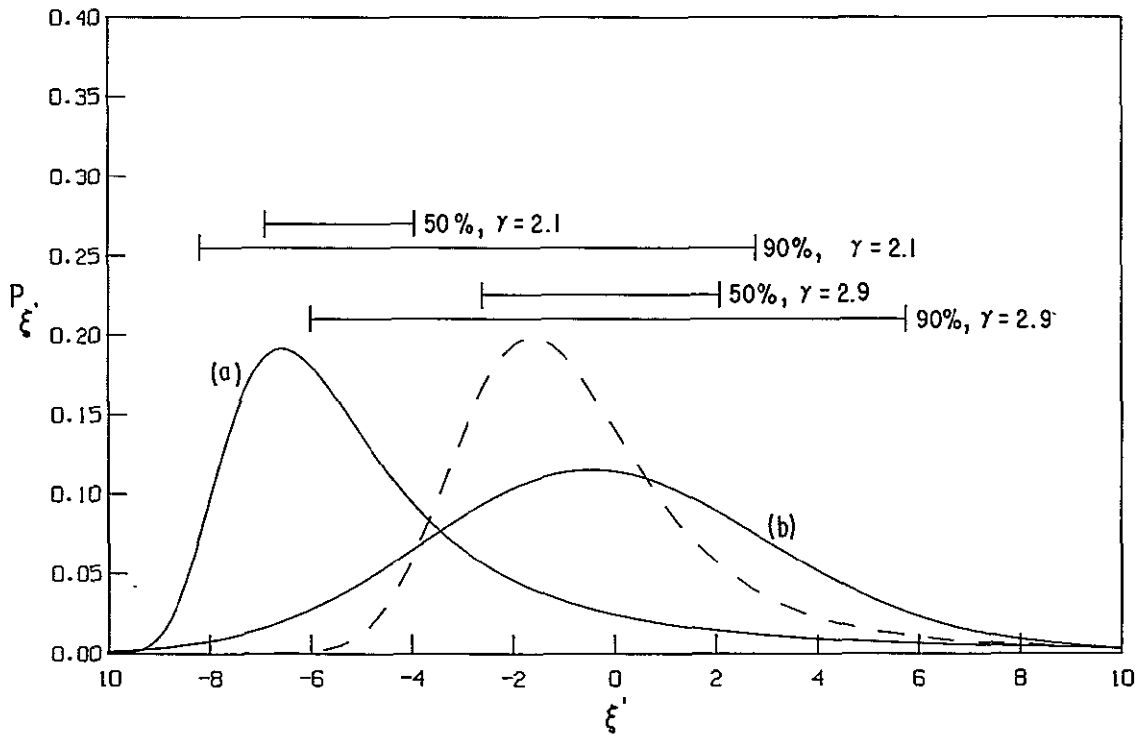


is outside of the range from 2 to 3. The reasons for this failure provide insight into the general process of calculating the $P_I$, distribution.

In Figure III.6(a) the trend for $\gamma \to 2.0$ is displayed. Essentially the peak of the distribution is being moved farther and farther from the mean. At $\gamma = 2$, the peak is infinitely far from the mean. This is the point where most of the mean for the distribution no longer comes from the bulk of the low intensity sources, but is dominated by the high intensity sources. So although $\gamma < 2$ is a solution to Olber's paradox, the mean is no longer infinite from the low-flux sources, and the utility of the "renormalization" of equation

Fluctuations Theory                    P(D) for Simple Models

Figure III.6: $P_{\xi'}(\xi')$ versus $\xi'$, extreme non-Euclidean models

(a) Solid: $\gamma = 2.02$, dashed: $\gamma = 2.1$.  (b) Solid: $\gamma = 2.98$, dashed: $\gamma = 2.9$.
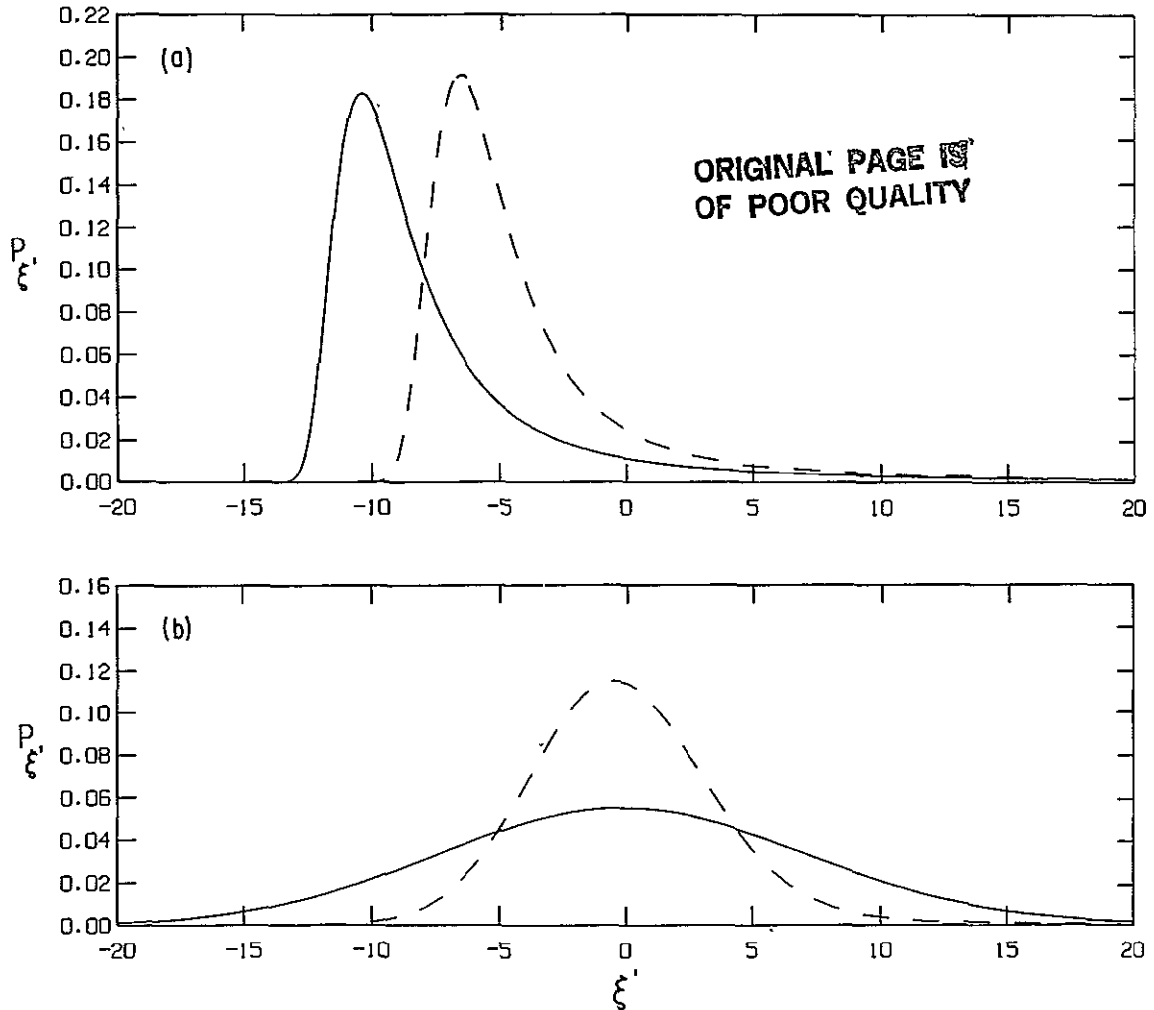


[III.55] is destroyed. The solution needs to include a cutoff in the high-flux sources. It turns out that the principal effect of such a high-flux cutoff is in the mean. For suitably high cutoffs any changes in the shape of the $P_{\xi'}$ curve are only significant at very high values of $\xi'$. For the models we actually use, the behavior of $P_{\xi'}$ in this region will not affect the fits of the models to the data.

Fluctuations Theory                                    P(D) for Simple Models

To briefly review the effect of the cutoff on the mean, we introduce a dimensionless parametrization of $N(S)$:

$$N_\zeta(\zeta) \, d\zeta = (\zeta)^{-\gamma} \, d\zeta, \qquad\qquad [III.58]$$

where

$$\zeta \equiv S \, / \, W \qquad\qquad\qquad\qquad\qquad . \quad [III.59]$$

by analogy to the dimensionless form of $n(I)$ using $\xi$. The mean is calculated as if $N_\zeta$ extends up to a sharp cutoff at $\zeta_u$. For $\gamma > 2.0$ the actual distribution <u>need</u> not have any cutoff. However, we have the <u>a priori</u> knowledge that there are no bright sources above a certain limit, say $S_{max}$. If we define $\zeta_{max} \equiv S_{max}/W$, the absence of sources above $S_{max}$ adjusts the mean by

$$\Delta\overline{I} = (2 - \gamma)^{-1} \, R_1 \, W \, \{ \, \zeta_{max}^{-\gamma+2} - \zeta_u^{-\gamma+2} \, \}. \qquad [III.60]$$

Using a finite value for $\zeta_{max}$ has a sizeable impact on the position of the mean.

In Figure III.7 we consider a somewhat idealized case. The detector used is the 1st layer of the HED 1, small field of view. As the adjustment to the mean, $\Delta\overline{I}$, is proportional to $R_1$ the correction is independent of any smearing of the detector. The dashed lines have the cutoff at 5 counts s$^{-1}$ cm$^{-2}$ (essentially infinite). The solid lines were made with a cutoff at the lower value of 0.025 counts s$^{-1}$ cm$^{-2}$. This corresponds roughly to the maximum observed flux of a high-latitude extragalactic source. For $\gamma$ near 2.0 the shift is substantial, and although reduced for the Euclidean case, it is still noticeable. This presages some difficulties with determination of the proper value of $\overline{I}$ to be used in the fits, which are considered in more detail later.

Condon and Dressel [1978] show that the integral solution [Condon 1974] for $P(D)$ can be extended to power law models with $\gamma$ less than 2. In this case the technique yields the distribution in terms of absolute intensities, $P_I$, rather than relative to the mean, $P_{I'}$.

Figure III.7: Effects of upper source cutoff on $P_\xi$,

(a) $\gamma = 2.02$, Solid: cutoff at $S = .0254$ counts $s^{-1}$ $cm^{-2}$, Dashed: cutoff at $S = 5.0$. (b) $\gamma = 5/2$, Solid and dashed as in (a).

ORIGINAL PAGE IS
OF POOR QUALITY
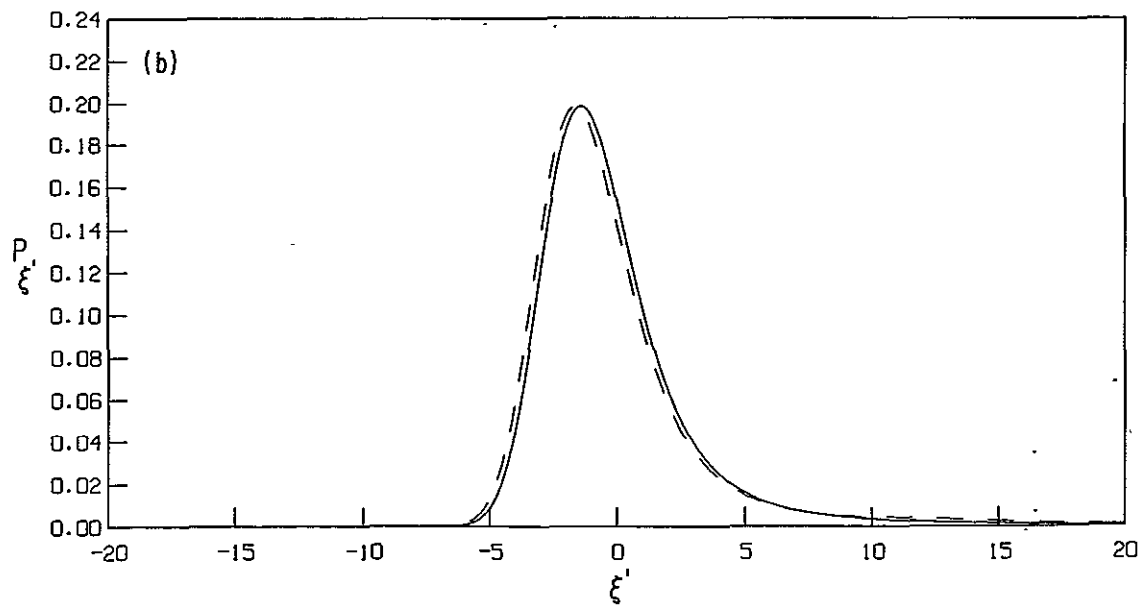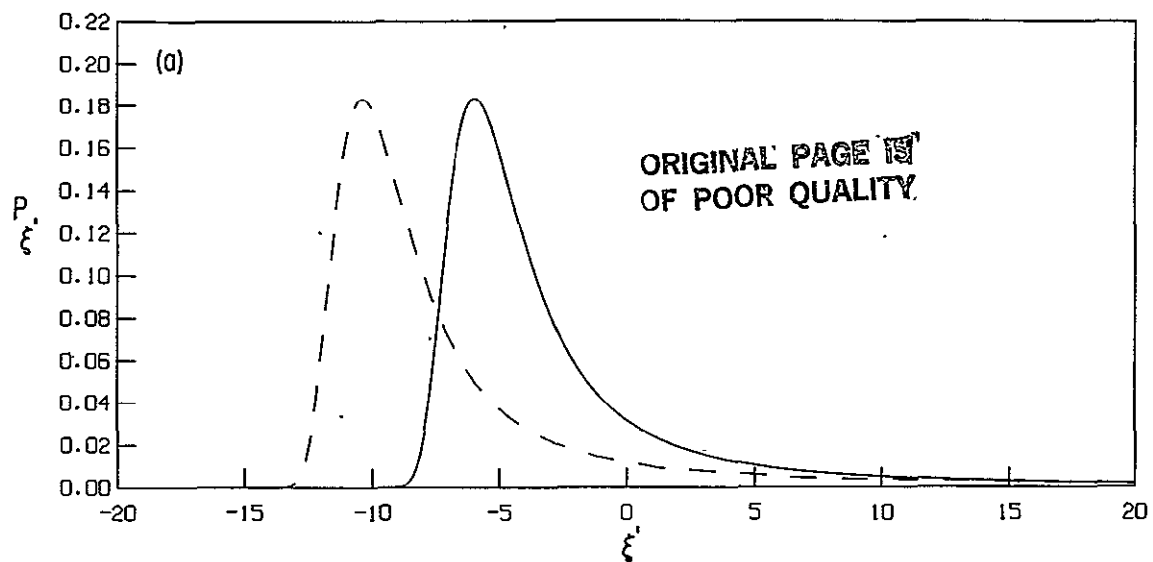
In Figure III.6(b) we plotted the trend as $\gamma$ approaches 3. The situation in this case has a similar flavor to that for $\gamma \to 2$. Now it is not the mean but the second cumulant, the variance, that is troublesome. As the power law steepens, the low-flux sources contribute more to the variance overtaking the high-flux sources in importance at $\gamma$ = 3. Unlike the high-intensity sources, whose contribution to the variance was through the high-intensity tail of the fluctuations' distribution, low-flux sources contribute to the variance through an ever widening Gaussian component. Renormalization might suggest itself, (this was how the infinite mean from low-flux sources was handled), but all the information of the model would be lost. For $\gamma > 3$, any finite value of $W_{eff}$ produces an infinitely wide Gaussian distribution unless a low-flux cutoff is imposed. This cutoff could affect more than just the variance, so we consider this addition to the model in more detail.

First let us consider an interesting trend. The problems at $\gamma$ near 2 and 3 stem from a shift in the flux regions that dominate the mean and the variance. One might expect that similar problems occur as $\gamma$ nears 4, where the third cumulant becomes dominated by the large number of low-flux sources rather than the high-flux tail. Does this presage a shift away from the highly Gaussian distributions when $\gamma$ is near 3? Actually, using a low-flux cutoff to avoid problems with the variance, we solve all future problems with higher cumulants. This is because the shape of the $P_{I'}$ curve is determined by the <u>reduced</u> cumulants, <u>i.e.</u> normalized by the variance. It is possible to show that even for very steep forms of $n(I)$ the reduced cumulants are vanishingly small. Thus except for the remnants of the effect of the high-flux tail, the $P_{I'}$ curve is well represented by a Gaussian for large values of $\gamma$.

## Power law $N(S)$ models with low-flux cutoffs

The effect of a sharp cutoff in $N(S)$ on $n(I)$, the source counts folded through the detector, was explicitly considered by Scheuer [1974], although he analyzed only the case of an un-smeared truly pyramidal detector response. We will consider the general solution for

any detector response function. We modify the model with a Heaviside function:

$$N(\zeta) \, d\zeta = \zeta^{-\gamma} \, \theta(\zeta - \zeta_0) \, d\zeta \qquad\qquad [\text{III.61}]$$

The low-flux cutoff is at $S_0 = W \, \zeta_0$. We parameterize in terms of the dimensionless $\zeta_0$ in order to maintain our procedure of using the Fourier technique to generate a $P_{\xi'}(\xi')$ once for a range of values for W. Folding through the detector using equation [III.32] gives

$$n_\xi(\xi) \, d\xi = \phi(\xi;\gamma,\zeta_0) \, \xi^{-\gamma} \, d\xi, \qquad\qquad [\text{III.62}]$$

where $\phi$ is the correction term due to the cutoff:

$$\phi(\xi;\gamma,\zeta_0) = \frac{\int dr \, \nu(r) \, r^{\gamma-1} \, \theta(\xi R_{\gamma-1}/r - \zeta_0)}{4\pi \, (R_{\gamma-1})^{\gamma-1}} . \qquad\qquad [\text{III.63}]$$

Note that $\phi$ is dimensionless. The Heaviside function in the integral can be rewritten as $\theta(\xi R_{\gamma-1}/\zeta_0 - r)$. From the definition of $R_{\gamma-1}$ (see equation [II.17]) we can see that $\phi(\xi \gtrless \xi_c)$ is identically 1, where the critical value of $\xi$ is

$$\xi_c = \zeta_0 . r_{max} / R_{\gamma-1} . \qquad\qquad [\text{III.64}]$$

$r_{max}$ is the maximum response of the detector. For an ideal flat-response detector, $\phi$ drops to zero below $\xi_c = \zeta_0 \, C^{-1/(\gamma-1)}$. For more realistic detectors, $\phi$ gradually rolls off to zero.

The formalism of equations [III.61] − [III.63] may be generalized to any functional form for a cutoff. Instead of the Heaviside function in [III.61] we use the general form $\Phi(\zeta;\gamma,\alpha,\beta,...)$, where $\alpha$, $\beta$, ... are dimensionless model parameters like $S_0/W$ in the above sharp-cutoff example. Then the definition of $\phi(\xi;\gamma,...)$ in [III.63] will have $\Phi(\xi R_{\gamma-1}/r;\gamma,...)$ in place of the sharp cutoff.

### Minimum cutoff from diffuse flux limit

In order to find the fluctuations distribution for steep power laws, $\gamma \gtrsim 3.0$, there must be a non-zero $\xi_c$. Just as the high-flux cutoff is automatically imposed by the observed lack of very bright sources, so is a minimum value of $S_0$ invoked to avoid saturation of the

Figure III.8: Effects of low-flux cutoff on $P_{\xi'}(\xi')$

(a) $\gamma = 2.98$. Dashed: no low intensity cutoff. Solid: low intensity cutoff set by diffuse sky limit for HED 1 unsmeared small FOV. (b) $\gamma = 5/2$. Dashed and solid lines defined as for (a).



ORIGINAL PAGE IS OF POOR QUALITY
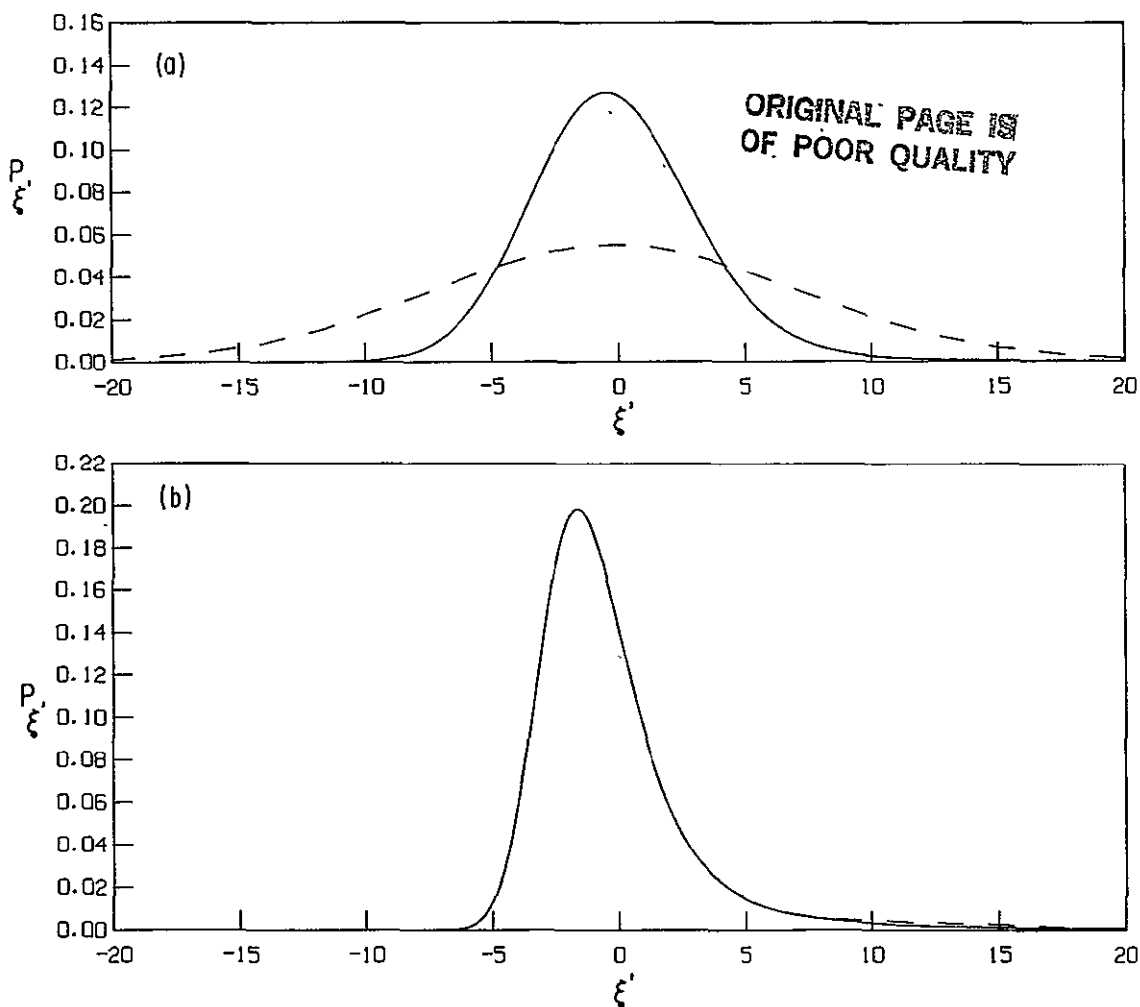
diffuse background. If the X-ray all-sky flux is $S_{as}$, then

$$S_o > W \; [ \; \xi_u^{-\gamma+2} + (\gamma-2) \; S_{as}/W \; ]^{1/(2-\gamma)} \; .$$

[III.65]

The inequality shows that this is a minimum required cutoff. If some fraction of $S_{as}$ were due to another population of sources or to a truly

Fluctuations Theory                                              P(D) for Simple Models

diffuse component, then the cutoff should be higher. Whether this minimum cutoff is sufficiently large to impact the form of the $P_{I'}(I')$ curves depends on $\gamma$ and the particulars of the measurement, especially the solid angle of the detector. Unfortunately the position of this cutoff cannot be made an invariant dimensionless quantity. Because $S_{as}$ is fixed, $\zeta_0$ must be adjusted when $W$ is varied for the models. So when we use the minimum cutoff, we will have to recalculate the shape as well as the size of $P_{I'}$ for every value of $W$.

Fortunately, although the cutoff is required for steep values of $\gamma$, it is at most a second order effect for near-Euclidean values. In Figure III.8 we apply the specific case of the HED 1 Layer 1 unsmeared small field of view. $S_{as}$ is approximately 60 counts $s^{-1}$ $cm^{-2}$. The value for $W$ used was ~0.06 counts $s^{-1}$ $cm^{-2}$, which is in good agreement with the resolved sources $N(S)$, as well as what we shall derive later from the fluctuations. When $\gamma$ is 2.98, the change is dramatic. The curve still looks very Gaussian, but its variance has been slimmed down noticeably. The critical value of $\xi$ is $\xi_c = 2.5 \times 10^{-4}$. This is actually a very low number, but for steep power laws, it is noticeable, mainly in terms of the missing variance. We will see that an indication of the significance of a cutoff to the shape of $P_\xi$, is the size of $\xi_c$. If $\xi_c$ is less than ~0.1, then the cutoff has little effect other than the missing variance; i.e. it is as if a purely Gaussian term were removed from the distribution.

For a Euclidean model $S_0$ is very low, ~$2.4 \times 10^{-7}$ counts $s^{-1}$ $cm^{-2}$, by virtue of the more gentle slope. In Figure III.8(b) the dashed and undashed curves are nearly identical. $\xi_c$ is ~$4 \times 10^{-6}$. Not even the variance is much missed. For models with $\gamma$ near 5/2 the minimum value of $S_0$ can be ignored.

"Pegged" models

The observed behavior of resolved sources is in agreement with the Euclidean model. The variations of the full power law model, with parameters $W$ and $\gamma$, that we will consider are "pegged" power law models.

That is, the basis for these models is the assumption that N(S) below the limit of observation joins smoothly to the known N(S) for resolved source counts. Although continuity is not strictly a requirement of N(S), it is expected any time N(S) is derived from a smoothly varying luminosity function. Let $S_p$ be the "peg" point, i.e. the flux where the resolved source and the model N(S) meet. Let the resolved N(S) be represented by a power law with index and width $\gamma_p$ and $W_p$. If the model is a power law below $S_p$, with some $\gamma$ and $W$, then for a particular value of $\gamma$ we have fixed $W$ to be

$$W = W_p^{[(\gamma_p-1)/(\gamma-1)]} \, S_p^{[(\gamma-\gamma_p)/(\gamma-1)]} \, . \qquad [III.66]$$

We will assume that $S_p$ is high enough so that the only contribution to the fluctuations from sources with $S > S_p$ is far up the high flux tail. The assumption is these sources have a significant effect only on the mean, and not on the shape of $P_\xi$,.

### First order modifications to simple power laws

Although a simple power law suffices to describe the resolved sources and although we might expect it to continue for some range in flux, an accurate model of N(S) will sooner or later involve details not represented by that single power law. The eventual roll-over will not be a sharp cutoff. There might be another population of sources. Evolution may produce a large number of very dim objects. Of course one can use any specific model to derive a $P_I$, curve that can be compared with observations of the fluctuations but this may be unnecessarily complicated. The vast bulk of models are well represented by a two component model: a single power law with or without a sharp low-flux cutoff; plus an additional number or a deficit of dim sources. We assume the additional component makes a significant contribution to N(S) only in the range of fluxes where there are already numerous sources. In this range the contribution of the new sources is essentially an additional Gaussian. They contribute to first order, only to the variance of the fluctuations.

More precisely, we split the total $N(S)$ into two parts:

$$N(S) \; dS = dS \; \{ \; W^{-1} \; (S/W)^{-\gamma} \; + A(S) \} \; .$$

[III.67]

$A(S)$ could be in principle a positive or negative quantity. We assume that the power law model given by $W$ and $\gamma$ give a good description of the higher flux sources which dominate the positive tail of the fluctuations. The impact of the function $A(S)$ is reduced to the single value:

$$\sigma^2_S = \int \; dS \; S^2 \; A(S) \; .$$

[III.68]

$\sigma^2_S$ will be referred to as the <u>invariant</u> excess. If it is negative, the $A(S)$ term is acting as a deficit, reducing the Gaussian part of $P_{\xi'}$ in a way similar to a sharp cutoff. Although $\sigma^2_S$ is not in itself a variance, it can be used to form an excess (or deficit) variance:

$$\sigma^2_\xi = (R_2 \; / \; R_{\gamma-1} \; )^2 \; \sigma^2_S \; / \; W^2 \; .$$

[III.69]

To obtain the net distribution, a Gaussian with this variance is convolved with the power law derived from $P_{\xi'}(\xi')$. One can also reverse the process, fitting an excess variance in $\xi'$ space and then using equation [III.69] to calculate the invariant excess. The invariant excess is more useful than the excess variance in that it can be used to test various forms of $A(S)$ using [III.68] directly, without referring to a particular collimator size.

The various power law models with the additional modelling parameter of the invariant excess cover the range of model behavior investigated in this dissertation.

———————————————

### 5: Comparisons between N(S) models

We derive a general tool, the likelihood difference, to tell how similar a test distribution is to the (presumed) "true" distribution. The relationship of this tool to the likelihood statistical tests allows us, when testing a parametric family of models, to estimate which range of parameters will, on average and for a given number of observations, be outside acceptable confidence contours. We

apply the tool to pure Gaussian distributions, and to the fluctuation models. The likelihood difference provides a quantifiable indicator of the sensitivity of the fluctuations to details of models for N(S). The ability to distinguish two models for the fluctuations is degraded in the presence of non-fluctuations noise. We find that the behavior of the lowest intensity sources can be well modeled as a purely Gaussian contribution.

In considering an ensemble of models, described by several parameters, it is important to ask just how "different" any two of these models are. Would we expect, given data of a certain quality and quantity, that we could distinguish between them? For models which are simple Gaussians, there are several statistical tools available to answer these questions. In our case, though, where the models have distinctly non-Gaussian aspects, we are led to develop a general indicator of "differentness", the likelihood difference. In particular, we are interested in how significant the behavior of the low intensity souces are for measurements of the fluctuations.

## Likelihood and likelihood difference

Consider a set of observations, $x_i$, drawn from a parent probability distribution $P_o(x) \, dx$. We have a model for these observations which we wish to study, represented by the distribution $P(x) \, dx$. We define the likelihood (or sometimes the log-likelihood) that an observation $x_i$ was drawn from $P(x)$ as

$$\lambda(x_i; P(x)) = -2 \ln P(x_i) \, . \qquad\qquad [III.70]$$

The smaller the value for $\lambda$, the more likely the distribution-observation relationship is said to be. As we will later see, the likelihood is used in fitting models and drawing confidence contours for model parameters. For now we need only note that the closer the model is to the true distribution, the lower the expected value of $\lambda$. The likelihood difference, $\Delta\lambda(P(x))$, is the mean value of the difference

between the model likelihood and the true distribution likelihood:

$$\Delta\lambda(P(x)) \equiv \langle \lambda(x;P(x)) - \lambda(x;P_0(x)) \rangle .$$

[III.71]

The angle brackets, $\langle \rangle$, indicate the expectation value of the quantity they contain as defined by the base distribution $P_0(x)$:

$$\Delta\lambda(P(x)) = \int dx \, P_0(x) \, [\lambda(x;P(x)) - \lambda(x;P_0(x))] .$$

[III.72]

As an example, let us use data drawn from a simple Gaussian distribution with mean $x_0$ and variance $\sigma_0^2$. The model distributions considered are also Gaussians with mean $\bar{x}$ and variance $\sigma^2$. The likelihood of such models is simply

$$\lambda(x_i;\bar{x},\sigma) = \frac{(x_i - \bar{x})^2}{\sigma^2} + \ln(2\pi \sigma^2) .$$

[III.73]

Except for the second, logarithmic, term this looks like the contribution of the ith measurement to a $\chi^2$ sum, strengthening the perception that $\lambda$ has relevance to goodness of fit. For the case where the model has the correct variance, the likelihood difference is

$$\Delta\lambda(\bar{x}) = (\bar{x} - x_0)^2 / \sigma^2 .$$

[III.74]

So the difference between two models is just the number of sigma, squared, separating their means, an answer in accord with our intuition. The likelihood difference is a minimum, and zero, only when $\bar{x} = x_0$.

Slightly less familiar, but more relevant to our studies of the fluctuations, is the case where the mean of the model is known to be correct but the variance is a parameter to be determined. Now

$$\Delta\lambda(\sigma^2) = \frac{\sigma_0^2 - \sigma^2}{\sigma^2} + \ln(\sigma^2/\sigma_0^2) .$$

[III.75]

This is still zero and minimized when $\sigma^2 = \sigma_0^2$. When $\sigma^2 \approx \sigma_0^2$, the

likelihood difference is

$$\Delta\lambda \approx \delta^2 \quad ; \quad \delta = \frac{\sigma^2 - \sigma_o^2}{\sigma_o^2} \, .$$

[III.76]

The numerical value of $\Delta\lambda$ has real significance. If the total number of observations is N, then, a model with m fitted parameters will be rejected (on average) at the 90% level when

$$N \, \Delta\lambda \gtrsim \chi^2_{m;.90} - m \, ,$$

[III.77]

where $\chi^2_{m;.90}$ is the 90% confidence upper limit of a variable from a chisquare distribution with m degrees of freedom. The demonstration of this statement is given in Section V:2. It is an approximate relationship, most valid in the limit of large N. We can rearrange the relationship in [III.77] to find roughly how many observations, N, are needed to exclude a model.

Applying these concepts to the case of the fluctuations, we first examine the case where $\gamma$ is 5/2, and try to see how $\Delta\lambda$ is distributed as a function of $W_{eff}/(W_{eff})_o$, where $(W_{eff})_o$ is the true effective width. We define the fractional difference of the two models' widths as

$$\delta_W = [W_{eff} - (W_{eff})_o] \, / \, (W_{eff})_o \, .$$

[III.78]

Note that the fractional difference for W is the same as for $W_{eff}$ since the two are proportional to each other as long as $\gamma$ is fixed. In Figure III.9(a) we plot $\Delta\lambda(\delta_W)$ versus $\delta_W$ for a $\gamma$ = 5/2 Euclidean model. With only about 60 observations with no extraneous noise we would expect a 90% confidence range in $\delta_W$ to be roughly the same size as the range bar shown in the figure. The exact values for the limits of the confidence range may be different (in fact, by definition there is a 10% chance of excluding the $\delta_W$ = 0 correct model) but the expected size of the range in is about ±10%.

Thus it seems that even with a relatively small number of observations the fluctuations are a powerful tool. In comparison, the precision of measurements based on the A-2 sample of ~60 resolved

Figure III.9: The likelihood difference, $\Delta\lambda(\delta_W)$ versus $\delta_W$

The outer y-axes are for $\Delta\lambda$, while the central axis is the number of observations needed to place, on average, models outside the 90% confidence range. (a) A log versus linear representation. The solid line is for positive $\delta_W$, the dashed for negative. A case for $\gamma = 5/2$ and no noise is plotted. The bar indicates the range of values that on average would not be excluded from the 90% range when there are 60 observations. (b) A log versus log representation. The positive and negative halves have been folded over onto this one plot. The 4 pairs of lines are for different amounts of noise, $\psi = \sigma_{noise}/W_{eff}$ : 0 (top/left pair), 2, 3, and 4 (bottom/right pair).
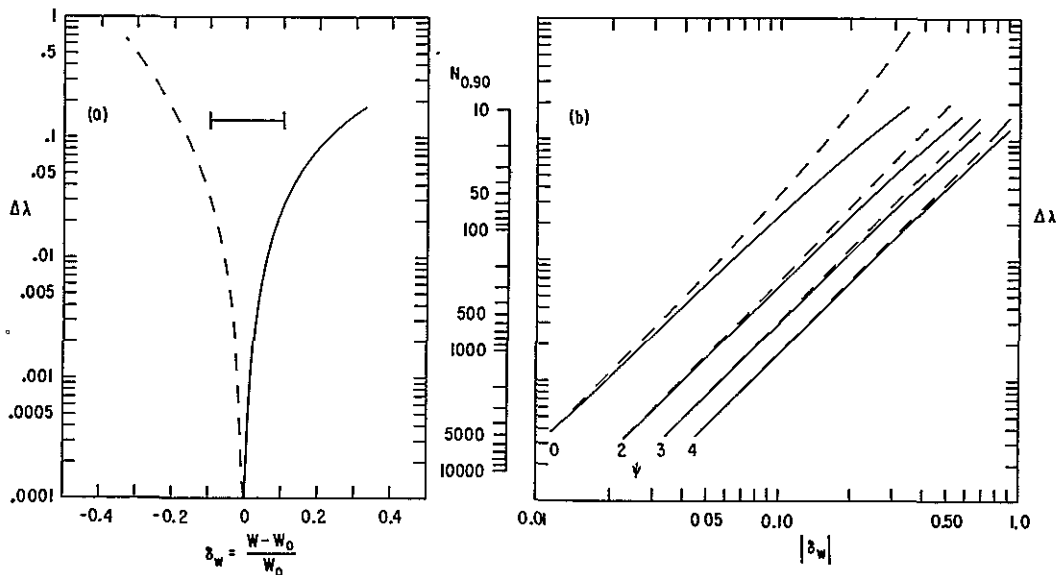


sources is about 17% [Piccinotti et al. 1982]. If we can increase the number of fluctuations measurements from 60 to hundreds we should do proportionately better.

The catch, of course, is that we do not have totally precise measurements of the fluctuations' intensity. As mentioned in the discussion of Figure III.4 in the previous section, noise can have a significant impact on the interpretation. We will continue to assume that this noise, which is due to counting statistics and uncertainties

Figure III.10: $P_{\xi'}(\xi')$ versus $\xi'$ with noise

The dashed line is for $\psi = 0.0$, no noise. The solid lines are for $\psi = 1.0$ (narrowest), 2.0, and 4.0 (the widest). The bars are the range containing 50% of the observations for the various solid line models.
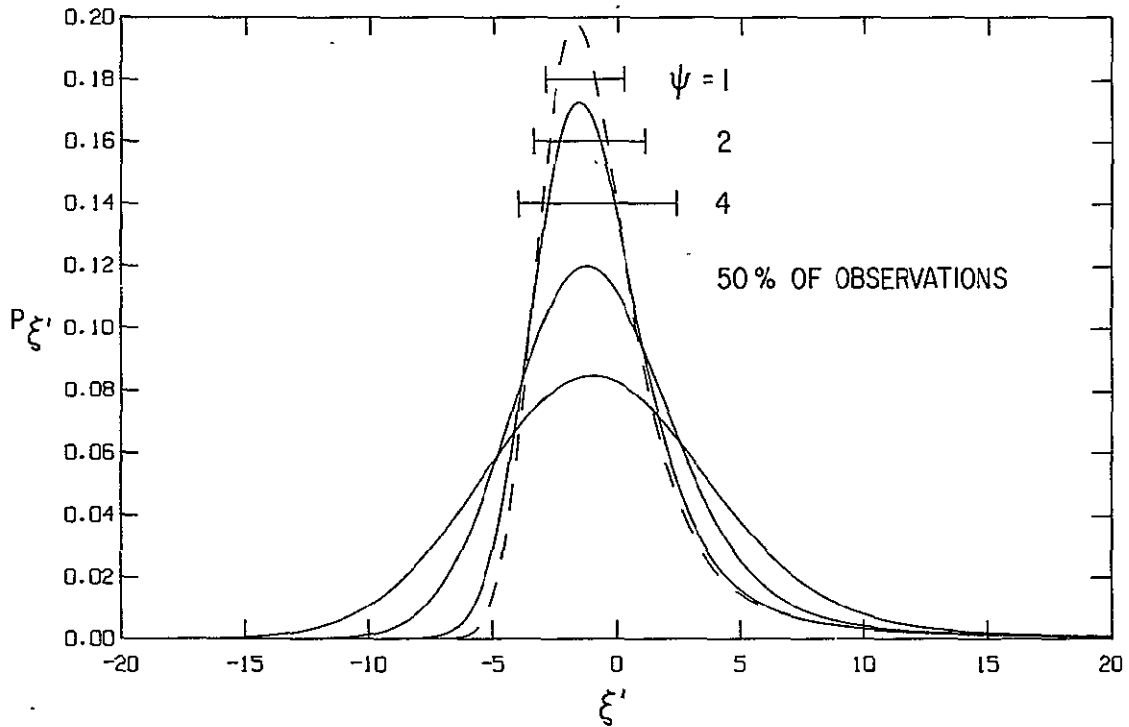


in the non-X-ray contributions to the signal, can be well represented by a Gaussian distribution, which we will parameterize by

$$\psi = \sigma_{noise} / W_{eff} \ . \qquad \qquad [III.79]$$

When $\psi$ is much less than about 1, the basic no-noise $P_\xi$, curve is a good approximation. When $\psi$ is much greater than 1, the curve looks more and more Gaussian with a variance equal to about $(\psi_f)^2 + \psi^2$ where $\psi_f$ is a constant $\sim 1$ to represent the intrinsic uncertainty contribution from the fluctuations (Figure III.10). At very high fluxes the power law tail will dominate the Gaussian noise, but this could be as high as the known source intensities if $\psi$ is large enough.

In Figure III.9(b) we plotted $\Delta\lambda(\delta_W)$ versus $\delta_W$ for a variety of values of $\psi$. The plot has been made log-log, with the negative branch of $\delta_W$ shown as a dashed line. Examination of the slopes of the lines will show that to first order

$$\Delta\lambda(\delta_W) \approx 2.9 \, \delta_W^2 \, / \, (1.1 + \psi^2) \, . \qquad\qquad [III.80]$$

In the presence of noise two models become more indistinguishable as their likelihood difference drops. However, given enough observations one can expect the confidence ranges to get arbitrarily small. It should be noted that for a fixed number of observations, we can reduce the expected size of the confidence range by reducing the noise. However there is an **irreducible** size to the confidence range, even for observations with no noise, generated by the fluctuations themselves.

### Limits on the fluctuations

In principle, an exact description of the fluctuations' distribution, $P_I(I)$, would totally determine the shape of $N(S)$ dS for all values of S, (given certain assumptions, e.g. that the sources are not clumped). However to successfully make such a total determination of $N(S)$ would require not only ideal noise-free observations, but an infinite number of them. In reality we are constrained by both finite-area imperfect detectors and the firm limit of $4\pi$ steradians in the sky which limits the number of independent measurements for a finite solid angle detector. The fluctuations are not equally generated by all sources, and so our determination of models for $N(S)$ will be dominated by a particular range of source fluxes.

The literature on the analysis of fluctuations, in particular Scheuer [1974], often suggests that the information concerning n(I) derivable from the fluctuations about n(I) extends only down to the intensity level where there is approximately one source per field of view. Scheuer's argument begins with the observation that a "typical" fluctuation is about 1.7 times the intensity of the brightest source in a typical beam. His estimated value of 1.7 is the expected variance due to all sources with intensity less than that of the brightest source. He argues, "Evidently the numerical constant (1.7) is of the order of 1

and not sensitive to the form of [n(I)] so long as [the variance] converges" which always happens when $\gamma$ is less than 3. This argument is correct but incomplete. The expected value of $\xi_1$, the value of the normalized intensity for the brightest source in the field of view, is

$$< \xi_1 > = (\gamma-1)^{-1/(\gamma-1)} \quad . \qquad\qquad\qquad \text{[III.81]}$$

For Euclidean models this is an intensity $\sim 0.76 \ W_{eff}$, so that 1.7 times this, or $\sim 1.3 \ W_{eff}$ is certainly the right magnitude for a typical fluctuation. Unfortunately, this begs the question of what intensity sources produce this typical fluctuation, and it implies we can ignore the variance of those sources with intensities less than $\xi_1$. Suppose there is a sudden cutoff in $n_\xi$ at $\xi_c$. The total variance from sources from $\xi_c$ up to $\xi_1$ is

$$\sigma^2{}_{\xi_c \to \xi_1} = (3-\gamma)^{-1} \ \xi^{3-\gamma} \ \Big|_{\xi_c}^{\xi_1} . \qquad\qquad \text{[III.82]}$$

The absence of any sources below $\xi_c$ will cause a deficit in the variance compared to the case where $n_\xi$ extends all the way to zero:

$$\frac{\Delta(\sigma_\xi{}^2)}{\sigma^2{}_{0 \to \xi_1}} = (\xi_c/\xi_1)^{3-\gamma} \ , \qquad\qquad\qquad \text{[III.83]}$$

whenever $\gamma < 3$. For the Euclidean model, the missing fractional variance goes like $(\xi_c/\xi_1)^{1/2}$. It is true that most of the variance comes from sources near $\xi_1$, but the square root dependence on $\xi_c$ means that a large fraction comes from the low intensity sources. This missing variance could be detected given enough observations, no matter how low the value of $\xi_c$. A rough indication of the likelihood difference between a model with a sharp cutoff at $\xi_c$ and one without a cutoff could be found by using the approximation of [III.76].

Rather than use any approximation, we will actually use the likelihood difference to tell us how sensitive we are to distinguishing a model with a sharp cutoff from a reality with no cutoff, i.e. if the data is drawn from a fluctuations distribution where N(S) continues unbroken to zero (ignoring Olbers paradox) then the likelihood

Figure III.11: $P_{\xi'}(\xi')$ versus $\xi'$, with a sharp cutoff in N(S)

$\gamma = 5/2$, see text for details of detector spatial response. (a) $\psi = 0.0$ (no noise). Solid: $\xi_c = 0.5$, dashed: $\xi_c = 0.0$. (b) $\psi = 2.0$. Solid: $\xi_c = 0.5$, 2.0, dashed: $\xi_c = 0.0$.
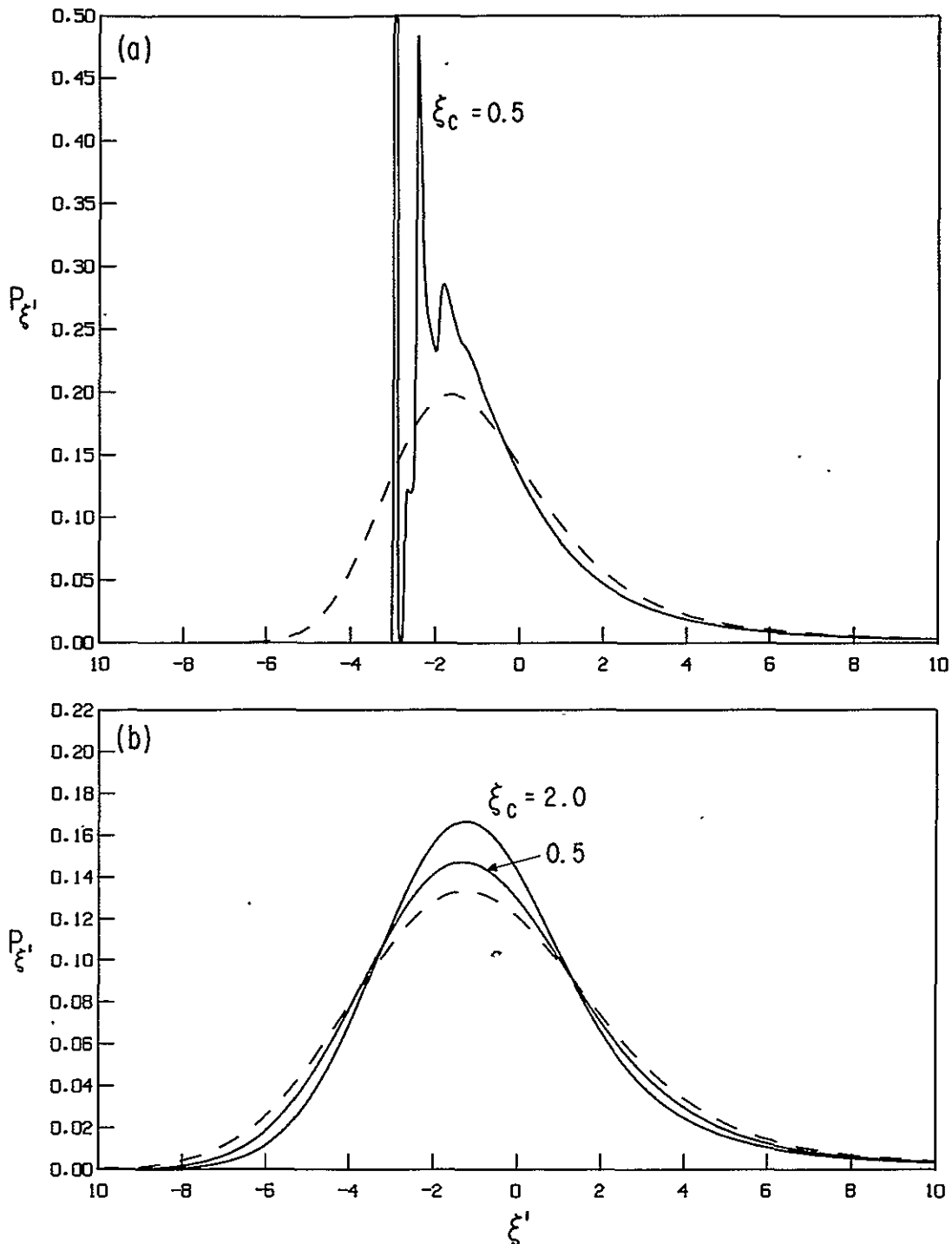
difference will indicate which models with sharp cutoffs are expected to
be outside the confidence range. In the previous section we saw that
the form of $n(\xi)$ for a sharp cutoff in $N(S)$ at low $S$ depended on the
detector spatial response through $\nu(r)$. Let us first consider the case
of a highly smeared detector, where the amount of smear in both
directions is much greater than the FWHM collimator sizes. This is a
good approximation to our flat-topped ideal detector. In Figure
III.11(a) we use this detector to plot $P_{\xi'}$ with a sharp cutoff at $\xi =$
0.5 versus $P_{\xi'}$ with no cutoff. We immediately see a striking difference
between the curves, which should be trivial to detect with only a few
observations. The single isolated sharp peak on the left for the $\xi_c =$
0.5 curve is actually a delta function, representing the significant
probability, ~0.15, of no sources at all being in the field of view.
All observations for this case have the same intensity. The strong bump
and decay to the right of the delta function cover the curve when there
is only one source in the field, the next bump when two sources are
present, and a slight break exists when three sources are in the field
of view. The regular spacing between these features occurs because the
minimum source intensity is 0.5. (We expect these features, in a true
flat-response collimator, to have very sharp rises, but here they are
smeared by the approximation to a flat collimator.) In Figure III.11(b)
we see that most of this structure is washed out in the presence of even
moderate amounts of noise, $\psi = 2.0$. The presence of the cutoff in $n_{\xi}$ is
now indicated only by the missing variance.

In Figure III.12 we plot the likelihood difference assuming that
the no-cutoff case is the true situation. The no-noise, $\psi = 0$, curve
indicates that cutoffs as low as 0.1 can be ruled out at the 90% level
with as few as 60 observations. This is about one seventh the level of
one source per field of view. In fact this cutoff corresponds to about
21 sources in the field of view. When noise is significant, the
likelihood difference is sharply reduced. When $\psi = 2$, we need a tenfold
increase in the number of observations to be able to exclude a $\xi_c = 0.1$

Figure III.12: Likelihood difference versus sharp $\xi$ cutoff, $\gamma = 5/2$

Reality assumed to be $\xi_c = 0$ (no cutoff). $\psi = 0$, 2, 3, 4.
Righthand axis gives the number of observations so that
the model will on average be outside the 90% confidence
contour. A flat spatial response detector is used.



ORIGINAL PAGE IS
OF POOR QUALITY
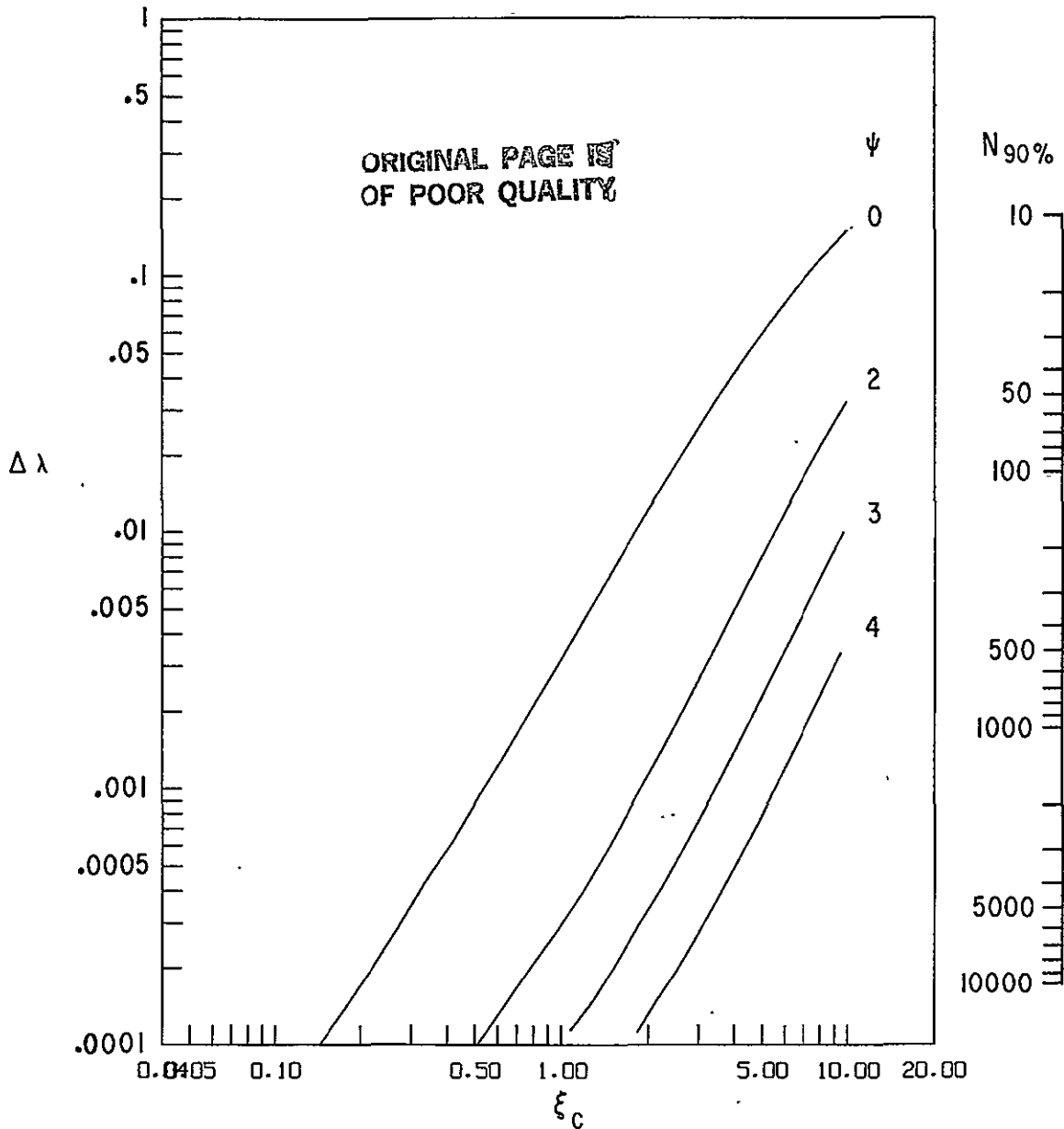
model. An examination of the figure shows that over a large range in

$\xi_c$, the likelihood difference is

$$\Delta\lambda(\xi_c) \approx 0.9 \; \xi_c \; / \; ((1.5)^2 + \psi^2)^2 \; . \qquad\qquad [\text{III.84}]$$

The fluctuations still provide an irreducible source of noise, now acting as if it has an effective "sigma" of ~1.5 $W_{eff}$. Again, with sufficiently many observations a confidence region of arbitrary width is obtainable, depending on the behavior of n(I) far below the level of the average brightest source per field of view.

The application of these limits to realistic data requires a value of $\psi$. The size of the noise variance is mostly a function of detector area, solid angle, and total exposure time per measurement. The last criterion is rarely constant for all measurements in a dataset. The A-2 data we will use for the fluctuations analysis has a typical value of $\psi$ ~ 1.2, although there is a fair amount of dispersion. Previous analyses of the fluctuations used data with $\psi$ in the range from 1.5 to 4.5 (see Section VI:6). The combination of low $\psi$ and several hundred observations make the A-2 data the current best all-sky dataset for the study of the fluctuations.

### Gaussian and non-Gaussian aspects

Notwithstanding the arguments above, it is unfortunately true that the fluctuations are relatively insensitive to the exact form of $n_\xi(\xi)$ for low-$\xi$ sources. Suppose there are n objects with reduced intensity $\xi$. They can be thought of as providing a modified Gaussian with variance $n \, \xi^2$. We can conceptually divide into two parts the contribution of these sources to the form of the fluctuations: a pure Gaussian and a portion that represents the non-Gaussian aspects of the Poisson distribution. As the number of objects, n, increases, equivalent to lower $\xi$ for power law distributions, the Poisson nature of their contribution to the fluctuations is represented better and better by a Gaussian.

What was the dominant aspect of the missing sources when we sharply cut off a power law distribution? In other words, for what range of values of $\xi$ is it acceptable to model the source contribution

only in terms of a Gaussian with the proper variance? In the $\gamma=5/2$ case we saw that over a large range the likelihood difference was proportional to the position of the cutoff, $\xi_c$. For this value of $\gamma$, the variance of these missing sources is $\propto (\xi_c)^{1/2}$ (see equation [III.83]). Thus there is a strong similarity between equations [III.84] and [III.76]. Most of the likelihood difference is due to the missing variance.

This suggests a way to show roughly where the non-Gaussian aspects of the fluctuations from low flux sources become important. We do this using the likelihood difference of a model with a sharp cutoff versus a parent distribution with no cutoff, but in addition we include in the model a purely Gaussian contribution of the missing sources using equation [III.82] for the range from $\xi = 0$ to $\xi_c$. The effect of this Gaussian distribution is to smooth the sharp peaks and the delta function of the $\xi_c = 0.5$, no noise, curve in Figure III.11(a). The likelihood difference drops to $\sim 9 \times 10^{-4}$ from a value greater than 1 when all the low flux sources contribution to the fluctuations are removed. Thus in order to be sensitive to the non-Gaussian aspects of the missing sources we need $\sim 2000$ observations, even in the absence of noise ($\psi = 0$). As plotted in Figure III.13, $\sim 400$ observations are needed to be sensitive to the non-Gaussian nature of sources below a $\xi_c$ of 1. As noted before, adding significant Gaussian noise makes the $P_\xi$ curves look essentially Gaussian even with very high cutoffs. We see in Figure III.13 that even for sources at $\xi$ as high as 5, their non-Gaussian contribution is unimportant for $\psi = 2$ unless there are over 300 observations.

This specific example can be extended to more general commentary on arbitrary $n_\xi(\xi)$ distributions. A similar likelihood difference curve would result if instead of cutting off the distribution at $\xi_c$ we doubled $n(\xi < \xi_c)$. In fact any additional component to $n(\xi)$ that is significant only below $\xi \sim 1$, when added to the power law $\xi^{-5/2}$ base, will have as its dominant effect an additional contribution to the Gaussian portion of the fluctuations, particularly in the presence of any non-fluctuations noise.

Fluctuations Theory                                        Comparing models

ORIGINAL PAGE IS
OF POOR QUALITY

Figure III.13: Sensitivity to non-Gaussian aspects of $P_{\xi'}$, measured by the likelihood difference

$\gamma = 5/2$, flat detector, reality assumed to have $\xi_c = 0$. Now the variance due to sources below the cutoff $\xi_c$ is included but not their contribution to the higher cumulants. Note the different range of the $\xi$ scale versus Figure III.12.
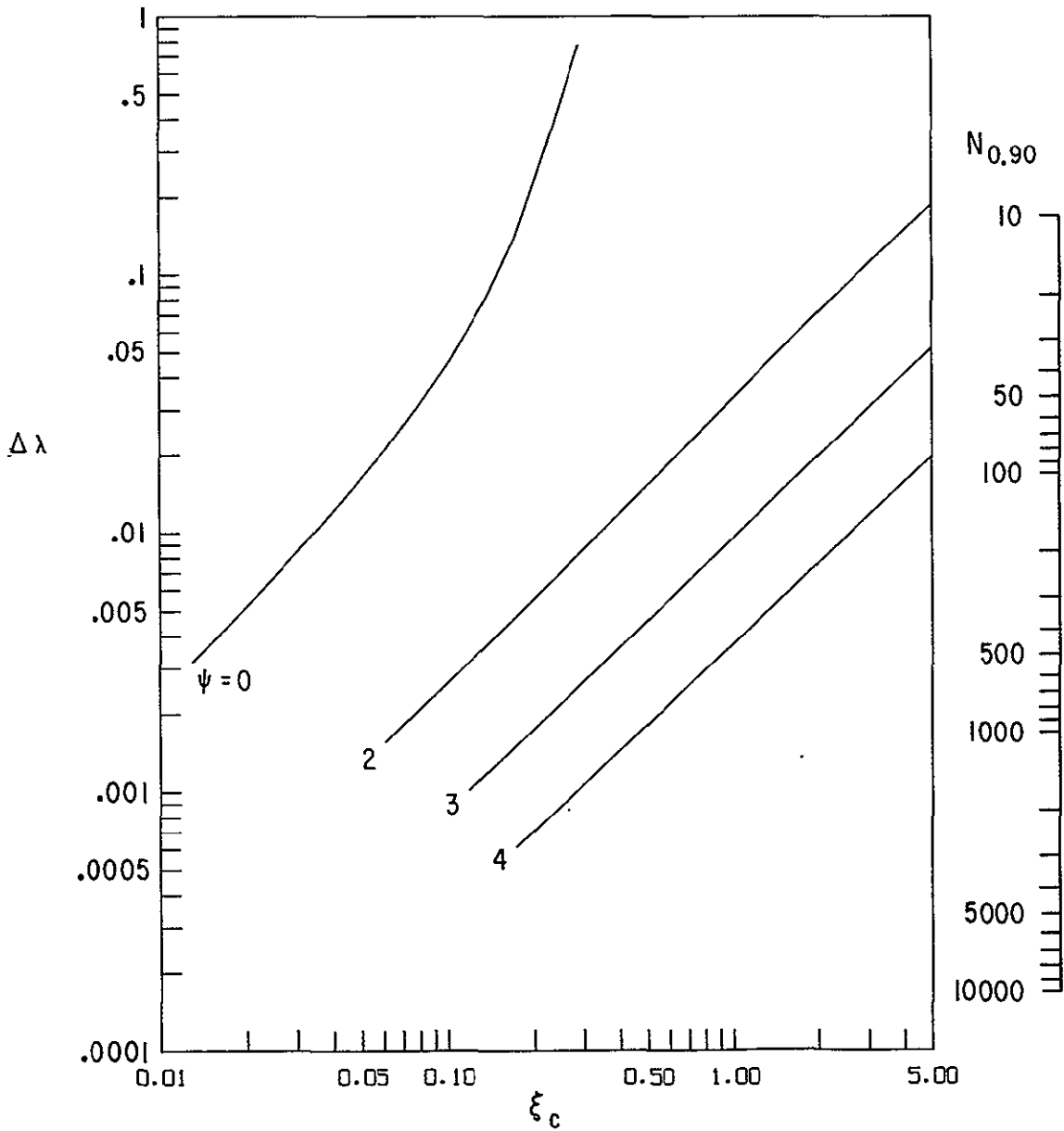
We may now with greater understanding and precision restate Scheuer's limit: when analyzing moderate numbers of observations, the information about $n(I)$ below the one-source-per-field-of-view limit, $\xi \sim 1$, derivable from the fluctuations, is restricted to the contribution made by such sources to the variance. What constitutes a "moderate" number of observations depends on the value of $\psi$, the noise-to-signal ratio of the fluctuations. For no noise, "moderate" is ~600 independent observations. For $\psi = 2$, moderate means up to $16 \times 10^3$ observations. Thus even this "limit" is not a hard one in that with a sufficient amount of data, the limit may be pushed to lower values of $\xi$. This understanding of the limits of the fluctuations for evaluating models of $N(S)$ is what allows us to reduce complicated models to a single power law with a Gaussian excess as we proposed in equations [III.68] and [III.69].

To round out the discussion of what the lowermost $\xi$ sources contribute to the fluctuations, we should investigate the effect of the very highest intensity sources. We could go through a similar likelihood difference analysis but this time intuition suffices. Clearly when $\xi$ is high enough that there is a negligible chance that such a source will appear in one out of N observations, bright sources can have only a marginal effect. We can ignore their impact on the shape of the fluctuations if we fit only observations with intensities below a certain upper cutoff. This cutoff must be high enough so that source intensities at or above it are still many times $W_{eff}$, and so that the observations are dominated by a single bright source. Care must be taken. Though the shape below the cutoff may not have much to do with those brighter sources, the placement of the mean with respect to the all-sky average intensity may still be impacted by their behavior (recall Figure III.7).

─────────────────────────

6: The Evaluation of Ideal Experiments for the Measurement of the

Fluctuations

We consider the competing effects of counting statistics,
number of observations, and variations in the non X-ray
count rate in the design and analysis of experiments to
measure the fluctuations. When we wish to measure the W
parameter of N(S), large sky coverage is preferred to a few
deep measurements. An optimal beam size for such
measurements exists, depending on the internal background
contribution to the noise. To measure the behavior of the
low flux sources through the invariant excess, $\sigma^2_S$, lower
noise is more important than sky coverage and number of
observations. Measuring the fluctuations with idealized
imaging experiments is briefly discussed. The ultimate
limitation of such experiments (for the fluctuations) comes
from spatial variations in the internal background.

In the last section we showed how the precision of any measurement
derived from the background depended on the number of observations and
on the ratio, $\psi$, of the other sources of noise to the effective size of
the fluctuations. For real measurements, neither parameter is
infinitely or even independently adjustable. For instance, an increase
in the number of observations is often gained at the cost of reduced
observation time per observation thereby increasing the size of the
photon statistics versus $W_{eff}$. The size and shape of the collimator
intrudes in a complicated way in this trade-off so as to impact optimal
collimator size for a particular experiment. In this section we
consider the interaction of these various parameters in the design of
experiments. Though we will cast our results in terms of "real"
detector characteristics, such as area, internal background, et cetera,
we will defer applying these calculations to known detectors, e.g. the
A-2 experiment. This is because of special details of the method used
by HEAO 1 to perform the all-sky survey.

We will try to illustrate qualitatively some general points about
experiments that measure the fluctuations. In addition, should a future
opportunity arise for further studies of fluctuations these equations
will be useful.

All information about a detector and its measurements of the diffuse background are tied up in its $R_a$s, the response moments defined in equation [II.17]. We remember that they may be divided into a part that depends on the angular response and a part that is related to the peak response to a point source:

$$R_a = C_a \, A \, t. \tag{III.85}$$

A and t are the detector area and the exposure integration time. For the ideal flat-response detector:

$$C_a = C_0^{1/a} . \tag{III.86}$$

$C_0$ is the total fraction of the sky to which the detector has non-zero response. When using the flat-response detector in an example we by convention replace all the $C_0$s with a simple C as an indication of that specific case.

Measurements with this detector will be subject to uncertainties. We will model the uncertainty by a total variance in the measurement which consists of three parts:

$$\sigma^2_{tot} = \sigma^2_f + \sigma^2_p + \sigma^2_{int}, \tag{III.87}$$

representing the contribution of the fluctuations, photon statistics, and internal variations. Though the fluctuations as we stressed in the last few sections, are not Gaussian, it will not be too inaccurate for present purposes to represent their contribution to the imprecision as

$$\sigma^2_f = (\psi_f \, W_{eff})^2 . \tag{III.88}$$

The coefficient, $\psi_f$, of the effective width is a number of order 1. Its actual value will depend on the kind of quantity being measured; e.g. in the last section we saw for measurements of $W_{eff}$ the imprecision contributed by the fluctuations was approximately of the form of equation [III.88] with $\psi_f \sim 1.1$.

The photon noise, modeled by a Poisson distribution, has a variance equal to the total number of counts in the intensity

measurement:

$$\sigma^2_p = I = C_1 \, A \, t \, S_{as} + B \, .$$                                    [III.89]

$S_{as}$ is the all-sky flux from X-rays and B is the number of internal background counts. The dependence of the internal background on the detector construction can be parameterized:

$$B = A \, t \, \beta$$                                                                   [III.90]

where $\beta$, the specific internal background, is the number of non-X-ray counts $s^{-1} \, cm^{-2}$.

$\sigma^2_{int}$ is a general term describing all other sources of variation. The most obvious such source is intrinsic variation in, or imprecision in, the estimate of the internal background. The variance it would contribute would be roughly proportional to $(A \, \beta)^2$. The dependence on the length of the integration time is more difficult to quantify without actual in-orbit study of the problem. $\beta$ could be have roughly periodic behavior with a characteristic period $t_b$. For integrations much longer than $t_b$, $\sigma^2_{int}$ will be independent of t, with magnitude roughly $t_b^2$ times the mean square amplitude of $\beta$. For short integrations, or the sum of a series of short integrations separated by intervals much longer than $t_b$, the variance will be proportional to the integration (or sub-integration) time squared. For an auto-regressive process [see Jenkins and Watts 1968; Tennant, Mushotzky, Boldt and Swank 1981], the variance will depend on a range of powers of t. For short integration times, it is proportional to $t^2$ and in the limit of long integrations it is proportional to t. Similar asymptotic behavior is observed for random processes described by shot noise [Sutherland, Weisskopf and Kahn 1978]. Purely for reasons of algebraic simplicity we will assume that

$$\sigma^2_{int} = b \, A^2 \, \beta^2 \, t$$                                             [III.91]

where b is a constant of proportionality. The analysis would only be affected in the event that the experiment has a severely high $\sigma^2_{int}$ and we were in the wrong integration time regime.

Measurement of point sources

Given a measurement of a point source with flux $S_{ps}$,

$$I_{ps} = A t S_{ps} ,$$ [III.92]

we see there are many ways to maximize the signal-to-noise ratio, $I/\sigma_{tot}$. We can increase the integration time without limit, unless $\sigma^2_{int}$ increases with t faster than $t^2$. We can decrease the solid angle of the detector, decreasing both $W_{eff}$ and the contribution to the photon noise due to $S_{as}$, until the internal background dominates through the photon statistics. Finally we can increase the area until the fluctuations dominate. We will study this irreducible complication of the fluctuations on resolved sources and other studies of the sky in more detail in the next section.

We can combine the photon and internal variances and identify them with the $\sigma^2_{noise}$ of the last section so that the total variance is

$$\sigma^2_{tot} = W_{eff}^2 ( \psi_f^2 + \psi^2 ) ,$$ [III.93]

where

$$\psi^2 = \frac{\sigma^2_p + \sigma^2_{int}}{W_{eff}^2} = \frac{C_1 S_{as} + (\beta + bA\beta^2)}{A t (C_{\gamma-1})^2 W^2} .$$ [III.94]

$\gamma$ and $W$ are the parameters that describe the fluctuations. We can rewrite this, retaining only the dependence on the experimentally adjustable values of t and the detector angular response:

$$\psi^2 = t^{-1} (C_{\gamma-1})^{-2} [C_1 \psi^2_{sky} + \psi^2_{int}] .$$ [III.95]

The values for $\psi_{sky}$ and $\psi_{int}$ now contain all the information about the sky brightness, size of the fluctuations, detector area, specific internal background, etc. They are not dimensionless, as $\psi_f$ and $\psi$ are. Note that the X-ray surface brightness and the internal background have been apportioned between them.

## Measurements of W

For the moment let us fix $\gamma$. From our analysis of the last section (see equations [III.77], [III.80], and Figure III.10) we can

relate the fractional size, squared, of the 90% confidence range in W to

$$\delta_W^2 \approx (1.7/2.9) \; N^{-1} \; (\psi_f^2 + \psi^2) \; , \qquad\qquad\qquad [III.96]$$

for the case of $\gamma = 5/2$. The coefficient 1.7 comes from $\chi^2_{1;.9}$ minus 1 for the single degree of freedom. The factor 2.9 is from an approximate numerical fit to the results displayed in Figure III.10. These same fits show that for determinations of W, $\psi_f \sim 1.1$.

We first note that if N, the number of independent observations, is fixed for some reason, then to minimize the confidence region we adjust the solid angle and the integration times to their highest possible values. This is because $\delta_W$ will be minimized when $\psi^2$ is minimized. For the flat-response detector, the solid angle response dependence of $\psi^2$ is

$$\psi^2 = t^{-1} \; [ \; c^{-1/3} \; \psi^2_{sky} + c^{-4/3} \; \psi^2_{int} \; ] \; . \qquad\qquad [III.97]$$

This shows that there will be basically two limiting regimes. For large enough solid angles,

$$c_1 > (\beta + b \, A \, \beta^2) \, / \, S_{as} \; , \qquad\qquad\qquad\qquad [III.98]$$

and the size of the confidence range will be dominated by $S_{as}$ through $\psi^2_{sky}$. For the smallest solid angles, the internal background will dominate either through photon statistics or the extra internal variance.

The above case, where N was fixed, is of course artificial. In the design of real experimental protocols, the number of observations will be a design parameter to be optimized. A more realistic constraint is T, the total time available for observation related to the single observation integration time by

$$t = T/N \; . \qquad\qquad\qquad\qquad\qquad\qquad [III.99]$$

Now we wish to minimize

$$\delta^2_W \propto \frac{1}{N} + \frac{Z}{T} \; . \qquad\qquad\qquad\qquad\qquad [III.100]$$

Z, which equals $\psi_f^{-2} \, t \, \psi^2$, contains the only C dependence. Now the optimal strategy is to increase the solid angle as much as possible while also maximizing the number of observations. Clearly this will run

into another constraint. There are, after all, at most $4\pi$ steradians in the sky available for observation, so that the product

$$C_{tot} = N\,C_o \qquad\qquad [\text{III.101}]$$

must be less than one. In fact after removal of contaminating resolved sources and the region at low galactic latitudes the constraint will be

$$C_{tot} \leqslant C_{max} = 1 - \text{contaminated sky fraction} . \qquad [\text{III.102}]$$

Now the optimal situation is where $C_{tot} = C_{max}$ and the detector solid angle response is set at some optimal size. For a flat-response detector in the sky-flux dominated region, this is

$$C = (\; C_{max}\psi^2{}_{sky} \;/\; 3\psi^2{}_f T \;)^{3/4} . \qquad\qquad [\text{III.103}]$$

As the sky brightness increases relative to the size of the fluctuations, or as the total time for the experiment is decreased, the optimal collimator size for the measurement of W is increased. This allows for an increased amount of time per observation which damps the size of the photon statistics versus the fluctuations. In the internal background dominated regime the optimal flat response detector size is

$$C = (\; 4\,C_{max}\;\psi^2{}_{int} \;/3\;\psi^2{}_f\;T)^{3/7} . \qquad\qquad [\text{III.104}]$$

### Measurements of low-flux behavior

Though an optimal measurement of W might argue for decreasing the size of the collimator to a certain point, and no farther, there is an independent reason for desiring as small a field of view as possible. Most of the fluctuations will be from sources with reduced intensities, $\xi$, ranging from a few tenths to a few tens. To determine the corresponding range of source fluxes, we modify equation [III.36]:

$$\langle S(\xi)\rangle = W\,\xi\;(C_{\gamma-1}/C_\gamma)^\gamma . \qquad\qquad [\text{III.105}]$$

$\langle S(\xi)\rangle$ is the expected flux of a source with reduced intensity $\xi$. For the flat-response detector this is just $W\,\xi\,C$. The smaller the field of view, the lower the values of S that we probe with our measurements of the fluctuations.

The sensitivity to the behavior of the very lowest flux sources is determined by how well we can measure the excess or deficit variance. In the last section we noted the similarity of equations [III.76] and [III.84] which indicated that being able to distinguish sharp cutoffs reduced to being able to distinguish the addition or deletion of the variance of a pure Gaussian distribution. If we are interested in setting an upper limit to any other source of fluctuations through their variance, $\sigma^2_\xi$, the size of the 90% upper limit should go as

$$\delta_{(\sigma^2_\xi)} \approx N^{-1/2} \left( \psi^2_f + \psi^2 \right) . \qquad [III.106]$$

For these measurements, $\psi_f$ is about 1.5. Note that increasing the number of observations is less important than decreasing the size of the noise term $\psi$. Because of this, it will turn out that even for a fixed detector angular size, there is an optimal number of observations for investigating any excess variance, in contrast to measurements of W where the more observations the better, subject to the limitations of the total amount of sky available. Of course if we are investigating the optimal-size field of view, we are interested in the upper limit that can be placed on the collimator-size-independent invariant excess, $\sigma^2_S$. The relationship between $\sigma^2_\xi$ and $\sigma^2_S$ was given in [III.69]. Since we are now investigating a quantity related to flux, rather than to the reduced intensity, $\xi$, we no longer use $\psi$ to characterize the noise. Instead,

$$\delta_{(\sigma^2_S)} \approx N^{-1/2} (C_2)^{-2}$$
$$\times \{(\psi_f C_{\gamma-1} W)^2 + t^{-1}[C_1 \phi_{sky} + \phi_{int}]\} , \qquad [III.107]$$

where

$$\phi_{sky} = S_{as}/A ; \quad \phi_{int} = (\beta + bA\beta^2)/A . \qquad [III.108]$$

The first term of the sum in [III.107] is the irreducible contribution that the fluctuations make to the upper bound, as indicated by its dependence on $W^2$.

For the $\gamma = 5/2$, sky-dominated, flat-response detector

$$\delta_{(\sigma^2_S)} \propto C^{1/6} .$$                                                [III.109]

This says that the smaller the collimator the better, but with an exponent of $1/6$, the gain in precision is very slow. The optimal number of observations for a particular value of C is

$$N = C^{1/3} \, T \, (\psi_f W)^2 \, / \, \phi_{sky} .$$                                  [III.110]

As the solid angle is decreased, the total observation time should be distributed among fewer observations so that photon statistics do not dominate the measurement of the invariant excess. When the non-fluctuations component of the noise is internal-background limited, the confidence range is proportional to $C^{-1/3}$. Smaller solid angles in this regime do not improve the upper limit of the invariant excess. Therefore the ultimate limit of a detector design for studies of the behavior of low-flux sources is determined by the size of $\phi_{int}$.

In the above case we also note that the optimal number of observations depends on the size of the underlying distribution of fluctuations from the power law N(S) component. Of course if we are measuring deeply enough we expect to run out of sources eventually, with the number-count gradually rolling over. A proper application of the above formula in those regimes will require the use of the proper index at S corresponding to $\xi$ of a few. In some cases, no good measurement of W, or $\gamma$, is possible at all. For those circumstances it is probably best to quantify the effects of the fluctuations purely as a variance term. In this case the first term in the sum of equation [III.107] must be replaced with the contribution due to that pure variance. Now the confidence range on the invariant "excess" (now with no power law base) is

$$\delta_{(\sigma^2_S)} \approx N^{-1/2} \, \{ \, \sigma^2_S + t^{-1} \, (C_2)^{-2} \, [C_1 \phi_{sky} + \phi_{int}] \, \} .$$     [III.111]

Further, the optimal number of observations, in the sky dominated regime, is

$$N = \sigma^2_S \, T \, / \, \phi_{sky} ,$$                                               [III.112]

independent of the size of the collimator. As a result, the confidence

range on the fluctuations is independent of the collimator size as long as the collimator size is small enough that $C_{tot} = N\,C$ is less than $C_{max}$, the total available fraction of the sky, and large enough to be outside the internal background regime.

## Experiments with imaging detectors

Imaging detectors have certain properties that affect the design of an optimal experiment. The two parameters which have the greatest effect on measurement of the fluctuations are T, the total experiment time, and $\phi_{int}$, the internal background contribution. The calculated impact of both is changed when the total exposure solid angle can be divided into pixels. Any concentrating detector removes the dependence of the internal background on the effective area of the experiment given in equations [III.90] and [III.91]. In addition, if the total imaging area is divided into m pixels, each one will be able to take data contemporaneously with the others. Thus everywhere that T appears in the above equations we may without loss of generality replace it with the product Tm. As it is always of benefit to increase T, we should design an imaging experiment so that the total image area $C_{im} = mC_0$ is as large as possible. $C_0$ is now the sky coverage of the pixel.

Given a particular value of $C_{im}$, the calculation of the optimal values for C are now changed. One of the major changes is in the scaling of the effects of the internal background. As well as removing the dependence on collecting area, the number of internal background counts in each pixel will be proportional to the fraction of the total image area it occupies:

$$B \sim t\ \beta\ C_0\ /\ C_{im}\ . \qquad\qquad [III.113]$$

Now the specific internal background rate, $\beta$, is the rate for the total image area. Therefore, as $C_{im}$ is divided into smaller and smaller areas, the internal background of each observation shrinks proportionately. Insofar as the measurements can be idealized as having flat spatial (solid angle) response, which is the case whenever the pixel size is greater than or the same order as the intrinsic spatial response of the imaging system, the ratio of the sky counts to the

background counts is fixed. There are no limitations to pixel size from a criterion based on the photon versus background counts.

The excess variance due to the internal background is another matter. At first it might seem that a similar argument would hold, that the decrease in the size of the pixel would quench $\sigma^2_{int}$ just as it quenched the magnitude of B. This is true for the excess due to the temporal variations in the internal background where $\sigma^2_{int} \propto C_o^2$. However, as indicated by the "~" in equation [III.113], there may be an additional source of variance due to spatial inhomogeneities of the internal background rate. How these scale with C depends on whether there are natural scale sizes to these inhomogeneities. If C is larger than these natural scales, $\sigma^2_{int} \propto C$, while if much smaller, $\sigma^2_{int} \propto C^2$. In principle, several deep exposures of several different fields of the sky should allow an estimate of these structures, assuming that they have a temporally invariant magnitude. Ironically, a real complication to such a program is the noise contributed by the sky fluctuations.

We write the two separate contributions to $\sigma^2_{int}$ as

$$\sigma^2_{int} = b' \, t \, \beta^2 \, C^2 \; + b'' \, t^2 \, \beta^2 \, C \; . \qquad \text{[III.114]}$$

The first term is from temporal variations in the internal background, while the second term is from small spatial variations. Now for the total noise we have three, rather than two, regimes. The first, the photon noise regime, is important when the solid angle is small enough so that internal variance from temporal instabilities is quenched:

$$C < [(S_{as}A) + (\beta/C_{im})] \; / \; b' \, \beta^2 \; . \qquad \text{[III.115]}$$

The second regime applies when the solid angle is too large to satisfy this condition. The third regime occurs when integration times are long enough that spatial uncertainties in the internal background are larger than photon statistics or temporal instabilities. This will happen for

times

$$t > \frac{S_{as}A + (\beta/C_{im}) + Cb'\beta^2}{b'' \beta^2} .$$                          [III.116]

The optimal size pixel for a measurement of W for a fixed $C_{im}$, is
small for the second and third regimes. If the integration times are
short enough and C small enough so that the photon statistics term
dominates, there will be an optimal size:

$$C = (4 C_{max} C_{im} \psi^2_{pho} / 3 \psi^2_f T)^{3/7}$$                            [III.117]

where

$$\psi^2_{pho} = \frac{S_{as} + (\beta/C_{im}A)}{A \bar{W}^2}$$                          [III.118]

is similar to $\psi^2_{sky}$ except it now includes the photon statistics
contribution of the internal background.

Measurements of the invariant excess are also changed for an
imaging detector. The procedure for finding the optimal pixel size is
much the same process as we outlined in the previous non-imaging case.
The size of the confidence range is still more sensitive to the noise
than to the number of observations. The optimal number of observations
is the same as given in equation [III.110] after replacing T with the
product Tm and including the counting statistics from both the sky and
internal background in $\phi_{sky}$.

---

7: Fluctuations and Other Measurements: Confusion Noise and Correlation
Studies

The fluctuations are an important consideration for other
measurement programs. Intensity measurements for low flux
sources will have a fluctuation contributed uncertainty in
addition to photon counting statistics. The correlation of
X-ray surface brightness measurements with other data are

similarly affected by the additional noise, though not if
the fluctuations are the signal that is being analyzed. For
data sets with substantial spatial overlap between
measurements, the fluctuations will also reduce the
effective number of independent measurements.

The fluctuations are an irreducible source of noise in studies of
the X-ray sky flux. In the devolopment of the $\Delta\lambda$ measure of sensitivity
of data to aspects of fluctuations models, we saw that even in the
absence of all other variation ($\psi = 0$), the variation from the
fluctuations are a source of uncertainty in evaluating $N(S)$ models.
Similarly, the fluctuations also impact other measurement programs. In
this section we estimate the magnitude of this impact for a variety of
situations.

### Single measurements of source intensities

We have already noted that the fluctuations provide an additional
uncertainty in measuring the intensity of a point source. Consider
first an example where the source intensity is determined by a single
measurement where the source position is known to be at the peak
response of the exposure, $r_{max}$. Let $I_{tot}$ be the total measured
intensity, and $I_s$ the sky flux. Therefore, $\overline{I}_s = R_1 S_{as}$ is the mean
contribution from the sky flux. The true source flux, $S_t$ is

$$S_t = (I_{tot} - I_s) / r_{max} , \qquad [III.119]$$

but because we can not know the true value of $I_s$ at that particular
position in the sky, even with no photon statistics uncertainty, there
is an intrinsic uncertainty in the estimated source flux

$$S_e = (I_{tot} - \overline{I}_s) / r_{max} . \qquad [III.120]$$

The approximate size of the uncertainty in $I_s$ is $\sim 2.5 W_{eff}$, so that the
minimum uncertainty in the source flux is

$$\delta_{S;fluct} \sim 2.5 \, W_{eff} / r_{max} . \qquad [III.121]$$

The factor of 2.5, we recall, is an approximation based on the size of
the 90% confidence range of the Euclidean $P_{I'}(I')$ curve. In Figure
III.14 we plot the size of other ranges. The asymmetric nature of the

Figure III.14: Integral probability ranges for $P_I$, for different
values of $\psi$

The solid curves are the minimum range of $\xi'$ that contain
the indicated percentage of the area. The dashed lines
are approximate fits to the behavior, assuming that the
fluctuations and the non-fluctuations noise, as
parameterized by $\psi$, add in quadrature. The fits were
performed only for $\psi$ restricted to the range 1 to 4.5.
The short dashed line is the distribution median.



$$\xi' = (W_{eff})^{-1}(I - \bar{I})$$
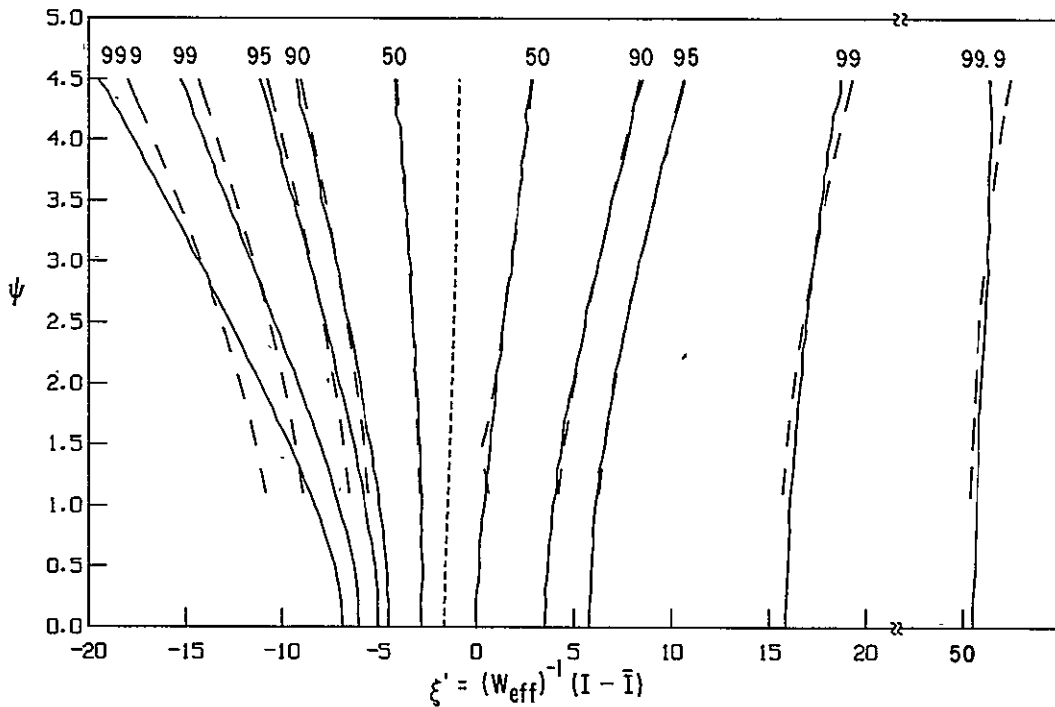
distribution is illustrated by the nonzero position of the median. In
addition the position of the range for areas greater than 90% show the
effects of the high flux power law tail of the distribution. The 99.9%
range runs from $-7$ $W_{eff}$ to over 50 $W_{eff}$. Figure III.14 also illustrates
the effect of additional noise, which is adds approximately in
quadrature to the fluctuations noise. This approximation does not work
that well for >95% limits. At the high end, the power law tail of the
distribution is essentially unperturbed by the presence of even very
large amounts of noise ($\psi \sim 5$).

Fluctuations Theory                                Confusion Noise

The statistical distribution of the source intensity is a mirror image of the fluctuations distribution, distributed about the mean $S_e$, and scaled by $W_{eff}$ / $r_{max}$. The reflected nature of the correction means that the small chance of a large positive fluctuation in the field of view corresponds to large negative correction to the source intensity. Reducing the measurement solid angle has the effect of reducing the amplitude of the fluctuations. For example, for the ideal flat response detector (see equation [II.21])

$$\delta_{S;fluct} \propto R_{\gamma-1} / r_{max} \propto C^{1/(\gamma-1)} .$$  [III.122]

C is the fraction of $4\pi$ steradians subtended by the detector. For a Euclidean power law N(S), $\delta^2_{S;fluct} \propto C^{4/3}$. The reduction of the solid angle will also reduce the contribution of the sky flux to the measurement. Therefore, the photon statistics uncertainty of the source flux is reduced. Again for the ideal detector this scales as

$$\delta^2_{S;pho} \propto C ,$$  [III.123]

assuming that the source intensity is small so that the sky intensity dominates the measurement. For a given integration time, t, as $C \to 0$ the contribution from the fluctuations will eventually fall below that of the counting statistics. On the other hand, if the measurement solid angle is fixed, then the integration time can be increased so that the fluctuations will be the predominant uncertainty, assuming there is no contribution to the uncertainty of the measurement due to variation in the internal background. The size of the photon variance is inversely proportional to the t, while $\delta_{S;fluct} \propto R_{\gamma-1}/r_{max}$ is independent of t. Therefore, the fluctuations provide the ultimate limit to the precision of flux measurement obtained with a perfect detector and infinitely long integrations.

## The fluctuations and baseline subtraction techniques

In this discussion we have assumed an a priori knowledge of the mean sky intensity, $\overline{I}_s$, in deriving the mean source flux in equation [III.120]. The value was derived from the all sky flux assuming that the sky flux was totally isotropic, apart from the fluctuations from

unresolved sources. If there is significant large scale structure, which we will see is true for the HEAO 1 A-2 HED and MED data only near the galactic plane, then a more local estimate of the mean sky contribution to the intensity is required. This is usually accomplished by a series of "off-source" measurements which will be called the baseline. For illustration we consider a baseline consisting of N nonoverlapping measurements with the same spatial response as the source measurement. Instead of using [III.120] to estimate the source flux we use

$$S_e = (I_{tot} - \overline{I}_{base}) / r_{max} .$$    [III.124]

$\overline{I}_{base}$ is the average baseline intensity

$$\overline{I}_{base} = (1/N) \sum_i I_{base;i} ,$$    [III.125]

where $I_{base;i}$ is the ith baseline measurement. If the baseline is a good estimator for the sky flux contribution to the source intensity measurement then the expected value of $(I_s - \overline{I}_{base})$ is zero. Since the expected value of $I_{tot}$ is $r_{max}S_t + I_s$, equation [III.124] is an unbiased estimator of the true source flux, $S_t$. However, the use of the baseline intensity to cancel $I_s$ introduces additional statistical uncertainties. The counting statistics contribution is straightforward to calculate but the fluctuations contribution to the uncertainty is more complicated. Both the baseline and $I_s$ are affected by the fluctuations. The value $-(I_s - \overline{I}_{base})$ can be treated as a single measurement. The spatial response function, $r(\theta,\phi)$, is positive over the baseline and negative over the source measurement. The net value for $R_1$ is zero, as required for the expected zero mean intensity. But with non-zero values for the other detector weighted responses, significant fluctuations are expected. The measurement is similar in these properties to the estimated internal background using the different detector fields of view (recall the discussion of Figures II.13 and II.14). Because of the negative $r(\theta,\phi)$ non-integral response weights are undefined, so that $W_{eff} = R_{\gamma-1} W$ can not be used to scale the size of the fluctuations. A sketch of the proper technique to calculate $P_{I'}(I')$ is outlined in Appendix A3. When both negative and positive values for $r(\theta,\phi)$

Fluctuations Theory                                        Confusion Noise

contribute to the measurement the resultant $P_{I'}(I')$ distribution looks different from the distribution when all values of r have the same sign. For example, if the baseline consists of a single measurement (N=1), the distribution of $P_{(I;s-I;base)}$ is symmetric. In addition to the high flux tail $\propto I^{-\gamma}$, there is a similar tail covering the negative intensities. Therefore there is equal chance for a very large positive fluctuation and for an identically large negative fluctuation in the total background for the point source flux measurement.

The contribution of the baseline to the fluctuations can be crudely estimated by considering the baseline portion of the response in isolation. Because of the averaging factor (1/N) in [III.125], the effective width of the baseline intensity fluctuations <u>decreases</u> with the inclusion of more observations in the baseline:

$$W_{eff;base} \propto N^{(2-\gamma)/(\gamma-1)} .$$                                   [III.126]

For Euclidean N(S) this is $\propto N^{-1/3}$. The decrease in the size of the baseline fluctuations is slower than the decrease for strictly Gaussian noise, such as photon statistics, where $\sigma_{gauss} \propto N^{-1/2}$. Therefore, the baseline size required to achieve a particular maximum contribution to the fluctuations is larger (for $\gamma=5/2$) than what would be needed if the fluctuations were Gaussian. Recall though that even for perfect knowledge of the baseline mean, <u>i.e.</u> $N \to \infty$, the sky fluctuations of the source measurement still contribute to the uncertainty of the source flux.

### The fluctuations and fits to source intensity

Our discussion so far has been limited to the idealized case where a source intensity was determined by a single on-source measurement with one or more off-source baseline measurements. In actuality, more sophisticated techniques are used to estimate source intensities. The data consist of a series of intensity measurements, $y_i$, with varying amounts of exposure to the source position, $r_i$. The source flux is

found by performing a weighted least squares solution to

$$\vec{y} = \vec{r} \, S_e + \overline{I}_s \, .$$                                    [III.127]

The error on the flux is estimated by inverting the correlation matrix (see Bevington [1969] for this and other details). The validity of this general technique requires that the source position be known, the weighting be determined by $1/\sigma^2_i$ where $\sigma^2_i$ is the Gaussian noise variance, and the model of a single source plus isotropic sky flux is the correct description of reality. All of these assumptions are flawed.

Good source positions depend on identifications with optical or radio counterparts or measurements with high spatial resolution X-ray experiments. When such information is lacking, the same data used to fit the source flux must also treat the source position as a free model parameter. The confidence range in the source flux is then best calculated using the likelihood statistic (equivalent to $\chi^2$ for Gaussian data) treating the source position as nuisance parameters (see Section II:5). The condition that the points are weighted properly is violated when the fluctuations are significant in comparison to the counting statistics. At the very least, an additional noise term with $\sigma \sim 2.5$ $W_{eff}$ should be added in quadrature to the photon counting statistics. The non-Gaussian nature of the fluctuations distribution may mean that $\chi^2$ is not the proper statistic to measure the source flux's likelihood. Also, in the presence of significant fluctuations, the hypothesis that all but the single source is a uniform baseline flux is an oversimplification, i.e. the baseline contains bumps and wiggles not included in the model.

The numerical value of the additional noise contributed by the fluctuations is difficult to calculate from first principles. The overlap in the sky coverage for individual measurements means that they are no longer statistically independent with respect to their variation due to the fluctuations. We will show (Section A2:2) that this is equivalent to a reduction in the number of measurements. As was derived in Section III:5, this reduction accentuates the uncertainty in the

Figure III.15: Histogram of fit flux values

The solid line histogram indicates the distribution of the derived flux values (in R15 units, see text) for 437 points chosen at random from areas of the sky containing no clearly detectable sources. The long dashed curve is a Gaussian with the width of the average derived uncertainty of the flux, $\sigma = 0.217$ R15. The short dashed curve is a Gaussian with the actual width of the histogram, $\sigma = 0.304$ R15.
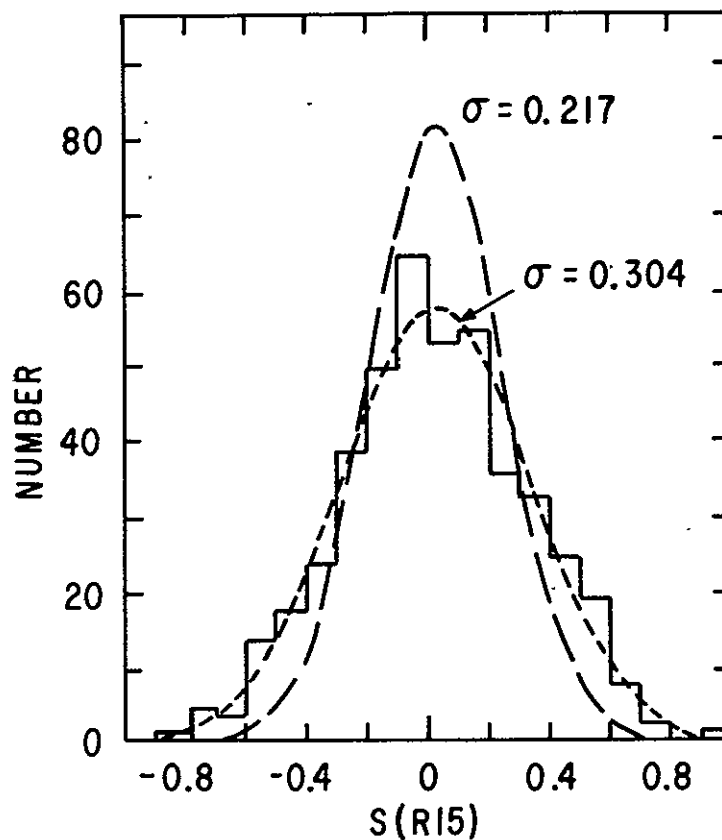


source flux. To account for this, the fit to the data explicitly tests a model in which the source flux produces an expected contribution to the different measurements as determined by the detector spatial response to a point source. Therefore, the fit is less sensitive to the bumps and wiggles of the fluctuations as long as they do not look like a point source.

Fluctuations Theory                                           Confusion Noise

Perhaps the most informative approach is to analyze sets of actual fits. As a crude measure of the impact of the fluctuations, we present in Figure III.15 the distribution of 437 fit point source intensities. The point source locations were randomly chosen spots covering high galactic latitude ranges, with no clearly detectable source in the HED 3 Small FOV. The derived source intensity should be distributed about zero, and the width of the distribution should give an indication of the intrinsic uncertainty of source intensities derived from such fits. The fluxes are measured in R15 units, which we will describe in more detail in Section VII:1. For now, let it suffice that R15 units are constructed out of several different Discovery Scaler windows of HED 3 and the MED, using the small field of view. The conversion for R15 flux units is ~623 R15 = 1 count $s^{-1}$ $cm^{-2}$ in the Hed 1 Layer 1 window. From the simple determination of the error using the inversion of the correlation matrix, the mean uncertainty in the flux was $\langle\sigma\rangle$ = 0.217. As Figure III.15 shows this is much narrower than the true distribution of fluxes which is well fit by a Gaussian with $\sigma$ = 0.304. Subtracting the two uncertainties in quadrature, we see that the the residual contribution to the source intensity variance is $\sigma_f$ ~ 0.21 R15. For comparison, we estimate that for W = 0.07 H1L1 (which is close to the best fit Euclidean models) and using a smeared detector of $\Delta\theta$ = 0.25° and $\Delta\phi$ = 2°, we predict that $W_{eff}$ should be ~0.1 R15 (based on estimates for the H1L1 window transformed to R15 units and using an unsmeared HED 3 SFOV detector response). These particular smear angles correspond to the size of the measurements used in deriving the distribution in Figure III.15. Thus if we use the 2.5 $W_{eff}$ definition of the effective $\sigma$ for the fluctuations, the expected $\sigma_f$ should be ~0.25 R15. This crude estimate is consistent with, although a little higher than, the above measured value for $\sigma_f$. However, as we do not expect the fluctuations to mimic the response of a point source, it is likely that our estimate is high. In any case, it is clear that the fluctuations can have a significant impact on the uncertainties of low intensity source fluxes.

### The fluctuations and other studies of the X-ray sky

Any analysis of the X-ray sky flux will be affected by the

fluctuations. When investigating correlations between the sky flux and other signals e.g. galactic structure or extragalactic source distribution, the fluctuations will be expected to contribute at least an additional noise term. In our analysis of such correlations for this dissertation we approximate this contribution as an additional Gaussian term. In addition, the use of overlapping data must be handled in a different fashion in order to account for the non-independence of the measurements. The exception to this general rule is for correlations that involve the sources that generate the fluctuations. An example is the analysis of Turner and Geller [1980], who examined the UHURU data of Schwartz [1980] (see the discussion of this data in Section VI:6) for correlations with the counts of bright galaxies or nearby clusters. Turner and Geller concluded from the lack of such correlation that sources associated with these objects could not produce the dominate fraction of the X-ray sky flux. If all of the fluctuations were generated by the sources physically associated with the optical counts used, then there would be no need to include an additional noise term from the fluctuations. It should be remembered that although the variance of the fluctuations is dominated by the bright and relatively nearby sources (which might be expected to occur in the optical data) a substantial noise term is generated by the weak sources, $d(\sigma^2) \propto I^{-1/2} \, dI$ for a Euclidean N(S) (see equation [III.56]). Thus unless the optical data used is very deep in its coverage of the X-ray sources, some additional noise term from the fluctuations due to excluded sources is necessary.

CHAPTER· IV

PREPARATION OF A DATABASE FOR STUDIES OF THE FLUCTUATIONS

In order to study the fluctuations we must prepare datasets
of independent measurements with good signal to noise. A
weighting correction is outlined that permits the use of
data where there is some overlap between two measurements.
Before these datasets are used for the study of the
fluctuations we estimate and remove other sources of
variation. Some of these other variations are
astrophysically interesting in their own right. Finally we
estimate from simulations the sensitivity of the experiment.
We choose for subsequent analysis data from HED 1 L+SFOV.,
smeared over $6^{\circ}$ in the scan direction.

- 1: The creation of the SCANER database

The SCANER database tapes contain all the Discovery Scaler data
taken while the experiment was not in "point" mode. The data taken
during a particular interval are superposed onto a single great circle.
In this section we outline the processes and programs that produced the
SCANER tapes from the XRATES tapes. For analysis of Discovery Scaler
data where temporal superposition is allowed, the SCANER tapes provide a
compact and complete representation of the X-ray sky. The organization
of the data base is detailed in Appendix A4.

Initial pass through XRATES: JUTE

The· production of the SCANER tapes requires two complete passes
through the XRATES data. The first pass is outlined in Figure IV.1. It
makes the initial definition of the mean spinaxis position for a given
range in time. The nominal spinaxis position was set by command,
changed by about one half degree every twelve hours to maintain the

Figure IV.1: Initial pass through XRATES: The JUTE program

ORIGINAL PAGE IS
OF POOR QUALITY
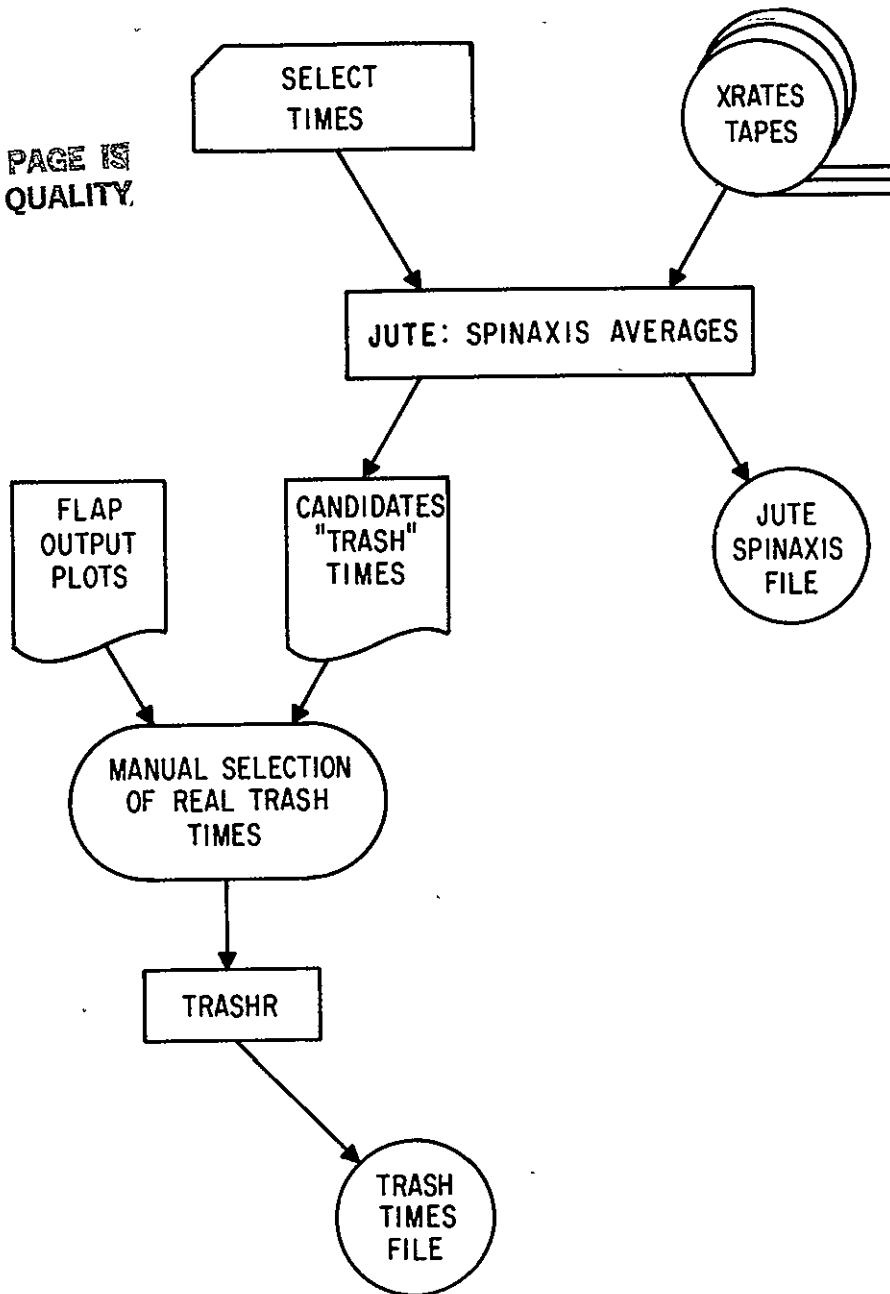
optimal orientation of the solar panels. The exact time of each "spinaxis select", including the beginning and end of each non-scanning "point" maneuver, was reported by the HEAO operations control team. The

data taken during a spinaxis select time are binned in a plane defined perpendicular to the mean spinaxis position.

The program JUTE calculates that mean spinaxis position. For a given spinaxis select period, the program checks each major frame, i.e. every 40.96 seconds, for the following conditions: satellite not in point mode; satellite not in South Atlantic or North Pacific magnetic anomalies; and at least one of the detectors "good", i.e. the CLEAN flag set and the Earth+200 km occult and electron contamination flags not set (see Section II:5 for the description of the XRATES database). From the qualifying major frames JUTE constructs the mean spinaxis position:

$$\hat{Z} = \vec{Z} \; / \; |\vec{Z}| \; ; \; \vec{Z} = \sum \vec{z}_i , \qquad\qquad [\text{IV.1}]$$

where $\vec{z}_i$ is the spinaxis position at the beginning of the ith allowed major frame. In addition, the angular separation of the y-axis over the time between the first and last exposure of the major frame, a period of 39.68 seconds, is used to calculate the mean angular spin rate. These data were written as a record of a Spinaxis file, which was then used as input in subsequent programs. A full description of the format of a Spinaxis-type record is given in Appendix A4. Initially part of the record is blank and is filled in by later dataset production programs.

When the Source-in-Field-of-View flags indicated that a particular detector was looking at the (more or less) isotropic diffuse sky, a second check on data quality was made. For each detector, the sum of the first four Discovery Scalers was compared on a 5.12 second basis with the expected rate from the nominal diffuse sky flux. Whenever the rate was 3.6 sigma outside of this nominal rate, a message was printed out flagging the time and position as a possible "trash time" event. These were often accidental statistical triggerings, at an expected rate of one accidental every 100 major frames, but whenever two detectors had two tiggerings per major frame, or when a single detector had three triggered intervals, the candidate trash time was marked for further investigation. Many times these were due to low intensity, uncatalogued sources or occasional high flux X-ray transients. Additional trash

times indicated intervals when the Discovery Scaler definitions were not
in their nominal modes; this sometimes happened after a point maneuver
where the telemetry format had been customized for the study of a
particular source.  However, by studying the plotted output of the
Discovery Scalers and the raw rates, including the layers used in anti-
coincidence, we were able to discover intervals where the data were
clearly contaminated by a charged particle event but, for some reason,
had not met the criteria for setting the electron contamination flag.
These occasional events, which usually took place just before the
beginning of a high intensity particle contamination period, were
flagged as real trash times, and a list of these time intervals along
with the affected detectors, was written in a disk file by the program
TRASHR.  Not all detectors were affected by all these events.  In
general HED 1 was the least affected, while HED 2 was the most sensitive
to these events.  On average, one of these events occurred every few
days, although during certain periods the satellite orbital path was
such that the events occurred once or twice an orbit.  When an event was
indicated, the total excluded time was enlarged far beyond the time of
the particular contaminated major frame which triggered the event.

<div align="center">Second pass through XRATES: BIGSCN</div>

The mean spinaxis position defined by the first pass through the
XRATES tapes allows us to define a scan plane coordinate system for the
period of a particular spinaxis select time.  The z-axis, $\hat{z}$, of such a
system is defined by the mean spinaxis.  The x- and y-axes of the scan
plane are defined by the following cross products:

$$\hat{x} \propto \hat{n} \times \hat{z} \qquad\qquad\qquad\qquad [IV.2.a]$$
$$\hat{y} \propto \hat{z} \times \hat{x} \qquad\qquad\qquad\qquad [IV.2.b]$$

where $\hat{n}$ is the position of the north ecliptic pole and the
proportionality sign indicates that the vectors must be renormalized to
unit length.  Ideally the spinaxis is pointed in the ecliptic equator so
that $\hat{n}$ and $\hat{y}$ are identical, but because of pointing jitter, the mean
spinaxis position may be as much as a few tenths of a degree out of the
ecliptic plane.

ORIGINAL PAGE IS
OF POOR QUALITY

---

Figure IV.2: Second pass through XRATES: The BIGSCN program

---
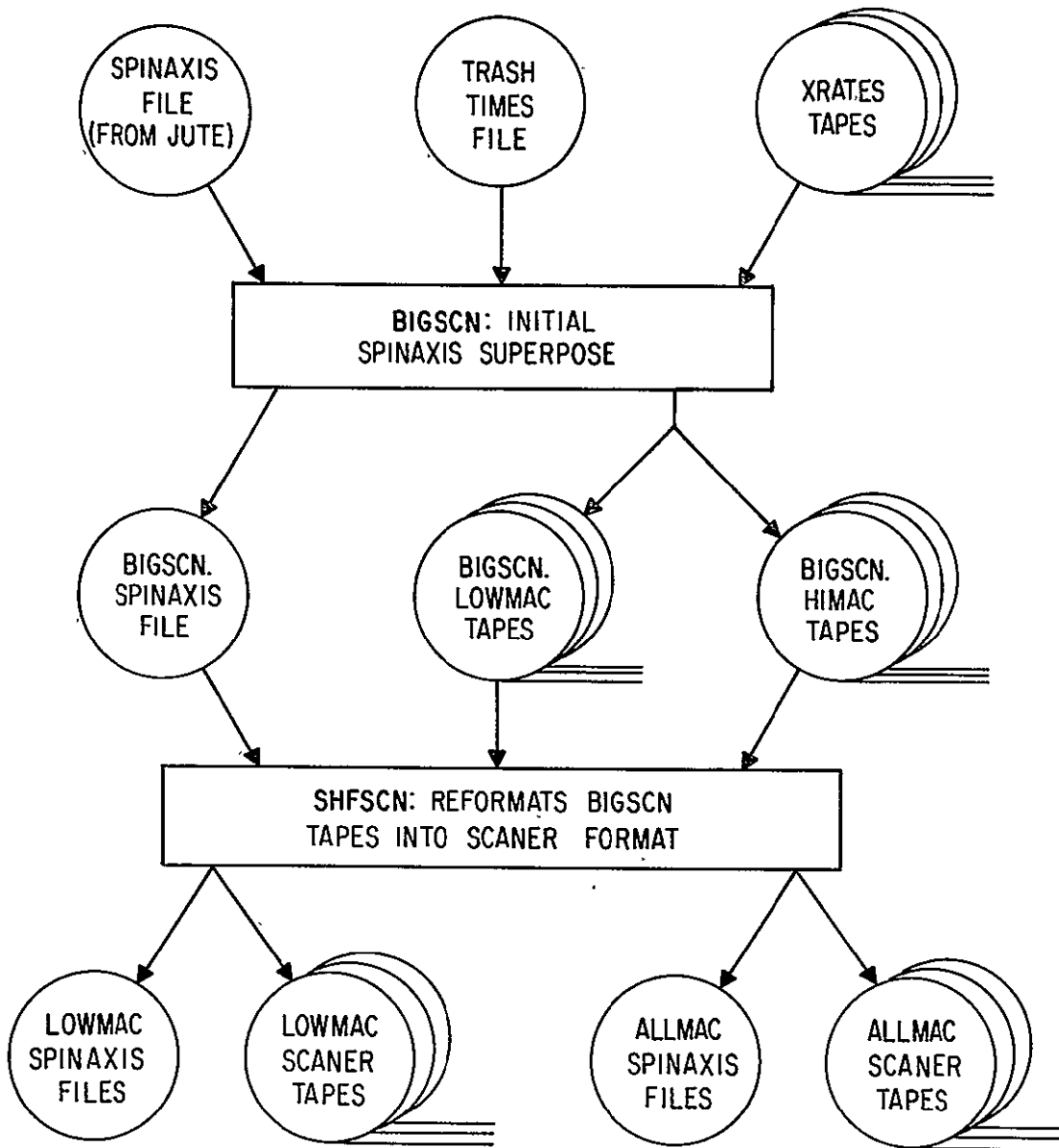


BIGSCN reads the XRATES tapes and makes the same data quality checks as the JUTE pass. In addition, the detector is not considered "good" if the major frame occurs during a trash time for that particular detector. The scan plane is divided into equal-angle bins, the standard bin size being one fifth of a degree, so that the total number of bins is 1800. For each major frame, the position of the Moon is found in

scan coordinates. If the Moon is within 3.5 degrees of the scan plane it could be a source of contamination. For each detector, a range of bins is excluded from the current major frame because an exposure centered on one of those bins might have a lower sky brightness due to shadowing by the Moon.

For each 1.28 second exposure, the position of a "good" detector's center is projected into the scan plane to find the corresponding bin number. The total telemetry allocated to the 1.28 second Discovery Scaler rates limited the maximum number of counts per exposure to 255. This could be exceeded for the very brightest sources, such as Sco X-1, Cyg X-1, the Crab Nebula, and the Galactic center region. This possibility was checked, and a final check for bit errors made, by summing a group of four 1.28-second exposures to see whether they equal the associated 5.12-second exposure counts. If an overflow is indicated, then those four 1.28-second exposures are not used for accumulations; however, each of their associated bins will have a flag set to show that overflow data were taken when the detector was pointed at this part of the sky. When no overflow is indicated, the number of exposures for that bin is incremented and the total number of counts as well as the sum of the counts squared are accumulated. The latter is useful for studies of time variability. Occasionally, the rates are so high that the sum of the squared count rates will overflow the available storage, in which case a flag is set so that analysis using the squared count rates will skip such bins.

In addition to the Discovery Scaler rates, the value of the McIlwain L parameter is accumulated. Also accumulated are the angle between the center of the detector and the scan plane and the square of this angle, so that the mean position and the rms amount of wobble can be calculated. These also serve to indicate the quality of the superposition of data from adjacent scans.

Flags are also set for each bin using the XRATES Source-in-Field-of-View flags for the major frame. Initially these one byte flags are incremented for every exposure where the Source-in-Field-of-View flag

was set. Because of the Fortran implementation, any non-zero value is equivalent to a logical TRUE value. However, because of possible overflow, the number of source-contaminated exposures is lost for subsequent processing, so that these flags contain only FALSE or TRUE (the latter indicating that at least one exposure came from a major frame where the Source-in-Field-of-View flag was set). Because these flags are calculated from the XRATES flags, they depend on the catalog used for the generation of that tape. In addition since these flags are calculated on a major frame basis they cover an interval during which the detector scans through eight degrees. Thus these flags are quite conservative, marking down areas which might in fact have no source contamination at all. For this reason, the use of these flags is good for only a crude selection of source-free surface brightness data, as they eliminate from consideration too large a fraction of the sky. For studies which require source-free data we will use a more careful and explicit calculation.

There are ten sets of accumulation arrays. These sets are divided into two groups of five, one for data where the McIlwain L value is less than 1.18, the other where it is greater. Each group of five gives the data for HED 1, HED 2, and HED 3, and the MED "super-clean" and "unclean" data. The "unclean" data meet only the less stringent of the two MED electron contamination flag criteria. At the end of the spinaxis select period that defines the current spinaxis and scan plane coordinate systems, the two groups of five are output on separate tapes, the low McIlwain L data on a tape with datasets named BIGSCN.LOWMAC and the other on BIGSCN.HIMAC. Each group is preceeded by an identifying header block containing general information, including the scan plane coordinate system axes. The header and the accumulated rates are output on a single file for each spinaxis select time period. At the same time a modified Spinaxis record is written. This record contains, in addition to the old information, the volume and file number of the BIGSCN.LOWMAC file associated with that time period. It also indicates the average deviation of the actual spinaxis position from the mean spinaxis postion calculated by JUTE and used to calculate the scan plane coordinate system.

Occasionally this average deviation is significant, more than half a degree. This can be due to periods inadvertantly including point mode data. Or a significant fraction of the spinaxis select time might have been during .a trash time that JUTE counted in calculating the mean -spinaxis, but that BIGSCN did not use in accumulating the data. Those scans where the average deviation was large enough were examined in further detail: the accumulated deviations of the detector position from the scan plane were used to determine a second mean spinaxis position. If this was more than a half degree from the JUTE mean spinaxis, then BIGSCN was re-run using a scan plane defined by the new mean. Four such sets of scans were re-run, all from the latter part of the mission.

## Reformatting the data: SHFSCN and SUSCAN

The output from BIGSCN was not the final desired form of the data. In particular, a dataset with all values of McIlwain L was desired, rather than just for L greater than 1.18. Similarly, a fifth dataset was needed, containing all the MED data that satisfied the less stringent electron flag, not just the data which satisfied only the less stringent and not the more stringent condition. The program SHFSCN performs this reformatting and outputs two sets of tapes in the final SCANER data form: a set of the low L data, and a set with data for all values of L. Each set has its own Spinaxis disk file, containing the volume and file number information for each superposed scan. The initial SCANER tapes output by SHFSCN have the finest spatial resolution, a fifth of a degree in the scan plane, and the temporal resolution of one superposed .scan per spinaxis select (nominally twelve hours corresponding to half a degree perpendicular to the scan plane).

Such resolution produces an enormous amount of data, difficult to handle all at once. Even to use the information from one superposed scan in a program imposes significant memory requirements. The program SUSCAN compresses the data .further, by re-superposing data from several spinaxis select periods on scan planes with coarser resolution. The usual mode for SUSCAN produces data from a two day period, re-binned at

one degree. The two day period corresponds to a smear perpendicular to the scan plane of at most 1.5 degrees at the ecliptic equator. SUSCAN can be used to produce a SCANER tape with any degree of re-binning desired.

----------------------------

## 2: Subsets of the SCANER database: Independent and Overlapping Data

The SCANER data format is not sufficiently processed for easy understanding of particular questions about the diffuse sky, including the analysis of the fluctuations. . Certain regions of the sky, such as the galactic plane and bright sources, should be avoided. In addition, the re-binning performed by SUSCAN may not include enough active time in each measurement to have good signal-to-noise. Further re-binning of the data by scan angle may be necessary. Finally, actual rates and other information must be. extracted from the accumulation arrays in the SCANER data. A particular issue to be faced is the problem of statistically independent measurements of the diffuse sky, what we call the problem of overlapping data.

### General region exclusion: NOTIND

Initially we will ignore the distinction of overlapping versus non-overlapping data. To first order, many structures or features in the X-ray sky may be studied with overlapping data. The program NOTIND will, for each detector, output records that contain the accumulated Discovery Scaler counts, McIlwain L value, deviation from the scan plane, and time and mean position for a group of SCANER bins which lie outside an excluded region. This excluded region is a rectangle in galactic longitude and latitude centered on the galactic center. Bins are also, excluded when they include data from certain other regions of the sky, e.g. positions of known sources.

The output format of the program is generically called the INDIFILE format, which is produced by several other programs. One such

program, closely related to NOTIND, is REBING. The records for NOTIND are all measurements from the same superposed scan, adjacent to each other in scan angle. REBING, while providing the same checks for contamination by sources and/or the galactic center or plane, re-bins the data in cells that include all the acceptable data centered on a particular region in the sky. The cells are defined with borders of constant galactic latitude and longitude. Thus cells near the ecliptic poles may contain the superposition of data from many different scans taken at markedly different times.

## Independent data and studies of the fluctuations

Trying to decide statistically whether a given set of measurements were randomly drawn from a parent distribution is greatly complicated if some of the measurements are statistically related. By such a relationship we mean that the position in the parent distribution of one measurement can be shown to be dependent on the position of another measurement in the set. This is quite unlike measurements being related to each other in a non-statistical sense.

In a general model for a set of measurements we can speak of three sources of variation. The first, non-stochastic variations, are predicted absolutely by the model; i.e. the model says non-probabilistically that the ith measurement has a particular non-stochastic component. The second, stochastic, source of variation is one where the model speaks only in a probabilistic way about the distribution of the measurement, but with the proviso that if the experiment were repeated, the ith measurement would have exactly the same value for the stochastic component. The final source of variation, pure statistical, is one that is inherently unpredictable from one repetition of the ith measurement to the next.

As an example of measurements with only non-stochastic and purely statistical variations, consider the common case in X-ray astronomy of spectral fitting. A set of counts from different pulse height channels is compared to a particular form predicted by a spectral model for the

source. The non-stochastic variations are the lines and wiggles of the
model. The pure statistical variations are Poisson counting statistics.
There is often an obvious relationship in the model from one channel to
the next, as when a spectral line spans several channels, but in a
statistical sense, each measurement is independent. Given that we have
the correct model, the fact that one channel is above or below the
predicted value has no impact on the behavior of adjacent channels. It
is this quality that allows us to treat the channels as independent
tests of the model in calculating various statistical tools such as the
$\chi^2$ statistic.

The diffuse sky might be idealized as an example of what we call
here a stochastic variation (ignoring for the moment counting statistics
and other structure in the sky). Consider a single measurement of the
X-ray surface brightness. If we repeat the measurement with the
collimator aligned exactly as before, we expect to get the same answer.
In the last chapter we saw how a large number of measurements is needed
to accurately determine the underlying source of the fluctuations in the
sky. However, repeating the same measurements will produce no gain in
information. Even if two measurements are not exactly co-aligned, they
are not expected to tell us as much about the distribution of the
fluctuations as two totally independent measurements (as long as those
cover substantially the same patch of the sky). The real problem of
such overlapping measurements is that there are few well-studied
techniques for explicitly handling them, particularly in astrophysics.
The vast majority of the statistical tools currently used explicitly
assume that the data points are statistically independent. In this
section we address the process of extracting subsets of the SCANER data
base, using the program NSCIND, which are totally internally
independent, i.e. no two members of a particular subset had any non-
zero response to the same point in the sky. We will later discuss how
answers from multiple subsets, though not totally independent, might be
combined in a useful fashion. We will also develop an approximate
technique for performing analysis with data having various amounts of
overlap that is particularly appropriate for studying the non-stochastic
behavior of the sky. The same technique, even more approximately, is

used in studying the fluctuations with overlapping data.

### Generating independent subset of data: NSCIND

By definition each smeared observation is of a size $\delta\phi$ by $\delta\theta$ in the directions perpendicular and parallel to the scan plane, respectively. These are not the smear angles, but the size of the sky

---

Figure IV.3: NSCIND: Generating independent subsets of SCANER data

(a): The selection of particular scans and their range of non-overlapping scan angles. The "X" indicates an ecliptic pole, through which essentially all scans pass. The width of each scan, for purposes of defining independent data, is given by $\delta\phi$. The scans are selected and processed in the numerical order indicated, with care taken that there is no overlap with previous selected scans. (b): The division of the observations within a single scan into internally independent subsets. The solid, dashed, and dotted lines outline schematically the smeared response function $(r(\theta,\phi=0))$ of observations that belong to the first (solid), second (dash-dotted), and third (dashed) subsets of the scan, respectively.

---
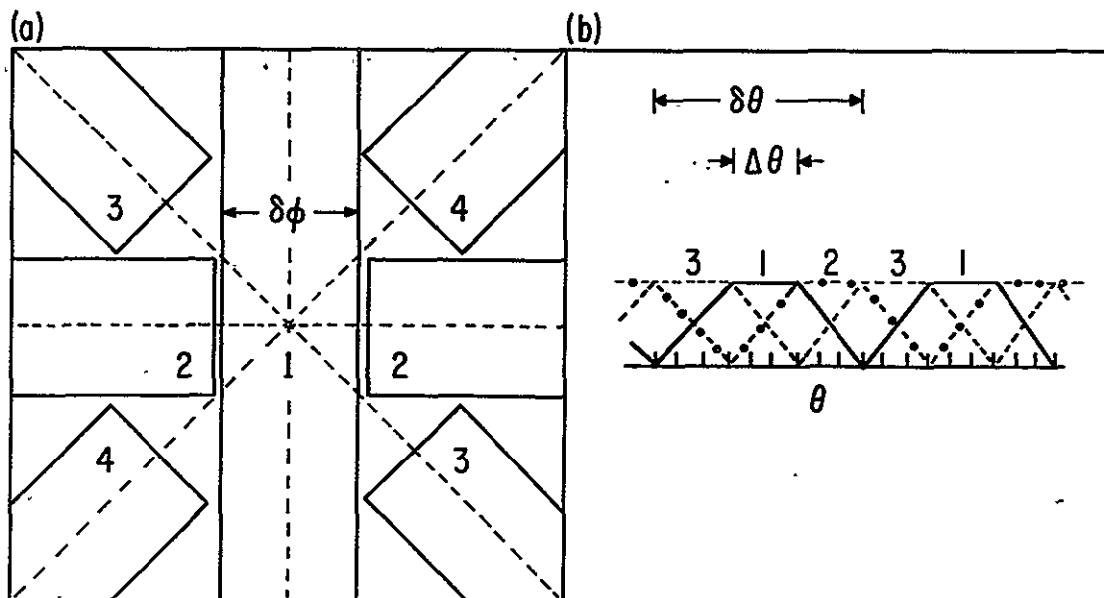


that each observation effectively covers. The first step is to define

for each superposed scan a range of scan angles which will not overlap
with similarly defined ranges for the other scans. This process is
outlined in Figure IV.3a. We start with a single scan, generally one
with the most exposure perpendicular to the galactic plane. This last
criterion maximizes the number of observations derived from this first
scan. All data outside the galactic plane, or uncontaminated by
exposure to selected sources, are used. The. next scan is chosen
essentially perpendicular. In addition to the above selection criteria,
data near the poles are not used because of possible overlap with data
taken in the first scan. The third and fourth selected scans are
positioned roughly between the first and second scans, and checked for
overlap with both of them. Because of this ordering of the scans, each
subsequent scan need only be checked for overlap with at most two
adjacent previous scans. Note, however that this may not be the most
efficient method in terms of sky coverage. Large fractions of the data
base will be marked unusable because some scans will be totally
overlapped by adjacent selected scans. Also, near the ecliptic poles
the fraction of the total sky area coverage is reduced because even with
the closest abutting scans, large areas are not included by any scan.
To maximize the total coverage there was some allowed leeway in the
selection of the follow-on scans, i.e. they did not have to be exactly
midway between two previously selected scans. The allowance of a few
degrees permitted the selection of the superposed scan with the most
active time that was nearby the otherwise optimal exact position. The
scans which made no contribution because of total overlap need not be
totally lost. A second, or even third, set of "independent scan ranges"
can be derived by running the scan selection process through the scan
range from the beginning multiple times. However, the statistics of
these subsequently defined data subsets will be affected by the
decreased coverage in the scans used.

The selection of a set of independent scan ranges guarantees that
the data will have no overlap from one scan to the next as set by $\delta\phi$.
In Figure IV.3.b we illustrate the similar selection technique for
handling overlap within the scan. Here the subsequent re-binned
observations will fall naturally into a series of subsets, each

internally independent. The number of such subsets depends on the size
of the re-binning and the parallel smear angel $\Delta\theta$, as compared to the
total non-zero response size, $\delta\theta$. For instance, if $\Delta\theta$ is one third of
$\delta\theta$, as is the case in Figure IV.3.b, there are three such subsets. At
least two such subsets are always generated, although they need not have
equal smear angle; the program NSCIND will only output the subsets
together when they do. Each observation produces an INDIFILE record
containing an associated subset number.

## XTRACT: Final data preparation

The INDIFILE records contain the accumulation data for all eight
Discovery Scalers. A final program is then needed to generate the
counts per exposure for a particular Discovery Scaler combination of
field of view and pulse height window. The simplest such program is
XTRACT, which produces an XTREC form record. Besides the rate
information, the XTREC record contains the estimated photon statistics
variance, the time associated with the superposed scan from which the
observation was taken, the mean value of McIlwain L, and a three-vector
containing the mean detector position in galactic coordinates. The
record also contains the subsample number, with the output file
generally sorted in ascending subsample number.

## XTROVL: XTRACT with calculation of overlap weights

There is a final slot in the XTREC record which originally held
the detector's rms excursion perpendicular to the scan plane. This
value was not used in any subsequently developed fitting programs as its
variation was generally small compared to the observation size $\delta\phi$. A
second version of XTRACT, named XTROVL, used that slot to store the
weight of the observation.. The weight is an estimated correction of
including data which might overlap other observations. The derivation
of the correction is outlined in Section A2:2; in this section we sketch
the algorithm and its limitations.

As we remarked above, two observations with the collimators

exactly overlapping, and no uncertainty due to photon statistics, are expected to have exactly the same intensity. If we wanted to include both measurements in a $\chi^2$ based estimator of the mean count rate or of some parameter of the model for the sky's non-stochastic variation, we could not treat the two measurements as two degrees of freedom. The individual contributions to the $\chi^2$ sum can be weighted by one half, so that the total contribution of each is one degree of freedom. Similarly, a set of observations consisting of n groups of m measurements that totally overlap must weight each observation by 1/m to give a total $\chi^2$ which is distributed as if its number of degrees of freedom were derived from n independent measurements.

The complications arise when the overlap is not total. We can show that in certain simple cases the dependence can be measured by the size of the cross-correlation of two measurements divided by the total variance. Let $\sigma^2_i$ be the pure statistical variation of the ith measurement. We assume that all measurements have the same shape for their spatial response function and that the fluctuations may be characterized as Gaussian with a variance $\sigma^2_D$. The contribution of the jth measurement to the ith measurement's weight is (to first order)

$$w_{ij} = \sigma_{ij} \ / \ (\sigma^2_j + \sigma^2_D),\qquad\qquad\text{[IV.3]}$$

where $\sigma_{ij}$ is the cross-correlation, calculated by an integral involving only the two spatial response functions. For two observations with no spatial overlap, $\sigma_{ij}$ and hence $w_{ij}$ reduce to zero. For two observations with total overlap, $\sigma_{ij}$ equals $\sigma^2_D$ so that when there are no counting statistics, $w_{ij}$ equals one. The total weight of the ith observation is

$$W_i = 1 \ / \ (1 + \sum_{j\neq i} w_{ij}).\qquad\qquad\text{[IV.4]}$$

For n observations with total overlap and no counting statistics the total weight is 1/n. The weighted number of observations,

$$N_w = \sum_i W_i,\qquad\qquad\qquad\text{[IV.5]}$$

is the total used in calculating the number of degrees of freedom for

the weighted $\chi^2$ sums.

The above prescription is exact for determinations of non-stochastic models using $\chi^2$ when the stochastic contribution is well-represented by a Gaussian with variance $\sigma^2_D$ and if all overlapping observations are in groups with equal weight. The more detailed derivation of this weighting is found in Section A2:2. We will assume the adequacy of this weighting for the real non-Gaussian fluctuations, but $\sigma^2_D$ will be found from an explicit Gaussian fit to the fluctuations and not equated to poorly determined variances of the power-law $N(S)$ model $P_I$ curves. Is the weighting appropriate for determinations of parameters of the stochastic variations, i.e. the parameters of the fluctuations' models? As we shall see, such parameters are determined using a likelihood statistic, rather than a $\chi^2$ distribution, but the close connection between the two might argue that this gap can be crossed.

The derivation in Section A2:2 is for the special case of the weighting of a least squares estimate of a distribution mean. The behavior of the likelihood difference measurement for W near its true value is similar to the behavior of the likelihood difference in determinations of the mean count rate (see equations [III.80] and [III.74]). This suggests that for determinations of W with the fluctuations, the proper weighting is approximately the same as the weighting for the least squares mean determination. The excess variance, however, has a different behavior for $\Delta\lambda$. Recalling that for good determinations of the excess variance a small size for the total variance was more important than the number of observations, we expect that the weighting might be less sensitive to overlap. In fact preliminary Monte Carlo work indicates that the proper weighting for overlapping data might use $w_{ij}^2$ in place of $w_{ij}$ in the sum for $W_i$ (equation [IV.4]). If this is the correct weighting, we could still use the unadjusted weighting of [IV.4], in which case confidence contours would be conservative estimators. For example, the "90%" confidence contour would be at least that secure and might actually represent a higher confidence limit.

Data Preparation                                        Database Subsets

Whether the above technique or any technique for handling overlapping data will give correct estimates of the confidence regions can be investigated by a Monte Carlo simulation of the fitting process using fake overlapping data — a project not covered in this dissertation. The preliminaty Monte Carlo results indicate that the weights outlined above will give reasonable answers. If one is less than happy with the admittedly sketchy arguments here and in Section A2:2 one can restrict consideration to results derived from the independent subsets output by NSCIND.

——                                      ——————————————————

### 3: Removal of Non-Stochastic and Internal Background Variations

In order to analyze and understand the variations in the X-ray sky flux due to unresolved sources, we must understand and remove any other variations in our measurements. In this section we determine the variation of the internal background as a possible source of noise. The internal background mean count rate is measured, and it is shown that for data selected for low values of McIlwain "L", the internal background is well represented by a constant rate modified by a slow secular drift. Any residual variation in the internal background is small. A much larger effect on measurements of the fluctuations is the contamination due to large scale structure in the sky. A thick disk of emission associated with the galaxy is the principal component. A residual "24-hour" signal is seen with only slight significance, ~95%. This signal is consistent in direction and magnitude with that from motion of the solar system with respect to the cosmic background as indicated by similar anisotropy measurements in the microwave.

Two sources of contamination must be examined before we are ready to fit our measurements to models for $P_{I'}$. An obvious source, the internal background, must be measured and an estimate of its variation made. This noise is usually assumed to be Gaussian and purely statistical, although neither assumption is generally checked. The second source of noise is large scale non-stochastic variation in the sky surface brightness. Radio measurements of the fluctuations minimize

this signal with sophisticated techniques that also remove drift in the
detector response [Condon and Dressel 1978; Wall, Scheuer, Pauliny-Toth,
and Witzel 1982]. In the X-ray regime, the original approaches to
analyzing the fluctuations in the X-ray regime restricted the area of
sky for analysis to high galactic latitudes, generally |b| > 20°. The
assumption was that anisotropies associated with the galaxy would be
small. This assumption has since been modified by detailed analysis of
the early all-sky satellite surveys [Warwick, Pye and Fabian 1980;
Protheroe, Wolfendale and Wdowczyk 1980]. The first results of fitting
the X-ray sky brightness for a galactic component for data from the HEAO
1 A-2 experiment, have been presented recently by Iwan et al. [1982].
In this section we will fit the class of models considered by Iwan et
al. to a slightly different A-2 experiment data base. We also consider
a possible additional source of large scale anisotropy, a "24 hour"
signal associated with the motion of the sun with respect to the rest
frame defined by the sources of the X-ray sky flux.

## Measurements of the internal background

Estimates of the internal background of the detector were made
using equation [II.12]. The data were taken from the 2-day superposed
scans produced by SUSCAN and were restricted to source-free regions with
|b| > 20°. By source-free, we mean a measurement contains no counts
from a given list of sources. The sources excluded from this high
latitude data were those catalogued by Piccinotti et al. [1982] plus an
additional complex of sources associated with the Magellanic Clouds.
The resulting list provided a complete sample of sources with fluxes
down to ~0.0022 counts $s^{-1}$ $cm^{-2}$ in the HED layer 1 window (~$3x10^{-11}$ ergs
$s^{-1}$ $cm^{-2}$ from 2-10 keV). Some additional bright sources with |b| $\leq$ 20°
were also included. Several datasets of internal background
measurements were created with varying amounts of smear along the scan
path.

The internal background was modeled by a component which varied
linearly with the value of the McIlwain L parameter and a component

which varied linearly with epoch:

$$B(t,L) = B_o + \alpha t + \beta L \qquad\qquad [IV.6]$$

The model parameters are: $B_o$, the mean internal background; $\alpha$, the temporal variation coefficient; $\beta$, the McIlwain L coefficient; and $\sigma^2_B$ the variance of the residual Gaussian noise. The residual Gaussian statistical component was introduced to estimate the residual variations produced by any unmodeled change in the background rate. In order to describe the temporal behavior of the experiment we define the time t in units of half-years from day 330 of 1977:

$$t = (DOY1977 - 330) / 182.6 . \qquad\qquad [IV.7]$$

With this definition $\alpha$ becomes the change per six month in the internal background counts per exposure. $B_o$ is the expected intensity at day 330. This definition of t is especially useful for HEAO 1 measurements of the X-ray sky because 6 months is the interval between scans of the same region. The linear dependence of the background on McIlwain L is consistent with the model that L is a measurement of the cut-off rigidity in incident cosmic rays, due to shielding from the Earth's magnetic field (see the discussion at the end of Section II:5).

The confidence limits of the parameters of such a model were generated using the likelihood statistic. This is discussed in more detail in Section II:5 so we simply state here that the best fit parameters are determined by minimizing the sum

$$\Lambda = \sum_i \left\{ \frac{(B_i - B_o - \alpha t - \beta L)^2}{\sigma^2_i + \sigma^2_B} + \ln(\sigma^2_i + \sigma^2_B) \right\} , \qquad\qquad [IV.8]$$

This is just the sum of the measurements' likelihoods (see equation [III.73]). $B_i$ and $\sigma^2_i$ are the ith measurement of the internal background and its variance due to counting statistics. The acceptability of an alternate model is determined by the difference between its value for $\Lambda$ and the minimum value. Changes in the value of $\Lambda$ are used to determine confidence contours for the model parameters in the same way that Lampton, Margon and Bowyer [1976] used the $\Delta\chi^2$ statistic for models where the variance is known.

Data Preparation                                          Removal of Variations

For both HED 1 and HED 3 the internal background decreased as the mission progressed. For HED 1 the best fit values were $\alpha = -0.085\pm0.007$ counts exposure$^{-1}$ per six months. A strong positive correlation of background versus L is found for data taken with all values of L. The best fit $\beta$ equals $0.92\pm0.12$. When the data are restricted to $L < 1.18$ no significant correlation is observed for either detector. / The best value for $B_0$ is $2.03\pm0.02$ counts per exposure. This is the internal background rate, per "side", i.e. the mean internal background for the large plus small field of view is twice $B_0$. Similar values for $\alpha$ and $\beta$ are found for HED 3. In addition, the HED 3 background exhibits a slight negative quadratic dependence on t (significant at the 99.9% level) not required in fits of HED 1 data. The magnitude of the quadratic coefficient for HED 3 is such that the deviation from a simple linear model was significant only late in the mission.

## Residual variations in the internal background

After modeling the internal background with the two components described above, there may be residual variations which will appear as noise terms in any measurement of the X-ray sky that does not use the large minus small field of view scheme for removing the background (as in [II.2]). In order to simplify the modeling of these residuals we make several assumptions. The most crucial is that the internal background is assumed constant for time scales up to 175 seconds, which correspond to scans of up to 30°. This means variations inferred from bins with 30° smears in the scan direction are equal to the variations expected for smaller bin sizes. This is important since the size of the excess variation is estimated for bins with large smears in order to avoid contamination from variations in the X-ray sky (recall the discussion associated with Figure II.13). This stability assumption is supported by the fact that the 175 second period corresponds to a change in the orbital position of the satellite of less than 10°, so that far from the magnetic anomalies and in regions where McIlwain L is low, the spacecraft geomagnetic environment is only slowly changing. In addition we will calculate the excess variance for a variety of bin sizes.

Figure IV.4: Excess variance of internal background for various amounts of smear

90% confidence ranges are presented for HED 1 (no restriction on McIlwain L), HED 1 (L<1.18), and HED 3 (L<1.18). The excess variance is calculated after removing the best fit models for the internal background variation. These models are: linear dependence on t and L for HED 1, all L; linear dependence on t only for HED 1, L<1.18; and linear and quadratic t dependence for HED 3. All data were taken before day 505 to avoid the "event" in the detector rates at day 544.
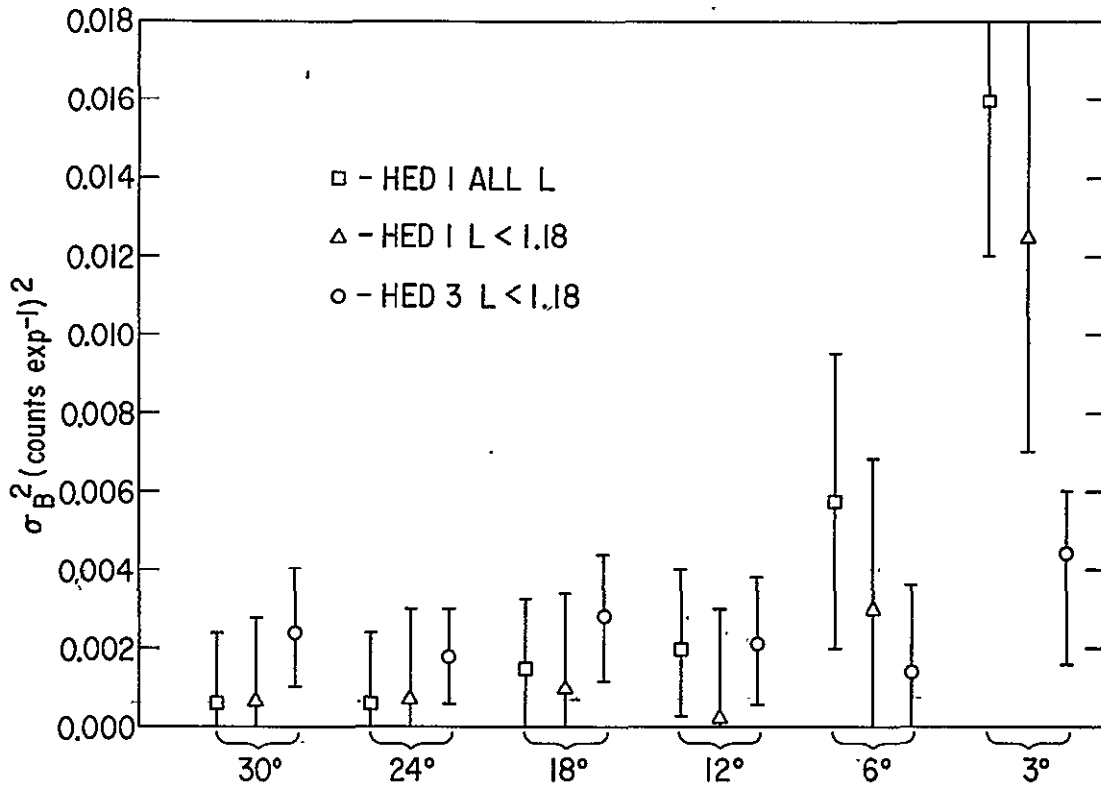


A second assumption we make is that the residual background variation may be modeled as a simple Gaussian with zero mean, meaning there is an extra variation in the measurements beyond counting statistics. This assumption allows us to use [IV.8]. The 90%

confidence region is defined as the range of values for $\sigma^2_B$ such that

$$\Lambda(\sigma^2_B) - \Lambda_{best\ fit} \leqslant 2.7 \ . \tag{IV.9}$$

The results of these fits are presented graphically in Figure IV.4. The HED 1 data for all values of L is consistent at the 90% level with excess variance until the bin size falls below about $6^o$, where the contribution from the sky fluctuations becomes significant. When the data are restricted to the low range of L, the best fit values for $\sigma^2_B$ are consistently lower. The actual excess variance may be smaller, but the confidence region is enlarged because of the increased uncertainty due to the counting statistics of the L-selected data. The excess variance may also be smaller for the L-selected data but not significantly so. However the low-L data are consistent with no variance at the 90% level for bin sizes as small as $6^o$. A reasonable value of the excess variance for the all-L HED 1 data is 0.001 (counts $exp^{-1})^2$, while the corresponding number for the low-L selected data is 0.0007 (counts $exp^{-1})^2$. The $\sigma'$s for these rates correspond to 1.5% and 1.3% of the total internal background, respectively. As a percentage of the total count rate, they are of course even lower: 0.3% and 0.25% respectively for the L+S mean rate of 21 counts $exp^{-1}$. Even if we consider the 90% upper bound values for the larger smears, $\sigma^2_B \sim 0.003$ (counts $exp^{-1})^2$, the size of the variance of the internal background amounts to about half a percent. In comparison, the typical values we find for $W_{eff}$ are of order $\gtrsim 0.1$ counts $exp^{-1}$. Clearly $W_{eff}^2$ is a factor of 10 larger than the best fit values for the internal background variance, and it is at least a factor of 3 greater than the 90% upper bounds.

HED 3 was not so quiet. Even for the low-L data a non-zero value for $\sigma^2_B$ is required at the 90% level, although the allowed range of values is large. A reasonable value for $\sigma_B^2$ is 0.0025 (counts $exp^{-1})^2$. This corresponds to almost 5% of the HED 3 internal background, or $\sim 0.8\%$ of the total L+SFOV flux. The 90% upper limit for $\sigma^2_B$ is roughly 0.004, or about 1% of the total flux.
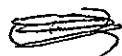
## Large scale structure in the X-ray sky: Galactic models

At energies below 1/4 keV the X-ray sky shows complicated structure associated with thermal emission from interstellar matter within ~100 pc of the solar system [see e.g. Fried, Nousek, Sanders and Kraushaar 1980]. At higher energies the picture is greatly simplified, with the largest anisotropy consistent with a disk of emission associated with our galaxy, beginning at energies as low as ~1 keV [Nousek, Fried, Sanders, and Kraushaar 1982]. The simplest model for such galactic disk emission is an infinite disk, which has intensity

$$I_{ID} \propto csc\ |b|\ . \qquad\qquad\qquad [IV.10]$$

As long as the plane of emission is infinite and the emissivity is independent of the planar coordinates, all models will follow the above cosecant law. Such models have no longitudinal variation in intensity. Iwan et al. [1982] showed that longitudinal variations were in fact observed in the HEAO 1 A-2 data, requiring a more complicated model. They were able to produce the observed longitude and latitude dependence with a finite-radius disk of emission having its center coincident with the galactic center. The X-ray volume emissivity interior to the disk has an exponential dependence on z, the distance out of the plane, while outside the disk radius, $R_d$, the emission is zero. The dependence of intensity on direction of observation, in galactic coordinates $\ell$ and b, is

$$I_{FD} \propto (h/R_{ge})\ csc|b|\ \{1 - exp[-x\ tan|b|\ R_{ge}/h]\}\ , \qquad [IV.11.a]$$

where the longitudinal dependence is embodied in

$$x \equiv cos\ \ell + [(R_d/R_{ge})^2 - sin^2\ell]^{1/2}\ . \qquad\qquad [IV.11.b]$$

$R_{ge}$ is the distance from the solar system to the center of the galaxy, approximately 10 kpc, and h is the scale height thickness of the disk. The infinite disk model of equation [IV.10] is a special case corresponding to the limit of infinite $R_d$. The longitudinal dependence of this model results from the variation of path length through the disk with direction of observation. The longest path through the disk (and therefore brightest surface brightness contribution from the disk) occurs for b = 0°, the shortest for b =180°.

Figure IV.5: Temporal change in X-ray sensitivity of HED 1

The ratios of flux measurements of particular parts of the source-free sky taken six months apart, showing a linear increase in sensitivity of ~0.7%.

X-RAY BACKGROUND (3-60 keV) : LONG-TERM STABILITY

[HEAO A-2 SCAN DATA. HED-I ·(3°x6°)-(3°x3°)]

R ≡ (EPOCH-2 INTENSITY)/(EPOCH-I INTENSITY)

[EPOCH-2→RESCAN OF EPOCH-I INTERVALS,6 MONTHS LATER]

DATA ACCUMULATION INTERVALS: EPOCH-I [(DAY OF 1977)-250]
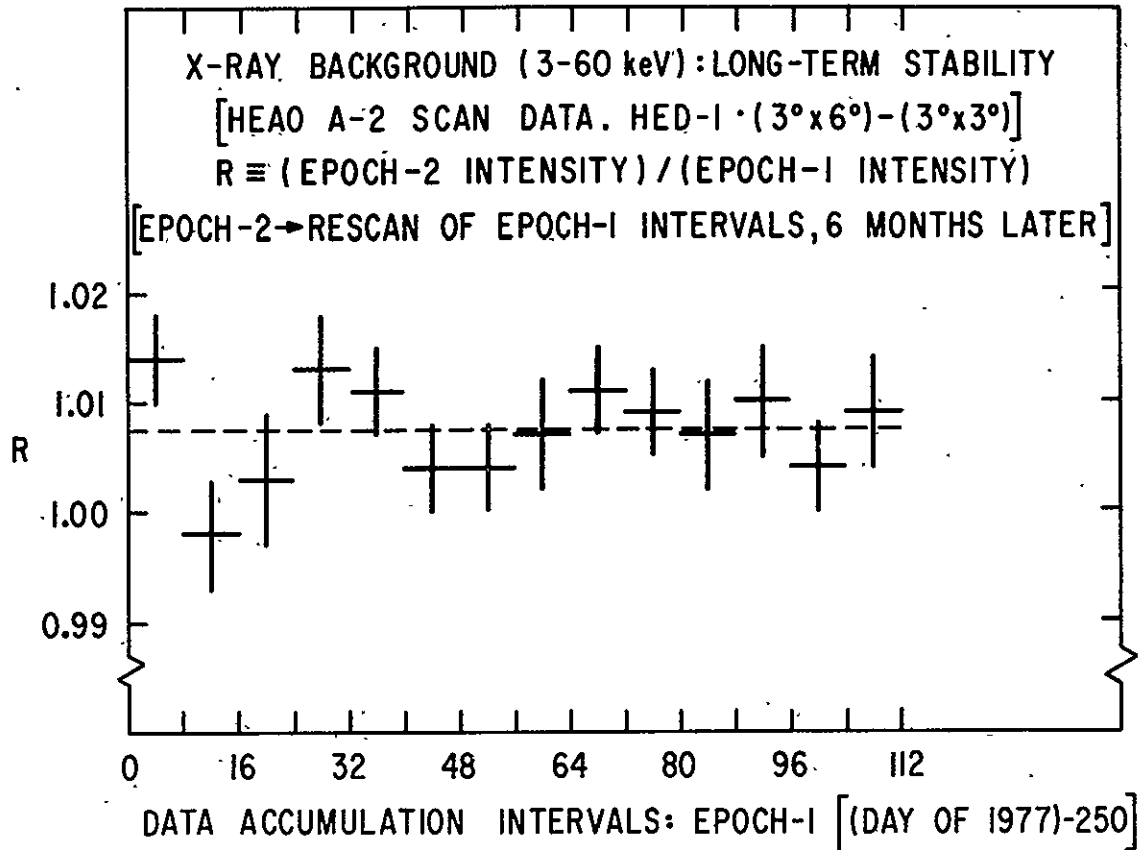
Temporal variability in the detector characteristics may induce erroneous large-scale structure in the sky. We saw above that the internal background data have a slow linear dependence on time. Studies of the internal-background-free large-minus-small field of view measurements show a linear increase in the sensitivity of the detectors to X-rays (see Figure IV.5).

Generating the data for studies of large-scale structure

In order to best utilize the all-sky data, we wish to use as much

Data Preparation                                    Removal of Variations

of each superposed scan as possible. This means using data cells that will sometimes overlap strongly with adjacent cells, particularly near the ecliptic poles where all the scans overlap. We must then weight the data properly see Section II:5 and equation [IV.4].

We make the simplification that the cross-correlation between measurements comes from Gaussian fluctuations with variance $\sigma^2_D$. The first step is to estimate the magnitude of such fluctuations. We first look at very high latitude data, $|b| > 45^\circ$, where the longitude dependence due to a finite disk model is small, and we fit a simple cosecant term for the galaxy, following equation [IV.10]. In addition, we allow for a secular change in the count rate proportional to t. We minimize a likelihood estimator similar to equation [IV.8] to find $\sigma^2_D$, the Gaussian variance over and above counting statistics. Because of the explicit assumption of a Gaussian form when the particular likelihood definition of [IV.8], $\sigma^2_D$ is not identifiable with the variance of the $P_I(I)$ fluctuations distribution. Because of the assumption of Gaussian behavior, its size will depend more on the quasi-Gaussian region of the fluctuations near the mean, and will be less determined by the position of a high flux limit or high intensity survey limit, as is the case for the variance of $P_I$, which is dominated by the high flux tail of the distribution. Because the data are overlapping we cannot use equation [IV.9] to estimate a confidence range on $\sigma^2_D$, but the value of $\sigma^2_D$ that produces the minimum in $\Lambda$ would be an unbiased estimator, if the fluctuating component were properly modeled, i.e. if it were Gaussian. The fact that the fluctuations are not Gaussian will probably bias this estimator, but the effect should be to overweight the overlapping data points. This will make the derived confidence contours conservative; i.e. if the data are overweighted the contour marked "90%" will in fact be more significant.

Using the derived value for $\sigma^2_D$ the data are prepared using the program XTROVL. For the remainder of this section we will restrict our consideration to HED 1 layer 1 L+SFOV data, for times t before day 505, $|b|>10^\circ$, and no contamination from sources with fluxes $\geq 0.0022$ counts $s^{-1}$ $cm^{-2}$. A complete list of sources for the lower galactic latitudes

not covered by the Piccinotti et al. survey was derived [Marshall, private communication]. The derived estimate for $\sigma^2_D$ is 0.145 (counts exp$^{-1}$)$^2$. Besides being used in the weighting process of equation [IV.4], $\sigma^2_D$ is added to the counting statistics variance of each point to get the total variance, which is then used in a $\chi^2$ procedure to determine the best model and the confidence contours for the model parameters. A total of 3835 measurements were made, but the weights reduced the equivalent number of independent measurements to 588.7.

The first model we will fit is a simple csc|b| infinite disk and a term with linear dependence on t. Let $\alpha$ and $\beta$ be the proportionality coefficients of the temporal and galactic components, respectively. The minimum $\chi^2$ for this model is 628.7, a significant drop from the value of 693.8 for a totally isotropic model. The best fit value of $\alpha$ is $-0.07\pm0.03$ counts exp$^{-1}$ per six months. In principle this could be derived from the temporal coefficients from the fits to the measurements of the background rate and X-ray sensitivity separately using the B and L-SFOV count rates, but it is more direct to measure $\alpha$ for the L+SFOV. The predicted value for $\alpha$ would be ~$-0.04$. Subsequent results do not seem to depend strongly on the exact value of $\alpha$. The best fit value for $\beta$ is 0.13 counts exp$^{-1}$.

When we replace the infinite disk model with the finite disk, $\chi^2$ is again significantly improved, dropping by 30.7 to 597.96, for a weighted equivalent of 583.7 degrees of freedom. The x% confidence region for the two parameters, h and $R_d$, is found by increasing $\chi^2$ from the minimum by less than the x% value for $\chi^2$ with 2 degrees of freedom [see e.g. Lampton, Margon, and Bowyer 1976]. $R_d$ and h are strongly correlated parameters, so we instead present the contours in terms of the more separable h/$R_d$ and $R_d$. $R_d$ is normalized by $R_{ge}$. The results are shown in Figure IV.6. The best fit value for $R_d$ is 1.8 $R_{ge}$, and the best fit value for h is 0.42 $R_{ge}$. The coefficient for the model is $0.77\pm0.08$, corresponding to a flux at the galactic poles of 0.32 counts exp$^{-1}$, or about 1.7% of the isotropic flux. The contours overlap the earlier result of Iwan et al. The fact that agreement is not stronger may be due to the different selection of data (Iwan et al. included

Data Preparation                                                    Removal of Variations

Figure IV.6: Finite disk model for galaxy, HED 1 layer 1 L+SFOV

Solid contours enclose the 70% and 90% confidence regions, and x marks the best fit position. The star and dashed contour indicate the best fit and 95% region from Iwan et al.
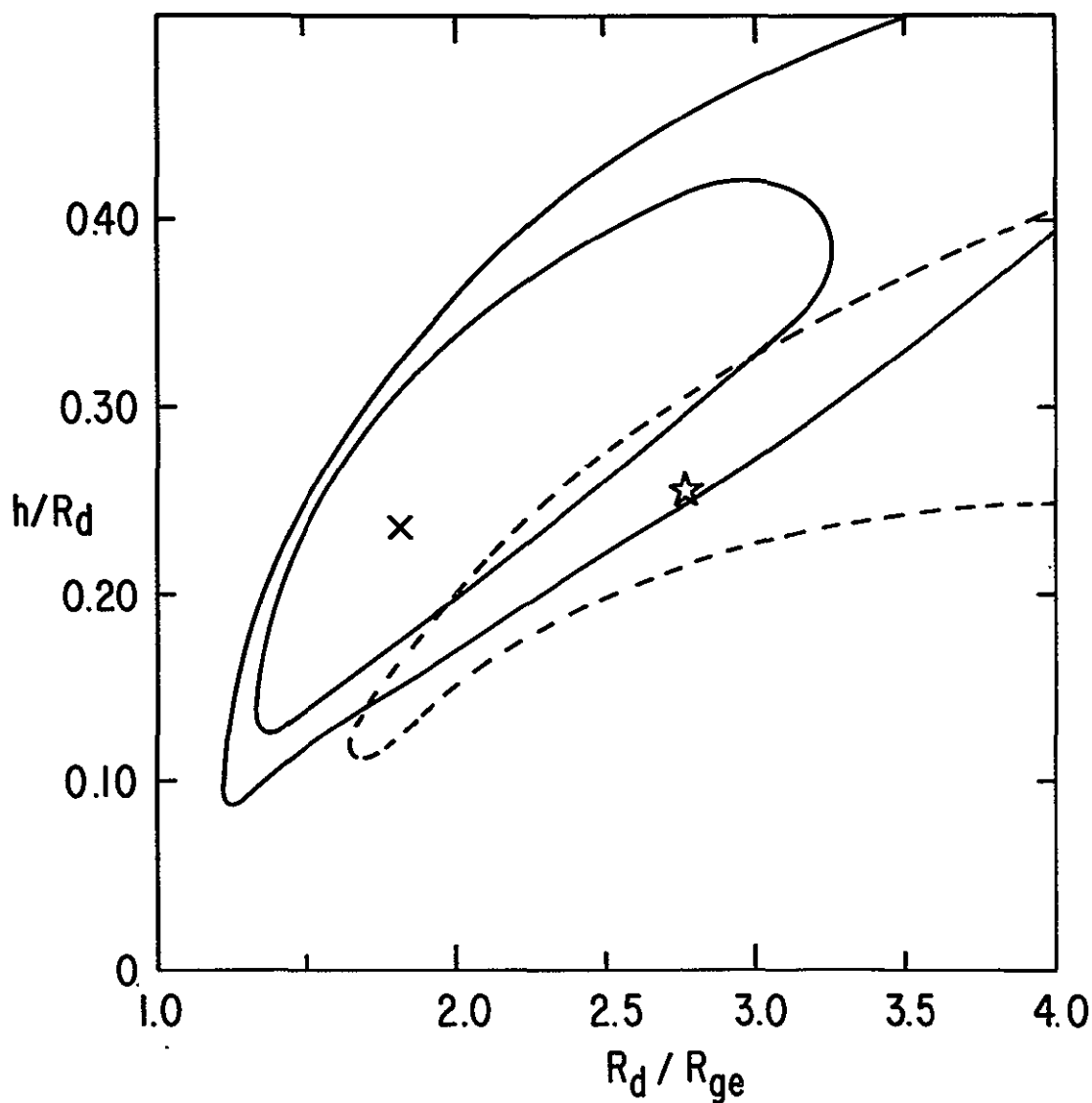
some data closer to the galactic plane), or the fact that the previous analysis did not consider impact of the overlap (although the variance due to the fluctuations was estimated and included in the total variance, as was done here). The smallest possible value for the scale

height is $\sim 0.1$ $R_{ge}$ (1 kpc) which, although smaller than the smallest value of Iwan et al., is still significantly thicker than most galactic disk phenomena. Possible interpretations and physical models for such a large scale height disk are presented in more detail in Iwan et al. Additional galactic components or more complicated models that account for the longitudinal variation may reduce the size of the scale height somewhat. (For an example of such a model applied to UHURU data see Protheroe, Wolfendale and Wdowczyk [1980].) The finite disk model was explicitly considered by Warwick, Pye, and Fabian [1980] using data from Ariel V. They concluded that there was an upper bound of $\sim 1$ kpc for h. However, their fits were restricted to a latitude range, $|b|>20^{\circ}$, versus our fits which include data down to $|b|=10^{\circ}$. The longitudinal dependence, which is the distinguishing property of the finite disk model, becomes stronger as the measurements approach the galactic plane.

### Additional large-scale variations: The Compton-Getting effect

The minimum $\chi^2$ per degree of freedom for the finite disk model is 1.03, indicating that the fit is acceptable without another component. However it has long been noted that if the bulk of the X-ray emission comes from outside the galaxy, the motion of the sun with respect to the rest frame defined by this emission should shift the spectrum of the X-ray sky depending on the direction of the measurement relative to the direction of motion. Measurements in the direction of the sun's motion are affected by a blueshift of the spectrum, while in the antipodal direction the spectrum is redshifted. If the spectrum is well-fit by a power law form with photon index $\Gamma$, such that $n(\nu)\ d\nu \propto \nu^{-\Gamma}\ d\nu$, then for $\Gamma$ greater than $-1$ the blueshifting will enhance the measured number of counts in the direction of motion. An additional effect is aberration which will increase the number of counts in the direction of motion irrespective of the spectrum. The total effect, the Compton-Getting effect, was first described in the context of anisotropies in the cosmic ray flux. The total angular dependence of the intensity is a simple cosine law,

$$I_{CG} = \overline{I}_{sky}\ (2+\Gamma)\ (v/c)\ \cos\theta\ , \qquad\qquad [IV.12]$$

where v is the velocity of motion and $\theta$ is the angle between the look

Data Preparation                                        Removal of Variations

axis and the direction of motion. $\bar{I}_{sky}$ is the mean sky intensity (not including internal background). Such variation is sometimes referred to as a "24 hour" anisotropy because of its $360^{\circ}$ periodicity. This cosine dependence is seen in the anisotropy of the microwave background. If the rest frames for the microwave background and the X-ray sky are the same, we would expect to see a signal in the same direction with a magnitude on the order of half a per cent.

A cosine signal of the appropriate magnitude is detected, but at a statistically marginal level. To make a simple search for the cosine signal we first fixed the temporal and finite galactic disk contributions to the sky flux. The best fit position for the direction of motion was then found to be $(\ell, b) = (282^{\circ}, +30^{\circ})$, with a best fit value for the magnitude of the signal of $0.09\pm0.03$ (one-sigma) counts $\exp^{-1}$. The X-ray background spectrum can be fit (over the range from which of the counts for the layer 1 HED window come) by a power-law with $\Gamma \approx 1.4$. Thus the coefficient corresponds to a velocity of $475\pm165$ km $s^{-1}$. This is roughly in agreement with the current microwave-deduced velocities of $\sim 340$ km $s^{-1}$. The confidence contours for the position are presented in IV.7. The acceptable positions cover a large fraction of the sky, roughly $(2/3)\pi$ steradians. A result similar in terms of direction, magnitude, and size of confidence region, was determined from UHURU data by Protheroe, Wolfendale and Wdowczyk [1980]. A "12 hour" anisotropy (i.e. one that is low at both poles and high around the equator) in the same direction was reported by Warwick, Pye and Fabian [1980] based in Ariel V data. This is not observed in our data. (For a discussion of the impact of the Compton-Getting signal on cosmological issues see Fabian [1981] and Section VII:6 of this dissertation.)

How significant is our detection? The new minimum $\chi^2$ is 589.7, a drop of only 8.2. Any time a new parameter is added to a model, the minimum $\chi^2$ is expected to drop, so the question is how unlikely is a drop of this magnitude when three new parameters are added (one for the velocity and two more for the direction). One statistical tool used to assess the significance of the addition of a new term to a model is the F-statistic (see e.g. Bevington[1969]. If $\Delta\chi^2$ is the change in $\chi^2$ with

Figure IV.7: 70% and 90% confidence contours for a Cos θ (dipole) signal

Confidence region for the dipole maximum is given in galactic longitude and latitude. Also marked are the galactic center and the position of the center of the Virgo Cluster, M87, which is also the center of the local supercluster. The 3 crosses are the one-sigma error bars for three different determinations of the microwave radiation. References: BCW - Boughn, Cheng, and Wilkinson [1981]; FGMN - Fabbri, Guidi, Melchiorri, and Natale [1980]; GS - Gorenstein and Smoot [1981].
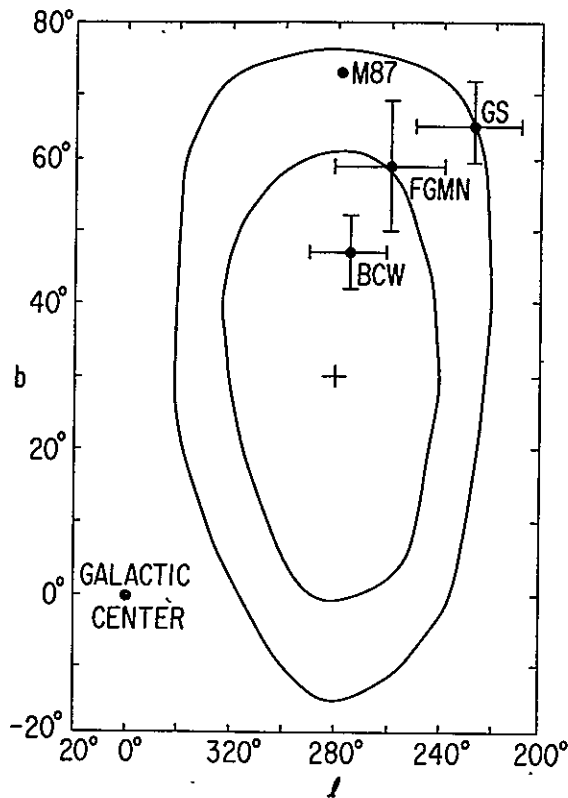


$v_1$ new parameters, and $\chi^2$ is the new minimum with $v_2$ degrees of freedom, then if the model was correct before the addition of new parameters, the quantity

$$F \equiv (\Delta\chi^2/v_1) \, / \, (\chi^2/v_2) \qquad\qquad [IV.13]$$

is distributed as an F-statistic with $(v_1, v_2)$ degrees of freedom. As an

Data Preparation                                     Removal of Variations

example, in going from the infinite to a finite disk we added two new parameters. The value of the F-statistic was 15. A value that large is expected very infrequently, less than one time in a million, so the addition of new parameters is considered significant (or rather, the old model was significantly worse). For the addition of the Compton-Getting term, the value of the statistic was 2.7, which for $\nu_1$ of 3 and $\nu_2$ of 580.7 is expected one time in twenty (95% confidence).

Apart from issues of statistical weighting there is the question of whether other large scale features may be interfering with the measurement of the Compton-Getting effect. Its low magnitude (~0.5%) means that it might be easily affected by other low intensity structures, such as second order structure associated with the galaxy, or perhaps large scale diffuse emission associated with the local supercluster. Also, an asymmetric galactic contribution to the 2-10 keV flux is expected at some level. On the other hand, it is suggestive that the most significant direction and magnitude for the Compton-Getting effect is roughly in agreement with the microwave result. We present further analysis of the dipole result and its comparison with the microwave derived velocity in Section VII:6.

Without dwelling on the reality of particular physical interpretations, the results of this section can still be used in the analysis of the fluctuations. That is, we can subtract the sky variations that we know are not associated with unresolved extragalactic point sources without identifying them with particular physical phenomena such as the galactic disk or the Compton-Getting effect.

———————————————

4: Optimal Bin Sizes for Subsets of Independent Data

The tradeoff between improved counting statistics, number of observations, and size of the fluctuations signal is considered for a simple model for the division of the superposed data scans into subsets of independent data. The likelihood difference technique is used to compare the

sensitivity of different subsets, depending on the amount of
smear along the scan direction (bin size) as well as various
detectors and their possible field of view combinations. It
is shown that the sensitivity for measurements of W is a
remarkably insensitive function of the bin size. For any
given detector the optimal combination of fields of view is
the L+SFOV. The two HEDs are roughly equivalent with HED 3
marginally superior. The optimal bin size for the L+SFOV is
a smear approximately equal to the FWHM of the unsmeared
detector.

In discussing the ultimate limitations to using the fluctuations
for understanding the behavior of low flux sources we introduced the
likelihood difference. This measures the difference between two models
and in addition allows us to predict the sensitivity of a set of
observations with a given number of measurements and a particular signal
to noise ratio. This was used in an earlier section to analyze the
design of ideal experiments. These experiments consisted of
measurements by a flat response detector which made pointed, i.e.
unsmeared, observations. In addition the point positions were assumed
to be selectable so as to totally divide the available sky into non-
overlapping independent measurements. Because of the scanning nature of
the A-2 experiment all-sky survey these assumptions do not apply. We
therefore customize the earlier analysis to apply to the special
circumstances. The derived sensitivity corresponds to the expected size
of the 90% confidence region for various parameters of N(S) models: W,
$\gamma$, and $\sigma^2_s$. This size is still an idealization, the actual confidence
regions may differ, but it allows us to optimize the choice of detector,
field of view combination and amount of data smear.

## Optimal bin size

The non-overlapping portions of the scans were selected by the
NSCIND program with various amounts of re-binning. The increase in the
number of $1^o$ bins per measurement will increase the total amount of time
the detector spent in making that measurement, and will therefore result
in smaller counting statistics. This reduces the size of $\psi$, helping to
increase the sensitivity of the subset. However at the same time, the

Data Preparation                                        Optimized binning

re-binning is equivalent to smearing out the detector which reduces the size of $R_{\gamma-1}$. This in turn reduces the size of the fluctuations signal, as measured by $W_{eff}$ (see equation [III.47]), so that $\psi$, the ratio of the non-fluctuations noise to the fluctuations signal may not decrease as fast as we might otherwise think. Offsetting the trend of decreasing $\psi$ is the fact that as the bins are enlarged, the number of possible independent measurements that may be created in a particular subset is decreased. This decreases the sensitivity.

The relevant parameters for determining the optimal bin size include: the total amount of scan angle in the non-overlapping portions of the scans selected by NSCIND (~1800 degrees); the total active time per scan (on average it was about one day for data unconstrained for L or ~1/2 day for data with L<1.18); the mean number of counts per exposure, used in calculating counting statistics (this depended on the detector/FOV combination, but recall that the L-SFOV has the same magnitude count rate for purposes of calculating the uncertainty for counting statistics); and finally (and most difficult to quantify) the basic detector collimator geometry, which is smeared by increasing amounts as the bin size is increased, and which determines the size of the fluctuations signal. The first parameter, in conjunction with the non-zero response size of the smeared detector, $\delta\theta$, fixes the number of independent observations. The bin size, total active time, and count rate allow an estimate of the size of the counting statistics. We fix the value of W to correspond to the prediction from the extension of the resolved source counts, 0.06 counts $s^{-1}$ $cm^{-2}$. The likelihood difference will show how far, on average, a test value of $W_{eff}$ can diverge from the true $W_{eff}$ and still be acceptable at the 90% level.

When we apply this simple model we find that there is a very broad minimum in the expected size of the 90% confidence range for W as a function of the re-binning size (see Figure IV.8). The very small binnings are not preferred because of their very small number of exposures per measurement, yielding large noise from counting statistics. Eventually the very largest binnings are not preferred when the number of observations per sub-sample of independent data becomes

Figure IV.8: Fractional size of 90% confidence of region for W as a
function of bin size

The parameters used are described in the text. The L-SFOV
uses the all-L data, while all other field of view
combinations use the restricted L data, which impacts the
total active time available per scan. The four field of
view combinations for HED 3 are plotted, as well as the
curve for the L+SFOV HED 1 rate.
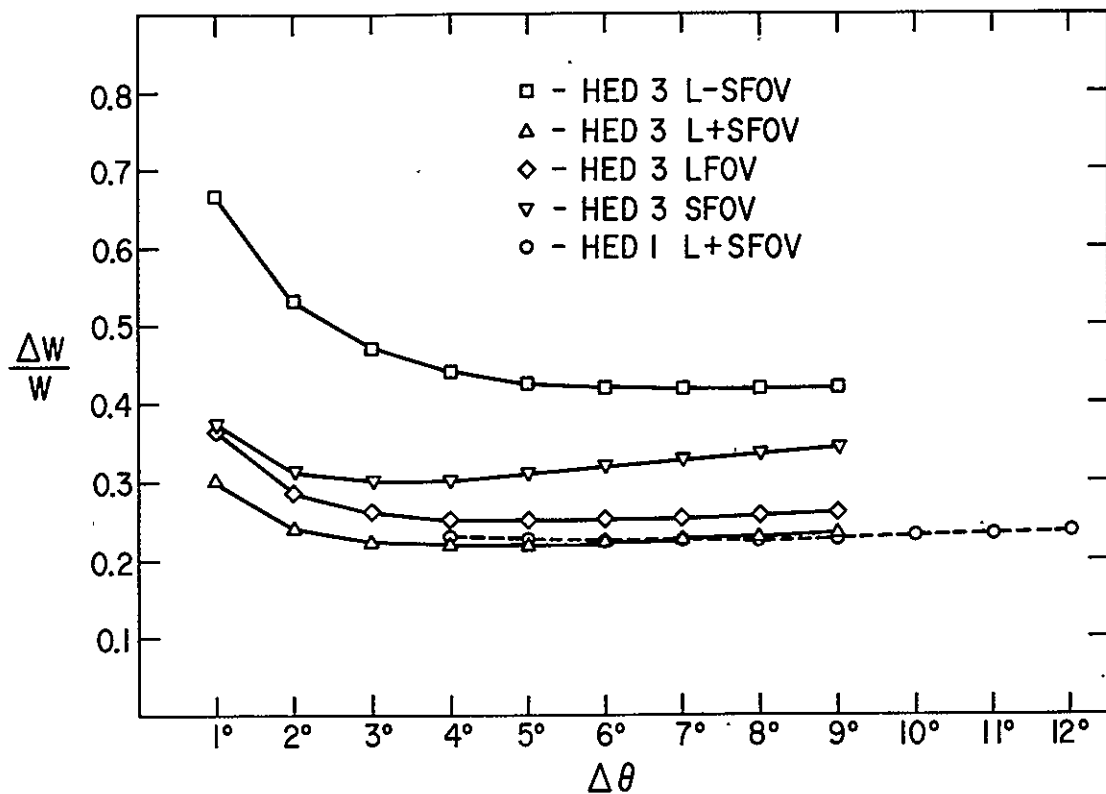


small. But as we wish to maintain our sensitivity to the behavior of
low flux sources we will prefer the smaller binning end of the broad
minimum. In comparing the different field of view combinations we see
that the increased number of exposures available to the L-SFOV data
(because of the acceptability of data with $L>1.18$) does not compensate
for the decrease in the size of $W_{eff}$. The smaller size of the SFOV
allows the available space to be divided into more observations, but at

Data Preparation                                            Optimized binning

the cost of poorer counting statistics compared to the L+SFOV. Therefore in general the L+SFOV is preferred. When we compare the relative merits of HED 1 and HED 3 we see that HED 3 is slightly preferred. Note, though, that the availability of the L+SFOV combination as an accurate, relatively uncontaminated measure of variations in the X-ray sky flux is due to our assessment of the stability of the internal background rate of the detectors. As HED 1 appears slightly "cleaner" than HED 3, we will concentrate on the analysis of the HED 1 L+SFOV data in the remainder of this dissertation. The optimal bin-size for this detetector and FOV combination is about the size of the unsmeared detector FWHM, ~6°. Such a choice allows us to divide the scans into 3 separate subsets.

If we search for the optimal bin size and detector for the determination of other model parameters, such as $\gamma$ and $\sigma^2_s$, the story is little changed. HED 3 and HED 1 are roughly equivalent. L+SFOV is the preferred field of view combination, and the optimal bin size is ~6°. The simple analysis outlined in this chapter suggests that with this binning such a dataset should be able to detect excess variances, $\sigma^2_s$, on the order of $4 \times 10^{-4}$ (counts $s^{-1}$ $cm^{-2}$)$^2$. This limit presumes that all other aspects of the model are known, i.e. the high flux behavior of N (W and $\gamma$) is known precisely. We will see that for the independent datasets described in this chapter, it is the imprecise knowledge of W and $\gamma$ that strongly limits the utility of non-overlapping datasets for constraining the excess variance.

CHAPTER V

HOW TO PREPARE FOR AND INTERPRET FITS OF OBSERVATIONS TO N(S) MODELS

———————————

1: Photon Statistics and the Probability Distribution used in Fitting

> The $P_I$, distributions cannot be directly compared to the distribution of measurements as they do not include the additional variation produced by counting statistics. The final distribution is a convolution of the $P_I$ distribution with a Gaussian representing the additional noise. A particular difficulty is encountered, though, when the size of the added variation changes from one measurement to the next. This means that each measurement is drawn from a unique distribution. A further refinement is necessary if in order to minimize the effect of bright sources on model fitting, a high intensity selection cutoff is imposed on the data. The final result of this section is a derivation of the probability distributions that properly describe the data set.

In Section III:3 we derived general tools to calculate the expected distribution of ideal intensity measurements of the sky: $P_I(I;\vec{\alpha})$. The $\vec{\alpha}$ represents a general vector of two kinds of parameters: those that scale and position the distribution, such as $W_{eff}$ and $\bar{I}$, and those that define the shape of the distribution, such as $\gamma$, $\xi_c$, and $\sigma^2_\xi$. In order to constrain the range of parameters we must compare a model distribution to the observed set of data points, $\{r_j\}$. The difference in notation, the rates $r_j$ versus the distribution of the intensity I, is deliberate. Because of the finite area of the detector it is not possible to perfectly and reproducibly measure the intensity. Counting statistics will impose an additional variation on the measurements beyond the model generated $P_I$. Accounting for this statistical component is relatively straightforward.

Let us consider the jth measurement, $r_j$. It is a random variable drawn from the "noise" distribution $P_{noi}(r;I,...)$, where I is the (unknown) true intensity. A common assumption is that this noise distribution is a Gaussian about I, with a known variance $\sigma^2_j$. This variance includes photon counting statistics as well as intrinsic non-sky variations from the internal background. The distribution, then, is given by:

$$P_{noi}(r;I,\sigma^2_j) = (2\pi\sigma^2_j)^{-1/2} \exp [ (r-I)^2 / 2\sigma^2_j ] .\qquad\qquad \text{[V.1]}$$

As the exact value of I is not known, but is described by its own probability distribution $P_I(I)$, $r_j$ is a random variable drawn from the convolution of the two distributions:

$$P_j(r) = \int_{-\infty}^{+\infty} dI\ P_I(I)\ P_{noi}(r;I,\sigma^2_j) .\qquad\qquad \text{[V.2]}$$

This is the probability that $r_j$ is in the interval from r to r+dr.

The assumption that $P_{noi}$ is a Gaussian implies, ignoring the variation in the internal background, that the total number of photons counted in the rate $r_j$ is large enough that the Poisson distribution is well-modeled by a Gaussian. This is generally true for measurements involving more than ~25 photons. If the measurements involve fewer counts, it is more appropriate to replace the continuous form of $P_{noi}$ with a discrete probability distribution and the integral in [V.2] with a sum.

Another assumption is that besides measuring the rate, $r_j$, we also measure (or are otherwise given) the variance, $\sigma^2_j$. If the rate of counts per exposure for the jth measurement is $r_j$ and the total number of exposures in making that measurement is $n_j$ then the counting statistics are described by (totally ignoring the internal background):

$$\sigma^2_j \approx r_j / n_j.\qquad\qquad \text{[V.3]}$$

This is an approximation since the true value of $\sigma^2_j$ is found by replacing $r_j$ with the true intensity I. To properly perform the

convolution in equation [V.2] to produce $P_j$, $\sigma^2_j$ is a function of the integrating variable I. Instead, estimating I with $r_j$ to find $\sigma^2_j$ is adequate if the range of values for I is small compared to its absolute value; that is, when the fractional size of the fluctuations is small. When the fluctuations cover a dynamic range comparable to the intensity itself then the explicit dependence of $\sigma^2_j$ on I should be properly included in the integral.

The indexing of the distribution $P_j(r)$ is intended to emphasize an important point. If the different measurements are made with varying numbers of exposures, the size of the $\sigma^2_j$ is not fixed. Each measurement is drawn from a different parent distribution. In cases where the distribution is a pure Gaussian, the traditional procedure has been to divide each measurement by its differing $\sigma$, thereby generating a set of random variables all drawn from a unit normal distribution. When the distribution is not Gaussian, as is the case for $P_j(r)$ even after convolving with the Gaussian noise, the unique shapes of the different distributions may not be easily transformed into a single equivalent form.

Standard tests of the goodness of fit of a model will not work; in particular the histogram technique of binning the data and comparing it to the number of observations expected in each bin is not valid. Though such a histogram is still useful in illustrating the range of intensities for which the model provides a good description of the data, all the observations must be drawn from the same parent distribution in order for the histogram to be statistically interpreted properly.

Still any collection of measurements drawn from a variety of probability distributions may be transformed into a unified description, the integral probability, which describes how well the distributions predict the observed measurements. Given a set of measurements and associated distributions, $\{(r_j, P_j(r))\}$, we can transform each pair into

a single number via

$$P(r_j; P_j(r)) = \int_{-\infty}^{r_j} dr \, P_j(r) \; .  \qquad [V.4]$$

We abbreviate $P(r; P_j(r))$ as $\mathbb{P}_j(r)$. Because of the normalization of the differential probability distribution $P_j$, $\mathbb{P}_j(r)$ is a number between zero and one. To say that $r_j$ is drawn from $P_j$ is equivalent to saying that $\mathbb{P}_j(r_j)$ is _uniformly_ distributed from zero to one. In subsequent sections we will examine how this transformation can be used to test the goodness of fit of a particular model.

### Censored data

When we derived the $P_I$ distribution for the fluctuations from the number versus flux models we explicitly included the effect of the bright sources above the resolution limit of the experiment. They contribute mostly to the high intensity tail of the $P_I$ distribution. Thus the measurement set should include data that might be contaminated by these same sources. But we are interested in learning as much _new_ information as we can about the N(S) distribution, so we want to reduce the impact on our model fitting of the already studied resolved sources. We therefore eliminate from consideration any measurement with a rate greater than some threshold value. A high flux source will only contribute to a counting rate (and not push it above the threshold) when it is positioned at the corner of the field, where the response is low. This rate threshold greatly reduces the sensitivity of the fitting process to the _shape_ of the high intensity $P_I$ distribution, and equivalently, the shape of the high intensity portion of the N(I) curve.

If we restrict the data, we must change the model distribution to reflect the process of selection. Here the selection is in terms of the observed rates (instead of a bright-source-in-field-of-view criterion). Let $r_u$ be the maximum allowed rate for an observation to be included. If $P_j$ is the old distribution, which may have included observations with

r greater than $r_u$, then the new distribution is

$$P'_j(r) = P_j(r) \,/\, P_j(r_u) \quad , \; r < r_u;$$

$$= 0 \qquad\qquad r > r_u \;.$$

<div align="right">[V.5]</div>

The shape of $P'_j$ is the same below $r_u$, but it has been renormalized so that $P'_j(r_u)$ is 1. The set of values of $P'_j(r_j)$ and $P'_j(r_j)$ will be used to fit models for the fluctuations.

---------------------

## 2: Statistical Techniques Used

Model fitting for the fluctuations requires care in the selection of the proper statistical tools. The integral probabilities, $P_j(r_j)$, may be used in a family of statistical tests that will tell the goodness of the fit of the model. For the determination of confidence contours the likelihood statistic is the preferred technique. We also outline a general rigorous approach to combine results from several non-independent determinations of a confidence contour based on Boole's Inequality.

We have turned the raw observations into differential and integral probabilities. Yet these numbers do not immediately tell us what we most want to know about model's representation of a set of data. Is the fit good? Do these particular values for the model parameters have a chance of being the true values? (These two questions are not identical, but are inter-dependent in subtle ways, as we will see.) In this section we outline the particular tools that we use for evaluating models in the remainder of the dissertation.

<div align="right">Measuring goodness of fit</div>

Given a particular model, with its parameters $\vec{\alpha}$ considered fixed, we can for our set of data generate the differential and integral probabilities of the observations, $P_j(r_j)$ and $P_j(r_j)$. (From now on we will refer to the distributions after renormalizing for the effects of the upper intensity cutoff and dispense with distinguishing the primed

from the unprimed distributions.) When we ask whether the model is a good fit, we ask if the observed collection of measurements is one that has a significant chance of occurring. The answer is obtained by distilling the observations in the context of that model into one or two numbers, i.e. statistics. This statistical distillation has the property that if the model is correct then we know the distribution properties of the statistic. Let $S(\{x_i\};\vec{\alpha})$ be such a statistic, written in this fashion to show that it is a function of the observations interpreted in terms of a particular model. From the distribution properties of S, we define an acceptable range of values for S. Let $P_S(S)$ be the differential probability of S and C be given by

$$C = \int_{S_{min}}^{S_{max}} dS \ P_{S(S)} \ .$$

[V.6]

Then if the value of S for a particular set of measurements falls within the range from $S_{min}$ to $S_{max}$ the model is said to be acceptable at the "C" level. For example, for a C of 0.90 we find an $S_{min}$ and $S_{max}$ that satisfy [V.6]. If a particular value of S lies outside this range we say the model is unacceptable. The chance that this is incorrect, i.e. that the model which correctly describes reality is mistakenly rejected, is only 1-C or 10%. The actual choice of the interval $S_{min}$ and $S_{max}$ is totally arbitrary, but generally it is chosen a priori to minimize the size of the range of model parameters that gives an "acceptable" fit.

For observations drawn from normal distributions, the $\chi^2$ statistic is often used for S, but for observations drawn from more complicated distributions, the approach is to transform the measurements. Using the latter method here, we estimate the goodness of fit not directly from the distribution of rates $r_j$ and their separate distributions $P_j$, but from the behavior of the integral probabilities, $P_j(r_j)$. If the model is a correct description of reality the $P_j$ should be drawn from a uniform distribution from zero to one. There are several statistical tools to test such a hypothesis. One which is often used in the astronomical literature is the Kolmogorov-Smirnov test (for an application see Piccinotti et al. [1982]). The difficulty with that particular test is that it measures almost exclusively the goodness of

fit. That is to say, if the fit is not quite good enough the K-S statistic contains little information about how or where the model might be changed to improve the fit, whereas a statistic like the $\chi^2$ statistic at least indicates the data points whose individual contributions may be the cause of an overall too high value.

We use an alternative set of statistics, the moment statistics. Given a set of observations, $\{u_i\}$, we define the nth moment statistic by

$$\mu_n \equiv (1/N) \sum_i (u_i - \langle u \rangle)^n \ . \qquad [V.7]$$

N is the total number of measurements and $\langle u \rangle$ is the expectation value of u. Given the cumulants of the supposed distribution for u we can calculate the cumulants for $\mu$ which in turn can be used in an asymptotic expansion to give $P_\mu$ (see equation 26.2.47 in Abramowitz and Stegun [1965]). If the $u_i$'s are drawn from a uniform distribution from zero to one the expected value of $\mu_n$ is zero for n odd, and $1/(n+1)$ for n even. In assesing the goodness of fit of a model of the fluctuations, we test the moment statistics for the $P_j$, to see if they are consistent with the expected. uniform distribution. Consistency is measured by the comparison of the observed moment statistic with the expected distribution, using the asymptotic expansion for the distribution.

The advantages of such statistics over the K-S test are ease in calculation and the interpretation of their values when a not-good fit is indicated. For example, if a set of measurements has $\mu_2$ significantly less than 1/3 it means that the $u_i$s are too concentrated towards the center, indicating that the model being examined might have too large a variance. If $\mu_1$ is much larger than zero then the mean of the model should be adjusted. The behavior of the $\mu_n$ does not directly correspond to a statement about the nth moment of the model, although it is an indicator. An additional caution: the $\mu_n$ are not independent measurements of the goodness-of-fit and can be highly correlated functions of model parameters. A disadvantage of the moment statistics, relative to the K-S statistic, is that they are sensitive to restricted aspects of the distribution. Clearly bad models might give acceptable

values for some of the $\mu_n$. For example, if the model variance is set too large all the $u_i$ will tend to 0.5, giving a perfect zero value for all the even $\mu_n$. (Of course the odd $\mu_n$ also tend to zero, which for them indicates a bad fit.)

## Parameter confidence regions

We consider a class of models described by a parameter vector $\vec{\alpha}$ and we assume that the true model is given by the parameter $\vec{\alpha}_0$. Then a 90% confidence region is a subset of the possible values of $\vec{\alpha}$ that has a 90% chance of containing $\vec{\alpha}_0$. Just as there was no unique measure of goodness of fit (or even of the acceptable value range of a particular statistic) there are many different ways to derive a confidence region. In our analysis of the fluctuations we use the likelihood technique to define our confidence regions. (For an introduction to the technique in an astrophysical setting see Cash [1979].)

Consider a model differential distribution $P(x;\vec{\alpha})$, the $\vec{\alpha}$ indicating the dependence of the distribution on the model parameters. If we have a particular set of $N$ observations, the likelihood that a model describes the observations is the product of the differential probabilities:

$$L(\vec{\alpha}) = \prod_i P(x_i) \ . \qquad \text{[V.8]}$$

As $L$ increases, we say the model is more likely. In very general situations, by maximizing $L$ we can calculate an efficient estimator of the true model's parameters $\vec{\alpha}_0$ [see Wilks 1963].

For purposes of defining confidence regions we use the log-likelihood,

$$\Lambda(\vec{\alpha}) \equiv -2 \ln L = -2 \sum \ln P(x_i;\vec{\alpha}) \ . \qquad \text{[V.9]}$$

With this definition, better models are found by minimizing $\Lambda$. Let $\Lambda_{min}$ be the minimum value for the entire range of values that $\vec{\alpha}$ can take. We can define a region in parameter space by all those models that have a

value less than or equal to $\Lambda_{min} + \Delta\Lambda$. Wilks [1938] showed that when the probability distribution and the parameter space satisy some very general constraints, the probability that the true parameter values, $\vec{\alpha}$, is contained in that region is distributed roughly as $\chi^2$ with m degrees of freedom, where m is the dimension of $\vec{\alpha}$. The formula is exact when the probability distribution is Gaussian and the variance is independent of the parameters. The formula is asymptotically correct for large number of observations. What constitutes a large number is not easily determined but $m \gtrsim 20$ is generally considered sufficient for most distributions. (However see Henry, Soltan, and Briel [1982] for an example where the confidence region was well determined for as little as five observations drawn from a very non-Gaussian distribution.) Strictly speaking, Wilk's theorem applies to observations which are all drawn from identical distributions, but relaxing this condition only slows the convergence of the distribution of $\Delta\Lambda$ to the asymptotic $\chi^2$ form. The use of the $\Delta\Lambda$ confidence region is similar to the use of the $\Delta\chi^2$ confidence test as outlined by Lampton, Margon, and Bowyer [1976].

Not all of the elements of the parameter vector may be of interest. We divide $\vec{\alpha}$ into two parts, an m'-dimensional vector of interesting quantities, $\vec{\alpha}'$, and the remainder nuisance parameters in the vector $\vec{\alpha}''$. $\Lambda_{min}$ is still defined as the minimum value of $\Lambda$ over all m parameters. The confidence region in the m'-dimension space is defined by all $\vec{\alpha}'$ such that at least for some value $\vec{\alpha}''$, $\Lambda(\vec{\alpha}',\vec{\alpha}'')-\Lambda_{min}$ is less than $\Delta\Lambda$. The probability that the true values for the interesting parameters, $\vec{\alpha}'_o$, is contained in this region is determined by the value for $\Delta\Lambda$ which by Wilks' [1935] theorem is asymptotically distributed as $\chi^2$ with m' degrees of freedom [Avni 1976; Cash 1976].

This derivation of a confidence region is the root of the behavior of the likelihood difference (see equation [III.71]). We claimed in equation [III.77] that a model with a likelihood difference $\Delta\lambda$ tested with a dataset containing N observations will in some average sense be outside the 90% confidence region when $N\Delta\lambda$ is greater than $\chi^2_{m;0.90} - m$, where m is the number of model parameters. The origins of this claim are now clear. $N\Delta\lambda$ is the expectation value for the difference between

the likelihood of the test model and the true model. The mean value of $\Delta\Lambda_o$, the change in likelihood from the minimum value to the value associated with true model, is the mean of a $\chi^2$ distribution with m degrees of freedom, which is m. The expected value for $\Delta\lambda$ for the test model is roughly $m + \Delta\lambda$, which if it is larger than the 90% value for $\chi^2$ with m degrees of freedom would place the test model outside the 90% region.

It is the strong similarity of the $\Delta\Lambda$ tool to the $\Delta\chi^2$ technique for defining confidence regions that suggests that non-independent data can be handled in a similar way, i.e. by weighting the contribution of each point to the sum:

$$\Delta\Lambda = \sum W_i \, \lambda_i \; , \qquad\qquad\qquad [V.10]$$

where the weights, $W_i$, are defined in [IV.4]. However one important difference in the use of $\chi^2$ and $\Lambda$ is that for normally distributed data $\chi^2$ can be used as an absolute measure of the goodness of fit. $\Lambda$ is defined to within an additive constant and its absolute magnitude is without statistical meaning. Only the difference of two models' log-likelihoods contains information.

A caveat concerning the application of $\Delta\Lambda$ (and $\Delta\chi^2$) to the derivation of confidence regions often appears with their use. The validity of the confidence region is predicated on the true model being one of the elements of the space spanned by the total range in $\vec{\alpha}$. This condition can be weakened somewhat. If we just insist that one of the models of the space is indistinguishable in the sense of the likelihood difference from the "true" model for the given number of observations then the confidence region may be said to include those models which provide a good representation of the data, even if they do not correspond exactly to the details of an underlying reality.

### Non-independent measurements and Boole's Inequality

Even if the use of weighted non-independent data seems undesirable because of the absence of a rigorous statistical justification we still are not restricted to considering only the small fraction of the sky

that lies in one subset of internally independent data. As we mentioned
in the previous chapter, we are able to extract from the database more
than one such subset. Each one is in some sense a new experiment, and
the results derived from each one contain in some measure new.
information. The problem, which we tried to address with the weighting
process outlined above, is that their results can not be considered
independent measurements. It is difficult to combine the separate
answers without a detailed analysis of the subsets' interdependence. In
some sense we can average the position of the best fits, and even the
confidence range limits (for at least single parameter models) to
describe a result typical of the several subsets. But such a result
could only be illustrative, and would not have a definable statistical
meaning. There is however a rigorous and correct method to construct a
combined confidence region from several different determinations without
analyzing their interdependence. Regions determined by this technique
are called Boole's Inequality confidence regions.

Boole's Inequality is a basic lemma of probability. Let A and B
be two events. They may be independant, completely correlated, totally
anti-correlated, or anything in between. Let $p_A$ and $p_B$ be the
respective probabilities with which A and B occur. Each event, A, has a
complementary event, A', the "event" that A does not occur. Its
probability is

$$p_{A'} = 1 - p_A .$$                                                  [V.11]

One can combine events in several ways: A ∪ B, the event where either A
or B occur; and A ∩ B, the event where both A and B occur. Boole's
Inequality states the following straight forward property of the union
event:

$$p_{AUB} < p_A + p_B .$$                                               [V.12]

If the sum of $p_A + p_B$ is greater than one, then Boole's Inequality is
trivially true. If A and B are totally correlated, then $p_{AUB} = p_A$. The
equality of equation [V.12] holds if they are anticorrelated to the
extent that the conditional probability of one given the other is zero.
If the two events are independent, i.e. $p_{A \cap B} = p_A p_B$, then we can show
that the exact form of $p_{AUB}$ is $p_A + p_B - (p_A p_B)$, which still satisfies

the inequality.

To see the utility of Boole's Inequality for determination of confidence regions we consider the simple case of two confidence regions labeled 1 and 2. Their associated confidence levels are C1 and C2. Let R1 be the event that $\vec{\alpha}_o$, the true model parameter vector, is contained in region 1. Therefore $P_{R1} = C1$. We form the complementary regions, 1' and 2'. The probability that $\vec{\alpha}_o$ is found in the 1' region is given by $P_{R1'} = 1 - P_{R1}$. Using Boole's Inequality for the union of the complementary regions,

$$P_{R1' \cup R2'} \leqslant P_{R1'} + P_{R2'} = 2 - C1 - C2 \; . \qquad [V.13]$$

Knowing that the intersection of two regions is the complement of the union of their complements we can rewrite this as

$$P_{R1 \cap R2} = 1 - P_{R1' \cup R2'} \geqslant 1 - (1-C1) - (1-C2). \qquad [V.14]$$

Thus we have a lower limit on the confidence associated with the intersection of the two regions. As an example, if we have two 95% confidence regions, their intersection will contain $\vec{\alpha}_o$ at least 90% of the time. This can be generalized to more than two regions, by

$$P_{\cap R_n} = 1 - \sum (1-C_n) \; , \qquad [V.15]$$

where $\cap R_n$ represents the intersection of n regions of individual confidences $C_n$.

We will not dwell on the nuances of the application of this statistic but we will emphasize three points: the region generated has a level of confidence at least as high as the one indicated; the utility of this tool is best for determining high level confidences, where the $(1-C_n)$ are small; and for any given confidence level there comes a point of vanishing returns where the addition of information from a new confidence region will not reduce the size of the net confidence region. However for all its weaknesses, the Boole's Inequality method for combining dependent measurements is a tool whose statistical interpretation is as certain as that of the regions derived from a single subset of independent measurements.

# CHAPTER VI
## RESULTS OF FITS TO THE DATA

We have examined the physical aspects of the A-2 experiment, shown how to model the effects of low intensity sources on fluctuations of the X-ray sky, constructed a database for examining such fluctuations, and developed tools for testing the models against the data. With our previous work as prologue we now perform that testing. We begin with a review of the data used and its characteristics. Our first fits are simple Euclidean power laws, where we show that the calculation of the mean count rate used with the fluctuations distribution is an important problem. In particular, if we use the naive value of the rms mean of the sky intensity we will introduce a significant bias in our results. After determining a way to isolate the mean from the other parameters, we examine the simple Euclidean fits in detail for a variety of sets of data. The sensitivity of our data to the behavior of low flux sources is probed by examining the addition of a sharp cutoff at low fluxes. The next, traditional, extension is to models with non-Euclidean power law indices. If we assume that the true N(S) is well modeled by a simple power law we see that the fluctuations strongly constrain the N(S) behavior over a factor of ten in flux below the A-2 experiment resolved source limit. A more general, and perhaps more realistic, parameterization of N(S) is to use a single Euclidean power law for the high flux behavior, while the only constraint on N(S) at low flux from the fluctuations is placed by the limits on the excess variance from a pure Gaussian component of the fluctuations. We pictorially consider the constraints this approach places on two component power law models, where the second power law represents an additional evolving component or the effects on a non-evolving model of the expected roll over from the Euclidean form at low fluxes. We conclude the chapter with a critique of previous measurements of the fluctuations, and a comparison of their fit results with ours. Although our results do not appear to be a significant improvement over the older reported measurements, in many cases their confidence regions or values are affected by

assumptions we have not made or by errors in their analysis.

———————————————

1: Recapitulation of Data Characteristics

We review the features of the data chosen for fitting
models. The various subsets of independent, and non-
independent data are described. The fitting process is
sketched.

We have chosen to analyze superposed HED 1 L+SFOV data. The data
came from scans superposed for periods up to two days, implying a smear
perpendicular to the scan of ~1.5°. The data in each scan were
collected in bins 6° in extent along the scan direction. Such a smeared
bin covers an area of 122 square degrees with non-zero response, but 90%
of the counts originate in a rectangle 11.2°x4.4°, for an area of 49
square degrees. The peak response of the smeared detector is 530 $cm^2$ s,
with an average response in the 90%-of-counts region of 220 $cm^2$ s. The
detector response weights of interest are: $R_0$ = 3x10$^{-3}$; $R_1$ = 0.292 $cm^2$
s; $R_{3/2}$ = 2.80 $cm^2$ s; and $R_2$ = 9.04 $cm^2$ s. $R_{5/2}$ has a value of 18.96
$cm^2$ s so that following equation [III.36] the mean flux of a source
drawn from a Euclidean N(S) observed with intensity I is <S(I)> =
3x10$^{-3}$xI counts s$^{-1}$ cm$^{-2}$.

The spectral window we have chosen is the entire Layer 1 window.
Fluxes from this window are measured in H1L1 units, where 1 H1L1 equals
1 count s$^{-1}$ cm$^{-2}$. This window in pulse height space has a total
efficiency greater than 10% of its peak value within the range 2.55-46.5
keV. The conversion coefficients from counts to ergs (2-10 keV) and the
range where 90% of the counts come from depend on the incident spectrum.
For the diffuse background spectrum, the coefficient is 1.28x10$^{-8}$ ergs
per detected photon. The value for typical extragalactic spectra is
1.35x10$^{-8}$ ergs, which we will generally use in this dissertation. 90%
of the counts of the X-ray sky spectrum originate in the range 2.5-13.3
keV.

The total intensity of un-resolved source contaminated data is 21.5 counts per 1.28 second exposure. After fitting the internal background using a special combination of the large and small fields of view (see Section II:2), we estimate the sky brightness to be 17.06 counts $\exp^{-1}$, corresponding to an all sky flux, $S_{as}$, of 58.4 counts $s^{-1}$ $cm^{-2}$ ($7.5 \times 10^{-7}$ ergs $s^{-1}$ $cm^{-2}$). When only data with L<1.18 are used, the typical photon noise for each bin is 0.23 counts $\exp^{-1}$. The expected size of the fluctuations, as characterized by the effective width $W_{eff}$, is ~0.2 counts $\exp^{-1}$. The ratio of noise to fluctuations signal, $\psi$, is 1.2.

---

Table VI.1: Summary of X-ray sky properties: fluxes and intensities

---

### Sky flux properties

| | |
|---|---|
| $S_{as}$ | 58.4 H1L1 |
| W (fluctuations scale) | 0.7 ($\gamma = 5/2$ model) |
| $\sigma^2_S$ (excess variance) | $\lesssim 10^{-3}$ H1L1$^2$ |

### Detector dependent measurements, H1L1 L+SFOV

| | |
|---|---|
| $\bar{I}_{tot}$ | 21.5 counts $\exp^{-1}$ |
| $\bar{I}_{sky}$ | 17.1 counts $\exp^{-1}$ |
| B (internal background) | 2.2 counts $\exp^{-1}$ (per side) |
| | = 4.4 counts $\exp^{-1}$ total for L+SFOV |
| $\sigma^2_B$ (excess variance due to internal background) | |
| | $\lesssim$ (0.003) (counts $\exp^{-1}$)$^2$ (per side) |
| | $\lesssim$ (0.012) (counts $\exp^{-1}$)$^2$ total |
| $W_{eff}$ | 0.2 counts $\exp^{-1}$ |
| $\sigma^2_I$ | $\lesssim$ 0.08 (counts $\exp^{-1}$)$^2$ |

(All fluctuations quantities are crude approximations of results presented in this and the succeeding chapter)

---

All data analyzed contain no contribution from sky with $|b| < 20°$, minimizing the effect of the galactic disk. In addition we exclude those parts of the sky contaminated by the brighter identified galactic

## Table VI.2: Excluded High Latitude Sources

The source and its position in right ascension and declination are presented. Data that included contributions from these sources were omitted from in the fluctuations analysis.

### Galactic Sources

| Source | RA | Dec |
|---|---|---|
| Sco X-1. . . . . . | 83.2° | -5.42° |
| Her X-1. . . . . . | 254.01 | 35.42 |
| AM Her . . . . . . | 273.75 | 49.85 |
| Ex Hyd . . . . . . | 192.5 | -29.0 |
| Orion/Trapezium . . . | 83.2 | -5.42 |
| NGC 1851 globular cluster. . | 78.1 | -40.1 |
| NGC 7078 globular cluster. . | 321.89 | 11.95 |

### Extragalactic Sources

| Source | RA | Dec |
|---|---|---|
| LMC X-1. . . . . . | 85.0 | -69.8 |
| LMC X-2. . . . . . | 80.25 | -72.0 |
| LMC X-3. . . . . . | 84.75 | -64.10 |
| LMC X-4. . . . . . | 83.25 | -66.40 |
| LMC transient. . . . | 72.25 | -55.0 |
| LMC transient. . . . | 84.0 | -66.9 |
| LMC N132d Supernova remnant . | 81.25 | -69.7 |
| LMC N63A Supernova remnant . | 84.0 | -66.0 |
| SMC X-1. . . . . . | 19.0 | -73.7 |
| SMC X-2. . . . . . | 13.25 | -74.0 |
| SMC X-3. . . . . . | 12.5 | -72.7 |
| M31 . . . . . . . | 10.0 | 41.0 |

X-ray sources. The regions around the Magellanic Clouds and the

Fit results                                                Recapitulation

Andromeda Nebula are also excluded (see Table VI.2). Measurements with
intensities greater than 23.0 counts $exp^{-1}$ are also excluded. This is an
excess intensity of 1.5 counts $exp^{-1}$. A single source in the field of
view with this intensity has, on average, a corresponding flux ~$6 \times 10^{-3}$
counts $s^{-1}$ $cm^{-2}$. The measured rates are adjusted by the best fit values
for large scale structure: the secular decrease, the galactic disk, and
the small 24-hour anisotropy.

The data were divided into subsets of independent measurements as
well as two sets that included overlapping data. One such set,
hereafter called the "all sky" dataset, contained all the data points
and was used to estimate the proper weighting function for the analysis
of overlapping data. The weighting term, as given in equation [IV.4],
depends on the size of the intrinsic variation associated with the
fluctuations. From the all sky data we estimate this variance, $\sigma^2_D$,
using a likelihood estimator as was computed for the excess variance
associated with the internal background (see equation [IV.8]). We
estimate $\sigma_D$ to be 0.38 counts $exp^{-1}$. As we describe in the next
section, the all sky data were then weighted and used to determine the
least squares mean count rate. For least squares determinations we can
not select the data on the basis of intensity, there being no clean way
to evaluate the impact of such a selection.

The number of individual measurements in the all sky data was very
large, but the total weighted equivalent number of individual
observations was not. Because of the computational expense of fitting
the models it would be economic to reduce the number of points. For
this reason we formed the "whole sky weighted" dataset, indicated in
plots and tables by the initials WW. The WW dataset covered the entire
sky, but was generated by the independent data subset creation program
NSCIND so that the data points were only partially overlapping. Instead
of using the true detector dimensions to calculate what constituted an
independent measurement, we set the detector equivalent size of $\delta\theta$ and
$\delta\phi$ to $\Delta\theta$, the bin size ($6°$), and $0.2°$ respectively. This size for $\delta\phi$
restricted the number of measurements from the north ecliptic pole, but
was still small enough that some data were selectable from almost all

scans. The weighted number of observations was only slightly less than the weighted number of observations of the all sky dataset. The size of $\delta\theta$ insured that all measurements in the independent parts of the scans, after selection by the $\delta\phi$ criterion, are part of the WW data set.

Sets of totally independent data were also created: two groups were generated, each subsequently divided into three subsets. These are denoted SS1 through SS6. We also combined the data into sets of non-independent data, weighting them appropriately, to form SS1-3(Wt.) and

Table VI.3: HED 1 datasets used in the study of the fluctuations

| Subset | Number of measurements | Weighted number |
|--------|------------------------|-----------------|
| WW | 2211 | 698.0 |
| SS1 | 101 | |
| SS2 | 99 | |
| SS3 | 102 | |
| SS1-3(Wt.) | 302 | 217.4 |
| SS4 | 93 | |
| SS5 | 88 | |
| SS6 | 87 | |
| SS4-6(Wt.) | 268 | 193.2 |

SS4-6(Wt.).

Each measurement consisted of a rate, a Gaussian noise variance (from counting statistics and excess variance in the internal background), and the value of the upper intensity cutoff. A model N(S) was parameterized by its size (W or $W_{eff}$), power law index ($\gamma$), a possible sharp cutoff ($S_o$ or $\xi_c$), and an excess or deficit variance to model the deviations from a simple power law ($\sigma^2_S$ or $\sigma^2_\xi$). The model number distribution was transformed into a predicted underlying distribution of intensities about some mean, $\bar{I}$. For each measurement the underlying distribution is convolved with the Gaussian noise from

counting statistics and internal background and the differential and integral probabilites of the measurement were calculated. The differential probabilities were used in a likelihood test to determine the best fit among the range of parameters and to estimate their confidence region. The moments of the distribution of the integral probabilities are examined to estimate the goodness of the fit. In reporting the results of the fits we will, for brevity, often restrict ourselves to the confidence regions for WW, SS1, and the combined result from subsets 1-3 using Boole's Inequality, denoted by the initials BI. This last region generally has the same size as, or is somewhat larger than, a single subset's region, but provides ,a compact and correct "average" contour for the three. The 90% BI contour encloses the intersection of the three subsamples' 96.66% confidence regions.

---

### 2: The Reconciliation of the Observed and Model Means

The fluctuations distribution, $P_{I'}(I-\bar{I})$, is calculated from the model $N(S)$ to give a distribution about the mean count rate. If in performing the fits the model mean is identified with the mean count rate of the observations, a systematic bias is introduced. We show in this section that the value for the size of the fluctuations, $W_{eff}$, can be strongly correlated with the value chosen for $\bar{I}$. The bias and correlations are due to the assumption that the model has the correct contribution from resolved sources, to the mean count rate. A simple correction term removes both bias and correlation, allowing the separation of the uncertainty in $\bar{I}$ from determinations of $W_{eff}$ of the uncertainty.

When we create the test distribution for the intensities, $P_{I'}$, (subsequently convolved with the noise distributions) we determine the distribution about some mean intensity, $\bar{I}$. $\bar{I}$ is essentially a model parameter like $W_{eff}$ and $\gamma$, but one that is of less interest. For one thing it includes non-X-ray sources of counts, such as the internal background. Fitting models to the data is time consuming, so it is preferable to know and fix $\bar{I}$ prior to the fluctuation fitting. Alternatively, if the results do not depend significantly on the value

of $\overline{I}$ we can dispense with considering it as a free parameter in our fits.

The simplest approach is just to measure $\overline{I}$. To zeroth order, $\overline{I}$ is the total number of counts observed by the detector during the mission, $\sim 2 \times 10^8$, divided by the total active time. More exactly we can take the same sky used in the study of the fluctuations, subtract our model for large scale structure, and find the mean from a least squares estimate. As the mean count rate does not depend on the stochastic behavior of the unresolved sources, it is reasonable to assume that we may use overlapping data. Though such use will yield an unbiased estimate of the mean, as we show in Section A2:2, this estimate is not the most efficient, nor is its statistical uncertainty knowable. To satisfy both aspects, the data should be weighted. We use the all sky dataset described in the previous section and perform a simple, weighted, $\chi^2$ fit to find that $\overline{I}_{obs}$, the observed mean count rate, is $21.096 \pm 0.016$ counts $\exp^{-1}$ ($1\sigma$). Using this value for $\overline{I}$, we then fit the data from the WW overlap set to a model of the fluctuations from a Euclidean N(S) model, ($\gamma = 5/2$). The only free parameter is the effective width, $W_{eff}$, which has a best fit value of $0.178(0.167-0.188)$ counts $\exp^{-1}$ (90%).

If we let $\overline{I}$ be a free parameter we learn that our simple identification of $\overline{I}_{obs}$ with the model mean is misleading. The 90% confidence region is contained in the upper contour of Figure VI.1. This contour is for only one interesting parameter, so that it constrains $W_{eff}$ or $\overline{I}$, separately, not jointly. If we are interested only in $\overline{I}$, its value is $21.169(21.128-21.222)$ counts $\exp^{-1}$. This region and the acceptable range for the least square determination are significantly different. Although the change from $\overline{I}_{obs}$ to $\overline{I}_{bestfit}$ is only 0.34%, by allowing $\overline{I}$ to vary we cause a shift in the best fit value of $W_{eff}$ of more than 10%, to $W_{eff} = 0.197(0.182-0.213)$ counts $\exp^{-1}$.

It is not obvious why the two estimates for the mean count rate should be so differant. After all, it is the same sky. A clue lies in the strong correlation between $\overline{I}$ and $W_{eff}$ as illustrated in the tilted aspect of the confidence contour. We saw in our theoretical description

Fit results                                                         Observed and Model Means

Figure VI.1: Fit Confidence Regions, $\overline{I}$ and $W_{eff}$

Euclidean model fits for weighted partially overlapping data. The dashed line denotes the one $\sigma$ range of $\overline{I}$ based on the weighted least squares estimate. The dotted error bar is the initial estimate of the 90% range in $W_{eff}$ holding $\overline{I}$ fixed and making no corrections for the behavior of resolved sources. The uppermost contour encloses the 90% confidence region when $\overline{I}$ is allowed to vary freely. The other two contours are cases where the effect of the bright sources on a determination of $\overline{I}$ is modeled and removed (see text). The contours are at $\Lambda_{min}+2.7$ (90% for only one interesting parameter.
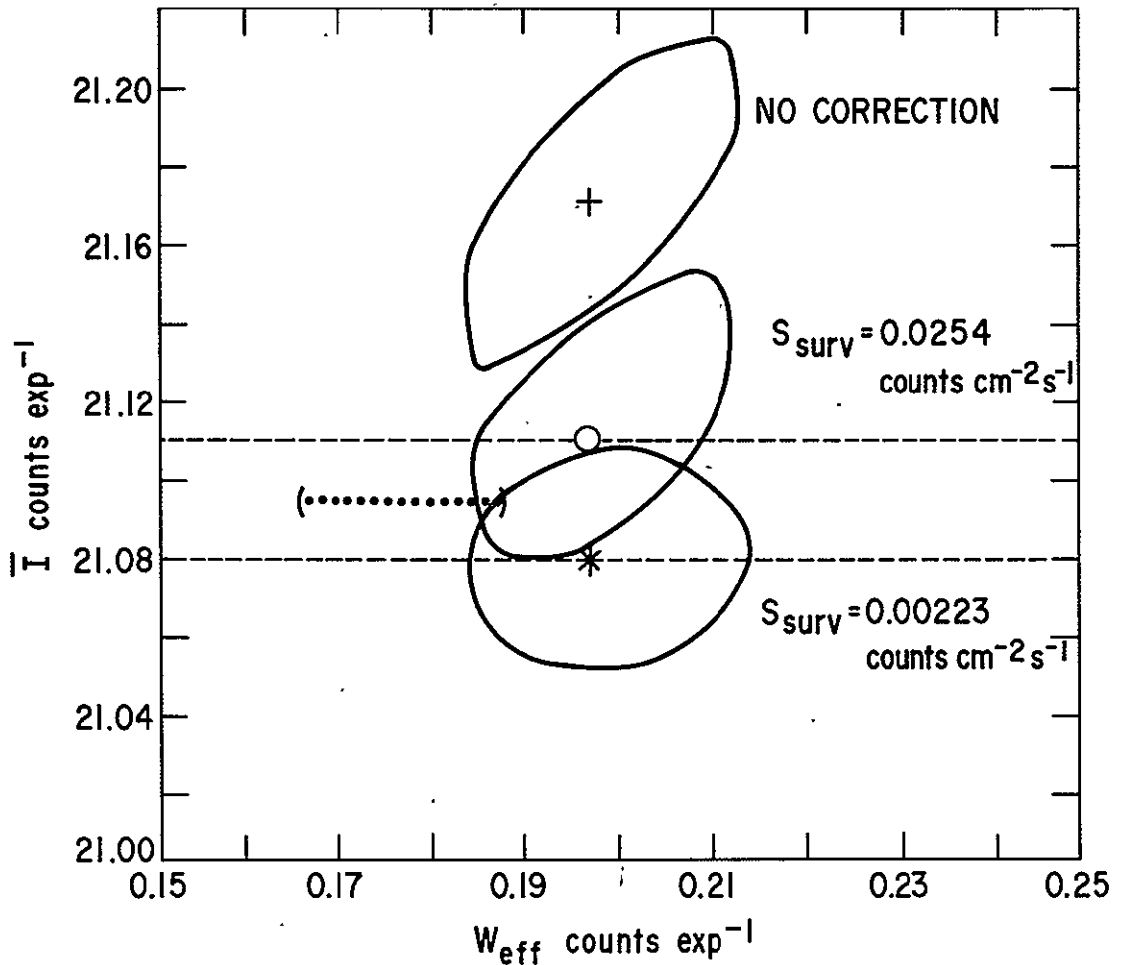


of the generation of $P_\xi$, distributions that the fluctuations may be shifted by changes in the model $N(S)$ at high flux levels. Figure

Fit results

Observed and Model Means

III.7(b) shows an example of such a shift when a sharp cutoff is imposed at about the level of the brightest observed sources. Although not as large a shift as when $\gamma$ is less than 5/2 (see Figure III.7(a)), the fits show that the shift is easily detectable, given the number of observations and the quality of data that we have. For the upper contour in Figure VI.1, marked "no correction", the upper limit was equivalent to infinity. The next contour, marked "$S_{surv} = 0.0254$", is when the correction to the model mean intensity is made following equation [III.60], with $S_{max}$ replaced by the survey limit, $S_{surv}$.

In this case the fluctuations-fitted mean count rate and the least squares value are consistent. However the confidence region still is in the form of an elongated and tilted ellipse, indicating continued correlation between the confidence range in $W_{eff}$ given a particular value of $\overline{I}$ (and vice versa). What is the source of this continued, albeit reduced, correlation? The reasonableness of the first correction depended on recognizing that the model for N(S) continued unchanged to fluxes higher than the brightest sources, contrary to the actual data from which $\overline{I}_{obs}$ was measured. But a similar presumption is in effect at slightly lower fluxes. We know something about the number and flux disribution of a complete sample of sources above some survey limit, $S_{surv}$, and a particular model will predict the contribution of sources in that flux range to a measurement of $\overline{I}$. The total integrated flux from sources above the survey limit is

$$S_{S>S_{surv}} = (4\pi/\Delta\Omega_{surv}) \sum S_i , \qquad [VI.1]$$

where the $S_i$ are the individual resolved source fluxes and $\Delta\Omega_{surv}$ is the total solid angle covered by the survey. For the survey presented in Piccinotti et al. [1982], $S_{S>S_{surv}}$ is 0.446 counts $s^{-1}$ $cm^{-2}$ for a survey limit of $2.23 \times 10^{-3}$ counts $s^{-1}$ $cm^{-2}$. The contribution of these resolved sources to the mean intensity is just $R_1 S_{surv}$ which for HED 1 L+SFOV is 0.13 counts $exp^{-1}$. If we limit the maximum flux of N(S) used in defining the mean of the fluctuations model in equation [III.60] then

the observed mean intensity may be split into two parts:

$$\overline{I} = \overline{I}_{S<S_{surv}} + R_1\ S_{S>S_{surv}} \cdot \qquad\qquad [VI.2]$$

The behavior of the resolved sources is totally acounted for in the second term and segregated from any other assumptions in the N(S) model, represented by the first term.

The result of such a separation is seen in the lowest contour in Figure VI.1. After applying equation [VI.2], the fluctuations analysis yields a best fit value for $\overline{I}$ that is consistent with the least squares observation. Better, the shape of the contour indicates that with this choice of $S_{surv}$, $S_{eff}$ and $\overline{I}$ are no longer correlated. As long as it is in the range of the best fit values of $\overline{I}$, the exact value chosen for $\overline{I}$ is unimportant in the determination of properties of N(S). It is important to keep in mind that this separation is possible only after explicitly removing changes in the mean from the model behavior at the levels of the resolved sources. Without such a correction, or an independent fitting of $\overline{I}$, the derived N(S) parameters may be systematically biased.

----------------------------

### 3: Fits of Euclidean Models

We test a power law model for N(S) with index $\gamma$ = 5/2, characteristic of a uniform distribution of sources in Euclidean space. The quality of the fit is good. The various subsets are used to determine the confidence region for the width, W. The answers from all the sets are in agreement. By fitting models with sharp cutoffs at low fluxes we are able to estimate the limitations of our models in describing the low flux region of N(S). The limit is set by a variance condition, rather than by the number of sources per field of view. The best-fit $\gamma$=5/2 model is used to define a fiducial N(S) to which subsequent models are compared.

Before fitting particular models to the fluctuations data base we already recognize that there is some deviation from perfect isotropy.

Table VI.4: Confidence range for W when $\gamma = 5/2$

The results for the various subsets are presented.  For the four entries associated with the WW set different correction terms were removed in order to assess their impact.

| Subset | | W(90%) counts $s^{-1}$ $cm^{-2}$ | $\Lambda_{min}$ |
|---|---|---|---|
| WW1 | all corrections | 0.0702(0.0650-0.0759) | 826.51 |
| WW2 | no $\sigma^2_B$ corrections | 0.0706(0.0654-0.0761) | 838.03 |
| WW4 | no large scale variations corrections | 0.0740(0.0685-0.0798) | 867.31 |
| WW5 | no corrections | 0.0742(0.0696-0.0788) | 878.82 |
| SS1 | | 0.0677(0.0558-0.0817) | |
| SS2 | | 0.0704(0.0552-0.0857) | |
| SS3 | | 0.0687(0.0573-0.0819) | |
| SS1-3 (Wt) | | 0.0687(0.0599-0.0785) | |
| SS1-3 (BI) | | (0.0538-0.0870) | |
| SS4 | | 0.0623(0.0504-0.0775) | |
| SS5 | | 0.0664(0.0554-0.0821) | |
| SS6 | | 0.0782(0.0640-0.0956) | |
| SS4-6 (Wt) | | 0.0684(0.0593-0.0789) | |
| SS4-6 (BI) | | (0.0599-0.0829) | |

The, moment statistics indicate that there must be some variation in addition to counting statistics.  We estimated a variation of $\sigma \sim 0.32$ counts $exp^{-1}$ when calculating the weights for the overlapping data in the WW set.  If we account for the variation by fitting it with the fluctuations distribution generated by a Euclidean power law, expected from a simple extrapolation of the resolved source counts, the first

Figure VI.2: 90% Confidence Range for W when $\gamma$ = 5/2

(See Table VI.4 for explanation of abbreviations)
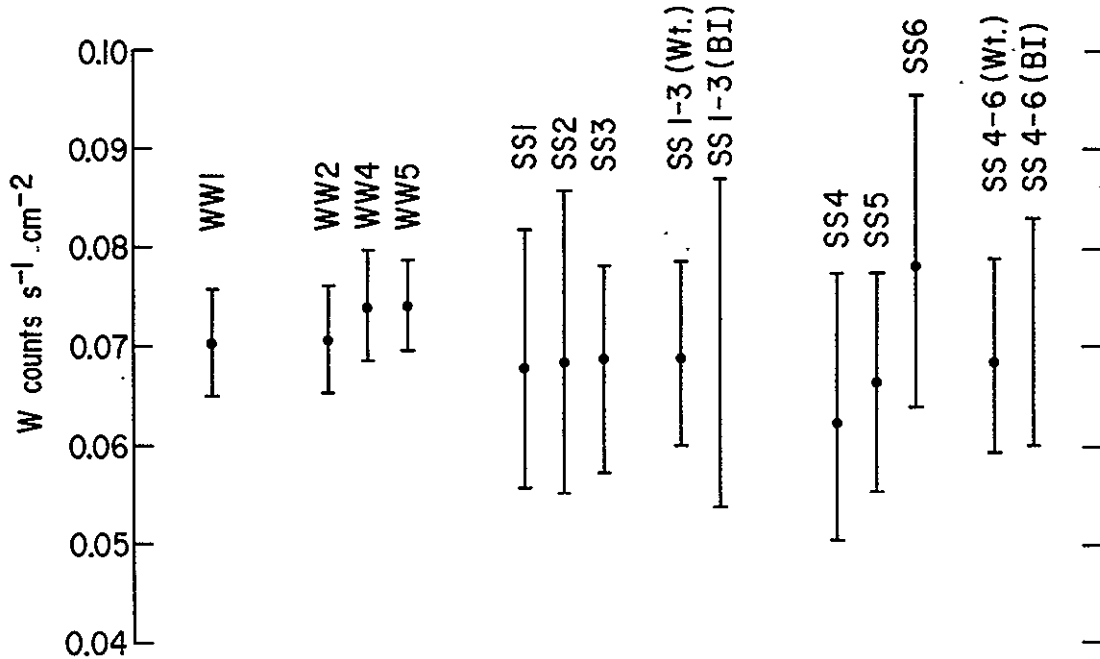


four moment statistics all indicate an acceptable fit. We present in Table VI.4 and Figure VI.2 the confidence region derived from the variation of $\Lambda$ as a function of the remaining model parameter, $W_{eff}$. We fixed $\overline{I}$ to 21.08 with a $S_{surv}$ of 0.00223 counts $s^{-1}$ $cm^{-2}$ and a $S_{S>S_{surv}}$ of 0.446 counts $s^{-1}$ $cm^{-2}$ as discussed in the previous section. The various subsets are in agreement with each other, and we will see later that the ranges here are also in agreement with the results from the resolved sources. We should note that the results presented for the Boole's Inequality combination of the subsets (indicated by the BI) does not include a best-fit value. Boole's Inequality speaks only about confidence regions and can not indicate the most likely model within those regions.

The four results presented for the whole sky weighted dataset (WW) are not different sets of data, but the same data analyzed under

different conditions. In particular, different corrections for other sources of variation are included. The change of $\Lambda_{min}$ indicates that the inclusion of these corrections is highly significant in a statistical sense, but the effect on the position of the confidence region is only a shift of ~5.4% when all the large scale structure effects are included. The inclusion of the excess variance of the internal background, $\sigma^2_B$, has a much smaller effect, ~0.4%. The exact value used for the internal background variance is therefore unimportant for determinations of W. Even the effect of including the galaxy is of the same order as the scatter between the subsets of independent data.

## Fiducial model

We will use the best-fit Euclidean model from the whole sky weighted data as a basis for comparing other models, as well as estimating the predicted N(S) behavior of populations of observed sources at lower fluxes. The best-fit value for W, 0.0702 counts $s^{-1}$ $cm^{-2}$, corresponds to a value for K of $1.480 \times 10^{-3}$. In comparing a general N(S) to this fiducial model it is convenient to factor out the Euclidean power-law component:

$$N(S) = 4\pi \ K(S) \ S^{-5/2} \ . \qquad\qquad [VI.3]$$

The comparison of a particular N(S) to the fiducial N(S) is equivalent to plotting the ratio of the associated K(S) to the fiducial K. We will refer to the fiducial (constant) K value as $K_{fid}$.

## The range of applicability

Though the Euclidean power-law models for N(S) used here are presumed to extend to infinitesimally small values of flux we have noted in previous chapters that the data are not uniformly sensitive to the models over the entire range of flux. For a first order understanding of the limitations of the models we can use Scheuer's criterion, which states that the limits of sensitivity are reached at the flux corresponding to one source per field of view. Because of the smeared, non-ideal, nature of the detector responses it is important to remember the distinction between intensity and flux. For the fiducial model, the

Fit results                                                  Fits of Euclidean Models

intensity where there is one source in the field of view is $0.76 \times W_{eff} =$ 0.15 counts exp$^{-1}$. The average flux of a source of this intensity is $4.5 \times 10^{-4}$ counts s$^{-1}$ cm$^{-2}$, about one fifth of the resolved sources survey lower flux limit. If we try to calculate the one source per field of view condition completely in terms of source fluxes, it is not appropriate to use $\Omega_{eff}$ as the size of field of view. If we use the FW90%DF definition of the field of view, or 49 degree$^2$, one source per field of view corresponds to 837 sources all sky. For the fiducial model this number of sources is reached at a flux of $6 \times 10^{-4}$ counts s$^{-1}$ cm$^{-2}$, roughly a factor of 3.7 below the resolved source flux cutoff.

We will see in the next section a demonstration that the data are most sensitive to the behavior of N(S) models at this flux level and a little lower. However, as we argued in our original discusion of Scheuer's limit, it does not provide a lower limit to the possible sensitivity of the model, at least when we consider the contribution to the fluctuations' variance of the very lowest flux region of N(S). To determine the sensitivity of our data to such low flux portions we consider models where we hold W fixed, and examine the range of acceptable models as we vary the position of a sharp cutoff in N(S). The cutoff flux is given by $S_o$. When W is the fiducial width, 0.072 counts s$^{-1}$ cm$^{-2}$, the best-fit cutoff is zero. The 90% upper bound on $S_o$ is $1.0 \times 10^{-5}$ counts s$^{-1}$ cm$^{-2}$ for the WW data set and $9.9 \times 10^{-5}$ for the SSI subset of independent data. As we discussed earlier, the cutoff could not be higher because the missing variance provided by sources below the cutoff is significant. Since it is the variance, rather than a cutoff per se, that we are sensitive to, we calculate our sensitivity limit in terms of the corresponding intensity variance contributed to the sources below the cutoff, $\sigma^2_I$. This limit is $9.5 \times 10^{-3}$ (counts exp$^{-1}$)$^2$ ($3.0 \times 10^{-2}$ (counts exp$^{-1}$)$^2$) for SSI). For the WW set, this corresponds to a sigma of about half a percent of the diffuse sky intensity, roughly the magnitude of the large scale cosine anisotropy.

The upper bound and best-fit values for the flux cutoff are functions of W. Not surprisingly, including a sharp cutoff allows larger of values of W. When $W_{eff}$ is fixed at 0.22 counts exp$^{-1}$ (when W

equals 0.0784 counts $s^{-1}$ $cm^{-2}$), the best-fit value of $S_o$ is $2.2 \times 10^{-5}$ counts $s^{-1}$ $cm^{-2}$. The 90% confidence range for the WW data set is $0.1-6.1 \times 10^{-5}$ counts $s^{-1}$ $cm^{-2}$. Sources between the best-fit value for $S_o$ and the 90% upper bound contribute $9.7 \times 10^{-3}$ (counts $exp^{-1})^2$ to the intensity variance. The similarity between this variance and the variance sensitivity limit for the fiducial model directly illustrates that the sensitivity of models to a cutoff is indeed a matter of the cutoff sources' contribution to the variance of the fluctuations.

We can use the likelihood difference to predict our sensitivity to changes in the variance, as we did with equation [III.106]. The sensitivity is a function of the number of observations as well as the ratio of noise variation to the fluctutations. With the results from the WW data set, the predicted size of the variance is $5.3 \times 10^{-3}$ (counts $exp^{-1})^2$, in order of magnitude agreement with the actual measurement of $9.6 \times 10^{-3}$.

-----------------------

## 4: Fits of Gamma and W

We consider non-Euclidean power law N(S) models, i.e. those with $\gamma$ not fixed at 5/2. We find that the WW (whole sky weighted) data-set provides much greater constraints on the allowed values for W and $\gamma$ than a dataset of independent measurements. Euclidean and near-Euclidean models provide a good description of the data. The acceptable models show that the N(S) is most tightly constrained at fluxes of $3 \times 10^{-4}$ counts $s^{-1}$ $cm^{-2}$, an order of magnitude below the limits of resolved source counts. A low flux cutoff, imposed by the requirement of not exceeding the total X-ray sky flux, has a significant impact by allowing with the confidence region high $\gamma$ models that otherwise would have been excluded. The technique used for calculating the mean count rate also has an impact. Pegged models, where N(S) is constrained at a given flux $S_p$, are examined.

### Issues in performing non-Euclidean fits

Figure VI.3 shows the 90% contours for fitting the WW set of data

Figure VI.3: Effects of model assumptions on 90% confidence regions, $W_{eff}$ and $\gamma$ plane

The WW set is fit under three sets of conditions: A(Solid curve) the model includes an automatic low flux cutoff imposed to avoid exceeding the sky flux; B(Dashed curve) no automatic flux cutoff; and C(Dotted curve) with an automatic cutoff, but with the mean count rate calculated assuming a sharp cutoff in N(S) above $S_{surv} = 0.0254$ counts $s^{-1}$ $cm^{-2}$. The best-fit positions are indicated by the symbols A, B, and C respectively.
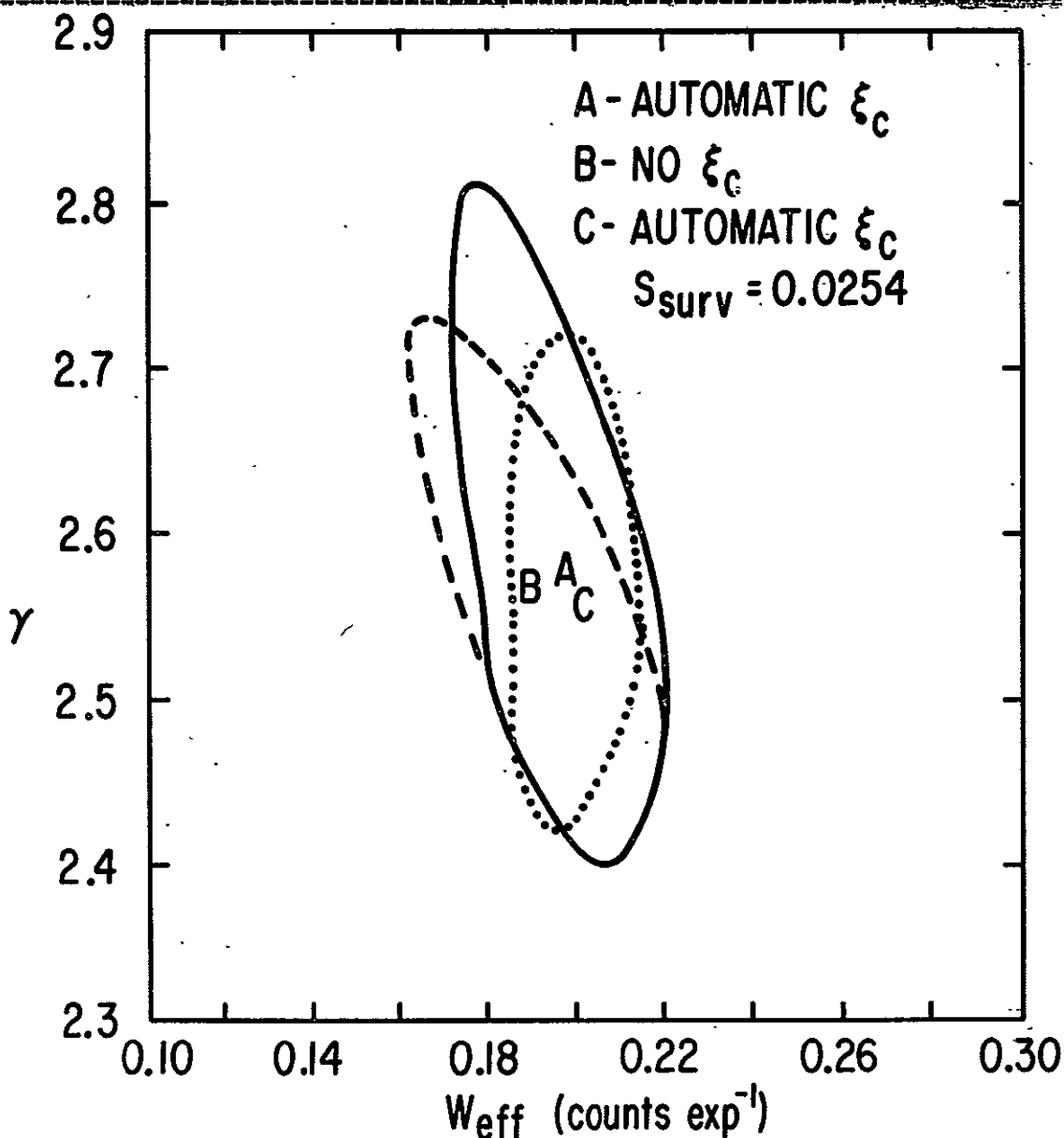
with non-Euclidean power law models. The confidence regions are defined by $\Delta\Lambda$ less than 4.60, so that both $\gamma$ and $W_{eff}$ are considered "interesting" parameters. Three separate contours are shown in the figure, illustrating different aspects of the fitting process. The A (solid) contour is for a model with a sharp low flux cutoff at the point where the model provides the total sky flux, while for the B (dashed) contour the model was extended to zero flux. For the B contour, models with $\gamma > 3.0$ were automatically rejected because of the infinite variance contributed by the lowest flux sources. In comparing the dashed and solid contours, we see that in the former case the models with higher values of $\gamma$ were incorrectly excluded. The low $\gamma$ models, where the two confidence regions strongly coincide, are less sensitive to the presence or absence of the low flux cutoff. Given that the presence or absence of the cutoff makes such an impact on the confidence region, it seems that the exact position of such a cutoff could be important. In particular higher cutoffs could, in principle, allow even steeper values of $\gamma$. Eventually, though, such models would be incompatable with the high flux behavior of N(S) and thus rejected. Rather than fitting regions where both $\gamma$ and $\zeta_o$ are free parameters, in the next section we fit models with an excess variance term. Because N(S) at low fluxes contributes an essentially pure Gaussian component to the fluctuations, constraints on such models can also be applied to determining the position of the cutoff.

For the C (dotted) contour in Figure VI.3, the low flux source behavior is the same as the A contour, with an automatic imposition of a sharp cutoff, but the treatment of <u>high</u> flux sources is different. This difference is not in the shape of the $P_{I'}(I')$ distribution, but in the way the mean count rate was calculated, and is related to the mean of the model distribution. The A curve followed the prescription of equation [VI.2], where the model was insulated from the contribution of sources with fluxes greater than the survey limit of 0.00223 counts $s^{-1}$ $cm^{-2}$. The C curve is insulated from the behavior of sources greater than 0.0254 count $s^{-1}$ $cm^{-2}$ (roughly the brightest resolved source flux). The actual mean count rate used in both curves was fixed to that determined by the best-fit Euclidean models for the two cases of 21.08

and 21.11 counts $\exp^{-1}$ respectively. The fact that the area enclosed by the C curve is smaller than that enclosed by the A curve is solely due to the behavior of sources between the two survey limits. This region is studied directly by the actual distribution of the resolved sources. Thus, if an independent assessment of N(S) is to be made, which is generally restricted to the flux region below that of the resolved sources, then the region enclosed by the A (solid) curve of Figure VI.3 should be used. In this chapter we will follow the techniques used in the generation of the A region, <u>viz.</u> automatic low flux cutoff and isolation from the resolved sources' behavior.

When analyzing strictly independent data, the automatic imposition of the low flux cutoff greatly increases the acceptable range in $\gamma$ (see Figure VI.4). In particular for the SS1 set, the upper range extends to values from 3.0 to 3.2, which would have been incorrectly removed from consideration without the automatic low flux cutoff. The other independent measurement datasets extend even farther, their average behavior represented by the Boole's inequality contour.

The curves of Figures VI.3 and VI.4 are in the $W_{eff}$, $\gamma$ plane. $W_{eff}$ parameterizes the number-intensity relationship n(I). To convert to the detector independent models, N(S), we change $W_{eff}$ to W, as illustrated in Figure VI.5. Note that the scale for W is logarithmic. The thin, elongated nature of the confidence region shows that the numerical values of W are strongly correlated with $\gamma$. It is difficult to use this region to see the restrictions the fluctuations place on the behavior of N(S). This is better seen in Figure VI.6 where we plot N(S) for each model acceptable at the 90% level, divided by the fiducial model N(S). Thus, the horizontal dashed line is the fiducial model of Euclidean index and W = 0.0702 counts $s^{-1}$ $cm^{-2}$. The "fish-tail" region on the right encloses the acceptable models. The right hand border of this region is roughly the survey limit of the resolved sources. Sources above this limit have some effect on the fit, for example when they contribute a small intensity to a measurement because they are just at the edge of a measurement bin where the spatial response is small. However, these sources can only have a relatively small constraint on

Figure VI.4: 90% confidence regions for $W_{eff}$ and Y

Contours for the SS1 data subset (dashed), the Boole's inequality region for SS1-3 (dotted), as well as for the WW (solid) data are given. The + symbol indicates the best-fit model for the SS1 dataset.



Figure VI.4: 90% confidence regions for $W_{eff}$ and Y
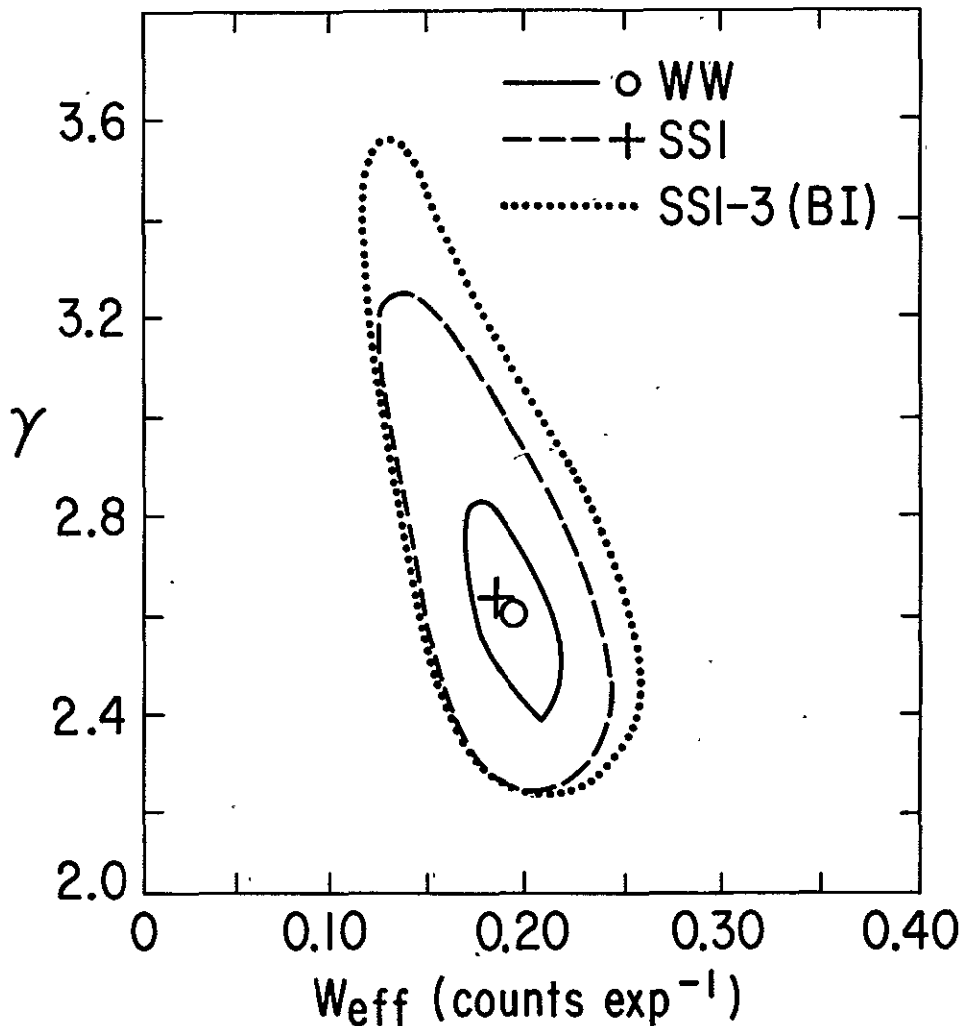
the fits. As the narrowing of the allowed models region indicates, the fits are most sensitive to sources at a flux about an order of magnitude below the limit of the resolved sources, at $\sim 3 \times 10^{-4}$ counts $s^{-1}$ $cm^{-2}$. Here the N(S) models all fall within $\pm 15\%$ of the fiducial model for the WW dataset ($\sim \pm 30\%$ for the SS1 dataset). The N(S) models are less

Figure VI.5: 90% confidence regions in the W, γ plane

The contours are as described for Figure VI.4, the small circle representing the WW best-fit model.
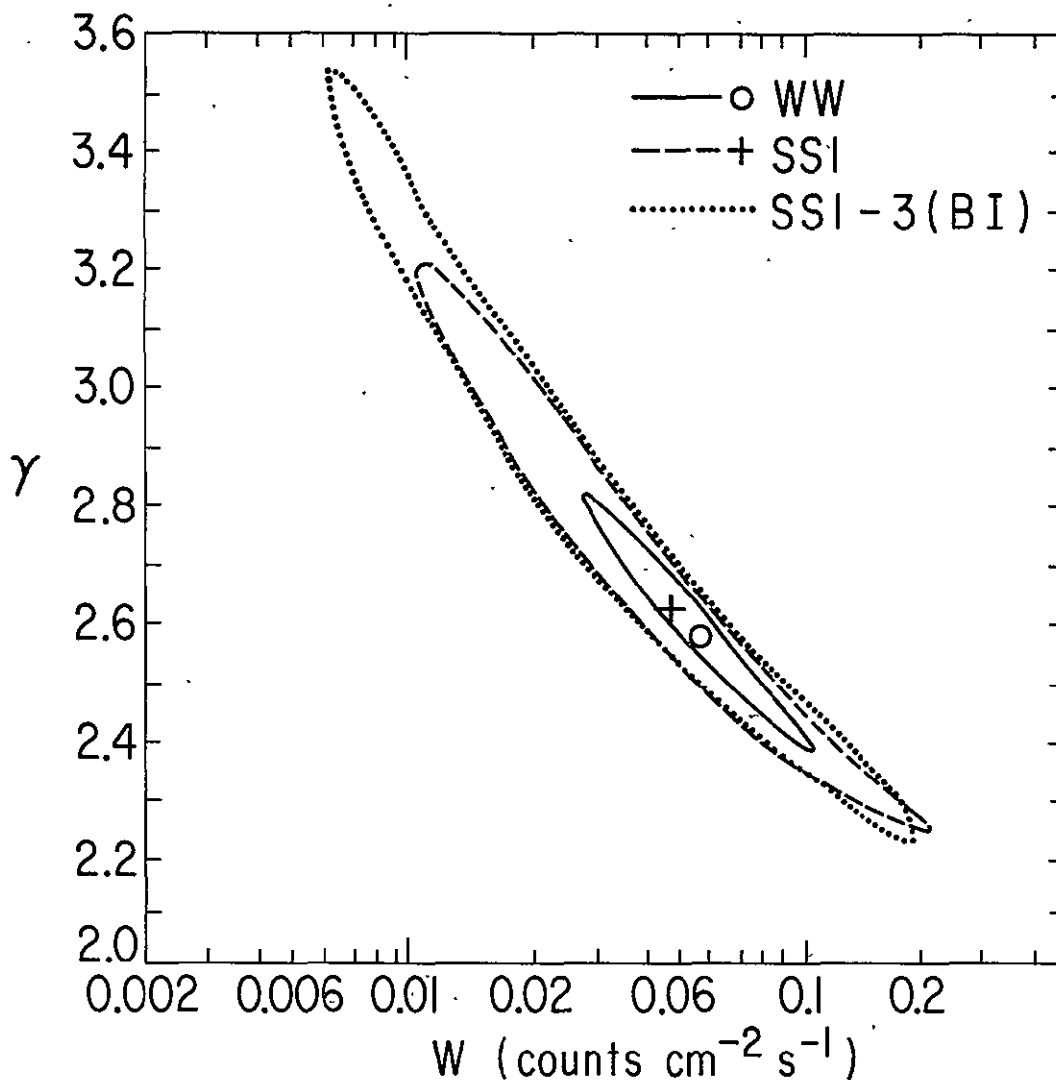


tightly constrained at higher and lower fluxes. In fact, in Figure VI.6(b) we see that the inclusion of models with relatively high values of γ (steep negative slopes in the fish-tail) for the SS1 dataset is allowed because the behavior of the higher flux sources is poorly constrained.

Figure VI.6: $N(S)/N(S)_{fid}$ for power law $N(S)$ models acceptable at the 90% level

(a) for the WW dataset; (b) for the SS1 dataset. A detailed explanation of the figure is given in the text.
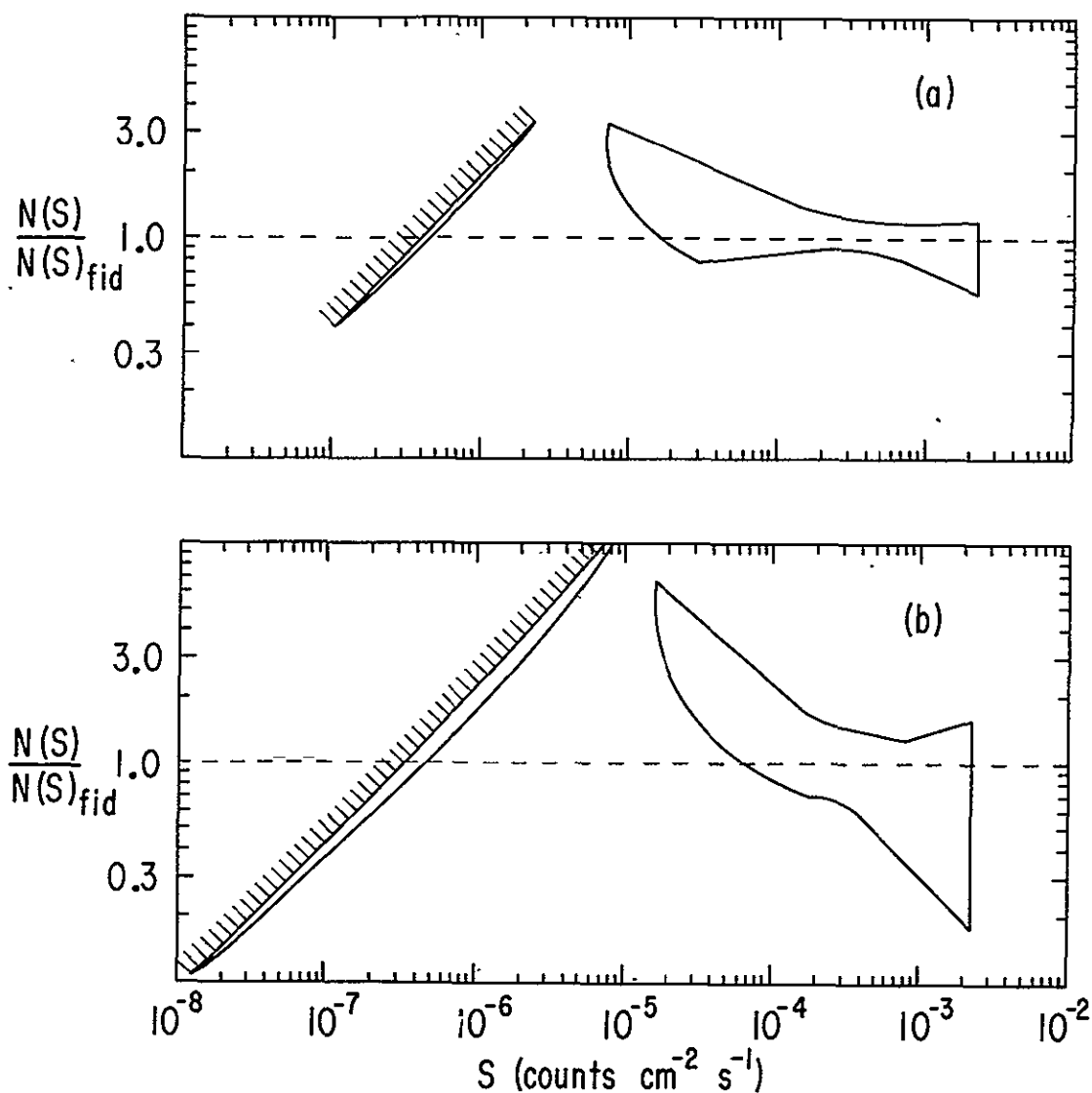


The contour in Figure VI.6 clearly shows what flux region is most constrained by the fluctuations. This disagrees with the folklore that "the fluctuations are dominated by the brightest sources, and insensitive to sources dimmer than the flux where one source is in the

field of view." The most tightly constrained region is not that of the brightest sources but rather roughly a factor of two below the limit of one source per field of view. For the smeared response function of our data the average flux of sources whose intensity is at the one source per field of view limit is $5 \times 10^{-4}$ counts $s^{-1}$ $cm^{-2}$. Thus, if we assume that N(S) follows a power law, we see that we can draw some conclusions about its behavior even below this limit. However, these conclusions are founded on the contribution to the distribution's variance by those low flux sources. We are not able to measure directly the power law index at these levels, as is possible in the study of the resolved sources. The constraint on the index rests on the assumption that a single index is a good fit over the entire range of flux.

The slender needle region on the left side of Figure VI.6, hatched on one side, shows where a particular power law N(S) model fills the total sky flux and must be terminated. The fact that the position where the total flux equals the sky flux is defined by a region, and not an exact function of N(S), is because it is an integral constraint. When the constraint is reached depends on the exact path of the N(S) model.

The left hand edge of the fish-tail region is defined so that the variance contributed by the sources below the edge is equal to the variance that was detectable when we included a sharp cutoff in the Euclidean models of the previous section. A sharp cutoff in the N(S) model between the left hand edge of the region and the point where the sky flux is exceeded is not distinguishable, with that set of data, from a continuation of the model to the lowest possible fluxes. Note that if the possibility of an early sharp cutoff were explicitly included in the fits the range of acceptable models would be modified. In particular models with even steeper values of γ would be acceptable if an earlier cutoff were allowed.

### Pegged models

Pegged models have N(S) fixed to some form above a certain flux and directly connected with a variable model below that flux. Given the

peg flux and the model above it, the condition of continuity is such that a power law model below the peg point is completely determined by a single parameter. Following equation [III.66], given $\gamma$ we can determine $W$.

In Figure VI.7 we show the derived confidence regions for four values for $W_p$, with the peg flux set to the resolved source survey limit. The four values, we will see, are of the same order as would be derived from the resolved source counts at the survey limit. For each value of $W_p$ we show the best-fit value and the 90% range. Now there is only one interesting independent parameter, so the confidence range is defined by $\Delta\Lambda < 2.71$. We see that the derived ranges are within the confidence regions of the un-pegged power law models. The SS1 data results are substantially improved in that the pegged models do not include high values of $\gamma$. This shows again that such high values were due to the data's poorly constraining the behavior of the higher flux sources. This is clearly seen in Figure VI.8.(b) where we plot $N(S)/N(S)_{fid}$ for the acceptable pegged models. The acceptable region is the union of models for all four values of $W_p$, as seen by the four wedges on the right hand side of the solid region. The pegged models significantly restrict the allowed models, depending mostly on the precision to which $N(S)$ for the resolved sources is known. The range of values used here for $W_p$ is representative of the uncertainty in this value. The similarity of the size of the solid and dashed regions in VI.8.(a) shows that for the whole sky weighted (WW) dataset the fluctuations alone provide roughly as strong a constraint on $N(S)$ near the survey limit as the resolved sources do.

Figure VI.7: Ranges of acceptable $W_{eff}$ and $\gamma$ for pegged models

The 90% ranges for the WW and SS1 sets are shown for four different values of $W_p$, the W of the Euclidean model assumed to hold for fluxes above the peg point, set at 0.00223 counts $s^{-1}$ $cm^{-2}$. For each value of $W_p$, the best-fit model for the WW and SS1 sets is indicated by a filled and open circle respectively. The inner and outer dashed curves are the 90% contours, unpegged, for the SS1 and WW datasets respectively.
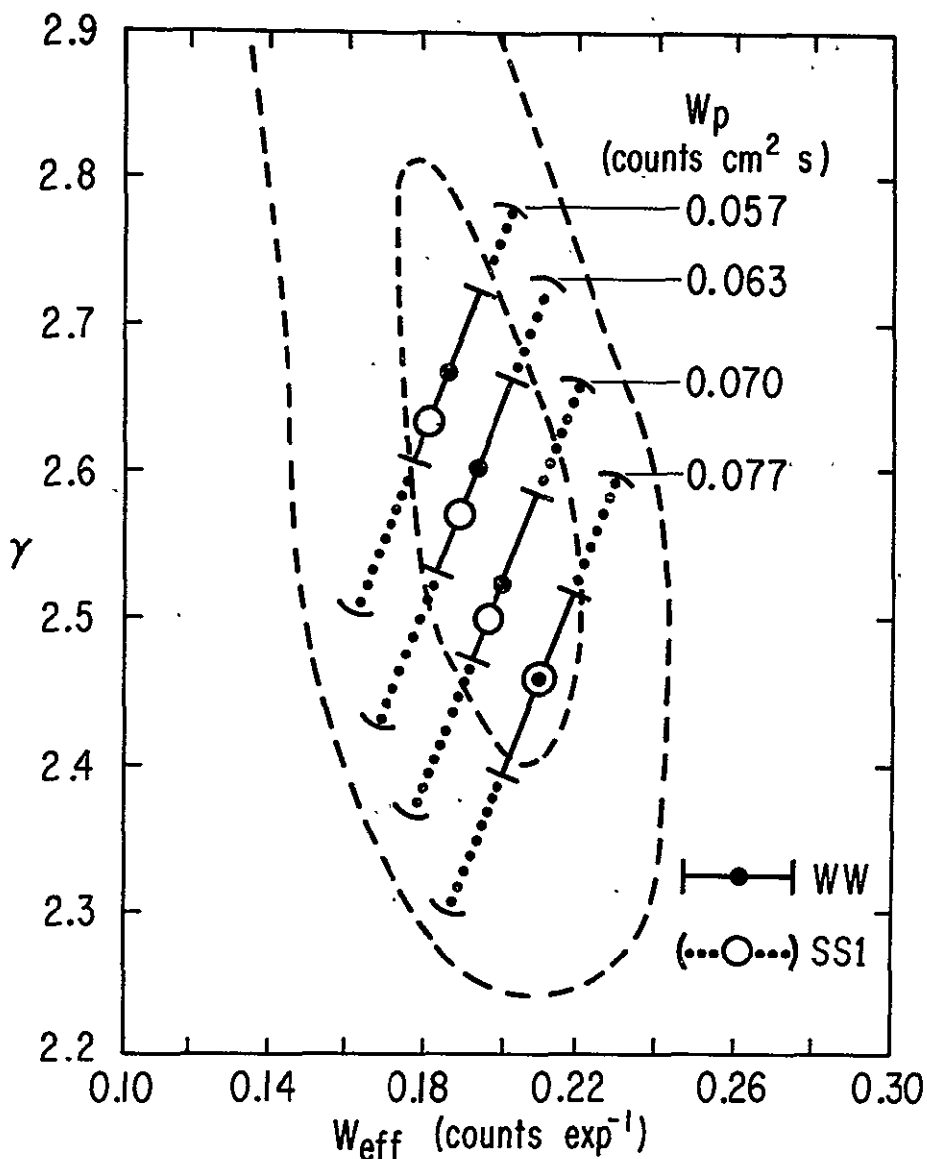
Figure VI.8: $N(S)/N(S)_{fid}$ for pegged models

(a). WW; (b) SS1. · The dashed outlines on the right are the
unpegged· 90% regions· of Figure VI.6.
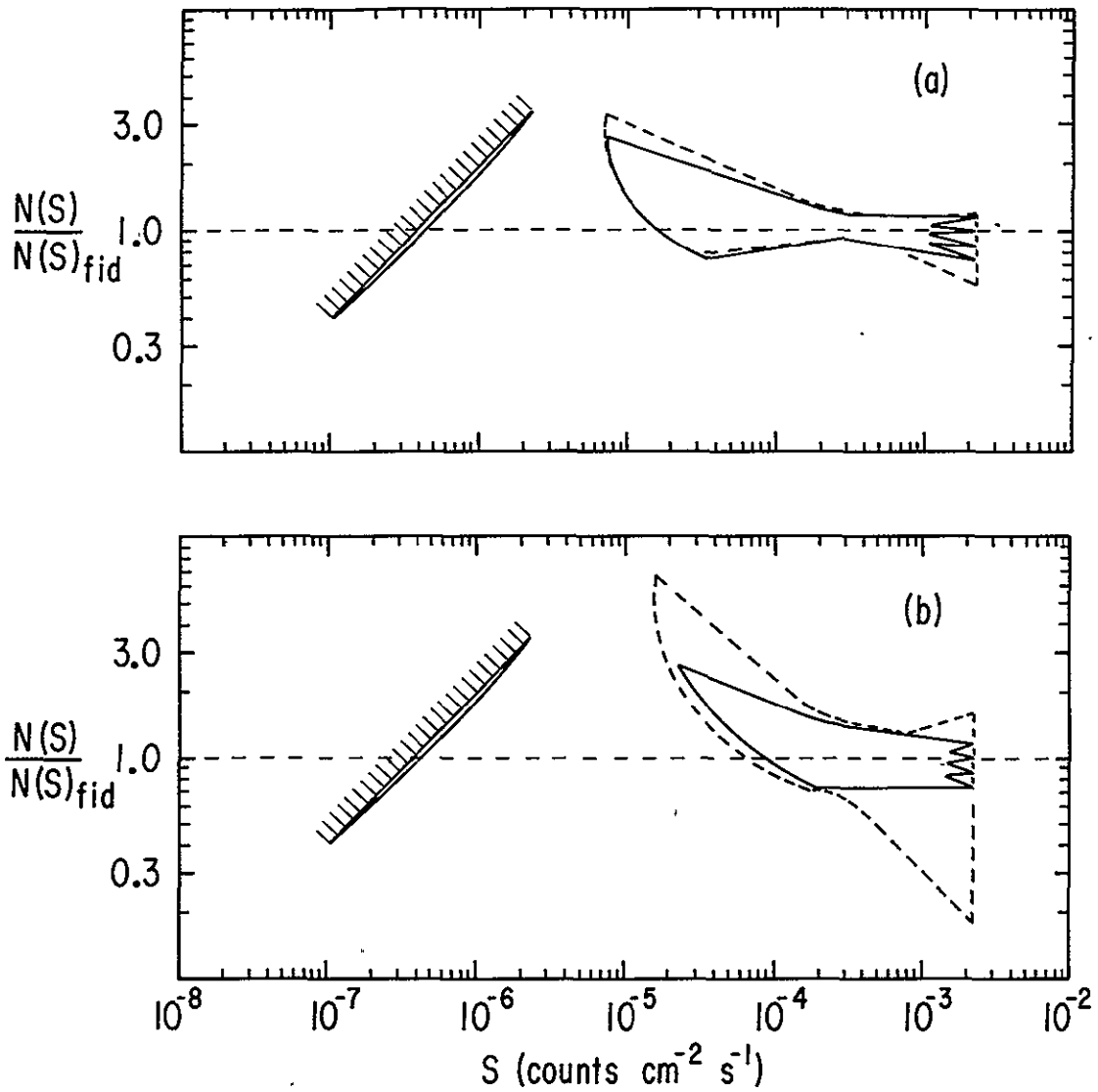


Fit results

Fits of Gamma and W

## 5: Fits of the Excess Variance and W

We describe fits when N(S) is parameterized as a pure Euclidean power law with an additional component whose contribution to the fluctuations is a pure Gaussian. The variance of the Gaussian, $\sigma^2_I$, and the value of W for the power law are strongly correlated. We present the 90% confidence contours for these parameters. The data are consistent with a zero excess variance, with a 90% upper bound to $\sigma^2_I$ of 0.064 (counts exp$^{-1}$)$^2$, or in collimator-size-independent units of flux, $\sigma^2_S$, 0.0008 (counts s$^{-1}$ cm$^{-2}$)$^2$. This limit is derived from the overlapping data WW dataset. The constraints imposed by the sets of independent measurements are much less severe. The limits on the excess variance is used to place limits on a broken power law model for N(S). We show how the confidence range would be modified if independent information could provide an exact value for W.

The validity of the N(S) region allowed by the fluctuations in Figure VI.6 depends on the assumption that the actual N(S) is indeed one of the models tested; i.e., it is a single power law extending from the survey limit to the point where the total flux of the sources equals that of the X-ray sky. This is a rigid and probably unrealistic condition. We may slightly relax this condition in that a sharp cutoff somewhere between equaling the sky-flux and the left-hand edge of the allowed N(S) region is indistinquishable, at some level, from the model N(S) that continues unbroken to the lowest flux values. Similarly, a doubling in N(S) to the left of the allowed region would produce an insignificant difference in $\Lambda$, making a model with such an excess indistinguishable from one of the models actually tested. It is difficult to assess whether a general N(S) is an acceptable model or not if it deviates from a single power law, particularly at fluxes higher than the left-hand edge of the allowed power law N(S) region. As we shall see, the expected behavior of N(S) for real source populations is not a single power law, so the parameterization of N(S) in terms of $\gamma$ and W as done in the previous section is not optimal for testing realistic N(S) models.

However, at high values of the flux, almost all source populations

asymptotically follow a Euclidean N(S). And we have discovered that a
low flux source's contribution to the fluctuations distribution can be
represented by a pure Gaussian term. Thus, a prefered parameterization
of the N(S) model is not in terms of $\gamma$ and W, but in terms of the W of
the asymptotic 5/2 Euclidean power law with an additional Gaussian
component of variance, $\sigma^2_I$ (which we call the excess variance). When
the fits are performed, the Gaussian is convolved with the Euclidean
power law generated $P_{I'}(I')$. If N(S) should fall below the Euclidean
law $\sigma^2_I$ is negative, indicating the deficit contribution to the variance
of the missing sources. Negative values for the excess variance can not
be handled directly in the convolution process. Instead a sharp low
flux cutoff is imposed on the Euclidean N(S) that is equivalent to the
removal of the given variance. As we saw in the discussion of Figure
III.13, such a sharp cutoff effects only the variance until the cutoff
value becomes quite high. The position of the cutoff is not, therefore,
important for interpreting the fits. In particular, the behavior of
N(S) above and below the Euclidean law that produces the excess variance
is not associated with any particular flux value, only that a divergence
from the Euclidean form is confined to the low fluxes where the sources
contribute principally to the variance.

Results of fits

The confidence regions for these fits are presented in Figure
VI.9. Not unexpectedly, the excess variance and the asymptotic value W
are strongly correlated. The parameterization of the Euclidean
component is given in terms of both W and K. The excess variance is
given in terms of the variance convolved with the intensity
distribution, $\sigma^2_I$, and in terms of $\sigma^2_S$. The two are related by

$$\sigma_I = R_2 \, \sigma_S \qquad\qquad\qquad [VI.4]$$

(see equation [III.69]). In turn, $\sigma^2_S$ is a solid angle independent
constraint on the non-Euclidean behavior of N(S) at low fluxes through
equation [III.68].

For each of the sets of data in Figure VI.9 there are two
confidence regions. The outer one is defined by $\Delta\Lambda < 4.60$ so that it

Fit results                                                    Fits of Excess Variance

Figure VI.9: Excess variance versus W

The 90% confidence regions for W or K and the excess variance, $\sigma^2_S$ or $\sigma^2_I$. The solid, dashed and dotted regions are for WW, SS1, and SS1-3 (Boole's Inequality) respectively. The inner of each pair of regions pertains when only one of the two parameters is considered interesting. The open circle and cross signify the best fit positions for the WW and SS1 datasets.
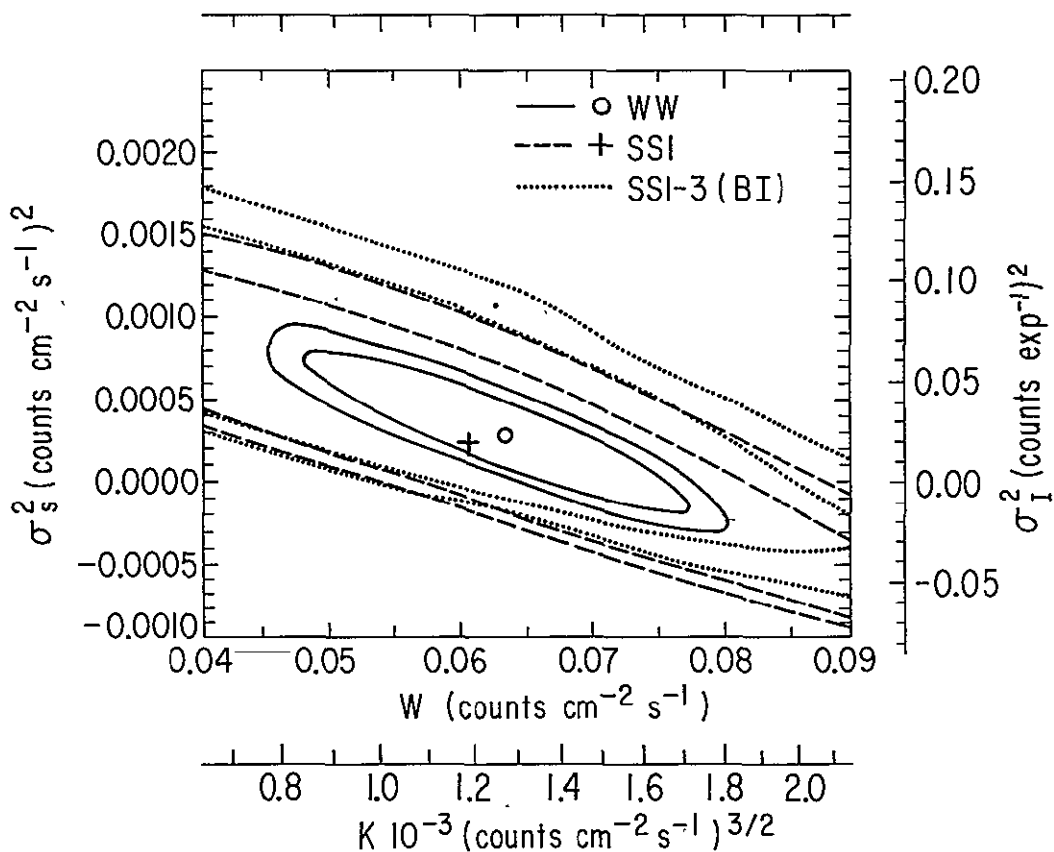
simultaneously contains the true value of both the variance and W 90% of the time. For this circumstance, where there are two "interesting" parameters, the delta likelihood range is defined with the $\chi^2$ distribution with two degrees of freedom. But if only one of the parameters were interesting, the relevant parameter space is indicated by the inner region, where $\Delta\Lambda \leqslant 2.71$, the 90% limit for $\chi^2$ with one degree of freedom. Thus we can conclude, using the WW dataset, either

that W ranges from 0.048 to 0.077 counts s$^{-1}$ cm$^{-2}$ <u>or</u> that $\sigma^2{}_I$ is in the range from −0.016 to 0.066 (counts exp$^{-1}$)$^2$. In terms of the fraction of the total sky flux, $\sigma_I/\bar{I}$ ranges from −0.74% to +1.49%.

Now the acceptable range for W is within an uncertainty of ±20% (90% confidence) from the current best fit position of 0.064 counts s$^{-1}$ cm$^{-2}$. This is roughly as accurate as the value derived from the resolved sources, but it is a significant degradation over the simplistic model which had a 90% range of 0.065 to 0.076 counts s$^{-1}$ cm$^{-2}$. Statistically there is no reason to prefer the model with the excess variance to the simple Euclidean model with no additional variance term. The models with zero variance provide acceptable fits. However, because of what is expected for actual source behavior the additional term is both justified and necessary.

To determine whether a particular model for N(S) is consistent with the limits imposed on the excess variance requires a check of the integral behavior of the model, as in equation [III.68]. We can crudely indicate the acceptable range of such behavior by modeling the N(S) as a Euclidean power law that breaks to another index at some low flux. Such a model must not exceed either the excess variance or the total sky flux. In Figure VI.10 we show the range of N(S) covered by models that either break to an index of 3.0, equivalent to strong source evolution, or to 2.0, roughly representing the rolloff from geometry and cosmological considerations. These actual values for the break are <u>ad hoc</u>, but they give a general idea of what constitutes acceptable behavior for N(S). Breaks that deviate further from a Euclidean index would indicate a somewhat different region, where the break occurs at a lower flux. To evaluate a particular model we should compare its behavior directly to the integral constraint of the excess variance.

These results were derived from the fits of the WW dataset. The sets of independent data provide much more loose constraints, in the same fashion that they failed to strongly limit the allowed values for Y. The independent sets include, or just barely exclude at the 90% level, models with a purely Gaussian component.

Fit results                                                     Fits of Excess Variance

Figure VI.10: Acceptable N(S) for broken power law

Models that break from the Euclidean value to either 3.0 or 2.0 are considered. The acceptable range of models is determined by the 90% limits on W and the excess variance from the WW dataset (both variables considered important, $\Delta\Lambda < 4.6$). The dashed region indicates the behavior of the W-$\gamma$ fits from the single power low models. The hatched line to the right shows where the total X-ray sky flux contributed by the model N(S) exceeds the observed flux.
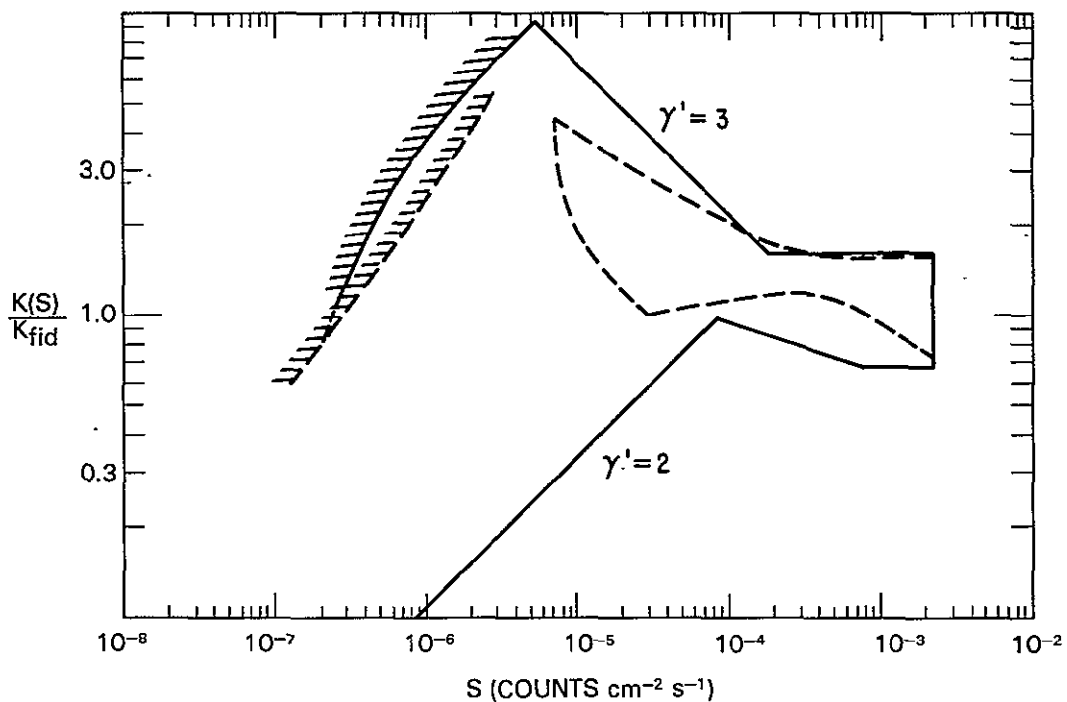


If we were given the exact value of W to use for the asymptotically Euclidean part of N(S) we could determine a signifcantly smaller region for the excess variance. This is illustrated in Figure VI.11. The confidence ranges shown there are drawn as a function of the external perfect determination of W. Note that (of course) all values of W now have an acceptable range of $\sigma^2_I$. The range is determined by fixing W, determining the minimum $\Lambda$ and finding the range of excess variance where $\Delta\Lambda < 2.71$. The case where the information about W is not

Figure VI.11: Confidence regions for excess variance, given W

The 90% confidence range is given provided that the exact value of W is known. The zigzag line gives the best fit excess variance as a function of the given W, for the WW dataset. All other symbols are the same as Figure VI.9.
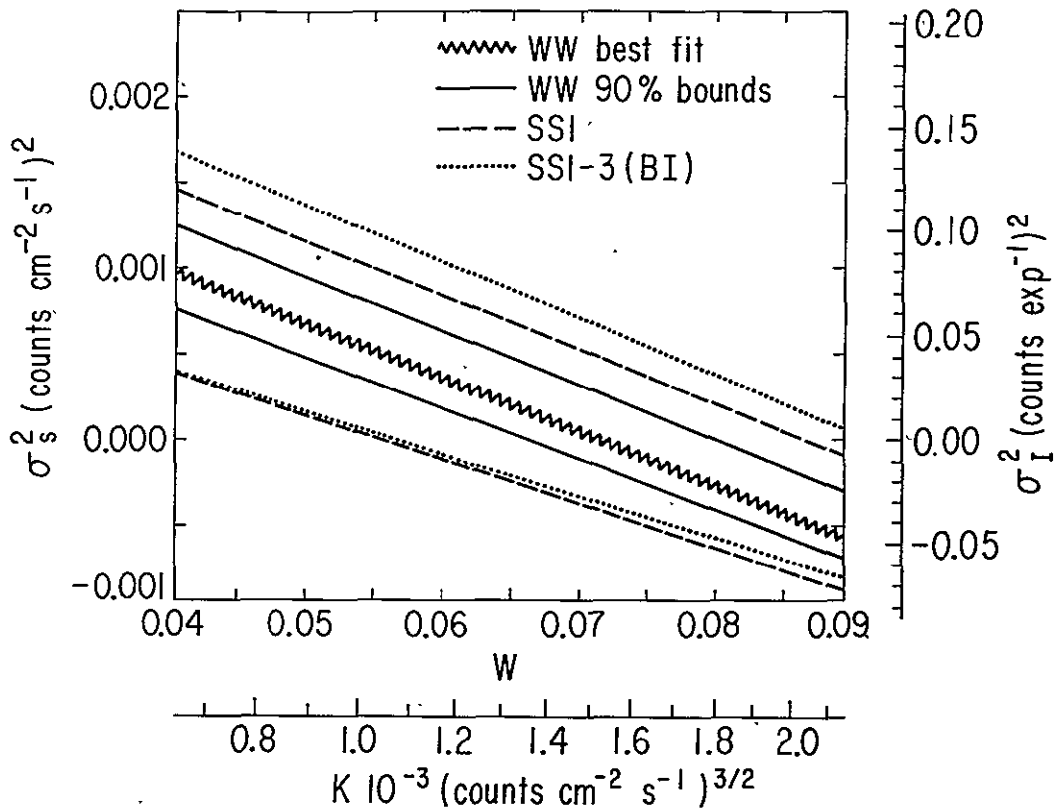


exact will be treated in the next chapter. However, it is worth noting that little improvement in the size of the acceptable range of the variance is expected for the WW dataset unless the outside information about W is better than the 20% uncertainty available directly from the fluctuations data.

6: Analysis of and Comparison with Other Experiments' Fluctuations
Results

Analyses of the fluctuations of the X-ray sky have appeared
in the literature using data from UHURU and Ariel V. One
analysis of the UHURU data measured only the variance of the
fluctuations. Given the limitations of such a technique it
is in general agreement with the results derived in this
chapter from the HEAO 1 A-2 data. An estimate of the excess
variance was made, but it is subject to several biases
weakening the published constraint. P(D) analyses were
performed on both UHURU and Ariel V data. However,
important effects, included in the A-2 analysis, were
ignored. The Ariel V published power law model confidence
region is in agreement with the A-2 region, although the
Ariel V results may be seriously flawed due to an arithmetic
error.

The fluctuations in the X-ray background have been a topic of
investigation since the era of sounding rockets [Schwartz et al. 1971;
Fabian and Sanford 1971; Fabian 1972]. An early satellite analysis was
performed by Schwartz [1970] using data from OSO III. These results
were generally upper bounds, where the total area, amount of active
time, and/or sky coverage were insufficient to provide a significant
measurement of any anisotropy in the unresolved high latitude X-ray
surface brightness. All three of these constraints were removed with
the first generation of all-sky survey instruments, starting with UHURU,
including Ariel V, and culminating in HEAO 1. Successful measurements
of the fluctuations have been made using datasets from each of these
instruments.

It is natural to compare the various results, if for no other
reason than to confirm the techniques used and the conclusions reached.
However there are three obstacles to any straightforward comparison:
(1) physical differences in the detector systems used, particularly
affecting the spectral response, which requires an unavoidably inexact
transformation of the measurements made with different units; (2)
differences in the fitting techniques used, ranging from minor
variations in assumptions and models tested to radical differences in

the quantity actually measured (e.g., an estimate of the variance of the fluctuations versus a fit of the actual $P_{I-\bar{I}}$ distribution); and (3) the presence of errors in the analysis requiring corrections or modifications before comparing two sets of results. (This difficulty is compounded by the inaccessibility of all the information available for the original analysis.) Keeping in mind these difficulties, we will briefly discuss four measurements of the X-ray sky fluctuations and compare their techniques and results with what we have presented in this dissertation. When assessing the different experiments, in terms of $\psi$, the ratio of the non-fluctuations noise to the fluctuations, the value for W we will use is the HEAO 1 best-fit Euclidean value of 0.0702 counts $s^{-1}$ $cm^{-2}$.

### UHURU as analyzed by Fabian [1975]

The UHURU satellite included two separate proportional counter detectors collimated by rectangular slats. The detector used in this and subsequent analysis of the fluctuations had a FWHM of $5.2^{\circ} \times 5.2^{\circ}$. The satellite made great circle scans of the sky, similar to HEAO 1, although not so methodical in coverage. The collimator slats were alligned parallel and perpendicular to the scan direction. (Many of the parameters of the UHURU experiment presented in this and following subsections are also based on information presented by Schwartz, Murray, and Gursky [1976] and Schwartz [1980].)

The active gas used in the detectors was argon, making it closer in energy response to the MED detector than to the xenon-filled HED 1 that we have used. The count rate used is often called a 2-6 keV measurement of the X-ray flux. The UHURU flux unit, or UFU, is defined by the flux of a source whose on-axis contribution to the intensity is one count per second. For typical spectra one UFU corresponds to $2.4 \times 10^{-11}$ ergs in the 2-10 keV band. The naive conversion to HEAO HED Layer 1 flux units (H1L1) is 1 count $s^{-1}$ $cm^{-2}$ is roughly equivalent to 560 UFU. Intensity and flux are measured in the same units. This definition normalizes the peak response of the unsmeared detector to 1. The response function, $r(\theta, \phi)$, and all of its weights are dimensionless.

Fit results                                    Comparisons with Other Experiments

The dataset used by Fabian in this analysis consisted of 105 measurments that had been smeared by $5^o$ along the scan direction. This corresponds to a weight for transforming a Euclidean power law model $W$ to $W_{eff}$ of $R_{3/2} = 5.3 \times 10^{-3}$, equivalent to an effective solid angle for a Euclidean power law of $\Omega_{eff} = 4.82$ msr (see equation [III.53]). This is slightly larger than the value calculated by Fabian, probably because he assumed the collimator size to be $5^o \times 5^o$ FWHM. The total count rate is 20.05 UFU, of which 17% is ascribed to internal background derived from measurements when the satellite was occulted by the Earth. This compares favorably to the level for the L+SFOV rates of the HED, in part because the larger UHURU solid angle increased the number of sky flux counts. In addition, UHURU's equatorial orbit helped minimize its internal background. Typical values for the counting statistics were 0.3 UFU. The value of $\psi$ is then 1.4, not including any contribution from variation in the internal background.

The data were binned in a frequency histogram of number of observations versus intensity. A $\chi^2$ test was used to fit power law $N(S)$ models to the data. The fluctuations' distribution was produced by Scheuer's closed form power series [Scheuer 1974], which is inherently restricted to power law indices between 2 and 3. The fluctuations distribution was convolved with a noise distribution representing the contribution to the fluctuations from counting statistics, and the resultant distribution was compared to the data histogram.

In preparing the data for analysis, Fabian made no correction for a galactic contribution or for possible noise from variation in the internal background. In addition the data used were contaminated by intervals where the detector was partially earth occulted [Schwartz, Murray, and Gursky 1976]. Arguably then, the results derived are in the nature of upper bounds on the behavior of the X-ray sky fluctuations. However, this is sufficient for the main conclusion of this paper that the fluctuations were too small when compared to the level expected from the resolved source counts derived by Matilsky et al. [1973].

A confidence region in K and Y was presented. However, it cannot

Fit results                                    Comparisons with Other Experiments

easily be compared with the HEAO 1 A-2 results because Fabian did not take several steps that we have found to be important. Fabian found the range in acceptable $\gamma$s to be roughly 2.28 to 2.80. We found with a dataset of similar size and precision (the SS1 set) that we were unable to constrain $\gamma$ better than the range 2.25 to 3.4. In his analysis, Fabian did not explicitly allow for a required low flux cutoff, nor did he include an explicit separation of the high flux sources when calculating the distribution mean. In fact, he did not try to fit the mean count rate with the fluctuations, but instead set the fluctuations' mean to be equal to the least squares mean. We have seen that this can introduce a significant bias. Even so, the best fit value derived of K = 25 $(UFU)^{1.5}$, $\gamma=2.5$, corresponds to a value of W, in HEAO units, of 0.083 counts $s^{-1}$ $cm^{-2}$. This is not far from the HEAO 1 confidence region.

In his analysis Fabian used the one source per beam criterion to estimate his lower limit of sensitivity at 0.17 UFU. That is, sources with intensities less than 0.17 UFU do not significantly affect his results. It should be remembered that this limit corresponds on average to a source with flux 2.32 times that intensity. Thus the limit after allowing for the smeared detector response is 0.39 UFU.

Fabian and Rees [1978], using the results of Fabian [1975], estimated an upper limit to the excess variance from evolved sources. Without giving any details, their stated upper bound was 1 per cent of the diffuse sky brightness. We have calculated that their smeared detector had a weighted response of $R_2$ = 0.0156, so we derive their limit on $\sigma^2{}_S \leq 112$ $(UFU)^2$, or in HEAO 1 detector units $3.6 \times 10^{-4}$ (counts $s^{-1}$ $cm^{-2})^2$. This is substantially smaller than the upper limit from the HEAO fits. However, without access to the details of the calculation of the Fabian and Rees limit, it is difficult to assess the accuracy. In particular, no details are given as to the determination of the size of N(S) from unevolved sources, an important aspect in fixing the magnitude of the excess. The best fit value of the fluctuations size by Fabian was actually larger than the best fit values derived from HEAO, so that if anything we would expect the upper bound of the excess variance from

the UHURU results to be larger.

## UHURU as analyzed by Schwartz, Murray, and Gursky [1976]

This analysis did not fit the distribution of the intensities, but instead estimated the size of the variance of the fluctuations distribution. If $\gamma$ is assumed to be 5/2 then equation [III.56] may be used to estimate the effective width:

$$W_{eff} \doteq [\ 0.5\ \delta^2_{\ I}\ I'^{-0.5}_u\ ]^{2/3}\ . \qquad\qquad [VI.5]$$

$I'_u$ is the upper intensity limit; all measurements with intensities greater than $\overline{I}+I'_u$ are not included in the determination of the variance. $\delta^2_{\ I}$ is the total variance of the distribution, in previous equations represented by $\sigma^2_{\ I'}$. The notation has been changed for this section to avoid confusion with the excess variance, $\sigma^2_{\ I}$.

The dataset consisted of 122 bins that were smeared by $10^o$ in the scan direction rather than by the $5^o$ bin smear that Fabian used. In general such an increase in the smear should decrease the uncertainty due to counting statistics, as the number of exposures per bin is increased. In this case, though, the rms deviation due to counting statistics increased to 0.42 UFU. Perhaps Schwartz, Murray, and Gursky used a more stringent data selection criterion than Fabian, or fewer scans were superposed. An estimate of the variation of the internal background was made by examining data taken when the detector was totally occulted by the earth, $\sigma_{bkg} = 0.22^\pm0.14$ UFU (one sigma). No model for galactic or any other large scale contamination was used. We have calculated the response moments for the UHURU detector smeared by $10^o$ in the scan direction and by $1^o$ in the perpendicular direction to account for uncertainty in the spinaxis pointing: $R_{5/2} = 0.0247$; $R_2 = 0.0131$; and $R_{3/2} = 0.0047$ (corresponding to an $\Omega_{eff}$ for a Euclidean N(S) of 4.05 msr). The value of $\psi$ is ~2.6.

All 122 observations were used to show that the data were inconsistent with total isotropy. A subset of 28 bins with no overlap in sky exposure was used to determine a best fit sky variation, $\delta_I$, of 0.49 UFU. A $\chi^2$ test was used to determine confidence bounds, although

Fit results                                          Comparisons with Other Experiments

this is strictly applicable only for purely Gaussian distributions. The 96% range in values was 0.18-0.84 UFU. Schwartz, Murray, and Gursky state that they exclude measurements with intensities of 3 UFU or greater; therefore, using equation [VI.5] we derive the 96% range for $W_{eff}$ of 0.044-0.3462 UFU. The corresponding value for K is 2.3-50 $UFU^{1.5}$. The W range is 9.44-73.7 UFU, or in HED 1 units 0.017-0.13 counts $s^{-1}$ $cm^{-2}$. This large confidence range is easily consistent with our results.

The value derived for W is a function of the upper level intensity cutoff, as well as the size of the fluctuations variance. Properly this limit is a criterion applied to the smeared bin intensities. However in the text of their paper, Schwartz, Murray, and Gursky state that the 3.0 UFU limit was applied on a level of individual exposures. If no single exposure could be more intense than 3.0 UFU, it is impossible that a bin consisting of the average of many exposures could ever be that bright. The equivalent intensity limit for the bins must be smaller than 3 UFU. A source of intensity 3.0 UFU in a single exposure has on average a flux of 5.9 UFU (see equation [III.36]). In the smeared bin a source with intensity 1.75 UFU will on average have the same flux. If we use 1.75 UFU as an estimate of the proper value for $I'_u$ we must rescale the above results for K by $(1.75/3)^{-0.5}$ and W by $(1.75/3)^{-1/3}$ giving 3-65 $UFU^{1.5}$ and 0.02-0.16 counts $s^{-1}$ $cm^{-2}$ respectively. Note that with this revised value the upper bound on K is now consistent with the value of 60 $UFU^{1.5}$ derived from the UHURU resolved source counts by Holt et al. [1974] based on the Matilisky et al. [1973] data.

### Analysis of the Ariel V data by Pye and Warwick [1979]

The Sky Survey Instrument (SSI), one of several experiments on the Ariel V satellite, had a 290 $cm^2$ proportional counter covering the energy range from 2-18 kev. As in the case of UHURU, both flux and intensity are measured in the same units, in this case SSI counts $s^{-1}$, which we will abbreviate as SSI. 1 SSI is roughly equivalent to $5.1 \times 10^{-11}$ ergs $s^{-1}$ $cm^{-2}$ (2-10 keV), depending on the spectrum. This corresponds to ~265 SSI for 1 count $s^{-1}$ $cm^{-2}$ in the xenon counter layer

Fit results                                       Comparisons with Other Experiments

one units of HEAO 1.  The mean sky intensity was 3.42 SSI.  The count
rate from particles (internal background) in a particular measurement
took a wide range of values, with 4.4–10 SSI being the range for one
scan of superposed data.

The detector had a rectangular collimation system with a FWHM of
$0.75^{\circ}$x$10.6^{\circ}$.  However, the detector was not placed on the satellite so
that its collimator was alligned with the scan direction.  That is, the
long axis was inclined $65^{\circ}$ from the scan plane.  For smearing purposes,
the detector was treated as equivalent to one of $9.6^{\circ}$x$0.97^{\circ}$ aligned with
the long axis perpendicular to the scan.  (This and other points not
covered in Pye and Warwick's published paper were provided by a private
communication.)  The data were superposed in $4.2^{\circ}$ bins.  If we allow for
$1.5^{\circ}$ of smear perpendicular to the scan plane the response moments are
$R_1 = 0.000227$, $R_{3/2} = 0.00187$ (equivalent to $\Omega_{eff} = 1.02$ msr), and $R_{5/2}$
$= 0.0108$.  The perpendicular smear is a small perturbation.

A typical value of the variance due to counting statistics is
~0.14 SSI.  An estimate of the variance due to uncertainties in the
particle contribution was 0.034 SSI.  The value for $\psi$ is 4.8.  A simple
model for the galaxy was subtracted, where $I_{gal} = 0.05$ cos(2b) (SSI).
The Pye and Warwick analysis was the first truly all–sky analysis of the
X-ray fluctuations.  There were 1110 data bins, but they were not
independent.  The bin weightings were estimated in a way philosophically
similar to equation [IV.4].  Instead of estimating the degree of
interdependence on a bin by bin basis, Pye and Warwick gridded the high
latitude sky into ~270 equal solid angle cells, each about 98 square
degrees in size.  If the center of the field of view was inside a cell,
a bin was counted as being in that cell.  Each bin was then weighted
inversely by the total number of bins that were in the same cell.
Therefore the weighted number of observations was about 270, the number
of cells.  Because the total sky area covered by a bin was actually
larger than the cell sky there was some correlation between the cells,
but this was counterbalanced by the probable overweighting of the
measurements within a cell, i.e. effectively treating all the bins as
if they exactly overlapped.  The fact that the best fit model had a $\chi^2$

Fit results                              Comparisons with Other Experiments

of 2.55 for 5 degrees of freedom indicates that the data were overweighted. The derived confidence region in this case would be conservative.

Figure VI.12: The 90% γ-W confidence region for Ariel V

We compare the results from Ariel V with our derived results from HEAO 1. The best fit value for the analysis by Fabian of the UHURU data is also shown by the + symbol.
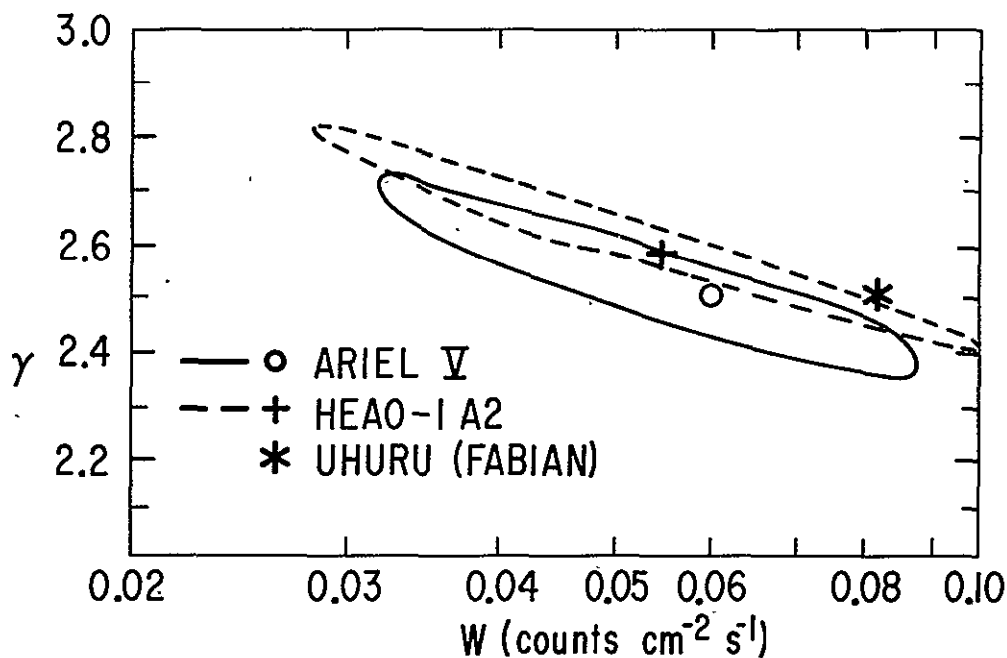


Other than including overlapping data and making an estimate of the internal background variance, the fitting technique used was similar to Fabian's [1975]. Again the fit was based on a $\chi^2$ test of a histogram of measurements versus a derived histogram for the fluctuations, using the Scheuer [1974] series expansion. Therefore the models were restricted a priori to power laws with index between 2 and 3. No attempt was made to fit the mean count rate or to correct for the modification in the calculation of the mean for high flux sources. As remarked above this could produce a significant bias in the results. Even so, the Ariel V confidence region is essentially coincident with

Fit results                                    Comparisons with Other Experiments

the HEAO 1 region (see Figure VI.12).

For the smeared detector response of the Ariel V bins, the average flux of a source is 6.6 times its contribution to the intensity. The intensity with one source in the field of view is roughly 0.035 SSI, corresponding to a flux of 0.23 SSI ($8.7 \times 10^{-4}$ counts $s^{-1}$ $cm^{-2}$). This is only a factor of 2 below the limit of resolved source counts, 0.5 SSI, from the Ariel V all-sky survey as analyzed by Warwick and Pye [1978]. However, as is the case for the HEAO 1 analysis, the Ariel V results probably apply to fluxes below the one source per field of view limit. This flux of the one source per field of view limit calculated here differs from the value presented in Pye and Warwick, as they did not make the distinction we have demonstrated between flux and intensity. Pye and Warwick did not fit data with intensities more than 0.38 SSI above the mean sky intensity. The actual upper flux limit, on average, is at 2.5 SSI. Pye and Warwick did not determine any limit on the excess variance with the Ariel V data.

The value we have calculated for $\Omega_{eff}$, 1.02 msr, is not the same as the value quoted in Pye and Warwick of 2.12 msr (assuming that $\gamma$ = 5/2). This discrepancy is serious. The Pye and Warwick value is almost certainly in error. It corresponds to an effective solid angle of 7.0 square degrees, which is larger than the effective solid angle of the unsmeared detector, $(16/25)\theta_0\phi_0$ = 6.0 square degrees. Any smearing can only reduce the magnitude of $\Omega_{eff}$. If Pye and Warwick actually used their published value in their analysis, and if the true value is as we have calculated it, then their results must be adjusted. The Euclidean values for W plotted in Figure VI.12 must be multiplied by a factor of 1.6, destroying the agreement between the Ariel V and HEAO 1 regions. K must be rescaled by a factor of 2.1, eliminating the smooth continuation of the Ariel V resolved source counts of Warwick and Pye [1978] to their fluctuations result. Similar rescalings are probably necessary for non-Euclidean values of $\gamma$, although exact numbers are impossible to give as Warwick and Pye gave only their Euclidean value for $\Omega_{eff}$. We are faced with explaining this large deviation from what was expected. The determination of the background variance is probably one of the weakest

Fit results                                              Comparisons with Other Experiments

steps in their analysis. If the true variance of the internal background were much larger than they have estimated, then the fluctuations would have measured values in excess of what could be caused by the unresolved sources. It is possible that the published number for $\Omega_{eff}$ was a typographical error, and that they used the proper value in their analysis. In this case the published analysis and results are correct.

## The excess variance from UHURU as analyzed by Schwartz [1980]

A value for the excess variance that has been used in the literature [e.g. Cavaliere, Danese, DeZotti, and Franceschini 1980]> is given by Schwartz [1980] based on his analysis of an UHURU all-sky database. He derived an upper bound to the excess variance of 1.3% of the sky flux in the UHURU $5.2^{o}$x$5.2^{o}$ detector. We will later see that the excess variance can be used to set a lower limit on the number of sources in the sky. Schwartz calculated that his bound corresponded to 7.5 million sources (all-sky). This means $\sigma^2_s = 4.4\text{x}10^{-4}$ (counts $s^{-1}$ $cm^{-2}$), a value signicantly better than the current HEAO 1 bound of $8\text{x}10^{-4}$. However the UHURU derivation is flawed at several key points.

The initial preparation of the data was identical to the techniques used in Schwartz, Murray, and Gursky [1976]. The data were superposed on scan great circles, each bin being $10^{o}$ in scan angle. Single exposures with intensities greater than 3.0 UFU were excluded (corresponding to an intensity limit for each bin of ~1.75 UFU). In order to use all the data, the measurements were rebinned into cells aligned in galactic coordinates of roughly $10^{o}$x$10^{o}$ in extent. Unlike the Ariel V case where the rebinning was only for purposes of determining the weight of a measurement, here the scan bins were summed to make a single measurement of the X-ray sky brightness. Although there was still some correlation between the galactic cells, Schwartz assumed that each measurement had equal weight. A more important omission was that he continued to use the smeared response function calculated for the original scan bins. By superposing the data into galactic cells additional smearing occurs. Calculating the correct

Fit results                                                     Comparisons with Other Experiments

value is difficult, as the exposure within a sky cell is not uniform, but for our purposes we will assume that the superposition into the galactic cells is equivalent to a <u>second</u> <u>additional</u> smearing of $10^o$ in the scan direction and a total smear of $5^o$ in the perpendicular direction. The weights for such a twice smeared detector are: $R_{5/2}$ = 0.0195, $R_2$ = 0.0107, and $R_{3/2}$ = 0.00408 (equivalent to $\Omega_{eff}$ for Euclidean N(S) of 3.28 msr). With this additional smearing the upper intensity limit, $I_u$ is reduced further to ~1.2 UFU. The accumulation of more exposures per measurement decreases the size of the noise due to both counting statistics and internal background variation, but the greater smearing also reduces the size of the fluctuations signal: $\psi$ = 1.8.

The measurements were combined into sets with the same absolute value in galactic latitude, and the variance due to the fluctuations was estimated in each set. This reduced the contribution of the galactic anisotropy without explicitly modeling its latitude dependence, although any longitude variation would not be excluded. The combined estimate for the size of the fluctuations variance, $\delta_I$ (not to be confused with the excess variance $\sigma_I$), is 0.30-0.45 UFU (90% confidence range). Schwartz reported a range for K of 9-22 $UFU^{1.5}$, which is actually close to the range 13-28 $UFU^{1.5}$ that is obtained with the correct values for the intensity limit and smeared response weights. This is fortuitous because he mixed the differential and integral forms of his N(S) relationship, which multiplied fluctuations result by a factor of 3/2, close to the factor of ~2 required to compensate for the wrong detector weights.

Our current bounds on the excess variance from HEAO 1 were found by directly fitting a Euclidean power law with an additional component. Schwartz only measures one aspect of the fluctuations, the variance of the distribution, and if he is to find a bound for the excess variance he must include some outside information. He uses the revised UHURU source counts which he states have a range of $K_{sc}$ of 12-22 $UFU^{1.5}$ (90%). $K_{sc}$ is the K value for the Euclidean N(S) law associated with the resolved source counts. Using equation [III.56] we find the predicted

Fit results                                        Comparisons with Other Experiments

range in the fluctuations variance is $\delta^2_{I;sc} = 0.123\pm0.023$ UFU$^2$ (1 sigma). If the 1 sigma range in the measured variance is given by $\delta^2_{I;meas} = 0.146\pm0.034$ then the 90% upper bound on the excess variance is $\sigma^2_I \lesssim 0.075$ UFU$^2$. We crudely test the magnitude of the effect due to the partial overlap between the presumed independent galaxy cells by assuming that the putative 90% range on $\delta^2_{I;MEAS}$ is only an 80% confidence region. In this case the 90% upper bound on the excess variance is $\sigma^2_I \lesssim 0.086$ UFU$^2$, a small effect.

For $\sigma^2_I$ of 0.075, an excess variance of 1.7%, we calculate the detector solid angle independent measure, $\sigma^2_S$, to be 661 UFU$^2$, or in our detector units $2.11\times10^{-3}$ (counts s$^{-1}$ cm$^{-2}$)$^2$. If we estimate the number of sources, all-sky, using the same technique (but with the right coefficients) as Schwartz did we find that his data are compatable with a lower bound of 1.6 million sources, or 38 per square degree. In the next chapter, we derive the HEAO 1 lower limit of 50 sources per square degree.

CHAPTER VII

ANALYSIS OF FLUCTUATIONS FIT RESULTS

Our confidence ranges for the abstract N(S) models tested in the previous chapter are now compared to other data and more realistic models of the universe.

----------------------------------

1: Comparison of Resolved Sources N(S) and the Fluctuations

The data of Piccinotti et al. [1982] for a complete sample of resolved sources are reviewed. The sources are binned by intensity and used to estimate K, the normalization coefficient, for an assumed Euclidean N(S). The treatment differs from that in Piccinotti et al. in that it includes identified galactic sources and excludes the highest flux sources as N(S) is expected to deviate from a single power law model at the highest fluxes. Corrections for the Malmquist bias are also estimated. The range of values for K is consistent with the Euclidean component of the fluctuations. We combine the resolved source fits with the fluctuations analysis to produce a significant reduction in the confidence regions for K (or W) and the excess variance, especially for the independent data subsets. The 90% confidence range for $\sigma^2_I$ is 0.0 to 0.05 (counts exp$^{-1}$)$^2$, (or $\sigma^2_S$ is 0 to 6x10$^{-4}$ (counts s$^{-1}$ cm$^{-2}$)$^2$) using overlapping data.

<u>The sample</u>

During the period from days 248-437 (1977), when the entire sky was scanned, a complete sample of sources was compiled. This period, referred to as the "first scan", was essentially contemporaneous with the measurements used in the fluctuations analysis. The survey did not cover regions within 20$^\circ$ of the galactic plane or within 6$^\circ$ of the Large Magellanic Cloud (LMC), leaving a total solid angle coverage of 8.23 sr (65.5% of the sky).

Source fluxes were measured using a combination of detector layers and fields of view: both layers of HED 3 with the small field of view plus the small field of view, Layer 2, of the MED. This combination gives fluxes in R15 units (see Marshall et al. [1979] for a complete description). Using the spectral response information in Tables &tbd2. and &tbd3., as well as the areas from II.1 we can construct the conversion coefficients from R15 to ergs $s^{-1}$ $cm^{-2}$. For thermal bremsstrahlung spectra, the conversion factors are $2.06 \times 10^{-11}$ and $2.47 \times 10^{-11}$ ergs $s^{-1}$ $cm^{-2}$ $(R15)^{-1}$ for kT = 40 and 6 keV respectively. For a single power law of photon index $\Gamma$ = 1.65 ($\alpha$=0.65), the conversion is $2.14 \times 10^{-11}$ ergs $s^{-1}$ $cm^{-2}$ $R15^{-1}$. The conversion to ergs varies by more than 10% for the spectra typical of extragalactic sources. This might cause difficulties in comparing results from the source survey, measured in R15, with the fluctuations fits. Fortunately we can use the same tables to estimate directly the conversion from R15 to HED 1 Layer 1 units, without going through ergs. Because in this chapter we will be comparing different measures of flux, the simple designation of "counts" is ambiguous without explicit reference to the detectors used. Therefore we will identify 1 count $s^{-1}$ $cm^{-2}$ in HED 1 Layer 1 as 1 H1L1. We estimate that 1 H1L1 is equivalent to 620, 623, and 622 R15 respectively for the three spectra used above. This conversion has very small dispersion. For our purposes we will use the general conversion 623 R15 $\equiv$ 1 H1L1. This conversion coefficient indicates that the R15 combination has an effective area equivalent to about 623 $cm^2$ of the first layer of a xenon (HED) detector. Thus the inclusion of the additional layers provided ~50% more counts than available from just the first layer, small FOV, of HED3.

For the first scan the survey was complete down to a level of 1.25 R15 ($2 \times 10^{-3}$ H1L1). (The completeness level was set so that all the sources were detected with a flux at least $5\sigma$, as defined by photon statistics, away from zero.) 76 sources were detected by this criterion, not including the local sources from M31 and the Magellanic Clouds, the high latitude globular clusters and the two brightest galactic binary X-ray sources, Sco X-1 and Her X-1 (see Table VI.2[*]).

[*]Two moderate flux galactic sources, EX Hyd and AM Her, were excluded

Only 2 of the 76 are currently without firm or suggested identifications, an identification rate in excess of 97%. Of the identified sources, 64 are extragalactic objects: viz. 30 clusters of galaxies, 33 active galactic nuclei (AGN), and 1 peculiar galaxy (M82). Of the AGN, 4 are BL Lac objects, 1 is a QSO (3C 273), and the remainder are generally type 1 or type 2 Seyfert galaxies although other forms of activity are also included (e.g. NGC 7172). (These classifications use additional information obtained since the publication of Piccinotti et al. Details are given in Appendix A5.)

### Fitting $N(S)$

Piccinotti et al. fit a power law index for $N(S)$ of the extragalactic sources using the "absolute maximum likelihood" method of Lightman, Hertz, and Grindlay [1980], a modification of the technique of Crawford, Jauncey, and Murdoch [1970]; (also Murdoch, Crawford, and Jauncey [1973]). This technique allows the use of unbinned data with errors; however it does not directly allow the calculation of the power law coefficient. They were able to show that the Euclidean power law index, $\gamma = 5/2$ provided an acceptable model for the extragalactic sources and unidentified objects. They did not fit an $N(S)$ model for their identified galactic sources. In the analysis of the fluctuations we are, of course, unable to distinguish variations produced by galactic sources from any fluctuations due to extragalactic objects. Therefore to compare properly the fluctuations results to the $N(S)$ of resolved sources we should not exclude the galactic sources.

Though we can construct likelihood techniques that analyze unbinned measurements for the simultaneous estimation of the norm and index of $N(S)$, we will use an easier-to-implement technique utilizing binned data. We divide the sources into M bins where the observed number of sources in the ith bin is $n_i$. This value is expected to be drawn from a Poisson distribution where the mean, $\bar{n}_i$, is determined by the model $N(S)$ being tested, so that $n_i$ is drawn from a probability

_____

from the fluctuations database, but are included in this sample as they are examples of cataclysmic variables (CV), other examples of which will contribute to the fluctuations.

Analysis of Results                                    Resolved Sources

distribution

$$P_i(n; \bar{n}_i) = (\bar{n}_i)^n \exp(-\bar{n}_i) / n! \ . \qquad\qquad [VII.1]$$

We can define confidence regions on the model parameters through the log-likelihood difference (see equation [V.9] and associated discussion). Ignoring constant terms the relevant statistic is

$$\Lambda(\{\bar{n}_i\}) = 2 \sum_i (\bar{n}_i - n_i \ln \bar{n}_i ) \ . \qquad\qquad [VII.2]$$

Other examples of the analysis of Poisson-distributed data using similar equations have been extensively examined by Cash [1979]. For large values of $\bar{n}_i$ the value for $\Lambda$ approaches the form of $\chi^2 = \sum (n_i - \bar{n}_i)^2/ \bar{n}_i$. However, when the expected number of sources per bin is small, $\lesssim 5$, the non-Gaussian nature of the Poisson distribution would significantly bias the best fit value. Worse, the confidence range estimated using $\Delta\chi^2$ will be seriously in error even for a mean of $\leq 15$ observations. Recall also that even the likelihood derived contours are only asymptotically correct, but will generally be accurate when there are as many as 10-15 sources in all the bins.

For a given N(S), the mean number of sources observed in an interval defined by the flux range $\Delta S_i$ is

$$\bar{n}_i = \int_{\Delta S_i} dS \ N(S) \ . \qquad\qquad [VII.3]$$

If we assume that the N(S) law is a power law with index $\gamma$ over the interval $S_\ell < S < S_u$, then the expected number of sources becomes

$$\bar{n} = \Omega \ K \ (S_\ell^{1-\gamma} - S_u^{1-\gamma}) / (\gamma-1) \ . \qquad\qquad [VII.4]$$

This can be inverted so that the observed number of sources can be used to estimate K.

So far we have assumed that we can determine perfectly the source flux so that we know exactly the flux range bin in which a particular source belongs. It has long been noted that if there are statistical or other uncertainties associated with a flux determination, corrections must be made. If N(S) is a power law $\propto S^{-\gamma}$, $\gamma>0$, then the expected number of observations in a bin is larger than the naive prediction, a

Table VII.1: Resolved source distribution, binned

| bin | $S_l$ R15 | $S_u$ R15 | $n_i$ extra- galactic | $n_i$ galactic | $n_i$ total | $\overline{\sigma}_i$ R15 | $f_i$ |
|-----|------|------|------|------|------|------|-------|
| 1 | 5.0 | $+\infty$ | 4 | 0 | 4 | 0.28 | 1.007 |
| 2 | 3.15 | 5.0 | 7 | 2 | 9 | 0.24 | 1.018 |
| 3 | 2.404 | 3.15 | 13 | 1 | 14 | 0.22 | 1.030 |
| 4 | 1.984 | 2.404 | 9 | 2 | 11 | 0.22 | 1.048 |
| 5 | 1.710 | 1.984 | 9 | 0 | 9 | 0.24 | 1.084 |
| 6 | 1.514 | 1.710 | 7 | 1 | 8 | 0.23 | 1.104 |
| 7 | 1.366 | 1.514 | 6 | 1 | 7 | 0.23 | 1.136 |
| 8 | 1.25 | 1.366 | 11 | 2 | 14 | 0.21 | 1.137 |

From bins 2-7, best fit and 90% range $\overline{n}$ = 9.53 (7.80,11.49).

bias known as the Malmquist effect. Simply stated, there are more sources with true fluxes below the bin which can be "scattered" by statistical imprecision to observed fluxes covered by the bin than there are sources in the bin to be scattered out of it. To quantitatively estimate the enhancement we assume that the statistical process is Gaussian with variance $\sigma^2$. The probability that a source with true flux S will have an observed flux S' is then

$$P(S';S,\sigma) \, dS' = dS' \, \exp[-(S'-S)^2 \, / \, 2\sigma^2] \, / \, (2\pi \, \sigma^2)^{0.5} \, . \qquad [VII.5]$$

In the presence of the Gaussian error we calculate that $N'(S')$, the number of sources with <u>observed</u> flux S' is

$$N'(S') \, dS' = dS' \int dS \, N(S) \, P(S';S,\sigma) \, . \qquad [VII.6]$$

Therefore to find the mean number of objects in a flux bin we replace N(S) in equation [VII.3] with $N'(S')$ to get the corrected $\overline{n}'_i$, which in turn replace the $\overline{n}_i$ in [VII.2] when calculating the model confidence region. The corrected mean, $\overline{n}'_i$, is a potentially complicated integral of the model N(S). However when we assume we know the <u>form</u> of N(S) and

wish only to fit its normalization, we can calculate the correction
factor

$$f_i = \overline{n}_i' / \overline{n}_i \quad . \qquad\qquad\qquad [VII.7]$$

Finding the confidence range for the model normalization now involves
replacing $\overline{n}_i$ in equation [VII.2] with $f_i \overline{n}_i$. For power law N(S) models
with $\gamma > 1$ the correction becomes infinite as S goes to zero. This is
avoided by imposing a low flux cutoff in N(S). It was to prevent the
position of the cutoff from having a significant impact on calculating
the correction that a requirement of $5\sigma$ detections was imposed on the
sample, following Murdoch, Crawford, and Jauncey [1973]. After all, the
cutoff is a property of N(S) far from the fluxes where the actual
sources are measured. (However, see Appendix A6 for further discussion
of this point.)

We divide the range above 1.25 R15 into 8 bins, with identical
expected number of sources in each bin using equation [VII.4]. Listed
in Table VII.1 are the flux range, the number of sources, the average
flux uncertainty, $\overline{\sigma}_i$, and the Malmquist bias correction assuming a
Euclidean power law form for the model N(S). The correction term is
largest for the low flux bins where the number of sigma from zero to the
lower flux edge of the bin is smallest.

We note that the first, high flux, bin has only 4 sources. Though
this is not more than $1.5\sigma$ below an expected 9.5 sources (not counting
any Malmquist correction) there is a reason to expect that the source
counts may not be strictly Euclidean in the higher flux bins so that the
first bin should be excluded. As will be discussed in the next section,
active galaxies have a luminosity function in the range $10^{43}$ to $10^{45}$
ergs $s^{-1}$ that is well described by a power law of index 2.75. The
contribution to source counts by a population with this luminosity
function, as noted in the discussion following equation [III.9], is
dominated by the low luminosity objects. 8 of the 28 AGNs have
luminosities $<10^{43}$ erg $s^{-1}$, with a mean of $\sim 4.5 \times 10^{42}$. For a spectrum
typical of a low luminosity active galaxy (see next section) a source of
intensity 5 R15 and luminosity $4.5 \times 10^{42}$ erg $s^{-1}$ will be at a redshift of

~0.0029. For comparison, the redshift of the Virgo Cluster is 0.0037.
If the universe is inhomogeneous on this scale, in particular if active
galaxies are under-represented in a sphere with radius $z \sim 0.003$ and
centered on our galaxy, then source counts at 5 R15 and brighter will be
below the Euclidean power law of lower fluxes. Because such
inhomogeneity is plausible, and to some degree observed in local galaxy
surveys (see e.g. Tully [1982]), we will only fit sources in bins 2-7,
with intensities from 1.25 to 5 R15. Number counts of clusters, which
have a flatter luminosity function, are dominated by high luminosity
objects, and are not expected to show a significant deficit in the high
flux bin. As it turns out, three of the four objects with $S > 5$ R15 are
indeed clusters.

We will fit a Euclidean power law model, so that each bin has
identical $\bar{n}$. The best fit value and 90% range is 9.53 (7.80,11.49). If
the Malmquist effect had been ignored the values would be roughly 8%
higher, while if all 8 bins had been fit the answer would be ~7% lower.
We can calculate an equivalent value of K by

$$K = 2.04 \; \bar{n} \; R15^{1.5} = 1.31 \times 10^{-4} \; \bar{n} \; (H1L1)^{1.5} \qquad [VII.8]$$

so that the values for K derived from the fit are $1.24(1.02,1.51) \times 10^{-3}$
$(H1L1)^{1.5}$. This is in agreement with, and smaller than, the 90% range
in K from the fluctuations fits, which is, regarding the excess variance
as a nuisance parameter, $0.8 \times 10^{-3}$ to $1.7 \times 10^{-3}$ $(H1L1)^{1.5}$ (see Figure
VII.1).

We can estimate K on a bin by bin basis using [VII.4]. To
properly place this estimate on Figure VII.1 we convert the K to the
K(S) form:

$$K(S)_i = K \; S^{2.5-\gamma} = S^{2.5-\gamma} \; \frac{(\gamma-1) \; (n_i/f_i)}{\Omega \; (S_\ell^{1-\gamma} - S_u^{1-\gamma})} \; . \qquad [VII.9]$$

$K(S)_i$ is the estimate based on the observation of $n_i$ sources in the ith
bin. When the assumed power-law within the bin is Euclidean, $K(S) = K$.

Figure VII.1: $K(S)/K_{fid}$ from resolved source counts

The hatched regions indicate the 90% range for K for a
Euclidean model as derived from source counts or the
fluctuations. The heavy horizontal lines are the inferred
value for K(S) in each bin based on source counts. The
central dashed line is the best fit K for the resolved
sources, with the upper and lower dashed lines indicating
the expected one sigma variation on K(S) as measured from
the seven low flux bins. The K(S) uncertainty for the
high flux bin is indicated separately. The horizontal
arrow on the upper flux bin illustrates that it contains
all sources brighter than 5 R15. The dotted continuation
of the lowest flux bin to below $2x10^{-3}$ indicates the
significant contribution to that bin's measurement from
sources with actual fluxes outside the bin's formal range
(Malmquist effect).
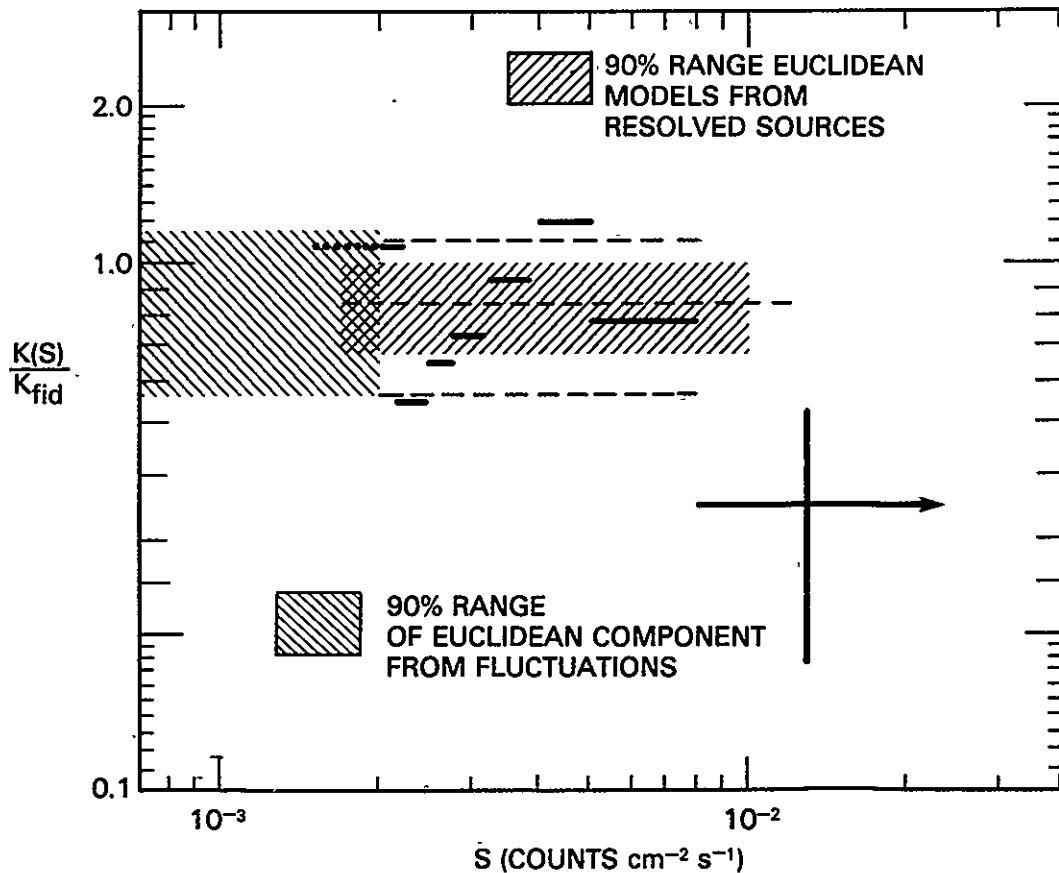
The alert reader will have noticed that the typical values for the source flux uncertainties from Piccinotti et al. as reported in Table VII.1 are small, ~ 0.23. From our discussion in Section III:7 we know that these numbers can not have included any correction for confusion to $\sigma_I$, as the minimum total uncertainty including counting statistics is ~0.3 R15. We can make an approximate correction to $\sigma_I$, however this will reduce the statistical confidence level of the intensity below 5 $\sigma$. Traditional calculations of the Malmquist correction will break down. In Appendix A6 we develop an extension to the technique for calculating $f_i$ that automatically imposes a reasonable low flux cutoff. The impact of including the approximate confusion noise to $\sigma^2_I$ and using our proposed calculation of the Malmquist bias is to reduce $\bar{n}$ to 8.85 (7.26-10.7). This is a 7% reduction below the range obtained from the values in Table VII.1.

In later sections we use the underlined uncorrected range to represent the resolved source counts. If we did use our estimated correction, the general result would be a slight ~7% decrease in the range for K when information from the resolved sources is included. Because K and the excess variance are so strongly coupled, the confidence range for $\sigma^2_S$ would also change. We can estimate that the effect will be to shift the allowed range for $\sigma^2_S$ upward by $\leq 1 \times 10^{-4}$ (counts $s^{-1}$ $cm^{-2})^2$. This correction is not miniscule, but it is less than the magnitude of most of the phenomena we will discuss in applying our limits on the excess variance.

## Combined likelihood estimation of excess variance

We noted when fitting the fluctuations to a Euclidean component and an excess variance that if we were able to obtain outside information about the size of the Euclidean component we would be able to reduce the confidence range for the excess variance. The resolved sources are such a source of information. If we treat the two measurements, the fluctuations and the resolved source counts, as statistically independent determinations of N(S), we can test models by

Analysis of Results                                              Resolved Sources

Figure VII.2: Combinded likelihood confidence region, excess variance versus W (or K)

Based on the combined likelihood of the fluctuations and the resolved sources. (Symbols same as in Figure VI.9)
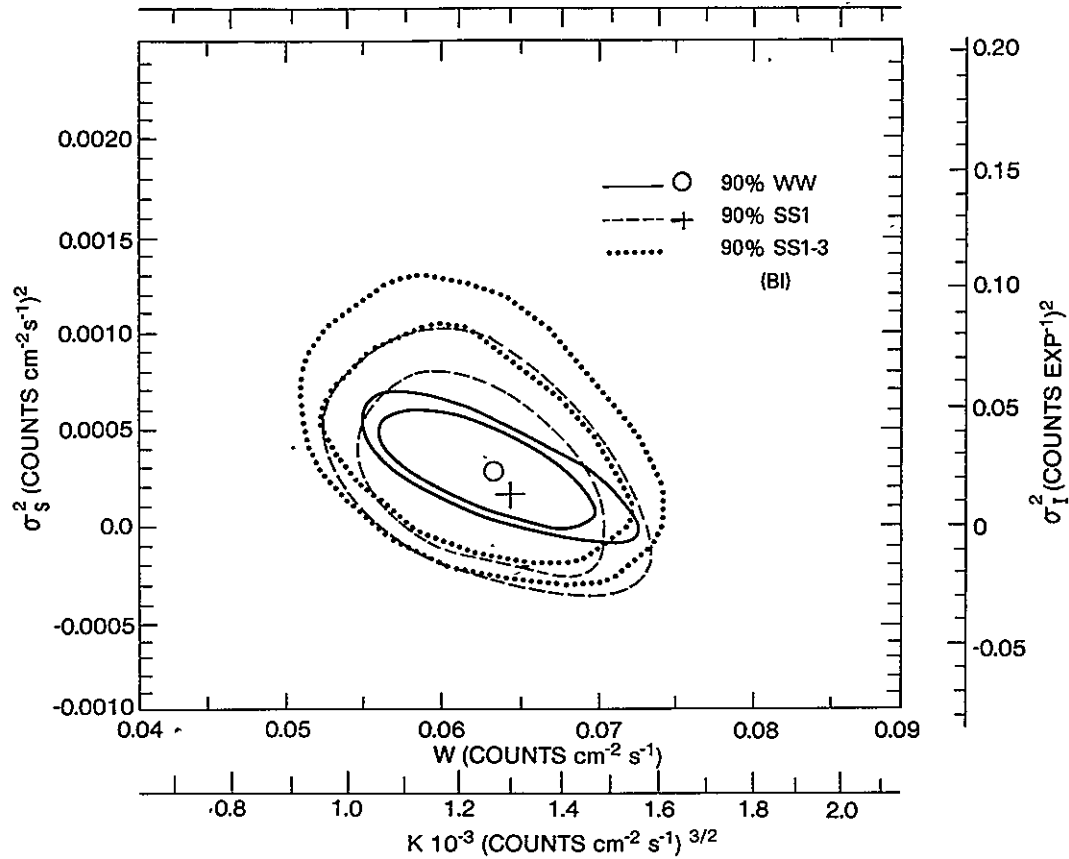


the combined likelihood test:

$$\Lambda = \Lambda_f(K, \sigma^2_S) + \Lambda_{RS}(K) \ .$$

[VII.10]

Confidence regions from the combined likelihood difference are plotted in Figure VII.2. When only one parameter is considered important, the range in K is essentially that found from the resolved sources alone: $1.05 \times 10^{-3}$ to $1.47 \times 10^{-3}$ (H1L1)[1.5] (using the WW set for the fluctuations). A noticeable improvement in the determination of the excess variance has been made (see Table VII.2). As a percentage of the mean sky intensity, the WW set determination ranges from 1.3% to 0.0%,

Analysis of Results                                    Resolved Sources

Table VII.2: 90% confidence range for invariant excess

$$\sigma^2{}_{S;max} \quad \sigma^2{}_{S;min}$$
$$10^{-3} \; (H1L1)^2$$

Using combined likelihood of resolved sources and fluctuations
(Only $\sigma^2{}_S$ considered interesting)

| | $\sigma^2{}_{S;max}$ | $\sigma^2{}_{S;min}$ |
|---|---|---|
| SS1 | 0.80 | −0.25 |
| WW | 0.60 | 0.0 |

From intersection of 95% confidence regions from sources and
fluctuations
(Boole's Inequality)

| | | |
|---|---|---|
| SS1 | 1.4 | −0.63 |
| WW | 0.88 | −0.20 |

now treating K as a nuisance parameter. At the 90% level, then, some excess variance is required. The improvement for the SS1 dataset is even more dramatic. Before including the information from the source counts the excess variance was relatively unconstrained. The acceptable range now is not far from the range previously attained by the WW set: 1.49% to −0.83%. The derived two parameter confidence region for the combined likelihood using the WW set we will refer to in following sections as our "best" region.

### Excess variance range from Boole's Inequality

The validity of the combined likelihood for forming the improved $K-\sigma^2{}_S$ region rests on the statistical independence of the two determinations. To some degree this is not the case. Though the fluctuations are dominated by sources at fluxes at a few times $10^{-4}$ H1L1, some of the observations are contaminated by sources which are much brighter, including possibly some of the resolved sources if they are near the edge of the smeared detector response. The resolved sources, formally limited to fluxes $>2x10^{-3}$, may through the Malmquist

effect include sources whose true intensity may be much lower (one of the sources in the lowest flux bin is expected to have a flux $<1.5 \times 10^{-3}$ H1L1). This was schematically indicated by the dotted continuation of the lowest flux bin's K(S) bar in Figure VII.1 Given that the fluctuations are dominated by sources so much below this, we see that the sources that dominate the two different measurements are still essentially seperated, so that a prudent observer may still conclude that to a large degree the two measurements are independent.

If we want to be certain, we can combine the information from the two fits by using Boole's Inequality, BI (see Figure VII.3). The 90% region is constructed from the intersection of the 95% range in K from the resolved sources ($\Delta\Lambda < 3.84$, giving $7.46 < \overline{n} < 11.90$) and the 95% region from the fluctuations (considering both parameters as significant). As might be expected the region is not as restrictive as the one produced by the combined likelihood. In fact the allowed range for the excess variance from the WW set is _larger_ than was obtained from the fluctuations alone. This is because the BI contour is necessarily one that allows a simultaneous determination of both parameters, and no additional precision is gained by labeling K as a nuisance parameter. Yet the BI technique still allows a delineation of the SS1 confidence region significantly better than what the fluctuations alone could indicate. From a standpoint of "provable" statistics, the preferred region for the excess variance is the SS1 result combined with the resolved sources via Boole's Inequality. This we will refer to as our "conservative" region or result.

To indicate the difference these ranges make for testing models of N(S) we show in Figure VII.4 the allowed range of models that break from a Euclidean index to $\gamma' = 3$ or 2. As was discussed for Figure VI.10, this representation is only illustrative of a particular class of models. The proper test of a particular model requires the explicit calculation of its excess variance. This is discussed in the next section.

Figure VII.3: 90% region for excess variance versus K (or W) for
intersection of resolved source and fluctuations 95% contours (Boole's
Inequality)

The bold regions indicate the intersection of the 95%
confidence regions from the fluctuations and the resolved
sources (RS) to yield at least a 90% region using Boole's
Inequality.
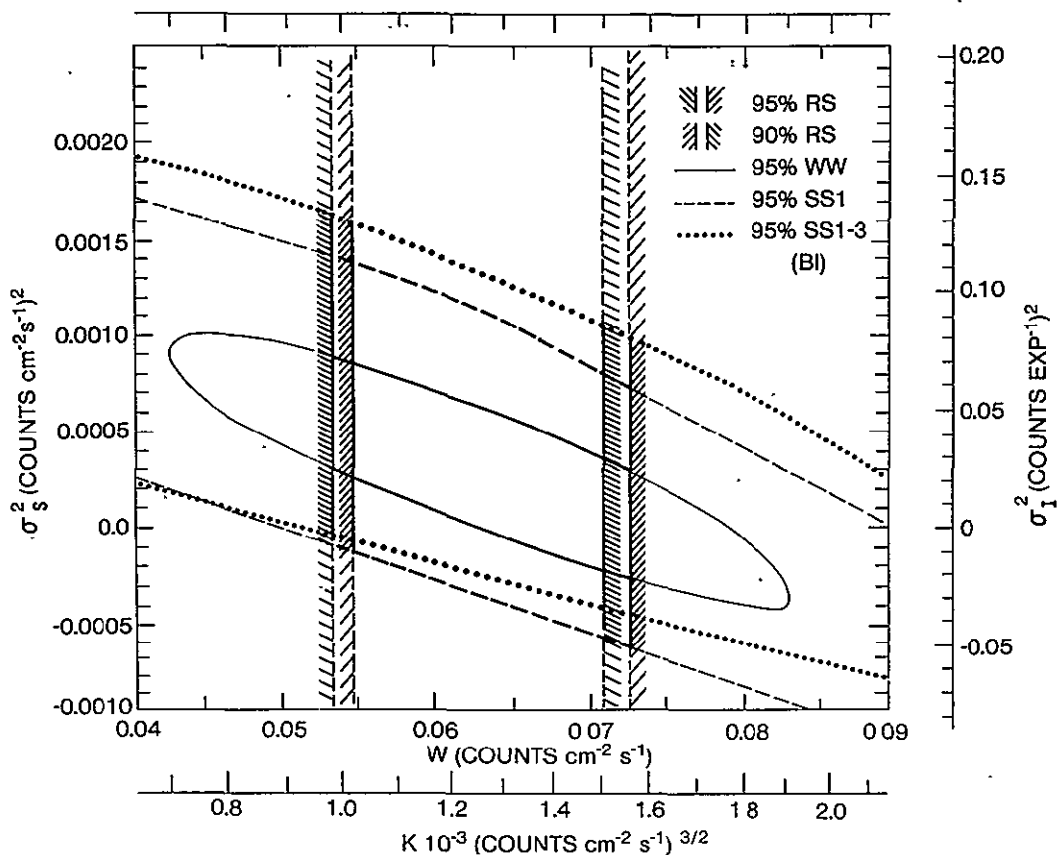
Figure VII.4: $K(S)/K_{fid}$ for broken power law

Outline of allowed models that break to aither 3.0 or 2.0 are constrained by the 90% limits on K and the excess variance from the combination of the fluctuations and the resolved sources. The $\gamma' = 3$ models have positive excess variance, $\sigma^2_S > 0$. The $\gamma' = 2$ models have $\sigma^2_S < 0$. The solid lines use the WW set and the combined likelihood method to include the resolved sources. The dashed outline is for the SS1 set, combined with the resolved sources' confidence region using Boole's Inequality. (See also Figure VI.10.)
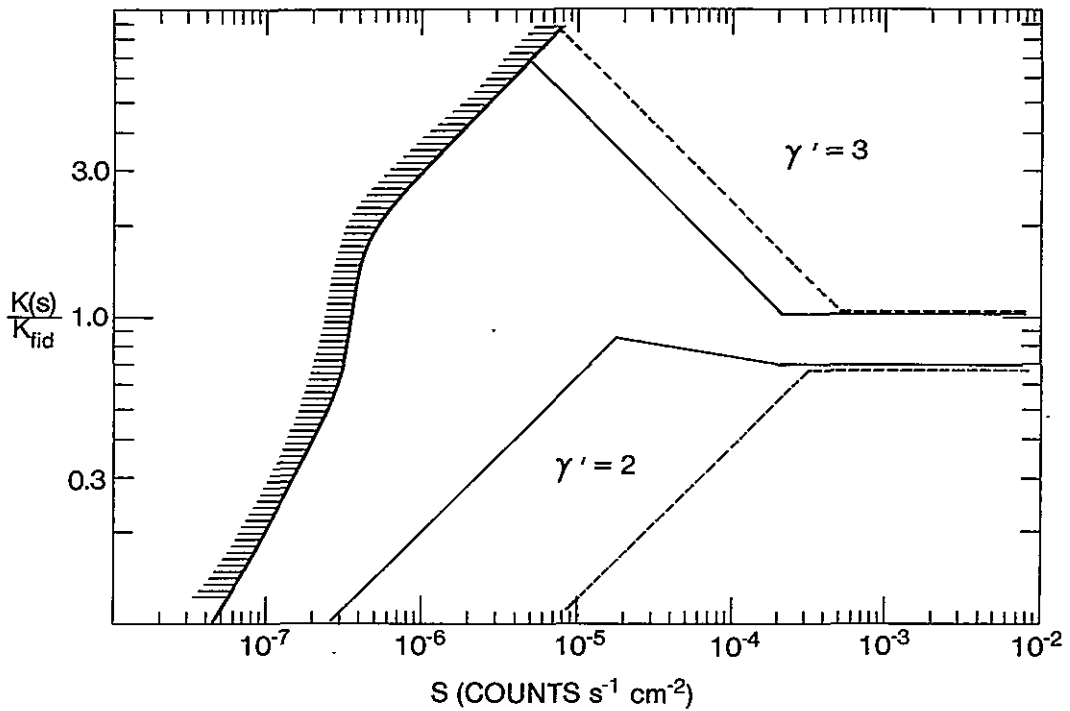


2: The Expected Excess Variance from Non-Evolving Populations

Analysis of the fluctuations has been reduced to applying the determination of two parameters: the strength of the fluctuations (measured by K) for a Euclidean N(S) component

and an excess variance characterizing the deviation from a
strictly Euclidean form. In this section we examine in
detail the populations of high latitude resolved sources as
studied in the complete sample of Piccinotti et al. [1982],
and predict their contribution to both the K values and the
excess variance, assuming no evolution. The total excess
variance for all the populations is $\sim-3.3 \times 10^{-4}$ (counts $s^{-1}$
$cm^{-2})^2$. This negative value is dominated by the
contribution from clusters of galaxies, due to the rapid
effect of the cosmological redshift on their contribution in
the HED 1 Layer 1 band. The value is significantly less
than that allowed by the fluctuations, except for our most
conservative estimate using non-overlapping data. The
fluctuations therfore require an additional source of
variance, perhaps a new population of X-ray sources or
evolution in one of the known populations.

We have fit the fluctuations in order to constrain the behavior of
N(S) at low flux values, and now we want to estimate the expected
behavior at these same fluxes of the observed source populations. Our
principal source of information for these predictions is the complete
sample of objects compiled in Piccinotti et al. [1982], hereafter
referred to as the XCS (X-ray Complete Sample). These are the same
sources that we used to determine the high flux N(S) in the last
section. Now we use their identifications, redshifts, and spectra to
predict their behavior at fluxes below $2 \times 10^{-3}$ counts $s^{-1}$ $cm^{-2}$. The
sample consists of 76 objects: 30 clusters of galaxies; 28 active
galactic nuclei (AGN); 10 moderate luminosity galactic sources; 4 BL Lac
objects; the peculiar active galaxy M82 (confused with the active
nucleus in M81); one QSO (3C 273); and 2 objects still without
identifications. The first three populations dominate the resolved
source counts and would be expected to similarly dominate the
fluctuations, in the absence of evolutionary effects.

Given a luminosity function for a population of sources, we have
shown (Section III:1 and Appendix A1) how to calculate the predicted
N(S). Ideally, to test whether this is consistent with the
fluctuations, one must combine this N(S) with those from all the other
populations and then, from the total N(S), derive the fluctuations
distribution (Section III:3). However, the sensitivity of our data to

the details of N(S) at low fluxes is restricted to measurements of a high flux power law component with a possible additional variance term. We therefore reduce the N(S) of a population to two numbers: equivalent value of K(S) at a high flux, say $S_h$, where the behavior is near Euclidean, and the invariant excess, $\sigma^2_S$ (see equation [III.68]). When combining two or more N(S) curves, the net $K(S_h)$ and $\sigma^2_S$ is the sum of the values for the constituent populations. An additional integral value used to constrain the acceptable behavior of an N(S) curve is the total flux contributed, $S_{as}$ (see equation [III.8]). Armed with the values for these three quantities we can test whether a given collection of N(S) curves for different populations is consistent with the resolved source counts $(K(S_h))$, the fluctuations $(\sigma^2_S)$ and the total sky brightness $(S_{as})$. When including cosmological effects, the value of $q_0$ is most important for determinations of $S_{as}$. The magnitude of all the integral values are independent of $H_0$, as long as the same quantity is used in calculating the N(S) curve and determining the luminosity function.

If the N(S) curves over the flux range of the XCS were identically Euclidean, the exact value of $S_h$ used to define $K(S_h)$ would be unimportant. However to allow for possible non-Euclidean behavior we chose our value with care. As most of the information about $K(S_h)$ comes from the fits to the resolved source counts, the value for $S_h$ that we will use is the expected mean flux of objects in the XCS, roughly $3 \times 10^{-3}$ counts $s^{-1}$ $cm^{-2}$, after including a slight Malmquist correction. We will use $K_h$ for $K(S_h)$.

In this section we consider only non-evolving populations. The invariant excess for such populations will be negative because they roll over at low fluxes from cosmological effects (see Appendix A1). The most extreme negative contribution possible from a population is when it has a sharp cutoff at a flux just below the sensitivity limit of the XCS. For a given value of $K_h$,

$$\sigma^2_S > -4\pi \cdot K_h \, S_h^{1/2} = -0.69 \, K_h \; . \qquad\qquad [VII.11]$$

The 30 clusters of galaxies in the XCS have luminosities between $28 \times 10^{44}$ ergs $s^{-1}$ (Abell 2142) and $0.24 \times 10^{44}$ (Virgo cluster). (All luminosities in this and subsequent sections are defined for the interval from 2 to 10 keV). The luminosity function is adequately described by a power law. The best fit value for the power law index, $\beta$, is 2.03, with the range of acceptable values being roughly 1.85 to 2.21. (The "error" terms presented in Piccinotti et al. are not strictly equivalent to one-sigma errors, or to a particular confidence level, but are still indicative of the range of models that provided a good fit to the data.) The flatness of the luminosity function, $\beta < 2.5$, indicates that the source counts and the Euclidean component of the

| Table VII.3: Luminosity versus cluster temperature used | | |
|---|---|---|
| $L_{min}$ | $L_{max}$ | T |
| $L_{44}$ | $L_{44}$ | keV |
| 0.1 | 2.6 | 3.3 |
| 2.6 | 8.4 | 5.9 |
| 8.4 | 17.5 | 7.9 |
| 17.5 | 30.0 | 9.8 |

fluctuations are dominated by the high luminosity objects. The luminosity function could not extend much above the most luminous observed clusters in the XCS without a break to an equivalent index greater than 2.5. Otherwise we would have expected to see even higher luminosity clusters in the XCS. We set the maximum luminosity to 30 $L_{44}$ (luminosity measured in units of $10^{44}$ ergs $s^{-1}$). The minimum luminosity is set at 0.1 $L_{44}$, the point where the total density of X-ray clusters is equal to the estimated density of Abell clusters of galaxies. When Piccinotti et al. calculated the norm of the cluster luminosity function, $\kappa$, they made no cosmological corrections. For sources with spectra as steep as clusters are in the HlLl band, these can be significant (see Equation [A1.18]). Because of this, when their luminosity function is used in a correct cosmological fashion it

Analysis of Results                                                Expected Excess Variance

predicts significantly fewer clusters than were actually seen in the XCS. For this reason we determined $\kappa$ by requiring that the predicted number of sources in the range from $2\times10$-3 counts $s^{-1}$ cm$^{-2}$ to $8\times10^{-3}$ counts $s^{-1}$ cm$^{-2}$ be equal to the 27 actually observed.

Table VII.4: Contribution of clusters to sky flux, fluctuations, and source counts

| $\kappa$ Mpc$^{-3}$ L$_{44}$$^{\beta-1}$ x10$^{-7}$ | $\beta$ | $q_0$ | $z_{max}$ | $S_{as}$ H1L1 | $K_h$ 10$^{-4}$ H1L1$^{3/2}$ | $\sigma^2{}_S$ 10$^{-4}$ H1L1$^2$ |
|---|---|---|---|---|---|---|
| 3.36 | 2.03 | 0.5 | 3.0 | 1.89 | 5.04 | -2.13 |
| 4.38 | 2.21 | " | " | 2.29 | 5.03 | -1.97 |
| 2.42 | 1.85 | " | " | 1.60 | 5.05 | -2.26 |
| 3.36 | 2.03 | 0.1 | " | 2.10 | 5.02 | -2.05 |
| 4.38 | 2.21 | " | " | 2.52 | 5.02 | -1.89 |
| 2.42 | 1.85 | " | " | 1.80 | 5.04 | -2.18 |
| 3.36 | 2.03 | 0.5 | 2.25 | 1.88 | 5.01 | -2.08 |
| " | " | " | 1.5 | 1.84 | 4.97 | -2.04 |
| " | " | " | 0.75 | 1.64 | 5.00 | -2.09 |

In calculating the form of the number counts from the luminosity function a primary cosmological effect is the redshifting of the source spectrum. We fold the redshifted spectrum through the detector response to get the observed flux in detector units, i.e. counts $s^{-1}$ cm$^{-2}$ in Layer 1 of HED 1. (In the following tables and discussion we abbreviate these units as counts $s^{-1}$ cm$^{-2}$, or H1L1.) A proper understanding of the source spectra is important.

The spectra of clusters of galaxies are well described by thin thermal bremsstrahlung with temperatures in the range 3 to 10 keV [see e.g. Mushotzky et al. 1978]. Furthermore, there is a strong correlation between temperature, $T$, and 2-10 keV luminosity [Mushotzky 1983]. The best fit relationship has the linear form $T = 2.63$ L$_{44}$$^{0.34}$ keV. For

Analysis of Results                                              Expected Excess Variance

simplicity we use the roughly equivalent equation $T = 3.42\ L_{44}^{1/3}$ keV.
We divide the range of the luminosity function into 4 bins, and assign
to each bin a single T, the mean temperature weighted by the cluster
contribution to the Euclidean regime source counts (see Table VII.3).

Figure VII.5: $K(S)/K_{fid}$ for clusters

The solid, long dashed, and short dashed curves are for
different values of the luminosity function index, $\beta$. The
two solid curves are for different values of $q_o$; the upper
one is for $q_o=0.1$, and all other curves in the figure
assume $q_o=0.5$.
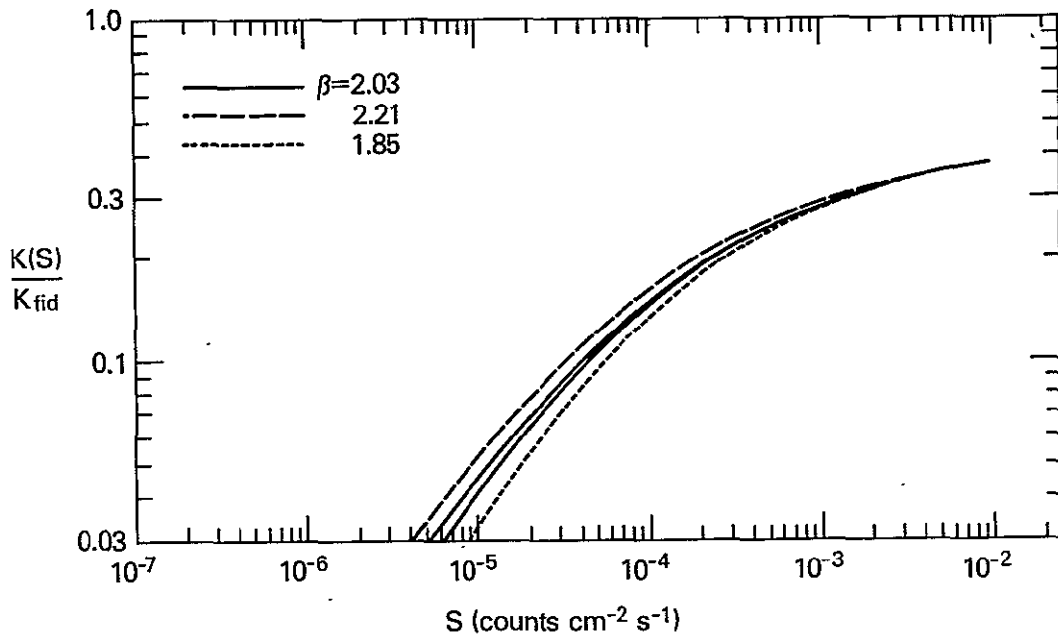


By examining Table VII.4 and Figure VII.5 we see that the cluster
source counts become significantly flatter than Euclidean at moderately
high flux levels, a few $10^{-4}$ counts $s^{-1}$ $cm^{-2}$. This is almost entirely
due to the redshifting of the thermal spectra out of the HED 1 bandpass.
Because of this the invariant excess is in the range $-2.1$ to $-1.9$ $x10^{-4}$.
This is roughly equivalent to a sharp cutoff in the cluster $N(S)$ at a
flux of $10^{-3}$ counts $s^{-1}$ $cm^{-2}$, not much below the resolved source flux
limit or $S_h$. Varying our assumptions about the luminosity function or

Analysis of Results                              Expected Excess Variance

Figure VII.6: $K(S)/K_{fid}$ for clusters with a sharp cutoff at moderate
redshift

$q_o = 0.5$
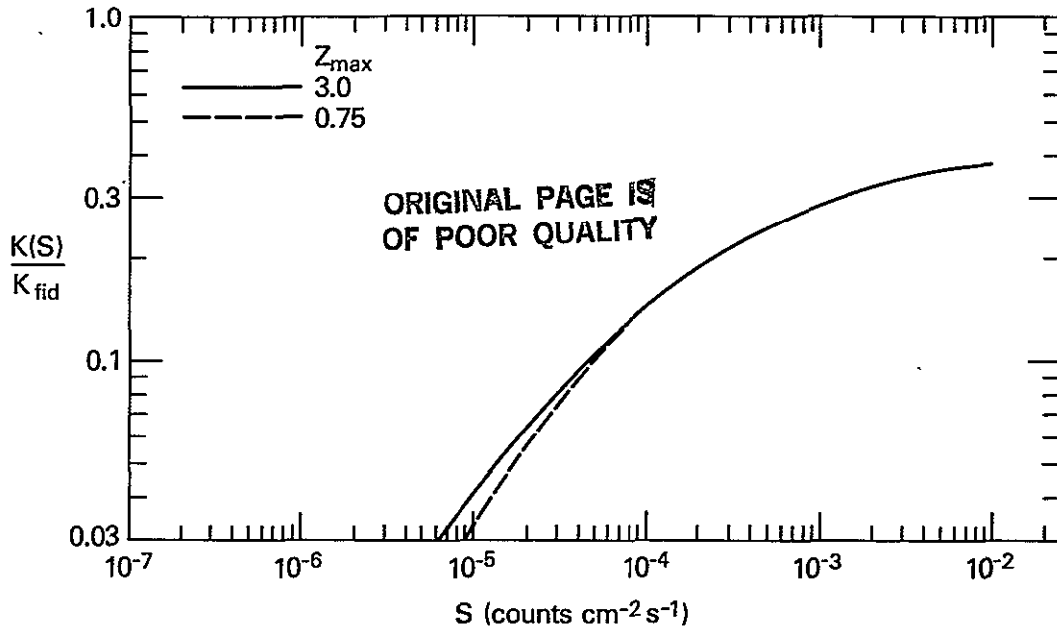


$q_o$ makes only a small difference. Even if we replace the variable temperature dependence outlined in Table VII.3 with a single 6 keV thermal spectrum there is no signficant difference in the excess variance.

The density of clusters is not expected to be uniform as a function of z. General models have been proposed where the temperature and luminosity of clusters is generally lower at earlier epochs [see e.g. Perrenod 1980]. However, Henry, Soltan, and Briel [1982] have examined a survey of high-redshift clusters observed with the Einstein Observatory and saw no indication for any evolution of the luminosity function index out to a redshift of. about 0.5. Most of the negative excess for clusters is generated by behavior at relatively low redshifts. This is seen in Figure VII.6, where the imposition of a sharp cutoff in the density of clusters at a redshift, $z_{max}$, of 0.75 produces

Analysis of Results                                          Expected Excess Variance

no increase in the deficit in $\sigma^2_S$ that is detectable within the precision of our calculations.

All the integral values derived from the luminosity function, i.e. $S_{as}$, $K_h$, and $\sigma^2_S$, are proportional to the luminosity function coefficient, $\kappa$. Therefore the fractional uncertainty in $\kappa$, roughly $N^{-1/2}$ where N is the number of sources observed in the XCS, is the same as the fractional uncertainty in the integral values. Thus the confidence range for $\sigma^2_S$, including the uncertainty in $\kappa$, is $-1.6$ to $-2.5 \times 10^{-4}$ H1L1$^2$.

### Active Galactic Nuclei (AGN)

The XCS includes 29 sources identified with various active galactic nuclei (excluding 3C 273 and the 4 BL Lac sources which we cover later in this section). 21 of these are classified as type 1 Seyferts, while the others are type 2 Seyferts, narrow emission line galaxies, N or other active galaxies. The range of luminosities is $0.0175$ $L_{44}$ (NGC 3227) to 13 $L_{44}$ (III Zw 2). The luminosity function is well fit by a power law with index, $\beta$, $\sim 2.75 \pm 0.15$. Cosmological corrections to the AGN N(S) over the XCS flux range are small. Therefore the coefficient reported by Piccinotti et al. is consistent with the number of observed objects in the XCS, after rescaling for the additional identifications made since the original analysis. $L_{max}$ is set at $15L_{44}$. The steepness of the index, $\beta > 2.5$, signifies that the source counts, and $K_h$, will be dominated by the portion of the luminosity function near $L_{min}$. Because of this, Piccinotti et al. were able to estimate how low in L the luminosity function can extend as a $\beta > 5/2$ power law and still not exceed the number of sources observed in the XCS. If the change in the luminosity function were a sharp cutoff at $L_{min}$ the estimated position would be between $0.0045$ and $0.025$ $L_{44}$, depending on $\beta$ and the degree of disagreement with the XCS. However, as we have remarked, below about $0.01$ $L_{44}$ all sources in the XCS would have redshifts $\lesssim 0.003$ where local inhomogeneities would be important, so that it would be difficult to correctly draw conclusions about the AGN luminosity function at lower luminosities from the XCS sample. In the

tables and figures below we assume that there is a sharp $z_{min}$ of 0.002. No source in the XCS other than M82 has a redshift less than this.

Using the _Einstein_ Observatory, Elvis and Van Speybrock [1982], detected very low luminosity X-ray activity in the nucleus of the nearby galaxy M81. Other than its magnitude, $L\sim0.0001 L_{44}$, the source has the general characteristics of a type 1 Seyfert nucleus. As Elvis and Van Speybrock noted, this single detection indicates that modeling the low luminosity behavior of the luminosity function as a sharp cutoff from the $\beta\approx2.75$ regime is a simplification. Recent preliminary investigations by Elvis, Soltan, and Keel [1983] show that the AGN X-ray luminosity function can be better represented with a break at $L_b \sim 3\times10^{42}$ erg s$^{-1}$ to $\beta'\lesssim2$, and a much sharper break to $\beta''\sim0$ at about $10^{40}$ erg s$^{-1}$. These results, based on indirect arguments from optical observations correlated with X-ray information from the _Einstein_ Obervatory, are difficult to connect accurately with the Piccinotti _et al._ results. However, we will assume that the AGN luminosity function is well represented by a two-index broken power law form, with $\beta = 2.75$ above the break luminosity $L_b$ and either 1.75 or 2.0 below the break.

The spectra of the XCS AGN show remarkable uniformity. Rothschild _et al._ [1983] showed that AGN are well represented by a single power law with energy spectral index, $\alpha \sim 0.68$ with only a small dispersion (see also Mushotzky [1982]). Low luminosity AGN show detectable absorption due to cold material [Mushotzky 1982; Lawrence and Elvis 1982], while high luminosity objects show comparatively little [see _e.g._ Holt _et al._ 1983]. To crudely model this effect we will assume that objects with luminosities greater than 0.2 $L_{44}$ have no intrinsic absorption while those with an _unabsorbed_ luminosity less than 0.2 $L_{44}$ are absorbed by the equivalent of a column of interstellar material containing $3.6\times10^{22}$ atoms cm$^{-2}$ of hydrogen in the line of sight. Most of the absorption is due to higher Z material, particularly oxygen, and was calculated assuming the interstellar abundances of Brown and Gould [1970], with iron adjusted following Fireman [1974]. The luminosity function is parameterized in terms of the unabsorbed 2-10 keV luminosity, rather than the actual observed luminosity after absorption. Note that there

Analysis of Results                                    Expected Excess Variance

is very little spectral information about the very lowest luminosity AGN, i.e. the objects of Elvis, Soltan, and Keel, which are not included in the A-2 XCS. We will assume that they are also characterized by a 0.68 power law with the same absorption as the A-2 sources with $L < 0.2 \, L_{44}$. However, Elvis and Van Speybroek derive a spectrum for M81 that is substantially steeper, $1.0 < \alpha < 1.6$ (90% confidence).

---

Table VII.5: Contribution of AGN to source counts, sky flux, and fluctuations

The form of the AGN luminosity function is a power law with index 2.75 and coefficient $\kappa = 3.267 \times 10^{-7}$ $Mpc^{-3}$ $(L_{44})^{1.75}$ from $L_b$ to $L_{max} = 15 \, L_{44}$, with a break to $\beta'$ between $L_{min}$ and $L_b$. (The first two models have no $\beta'$, instead having a sharp cutoff at $L_{min}$.) The redshift range is from 0.002 to 3.5. The quantity $N_{as}(XCS)$ is the number of sources, all sky, in the range $2 \times 10^{-3} < S < 8 \times 10^{-3}$ HlLl. For comparison, the number actually seen was 41.2, after correcting for sky coverage, with a $2\sigma$ range of about 25 to 61. The models marked by the asterisk (*) are identical except for the value of $q_o$. These values for $L_{min}$, $L_b$, and $\beta'$ consititute the canonical AGN model used for most subsequent discussion.

| $L_{min}$ $L_{44}$ | $L_b$ | $\beta'$ | $N_{as}(XCS)$ | $S_{as}$ HlLl | $K_h$ $10^{-4}$ HlLl$^{3/2}$ | $\sigma^2_S$ $10^{-4}$ HlLl$^2$ |
|---|---|---|---|---|---|---|
| | | | $q_o = 0.5$ | | | |
| 0.035 | — | — | 40.1 | 10.1 | 4.9 | −0.86 |
| 0.015 | — | — | 51.2 | 18.4 | 6.3 | −0.93 |
| 0.001 | 0.25 | 2.0 | 33.5 | 10.5 | 4.1 | −0.59 |
| 0.0003 | 0.10 | " | 42.0 | 17.7 | 5.2 | −0.48 |
| * 0.001 | 0.10 | 1.75 | 39.4 | 13.3 | 4.8 | −0.65 |
| 0.0003 | 0.03 | " | 51.8 | 26.8 | 6.4 | −0.36 |
| | | | $q_o = 0.1$ | | | |
| * 0.001 | 0.10 | 1.75 | 39.4 | 14.7 | 4.8 | −0.65 |

---

In Table VII.5 we calculate the integral quantities for a variety

Analysis of Results                    ,              Expected Excess Variance

Figure VII.7: $K(S)/K_{fid}$ for AGN model luminosity functions, $q_o = 0.5$
(See Table VII.5)
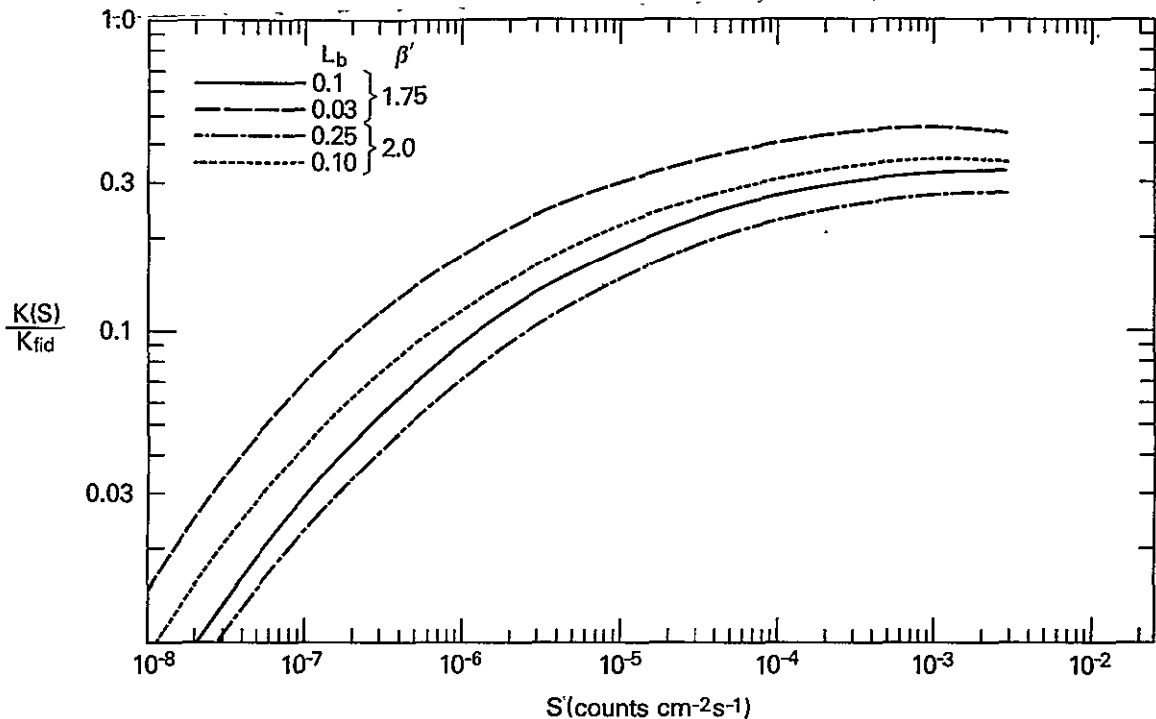


of models, all of which are consistent with the Piccinotti et al. derived luminosity function. The first two models have a sharp cutoff at $L_{min}$ while the latter four models have breaks to a new power law index roughly consistent with preliminary data for the low end of the luminosity function [Elvis, private communication]. The table indicates that the integral quantities take on a wide range of values. Even without evolution the AGN contribute a significant fraction of the total all-sky flux, ranging from 18% to 46% depending on the luminosity function model. This is much more than the ~3% contribution from clusters of galaxies, even though the two populations contribute an equal number of sources to the XCS. This is because the AGN contribution is dominated by the low luminosity objects and is also less affected by the redshifting. Therefore the source counts persist at near Euclidean levels to much lower fluxes (see Figure VII.7). In turn

Analysis of Results                                    Expected Excess Variance

the negative invariant excess is much less for AGN than for clusters.

Most of the imprecision comes from the difficulty in accurately describing the behavior of the luminosity function near the break to the flatter index, at L ~ 0.1 $L_{44}$. A preliminary, non-parametric (i.e. not depending explicitly on any model form) estimate of the contribution of the AGN to the all sky flux, made by Elvis, Soltan and Keel [communicated in advance of publication by Elvis], is ~28% with large uncertainties. Reducing $q_0$ from 0.5 to 0.1 will increase $S_{as}$ by about 10% but will otherwise have no significant effect on the resolved source counts or the invariant excess. We use the starred model from Table VII.5 (hereafter the canonical AGN luminosity function) for more detailed analysis of AGN, particularly in assessing the effects of evolution of the luminosity function. The break to $\beta'=1.75$ below 0.1 $L_{44}$, besides being in agreement with the preliminary curve of Elvis, Soltan, and Keel, is also roughly in agreement with the low end of the optical AGN luminosity function of Huchra [1977; see also Huchra and Sargent 1973]. This canonical model for the AGN luminosity function has a total volume density of $1.88 \times 10^{-3}$ objects per cubic megaparsec. For comparison, we estimate that the total number of galaxies with $M_B < -17.5$ is $7 \times 10^{-3}$ $Mpc^{-3}$ [Felten 1977; Kirshner, Oemler, and Schechter 1979]. (The choice of the magnitude limit is based on the observation by Balick and Heckman [1982] that no Seyfert type activity has been seen in a galaxy less luminous than $M_B$ ~ -18. (See Bothun et al. [1982] for an example of AGN behavior in a low luminosity host galaxy.) Thus with this model a significant fraction of those objects that could contain nuclear activity is included in the X-ray AGN luminosity function.

The distinction between quasars and other AGN appears to be one of degree (i.e. luminosity) rather than of kind, at least in terms of optical spectral characteristics. Woltjer and Setti [1982] make the observation that quasars and Seyfert 1 galaxies differ in no important way except in absolute magnitude. The selection is due principally to the difficulty of detecting a galaxy surrounding a luminous nucleus. They arbitrarily use Fairall 9 as the demarcation between "fuzzy" quasars and Seyferts. Using this criterion, the AGN in the XCS divide

into 21 low luminosity objects and 5 quasars (including Fairall 9 and 3C 273). Four of the latest identifications could not be classified since we have not yet found their redshifts. In analyzing the impact of the AGN on the source counts, fluctuations, and all sky flux, we set $L_{max}$ to 15 $L_{44}$. Only 3C 273 is above this limit. If there were no evolution of the luminosity function, the high value of the index would indicate that any very luminous objects above $L_{max}$ could make only a small contribution to the integral quantities compared to the contribution of the lower luminosity objects near the break, $L_b$. However, in some evolutionary scenarios this region of the luminosity function could become dominant. The details and ramifications of this are discussed in the next section.

BL Lacertae objects are another class of extraglactic X-ray sources probably associated with AGN. BL Lacs are strong X-ray emitters that have been the target of several detailed studies [Margon et al. 1976; Mushotzky et al. 1978b; Ku 1980; Maccagni and Tarenghi 1981; Schwartz and Ku 1983]. However several difficulties preclude a determination of a predicted N(S) curve for them to the same accuracy as for clusters and other AGN. The fact that only four objects were observed in the XCS makes the determination of the norm of their luminosity function difficult, let alone its slope. (Cf. Schwartz and Ku [1983] for an attempt in the bandpass of the Einstein Observatory.) In addition, the spectra are exceptional in their deviation from the ~0.68 index power law spectrum typical of most other AGN [Mushotzky et al. 1978b; Riegler, Agrawal, and Mushotzky 1979; Worrall et al. 1981; Urry and Mushotzky 1982; Urry et al. 1982]. From the fact that four BL Lacs are in the XCS we can estimate $K_h$ to be $0.7 \times 10^{-4}$ HlLl$^{1.5}$. Using equation [VII.11], we can set a minimum to the invariant excess of $\sigma^2_s \gtrsim -0.48 \times 10^{-4}$ HlLl$^2$. However, the BL Lac counts are expected to extend significantly below the flux limit of the XCS, and have been detected at much lower fluxes in the Einstein studies [see e.g. Schwartz and Ku 1983]. If the form of the N(S) curve for BL Lacs approximates that of other AGN, then $\sigma^2_s \sim -0.09 \times 10^{-4}$. HlLl$^2$, while if, because the BL Lac source counts are dominated by the more luminous objects or because steep spectra are affected by redshifts, the N(S) shape is more like

Analysis of Results                                    Expected Excess Variance

that of clusters, $\sigma^2_S \sim -0.3 \times 10^{-4}$ HlLl$^2$. Their possible contribution to
the total $S_{as}$ has a similarly wide range but in the absence of strong
evolution it is not more than about a percent or two. BL Lac source
counts in the optical [Woltjer and Setti 1982] and the X-ray [Maccacaro
and Gioia 1983] are consistent with little or no evolution.

### Galactic Sources

Of the 14 sources in the XCS identified with galactic sources,
four are of high luminosity, $L > 10^{34}$ ergs s$^{-1}$, two X-ray binaries and
two sources in globular clusters. We are reasonably sure that we have
counted all such objects above the galactic latitude $b = 20^o$, and thus
we do not further consider their impact on the fluctuations. The
remaining ten consist of moderate luminosity objects, $L \sim 1$–100 $L_{30}$.
($L_{30} = 10^{30}$ ergs s$^{-1}$.) These have been identified with 3 RS Canus
Venaticorum (RS CVn) stars and 6 cataclysmic variable (CV) stars
[Worrall and Marshall 1983] (see Appendix A5). The tenth object,
H2215-08, has no clear classification[*]. RS CVn stars have been
identified as neutron stars or white dwarfs associated with G giants
[see e.g. Swank, White, Holt and Becker 1981]. The typical scale height
for such objects is $h \sim 110$ pc. CVs, on the other hand, have a scale
height of $\sim$300 pc, indicative of their association with dwarf K stars.
X-ray characteristics of CVs have been reviewed by Cordova and Mason
[1982].

As we saw in Figure III.2, the N(S) curve for a population of
galactic sources follows a Euclidean form at high fluxes but then rolls
over for fluxes corresponding to distances substantially out of the
plane. The critical flux is a function of L, h, and the direction of
observation measured by the sine of the galactic latitude (see
[III.16]). For a CV with a luminosity of 10 $L_{30}$, this point occurs at
$1.36 \times 10^{-4}$ HlLl if one looks perpendicular to the galactic plane, or at
$0.25 \times 10^{-4}$ HlLl for the lower latitude of 25$^o$. RS CVns have a lower

---

[*]Subsequent optical work [Patterson and Steiner 1983] has identified
this object with an uncatalogued cataclysmic variable. This
identification increases our derived N(S) curves for CVs in this section
by 8%, not important for our conclusions.

Analysis of Results                                        Expected Excess Variance

scale height, increasing the critical flux point, but this is partially
compensated by their slightly lower luminosities. To calculate the
integral values $K_h$ and $\sigma^2{}_s$, we average N(S) over the look directions for
the fluctuations data, i.e. $23^\circ < |b| < 90^\circ$.

| Table VII.6: Model luminosity functions for galactic sources | | | | | | |
|---|---|---|---|---|---|---|
| h | β | κ $10^{-5}$ | $L_{min}$ | $L_{max}$ | $K_h$ $10^{-4}$ | $\sigma^2{}_s$ $10^{-4}$ |
| pc | | $L_{30}{}^{\beta-1}$ pc$^{-3}$ | | $L_{30}$ | H1L1$^{3/2}$ | H1L1$^2$ |
| | | | | RS CVn | | |
| 110 | 2 | 1.6 | 2 | 100 | 0.7 | −0.39 |
| " | 3 | 18.5 | " | " | 0.7 | −0.30 |
| | | | | CV | | |
| 300 | 2 | 1.2 | 2 | 300 | 1.2 | −0.52 |
| " | 3 | 25.1 | " | " | 1.2 | −0.31 |
| | | Worrall and Marshall combined luminosity function | | | | |
| 250 | 2 | 1.75 | 2 | 440 | 1.8 | −0.94 |
| 125 | " | 2.43 | " | " | 1.8 | −1.16 |

The accurate determination of a luminosity function for these
sources in the HEAO 1 A-2 bandpass is limited by the relatively small
numbers of sources. However, we can estimate the general properties of
the N(S) curve and the invariant excess for galactic sources by
examining several models consistent with what we now know. Some such
models are tabulated in VII.6. We will assume the luminosity function
is well represented by a power law of index 2 or 3. The luminosity
limits are roughly indicated by the range observed, with CVs extending
to higher luminosities than are observed for RS CVns. The luminosity
function is normalized to give the number of observed objects in the
XCS, assuming that one of the two remaining unidentified objects is a
galactic source. In Figure VII.8 we see that the CVs dominate the
source counts, and by virtue of their higher scale height will continue
their near-Euclidean behavior to lower fluxes. If β is 3, as we noted

Figure VII.8: K(S)/$K_{fid}$ for CV and RS CVN

See Table VII.6
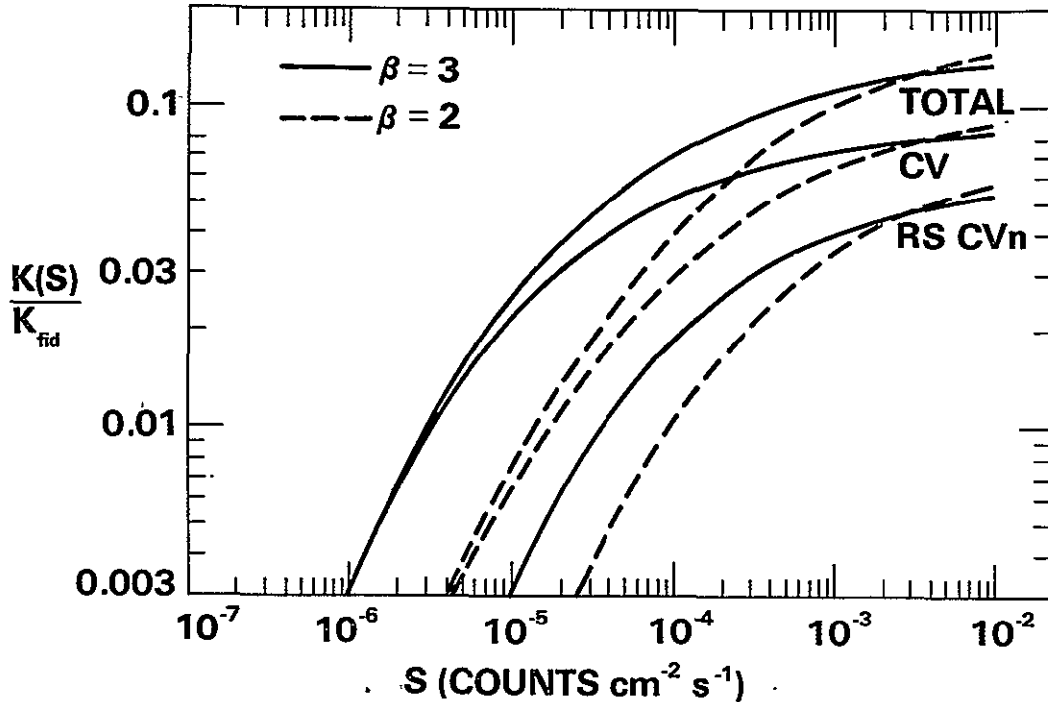


in discussing Figure III.2, the number of low flux sources is enhanced, reducing the magnitude of the negative invariant excess.

Worrall and Marshall [1983] estimated the combined volume emissivity of RS CVns and CVs, in order to estimate their contribution to a small scale height disk of emission associated with the galactic plane [see e.g. Worrall, Marshall, Boldt, and Swank 1982]. Their estimate was consistent with the h=250 pc model labeled WM in Figure VII.8 and VII.6. They assumed that β was 2 and that a single luminosity function described both classes of objects. The coefficient they derived predicts the correct number of sources in the XCS for the 250 pc scale height model, but must be significantly corrected for the 125 pc scale height. For all the models for the galactic sources the magnitude of $\sigma^2_S$ is on the same order or larger than the magnitude derived from the AGN models, although the galactic contribution to the XCS counts is

Figure VII.9: $K(S)/K_{fid}$ for total galactic sources

WM indicates the models of Worrall and Marshall. See Table
VII.6 for details.
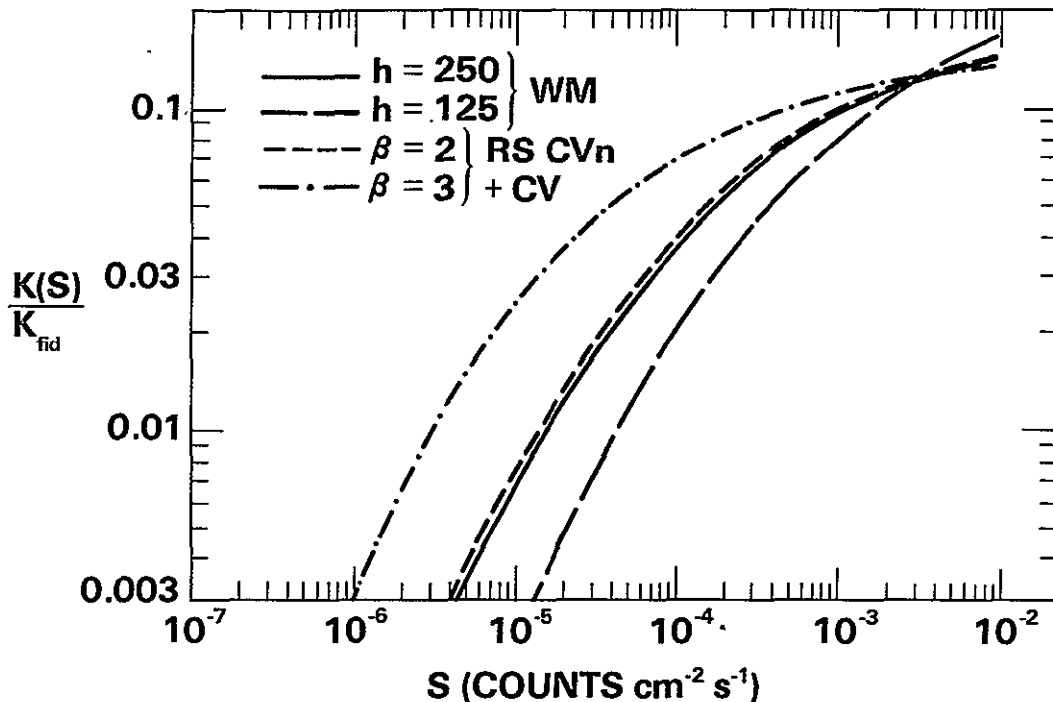


much less.

<u>Ordinary galaxies and other sources</u>

At some level ordinary galaxies will contribute to the X-ray sky
flux. That such objects are not well represented in the XCS is
understood from their low X-ray luminosity, $L \lesssim 10^{39}$ ergs $s^{-1}$ (2-10
keV). Actually M31 would have been included in the XCS save that it was
explicitly excluded by Piccinotti <u>et al.</u> along with the Magellanic
Clouds for being too local. Worrall, Marshall, and Boldt [1979] made a
survey of nearby galaxies with the A-2 experiment and estimated that the
total brightness of such objects was $\lesssim 1\%$ of the total sky flux. Setti
and Woltjer [1982] argued, based on data from Long and Van Speybroeck
[1983; see also Long <u>et al.</u> 1982], that the average galaxy luminosity

Analysis of Results                                    Expected Excess Variance

was substantially larger than what Worrall, Marshall, and Boldt estimated, so that as much as 9% of the sky flux could be due to normal galaxies. However, some of these more luminous "normal" galaxies may be already included in our analysis, through the luminosity function of Elvis, Soltan, and Keel [1983], as very low luminosity AGN. Without more detailed observations in the same spectral bandpass as the A-2 experiment it is difficult to estimate from the Einstein observations the galaxy contribution to the invariant excess but it is expected to be small, smaller than that of the AGN, as always assuming no evolution.

The correlation study of Turner and Geller [1980] places a significant upper bound on the possible contirbution of normal galaxies, or current epoch objects that cluster with galaxies, to the total sky flux. By comparing the UHURU surface brightness fluctuations with variations in the number of galaxies in different complete collections of galaxies, Turner and Geller were able to state that such objects could be responsible for $\lesssim 25\%$ of the X-ray sky flux. (However, see our discussion at the conclusion of Section III:7 about the impact of the fluctuations on their analysis.) A similar study using the HEAO 1 A-2 database may prove useful, not only to place an upper bound on the contribution of normal galaxies, but also to place a limit on the current epoch volume emissivity of the low luminosity AGN.

---

## Comparison with fluctuations results

In the last section we found that the 90% confidence region for the $\sigma^2_S$ was always greater than zero, considering the excess as the single interesting parameter and using the fluctuations data from the overlapping WW set combined in an independent fashion with the source counts. Given that any non-evolving population of sources is expected to have a negative invariant excess, then this alone argues for the existence of an additional source of variation, either evolution of the source populations, or non-Poisson distribution of the sources.

This conclusion is strengthened by the results of this section, showing that for the luminosity functions derived from the XCS the

Figure VII.10: $K(S)/K_{fid}$ total source populations compared with fluctuations

See text for details, $q_o$ =0.5 for clusters and AGN.
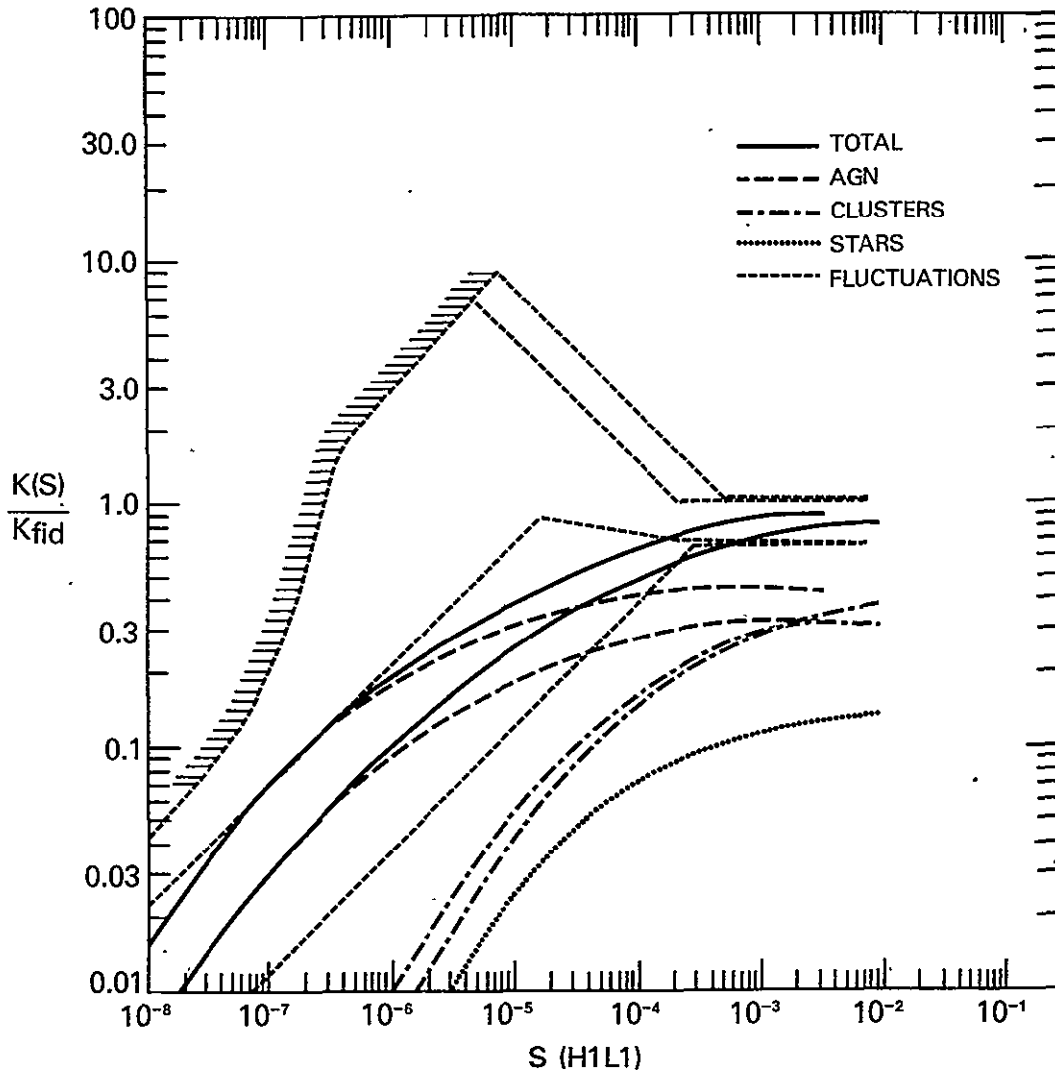


expected value for $\sigma^2{}_S$ is in the range $-3.7 \times 10^{-4}$ to $-2.9 \times 10^{-4}$ $H1L1^2$, depending on the exact forms chosen for the luminosity functions. In Figure VII.10 we plot the total N(S) curves for the galactic sources, assuming $\beta = 3.0$, clusters with $\beta = 2.03$, and the canonical AGN model (indicated by the asterisk in Table VII.5). To indicate the effect of using the more extreme acceptable models for N(S) the long dash curves

Analysis of Results                                    Expected Excess Variance

Figure VII.11: $K_h$ and $\sigma^2_s$ for total source populations

The lower box is the region of expected values for the non-evolving populations based on the XCS derived luminosity functions. The points U and L correspond to the upper and lower total N(S) curves in Figure VII.10 The solid and dashed contours are the best and conservative 90% regions based on the fluctuations and the source counts. The confidence regions assume that both parameters are interesting.
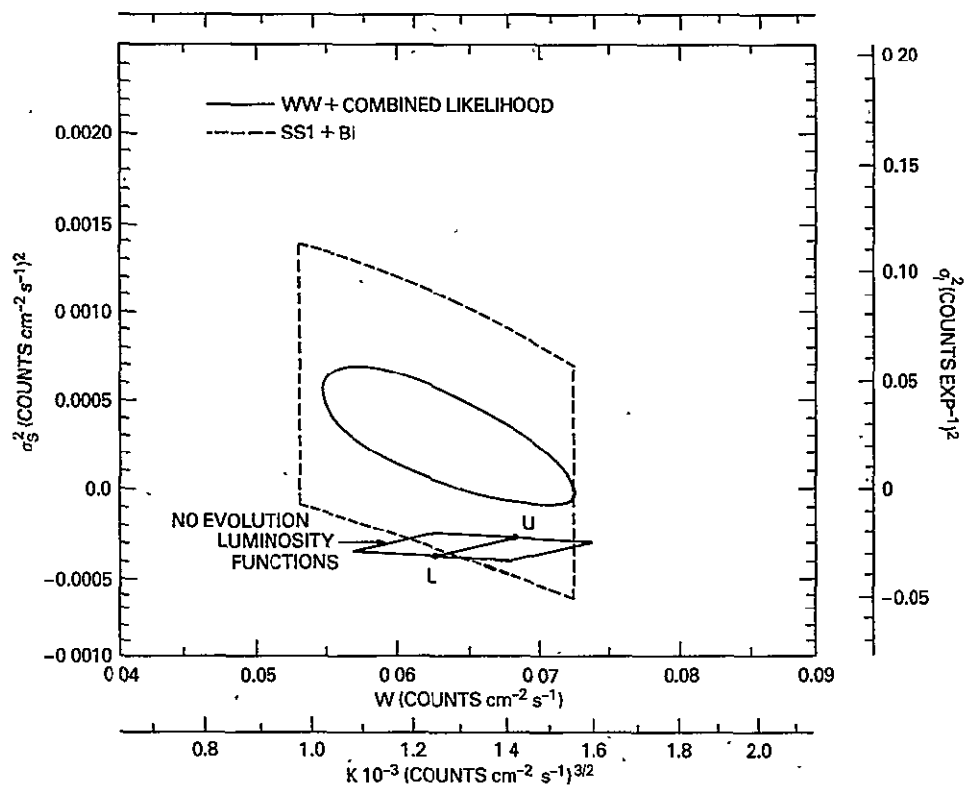


indicate the effect of modifying the cluster model by using the $\beta = 2.21$ form and the AGN model by moving $L_h$ to $0.03\ L_{44}$. This combination leads to the smaller absolute value for the invariant excess presented above, $-2.9 \times 10^{-4}$.

We can compare these model N(S) curves with the behavior allowed by the fluctuations using the broken power law representations for N(S) presented in Figure VII.4 and illustrated by the short dashed outlines

Analysis of Results                              Expected Excess Variance

in Figure VII.10. The inner outline is from our best confidence region, while the outer is that derived from our conservative result using only independent data for the fluctuations combined with the source counts information using Boole's Inequality. Though the models seem acceptable to the conservative region, they are below the allowed range for the best results, illustrating the too negative $\sigma^2_S$ of·the models. In figure VII.10 we did not include any possible contribution from BL Lac objects or other populations.

Of course the N(S) plots do not properly show whether a particular model is consistent with the fluctuations and source counts. For that we plot the total model values of $K_h$ and $\sigma^2_S$ and compare them directly to the confidence regions for those parameters. This is what is shown in Figure VII.11, now including BL Lacs, but only in the determination of $K_h$. Allowing for the coupling between K and $\sigma^2_S$ in the confidence region strengthens the disagreement between the models and the fit results.

In estimating the significance of this disagreement we reconsider our estimates of the internal background variance. Replacing our best fit estimate of $\sigma^2_B$ of 0.001 (counts $\exp^{-1})^2$ with the 90% upper limit of 0.003 (see Figure IV.4) corresponds to subtracting $0.5 \times 10^{-4}$ $H1L1^2$ from the invariant excess derived from the fluctuations. Based on our best ·confidence region, a significant source of excess variance must be provided, either by evolution of the known sources, the introduction of a new population, or non-Poisson structure of the X-ray sky.

———————————————

### 3: Sources of Excess Variance: New and Evolving Populations

New or evolving populations of sources may provide the amount of invariant excess required by the fluctuations analysis, compensating for the deficit calculated from the non-evolving luminosity functions of the known sources. The upper limit on the invariant excess places constraints on any new population that could provide the bulk of the sky flux. There must be at least 50 such new sources per square

degree, with an average flux of no more than $\sim 3 \times 10^{-13}$ ergs
$s^{-1}$ $cm^{-2}$. Any evolution of the AGN luminosity function
would also affect the AGN contribution to the excess and the
sky flux. The requirement that the observed sky flux not be
exceeded places the strongest constraints on the amount of
such evolution. For pure luminosity evolution, it is
difficult to provide the missing invariant excess without
overproducing the sky flux, unless moderate luminosity
objects ($\gtrsim 0.1$ $L_{44}$) undergo more-evolution than the lower
luminosity sources. Models of luminosity-dependent density
evolution (index evolution) are also considered, and are
consistent with the required behavior. The possible
presence and form of a high luminosity extension of the AGN
luminosity function can affect the high flux N(S). While
observations of the resolved sources cannot constrain a non-
evolving extension of the luminosity function, they do rule
out strong luminosity evolution by these high luminosity
sources.

In the last section, we were able to give a reasonably complete
description of the objects that make up the X-ray Complete Sample of the
A-2 sky survey. We showed how the derived luminsity functions were
consistent with the source counts and the high-flux Euclidean N(S)
component fit from the fluctuations. However, these same luminosity
functions show that without evolution the sources that make up the XCS
could account for only $\sim 25\%$ of the total all-sky flux, $S_{as;tot} = 58$
counts $s^{-1}$ $cm^{-2}$ (HlL1). Admittedly, this is a fairly imprecise
estimate, particularly in the contribution of the very low luminosity
AGN, but it leaves the dominant fraction of the total sky flux to be
explained by a new population of sources, the evolution of a known
population, or some truly diffuse process not associated with point
sources. The first two possibilities, in addition to providing the bulk
of total sky flux, would make an additional contribution to the
invariant excess, perhaps sufficient to counter the negative
contribution from the known sources and thereby restore agreement with
the positive excess derived from the fluctuations.

Such new sources of X-ray emission are doubly constrained: not to
exceed the total $S_{as}$ and not to exceed the upper bound on the invariant
excess. The latter limit, for the new sources, is $\sigma^2{}_S \lesssim 9 \times 10^{-4}$ (HlL1)$^2$

after correcting the total upper bound of $6 \times 10^{-4}$ $(H1L1)^2$ by the estimated negative contribution of the unevolved populations of $\sim -3 \times 10^{-4}$ $(H1L1)^2$. Here, and throughout this section, we use the best 90% confidence region derived from the WW dataset to define constraints on the invariant excess. Use of one of the more conservative regions reduces the statistical significance of the conclusions we draw.

### General constraints on a new population

The upper limit on the excess variance can place a general constraint on the properties of any new population that makes a significant contribution to the total sky flux. Let $N'(S)$ be the number versus flux for such a new population. An evolving source population may also be constrained by letting $N'(S)$ be the difference between the evolved and unevolved number versus flux distributions. If $N^*$ is the total number of sources seen in the sky, then their contribution to the all-sky flux is

$$N^* \; \overline{s} = S_{as}^* = f \; S_{as;tot} \; , \qquad\qquad [VII.12]$$

where

$$\overline{s} = (N^*)^{-1} \int \; dS \; N'(S) \qquad\qquad [VII.13]$$

is the mean flux of the population and $f$ is the fractional contribution to the total all-sky flux. The invariant excess of such a population is calculated to be

$$\sigma^2{}_S^* = \int \; dS \; S^2 \; N'(S) \geqslant N^* \; \overline{s}^2 \; . \qquad\qquad [VII.14]$$

The equality holds when all the sources are of equal flux, i.e. $N'(S)$ is a delta function at $\overline{s}$. We can now solve for a lower limit to the number of sources

$$N^* \geqslant (f \; S_{as;tot})^2 \; / \; \sigma^2{}_S \approx 2.1 \times 10^6 \; (f/0.75) \; . \qquad\qquad [VII.15]$$

This is a minimum density of 51 sources per square degree, given that the sources provide the remaining 75% of the sky flux. In comparison, the extrapolated number of QSOs per square degree determined by Koo and Kron [1982] brighter than J magnitude of 22 is ~150, based on color-color identification techniques on 0.3 square degree field. There are

further counts at dimmer magnitudes, ~350 per square degree for J<23, with large uncertainties. These values suggest a convergence in the quasar contribution to the <u>optical</u> sky flux, $N(S) \propto S^{-2}$, in this magnitude range. The minimum number of sources required by the excess variance limit is well within the QSO counts. We can also place an upper limit on the mean flux of the sources:

$$\bar{s} < \sigma^2{}_S \;/\; (f\; S_{as;tot}) \approx 2x10^{-5}\; (0.75/f)\; HlLl\; .$$  [VII.16]

If the sources have a spectrum typical of the diffuse background this corresponds to $\sim 3x10^{-13}$ ergs $s^{-1}$ $cm^{-2}$ (2-10 keV). Thus if all the sources have the same flux, they could be as bright as one hundredth of the lower flux limit of the XCS.

These general bounds on $N^*$ and $\bar{s}$ can be tightened if we know more of $N'(S)$. The proper technique, if the sources are <u>not</u> all of the same flux, is to explicitly perform the integral in equations [VII.13] and [VII.14] for a particular model $N'(S)$. $N^*$ is then scaled by $<s^2> \;/\; \bar{s}^2$, where

$$<s^2> = N^{*-1} \int dS\; S^2\; N'(S)$$  [VII.17]

is the population's mean value of the flux squared, which is always greater than $\bar{s}^2$. As an example, assume that the new population has a range of fluxes, covering one decade in S, and that they are distributed roughly $\propto S^{-5/2}$, so that $<s^2> \;/\; \bar{s}^2 = \sim 1.5$. This increases the number of sources in equation [VII.15] by a factor of $\sim 1.5$, and correspondingly decreases the mean flux by $\sim 2/3$.

### Evolution of AGN

As has been noted in discussions of the X-ray sky flux, there is no need to postulate the existence of a "new" population of sources. Measurements of the distribution of quasars, both in redshift and in N(S), have long indicated that their distribution is non-uniform [e.g. Schmidt 1968; Braccesi, Zitelli, Bonoli, and Formiggini 1980], and that these same, or similar, objects are copious producers of X-rays [e.g. Tananbaum et al. 1979; Zamorani et al. 1981]. If such behavior were typical of AGN of all luminosities, then it would be easy to produce the

total sky flux [Avni 1978]. However, to calculate the total
contribution to the sky flux as well as to the invariant excess, the
details of the evolution are important.

When considering evolution models in this section, we will assume
that the unevolved luminosity function is given by the canonical form of
last section, i.e. $\kappa = 3.26 \times 10^{-7}$, $L_{min} = 3 \times 10^{40}$, $L_{max} = 15 \times 10^{44}$, and $\beta = 2.75$
with a break to 1.75 for luminosities less than $0.1 \times 10^{44}$. We will refer
to sources whose unevolved luminosities are between the break luminosity
and 15 $L_{44}$ as in the "medium" range, while sources below the break are
"low" luminosity sources. We will investigate possible extensions of
the luminosity function to higher luminosities, particularly the range
from 15 to 100 $L_{44}$. These we will refer to as "high" range sources.
The XCS has no direct information on such sources so our discussion will
be of the nature of illustrating the effects of possible extensions.
The two cases considered in any detail are simple extensions of the
luminosity function, in one case with the same index, $\beta_h = 2.75$, and in
the other with a break in the index to $\beta_h = 3.75$ at the transition
luminosity, 15 $L_{44}$.

Remember that all luminosities are the inferred 2-10 keV
integrated emission, corrected for absorption. AGN with $L > 0.2$ $L_{44}$ are
essentially unabsorbed. Those with $L < 0.2$ $L_{44}$ have absorption from a
column of matter associated with $3.6 \times 10^{22}$ atoms of hydrogen $cm^{-2}$ (see
the discussion in the last section). One of the evolution models we
consider below, luminosity evolution, is equivalent to the gradual
increase of the source luminosities. In these models we assume that the
presence or absence of absorption is directly tied to the physical value
of the luminosity, so that sources with emitter-frame luminosities
greater than the critical luminosity, $0.2$ $L_{44}$, are unabsorbed.

A reminder about AGN spectra: as before, we continue to assume
that AGN are well described by a single power law spectrum with energy
index $\alpha \approx 0.68$. Some of the evolutionary models developed below match, or
even exceed, the total all-sky flux. As we and others have emphasized,
this necessitates spectral evolution of the sources as well as simple

evolution of their number and/or 2-10 keV luminosity. We do not make an explicit provision for such spectral evolution; for general conclusions about possible forms of evolution our current, admittedly unsophisticated, analysis should be sufficient.

## Parametric models for evolution

The evolution of AGN is expressed as a change in the luminosity function with epoch or redshift. Such change is normally discussed in terms of a few simple parameterizations, usually ad hoc modifications in the luminosity function, not directly related to specific causal physical processes (cf. Stocke and Perrenod [1981] or Meier et al. [1979]). In Figure VII.12 we have a schematic presentation of three such parameterized models. The earliest models were simple density evolution, where the normalization of the luminosity function is changed:

$$\phi(L;z) = f_\rho(z) \, \phi_o(L) \, . \qquad\qquad [VII.18]$$

$\phi_o$ is the current epoch luminosity function. An alternative is pure luminosity evolution, where the number of objects is conserved, but the function is shifted along the luminosity axis. This could be due to either a change in the luminosity of each object or an ensemble change of the mean luminosity:

$$L(z) = f_L(z) \, L(z=0) \qquad\qquad [VII.19.a]$$

so that

$$\phi(L;z) = (1/f_L(z)) \, \phi_o( \, L \, / \, f_L(z) \, ) \, . \qquad\qquad [VII.19.b]$$

The additional factor of $1/f_L$ corrects the differential dL. In figure VII.12 we plotted luminosity evolution for the number of objects per decade in luminosity, $L\phi(L)$, which compensates for this additional factor. Both pure luminosity and pure density evolution will have an identical effect on the evolution of the volume emissivity, and therefore equal contribution to the sky flux (see equation [III.6] et seq.). However, as we shall see, the N(S) curve will be quite different for the two models, with a more peaked N(S) at middle fluxes and a

Figure VII.12: Schematic examples of luminosity function evolution

The solid curve represents the current epoch, i.e. an unevolved, luminosity function. The short dashed, long dashed, and dashed-dot lines are examples of pure density, pure luminosity, and index evolution respectively. $L_0$ is the pivot point for index evolution. The ordinate is the logarithm of $L^X\phi(L)$, which is the number of objects per decade of luminosity. This simplifies the presentation of pure luminosity evolution. See text for details.
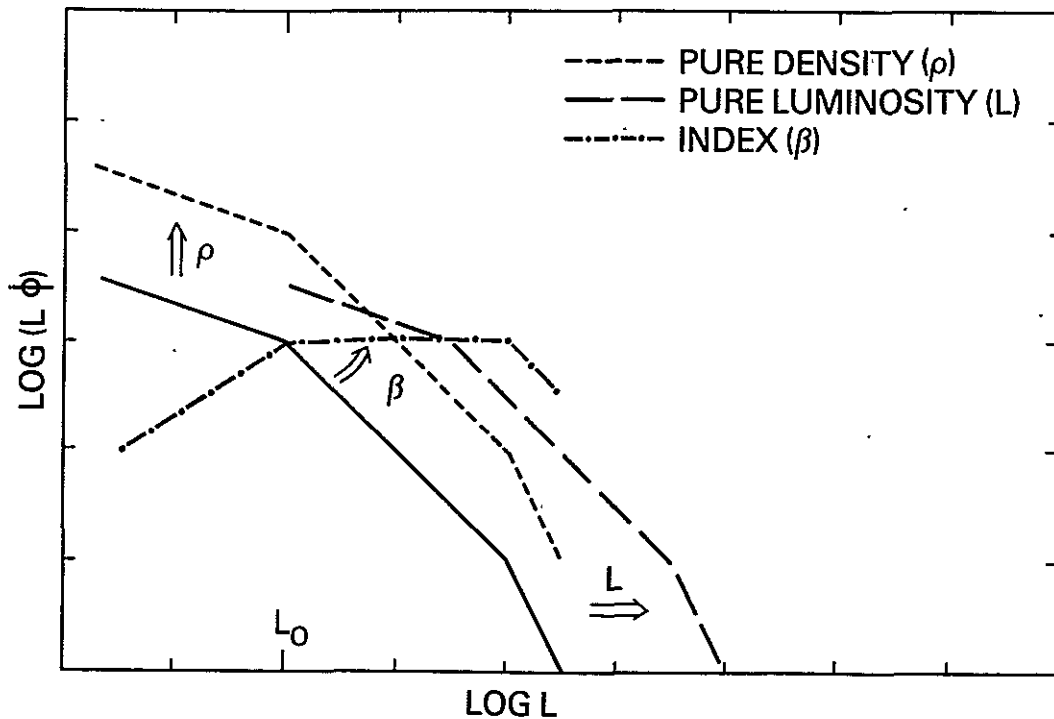


sharper roll off at the lower flux levels for pure luminosity evolution. This appears to be in better agreement with observed optical source counts. For this reason authors currently prefer pure luminosity evolution to pure density evolution [e.g. Cheney and Rowan-Robinson 1981a; Chanan 1982; Marshall, Avni, Tananbaum, and Zamorani 1983b], although Cheney and Rowan-Robinson [1981a] argue that a combination of pure luminosity and pure density evolution is also in agreement with the optical data. We will not further consider pure density evolution. (Note, though, Meier et al. [1979] show that the evolution of radio galaxies at moderate reshifts is not consistent with either pure

Analysis of Results                                    Evolving Populations

luminosity or pure density evolution.)

The general property of pure luminosity evolution that provides the desired behavior for the optical source counts is that high luminosity objects undergo more rapid evolution than low luminosity objects, at least for steep index luminosity functions. Schmidt and Green [1982; 1983] prefer another form of evolution with similar behavior, which they call luminosity-dependent density evolution:

$$\phi(L;z) = (L/L_0)^{f_\beta(z)} \phi_0(L) \ .$$

[VII.20]

$L_0$ is a fixed luminosity, given by the model, where there is no evolution. If the luminosity function is a power law form, with current epoch index $\beta_0$, then an effect of the evolution is to modify the index

$$\beta(z) = \beta_0 - f_\beta(z) \ .$$

[VII.21]

We refer to evolution of this form as index evolution. In VII.12 we see that index evolution is equivalent to a "rotation" of the luminosity function (in log-log coordinates) about the $\phi_0(L_0)$ point. There is a _decrease_ in the number of objects with $L < L_0$ so that for some amounts of evolution the volume emissivity can be less than for no evolution. The value chosen for $L_0$ has a major impact on the contribution to the sky flux and invariant excess.

An early form commonly used for the "pure" evolution functions, $f_\rho$ and $f_L$, was $(1+z)^k$ where k was the model parameter. However, because of divergences at high redshifts which produced a strong dependence on the redshift cutoff, $z_{max}$, the current preferred form is an exponential in the look-back time:

$$f_L = \exp[Q_L \tau(z)] \ ,$$

[VII.22]

where $Q_L$ is the constant model parameter and $\tau$, a function of redshift, gives the time from the emission to the current epoch as a fraction of the current age of the universe. For a $q_0 = 0$ cosmological model, $\tau = z/(1+z)$. An analogous parameter, $Q_\rho$, may be defined for density evolution. For the models tested by Schmidt and Green [1983] the index

Figure VII.13: K(S) for density, luminosity and index evolution

The unevolved population is the high luminosity end of an extension of the canonical AGN luminosity function, $\beta_h$ = 2.75. $q_o$ = 0.1. $z_{max}$ = 3.5.
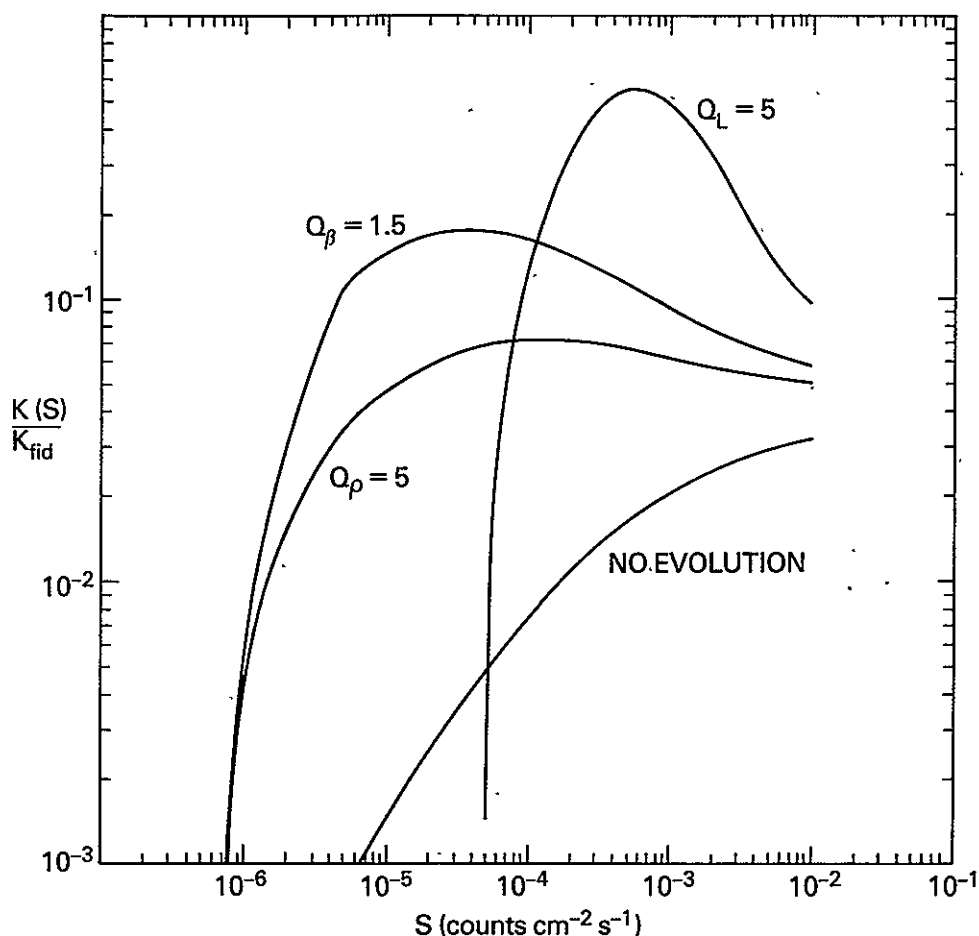


evolution function is proportional to $\tau$:

$$f_\beta = Q_\beta \tau .$$ [VII.23]

In Figure VII.13 we plot examples of such evolutionary models. We use the $\beta_h$ = 2.75 high range because the relatively narrow range in luminosities, 15<L<100 $L_{44}$, better illustrates the differences of the various models. As mentioned above, the pure luminosity model has a more peaked N(S) curve than the pure density evolution curve. Thus, though both have identical contribution to $S_{as}$, roughly a 14-fold

increase from the contribution of the unevolved form of the luminosity function, the luminosity model has a much larger contribution to $\sigma^2{}_S$ than the pure density model. The particular index evolution model shown has more contribution to the sky flux, but is midway between the other two in its contribution to the invariant excess. The curves in Figure VII.13 et seq. were constructed following the prescription in Appendix A1 for calculating N(S) with the inclusion of cosmological effects.

---

Figure VII.14: Effect of $q_0$ on luminosity evolution

The unevolved population is the canonical AGN luminosity function. Only the middle luminosity sources are included ($0.1 < L < 15 \ L_{44}$).



---
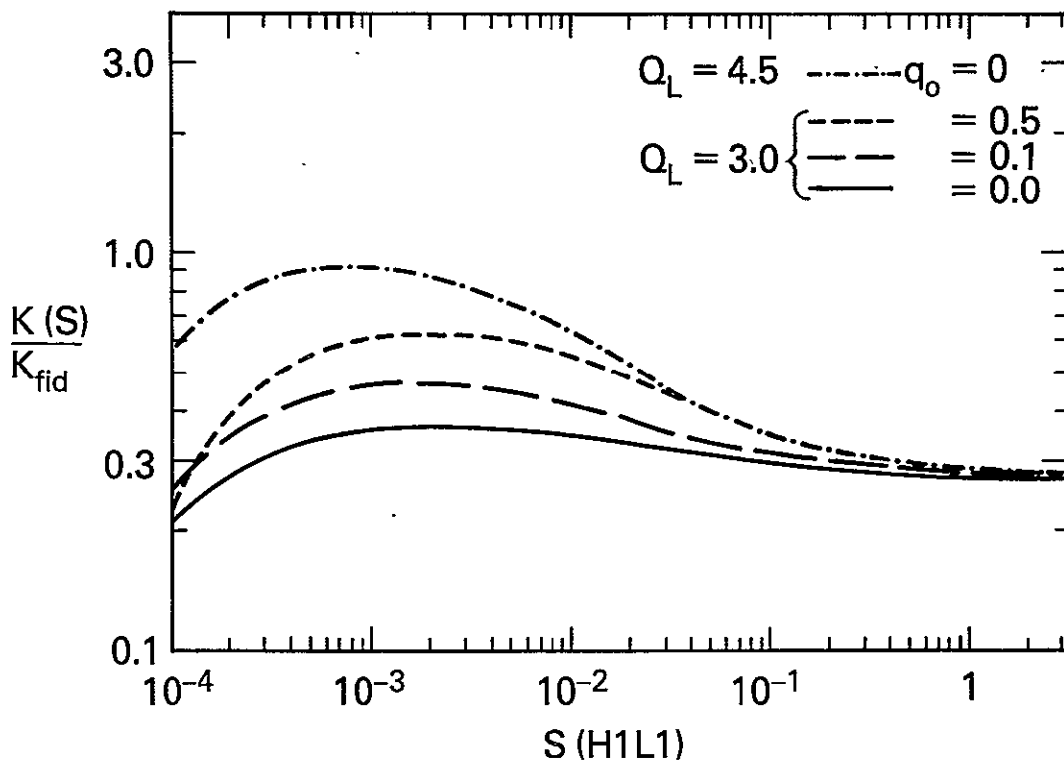
When the evolution models are directly parameterized in terms of the fractional look-back time, $\tau$, the behavior of the models has a strong dependence on the value chosen for $q_0$. This is most apparent for measurements dominated by the behavior at small or moderate redshifts.

Analysis of Results                                    Evolving Populations

For example, to get similar behavior at low z the value of Q for $q_o=0$ should be 3/2 that of a $q_o=0.5$ model. This is the ratio of the total look-back times of the two models. For $S_{as}$, those models with sufficient evolution that most of the contribution to $S_{as}$ is at high redshifts where the behavior of $\tau$ as a function of z is similar, there is less dependence on $q_o$. The contribution to $\sigma^2_S$, on the other hand, does show a significant dependence on $q_o$.

---

Figure VII.15: K(S) for luminosity evolution with various $z_{max}$

The unevolved population is the high luminosity extension with $\beta_h=3.75$.

---
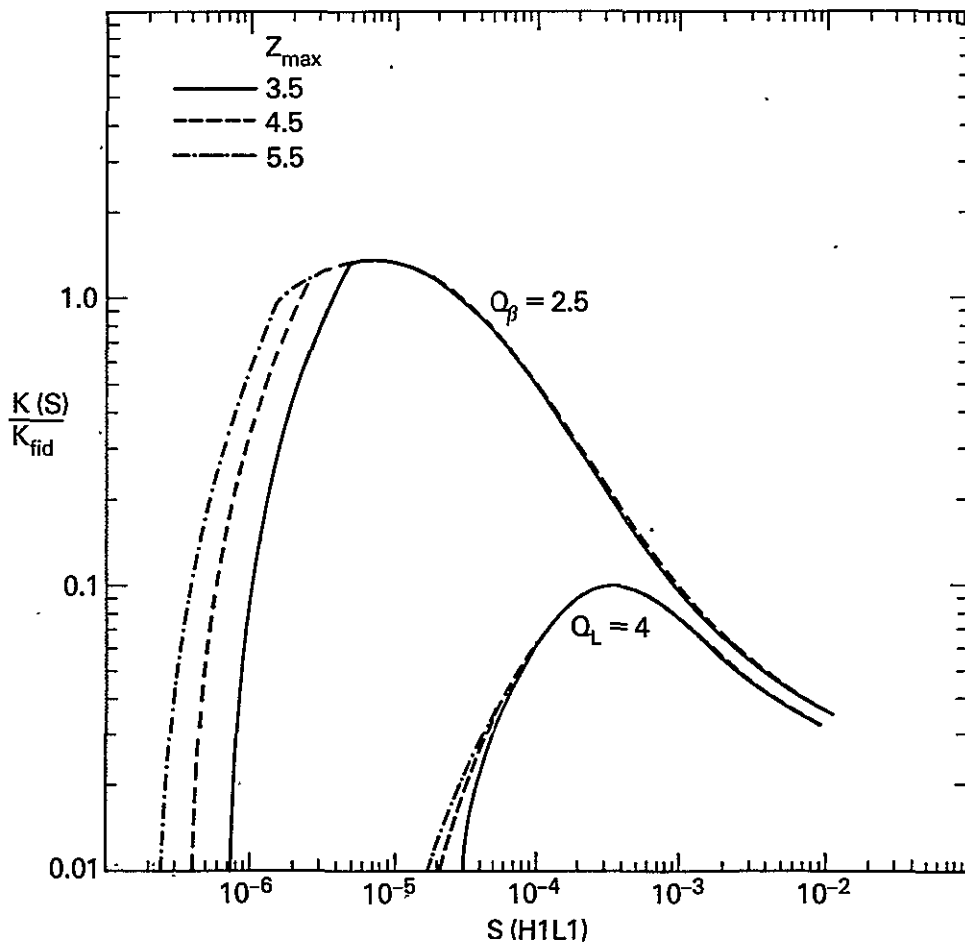


As we mentioned above, one reason for the particular choice of the

exponential form for the evolution function is the non-divergent behavior at high redshifts, unlike the $(1+z)^k$ laws. The exponential pure luminosity laws have only a small dependence on the actual choice of a value for $z_{max}$, the maximum redshift for the existence of AGN. For the pure luminosity evolution model shown in Figure VII.15, with $Q_L=4$, the contribution to both $S_{as}$ and $\sigma^2_S$ varied by less than 1%, although stronger evolution naturally produces a more significant difference. The index evolution model shows a much stronger dependence, a variation of 5% in $\sigma^2_S$ and 65% in $S_{as}$ in going from $z_{max}$ of 3.5 to 5.5. This is a consequence of using the high luminosity AGN range in the example of Figure VII.15. The effect for the low and medium luminosity range objects is significantly smaller. Also, models with larger values of $q_0$ show less dependence on $z_{max}$ because they reach the asymptotic limit for the look-back time, $\tau \to 1$, at lower redshifts.

For the rest of our discussion we will assume that $z_{max}$ is 3.5, following the indication by Osmer [1982] of a decrease in the density of quasars at about that redshift. Osmer's conclusion is independent of any evolutionary laws, and would be strengthened in the event of significant luminosity (or index) evolution in the redshift ranges 2.5 to 3.5 and 3.7 to 4.7. However, this result does depend on the value for the low redshift quasar density derived by Hoag and Smith [1977]. For a critique of Osmer's use of this value, and his general conclusion of the need for a cutoff, see Mathez and Nottale [1982]. We also note that the redshift cutoff in the quasar luminosity function can not be total, as there are examples observed out to redshifts of $z = 3.78$, the record currently held by PKS 2000-330 [Peterson, Savage, Jauncey, and Wright 1982].

In the discussion that follows it is important to keep in mind the ad hoc nature of the models presented. We have already mentioned the assumption of no spectral evolution. No less a difficulty is the imprecision in the current epoch luminosity function, particularly the position and size of the break in the index between the low and middle luminosity sources. Rather than treat the evolution models and luminosity function as true representations of reality, let their

limitations and successes serve to suggest the assuredly more complicated behavior of the AGN luminosity function. The two constraints that we will use are the total all-sky flux and the acceptable range for $\sigma^2_S$ for AGN. The upper bound to $\sigma^2_S$ is $\sim 8.5\times10^{-4}$ HlLl$^2$. The lower bound of $2.5\times10^{-4}$ is softer because there may be some other sources of excess variance (see particularly the discussion in Section VII:6). This differs from the allowed range for the excess variance of a totally new component, used above, since for that case the contribution of AGN was fixed to be $\sim -0.5\times10^{-4}$ HlLl$^2$. This range of values is derived assuming that $\sigma^2_S$ is the only interesting parameter of the fit.

### Pure luminosity evolution

First consider the case where both the low and middle luminosities evolve following the same rule. When calculating K(S) we find that for even the high flux range of the XCS, evolution has a noticable impact, particularly for the larger values of $Q_L$. The effect is generally not a strong departure from Euclidean behavior, but rather a general increase in the norm, as represented by $K_h$. (Recall $K_h = K(S_h = 3\times10^{-3}$ HlLl).) It is basically a low redshift effect, with the excess contribution from the evolution of the moderate and higher luminosity objects more than compensating for cosmological volume and K-correction redshifting of the object's spectrum effects. For $q_o = 0.1$ and $Q_L = 3$ the enhancement of $K_h$ is about 17%, increasing to 29% for $Q_L = 4.5$. Thus for evolution at such levels the current epoch luminosity function should be modified. We will confine our correction to the normalization, adjusting $\kappa$ so that the evolved $K_h$ is the same as the unevolved $K_h$. A proper correction would also include a modification in the luminosity function index, $\beta$, although the exact form of the correction is difficult to determine. The effect originates primarily in the higher luminosity AGN, so that we might expect a steepening in the index for the highest luminosity objects and a slight flattening at moderate luminosities. The exact magnitude of the correction is difficult to estimate as it depends on the actual index in a non-linear fashion, i.e. as the true index steepens, the magnitude of the correction to both $\kappa$ and $\beta$ is reduced.

ORIGINAL PAGE IS
OF POOR QUALITY

The XCS alone probably would not be able to constrain any but the most extreme evolutionary models.

Figure VII.16: K(S) for AGN with luminosity evolution

Both low and middle luminosity objects are evolved, using the indicated amount of pure luminosity evolution. $q_0=0.1$. The long dashed lines give the total K(S), including non-evolving clusters and the estimated galactic source contribution. The short dashed lines are range of behavior for broken power law models derived from the best confidence region on K and $\sigma^2_S$.
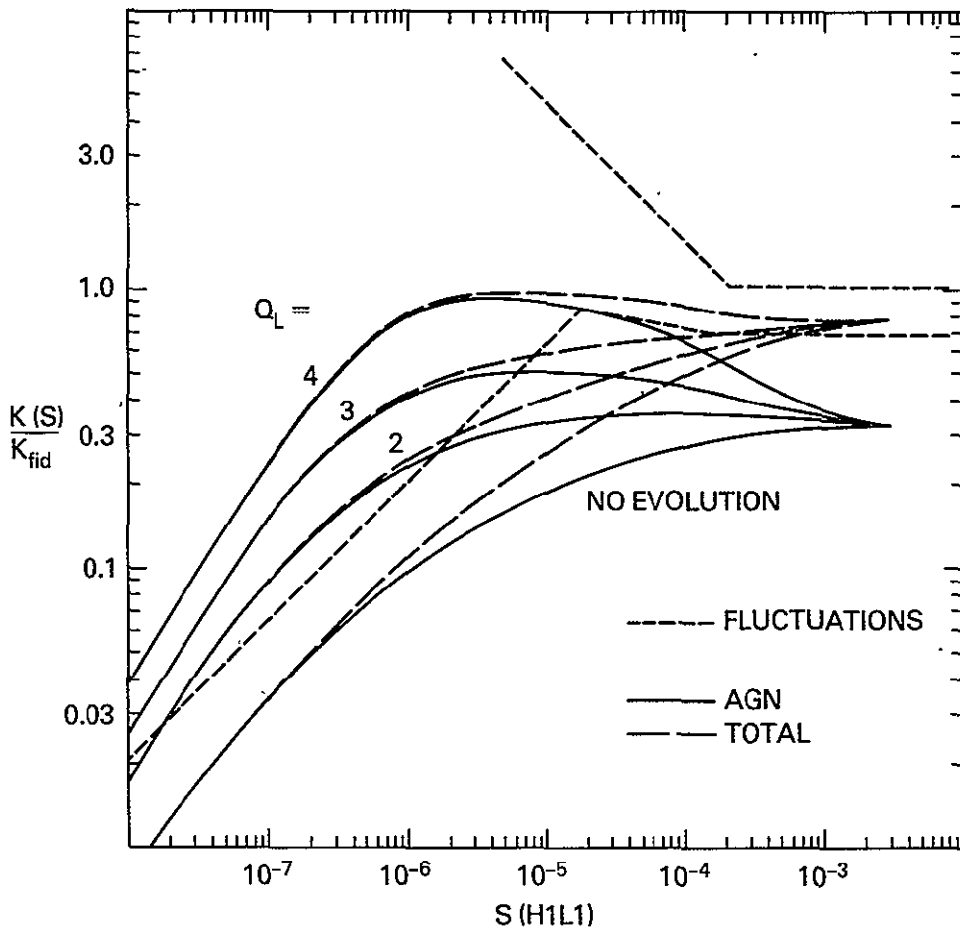


In Figure VII.16. we present K(S) curves for pure luminosity evolution, where both low and middle luminosity objects are evolved. We

Figure VII.17: Contributions to $S_{as}$ and $\sigma^2_s$ for AGN versus $Q_L$

Curves for both $q_o$=0.1 and 0.5 are plotted.. Curves marked "Evolving Mid + Evolving Low" have evolution for both middle and low luminosity objects. Curves marked "Evolving Mid + Nonevolving Low" include both middle and low luminosity sources but the evolution law applies only to the middle luminosity objects. The curve marked "High" indicates only the contribution of the high luminosity extension ($\beta_h$ = 3.75) to the invariant excess. The contribution of such sources to $S_{as}$ is negligible.
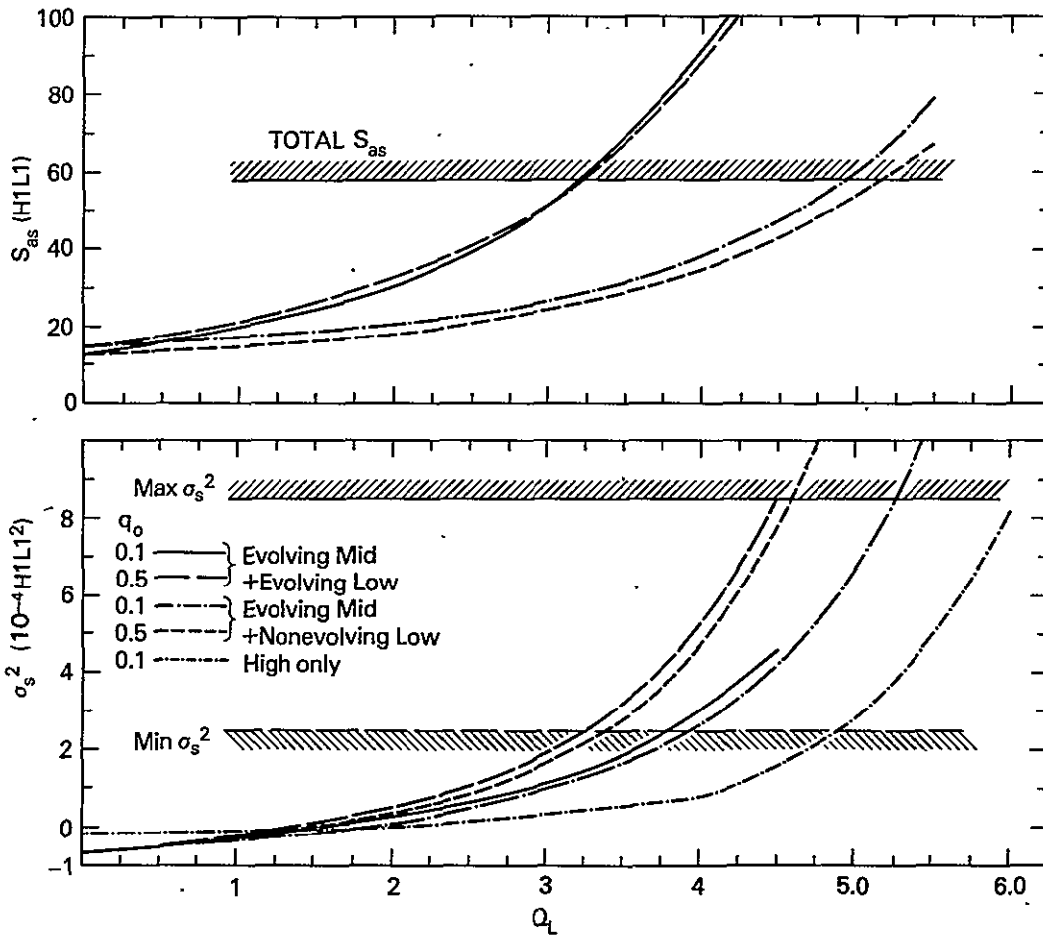


see that the excess "hump" of the luminosity evolution curve in Figure VII.13 has been broadened to a shelf by the extension of the range covered by the luminosity function. The form of the "Total" curve shows us that with evolution of the AGN with $Q_L{\sim}3.5$, the total N(S) has a near

Euclidean behavior continuing for almost 3 1/2 decades below the resolved sources. Though Figure VII.16. allows us to compare the K(S) curve with broken power law constraints set by the fluctuations and the fits to the resolved sources, we know that the better test of a model is a direct comparison with the limits set on the fit parameters of K and $\sigma^2_S$. Because of the renormalization of the curves, the value of K is fixed, by definition, at an acceptable level. Figure VII.17 illustrates the contribution to the invariant excess as well as the contribution to the all-sky flux. We see here that the contribution to $S_{as}$ is not strongly affected by the value of $q_0$, and that we generally constrain $Q_L \lesssim 3.25$ to avoid exceeding the observed total sky flux. The contribution to the invariant excess, however, has a stronger dependence on $q_0$. As an absolute limit on the evolution, $\sigma^2_S$ is less of a constraint than the total sky flux: $Q_L \lesssim 4.5$ for $q_0=0.5$; $Q_L \lesssim 5.2$ for $q_0=0.1$. More interesting perhaps, is the lower limit placed on $Q_L$ if we require that the AGN supply the "missing" excess variance. For $q_0=0.5$ this would imply that $Q_L \gtrsim 3.25$, so that only a very narrow range of values $\approx 3.25$ can provide the required $\sigma^2_S$ without exceeding the limit on $S_{as}$. For $q_0=0.1$ the difficulty is severe. At the point where $S_{as}$ is exceeded $\sigma^2_S \approx 1.5 \times 10^{-4}$ $H1L1^2$, significantly less than the required $2.5 \times 10^{-4}$. The difficulty is not resolved by having some fraction of the evolution as pure density evolution, such as was suggested by Cheney and Rowan-Robinson [1981a]. As mentioned in the discussion of Figure VII.13, for an equal contribution to the all-sky flux, pure density evolution will contribute less to the excess variance than pure luminosity evolution. Keep in mind that the minimum $\sigma^2_S$ requirement need not be met solely by evolution. As we shall demonstrate in Section VII:6 there are equally plausible processes that generate significant additional variance, such as clumping of sources or other large scale structure in the X-ray sky.

The difficulty is to produce a significant contribution to the invariant excess without exceeding the total $S_{as}$. We recall that the bulk of the contribution to the all-sky flux, for both the evolved and unevolved models, comes from sources in the neighborhood of the division between the low and middle luminosity sources, where the luminosity

Figure VII.18: K(S) for evolving middle luminosity AGN

The curves compare K(S) for AGN where both low and middle luminosity objects evolve (dashed) to that where only the middle luminosity objects evolve (solid). $q_o = 0.1$.
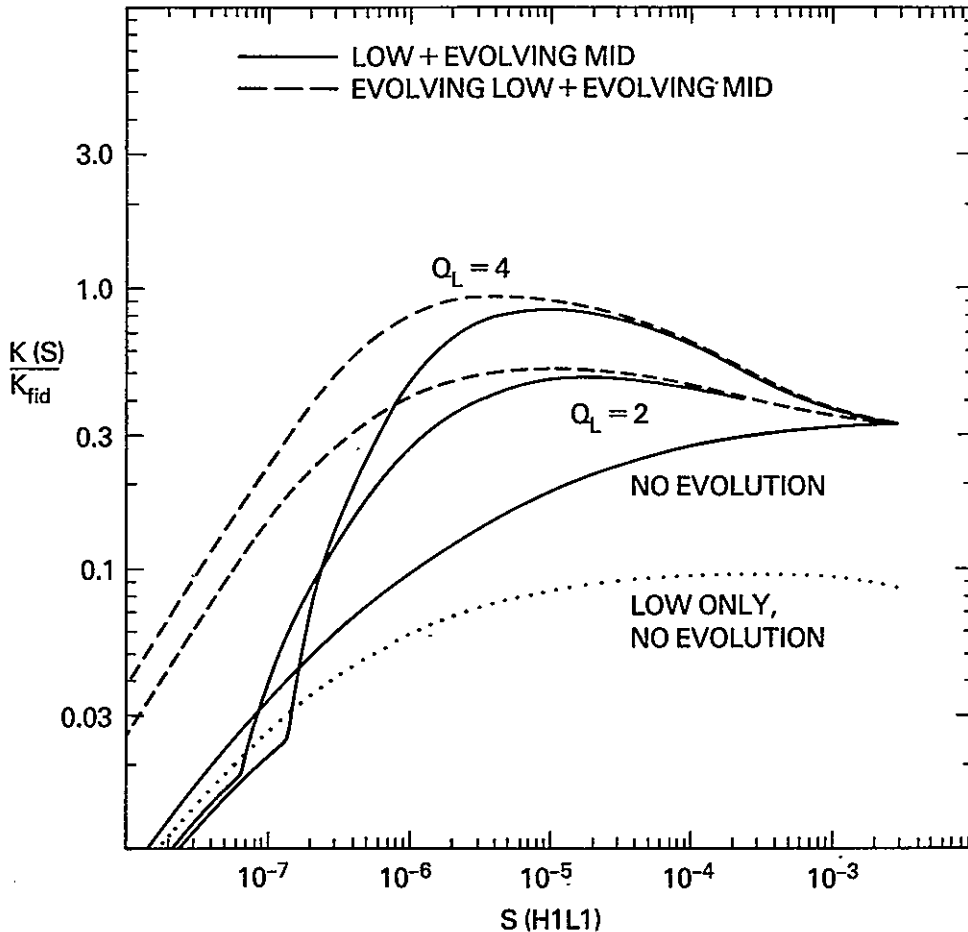


function index breaks from 2.75 to 1.75, at $10^{43}$ ergs $s^{-1}$. The most significant contribution to $\sigma^2_S$, on the other hand, is made by sources that produce the high flux edge to the enhanced plateau of K(S), which are the higher luminosity sources. To show this, Figure VII.18 compares models where only the middle luminosity objects are evolved to models where all AGN are similarly evolved. By only evolving the middle luminosity sources the enhanced plateau has been significantly shortened. Examination of the "mid only" curves in Figure VII.17 shows

Analysis of Results                                   Evolving Populations

that this does have the desired effect. The sky flux is not exceeded until $Q_L \approx 5$ for $q_o=0.1$ ($\sim 5.25$ for $q_o=0.5$). The contribution to $\sigma^2_{\#S}$ is only slightly affected, both the high and low limits on $Q_L$ are increased by $\sim 0.1$. Therefore for $q_o=0.5$ the strong upper limit of $\sim 4.6$ on $Q_L$ comes from the upper limit on the invariant excess. If the AGN evolution provides the missing invariant excess then $3.4 \lesssim Q_L \lesssim 4.6$. For $q_o=0.1$, larger values of $Q_L$ are preferred: $3.9 \lesssim Q_L \lesssim 5.0$, the upper limit imposed this time by the $S_{as}$ condition.

Though now we have a model that can provide the required amount of $\sigma^2_S$ without exceeding the diffuse sky flux, we recognize the arbitrary nature of the change from the all source evolution model. There is no particular reason that sources below $10^{43}$ erg s$^{-1}$ did not evolve while those sources currently above that level have evolved. Also, such a model would have at large redshifts an empty gap in luminosity between the top of the unevolved sources and the bottom of the evolved sources. For $Q_L=4$ the gap could span a decade and a half. However, the results for this admittedly simple model indicate that if AGN evolution is to provide the missing invariant excess then the preferred form for the evolution is one where the large luminosity objects undergo stronger evolution than the lower luminosity objects.

### Index evolution

The luminosity-dependent density evolution proposed by Schmidt and Green [1983] is one attempt to parameterize a model with explicitly stronger evolution for the higher luminosity AGN. In fact, for strong evolution the index evolution models will completely invert the luminosity function so that not only $\sigma^2_S$ but the contribution to the sky flux and even the total number of objects will be completely dominated by the very high end of the luminosity function. For $q_o=0.1$, the fractional look-back time at $z=3.5$ is $\sim 0.8$. If the current epoch luminosity index is 2.75, by the time we reach $z_{max}$ the effective index is $2.75 - 0.8\, Q_\beta$, so that for $Q_\beta \gtrsim 0.9$ the contribution to the volume emissivity at $z_{max}$ is dominated by the high luminosity objects. Initially, however, the enhancement is a slow function of $\tau$, so that for

moderate values of $Q_\beta$ the high flux region does not show the sharp rise in $K(S)$ seen in the luminosity evolution models. The enhancement to $K_h$ is less for a given contribution to $S_{as}$ compared to pure luminsity evolution. For $q_o=0.1$, $L_o=0.5$ $L_{44}$, and $Q_\beta=3.5$, the enhancement is ~13%.

---

Figure VII.19: $K(S)$ for luminosity function index evolution of AGN

Both low and middle luminosity AGN are subject to evolution: $q_o=0.1$, $L_o=0.5$ $L_{44}$. The long dashed lines give the total $K(S)$, the short dashed lines are derived from the confidence region on $K$ and $\sigma^2_S$ (see Figure VII.16.).

---
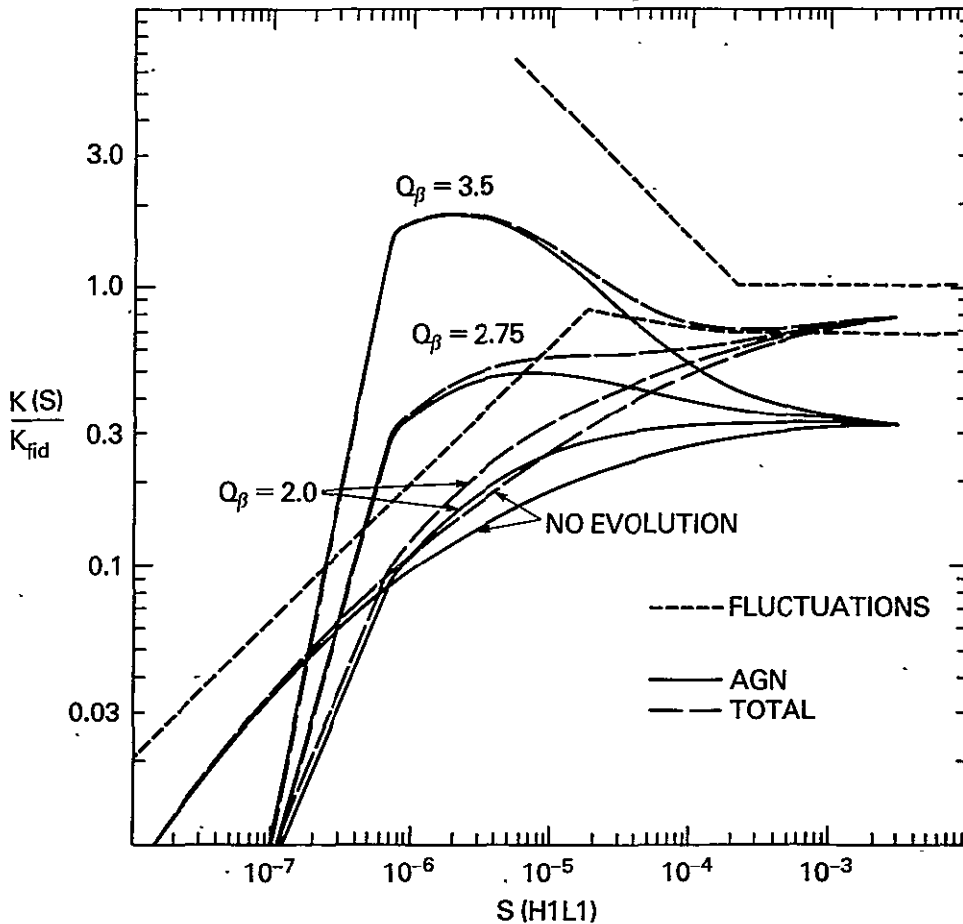


The slow deviation from the unevolved $N(S)$ is illustrated in the

Analysis of Results                                    Evolving Populations

Figure VII.20: Contribution to $S_{as}$ and $\sigma^2_S$ for AGN versus $Q_\beta$

See description of Figure VII.17. Note that the contribution of the high luminosity sources ($\beta_h$ = 3.75) to $S_{as}$ is now significant, and plotted. $L_o$=0.5 $L_{44}$.
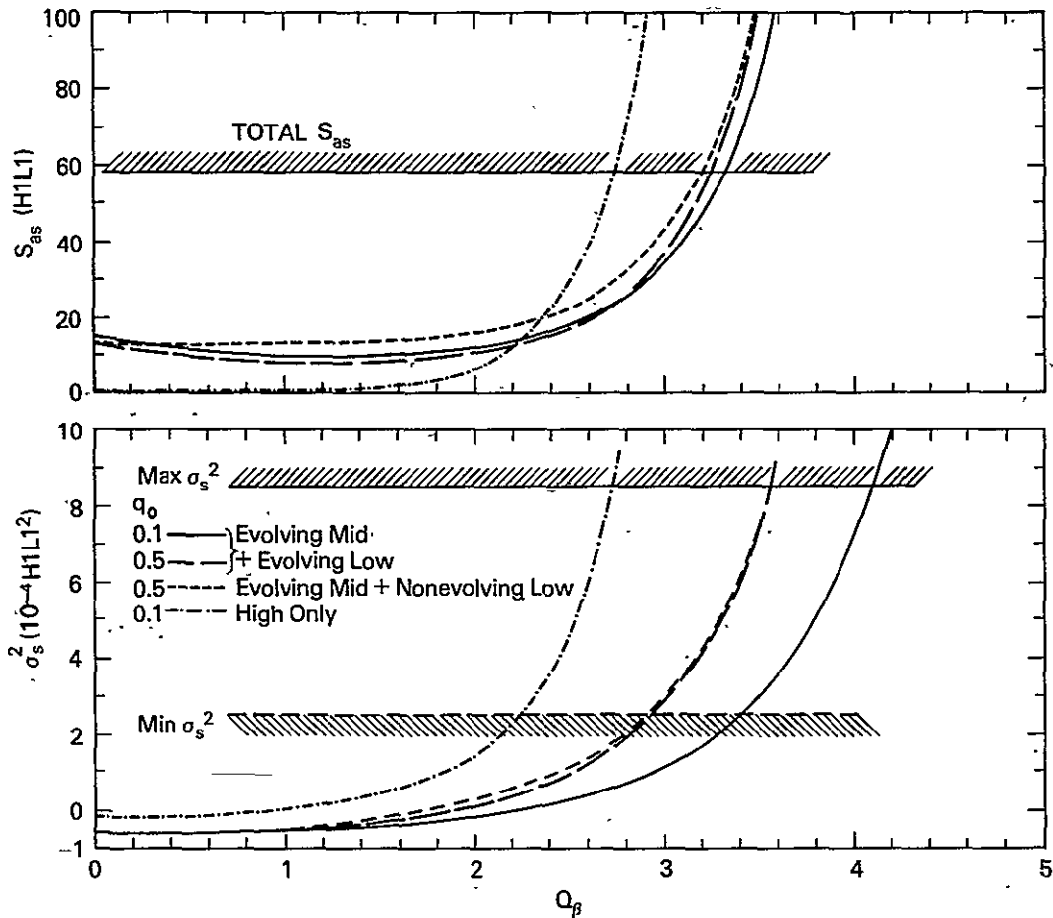


K(S) plots of Figure VII.19. For moderate amounts of evolution, $Q_\beta \lesssim 2$, the principal effect is the deficit of low flux sources. This deficit is maintained for the stronger amounts of evolution, and is due to the reduction of sources with luminosities less than $L_o$=0.5 $L_{44}$. In Figure VII.20 we see that the contribution to the all-sky flux is exceeded for $Q_\beta \approx 3.25$. There is little dependence on $q_o$. For intermediate amounts of evolution, i.e. $Q_\beta \sim 1.5$, the contribution to $S_{as}$ is actually decreased by the suppression of low luminosity sources. If the evolution is

Analysis of Results                                    Evolving Populations

confined only to the medium luminosity sources this suppression is reduced. However for larger amounts of evolution the contribution to $S_{as}$ is totally dominated by the high end of the medium luminosity objects. The contribution to $\sigma^2_S$ is similarly dominated by the middle luminosity sources so that the "middle only" curve is indistinguishable from the case where evolution applies to both middle and low luminosity objects.

Though the evolution is dominated by the changes at the high end of the luminosity function, the strong deviation from a Euclidean N(S) does not occur until relatively small values of the flux. Therefore, index evolution models with $L_o = 0.5 \ L_{44}$ do not make much contribution to the invariant excess even for values of $Q_\beta$ large enough to exceed $S_{as}$. Considering only the invariant excess we find that if $q_o = 0.5$ then $2.9 \lesssim Q_\beta \lesssim 3.6$. The bottom limit is the soft requirement that the evolving population produces the required positive excess variance, while the upper limit is from the hard maximum allowed excess. The range of allowed models is quite small. For $q_o = 0.1$ the limits placed by the invariant excess are $3.4 \lesssim Q_\beta \lesssim 4.1$ so that formally no model can produce the needed positive excess without exceeding the total sky flux.

These ranges were calculated for a single value of $L_o$, $0.5 \ L_{44}$. The contribution to both $S_{as}$ and $\sigma^2_S$ depends strongly on the value chosen for $L_o$, so the ranges should only be considered indicative of the general behavior of index evolution models. For example, if we fix $Q_\beta = 3$ and allow $L_o$ to vary from 0.25 to 1.0 $L_{44}$, the contribution to $S_{as}$ runs from ~130 to less than 15 H1L1 while $\sigma^2_S$ covers the range from 4.5 to $-0.2 \ 10^4$ H1L1$^2$. When Schmidt and Green performed their fits to the optical data their best fit values for $L_o$ ranged over a factor of 3, depending on the value chosen for $q_o$ and the number of AGN in other low-flux surveys. Though modifying the value for $L_o$ will change the range of acceptable values for $Q_\beta$, this has only a slight impact on the difficulty of providing the minimum positive excess variance, ~$2.5 \times 10^4$, without exceeding the diffuse background. This is especially true for the $q_o = 0.1$ case.

Analysis of Results                                              Evolving Populations

## High luminosity extensions of AGN luminosity function

Given the strong dependence on $L_o$, we might think that to make a statement about index evolution models we should replace Figure VII.20, which was created holding $L_o$ fixed at 0.5 $L_{44}$, by a contour plot of $S_{as}$ and $\sigma^2_s$ in the $L_o-Q_\beta$ plane. However, we will now show that there is a more significant source of imprecision in the contribution to the integral quantities, particularly for the larger values of $Q_\beta$. Most of the contribution to $S_{as}$ and $\sigma^2_s$ comes from the sources near the high luminosity limit of 15 $L_{44}$. This limit is arbitrary, and its position was more a statement of the limitations of the XCS for the study of high luminosity AGN. In fact, the A-2 experiment has observed two AGN with luminosities significantly greater than this:  3C 273 at 82 $L_{44}$ and NRAO 140 at 400 $L_{44}$ [Marscher et al. 1979; see also Holt et al. 1983; Halpern 1982]. The former was included in the XCS, while the latter was excluded because its intensity of 1.02 RI5 ($\approx 1.6 \times 10^{-3}$ HIL1) was below the XCS lower flux limit, as well as being at a galactic latitude outside the solid angle covered by the XCS. The existence of these objects, and extrapolations based on detections performed on other high luminosity QSOs with the Einstein observatory, indicate that the current epoch luminosity function must be extended above 15 $L_{44}$. (It is interesting to note that both 3C273 [Pearson et al. 1981] and NRAO 140 [Marscher and Broderick 1982] show superluminal expansion in the radio.)

Though we have no information from the XCS about the form such an extension should take, we will here consider the two examples we mentioned above:  (1) a luminosity function with index $\beta_h$ continuing at 2.75 between 15 and 100 $L_{44}$, and (2) one where $\beta_h$ steepens to 3.75 in that same luminosity range. In Figure VII.20 we also showed the contribution of the high luminosity sources to $S_{as}$ and $\sigma^2_h$ as a function of $Q_\beta$, with $\beta_h=3.75$. The curves shown give the contribution from the high luminosity sources only. The high sources always dominate the invariant excess, and in fact contribute more to $S_{as}$ than all other AGN for $Q_\beta \gtrsim 2.2$ ($q_o=0.1$, $L_o=0.5$ $L_{44}$). With a value of $Q_\beta$ as small as 2.2 the high luminosity objects can provide the needed minimum $\sigma^2_s$, even for $q_o=0.1$, with as little as ~50% of the total sky flux accounted for by

Figure VII.21: Effects on AGN K(S) for index evolution of a high
luminosity extension of the luminosity function

The three solid curves give K(S) of all AGN for, from the
bottom: no high luminosity objects, $\beta_h$=3.75, and $\beta_h$=2.75.
$Q_\beta$=2.5, $q_o$=0.1 and $L_o$=0.5 $L_{44}$.
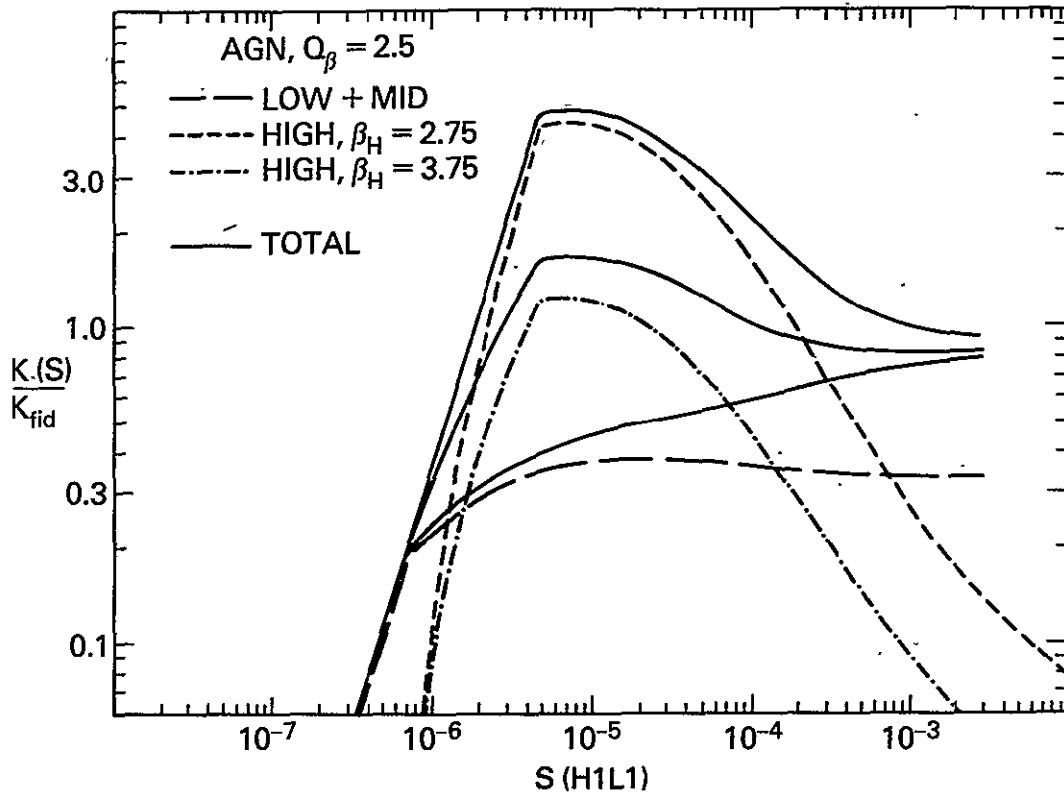


all AGN. With the flatter extension of the luminosity function,
$\beta_h$=2.75, it is even easier for small amounts of index evolution to
provide the needed invariant excess while making a relatively small
contribution to $S_{as}$. However, the flatter extension will quickly exceed
the maximum allowed excess and sky flux.

In figure VII.21 we show the contribution to the total K(S) made
by the two extensions, with $Q_\beta$=2.5. The $\beta_h$ curve has a larger, peaked,
contribution at fluxes just above the steep break. This is because such
a model has more high luminosity objects at the current epoch.

Analysis of Results                                    Evolving Populations

Therefore, for an understanding of particular index evolution models and their contribution to N(S), particularly at fluxes near $\sim 10^{-6}$ HlLl, a knowledge of the highest end of the current epoch luminosity function is essential. Note however that all index evolution models have the general property that below the enhancement there is a sharp characteristic drop in K(S). The size of the drop depends on the behavior of the luminosity function at the low luminosity end. By this we mean that if the low luminosity sources $\lesssim L_o$ are not subject to index evolution, then K(S) will drop a few decades to meet the less steeply rolling over contribution to K(S) from the unevolved low luminosity objects. A similar sharp drop is seen for the pure luminosity evolution models but, again, only when the low luminosity sources do not evolve (see Figure VII.18).

In Figure VII.22(a) we see that no matter how strong the evolution or what the details of the luminosity function at the high end, the major effect of index evolution is on fluxes far below the range of the XCS. This is because index evolution enhances the number of sources, not their luminosity. To first order, increasing $Q_\beta$ just increases the pile-up at the low flux range.

Pure luminosity evolution has a very different interaction with the current epoch highest luminosity sources. To begin with, for suitably large values of $Q_L$ there is no requirement to have _any_ extension of the luminosity function above $\sim 15\ L_{44}$. That is, 3C273 and NRAO 140 are both observed at high enough redshifts that already significant evolution has occurred. For example, at a $z \sim 1$ and $Q_L \sim 3$ the 15 $L_{44}$ upper luminosity limit has been shifted to $\sim 70\ L_{44}$. For $Q_L$ of 6 it is at $\sim 300\ L_{44}$.

The existence of a high luminosity extension to the luminosity function has only a small impact on the total AGN contribution to $S_{as}$. For example, for $Q_L=6$ the high luminosity sources would contribute 2.4 HlLl or 1.3 HlLl for $\beta_h$ of 2.75 or 3.75 respectively. As illustrated by the curve marked "High luminosity only" in Figure VII.17(b), if the high luminosity sources evolve at the same rate as both the low and middle

Figure VII.22: K(S) for only high luminosity AGN

Solid and long dashed curves are for the indicated values
of $\beta_h$. The short dashed curve is the canonical total K(S)
for unevolved populations. No renormalization to the XCS
and fluctuations-derived value for $K_h$ was made for the
high luminosity extensions. The flux range of the XCS,
$2 \times 10^{-3}$ to $8 \times 10^{-3}$ H1L1, is centered on the high flux end of
the short dashed curve. $q_0 = 0.1$. (a) Index evolution,
$L_0 = 0.5$ $L_{44}$. (b) Pure luminosity evolution.
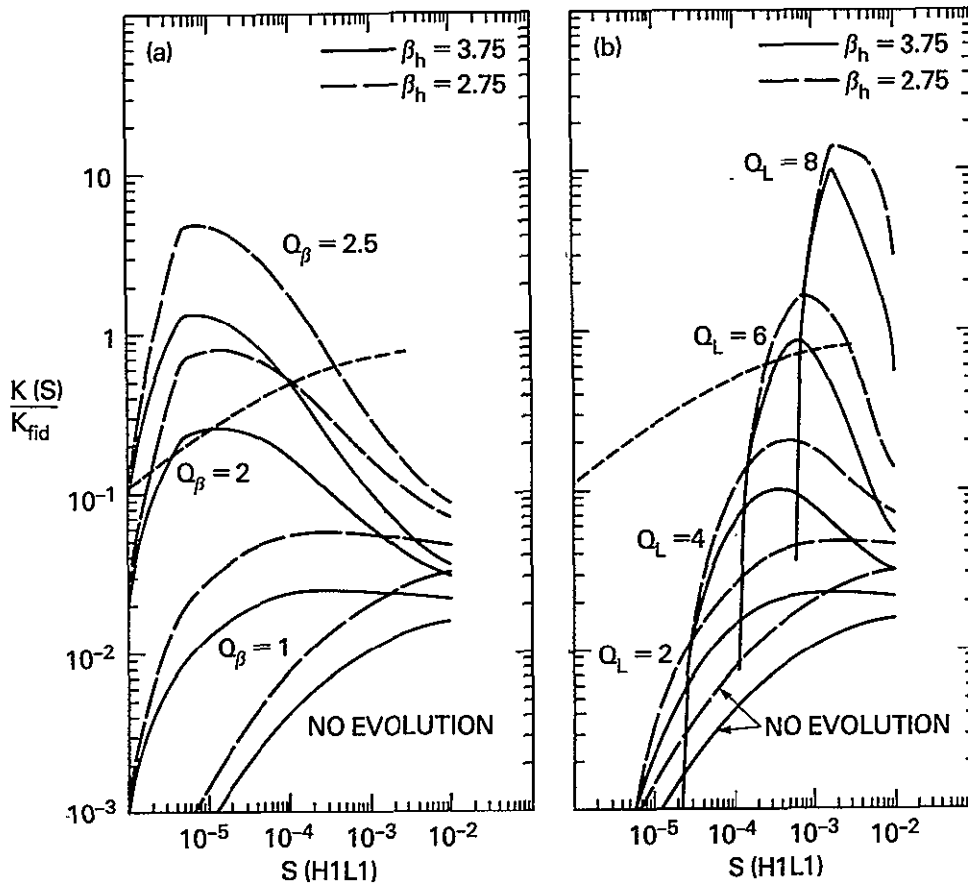


luminosity objects, ($Q_L \lesssim 3.2$) then the additional contribution to the
invariant excess from the high range is small. If the low luminosity
sources do not evolve, on the other hand, the contribution to $\sigma^2_S$ from
the high luminosity extension can be a significant fraction of the total
when the high and middle sources have the same $Q_L$. The relative
contribution can be further enhanced if the high range undergoes

Analysis of Results                                    Evolving Populations

stronger evolution than the middle range. For example, if $Q_{L;mid}$ is 3 and $Q_{L;high}$=4.5 then the total $\sigma^2{}_S$ is $2.5\times10^{-4}$ HlLl$^2$ and $S_{as}$=25 HlLl ($q_o$=0.1, $\beta_h$=3.75). Such a model would therefore provide the missing excess variance and only contribute 45% of the total sky flux. There are limitations, however, on $Q_{L;high}$. An examination of VII.22(b) shows that for $Q_L\geq6$ ($q_o$=0.1, $\beta_h$=3.75) the high luminosity objects will make a significant contribution to the source counts in the flux range of the XCS. Given that 3C273 is the only high luminosity object in the XCS, either such extreme evolution is rejected or the power law index of the extension is even steeper (equivalently the extension does not continue out to 100 $L_{44}$ as we have assumed here). For the flat high luminosity extension, $\beta_h$=2.75, the maximum allowed evolution is $Q_L\sim5$.

One might argue on purely physical grounds against the largest amounts of pure luminosity evolution. Rather than allowing the luminosity of an object to become arbitrarily large there could be intrinsic limits on the maximum possible luminosity. The Eddington limit, for example, imposes a maximum total luminosity for steady state emission associated with accretion onto an object. Luminosities above $1.3\times10^{38}$ ergs s$^{-1}$ for a one solar mass object will disrupt the object or inhibit the accretion. The details of this limit depend on geometric considerations, the above limit being for the case of a spherically symmetric system. In Figure VII.23 we examine the effects on luminosity evolution assuming that no object can exceed a luminosity of $1.3\times10^{47}$ ergs s$^{-1}$, corresponding to a $10^9$ solar mass central engine for AGN. All sources that would otherwise exceed this limit are piled up at the maximum luminosity. In the figure we see that for $Q_L$ of 5 ($q_o$=0.1) the effects on the K(S) curve for the high flux sources is small, though not negligible. Larger amounts of evolution would have a correspondingly greater impact; however in general such a limit will not interfere with the conclusion that really strong amounts of evolution are incompatible with the lack of high luminosity objects observed in the XCS.

For pure luminosity evolution, as well as index evolution, the extension of the luminosity function can have important ramifications, particularly for the contribution to the excess variance. There is,

Figure VII.23: Effects of absolute maximum in luminosity on K(S)

The curves are for the high luminosity extension of the AGN luminosity function undergoing $Q_L$=5 pure luminosity evolution, with and without an absolute maximum luminosity of 1.3x10[47] ergs s[-1]. This corresponds to an Eddington limited luminosity for accretion on a 10[9] solar mass object. $q_o$=0.1.
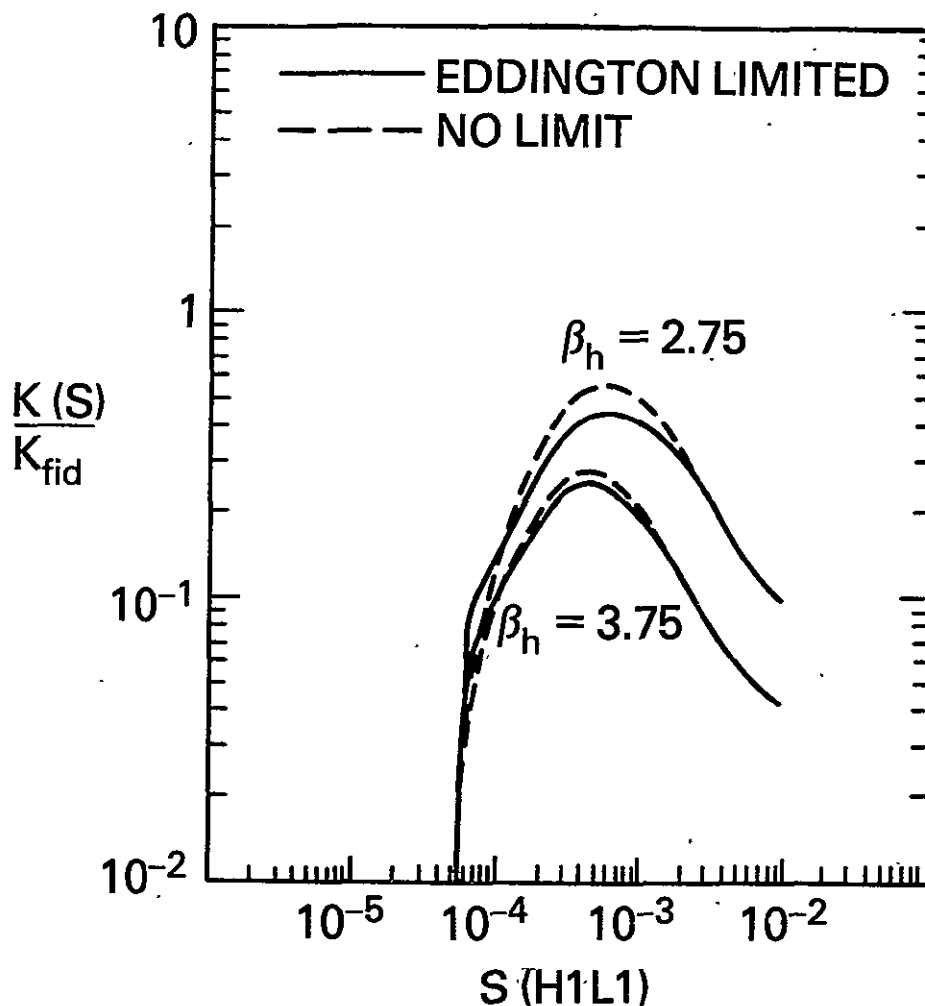


however, a problem in determining the form of such an extension, as the high luminosity objects will not be observed, on average, except at relatively large distances. By "relatively large" we mean compared to the distances for the bulk of the sources in the XCS, so that the high

luminosity sources may already have been subject to significant amounts of evolution. Therefore a determination of the "current epoch" luminosity function will depend on the assumed model for the evolution.

Summary

The evolution of AGN can provide a significant contribution to the excess variance and to the total sky flux. The total sky flux provides a strong constraint on most forms of evolution. Only a limited range of models can provide the desired amount of excess variance without exceeding the sky flux condition. However, large uncertainties exist in the contribution from both the very lowest and the very highest luminosity sources. A summary of the numerical results from this section is presented as part of Chapter IX.

---

4: Results from the Einstein Observatory: Source Counts

We review the results of the low flux source counts obtained from the Einstein Observatory. Though the differing energy bands of Einstein and the A-2 experiment make comparisons strongly dependent on spectral assumptions, the "Medium Survey" of serendipitous sources covers a range in flux roughly equivalent to the range of greatest sensitivity of the fluctuations. The N(S) curves derived are in agreement, again depending on spectral assumptions. The identification of the sources which constitute the Medium Survey implies some evolution by AGN from the luminosity function observed in the XCS. If the evolution is pure luminosity evolution, where $L(\tau) = L(0) \exp(Q_L \tau)$, ($\tau$ is the fractional lookback time), then the Medium Survey source counts and the XCS derived current epoch luminosity function predict $2 \lesssim Q_L \lesssim 4$, $q_0 = 0.1$. If a substantial number of the low luminosity AGN are not absorbed or are only partially covered with absorbing material, the allowed range must be significantly reduced. Similar results are obtained for the original "Deep Survey" of Giacconi et al. [1979]. The analysis of these sources is sensitive to the behavior of N(S) in a narrow flux range, and its interpretation depends on the assumed power law index of N(S). No conclusion about the presence or absence of evolution for BL Lacertae objects can

be drawn solely from their relative frequency in the XCS and Medium Survey. The promise, and difficulties, of an analysis of the fluctuations using _Einstein_ Observatory data are outlined.

Though the fluctuations do make a case for the existence of evolution, we have seen that the range of allowed models is not well determined. In addition, the required excess variance need not be supplied solely, or at all, by such evolution. As we will discuss, clumping of the sources or other intermediate scale structures in the sky flux may also be responsible. Explicit study of AGN over a range of fluxes and redshifts would better allow us to decide on the amount and form of the evolution.

Such a study is being conducted by several authors using data obtained by from the HEAO 2 satellite _Einstein_ Observatory. The imaging capabilities of the _Einstein_ X-ray telescope allowed observations of sources with much lower fluxes than could be studied with HEAO 1 A-2. Unfortunately, the grazing incidence nature of the X-ray focusing optics, when coupled with the relatively short focal length of the telescope, meant that the energy band was substantially below the bandpass covered by the H1L1 measured fluxes. A comparison of results obtained in the two different bandpasses therefore depends on the spectral properties of the sources studied. In addition, the lower energy band of _Einstein_ is more sensitive to self-absorption of sources, common in low luminosity AGN. In principal, spectral data from the _Einstein_ observations could be used to estimate the self-absorption, however many of the sources studied are of such low flux that the interpretation of the spectral data is ambiguous. In this section we will review the N(S) relation derived from _Einstein_ studies. Though the _Einstein_ results are justifiably celebrated for the information they provide about the low flux sources, we will here examine the unavoidable ambiguity of spectral effects, particularly when analyzing the origin of the total X-ray sky flux observed at higher energies.

Most source count studies from _Einstein_ used the Imaging, Proportional Counter (IPC). A first order analysis relating broad band

Analysis of Results                              Source counts from _Einstein_

flux measurements obtained with the IPC to those from HEAO 1 is outlined in Section A7:2. Although we cannot provide total accuracy in the conversion, the general trends illustrated in this section are sufficient for our purposes.

### The Medium Survey: Maccacaro et al. [1982]

With the Einstein Observatory both galactic and extragalactic X-ray sources have been extensively studied. Most of these objects were targeted on the basis of their properties at other wavelengths, particularly in the optical and/or radio. Selection biases make it difficult to use such studies to directly assess the contributions of various classes of objects to source counts and the all sky flux. Ideally we should use an X-ray selected collection of objects. One such sample is the Medium Sensitivity Survey of Maccacaro et al. [1982], hereafter referred to as the Medium Survey or MS. The Medium Survey resulted from the analysis of over 110 high latitude IPC fields for serendipitous sources. By serendipitous we mean the source was discovered by virtue of its fortuitous presence in the same region of sky as the (ususally much brighter) targeted source.

Source fluxes were reported in Maccacaro et al. in ergs $s^{-1}$ $cm^{-2}$ from 0.3 to 3.5 keV. These fluxes were calculated assuming each source had a spectrum with energy index $\alpha = 0.4$ and absorption only due to matter in our galaxy. The column density of hydrogen used was typical of high latitude observations out of the galaxy: $N_H = 3 \times 10^{20}$ $cm^{-2}$. However, we wish to make comparisons of the observed counts with predictions from populations with greatly differing spectra, i.e. clusters of galaxies, unabsorbed AGN, and absorbed AGN. To do this properly we should fold the various source spectra through the detector response and make the comparison in terms of observed counts. Therefore we convert the ergs back to IPC counts over the equivalent 0.3 to 3.5 keV energy band. Such units are indicated by IPCMS. To make the back conversion, our analysis of the IPC spectral response (see Section A7:2) produces a conversion factor of 1 IPCMS = $3 \times 10^{-11}$ ergs $s^{-1}$ $cm^{-2}$ (0.3-3.5 keV)[*]. For an incident power law spectrum with energy index $\alpha$=0.4, 90%

Analysis of Results                                         Source counts from Einstein

of the photons in an IPCMS measurement originate from the range $\Delta E_{90}$ = 0.5 to 2.6 keV (see the discussion following equation [II.25]). This particular power law index is a good approximation in the IPCMS energy range to an extrapolation of the 40 keV thermal bremsstrahlung spectrum characteristic of the X-ray sky flux observed at the higher energies of HEAO 1. Recall that the $\Delta E_{90}$ of the HlLl window is 2.5-13.3 keV for such a spectrum.

The different exposures that made up the Medium Survey constituted ~50 square degrees in individual observations distributed over the sky with varying levels of sensitivity, depending on the total active time and internal background rate of the particular observation. The lowest flux detectable at the $5\sigma$ level was 0.0023 IPCMS, while the median sensitivity was 0.016 IPCMS. In order to concentrate on the behavior of N(S) at the lower fluxes, sources brighter than 0.17 IPCMS were explicitly excluded from the sample. 63 sources were detected, 51 of which had positive identifications based on optical spectroscopic work. Forty of these were identified with extragalactic objects: 31 AGN, 7 clusters, 1 BL Lac, and 1 normal galaxy. Of the unidentified objects, 8 are tentatively identified as extragalactic, as their ratio of X-ray flux to optical flux is significantly larger than is typical of galactic sources.

As was done by Piccinotti et al. [1982] for the XCS, the Medium Survey source fluxes were used to fit a power law index for N(S), following the prescription of Murdoch, Crawford, and Jauncey [1973]. The best fit value and 1 $\sigma$ range for the power law index $\gamma$ is 2.53 (2.37-2.69). Application of the K-S test to the distribution of fluxes showed that single power law models were consistent with the data. The normalization of the N(S) law is not directly obtainable by the technique used. A least squares fit of the sources to the given best fit power law yielded a value for K of 5.2 $(IPCMS)^{1.53}$. Though a

---

*The conversion from ergs to counts used by Maccacaro is 1 IPCMS = $2.5 \times 10^{-11}$ ergs $s^{-1}$ $cm^{-2}$ [private communication]. The discrepency between these two numbers indicates the precision attained for much of the following discussion of spectral dependencies.

Analysis of Results                              Source counts from Einstein
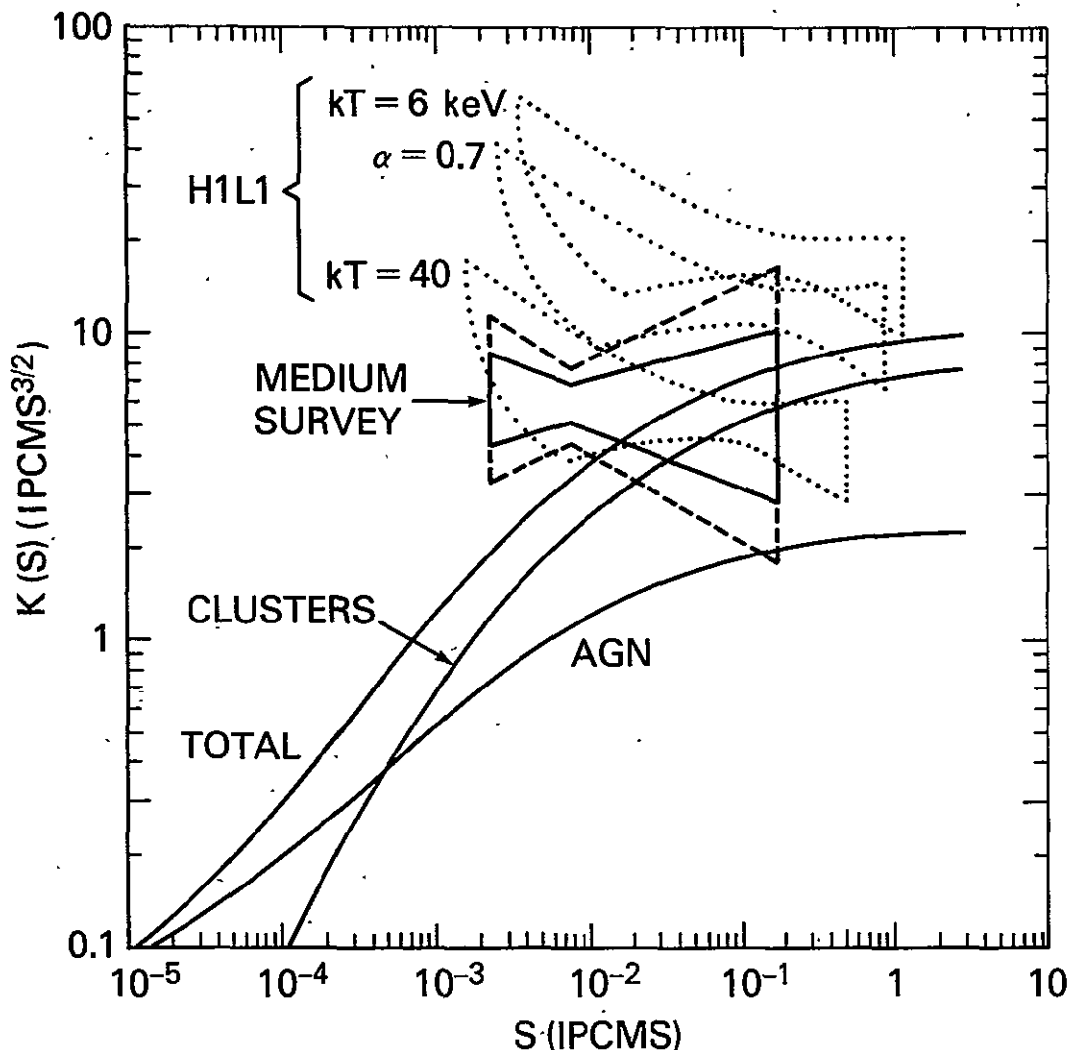
Figure VII.24: K(S) for the Einstein Medium Survey

The solid and dashed regions are the 50% and estimated 90%
confidence regions for the Einstein Observatory. Medium
Survey. The three dotted regions marked "H1L1" are the
regions from the analysis of the HEAO 1 A-2 fluctuations
(assuming a single power law model N(S)), transposed to
the IPCMS flux units assuming that all the sources of the
fluctuations had the indicated spectra. The overlap of
the kT = 40 keV H1L1 contour and the Medium Survey contour
is not evidence that the source spectra are 40 keV thermal
spectra. See the text for further important
considerations in comparing these contours and the MS
contours. The K(S) curves are from the luminosity
functions derived from the XCS of Piccinotti et al.

Analysis of Results                          Source counts from Einstein

| [1982], with no evolution. |

confidence range in K cannot be derived from this estimate, the normalization can be determined no better than the total number of sources in the sample, which is uncertain due to Poisson statistics at the very least. Thus, since the Medium Survey consists of 48 confirmed or suspected extragalactic objects, the 1 σ range in K is at least ±14% = $48^{-1/2}$. Given the normalizations for the extreme allowed values for γ, Maccacaro et al. report that the associated N(S) curves cross at $S_x$ = $7.5 \times 10^{-3}$ IPCMS. In Figure VII.24 the solid contour labeled "Medium Survey" outlines the K(S) for these models, assuming that the normalization and the power law slope are decoupled at the crossover flux. Since the ranges in γ and $K(S_x)$ were from their separate 1 σ confidence regions (i.e. 68%), the solid contour has a 46% chance of containing the correct N(S) model, assuming that it is indeed a single power law over the given range. The dashed outline is an estimate of the 90% region, formed by scaling the given 1 σ ranges for γ and K to the equivalent 95% ranges (1.96 σ) of 2.22–2.84 and ±28%. The upper and lower flux limits in the contours are based on the total range of sensitivity of the Medium Survey. The actual behavior at either extreme is poorly constrained due to comparatively little data at these points.

The three dotted regions in Figure VII.24 marked "H1L1", derived from the single power law fits to the A-2 fluctuations data, illustrate the difficulty of making direct comparisons of results derived from A-2 and the IPC. The fluctuations result used is the single power law fit to the WW dataset. There has been no inclusion of information from the resolved source counts from the XCS, which would favor the steeper (negative slope) models. The three regions follow from assuming that all the sources which produce the fluctuations have a different, single, spectral form: a power law of index α = 0.7, thermal bremsstrahlung with kT = 6 keV, or one with kT = 40 keV. There is no simple model-independent way to give similar contours when the sources have different spectral characteristics, but crudely we can imagine that the behavior is some average of the given contours. The flux range that the fluctuations analysis is most sensitive substantially overlaps the range

Analysis of Results                                          Source counts from Einstein

covered by the Medium Survey. When comparing the values for K(S) from
the two experiments we might argue that, except for the kT=40 keV
spectrum, there is an indication of a slight rollover. Keep in mind,
however, that the fluctuations include as large as a ~13% contribution
from cataclysmic variables and other galactic sources while the Medium
Survey region explicitly excludes galactic sources. Also, the
contribution of the absorbed low luminsity AGN is much larger in the
H1L1 band than in the IPCMS measurements, as we will see in more detail.
Therefore, a comparison of the H1L1 contours and the Medium Survey
contours can not be used to conclude either the existence of a rollover,
or that the sources have a kT = 40 keV spectrum, as such conclusions
likely are artifacts of the inexact spectral correspondence between the
two surveys.

The uncertainty due to spectral band effects can be reduced, or at
least explicitly addressed, by calculating N(S) for source populations
with specific spectra. The curves in Figure VII.24 plot the associated
K(S) for the luminosity functions derived from the XCS, assuming no
evolution. The spectra for clusters of galaxies are the luminosity
dependent forms outlined in Table VII.3. The AGN curve is for the
canonical luminosity function (see Table VII.5) with the standard
spectral model: $\alpha$=0.68, $N_H$=3.6x10$^{22}$ cm$^{-2}$ for $L_{44}$<0.2. All observations
are assumed to have an additional absorption of $N_H$=3x10$^{20}$ atoms cm$^{-2}$ in
the observer frame, i.e. not subject to the effects of cosmological
redshifting. This amount of absorption is typical of the column density
due to gas in our galaxy for high galactic latitude observations, and is
important for the lower energy bandpass of the IPC but is undetectable
in the H1L1 band. The "total" curve in Figure VII.24 generally falls
far below the acceptable range for the Medium Survey, which argues
directly for some amount of evolution. The composition of the Medium
Survey also suggests evolution is necessary. One aspect of the change
from the higher energy bandpass of H1L1 to the Medium Survey is seen in
the K(S) calculated for clusters. At the XCS fluxes in the high energy
bandpass the numbers of clusters and AGN were roughly equivalent. At
the analagous flux levels in the the low energy bandpass the predicted
ratio of clusters to unevolved AGN was near 3. In the high energy

bandpass the cluster contribution to N(S) quickly flattens from the Euclidean 5/2 index power law. In the low energy bandpass this occurs at a lower flux, as higher redshifts are required to produce the same reduction in the observed flux. In fact the N(S) curve for clusters does not fall below the N(S) for unevolved AGN until S $\lesssim 10^{-3}$ IPCMS. From the relative strength of the two curves for AGN and clusters in Figure VII.24 we can crudely estimate that the ratio of clusters to AGN in the Medium Survey should be 2 to 1. Instead, clusters are significantly outnumbered by the AGN in the Medium Survey, constituting 17% of the sample, and not 67%. This indicates that some evolutionary change from the XCS-derived luminosity functions has taken place, producing the ratio observed in the Medium Survey.

This argument for AGN evolution depends to a large degree on our model for the AGN spectra, particularly that most of the low-luminosity AGN have significant intrinsic absorption, which reduces their contribution to N(S) at a given flux. If the effects of the absorption are reduced, then the AGN N(S) curve is increased with respect to the cluster AGN, and the required amount of evolution is decreased. On the other hand, the case for evolution is strengthened if there are substantial numbers of low temperature clusters, kT$\lesssim$3 keV, which would not contribute significantly to the luminosity function determined by high and medium energy detectors of HEAO 1 A-2. The expected number of clusters is similarly enhanced by the existence of a low temperature component in some clusters, which has a contribution to the flux in the low energy bandpass comparable to the contribution from the high temperature component in the high energy bandpass [e.g. Mushotzky et al. 1978; Pravdo et al. 1979; Mitchell and Mushotzky 1980; Reichert et al. 1981]. Maccacaro et al. had also concluded that the low fraction of clusters indicated evolution, but only by direct comparison to the fraction in the XCS resolved source counts, with no corrections for redshift or the differing bandpasses.

To produce the observed ratio of AGN to clusters in the MS we could either increase the number of AGN and/or decrease the number of clusters with redshift or look-back time. In fact, as illustrated in

Figure VII.25: K(S) for clusters in the Medium Survey

The curves compare K(S) for the canonical XCS cluster luminosity function with cutoffs at $z_{max}$= 3.0 and 0.75, as well as a model with $z_{max}$ at 3.0 but also negative density evolution, $Q_\rho$=-2. $q_o$ = 0.1.
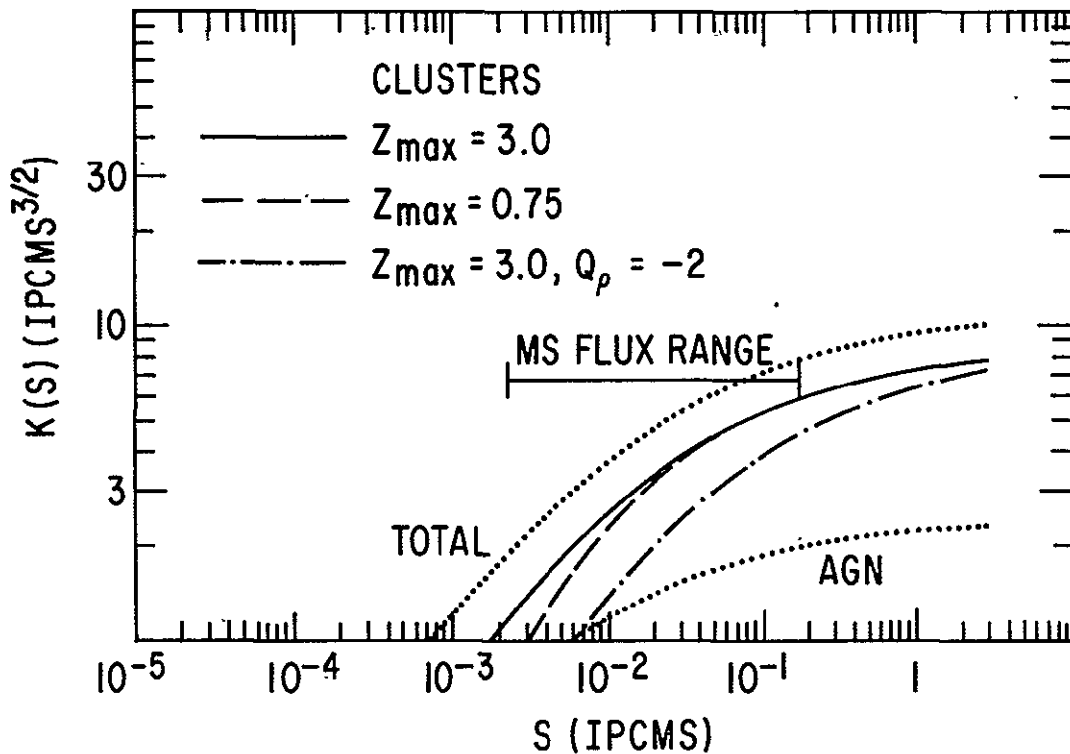


Figure VII.25, the study of clusters over the flux range of the Medium Survey may yield useful information about the evolution of the cluster luminosity function. However, 7 clusters are too few to lead to any strong conclusions; an extension of the serendipitous objects survey to include more observations is needed to increase the number of clusters.

Limits on AGN evolution from the Medium Survey

In Figure VII.26 we estimate the form for K(S) when the AGN are allowed to evolve following the models discussed in the previous section, with $q_o$=0.1 and evolution restricted to the medium luminosity AGN only. For this choice of assumptions we see that the source counts

Analysis of Results                    Source counts from Einstein

Figure VII.26: K(S) from AGN evolution compared to Medium Survey results

The Medium Survey derived 47% and 90% confidence ranges for single power laws are indicated by the solid and dashed regions. Only medium luminosity AGN, $0.1 < L < 15$ $L_{44}$, are subject to evolution and $q_o = 0.1$. (a) Pure luminosity evolution. (b) Index evolution, $L_o = 0.5$ $L_{44}$.
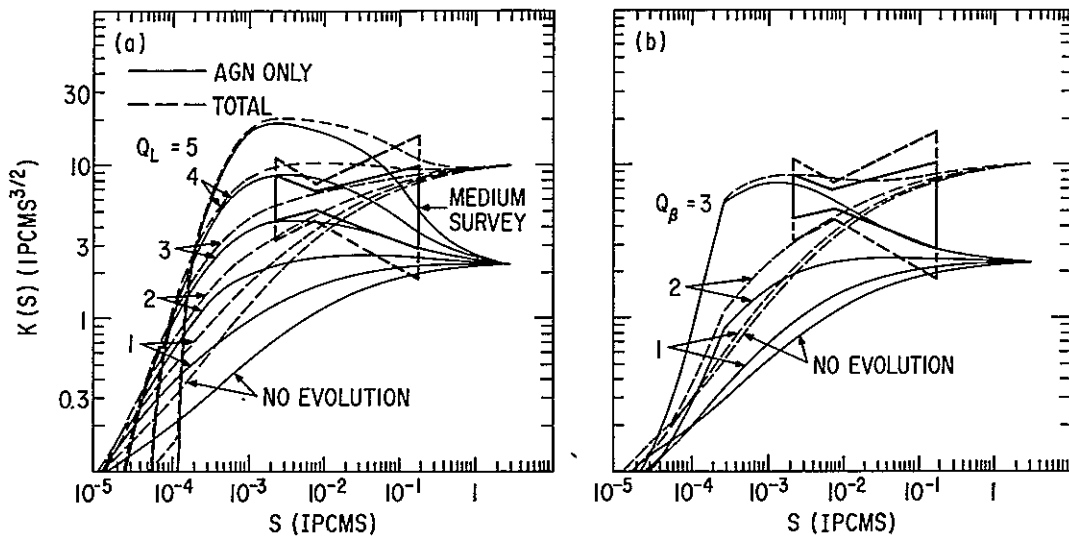


of the Medium Survey provide a significant restriction on the amount of evolution allowed. From VII.26(a) the preferred range in $Q_L$ (pure luminosity evolution) is from 2 to 4, while by VII.26(b) the range in the index evolution parameter is $2.0 \leq Q_\beta \leq 3.1$.

Later work by Maccacaro et al. [1983a] used the redshift and luminosity data of the MS, instead of the source counts, to explicitly test for pure luminosity evolution. Their best fit value for the evolution of $Q_L$ was 5.7 with a 90% confidence range of roughly 4 to 6. Although they assumed $q_o = 0$, which indicates that their values should be slightly decreased to compare with our evolution models where $q_o = 0.1$, these results are generally larger than what appears to be allowed by the source counts as displayed in Figure VII.26(a). However, there is a substantial difference in the underlying model for the luminosity

Analysis of Results                              Source counts from Einstein

function that makes direct comparison difficult. Maccacaro et al. [1983a] assumed that both low and high luminosity objects evolved. The limits derived when $Q_L$ has its best fit value are $L_{min} \approx 0.01\ L_{44}$ and $L_{max} \approx 2\ L_{44}$. (All luminosities discussed here are derived from the values for the differential luminosity at 2 keV reported by Maccacaro et al. [1983a] assuming an $\alpha=0.65$ power law spectrum and integrating from 2 to 10 keV.) The extension of the evolution to lower luminosities makes a relatively small difference to the source counts in the Medium Survey range, in part because of their high absorption, but also because the of evolution of the low flux sources does not have a large effect except at the lowest fluxes (see Figure VII.18). However, in the H1L1 band, evolution at these levels by the low flux sources has a great impact on the contribution to the sky flux, severely overproducing the total sky flux. The value of $L_{max}$ is more important in determining the behavior of N(S) over the range of the Medium Survey. To provide agreement with the low redshift, high luminosity, AGN observed in the XCS, $L_{max}$ should be extended to at least ~15 $L_{44}$. When $L_{max}$ is increased to ~100 $L_{44}$ Maccacaro et al. [1983a] report that the best fit $Q_L$ drops to 4.2, although this model then fails the goodness of·fit criterion, based on a K–S test of the model. The poor fit could be caused by their assumption that the luminosity function of AGN is a single power law from $L_{min}$ to $L_{max}$. The observations of the XCS with the preliminary results of Elvis, Soltan, and Keel [1983] indicate that this is not the case at the lower luminosities. We recall from the last section how the luminosity function might be extended from 15 to 100 $L_{44}$, but only if the evolution were slight enough ($Q_L \lesssim 4$) to avoid predicting significant numbers of high luminosity sources in the XCS.

We have made different spectral assumptions than Maccacaro et al. [1983a], which create additional difficulties when comparing their fits to the limits we derive from Figure VII.26(a). They assume that all sources are described by the same spectral form: a single power law with index $\alpha=0.5$ and absorption only from our galaxy. Strong absorption below 0.2 $L_{44}$, as we have assumed, will produce a strong break in the luminosity function when measured in the IPCMS bandpass without correcting for that absorption. Although only the four lowest redshift

Analysis of Results                                        Source counts from Einstein

objects of the Medium Survey have $L < 0.2\ L_{44}$, this effect can have significant impact on the best fit values for $Q_L$, as well as on the

Figure VII.27: Effect of absorption in AGN on the K(S) in the Medium Survey

AGN are separated into low and middle luminosity ranges. For purposes of illustration, the low range sources have either a value for $N_H$ typical of the low luminosity XCS sources, or (arbitrarily) half that absorption. Most of the sources in the middle range have comparatively little absorption. $q_o = 0.1$, no evolution.
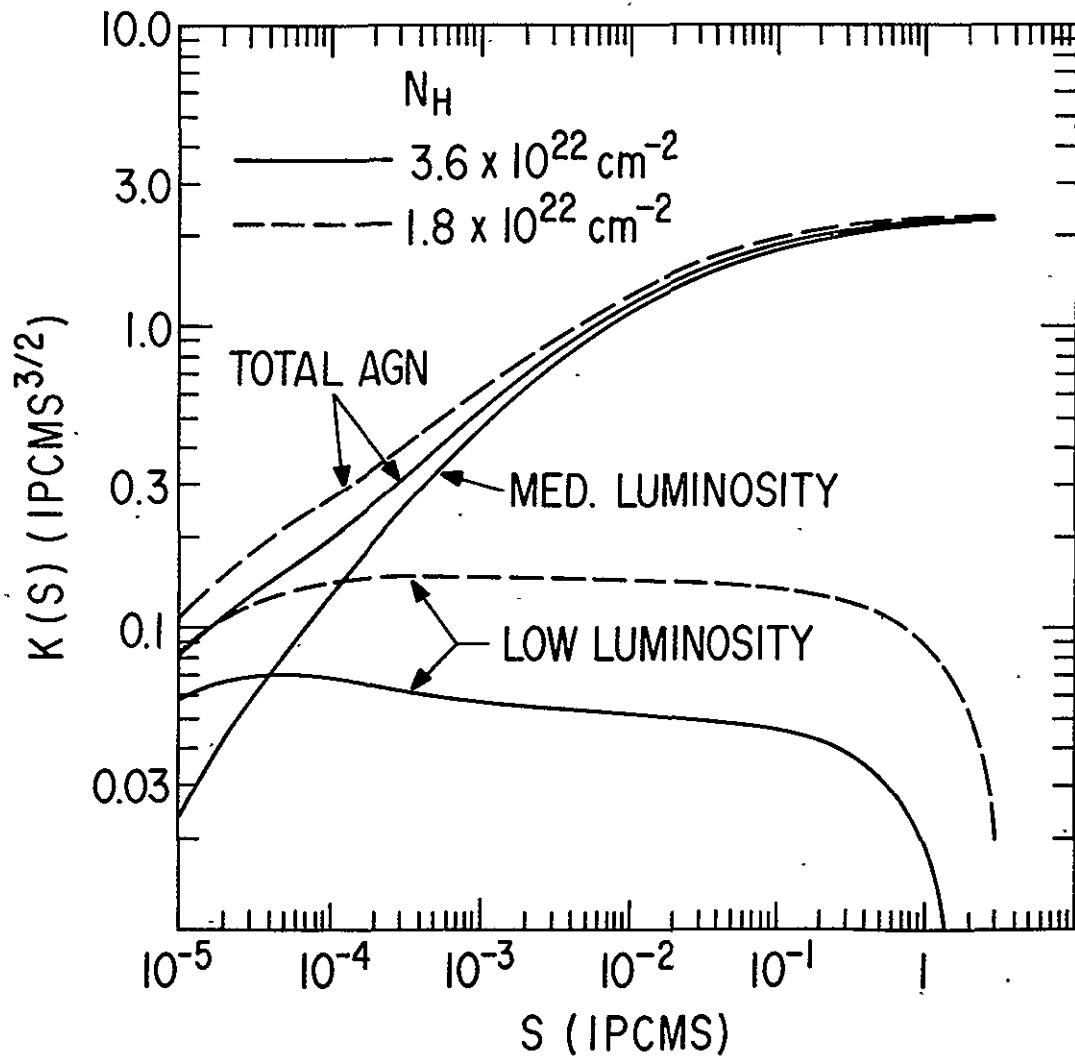


goodness of fit. Assumptions about absorption have just as strong an

Analysis of Results                    Source counts from Einstein

impact on our estimates of the source counts. We have assumed that the
low luminosity objects did not evolve, but modifications in the amount
and form of the absorption can have a significant impact on the
predicted source counts. As seen in Figure VII.27, those sources which
are absorbed make almost no contribution to the source counts, at least
over the flux range of the Medium Survey. Even halving the amount of
absorption makes little difference. Therefore, moving the critical
luminosity that divides absorbed from less absorbed sources can produce
a large change in the predicted numbers of AGN. The leverage of this
effect is generated by the steepness of the AGN luminosity function.
Thus if a significant fraction of the AGN above 0.2 $L_{44}$ are absorbed,
then larger amounts of evolution would be consistent with the numbers of
the Medium Survey and the XCS. Though Lawrence and Elvis [1982] infer
significant absorption for a few moderate luminosity AGN, their
conclusions are based on comparisons of non-contemporaneous measurements
of known variable sources. Variablity could easily mimic the effects of
high column density, though not systematically.

Probably absorption is not a simple on/off aspect of AGN
luminosity. Holt et al. [1980] have proposed a model for the
absorption in NGC 4151 which assumes a partial covering factor, that is,
a fraction of the flux comes through essentially unabsorbed while the
remainder is absorbed. The fraction depends on a detailed interaction
between the X-ray continuum luminosity and the geometric and density
characteristics of the clouds that produce the broad emission lines
typical of many AGN (see Ferland and Mushotzky [1982]). Measurements
made in the bandpass of the A-2 MED and HED detectors can not easily
distinguish a gradual change in the covering fraction from a sudden
onset of absorption. For example, in A-2 spectra, a source with 80%
covering fraction is indistinguishable from one totally covered with
slightly lower column density, while a source with 20% covering fraction
may be consistent with no absorption. In the HILl window, 20% coverage
by $N_H = 3.6 \times 10^{22}$ $cm^{-2}$ material reduces the flux slightly to 97% of the
unabsorbed flux; with 80% coverage it is reduced to 89%; and with full
coverage to 86%. Thus even total coverage for an individual source is
difficult to measure with the HED detectors, although this ability can

Analysis of Results                              Source counts from Einstein

be enhanced when the higher precision PHA spectra are analyzed. In the
IPC bandpass, however, there are strong differences between changes in
the partial covering fraction and a sudden onset of total coverage.
Even 20% coverage by material with $N_H = 3.6 \times 10^{22}$ cm$^{-2}$ material produces
a significant absorption, reducing the flux to 82% of its unabsorbed
value. Full coverage reduces the flux to 11%, yet even partial covering
fractions of 80% will have fluxes almost three times that, i.e. 29% of
the unabsorbed flux. The partial covering fraction model makes it
important to have high quality broad band spectra of the very low
luminosity objects, such as M81, in order to properly compare their
luminosity function derived from IPC fluxes by Elvis, Soltan, and Keel
[1983] to the higher bandpass luminosity function for the moderate
luminosity AGN derived from the XCS. On balance, the expected impact of
a partial covering fraction model for the absorption is to allow more
contribution to the source counts from sources below 0.2 $L_{44}$. This
reduces the allowed values of $Q_L$ for pure luminosity evolution below
those indicated by Figure VII:26.

Not all AGN have simple $\alpha \sim 0.7$ power law spectra. Some have
additional, steep ($\alpha \gtrsim 2$) spectral components which dominate at low
energies [e.g. Pravdo et al. 1981]. The existence of these AGN tends to
reduce the amount of evolution of the XCS luminosity function that is
compatible with N(S) derived from the Medium Survey. Though these
phenomena may not be common, the possible existence of such soft
components and potentially complicated effect of the absorption on low
energy spectra indicate the inherent difficulties of extrapolating low
energy results to higher energy regimes. Detailed spectral studies in
the low energy band may help [see e.g. Holt et al. 1983] but one would
ideally be a study of sources of the Medium Survey flux range in the
higher energy A-2 band. As AGN appear well fit there by the single
power law spectrum, $\alpha \sim 0.7$, this band provides a cleaner environment
for statistical studies. Still, though the detailed aspects of the
typical spectra of AGN, their luminosity function, and its evolutionary
form are poorly known, the very presence of a number of AGN
substantially in excess of that predicted from the XCS determined
luminosity function is a persuasive argument for significant evolution

Analysis of Results                                              Source counts from Einstein

of AGN over cosmological timescales.

## Evolution of BL Lac objects

BL Lacertae objects are currently thought to be another form of active galactic nucleus, but are usually considered separately from the more traditional AGN such as Seyfert galaxies and quasars. Comparisons among these objects may help determine what physical processes are responsible for the observed differences between them. It has been argued on the basis of source counts versus flux in the optical [e.g. Woltjer and Setti 1982; see also Véron 1979] that BL Lacs undergo little, if any, evolution in contrast to the behavior of quasars and other AGN. The optical source counts may be significantly incomplete, however. Maccacaro et al. [1982] argue that comparison of the relative numbers of BL Lacs and AGN in the Medium Survey and the XCS is evidence based on complete X-ray slected samples for a similar conclusion that BL Lac objects do not evolve in comparison to other AGN. We examine this claim and find that in the light of subsequent identifications in the XCS there is no statistically significant difference between the two samples.

We now combine the previously separate catagories of AGN and BL Lac into an All Active Galactic Nuclei (AAGN) designation. Let P be the probability that an AAGN will be classified as a BL Lac. In the XCS, 4 of 34 AAGN are BL Lacs. Thus the measured value for P in the XCS is $P_X$ = 0.118. In the Medium Survey only 1 of 32 AAGN was a BL Lac, so that $P_M$ = 0.031, substantially less than $P_X$. Maccacaro et al. [1982] argued that this is expected if BL Lac sources undergo less evolution than the AGN. If BL Lac sources evolve as much as AGN, then $P_X$ and $P_M$ should be the same. The key issue in this argument is whether $P_M$ and $P_X$ are significantly different. Maccacaro et al. calculated that with no evolution, and extrapolating from the results of the XCS, only ~ 1 BL Lac would be expected in their survey. Thus, their data are consistent with the no evolution hypothesis. The question remains whether their data explicitly rule out BL Lac evolution similar to that of the AGN.

Analysis of Results                                        Source counts from Einstein

We will calculate a statistical measure of the null hypothesis that in going from the XCS to the Medium Survey there has been no change in the parent distribution P, of which $P_X$ and $P_M$ are individual estimators. If we construct the ratio

$$r_1 = P_M \; / \; P_X \; , \qquad\qquad\qquad\qquad [VII.24]$$

for no evolution the expected value of $r_1$ is 1. Statistical variations in the observed quantities $P_M$ and $P_X$ allow $r_1$ to have other values, thus the test of the null hypothesis reduces to estimating for a given value of P the probability that $r_1$ would be less than or equal to the observed value of 0.266. The most likely value for P is derived from the combined set of objects in the XCS and Medium Survey, P = 5/66 = 0.076. Based on the statistics of the binomial distribution applied separately to the AAGN of the two surveys, $r_1$ is <0.266 13% of the time. $r_1$ is .actually a "one-sided" statistic, in that we assumed that the only interesting change in P was a reduction in going from the XCS to the lower flux (and more evolution sensitive) Medium Survey. If we also consider excursions in the estimators of P where $P_M$ might be <u>larger</u> than $P_X$ we should use a two-sided test based on

$$r_2 = \min\{P_M, P_X\} \; / \; \max\{P_M, P_X\} \; . \qquad\qquad [VII.25]$$

For our numbers, $r_2$ has the same value as $r_1$. However, under the null hypothesis with P=0.076, we would expect to see values of $r_2$ this low or lower 25% of the time.

Of course, as we have repeatedly pointed out, there are substantial differences in measurements made in the respective bandpasses of the two surveys. A major difference is that low luminosity AGN are under-represented in the Medium Survey, most likely by virtue of substantial absorption. However, even if we restrict the sample to AGN with L>0.2 $L_{44}$, which are expected to have negligible absorption, we still do not find a statistically significant change in the fraction of BL Lac between the two samples. For this restricted population the fraction of the Medium Survey becomes 1 BL Lac out of 28 moderate or high luminosity AAGN, $P'_m$ = 0.036. At the present time we do not have luminosities for all the newly identified AGN in the XCS so we cannot clearly divide the entire sample into luminosity classes L>0.2

$L_{44}$ and $L < 0.2\ L_{44}$. For the objects with known luminosities, 14 are in the first class and 10 are in the second, so if the unknown luminosities have a similar distribution, the revised $P'_x \approx 4/24 = 0.17$, yielding a value of $r_1 = 0.214$. The most likely value for the null hypothesis constant fraction of high luminosity AAGN that are BL Lacs is $P' = 0.097$. Given this value, the one and two-sided probabilities for the measured r value are 9% and 18%. If we were to restrict our consideration to even higher luminosity AGN, as is done in the discussion of Maccacaro and Gioia [1983b], statistically convincing numbers might be found for rejecting the null hypothesis. It is difficult, however, because of the evolution of the AGN luminosity function, to match a sub-population of AGN in the Medium Survey with the corresponding sources in the XCS. The high luminosity objects in the Medium Survey are all at appreciable redshifts.

An additional effect, pointed out by Maccacaro et al. in their comparison of the two surveys, is that the steep spectra typical of BL Lacs should make them significantly easier to detect in the low energy bandpass of the IPC than in the XCS. We have not considered this aspect in our analysis. However, the simple examination of the number of BL Lacs in the Medium Survey compared to the XCS is not sufficient to rule out similar amounts of evolution for both BL Lacs and AGN. Neither is there evidence that BL Lac objects do evolve. Models where BL Lac sources undergo no evolution are also consistent with the Medium Survey and XCS source counts.

### The Deep Survey: Giacconi et al. [1979]

The imaging experiments of the Einstein Observatory were also used to make long exposures of selected fields, in contrast to the serendipitous sources of the Medium Survey which were achieved by generally shorter exposures. The hope was that the deep exposures could resolve the sources that produce the bulk of the X-ray all sky flux. Analysis of these deep exposures for the low flux form of N(S) is given in Giacconi et al. [1979], hereafter called the Deep Survey. IPC flux measurements were made over a different band than was used for the

Analysis of Results                                    Source counts from Einstein

Medium Survey. The nominal range was 1 to 3 keV; however, errors in the original calibration meant that the channels used probably included significant response to higher energy photons. The $\Delta E_{90}$ for an $\alpha=0.4$ power law with low absorption is 1.1-3.1 keV (see Section A7:2). Fluxes from the Deep Survey will be quoted in IPCDS units, 1 IPCDS equals 1 count per second in the energy range of the sample. The conversion between IPCDS and IPCMS measurements is spectrally dependent. For $\alpha$ = 0.4 and galactic absorption, 1 IPCDS = 1.4 IPCMS. As $\alpha$ ranges from 0 to 1 the conversion factor varies from 1.3 to 1.7. For strongly absorbed sources, $N_H \gtrsim 3.6 \times 10^{22}$ cm$^{-2}$ with complete covering, the conversion is essentially 1 for all spectra with $\alpha \lesssim 2$. Comparing H1L1 measurements to the Deep Survey is difficult for the same reasons that were described above for the Medium Survey comparison, but to a lesser degree because of the higher energy bandpass of the Deep Survey.

The Deep Survey sample was derived from observations of two separate fields, one in the constellation Eridanus, the other in Draco. The total solid angle covered was 0.17 msr or $1.38 \times 10^{-5}$ of the total sky solid angle. The 5 $\sigma$ sensitivity level was different for the two fields: $1.0 \times 10^{-3}$ IPCDS for Draco and $1.5 \times 10^{-3}$ for Eridanus. Fourteen sources satisfying the 5 $\sigma$ significance criterion were found in these fields. Four of these were identified as stars on the basis of their optical spectra, and were eliminated from further consideration. Each of the ten remaining objects has several suggested identifications with a variety of objects, but are expected to include few galactic sources. Seven of the ten were brighter than $1.5 \times 10^{-3}$ IPCDS. Because of the reduced solid angle coverage for the remaining three sources (only one field was sensitive to these low fluxes), their contribution to the number of sources expected in the total solid angle is doubled, yielding an estimate of 13 objects with flux greater than $10^{-3}$ IPCDS in a solid angle of 0.17 msr. Giacconi et al. further estimate that source confusion and the Malmquist effect should produce a bias of ~20%, giving a best value of $6.3(\pm 2.6) \times 10^4$ sources sr$^{-1}$. The indicated range is an estimate of the 1 $\sigma$ (68%) confidence region of the number of sources.

Given this corrected number of sources per unit solid angle we can

Analysis of Results                                    Source counts from Einstein

Figure VII.28: K(S) for the Deep Survey

The solid open box is the one sigma range from the Deep
Survey counts, assuming a Euclidean form. The dashed
region is the estimated 48% confidence region if γ is not
restricted to 5/2. The two dotted regions are the
estimated 90% confidence regions from the Medium Survey,
assuming the indicated spectral index and low absorption.
The indicated K(S) curves are from the non-evolving
canonical luminosity functions derived from the XCS.
$q_o$=0.1.
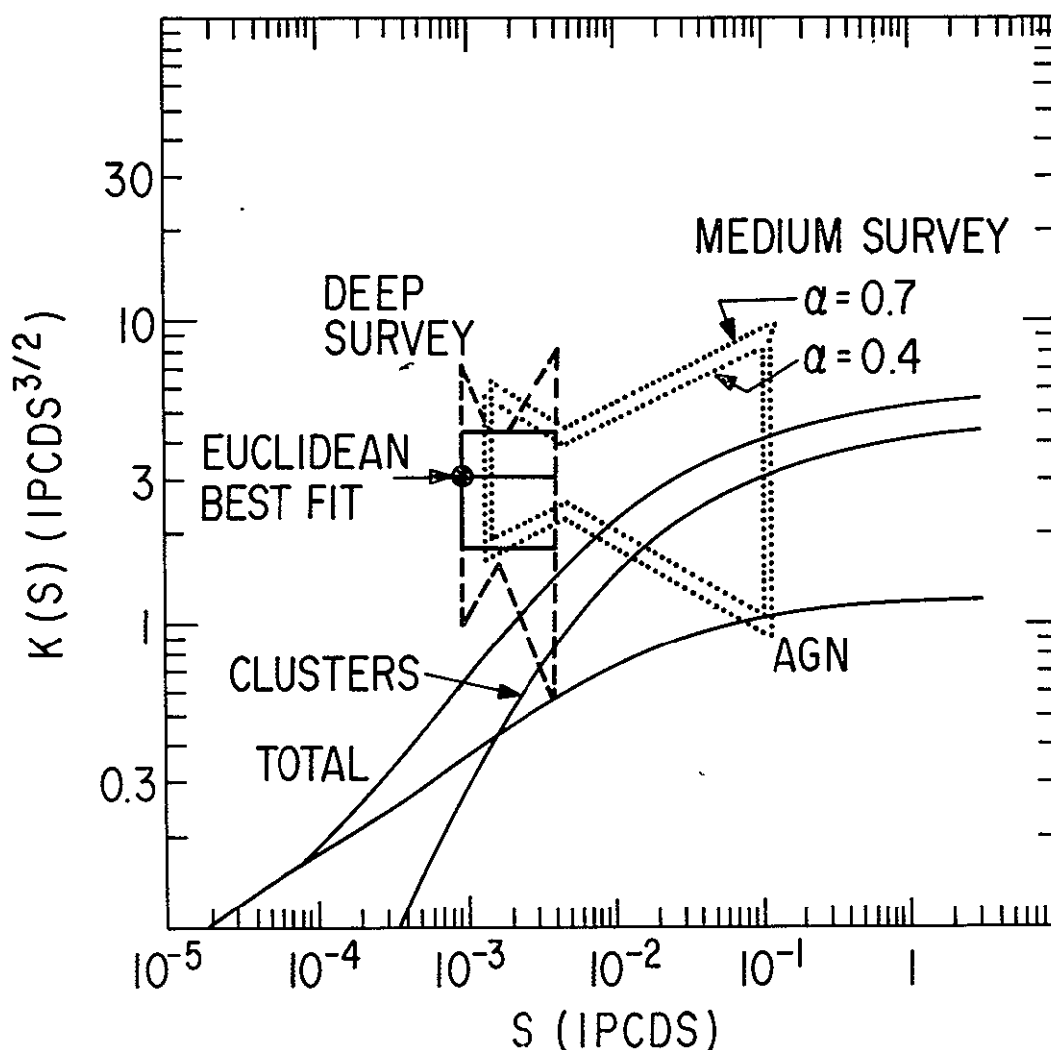


use equation [VII.9] to estimate the behavior of K(S) over the flux
range of the Deep Survey. This range is roughly from the lowest

Analysis of Results                    Source counts from Einstein

sensitivity limit of $10^{-3}$ IPCDS to the flux of the brightest source in the sample of $4.2 \times 10^{-3}$ IPCDS. Equation [VII.9] requires an assumption about the effective power law index of N(S) over the bin's flux range. In our analysis of the XCS source counts we assumed that $\gamma$ had the Euclidean 5/2 value based on the argument that most of the sources were near enough that the universe was still essentially Euclidean in geometry. Giacconi et al. estimated directly from the number of sources that they were consistent with a single Euclidean power law continuation of N(S) from the high flux studies of Uhuru and Ariel V. They then fixed the value of $\gamma$ at 5/2 for the remainder of their analysis, equivalent to the solid contour shown in Figure VII.28. Even ignoring the issue of the spectral complications in comparing source counts in different energy bandpasses, such an assumption is restrictive and unwise. In many of the model N(S) curves we have calculated, the effective power law index may undergo large changes over relatively short ranges in log S. This is particularly true of cases with significant index or pure luminosity evolution, where the roll-over effects of reaching the edge of the observable universe are concentrated in a rather narrow flux range.

Direct study of the fluxes of the ten probably extra-galactic sources could be used to constrain the N(S) index over the Deep Survey flux range, in a manner similar to the analysis of the Medium Survey and the XCS. Without an analysis of that sophistication, however, we can place a crude limit on the allowed range of values for $\gamma$, using the sources binned into a high and low flux range. If $N_h$ and $N_\ell$ are the number in each range, corrected for solid angle exposure, then the most likely value for $\gamma$ is

$$\gamma = 1 + \frac{\ln(1 + N_\ell/N_h)}{\ln(S_h/S_\ell)} , \qquad\qquad \text{[VII.26]}$$

where $S_h$ and $S_\ell$ are the lower flux limits of each range, $1.5 \times 10^{-3}$ and $1.0 \times 10^{-3}$ IPCDS respectively. For $N_h$ and $N_\ell$ of 7 and 6 respectively the corresponding value of $\gamma$ is 2.53, providing apparent confirmation of the suitability of the Euclidean form for the Deep Survey analysis. This

Analysis of Results                              Source counts from Einstein

includes no Malmquist or confusion correction to the numbers in each bin. As confusion is expected to be a larger problem for the lower flux bin, such a correction would lead to smaller values of $\gamma$.

The confidence range for $\gamma$, however, is quite large. We can crudely estimate the acceptable range for $\gamma$ by examining the log-likelihood for a binomial distribution of the ten sources into the two flux bins. The approximate 1 $\sigma$ range for $\gamma$ derived via this technique is 1.85 to 3.5. The dashed extension in Figure VII.28 is the estimated confidence region combining the possibility of variations in both $\gamma$ and the expected number of sources. As was the case for the similarly formed contours of the Medium Survey, this jointly defined region is approximately a 48% confidence region. This analysis does not fully exploit the information available from the Deep Survey in the way that using the unbinned data would. For example, Henry, Soltan, Briel, and Gunn [1982] were able in an analagous problem to determine a luminosity function index to within ~±18% with only 5 unbinned objects, versus the ±30% uncertainty in the N(S) index we achieved above with 10 objects binned. To forestall a simplistic scaling of our derived region, we note that their 5 objects had an advantage in dynamic range, being distributed over a decade in luminosity while the Deep Survey sources were restricted to a factor of ~4, or 0.6 decades, in flux. Thus the dashed region of Figure VII.28 is in some sense <u>at least</u> a 48% confidence region. Even without ascribing a specific confidence level to the region, however, it is clear that less than total knowledge about the power law index enlarges the range of allowed behavior for N(S).

The dotted regions in Figure VII.28 indicate the allowed regions derived from the Medium Survey. Even in moving from the closely overlapping bandpasses of the two IPC surveys, the assumed spectra of the sources has at least a small affect. We see that there is substantial overlap between the upper flux range of the Deep Survey and the lower end of the Medium Survey flux range. The precision of the Medium Survey appears even better than the Deep Survey, even with the $\gamma=5/2$ assumption, but only if N(S) is described by a single power law over the entire Medium Survey flux range. The model K(S) curves for

ORIGINAL PAGE IS
OF POOR QUALITY

Figure VII.29: K(S) from AGN evolution compared to Deep Survey results

The curves show the AGN and total K(S) for the indicated amounts of pure luminosity evolution for the middle luminosity AGN. $q_o$=0.1.
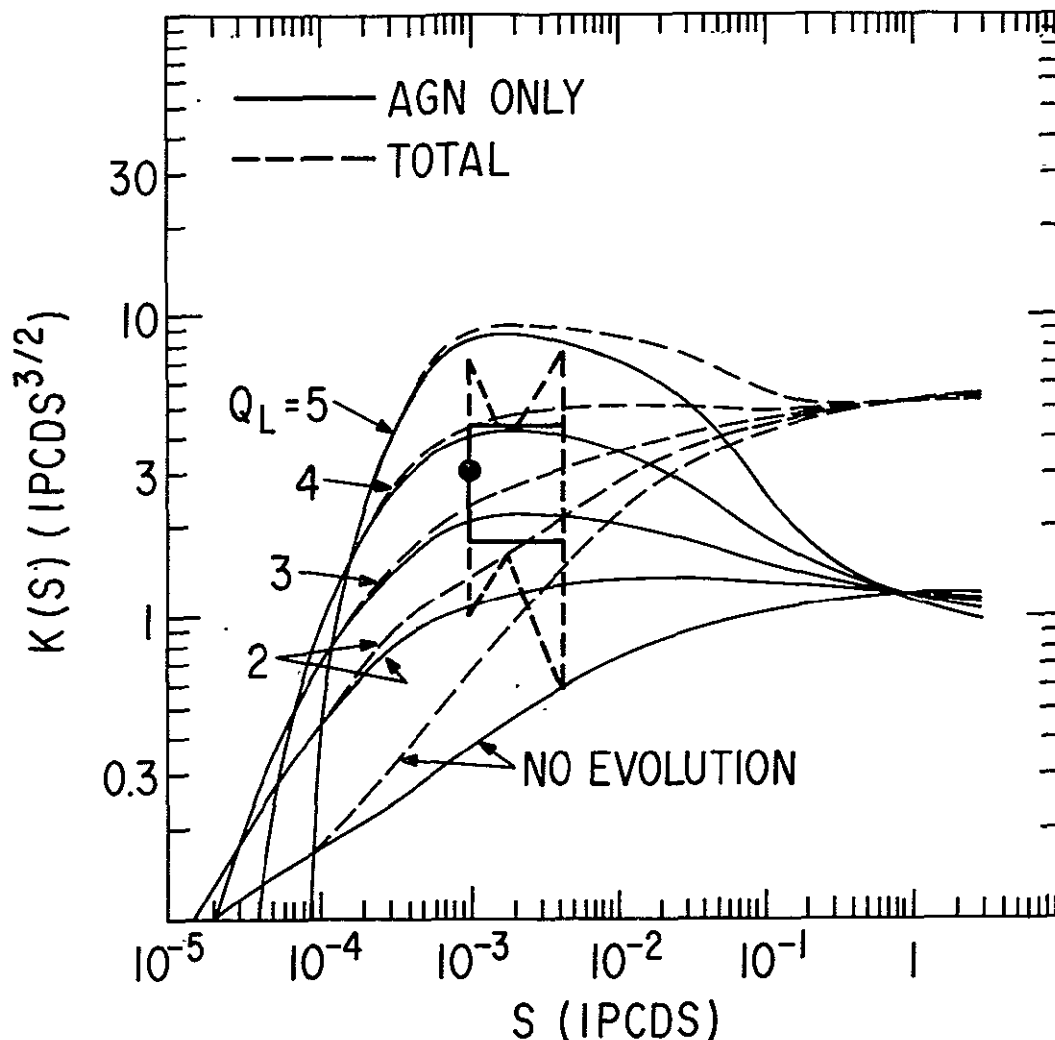


Figures VII.28 and VII.29 show that the limits on evolution from the Deep Survey are similar to the conclusions reached above from the Medium Survey. Few if any of the Deep Survey sources are expected to be clusters of galaxies, and significant evolution is required on the part of the AGN in order to make up the total observed source counts.

Analysis of Results                                    Source counts from Einstein

Many authors have used the Deep Survey results to estimate the minimum fraction of the total sky flux that is due to point sources, in contrast to models involving emission from a diffuse intergalactic medium. However, the value of this contribution is substantially model dependent. The usual calculation is made by integrating the derived Euclidean form N(S) laws from the limiting flux of the Deep Survey to infinity. This yields a total flux contributed over the Deep Survey solid angle of $S_{\Delta\Omega;model}$ = $33 \pm 14 \times 10^{-3}$ IPCDS. However, a model independent assessment, i.e. directly summing the fluxes of the Deep Survey sources, yields an unexpectedly lower value of $23 \times 10^{-3}$ IPCDS. This is equivalent to saying that the average source flux is less than what a Euclidean distribution of fluxes would predict. In fact the measured mean flux corresponds to the expected value for a $\gamma \approx 3$ power law N(S) model.

The IPC was unable to measure the total sky flux directly, and previous estimates of the fraction of the diffuse sky flux contributed by sources as studied in the Einstein spectral band have relied on extrapolations of the measured flux sky flux made in higher energy bands, using a particular assumed spectral form. From our sky flux of 58 HlLl in the high energy band and assuming a kT=40 keV spectrum, the extrapolated value in the Deep Survey bandpass is 9700 IPCDS. The Euclidean model and summed fluxes correspond to all sky fluxes of $2.4 \times 10^3$ and $1.7 \times 10^3$ IPCDS. The Deep Survey sources' summed contribution is 25%±10% (17% by direct summation) of the extrapolated flux. The danger of such extrapolations is that some significant modification in the makeup of the total sky flux may occur in moving to the lower energy bandpass. In fact, there are preliminary measurements from the LED detectors of HEAO 1 A-2 that indicate there is an excess in the extragalactic sky flux above the value obtained by simple extrapolation [Garmire, private communication]. This excess could be due to the contribution of resolved sources measured by the Deep Survey, in addition to some other component that produces the 40 keV thermal spectrum dominating the higher energy bandpasses.

We can invert the extrapolation problem. Rather than estimating

the total sky flux in the IPCDS band from higher energy measurements we convert the Deep Survey sky flux contributions to the higher energy bandpass units. Since the objects in the Deep Survey are dominated by significantly evolved AGN, we assume that all such sources are characterized by a single spectral form. The exact nature of that form is, of course, critical. If the sources have a kT=40 keV thermal spectrum, their conversion to the H1L1 band is the inverse of the calculation of the extrapolated IPCDS sky flux. The fraction of the sky flux is then the same, 25%, corresponding to 14 H1L1 out of the total 58 H1L1. (Henceforth we will use the best fit Euclidean model derived values for the Deep Survey sky flux contribution.) If the Deep Survey sources still have the 0.7 power law spectrum characteristic of unevolved AGN then their sky flux contribution drops to 9 H1L1 or 16%, asssuming high latitude galactic levels of absorption. If there are significant numbers of sources in the Deep Survey with soft excesses, such as are typical of BL Lacs, so that the effective spectral energy index rises to 1.0, then the sky flux contribution falls further to 6.3 H1L1, or 11%. Use of the summed fluxes contribution to the sky flux will produce even lower percentages. The Deep Survey may have directly imaged an important component of the total sky flux, but the largest fraction still remains to be explored.

Again, assuming that the Deep Survey sources are representative of a homogenous population with similar spectra, it makes more sense to convert the Deep Survey source counts to the H1L1 band for comparison with the fluctuations models. In Figure VII.30 we plot the Deep Survey derived counts for several assumed power law spectral indices. It clearly shows the wide range of possibilites for future low flux studies in the H1L1 band.

It cannot be disputed that the Deep Survey samples the behavior of N(S) at unprecedently low fluxes. Converting the lower flux limit of the Deep Survey into H1L1 units we see that the sample probes down to 6, 4, or $3 \times 10^{-6}$ H1L1 depending on whether the source spectrum is kT=40 keV, $\alpha=0.7$, or $\alpha=1.0$ respectively. These values are significantly below $20 \times 10^{-6}$ H1L1, the upper limit to the average flux of the population that

Figure VII.30: K(S) for Deep Survey converted to the H1L1 band

The Deep Survey ~47% confidence regions are shown assuming
.all sources in the sample have the identical indicated
spectral index. The four different plots were divided
between the (a) and (b) panels for improved legibility.
The figure also shows the best fluctuations contour for
the single power law N(S) model (dotted line) and the
broken power law limits from the combined fluctuations and
XCS resolved source confidence region for the excess
variance (long dashed line). The K(S) curves are the
cluster plus AGN contribution, $q_o$ = 0.1, assuming no
evolution and pure luminosity evolution ($Q_L$=4) for the
middle luminosity AGN. Remember that the fluctuations
results include a contribution from galactic sources.
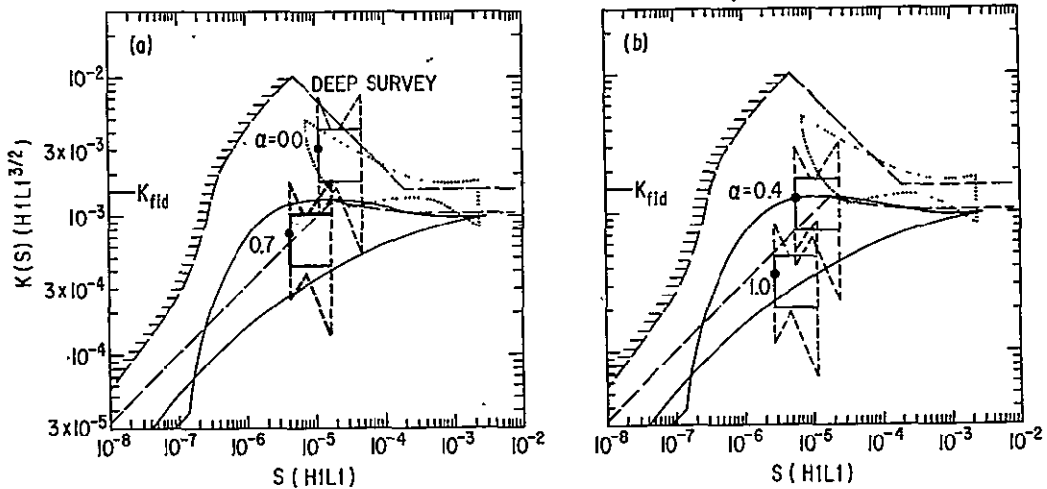


dominates the total sky flux, as derived from the upper limits to the
fluctuations' invariant excess. The Deep Survey results are potentially
stronger constraints on the sources that could contribute the remainder
of the total sky flux. We can place a lower limit on the all sky number
of sources

$$N^* \geqslant f \, S_{as} / S_{DS} ,$$           [VII.27]

where f is the remaining fraction of the sky flux produced by the
sources, and $S_{DS}$ is the Deep Survey lower flux limit. All these
quantities are evaluated for the H1L1 band, where we can directly

Analysis of Results                          Source counts from <u>Einstein</u>

measure $S_{as}$. The actual values of f and $S_{DS}$ depend on spectral assumptions about the Deep Survey sources. The three spectral forms used above translate into lower limits of 180, 300, and 460 sources per square degree, respectively, making up the remainder of the total sky flux. These values are larger than the estimates of Koo and Kron [1982] for counts of quasars with J<22, 136±22 per square degree, but are comparable to their lowest flux estimates of 358±35 for J<23. (Remember that J magnitudes are close to the B band measurements.) These values assume that all the sources have fluxes just below the Deep Survey limit. If they are instead spread over a range of fluxes, their required number must be increased by $S_{DS}$ / $\overline{S}$, a factor of 4.7 for a decade of Euclidean behavior, which is 850, 1400, and 2200 sources to make up the H1L1 total sky flux. The optical source counts must then continue at least one to two magnitudes beyond current observational limits. Note that, as expected, all of these limits are significantly larger than the lower limit of ∼50 sources per square degree derived from the upper limit on the fluctuations' invariant excess.

The behavior of the source counts in the decade just below the Deep Survey sensitivity limit would have important information on the evolution of such sources as well as their possible contribution to the total sky flux. The results of a survey in the constellation Pavo [Griffiths, Maccacaro, Murray, and Giacconi 1981] may provide some information in this regard as it promises to probe the flux range down to ∼0.5x10$^{-3}$ IPCDS, half the flux limit of the Giacconi et al. Deep Survey. The preliminary results of Griffiths et al. [1981] suggest that there are as many as 19 extragalactic sources in a field of 0.135 msr, though at the time only a few of the identifications could be definitely classified as extragalactic. If the flux limit of the Pavo survey is equivalent to 0.5x10$^{-3}$ IPCDS then a simple Euclidean extension of the Deep Survey counts predicts that 24 objects should be seen, not including the enhancement due to Malmquist bias and confusion effects. However, the Pavo survey fluxes are measured by the High Resolution Imager, HRI, instrument of the Einstein Observatory instead of the IPC. The HRI has a lower energy reponse than the IPC, and HRI fluxes contain substantial contribution from below 1 keV. Therefore any comparisons

with measurements in the higher energy bandpasses will have even stronger dependence on the assumed spectral form. Even the change from HRI fluxes to IPCDS fluxes could modify the predicted number of sources in the Pavo survey by ~30%, depending on the source spectra. If sources are steeper than the assumed spectral form used to convert the Pavo HRI flux limit into $0.5 \times 10^{-3}$ IPCDS, then a correct conversion would increase the predicted number of sources expected in the Pavo survey if N(S) had a Euclidean continuation below the Deep Survey limit. If the source spectra are steep enough that the Pavo survey corresponds to a prediction of 31 sources for a straight Euclidean continuation, then the the observed upper bound of 19 sources would indicate that N(S) must break from the Euclidean form.

<u>Fluctuations Analysis with the IPC</u>

Fluxes below the limit of the Deep Survey can also be probed by studies of the fluctuations, in principle. In discussing the possibilities of such analysis we will use the parameters of a study by Stewart and Fabian, as outlined by Fabian [1981]. Their preliminary result is that the fluctuations measured by the IPC are consistent with the variation due to counting statistics alone. Single power law continuations of N(S), pegged to the Deep Survey results, and making up the remainder of the sky flux, are not consistent with the fluctuations distribution, according to Stewart and Fabian. Such a result signifies a sharp turnover in N(S) somewhat below the Deep Survey flux limit. In this section we will mention some of the pitfalls of such an analysis, though without further details from Stewart and Fabian we cannot say to what degree their result is affected. It is expected that a simple continuation of the Deep Survey source counts to a level where those sources provide the total sky flux would produce fluctuations detectable in the data used by Stewart and Fabian. That such fluctuations were not seen can have important ramifications in estimates of the fraction of the total sky flux that can be contributed by the sources dominating the Deep Survey source counts.

The data were taken from a $10^4$ second IPC exposure. The central

Analysis of Results                                      Source counts from <u>Einstein</u>

portion of the image plane was divided into 100 observations of $2'$ on a side, a solid angle of $1.11 \times 10^{-3}$ $deg^2$. The 1-3 keV bandpass of the IPC was used, giving counts roughly equivalent to IPCDS units. The extrapolated all sky flux was $9.7 \times 10^3$ IPCDS, equivalent to a mean count rate of 2.6 counts in the $10^4$ second exposure. The internal background was substantially larger, 15 counts per exposure, giving a total $\sigma_{noise}$ of 4.2 counts $exp^{-1}$. For the best fit Euclidean model of the Deep Survey, $W = 11$ IPCDS. Treating the observation bins as ideal flat response detectors with $2' \times 2'$ solid angle, the weighted response is $R_{3/2}$ = $9 \times 10^{-2}$ (seconds $\cdot exp^{-1}$), when S is measured in IPCDS and the intensity, I, is measured in counts per $10^4$ second exposure. Therefore $W_{eff} \approx 1$ count $exp^{-1}$.

A definitive analysis of the fluctuations can usefully constrain the behavior of sources with $I \sim W_{eff}$. Such sources have fluxes $\sim 10^{-4}$ IPCDS, or an order of magnitude below the Deep Survey flux limit. The Deep Survey best fit power law will actually saturate the total sky flux at intensities of only 0.6 counts $exp^{-1}$, which is just less than the one source per observation intensity level. However, the ease with which we can actually extract useful information at such low flux levels depends on the quality and quantity of the data. The high internal background rate and relatively short exposure time mean that $\psi = \sigma_{noise} / W_{eff} =$ 4.6. This is much larger than the values for the fluctuations analysis of UHURU, and our A-2 data, where $\psi$ is 1.4 and 1.2, although it is comparable to the Ariel V $\psi = 4.8$. The problem of too high $\psi$ (too small signal to noise) is exacerbated by the small number of observations, only 100 compared to effectively 270 for Ariel V. We can estimate the size of the 90% confidence region for any determination of $W_{eff}$ using the likelihood difference technique (see Section III:5, particularly equations [III.77] and [III.80] and Figures III.9 et seq.). With these estimates of $\psi$ and the number of observations, the approximate derived size of the 90% confidence region is $\delta_W \approx 0.4$. From this we can infer that a model with no fluctuations is statistically distinquishable from a reality where the fluctuations are generated at the level expected. More simply, the non-detection of the fluctuations by Stewart and Fabian is significant.

Analysis of Results                                Source counts from Einstein

.There are complications. The above estimates, and the analysis of
Stewart and Fabian, use the assumed best fit values from the Deep Survey
for the peg point of the N(S) low flux extension. As we saw in Figure
VII.28, there is a wide latitude in the allowed position of the peg
point, particularly if N(S) is significantly non-Euclidean over the Deep
Survey flux range. For example, if we take the bottom 1 $\sigma$ limit for the
Euclidean models, the predicted level for $W_{eff}$ for a Euclidean extension
drops to 0.7 counts $\exp^{-1}$ and $\psi$ increases to ~6. The fractional size of
the 90% $W_{eff}$ confidence contour is now 0.5 so that we would still expect
to see some fluctuations if the low flux extension is Euclidean. Non-
Euclidean models with $\gamma$~2 would be more consistent with the non-
detection of the fluctuations, though we will not explicitly estimate
here the expected likelihood difference. We have also assumed that the
point source spreading of the _Einstein_ Observatory telescope was much
less than the bin size of 2′. This is not quite true, since ~10% of the
counts from a point source can be found outside a circle of 3′ radius.
This could produce a slight reduction in the weighted response, $R_{3/2}$,
and therefore in the expected level of the fluctuations through $W_{eff}$.
Lastly, the likelihood difference estimates assumed that the analysis is
performed on unbinned data using the $\Delta\lambda$ likelihood statistic, while
Stewart and Fabian examined data in intensity bins, using the $\chi^2$
statistic on the histogram. This would have a measurable reduction in
the sensitivity of the data to the existence of fluctuations. On the
other hand, the Stewart and Fabian results are strengthened by the fact
that they took no account of any other source of variation, e.g.
variations across the IPC field of view due to changes in the gain.
Likewise, the fluctuations should contain a component due to galactic
sources, not included in the extension from the Deep Survey
extragalactic objects. The expected contribution of these variations
coupled with Stewart and Fabian's inability to detect of any
fluctuations in excess of counting statistics strengthens their
conclusion that N(S) must have a significant break near the Deep Survey
flux limit. These preliminary results certainly merit a more detailed
analysis. They would be improved by the use of additional IPC exposures
to enlarge the number of observations, as well as an unbinned data
analysis, similar to that used in this dissertation. The results of

Analysis of Results                                    Source counts from _Einstein_

such a project would also have an impact on the understanding of the source counts and identifications from the Pavo survey.

———————————————

## 5: Comparisons of AGN models with Optical Data

Studies of X-ray sources in other spectral bands, besides having an intrinsic value for elucidating the physical mechanisms of the emission, can provide indirect information about the behavior of those sources in the X-ray band. For example, with data on the connection between optical and X-ray fluxes, optical source counts can be used to predict the contribution of AGN to the total X-ray sky flux. Such a program is complicated by nonlinear correlations between optical and X-ray luminosities. Many authors have found that increases in the optical luminosity produce a smaller increase in the X-ray luminosity than strict proportionality would predict. That is, $L_x \propto (L_{opt})^\delta$, with $\delta < 1$. However, this is difficult to reconcile with the fact that the X-ray luminosity function is flatter than the optical luminosity function. The optically derived-parameters for evolution must be corrected by the deviations from simple proportionality, in order to apply them to the X-ray band. The observed correlations between X-ray and optical luminosity predict that $Q_{L;x}$ must be much less than $Q_{L;opt}$, the exact value depending on the actual form of the correlation. The use of the optical source counts to predict the X-ray sky flux involves several important extrapolations.

The message of the previous sections is that although the fluctuations, and more directly the IPC studies, tell us much about the behavior at lower fluxes of the sources that make up the XCS, there is still a great deal of ambiguity and confusion. AGNs evolve, though the form and amount is difficult to specify. From the Deep Survey counts we know they contribute substantially to the total sky flux, but the exact magnitude is poorly determined, even without extrapolatian below the Deep Survey limit. It would appear that our current information is more tantalizing than definitive, and that there are slim prospects for improvement in the quantity and quality of this information in the immediate future. However, this prognosis holds only for the direct X-

ray data.  X-ray sources are known to be emitters at other frequencies and by studying their behavior at those frequencies we may be able to draw more accurate conclusions about their X-ray behavior than can currently be done directly.

## Optical studies of AGN: advantages and problems

These non-X-ray studies should be particularly useful in the case of AGN, which are strong emitters in the ultraviolet, optical, infrared, and sometimes radio, as well as the X-ray.  The original definition of an AGN was based on behavior in the other bands:  broad emission lines, a blue stellar nucleus, etc.  An examination of the optical data on AGN offers distinct opportunities.  Current optical data probes to deeper equivalent fluxes than X-ray data, with indications that current observations are at or near the point of convergence, where the effective N(S) index drops just below 2 [Koo and Kron 1982].  There are also a larger number of sources discovered and studied in greater detail in the optical band.  This allows better coverage of the total luminosity-redshift plain, which is necessary in order to have a non-extrapolative picture of the evolution of the luminosity function.

With all the promise of the optical studies, there are still significant difficulties.  A general problem is identification.  The uniform characteristics of AGN in the X-ray band above 4 kev, so far as we currently know, make it fairly certain that we can completely and correctly identify the vast majority of AGN in a flux limited survey (e.g. the XCS).  In contrast the completeness of any optically selected sample is more difficult to ascertain (though this may change with future technical progress).  Recent studies have found X-ray selected AGN that would have failed traditional optical criteria such as blue excess or emission line width [e.g. Stocke et al. 1982; Huchra, Wyatt, and Davis 1982].  It is estimated that from 10-20% of all AGN would have been missed by such criteria.  Griffiths et al. [1980] made a preliminary estimate that as an upper bound no more than 20% of the Pavo survey objects detected in the X-ray would have been missed in an

Analysis of Results                                      Optical Comparisons

optical program. The problem of identification is exacerbated for very low luminosity AGN, such as M81 [Elvis and Van Speybroeck 1982], which are prone to selection difficulties in the soft X-ray band where absorption may have a significant impact. On the other hand, it is possible that some fraction of the very low flux objects identified on the basis of color may not be AGN at all [e.g. Bonoli et al. 1980; cf. Veron and Veron 1982]. A related problem in making comparisons between optical and X-ray samples is the probably misleading distinction between quasars and Seyferts. (Indeed, the determination of which AGN are Seyferts, and their type, is also an ambiguous process.) A Seyfert galaxy at a sufficiently high redshift so that the non-nuclear component is not easily visible is often classified as a QSO, while it would not be if the galaxy were detected. Such a selection effect can produce a spurious enhancement in the optical estimates of QSO evolution [Hawkins and Stewart 1981].

AGN have very rich spectra in the optical band, exhibiting varied and complex behavior. While the presence of broad and narrow emission lines, absorption complexes, recombination blends, and possible thermal components [e.g. Malkan and Sargent 1982; Grandi 1982] makes possible studies of a wide range of physical processes in and near the AGN, they complicate measurements of the nonthermal optical continuum thought to be directly associated with the power law component observed in the X-ray band. The difficulty is compounded because the optical fluxes are generally not narrow band continuum measurements, but relatively broad band measurements that may encompass spectral features from many of the above phenomena. As discussed by Schmidt and Green [1983], this is a particular problem for samples of objects spread over a wide range of redshifts. The redshift corrections needed to produce a predicted flux at a particular emitter frame frequency far from the observed broad band measurement become very uncertain.

## Relationship between optical and X-ray flux

The connection between X-ray observations and measurements taken

in other bands is often parameterized by

$$\alpha_{yx} \equiv - \frac{\log( \ell_x / \ell_y)}{\log( \nu_x / \nu_y )} \, . \hspace{4cm} [VII.28]$$

The "y" in $\alpha_{yx}$ is a free symbol for the band of comparison, i.e. $\alpha_{Rx}$ for the Radio to X-ray, $\alpha_{ox}$ for optical, etc. $\ell_x$ and $\ell_y$ are the differential luminosities at the indicated frequencies $\nu_x$ and $\nu_y$. $\alpha_{ox}$ can be thought of as the effective energy spectral index of an idealized power law connection between the two bands. Because of the large difference in the frequencies ($\nu_{opt}$ and $\nu_x$ are separated by over 2.6 decades), even a small difference in $\alpha_{ox}$ is lever-armed into a very large difference in the ratio of luminosities: for a fixed $\ell_{opt}$, $\Delta\alpha_{ox}=0.1$ multiplies $\ell_x$ by 1.8, and by 3.3 for $\Delta\alpha_{ox}=0.2$.

The observations of and conclusions about $\alpha_{ox}$ for collections of AGN have become quite common, with results presented from a wide range of perspectives. The large variation in methodology complicates a synthesis of the results of different authors. The AGN samples can be defined in various ways: by previous lists of optical objects [Tananbaum et al. 1979; Zamorani et al. 1981]; by complete samples selected in the optical [Marshall et al. 1983a] or the radio [Tananbaum, Wardle, Zamorani, and Avni 1983]; or even by an X-ray selected sample from identified serendipitous objects [Chanan, Margon, and Downes 1981; Reichert, Mason, Thorstensen, and Bowyer 1982; Kriss and Canizares 1982]. (The references presented in this section are in no way a complete list of the literature.) Note, however, that an X-ray selected sample of serendipitous objects without a complete set of identifications may still have optical selection effects present [see Reichert et al. 1982]. For example, objects whose optical counterparts are dimmer than a certain threshold magnitude may be excluded. When comparing the results from these variously defined samples it is important to remember that the derived distributions for $\alpha_{ox}$ can depend on the selection criteria. This is particularly true for non-parametric analysis, i.e. where no attempt at an a priori assumption of the form of the $\alpha_{ox}$ distribution is made (see Condon, O'Dell, Puschell, and Stein

[1981], especially their Appendix 2).

The variation in what is meant by "optical" creates further problems. A reference wavelength often chosen is 2500 Å. It has the advantage that it is directly observable for many high redshift objects, though not at low redshifts without direct ultraviolet data. The flux in this region of the spectrum is dominated by the 3000 Å blue bump, the emission of which is dominated by line processes and a possible black-body component [e.g. Malkan and Sargent 1982]. Other groups use 4400 Å or 5500 Å, the center of B or V magnitude measurements [Allen 1973]. These measurements are often from broad band photometry, which is prone to inaccurate redshift corrections. Even when examining low redshift objects the broad band measurements will have strong components from emission lines: for typical AGN spectra, $H_\gamma$ affects the B magnitudes, and $H_\beta$ and [OIII] influence the V magnitudes. Direct measurements of the nonthermal continuum can be made, but some effort is required [e.g. Blumenthal, Keel, and Miller 1982]. The X-ray band flux measurements have similar problems. Ideally, the measurements should be taken at the medium energies of the A-2 experiment, where complications from absorption and possible soft components are minimized. Instead, all the studies referenced above are based on Einstein Observatory IPC or HRI flux measurements. These broad band measurements are converted to the differential 2 keV luminosity. In interpreting the broad band flux and in correcting for redshift, the X-ray spectrum is assumed to have a particular form, generally a power law with $\alpha$ = 0.5 and galactic absorption. Deviations in either the spectral index, for high redshift sources, or $N_H$, for low redshift sources, can have pronounced effects.

From the measurement uncertainties alone, it is not surprising that there is a large scatter in the values of $\alpha_{ox}$. In fact there appears to be intrinsic scatter in the distribution, with rms range in $\alpha_{2500;x}$ of 0.2 [Avni and Tananbaum 1982]. The typical value for $\alpha_{2500;x}$ in the same analysis was ~1.5.[*] In all studies, strong correlations

[*]As a crude point of comparison, Koo and Kron estimated that the total surface brightness from QSOs at 4500 Å is ~$1.8 \times 10^{22}$ erg s$^{-1}$ cm$^{-2}$ sr$^{-1}$ Hz$^{-1}$. Using the extrapolated X-ray sky flux at 2 keV we can estimate that $\alpha_{4500;x}$ ~1.3, though this makes no allowance for possible

between the X-ray and optical luminosities are observed, encouraging the
supposition that phenomena in the two bands are physically related. The
earlier studies assumed that the two values were in fact strictly
proportional, i.e. $L_x \propto L_{opt}$. Further work has lead many authors to a
modification of this relationship:

$$L_x \propto (L_{opt})^{\delta} . \qquad\qquad\qquad [VII.29]$$

The derived values for $\delta$ were significantly less than 1, so that a given
increase in the optical luminosity produced on average a lesser increase
in the X-ray band. Typical best fit values are $0.66 \pm 0.06$ [Zamorani
1982a] and $0.47 \pm 0.15$ [Tananbaum et al. 1983]. One model for the non-
proportionality has been tentatively ascribed to the division of the
optical flux into the nonthermal continuum, which is strictly
proportional to the X-ray flux, and the lines and other components,
which are not [Malkan and Sargent 1982]. Note, however, that the low
redshift ($z \lesssim 0.4$) continuum measurements of Blumenthal, Keel, and Miller
[1982] are consistent with $L_x \propto (L_{opt;cont})^{0.5}$, when an index of one is
expected, though there is large scatter. The result that $\delta < 1$ has been
used by many authors [e.g. Zamorani 1982a; Reichert et al. 1982] to
explain the observed difference of the redshift distribution of an X-ray
selected sample from an optically selected one [e.g. Margon, Chanan,
and Downes 1982].

Because the luminosity of an AGN drawn from a flux limited sample
is strongly correlated with its redshift, it is possible that the data
indicate not a variation in $\alpha_{ox}$ with optical luminosity, but rather a
dependence on redshift. We can generalize [VII.29] to

$$L_x = A f_z(z) (L_{opt})^{\delta} , \qquad\qquad\qquad [VII.30]$$

where A is a constant of proportionality. Reichert et al. [1982] have
fit their sample of X-ray selected and other AGN to three different
forms of $f_z(z)$. Though the details differ from model to model, they
find that for all forms the best fit parameters imply significant

differences in the redshift correction in the two bands. If the
observed uncorrected values for $\alpha$ for QSOs are substantially greater
(steeper) than this, then QSOs could not make up the bulk of the 2 keV
background.

correlation with <u>both</u> luminosity and redshift, though the correlation
with luminostiy is the more significant of the two. They point out that
being able to distinguish between redshift and luminosity dependences is
important for an interpretation of the effect. Redshift dependence,
instead of being directly connected to the emission processes as
outlined above, may indicate explicit differential evolution in the two
bands or may be due to incorrect spectral assumptions. (<u>e.g.</u> Cavaliere
<u>et al</u>. [1980]). An example of the latter would be if the true spectral
index, $\alpha$, of the sources was, on average, different from the assumed
value used for redshift corrections, $\alpha_m$. The relationship between the
true luminosity, $L_t$ and the "measured" luminosity, $L_m$, used in
performing fits of equations [VII.29] and [VII.30] is

$$L_m = L_t \; (1+z)^{\Delta\alpha} \; , \hspace{3cm} \text{[VII.31.a]}$$

$$\Delta\alpha \equiv \alpha_m - \alpha \; . \hspace{3cm} \text{[VII.31.b]}$$

In the presence of such errors in the redshift correction, there will be
a z-dependent correlation between the measured X-ray and optical
luminosities of the form

$$f_z(z;\upsilon) = (1+z)^{\upsilon} \; . \hspace{3cm} \text{[VII.32]}$$

The parameter $\upsilon$ is determined by the errors in the spectral assumptions
in both bands:

$$\upsilon = \Delta\alpha_x - \delta \; \Delta\alpha_{opt} \; . \hspace{3cm} \text{[VII.33]}$$

Reichert <u>et al</u>. derived best fit values for $\delta$ of ~0.6, with the
99% confidence range from ~0.35 to ~0.75, depending on the form of $f_z$
being tested. Equation [VII.32] represents one model they tested, with
$\upsilon = 1.4$ being the best fit. The confidence region indicated strong
correlation between the values of $\delta$ and $\upsilon$. $\upsilon = 0$ is barely acceptable
at the 99% level. In contrast, Tananbaum <u>et al</u>. [1983] found no
correlation of redshift and $L_x/L_{opt}$ for radio-selected quasars. For
optically selected objects, Avni and Tananbaum fit the model

$$f_z(z;\eta) = \exp( \; \eta \cdot \tau(z) \; ) \hspace{3cm} \text{[VII.34]}$$

where $\tau$ is the fractional look back time (see Appendix A1). They

derived a best fit value of $\eta=0$, _i.e._ no z dependence, and $\delta = 0.7$. The allowed range for the parameters again showed strong interdependence. In fact a model with no explicit luminosity dependence, $\delta = 1$, was acceptable at the 95% confidence level but only with substantial redshift dependence, $\eta = 3.6$[*].

In contrast to all the above results, where the ratio of X-ray to optical luminosity decreases as the optical luminosity increases ($\delta<1$), Kembhavi and Fabian [1982] have come to the directly opposite conclusion. From their distribution of $\alpha_{ox}$ for optically bright radio-quiet quasars, $m_v < 18$, they predict that 6 out of the 55 low optical flux objects, $m_v > 18$, should have been detected in the X-ray. Only one source was detected. Because the average redshifts of the low and high optical flux groups are similar, the principal difference between the two samples must be that the low flux objects have on average correspondingly lower optical luminosities. Therefore Kembhavi and Fabian conclude that the low rate of X-ray detections for the low luminosity objects indicates a lower value of $L_x/L_{opt}$, equivalent to models of the form of [VII.29] with $\delta > 1$. The contradiction with the conclusions of Zamorani et al. [1981], who provided much of the data used by Kembhavi and Fabian, has been examined by Reichert [1982]. She points out that if the Kembhavi and Fabian objects are grouped by constant optical magnitude, the higher X-ray flux objects in each group are at significantly lower redshifts than the lower X-ray flux sources. If we ignore possible explicit redshift correlations, this is consistent with $L_x/L_{opt}$ increasing for the lower values of $L_{opt}$. This suggests that the distribution of $\alpha_{ox}$ that Kembhavi and Fabian derived from the high optical flux sources, of which roughly half were not detected in X-rays, is in error, and that the true distribution would predict fewer X-ray detections of the low flux sources.

---

[*]Avni and Tananbaum used a different, though equivalent parameterization, $A_o$ and $A_z$. These can be related to our parameters by $\delta = 1 - 2.605\ A_o$, and $\eta = -6.0\ A_z$.

Analysis of Results                                            Optical Comparisons

## Comparisons: X-ray and optical luminosity functions

If X-ray and optical luminosities and fluxes are not strictly proportional, there will be differences between the statistical properties of AGN as determined separately in the two bands. For example, the shape of the luminosity function of AGN will differ from one band to the other. Restricting ourselves to low redshift samples, where any effect from any redshift dependence is expected to be minimal, a luminosity-dependent relationship of the form given in [VII.29] will modify the power law index of the luminosity function in the following way:

$$\beta_{opt} - 1 = \delta \, (\beta_x - 1) \, . \hspace{3cm} [VII.35]$$

As pointed out by Maccacaro et al. [1983a], the derived values of $\delta < 1$ predict that the optical AGN luminosity function is flatter than the X-ray luminosity function, i.e. $\beta_{opt} < \beta_x$. This is not consistent with current estimates of the luminosity function in the two bands. From the XCS, Piccinotti et al. [1982] find that $\beta_x = 2.75\pm0.15$. Even with the largest allowed value, $\beta_x = 2.9$, the predicted value of $\beta_{opt} \approx 2.3$ for $\delta = 0.7$, and if $\delta = 0.5$ as suggested by some samples [e.g. Tananbaum et al. 1983] then $\beta_{opt} \approx 2$. In contrast, the optical luminosity function appears to be steeper than the X-ray. Véron [1979] estimated that for low redshift Seyfert 1 Nuclei, the observed $\beta_{opt} \approx 3$. A $\beta_{opt}$ of 3 corresponds to a $\delta$ of 1.05. Although Huchra and Sargent [1973] (see also Huchra [1977]) see a flattening of the optical Seyfert galaxy luminosity function, this occurs at a point corresponding to an X-ray luminosity below the sensitivity limit of the XCS, $\leq 10^{43}$ ergs $s^{-1}$. A corresponding break in the X-ray luminosity function has been suggested by the observations of Elvis, Soltan, and Keel [1983].

The above disagreement between the slopes of the luminosity functions and the derived values for $\delta$ might be due to the different X-ray bands used in the measurements. The luminosity function of the XCS is based on measurements in the 2 to 15 keV energy range, while the $\alpha_{ox}$ distribution studies use the 1 to 3 keV information from the Einstein observation programs. However, the analysis of Maccacaro et al. [1983a], using IPC fluxes for the Medium Survey objects, derived a 90%

confidence range for $\beta_x$ of 2.9 to 2.3, with a best fit value of 2.6. Any reduction of $\beta_x$ below 2.9 greatly increases the difficulty of reconciling the determinations of $\delta$. The Medium Survey based determination of $\beta_x$, however, will be strongly affected by the presence of absorption in low luminosity objects, as well as by an expected significant contribution by sources at or below the break in the luminosity function. In addition, the Medium Survey contains sources at significant redshift so that corrections for evolution are a possible problem. We can compare the Medium Survey luminosity index with the optical value from sources covering a similar range in luminosity and redshift using the results of Marshall et al. [1983b]. Their 90% confidence range for $\beta_{opt}$ is still steeper, 3.1 to 4.5 after correction for evolution, so the disagreement is maintained or even worsened.

If we find the luminosity function argument that $\delta \gtrsim 1$ convincing, then the survey of $\alpha_{ox}$ results can only be consistent with a large explicit redshift dependence of $L_x/L_{opt}$, assuming no systematic error in the fitting process. As we remarked earlier, the thin tilted ellipse for the confidence region for $\delta$ and the z dependence parameter, $\eta$ or $\upsilon$, shows the strong coupling of these two parameters. The semi-major axis of this ellipse can approximately relate the best fit value of one parameter given the other, and vice-versa. For the $(1+z)^{\upsilon}$ model, the Reichert et al. analysis indicates that the best fit values are related by

$$\delta = -0.12 \ \upsilon + 0.78 \ . \tag{VII.36}$$

Models with $\delta = 1$ are outside their 99% confidence region, although their best model with this value has $\upsilon = -1.8$. The Avni and Tananbaum study yields

$$\delta = -0.09 \ \eta + 0.7 \ . \tag{VII.37}$$

If $\delta = 1$ then the most likely value for the redshift dependence is $\eta \approx -3.5$, which is marginally acceptable at their 95% level. The statistical uncertainty in $\eta$, for a given value of $\delta$, is $\sim \pm 0.6$ (1 $\sigma$), estimated from the maximum width in $\eta$ of the $\Delta\Lambda = 1$ confidence contour. Therefore, if we have other information that $\delta$ must be $> 1$, then if the Avni and Tananbaum parameterization is correct, their fits require a

strong redshift dependence, $\eta < -2.9$ (84% confidence upper bound).

<u>Comparisons</u>: <u>amounts of evolution in the X-ray and optical</u>

Pure density evolution would be idential in any band, as it corresponds to a change in the total number of objects no matter what their luminosity. For other forms of evolution, where the changes depend on luminosity, possible variation in $L_x/L_{opt}$ will modify the impact of the evolution in moving from one band to the other. When the redshift dependence of $L_x/L_{opt}$ is given by the exponential look-back-time model of Avni and Tananbaum (equation [VII.34]), then for pure luminosity evolution equation [VII.30] leads to

$$Q_{L;x} = \delta \, Q_{L;opt} + \eta \, ,\qquad\qquad\qquad\text{[VII.38]}$$

where $Q_{L;x}$ and $Q_{L;opt}$ are the evolution parameters in the two bands. If we assume no explicit redshift dependence, $\eta = 0$, then for $\delta < 1$, the amount of evolution seen in the X-ray band will be less than what would be inferred based on optical data. Derived values for $Q_{L;opt}$ are 8 [Braccesi et al. 1980; Koo and Kron 1982], 6 [Cheney and Rowan-Robinson 1981a; Schmidt and Green 1983] and the range 6-8 (90%) [Marshall et al. 1983a]. Maccacaro et al. [1983] determine that $Q_{L;x}$ is in the range 4 to 6, although recall that in the last section we showed that $Q_{L;x}$ could not be too much greater than 4 without exceeding the observed Medium Survey source counts. (For most of the results reviewed here, $q_0$ was assumed to be 0 or near 0.1) If we fix $Q_{L;x} = 4.5$ and $Q_{L;opt} = 7$ then $\delta$ ~0.6, consistent with many of the $\delta$ fits [Maccacaro et al. 1983].

Observational evidence that $Q_{L;x}$ is less than $Q_{L;opt}$ is not necessarily confirmation that $\delta < 1$, if the possibility that $\eta < 0$ is allowed. By including the regression relationship of [VII.37] in equation [VII.38], we find

$$Q_{L;x} - 7.9 \approx \delta \, (Q_{L;opt} - 11.3) \, .\qquad\qquad\text{[VII.39]}$$

Now if $Q_{L;x} = 4.5$ and $Q_{L;opt} = 7$, the most likely values in the light of the Avni and Tananbaum analysis are $\delta \sim 0.8$ and $\eta \sim -1$. When nonzero values of $\eta$ are allowed, reducing $Q_{L;x}$ for fixed $Q_{L;opt}$ is compensated

by the decrease in $\eta$ to produce an increase in the most likely value of $\delta$, exactly opposite to the case where $\eta$ is fixed equal to 0. For example, reducing $Q_{L;x}$ to 3.5 increases the most likely $\delta$ to 1, though the allowed range for $\delta$ will still include values less than 1. A crude estimate of the 1 $\sigma$ range is $.84 \leq \delta \leq 1.16$, treating $\eta$ as uninteresting. Similar general trends will also arise for the other possible forms for $f_z$, though the numerical parameterization will be less straightforward.

We therefore have three sets of observations in the optical and X-ray bands: the current epoch luminosity function index ($\beta$), the estimated pure luminosity evolution coefficients ($Q_L$), and the correlation of $\alpha_{ox}$ with $L_{opt}$ and $z$ ($\delta$). Considered pairwise, the $\beta$ and $Q_L$ relations are consistent, as are the $Q_L$ and $\delta'$ derived conditions. However, it is difficult to combine all three together, assuming that the various determinations are systematics free. If we assume that the $\delta \geq 1$ conclusion drawn from $\beta_{opt} \geq \beta_x$ is correct, then we require that that there is a strong explicit redshift dependence of $\alpha_{ox}$ and that if the evolution is pure luminosity form, $Q_{L;opt}$ is significantly larger than $Q_{L;x}$. That is, the optical luminosity evolves to a significantly greater extent than does the X-ray luminosity.

The relationship between the optical and X-ray bands for the index evolution models of Schmidt and Green [1983] can be more complicated, since these models are described by two parameters, $Q_\beta$ and $L_o$, rather than the single parameter of pure luminosity evolution. The evolution of the luminosity function index (see equation [VII.21] and the discussion following) is actually very simple, as at any given epoch the relationship between the X-ray and optical indices is determined by equation [VII.35]. This means that independent of any explicit redshift correlations

$$Q_{\beta;x} = \delta^{-1} Q_{\beta;opt} \cdot \qquad\qquad [VII.40]$$

If $\eta = 0$, then for $\delta < 1$ the amount of X-ray evolution is _stronger_ than in the optical band, in contrast to pure luminosity evolution (equation [VII.38]). The best fit values for $Q_{\beta;opt}$ derived by Schmidt and Green depend on the interpretation of the low flux number counts and on $q_o$.

Analysis of Results                                         Optical Comparisons

For $q_o$ = 0.1 they find $Q_{\beta;opt}$ = 3.4 to 4.5. For $q_o$ = 0.5 the best fit
values range from 2.4 to 3.4. The corrected value for $\beta_{opt}(\tau=0)$ depends
strongly on the actual model used, particularly for the higher
luminosities. At the low end of the luminosity function, determined
mostly by low redshift sources, the evolutionary corrections are small.
For $-23.5 > M_B > -24.5$ the effective optical index is ~3.3, with significant
steepening at higher luminosities. Given the typical value of
$\alpha_{4400\text{Å};2\text{ keV}}$ = 1.4 [Schmidt private communication] this magnitude range
corresponds to 1.4 to 3.6 $L_{44}$. Over this range the XCS-determined
maximum in $\beta_x$ is $\leq$ 2.9. From this value and equation [VII.35] we derive
$\delta$ = 1.2 and $Q_{\beta;x}$ equals 2.8, 3.7, 2.0, and 2.8 respectively for the four
models presented by Schmidt and Green. Their best fit for $M_{o;opt}$, the
absolute B magnitude corresponding to $L_{o;opt}$, took on values ranging
from -20.1 to -23.3.

Given the strong dependence on $L_o$ for the allowed range of $Q_{\beta;x}$
from the fluctuations analysis, it is difficult to make a useful
comparison of the fits in the two bands. In fact, it is very difficult
to provide a simple transformation of the model from one band to the
other, particularly the conversion of the pivot luminosity, $L_o$, from
optical to X-ray units, if there is any explicit redshift dependence of
$L_x$ versus $L_{opt}$. If $L_o$ is a single well defined quantity in one band,
its value will vary with epoch in the other. For instance, if we assume
that the optical $L_{o;opt}$ is fixed, then

$$L_{o;x}(\tau) = L_{base;x} \exp\left[ \frac{-\eta \ (\beta_x(\tau)-1)}{Q_{\beta;x}} \right] , \qquad\qquad \text{[VII.41.a]}$$

where

$$L_{base;x} \equiv A'(L_{o;opt})^\delta \qquad\qquad\qquad \text{[VII.41.b]}$$

is the simple transformation of $L_{o;opt}$ for z=0 using equation [VII.30].
If there is no explicit redshift dependence ($\eta$ = 0), then $L_{o;x}$
isconstant and equal to $L_{base;x}$. The above values for $M_o$ correspond to
values of $L_{base;x}$ ranging from 0.06 to 1.2 $L_{44}$. Note that $L_{o;x}(\tau)$ is
not the simple transformation of $L_{o;opt}(\tau)$. Also, the variation of $L_{o;x}$
with epoch does not explicitly depend on $\delta$, as both the numerator and

denominator of the exponential transform identically in moving from the
X-ray to the optical bands (equations [VII.35] and [VII.40]), though the
transformation from $L_{o;opt}$ to $L_{base;x}$ does have a $\delta$ dependence. In
addition, estimates of $\eta$ are dependent on the assumed value of $\delta$.

The movement of the index evolution pivot point, $L_o$, has strong
repercussions. For an example we use our earlier estimate of $\delta \approx 1.2$
and one of the Schmidt and Green best fit models (their model HH1) which
have corresponding X-ray models of $Q_{\beta;x} = 2.8$ and $L_{base;x} = 0.3\ L_{44}$ ($q_o$
$= 0.1$). Though Avni and Tananbaum find that models with $\delta$ this high are
specifically excluded (95% confidence), we use their correlation between
acceptable values of $\delta$ and $\eta$ to estimate a best fit value $\eta \approx -5.5$.
This very strong epoch dependence means that for low redshifts the
effective X-ray pivot point, $\sim 13\ L_{44}$, is above most of the range of the
luminosity function. Thus the evolution would tend to decrease the
contribution to the sky flux and invariant excess. However, at higher
redshifts $L_{o;x}$ will decrease until $\tau \sim 0.7$, ($\beta_x(\tau) \approx 1$), where it is equal
to $L_{base;x}$. Finally at $z = 3.5$ it is equal to $\sim 0.15\ L_{44}$, enhancing the
effects of evolution. Because the original luminosity function is so
steep, even a small change in $L_o$ will have an impact. The result of the
general decrease in $L_o$ with increasing redshift is to depress the effect
of evolution at high and moderate fluxes and enhance it at the very
lowest fluxes, making simple extrapolations from the moderate fluxes
hazardous.

These difficulties decrease the utility of simply mapping optical
fits of index evolution models into the X-ray band. If index evolution
is found to be a better description than pure luminosity evolution, it
is probably better to directly apply the models to even the currently
limited X-ray data than to transform the best fitting optical
parameters. At least that would remove the uncertainty due to possible
explicit redshift dependence of the optical to X-ray ratio.

### Use of optical AGN counts to estimate the contribution to $S_{as}$

Do quasars, or rather AGN, contribute the dominant fraction of the

X-ray sky flux? Even the best analysis efforts using Einstein Observatory data can make no clear statement based solely on X-ray measurements, but the optical source counts can help. These extend to much deeper levels than are currently available in the X-ray, and may go deep enough that the contribution to the sky flux converges. Adding up all the optical contribution to the sky flux from these sources, and multiplying by a "typical" conversion factor to get an X-ray flux, we can compare the predicted X-ray contribution of AGN to the total sky flux. Of course, as we have already discussed in some detail, there are complications at each step. One indication of the difficulties is the wide range of estimates extant in the literature, from 30% to over 100% of the total sky flux, as variations on the basic technique are applied by different authors [e.g. Cavaliere et al. 1981; Zamorani et al. 1981; Cheney and Rowan-Robinson 1981; Kembhavi and Fabian 1981; Zamorani 1982b]. We will not perform such an analysis here, but will touch instead on the more obvious weak points and assumptions. A good application of the technique, and its results in the context of other information on the diffuse background, has been presented by Zamorani [1982b].

A key difficulty is that there is no single correct value of $\alpha_{ox}$ to tranform an optical flux, or luminosity, into an X-ray one. Even a distribution of values is inaccurate because the data show strong correlations with optical luminosity and/or redshift. The $\alpha_{ox}$ distribution is usually simplified to a single number, or function of $L_{opt}$ (an effective $\alpha_{ox}$), but this requires significant assumptions about the underlying parent distribution. Current work [Avni and Tananbaum 1982] indicates that such assumptions are consistent with and justified by the data; however, remember there are unavoidable, albeit correctable, biases introduced when using any flux-limited sample [Condon et al. 1981].

The common practice is not to work directly with the optical counts, but to fit an evolutionary model for the optical luminosity function to the counts, to transform this fitted model to the X-ray, and thereby determine the total sky flux. The optical fits depend

critically on the behavior at the lowest fluxes, and are purely based on comparing the model N(S) to the observed counts. The redshift distribution at these low fluxes would materially help assess the validity of these fits. However, only recently has it become possible to obtain this distribution, and it is not yet available. The difficulty of mapping the evolutionary model from the optical to the X-ray has been examined in some detail above. However, we reemphasize the sensitivity of the process to a possible redshift dependence, rather than a sole dependence of $\alpha_{ox}$ on $L_{opt}$, as is usually presumed. Note in particular that the optical source counts do not, and cannot in the absence of redshift information, make any correction for spectral redshifting. This can be approximated from a particular model, but it is another source of error, particularly given the rich structure of AGN optical spectra.

The greatest contribution to the sky flux will come from the sources corresponding to the deepest optical source counts ($J \gtrsim 22$). In fact, a "reasonable" continuation of the X-ray counts below the Deep Survey limit (see the last section, particularly the discussion following equation [VII.27]), implies that a significant contribution ($\gtrsim 25\%$) to the sky flux will be made by sources dimmer then even the best current optical survey limits. Therefore all estimates of the optical sky flux require that the fit made to the optical counts is an accurate model of the luminosity function and its evolution. The redshifts of the lowest intensity sources are important to validate particular models, but currently we do not have this information. Direct estimates of $\alpha_{ox}$ for such sources would also be helpful, but we do not have that information either. Without this information, we base our conclusions primarily on objects covering a much different region of the redshift-luminosity plane. Calculations restricted solely to that part of the redshift-luminosity plane that is covered are more believable, but less useful for ascertaining the total contribuiton of AGN to the sky flux. For example, Marshall et al. [1983a] estimate from their X-ray observations of an optically complete sample of ten QSOs that between 8% and 24% of the 2 keV sky flux comes from quasars with B<19.2 and z<2.2. To be fair, the current consensus is that an extrapolation from the

Analysis of Results                                      Optical Comparisons

known sources to the unknown properties of the low flux sources is consistent with all available information so that we do not <u>require</u> a major change in the correlations and other properties.

Additional important information is available from radio observations of the AGN. When dividing the sources into radio-loud and radio-quiet objects many authors [<u>e.g.</u> Zamorani <u>et al.</u> 1981; Reichert <u>et al.</u> 1982] have noted that the radio-loud objects will have roughly three times the X-ray luminosity compared to a radio-quiet source with the same optical luminosity. The effective $\alpha_{ox}$ of the two groups of objects are calculated separately, and applied to the evolving luminosity function model, weighted by their presumed fraction. Usually 10% of all AGN are considered radio-loud. Though this is a substantial improvement over an identical treatment of all objects, it requires another host of assumptions. Particularly, we must assume that the fraction of radio-loud objects, as well as their relative X-ray flux enhancement, are substantially unaffected by evolution. In addition, a radio-loud/radio-soft bifurcation may be a critical oversimplification. Condon <u>et al.</u> [1981] indicate a continuous distribution in the ratio of radio to optical luminosities. The creation of a truly trivariate radio-optical-X-ray distribution may be important for an accurate comparison of deep source counts in the three different bands.

<div align="right"><u>Summary</u></div>

All of this discussion should be considered illustrative of the difficulties of applying optical evolution models to the X-ray band. Naive transformations can produce unlikely or even unphysical behavior based on admittedly sparse, and possibly incomplete, optical data. This is not to say that comparisons among phenomena in the X-ray, optical, and radio bands are useless. Particular studies of correlations in observed sources obviously can and do lead to important conclusions about the physical processes at the heart of AGN. Comparisons of the source counts in the different bands can and do yield information about the sources' evolutionary behavior that would be unavailable from studies of a single band. Though the extension of these studies to the

very highest redshift, low flux, sources may be no more unwarranted than
our extrapolations from the XCS-derived luminosity functions, we should
continue to be aware that they are extrapolations, each with its
concomitant (perhaps reasonable) assumptions.     One particular
assumption, always relevant when discussing Einstein results, concerns
the X-ray spectrum of the high redshift AGN. As we have shown, the
conversion between IPC and HED fluxes can be uncertain by a factor of
two.

------------------------

### 6: Excess Variance from Non-Poisson Fluctuations

Non-Poisson distributions of X-ray sources can produce
significant additional contribution to the fluctuations. We
present estimates of the amplification of the fluctuations
due to sources clumped on supercluster scales. We also
sketch the utility of studies of the X-ray anisotropy for
probing the general distribution of matter. On the very
largest angular scales, the comparison of the Compton-
Getting (dipole) signal in the X-ray and microwave bands
supports the hypothesis that the overdensity responsible for
our observed velocity is "local" compared to the scale of
the sources that produce the X-ray sky flux. The limits of
the excess variance constrain the magnitude of lumps of X-
ray volume emissivity, the size of the limit depending on
the lump size. For data from our HED 1 smeared-detector,
the strongest constraints are placed on structures larger
than ~400 Mpc. These constraints can be translated to
limits on matter clumping of the universe at scales
otherwise not easily accessible. Using a naive Euclidean
approximation, the fluctuations do not permit highly clumped
structures ($\delta\rho/\rho > 1$) larger than ~100 Mpc.

An important assumption in our analysis of the fluctuations is
that they are due to point sources whose spatial distribution is
generated by a Poisson process; that is, there is no explicit
correlation in the position of two objects. An accurate treatment of
any other cases requires new tools of scope and complexity similar to
what has already been presented. In place of such a complete treatment,
we will present estimates of the effect of non-Poisson and other clumped

behavior on fluctuations measurements.

It is often stated that the Poisson assumption leads to minimum estimates of the fluctuations, that any other source distribution will increase the variance of the fluctuations distribution. The Poisson distribution of sources means that each source is placed randomly in space, without consideration of the actual position of sources placed earlier. Any structure is produced by structure in the parent density distribution. If the placement of the source is more likely near the physical position of other sources (not just due to enhancement of the parent density), then the resultant, non-Poisson, distribution will indeed have its variance enhanced. However, in contrast to these "clumped" distributions, we can imagine "anti-clumped" behavior, where sources are preferentially placed away from the location of other sources. Such a distribution of sources would have a fluctuations distribution narrower than the Poisson model. An extreme example of anti-clumping is where the sources have a coherent structure with even spacing. For statistical purposes, all the sources within one coherence length are effectively equivalent to a single source with the summed intensity of its components. If the coherence length is large enough, so that the anti-clumping occurs at all scales, then the sky will appear very isotropic. Admittedly, such structures seem unlikely according to our current intuition of the origin and evolution of large scale structure of the universe. Most currant models for clumping produce enhancements to Poisson fluctuations.

In this section we consider the effects of structure in the universe in two different scale regimes. The distinction of the scales is defined by the angular size of our measurements. Small scale variations, in our context, are those with small angular extent compared to the measurement, while large scale variations are comparable to or much larger than the measurement. Determining the "size" of the measurement, with proper allowances for the spread in the spatial response function, is one of the difficulties in performing a proper analysis. The measurements, as we have remarked, have non-zero response to an area in excess of 100 square degrees, but 90% of the total $R_1$

weighted response comes from a rectangle of $11.2^{\circ} \times 4.4^{\circ}$. The area of the sky that dominates the measurements of the excess variance is even more compact, with 90% of the contribution to $(R_2)^2$ coming from an area of ~26 square degrees. Let us consider the linear dimensions that correspond to an angular scale of $5^{\circ}$. We will later show how, crudely, a source with radius $\lesssim 3.7^{\circ}$ can be treated as a point source. The linear size that this angular scale corresponds to is, of course, a function of distance and cosmological assumptions:

$$x \equiv \theta \, d_A = \theta \, d_\ell / (1+z)^2 \, , \qquad\qquad [VII.42]$$

where $d_A$ is the effective angle distance (see e.g. Weinberg [1972]) and $d_\ell$ is the luminosity distance, described in Equation [A1.6] et seq.. The absolute value of this is set by the Hubble length, $\lambda_H = c \, / \, H_0$. We continue to assume that $H_0 = 50$ km s$^{-1}$ Mpc$^{-1}$, corresponding to $\lambda_H = 6$ Gpc. For different values of H, all distances in this section should be scaled by $50 \, / \, H_0 = h_{50}^{-1}$, e.g. by one half for $H_0 = 100$ km s$^{-1}$ Mpc$^{-1}$. Most current models for the origin of the X-ray background have the bulk of the emission originating at large redshifts, $z \gtrsim 1$. For $q_0 = 0.1$, $5^{\circ}$ is subtended at $z = 1$ by 180 Mpc. At $z = 2$ the luminosity distance has increased by a factor of 2.5, but geometric corrections to the angular scale substantially reduce the increase in $d_A$, so that this separation is 200 Mpc. As is well known, the behavior of the angular scale over this moderate redshift range is a sensitive function of $q_0$. If $q_0 = 0.5$, $z = 1$ corresponds to 150 Mpc, but at $z = 2$ the spatial scale for $5^{\circ}$ has decreased to 74 Mpc. All these scales refer to the physical size at the epoch of emission. If the structure undergoes collapse its size in the current epoch may be significantly smaller. Gravitationally bound stable objects will maintain their size through the intervening epochs, while the lowest over-density structures are generaly unbound and will participate in the general cosmological expansion. The size of such sources increases from the epoch of emission to the current epoch by a factor $\lesssim 1+z$.

Studies of the distribution of galaxies indicate significant correlations out to at least 10 Mpc [e.g. Peebles 1980], while studies of individual superclusters show evidence for individual structures in

the 10-60 Mpc range [Chincarini and Rood 1979; Gregory, Thompson, and Tifft 1981; Tully 1982] with extended filaments perhaps as long as 120 Mpc [Davis, Huchra, Latham, and Tonry 1982]. The separations between the superclusters are perhaps even more spectacular, showing voids and rifts of 120 to >300 Mpc in size [Kirshner, Oemler, Schechter, and Shectman 1981; Bahcall and Soneira 1982]. An early estimate of the mean free path between superclusters is ~400 $h_{50}^{-1}$ Mpc [Gregory, Thompson, and Tifft 1981].

We will begin by analyzing the additional contribution to $\sigma^2{}_S$ that clustered point sources will make, assuming that the clustering scale corresponds to angular dimensions small compared to the detector size. We next consider the effects of a single lump of emission, where it totally covers the measurement FOV. The distinction between the actual angular dimensions of a measurement and the quantity $\Omega_{eff}$, often used in other authors' discussion of the fluctuations, is outlined. The dipole signal of the Compton-Getting effect constitutes a very large angular scale phenomenon. We illustrate how comparisons between measurements of the dipole fits made in the X-ray and the microwave bands can provide information about the origin of the indicated peculiar velocity, but exact values are difficult to obtain. From individual lumps we move to a collection of single scale lumps with a Poisson spatial distribution. More complicated structure is handled by Fourier decomposition techniques, which require the calculation of the expected fluctuations from a arbitrary phase plane wave. Our observed upper limit to any other source of fluctuations, $\sigma^2{}_I$ in intensity units, can place interesting limits on structure at very large scales otherwise not easily accessible.

### AGN in superclusters: additional $\sigma^2{}_S$

Oort, Arp, and de Ruiter [1981] analyzed 12 close pairs of QSOs with similar redshifts and showed that their existence was consistent with the location of the objects in superclusters spanning 10 to 60 Mpc. Similar results were obtained by de Ruiter and Zuiderwijk [1982] who applied particular models for the correlation to an extended set of

data. If the X-ray sky is dominated by objects that are similarly distributed, their contribution to$\kappa$ the excess variance will be significantly enhanced.

The stochastic effects of clumping are reduced for the nearby objects, where the supercluster scale is much larger than the measurement scale. We concentrate on the evolved population (that in excess of the expected model N(S) from nonevolving AGN and cluster luminosity functions), whose average number versus flux relationship is N'(S) (see equations [VII.13] et seq.). We assume that for the redshifts at which these sources dominate the contribution to the sky flux and variance, the supercluster scale is less than the measurement scale. In this case we can let N'(S) represent the distribution of the fluxes for the superclusters as a whole, with the individual AGN only determining where a particular supercluster is on a supercluster luminosity function. A naive application of the limits on the number of sources (equation [VII.14]) determine that there must be at least 50 superclusters per square degree, but this assumes that they produce 75% of the sky flux and makes no correction for the finite angular extent of the closer examples (see the estimates of this latter in this section). Difficult as it is to estimate the total surface density of AGN, the equivalent project for superclusters would have significant additional problems. Determining the number of superclusters in a deep X-ray exposure suffers from identification problems, · while current optical exposures could not extend supercluster surveys to the low fluxes accessable with the AGN counts. In this subsection we will estimate how clumping modifies the required number of AGN, or to invert the problem, show how the number of AGN and the excess variance can constrain the amount of clumping.

The small angular scale condition also allows us to use the weighted response $R_2$ to transform from the measured variance $\sigma^2_I$ to a solid angle independent $\sigma^2_S$ to be compared with model predictions. We also assume that all the evolved sources are actually contained in superclusters. The superclusters are· assumed to be distributed throughout the universe following a Poisson distribution. The number of

sources in each supercluster is also determined by a Poisson
distribution, where $\mu_i$ is the mean number of sources in the ith
supercluster. $\mu_i$ need not be identical for all superclusters,
reflecting intrinsic dependence on supercluster properties such as size.
$\bar{\mu}$ is the mean over all superclusters.

Let n be the average number of sources in a measurement. The
variance will be enhanced by a factor, c,

$$\text{var}(n) = (1 + c)\, n \ . \qquad\qquad [VII.43]$$

The factor n on the right hand side of the equation is from to the
variance of a purely Poisson distribution, $\text{var}(n)_{Poisson} = n$. De Ruiter
and Zuiderwijk showed that the multiple Poisson process insures that

$$c \geq \bar{\mu} \ , \qquad\qquad [VII.44]$$

where the equality holds only for the case where all superclusters have
the same expected mean number of sources, i.e. $\mu_i = \bar{\mu}$. The variation
in the measured intensity is difficult to calculate directly because of
cross correlations between n(I) at two different intensities. In
general,

$$\sigma^2_{I} \geq \langle I^2 \rangle\, \text{var}(n) \ , \qquad\qquad [VII.45]$$

where $\langle I^2 \rangle$ is the mean square source intensity. The relationship is an
equality if all the sources in a particular supercluster have the same
intensity, i.e. identical luminosity. Converting to detector
independent quantities we can rearrange equation [VII.15] to give the
minimum number of sources,

$$N^* \geq 2.1 \times 10^6 \ (1+c) \ (f/0.75)^2 \ (\langle s^2 \rangle / \bar{s}^2) \ (9 \times 10^{-4} / \sigma^2_S) \ . \qquad [VII.46]$$

The factor f is the fraction of the sky flux due to sources clumped in
superclusters. The clumping factor (1+c) increases the number of
sources from the non-clumped minimum of 51 sources per square degree.
For the $\mu_i = \bar{\mu}$ case, the de Ruiter and Zuiderwijk analysis of close
pairs of quasars led to the estimate c = $\bar{\mu}$ ≈ 0.4. For a model
distribution of $\mu_i$, their data were consistent with $\bar{\mu}$ = 0.12, but the
effect of clumping was enhanced so that c = 0.6. One difficulty in
using these estimates in equation [VII.46] is that they are based on a
population that may be only a small fraction of the total evolved

Analysis of Results                                        Non-Poisson Fluctuations

population. In other words, to properly use [VII.46] with the derived values for $c_{QSO}$ we must replace f with the fraction of the all sky flux due to AGN that satisfy de Ruiter and Zuiderwijk's selection criteria, $f_{QSO}$, which is probably much less than $f_{AGN}$. In any case the numbers are still far below the minimum number of AGN inferred from the Einstein Deep Survey limits of at least 180 sources per square degree (equation [VII.27]) and the actual optical counts of AGN, $\geq 360$ deg$^{-2}$.

Rather than use the possibly incompatible estimates of c to provide the required number of objects, we can instead invert the problem to estimate from the observed number of AGN their excess contribution to the variance.

$$1 + c \leq f_{evAGN}^{-2} \ (<s^2>/\bar{s}^2)^{-1} \ N_{evAGN} \ (\sigma^2{}_S \ / \ S_{as}{}^2) \ . \qquad [VII.47]$$

The "evAGN" subscripts indicate that we are estimating the properties of the evolved AGN, whose N(S) and contribution to $S_{as}$ and $\sigma^2{}_S$ are in excess of the estimates assuming no evolution. If we assume that are 500 such sources per square degree, $2.1 \times 10^7$ all sky, that they are distributed roughly in a Euclidean manner over a decade of flux ($<s^2>/\bar{s}^2 \approx 1.5$), and that they provide the missing 75% of the total sky flux, then

$$1 + c \leq 6.54 \ (0.75/f_{evAGN})^2$$
$$\times \ (1.5 \ \bar{s}^2 \ / \ <s^2>) \ (N_{evAGN}/2.1 \times 10^7) \ . \qquad [VII.48]$$

We have assumed that $\sigma^2{}_S$ is equal to our upper bound of $9 \times 10^{-4}$ (H1L1)$^2$. For this choice of parameters, then, a strong upper bound on $\bar{\mu}$ is 5.54, assuming that all superclusters have the same number of AGN. If a more realistic distribution is similar to the models discussed by de Ruiter and Zuiderwijk then $\bar{\mu}$ is $\leq 1.4$.

How does this compare with expected numbers of AGN in a supercluster? From the AGN luminosity function we estimate that $4.5 \times 10^{-3}$ of all galaxies with $M_B < -18$ will be an AGN with $L > 0.1 \ L_{44}$, and therefore a candidate for significant evolution. The local supercluster contains ~1500 galaxies, extrapolating from the sample of Tully [1982], so that roughly seven moderate luminosity AGN are

expected. This estimate of $\bar{\mu} \approx 7$ contradicts our upper bound based on the fluctuations. The local supercluster is considered a small example of the supercluster phenomenon so that $\bar{\mu}$ might be substantially larger. If the number of AGN is reduced, or if they are spread over a still larger flux range, the disagreement with the prediction of equation [VII.48] is strengthened. If we allow $f_{evAGN}$ to vary and hold to our original estimates for the other parameters, and if the true value for c is $\sim 4 \times \bar{\mu} \gtrsim 4 \times 7$ then

$$f_{evAGN} \lesssim 0.4 \ . \qquad\qquad\qquad\qquad [VII.49]$$

In this crude picture at most $\sim 1/2$ of the sources that produce the X-ray sky flux can be in superclusters. Admittedly, there are large uncertainties in this calculation, not the least of which are the assumption that supercluster size is $\lesssim 5^{\circ}$ and the inability to handle the clustering at larger angular scales. In addition, the very concept of the number of AGN "per supercluster" implies that superclusters exist as definable individual entities. Even on the smaller scale of clusters of galaxies there are examples of complicated structures that have been variously identified as several objects or as a single object with internal structure. However, this discussion shows how studies of optical counts, and optical superclustering, when joined to the upper limits on the X-ray excess variance using more sophisticated models, would place limits on the contribution of evolved sources in superclusters to the sky flux. Inverting the problem, when future experiments give unambigous estimates of the contribution such sources make to the H1L1 band sky flux the HEAO 1 excess variance limits constrain their participation in non-Poisson clumping.

## Large scale structure: a single lump

We now investigate measurements from structures that have a large angular size, covering areas large with respect to the detector solid angle. The remainder of this section is heavily grounded on previous analysis by others [e.g. Rowan-Robinson and Fabian .1974; Rees 1980; Fabian, Warwick, and Pye 1980; McKee 1980; Warwick, Pye, and Fabian 1980; Fabian 1981]. However, the discussion will be complementary to our earlier derivation of the point-source fluctuations in order to

illustrate how and where the measurement solid angle response enters. We will generally concentrate on simple examples in a finite Euclidean space; more accurate models are presented in the above references.

Consider a lump of excess volume emissivity, $\delta\Lambda$. Let $\lambda$ be the radial distance coordinate from the observer to the source. The lump will produce an excess of surface brightness

$$\delta\Sigma = \int_{lump} d\lambda \; \delta\Lambda \; . \qquad\qquad [VII.50]$$

Let L be the total integrated luminosity of the lump and $\lambda_s$ be its radius. We can approximate $\delta\Sigma$ by the average surface brightness over a circle with angular radius $\theta_s = \lambda_s / \lambda$,

$$\delta\Sigma = L \; / \; 4\,\pi^2\,\lambda_s^{\;2} \; . \qquad\qquad [VII.51]$$

Euclidean geometry assures that the surface brightness is constant. The contribution to the intensity, $\delta I$, depends on the detector spatial response function (see equation [II.10]). We model a circular detector with flat response, R, and a solid angle of $\pi\,\theta_d^{\;2}$. We can define two natural regimes for $\delta I$, depending on the distance to the lump. When the lump is so close that its angular size is much larger than the detector solid angle, then if the angular position of the lump is close enough to the detector center to be detected at all by a measurement ($\theta < \theta_d + \theta_s$) it usually completely fills the detector and contributes a constant intensity. If the lump subtends a small angle, it will be essentially point like and its intensity will follow the inverse-square law.

$$\delta I \approx R\,\pi\,\theta_d^{\;2}\,\delta\Sigma \qquad , \; \lambda \ll \lambda_d \; ; \qquad\qquad [VII.52.a]$$

$$\delta I \approx R\,L \; / \; 4\,\pi\,\lambda^2 \; , \; \lambda \gg \lambda_d \; ; \qquad\qquad [VII.52.b]$$

where

$$\lambda_d \equiv \lambda_s \; / \; \theta_d \qquad\qquad [VII.53]$$

is the distance where the lump would exactly fill the detector response. For our surface brightness measurements, we approximate R as 220 cm$^2$ s, and $\theta$ as 4.2°. Lumps in the $\lambda \ll \lambda_d$ range will be called resolved sources, while lumps in the further regime are point-like. Note the difference between resolved and "detected". A lump may cover several measurements, but still be of such low surface brightness that it could not · be

detected as an individual feature. We will see, though, that undetected
lumps, like undetected point sources, have measurable effects on the
fluctuations measurements.

It is plausible to assume that the lump's large scale excess in
the volume emissivity is indicative of a similar mass density
concentration, i.e. that on these large scales the mass to X-ray light
ratio is constant. If the lump is in the detector covering regime,
$\lambda \langle \lambda_d$, then the overdensity is

$$\delta\rho \propto \delta\Lambda \approx \delta_I / \lambda_s \; \pi \; \theta^2 \; . \qquad\qquad\qquad [VII.54]$$

The same sources of the X-ray emission clumped in the lump also
contribute to the total X-ray intensity, $\overline{I}$. If they are otherwise
uniformly distributed out to the Hubble length then the average mass
density is

$$\overline{\rho} \propto \overline{\Lambda} \approx f_c \; \overline{I} / \lambda_H \; \pi \; \theta^2 \; . \qquad\qquad\qquad [VII.55]$$

Therefore the fractional size of the X-ray intensity fluctuation from
the lump is [Fabian 1981]

$$\delta_I / \overline{I} \approx f_c \; ( \; \delta\rho / \overline{\rho} \; ) \; (\lambda/\lambda_H) \; . \qquad\qquad\qquad [VII.56]$$

The value to be used for $f_c$, the fraction of $\overline{I}$ due to sources similar to
those in the lump, depend on models for the X-ray sky flux and the
position of the lump. For instance, if we assume that 75% of the sky
flux is due to highly evolved sources at large redshifts, and the lump
is at a similar distance then $f_c \sim 0.75$. (If the redshift region that
dominates this emission is sufficiently narrow, $\lambda_H$ may need to be
rescaled in [VII.56].) On the other hand, if the lump is nearby then $f_c$
$\sim 0.25$, indicative of the contribution from the non-evolving population
or other sources of emissivity typical of space in general.

### Large scale structure and the Compton-Getting effect

A large scale overdensity of matter would have other effects
besides an enhancement of the X-ray surface brightness. It would
perturb nearby objects and impart a peculiar velocity with respect to
the cosmological rest frame. Just such a lump may be responsible for
our own peculiar velocity inferred from the microwave background dipole

anisotropy [e.g. White 1979]. For a given distribution of mass fluctuations, one particular scale is expected to dominate the process, with our velocity principally generated by the nearest such lump. (For a more detailed picture of the origins of the microwave dipole see Wilson and Silk [1981], and Silk and Wilson [1981].) The presence, or lack, of a similar large scale structure in the X-ray sky can provide insight into the location and magnitude of such a lump [e.g. Fabian, Warwick, and Pye 1980; Warwick, Pye, and Fabian 1980], and in a general sense the magnitude or upper bound on the signal can lead to general conclusions about the structure and density of the universe (see the previous references; also Wolfe [1970], Rees [1979], Fabian [1980], and Kaiser [1982]).

In Figure VII.31.(a) we illustrate schematically the effects of a very large lump located predominantly outside the X-ray sources. Because of its large scale, the lump will accelerate all of the X-ray sources, however not all sources will have the same acceleration or current velocity [Warwick, Pye, and Fabian 1980]. In this crude picture, the relative velocity observed by the current epoch observer would have a general "12-hour" characteristic: at the poles, oriented towards and away from the lump, a perceived outward flow would cause an intensity decrement; and around the equator there would be a general inward flow, producing an excess intensity (see equation [IV.12]). If the distance scale to the lump were even larger all X-ray sources would be similarly accelerated, reducing the strength of the large scale feature. In both cases the generated peculiar velocity would still induce the observed dipole signal in the microwave background. However, the strength of a 12-hour signal from a large lump could be significantly altered when accurate cosmological adjustments are applied to these qualitative arguments. The largest relative velocity sources are at the edges of the X-ray emission region, i.e. at the highest redshifts. At the epoch of emission these sources would be relatively closer to the lump, thereby increasing the amount of acceleration, but all sources would similarly benefit. On the other hand the high redshift objects are observed when the amount of elapsed time for acceleration is less. Possible evolution of the lump mass further

Figure VII.31: Large scale anisotropies from an accelerating lump

Schematic picture of a possible origin for the observed
peculiar microwave velocity (after Warwick, Pye, and
Fabian [1980]). The arrows represent the induced peculiar
velocity at that point. The central filled dot, labelled
"observer", represents the current epoch observer. Two
cases are illustrated: (a) a very large lump, existing at
the edge of where most of the X-ray emission is produced,
resulting in a "12-hour" tidal distortion in the X-ray
surface brightness; and (b) a smaller more local lump
which does not produce a significant acceleration of most
of the X-ray sources.
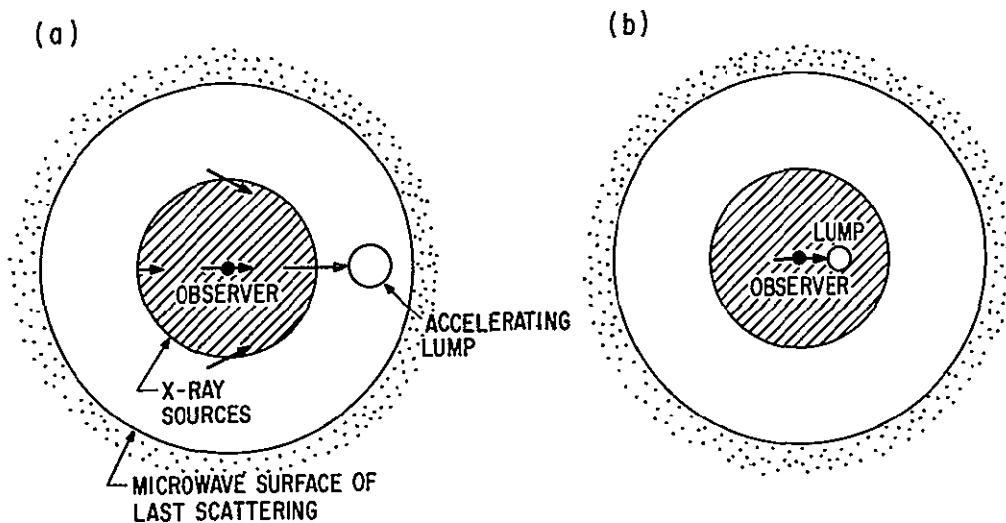


complicates the picture. The 12-hour intensity amplitude would also be
reduced to the extent that the X-ray emission is produced over a
significant range of redshifts. In Figure VII.31.(b) the picture is
less complicated. The differential velocity is due to accelerations
induced by a comparatively nearby lump, which does not itself strongly
perturb the velocity field of the bulk of the X-ray emitters. In this
case both the X-ray and microwave sky should exhibit similar dipole
anisotropy.

If there is a surface brightness enhancement associated with the

lump, the X-ray signal will have a larger amplitude than would otherwise be expected. Consider a circularly symmetric surface brightness enhancement, $\delta I_{lump}$, which is much larger than the detector angular extent. If we unknowingly try to fit this enhancement with a dipole signal, the least squares solution will be one with amplitude $\delta I_{lump}$ centered on the enhancement. To first order the total dipole signal indicated by the fit in the X-ray band will be

$$\delta I_{dipole} \approx \delta I_{CG} + \delta I_{lump} \ . \qquad\qquad [VII.57]$$

$\delta I_{CG}$ is the exact dipole signal from the peculiar velocity (equation [IV.12]). Wilkinson [1983], in a recent synthesis of several groups' microwave observations, estimated that the velocity* is $v = 372\pm25$ km $s^{-1}$. This corresponds to a predicted X-ray magnitude of $\delta I_{CG} \approx 0.072$ counts $exp^{-1}$. This is less than the best fit X-ray dipole amplitude of $0.09\pm0.03$ counts $exp^{-1}$, but not significantly so. The microwave dipole direction, in galactic coordinates $(\ell,b)$, is $(265^{\circ},+50^{\circ})$, which is well within the 90% contour for the X-ray direction (Figure IV.7).

From the 1 $\sigma$ range in $\delta I_{dipole}$ we estimate an upper limit of $\delta I_{lump} \lesssim 0.05$ counts $exp^{-1}$. As the lump is assumed to fill the detector, its average surface brightness is $\Sigma_{lump} \lesssim 1.4\times10^{-2}$ H1L1 $sr^{-1} \approx 1.8\times10^{-10}$ ergs $s^{-1}$ $cm^{-2}$ $sr^{-1}$. Then from equation [VII.56],

$$(\delta\rho/\rho)_{lump} \lesssim 3\times10^{-3} \ f_c^{-1} \ (\lambda_H/\lambda_s) \ . \qquad\qquad [VII.58]$$

A cosmologically correct derivation could modify this limit by a factor of a few; however the greatest uncertainty is in $f_c$.

We can estimate the magnitude of the peculiar velocity, $V_p$, if we assume that we are near the edge of a uniform spherical lump that has

---

*The use of the microwave dipole signal to derive the velocity assumes there is no similar surface brightness enhancement from the lump, perhaps due to microwave or infrared sources associated with galaxies. Our own galaxy does emit enough radiation to require significant corrections to some microwave anisotropy measurements. The possibility of such contamination in the dipole magnitude can be limited by searches for fluctuations in the microwave band on moderate angular scales and by checking that the variation in the dipole amplitude versus spectral frequency is consistent with a $\sim3^{\circ}$ black body spectrum. Current limits on both tests support this assumption.

been acting on us for about the Hubble time:

$$V_p \sim G \; \delta\rho \; \lambda_s \; \lambda_H \; . \qquad\qquad\qquad\qquad [VII.59]$$

As pointed out by Fabian [1981] we can combine this estimate with our limit on $\delta\rho/\rho$ and the observed microwave dipole velocity to place a lower limit on the average mass density of the universe.

Before using the observed velocity for $V_p$, we must correct for the rotational motion of the galaxy, 220 km s$^{-1}$ in the direction $(90^\circ, 0^\circ)$, so that the adjusted direction of $V_p$ is $(267^\circ, +32^\circ)$, and the magnitude is 540 km s$^{-1}$. This new direction is even closer to the best fit X-ray dipole position of $(282^\circ, +30^\circ)$. As the direction for $\delta I_{lump}$ is not affected by galactic rotation, it is more likely that the X-ray dipole signal might contain some contribution from the lump surface brightness, but strong statements are precluded by the large confidence region for the X-ray dipole position. $\rho$ is often parametrized in terms of the critical density for closure in the Friedmann cosmological models:

$$\Omega \equiv (8/3) \; \pi \; G \; \rho \; / \; H_o^2 \; , \qquad\qquad\qquad [VII.60]$$

so that $\rho = 4.7\mathrm{x}10^{-30} \; \Omega \; h_{50}^2$ g cm$^{-3}$. Our crude lower limit is

$$\Omega > \sim 0.6 \; f_c \; . \qquad\qquad\qquad\qquad\qquad . \; [VII.61]$$

Other estimates of $\Omega$ cover a wide range, $0.1 < \Omega < \sim 1$, so our limit could be interesting. However, cosmological corrections to equations [VII.59] and [VII.56] can be as much as an order of magnitude.

Moreover, the utility of [VII.61] hinges on a good lower bound for $f_c$. If the lump is nearby, so that the lumped emission comes from at least some sources typical of the XCS, i.e. the unevolved populations of clusters and AGN, then we might expect that $f_c \gtrsim 0.2$, the lower estimate of the unevolved populations' fractional contribution to the X-ray sky flux. However, several authors have suggested that the peculiar velocity is due to our infall into the local supercluster [e.g. White 1980; Davis, Tonry, Huchra, and Latham 1980; Davis and Tonry 1981]. The direction of $V_p$ is generally towards the supercluster, although there is a significant velocity component orthogonal to the direction of the Virgo Cluster and M87 in the supercluster center. (This is addressed in

the above references; see also .White and Silk [1979]). Although ·
previous searches for emission associated with the local supercluster
have failed, the derived upper limits for the surface brightness are
above our value inferred for $\delta I_{1ump}$ [e.g. Schwartz 1980]. If the local
supercluster is indeed the velocity-inducing lump, this would have a
major impact on the proper value. for $f_c$ to use in the limits [VII.58]
and [VII.61]. This is because many of the unevolved sources in the
local supercluster were actually detected in the XCS and therefore
explicitly removed from the data used to fit the dipole signal. Sources
in the low luminosity extension of the AGN luminosity function will
still contribute to $\delta I_{1ump}$ as few were detected in the XCS. In general,
though, $f_c$ could not be much more than 0.1. However, the nearness of
the local supercluster as well as detailed studies of its optical
morphology [Tully 1982] hold out the prospect for a more accurate
determination of $\delta I_{1ump}$, justifying the effort required for a
cosmologically proper derivation of [VII.56].

### Multiple lumps and the excess variance

The preceeding subsections outlined the properties of a single
lump, and showed what information could be gained from the detection of
(or upper bound to) the surface brightness enhancement from that lump.
In the same way that the fluctuations placed bounds on the unresolved
point sources, the excess variance from the fluctuations can yield
information about the general properties of individually undetectable
lumps. As an example of this ability we consider a Euclidean universe,
of radius $\lambda_H$, consisting entirely of lumps of matter with radius $\lambda_s$.
The lumps are not considered hard spheres, but are allowed to overlap so
that some regions will have the mass density of several times the
density from a single lump. Let $\eta$ be the mean number density of lumps
in the universe. The filling factor,

$$\Phi \equiv \eta \ (4/3) \ \pi \ \lambda_s^3 \tag{VII.62}$$

is the mean number of lumps that enclose a point chosen at random. If
the contribution to the mass density of a single lump is $\delta \rho_1$ then the

mean mass density is

$$\rho = \Phi \ \delta\rho_1 \ . \qquad\qquad\qquad\qquad [VII.63]$$

Therefore the fractional size of the lump density is $\delta\rho_1/\rho = \Phi^{-1}$. In contrast, the rms variation in the density through the universe is

$$\delta\rho_{rms} \ / \ \rho = (\Phi \ \delta\rho_1{}^2)^{1/2} \ / \ \rho = \Phi^{-1/2} \ , \qquad\qquad [VII.64]$$

assuming a Poisson spatial distribution for the lumps.

For distances where the lump extent completely covers the detector, $\lambda < \lambda_d$, a lump will be detected if its center is within $\lambda_s + \theta_d \ \lambda \approx \lambda_s$ of the center of the observation. The mean number of sources expected over a range $d\lambda$ is then

$$dN \approx \eta \ \pi \ \lambda_s{}^2 \ d\lambda \ . \qquad\qquad\qquad [VII.65]$$

From the constant contribution such sources make to the measured intensity (see equation [VII.52]), their expected contribution to the excess variance is

$$d\sigma^2{}_I = (\delta I)^2 \ dN$$

$$\approx A \ \lambda_H{}^{-1} \ L^2 \ \eta \ \ell_s{}^{-2} \ d\ell \ . \qquad\qquad [VII.66]$$

The coefficient

$$A \equiv R^2 \ (\theta_d)^4 \ / \ 16 \ \pi \qquad\qquad\qquad [VII.67]$$

contains the detector parameters. We have also renormalized our linear variables by the Hubble length, $\ell_s \equiv \lambda_s \ / \ \lambda_H$ ; $\ell \equiv \lambda \ / \ \lambda_H$. Sources at distances $\lambda > \lambda_d$ can be approximated as point sources so that

$$d\sigma_I{}^2 \approx R_2{}^2 \ (L/4\pi\lambda_s{}^2)^2 \ \eta \ 4\pi \ \lambda^2 \ d\lambda \qquad\qquad [VII.68]$$

(see equation [III.28]). $R_2$ is the second weighted detector response (equation [II.17] et seq.), which for a flat response circular collimator is

$$R_{2;est} = R \ \theta_d \ / \ 2 \ . \qquad\qquad\qquad [VII.69]$$

For our approximation of the smeared detector with $R = 220 \ cm^2$ s and $\theta_d = 4.2°$, we have $R_{2;est} = 7.7 \ cm^2$ s. Calculations using the actual response produce a value of $9.04 \ cm^2$ s, an indicator that the peaked properties of the detector enhance the variance over a flat response.

If we define a scaling factor

$$T \equiv R_{2;true} / R_{2;est} = 1.2 \qquad\qquad [VII.70]$$

we can rewrite [VII.68] for the point-like contribution of lumps to the excess variance as

$$d\sigma^2{}_I \approx A \; \lambda_H^{-1} \; L^2 \; \eta \; (T/\theta_d)^2 \; \ell^{-2} \; d\ell \; . \qquad\qquad [VII.71]$$

We see that the two estimates for $d\sigma^2{}_I$ are equal at the transition point

$$\lambda_{trans} = (T/\theta_d) \; \lambda_s = T \; \lambda_d \; , \qquad\qquad [VII.72]$$

which is equivalent to the distance where the lump of size $\lambda_s$ can exactly fill a detector with angular radius $\theta_d/T$. Therefore, our smeared detector is effectively one with a $3.7^\circ$ radius in terms of the transition from point sources to resolved lumps.

As we consider nearby extended sources, $\lambda < \lambda_{trans}$, we see by [VII.65] that the number seen at a particular distance is enhanced over that expected for the same density of point sources, $dN = \eta \; \pi \; \theta_d^2 \; \lambda^2 \; d\lambda$. Yet as the contribution of extended sources saturates, the contribution to the variance of the fluctuations is convergent even for the very closest sources. Thus there is no need to have an intensity cutoff, $I_u$, as was necessary in the variance analysis of the point source fluctuations (see equation [III.54] et seq.). No longer is it appropriate to say that the variance is dominated by the nearest source, but rather it is affected equally by all the sources that cover the detector. In addition the fact that there are several sources with the brightest contribution to the intensity also reduces the non-Gaussian aspect of the fluctuations distribution. Therefore the excess variance is sufficient to fit the fluctuations from a lump population that contains a number of resolved extended sources. Another effect of large scale structure is to produce additional correlations between overlapping and adjacent measurements. These correlations will, as we have discussed, introduce a reduction in the effective number of independent measurements. This will lower the confidence of the particular $\Delta\Lambda$ contours, unless we re-weight the measurements (Section A2:2). Note that this correlation does not bias the measured enhancement. The correction to the weighting should be small compared

to that already calculated from the point sources which dominate the magnitude of the fluctuations.

The total variance is calculated by integrating $d\sigma^2_I$ out to the Hubble distance, $\ell = 1$:

$$\sigma^2_I = \lambda_H^{-1} \, A \, F(\ell_s) \, L^2 \, \eta \, , \tag{VII.73}$$

where $F(\ell_s)$ contains all the dependence on the physical size of the lumps. The total intensity is

$$\bar{I} = \lambda_H \, R \, \pi \, \theta_d^2 \, L \, \eta \, . \tag{VII.74}$$

In equation [VII.15] we were able to use the upper limit to the invariant excess as a fraction of the total sky flux in order to place a lower limit on the total number of point sources. In similar fashion equations [VII.73] and [VII.74] can be combined to derive a lower limit on $\eta$, the density of clumps with a given radius $\lambda_s$. We can then use [VII.62] to derive an upper bound on

$$\Phi^{-1} = \frac{12 \, \pi^2 \, p^2}{\ell_s^3 \, F(\ell_s)} \, , \tag{VII.75}$$

$$p^2 \equiv \sigma^2_I \, / \, \bar{I}^2 \, . \tag{VII.76}$$

The value of $\Phi^{-1}$ directly limits the single lump density as well as the general rms density variation (equation [VII.64]).

When we calculate $F(\ell_s)$ we note there are two regimes, depending on the transition distance, $\ell_{trans}$, which is a function of the lump size. If $\lambda_{trans} > \lambda_H$, the limit of the integration, then all the lumps will be resolvable and

$$F(\ell_s) = \ell_s^{-2} \, ; \, \ell_s > \ell_{crit} \equiv \theta_d \, / \, T \, . \tag{VII.77}$$

$\ell_{crit}$ marks the critical lump radius where the lump would just fill the effective detector at the Hubble distance. Once a detector has a large enough angular size that $\ell_s$ is greater than $\ell_{crit}$, further increases in opening angle will not modify F, and therefore the expected fractional

Analysis of Results                                    Non-Poisson Fluctuations

C - S

size of the fluctuations. For these very large scale lumps,

$$\delta\rho_{rms} / \rho \approx 11 \cdot p \ (\lambda_H/\lambda_s)^{1/2} \ . \tag{VII.78}$$

Similar behavior for large lumps was found by Fabian [1981], though his limit on $\delta\rho/\rho$, omitting the factor of 11, is an order of magnitude more stringent for a given value of p. As the size of the lumps decreases, more and more of the distant sources will behave as point sources and their contribution to the fluctuations variance reduced:

$$F(\ell_s) = (2 \cdot / \ \ell_s \ \ell_{crit}) - (1 \ / \ \ell_{crit}^2) \ . \tag{VII.79}$$

This weakens the restriction on the filling factor, allowing larger density. For the smallest lumps the maximum allowed $\delta\rho_{rms}/\rho \propto \ell^{-1}$. The change in the index from $-1$ to $-1/2$ occurs for lumps $\sim\lambda_{crit}$ in size. For comparison, note that the limit in equation [VII.58] derived from the upper limits to the lump surface brightness associated with the Compton-Getting effect, $(\delta\rho/\rho)_{lump} \propto \lambda_s^{-1}$, is a limit on the lump excess density, and not on the rms universal variation. If we were to estimate the rms density from the Compton-Getting single lump limit, given that $\delta\rho_{rms}/\rho = (\delta\rho_1/\rho)^{1/2}$ (see equation [VII.64]), we would see that it would be $\propto \lambda_s^{-1/2}$. This behavior is the same as for the always resolved sources, consistent with the assumption behind the derivation of [VII.58] that the Compton-Getting lump was resolved.

Our model is admittedly crude, particularly in neglecting cosmological effects, which are undoubtedly significant. For example, surface brightness is not a constant. For an object of fixed size and luminosity, $\Sigma \propto (1+z)^{-4}$. On the other hand, if the lumps do not have a fixed absolute size, but rather maintain their co-moving volume, $\Sigma \propto (1+z)^{-2}$. Similar modifications would be necessary if the lump luminosity were redshift dependent. As an example, if the lump emits via thermal bremsstrahlung, then the luminosity at a fixed frequency is $\propto (1+z)^2$ (ignoring spectral redshift corrections). For such objects the surface brightness is redshift independent. Of course there are other geometric corrections as well.

A modification to our simple Euclidean model can simulate the effect of cosmological source evolution. We assume that X-rays are

emitted only by lumps confined to the outermost part of the universe-sphere in a shell of thickness $\Delta$ $\lambda_H$. The total intensity will simply scale with $\Delta$, so that a factor of $\Delta^2$ appears in the numerator of equation [VII.75]. The effect on the variance is more complicated, necessitating the replacement of $\bar{F}(\ell_s)$ with $F(\ell_s, \Delta)$. For the always resolved large scale lumps, $\ell_s > \ell_{crit}$, and $F = \Delta / \ell_s$, so that the density limit

$$\delta\rho_{rms}/\rho \approx 11\ \Delta^{1/2}\ _p\ \ell_s^{-1/2}\ ,  \qquad\qquad\qquad\text{[VII.80]}$$

is strengthened by increasing amounts of evolution ($\Delta \rightarrow 0$). For sources with $\lambda_s < (1-\Delta)\ \lambda_{crit}$, all these smaller lumps are distant enough that none are resolved. Therefore $F = (T/\theta_d)^2\ \Delta\ /\ (1 - \Delta)$ is, as expected, independent of the lump size. The intensity variance is similarly unaffected, but

$$\delta\rho_{rms}/\rho \propto \ell_s^{-3/2}\ .  \qquad\qquad\qquad\qquad\qquad\text{[VII.81]}$$

For intermediate scales $\lesssim \lambda_{crit}$ evolutionary models will have more stringent limits on density fluctuations for a given observed limit on the intensity fluctuations. However the steeper index insures that at some small scale, the evolutionary model constraint will be less severe than a similar scale no-evolution model. More complicated intermediate behavior is seen for lumps with $(1-\Delta)\ \lambda_{crit} < \lambda_s < \lambda_{crit}$, where there are both point-like and resolved lumps.

To get a feel for the limits we might obtain later with more rigorous models, we present in Figure VII.32 the limits on $\delta\rho_{rms}/\rho$ for the simple Euclidean model we have examined here. For our effective detector size, the critical scale size delimiting all-resolved and some point-like lumps is $\lambda_{crit} \approx 390\ h_{50}^{-1}$ Mpc. A detector response that covered less solid angle would lead to more stringent limits on $\delta\rho_{rms}$ at scales below 390 Mpc. The limit for $\sigma^2_I$ that we have used, $\approx 0.07$ (counts $\exp^{-1})^2$, has not been corrected for AGN evolution, or for the possible clumping of point sources. In this subsection we have not made any adjustment for the possibility that some fraction of the X-ray emissivity from a given lump is due to point sources. As we saw earlier in this section, the application of Poisson statistics to such sources would enhance the fluctuations and decrease the allowed amount of

Figure VII.32: Limits on large scale structure imposed by X-ray fluctuations bounds

We plot the upper bounds on the rms density variations due to a single scale of lumps versus the lump radius. No cosmological corrections are made. $\sigma_I < 1.6\%$ of $\bar{I}$. For the evolutionary model, $\Delta = 0.2$. See text for further details.
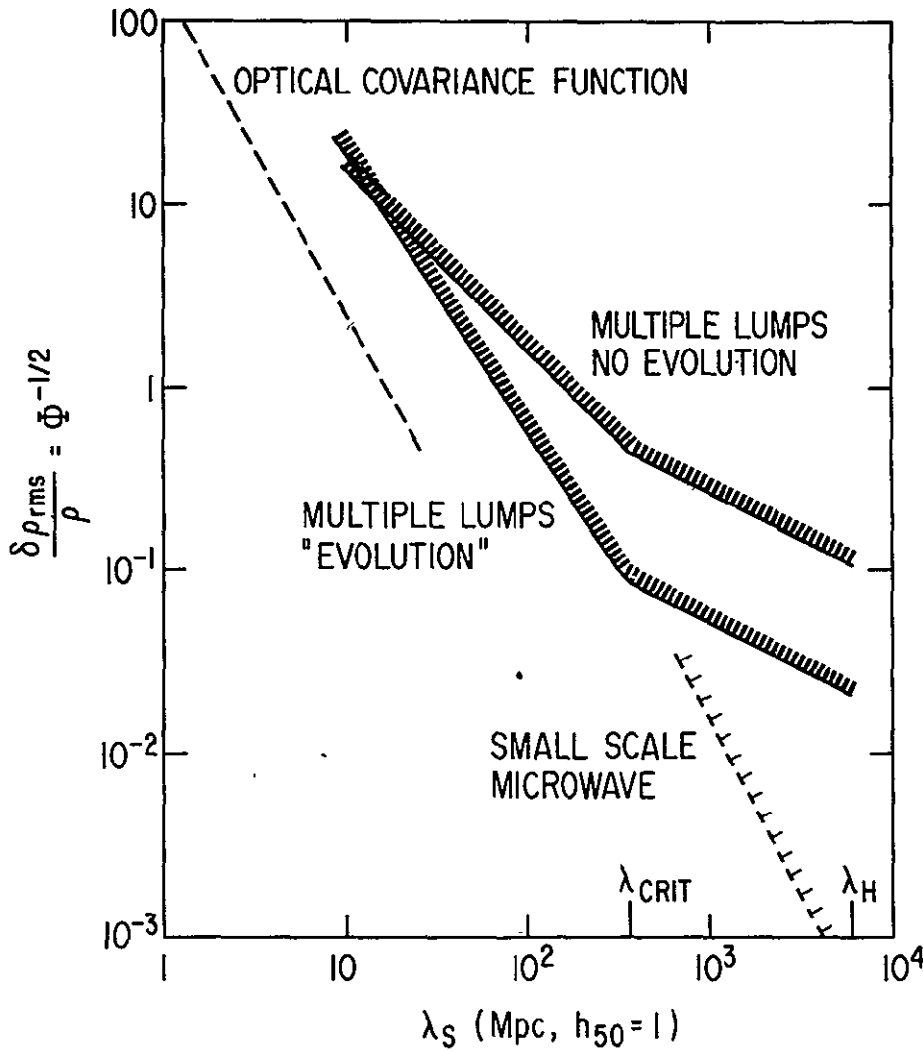


$\delta\rho_{rms}/\rho$. Figure VII.32 presents a limit for no evolution and for the simple evolution model with $\Delta = 0.2$. The other dashed lines were taken from a similar figure by Fabian [1981], where the optical correlation

Analysis of Results                                    Non-Poisson Fluctuations

line indicates actual clumping deduced from the observed angular distribution of galaxies [Peebles 1974]. The limit at the very largest linear scales is based on limits to the small scale fluctuations of the microwave background [Partridge 1980]. Even with our crude techniques and relatively conservative limit on the excess variance we see that the X-ray sky fluctuations provide an interesting probe of structure of the universe over a large range in scales. The region from ~100 Mpc to 6000 Mpc is particularly difficult for techniques other than the X-ray sky variations, although Osmer [1981] hopes to cover some fraction of this range through searches for large angular scale clumping of QSOs.

## A note on appropriate detector size

Previous authors' techniques for calculating the excess variance due to clumping have involved more accurate estimates of the cosmological effects (see particularly McKee [1980]) than we have used here. However, a certain imprecision of language has led to the use of the wrong values for the detector size in applying their formulae. Typically the detector size was based on the so-called effective solid angle, $\Omega_{eff}$ (equation [III.53]). According to our derivation, $\Omega_{eff}$ includes a dependence on the units of the detector spatial response function, $r(\theta,\phi)$ in addition to units of solid angle. This is normally overlooked in most analyses of the fluctuations, where the same units are used for intensity and flux so that the response is dimensionless. To see that $\Omega_{eff}$ is indeed not the appropriate variable to use, we consider the effects of smearing the detector response. As the solid angle covered by the detector is increased, the average response must decrease. The ratio of surface brightness to measured intensity is thus constant, but the effective solid angle actually decreases. For example, we have commented on two sets of UHURU analysis: Fabian [1975] smeared by $5°$ the $5.2°\text{x}5.2°$ detector, yielding $\Omega_{eff} \approx 5$ msr; while Schwartz et al. [1975] smeared the detector by $10°$, to get a smaller $\Omega_{eff} \approx 4$ msr. If $\Omega_{eff}$ is interpreted as the effective size of the collimator, then by spreading the detector around we are presumably becoming sensitive to structure on smaller scales. An additional erroneous aspect to using $\Omega_{eff}$ in calculating the fluctuations variance

is that it applies to the case of sources distributed in intensity $\propto I^{-5/2}$, and that only sources below a certain intensity limit contribute (see [III.56] et seq.). For a population of lumps, none of which have been individually detected, one or both of these assumptions will not apply.

We can make a better estimate of an effective detector size using the weighted detector responses we have developed. As noted before, $R_2$ is appropriate for measurements of the fluctuations from any distribution of point sources, when not selected by a critical upper intensity value. For resolved sources, the appropriate weight to use is $R_1$ which converts uniform surface brightness to intensity. In our simplified model for the contribution to lumps, the effective size of the detector should correspond to the angular size of a source at the demarcation point between resolved and point-like behavior in its contribution to the excess variance. In our analysis, the transition point is reached when a source radius subtends an angle

$$\theta_{eff} = 2 \, R_1 \, / \, R_2 \, . \qquad\qquad\qquad\qquad [VII.82]$$

For the smeared detector we used this is the $3.7^o$ size derived above. For comparison, the smeared UHURU detector of the Fabian analysis has $\theta_{eff} = 4.8^o$, while the rebinned data Schwartz [1980] used in his estimate of the excess variance has $\theta_{eff} = 7^o$. These values, crude as they are, give a feeling for the range of structure most constrained by a limit on the excess variance, i.e. those with

$$\lambda_s > \lambda_{crit} = \theta_{eff} \, \lambda_H \, . \qquad\qquad\qquad [VII.83]$$

### Fluctuations from continuous density ripples

In placing limits on the density enhancements of lumps of matter we have so far assumed that all such lumps are the same size. If the lumps were to have a range of sizes, the limit from the excess intensity on the amount of structure allowed at a particular scale would be stronger. Such a range in structure size is assumed by many studies of large scale phenomena in both the early and current epochs [e.g. Peebles 1980]. Rather than being a set of spherical lumps, the universe

is decomposed into a set of plane waves, of wavelength $\lambda$ and amplitude $\delta\rho(\lambda)$, that often have a power-law form for their power density spectrum.

A detailed analysis of such phenomena, even for the finite Euclidean universe we used above, is beyond the scope of this exposition. We note, without proof, a few general properties that have been worked out in preparation for future analysis. The total excess variance from all scales is the integral of the contribution to the excess at each scale weighted by $(\delta\rho(\lambda))^2$. For fixed amplitude the excess variance is proportional to $\lambda^2$, equivalent for a Poisson distribution of lumps in the strongly constrained always-resolved regime. However, the contribution to the fluctuations from a given $\delta\rho$ is generally less than from the same $\delta\rho_{rms}$ of the Poisson lumps. This is because the structure is strongly regular and correlated, so that the fluctuations are reduced, as we discussed in the beginning of this section. The contribution is equivalent to only one or a few Poisson lumps in the field of view. Small solid angle detectors will have larger fractional variance, at all scales, although the calculation is complicated. The effects of evolution are much reduced unless the scale lengths of the evolutionary behavior are comparable to or smaller than $\lambda$.

The examination of realistic models of multiple scale variations, in light of the X-ray excess variance limit, probes the universe near the critical epoch of galaxy formation.

CHAPTER VIII .

THE X-RAY SKY FLUX: SPECTRUM, MODELS, AND FUTURE PROSPECTS

Previous chapters have discussed the anisotropy of the X-ray sky, but there is a separate approach that gives important information. A successful model of the sky flux, besides describing its spatial distribution, must also explain the observed spectrum. In this chapter we will summarize the spectral properties of the X-ray sky, as well as the various physical processes that might produce such a spectrum. Several models for the origin of the X-ray sky are examined in the context of the spectrum. One model, with substantial flux generated by a hot intergalactic medium (IGM), is considered in more detail as the observed limits on the excess variance have an explicit role in constraining variaitons of the model. This chapter also includes a discussion of the results that future, more detailed, examinations of the HEAO 1 A-2 database should produce. Future studies of the sky flux are outlined and contrasted.

———————————————

1: The Observed X-ray Sky Spectrum and Physical Processes

Recent measurements of the X-ray sky spectrum from 3 keV to 400 keV are reviewed. A thin thermal bremsstrahlung spectrum with $kT \approx 40$ keV is a good description of the total spectrum from 3 to 100 keV. Other possible interpretations of the data are presented. The effective spectrum, after removal of an estimated component due to known sources, is described. We summarize briefly the physical mechanisms that can produce X-ray emission.

Our understanding of the X-ray "diffuse background" and the astrophysical implications of its flux has evolved over the past decade (compare reviews by Silk [1973] and Rees [1980]). Two basic criteria are still used to evaluate and compare the proposed origins for the flux: limits on the small scale fluctuations and the spectrum. We have

already seen how our knowledge of the fluctuations can constrain models
for the number versus flux of unresolved point sources. In addition the
residual fluctuations are a potential probe of structure in the universe
at the largest scales. Unfortunately, the restriction of these
measurements to a single narrow portion of the X-ray spectral band makes
it difficult to determine the physical processes that produce the X-
rays. On the other hand, a detailed moderate resolution measurement of
the sky spectrum over the entire X-ray band can give clues of the origin
of the emission. The photon distribution predicted for a particular
process or actually observed for a class of sources can be compared to
the observed X-ray sky spectrum to place limits on the possible
contribution by such a process or class. No model for the X-ray sky
flux can be considered successful unless its expected contribution over
a broad energy range is compatible with the spectral observations.

### Spectral characteristics of X-ray emission processes

The interaction of energetic electrons with their environment is
at the heart of the models which dominate continuum emission in the X-
ray band. Specific models differ in several ways. The electrons can be
in thermal equilibrium or in a nonthermal power law distribution. An X-
ray photon can be created by accelerating an electron or a low energy
photon can be upscattered to X-ray energies by $\gamma$-e interactions. Those
processes associated with atomic interactions (atomic absorption, line
emission from atomic transitions, ionization, and recombination), though
of intense astrophysical interest and exhibited by a wide range of X-ray
sources, are not thought to be important for the X-ray sky spectrum.

We will outline the observational characteristics of the basic
processes invoked in models of the X-ray sky flux. A more detailed
treatment of the physics may be found in Tucker [1975] and Rybicki and
Lightman [1979] as well as specific references that follow. A review of
emission processes in the context of current observations of X-ray
sources is found in Holt and McCray [1982].

Two typical processes for X-ray emission from a relativistic

electron, $\gamma \equiv E_e / m_e c^2 \gg 1$, are synchroton emission and inverse Compton scattering. In the former, acceleration of the electron by a cosmic magnetic field will produce a quasi-continuous·photon emission which peaks at an energy $E = 5.2 \times 10^{-12} \gamma^2 B$ keV, with B measured in Gauss. Therefore the production of a 20 keV photon by the synchrotron process requires very large fields or ultrarelativistic electrons. Inverse Compton scattering, where an initially low-energy photon with energy $E_\ell$ gains energy from the electron via Compton scattering, is able to produce X-rays with less energetic electrons: $E \approx \gamma^2 E_\ell$. Therefore electrons with $\gamma \sim 3 \times 10^3$ will produce 20 keV photons by scattering off the microwave background. If the soft photons are in the optical band, e.g. starlight, then $\gamma \sim 30$. Both processes, synchrotron and inverse Compton, are combined in some compact sources, and the overall process is called synchrotron self-Compton (SSC) [Jones, O'Dell, and Stein 1974]. The electrons interact with the magnetic field to produce radio through optical synchrotron photons. Some of these photons are then scattered up to X-ray and higher energies by the same electron population. This mechanism has had some success in providing a unified spectral picture of emission from Seyferts [Mushotzky 1977] and, with modifications, from BL Lac objects [Urry and Mushotzky 1982]. The identical $\gamma^2$ dependence of both the inverse Compton and synchrotron processes shows that for a given population of electron energies, these processes will produce photon spectra with essentially the same shape. A power law distribution of electrons,

$$dN_e \propto \gamma^{-n} d\gamma , \qquad \text{[VIII.1]}$$

produces a power law photon spectrum with energy index

$$\alpha = (n - 1) / 2 \qquad \text{[VIII.2]}$$

(recall that the index is defined so that the flux is proportional to $\nu^{-\alpha}$).

If radiative losses dominate, the electron energy loss rate is proportional to $\gamma^2$, so that large $\gamma$ electrons are affected most. The

time scale for energy loss is

$$t_c \sim 3 \times 10^7 / U \gamma \text{ seconds,} \qquad\qquad \text{[VIII.3]}$$

where U is the energy density, in ergs $cm^{-3}$, of the photon field for inverse Compton emission or of the magnetic field for synchrotron emission. This leads to a fairly sharp break in the electron energy spectrum at $\gamma_b$, above which the electrons are strongly affected by radiative losses. The break energy will move to lower and lower energies as the electron population ages. If t is the age of the source, $\gamma_b(t) \sim 3 \times 10^7 / (U t)$. This will in turn affect the emitted spectrum at energies greater than the corresponding photon energy, $E_b(t)$. For inverse Compton emission this is

$$E_b(t) \approx 2.4 \, E_\ell \, (\gamma_b(t))^2 . \qquad\qquad \text{[VIII.4]}$$

If the electrons are continuously replenished by new electrons also having index n, then above the break energy the population reaches equilibrium with index n+1. Therefore the emitted spectrum energy index breaks to

$$\alpha' = \alpha + 1/2 . \qquad\qquad \text{[VIII.5]}$$

If the population is not replenished the observed break is sharper. For synchrotron emission the index breaks to [Tucker]

$$\alpha' \geqslant (4/3)\alpha + 1 . \qquad\qquad \text{[VIII.6]}$$

The equality holds only when the initial distribution of pitch angles, the angle the electron velocity makes with the B field, is isotropic and there is no mixing of the populations with differing angles as the electrons age. The higher pitch angles are preferentially depopulated as those electrons have larger energy loss rates. Synchrotron emitters that continuously re-isotropize the pitch angle distribution or inverse Compton emitters in an isotropic soft photon bath may have essentially no flux above the photon break energy, $E_b(t)$.

In the above example of X-ray production by inverse Compton, the photon was energized by a single scattering with the relativistic electron. If the medium is optically thick to Compton scattering (but thin to absorption) multiple inverse Compton interactions can increase the energy of soft incident photons to the typical electron energy. X-

rays may be produced in such cases from soft photons by a distribution of non- or near-relativistic electrons. For example, in unsaturated Comptonization soft photons are scattered through a thermal electron distribution which is only a few Compton optical depths in extent [Shapiro, Lightman, and Eardley 1976]. If the electron temperature is T, then the low energy emitted photons, $E \lesssim kT$, have a power law form, the index determined by details of the Compton optical depth. At higher energies, $E \sim kT$, the spectrum includes an exponential cutoff, $\exp(-E/kT)$, reflecting the thermal electron distribution.

A thermal distribution of electrons can produce X-rays directly through bremsstrahlung interactions with other constituents of the plasma. Though e-e interactions may make a significant contribution, particularly at higher temperatures, $kT \gtrsim 100$ keV [Maxon 1972], the electron-ion interactions dominate at lower temperatures. For a thin (to Comptonization) thermal plasma the emissivity is [Tucker 1975]

$$j(E) \approx 6.8 \times 10^{-38} \left( \sum_i z_i^2 N_i \right) N_e \, g(E,T)$$
$$\times \ T^{-1/2} \exp(-E/kT) \ \text{ergs cm}^{-3} \text{ s}^{-1} \text{ Hz}^{-1}, \qquad \text{[VIII.7]}$$

$N_e$ is the number density of electrons; and $N_i$ and $Z_i$ are the number density and charge of the ith ion species. For a completely ionized gas with $N_{He} = 0.1 \, N_H$ (He 28% of total mass) and no higher Z constituents the sum reduces to $1.2 \, N_e$. For material containing typical cosmic abundances of high Z species, the sum is $\sim 1.4 \, N_e$. The effective Gaunt function, $g(E,T)$, can be treated as a general correction term. An approximate form, derived by Mätzler, Bai, Crannel, and Frost [1978] is

$$g(E,T) \approx (E/kT)^{-(\alpha_g)}, \qquad \text{[VIII.8.a]}$$
$$\alpha_g \equiv 0.37 \, (30 \ \text{keV}/kT)^{0.15}. \qquad \text{[VIII.8.b]}$$

However, this is valid only for moderate temperature plasmas, $15 < kT < 50$ keV. Second order radiation processes, relativistic corrections to the Maxwellian thermal distribution of the particles, and contributions from electron-electron scatterings, already important at these temperatures, require increasingly accurate calculations at higher temperatures. The problem has been addressed by Gould and others (Gould [1982] and references therein). In general, as temperatures increase,

g(E,T) becomes a flatter function of energy. Like unsaturated Comptonization, the spectrum produced by thin thermal bremsstrahlung is approximately a power law at low photon energies with an exponential cutoff at E ~ kT.

### Characterizations of the sky spectrum

Prior to HEAO 1, the spectrum over the 3 to 100 keV range was measured in separate portions by several groups. Early results from any one band were generally consistent with a smooth, featureless power law. Trying to combine these different measurements proved difficult as the experiments had fundamental differences in sensitivity and internal background. Even two early measurements that covered roughly the same energy band with similar detectors seemed to differ; the reported value for the spectral energy index, $\alpha$, ranging from 0.4 [Boldt, Desai, Holt, and Serlemitsos 1969] to 0.7 [Gorenstein, Kellogg, and Gursky 1969]. It was eventually recognized that the power law index steepened significantly at ~20 keV, though it was not until the results of Schwartz and Peterson [1974] that this rollover was detected by a single experiment.

Two experiments on HEAO 1 were together able to cover the entire range of the "hard" X-ray sky spectrum. The high energy ranges, 12-100 keV [Rothschild et al. 1983] and 60-400 keV [Matteson et al. 1979], were covered by the scintillation counters of the A-4 experiment. The MED and HED proportional counters from the A-2 experiment covered the range 3 to ~50 keV [Marshall et al. 1980]. Using the internal background subtraction (L-SFOV) and the electron rejection (propane veto layer) techniques (see Sections II:1 and II:2), the A-2 experiment was able to obtain a relative accuracy of a few per cent. The absolute precision of the normalization is estimated to be ±10%. This produced a spectrum with an accuracy rivaled by few other source spectra in X-ray astronomy. Since the final calibration of both experiments involved in-orbit examinations of the Crab Nebula it is not surprising that they give fairly good agreement over the regions of overlap (Figure VIII.1).

The X-ray Sky Flux                                                    The Spectrum

Figure VIII.1: The cosmic X-ray sky spectrum from HEAO 1

The error bars on the solid curve from Marshall et al. are generally smaller than the line thickness. The A-4 results are from preliminary fits. The very highest energy results (0.3-10 MeV) were obtained from scintillation experiments flown on several Apollo missions [Trombka et al.UNL 1977]. (Based on a figure in Rothschild et al. [1983]. The dashed AGN contribution lines are from two models of the AGN luminosity function considered by Rothschild et al., with no evolution.)
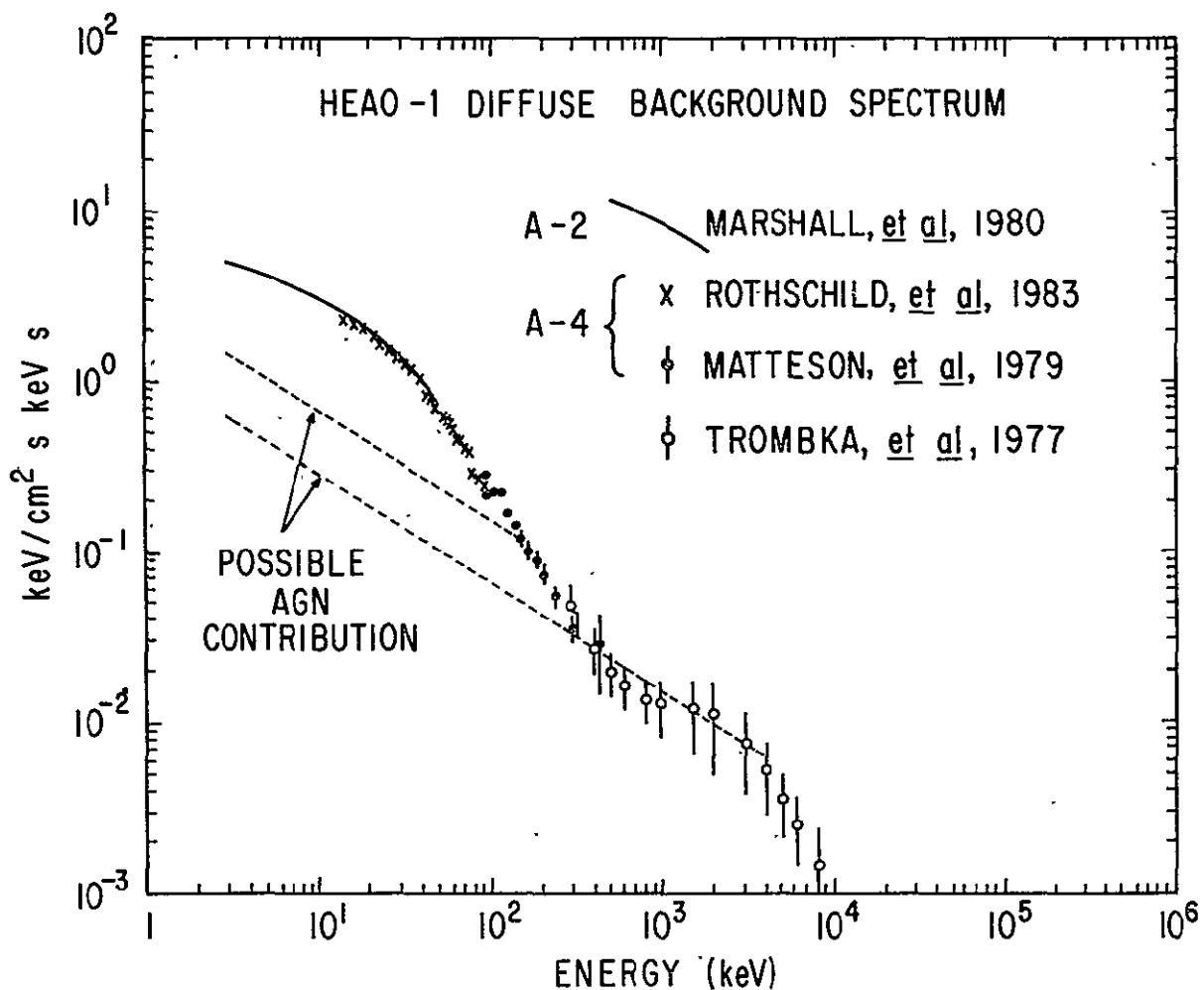


Marshall et al. were able to reject decisively a single power law model for the total A-2 band. Models with an exponential form provided

a significantly better fit. Because of previous theoretical suggestions, a thin thermal bremsstrahlung model was examined. First order corrections to the simple dipole model were included. Such a model, with kT = 40±5 keV, gave an excellent fit. Rothschild et al. fit the 12–100 keV range with kT = 45, but their model omitted some of the correction terms used by Marshall et al. Rothschild et al. estimate that the inclusion of these terms will shift their best fit value to kT ~ 40 keV, in excellent agreement with the A-2 result. Rothschild et al. show that the 12-100 keV thermal component smoothly joins the results for the 60-400 keV range, from earlier analysis of A-4 data by Matteson et al., where a simple power law with $\alpha$ = 1.67 is a good fit. However this fit could be, especially at lower energies, substantially contaminated by the continuation of the 40 keV exponential. If the exponential component were removed, the index of the remaining 60-400 keV flux would be substantially flatter.

The measured spectrum need not be produced by a thermal or exponential component. Superpostions of evolving non-thermal sources can mimic the single temperature specturm. Following the example of earlier work [Cavaliere et al. 1979], de Zotti et al. [1982] fit the A-2 3-50 keV data with various nonthermal source spectra models. The forms chosen were ad hoc in that they are not grounded in the observation of any sources with these spectra, but they are not unreasonable. De Zotti et al. assumed $q_0$ = 0.5 and allowed for evolution of the co-moving volume emissivity, $\Lambda \propto (1+z)^m$. In one model the source spectrum is a broken power law, with $0.4 < \alpha_1 < 0.5$ below a break at 50-55 keV and $1.2 < \alpha_2 < 1.4$ above it. Models with significant evolution, $m \gtrsim 3$, were preferred. Single power law sources with a sharp high energy cutoff were also fit, although the position of the cutoff must either undergo strong evolution or be drawn from a broad distribution. However, the higher energy A-4 data (not included in the de Zotti et al. fits) may strongly influence the acceptability of these models.

Apart from the general exponential roll off, there is little obvious structure in the spectrum. The residuals to the A-2 thermal fit are ≤ 1% over the 3 to 20 keV range. Marshall et al. specifically

examined the A-2 data for indications of a possible edge due to iron emission. Their upper bounds are consistent with the expected contribution from a hot intergalactic medium (IGM, see next section), or from clusters.

If the thermal interpretation of the spectrum is correct, and if the origins of the X-ray sky flux are at high redshift, then the current observed X-ray temperature indicates quite high electron temperatures at the time of emission, $kT_{em} \approx 40 (1+z_{em})$. For compact thermal plasmas with $kT \gtrsim 160$ keV ($z \gtrsim 3$), significant numbers of electron-positron pairs would be produced, principally by 2-photon processes. Annihilation radiation from these pairs would produce a significant feature. The details of the calculations are complicated, but the failure to detect a significant perturbation to the thermal fits at $\lesssim 100$ keV constrains the physical characteristics of the emitting sources (see Kazanas and Shafer [1983]).

------------------------

## 2: The X-ray Sky: the Problem, and Model Solutions

The subtraction of estimated contributions from clusters and AGN, without evolution, modifies the simple thermal spectral fit. QSOs certainly may provide sufficient flux, but leave important spectral issues unanswered. Other models specifically address the spectral constraints, but lack specific source identifications. Inverse Compton models are shown to be incompatible with the residual sky spectrum. Thermal bremsstrahlung from a hot intergalactic medium is considered in some detail, though a simple uniform medium has significant difficulties. The limits on the excess variance do not rule out significant clumping, in the context of our simple Euclidean analysis of the last chapter.

The problem

Observations indicate the existence of substantial emission in the range from 3-100 keV, most of which (certainly at higher energies)

cannot be ascribed to observed sources without some extrapolative assumptions. The total all sky flux from 2 to 100 keV is $2.3 \times 10^{-6}$ ergs $s^{-1}$ $cm^{-2}$. This corresponds to an energy density, $U_{o;x} = S_{as} / c = 7.8 \times 10^{-17}$ ergs $cm^{-3}$. In comparison, the microwave background energy density is $4 \times 10^{-13}$ ergs $cm^{-3}$. If we ignore sinks and sources of either kind of radiation, the ratio of the two is independent of redshift, and $U_x = 2 \times 10^{-4}$ $U_\mu$. The X-ray photon density is $4.7 \times 10^{-9}$ photons $cm^{-3}$. The spectral form of the emission is smooth, with a distinct exponential rollover at 40 keV. At higher energies ($\gtrsim$ 100 keV) there is a power law component, $\alpha_3 \sim 1.7$, extending to low $\gamma$-ray energies. However, in this section we do not treat both components, and instead concentrate on the exponential component that dominates the sky flux in the A-2 bandpass. This component may be of entirely different origin. A comprehensive and coherent model for the origins of the X-ray sky flux must exhibit the following characteristics: (1) it must provide the energy source(s) sufficient for the observed flux; (2) the total emitted spectrum predicted by the model, after adjusting for cosmological effects, must roll over at E $\sim$ 40 keV; (3) the contribution of known populations of X-ray emitters must be allowed for, not only in the total flux but also in the spectral shape; and (4) finally the sources cannot cause additional variance in the H1L1 band which would exceed our upper bound for $\sigma^2_s$.

The energetics for producing the X-ray sky are not intrinsically overwhelming. The current energy density $U_{o;x}$ correponds to 17 $\Omega_b^{-1}$ $h_{50}^{-2}$ eV per baryon, where $\Omega_b$ is the baryon density of the universe in units of the critical density (equation [VII.60]). In comparison, nuclear fusion can liberate as much as 5 MeV per nucleon, while gravitational collapse can produce significantly more. If $\Omega_b = 0.1$, then we need "burn" only 0.003% of the hydrogen of the universe into iron. One problem, of course, is that normal thermonuclear processes are not at all efficient in producing X-rays, most of the radiation escapes at wavelengths in or near the optical band. In terms of discrete objects, the total emission required is $3 \times 10^{59}$ ergs for every galaxy brighter than $M_B$ -17.5. Bookbinder et al. [1980] estimate that an average galaxy has produced $\sim 5 \times 10^{61}$ ergs in supernovae, based on estimates of the metal content of the old stellar population of

elliptical and spiral galaxies (see also Schwarz, Ostriker, and Yahil [1975]). If AGN are the source of the X-ray sky flux then we see from our canonical luminosity function of AGN (see Table VII.5) that $\sim 10^{60}$ $h_{50}^{-3}$ ergs per AGN are required. (This estimate is independent of any evolution if the evolution conserves the total number of AGN, such as pure luminosity evolution.) As we know, AGN are efficient producers of X-rays. If we assume an average luminosity of only $10^{43}$ ergs s$^{-1}$ (2-100 keV), then we could generate $U_{o;x}$ in $\sim 3 \times 10^{9}$ years. This is crude confirmation that with some evolution, AGN could produce most of the observed sky flux.

Cosmological considerations modify these simple numbers. If the observed energy density were created at an earlier epoch corresponding to a redshift $z_c$, then the X-ray energy density at that epoch was $U_x(z_c)$ = $(1+z_c)^4 U_{o;x}$. Since the epoch of creation, the radiation has been redshifted, causing an adiabatic cooling. Of course, the density of the sources of the radiation was also greater at earlier epochs, since for matter $\rho \propto (1+z)^3$. A more helpful measure of these quantities is the co-moving density:

$$U^* \equiv (1+z)^{-3} U(z) .$$
[VIII.9]

If there is no creation or destruction of sources, their co-moving number density is constant. For photons the evolution of the co-moving energy density, $U^* \propto (1+z)$ reflects the higher energy each photon had at the earlier epoch. As an example, if the co-moving rate of emission in the appropriate band is constant, then the total required co-moving energy density required to produce what is seen today is ($q_o = 0.5$)

$$U^*_T = (7/3) U^*_{o;x} .$$
[VIII.10]

Either the numbers or luminosity of the sources would have to adjusted. This does not include effects due to spectral redshifting of the energy source, so that flat spectrum sources ($\alpha < 1$) such as AGN would require less of a correction to their emission as measured at a fixed frequency in the emitter frame.

Models from known X-ray sources

Since they were first detected, clusters of galaxies have been proposed as significant contributors to the X-ray sky flux. Further spectral investigation has shown that clusters have thin thermal spectra, but distinctly cooler than the 40 keV sky. Our current estimate, based on adjustments to the Piccinotti et al. [1982] luminosity function, is that clusters can make up as much as 4.5% of the H1L1 flux (roughly 2.5 to 13.3 keV). Previous percentages, based on incomplete estimates of spectra, luminosity function, and the impact of cosmological corrections, were substantially larger. Without a substantial change in cluster properties at large redshifts or low luminosities there is small chance of increasing their contribution from current estimates by more than a factor of ~2.

The contribution of AGN to the sky flux has grown in importance since their original discovery as sources of X-rays. (We continue to include QSOs under the category of AGN.) But early and simple arguments based on data in other·bands also argued that this would be the case [e.g. Setti and Woltjer 1973]. We have seen from the last chapter that without evolution AGN may provide ~25%, or more, of the sky flux, and with evolution consistent with extrapolations from optical data, AGN can entirely make up the total.

There have been arguments, based in part on X-ray fluctuations measurements, that place upper limits of ~50% on the contribution of AGN to the background [Cavaliere et al. 1980; 1981]. However, our own analysis leads to no such conclusion. As we saw in Figures VII.17 and VII.19, there exist evolutionary models for AGN that can produce all of the background without exceeding the excess variance limits. Our disagreement with the earlier work is due in part to their use of an inaccurate estimate of the excess variance based on UHURU data. In addition they took no account of the substantial negative excess variance that is expected to be contributed by the rollover in N(S) of clusters.

Given the present analysis of HEAO 1 A-2 data, especially with the confirmation of significant AGN X-ray evolution from Einstein observations, many observers feel that it is all but clinched that AGN provide the dominant fraction of the sky flux. However, some caution is necessary. Much of the argument depends on extrapolation. The lower energy spectral band covered by Einstein, compared to the higher energy of the bulk of the sky flux, is a particular source of bedevilment. For example, the calculation of the contribution of the observed Einstein Deep Survey sources to the sky flux at higher energies makes the possibly circular assumption that the sources have the same spectrum as the 2-10 keV sky. Even the contribution of an unevolved population of AGN is dominated by the lowest luminosity sources which are not directly observed by the A-2 experiment. Our estimates have relied heavily on the form of the continuation to lower luminosities, which in turn was based on IPC measurements.

If the evolution of AGN allows them to be responsible for most of the total sky flux, there are important ramifications for our understanding of the X-ray producing physical processes in AGN. In general these processes are similar to those we have discussed in connection with the sky spectrum in the last section (see also Rees [1980]).

Our strong conclusion: if AGN do provide the bulk of the sky flux then they must undergo spectral evolution, to a form distinctly different from what the moderate luminosity, low redshift AGN are observed to have by the A-2 and A-4 experiments. (Recall the observation in Chapter I that no source studied in medium- or high-energy X-rays has the appropriate spectral shape to mimic the total sky flux.) In particular, the effective spectral index over 10-40 keV must flatten, changing from ~0.7 at low redshifts to $\lesssim 0.4$ at higher redshifts (or luminosities). De Zotti et al. [1982] specifically tested for the amount of allowed variation in the sources' power law indices, if power law source spectra were to constitute the sky spectrum. When this variation was modelled as a Gaussian distribution of the index, with all sources having identical differential luminosities at 5 keV, they were

able to show that the allowed $\sigma_\alpha \leqslant 0.25$ (90%). They estimated that based on this limit there could be at most a 30% contribution to the sky flux, presumably as measured at 5 keV, from AGN or other objects with $\alpha \approx 0.7$. In this model, little or no AGN evolution is allowed. De Zotti et al. also considered a variant of their single power law model with an evolving sharp cutoff at high energy, now allowing evolution of the power law index. For an exponential evolution of the volume emissivity, $Q_L = 4$, the low energy data is fit by sources with $\alpha = 0.7 \exp[-1.35\tau]$. However, as de Zotti et al. note, such models provide a poor fit to the 40 keV exponential rollover at the high energy end. This kind of evolution of the power law index would result in significant changes at relatively low redshifts. At $z = 0.5$ the index has flattened to 0.45, and by $z=1.5$ to 0.31. As is the case for most cosmological evolution models, this proposed form for the variation in $\alpha$, though ad hoc, gives a rough feeling for the sort of behavior required by the constraints of the X-ray sky spectrum.

Though some such spectral evolution is allowed by the sky flux data, and is required if AGN do indeed provide the sky flux, the physical basis for such change is an outstanding question. It would help to find some change in AGN properties at high redshift coincident with the change in X-ray effective index. One obvious change is in the average source luminosity. An additional indicator is the possible shift in $\alpha_{ox}$ with either luminosity or redshift. Indeed, the supposed decrease in the ratio $L_{2 \; kev} / L_{opt}$ could be caused in part by a systematic change with redshift of the X-ray spectral index in the lower flux Einstein sources. On the other hand, we would be equally justified in positing an opposite evolutionary trend in $\alpha$, i.e. where the high redshift objects have preferentially steeper spectra than are observed in the A-2 and A-4 low redshift AGN spectra. The variation in $\alpha_{ox}$, as measured at ~ 2 keV, can be accommodated by either a systematic steepening or flattening of $\alpha$, depending on whether the X-ray energy where $\alpha_{ox}$ is not a function of redshift is below or above the Einstein bandpass. We have discussed how the extrapolation of the Einstein source counts to the counts in a higher energy band is greatly affected by the source spectral index.

The contribution of AGN to the higher energy sky flux, $\gtrsim$ 100 keV, is also an important issue. It has been suggested by Boldt [1981b] and others that the sky flux at $\gtrsim$ 500 keV can be produced by AGN with little or no evolution. This conclusion is predicated on the continuation of the AGN spectra, with $\alpha \sim 0.7$, out to these high energies. Also the limitations on the evolution depend critically on the AGN luminosity function at the low luminosity end. Both issues have been considered in more detail by Rothschild et al. [1983]. Their data do not require any break in $\alpha$ at high energies, but because they have few counts above 100 keV, one is certainly allowed. Eventually, the sources must break to follow the general downturn of the $\gamma$-ray sky flux. In addition, the existence of such a break in the spectrum of 3C273 between 100 keV and 50 MeV can be inferred from observations taken by COS-B [Hermsen et al. 1981] and simultaneous lower energy data from HEAO 1. At the other end of the luminosity function, the radio galaxy CEN A has been studied at high energies by Baity et al. [1981]. When the data taken from the A-2 and A-4 experiments is compared with non-contemporaneous measurements at higher energies there is again evidence for a steepening in the spectral index as the energy increases. Rothschild et al. investigate models with spectral breaks and sharp low luminosity cutoffs in the AGN luminosity function, assuming no evolution. Other discussions of the observational constraints can be found in Setti and Woltjer [1979; 1982]. SSC models for X-ray and $\gamma$-ray production in AGN (see the previous section) provide a theoretical basis for a gradual steepening in the spectral index with increasing energy [Mushotzky 1976] (see also the discussion in Baity et al. [1981]). Another theoretical model for such structure is found in Protheroe and Kazanas [1983], and the ramifications for the high energy sky spectrum are outlined in Kazanas and Protheroe [1983].

It is obvious that further spectroscopy is mandatory before the solution of the X-ray sky problem is in hand. If spectral evolution is to allow AGN to provide the bulk of the sky flux, then from the models of de Zotti et al. moderately accurate 1-10 keV continuum spectra of AGN at $z \sim 0.5$ should be distinguishable from the current epoch spectra. Such spectra would also make possible a more accurate extrapolation of

the Deep and Medium Sensitivity Surveys to the H1L1 band. These observations will be made by the next generation of X-ray detectors, such as the Broad Band X-Ray Telescope [Serlemitsos 1981] and others presented in Holt [1981]. If, even after spectral and luminosity evolution, AGN still do not provide the total X-ray sky flux, then we must look for explanations of the residual emission. An accurate calculation of the residual will, of course, require a good model for the AGN contribution. Not only will we need the 3-50 keV spectral results obtained at moderate redshifts, but also the evolution models and the Einstein survey counts must be included. In addition, to measure accurately the position and shape of the rollover at $\lesssim$ 40 keV we must have some understanding of the location and extent of any high energy spectral break for the AGN.

Preliminary estimates of the residual sky spectrum have been made by several authors, each with their own assumptions. Leiter and Boldt [1982] (see especially their Appendix D) fit the residuals of the A-2 3-50 keV sky spectrum to a generalized exponential and power law:

$$s(E) \propto E^{-\alpha} \exp(-E/E_o) . \qquad\qquad\qquad [VIII.11]$$

The e-folding energy $E_o$ is equivalent to the plasma temperature for thin thermal bremsstahlung and unsaturated Comptonization models. The components subtracted to obtain the residual include 4% (at 3 keV) from clusters (kT = 6.9 keV), 20% due to "Seyfert" galaxies (equally divided between $\alpha$ = 0.6 and 0.7), and a component due to "quasars". This last fraction is assumed to be steeper than Seyferts, $\alpha$ = 0.9 or 1.4, and it contributes between 14% and 17% of the 3 keV sky flux. Though, this model can be said to use our best current understanding of AGN, there are large uncertainties in our knowledge, particularly its contribution from quasars. Even so, the model indicates the trends in the residual spectrum. The values for $\alpha$ in equation [VIII.11] are fairly flat, showing that the low energy range of the residual is considerably flattened by the removal of the known sources. The exact value is strongly correlated with the value of $E_o$:

$$\alpha \approx 0.8 \ [1 - (23 \ keV/E_o)] , \qquad\qquad\qquad [VIII.12]$$

with good fits to the residual spectrum if $\alpha$ < 0.2. In particular, a

pure exponential, $\alpha = 0$, is a good fit with $E_o = 23$ keV. This illustrates that the residual spectrum has an essentially lower observed temperature, from 23 to 31 keV. A superposition of several exponentials, for example $E_o = 13$ and 33 keV with $\alpha = 0$, can also describe the residual spectrum. If we were to fit the residual spectrum to the broken power law models of de Zotti et al. similar trends would be observed. The index below the break would flatten, the break point would move to a lower energy, and the index above the break would steepen. Fabian [1981] and Fabian and Kembhavi [1982] (hereafter FK) fit the residuals for the entire 3-300 keV range. They obtain their residual spectrum by subtracting only an AGN component with $\alpha = 0.7$ that contributes 20% of the 3 keV sky flux. They fit the remainder with thin thermal bremsstahlung models and report that no single temperature model provides a good fit. They are able to fit a range of temperatures following a model for a cooling IGM which we discuss in more detail below.

Boldt has pointed out [private communication] that if the origin of the residual flux is indeed thermal bremsstrahlung, the flat power law-like component at the lower energies can contain very useful cosmological information about the origins of this flux. As we remarked in the last section (see the discussion of equation [VIII.8]), the effective Gaunt factor index is a function of the emitting-plasmas' temperatures, with the generally flatter $\alpha_g$ corresponding to hotter temperatures. Therefore, by fitting for the effective Gaunt factor, we can in principle measure the emitter frame temperature. With the observation of the current epoch $E_o$, we can then determine the redshift of the origin of the bulk of the sky flux:

$$1+z \approx kT_{em} / E_o .$$                                    [VIII.13]

The very flat residual spectrum obtained by Leiter and Boldt indicates that thin thermal bremsstrahlung models must originate at appreciable redshifts, $z \gtrsim 3$. The quality of this limit, however, depends on the assumptions used in obtaining the residual. In a similar manner, the determination of $\alpha$ can reflect the physics of unsaturated Comptonization models. However, as the effective index is mostly a measure of the

optical depth of the emitting plasma, rather than its temperature, we
cannot use it to determine the redshift of the emitter.

<div align="right">Inverse Compton models for the residual</div>

The production of the X-ray sky flux by inverse Compton scattering
of the microwave background photons off relativistic electrons has been
among the earliest models proposed. The model was originally proposed
by Felten and Morrison [1966] and elaborated by Brecher and Morrison
[1968]. Electrons were either produced in ordinary galaxies and then
leaked into intergalactic space, or injected by the processes that
produced the large scale structures associated with radio galaxies.

For some versions of this model the cosmological corrections are
straightforward. Because the seed photons come from the microwave
background, and are affected by the cosmological expansion in the same
way as the X-ray sky flux, the electron energy needed to produce what is
observed today as a 20 keV X-ray is unaffected by the epoch of
production, i.e. $\gamma_x \sim 3 \times 10^3$ at all z. Because the rate of scattering
is proportional to the energy density of the soft photons, which is
proportional to $(1+z)^4$, inverse Compton processes offer increased
efficiencies at earlier epochs:

$$dN^*_x = 7.3 \times 10^{-12} \; s^{-1} \; (1+z)^3 \; N^*_e(\gamma;z) \; dt \; .$$
                                                          [VIII.14]

$N^*_x$ is the co-moving density of X-ray photons and $N^*_e(\gamma;z)$ is the co-
moving density of electrons at redshift z and energy $\gamma$. The coefficient
is proportional to the co-moving number density of the soft photons.
The numerical value used here assumes that these are the $\sim 280 \; cm^{-3}$
photons of the 2.7 Kelvin microwave background.

This model was originally suggested at a time when the details of
the X-ray sky spectrum were very poorly known, particularly the need for
a break around 20 to 40 keV. As discussed in the last section, a power
law distribution of relativistic electrons can easily produce a power
law distribution of X-ray counts. However, as the sophistication and
accuracy of the spectral measurements increased, theoretical models

needed to reflect the observed break. Cowsik and Kobetich [1972] argued that inverse Compton from a blackbody spectrum was unable to match the observed sharpness of the break, examining a particular model for the sources' electron distributions. Brecher [1973], in rejoinder, pointed out that much of the width of their break resulted from Cowsik and Kobetich's assumption of a distribution in the break point for the equilibrium electron spectrum. Brecher maintained that if all electron spectra have a universal break energy, then their superposition would give a good description of the X-ray observations. To illustrate this possibility he presented several radio source spectra which matched the curvature in then current measurements of the X-ray sky spectrum. Because these source spectra were presumably generated by synchrotron processes, which are inherently broader than inverse Compton from a black-body, there must exist some electron spectrum which could reproduce the X-ray curvature.

In order to compare observations of individual sources to the sky spectrum we must make specific assumptions, in particular that the snapshot spectrum of a source at one instance is equal to the total spectrum produced over the lifetime of the source. Unfortunately, for the inverse Compton process, the energy loss times of the electrons responsible for the X-ray emission are short on a cosmological timescale. From equation [VIII.3] we see that electrons with energy $\gamma_x$ have lifetimes

$$t_c \; \bar{H}_o \sim 0.04 \; (1+z)^{-4} \; h_{50} \; . \qquad\qquad [VIII.15]$$

Electrons produced even at the relatively nearby redshift of 0.5 have the comparatively short lifetime of ~150 million years. All electrons with $\gamma > \gamma_x$ at injection, after time ~ $t_c(\gamma_x)$, will have decayed to energies less than $\gamma_c$ and therefore no longer contribute photons to the X-ray band. Hogan and Layzer [1977] derived an approximate relationship between the electron injection spectrum and the sum of the emitted X-ray photons, assuming that inverse Compton dominates the electron loss rates. In their example, the electrons were produced during a single cosmologically short interval. We can generalize their result to allow for electron production throughout the history of the universe. Let

$N^*_{e;tot}(\gamma)$ d$\gamma$ be the total co-moving injection spectrum. If $U_{o;x}(E)$ is the current epoch X-ray energy density between E and E+dE, then

$$N^*_{e;tot}(\gamma) \, d\gamma \approx -2 \, E_\ell \, d[\gamma \, U_{o;x}(E_\ell\gamma^2)] \,. \qquad [VIII.16]$$

The accuracy of this formula rests on the fact that the decay times are short in comparison to the expansion time of the universe. Also, we have assumed that the distribution of inverse Compton photons from an electron of a given energy is a delta function, $\delta(E - \gamma^2 E_\ell)$. Originally inverse Compton models were used to explain the sky flux from 1 keV to well over an MeV. We can show, however, that these models do not explain the residual component of the X-ray sky flux after subtraction of the expected AGN contribution. If the spectral form of the residual is well described by equation [VIII.11], then the total electron injection spectrum is

$$N^*_{e;tot}(\gamma) \propto \gamma^{-2\alpha} \, \exp(-\gamma^2 E_\ell/E_o)$$
$$\times \, [ \, 2 \, \gamma^2 \, (E_\ell/E_o) - (1-2\alpha) \, ] \,. \qquad [VIII.17]$$

If $\alpha \gg 0.5$ then $N^*_{e;tot}$ is well defined, following the general form as the X-ray spectrum: a power law with an exponential (in $\gamma^2$) cutoff. However as $\alpha \rightarrow 0.5$, the electron spectrum begins to require a roll over at low $\gamma$. Finally for $\alpha < 0.5$, the X-ray spectrum requires <u>negative</u> numbers of low $\gamma$ electrons, <u>i.e.</u> no realistic distribution of electrons can provide such an X-ray spectrum. The fits to the X-ray sky spectrum had $\alpha \leq 0.4$, even before removal of any AGN contribution, so it is easy to see that these models are not viable. Inverse Compton models could still be possible if radiative losses were not the dominant mode of electron energy loss. One possible approach might be adiabatic expansion losses, but this needs further investigation. Even if it is possible to reproduce the observed structure, a specific break energy is required in the electron injection spectrum. The model is incomplete without some physical understanding of why all the sources of electrons should have such a feature at nearly identical energies.

## Thin thermal bremsstrahlung models for the residual

At the same time that Cowsik and Kobetich objected to inverse

Compton models they proposed that the curvature in the sky flux spectrum could be fit by the exponential characteristic of thin thermal bremsstrahlung. This replaces the problem of determining a unique feature in the injection electron spectrum with that of postulating a universal temperature, or a moderately narrow range of temperatures. Cowsik and Kobetich answer this problem with the suggestion that the gas is a uniform hot IGM. The amount of material required was significant on a cosmological scale, $\Omega_{IGM} \lesssim 1$, so that it was important in the missing mass controversy. This idea was examined in more detail by Field and Perrenod [1977] who proposed a specific thermal history of the universe. As mentioned earlier, we will discuss the ramifications of the Fabian and Kembhavi [1982] (FK) numerical fit to the residual spectrum. In the FK model, the IGM is abruptly heated at a redshift $z_m$. Thereafter the gas cools, principally due to adiabatic cooling from the cosmological expansion, so that the thermal history of the IGM is

$$T = (1+z)^2 \, T_{oIGM} \, .  \qquad\qquad [VIII.18]$$

$T_{oIGM}$ is the current epoch gas temperature. The thermal bremsstrahlung emissivity of the gas is [Tucker 1975]

$$dU_x \approx 2.4 \times 10^{-27} \, N_e^2 \, T^{1/2} \, dt \text{ ergs cm}^{-3}. \qquad [VIII.19]$$

$N_e$ is the number density of electrons. The numerical coefficient is calculated only for the simple case of dipole electron-ion emission, which at the higher temperatures may significantly underestimate the total emissivity. From the effects of expansion on $N_e$ and $U_x$ we can find the current X-ray energy density:

$$U_{o;x} = \int dt \, (1+z)^3 \, 2.4 \times 10^{-27} \, N_{o;e}^2 \, T_{oIGM}^{1/2} \, . \qquad [VIII.20]$$

We see that just like the inverse Compton model, thermal bremsstrahlung is a more efficient contributor to the sky flux at the larger redshifts. The equation also indicates that the emissivity is proportional to the square of the electron density.

The hot IGM may not be homogenous, but clumped. When considering averages over all space we can replace $\langle N_e^2 \rangle$ with $C \langle N_e \rangle^2$, where C is

the clumping factor:

$$C \equiv \frac{\langle \rho_{IGM}^2 \rangle}{\langle \rho_{IGM} \rangle^2} = \frac{(\delta \rho_{rms})^2}{\rho_{IGM}^2} + 1 .$$  [VIII.21]

If we assume that the IGM is clumped into spherical, possibly overlapping, regions ("lumps") with constant surface brightness, then we can use the crude results of Section VII:6 to estimate an upper bound for the clumping factor. Our upper limit on $\sigma^2_I$ provides an upper bound on the filling factor, $\Phi$ (see equation [VII.62]), given the lump size $\lambda_s/\lambda_H$. In turn, this is related to the clumping factor by (see equation [VII.64])

$$C = \Phi^{-1} + 1 .$$  [VIII.22]

The use of the Section VII:6 limits assumes that the hot IGM is the dominant source of the total sky flux and the excess variance.

The FK fits assumed $q_0 = 0.5$. Their derived contribution of the X-ray emitting IGM to the mass density, in units of the critical density $\rho_{crit}$, $4.7 \times 10^{-30} h_{50}^2$ ergs cm$^{-3}$, is

$$\Omega_{IGM} \approx 0.6 \ (1+z_m)^{-3/2} \ C^{-1/2} \ h_{50}^{-3/2} .$$  [VIII.23]

A significant clumping factor and/or heating of the gas at a higher reshift increases the emissivity of the gas, thereby reducing the required amount. FK do not report a range of acceptable values of $z_m$, but the value 5 provides a good fit to their residual spectrum so that $\Omega_{IGM} \sim 0.04 \ C^{-1/2}$. The Field and Perrenod estimate, with a $z_m = 3$, is substantially larger, $\Omega_{IGM} \sim 0.3$ [Rees 1980]. The decrease in the required amount of hot gas is partially due to the heating at an earlier epoch, but is also due to the. FK inclusion of higher order bremsstrahlung processes, as well as their removal of the AGN contribution to the sky flux. The flatter $\alpha$ of the residual spectrum may be responsible for the higher $z_m$ used by Fabian and Kembhavi.

The mean electron density in the universe is

$$\langle N_e \rangle = \Omega_{IGM} \ \rho_{crit} \ (0.85/m_H) \ (1+z)^3 .$$  [VIII.24]

The numerical factor is derived from the fraction of He in the universe

and $m_H$ is the mass of a hydrogen atom. Using the FK values this corresponds to

$$\langle N_e \rangle = 1.4 \times 10^{-6} \ C^{-1/2} \ h_{50}^{1/2} \ (1+z_m)^{-3/2} \ (1+z)^3 \ . \qquad \text{[VIII.25]}$$

The average density decreases as the medium becomes more clumpy, although the mean density <u>within</u> each clump increases,

$$\langle N_e \rangle_{cl} = \langle N_e \rangle \ ( \ \Phi^{-1} + 1) = C \ \langle N_e \rangle \ , \qquad \text{[VIII.26]}$$

so that $\langle N_e \rangle_{cl} \propto C^{+1/2}$. The temperature of the medium is

$$T_{IGM} = T_o \ (1+z)^2 \ / \ (1+z_m) \ . \qquad \text{[VIII.27]}$$

$T_o$ is the current epoch temperature of the X-rays emitted just at $z_m$, $kT_o = 50$ keV, or $6 \times 10^8$ Kelvin. $T_{o\,IGM} = T_o \ / \ (1+z_m)$ is the current epoch temperature of the gas.

Perhaps the most frequently heard objection to a hot IGM model for the production of the X-ray sky flux is its extravagant energy requirements. Not only must we find a source for the energy of the X-ray photons, but also energy to heat the medium to the initial temperature, $T_i = T_o \ (1+z_m)$. This requirement is more severe:

$$U^*_{heat} \approx 3 \ N^*_e \ kT_i$$

$$\approx 6000 \ U_{o;x} \ (1+z_m)^{-1/2} \ C^{-1/2} \ h_{50}^{1/2} \ . \qquad \text{[VIII.28]}$$

The energy input required per baryon is now $0.1 \ \Omega_b^{-1} \ (1+z_m)^{-1/2} \ C^{-1/2}$ $h_{50}^{-3/2}$ MeV, or 0.4 Mev for a smooth universe heated at $z_m = 5$ and $\Omega_b = 0.1$. That is, 8% of the universe has been converted to iron, if the energy comes from thermonuclear reactions. This is a strong argument that if a hot IGM is the origin of the X-ray sky flux, then some other efficient, powerful source of energy is needed. Even AGN may have difficulties providing this much. The average energy requirement of $2.5 \times 10^{63}$ ergs per AGN corresponds to an average luminosity of 60 $L_{44}$ over the entire life of the universe ($q_o = 0.5$). As this calculation includes the low luminosity extension of the AGN luminosity function we see that strong evolution is required. For instance, the average AGN luminosity with a $Q_L = 4$ evolution of the entire canonical AGN luminosity function is $\sim 10^{43}$ ergs s$^{-1}$. Of course, this figure considers solely the energy output as X-rays. The non-electromagnetic fraction of

the AGN total luminosity, such as kinetic energy in jets or beams, may
be much higher (and of course if there is any residual spectrum to be
explained by a hot IGM, the bulk of the gas-heating energy is probably
not in the X-ray band).

The relative inefficiency of thermal bremsstrahlung means that
little of the original energy required to heat the material will be
emitted as X-rays. The characteristic loss time, $t_b$, from thermal
bremsstrahlung can be found from equations [VIII.19] and [VIII.28]:

$$t_b \approx 1.8\times10^{11}\ T^{1/2}\ /\ N_e \text{ seconds .} \qquad [\text{VIII.29}]$$

At the heating epoch the loss time is hundreds of times the present age
of the universe,

$$t_b(z_m)\ H_o \approx 4900\ C^{-1/2}\ h_{50}^{1/2}\ (1+z_m)^{-1} . \qquad [\text{VIII.30}]$$

(The product $t\ H_o$ is dimensionless, and measures the time scale in units
of the Hubble time, which equals $2\times10^{10}\ h_{50}^{-1}$ years.) The dominant
energy loss rate of the hot IGM is adiabatic loss from the expansion of
the universe:

$$(dU^*/dt)_{ad} \approx 3\ N^*_e\ k\ dT/dt , \qquad [\text{VIII.31}]$$

which for a $q_o = 0.5$ cosmology is

$$(dU^*/dt)_{ad} \approx 6\ N^*_e\ kT_{oIGM}\ H_o\ (1+z)^{7/2} . \qquad [\text{VIII.32}]$$

From this the characteristic loss time due to adiabatic expansion is

$$t_{ad}\ H_o = 0.5\ (1+z)^{-3/2} , \qquad [\text{VIII.33}]$$

which at a redshift of 5 equals 0.03 Hubble times. At very high
redshifts Compton cooling from the microwave background can be
important:

$$(dU^*/dt)_{cc} = (4\sigma_t/m_ec)\ N^*_e\ (1+z)^4\ U_{o;\mu}\ kT_{IGM} . \qquad [\text{VIII.34}]$$

$\sigma_t = 6.65\times10^{-25}$ cm$^2$ is the Thomson cross section. However, Compton
cooling is small compared to adiabatic expansion losses for $z \lesssim 280$, $q_o$
= 0.5. Even if the effect on the energy budget of the IGM is small,
Compton cooling can have a detectable impact on the spectrum of the
microwave background. This has been discussed in more detail by Field
and Perrenod, and Wright [1979]. Other ramifications of a hot IGM are

reviewed by FK and by Field [1978]. Possible scenarios for the heating of the IGM and their observational consequences have been considered in detail in a series of papers by Sherman [e.g. 1979, 1980] including examples where the IGM is explicitly assumed not to contribute substantially to the residual X-ray sky flux [Sherman 1982].

As pointed out by Fabian and Kembhavi, there is an additional difficulty with the hot IGM model: their calculations assume that the medium is in thermal equilibrium but "the electron-ion coupling time generally exceeds the age of the hot gas [so that] the assumption of an underlying Maxwellian distribution may fail." We will consider this aspect in a little more detail. We investigate the constituents of the ionized medium separately, ignoring nuclei heavier than hydrogen. A good measure of the timescales required for the electrons or protons to come to a thermal, Maxwellian distribution is their self-collision time [Spitzer, 1962, his equation (5-26)]. For electrons this is

$$t_{ee} = 0.266 \ T^{3/2} \ / \ (N_e \ \ln \Lambda) \quad \text{seconds} \ . \qquad \text{[VIII.35]}$$

$\ln \Lambda$ is the electron Coulomb screening factor [Spitzer, equation (5-14) et seq.], which at the epoch of heating, $z_m$, is a slow function of our model parameters, i.e. $\Lambda \propto C^{-1/4} \ (1+z_m)^{1/2}$. For $z_m = 5$, the FK fits give $\ln \Lambda = 40$. Other than this slow dependence through $\ln \Lambda$, $t_{ee}$ at $z_m$ is independent of the value of $z_m$:

$$t_{ee} \ H_o \approx 0.08 \ h_{50}^{1/2} \ C^{-1/2} \ . \qquad \text{[VIII.36]}$$

This is longer than the adiabatic cooling time, but is sufficiently short so that the electrons should be able to come to a Maxwellian distribution within a reasonable length of time to provide the observed thermal characteristics of the X-ray sky spectrum. In the equations of thin thermal bremsstrahlung it is the electron distribution and temperature that determines the emission characteristics.

It is much less clear that the protons can be treated as having reached thermal equilibrium. If we assume that the proton and electron temperatures are the same, then the self-collision times scale inversely

as the particle velocities, or

$$t_{pp} = (m_p/m_e)^{1/2} t_{ee} \approx 43\ t_{ee} \approx 3\ H_o^{-1} \ . \qquad\qquad [VIII.37]$$

The estimates of the energy requirements for heating the IGM assumed equipartition of the energy between the electrons and protons. The initial mechanism need not be so impartial; for example the energy may be preferentially deposited in the protons, as in mechanical shock heating (see the discussion in Hamilton, Sarazin, and Chevalier [1983], also Protheroe and Kazanas [1983]). (However McKee [1974] argues that plasma instabilities at the time of heating may provide equipartition.) Spitzer (his equation 5-30 et seq.) estimates that the characteristic time for achieving equipartition is

$$t_{eq} = 946\ [1 + (T_p/1836T_e)]^{3/2}\ t_{ee} \ . \qquad\qquad [VIII.38]$$

Crudely, if initially the electron temperature is near zero, $T_e \ll 5\times10^{-4}\ T_p$, then even for the low densities and high temperatures of the IGM there is good p-e coupling, $t_{eq} \sim 10^{-3}\ H_o^{-1}$. As the electrons heat up, the equipartition time increases greatly. If $T_e$ is set to the temperature for the IGM at $z_m = 5$ inferred by FK, then as long as $T_p < \sim 200\ T_e$, $t_{eq} \sim 75\ H_o^{-1}$. Even without a detailed analysis we see that it is possible for the proton temperature to be orders of magnitude larger than $T_e$. The amount of energy required to heat the IGM would be multiplied by a similar factor. If it seemed difficult to obtain 0.4 Mev baryon$^{-1}$ before, the prospect of 4 to 40 MeV baryon$^{-1}$ is even more daunting. If on the other hand, the mechanism that energized the IGM preferentially accelerated electrons, then equation [VIII.38] indicates that in terms of equipartition, the protons essentially are totally decoupled and will not acquire any large fraction of the electron energy for time scales much longer then the period of emission. This would halve the energy required to heat the IGM.

The appearance of the clumping factor C in equations [VIII.28] and [VIII.36] shows that both the energy requirement and the coupling timescales are proportional to $C^{-1/2}$. For clumping to have a significant impact, $C \gg 1$ and therefore the filling factor must be small, $\Phi \ll 1$. Our simple examination of the excess variance limits implies that to acheive this the clump size must be small, $\lambda_s \ll \sim 100\ h_{50}^{-1}$ Mpc,

the exact limit depending on the amount of evolution in the volume emissivity (see Figure VII.32). For $\lambda_s \sim 10~h_{50}^{-1}$ Mpc the maximum allowed value for the clumping factor is $\sim 100$. Field and Perrenod, and McKee [1980], have examined the limits on clumping, assuming that the clumps were isothermal spheres. However, Field and Perrenod do not properly calculate the expected variance from the spheres and McKee, as we have commented, uses an erroneous expression for the detector size. The analysis presented here is incomplete as it fails to include important cosmological corrections; however we will outline the general argument and derive limits appropriate to our simple cosmological picture. Following the previous authors, we examine the condition that our spherical lumps be gravitationally bound:

$$GM_s/\lambda_s > 2kT_{IGM} / \mu ~.\qquad\qquad [VIII.39]$$

$M_s$ is the mass of the clump and $\mu$ is the mean molecular weight, $\sim 500$ MeV. Using equations [VII.60] and [VII.63] we can rewrite this limit as

$$\Phi^{-1} ~(\lambda_s/\lambda_H)^2 ~\Omega_{IGM} > 4kT_{IGM} / \mu c^2 ~.\qquad\qquad [VIII.40]$$

We can think of this condition as requiring a minimum size for the sphere to be bound. As an example we consider lumps at the current epoch, using the FK model for the hot IGM with $z_m = 5$:

$$[\Phi^{-1} / (\Phi^{-1}+1)^{1/2}] ~(\lambda_s/\lambda_H)^2 \gtrsim 1.2 \times 10^{-3} ~.\qquad\qquad [VIII.41]$$

If $\Phi^{-1} = 1$, the lower limit on a bound sphere is $\lambda_s \gtrsim 7~h_{50}^{-1}$ Mpc. Larger values of $\Phi^{-1}$, corresponding to larger clumping factors, allow even smaller spheres. An examination of Figure VII.32 shows that a 7 Mpc sphere and $\Phi^{-1} = 1$ are compatible with the limits from the X-ray data. In fact at that radius, the excess variance limits allow $\Phi^{-1}$ to be as large as $\sim 100$, although the limit derived from the dipole signal analysis limits $\Phi^{-1} \leq 10$. If the larger values are appropriate, $\Omega_{IGM}$ and the e-e self-collision times could be reduced by a factor of 10. Thus these large scale stable structures, which are allowed by this simple analysis, can have a significant impact on the requirements of a hot IGM model for the production of the X-ray sky flux. This is in contradiction to the earlier results which prohibited the existence of significant clumping at scales large enough to be bound (see especially McKee), so that a more rigorous examination of the situation will be

performed in a future project. The evolution of the volume emissivity and clump size for a hot IGM model can have important effects, allowing larger clumping for small $\lambda_s$ and requiring less clumping for larger $\lambda_s$ (see the "evolution" curve on Figure VII.32). Note that, even with clumping factors on the order of 100, the e-p equipartition time is still quite long, ~7 Hubble times. A possible weak point of this analysis is the assumption that $T_{IGM}$ on the right hand side of our limits for bound spheres is the same as the electron temperature derived from the FK fits to the X-ray sky spectrum. If $T_p$, the temperature appropriate for the stability of the spheres, is much greater than $T_e$, then a larger sphere radius is required, $\lambda_s \propto T_{IGM}^{1/2}$.

Future analysis may place tighter constraints on the existence of gravitationally bound lumps, however even given such limitations smaller, highly clumped, regions can still exist. Such structures could be unbound, or pressure confined, instead of gravitationally bound. The unbound structures would have a short lifetime for substantial X-ray emission, as their expansion and dissipation would quench their contribution. The reduced lifetime can have a strong impact. The ratio of the thermal energy required to the X-ray energy produced can be higher than would otherwise be the case for the same amount of clumping. The higher required densities, however, would also be responsible for shorter coupling times.

In this context the model of Bookbinder et al. [1980] can be seen as an extreme variation of the hot IGM model. In their approach the X-ray sky spectrum is produced by intragalactic gas heated by a burst of supernovae during the earliest stages of the history of the galaxies. Such sources would appear point-like, and if the emission period is long enough there will certainly be more than ~50 per square degree, so that the excess variance constraint is satisfied. The densities are more typical of the interstellar medium, that is, ~$10^4$ times that of the uniform IGM models. This is equivalent to a clumping factor of $C \sim 10^8$. The coupling times are also reduced by at least a factor of $10^4$. As the material is originally heated by the interaction of several supernova-produced shockfronts, the densities at thermalization are even higher,

essentially meeting any objections from the long coupling times.

Small isolated clumps have a possible disadvantage in making up the observed X-ray sky flux. The hot uniform IGM model has only a single temperature for a given epoch. In addition the decrease in the emissivity with increasing epoch selects a relatively narrow range of redshifts to dominate the total spectrum. This explains our observation today of a signal that from 3-50 keV is well fit by a single temperature exponential form. With isolated clumps, some mechanism that selects a particular temperature or relatively narrow range of temperatures is needed. Bookbinder et al. fix the temperature in their model by the supernova shock velocity. If $z_i$ is the redshift corresponding to the epoch of galaxy formation and the initial generation of supernovae, then the current epoch X-ray temperature is (see their equation (24b))

$$kT_x = 95 \text{ keV } V_{S9}^2 / (1+z_i) \text{ keV} ,  \hspace{2cm} [VIII.42]$$

where $V_{S9}$ is the shock velocity in units of $10^9$ cm s$^{-1}$. They regard $0.5 \rightarrow 1$ to be a plausible range for the value of $V_{S9}$. If this range actually reflects the conditions in newly formed galaxies, then there is a spread in $kT_x$ over a factor of 4, not including possible additional smearing due to emission from a range of redshifts. A spread in the observed $kT_x$ must encompass any intrinsic deviations from a single temperature thermal bremsstrahlung within each galaxy. Should future experiments actually observe such emission from young galaxies, it would be important to fit the A-2/A-4 residual spectrum to learn what restrictions can be placed on the allowed behavior of these early supernovae. The observed residual spectrum temperature, $23 \lesssim kT_x \lesssim 30$ keV, and the suggested range on $V_{S9}$, indicate that the epoch of X-ray emission and galaxy formation is relatively recent, $z_i \lesssim 3$. However, the flatness of the effective Gaunt factor of the residual spectrum may not be consistent with thermal bremsstrahlung emission at such low emitter-frame temperatures, $kT \sim 90$ keV. Analysis of this possible difficulty requires further work.

Other models

Carr [1979; 1980] has developed a thermal history of the universe

whereby accretion onto primordial black holes ionizes the early IGM.
The IGM temperature never gets high enough to produce significant X-ray
emission directly. Instead, the accretion at the largest end of the
mass range he considered, $M \sim 10^9$ solar masses, contributes to the X-ray
sky flux through thermal bremsstrahlung in the accretion disk. The
details of accretion disk formation and the black hole mass dictate the
emission temperature, although during the evolution of the accretion
environment a range of temperatures will be produced. The epoch of peak
contribution to the X-ray energy density corresponds to large redshift,
$z \sim 12$. The emitter frame temperatures of the accretion disk are in the
transrelativistic regime, $kT \sim 400$ keV. Carr did not explicitly try to
fit the spectrum from such a model to the observed sky spectrum, and in
particular did not calculate the effect of pair creation. At these
temperatures a compact emission region, such as an accretion disk, can
suffer significant distortion of the emitted spectrum [Kazanas and
Shafer, 1983].

Carr hypothesized that the primordial black holes would go on to
form the central engines in AGN. Building on this suggestion, Boldt and
Leiter [1981] (see also Leiter and Boldt [1982]) have developed a
unified model for the explanation of both the thermal residual component
and the nonthermal sky flux in the $\gtrsim 100$ keV regime. In essence it is a
return to the first model considered in this section: AGN are
responsible for the total background. However, their life cycle is
divided into two phases, an early pre-active galaxy (PAG) era when
thermal emission dominates, and a later era when emission is in the form
of the power law spectra observed in current epoch AGN. An accretion
disk around a large black hole is again the origin of the PAG thermal
emission, but this time instead of thin thermal bremsstrahlung, the
process is unsaturated Comptonization of soft photons produced in the
outer regions of the disk. The unique temperature of the sky flux is
determined by a thermostatic process of pair production within the disk
[Lightman 1982] and additional cooling from unsaturated Comptonization,
holding the temperature in the range 90–180 keV. An additional feature
of the Leiter and Boldt model is that the black holes are assumed to
have little or no angular momentum initially. However, as accretion

progresses and the hole is spun-up to a canonical Kerr black hole, the
model predicts a distinct change in the nature of the emission processes
from thermal photons to the production of relativistic electrons. These
in turn are responsible for the AGN synchrotron and SSC emission
observed in the current epoch. The thermal PAG era is relatively short,
$\sim 10^8$ years. The expected surface density of sources is large, 200-1000
$deg^{-2}$, well within current requirements for those sources responsible
for the bulk of the X-ray sky flux.

—————————————————

3: Future Analysis of the X-ray Sky with A-2 Data and Other Experiments

> Several important questions about the origins and properties
> of the X-ray sky remain to be answered by future studies.
> These may utilize the A-2 database as well as future
> generation experiments.

### Different angular binnings from the A-2 database

The results of this dissertation are only a fraction of the
information available from the A-2 database. Limiting ourselves to
information available from the fluctuations, we can apply the techniques
we have developed to different detectors (e.g. HED 3), rebinned data,
new FOV combinations, and different Discovery Scaler spectral choices.

Fitting P(D) models to HED 3 data by itself should simply confirm
our analysis of the HED 1 dataset. Any combined answer from the two
detectors as separate datasets would require the use of Boole's
Inequality, as they are not independent estimates of the statistical
properties of the fluctuations. Data from both detectors can be
combined and treated as coming from a single detector with effective
response equal to the sum of the spatial responses of the separate
detectors. This is the same process as combining the two fields of view
of a single detector to make the L+SFOV measurements. One advantage of
such a direct combination is a higher count rate per combined exposure.
However, because of the small solid angle of HED 3, adding HED 3 data to

HED 1 data will only reduce the $\sigma_{photon}$ by ~18%. A second advantage is that the HED 3 contribution will concentrate the net response to the center of the measurement, decreasing the overlap between adjacent data points, and thereby increasing the effective number of measurements in the overlapping datasets. A difficulty in the combination process is that the two detectors may have different levels of internal background, B, or additional variance, $\sigma^2_B$, due to temporal changes in the internal background rate. Also, the ratio of the numbers of exposures from each detector will vary from one position in the sky to the next, complicating the calculation of either the effective detector spatial response or the combined measurement photon counting statistics.

A very real difficulty in interpreting the excess variance is that there are multiple explanations of the phenomenon: source evolution, non-Poisson clumping of point sources, or a general large scale structure in a truly diffuse component. We could begin to distinguish among these possibilities if we had accurate measurements of the excess variance from datasets with different angular scales. For example, the evolution of point sources will cause $\sigma^2_I$ to vary proportionally with the change in the weighted response $R_2{}^2$. The variation due to large clumps will have a different signature, the details depending on the actual model for the cosmology and the clumping. In practice, the large uncertainty in the value of the excess variance would mask the source of $\sigma^2_I$. Still, the upper limit on any excess for smaller binnings of the data may allow for more stringent limits on large scale clumping as in the simple Euclidean universe model of Section VII:6 (see especially Figure VII.32). As an example, compare the effects of the different sky fields of view for HED 1 and HED 3. The transition size between large angular scale sources and sources that are essentially point like, $\lambda_{crit}$, (see equations [VII.75], [VII.82] et seq.), is a function of the angular scale of the measurement. For the HED 1 L+SFOV with 6° smear data that we have used, $\lambda_{crit} \approx 390 \ h_{50}{}^{-1}$ Mpc. For data using just the SFOV of HED 3 and smeared by 3°, the critical scale drops to 240 Mpc. If the two sets of data had similar upper bounds for $\sigma^2_I$, then the HED 3 limit places a more severe restriction on the allowed amount of smaller scale clumping. Unfortunately we do not expect a priori that the limits

would be the same magnitude. Although the smaller angular size of the HED 3 detector would allow for a larger number of independent measurements, the signal-to-noise is quenched substantially.

An analysis of the fluctuations using small angular scale measurements can also be approached using the "-Y" pointing data. This is data taken during a series of pointing operations in which target positions were selected to benefit the A-1 experiment detectors oriented opposite to the A-2 detectors, so that the A-2 data reflect "random" positions on the sky and can therefore be used as an unbiased estimate of the sky flux. Although the number of independent data points is small, ~50 at high galactic latitudes, the total exposure time for each point is high, ~$10^4$ seconds. Even for the HED 3 SFOV rate alone, the photon statistic noise is significantly smaller than the fluctuations, $\psi$ ~ 0.4. In addition, the smearing due to jitter during the point operation is small. The critical lump size drops further to ~180 Mpc. For this reason, and the fact that precision in a measurement of $\sigma^2_I$ depends more on $\psi$ than on the number of independent observations, stringent constraints on large scale structure may be obtained from the analysis of the -Y point data fluctuations.

The possible existence of large scale structure may be investigated directly in the A-2 all sky database by looking for significant power in the spatial autocorrelation function. By significant we mean over and above the expected power from a Poisson distribution of point sources folded through the finite angular size of the detector. One approach involves fitting the higher order spherical harmonics in addition to the simple dipole of the Compton-Getting effect, as Fabian, Warwick, and Pye [1980] did with Ariel V data. Perhaps the cleanest approach to high angular resolution would be to examine the one-dimensional autocorrelation functions of individual superposed scans (for an early example of this see Schwartz et al. [1971]). Such scans allow the detector response to be cleanly removed using deconvolution techniques. Spurious structure due to variations in the detector internal background can be addressed by examining the cross correlation between scans of the same physical region of sky taken six

months apart. Care must be used in interpreting the statistical
significance of any result obtained from several scans that may overlap.
If an actual signal appears, rather than an upper bound, it could be
worthwhile to analyze the entire dataset using two-dimensional
autocorrelation techniques. A similar program using optical data has
been presented by Schectman [1974].

## The fluctuations in other spectral bands

Another aspect of the A-2 data that has not been investigated is
the spectral information available from the various Discovery Scaler
windows. Analysis of the fluctuations for windows effectively lower in
energy than the H1L1 rates, such as the MED 2A window, may have a
significant additional variance due to the enhanced galactic
contribution in that band. A higher energy band, e.g. the HED 1 Layer 1
C+D windows (hereafter the H1CD rate), will have a decreased galactic
contribution to the fluctuations. This band covers 5.6 to 17 keV for a
kT = 40 keV thermal bremsstrahlung spectrum (see Section II:3 and tables
in Section A7:1).

If we know the spectrum of the sources which cause the
fluctuations in the H1L1 band, we can predict the level of the
fluctuations in any other spectral band. As an example, we calculate
the expected results for the H1CD band. Using our results of Table
VII.5 we divide the sources that contribute to the high flux N(S) into
AGN, clusters of galaxies, and local galactic sources. From the values
of $K_h$ in Tables VII.4, VII.5, and VII.6 we can calculate (using equation
[III.45]) $W_{h;AGN}$, $W_{h;cl}$, and $W_{h;gal}$, in units of H1L1. We assume that
the spectral form for AGN is a power law with $\alpha = 0.65$. From the
conversion coefficients of the various window combinations in Section
A7:1,

$$W_{h;AGN} = 0.033 \text{ H1L1} = 0.014 \text{ H1CD} . \hspace{2cm} [VIII.43]$$

This is a slight underestimate as it does not take into account the
effect of absorption for the low luminosity AGN in the H1L1 band. The
presence or absence of absorption has a negligible effect for the higher

energy H1CD band. We also assume that clusters and galactic sources are described by a low temperature thermal spectrum, kT = 6 keV:

$$W_{h;cl} = 0.034 \; H1L1 = 0.009 \; H1CD \; ; \qquad\qquad [VIII.44]$$

$$W_{h;gal} = 0.017 \; H1L1 = 0.005 \; H1CD \; . \qquad\qquad [VIII.45]$$

To find the total W we convert the individual components to their K values to find $K_{h;tot} = 2.3 \times 10^{-4} \; (H1CD)^{3/2}$. The contribution from AGN is 60% of this total, up from only 41% of the H1L1 total. The value for $W_{h;tot}$ is therefore 0.02 H1CD. In the wider energy band of H1L1, $W_{h;tot}$ was 0.06 H1L1. The transformation of the all sky flux, $S_{as}$, is from 58 H1L1 to 26 H1CD. Because the fractional reduction in W is greater than for $S_{as}$, the size of the fluctuations as a percentage of the surface brightness is smaller in the H1CD band than the H1L1 band. This is the expected consequence of the sky flux spectrum being flatter than known source spectra. This trend would be reversed in bands that cover energies higher than the 40 keV exponential rollover of the sky spectrum. Because it is flatter than the sky flux, the internal background is a larger fraction of the total count rate in 1C+D intensities, $B = 1.6$ counts $\exp^{-1}$. A large fraction of this is due to background counts in the highest PHA channels where few real X-ray counts are accumulated for a 40 keV thermal spectrum. The internal background could have been reduced substantially if there had been a high energy Discovery Scaler window that excluded counts above 20 keV. Such a rate can be constructed from the 1C window, but this was usually not available under the Discovery Scaler telemetry definitions. For the 1C window $B = 0.95$ counts $\exp^{-1}$.

There are several possible advantages to the analysis of the higher energy band. Because this band is dominated more by AGN the analysis would be more sensitive to their possible evolution. This is particularly true if the spectral form of AGN has already begun evolving in conformance with the requirements of the total sky flux spectrum. Similarly, if there is a true excess variance due to the clumping of a hot IGM, this signal will be enhanced with respect to the fluctuations from the nearby AGN and clusters. However, the total reduction in the

number of counts in this band produces an increase in the size of the counting statistics compared to any fluctuations signal. Thus, the analysis of data from this band may be improved by combining HED 1 and HED 3 data, or by using the -Y point data. Measurements of the excess variance suffer from the additional difficulty that the smallest confidence regions require data from the resolved source counts. This would necessitate an extrapolation from the XCS R15 count rates, making assumptions about the source spectra. An alternative would be to redo the analysis of Piccinotti et al. [1982] restricted to the H1CD band. Because of the poorer counting statistics, we would expect that the number of sources in such a flux limited sample would be sharply reduced. Therefore the expected upper bound in the excess variance would be larger. A measurement of the Compton-Getting dipole signal may be improved by using the H1CD band. Not only is the galactic signal slightly suppressed, but the size of the fluctuations, a significant noise term for the dipole measurement, is reduced.

## Study of sky flux from a future large area, low resolution X-ray experiment

Moving beyond the all sky survey of HEAO 1, we consider several model experiments. The first two concentrate on surface brightness studies, similar in form to our analysis of the A-2 data. These studies are presented in outline only, and are not even optimized for the study of the fluctuations as we discussed in an earlier section on ideal experiments. The requirements of these experiments are not far from current technology. We will not refer to any specific proposal for a future experiment although similar experiments have been proposed.

The 1C+D window of the A-2 experiment HEDs suffers from a low count rate because it is a high energy band. Our first hypothetical experiment attempts to overcome this restriction by using very large area proportional counters. We assume the detectors used are xenon proportional counters similar to the HED design. Let $A = 8000$ cm$^2$, with rectangular collimators with FWHM = 1$^\circ$. The smaller solid angle in comparison to the A-2 collimators reduces the number of sky counts, and

diminishes the signal to noise ratio; howsever it also allow the easy investigation of possible clumping or other large scale structure at smaller angular scales, $\lambda_{crit} \approx 100$ Mpc. In place of the kind of all sky survey performed by HEAO 1, we consider what a limited observation period of a limited region of sky can tell us. Let the total observation time be $5\times10^5$ seconds (~5.8 days), divided into 500 exposures. If the exposures are only $1^0$ apart we can perform a raster scan on a 500 square degree area, which can be divided into 4 subsets of 125 non-overlapping measurements. In the spectral range equivalent to the 1C+D window the expected value for $W_{eff}$ is 100 counts in the $10^3$ second exposure. $\bar{I}$ is expected to be 5000 counts $exp^{-1}$, so that in the absence of any internal background the total noise variance is 71 counts $exp^{-1}$. Therefore $\psi \sim 0.7$, indicating that the properties of the fluctuations can be measured easily. Sources with intensities on the order of $W_{eff}$ have corresponding fluxes of about $3\times10^{-5}$ H1CD $\approx 10^{-4}$ H1L1. This flux range is expected to be dominated by moderate redshift AGN.

The small angular scale of the collimators also improves the sensitivity of the measurements to source evolution. For $\sigma^2_S = 10^{-4}$ (H1CD)$^2$, corresponding to an evolved population with $\gtrsim 130$ sources per square degree, the expected excess variance is ~260 counts $exp^{-1}$. As this is larger than $W_{eff}$ we might expect a significant improvement in the invariant excess limits over the A-2 data, equivalent to ~$2\times10^{-3}$ (H1CD)$^2$. However, the promise of such an observational program is spoiled, at least in part, by the internal background. Because of the small collimator field of view the internal background is larger than the sky contribution to the total intensity. Even for the 1C window the specific background is ~$2\times10^{-3}$ counts $s^{-1}$ $cm^{-2}$ (extrapolating from the A-2 detectors), which will produce 15600 counts $exp^{-1}$. Combining this rate with the sky flux contribution to the intensity, the revised photon statistics noise is 144 counts $exp^{-1}$, now larger than the fluctuations $W_{eff}$. Furthermore, temporal instability of the internal background rate can introduce even more additional variation. If the best fit estimate for $\sigma_B$ derived from HED 1, 1.3% of the background rate or 203 counts $exp^{-1}$, is any guide, this term overwhelms the counting statistics as

well as the fluctuations distribution. The total noise variation, $\sigma_{noise} \sim 250$ counts $\exp^{-1}$, produces a $\psi \sim 2.5$. Examination of Figure III.9 shows that even with this poor a signal to noise, with $\sim 125$ independent measurements we are able to determine $W$ to within $\sim 20\%$ (90% confidence). However, this presumes that we know $\gamma$ and the exact value of the background variance. Imprecision in $\sigma_B$ of $\sim 30\%$ will make it essentially impossible to accurately estimate the power law component of the fluctuations. Therefore, for any studies of the sky flux on scales of a degree or so using standard proportional counter technology, an accurate understanding of the internal background and its variations is required.

### A moderate area, moderate resolution imaging experiment

The answer to this problem is to break the proportionality between sky counts and internal background counts by using the concentrating aspect of an imaging detector. We consider a system with an area equivalent to 500 $cm^2$ of the first layer of a xenon detector. For illustration we will assume that the spectral response is similar so that all fluxes are measured in HlLl units. We assume that the total image plane is 0.25 square degrees, divided into 25 elements of $6'x6'$ in extent. Again the total observation time is $5x10^5$ seconds, but now we survey only 125 square degrees. Because of the multiplexing effect of the imaging system, the total active time for each element is $10^3$ seconds, providing 1250 elements with exposures of $5x10^5$ $cm^2$ s. We model the response for each element with a flat profile, so that the various weighted response moments are: $R_1 = 0.12$ $cm^2$ s, $R_{3/2} = 19.4$ $cm^2$ s, and $R_2 = 246$ $cm^2$ s. The mean X-ray sky intensity per element is 7 counts $\exp^{-1}$, while the expected value for $W_{eff}$ is 1.2 counts $\exp^{-1}$. Though $W_{eff}$ is smaller than the expected photon counting statistics, the large number of observations promises that the fluctuations will provide significant information down to and below the one-source-per-field-of-view limit, $\sim 2x10^{-6}$ HlLl, which is comparable to the fluxes of the sources in the Einstein Deep Survey. The small angular scale of the measurement means that excess variance results will help constrain the behavior of clumping down to "small" scale sizes, $\lambda_{crit} \sim 10$ Mpc.

Again, the internal background can have a major impact. However, estimating the background is less certain than in the first example. This is in part due to the relatively recent development of imaging detector technology in comparison to traditional proportional counters. Also variables of the telescope plate scale, spectral band width, satellite orbit, etc., play an important role. If we use as an estimate the IPC background as determined by Stewart and Fabian [Fabian 1981], suitably scaled for solid angle, integration time, and bandpass, $B \approx 40$ counts $\exp^{-1}$. Thus, from this estimate the intensity measurements are internal background limited. However, future refinements of the technology should produce a striking reduction, perhaps by as much as an order of magnitude. In this case the background will be less than the X-ray contribution to the intensity. Note that even for the high background rate, a 1.3% variation corresponds to a non-X-ray fluctuation of 0.5 counts $\exp^{-1}$, less than the expected magnitude for $W_{eff}$. Of course, as we have pointed out, an additional consideration for imaging systems is the additional variation from spatial as well as temporal inhomogeneities in the internal background.

If the internal background is reduced to a level less than the surface brightness of the X-ray sky, then the imaged survey of the 125 square degrees can provide an additional dividend: an X-ray selected sample of $\sim$85 sources down to fluxes 40 times deeper than the XCS. The broad band study of such sources can make the important connection between the high energy band with the results of the _Einstein_ Medium Survey.

<u>Imaging the X-ray sky with future instruments</u>

Current observations are not sufficient to determine decisively if most of the X-ray sky flux originates in point sources (<u>e.g.</u> QSOs or young galaxies) or if a significant fraction is produced by a truly diffuse component such as a hot IGM. The _Einstein_ Deep Survey observation directly images $\sim$25% of the extrapolated sky flux, but this is open to argument because of ignorance of the Deep Survey sources' spectra. A definitive answer will probably be reached only when a high

spatial resolution imaging instrument is able to continue the 2-10 keV band source counts down to a flux level where the dominant fraction of the total flux is accounted for. We consider the requirements for such a system under three ad hoc but plausible models for the source counts in the HILI band. As our first model, let N(S) follow the Euclidean $S^{-5/2}$ form until the total sky flux is saturated. We assume that W is the same as our fiducial value derived from the fluctuations, 0.0702 H1L1. This is slightly higher than the value we derive from the fits of a Euclidean model with an excess variance, including the information from the resolved source counts from the XCS, but may be closer to the correct value of N(S) over the flux range below the XCS. The sky flux is saturated at $S_o = 4.1 \times 10^{-7}$ H1L1 $= 5.6 \times 10^{-15}$ ergs $s^{-1}$ $cm^{-2}$. In the absence of any background, we would require exposures on the order of $5.6 \times 10^6$ $cm^2$ s for a detector with response similar to the HEDs. This will produce at least one photon from 90% of the lowest flux sources. The total number of sources, all sky, required to make up the sky flux is $4.7 \times 10^7$, or one per pixel of $(1.8')^2$. To reduce problems from confusion we note that if we have pixel sizes of $(47'')^2$ so that the expected number of sources per pixel is 0.2, then over 90% of the pixels with sources will contain only one source. Therefore, the X-ray sky flux can be resolved with relatively moderate angular resolution. However, in the presence of significant specific internal background, clean detections may require much higher resolution, particularly if extraordinarily long exposures are to be avoided. Smaller sized pixels will have fewer background counts for a fixed level of specific background. A similar improvement in the signal to noise could be achieved by increasing the collecting area of the X-ray telescope. For example, a pixel of $(30'')^2$ with internal background similar to the IPC will have ~0.3 counts in $10^3$ seconds. Therefore a 1000 $cm^2$ collecting area telescope will have ~1.6 background counts compared to an average of 2.4 counts from sources with flux $S_o$ during an integration time of 5600 seconds.

As our second model, we assume that N(S) breaks from the Euclidean form at the point where 20% of $S_{as}$ is accounted for. This is at $1 \times 10^{-5}$ H1L1 $= 1.3 \times 10^{-13}$ ergs $s^{-1}$ $cm^{-2}$, roughly the level of the Deep Survey.

The remaining 80% of $S_{as}$ can be accounted for if N(S) breaks to a power law form with index $\gamma' = 1.88$. Now at the flux level, $S_0$, where formally the total sky flux is accounted for, the accumulated sources provide only 47% of the total. Our studies at this flux level must include the possible effects of confusion with the sources below $S_0$. If the background permitted, we could extend our knowledge about N(S) below $S_0$ using the fluctuations techniques derived in this dissertation. However, the flatness of N(S) greatly increases the sensitivity of the fluctuations to only the highest flux unresolved sources, $\sim S_0$. The expected $\sigma$ for the $(47")^2$ pixel size is about 14% of the intensity of a source with flux $S_0$. With exposures of $5.6 \times 10^6$ cm$^2$ s the ratio of the counting statistics $\sigma$ to the fluctuations $\sigma$ is $\sim 4.5$. A study of the fluctuations will require very long exposures or huge collecting areas.

If point sources provide only $\sim 30$% of $S_{as}$ then N(S) down to $S_0$ should exhibit a very sharp break. Such behavior could be interpreted as evidence for the hot IGM model. This evidence is only circumstantial, however, since a new population of low flux point sources may provide the bulk of $S_{as}$ below $S_0$. If this is the case, they must have a very high surface density, $\gtrsim 2400$ sources per square degree. Recall that the Leiter and Boldt PAG model estimated that the surface density is $\lesssim 1000$ per square degree, which would have an average flux $\gtrsim 10^{-6}$HlLl = $1.2 \times 10^{-14}$ ergs s$^{-1}$ cm$^{-2}$, only an order of magnitude below the Deep Survey limit. Models associated with young galaxies are expected to give similar surface densities, although these calculations are very sensitive to the lifetime of the thermal X-ray emission epoch. If the additional density of sources is $\sim 1000$ per square degree, pixel sizes of $\sim (40")^2$ are sufficient to resolve the sky flux, in the absence of appreciable internal background.

The X-ray Sky Flux                                                  Future Work

CHAPTER IX

CONCLUSIONS

The general question addressed by this dissertation, the nature of the unresolved X-ray sky flux, especially as examined through its spatial fluctuations, is easily stated. It is less easy, however, to give a succinct definitive answer. Many of our results are of a procedural, rather than observational or theoretical, nature. By the close examination of the tools used for the study of the sky fluctuations we have a much better understanding of the limitations of previous work, the implications of our present results, and the possibilities for future understanding. Our observational results, though significant in their own right, are just one step in the eventual solution of the X-ray sky.

Our results, in summary:

Chapter II:

The multiple field of view construction of the A-2 experiment provides a direct estimate of the internal background, which is important for distinguishing variations in the sky flux from spurious signals. The distinction between the flux of a source and its contribution to an intensity measurement as determined by its position in the detector field of view, indicates the importance of the proper treatment of the detector spatial response function.

In particular, the weighted detector responses express the effect of the collimator size on fluctuations measurements in a less confusing way than the "effective solid angle."

Chapter III:

A model for the number of sources versus flux produces a predicted distribution of sky intensity measurements. For power law models of $N(S)$, the width of the resulting distribution scales with the effective

width, $W_{eff}$. This quantity provides a better characterization of the size of the fluctuations than the variance or other moments, which are formally infinite due to the high intensity tail of the fluctuations distribution.

As developed in this dissertation, the likelihood difference measures the extent to which two model distributions are distinguishable. This allows us to investigate quantitatively the sensitivity of a fluctuations analysis to the behavior of N(S) at different flux levels. In principle, we can determine N(S) for all fluxes, though for the lower fluxes this ability is degraded given a finite number of observations or the presence of significant non-fluctuations noise. For our data, the behavior of sources below the one-source-per-field-of-view level is evidenced primarily by their contribution to the variance of the fluctuations distribution. For this reason, a general model consisting of a Euclidean index power law component with an additional pure Gaussian component is an adequate description of all realistic N(S) models used to fit the HEAO 1 A-2 data.

The likelihood difference also allows us to evaluate the competing factors in the design of ideal experiments for the measurement of the fluctuations.

The existence of the fluctuations can have substantial effects on the measurement of other quantities, particularly low flux sources and surface brightness correlation studies. For a given angular collimation size, the fluctuations provide an irreducible source of noise to many measurements, including measurements of the properties of the fluctuations distribution.

Chapter IV:

The analysis of the fluctuations requires that the individual measurements either be strictly independent (no overlap) or that they be weighted. We have estimated what the weighting correction for overlapping data should be. This estimate is exact in the limit that the fluctuations are Gaussian.

The internal background variations for the A-2 detectors are very low. For HED 1 the variation is consistent with zero.

A thick ($\gtrsim$ 1 kpc) finite radius disk provides a good description of the component of the X-ray sky flux associated with the galactic plane.

An additional large scale cosine feature is detected with significance ~95%. The best fit direction is ($\ell$,b) = ($282^\circ$,$30^\circ$), with an amplitude of 0.53$\pm$0.18% of the total X-ray sky brightness. This is consistent with the interpretation that the signal is from the Compton-Getting effect with a velocity of 475$\pm$165 km s$^{-1}$, in agreement with the direction and velocity derived from the microwave background dipole signal.

Chapter V:

When fitting a model distribution of the fluctuations to the A-2 database, the traditional histogram statistics measuring the goodness of fit are inappropriate because different measurements have different amounts of noise from photon statistics. We have developed a set of statistics that provide a general indication of the goodness of fit of a model under a wide variety of situations. Confidence contours are found from the likelihood statistic. The information from different determinations of a confidence region may be combined using Boole's Inequality, even if the determinations are not statistically independent.

Chapter VI:

The data used to fit the fluctuations models were from the Large plus Small Field Of View combination of HED 1, binned by $6^\circ$. The resultant region of the sky covered by each measurement is effectively $11.2^\circ$x$4.4^\circ$. The chosen rate, the entire layer 1, is dominated by counts from 2.5 to 13.3 keV for an incident spectrum like the X-ray sky flux.

The mean count rate is an important parameter. A significant bias can be introduced if we blindly assume that the root-mean-squared count rate is equivalent to the mean of the fluctuations model distribution. This can be seen when the value of the model mean is treated as a free parameter. There is a strong interdependence between the model mean and $W_{eff}$ unless we specifically truncate the contribution of sources brighter than ~$3$x$10^{-11}$ ergs s$^{-1}$ cm$^{-2}$ (2-10 keV) to the mean.

The fluctuations distribution derived from a single Euclidean power law model for N(S) provides an acceptable fit to the data according to our goodness of fit criteria. The quality of the fits for models with sharp cutoffs at low flux levels explicitly demonstrates that the data is sensitive to the behavior of such low flux sources only in terms of their contribution to the variance of the fluctuations distribution. The confidence regions for W derived from individual sets of non-overlapping data are consistent with each other and with the range from the overlapping dataset. The 90% range for W from the overlapping data is $9\times10^{-10}$ to $10\times10^{-10}$ ergs(2-10 keV) $s^{-1}$ $cm^{-2}$. W determines the normalization of the N(S), N(S) dS = $(S/W)^{-5/2}$ d(S/W). The parameterization can also be in terms of K, N(S) = $4\pi K$ $S^{-5/2}$, so that $W = (4\pi K)^{2/3}$.

When non-Euclidean power law forms of N(S) are fit to the fluctuations, the Euclidean value of $\gamma$ = 5/2 for the index is always within the allowed confidence region. For the independent datasets, a wide range of $\gamma$ may be allowed, from 2.2 to over 3.4. The overlapping data is more narrowly constrained to the range from 2.4 to 2.8. When we plot the N(S) curves for the allowed models we see that the N(S) is most constrained for sources with fluxes $\sim4\times10^{-12}$ ergs $s^{-1}$ $cm^{-2}$. This is roughly an order of magnitude below the limits of the resolved source count surveys attained with the HEAO 1 A-2 data. It is also a factor of two below the level where one source is in the measurement field of view, illustrating that this traditional lower limit of sensitivity for the fluctuations is only approximate.

Providing an automatic low flux cutoff in N(S) when the total sky flux is saturated is important when testing power law models with high values of $\gamma$.

The decomposition of N(S) into a Euclidean high flux component and an additional term specified only by its contribution to the variance of the fluctuations is a general and flexible form for comparing the results from the fluctuations to realistic N(S) models. From the overlapping data the upper bound on the excess variance is 1.5% of the total sky flux. In detector-solid-angle-invariant terms this limit is $\sigma^2_s \leq 15\times10^{-20}$ (ergs $s^{-1}$ $cm^{-2}$)$^2$. Negative excess variances are also allowed, corresponding to models with a deficit of low flux sources

compared to the simple Euclidean power law with no low flux cutoff. The actual size of the excess variance is strongly coupled to the strength of the Euclidean power law component of the fluctuations, W. The confidence region derived from a set of independent non-overlapping data is essentially unable to constrain the excess variance without some additional information about the proper value for W.

The limits on the excess variance from the overlapping data can be derived solely from the analysis of the fluctuations. Previous bounds on the excess variance have required information from resolved source counts.

The results obtained for the W-$\gamma$ confidence region are in essential agreement with previous results obtained by other experimenters using data from the UHURU and Ariel V satellites. The published size of the early confidence regions are similar to or smaller than our own, but this is due to assumptions made in the early analyses, particularly in not automatically imposing a low flux cutoff in N(S) when the total sky flux is exceeded. Early upper bounds on the excess variance almost certainly should be revised upward substantially.

Section VII:1:

We have determined the coefficient of the high flux Euclidean component of N(S) using the X-ray selected sample of resolved sources of Piccinotti et al. [1982], which is complete down to fluxes $\sim 3 \times 10^{-11}$ ergs s$^{-1}$ cm$^{-2}$. We used binned data, corrected for the Malmquist bias, and the likelihood of the Poisson distribution to determine the best fit values and confidence ranges for the normalization coefficient. This was in excellent agreement with the confidence regions for the two component fluctuations model of a Euclidean N(S) plus an excess variance.

By combining information from both the fluctuations and the resolved sources, the size of the excess variance confidence range was improved. Interesting constraints then resulted, even from the fits derived from non-overlapping data. When overlapping data was used, the 90% range on the excess variance was between 0.0% and 1.3% of the mean sky intensity, or $0.0 < \sigma^2_S < 1.1 \times 10^{-16}$ (ergs s$^{-1}$ cm$^{-2}$)$^2$.

Section VII:2:

   We used the resolved sources to derive the local luminosity functions. This, in turn, allowed us to predict the form for N(S) below the resolved source flux limit. Cosmological effects produce a rollover from the Eulidean 5/2 power law. When the expected contributions to N(S) from clusters of galaxies, AGN, and moderate-luminosity galactic sources (CV and RS CVn stars) are summed, assuming no luminosity function evolution, the sources are unable to provide the observed total sky flux, $S_{as}$. Therefore, evolution of known sources or a new source of X-ray emission is required. The rollover in the total N(S) is sharp enough that a significant negative excess variance is expected, $\sim -6 \times 10^{-20}$ (ergs s$^{-1}$ cm$^{-2}$)$^2$. This is not consistent with the overlapping data lower bound on $\sigma^2{}_S$. This disagreement can be resolved by providing additional sources of fluctuations, such as non-Poisson behavior in the distribution of the non-resolved sources. An alternative source of excess variance might be evolution of an existing population of sources or the existence of a new, evolving, population of souces. One of these possibilities is required if point sources produce the total sky flux, in any case.

   Because of their soft spectra and flat luminosity function, clusters of galaxies roll over, that is depart from Euclidean behavior, at a particularly high flux. They contribute only $\sim 3.5\%$ of $S_{as}$, and are responsible for the bulk of the expected deficit for the excess variance. These conclusions are relatively insensitive to details of the cluster luminosity function normalization, or its low luminosity cutoff. Non-Euclidean corrections are important for clusters even for the resolved sources, requiring a recalculation of the cluster luminosity function normalization.

   AGN are expected to dominate the source counts at low fluxes. However the details of their behavior and their total contribution to $S_{as}$ are subject to significant uncertainties as the form of the low end of the AGN luminosity function is poorly known. X-ray sources with luminosities as low as $10^{40}$ ergs s$^{-1}$ have been observed by the <u>Einstein</u> Observatory, but $1.75 \times 10^{42}$ ergs s$^{-1}$ is the lowest luminosity object in the A-2 bandpass complete survey. Preliminary studies of these low luminosity AGN by Elvis, Soltan, and Keel [1983] indicate that the

luminosity function flattens at lower luminosities, beginning somewhere in the range from $10^{42}$ to $10^{43}$ ergs $s^{-1}$. The exact location and magnitude of the flattening have a large impact on the expected contribution of a non-evolving AGN population to $S_{as}$. Crude estimates range from 18% to 46% of the total sky flux. Our canonical model luminosity function produced ~25%. Similar imprecision exists with respect to the contribution of AGN to the negative excess variance, but this is overshadowed by the much larger contribution from clusters. One difficulty in determining the location of the flattening is the lack of quality broad band spectra of the low luminosity sources to accurately connect the luminosity functions determined in the different bandpasses of the A-2 experiment and the Einstein Observatory.

CV and RS CVN stars make a significant contribution, ~15%, to the source counts at the higher fluxes, $S \gtrsim 10^{-12}$ ergs $s^{-1}$ $cm^{-2}$. As the fluctuations analysis cannot discriminate between galactic and extragalactic sources, the expected contribution of galactic sources must be estimated. They are expected to contribute more to the negative excess variance than AGN, though less than clusters. The exact value depends on their typical scale height and the steepness of their luminosity function.

Section VII:3:

Any new population of sources is constrained to not make a significant contribution to the resolved source counts and to make a contribution to the excess variance from $5 \times 10^{-20}$ to $16 \times 10^{-20}$ (ergs $s^{-1}$ $cm^{-2})^2$. The lower limit on the excess variance is soft, as there may be other ways of producing the observed excess variance, e.g. evolution by the known sources, or source clumping. For the same reason the upper limit on the excess variance from a new population of sources may be reduced when further observations constrain the other possible sources of fluctuations. With the current limits, if we assume that the new population is responsible for the 75% of the total sky flux otherwise unaccounted for, then its surface density is $\gtrsim 50$ sources per square degree with an average flux of $\lesssim 3 \times 10^{-13}$ ergs $s^{-1}$ $cm^{-2}$. Stronger conclusions can be reached using results from the Einstein Observatory Deep Survey counts (see VII:4 below) but these are subject to

uncertainties dependent on the Deep Survey source spectra.

Evolution by AGN can provide both the required sky flux and the excess variance. However details of the form of the evolution and cosmological parameters can have a strong impact. We have presented a discussion of different models for evolution, pure density, pure luminosity, and index evolution, and illustrated their general forms under a variety of assumptions. Conclusions drawn from the study of a particular class of models depend on the current epoch luminosity function used, as well as the value chosen for $q_o$.

We have particularly examined pure luminosity evolution models where the AGN luminosity effectively evolves as $L(\tau) = L(0) \exp(Q_L \tau)$, where $\tau$ is the fractional look-back time. For these models, the amount of evolution is usually constrained most by the total sky flux: $Q_L \lesssim 3.2$, essentially independent of $q_o$. This assumes that both low and medium luminosity sources evolve with the same value for $Q_L$. This upper limit is much less than values for the evolution derived for optical studies of quasars or even from the original analysis of AGN evolution from the _Einstein_ source counts. With such small values of $Q_L \lesssim 3.2$, AGN have little significant contribution to the excess variance. For small values of $q_o$, the contribution is much less than the soft lower limit to $\sigma^2_S$ required by the fluctuations. More evolution is allowed ($Q_L \lesssim 5.2$) if only medium luminosity AGN evolve ($L > 10^{43}$ erg s$^{-1}$). Now a significant contribution to the excess variance can be made. For $q_o = 0.1$, the excess variance upper limit becomes a stronger constraint ($Q_L \lesssim 4.5$) than the total sky flux limit. Pure luminosity evolution models where only the medium luminosity objects evolve can have steep rollovers in N(S) at $\sim 10^{-14}$ ergs s$^{-1}$ cm$^{-2}$, only a factor of two lower than the _Einstein_ Deep Survey lower flux limit.

For the pure luminosity evolution models, the current epoch high luminosity sources make no significant contribution to the sky flux. Because of the steepness of the current epoch luminosity function, such an extension cannot be directly studied using the A-2 resolved sources. However, if the high luminosity sources underwent strong evolution they would dominate the source counts at high fluxes. Given that only a single high luminosity object (3C273) is observed in the A-2 X-ray Complete Sample we can conclude that a high luminosity extension of the

Conclusions

AGN luminosity function out to $10^{46}$ ergs s$^{-1}$ can evolve only with $Q_L \leq$ 5. If the luminosity function steepens around $1.5 \times 10^{45}$ ergs s$^{-1}$, then $Q_L \leq 6$.

Luminosity-dependent density evolution (index evolution) can also provide the total sky flux and make a significant contribution to the excess variance. However these models have an additional parameter, $L_o$, which strongly influences the allowed range in parameters. Also, to account for the observed high luminosity AGN at high redshifts, the luminosity function must have a high luminosity extension, $L > 1.5 \times 10^{45}$ ergs s$^{-1}$. For many values of the evolution parameter, $Q_\beta$, the high luminosity sources will dominate the contribution to both the sky flux and the excess variance.

Section VII:4:

A direct comparison of the inferred behavior of the source counts from the fluctuations analysis and the Einstein Observatory Medium Sensitivity Survey and Deep Survey source counts is complicated by the different spectral bandpasses. In particular, a large component of the A-2 source counts are sources with significant absorption that make a much reduced contribution to the Einstein surveys' counts. One way of comparing results between two bands is to make specific assumptions about the spectra of the sources, and fold the model spectra through the separate detector spectral responses to derive the observed count rate.

The Medium Survey covers roughly the range of fluxes to which the A-2 measurements of the fluctuations are most sensitive. With no evolution, the XCS luminosity functions predict that clusters should substantially outnumber AGN in the MS sample, contrary to observation. This is persuasive evidence that the AGN X-ray luminosity function evolves over cosmological timescales. When specific models of pure luminosity evolution (for medium luminosity AGN) are examined the MS source counts prefer $2 \leq Q_L \leq 4$. These values are much smaller than indicated by the values derived from a specific examination of the source redshift distribution, although that analysis made different assumptions about AGN spectra and the luminosity range of the current epoch luminosity function. Partial covering fraction models for the absorption in low luminosity AGN can require a further reduction in

acceptable values for $Q_L$.

The change in the fraction of BL Lac sources in going from the XCS to the Medium Survey is not, by itself, significant enough to conclude that BL Lacs do not evolve to the same degree as other AGN. The observed difference would occur at least 18% of the time by chance. This calculation does not use the additional information of the known spectral differences between the two populations, which may increase the significance of the no-BL-Lac-evolution hypothesis.

There is substantial overlap between the flux range covered by the Deep Survey and the lower flux range of the Medium Survey. Allowing non-Euclidean behavior for N(S) greatly increases the range of models consistent with the Deep Survey source counts, although crude measures of the fit indicate that the best models are Euclidean. The results of comparisons with the N(S) models from the XCS are similar to those derived from the Medium Survey.

The fraction of the total all sky flux directly contributed by sources in the Deep Survey depends on their spectra, as well as details of calculation. If the spectrum is similar to the sky flux spectrum, then the total is $\sim 25\% \pm 10\%$. Steeper spectra, typical of low redshift AGN, can reduce this to $\sim 16\% \pm 6\%$. Given the large statistical and systematic uncertainties, as well as the behavior of some AGN evolution models with sharp low-flux rollovers in N(S), current X-ray data alone cannot determine whether the Einstein low flux sources are responsible for the bulk of the X-ray sky flux, particularly in the $\gtrsim 3$ keV band.

The Deep Survey counts can set lower limits on the number of sources responsible for the remainder of the sky flux, the exact value again depending on spectral assumptions. For all reasonable assumptions, these limits are substantially more stringent than those derived from the excess variance upper limit. If the sources are spread over at least a decade below the Deep Survey lower flux limit, at least 850 sources per square degree are required, larger than the estimates from current deep optical quasar counts.

The analysis of the fluctuations in a deep exposure IPC field by Stewart and Fabian may have indicated that the actual source counts break near the Deep Survey flux limit. This conclusion merits further investigation, particularly including the results of additional fields

to improve the statistical sensitivity of the analysis.

Section VII:5:

Low flux optical counts of AGN can be used to predict the behavior of the X-ray N(S) below the Deep Survey. The problem is complicated because the optical and X-ray luminosities may not be strictly proportional. Work by many authors indicates that the ratio of $L_x/L_{opt}$ could be a decreasing function of $L_{opt}$ (for the particular parametric model considered, $L_x \propto (L_{opt})^\delta$, $\delta < 1$) or an increasing function of z. Incorrect spectral assumptions used in making redshift corrections to the observed fluxes could be one explanation for a redshift dependence. The evolution of the AGN spectral index, required if AGN produce the bulk of the X-ray sky flux over the 3-100 keV range, can also generate a redshift dependence of $L_x/L_{opt}$.

A comparison of the power-law indices of the AGN luminosity function in the X-ray and optical bands is, however, <u>incompatible</u> with $\delta < 1$ models. The strong correlations observed by many authors might then be interpreted as requiring a strong explicit redshift dependence for $L_x/L_{opt}$. The net result on pure luminosity evolution models is that $Q_{L;x}$ is expected to be much less than $Q_{L;opt}$, consistent with the observed ranges in both bands. Although the indicated values for $\delta$ and the strength of the redshift dependence are outside the confidence regions obtained by other authors (who do not address the problem posed by the differing luminosity function indices), the wrong parametric form for the $L_{opt}$ and z dependence may have been chosen.

The stronger the explicit redshift dependence, the smaller $Q_{L;x}$ is expected to be for a given value of $Q_{L;opt}$. This has important ramifications for estimates of the total AGN contribution to the sky flux derived from the low flux optical counts.

The presence of a strong redshift dependence of $L_x/L_{opt}$ complicates comparing luminosity function index evolution models derived in the optical band with the X-ray models. An additional problem is that a significant fraction of the AGN contribution to the sky flux must come from sources fainter than the faintest current optical source counts (see Section VII:4).

Section VII:6:

Non-Poisson distribution of sources or general large scale structure may make a significant contribution to the excess variance. This weakens our requirement of evolution based on the fluctuations-derived lower limit for $\sigma^2_S$. Our upper bound on $\sigma^2_S$, however, can be used to place limits on the large scale structure. A proper assessment of the value of these limits must be done more rigorously.

A crude analysis of point sources clumped into superclusters of "small" angular size ($<<3.7^\circ$) leads to the estimate that either the mean number of sources per supercluster is small ($<<5.5$), or alternatively such clumpy sources can only provide a fraction of the total background.

General large scale structure of the X-ray volume emissivity, in the form of spherical lumps in a finite Euclidean universe, are examined. The data places its strongest constraints on the density amplitude of such lumps when the lump size is greater than $\sim 400$ $h_{50}^{-1}$ Mpc. The constraint is tighter if there is strong evolution of the X-ray emissivity.

The effective solid angle is an inappropriate measure of the angular extent of a measurement when distinguishing between point-like and extended sources. The effective measurement size for the data examined in this dissertation has a radius of $3.7^\circ$.

The analysis of the very largest scale anisotropies in the X-ray sky (dipole and quadrupole) and their comparison to the microwave dipole signal is a crude indicator of the matter overdensity responsible for the observed peculiar velocity. The current picture is consistent with a nearby location of the density enhancement, including Virgo infall models.

Chapter VIII:

Evolutionary models for AGN that provide most of the high energy X-ray sky flux (3-100 keV) are consistent with our current understanding, but the evidence is not conclusive. Even if future observations provide proof of this hypothesis, in a sense only the astronomical question of origin will have been answered. The more interesting issue of physical mechanism must then be considered. Given the fundamental differences between current epoch, moderate luminosity

AGN spectra and the X-ray sky flux spectrum, an important change in the AGN emission mechanism is indicated.

The exact form of the sky flux spectrum to be explained depends on the subtraction of a contribution from the known sources, in particular those AGN whose spectra have not evolved. In general, the residual spectrum is consistent with a very flat power law with an exponential high energy cutoff. Emission processes involving non-thick hot thermal plasmas produce spectra that have these characteristics.

Non-thermal inverse Compton models are unable to produce the observed spectrum if the dominant energy loss mechanism for the emitting relativistic electrons is radiation. This conclusion applies to both diffuse emission and compact sources, though the latter may still be allowed if other loss mechanisms are important.

Thin thermal bremsstrahlung from a hot IGM has the proper form for the residual spectrum. The flatness of the residual spectrum at low energies suggests that the bulk of the emission from thin thermal models must be at high redshifts, although the exact limits depend on how the residual spectrum is obtained. However if the hot gas is uniformly spread through the universe, the total energy required to heat the gas is many times that required to produce the X-rays directly. The densities required, $\Omega_{IGM} \sim 0.04$, though significant on a cosmological scale, are so low that the electron-ion coupling times for attaining a Maxwellian thermal distribution is many times the age of the universe, though the electrons may form an essentially decoupled thermal distribution in the relatively short timescale of 0.08 $H_o^{-1}$. The long equipartition time between the proton and electron populations will exacerbate the high energy requirements if the gas heating mechanism is initially most efficient for protons.

Clumping of the gas can reduce both the energy requirements and the coupling timescales. The simple analysis of the limits on clumping placed by the upper bound on the excess variance indicates that the clumps must be $\ll$ 100 Mpc, if the effect is to be significant. Isothermal bound lumps as small as 7 Mpc are consistent with current anlysis. Such lumps provide a factor of 10 reduction on the magnitude of both the energy requirement and coupling timescale problems. Non-gravitationally bound objects can be even smaller. Very small sources

become equivalent to point-sources for models of the sky flux.

The physics of any point-source model should specifically provide for the observed exponential rollover in the sky flux spectrum. The relative sharpness of the observed feature argues persuasively that the location of the rollover should be essentially invariant over the entire range of plausible model behavior.

An important future observation for distinguishing among the competing models and for resolving some of the current ambiguities is the determination of the 2-10 keV spectral index, and possible absorption, for the Einstein Observatory detected AGN, including both the distant high luminosity sources and the sources at the low end of the luminosity function.

If an evolving population of point sources does make up the total sky flux, then an experiment with relatively modest angular resolution (30") and large collecting area/low specific internal background may be expected to directly resolve the entire sky flux.

Excellant data quality and care in the analysis allow us to conclude from a study of the fluctuations that the N(S) source counts continue as a Euclidean form power law for about a decade below what can be studied directly from the resolved souce counts in the 2.5-13 keV band. Studies of the locally derived luminosity function and the excess variance show that in the absence of source evolution there is an additional component of the fluctuations to be explained. However, this study, and all other currently available information on sources of the X-ray background, still have not completed the explanation of the origins and properties of the X-ray sky flux. Enormous strides have been made in the last two decades, particularly with the launches of HEAO 1 and HEAO 2, but a complete, definitive picture has not yet emerged. The next decade promises continued progress, and possibly an answer to the origin question. Even then, the problem of generating the observed amount of flux with the proper spectrum promises that further studies of the X-ray sky will be an important part of high energy

astrophysics.    Studies   of   the   fluctuations   and   other   larger-scale anisotropies will offer an important measure of large scale structure of the universe.

APPENDIX A1

COSMOLOGICAL CORRECTIONS TO THE DERIVATION OF N(S) FROM A LUMINOSITY
FUNCTION

We address three technical difficulties in this appendix: the effects
of non-ideal detector spectral response and the redshifting of source
spectra; the non-Euclidean aspects of the Friedmann cosmologies; and
possible difficulties of evolving luminosity functions, particularly
when luminosity dependent variation in the sources' spectra might be
involved. Applications and examples of this analysis are presented in
the main body of the text (see Section VII:2).

The material presented here is based in large part on Weinberg's
[1972] Chapter 14 on Cosmography. The principle deviation here is in
the treatment of non-zero bandwidth detectors and source spectra not
restricted to simple single power laws. The metric of the universe is
presumed to be the Robertson-Walker metric: $ds^2 = -dt^2 + a^2(t)(d^3\sigma)^2$.
The spatial coefficient, $a(t)$, is the cosmic scale parameter. $a(t_0)$ is
its current value. The Hubble "constant" is given by $H_0 = \dot{a}(t_0)/a(t_0)$.
The dot in $\dot{a}$ signifies differentiation by t. $d^3\sigma$ represents the spatial
part of the metric, given by

$$(d^3\sigma)^2 = \frac{d^2r}{1 - kr^2} + (r^2 \, d^2\theta + r^2 \, \sin^2\theta \, d^2\phi) \ . \qquad [A1.1]$$

r is a dimensionless coordinate whose scale is set by fixing the value
of k to +1, 0, or -1. Solving the Einstein equations, (with no
extraneous terms like a cosmological constant) for a homogenous perfect
fluid with negligible pressure yields the Friedmann solutions. This
solution for $a(t)$ allows us to relate the position of an object in terms
of r, to observables such as the redshift $z = [a(t_0)/a(t)] -1$. If the

observer is at r=0, then the emitter's position is

$$r_1 = \frac{zq_0 + (q_0-1)(-1 + (2q_0z+1)^{1/2})}{h_0 \, a_0 \, q_0^2 \, (1+z)}$$    [Al.2]

independent of the value of k. $q_0$ is the dimensionless deceleration parameter related to the solution for a(t) by

$$q_0 = - \ddot{a}(t_0) \, a(t_0) \, / \, \dot{a}^2(t_0).$$    [Al.3]

We have examined the cosmological impact on observations for both low and high values of $q_0$, 0.1 and 1. Observations of $q_0$ are quite imprecise and both values are consistent with our current understanding.

## Determining observed S

We consider a class of objects at a redshift z with a rest frame luminosity spectrum $\ell(E;x)dE$. The variable x represents some general intrinsic parameter of the emitter used to parameterize the class distribution function $\phi(x;\vec{r}) \, dx$, the density of objects at position $\vec{r}$ with values of x from x to x+dx. A common example for x is the luminosity, so that $\phi(\ell(E))$ is the luminosity function. If the spectra of all the objects have the same shape, then luminosity defined over a particular rest-frame energy range $\Delta E$, $L(x) = \int_{\Delta E} \ell(E;x)dE$, is equivalent to the differential luminosity, $\ell$. This is a moderately good assumption in the case of H1L1 band emission from clusters and to a lesser degree from active galactic nuclei. We can replace $\ell(E;x)$ with $L \, \eta(E)$, where $\eta(E)$ has been normalized so that

$$\int_{\Delta E} \eta(E) \, dE = 1.$$    [Al.4]

We are interested in what the observed flux will be after redshifting and geometrical focusing:

$$s(E) \, dE = \frac{\ell((1+z)E)}{4\pi \, r_\ell^2} \, (1+z) \, dE \, .$$    [Al.5]

The factors of (1+z) that precede E and dE are due to the spectral

redshift while $r_\ell$, the luminosity distance

$$r_\ell = a_o (1+z) r_1 ,$$                              [Al.6]

incorporates the geometrical modifications. This definition and the solution for $r_1$ in [Al.2] eliminate any dependence on the unobservable $a_o$:

$$r_\ell = 5.99 \, h_{50}^{-1} F(z;q_o) \text{ Gpc},$$      [Al.7]

where $F(z;q_o)$ is set by the solution of $r_1$ to be

$$F(z;q_o) = z \left\{ 1 + \frac{z (1-q_o)}{(1+2q_o z)^{1/2} + 1 + q_o z} \right\} .$$    [Al.8]

The appearance of $h_{50}^{-1}$ in [Al.7] shows that the given value of an object's luminosity, or $\ell(E)$, as inferred from its observed $s(E)$, scales as $H_o^2$.

For low redshifts, $F(z;q_o)$ is well approximated by $z$ so that [Al.5] simplifies to the Euclidean form (see equation [III.1]) with $s(E)$ inversely proportional to $z^2$. Even at high redshifts, $F(z;1) = z$ when $q_o = 1$ so that the inverse square relationship is maintained for geometric focusing. The additional effect of spectral redshifting must still be included. For values of $q_o > 1$, F and therefore $r_\ell$ reaches a maximum at some redshift. This means that a given luminosity object would have a minimum observed flux, again neglecting the effects of spectral redshifting.

So far we have not folded in the physical detector response to move from the incident energy flux to the observed count rate in the detector. To find the count rate in a given pulse height window we use the quantum efficiency, $Q(E)$, defined in [II.23] (see Section II:3) for that window.

$$S = \int Q(E) E^{-1} s(E) dE.$$                        [Al.9]

The factor of $E^{-1}$ converts the energy flux of $s(E)$ into photons as $Q(E)$ is defined by the probability that a photon at energy E is detected in

the particular pulse height range. With our expression for s(E) dE (equation [A1.5]) we can now relate observed count rate with cosmological and distributional quantities: $S(x,z;q_o)$, or if L is used $S(L,z;q_o)$.

### Cosmological corrections to Euclidean volume

We again start with an emitter at redshift z, which has the r-coordinate position $r_1$. The differential volume of all such emitters in a shell surrounding the observer is

$$dV = 4\pi \, a_o^3 \, (1+z)^{-3} \, [1-kr_1^2]^{-1/2} \, r_1^2 \, dr_1. \qquad [A1.10]$$

Using the Friedmann solution eliminates any explicit k and $a_o$ dependence,

$$dV = 4\pi \, H_o^{-3} \, (1+z)^{-6} \, (1 + 2q_o z)^{-1/2} \, F^2(z;q_o) \, dz. \qquad [A1.11]$$

Again at low z the volume element reduces to a Euclidean $z^2$ dz. But the smaller physical size of the universe at the early epochs begins to constrain the available volume as seen in the $(1+z)^{-6}$ term.

Because of the expansion of the universe, the density of objects should increase when the size of the universe is smaller, i.e. the density is proportional to $(1+z)^3$ even if a class of objects underwent no other changes. To better measure real changes in the number of sources we define the comoving density function $\phi_c$:

$$\phi_c(x;z) = (1+z)^3 \, \phi(x;z). \qquad [A1.12]$$

Any change in $\phi_c$ indicates a physical change in the distribution of the objects, i.e. some evolutionary change in number or luminosity. The number of objects observed in a particular redshift range is

$$\phi(x;z) \, (dV/dz) \, dz = \phi_c(x;z) \, (dV_c/dz) \, dz \qquad [A1.13]$$

where $dV_c$ is the comoving volume element given by

$$dV_c = (1+z)^3 \, dV$$

$$= 4\pi \, H_o^{-3} \, (1+z)^{-3} \, (1+2q_o z)^{-1/2} \, F^2(z;q_o) \, dz. \qquad [Al.14]$$

Again, $H_o$ sets the scale size of the universe. A convenient characteristic volume is the Hubble volume,

$$V_H = (4/3)\pi \, H_o^{-3} = 900 \, h_{50}^{-3} \, Gpc^3. \qquad [Al.15]$$

The true value of $h_{50}$ is unimportant when calculating $N(S)$ or $S_{as}$ as long as the value used in $dV_c$ is the same as when calculating $\phi$. This is because the product $\phi \times dV/dz$ is independent of the exact value of $H_o$.

<div align="right">Calculating the number counts</div>

Given the relationship $S(x,z)$, the integral number of sources with fluxes above S is

$$N(>S') = \int dz \, (dV_c/dz) \int dx \, \phi_c(x;z) \, \theta(S(x,z)-S') \, . \qquad [Al.16]$$

For intervals of x where $\ell(x)$, the shape of the emitted energy spectrum, is constant, x may be replaced with the luminosity L. In addition, for a particular redshift S is proportional to L so that we can rewrite the Heaviside function as $\theta(L(S,z)-L(S',z))$. For simple forms of $\phi_c$ the inner integral can be done analytically. Only the dz integral need be done numerically. The differential $N(S)$ can be found by a limiting difference process:

$$N(S) = \lim_{\delta S \to 0} \frac{N(>S+\delta S)-N(>S)}{\delta S} \, . \qquad [Al.17]$$

Again for cases of constant spectral shape the inner integral of [Al.16] reduces to an integral from $L(S,z)$ to $L(S+\delta S,z)$. Other interesting quantities that can be calculated from variants of [Al.16] are $\bar{z}(S)$, the mean redshift of all objects with observed flux S, and $\bar{L}(S)$, the mean luminosity of such objects. An examination of such quantities illustrates how different flux-selected samples are dominated by different regions of redshift or luminosity space.

<div align="right">Cosmological N(S)</div>

A first order correction to the Euclidean model that is valid for low redshifts can be calculated when the source spectra may be represented by a simple power law with energy index $\alpha$ and the luminosity function is a power law of index $\beta$ (see Piccinotti et al. [1972] equation 8 taken in turn from Weinberg [1972] equation 14.7.34). We ignore the effects of detector spectral response so that S and L are the energy flux and luminosity defined over the same bandpass $\Delta E$. In this simple case the source counts are proportional to

$$N(S) \propto S^{-5/2} \times$$

$$\{1 - 3.7 \times 10^{-27} \, h_{50} \, (1+\alpha) \, \frac{(2.5-\beta) \, \Delta(L^{3-\beta})}{(3-\beta) \, \Delta(L^{2.5-\beta})}\} \qquad [A1.18]$$

where

$$\Delta(L^{-\alpha}) \equiv L_{max}^{-\alpha} - L_{min}^{-\alpha} \qquad [A1.19]$$

and $L_{max}$ and $L_{min}$ are the upper and lower cutoffs of the luminosity function. From equation [A1.18] we see that at the highest flux values the source counts initially follow the Euclidean 5/2 index power law, but at low enough fluxes there will be a flattening in the curve as the correction term begins to be significant. The appearance of the source spectra through $(1+\alpha)$ indicates that this correction involves both the redshifting and geometric effects. Note, however, there is no explicit dependence on $q_o$ in the first order correction. The approximations used in deriving [A1.18] are accurate over values of $S \gg H_o^2 \, L \, / \, 4\pi$. Equation [A1.18] also illustrates a general property of the N(S) curves for all non-evolving populations with suitable spectra (a noninverted photon spectrum, $\alpha > -1$): the cosmological corrections flatten N(S) below the Euclidean 5/2 power law. Source counts steeper than Euclidean provide strong evidence for evolution of $\phi_c$.

Particular models for the evolution of the luminosity function are discussed in Section VII:3. Current models are usually parameterized in terms of redshift or emission epoch. The explicit redshift evolution forms, usually involving powers of (1+z), have no dependence on $q_o$ but often require an explicit high z cutoff to avoid infinite sky flux (Olber's paradox). The models which are explicitly dependent on epoch,

rather than redshift, usually use $\tau$, the fractional look-back time

$$\tau = 1 - (t_{em}/t_o) , \qquad\qquad\qquad [A1.20]$$

where $t_{em}$ is the epoch of emission and $t_o$ is the current epoch of detection. The value of $\tau$ as a function of redshift can be calculated by integrating the differential relationship between redshift and epoch for a Friedmann solution:

$$dt = - H_o^{-1} \frac{dz}{(1 + z)^2 (1 + 2q_o z)^{1/2}}. \qquad\qquad [A1.21]$$

Solutions exist for $t(z)$ with the integral performed. However, they have forms dependent on the value of k. A simple case is where k = 0, equivalent to $q_o = 1/2$:

$$t_{LB} = t_o \tau = (2/3) H_o^{-1} (1 - (1+z)^{-3/2}). \qquad\qquad [A1.22]$$

APPENDIX A2

DISTRIBUTIONS, PROBABILITIES, AND TRANSFORMS

This appendix is a general description of distributions, with an application to the evaluation of the effects of overlapping data. The Mellin transform is also presented as a possible tool for future analysis of complicated fluctuations models.

————————————————

1: Characteristic Function and Cumulants of a Distribution

We outline several ways the distribution of a random variable may be described. The characteristic function of a distribution provides a simple way to describe the addition of independent random variables through the convolution theorem. The characteristic function can also be used to define the cumulants of a distribution. Expressions for the fifth and sixth cumulants are presented.

This section is a ready reference for statistical concepts mentioned in the main body of the thesis. A description of cumulants can be found in almost any intermediate text book on distributions, probability, or transforms (see for example Wilks [1963]). In addition this section presents formulae for the 5th and 6th cumulants, which are not found in most accessible references, as well as the general algorithm for the derivation of any higher cumulants needed. Other formulae not derived here are collected in Abramowitz and Stegun [1965].

Distributions, moments, and the characteristic function

The behavior of a continuous random variable is often described in terms of a differential probability density, the distribution function $p_x(x)$, which is the probability that x is in the interval x to x+dx. $p_x(x)$ is normalized so that $\int p_x(x) \, dx = 1$. Given any function of x,

f(x), its expectation value, $\langle f(x)\rangle$, is just

$$\langle f(x)\rangle = \int_{-\infty}^{+\infty} ds\ p_x(s)\ f(x)\ . \qquad [A2.1]$$

(All integrals henceforth are presumed to be definite integrals over the entire range of the variable of integration.) Several sets of expectation values are used to characterize $p_x$. The nth moment about the origin is:

$$\mu'_n \equiv \langle x^n\rangle\ . \qquad [A2.2]$$

$\mu'_1$ is the mean of the distribution. The central moments are the moments about the mean, $\mu_n = \langle(x-\mu'_1)^n\rangle$. The second and third such moments are the variance, $\sigma^2$, and the skewness of the distribution, respectively. The mean, variance, and skewness of a distribution are usually a sufficient description of a distribution. Of course they cannot tell the whole story, particularly when the distribution has an unusual shape. In some cases, as with the fluctuations, the values may be formally undefined because the integral in [A2.1] is infinite.

All the information of the distribution is contained in another expectation value, or rather a function produced from expectation values, the characteristic function:

$$\phi_x(t) \equiv \langle\exp(itx)\rangle = \int dx\ p_x\ e^{itx}\ . \qquad [A2.3]$$

This is closely related to the Fourier transform of $p_x(x)$,

$$\phi_x(t) = F(p_x,x;\ -t)\ . \qquad [A2.4]$$

The minus sign is due to the convention where the forward transform is

$$F(f(x),x;t) \equiv \int dx\ f(x)\ e^{-ixt}\ . \qquad [A2.5]$$

and the inverse transform given by

$$F^{-1}(g(t),t;x) = (2\pi)^{-1} \int dt\ g(t)\ e^{+ixt}\ . \qquad [A2.6]$$

Thus in a similar way we can invert the characteristic function and re-

derive the probability density:

$$p_x(x) = \mathbf{F}^{-1}(\phi(t),-t;x) \; .$$                         [A2.7]

So in this way we see that $\phi_x(t)$ contains all the information of $p_x(x)$. Both characterize the random variable, x.

To make this more explicit we consider $\phi_x^{(n)}(t)$, the nth derivative of $\phi_x(t)$:

$$\phi_x^{(n)}(t) = \frac{d^n \phi_x(t)}{dt^n} = \int dx \; p_x(x) \; (ix)^n \; e^{itn} \; .$$   [A2.8]

Therefore

$$\phi_x^{(n)}(t=0) = i^n \langle x^n \rangle \; ,$$                    [A2.9]

relating the characteristic function to the moments of the random variable.

The utility of characteristic functions is particularly easy to appreciate when we form combinations of random variables. This is due to the convolution property of Fourier transforms. We define the convolution $f*g \; (x)$ of two functions $f(x)$, $g(x)$ as

$$f*g \; (x) = \int dx\{ \; f(x') \; g(x-x') \; \} \; .$$               [A2.10]

The convolution theorem states that

$$\mathbf{F}(f*g \; (x),x;t) = \mathbf{F}(f,x;t) \times \mathbf{F}(g,x;t) \; .$$   [A2.11]

To demonstrate this we use characteristic functions, for the example of two independent random variables, x and y. The independentness means that the joint probability distribution is separable, i.e. $p_{x,y}(x,y) = p_x(x) \; p_y(y)$. The expectation of a function $g(x,y)$ is

$$\langle g(x,y) \rangle = \int dy \; p_y \int dx \; p_x \; g(x,y) \; .$$   [A2.12]

The distribution of a new random variable $z \equiv x+y$ is given by the

convolution of the distribution of x and y:

$$p_z(z) = \int dx\, p_x(x)\, p_y(z-x) \, .$$                            [A2.13]

The characteristic function of z, $\phi_z(t)$ is easily calculated, given that x and y are independent:

$$\begin{aligned} \phi_z(t) &= \langle \exp(izt) \rangle = \langle \exp(i(x+y)t) \rangle \\ &= \langle \exp(ixt) \rangle \langle \exp(iyt) \rangle \\ &= \phi_x(t)\, \phi_y(t) \, . \end{aligned}$$                            [A2.14]

By the relationship between characteristic function and Fourier transform we see that this is an equivalent form of the convolution theorem. In general given a sum of independent variables, $x_i$, we can describe the characteristic function of the sum by

$$z = \sum_i x_i$$                                                      [A2.15a]

$$\phi_z(t) = \prod_i \phi_{x_i}(t) \, .$$                              [A2.15b]

### Cumulants and reduced cumulants

The natural logarithm of $\phi(t)$ may be expanded in a Taylor series in (it) about t=0:

$$\ln \phi_x(t) = \sum_{n=0}^{\infty} \kappa_{n;x} \frac{(it)^n}{n!} \, .$$                            [A2.16]

The nth coefficient of the expansion, $\kappa_{n;x}$, is the nth cumulant of x. If we form a sum of random variables, $z = \sum x_i$, then since $\phi_z(t)$ was formed by the product of the characteristic functions, the cumulants of z are just sums of the cumulants of the $x_i$. The cumulants, like the characteristic function, contain all the information about the distribution of the random variable, as long as the cumulants are well defined (i.e. when the Taylor series expansion is valid).

Table A2.1: First six cumulants in terms of moments about zero

$\kappa_1 = \langle x \rangle$ $\qquad\qquad$ = mean

$\kappa_2 = \langle x^2 \rangle - \langle x \rangle^2$ $\qquad\qquad$ $= \sigma^2$ = variance

$\kappa_3 = \langle x^3 \rangle - 3\langle x \rangle\langle x^2 \rangle + 2\langle x \rangle^3$ $\quad = \mu_3$

$\kappa_4 = \langle x^4 \rangle - 4\langle x \rangle\langle x^3 \rangle - 3\langle x^2 \rangle^2 + 12\langle x \rangle^2\langle x^2 \rangle - 6\langle x \rangle^4$

$\qquad\qquad\qquad\qquad\qquad = \mu_4 - 3\sigma^2$

$\kappa_5 = \langle x^5 \rangle - 5\langle x \rangle\langle x^4 \rangle - 10\langle x^2 \rangle\langle x^3 \rangle + 20\langle x \rangle^2\langle x^3 \rangle + 30\langle x \rangle\langle x^2 \rangle^2$

$\qquad\qquad\qquad\qquad - 60\langle x \rangle^3\langle x^2 \rangle + 24\langle x \rangle^5$

$\kappa_6 = \langle x^6 \rangle - 6\langle x \rangle\langle x^5 \rangle - 15\langle x^2 \rangle\langle x^4 \rangle - 10\langle x^3 \rangle^2 + 30\langle x \rangle^2\langle x^4 \rangle$

$\qquad\qquad\qquad + 120\langle x \rangle\langle x^2 \rangle\langle x^3 \rangle + 30\langle x^2 \rangle^3$

$\qquad\qquad\qquad - 120\langle x \rangle^3\langle x^3 \rangle - 270\langle x \rangle^2\langle x^2 \rangle^2$

$\qquad\qquad\qquad + 360\langle x \rangle^4\langle x^2 \rangle - 120\langle x \rangle^6$

If we define $\tau \equiv it$, then the nth cumulant is

$$\kappa_n = \frac{\partial^n \ln \phi(i^{-1}\tau)}{\partial \tau^n} \Bigg|_{\tau=0} , \qquad\qquad\qquad [A2.17]$$

where for the time being we drop the subscript ";x". Recalling [A2.9] and knowing that $\phi(t=0) = 1$, we can generate relationships between the cumulants and the moments about zero. Inspection of Table A2.1 shows that the second and third cumulants are identical to the second and third moments about the mean. Thus the well known additivity of variances is a result of the equivalence of the variance and the second cumulant. This property does not extend higher than the third moment about the mean.

The cumulants of the Gaussian distribution, other than the mean and variance, are all zero. The cumulants of the Poisson distribution are all equal.

The nth cumulant is a linear combination of the first n moments. This may be inverted, as we show in Table A2.2, giving the nth moment about the origin in terms of the first n cumulants. This closely

---

Table A2.2: Moments about the origin of a random variable

in terms of its cumulants.

---

$\langle x \rangle = \kappa_1$

$\langle x^2 \rangle = \kappa_2 + \kappa_1^2$

$\langle x^3 \rangle = \kappa_3 + 3\kappa_1\kappa_2 + \kappa_1^3$

$\langle x^4 \rangle = \kappa_4 + 4\kappa_1\kappa_3 + 6\kappa_1^2\kappa_2 + 3\kappa_2^2 + \kappa_1^4$

---

coupled relationship between moments and cumulants allows the calculation of the cumulants for new distributions defined as functions of variables with known cumulants. For example, consider $y \equiv x^2$. The first cumulant of $y$ is $\kappa_{1;y} = \langle y \rangle = \langle x^2 \rangle = \kappa_{2;x} + \kappa_{1;x}^2$. The variance of $y$ is $\kappa_{2;y} = \langle x^4 \rangle - \langle x^2 \rangle^2 = \kappa_{4;x} + 4\kappa_{1;x}\kappa_{3;x} + 4\kappa_{1;x}^2\kappa_{2;x} + 2\kappa_{2;x}^2$. Linear combinations of variables can be constructed. As an example, consider $\chi^2$ which is the sum of $\nu$ independent variables $X_i^2$, each of zero mean and unit variance ($\nu$ is the number of degrees of freedom). Therefore $\kappa_{1;X} = 0$ and $\kappa_{2;X} = 1$, so that we can show that $\kappa_{1;\chi^2} = \nu [ \kappa_{2;X} + \kappa_{1;X}^2 ] = \nu$. All higher cumulants of X are zero, therefore the variance of the $\chi^2$ distribution is $2 \kappa_{2;X}^2 = 2\nu$. This could be found directly from the series expansion of the characteristic function for $\chi^2$, but using the moments to derive the cumulants is in some circumstances computationally easier.

In the body of the text we introduced the dimensionless reduced cumulants (equation [III.21]) which are quantitative measurements of the distribution's deviation from a pure Gaussian. The differential distribution, the integral probability, and the inverse of the integral probability for any distribution can be approximated with a power series expansion in the reduced cumulants. The form of this expansion is given by Abramowitz and Stegun [1965] in their equations 26.2.47–50.

---

2: Overlapping Data and Corrections to Data Weighting

The effective number of measurements for overlapping data is investigated using the example, of non-independent measurements drawn from a Gaussian distribution.

In Section IV:2 we outlined a correction for $\chi^2$ determinations that used overlapping data. This correction weighted the individual measurements, using equation [IV.4]. If the ith observation has a weight $W_i$ equal to 0.7, for example, then the observation is counted as equivalent to 0.7 of an independent observation. The contribution of this observation to $\chi^2$ is weighted by this factor, and by analogy so is its contribution to the likelihood estimators and moment statistics of the fluctuations fits. The form for the weighting factor we have suggested is

$$W_i = 1 \;/\; (1 + \sum_{j \neq i} w_{ij}) \qquad\qquad\qquad [A2.18]$$

where $w_{ij}$ is the contribution to the ith measurement's weight due to correlation with the jth measurement. This depends on the size of the covariance coefficient, $\sigma_{ij}$, relative to the size of the jth observation's variance:

$$w_{ij} = \sigma_{ij} \;/\; \sigma^2_j \;. \qquad\qquad\qquad\qquad [A2.19]$$

Note that for the discussion in this appendix, we assume that $\sigma^2_j$ is the total variance, which includes a contribution from fluctuations, $\sigma^2_{j;f}$, which may be correlated with other overlapping measurements, and a contribution from photon counting statistics, $\sigma^2_{j;p}$. The latter variation is statistically independent of the other measurements.

In this section, we present plausibility arguments that such a weighting is appropriate. For a few simple cases we can show that this weighting has exactly the correct properties. Slightly more complicated examples have been investigated, and the appropriateness of the weighting confirmed, using Monte Carlo techniques. However, we have not yet performed a full examination of the most complicated situation: the analysis of N(S) models from the fluctuations in a manner similar to

that used in this dissertation. This weighting is an improvement over previous attempts to use overlapping data when analyzing the fluctuations, but some additional examination may be useful.

Consider a series of observations $y_i$, not necessarily independent, each with a variance $\sigma^2_i$. All the observations have the same (unknown) mean, $\overline{y}$. N is the total number of observations. The observations can be used to estimate $\overline{y}$ by minimizing

$$E(y) = \sum_i \frac{(y_i - y)^2}{\sigma^2_i} . \qquad\qquad [A2.20]$$

We note that when the $y_i$ are independent, $E(\overline{y})$ is distributed as $\chi^2$ with N degrees of freedom. Let $y_o$ be the estimator of $\overline{y}$ obtained by minizing $E(y)$. Then we can show that

$$y_o = A \sum (y_i/\sigma^2_i) , \qquad\qquad [A2.21]$$

where

$$A \equiv 1 / \sum (1/\sigma^2_i) . \qquad\qquad [A2.22]$$

The expectation value for the estimator, $y_o$, is

$$\langle y_o \rangle = A \sum (\langle y_i \rangle/\sigma^2_i) = \overline{y} \qquad\qquad [A2.23]$$

since all the observations have the same mean, $\langle y_i \rangle = \overline{y}$. Therefore, $y_o$ is an unbiased estimator of $\overline{y}$. The efficiency of $y_o$ as an estimator is measured by the variance of $y_o$, $\sigma^2_{y_o} = \langle y_o^2 \rangle - \langle y_o \rangle^2$. By squaring both sides of equation [A2.21], and finding the expectation value, we can derive

$$\sigma^2_{y_o} = A + A^2 \sum_i \{ \frac{1}{\sigma^2_i} \sum_{j \neq i} \frac{\sigma_{ij}}{\sigma^2_j} \} . \qquad\qquad [A2.24]$$

The first term, A, gives the variance when there is no correlation between the observations. The second term was derived using the relationship $\langle y_i y_j \rangle = \sigma_{ij} + \overline{y}^2$.

If we consider first the case where there is no correlation, i.e. $\sigma_{ij} = 0$ for all i,j, and where the observations all have identical

variances, $\sigma^2$, then the uncertainty in the $y_o$ estimator for $\overline{y}$ reduces to

$$\sigma^2{}_{y_o} = A = \sigma^2 / N .$$                                      [A2.25]

Thus increasing the number of observations decreases the variation in $y_o$. As we discussed in Section IV:2, additional observations with total correlation to one of the original set of independent measurements ($\sigma_{ij} = \sigma^2$) cannot add any new information. The function $E(\overline{y})$ is no longer distributed as $\chi^2$. The addition of such observations in equation [A2.21] will not decrease the uncertainty in $y_o$ through the increase in the second term of equation [A2.24]. If the total set of N observations consists of n groups of M totally correlated measurements then

$$\sigma^2{}_{y_o} = A^2 \sum M/\sigma^2{}_i = \sigma^2 / n .$$              [A2.26]

If the groups did not each have the same number of observations then $\sigma^2{}_{y_o}$ would actually be increased, so that the utility of equation [A2.21] as an estimator of $\overline{y}$ would degraded by the additional observations. The correct solution is to weight the observations in each group so that they will contribute the equivalent of a single independent observation in total. Thus, if an observation is a member of a group with M totally correlated measurements then its contribution to the sum in equation [A2.21] should be weighted by 1/M. A similar weighting of the terms of the sum for $E(y)$ in equation [A2.20] will restore it to a $\chi^2$ distribution with a number of degrees of freedom determined by the number of groups.

When the correlation is less than total the correct weighing is not intuitively obvious. Consider a generalization of the least squares measure in equation [A2.20]

$$E(y) = \sum_i W_i \frac{(y_i - y)^2}{\sigma^2{}_i} .$$                       [A2.27]

Minimization of this weighted sum is

$$y_o = A' \sum (W_i y_i / \sigma^2{}_i)$$                                    [A2.28]

with

$$A' \equiv 1 \; / \; \sum_i (W_i/\sigma^2_i) \; . \tag{A2.29}$$

The estimator $y_o$ is still unbiased, _i.e._ $\langle y_o \rangle = \bar{y}$. The variance in the estimator is

$$\sigma^2_{y_o} = (A')^2 \; \sum_i \frac{W_i^2}{\sigma^2_i} \; B_i \tag{A2.30}$$

where the effects of measurement correlation are expressed in the term

$$B_i = 1 + \sum_{j \neq i} \frac{\sigma_{ij}}{\sigma^2_j} \frac{W_j}{W_i} \; . \tag{A2.31}$$

So far in this discussion we have not determined what values are to be used for the $W_i$ weighting parameters. The expression in equation [A2.30] is correct no matter what particular weights are chosen. (Note that it reduces to equation [A2.24] when the $W_i$ are fixed equal to one.)

The optimal values for $W_i$ are clearly those that will minimize the variance in the estimator $y_o$. These values can be found from partial derivatives of equation [A2.30]. It can be shown that such optimal weights must satisfy the relationship

$$W_i = (B_i)^{-1} = 1 \; / \; \{ 1 + \sum_{j \neq i} \frac{\sigma_{ij}}{\sigma^2_j} \frac{W_j}{W_i} \} \; . \tag{A2.32}$$

When the $W_i$ satisfy this condition, the weighted least squares sum in equation [A2.27] is distributed (to first order) as $\chi^2$ with $N_w$ degrees of freedom where $N_w$ is the weighted number of observations

$$N_w = \sum W_i \; . \tag{A2.33}$$

Unfortunately, equation [A2.32] does not directly determine the correct value for $W_i$ as it contains $W_i$ on both sides of the equation. The exact form requires an iterative solution of equation [A2.32]. An approximate solution is the weighting of equation [A2.18], where we have assumed that whenever two observations are correlated then $W_i \approx W_j$. This is the weighting that we used in the analysis for this dissertation. It is

exactly true for the example of the groups of totally correlated observations.

To calculate even the approximate weight for the overlapping data requires an estimate of the correlation coefficient, $\sigma_{ij}$. Here is where we must approximate the fluctuations as a Gaussian with a size $\sigma^2_{i;f}$ for the ith observation. This should not be identified with the true variance of the $P_I$ distribution, but instead with some crude measure of the size of the quasi-Gaussian aspect of the fluctuations away from the high intensity tail of the distribution. If the fluctuations are Gaussian, and if $r_i(\theta,\phi)$ and $r_j(\theta,\phi)$ are the response functions of the ith and jth measurment, then

$$\frac{\sigma_{ij}}{\sigma^2_{i;f}} = \frac{\int d\Omega \; r_i \; r_j}{\int d\Omega \; r_i^2} \; . \qquad\qquad [A2.34]$$

-------------------------

### 3: The Mellin Transform and Calculation of n(I)

An algorithm for folding a general N(S) function through the detector response is presented using Mellin transforms.

Equation [III.33] gives the general integral form for using the detector response frequency, $\nu(r)$, to transform the detector independent function N(S), to the detector dependent mean number of sources observed in the field of view of a single measurement, n(I). The process of taking two functions, $\nu$ and N, to create a third, n, is reminiscent of the convolution process (equation [A2.10]), although now the combination is a multiplication, I = rS, rather than a sum, z = x+y. The standard convolution could be quickly performed in terms of Fourier components, so that the convolution integral reduces to a calculation of forward and inverse fast Fourier transforms. The "multiplicative" convolution also has a transform form, using Mellin transforms. Mellin transforms and their properties are presented in more detail in many books on general

mathematical transforms [e.g. Bracewell, 1978]. This section will touch on some of these properties, usually without proof, and present an outline of an algorithm to convert Mellin transforms to a Fourier form, so that the fast Fourier transform computational tools can be used for efficient calculation. In analyzing the fluctuations using the A-2 data, we had no need to use these sophisticated techniques, however in the future when $N(S)$ is investigated at lower fluxes, the more complicated $N(S)$ models that will be required may benefit from the techniques we present here.

The Mellin transform of a real function $f(x)$ is the integral operator

$$M(f(x),x;s) = \int_0^\infty dx\ x^{s-1}\ f(x)\ ,\qquad\qquad [A2.35]$$

where $s$ is the new independent complex variable. The integral may not converge, depending on the behavior of $f(x)$, for certain values of $Re(s)$, the real part of $s$. Note the range of the integral, from 0 to infinity, means that the transform is sensitive only to the positive behavior of $f(x)$. It can be shown [Bracewell] that the inverse transform is

$$M^{-1}(g(s),s;x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\omega\ g(r+i\omega)\ x^{-(r+i\omega)},\qquad\qquad [A2.36]$$

where the complex variable $s$ has been broken into real and imaginary parts

$$s = r + i\omega\ .\qquad\qquad [A2.37]$$

The real component, $r$, is held constant, and is chosen so that the path integral of the inverse transform, $-\infty < \omega < \infty$, is over a domain where $g(r+i\omega)$ is always defined. If this is true then the inverse transform will retrieve the original function:

$$f(x') = M^{-1}(\ M(f(x),x;s),\ s;\ x')\ .\qquad\qquad [A2.38]$$

Consider three functions, $f$, $g$, and $h$, related by the

multiplicative convolution

$$h(x) = \int_0^\infty du\ f(x/u)\ g(u)/u\ .$$                               [A2.39]

This is the same form as equation [III.33], with $h(x) = n(I)$, $g(u) =$
$\nu(r)/4\pi$, and $f(x/u) = N(S)$. Performing Mellin transforms on all three
functions, it can be shown that a property of the Mellin transform is

$$M(h(x),x;s) = M(f(x),x;s)\ ^\times\ M(g(x),x;s+1)\ .$$                    [A2.40]

Thus, the Mellin transform plays a similar role for multiplicative
convolutions that the Fourier transform plays for the more ordinary
additive convolution. Some care is needed in selecting the constant r
for the real part of s in the inverse transform path integral. It must
have the property that $M(f;r+i\omega)$ and $M(g;(r+1)+i\omega)$ are defined over the
entire range $-\infty < \omega < +\infty$.

To have any computational advantage, we must have a quick and
efficient way to calculate the Mellin transform and its inverse. We
accomplish this by rewriting the Mellin transform as a Fourier
transform:

$$M(f(x),x;r+i\omega) = \mathbb{F}(\ f(e^{-t})e^{-rt}\ ,\ t;\ \omega)\ ,$$     [A2.41]

and

$$M^{-1}(g(r+i\omega),\ r+i\omega;\ x) = x^{-r}\ \mathbb{F}^{-1}(\ g(r+i\omega),\ \omega;\ -\ln x)\ .$$   [A2.42]

The Fourier transforms may be evaluated with fast Fourier algorithms
(e.g. Bergland [1969]). In preforming the FFTs, note that there is
significant signal for both the positive and negative values of the
Fourier variable $t = -\ln x$. Because of this change of variables, the
$N(S)$, $\nu(r)$, and resultant $n(I)$ functions are not sampled at equally
spaced intervals, but rather at points that are more closely spaced at
the low values. Given that the strongest variations in the functions is
expected at these low values, this assists an accurate calculation of
$n(I)$. However, interpolation is needed to convert the unequally spaced
values of $n(I)$ to an equally spaced estimate to continue with the
calculation of the fluctuations distribution (see Appendix A3).

APPENDIX A3

COMPUTATIONAL ISSUES IN CALCULATING AND FITTING FLUCTUATIONS MODELS

We outline the general techniques used in the actual
construction of the $P_{\xi'}(\xi')$ curve and the subsequent fit to
data. The finite discrete Fourier transforms used have
important properties that are allowed for. A proposal is
made to extend the application of these techniques to cases
where the detector spatial response includes significant
regions with negative response.

The equations which determine the calculation of a model
fluctuations distribution and how it should be tested against
observations have been detailed in the body of the dissertation (see
equations [III.43], [III.46] et seq., [III.61] et seq., [V.2], and
[V.4], and associated discussions). The implementation of these
equations in a computer program raised many issues. The problem of
determining what constitutes acceptable accuracy led eventually to the
concept of the likelihood difference of two models (Section II:5). The
expected quality of the A-2 data and the number of observations in the
WW dataset required that care be used in turning the general formulae
into an algorithm. The eventual result was a body of subroutines, each
covering a single phase of the process that turned data plus model
parameters into a set of statistics for determining the acceptability of
the fit and the parameter confidence region. The subroutines were
written in FORTRAN IV (using the IBM H Extended Enhanced compiler
dialect). Rather than present a listing of the programs, this appendix
consists of brief notes on the general issues of technique. The actual
source code is available on request. Data with lower signal to noise or
fewer observations than the A-2 data may not require the care used in
these subroutines, and might instead be analyzed with more efficiency by
a simpler implementation.

Calculating $n(\xi)$

The basic model tested is a power law $n(\xi)\,d\xi$, with a roll over

Calculating $P(D)$

beginning at $\xi_c$ (corresponding to a sharp cutoff at some low flux $S_o$, see equation [III.61] et seq.). The fluctuations from this distribution may also include an additional pure Gaussian component of variance $\sigma^2_\xi$. The result is a distribution in the dimensionless value $\xi$, which is related to the observed intensities by $W_{eff}$. This approach provides some economy in computation, in that the fluctuations distribution need be calculated only once for a given value of $\gamma$, $\xi_c$, and $\sigma_\xi$.

The discrete nature of numerical calculations with computers, as opposed to exact algebraic closed form solutions, requires that the models and resulting distributions exist as a series of sample points. Therefore $n(\xi)\ d\xi$ is replaced by an array of points $N_j$. The size of the array is a power of two, usually 2048 points. The array index, $j$, is proportional to the dimensionless intensity $\xi$ associated with a particular array element. $j1$, which need not be an integer, corresponds to the point where $\xi = 1$. $N_j$ is the mean number of sources expected with intensities

$$\xi_j = j\ /\ j1\ ,  \hspace{3cm} [A3.1]$$

so that

$$N_j \approx n(\xi_j)\ \Delta\xi\ ,  \hspace{3cm} [A3.2]$$

$$\Delta\xi \equiv 1\ /\ j1\ .  \hspace{3cm} [A3.3]$$

The exact value chosen for $j1$ is unimportant. For the fits presented in this dissertation, we used $j1 = 30$. Values much smaller than 10 did not allow sufficient room to model accurately the form of the low intensity roll off. Values much larger than 50 tended to have very large mean intensities for the fluctuations distribution. Such distributions suffered from precision difficulties. The zeroth element of the array, $N_0$, corresponded to the number of sources with no intensity. By convention this was set to zero.

The cumulants of the $N_j$ array (see Section A2:2 and Section III:2) are easily calculated. Similarly accurate calculations can be made for the cumulants of the exact $n(\xi)$ model distribution and compared with the array values. Because of the discrete nature of the array (and the

approximation inherent in [A3.2]) the two will'not be in agreement. For
the modified power law $n(\xi)$ models we fit, most of the disagreement
originates where the model is changing rapidly from one array element to
the next, i.e. at the lowest few elements. Thus the formal divergence
of the higher model cumulants can be avoided by restricting the
comparison of the cumulants contributed by the elements from 1 to j1
versus the model range $\xi = 0 \to 1$. By adding or subtracting sources from
the first few (low intensity) elements of the array, the third cumulant
of the $N_j$ can be adjusted to exactly match the corresponding value for
the model distribution. Any residual difference in the second cumulant
(variance) is then calculated, $\Delta\sigma^2_j$. The array mean intensity, $\bar{j}$, is
calculated for the entire array. This mean could be much larger than
the size of the array, especially if j1 is large so that there are a
large number of low flux sources in the array. (Remember that the mean
of the fluctuations depends on the behavior of the low flux sources for
$\gamma > 2$.)

### Calculation of the fluctuations array

The $N_j$ array is used to calculate the fluctuations distribution
using the Fourier techniques of equation [III.43]. Because the array
size was chosen to be a multiple of two and the values are all real, the
radix 8 fast Fourier transform technique [Bergland 1969] offers an
efficient realization of the transforms. Double precision calculations
of the transforms proved necessary to avoid unacceptable accumulation of
round-off error. Because of the nature of the floating point hardware
in the IBM machines used in the analysis (360 Model 91 and 3081) there
was little computational penalty in this, except for the calculation of
the complex exponentials.

The intrinsic Gaussian variation, both from the model $\sigma^2_\xi$ and $\Delta\sigma^2_j$
(the correction for any difference between the model $N_j$ low intensity
end variance), can be easily included in this step. The convolution
with the additional Gaussian is equivalent to multiplying the
transformed fluctuations distribution with the Gaussian before applying
the inverse transform. (As the magnitude of the counting statistics

variance changed from one data point to the next, that variation is not included in this step.  Such inclusion would also destroy the scale free nature of the resulting distribution, where all scale information is determined by the value of $W_{eff}$.)

The final inverse transform (Fourier synthesis) returns an array $P_j$.  The array is normalized so that $P_j$ is the probability of an observation with intensity $\xi_j \pm k \times M$, where k is some integer and M is the array size.  The imprecision is due to the finite Fourier transform's treatment of the $N_j$ as a periodic function with period M. The mean of the distribution will correspond to the bin position $\bar{j}$ modulus M.  The distribution may continue past the physical edge of the array by wrapping around, i.e. bin zero is between bin M-1 and bin 1. This wrap-around can actually contaminate the probabilities near the mean of the distribution with an additional contribution from the power-law high intensity tail.  This can be explicitly calculated and removed from the distribution.  Alternatively, the behavior of the high intensity sources responsible for the tail can be excluded in the initial $N_j$ array by zeroing a portion of the high intensity elements. The existence of a high intensity cutoff in $N_j$, either from the explicit zeroing or from the top of the array, must be included in calculating the proper correction to the array distribution mean $\bar{j}$ for comparison with the true intensity mean (equations [III.56] and [VI.2]).

## Calculating the integral and differential probabilities

In fitting the newly derived fluctuations distribution to observations the array must be scaled by $W_{eff}$ and convolved with each observations photon statistics distribution.  The convolutions are necessary to calculate the differential probability ([V.2]) for likelihood estimates and the integral probability ([V.4]) for the goodness-of-fit estimates.  We assumed that the photon statistics were Gaussian and independent of the model intensity.  The computation was greatly speeded up by replacing the evaluation of the Gaussian with a look-up in a pre-calculated table.  For the integral probability the order of integration was reversed so that the inner integral could also

be handled by table look-up. Thus, only multiplications and adds were needed for most of the performance of the noise convolution integral. Cubic interpolation was used for the $P_j$ array, with linear interpolation for the noise distributions.

## Negative response calculations

The techniques that we have outlined above can be easily modified for use in situations where the spatial response function, $r(\theta,\phi)$, includes positive and negative response regions. Examples of such an effective detector response are: the internal background measuring FOV combination $r_B$ (equation [II.13]); baseline subtraction estimates of the sky flux (Section III:7); and, in the radio, phase-switching interferometers. As discussed at the end of Section II:2, the general approach is to divide the response into two portions, covering the positive and negative response regions respectively. The $\nu_+(r)$ response frequency is used to derive an $N_j$ array as described earlier in this Appendix. The wrap-around character of the array is now exploited. From $\nu_-(r)$ we can calculate a similar $N^{(-)}{}_j$, but the value of the scale point bin is fixed at

$$j1_- = j1 \; (R_{-;(\gamma-1)} \; / \; R_{+;(\gamma-1)}) \qquad\qquad [A3.4]$$

where $R_{\pm;(\gamma-1)}$ are the indicated detector response weights for the $\nu_\pm(r)$ response frequencies. $N^{(-)}{}_j$ is then added to the M-j element of the $N_j$ array. The derivation of $P_j$ is the same as before, though we note that the variance corrections of the two regions are added together. When scaling the resultant $P_j$ array for the fits, we must use $W^{(+)}{}_{eff}$, the effective width from the positive response portion of the detector.

APPENDIX A4

THE SCANER DATABASE

The format of the SCANER database is given, with a detailed
description of the header blocks and the SPINAXIS file used
to access the data.

The SCANER database is the basic HEAO 1 database used in the
analysis of the X-ray sky. Its creation is detailed in Section IV:1.
The SCANER data exist as tape files. Those produced directly by the
program SHFSCN have the best resolution, each scan corresponding to one
spinaxis select period (nominally 0.5 days) with the scan divided into
1800 bins ($0.2°$ resolution). Lower resolution scans are produced by the
program SUSCAN, which combines adjacent scans and also can produce a
coarser binning.

The SCANER tape format

Each scan, or superimposed scan, is a single file. Each file is a
series of 500 byte records. The first is the header record, containing
information on the organization and the accumulation of the data:

| | | |
|---|---|---|
| iafile | I*2 | The output file number of the BIGSCN tape. |
| irun | I*2 | The run number of the BIGSCN run that produced this data. |
| nspin | I*2 | The number of good major frames for this scan. |
| ftime | R*8 | The start time of the scan (in DOY 1977). |
| stime | R*8 | The stop time of the scan. |
| xax | R*8(3) | The scan plane x-axis as a real triple. |
| yax | R*8(3) | The scan plane y-axis as a real triple. |
| zax | R*8(3) | The scan plane z-axis, i.e. the mean spinaxis, as a real triple. |
| spnper | R*4 | The mean spin period (in minutes) for the data. |
| spnexc | R*4 | The mean variation in the spin period (in minutes). |
| excurs | R*4 | The approximate mean excursion of the spinaxis from its mean position during data acumulation. |

SCANER Database

| amxmac | R*4 | The maximum value of Macilwain L, for LOWMAC files only (see following description of McIlwain L classification), ALLMAC files this value is meaningless. |
|--------|-----|-------------------------------------------------------------------------------------------------------------------------------------------------------|
| dattap | R*8 | The name (a8) of an input data XRATES tape (for some part of the data used the scan). |
| darch | R*8 | The name (a8) of the output BIGSCN tape for low McIlwain data |
| nbin | I*4 | The number of bins into which the scan plane is divided, less than or equal to 1800. |
| nrec | I*4 | Spare. |
| darmak | R*8 | The name (a8) of the output BIGSCN tape for the high McIlwain L data. |
| darmak | R*8 | The name (a8) of the original SHFSCN tape used for this scan. |
| dmemo | R*8 | The name (a8) of the SPINAXIS member that originally had a record that points to this file. |
| iadmac | I*4 | The MacIlwain L index used for selecting LOWMAC data, an integer value derived from amxmac above. |
| naxott | I*4 | The total number of major frames that had acceptable data for at least one detector. |
| nspnml | I*2 | The number of major frames that did not meet the McIlwain L selection criterion. |
| nspnuc | I*2 | The number of major frames that produced not electron superclean MED data. |
| mfilef | I*4 | The file number of the original SHFSCN tape. |
| qflag | L*1(8) | A flag array specifying the qualities of this tape: |

(1) If True, indicates includes all values of McIlwain L; if False, the data was selected by the criterion indicacted in amxmac.

(2) If True, indicates the scan is a superposition of many scans, if False, the data was produced directly by SHFSCN from the BIGSCN output tapes.

(3-8) Spare.

qpad          L*1          Sufficient bytes to pad out to total record size of 500. For For superposed scans produced by SUSCAN this array contains information about where the data used in superposition came from.

Each bin on the scan plane, 1800 of them for the SHFSCN output scans, usually 360 for the SUSCAN output files, is contained on one of the following 500 byte records. Each record consists of 5 100 byte subrecords, one for each of the five detectors: HED-1 HED-2, MED(superclean), HED-3, MED(all data). The subrecords organization is:

lds           I*4(8)       LDS(n) is the sum of the counts of the nth Discovery Scaler.

numexp        I*4          The number of exposures.
                           CAUTION: Many subroutines, in particular GTRATL, which returns a particular rate, is passed LDS as an array and assumes that NUMEXP is the 9th entry of that array. As long as the records from the tape are read using FREAD onto a common block and not moved around, this is valid. To avoid problems, NUMEXP should be picked up as the 9th entry for LDS whenever LDS is copied to a new location.

ldssqr        L*4(8)       LDSSQR(n) is the sum of the squared counts of the nth Discovery Scaler.

ldscrs        I*4(4)       LDSCRS(n) is a sum of the cross product of the counts in the Nth pair of Discovery scalers, e.g. LDSCRS(1) is the sum of the product of counts in Discovery Scalers 1 and 2.

ldelz         I*4          The sum of the excursion from the scan plane, in units of 0.05 degrees.

ldelz2        I*4          The sum of the square of the excursions from the scan plane.

macsum        I*4          The sum of L, where L is (100. * (McIlwain L -1)).

qdatfl        L*1(4)       A set of logical data quality flags.
                           (1) If true, then the detector was pointing to this bin during an exposure when the Discovery

Scaler LSB register had an overflow, so that no 1.28 second data was provided. This is usually due to a bright source in the field of view. There is no attempt to estimate the correct data and the number of exposures was not incremented.

(2) If true, then the sum of counts squared became so large that some of the entries in LDSSQR and LDSCRS may be invalid. However all of the other accumulations were made and the number of exposures was incremented.

(3) If true then some of the data in this bin were taken during a major frame where the detector's Source-in-Small-Field-of-View flag had been set.

(4) If true then as for (3), save an indicator that the Source-in-Large-Field-of-View flag had been set.

The SCANER files are usually files on 6250 bpi density tapes. There are a set of tapes for all the data, and another set that is for Low McIlwain L data only. All low L data was taken when L was less than 1.18.

### The SPINAXIS file format

The SCANER tape files are approximately in time order. However due to some correction programs, some of the scans were placed, out of order, on an auxillary tape. To determine the tape volume and file number associated with a particular time interval, or spinaxis position, there are the SPINAXIS files. These members of various PDSs contain a short summary record for each SCANER file:

ftimes   R*8       Begin time for scan accumulation.

stimes   R*8       End time for scan accumlation.

spinax   R*8(3)    Vector triple for the spinaxis position for the scan.

SCANER Database

| sspnpr | R*4 | Scan period. |
|--------|-----|--------------|
| sspnex | R*4 | Mean variation in scan period. |
| sexces | R*4 | Mean deviation from position given in spinax (in degrees). |

(It is often wise to discard data with too large a value for SEXCES, as this is an indication of possibly abnormal conditions during data accumulation. In particular some data near a pointing period might have been included with a particularly large displacement from the nominal scan plane.)

| itflag | L*4 | Spare |
|--------|-----|-------|
| tname | R*8 | Tape volume name (a8) of associated SCANER file. |
| ifile | I*2 | File number of associated SCANER file. (CAUTION: As this is a two byte integer it CANNOT be used as an argument to the FTIO routine MOUNT and POSN which expect a full word integer. It should be converted by an assignment "IFULL=IFILE" where IFULL is I*4 and then use IFULL as the argument.) |
| nspins | I*2 | The number of major frames that contributed data to the scan. This is a good indication of the amount of good data avalable in this scan. |

The SPINAXIS members are time ordered, although it may take one or more members to cover an entire set of SCANER files. The current SPINAXIS files available are: ALLMAC.SPINAXIS, the master all-McIlwain-L data produced by SHFSCN; LOWMAC.SPINAXIS, the low-McIlwain SHFSCN produced data; SUSALL.SPINAXIS, the superposed data (1 degree by 2 days) for all-McIlwain data produced by the program SUSCAN; and SUSLOW.SPINAXIS, the superposed low-McIlwain data.

### Programming hints

To change the data array into rates, look at XTRACT and in particular the subroutine GTRATL. They are fairly well documented.

GTRATL is used to convert the summed counts into a rate of a given type, i.e. LFOV, SFOV, L+S, L-S, and Internal Background as well as to return the photon variance.

.    It should be noted that it is fairly expensive to run through even 6 months of the data base, so a fair amount of debugging should be done on a small subset. It is also very expensive to keep an entire SHFSCN file (1800 bins by 500 bytes) in core, so a bit of economy here goes a long way.

APPENDIX A5

NEW SOURCE IDENTIFICATIONS FOR THE X-RAY COMPLETE SAMPLE

Sources from the X-ray Complete Sample of Piccinotti et al.
[1982] are discussed. New identifications are presented, as
well as the classification of the galactic sources.

New identifications

The identifications are presented here as a public service. We
also make additional notes about the characteristics of some of the
previously identified sources. The order of the sources is the same as
Table 1 in Piccinotti et al. Information from "A-1" is usually from the
catalogue of sources compiled with the A-1 experiment on HEAO 1 [Wood et
al. 1982, NASA data center]. Improved error boxes are also provided
for some sources by the third Ariel V catalogue of high latitude sources
[McHardy, Lawrence, Pye, and Pounds 1981] indicated by the prefix "3A".
Other sources of identifications are the HEAO 1 A-3 experiment scanning
modulation collimators and the imaging detectors from the Einstein
Observatory.

The sources are:

H0111-149:  Identified with Markarian 1152 [Ward, Elvis, Lawrence et
    al., in preparation]. Seyfert 1.

H0235-52 (2A0235-52): Identified with ESO 198-624 [D. Schwartz, private
    communication 1982]. Seyfert 1.

H0316-443 (2A 0316-443): Identification with the cluster source
    associated with PKS 0316-443 is strengthened by smaller A-1 and
    Ariel V (new designation 3A 0316-442) error boxes.

H0917-074:  Currently  unidentified.  Though  a  tentative  A-1
    identification was made with the cluster of galaxies Abell 780, a
    GSFC IPC field shows no flux from the cluster. No additional
    Abell clusters are in the detector field of view. The source
    intensities were 1.36±0.28 R15 and 1.21±0.3 taken during
    observations six months apart, consistent with no variability.

H0952+699 (2A 0954+700 and others): Although identified with the peculiar galaxy M82, there is the possibility for confusion with the active galactic nucleus in M81. Based on the A-2 2-10 keV luminosity of 0.0024 $L_{44}$ and the estimates of the luminosity of M81 of 0.0001 $L_{44}$ [Elvis and van Speybroeck 1982], the amount of contamination is less than 10%; however, these two measurements were not simultaneous. The X-ray spectrum observed with A-2 for this source is consistent with a 5 keV thermal [Mushotzky].

H1257-042: Probable identification with Abell 1651 is strengthened by smaller A-1 error box (new designation H1257-042).

H1325-020: Currently no identification. The first scan intensity was 1.47±0.25 R15, while six months later the 95% confidence upper bound was 1.0 R15. This probably variable behavior is inconsistent with a cluster identification. A galactic source, or perhaps an AGN are therefore more likely possibilities.

H1344-333 (2A 1344-325): Possible identification with the cluster SC 1344-32 is improved by A-1 (H1834-326).

H1630+057 (2A 1630+057): Possible identification (A-1) with the distance class 5, richness 3, Abell cluster of galaxies 2204 [Johnson et al. 1983]. With large uncertainties, the derived luminosity of 50 $L_{44}$ makes this object the most luminous cluster in the XCS. The A-2 spectrum, obtained from point data, is best fit with a 6.5 keV thermal bremsstrahlung [Mushotzky].

H1829-591 (4U 1830-60): Identified by J. Patterson with a narrow emission line galaxy, classifying this source as an AGN.

H1834-653: The identification with the Seyfert 1 galaxy ESO 103-G35 is strengthened by the smaller A-1 error box (H1832-652).

H1846-786 (1M 1849-781): Identified by Remilard and Bradt with a Seyfert 1.

H2154-304 (2A 2151-316): Piccinotti et al. identify this object with the BL Lac object PKS 2155-304. On the basis of the redshift used, z = 0.17, this object is the most luminous in the XCS, 150 $L_{44}$. There is some disagreement over the validity of this redshift (see the discussion in Urry and Mushotzky [1982] who also provide a detailed analysis of the emission from this object).

H2151-605 (2A 2155-609): The possible identification with the cluster

New Identifications

Str 2159-602 is improved by the observation of a thermal X-ray spectrum [Mushotzky] and small A-1 error box.

H2158-321 (2A 2151-316):The possible identification with NGC 7172 has been confirmed by an IPC detection [M. Elvis]. The source is probably more unusual than a "normal galaxy", as identified in Piccinotti et al. The X-ray luminosity of 0.18 $L_{44}$ is much larger than that expected from a normal galaxy. Longmore and Sharples [1982] have detected strong, variable, infrared emission. Subsequent work [Longmore] indicates emission lines from the nucleus, so that the identification probably should be revised to "other AGN".

H2216-027 (2A 2220-022): A possible identification with the AGN 3C 445 was made. The source was subsequently confirmed as an X-ray emitter by an IPC observation, but with greatly reduced flux. The high luminosity from A-2, 4.7 $L_{44}$, makes it unlikely that the discrepancy is due to absorption, so that the difference requires marked variability. A-2 saw a 37% reduction in the flux between two observations seperated by six months. Some fraction of the flux attributed to this object is probably due to Abell 2440, whose flux is ~0.35 R15, based on an IPC measurement.

## Classification of high-latitude galactic sources

The 10 galactic sources analysed in Section VII:2 consist of 3 RS CVn stars and 7 cataclysmic variables. These are presented in Table 2 of Piccinotti et al. The RS CVn stars are: H0123+075 (a transient, HD 8357); 4U 0336+01 (HR 1099); and 2A 1052+60 (BD 61+1211). The cataclysmic variables are: 2A 0311-22; 2A 0526-32; H0751+22 (U Gem); 2A 1249-28 (EX Hya); 2A 1815+50 (AM Her); H2215-08 (incorrectly identified as Wolf 1561, see Patterson and Steiner [1983]); and H2252-035. See Worrall and Marshall [1983] for a further analysis of the cataclysmic variables. Two other sources in Table 2 of Piccinotti et al., are not used in our analysis as their first scan fluxes are below the 1.25 R15 completeness limit. These are H0328+05 and 2A 1704+24. Doxsey et al. [1983] have improved the X-ray position for H0328+05, redesignating it H0323+022. They discuss several interesting

"Page missing from available version"

# APPENDIX A6

## THE MALMQUIST CORRECTION AND THE FLUCTUATIONS

We examine the effects of the fluctuations on the Malmquist correction technique. Though the intensity measurements may no longer have the $5\sigma$ significance, if the variation is predominatly confusion noise from the fluctuations, there is a natural low flux cutoff to the probability distribution that removes the divergence. We estimate the correction to the resolved source $N(S)$ fits due to an additional confusion noise term.

The problem

As we remarked in Section VII:1, the analysis of Piccinotti et al. [1982] may have not properly calculated the total uncertainty of a source intensity measurement. This was based on the observation that the average source flux error was less than the total error observed when least-square fits were made of random positions in the sky (see Section III:7). It is certainly true that the Piccinotti et al. analysis did not explicitly include the effects of confusion. It is less certain whether the correct magnitude of the confusion uncertainty is the extra variance derived in Section III:7 (see the discussion of Figure III.15). This is because the error term derived by Piccinotti et al. was determined by a $\delta\chi^2$ fit, treating the position of the source to be an (uninteresting) unknown parameter. In contrast the extra variance of $\delta_{I;fluct}$ of Section III:7, did not vary the position and calculated the error by inverting the correlation matrix. It might be argued that the $\delta\chi^2$ errors are less prone to confusion noise because they are determined by a process that is explicitly testing for a signal shaped like the detector response. A proper test of the magnitude of confusion noise contribution requires a Monte Carlo simulation following the exact technique used by Piccinotti et al. In the absence of such a test, in this section we will assume that each of the Piccinotti et al. source flux errors must be increased by an additional variance term due to confusion, $\sigma^2_{S;fluct} = (0.2 \ R15)^2$. The exact value is not important, as the purpose of this appendix is to estimate the possible effect. The

Malmquist Bias Reconsidered

total σ, then ranges from 0.29 to 0.34 R15, depending on the source (see Table VII.1).

In trying to estimate the impact of this additional uncertainty on the Malmquist correction, (so that we can calculate the resolved source Euclidean N(S) normalization, see Section VII:1), we are faced with a limitation of our algorithm. The lowest flux sources, with S ~ 1.25 R15 are no longer formally significant detections at the 5σ level. As determined by Jauncey, Crawford, and Murdoch [1973], the 5σ criterion is needed to avoid divergences in evaluating the Malmquist correction. For sources with lesser significance, the power law of the underlying N(S) overwhelms the Gaussian error term in the convolution in equation [VII.9]. A low flux cutoff will eliminate the divergence, however for sources with S<5σ, the size of the Malmquist correction depends to an unacceptable degree on the position and form of the cutoff. An attempt to determine the N(S) for the resolved sources should not rely on the form of N(S) at much lower fluxes.

The partial solution to this problem requires a more careful study of the · Malmquist correction equations [VII.8] and [VII.9]. The difficulty is that there are an enormous number of low flux sources, which by equation [VII.8], have a small individual chance of being detected with a much larger flux. This picture is inaccurate. Let us assign the fit source flux to the brightest source in the field of view. The effect of all dimmer sources is only through the confusion noise. Therefore, although there may be an enormous number of low flux sources, the probability that one of them is the brightest source is vanishingly small. We therefore replace the N(S) distribution in [VII.9] with $P_b(S)$, the differential probability distribution of the brightest source. This is found by

$$P_b(S) = d\mathbb{P}_b(S) \, / \, dS \qquad\qquad\qquad [A6.1]$$

where $\mathbb{P}_b$ is the integral probability, found by

$$\mathbb{P}_b(S) = 1 - \exp\left( \int_\infty^S n(S) \, dS \right) . \qquad\qquad [A6.2]$$

n(S) is the expected number of sources in the field of view of the

measurement with flux S. (For the exposition in this section we ignore the nicety of the distinction between source intensity and flux, though care has been taken to make the proper correction in the results reported later in this appendix.) The corrected number of sources in the sample is found by multiplying the distribution $P_b$ by the total number of observations. This is equivalent to the correction by the survey solid angle in equation [VII.4]. $P_b$ has the expected behavior at both limits of S. The order of the limits of the integral in equation [A6.2] insures that the argument of the exponential is less than zero, so that P has the property of being initially zero for high values of S, and then approaching 1 as S goes to zero. In addition at high fluxes, where it is almost certain that a source of flux S will be the brightest source in the measurement, $P_b$ is proportional to N(S), so that the correction is the same as the usual Malmquist factor we calculated in Section VII:1. At lower flux values $P_b$ rolls over so that a natural cutoff is imposed. Unfortunately, where before the normalization of the N(S) model was unimportant for the calculation of the correction term, now the cutoff does depend on the normalization chosen. On the surface, it appears that as the correction term now depends on a detail of the model (the normalization) that this is similar to the original correction which strongly depended on the position and form of the cutoff for the $<5\sigma$ significant sources. It turns out, though, that because the particular, natural, form of the $P_b$ cutoff the correction term varies only slowly as a function of $W_{eff}$, the N(S) normalization even for the sources with $\sim3\sigma$ detections.

Table A6.1: Estimated new Malmquist corrections

The bins are the same as used in Table VII.1. The $\sigma_{tot}$ is the total uncertainty, with a confusion noise of 0.2 R15 added in quadrature to the original uncertainty. $f_{old}$ is the old Malmquist correction. $f_{new}$ is the new correction using the larger photon uncertainty, and N(S) replaced with $P_b$. The calculation of $P_b$ assumed a W of 0.062 H1L1, equivalent to a $W_{eff}$ of 0.066 R15.

| Bin | $\sigma_{old}$ | $\sigma_{new}$ | $f_{old}$ | $f_{new}$ |
|-----|------|------|-------|-------|
| 1 | 0.28 | 0.34 | 1.007 | 1.010 |
| 2 | 0.24 | 0.31 | 1.018 | 1.029 |
| 3 | 0.22 | 0.30 | 1.030 | 1.055 |
| 4 | 0.22 | 0.30 | 1.048 | 1.092 |
| 5 | 0.24 | 0.31 | 1.084 | 1.149 |
| 6 | 0.23 | 0.30 | 1.104 | 1.194 |
| 7 | 0.23 | 0.30 | 1.136 | 1.273 |
| 8 | 0.21 | 0.29 | 1.137 | 1.339 |

APPENDIX A7

FLUX CONVERSIONS

Using the concepts and formulae developed in Section II:3, this appendix presents the conversion coefficients from ergs to detector counts and the effective spectral range of various A-2 counters and idealized windows of the Einstein Observatory IPC.

———————————————

1: Conversions for Selected Discovery Scaler Windows of the A-2 Experiment

We present conversion tables for selected windows of the HED 1, MED, and HED 3.

The tables include: the Discovery Scaler window, ·with L1 or L2 representing the total first or second layer response; and $\Delta E_Q$, the energy range in keV that includes all values where the total efficiency, $Q(E)$, is greater than 10% of its peak value (see equation [II.23]). Then for each of three incident spectra (two thin thermal bremsstrahlung, and a single power law model, all with no absorption) we tabulate the energy range where 90% of the detected counts originate ($\Delta E_{90}$, see discussion of equation [II.25]); and the conversion coefficient from counts to $10^{-8}$ ergs $s^{-1}$ $cm^{-2}$, 2-10 keV ($\varepsilon \times 10^8$, see equation [II.27]). All values presented use the detector response matrices as determined around day 257 of 1977. See Section II:3 for further details.

Table A7.1: Spectral response for HED 1

Taken from detector matrices H111095C and H121095C

| Window | $\Delta E_Q$ keV | Incident Spectrum | | |
| --- | --- | --- | --- | --- |
| | | kT = 6 keV | kT = 40 keV | $\alpha$ = 0.65 |
| L1 | 2.6–46.5 | | . | |
| | $\Delta E_{90}$ | 2.6–9.9 | 2.5–13.3 | 2.2–12.7 |
| | $\varepsilon \times 10^8$ | 1.54 | 1.27 | 1.33 |
| 1A+1B | 2.6–36 | | | |
| | $\Delta E_{90}$ | 2.5–6.4 | 2.6–6.6 | 2.6–6.6 |
| | $\varepsilon \times 10^8$ | 2.11 | 2.32 | 2.28 |
| 1C+1D | 5.7–46.5 | | | |
| | $\Delta E_{90}$ | 5.8–12.8 | 5.75–46.5 | 5.4–17.4 |
| | $\varepsilon \times 10^8$ | 5.69 | 2.84 | 3.18 |
| L2 | 3.85–64.5 | | | |
| | $\Delta E_{90}$ | 3.6–15.1 | 3.7–25.4 | 3.5–30.4 |
| | $\varepsilon \times 10^8$ | 21.6 | 8.88 | 9.68 |

Table A7.2: Spectral response for MED

Taken from detector matrices M11N and M12N

| Window | $\Delta E_Q$ keV | Incident Spectrum | | |
|---|---|---|---|---|
| | | kT = 6 keV | kT = 40 keV | $\alpha$ = 0.65 |
| L1 | 1.7–12.2 | | | |
| | $\Delta E_{90}$ | 1.5–6.2 | 1.5–7.5 | 1.5–7.1 |
| | $\varepsilon \times 10^8$ | 1.60 | 1.84 | 1.75 |
| L2 | 1.9–16.4 | | | |
| | $\Delta E_{90}$ | 1.7–8.0 | 1.7–10.1 | 1.5–9.5 |
| | $\varepsilon \times 10^8$ | 3.32 | 3.35 | 3.31 |
| 2A | 1.8–3.2 | | | |
| | $\Delta E_{90}$ | 1.8–3.1 | 1.8–3.1 | 1.8–3.1 |
| | $\varepsilon \times 10^8$ | 9.96 | 15.3 | 12.6 |
| 2B | 2.6–18 | | | |
| | $\Delta E_{90}$ | 2.5–8.8 | 2.8–11.2 | 2.6–10.6 |
| | $\varepsilon \times 10^8$ | 4.99 | 4.28 | 4.06 |

Table A7.3: Spectral response for HED 3

Taken from detector matrices H31257c and H321095C

| Window | $\Delta E_Q$ keV | Incident Spectrum | | |
|---|---|---|---|---|
| | | kT = 6 keV | kT = 40 keV | $\alpha$ = 0.65 |
| L1 | 2.55–46.5 | | | |
| | $\Delta E_{90}$ | 2.6–9.7 | 2.6–13.4 | 2.2–12.6 |
| | $\epsilon \times 10^8$ | 1.52 | 1.26 | 1.32 |
| 1A+1B | 2.55–36.0 | | | |
| | $\Delta E_{90}$ | 2.6–7.0 | 2.8–7.4 | 2.7–7.3 |
| | $\epsilon \times 10^8$ | 1.86 | 1.95 | 1.95 |
| 1C+1D | 6.85–46.5 | | | |
| | $\Delta E_{90}$ | 6.4–13.2 | 6.4–18.0 | 6.6–19.3 |
| | $\epsilon \times 10^8$ | 8.34 | 3.60 | 4.06 |
| L2 | 3.85–64.5 | | | |
| | $\Delta E_{90}$ | 3.0–14.4 | 3.7–35.4 | 3.5–30.4 |
| | $\epsilon \times 10^8$ | 21.4 | 8.85 | 9.64 |

——————————

## 2: IPC Spectral Response and Flux Conversion

The approximate form of the Einstein Observatory IPC
spectral response is presented. This is used to calculate
the effective bandpasses for the Deep and Medium Sensitivity
Surveys. Conversion coefficients to absolute units (ergs
2–10 keV) and HEAO 1 flux units (H1L1) are estimated. The
exact value of the conversion is strongly dependent on the
assumed incident spectrum.

The spectral response of the IPC depended on both the intrinsic
response of the proportional counter and the telescope imaging
efficiency. The physics of the latter included the reflectivity of

nickel for grazing incidence X-rays, with corrections for scattering of
the photons by small and large scale inhomogeneities in the scattering
surface. For the HEAO detectors we worked with the total response
function R(h,d) (see section II:3). We do not have available a fully
detailed response for the IPC, but scientists at the Harvard-Smithsonian
Center for Astrophysics have distributed the effective area as a
function of energy, A(E). This is similar to our total efficiency
(equation [II.23]) multiplied by the concentrating area of the imaging
system. Thus A(E) has units of $cm^2$. If s(E) is the incident
differential energy flux, then the expected count rate in a pixel is

$$S_{counts} = \int dE \ E^{-1} \ s(E) \ A(E) \qquad \qquad [A7.1]$$

(compare equation [II.24]). All information in this section is taken
from the memo of 29 September 1982 from D. Fabricant. The spectral
response is calculated for image bins of 3' radius, using the on-axis
spatial response. A plot of A(E) is presented in Figure A7.1(b). The
response is separated into two regions by the edge at 0.284 keV. A
substantial number of counts can be contributed by the below 0.284 keV
region. This can be seen in Figure A7.1(a), where the probability
distribution of incident photons is plotted, assuming a 0.7 power law
spectrum and low absorption. 90% of the photons across the entire IPC
bandpass come from 0.2-2.45 keV. This is almost totally disjoint with
the H1L1 bandbass.

There is increased noise in the very lowest channels, therefore
for both the Medium and Deep Surveys, the investigators restricted their
consideration to a subset of the IPC pulse height channels. The intent
was to have a well defined energy bandpass. Unfortunately, the IPC
channel bandpass corresponding to a particular energy range was strongly
variable, particularly early in the life of the instrument. Thus the
calibration of the gain was needed to choose the particular pulse height
boundaries for a particular observation. As an additional complication,
the initial calibration used in both the Deep Survey [Giacconi et al.
1979] and the Medium Survey [Maccacaro et al. 1982] has since been
found to be wrong. Using the improved calibration [Harnden, private
communication] we estimate that the Deep Survey channels used actually

Figure A7.1: P(E) and A(E) for IPC

(a)    The probability distribution of counts for an incident $\alpha$ = 0.7 spectrum with $N_H$ = $3 \times 10^{20}$ cm$^{-2}$ (galactic absorption).   The distributions are shown for the total IPC response, as well as idealized estimates for the channel ranges selected for the Deep and Medium Surveys. The distribution for the H1L1 window is also shown for comparison.    (b) The effective area A(E), with the indicated ranges defining the Deep and Medium Survey bandpasses.
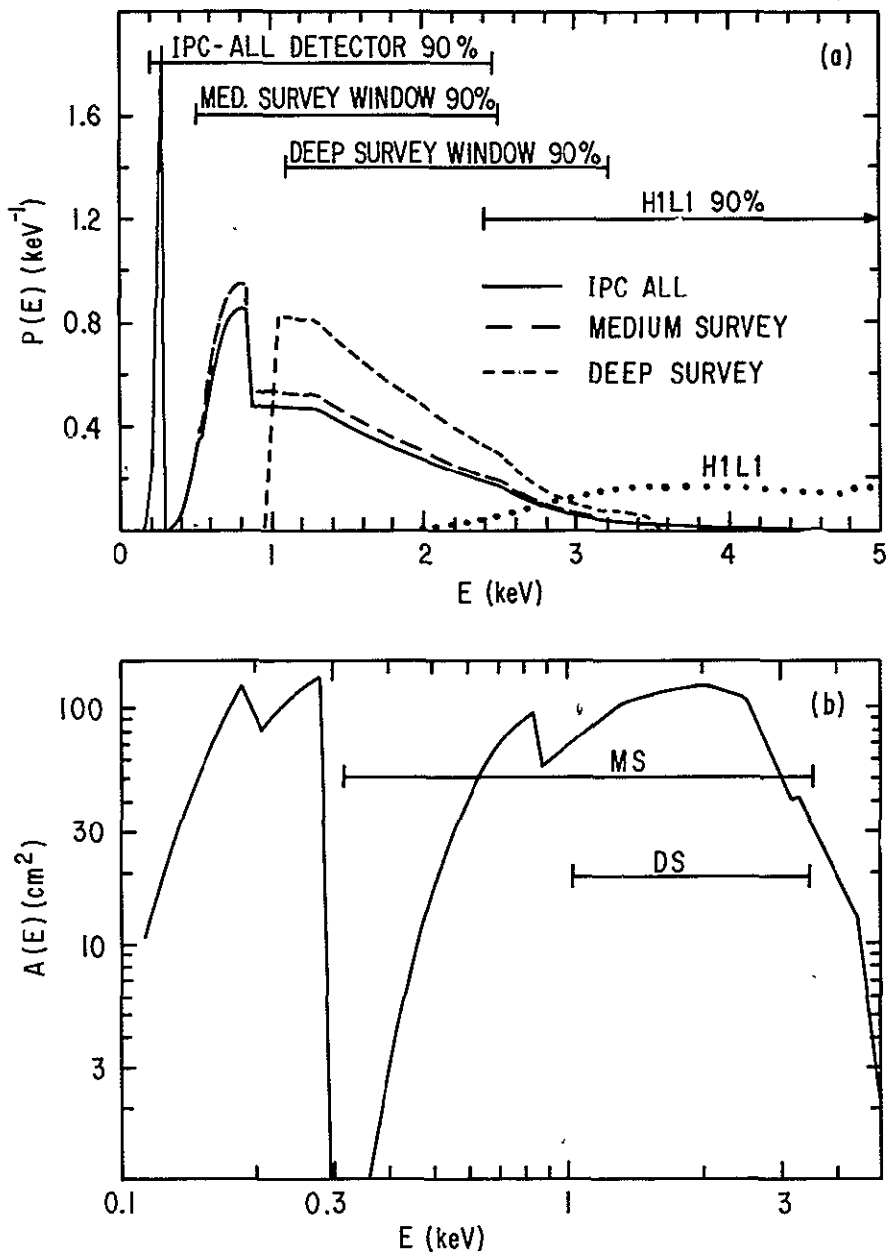
correspond to slightly different energies than the 1 and 3 keV boundaries desired. The exact impact depends on the particular observation out of the four that make up the Deep Survey. We estimated that in an average sense the effect is to raise the true Deep Survey band pass from 1-3 kev to 1.1-3.3 keV. A similar shift can affect the Medium Survey band (nominally 0.3-3.5 keV), but we have not estimated the magnitude.

In Figure A7.1(a) we plot the P(E) curves for our estimated responses of the Medium and Deep Survey windows. The calculation of these curves, and the response actually used for the MS and DS windows, are based on the A(E) function truncated at the low and high energies of the particular bandwidth. This is an approximation. The true window definition is in terms of pulse height channels, rather than energies (see equations [II.23] and [II.24]). Because of the stochastic nature of the number of electron-ion pairs generated by an incident X-ray in a proportional counter, there is significant spread in the pulse height distribution for a given energy. The effect is greatest for low energy photons. For the IPC, the spread for energies $\lesssim$ 1 keV is ~100%. Thus, contrary to our approximation for Figure A7.1, an incident photon of energy 0.7 keV has a non-negligible chance of being detected in the Deep Survey pulse height range. Because of the intrinsic falloff of the detector response at the high end of the bandpasses used, this approximation is not important at that end. For our purposes, this approximation means that we have underestimated the size of the expected count rate for a given incident spectrum. For figures where we have plotted the K(S) curves based on the XCS derived luminosity functions (e.g. Figures VII.24 et seq.) the use of a more correct version of the IPC response function is expected to shift the K(S) curves slightly up and to the right. Therefore, the upper limits on AGN evolution consistent with the Medium and Deep Survey counts are conservative ones. Similarly when transforming the IPC results to the higher energy HIL1

band (Figure VII.30 and the discussion of the contribution to the total sky flux from the Deep Survey sources) the correct conversion will have slightly fewer H1L1 counts per IPC count. This lowers the estimate of the contribution of the Deep Survey sources to the sky flux.

Figure A7.2: Conversion of IPC fluxes to absolute and H1L1 values

The abcissa is the power law energy index of the incident spectrum. Two values of the absorption are used, both assuming no intrinsic redshift in the absorption energy. "Low absorption" is $N_H = 3 \times 10^{20}$ atoms $cm^{-2}$, "high absorption" is $3.6 \times 10^{22}$ $cm^{-2}$. (a) The corresponding number of ergs (2–10 keV) for one IPC count. (b) The ratio of IPC counts to H1L1 counts. The dashed horizontal and vertical lines indicate the ratios and equivalent power law indices for thin thermal spectra with kT = 6 and 40 keV.
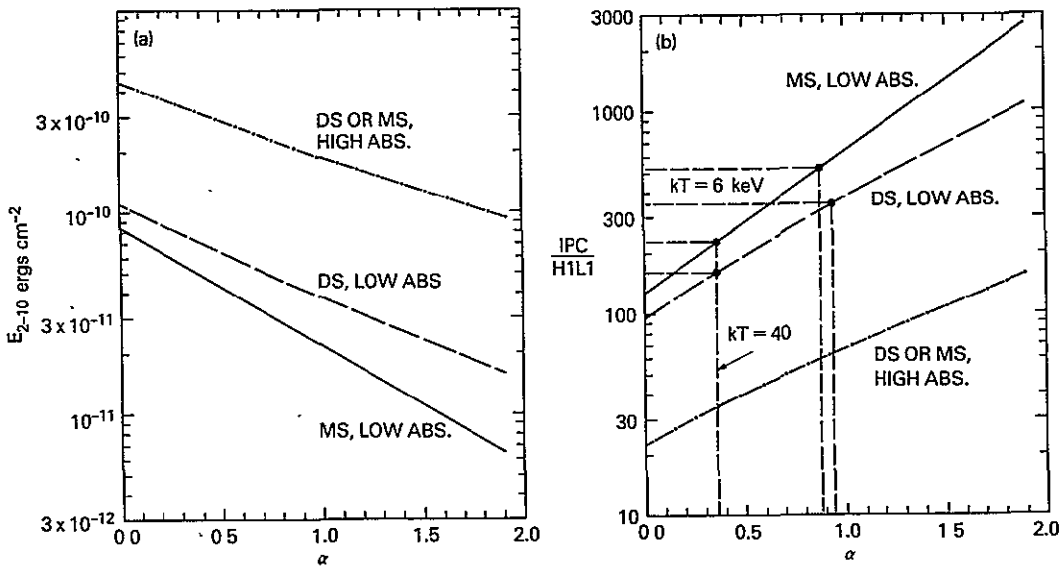


Figure A7.2 presents the conversion coefficients for the Deep Survey and Medium Survey IPC windows. A7.2(a) gives the equivalent number of ergs $cm^{-2}$ in the 2–10 keV range (see equation [II.27]). A7.2(b) presents the ratio of the expected IPC counts to H1L1 units. From this ratio we see, for example, that an α = 1.0 spectrum (low

absorption) produces a count rate in the IPCMS band equal to 650 cm$^2$ of an H1L1 detector. The horizontal dashed lines indicate the ratio for a thermal bremsstrahlung spectrum of either kt = 6 keV or 40 keV. The associated vertical dashed lines show the equivalent power law index. For example, a kT = 40 keV spectrum is equivalent, in converting to the IPCMS band, to an $\alpha$ = 0.36 power law.

The spectral dependence of the IPC flux rates applies not only to conversion to HEAO 1 flux rates, but also to ideal differential rates. For example, the differential X-ray flux at 2 keV, $s_2$, often used in calculating $\alpha_{ox}$, is not directly measurable from the IPC rates but must be estimated using an assumed incident spectral form. For example, a 1 IPCMS source with low absorption and $\alpha$ = 0.4 will have $s_2$ = 1.84x10$^{-11}$ ergs s$^{-1}$ cm$^{-2}$ keV$^{-1}$. If $\alpha$ flattens to 0.0, $s_2$ for 1 IPCMS increases by 15%, while if the spectrum steepens to 0.7, it decreases by 12%. Thus, if the true source spectrum is not known, the above range in the value of $\alpha$ corresponds to uncertainty in $\alpha_{ox}$ of ~±0.02. If the true $\alpha$ is outside the range 0.0-0.7 or if there is significant absorption the uncertainty in $\alpha_{ox}$ is magnified.

REFERENCES

Abramowitz, Milton, & Irene Stegun, eds., 1965 (fifth printing, corrected edition, 1968), Handbook of Mathematical Functions, (Dover:New York; also U.S. Government Printing Office).

Allen, C.W., 1973, Astrophysical Quantities, (Athlone Press:London).

Anderson, R.C., W.H. Brune, R.C. Henry, P.D. Feldman, & W.G. Fastie, 1979, Ap.J. (Lett.), 233, L39, "The Spectrum of the Diffuse Cosmic Ultraviolet Background".

Avni, Y., 1976, Ap. J., 210, 642.

_____, Astr. Ap., 63, L13, "On the Contribution of Active Galactic Nuclei to the Diffuse X-ray Background".

Avni, Y., and H. Tananbaum, 1982, Ap. J. (Lett.), 262, L17, "On the Cosmological Evolution of the X-ray Emission from Quasars".

Bahcall, John N., and Raymond M. Soneira, 1980, Ap. J. (Lett.), 238, L17, "Star Counts as an Indicator of Galactic Structure and Quasar Evolution".

Bahcall, N.A., 1979, Ap. J., 232, 689.

Bahcall, Neta A., and Raymond M. Soneira, 1982, Ap. J., 262, 419, "A ~300 Mpc Void of Rich Clusters of Galaxies?"

Baity, W.A., R.E. Rothschild, R.E. Lingenfelter, W.A. Stein, P.L. Nolan, D.E. Gruber, F.K. Knight, J.L. Matteson, L.E. Peterson, F.A. Primini, A.M. Levine, W.H.G. Lewin, R.F. Mushotzky, and A.F. Tennant, 1981, Ap. J., 244, 429, "Centaurus A (NGC 5128) at 2 keV-2.3 MeV: HEAO 1 Observations and Implications".

Balick, Bruce, and Timothy M. Heckman, 1982, Ann. Rev. Astr. Ap., 20, 431, "Extranuclear Clues to the Origin and Evolution of Activity in Galaxies".

Bergland, Glenn D., 1969, IEEE Trans. Audio, AU-17, 138, "A Radix-Eight Fast Fourier Transform Subroutine for Real-Valued Series".

Bevington, P.R., 1969, Data Reduction and Error Analysis for the Physical Sciences, (McGraw-Hill:New York).

Blumenthal, George, R., William C. Keel, and Joseph S. Miller, 1982,

Ap. J., 257, 499, "X-ray, optical, and radio properties of QSOs".

Boldt, E., U. Desai, S. Holt, P. Serlemitsos, 1969, Nature, 224, 677.

Boldt, Elihu, 1971, Bull. APS, 224, 534, "Nonthermal Background radiation $10^8$–$10^{18}$ Hz", (abstract).

_____, 1974, in McDonald & Fichtel (eds.), High Energy Particles and Quanta in Astrophysics, (MIT Press:Cambridge).

_____, 1981a, J. Wash. Acad. Sci, 71, (Proceedings of the UHURU Memorial Symposium), 24. "The High Energy Astronomy Observatory: HEAO 1".

_____, 1981b, Comments Ap., 9, 97, "The Cosmic X-ray Background".

Boldt, Elihu, and Darryl Leiter, 1981, Nature, 290, 483, "Precursor active galaxies and the cosmic X-ray background."

Bonoli, F., A. Braccesi, B. Marano, R. Merighi, and V. Zitelli, 1980, Astron. Ap., 90, L10, "On the Nature of the Faint (B ≈ 20) Ultraviolet Excess Objects and the Problem of the X-ray Background".

Bookbinder, J., L.L. Cowie, J.H.Krolik, J.P. Ostriker, M.J. Rees, 1980, Ap. J., 237, 647, "The Contribution of Young Galaxies to the X-ray Background".

Bothun, Gregory D., Jeremy Mould, Timothy Heckman, Bruce Balick, Robert A. Schommer, Jerome Kristian, 1982, Astron. J., 87, 1621, "0351+026: A QSO spawned by interacting galaxies?"

Boughn, Stephen P., Edward S. Cheng, and David T. Wilkinson, 1981, Ap. J. (Lett.), 243, L113, "Dipole and Quadropole Anisotropy of the 2.7 K Radiation".

Box, G.E.P., and G.M. Jenkins, 1976, Time Series Analysis: Forecasting and Control (Holden-Day:San Francisco).

Boynton, Paul E., and R.B. Partridge, 1973, Ap. J., 181, 243, "Fine-Scale Anisotropy of the Microwave Background: An Upper Limit at λ = 3.5 Millimeters".

Braccesi, A., V. Zitelli, F. Bonoli, L. Formiggini, 1980, Astron. Ap., 85, 80.

Bracewell, R.N., 1978, The Fourier Transform and its Applications, (McGraw Hill:New York).

Brecher, K., and P. Morrison, Phys. Rev. Lett., 23, 802.

Brecher, Kenneth, 1973, Ap. J., 181, 155, "On Compton Models of the Isotropic X-ray Background".

Brown, R.G., and R.J. Gould, 1970, Phys, Rev. D., 1, 2252.

Brück, H.A., Coyne, and Longair (eds.), 1982, "Cosmology and Fundamental Physics", Pontificiae Acad. Scientiarum Scripta, 48.

Carr, B.J., 1979, Mon. Not. R. astr. Soc., 189, 123, "The contribution of accreting black holes to the background radiation density".

_____, 1980, Nature, 284, 326, "Can pregalactic black holes produce the hard X-ray background?"

Cash, W., 1976, Astron. Ap., 52, 307.

_____, 1979, Ap. J., 228, 939.

Cavaliere, A., L. Danese, G. de Zotti, and A. Franceschini, 1979, Astron. Ap., 79, 169, "Constraints to the Non-thermal Source Synthesis of the X-ray Background".

_____, 1980, Astron. Ap., 85, L9, "X-ray Background and Discrete, Evolving Sources".

_____, 1981, Astron. Ap., 269, "Constraints to the QSO Contribution to the X-ray Background".

Cavallo, G., H.M. Horstman, E. Moretti-Horstman, 1977, Astron. Ap., 59, 405, "On the Measurement of the Diffuse X- and γ-Background as Seen by Non-Directional Detectors".

Chanan, Gary A., Bruce Margon, and Ronald A. Downes, 1981, Ap. J. (Lett.), 243, L5, "Optical Identification of Serendipitous Einstein sources: 19 X-ray-selected AGN".

Chanan, Gary A., 1982, Ap. J., 252, 32, "Do Quasars Have Cosmologically Long Lifetimes?".

Cheney, Janet E., and Michael Rowan-Robinson, 1981a, Mon. Not. R. astr. Soc., 195, 497, "The Interpretation of Optical Counts of Quasars".

_____, 1981b, Mon. Not. R. astr. Soc., 197, 313, "The Contribution of Quasars to the 2 keV-100 MeV Background Radiation and the X-ray Source Counts at 2 keV".

Cheng, Edward S., Peter R. Saulson, David T. Wilkinson, and Brian E. Corey, 1979, Ap. J. (Lett.), 232, L139, "Large Scale Anisotropy in the 2.7 K Radiation".

Chincarini, Guido, and Herbert J. Rood, 1979, Ap. J., 230, 648, "Space Distribution of Sc Galaxies: Clues to Large-Scale Structure".

Condon, J.J., 1974, Ap. J., 188, 279, "Confusion and Flux-density Error Distributions".

Condon, J.J., and L.L. Dressel, 1978, Ap. J., 222, 745, "Statistical Counts of Faint Sources at 2380 MHz".

Condon, J.J., S.L. O'Dell, J.J. Puschell, and W.A. Stein, 1981, Ap. J., 246, 624, "Radio Emission from Bright, Optically Selected Quasars".

Cordova, F.A., and K.O. Mason, 1982, in Lewin and van den Heuvel (eds.), Accretion Driven Stellar X-ray Sources, (Cambridge Univ. Press: Cambridge).

Cowsik, R., and E.J. Kobetich, 1972, Ap. J., 177, 585, "Comment on Inverse Compton Models for the Isotropic X-ray Background and Possible Thermal Emission from a Hot Thermal Gas".

Crawford, D.E., D.L. Jauncey, and H.S. Murdoch, 1970, Ap. J., 162, 405.

Davis, Marc, John Tonry, John Huchra, and David W. Latham, 1980, Ap. J. (Lett.), 238, L113, "On the Virgo Supercluster and the Mean Mass Density of the Universe".

Davis, Marc, John Huchra, David W. Latham, and John Tonry, 1982, Ap. J., 253, 423, "A Survey of Galaxy Redshifts. II. The Large Scale Space Distribution".

de Ruiter, H.R., and E.J. Zuiderwijk, 1982 Astron. Ap., 105, 254, "Quasar-generating Superclusters: An Explanation for a Clumpy Quasar Sky?"

de Zotti, G., 1982, Acta Cosmol., 11, 65, "Discrete Source Contributions to the X-ray Background".

de Zotti, G., E.A. Boldt, A. Cavaliere, L. Danese, A. Franceschini, F.E. Marshall, J.H. Swank, and A.E. Szymkowiak, 1982, Ap. J., 253, 47, "On Synthesis of the X-ray Background with Power-Law Sources".

Doxsey, R., H. Bradt, J. McClintock, L. Petro, R. Remilard, G. Ricker, D. Schwartz, and K. Wood, 1983, Ap. J. (Lett.), 264, L43,

"H0323+022: A Puzzling High-Latitude X-Ray/Optical/Radio Source".

Dube, Roger R., William C. Wickes, and David T. Wilkinson, 1979, Ap. J., 232, 333, "Upper Limit on Extragalactic Background Light".

Elvis, M. and L. Van Speybroeck, 1982 Ap. J. (Lett.), 257, L51, "The Nuclear X-ray Source in M81".

Elvis, M., A. Soltan, and W. Keel, 1983, Ap. J., (Submitted).

Fabbri, R., I. Guidi, F. Melchiorri, and V. Natale, 1980, Phys. Rev. Lett., 44, 1563, (Erratum in 45, 401), "Measurement of the Cosmic-Background Large-Scale Anistropy in the Millemetric Region".

Fabian, A.C. and P.W. Sanford, 1971, Nature (Phys. Sci.), 234, 20.

Fabian, A.C., 1972, Nature (Phys Sci.), 237, 19, "Analysis of X-ray Background Fluctuations".

_____, 1975, Mon. Not. R. astr. Soc., 172, 149, "X-ray Source Counts and the X-ray Background from UHURU".

Fabian, A.C., and M.J. Rees, 1978, Mon. Not. R. astr. Soc., 185, 109, "Predictions and Strategies for Deep X-ray Surveys".

Fabian, A.C., R.S. Warwick, and J.P. Pye, 1980, Physica Scripta, 21, 650, "Large Scale Shearing of the Universe from ARIEL V X-ray Data".

Fabian, A.C., 1981, in Ramaty and Jones (eds.), Tenth Texas Symposium on Relativistic Astrophysics, Ann. N.Y. Acad. Sci., 375, 235, "The X- and Gamma-Ray Backgrounds".

Fabian, A.C., and A.K. Kembhavi, 1982, in Heeshcen and Wade (eds.), IAU Symposium 97, Extragalactic Radio Sources, 453, "The Intergalactic Medium".

Felten, J.E., and P. Morrison, 1966, Ap. J., 146, 686.

Felten, James E., 1977, Astron. J., 82, 861, "Study of the Luminosity Function for Field Galaxies".

Ferland, G.J., and R.F. Mushotzky, 1982, Ap. J., 262, 564, "Broad Line Region Clouds and the Absorbing Material in NGC 4151".

Fichtel, Carl E., and Jacob I, Trombka, 1981, Gamma Ray Astrophysics, NASA Publication SP-453, (Government Printing Office: Washington D.C.).

Field, George B., and Stephen C. Perrenod, 1977, Ap. J., 215, 717, "Constraints on a Dense Hot Intergalactic Medium".

Field, George B., 1978, "Intergalactic Matter and the Evolution of Galaxies", Harvard-Smithsonian Center for Astrophysics Preprint NO. 1132, also in the Proceedings of the Astronomische Gesellschaft, Schwarzschild Lecture of 5 Sept. 1978, Tubingen, Germany. Fireman, E.L., 1974, Ap. J., 187, 57.

Franceschini, A., 1982, Ap. Space Sci., 86, 3, "Source Counts at and Beyond the Confusion Limit".

Fried, P.M., J.A. Nousek, W.T. Sanders, and W.L. Kraushaar, 1980, Ap. J., 242, 987, "The Soft X-ray Diffuse Background and the Structure of the Local Interstellar Medium".

Frootmig, Hurling, in Lig Lury, Jr. (ed.), The Hitch Hiker's Guide to the Galaxy, (Megadodo:UMi β), The Cover.

Giacconi, Riccardo, Herbert Gursky, Frank R. Paolini, and Bruno B. Rossi, 1962, Phys. Rev. Lett., 9, 439, "Evidence for X-rays from Sources Outside the Solar System".

Giacconi, R., J. Bechtold, G. Branduardi, W. Forman, J.P. Henry, C. Jones, E. Kellogg, H. van der Laan, W. Liller, H. Marshall, S.S. Murray, J. Pye, E. Schreir, W.L.W. Sargent, F. Seward, and H. Tananbaum, 1979, Ap. J. (Lett.), 234, L1, "A High-Sensitivity X-ray Survey Using the Einstein Observatory and the Discrete Source Contribution to the Extragalactic X-ray Background".

Giacconi, Ricardo, 1981, in Ramaty and Jones (eds.), Tenth Texas Symposium on Relativistic Astrophysics, Ann. N.Y. Acad. Sci., 375, 210, "Extragalactic X-ray Astronomy with the Einstein Observatory".

Gorenstein, M.B. and G.F. Smoot, 1981, Ap. J., 244, 361, "Large-Angular-Scale Anisotropy in the Cosmic Background Radiation".

Gorenstein, P., E. Kellogg, and H. Gursky, 1969, Ap. J., 156, 315.

Gould, Rober J., 1982, Ap. J., 258, 131, "The Temperature of Thermal X-ray and γ-ray Sources".

Grandi, Steven A., 1982, Ap. J., 255, 25, "The 3000 Å Bump in Quasars".

Green, Richard F., and Maarten Schmidt, 1978, Ap. J. (Lett.), 220, L1, "Evidence for Nonuniform Radial Distribution of Quasars, Regardless of the Nature of Their Redshifts".

References

Gregory, Stephen A., Laird A. Thompson, William G. Tifft, 1981, Ap. J., 243, 411, "The Perseus Supercluster".

Griffiths R.E., T. Maccacaro, S.S. Murray, and R. Giacconi, 1981, Space Sci. Rev., 30, 67, "Optical Identification of X-ray Sources in the Einstein Observatory Medium and Deep Surveys".

Halpern, Jules, 1982, Ph.d. Dissertation, Harvard University Department of Astronomy.

Hamilton, Andrew J.S., Craig L. Sarazin, and Roger A. Chevalier, 1983, Ap. J. Suppl., 51, 115, "X-Ray Line Emission from Supernova Remnants. I. Models for Adiabatic Remnants".

Hawkins, M.R.S., and N.J. Stewart, 1981, Ap. J., 251, 1, "The $V/V_m$ Test for Quasars: A New Interpretation".

Henry, J. Patrick, Andrezej Soltan, Ulrich Briel, and James E. Gunn, 1982, Ap. J., 262, 1. "Evolution of the Cluster X-ray Luminosity Function Slope".

Hermsen, W., K. Bennett, G.F. Bignami, J.B.G.M. Bloemen, R. Buccheri, P.A. Caraveo, G. Kenbach, J.L. Mesnou, H.A. Mayer-Hasselwander, and J.A. Paul, 1981, 17th ICRC (Paris), 1, 230.

Hoag, A.A., and M.G. Smith, 1977, Ap. J., 217, 362.

Hogan, Craig, and David Layzer, 1977, Ap. J., 212, 360, "Origin of the X-ray Background".

Holt, S.S., E.A. Boldt, P.J. Serlemitsos, S.S. Murray, R. Giacconi, and E.M. Kellog, 1974, Ap. J. (Lett.), 188, L97, "On the Nature of the Unidentified High-latitude UHURU Sources".

Holt, S.S, R.F. Mushotzky, R.H. Becker, E.A. Boldt, P.J. Serlemtisos, A.E. Szymkowiak, and N.E. White, 1980, Ap. J. (Lett.), 241, L13, "X-ray Spectral Constraints on the Broad-line Cloud Geometry of NGC 4151".

Holt, Stephen S. (ed.), 1981, X-ray Astronomy in the 1980s, NASA Technical Memorandum 83848, Goddard Space Flight Center.

Holt, Stephen S., and Richard McCray, 1982, Ann. Rev. Astron. Ap., 20, 323, "Spectra of Cosmic X-ray Sources".

Holt, S.S, J.H. Krolik, R.F. Mushotzky, R. Petre, 1983, in preparation.

Huchra, J.P., 1977, Ap. J. Suppl., 35, 171.

Huchra, J.P., and W.L.W. Sargent, 1973, _Ap. J._, _186_, 433.

Huchra, J.P., W.F. Wyatt, and M. Davis, 1982, _Astron. J._ _87_, 1628, "New Bright Seyfert Galaxies".

Iwan, DeAnne, F.E. Marshall, E.A. Boldt, R.F. Mushotzky, R.A. Shafer, and A. Stottlemyer, 1982, _Ap. J._, _260_, 111, "A Large Scale Height Galactic Component of the Diffuse 2-60 keV Background".

Jauncey, D.L., (ed.), 1977, _Radio Astronomy and Cosmology, IAU Symposium No. 74_, (Reidel:Amsterdam).

Jenkins, Gwilym M., and Donald G. Watts, 1968, _Spectral Analysis and Its Applications_, (Holden-Day:San Francisco).

Johnson, M.W., R.G. Cruddace, M.P. Ulmer, M.P. Kowalsky, and K.S. Wood, 1983, _Ap. J._, in press.

Jones, T.W., S.L. O'Dell, and W.A. Stein, 1974, _Ap. J._, _188_, 353.

Kaiser, N., 1982, _Mon. Not. R. astr. Soc._, _198_, 1033, "Background Radiation Fields as a Probe of the Large-scale Matter Distribution in the Universe".

Kazanas, D., and R.J. Protheroe, 1983, _Nature_, in press, "Active Galaxies and the Diffuse Gamma-Ray Background".

Kazanas, D., and R.A. Shafer, 1983, in Burns, Harding, and Ramati (ed.s), Conferenece on e+e- in Astrophysics.

Kembhavi, A.K. and A.C. Fabian, 1982, _Mon. Not. R. astr. Soc._, _198_, 921, "X-ray Quasars and the X-ray Background".

Kirshner, Robert P., Augustus Oemler, Jr., and Paul L. Schechter, 1979, _Astron. J._, _84_, 951.

Kirshner, Robert P., Augustus Oemler, Jr., Paul L. Schechter, and Stephen A. Schectman, 1981, _Ap. J. (Lett.)_, _248_, L57, "A Million Cubic Megaparsec Void in Bootes?"

Koo, D.C., and R.G. Kron, 1982, _Astron. Ap._, _105_, 107, "QSO Counts: A Complete Survey of Stellar Objects to B = 23".

Kriss, Gerard A., and Claude R. Canizares, 1982, _Ap. J._, _261_, 51, "Optical and X-ray Properties of X-ray Selected AGN".

Ku, W. M.-H., 1980, in Wayman (ed.), _Highlights of Astronomy_, _5_, (Reidel:Dordrecht), 677.

Lawrence A., and M.S. Elvis, 1982, Ap. J., 256, 410.

Lampton, M., B. Margon, and S. Bowyer, 1976, Ap. J., 208, 177.

Leiter, Darryl, and Elihu Boldt, 1982, Ap. J., 260, 1, "Spectral Evolution of Active Galactic Nuclei: A Unified Description of the X-ray and Gamma-Ray Backgrounds".

Lightman, A., P. Hertz, and J.E. Grindlay, 1980, Ap. J., 242, 492.

Lightman, Alan P., 1982, Ap. J., 253, 842, "Relativistic Thermal Plasmas: Pair Processes and Equilibria".

Long, Knox S., and Leon P. Van Speybroeck, 1983, in Lewin and Van den Heauvel (eds.), Accretion Driven Stellar X-ray Sources, (Cambridge University Press), in press, "X-ray Emission from Normal Galaxies".

Longair, M.S., 1978, in Gunn, Longair, and Rees, Observational Cosmology, (Geneva Observatory:Sauverny Switzerland).

Longmore, A.J. and R.M. Sharples, 1982, Mon. Not. R. astr. Soc., 201, 111, "Infrared Observations of Early-Type Galaxies with Dust Lanes".

Maccacaro, Tommaso, E.D. Feigelson, M. Ferner, R. Giacconi, I.M. Gioia, R.E. Griffiths, S.S. Murray, G. Zamorani, J. Stocke, and J. Liebert, 1982, Ap. J., 253, 504, "A Medium Sensitivity X-Ray Survey Using the Einstein Observatory: The Log N-log S Relation for Extragalactic X-ray Sources".

Maccacaro, T., Y. Avni, I.M. Gioia, P. Giommi, J. Lieger, R.E. Griffiths, J. Stocke, and J. Danziger, 1983, Ap. J. (Lett.), 266, L73.

Maccacaro, T., and I.M. Gioia, 1983b, in Abell and Chincarini (eds.), Early Evolution of the Universe and its Present Structure, IAU Symposium 104, (Reidel:Dordrecht), in press, "X-ray Sources of Cosmological Relevance".

Maccagni, D. and M. Tarenghi, Ap. J., 243, 42, "X-ray Observations of Six BL Lacertae Fields".

Malkan, Matthew A., and Wallace L.W. Sargent, 1982, Ap. J., 254, 22, "The Ultraviolet Excess of Seyfert 1 Galaxies and Quasars".

Margon, B., S. Bowyer, T.W. Jones, Arthur Davidson, Keith O. Mason, and Peter W. Sanford, 1976, Ap. J., 207, 359, "An X-ray Survey of BL

Lacertae Objects".

Margon, Bruce, Gary A. Chanan, and Ronald A. Downes, 1982, Ap. J. (Lett.), 253, L7, "The Luminosity of Serendipitous X-ray QSOs".

Marscher, A.P., F.E. Marshall, R.F. Mushotzky, W.A. Dent, T.J. Balonek, and M.F. Hartman, 1979, Ap. J., 233, 498, "Search for X-ray Emission from Bursting Radio Sources".

Marscher, Alan P., John J. Broderick, 1982, in Heeschen and Wade (eds.), IAU Symposium 97: Extragalactic Radio Sources, (Reidel:Dordrecht), 359, "Superluminal Motion in NRAO 140 and a Possible Future Method for Constraining $H_o$ and $q_o$".

Marshall, F.E., E.A. Boldt, S.S. Holt, R.F. Mushotzky, S.H. Pravdo, R.E. Rothschild, P.J. Serlemitsos, 1979, Ap. J. Suppl., 40, 657.

Marshall, F.E., E.A. Boldt, S.S. Holt, R. Miller, R.F. Mushotzky, L.A. Rose, R. Rothschild, P. Serlemitsos, 1980, Ap. J., 235, 4.

Marshall, F.E., 1981, private communication, "Low galactic latitude A-2 source positions".

Marshall, H.L., H. Tananbaum, G. Zamorani, J.P. Huchra, A. Braccessi, and V. Zitelli, 1983a, Ap. J., submitted.

Marshall, H.L., Y. Avni, H. Tananbaum, and G. Zamorani, 1983b, Ap. J., submitted, "Analysis of Complete Quasar Samples to Obtain Parameters of Luminosity and Evolution Functions".

Mathez, G., and L. Nottale, 1982, Astron. Ap., 113, 336, "On the Behavior of QSO Space Density Beyond z = 3.5".

Matilsky, T., H. Gursky, E. Kellogg, H. Tananbaum, S. Murray, R. Giacconi, 1973, Ap. J., 181, 753.

Matteson, J.L., D.E. Gruber, P.L. Nolan, L.E. Peterson, and R.L. Kinzer, 1979, B.A.A.S., 11, 653, abstract.

Mätzler, Christian, Taeil Bai, Carol Jo Crannell, and Kenneth J. Frost, 1978, Ap. J., 223, 1058, "Adiabatic Heating in Impulsive Solar Flares".

Maxon, Stephen, 1972, Phys. Rev. A, 5, 1630, "Bremsstrahlung Rate and Spectra from a Hot Gas (Z=1)".

McHardy, I.M., A. Lawrence, J.P. Pye, and K.A. Pounds, 1981, Mon. Not.

R. astr. Soc., 197, 893, "The Ariel V (3A) catalogue of X-ray sources -- II. Sources at high galactic latitude (|b| > 10°)".

McIlwain, C.E., 1961, J. Geophys. Rsch., 66, 3681.

McKee, C.F., 1974, Ap. J., 188, 335.

McKee, Christopher, F., 1980, Phys. Scr., 21, "The Contribution of Emission from a Hot Intergalactic Medium to the X-ray Background".

McKee, J., R. Mushotzky, E. Boldt, S. Holt, F.E. Marshall, S. Pravdo, and P. Serlemitsos, 1980, Ap. J., 242, 843.

Meier, David L., Marie-Helene Ulrich, R. Fanti, I. Gioia, and C. Lari, 1979, Ap. J., 229, 25, "A New Determination of the Luminosity Function of Radio Galaxies and an Investigation of the Evolutionary Properties of the Radio Galaxy Population in the Recent Past".

Mitchell, R., and R. Mushotzky, 1980, Ap. J., 236, 730, "HEAO A-2 Observations of the X-ray Spectra of the Centaurus and A1060 Clusters of Galaxies".

Murdoch, H.S., D.E. Crawford, and D.L. Jauncy, 1973, Ap. J., 183, 1.

Mushotzky, Richard Fred, 1976, Ph.D. dissertation in physics, University of California at San Diego.

_____, 1977, Nature, 265, 225.

Mushotzky, R.F., P.J. Serlemitsos, B.W. Smith, E.A. Boldt, and S.S. Holt, 1978, Ap. J., 225, 21, "OSO 8 X-Ray Clusters of Galaxies. I. Observations of Twenty Clusters: Physical Correlations".

Mushotzky, R.F., E.A. Boldt, S.S. Holt, S.H. Pravdo, P.J. Serlemitsos, J.H. Swank, and R.E. Rothschild, 1978, Ap. J. (Lett.), 226, L65, "X-Ray Observations of BL Lac Objects".

Mushotzky, R.F., F.E. Marshall, E.A. Boldt, S.S. Holt, and P.J. Serlemitsos, 1980, Ap. J., 235, 377, "HEAO 1 Spectra of X-Ray Emitting Setfert 1 Galaxies".

Mushotzky, R.F., 1982, Ap. J., 256, 92, "The X-ray Spectrum and Time Variability of Narrow Emission Line Galaxies".

_____, 1983, Phys. Scr., in press.

Nousek, J.A., D.M. Fried, W.T. Sanders, W.L. Kraushaar, 1982, Ap. J.,

258, 83, "On the Origin of the 1 keV Diffuse X-Ray Background".

Oort, J.H., H. Arp, H. De Ruiter, 1981, Astron. Ap., 95, 7, "Evidence for the Location of Quasars in Superclusters".

Osmer, Patrick S., 1981, Ap. J., 247, 762, "The 3-Dimensional Distribution of Quasars in the CTIO Surveys".

_____, 1982, Ap. J., 253, 28, "Evidence for a Decrease in the Space Density of Quasars at z ≳ 3.5".

Paresce, Francesco, Christopher F. McKee, and Stuart Bowyer, 1980, Ap. J., 240, 387, "Galactic and Extragalactic Contributions to the Far Ultraviolet Background".

Partridge, R.G., 1980, Phys. Scr., 21, 624.

Patterson, Joseph, and J.E. Steiner, 1983, Ap. J. (Lett.), 264, L61, "H 2215-086: King of the DQ Herculis Stars".

Pearson, T.J., S.C. Unwin, M.H. Cohen, R.P. Linfield, A.C.S. Readhead, G.A. Seielstad, R.S. Simon, and R.C. Walker, 1981, Nature, 290, 365.

Peebles, P.J.E., 1974, Ap. J., 32, 197.

_____, 1980, The Large-Scale Structure of the Universe, (Princeton Univ.:Princeton).

Perrenod, S.C., 1980, Ap. J., 236, 373.

Peterson, Bruce, Ann Savage, David L. Jauncey, and Alan E. Wright, Ap. J. (Lett.), 260, L27, "PKS 2000-330: A Quasi Stellar Radio Source with a Redshift of 3.78".

Piccinotti, G., R.F. Mushotzky, E.A. Boldt, S.S. Holt, F.E. Marshall, P.J. Serlemitsos, and R.A. Shafer, 1982, Ap. J., 253, 485, "A Complete X-ray Sample of the High Latitude ($|b| > 20^{\circ}$) Sky from HEAO-1 A-2: Log N-Log S and Luminosity Functions".

Pravdo, Steven Howard, 1976, Ph.D. dissertation, Physics Department, University of Maryland, GSFC document X-661-76-280.

Pravdo, S.H., E.A. Boldt, F.E. Marshall, J. McKee, R.F. Mushotzky, B.W. Smith, and G. Reichert, 1979, Ap. J., 234 234, 1, "A Search for X-ray Emission from Rich Clusters, Extended Halos Around Clusters and Superclusters".

Pravdo, S.H., J.J. Nugent, J.A. Nousek, K. Jensen, A.S. Wilson, R.H. Becker, 1981, Ap. J., 251, 501, "Discovery of a Seyfert 1 Galaxy with an Unusually Soft X-Ray Spectrum".

Primini, F.A., B.A. Cooke, C.A. Dobson, S.K. Howe, A. Scheepmaker, W.A. Wheaton, W.H.G. Lewin, W.A. Baity, D.E. Gruber, J.L. Matteson, and L.E. Peterson, 1979, Nature, 278, 234.

Protheroe, R.J., A.W. Wolfendale, and J. Wdowczyk, 1980, Mon. Not. R. astr. Soc., 192, 445, "A Galactic Component of the Diffuse X-ray Flux in the Range 2-7 keV".

Protheroe, R.J., and D. Kazanas, 1982, Ap. J., 265, 620, "On the Origin of Relativistic Particles and Gamma Rays in Quasars".

Pye, J.P., and R.S. Warwick, 1979, Mon. Not. R. astr. Soc., 187, 905, "X-ray Source Counts at High Galactic Latitude from measurements of fluctuations in the X-ray background".

Rees, Martin J., 1980, in Abell and Peebles (eds.), I.A.U. Symposium 92: Objects at High Redshifts, 207.

Reichert, G., K.O. Mason, S.M. Lea, P.A. Charles, S. Bowyer, and S. Pravdo, 1981, Ap. J., 247, 803, "Low Energy X-ray emission from Five Galaxy Cluster Sources".

Reichert, G.A., K.O. Mason, J.R. Thorstensen, and S. Bowyer, 1982, Ap. J., 260, 437, "Optical Identification of Serendipitous Einstein X-ray Sources: Luminosity and Redshift Relations in the X-ray and Optical emission from Active Galaxies and Quasars".

Reichert, Gail Anne, 1982, Ph.D. dissertation in physics, University of California at Berkeley.

Riegler, G.R., P.C. Agrawal, and R.F. Mushotzky, 1979, Ap. J. (Lett.), 233, L47.

Rosner, R., Y. Avni, J. Bookbinder, R. Giacconi, L. Golub, F.R. Harndon Jr., C.W. Maxson, K. Topka, and G.S. Vaiana, 1981, Ap. J. (Lett.), 249, L5.

Rothschild, R., E. Boldt, S. Holt, P. Serlemitsos, G. Garmire, P. Agrawal, G. Riegler, S. Bowyer, and M. Lampton, 1979, Space Sci. Instrum., 4, 269, "The Cosmic X-ray Experiment Aboard HEAO-1".

Rothschild, R.E., R.F. Mushotzky, W.A. Baity, D.E. Gruber, and J.L. Matteson, 1983, Ap. J., in press, "2-165 keV Observations of Active Galaxies and the Diffuse Background".

Rowan-Robinson, M., and A.C. Fabian, 1974, Mon. Not. R. astr. Soc., 167, 419.

Rybicki, George B., and Alan P. Lightman, 1979, Radiative Processes in Astrophysics, (Wiley:New York).

Scheuer, P.A. G., 1957, Proc. Cambridge Phil. Soc., 53, 764.

_____, 1974, Mon. Not. R. astr. Soc., 167, 419, "Fluctuations in the X-ray Background".

Schmidt, Maarten, 1968, Ap. J., 151, 393.

Schmidt, M., and R.F. Green, 1982, in Brück, Coyne, and Longair (eds.).

_____, 1983, Ap. J., in press, "Quasar Evolution Derived from the Palomar Quasar Survey and Other Complete Quasar Surveys".

Schwartz, D.A., 1970, Ap. J., 162, 439.

Schwartz, D.A., E.A. Boldt, S.S. Holt, P.J. Serlemitsos, and R.D. Bleach, 1971, Nature Phys. Sci., 233 110.

Schwartz, D., and L. Peterson, 1974, Ap. J., 190, 297.

Schwartz, D.A., S.S. Murray, and H. Gursky, 1976, Ap. J., 204, 315, "A Measurement of Fluctuations in the X-ray Background by UHURU".

Schwartz, D.A., 1978, Ap. J., 220, 8, "X-ray Luminosity Function of Abell Clusters".

_____, 1980, Phys. Scr., 21, 644, "Irregularities in the X-ray Background".

Schwartz, Daniel A., and William H.-M. Ku, 1983, Ap. J., in press, (March 15), "Studies of BL Lac Objects with the Einstein X-ray Observatory: The Absolute Volume Density".

Schwarz, J., J.P. Ostriker, and A. Yahil, 1975, Ap. J., 202, 1.

Serlemitsos, P.J., 1981, in Holt (ed.), X-ray Astronomy in the 1980s, NASA Technical Memorandum 83848, 441.

Setti, G., and L. Woltjer, 1973, in Bradt and Giacconi (eds.), X-ray and Gamma-ray Astronomy, (Reidel:Dordrecht).

_____, 1979, Astron. Ap., 76, L1.

_____, 1982, in Brück, Coyne, and Longair (eds.), "The Origin of the X- and Gamma-ray Backgrounds".

Shapiro, S.L., A.P. Lightman, and D.M. Eardley, 1976, Ap. J., 204, 187.

Shafer, R.A., E.A. Boldt, S.S. Holt, P.J. Serlemitsos, 1977, Bull.A.A.S., 9 616, "Fluctuations in the High Galactic Latitude X-ray Sky" (abstract).

Shectman, Stephen A., 1974, Ap. J., 188, 233, "The Small-scale Anisotropy of the Cosmic Light".

Sherman, Richard D., 1979, Ap. J., 232, 1, "Theory of the Intercluster Medium".

_____, 1980, Ap. J., 238, 794, "Quasi-Stellar Objects in the IGM: Source for the Cosmic X-Ray Background?"

_____, 1982, Ap. J., 256, 370, "Cosmological Constraints on Hot Plasma in a Closed Universe".

Shulz, M., and L.J. Lanzerotti, 1974, Particle Diffusion in the Radiation Belts, (Springer-Verlag:New York).

Silk, Joseph, 1973, Ann. Rev. Astron. Ap., 11, 269, "Diffuse X- and Gamma-Radiation".

Silk, Joseph, and Michael L. Wilson, 1981, Ap. J. (Lett.), 244, 137, "Large-Scale Anisotropy of the Cosmic Microwave Background Radiation".

Spitzer, Lyman, Jr., 1962, Physics of Fully Ionized Gases, (Wiley:New York).

Stocke, John T., and Stephen C. Perrenod, 1981, Ap. J., 245, 375, "Quasar Number Density Evolution".

Stocke, John, James Liebert, T. Maccacaro, R.E. Griffiths, and J.E. Steiner, 1982, Ap. J., 252, 69, "Discovery of a Narrow Line Quasar".

Suterland, P.G., M.C. Weisskopf, and S.M. Kahn, 1978, Ap. J., 219, 1029, "Short-term Variabiltiy of Cygnus X-1. II".

Swank, J.H., N.E. White, S.S. Holt, and R.H. Becker, 1981, Ap. J., 246, 208.

Tananbaum, H., Y. Avni, G. Branduardi, M. Elvis, G. Fabbiano, E.

Feigelson, R. Giacconi, J.P. Henry, J.P. Pye, A. Soltan, and G. Zamorani, 1979, Ap. J. (Lett.), 234, L9, "X-ray Studies of Quasars with the Einstein Observatory".

Tananbaum, H. J.F.C. Wardle, G. Zamorani, and Y. Avni, 1983, Ap. J., submitted, "X-ray Studies of Quasars with the Einstein Observatory (III): The 3CR Sample".

Tennant, A.F., R.F. Mushotzky, E.A. Boldt, and J.H. Swank, 1981, Ap. J., 251, 15, "Rapid X-ray Variability in the Seyfert Galaxy NGC 6814".

Terrell, James, 1977, Am. J. Phys., 45, 869, "The Luminosity Distance Equation in Friedmann Cosmology".

Tonry, John L, and Marc Davis, 1981, Ap. J., 246, 680, "Velocity Dispersions of Elliptical and SO Galaxies. II. Infall of the Local Group to Virgo".

Trombka, J.I., C.S. Dyer, L.G. Evans, M.J. Bielefeld, S.M. Seltzer, and A.E. Metzger, 1977, Ap. J., 212, 925, "Reanalysis of the Apollo Cosmic Gamma-ray Spectrum in the 0.3 to 10 MeV Region".

Tucker, Wallace H., 1975, Radiation Processes in Astrophysics, (MIT Press:Cambridge Mass.).

Tully, R. Brent, 1982, Ap. J., 257, 389, "The Local Supercluster".

Turner, Edwin L, and Margeret J. Geller, 1980, Ap. J., 236, 1, "The Correlation of the Cosmic X-ray Background with the Light of Galaxies".

Urry, C. Megan, and Richard F. Mushotzky, 1982, Ap. J., 253, 38, "PKS 2155-304: Relativistically Beamed Synchrotron Radiation from a BL Lacertae Object".

Urry, C.M., R.F. Mushotzky, Y. Kondo, K.R.H. Hackney, and R.L. Hackney, 1982, Ap. J., 261, 12, "Ultraviolet and X-ray Observations of the BL Lacertae Object PKS 0548-322".

Veron, P., 1979, Astron. Ap., 78, 46, "The Luminosity Funciton of Seyfert 1 Galaxy Nuclei and BL Lac Objects, and the X-ray Background".

Veron, P. and M.P. Veron, 1982, Astron. Ap., 105, 405, "On the Quasar Surface Density".

Wall, J.V., P.A.G. Scheuer, I.I.K. Pauliny-Toth, adn A. Witzel, 1982, Mon. Not. R. astr. Soc., 198, 221, "A Revised Statistical Estimate

of the Counts of Faint Radio Sources at 5 GHz".

Warwick, R.S. and J.P. Pye, 1978, Mon. Not. R. astr. Soc., 183, 169.

Warwick, R.S., J.P. Pye, and A.C. Fabian, 1980, Mon. Not. R. astr. Soc., 190, 243.

Weinberg, Steven, 1972, Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity, (Wiley:New York).

Weiss, Rainer, 1980 Ann. Rev. Astron. Ap., 18, 489.

White, Simon D.M., and Joseph Silk, 1979, Ap. J., 231, 1, "The Growth of Aspherical Structure in the Universe: Is the Local Supercluster an Unusual System?"

White, Simon D.M., 1980, Phys. Scr., 21, 640, "Dynamical Inferences from the Dipole Anisotropy in the Microwave Background".

Wilkinson, David T., 1983, in Abell and Chincarini (eds), I.A.U. Symposium 104: Early Evolution of the Universe and its Present Structure, (Reidel:Dordrecht).

Wilks, S.S., 1938, Ann. Math. Stat., 9, 60.

_____, 1963, Mathematical Statistics, (Princeton University Press:Princeton), see Chapter 13.

Wilson, M.L., and Joseph Silk, 1981, Ap. J., 243, 14, "On the Anisotropy of the Cosmological Background Matter and Radiation Distribution. I. The Radiation Anisotropy in a Spatially Flat Universe".

Wolfe, A.M., 1970, Ap. J., 159,

Woltjer, L. and G. Setti, 1982, in Brück, Coyne, and Longair (eds.), "Quasars in the Universe".

Worrall, D.M., F.E. Marshall, and E.A. Boldt, 1979, Nature, 281, 127, "A Limit to the X-ray Luminosity of Nearby Normal Galaxies".

Worrall, D.M., R.F. Mushotzky, E.A. Boldt, S.S. Holt, and P.J. Serlemitsos, 1979, Ap. J., 232, 683, "The X-ray Spectrum of 3C273".

Worrall, D.M., E.A. Boldt, S.S. Holt, and P.J. Serlemitsos, 1980, Ap. J., 240, 421, "The X-ray Spectrum of QSO 0241+622".

_____, 1981, Ap. J., 243, 53, "Variable X-ray Spectra of BL Lac Objects: HEAO 1 Observations of PKS 0548-322 and 2A 1219+305".

Worrall, D.M., F.E. Marshall, E.A. Boldt, and J.H. Swank, 1982, Ap. J., 255, 111.

Worrall, D.M., and F.E. Marshall, 1983, Ap. J., 267, 691. "Stellar Contributions to the Hard X-ray Galactic Ridge".

Wright, Edward L., 1979, Ap. J., 232, 348, "Distortion of the Microwave Background by a Hot Intergalactic Medium".

Zamorani, G., J.P. Henry, T. Maccacaro, H. Tananbaum, A. Soltan, Y. Avni, J. Lieber, J. Stocke, P.A. Strittmatter, R.J. Weymann, M.G. Smith, J.J. Condon, 1981, Ap. J., 245, 357., "X-ray Studies of Quasars with the Einstein Observatory. II.".

Zamorani, G., 1982a, Ap. J. (Lett.), 260, L31, "The Redshift Distribution of X-ray Selected Quasars, is it Really Anomalous?"

_____, 1982b, in Wolfendale (ed.), Progress in Cosmology, 1981 Oxford International Symposium, (Reidel:Dordrecht), 203, "The Origin of the X-ray Background".

# BIBLIOGRAPHIC DATA SHEET

| 1. Report No.<br>TM 85029 | 2. Government Accession No. | 3. Recipient's Catalog No. |
|---|---|---|
| 4. Title and Subtitle<br><br>Spatial Fluctuations in the Diffuse Cosmic X-Ray Background | | 5. Report Date<br>May 1983 |
| | | 6. Performing Organization Code<br>661 |
| 7. Author(s)<br>Richard Arrick Shafer | | 8. Performing Organization Report No. |
| 9. Performing Organization Name and Address<br>Code 661<br>Laboratory for High Energy Astrophysics<br>NASA/Goddard Space Flight Center<br>Greenbelt, Maryland 20771 | | 10. Work Unit No. |
| | | 11. Contract or Grant No. |
| | | 13. Type of Report and Period Covered |
| 12. Sponsoring Agency Name and Address | | |
| | | 14. Sponsoring Agency Code |

16. Abstract

Studies of the bright, essentially isotropic, X-ray sky flux above 2 keV yield information on the universe at large distances. However, a definitive understanding of the origin of the flux is lacking. Some fraction of the total flux is contributed by active galactic nuclei and clusters of galaxies, but less than one percent of the total is contributed by the $\gtrsim$ 3 keV band resolved sources, which is the band where the sky flux is directly observed.

Parametric models of AGN (quasar) luminosity function evolution are examined. They are most constrained by the total sky flux. The acceptability of particular models hinges on assumptions currently not directly testable. The comparison with the _Einstein_ Observatory 1-3 keV low flux source counts is hampered by spectral uncertainties.

A tentative measurement of a large scale dipole anisotropy is consistent with the velocity and direction derived from the dipole in the microwave background. The impact of the X-ray anisotropy limits for other scales on studies of large-scale structure in the universe is sketched. Models of the origins of the X-ray sky flux are reviewed, and future observational programs outlined.

| 17. Key Words (Selected by Author(s)) | | 18. Distribution Statement | |
|---|---|---|---|
| 19. Security Classif. (of this report)<br>U | | 20. Security Classif. (of this page)<br>U | 21. No. of Pages | 22. Price* |