THE ASTROPHYSICAL JOURNAL, 336:896–919, 1989 January 15


# ANALYZING GAMMA-RAY BURST SPECTRAL DATA

THOMAS J. LOREDO
Department of Astronomy and Astrophysics, University of Chicago
AND
RICHARD I. EPSTEIN
Los Alamos National Laboratory
*Received 1988 January 12; accepted 1988 July 6*

## ABSTRACT

We describe an improved method for calculating $\gamma$-ray spectra of $\gamma$-ray bursts. Unlike methods currently in use, this method is independent of any assumed model for the incident spectrum. In many cases it allows one to perform $\chi^2$ model fitting without detailed knowledge of the instrument response functions. We first motivate the use of direct, model-independent inverse methods in general, and linear inverse methods in particular. We then explore the use of a particular linear technique, the Backus-Gilbert technique. We present the results of extensive calculations using simulated data and realistic response functions; the Backus-Gilbert technique produces model-independent spectral estimates that are quantitatively accurate and easy to calculate. Moreover, each estimate has a well-defined uncertainty and resolution that can be accurately calculated and straightforwardly displayed. Finally, we show how these estimates can be used for $\chi^2$ model fitting without detailed knowledge of the spectrometer response functions.

*Subject headings:* gamma rays: bursts — radiation mechanisms

## I. INTRODUCTION

Spectra of $\gamma$-ray bursters are usually observed with satellite-borne scintillators. The response functions of these devices are complicated; the data they produce are related to the incident spectrum in a complicated manner. Estimating the incident spectrum, given the data and the response functions, is thus a nontrivial inverse problem.

The method currently used to calculate $\gamma$-ray burst spectral estimates is based on model fitting. The resulting estimates are well-known to be model dependent in an "obliging" fashion, in the sense that the estimates tend to follow the model being fitted, even when the model is poor (Fenimore, Klebesadel, and Laros 1984; Teegarden 1984). Further, these estimates cannot be used by researchers who do not have intimate knowledge of the detectors to assess alternative models quantitatively. Currently, such assessments can only be made by using the raw data and the full response functions of $\gamma$-ray spectrometers. Unfortunately, the raw data are seldom published, and response functions never presented. This is because the response functions vary from burst to burst due to geometrical effects, making it impractical to present them. In addition, many or even most $\gamma$-ray bursts have been observed with instruments whose response functions are unavailable. These problems have made the topic of $\gamma$-ray burst spectra somewhat controversial, with investigators differing on such basic issues as the shape of the continua and the presence of line features (see the review by Harding, Petrosian, and Teegarden 1986, and references therein).

In this paper we describe a new method for inverting $\gamma$-ray burst spectral data that is free of many of the weaknesses of the currently used method. This method is based on the Backus-Gilbert inverse technique which has enjoyed a long history of success in analyzing complicated geophysical data. We show that it provides accurate spectral estimates which are model independent and have well-defined resolution and uncertainty. Also, we demonstrate that in many cases these estimates can be used for model fitting without detailed knowledge of the spectrometer response functions. Our method is not as accurate as rigorous model fitting using raw data and full response functions; however, it is a useful alternative when the response functions are not available or high accuracy is not needed.

In the next section, we set the stage for our discussion by formulating mathematically the inverse problem of determining a $\gamma$-ray spectrum from spectral data. There we also describe the currently used method and analyze its weaknesses. In § III we briefly discuss alternative inverse methods in general, and then motivate the selection of a particular one, the Backus-Gilbert method. In § IV we describe the application of this method to the problem at hand. We discuss our results in § V. Mathematical details regarding the method and our model response functions appear in two appendices.

## II. THE GAMMA-RAY SPECTRUM INVERSE PROBLEM

A $\gamma$-ray photon hitting a spectrometer produces a signal with a magnitude (or "pulse height") that is approximately proportional to the energy lost by the photon in the detector. We take $R(E, h)dh$ to be the probability that a $\gamma$-ray photon of incident energy $E$ generates a pulse with a pulse height between $h$ and $h + dh$; $R(E, h)$ is called the *response function* of the spectrometer. A typical response function is shown in Figure 1 as a function of pulse height evaluated at three different incident photon energies. In displaying this figure, we have followed the common practice of scaling the pulse height so that it is in units of energy, the scale being set so that $h$ at the rightmost peak (the "photopeak") is equal to the full energy of the incident photon. Scaled in this manner, the pulse height is a "blurred" measure of the energy lost by an incident photon in the detector. Accordingly, the abscissa is labeled "Energy loss." The distinction between incident photon energy, energy loss, and pulse height is elaborated on in Appendix A, where the model response function used in this work is discussed in detail.

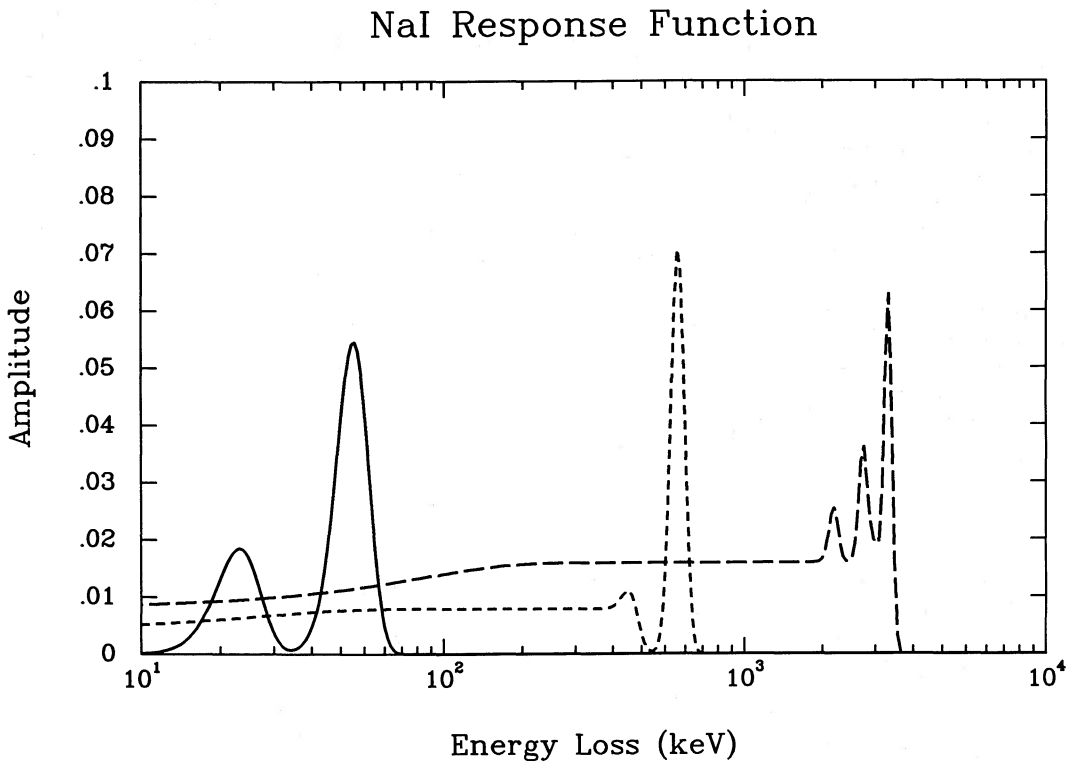


FIG. 1.—The response function of a $\gamma$-ray spectrometer. The function $R(E, h)$ is plotted as a function of pulse height $h$ for three constant values of incident photon energy, $E$. Solid curve is for $E = 50$ keV; the peak near 20 keV is due to the escape of flourescence photons. Short-dashed curve is for $E = 600$ keV; it has been multiplied by a factor of 10 relative to the solid curve. The Compton continuum is apparent, as is the Compton edge feature near 400 keV. Long-dashed curve is for $E = 3$ MeV; it has been multiplied by a factor of 100 relative to the solid curve. The two peaks below 3 MeV correspond to the loss of one or two pair quanta, and the low-energy tail is due to Compton scattering. These curves were plotted using the function described in Appendix A.

As is evident in the figure, $\gamma$-ray spectrometer response functions are highly structured. There is usually a prominent photopeak corresponding to deposition of all the incident photon's energy, but a significant fraction of the response function's area lies in a broad continuum, and possibly in additional peaks which correspond, for example, to the escape of pair quanta or of a fluorescence photon or other processes. It is the complexity of typical response functions that makes the interpretation of $\gamma$-ray spectral data nontrivial.

On exposure to some incident photon number spectrum $s(E) = dN_\gamma(<E)/dE$, where $N_\gamma(<E)$ is the number flux of photons with energy less than $E$, a detector with response function $R(E, h)$ produces a number of detection events, $n(h)dh$, with pulse height between $h$ and $h + dh$ according to

$$n(h) = \int_0^\infty R(E, h)s(E)dE + g(h) . \quad (1)$$

Here and throughout this paper we suppress trivial time and area integrals. The function $g(h)$ is an unknown function included to represent the uncertainty in the measurement of $n(h)$ due to photon statistics, noise, or other phenomena; $g(h)$ can be positive or negative. Since a photon detection event is commonly called a "count", we refer to the function $n(h)$ as a *continuous counts spectrum.* It is an energy *loss* spectrum, and is not an adequate estimate for the incident photon energy spectrum $s(E)$. For example, the curves shown in Figure 1 are the continuous counts spectra that would result from the observation of a perfectly monoenergetic source; clearly they are not good estimates of Dirac $\delta$-functions.

Equation (1) is an integral equation of Fredholm type, and the problem of solving it for the function $s(E)$, given knowledge of the other functions, has been studied in science and applied mathematics as one of a class of problems known collectively as *inverse problems.* This problem is known to be ill-posed in the sense that its solution lacks both uniqueness and stability (see Sarkar, Weiner, and Jain 1981; Jeffrey and Rosner 1986, and references therein). These problems arise because the smoothing action of the response function and observational uncertainties destroy information about finely detailed features of the spectrum.

For data from real spectrometers, the problem of uniqueness in the solution for the incident spectrum is compounded because the continuous function $n(h)$ is not measured. Rather, counts are binned into various ranges of pulse height referred to as *channels.* Thus, $\gamma$-ray spectral data consist of a finite number $N$ of quantities $n_i$ which comprise a *discrete counts spectrum*:

$$n_i = \int_{h_i}^{h_{i+1}} n(h)dh , \quad i = 1\text{–}N . \quad (2)$$

Here the $N + 1$ numbers $h_i$ are the boundaries of the pulse height bins that define the $N$ channels. Using equation (1), we may rewrite equation (2) as

$$n_i = \int_0^\infty R_i(E)s(E)dE + g_i , \quad (3)$$

where we have defined the *channel response function* $R_i(E)$ by

$$R_i(E) = \int_{h_i}^{h_{i+1}} R(E, h)dh , \quad (4)$$

and the $g_i$ represent the uncertainties which we shall assume are the result of random processes with zero means, known variances, and no correlations.

Some examples of channel response functions are shown in Figure 2. Typically channel response functions have a peak approximately bounded by energies equal to the scaled pulse heights that define the channel, and amplitudes that essentially vanish at energies below this peak. However, over half of the area of the channel response functions lies at energies above the main peak; high-energy photons can "masquerade" as low-energy photons in the energy loss spectra provided by $\gamma$-ray spectrometers.

Since the data only provide a finite number of constraints on the continuous function $s(E)$, the problem of determining $s(E)$ is underdetermined and hence has no unique solution. Because of the ill-posed nature of the $\gamma$-ray spectrum inverse problem (GRIP), the incident photon spectrum cannot be determined, but only estimated. We will designate such a spectral estimate by the symbol $\hat{s}(E)$. We may now state precisely the $\gamma$-ray spectrum inverse problem. Given the discrete counts spectrum $n_i$, the channel response functions $R_i(E)$, and some knowledge about the uncertainties $g_i$, use the $N$ equations (3) to arrive at an estimate $\hat{s}(E)$ for the true incident photon spectrum $s(E)$.

The usual technique for estimating a spectrum employs an *assumed* parametrized form for the spectrum, $\hat{s}_P(E)$, where the subscript $P$ refers to the set of parameters (Fenimore *et al.* 1982*a*; Mazets *et al.* 1983). Using this spectral estimate one calculates an expected counts spectrum, $\hat{n}_i$, according to

$$\hat{n}_i = \int R_i(E)\hat{s}_P(E)dE \,. \tag{5}$$

The parameters are adjusted so that the predicted counts spectrum $\hat{n}_i$ "matches" the data in some well-defined sense (e.g., by minimizing a $\chi^2$ parameter); $\hat{P}$ denotes the set of parameters which provides the best spectral estimate. The inverted spectrum is plotted as $N$ points at the energies $E_i = \frac{1}{2}(h_i + h_{i+1})$ with photon fluxes $\hat{s}_i$ derived from the data $n_i$ by scaling the latter according to $\hat{s}_P$. Different investigators do this scaling differently; two common choices are (Fenimore *et al.* 1982*a*; Mazets *et al.* 1983)

$$\hat{s}_i = \frac{n_i}{\hat{n}_i}\,\hat{s}_P(E_i)\,, \tag{6}$$

and (Matz 1986*b*)

$$\hat{s}_i = \frac{n_i}{\hat{n}_i}\,\langle \hat{s}_P \rangle_i\,, \tag{7}$$

where

$$\langle \hat{s}_P \rangle_i = \frac{\int_{E=h_i}^{E=h_{i+1}} \hat{s}_P(E)dE}{h_{i+1} - h_i}\,.$$

At each of these $N$ points, a vertical error bar is plotted, scaling the counts error $\sigma_{n_i}$ (often just $n_i^{1/2}$) by the same factor as $n_i$ was scaled by to calculate $\hat{s}_i$. A horizontal error bar is plotted at each point extending from $E = h_i$ to $E = h_{i+1}$.

One problem with the usual technique is that one must assume some parametrizable model for the true incident spectrum, such as thermal bremsstrahlung or synchrotron. Our ignorance of the physical nature of the sources of $\gamma$-ray bursts makes such assumptions unjustifiable, and the results of the usual technique can depend sensitively on the choice of model

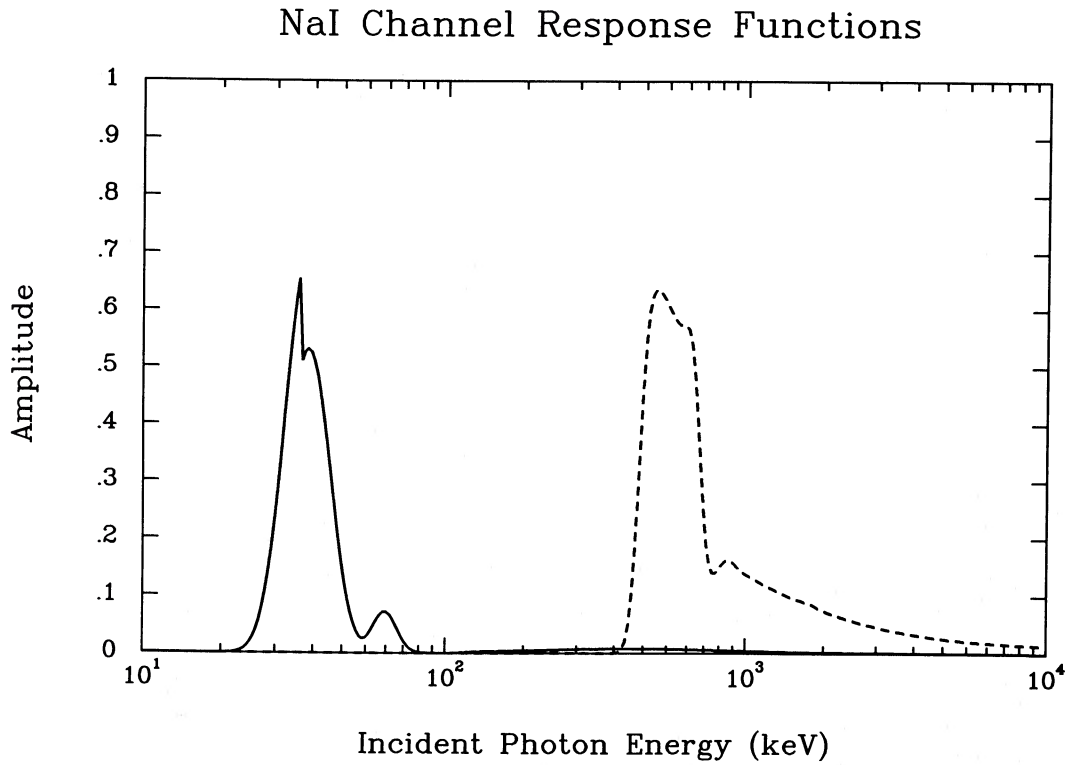


FIG. 2.—Channel response functions for a $\gamma$-ray spectrometer. The function $R_i(E)$ is plotted for two channels extending from a scaled pulse height of 30–40 keV (*solid line*) and 440–700 keV (*dashed line*). Note that the abscissa is logarithmic, so that for each function the area outside the main peak is significantly larger than it may appear.

(Fenimore *et al.* 1982*b*; Fenimore, Klebesadel, and Laros 1983; Teegarden 1984); the results are "obliging" in the sense that the spectral estimate follows the model, even if the model is poor (Fenimore, Klebesadel, and Laros 1983). An unfortunate consequence of this is that a spectral estimate produced with this method is not suitable for assessing spectral models differing from that used to produce the estimate.

A second problem with the usual technique is that, when viewed as an inverse technique, *it does not estimate the incident spectrum* $s(E)$, but rather, the *parameters* of the model chosen to represent $s(E)$ (Menke 1984). Ambiguity and arbitrariness can enter into the process by which parameter estimates are converted to spectral estimates $\hat{s}_i$; there is no unique method for determining $\hat{s}_i$ from the data. One particular consequence of this problem is that *the horizontal error bars plotted on spectral estimates produced by the usual method have no direct relationship to the resolution of the spectral estimate*, but typically indicate the bin size (that is, the resolution of photon energy *loss*). Section IV shows that the resolution obtainable with a given set of data is often significantly worse than error bars derived from bin size would imply.

The weaknesses of the usual technique can be avoided by using a technique that estimates photon spectra directly, thus eliminating the problems associated with choosing a parametrized model and converting parameter estimates to spectral estimates. We will refer to such a technique as a *direct inversion*. In the remainder of this paper we explore the application of direct inversion to the GRIP.

## III. THE BACKUS-GILBERT INVERSE

Decades of research on inverse techniques have demonstrated the importance of using *a priori* knowledge in constructing a solution to equations like (1) and (3) (Frieden 1975, 1983, 1984; Chambless and Broadway 1981). For example, by incorporating knowledge of the positivity of a spectrum (Frieden 1975; Chambless and Broadway 1981; Jansson 1984) or the boson nature of photons (Frieden 1983, 1984), inversion techniques have been developed which provide significant gains in resolution over simpler techniques. These techniques, such as maximum likelihood, maximum entropy, and maximum information, tend to be rather sophisticated and highly nonlinear (Frieden 1983; Narayan and Nityanada 1986).

Although research in inverse theory has moved toward sophisticated techniques, the geophysics community has consistently relied on linear techniques that take advantage of only the most basic *a priori* knowledge. This is because geophysical data (for example, seismic travel times) are related to the quantities of interest (for example, the density run with depth) in a very complicated manner, necessitating the use of inverse methods that clearly indicate what information may be obtained about the quantities of interest from the available data. Linear techniques are useful tools for extracting this information because their simplicity allows one to understand clearly how the inverse is related to the quantities of interest, and their linearity allows straightforward propagation of error from the data to the estimates.

Gamma-ray spectral data share the complexity of geophysical data, and so the criteria for selecting a solution to the GRIP are very similar to the criteria that have guided geophysicists in selecting inverse methods. In particular, a solution to the GRIP should provide accurate quantitative measures of the resolution and uncertainty attained in the solution. Moreover, a solution to the GRIP must accurately invert the smooth continuum as well as the features in a spectrum. Many nonlinear techniques are known to be unreliable when applied to smoothly varying functions, despite their ability to identify and enhance pointlike features (Frieden 1976; Frieden and Wells 1978; Schwarz 1979; Narayan and Nityanada 1986). Further, these techniques often rely on significant oversampling of response functions, i.e., pulse height bins which are narrow compared to the width of the response function (Frieden and Swindell 1976). The response functions of $\gamma$-ray spectrometers are seldom oversampled.

We now describe the Backus-Gilbert technique. This technique estimates the incident spectrum with a linear combination of the data:

$$\hat{s}(E_j) = \sum_{i=1}^{N} a_i(E_j) n_i \tag{8}$$

$$= \int \sum_{i=1}^{N} a_i(E_j) R_i(E) s(E) dE \,, \tag{9}$$

where $\hat{s}(E_j)$ is an estimate of the incident spectrum at some energy $E_j$, and the coefficients $a_i$ depend on the energy at which the spectrum is being estimated. The linear inverse problem reduces to determining the best set of coefficients $a_i$.

Equation (9) may be suggestively rewritten as

$$\hat{s}(E_j) = \int \hat{\delta}(E, E_j) s(E) dE \,, \tag{10}$$

where we define the *resolution function* $\hat{\delta}(E, E_j)$ by

$$\hat{\delta}(E, E_j) \equiv \sum_{i=1}^{N} a_i(E_j) R_i(E) \,. \tag{11}$$

The resolution function is a linear superposition of channel response functions. The form of equation (10) immediately suggests a criterion for choosing the $a_i$; if they could be chosen to make $\hat{\delta}(E, E_j)$ a Dirac $\delta$-function $\delta(E - E_j)$, then our spectral estimate would perfectly represent the incident spectrum. In practice, the set of response functions is not complete, and no set of coefficients can make the resolution function a $\delta$-function. But we can define some quantitative measure of "$\delta$-function-likeness" and choose the $a_i$ to make $\hat{\delta}(E, E_j)$ most like $\delta(E - E_j)$ by this criterion.

One scheme for choosing the $a_i$ that is popular among geophysicists was suggested by Backus and Gilbert (Backus and Gilbert 1970; see Parker 1977 for an excellent review). Define the *spread* of the resolution function by

$$r(E_j) \equiv \int (E - E_j)^2 [\hat{\delta}(E, E_j)]^2 \, dE \,. \tag{12}$$

This functional of $\hat{\delta}$ vanishes if $\hat{\delta}$ is a $\delta$-function and grows larger the more $\hat{\delta}$ is spread from the energy $E_j$. We thus choose the $a_i(E_j)$ so as to minimize the spread $r(E_j)$, subject to the normalization constraint,

$$\int \hat{\delta}(E, E_j) dE = 1 \,. \tag{13}$$

Intuitively, one might suppose that this procedure generates a resolution function that gives the most localized estimate of $s(E_j)$. Using techniques of functional analysis, Backus and Gilbert (1970) have shown this supposition to be rigorously true in the sense that $\hat{s}(E_j)$ will be close to $s(E_j)$ (in the sense of the $L_2$ norm) only if $r(E_j)$ is small. In any case, a plot of $\hat{\delta}$ would display exactly how localized the spectral estimate is.

Because the data have uncertainties associated with them, any spectral estimate will be uncertain. The propagated variance of the estimate, $\sigma_{\hat{s}}^2$, can be calculated from equation (8); if we assume that the data are uncorrelated, then

$$\sigma_{\hat{s}}^2(E_j) = \sum_{i=1}^{N} a_i^2(E_j)\sigma_i^2 \ . \tag{14}$$

(The more general case of correlated errors is discussed in Appendix B). This formula suggests an alternative criterion for the selection of the coefficients $a_i$. If we would like to be certain of our spectral estimate, we should choose the $a_i$ so as to minimize $\sigma_{\hat{s}}^2$ (subject, of course, to the $\delta$ normalization constraint mentioned above). However, the set of coefficients that minimizes the propagated uncertainty is *not* the set which minimizes the spread functional. Note, in particular, that the $a_i$ which minimize equation (14) will be independent of energy. In fact, the minimum spread solution may be characterized by very large error propagation, and the minimum uncertainty solution can have very poor resolution (very large spread); *resolution and error propagation compete*. We are forced to effect some compromise between spread and variance. We will see below that this is not the only way in which resolution and error propagation can compete (see § IV*b*).

Backus and Gilbert suggested effecting this compromise by choosing the $a_i$ that minimize the following *weighted sum* of spread and variance:

$$w(E_j, \theta) = r(E_j) \cos \theta + v\sigma_{\hat{s}}^2 \sin \theta \ , \tag{15}$$

where $\theta$ parametrizes the tradeoff between resolution and variance and may take on any value from 0 to $\pi/2$ (Backus and Gilbert 1970; Parker 1977). (The constant $v$ is arbitrary and merely selects how $\theta$ parametrizes the inverses obtained by minimizing $w$; see Appendix B.) The virtue of this method of tradeoff is that Backus and Gilbert have shown that, for the definition of the spread functional given by equation (15), *the inverse obtained by minimizing $w(E_j, \theta)$ is the best possible linear inverse*, in the sense that the spread of such an inverse is the lowest possible, given its variance, and simultaneously, its variance is the lowest possible, given its spread. Moreover, the solution for the $a_i$ is analytical and simple; it is derived in Appendix B.

The manner in which minimizing $w(E_j, \theta)$ gives the best inverse is illustrated in Figure 3, which shows a schematic *tradeoff curve*—a plot of variance *versus* spread for the inverses found by minimizing $w(E_j, \theta)$ at a particular $E_j$ as $\theta$ varies from 0 to $\pi/2$. The shaded portion of the graph is the region of all possible linear inverses where $w(E_j, \theta)$ is not necessarily a minimum. The tradeoff curve bounds this region where the spread and variance are both smallest.

The Backus-Gilbert technique does not provide a unique inverse, but rather provides a one-dimensional family of inverses parametrized by $\theta$. But though the inverse is nonunique, it is not poorly defined or ambiguous. The choice of $\theta$ merely specifies a particular representation of the inverse. All the possible Backus-Gilbert inverses are quantitatively equivalent. For example, if any inverse representation is used to test a given theoretical model it will yield the same $\chi^2$ value (see § IV*a*), and any inverse solution can be transformed to any other one (or back to the initial data set) by a linear transformation. One cannot be misled by an inverse produced with this method; no conclusion can be drawn from the inverse that cannot be supported by the original data. Thus every Backus-

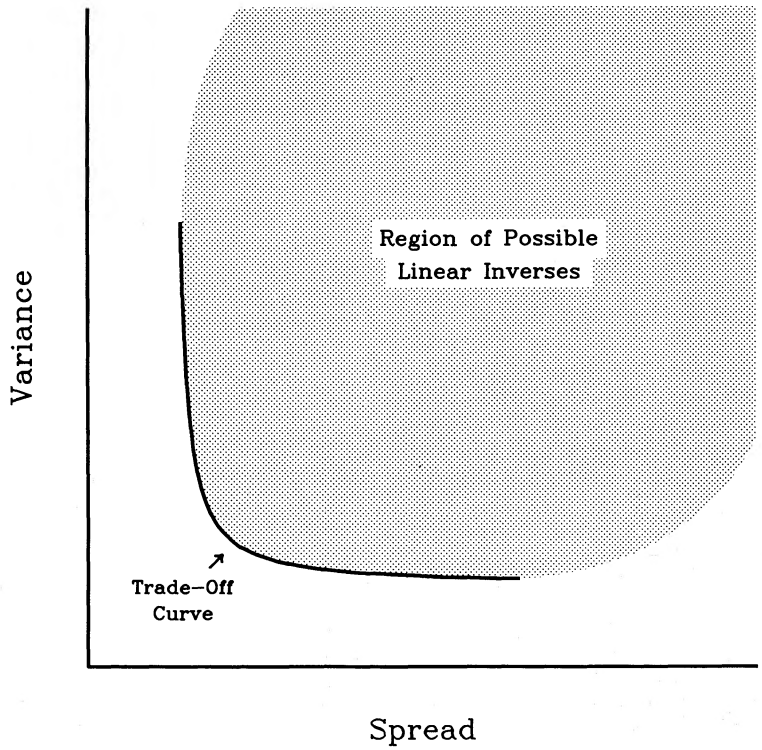


FIG. 3.—Schematic Backus-Gilbert tradeoff curve, after Parker (1977). A point on this plot represents a solution to the inverse problem, parametrized by its variance (*vertical axis*) and spread (*horizontal axis*). Shaded region shows schematically the region containing all possible linear inverses. Solid curve bounding this region near the origin—the tradeoff curve—shows the inverses selected by the Backus-Gilbert method.

Gilbert inverse is *meaningful*, but not all are *useful*. A particular solution may be subjectively superior for addressing a particular question, according to whether one is most concerned about resolution or uncertainty. For example, choosing an inverse with high resolution ($\theta$ close to 0) gives well-localized $\hat{\delta}$ functions and values of $\hat{s}(E_j)$ that are close to $s(E_j)$. We will show that these solutions convey the spectral informaton in a way that is appropriate for graphic presentation and that can be used for $\chi^2$ analyses without detailed knowledge of the response functions. But such a solution is not any more or less "correct" that any other one. The problem is not dissimilar to what one encounters when expanding a function in a complete set of basis functions. While all such expansions are equivalent and equally correct, considerations such as symmetries or which quantities are to be computed dictate which representation is most appropriate. These points are elaborated on in § IV*c*, where we discuss the selection of $\theta$ when inverting $\gamma$-ray burst spectral data.

Before describing the application of the Backus-Gilbert method to $\gamma$-ray spectroscopy, we point out that it is not the only linear inverse method available. Other methods exist which employ spread functions that differ from equation (12) (see Parker 1977). Also, there are methods which estimate spectra at different energies jointly, using smoothness and $\chi^2$ constraints on the entire solution set instead of resolution and uncertainty constraints on individual estimates (see, e.g., the Phillips-Twomey method described in Frieden 1975 and generalized in Jeffrey 1988). In addition, an entirely different class of linear inverses exists based on an eigenfunction analysis of the system of channel response functions. These methods include the singular value decomposition used in X-ray spectroscopy (Kahn and Blissett 1980; Kahn *et al.* 1980), and the closely related spectral expansion method developed by Gilbert (1971) and used extensively by geophysicists (Parker 1977) and more recently by solar physicists (Jeffrey and Rosner 1986).

We have examined the applicability of these various methods to the GRIP and found that they are not as useful as the Backus-Gilbert method (with the modifications that we describe in the next section). Although computationally efficient, the Phillips-Twomey type of methods do not give well-localized estimates of $s(E)$; i.e., they have poor energy resolution. This is because they do not allow the use of *a priori* knowledge that varies from estimate to estimate; such knowledge is essential in constructing well-localized inverses of $\gamma$-ray spectral data (see § IV*a*). Gilbert's spectral expansion method gives inverses that are unstable in the presence of noise and which often produce spurious linelike features. This result has also been found by other investigators using the singular value decomposition to invert $\gamma$-ray spectral data (Chambless and Broadway 1981; Fenimore 1986). This illustrates a lesson learned long ago in inverse method research: *the usefulness of a particular technique depends on the nature of the data being inverted, in particular on the shapes of the response functions and on the nature of the noise* (Frieden 1984; Jeffrey and Rosner 1986). Unfortunately, there is no inverse method appropriate for all problems.

## IV. APPLICATION TO $\gamma$-RAY SPECTROSCOPY

One test of the ability of a particular inverse method to provide accurate solutions to a given problem is extensive simulation. In this section we discuss the details of the application of the Backus-Gilbert method to the GRIP and present the results of inversions of simulated data. We first discuss some general properties of the resolution functions for $\gamma$-ray spectral inverses. Then we discuss three issues that must be decided before inverses of simulated data can be calculated and displayed: the determination of the number and locations of the energies $E_j$ at which to calculate spectral estimates, the choice of the tradeoff parameter to use in constructing each estimate, and the calculation of a quantitative measure of the resolution of an estimate to plot as a horizontal error bar. We then present the results of extensive simulations that demonstrate the accuracy and usefulness of the Backus-Gilbert method.

### *a) Resolution Functions*

One of the great virtues of the Backus-Gilbert method is that it provides a resolution function that reveals precisely how a spectral estimate is related to the unknown incident spectrum. A comparison of the resolution functions with the channel response functions illustrates the extent to which the Backus-Gilbert estimate is superior to the data. We will begin our discussion of the Backus-Gilbert method to $\gamma$-ray spectroscopy by examining the resolution functions it produces.

Figures 4*a* and 4*b* show two resolution functions, $\hat{\delta}(E, E_j)$, for inverses calculated using a system of 19 channel response functions with bin boundaries $h_i$ chosen to be similar to those employed in actual satellite experiments (see Appendix A for a detailed discussion of channel response functions). The arrow on each plot indicates the energy $E_j$ of the spectral estimate this resolution function was calculated for. The tradeoff parameter $\theta$ was set to zero, corresponding to the highest resolution inverse. The channel response function whose peak encloses $E_j$, normalized to have unit area, is also shown on each plot (*dashed curves*). Comparison of the resolution function and the channel response function shows that a significant gain in resolution has been achieved. Where the original response function in Figure 4*a* had only $\sim$24% of its area in the peak at $E_j$; the resolution function has 65% of its area in that peak (note that the horizontal axis is logarithmic). Similarly, the channel response function in Figure 4*b* has only 32% of its area in its peak, whereas the resolution function has $\sim$78% of its energy in that peak. In these fairly typical examples the significance of the central peak in the resolution function is more than twice that of the channel response function.

These two figures also illustrate an undesirable consequence of naively applying the Backus-Gilbert method. Because the spread function (12) equally weights the parts of the resolution function on each side of $E_j$, the functions obtained by minimizing the spread tend to be symmetric about $E_j$, to the extent permitted by the original channel response functions. This symmetry is apparent in Figures 4*a* and 4*b*. The resolution functions are significantly better than the response functions at energies above $E_j$, but are worse than them at energies below $E_j$. The resolution function has a small positive amplitude at energies below the main peak where the channel response function essentially vanished. This is undesirable, because typical GRB spectra fall off steeply and the small amplitude of the resolution function at energies far below $E_j$ can cause the estimate to be significantly contaminated by low-energy photons.

A slight modification of the standard Backus-Gilbert technique readily cures the low-energy contamination. To circumvent this problem, one omits some of the channel response functions when calculating a resolution function. Since channel response functions have fairly sharp low-energy boundaries (see Fig. 2), the Backus-Gilbert method is only needed to

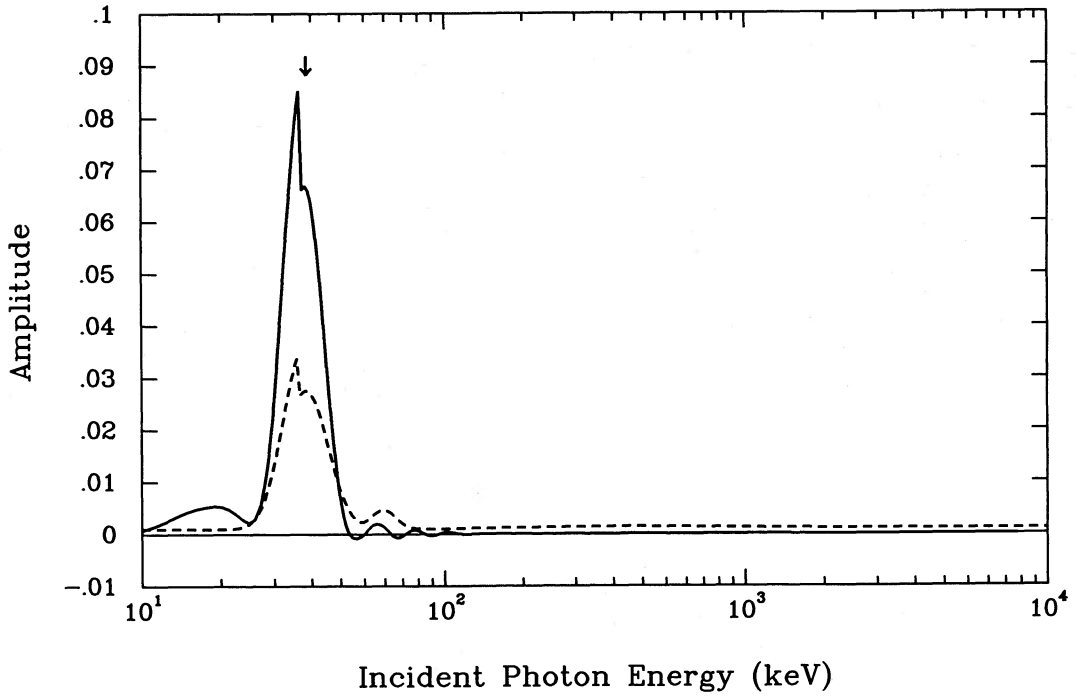


FIG. 4*a*

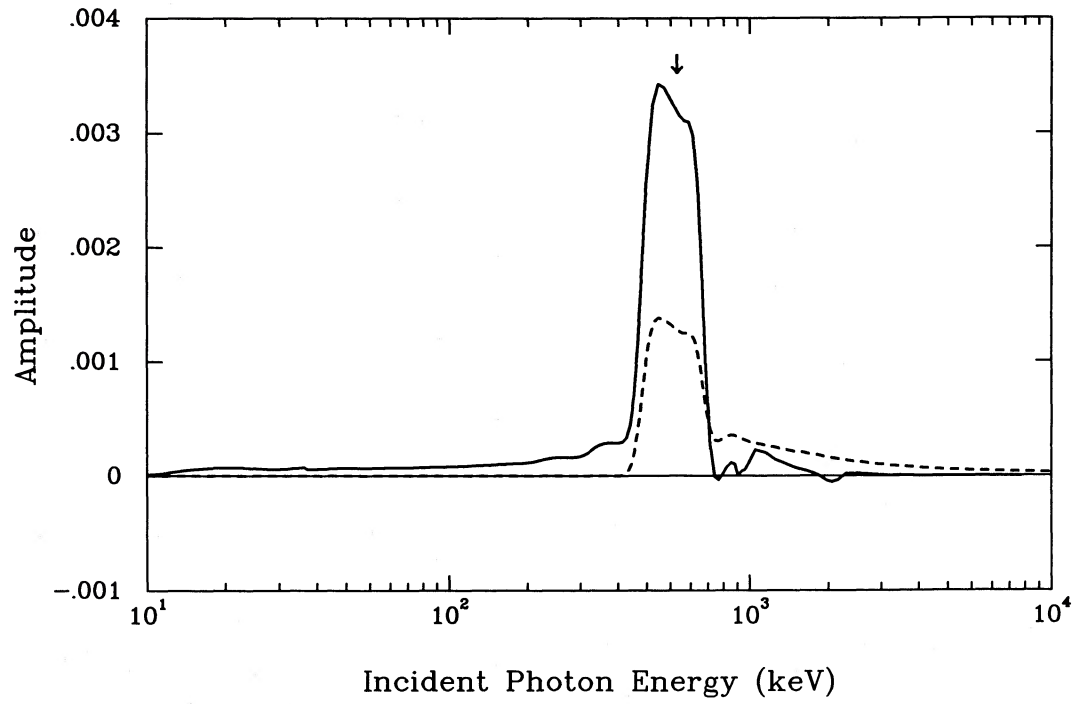


FIG. 4*b*

FIG. 4.—Typical resolution functions produced by the Backus-Gilbert method for estimating $\gamma$-ray spectra. (*a*) Resolution function (*solid curve*) for an estimate at $E_j = 35$ keV (indicated by the arrow), calculated using all the channel response functions. For comparison, the dashed curve shows the channel response function whose peak contains $E_j$, normalized to have unit area; its bin boundaries extend from a scaled pulse height of 30 to 40 keV. Note that the horizontal axis is logarithmic, so that the actual area under the curves at high energy is larger than the geometrical area on these plots. (*b*) As in (*a*), for an estimate at $E_j = 570$ keV, with the channel response function with bin boundaries extending from 440 to 700 keV, normalized to have unit area. (*c*), (*d*) Same as (*a*) and (*b*), respectively, except that all channel response functions with peaks at energies less than $E_j$ were omitted from the calculation of the resolution functions.

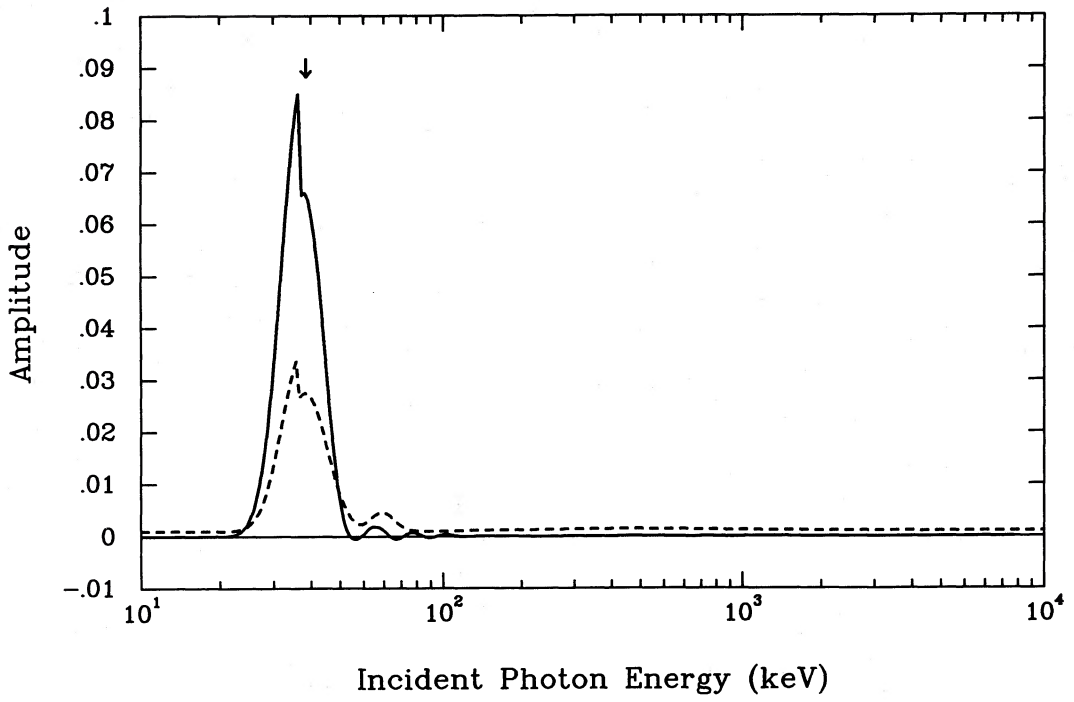


FIG. 4*c*

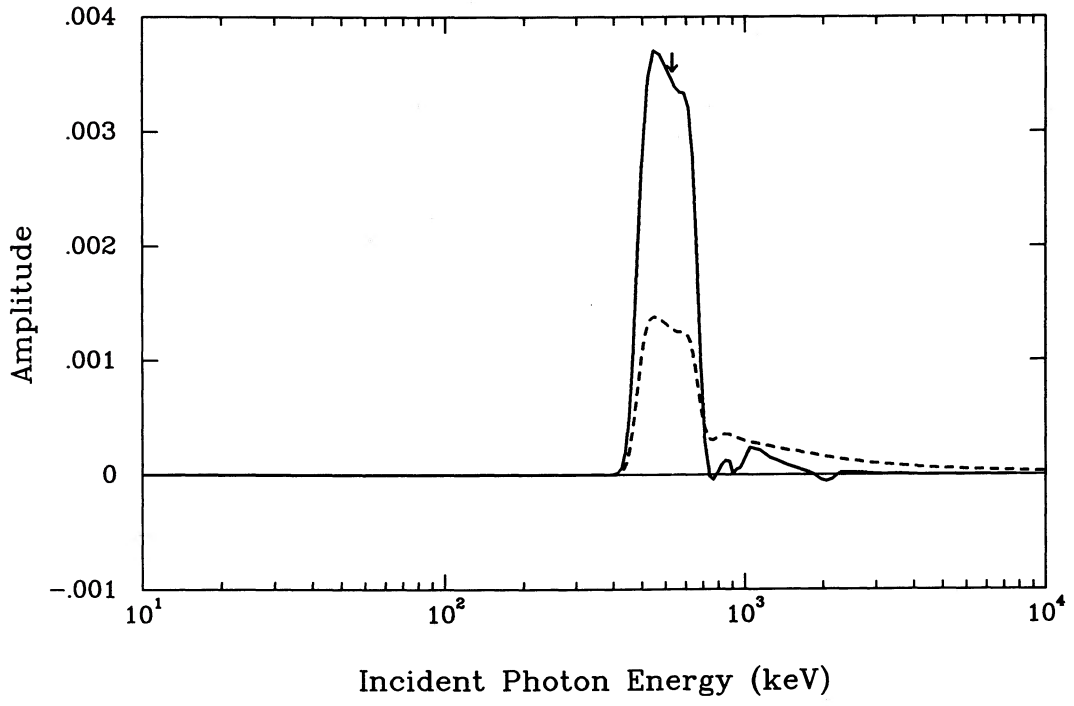


FIG. 4*d*

improve the high-energy part of the response. Thus, when estimating the incident spectrum at an energy that falls within the peak of the $k$th channel response function, one should use only the higher energy channels $k$ through $N$ to construct the inverse. This corresponds to changing the lower bound on the sums in equations (8), (9), (11), and (14) to $i = k$. Applying this technique to the response functions used to produce Figures 4*a* and 4*b* gives the resolution functions shown in Figures 4*c* and 4*d*. The amplitudes of these functions essentially vanish at low energies. In addition, the fraction of the area in the main peak of the resolution function has increased to $\sim 83\%$ in Figure 4*d*; this represents almost a factor of 3 improvement over the channel response function.

The Backus-Gilbert resolution function need not always be better than the channel response function. For example, if one considers a fictitious set of channel response functions that are

nonoverlapping, then the Backus-Gilbert resolution functions have to be broader than the individual channel response functions.

For current $\gamma$-ray spectrometers the great improvement of the Backus-Gilbert resolution functions over the channel response functions is due to the particular form of the response functions: a peak with a sharp low-energy cutoff and a fairly flat Compton tail. If future $\gamma$-ray spectrometers have significantly different characteristics from contemporary devices, the improvement in resolution may be significantly changed from that shown here.

### b) *Selecting the Energies $E_j$*

To address the question of where to estimate the incident spectrum, in Figures 5*a* and 5*b* we have plotted the spread and the variance of an inverse as a function of the energy $E_j$ inverted for, with $E_j$ varying across the range of two channels. The channel boundaries are indicated by dashed lines. Figure 5*a* shows that resolution is best for inverses at bin centers, as one might intuitively expect. Figure 5*b* shows that propagated variance is best at channel *boundaries*. We thus find the interesting but not surprising result that resolution and uncertainty compete in the selection of the point at which to estimate the incident spectrum. In all of our applications, the importance of resolution exceeded that of uncertainty enough that it was always proper to invert at the center of a channel. Accordingly, we always put $E_j = (h_j + h_{j+1})/2$.

We now must determine the number of places, $M$, at which to calculate $\hat{s}(E_j)$ ($j = 1$–$M$). It is clear that it is not meaningful to take $M > N$, because the data provide only $N$ independent constraints on the spectrum, so that at most $N$ estimates of the spectrum can be linearly independent. For the calculations presented here, we have put $M = N$, on the grounds that the $N$ statistically independent data should allows us to place $N$ constraints on the incident spectrum. Also, the choice $M = N$ is required if spectral estimates are to be compared with theoretical models using $\chi^2$ techniques (see § IV*g*).

### c) *Selecting the Tradeoff Parameter, $\theta$*

Once we have selected the energy at which to calculate the spectral estimate, we must choose how to compromise between resolution and uncertainty in the calculation, that is, we must select a value for the tradeoff parameter, $\theta$. The most rigorous way to go about this is to develop some criterion for selecting a point on the tradeoff curve, calculate tradeoff curves for every energy at which the spectrum is being estimated, and use the criterion to choose $\theta$ for each energy. Since this is expensive computationally, we simply use the highest resolution inverse ($\theta = 0$) unless its propagated error is not acceptable. Use of the highest resolution inverse is particularly economical when several inverted spectra of the same source must be calculated, as in the case of time-resolved spectroscopy, because the coefficients defining this inverse are independent of the data (see Appendix B). However, when the error amplification of the highest resolution inverse is unacceptable, one should increase $\theta$ until acceptable uncertainties are obtained. This may be necessary when the signal-to-noise ratio of the data is small.

Although this method for selecting $\theta$ is computationally economical, when economy is not an issue it can pay to be more careful about choosing $\theta$. Figure 5*c* shows a typical tradeoff curve for an inverse of simulated GRB data. It has the desirable property of being nearly parallel to the axes near its endpoints (this is true for all tradeoff curves, by definition). This means that one need sacrifice little resolution compared to that

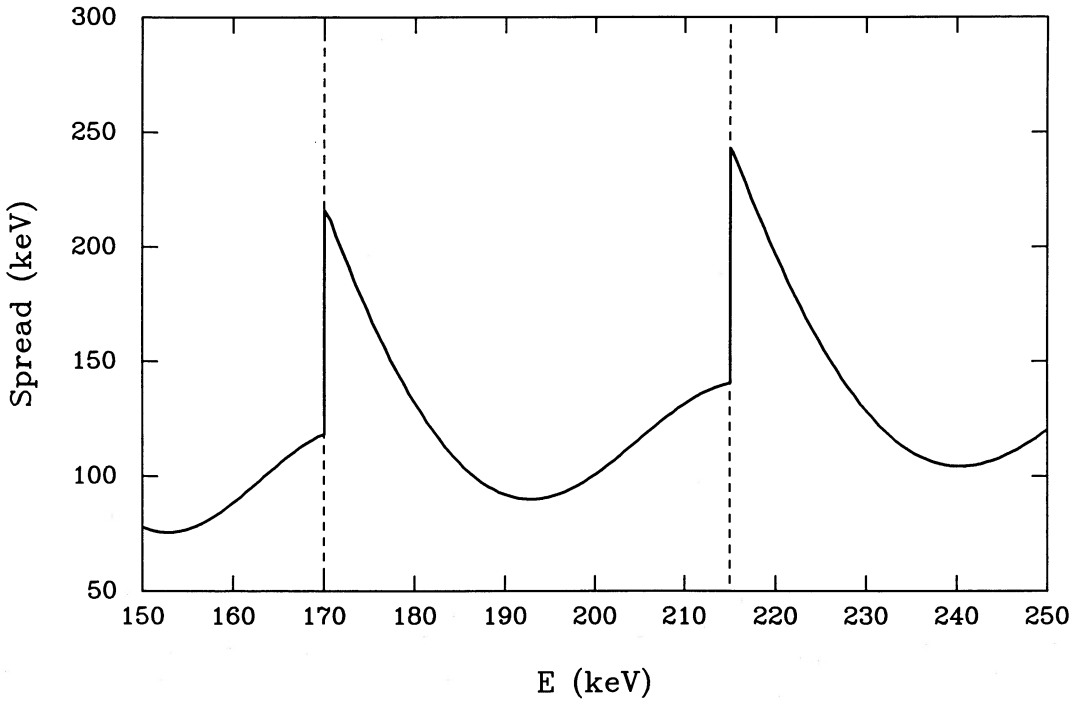


FIG. 5*a*

FIG. 5.—The competition between resolution and uncertainty in selecting an inverse. (*a*) Spread of an inverse vs. the location. Dashed lines show the scaled pulse height boundaries defining the channel response function whose peak is near the location of the inverse. (*b*) Variance of an inverse vs. its location. (*c*) Tradeoff curve for inverses at $E_j = 192.5$ MeV. The inset shows the tradeoff curve on a linear scale.

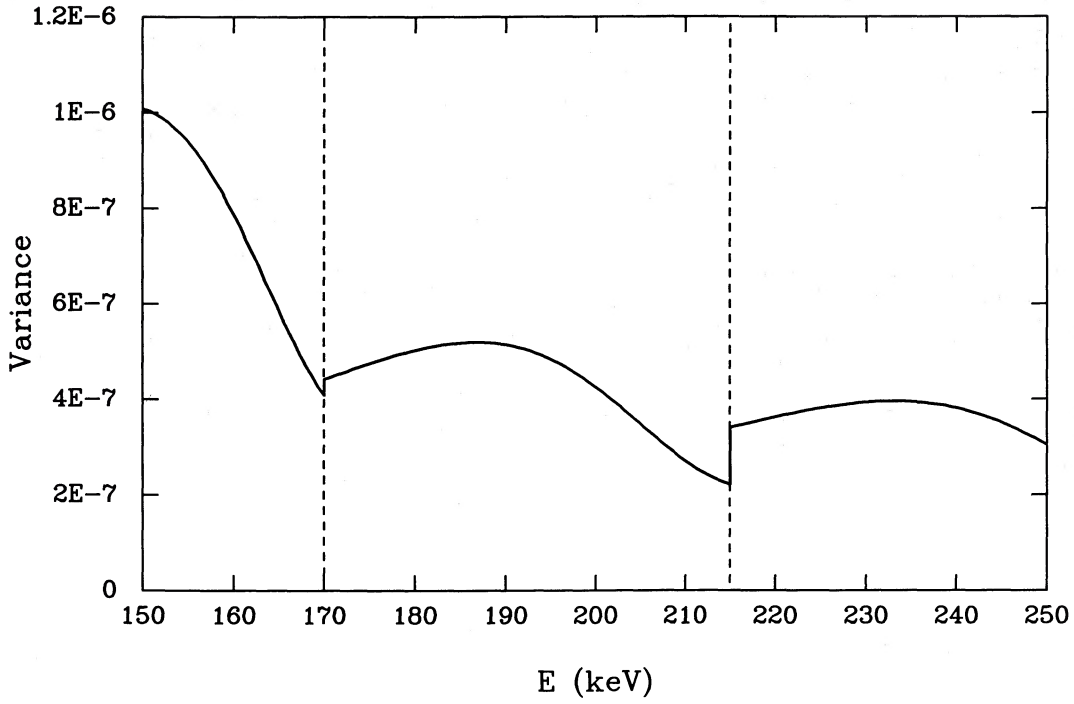


FIG. 5b

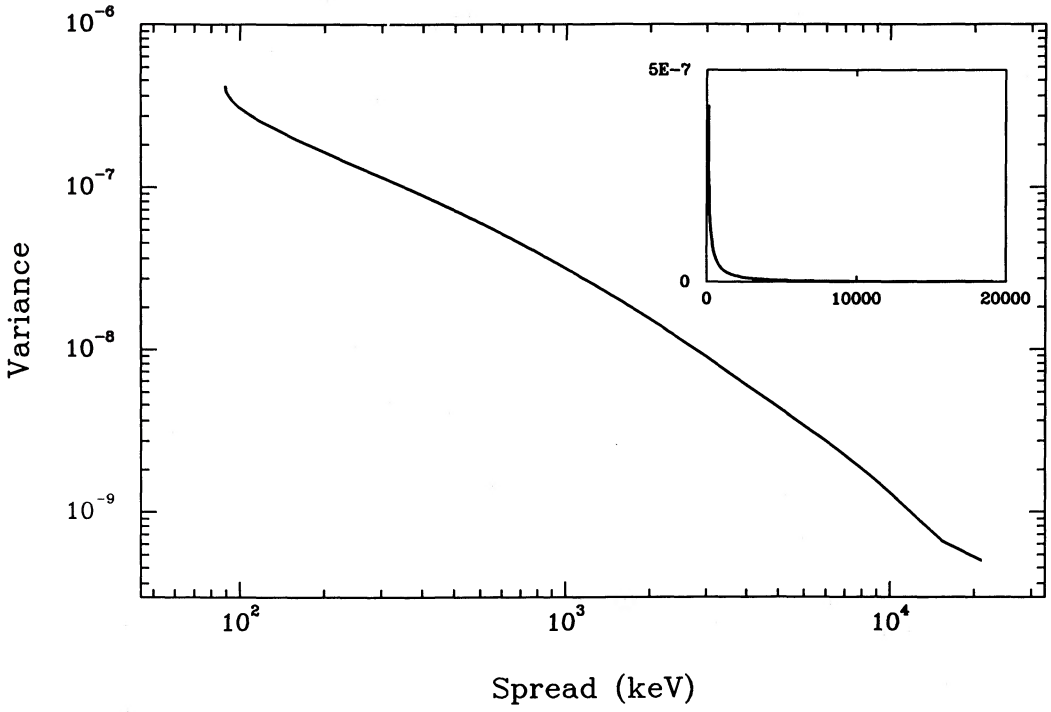


FIG. 5c

achievable with the highest resolution inverse in order to obtain error propagation significantly better than that of the highest resolution inverse.

We emphasize again that *all* values of $\theta$ lead to *meaningful* inverses, although not all values lead to *useful* inverses. The inverses are meaningful in the sense that one cannot be misled by them, even in the low signal-to-noise case. Most inverse techniques do not provide meaningful measures of the uncertainty associated with the estimates they provide; excessively noisy data can cause such techniques to produce spurious, misleading estimates. But the Backus-Gilbert method allows one to calculate the uncertainty of the estimates by propagating the uncertainty of the data. When the noise in the data is excessive, the propagated uncertainty will be large, preventing an investigator from drawing conclusions that are not substantiated by the data.

But though any Backus-Gilbert inverse is meaningful, an inverse is useful only to the extent that its characteristics allow one to address particular questions about $\gamma$-ray burst spectra. The practical $\theta$ criterion described above produces spectral estimates that, being well localized, are generally useful for addressing the questions many investigators have about burst spectra. But for particular applications, other criteria may be preferred.

### *d) Measuring Resolution*

It is useful to have a numerical measure of the resolution, especially for $\chi^2$ evaluations of models (see § IV*g*) and for the graphical presentation of inverted spectra, where this measure would take the form of a horizontal "error bar." One commonly used measure is the spread of the resolution function, as defined by equation (12), with an appropriate normalization. We prefer a more straightforward characterization of the resolution. Calculate the region in energy about $E_j$ from which some fraction $f$ of the detected photons comprising the estimate $\hat{s}(E_j)$ came from. The lower and upper bounds of this region $E_{\rm lo}$ and $E_{\rm hi}$ satisfy

$$\frac{1-f}{2} = \frac{1}{\hat{s}(E_j)} \int_0^{E_{\rm lo}} \hat{s}(E)\hat{\delta}(E, E_j)dE \,, \tag{16}$$

and

$$\frac{1-f}{2} = \frac{1}{\hat{s}(E_j)} \int_{E_{\rm hi}}^{\infty} \hat{s}(E)\hat{\delta}(E, E_j)dE \,. \tag{17}$$

The continuous function $\hat{s}(E)$ is calculated by interpolating between the finite number of spectral estimates $\hat{s}(E_j)$ that have been calculated. For the error bars plotted in this paper, we used $f = 0.90$.

We note that this measure of resolution is most meaningful when the resolution functions are well localized, lacking the strong sidelobes often produced by linear inverse methods. In all our caculations, the amplitude of the largest sidelobe of a resolution function never exceeded a few percent of its amplitude at $E_j$.

### *e) Sample Inverses*

Figures 6*a*–6*d* illustrate the application of the Backus-Gilbert method to the inversion of simulated $\gamma$-ray spectral data using the response functions described in Appendix A. For Figure 6*a*, we take $s(E)$ to be a broken power law, with $s(E) \propto E^{-1}$ for $E < 300$ keV and $s(E) \propto E^{-2}$ for $E > 300$ keV. This input spectrum is convolved with the response functions of a 16 channel spectrometer to give a simulated data set with Poisson noise. From these data a spectrum $\hat{s}(E_j)$ was estimated at 16 energies $E_j$ selected as described above. The bin boundaries of the channels were chosen to be similar to those used by the *Konus* collaboration in processing and displaying their GRB data (these results are only suggestive of what may be attainable with the actual *Konus* response functions). At each point, a vertical error bar is plotted showing 1 $\sigma$ confidence regions corresponding to the propagated variances calculated from equation (14). A horizontal error bar is plotted indicating the regions where 90% of the photons contributing to the estimate came from as described above. The tradeoff parameter $\theta$ was set to zero for all the estimates. Figure 6*b* shows the results of the same process with an input spectrum that is an approximate optically thin thermal bremsstrahlung spectrum (with constant Gaunt factor) with an absorption feature at 45 keV. This spectrum is similar to those used by the *Konus* collaboration in fitting many of their burst spectra (Mazets *et al.* 1981).

Figures 6*c* and 6*d* show the results of similar calculations with different channel response functions. The energy range and bin boundaries for these figures were chosen to simulate qualitatively the channel response functions for the High-Energy X-ray and Low-Energy Gamma-Ray Experiment on *HEAO 1* which observed a low-energy absorption feature in GB 780325 (Heuter 1984; note that this instrument employs active shielding to decrease the amplitude of the Compton continuum in the response function). For Figure 6*c*, a thermal synchrotron-like spectrum with a shallow absorption feature at 55 keV was used to produce the simulated data, and then 19 spectral estimates were calculated and displayed as in the previous figures. For Figure 6*d*, the simulated data were produced using the broken power-law spectrum described above, with Gaussian random noise added at a constant signal-to-noise ratio of 5, mimicking the effect that high background or detector noise might have on the data.

The accuracy of these inverses is very satisfying, and is typical of that achievable with this method applied to a wide variety of spectral shapes. It is interesting to note that the horizontal error bars plotted in these figures overlap, indicating that the resolution of these estimates is significantly worse than error bars defined from bin sizes (as they are in usual spectral inversion techniques) would imply. The estimates near the features in Figures 6*b* and 6*c* clearly show the effect of the finite resolution of the detector on it ability to distinguish narrow features; one would probably not conclude that significant line features are evident in the inverted spectra if the incident spectrum curves were not present to guide the eye. It is clear that a straightforward measurement of the resolution of each estimate, as is provided by the method we have presented here, is very helpful for properly interpreting the data.

### *f) Monte Carlo Results*

To characterize the statistical properties of these Backus-Gilbert inverses, we have performed extensive Monte Carlo calculations. The results of four of these are presented in Figure 7. Figure 7*a* indicates the results of inversions of 100 sets of simulated data like that used to produce Figure 6*a*. At the average value of the 100 inverses calculated at each $E_j$, a horizontal error bar is plotted that is the average of the 100 error bars calculated. *Two* vertical error bars are plotted for each horizontal one. The left vertical bar indicates the average value of the propagated error for the 100 inverses, and the right one indicates the standard deviation of the mean of the 100 estimates. The agreement between these two error bars indicates that the error is being accurately propagated.

Figures 7*b*, 7*c*, and 7*d* show similar averages for data like that used to produce Figures 6*b*, 7*c*, and 6*d*. Together with other calculations, not displayed here, these figures demonstrate that the Backus-Gilbert method can produce accurate, meaningful inverses of GRB spectral data, and that the quality of these inverses is insensitive to the shape of the incident spectrum, the character of the noise associated with the data, or the locations of the bin boundaries of the channel response functions.

### *g) Evaluating Models and Estimating Parameters*

The techniques described so far allow one to obtain model-independent estimates of the spectral flux $\hat{s}_k \equiv \hat{s}(E_k)$ at energies

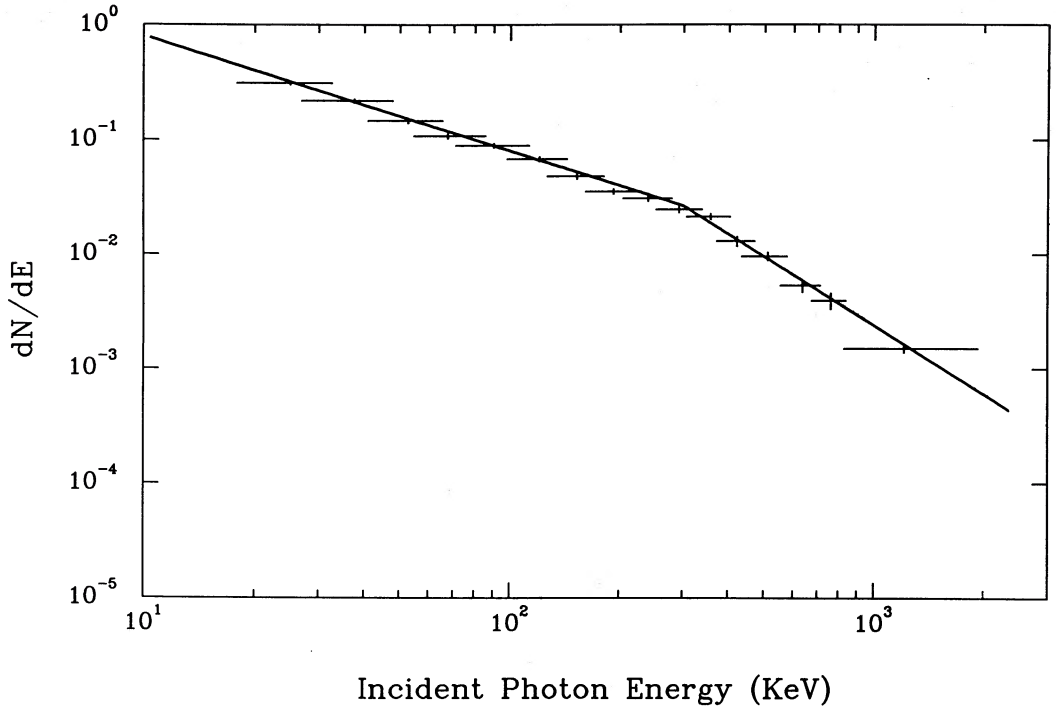


FIG. 6a

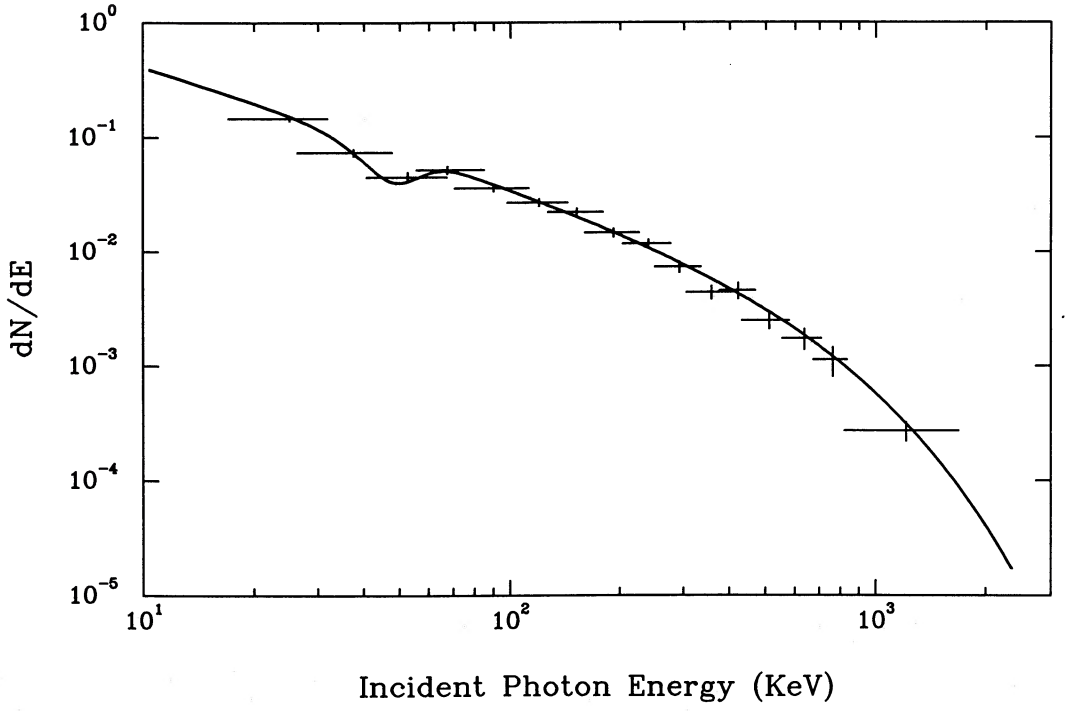


FIG. 6b

FIG. 6.—Inversions of simulated data. (*a*) Broken power-law spectrum (*solid line*) was convolved with 16 channel response functions, and Poisson noise was added to produce simulated data. The results of Backus-Gilbert inversions at 16 energies are shown. Vertical error bars show 1 $\sigma$ confidence regions calculated from the propagated variance of the inverse; horizontal error bars show the region from which 90% of the photons comprising the estimate came from. See text for details. (*b*) As in (*a*), but here the incident spectrum is optically thin thermal bremsstrahlung with an absorption feature at 45 keV. (*c*) As in (*a*), but here the input spectrum is a thermal synchrotron-like spectrum with an absorption feature at 55 keV, and a 19 channel response system was used. (*d*) As in (*c*), using a broken power-law input spectrum with Gaussian noise added to the simulated data with a constant signal-to-noise ratio of 5.

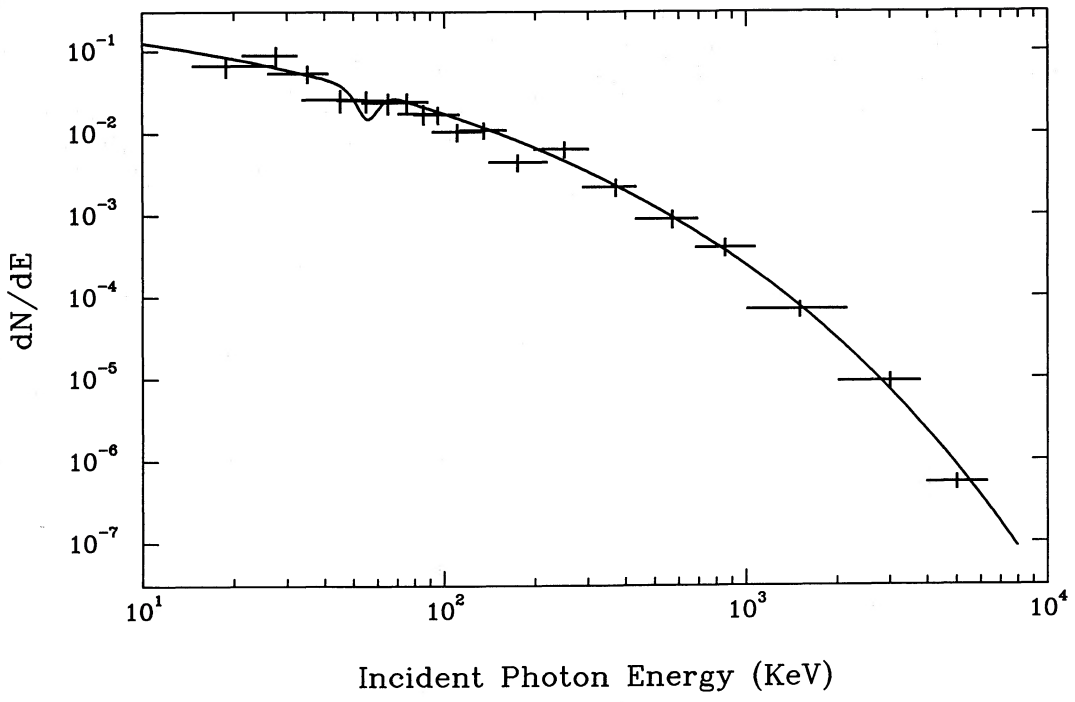


FIG. 6c

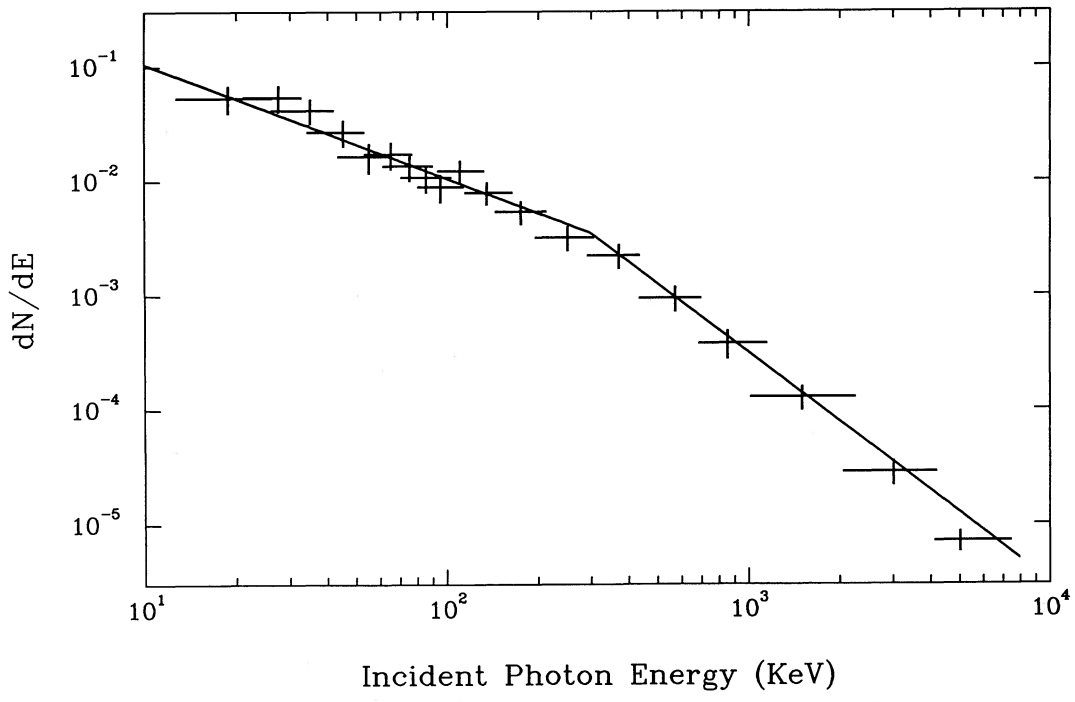


FIG. 6d

$E_k$ as well as the uncertainty in $\hat{s}_k$ and the resolution about $E_k$. These quantities are suitable for graphical display of spectral data. It is also important to be able to evaluate quantitatively how well a given theoretical model $\hat{s}_P(E)$ with parameters $P$ fits the data, and to determine the acceptable range for the parameters. These questions are conveniently treated by $\chi^2$ analyses; for a good discussion of these techniques see Lampton, Margon, and Bowyer (1976). Here we describe how one obtains the $\chi^2$ functionals in terms of the deviations between the flux estimates $\hat{s}_k$ and the model estimates $\hat{s}_{Pk}$, defined by

$$\hat{s}_{Pk} \equiv \int \hat{\delta}(E, E_k) s_P(E) dE \, .$$

First let us review how the $\chi^2$ functional is obtained in terms of the instrument channels. The counts $n_i$ from the $N$ instru-

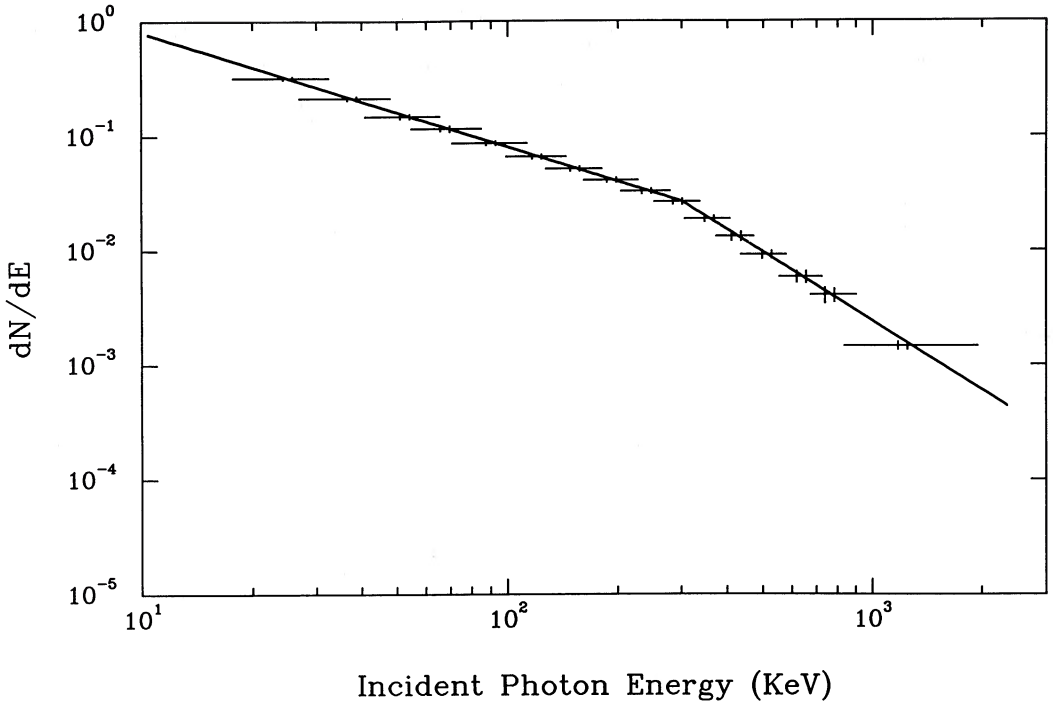


FIG. 7a

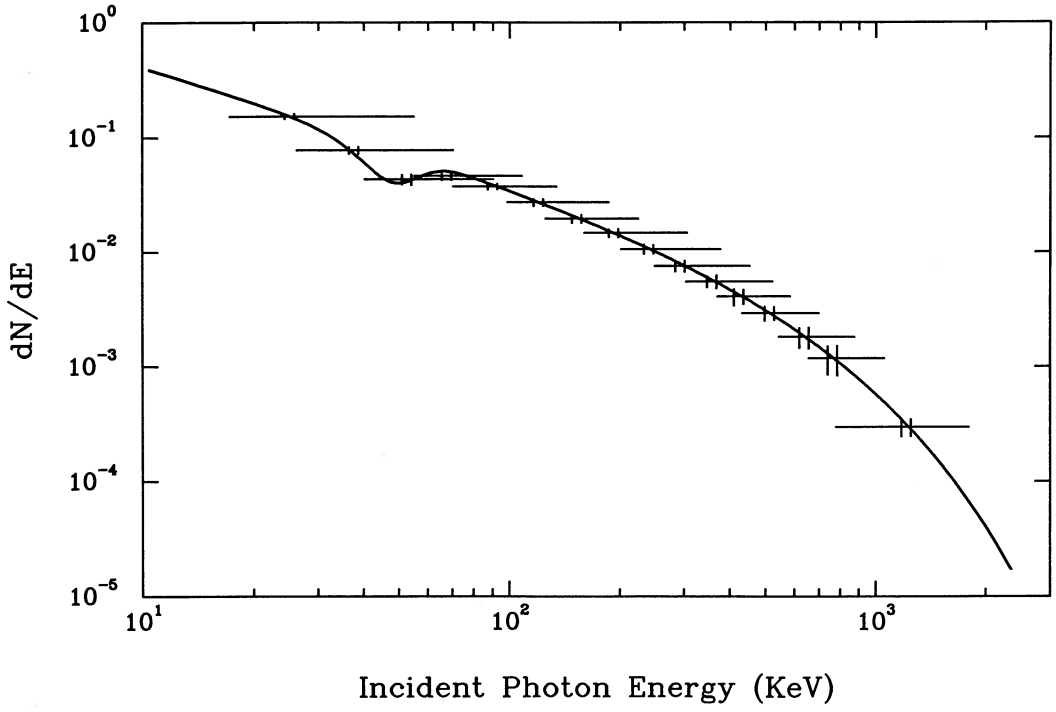


FIG. 7b

FIG. 7.—Monte Carlo results for Backus-Gilbert inversions of simulated data. (*a*) The result of 100 calculations as in Fig. 6*a*. The mean of the 100 estimates at each of the 16 energies is plotted, with a horizontal error bar indicating the mean of the 100 error bars. Two vertical error bars are plotted for each point. The one to the left of the mean estimate shows the mean of the propagated standard deviations of the estimates; the one to the right shows the standard deviation of the 100 estimates from their mean. (*b*), (*c*), (*d*). As in (*a*), for 100 calculations like those in Figs. 6*b*, 6*c*, and 6*d*, respectively.

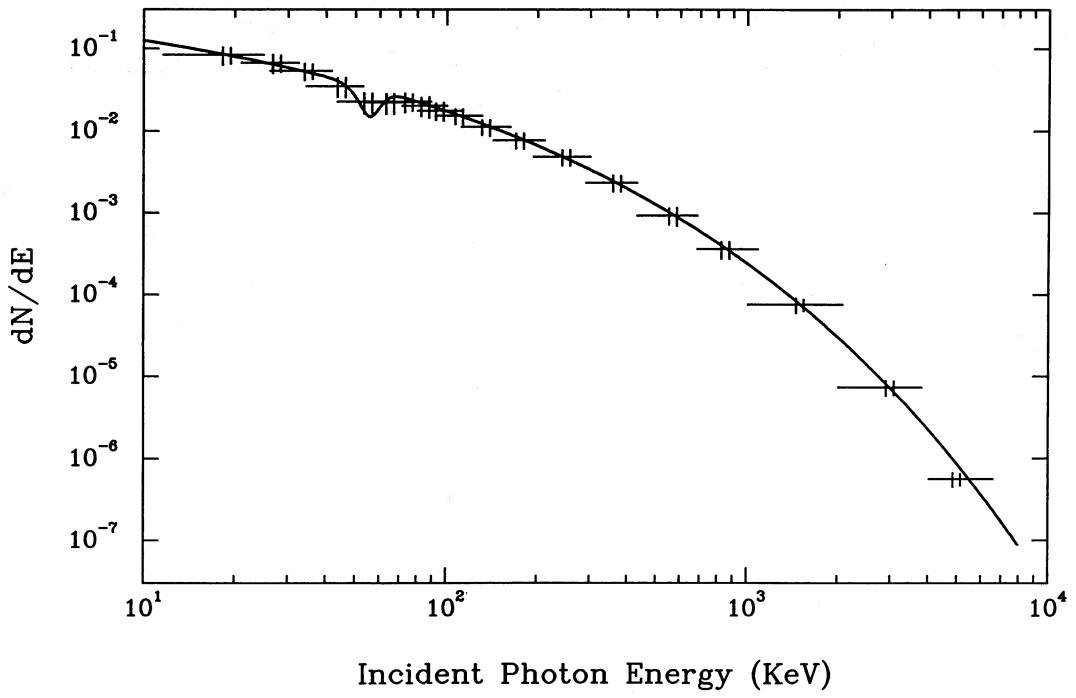


FIG. 7c

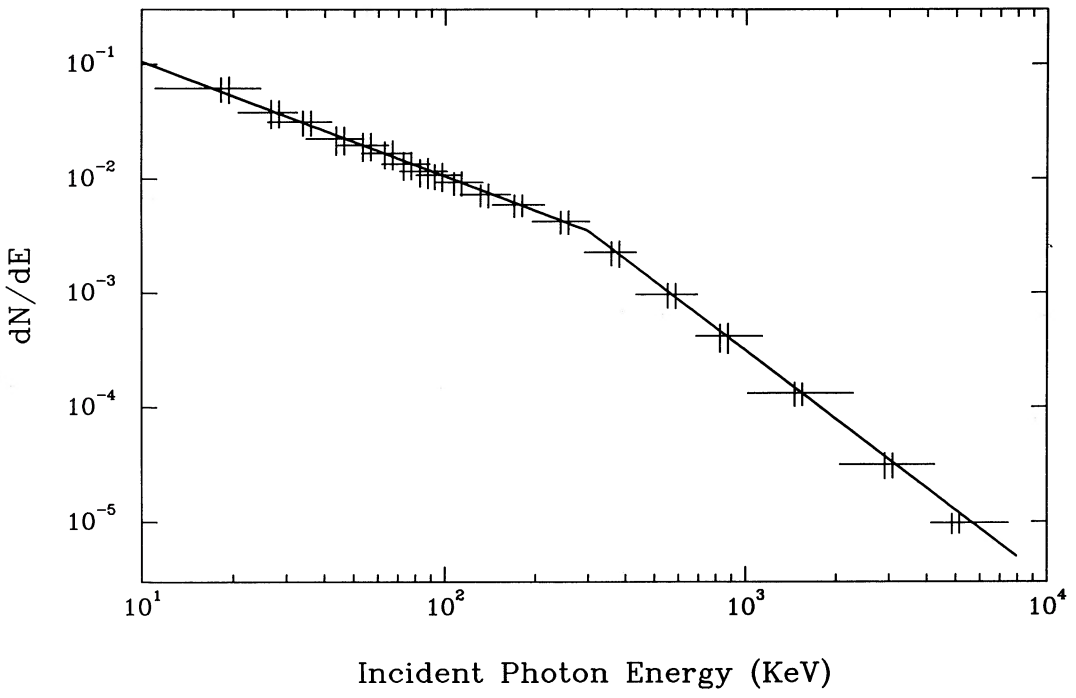


FIG. 7d

ment channels are uncorrelated so the variances $\sigma_i^2$ are independent and $\chi^2$ is

$$\chi^2 = \sum_{i=1}^{N} \frac{(\hat{n}_{Pi} - n_i)^2}{\sigma_i^2} . \tag{18}$$

Here $\hat{n}_{Pi}$ is the expected counts in the instrument channel for the theoretical model. This expression can be written in matrix notation. Let matrices $\boldsymbol{E}$ and $\boldsymbol{D}$ have components

$$E_{ij} = \sigma_i^2 \delta_{ij} \quad \text{(no sum)} , \quad D_{ij} = (\hat{n}_{Pi} - n_i)(\hat{n}_{P_j} - n_j) ;$$

then,

$$\chi^2 = \text{Tr} \, (\boldsymbol{E}^{-1}\boldsymbol{D}) , \tag{19}$$

where Tr indicates the matrix trace operation. Now consider these quantities for the Backus-Gilbert resolution functions. The transformation of $\boldsymbol{D}$ follows directly from equation (8):

$$\hat{D}_{ij} = (\hat{s}_{Pi} - \hat{s}_i)(\hat{s}_{Pj} - \hat{s}_j) = a_k(E_i)a_m(E_j)D_{km} ,$$

where a carat over a matrix denotes quantities derived from the Backus-Gilbert resolution functions, and repeated indices

are summed over. If we define a matrix $A$ with components $A_{ij} = a_j(E_i)$, and if $A$ is nonsingular we have

$$\hat{D} = ADA^{\mathrm{T}} .$$

Now let

$$\hat{E} \equiv AEA^{\mathrm{T}} .$$

The diagonal components of $\hat{E}$ are $\sigma_{\hat{s}}^2(E_j)$ as given by equation (14). We now find that

$$\begin{aligned} \hat{\chi}^2 &\equiv \mathrm{Tr}\,(\hat{E}^{-1}\hat{D}) = \mathrm{Tr}\,[(AEA^{\mathrm{T}})^{-1}(ADA^{\mathrm{T}})] \\ &= \mathrm{Tr}\,[(A^{\mathrm{T}})^{-1}E^{-1}DA^{\mathrm{T}}] \\ &= \chi^2 . \end{aligned} \tag{20}$$

The last step follows because the trace is invariant under a similarity transformation. The expression (19) for the $\chi^2$ functional is therefore invariant for the linear transformations that we are using. In component notation equation (20) is

$$\chi^2 = \sum_i \sum_j (\hat{E}^{-1})_{ij}(\hat{s}_{Pj} - \hat{s}_j)(\hat{s}_{pi} - \hat{s}_j) . \tag{21}$$

Now let us consider the practical advantages of using equation (21) rather than equation (18). The experimental group that measures a given $\gamma$-ray spectrum has the fullest possible knowledge of the response functions $R_i(E)$. For this group it is easier to use equation (18) than equation (21); they are entirely equivalent, and the former involves only a single summation. However, independent researchers not connected with the experimental groups can hardly be expected to implement equation (18). The problem is that the response functions $R_i(E)$ are complicated functions of energy which depend on the orientation of the source relative to the detector and which change during the lifetime of the instrument as a result of recalibration, gain shifts, and changes in the background. The independent researcher would have a phenomenally difficult time computing the model expectations $n_{Pi}$. In addition, the experimenter does not customarily publish the "raw" counts $n_i$, since these are not thought to be useful because of their complicated relationship to the incident spectrum.

The principal advantage of equation (21) is that the Backus-Gilbert resolution functions $\hat{\delta}(E, E_k)$ are sufficiently narrow that they can be usefully approximated by simple analytic functions for computing the model expectations $\hat{s}_{Pk}$. In addition, the transformed covariance matrix $\hat{E}$ may be well approximated by a diagonal matrix, simplifying the calculations even further.

We have performed numerous Monte Carlo experiments to test the reliability and stability of approximate versions of equation (21). We produced simulated data sets using a particular input spectrum, inverted these data, and then fit the resulting estimates to various models by minimizing approximations of $\chi^2$. We considered five levels of approximation. (1) We first performed the calculation of $\chi^2$ with equation (21) and the exact Backus-Gilbert resolution functions $\hat{\delta}_i(E)$. (We checked our numerics by assuring that this gave the same result as eq. [18] with the instrument response functions.) This calculation represents a bench mark with which we compare the approximations. (2) For the next level of approximation we approximated the Backus-Gilbert resolution functions $\hat{\delta}(E, E_k)$ by normalized "box" functions with the same width as the horizontal error bar associated with that resolution function, as defined by equations (16) and (17). (3) For the next coarser level of approximation, the resolution functions $\hat{\delta}(E, E_k)$ are treated as true $\delta$-functions, i.e., $\hat{\delta}(E, E_k) = \delta(E - E_k)$. (4) For the first three cases the full $\hat{E}$ matrices were used. For the next level of approximation, we used the box function approximation for the resolution functions, and took $\hat{E}^{-1}$ to be diagonal with components $(\hat{E}^{-1})_{ij} = [1/\sigma_{\hat{s}}^2(E_i)]\delta_{ij}$ (no sum; see eq. [14]). (5) Finally, we approximated $\chi^2$ using $\delta$-functions for the resolution functions and the diagonal $\hat{E}^{-1}$.

A typical example of the results of these calculations is presented in Figure 8. This figure shows the results of fits of various models to simulated data produced with an optically thin thermal bremsstrahlung spectrum (without Gaunt factor) of the form,

$$s(E) = AE^{-1} \exp(-E/E_{\mathrm{brems}}) , \tag{22}$$

with $E_{\mathrm{brems}} = 505$ keV and $A$ chosen to simulate bursts of varying fluence and hence varying statistical significance. The same 16 channel response system used in the calculations for Figure 6*a* was used to generate the simulated data with Poisson statistics. These data were then inverted, and the resulting estimates were fit to three model functions using the various $\chi^2$ approximations. The models included a broken power-law spectrum (BPL) of the form

$$s(E) = \begin{cases} A\left(\dfrac{E}{E_{\mathrm{break}}}\right)^{-\alpha_1}, & E \le E_{\mathrm{break}} , \\ A\left(\dfrac{E}{E_{\mathrm{break}}}\right)^{-\alpha_2}, & E > E_{\mathrm{break}} , \end{cases} \tag{23}$$

an optically thin thermal synchrotron-like spectrum (TS) of the form

$$s(E) = A \exp[-(4.5E/E_{\mathrm{sync}})^{1/3}] , \tag{24}$$

and the thermal bremsstrahlung spectrum (TB) of equation (22). The best-fit parameters were recorded, as was the lowest value of $\chi^2$ and the corresponding value of the function $Q(\chi^2 \mid \nu)$ which is the probability that the calculated $\chi^2$ would be exceeded by chance by a realization of the data based on the model, given $\nu$ degrees of freedom. This procedure was done 50 times for each of five different fluences, and at the end the average value of $Q$ and the best-fit parameters were calculated for each $\chi^2$ approximation at each fluence. The results are plotted as a function of fluence in Figure 8. Figure 8*a* shows the average $Q$ values at each fluence for fits using the exact $\chi^2$ (*points connected by the solid lines*), the box approximation (*dashed lines*), and the $\delta$-function approximation (*dotted lines*). The vertical error bars indicate the standard deviation of the 50 $Q$ values obtained using the exact $\chi^2$. Figure 8*b* shows the best-fit amplitudes of the fits to the three models using the same approximations, and Figure 8*c* shows the best-fit values of the parameters $\alpha_1$, $E_{\mathrm{sync}}$, and $E_{\mathrm{brems}}$; in these figures the error bars show the average 1 $\sigma$ confidence regions of the 50 true $\chi^2$ fits. Figures 8*d*, 8*e*, and 8*f* are similar, except that the dashed and dotted lines indicate results obtained using the box function and diagonal $\hat{E}$ approximation and the $\delta$-function and diagonal $\hat{E}$ approximation, respectively. We have also performed similar calculations with input spectra and detector parameters corresponding to Figures 6*b*, 6*c*, and 6*d*, with very similar results.

These figures show that all of the $\chi^2$ approximations considered provide acceptable estimates of the goodness of fit of a model and the best-fit parameters for most realistic applications. When the true $\chi^2$ is high, the estimates of $Q$ are inaccurate because the function $Q(\chi^2 \mid \nu)$ is sensitive to small variations in $\chi^2$. Fortunately this occurs only when $\chi^2$ is so

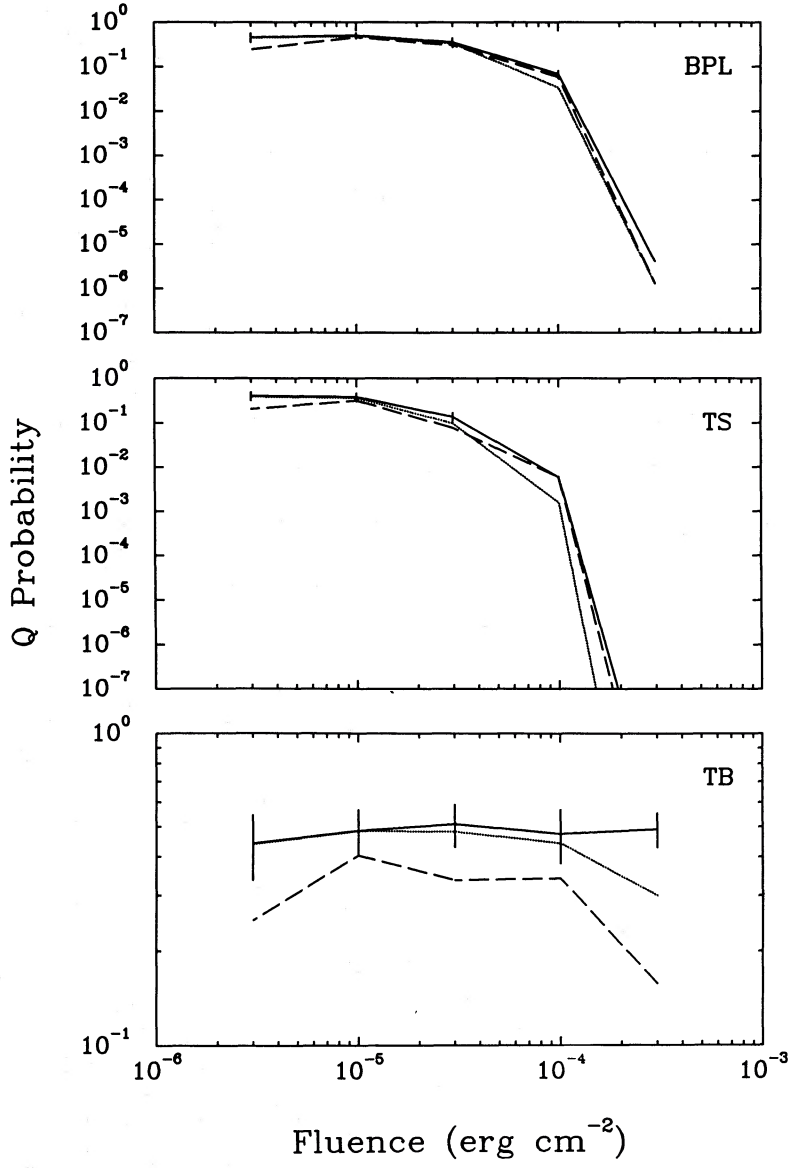


FIG. 8a

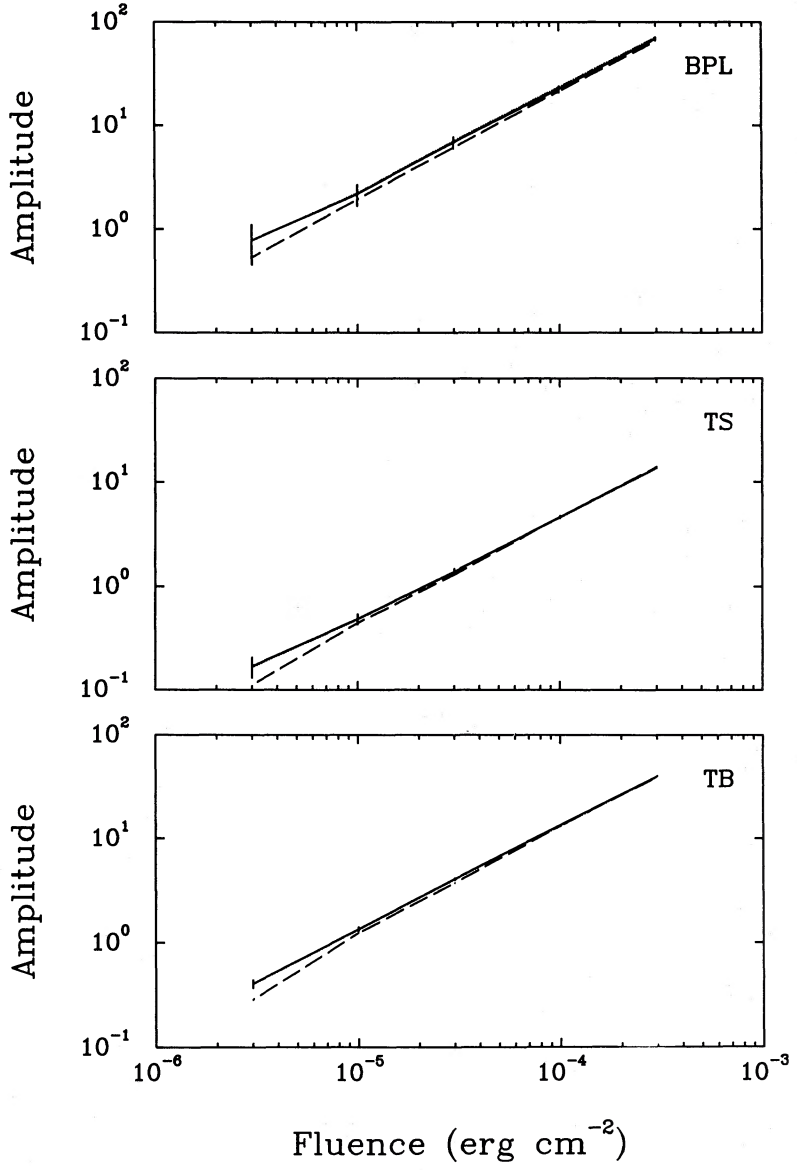


FIG. 8b

FIG. 8.—Monte Carlo results for model fitting approximations. The incident spectrum was thermal bremsstrahlung, and a 16 channel response system was used. (*a*) Average $Q$ probabilities for the best fits of 50 sets of simulated data at each of five fluences to three models, broken power-law (BPL), thermal synchrotron (TS), and thermal bremsstrahlung (TB). Solid line connects points obtained using the true $\chi^2$, dashed line connects points obtained using a box function approximation, and dotted line connects points obtained using a $\delta$-function approximation (the full covariance matrix is used for the two latter calculations). Error bars indicate the standard deviations of the values obtained using the true $\chi^2$. (*b*), (*c*). The best-fit parameters for the fits. Error bars indicate the average 1 $\sigma$ confidence regions of the 50 fits that used the true $\chi^2$. (*d*), (*e*), (*f*). As in (*a*), (*b*), and (*c*), exept here the dashed curves connect points obtained using the box function approximation with a diagonal covariance matrix, and the dotted curve connects points obtained using the $\delta$-function approximation with a diagonal covariance matrix. See text for details.

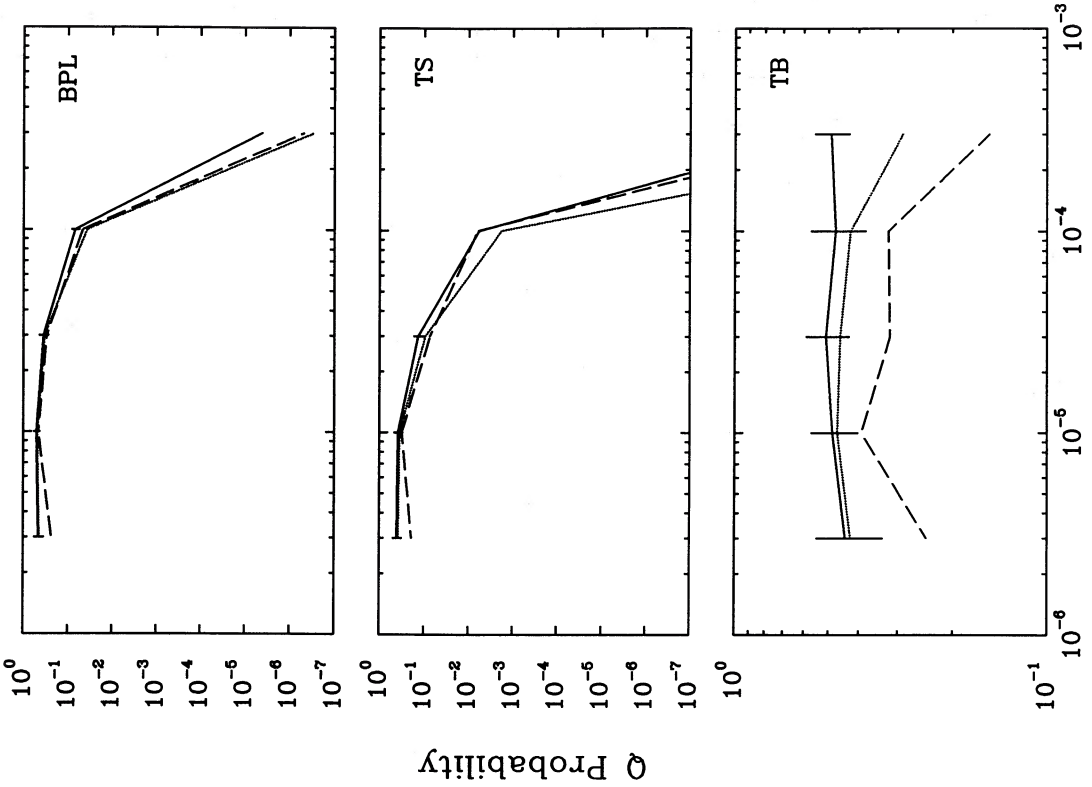


FIG. 8d

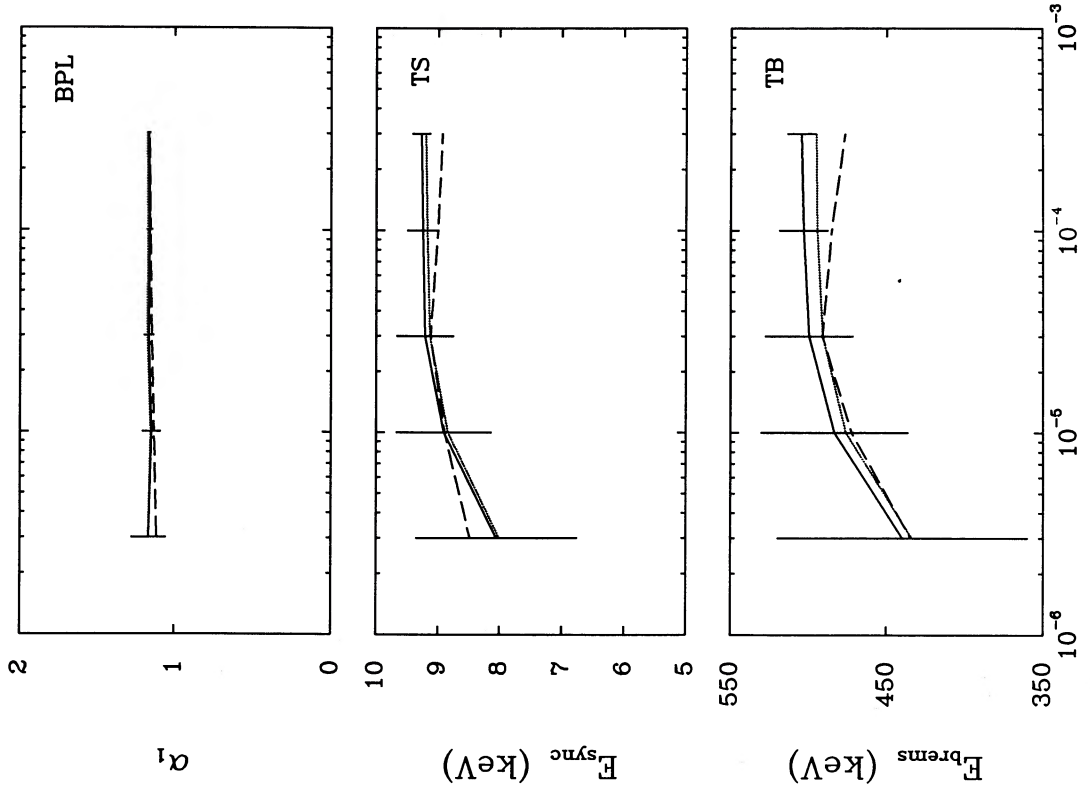


FIG. 8c

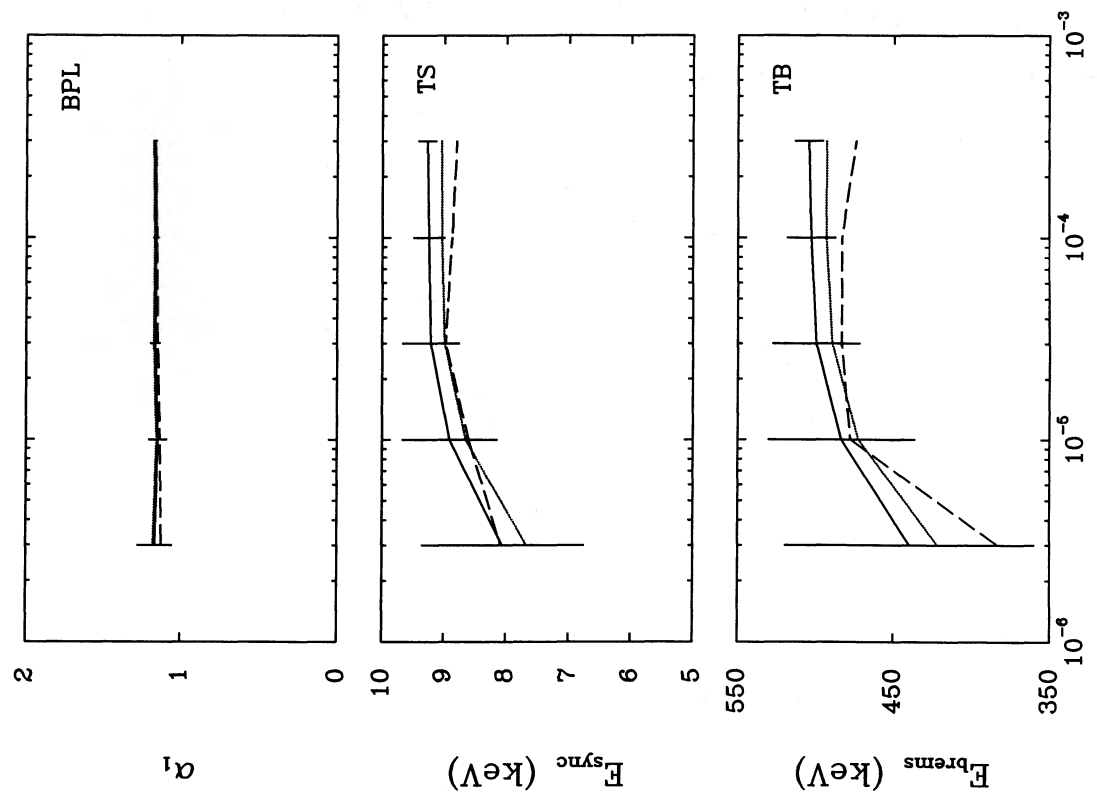


FIG. 8f

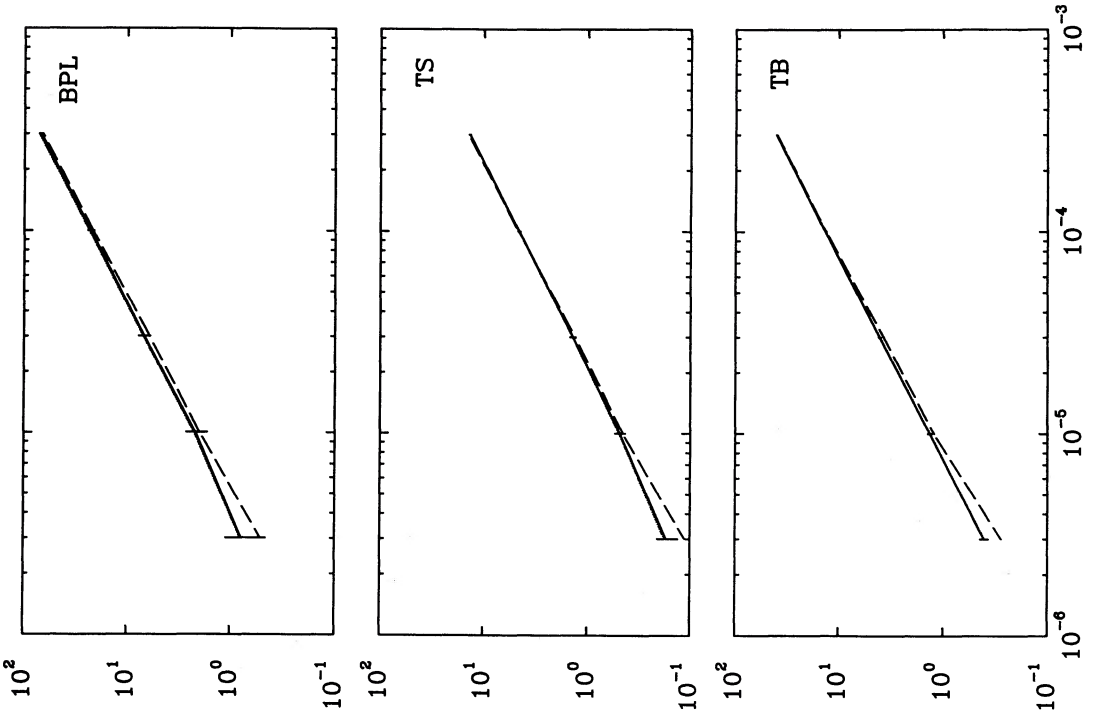


FIG. 8e

large that the model would be excluded anyway. All the approximations give estimates of the best-fit parameters that are within the statistical uncertainties except at the highest fluences considered. At the highest fluences a few of the estimated parameters are off by as much as three standard deviations.

Surprisingly, these calculations do not show any of the four approximations we considered to be preferable to the others. Simplicity would then dictate that the $\delta$-function approximation with diagonal $\hat{\boldsymbol{E}}$ be used. It is likely, however, that a box function would be preferable to a $\delta$-function if the model included narrow-line features that were comparable in width to the resolution of the estimates. The models we considered were smooth, so that the value of the model at any point is a good estimate of its average over a localized region, but this would not be true of a spectrum with a narrow-line feature.

In practice, the accuracy of approximating $\chi^2$ will have to be determined by the experimenter for each burst. The experimenter must fit a model to the inverted spectrum using the true $\chi^2$ and one or more of the approximations in order to determine if the statistical significance of the data permits $\chi^2$ to be accurately approximated.

### h) *Integration Limits*

A final technical point must be addressed: lower and upper limits for the energy integrals in equations (9), (10), (12), (13), (16), and (17) must be specified. Formally these limits are 0 and $\infty$, but the response functions are not known over this full interval. As mentioned above, channel response functions have fairly sharp low-energy cutoffs, so the lower limit is no problem; we set it to the energy where the first channel response function, $R_1(E)$, just starts to rise above zero. However, all channel response functions have significant amplitudes out to very high energies, so determining the upper limit is not as straightforward. Fortunately, typical GRB spectra fall rapidly enough at high energy that one does not err significantly by setting the upper limit to some arbitrary energy somewhat higher than the energies of interest, where the response functions are measurable. In the present calculations we used an upper limit equal to an energy about one channel width above the boundary of the peak of the highest channel response function.

This argument suggests a method by which the calculation of inverses may be shortened when there are many channels of spectral data, as may be the case for future generations of $\gamma$-ray spectrometers. Because GRB spectra fall rapidly with increasing energy, it may not be necessary to include the highest energy channels in the calculation of a spectral estimate, so that all of the sums presented in this paper could be truncated at some channel number less than $N$ when inverting at an energy well below the peaks of the channel response functions of higher numbered channels. A technique like this has been used to apply the Backus-Gilbert method to optical images, where typical response functions ("point spread functions") are well-localized (Saleh 1974).

## VI. CONCLUSION

In this paper, we have examined methods for analyzing $\gamma$-ray spectral data. The usual techniques have several weaknesses that unnecessarily complicate the interpretation of $\gamma$-ray spectra as they are usually presented. These weaknesses arise mainly from the model dependence of the current fitting estimates, which makes it difficult to identify model-independent characteristics of $\gamma$-ray spectra. In addition, the usual procedure for plotting spectra is somewhat arbitrary. For example, the horizontal error bars presented in published GRB spectral estimates are not directly related to the resolution of those estimates. These considerations have led us to explore the use of model-*independent* or "direct" inversion strategies for analyzing $\gamma$-ray spectral data.

We have examined the linear Backus-Gilbert inverse method because this method provides straightforward quantitative measures of the resolution and uncertainty of the estimates it produces. Our results indicate that the Backus-Gilbert inverse method can be successfully applied to $\gamma$-ray spectral data. The estimates it produces are quantitatively accurate and easy to interpret. The spectral energy resolution and uncertainty can be accurately calculated and straighforwardly displayed, and the estimates can be used for simple, accurate $\chi^2$ analysis of parametrized models.

We envision this method being used in the following manner. The observer of a $\gamma$-ray burst would invert the measured spectrum using the method described here, and then plot and tabulate the resulting spectral estimates. The tabulated information would include the energies $E_j$ of each estimate, the estimated spectrum value $\hat{s}(E_j)$, the propagated errors $\sigma_{\hat{s}}^2(E_j)$, and the bounds for the horizontal error bars for each estimate, $E_{\rm lo}$ and $E_{\rm hi}$. If the observer wishes to fit models to the data, this would be done as it always has been, using the $\chi^2$ equation (18) with the actual channel response functions. The results of the fit would be stated in terms of the parameters of the model and the $\chi^2$ of the fit; in addition, the best-fit models could be shown on the plot of the inverted spectrum. The observer could also repeat the model fitting using the tabulated spectral estimates and an approximate version of the transformed $\chi^2$ equation (21) to assess the accuracy of approximate model fitting for a particular burst. An independent investigator could then use the spectral estimates to assess alternative models using an approximate version of equation (21) to estimate $\chi^2$. We believe that the application of this method to GRB spectroscopy will help clarify the presentation and interpretation of $\gamma$-ray spectral data.

We would like to thank Jørgen Christensen-Dalsgaard who suggested we look at the Backus-Gilbert technique, and Ray Klebesadel, Nobuyuki Kawai, and especially Ed Fenimore, Bob Rosner, Bill Jeffrey, and Don Lamb for very valuable discussions. We would also like to thank the participants of the 1986 Taos workshop on gamma-ray stars, with special thanks to Steven Matz and U. Desai, and the members of the Gamma-Ray Observatory Burst and Transient Source Experiment collaboration, especially Gerald Fishman and William Paciesas, for many useful comments about earlier presentations of this work. We also thank Don Lamb for making available the computational facilities of the University of Chicago Physical Sciences Numerical Computation Laboratory used to perform the calculations presented here. One of us (T. J. L.) thanks the ESS-9 group of the Los Alamos National Laboratory for their hospitality to him as a graduate research assistant during the time when much of this work was done. This work was supported in part by NASA grants NGT-50189 and NAGW-830, and carried out in part under the auspices of the US Department of Energy.

## APPENDIX A

## GAMMA-RAY SPECTROMETER RESPONSE FUNCTIONS

This Appendix describes the $\gamma$-ray spectrometer response function model used for the calculations presented in the body of this paper. This model is based on calculations performed by Berger and Seltzer (1972) and Israel, Lier, and Storm (1971), and on measurements reported by Matz (1986*a*). This model cannot replace response functions generated by Monte Carlo simulations of a specific detector, but it accurately displays all of the salient features of such calculations.

The heart of most $\gamma$-ray spectrometers is an inorganic scintillator crystal which produces a very brief flash of optical light (a "scintillation") when a $\gamma$-ray interacts with it. The intensity of this optical scintillation is proportional to the energy lost by the $\gamma$-ray in the detector, and thus can be used as a measure of this energy loss. Although some $\gamma$-ray spectrometers use semiconducting detectors rather than inorganic scintillators, their operation is totally analogous to that of inorganic scintillators. The remainder of this Appendix will discuss the response function of an inorganic scintillator.

The detection of a $\gamma$-ray by a scintillator involves two distinct processes. First, the $\gamma$-ray loses some or all of its energy to electrons in the crystal. Second, the energy of these electrons is converted to a pulse of optical light through the (poorly understood) scintillation process, and the intensity of this pulse (the pulse height) is measured by appropriate instruments.

The $\gamma$-ray interaction can be characterized by an energy-dependent efficiency, $\eta(E)$, and an *energy deposition spectrum*, $D(E, \epsilon)$. The efficiency $\eta(E)$ is the probability that an incident $\gamma$-ray of energy $E$ will interact with the detector; it is also a function of the angle of incidence. The energy deposition spectrum $D(E, \epsilon)d\epsilon$ is the probability that a $\gamma$-ray of energy $E$ will deposit an amount of energy between $\epsilon$ and $\epsilon + d\epsilon$ when it interacts with the detector. The scintillation and detection processes can be collectively characterized by a *scintillator resolution function*, $G(\epsilon, h)$, defined so that $G(\epsilon, h)dh$ is the probability that an amount of energy $\epsilon$ deposited in the crystal gives rise to a pulse with a measured pulse height between $h$ and $h + dh$. In terms of the efficiency, the energy deposition spectrum, and the scintillator resolution function, the response function described in § II can be written

$$R(E, h) = \eta(E) \int_0^\infty D(E, \epsilon)G(\epsilon, h)d\epsilon \, . \tag{A1}$$

Since $\gamma$-rays can interact with the detector via several different processes, the energy deposition spectrum is quite complicated. The most important interaction processes are the photoelectric effect and $e^+$–$e^-$ pair production, by which an incident photon loses all of its energy to energetic particles, and Compton scattering, by which the incident photon may lose only a fraction of its energy to energetic electrons. The energy deposition spectrum that results from these processes is further complicated by the escape of the energetic particles or secondary photons (from $e^+$ annihilation or K-shell fluorescence) produced in the interaction, by Compton-scattered photons interacting further with the scintillator, and by bremsstrahlung losses. It is also important to realize that all of these processes occur in the material around the scintillator as well as in the scintillator itself, and that the incident spectrum can be contaminated by radiation scattered into the scintillator from its environment.

Despite the phyical complexity of the $\gamma$-ray and energetic particle interactions that produce the energy deposition spectrum, it can be well approximated by a set of four lines and a continuum (Berger and Seltzer 1972):

$$D(E, \epsilon) = C(E, \epsilon) + P_0 \, \delta(\epsilon - E) + P_1 \, \delta(\epsilon - E + mc^2) + P_2 \, \delta(\epsilon - E + 2mc^2) + P_3 \, \delta(\epsilon - E_f) \, . \tag{A2}$$

The lines are weighted by the relative probabilities of their occurrence. Thus $P_0$ is the probability that an incident photon will deposit *all* of its energy upon interaction with the detector. The quantity $P_0$ is called the photofraction, and the $P_0$ term in equation (A1) gives rise to a peak in the continuous counts spectrum that is called the photopeak. The quantities $P_1$ and $P_2$ are the probabilities that an incident photon will interact via pair production with the ensuing escape of one or two pair quanta, respectively, and $P_3$ is the probability that a photon deposits all of its energy except for an amount $E_f$ due to the escape of a K-shell fluorescence photon. The function $C(E, \epsilon)$, which represents the Compton continuum, can be well represented by a weak line feature at the Compton edge (at an energy $E_C = 2E^2/[mc^2 + 2E]$, corresponding to 180° scattering) and a flat continuum at energies below this feature (Matz 1986*a*):

$$C(E, \epsilon) = P_4 \, \delta(\epsilon - E_C) + \frac{P_5}{E_C} \, \Theta(E_C - \epsilon) \, . \tag{A3}$$

Here $\Theta(E)$ is the unit step function. All the coefficients $P_n$ are functions of energy, and $\sum_{n=0}^{5} P_n = 1$. We use values calculated by Berger and Seltzer (1972) and reported by Matz (1986*a*) for a 3″ × 3″ cylindrical NaI scintillator with the $\gamma$-ray incident direction parallel to the crystal symmetry axis; these values are supplemented at low energies by the calculations of Israel, Lior, and Storm (1971). The resulting response function is similar to that used to invert $\gamma$-ray burst data from the *Solar Maximum Mission* Gamma-Ray Spectrometer (the *SMM* experiment, however, used active Compton shielding to reduce the amplitude of the Compton continuum). It should also be similar to the response function of the *Konus* instrument, which used somewhat smaller NaI scintillators (Mazets and Golenetskii 1981).

The scintillator resolution function is well approximated by a Gaussian with $\epsilon$-dependent width $\sigma(\epsilon)$,

$$G(\epsilon, h) = \frac{1}{\sigma(\epsilon)\sqrt{2\pi}} \exp\left[ -\frac{(h - \bar{h})^2}{2\sigma^2(\epsilon)} \right] . \tag{A4}$$

The materials used as scintillators in $\gamma$-ray spectroscopy are chosen because to a good approximation the mean of this Gaussian is proportional to the energy deposited. Thus by appropriate scaling we may put $\bar{h}(\epsilon) = \epsilon$. It is this property of scintillators that allows

counts spectra to be displayed as a function of an independent variable with units of energy. We emphasize, however, that $h$ is a measure of the energy *lost* by the photon to the detector, and not the full energy of the photon that produced the count. For the sake of generality, we include the function $\bar{h}(\epsilon)$ explicitly in the calculations presented in this Appendix. For the calculations that appear in the body of this paper, we set $\bar{h}(\epsilon) = \epsilon$.

Empirically, the width of the Gaussian in equation (A2) obeys a power-law,

$$\sigma(\epsilon) = \sigma(\epsilon_0)\left(\frac{\epsilon}{\epsilon_0}\right)^n \tag{A5}$$

We use the value $\epsilon_0 = 662$ keV, $\sigma(\epsilon_0) = 29$ keV, and $n = -0.34$ cited by Berger and Seltzer (1972).

Using equations (A2)–(A4), we perform the integral in equation (A1) to arrive at an analytical approximation for a scintillator response function. All the $\delta$-function terms in the energy deposition spectrum can be trivially integrated. Integrating the step function term in the Compton continuum is not as straightforward, because of the $\epsilon$-dependent functions $\bar{h}(\epsilon)$ and $\sigma(\epsilon)$ in the scintillation resolution function. However, the nonlinearity represented by $\bar{h}(\epsilon)$ is only significant at low energies, where the effects of Compton scattering are small ($<1\%$), so we may set $\bar{h}(\epsilon) = \epsilon$. Also, because the only "feature" in the step function that is being smoothed by the resolution function is at the Compton edge, we may set $\sigma(\epsilon) = \sigma(E_C)$ for this term in the integral. We then find

$$R(E, h) = \eta(E)\left[\sum_{n=0}^{4} \frac{P_n(E)}{\sigma(\epsilon_n)\sqrt{2\pi}} \exp\left\{-\frac{[h - \bar{h}(\epsilon_n)]^2}{2\sigma^2(\epsilon_n)}\right\} + \frac{P_5(E)}{2\epsilon_4}\left\{\mathrm{erf}\left[\frac{h}{\sigma(\epsilon_4)\sqrt{2}}\right] - \mathrm{erf}\left[\frac{h - \epsilon_4}{\sigma(\epsilon_4)\sqrt{2}}\right]\right\}\right], \tag{A6}$$

where we have defined characteristic energy losses, $\epsilon_n(E)$, according to

$$\epsilon_0 = E\,, \quad \epsilon_1 = E - mc^2\,, \quad \epsilon_2 = E - 2mc^2\,, \quad \epsilon_3 = E - E_f\,, \quad \epsilon_4 = E_C = \frac{2E^2}{mc^2 + 2E}\,. \tag{A7}$$

The channel response functions may be found by integrating this according to equation (7), giving

$$R_i(E) = \frac{\eta(E)}{2}\sum_{n=0}^{4}\left[P_n \,\mathrm{erf}\,(e_{i,n}) - P_n \,\mathrm{erf}\,(e_{i+1,n})\right] + \frac{\eta(E)P_5(E)\sigma(\epsilon_4)}{\sqrt{2}\,\epsilon_4}\sum_{n=5}^{6}(-1)^{n+1}\left\{e_{i+1,n}\,\mathrm{erf}\,(e_{i+1,n}) - e_{i,n}\,\mathrm{erf}\,(e_{i,n}) + \frac{1}{\sqrt{\pi}}\left[\exp\left(-e_{i+1,n}^2\right) - \exp\left(-e_{i,n}^2\right)\right]\right\} \tag{A8}$$

where we have defined dimensionless scaled pulse heights, $e_{i,n}(E)$, according to

$$e_{i,n} \equiv \begin{cases} \dfrac{\epsilon_n - h_i}{\sigma(\epsilon_n)\sqrt{2}} & n \le 4, \\ \dfrac{\epsilon_n - h_i}{\sigma(\epsilon_4)\sqrt{2}} & n = 5,\ 6, \end{cases} \tag{A9}$$

with $\epsilon_5 = 0$ and $\epsilon_6 = \epsilon_4$.

## APPENDIX B

## DERIVATION OF THE BACKUS-GILBERT INVERSE

Here we derive an expression for the coefficients $a_i$ which minimize the weighted sum of spread and variance (eq. [15]), subject to the normalization constraint equation (13) (Backus and Gilbert 1970). Using the definition of the resolution function given in equation (11), equation (12) for the spread can be written,

$$r(E_j) = a_i S_{ik} a_k\,, \tag{B1}$$

where we define the *spread matrix* according to

$$S_{ik} \equiv \int (E - E_j)^2 R_i(E) R_k(E) dE\,. \tag{B2}$$

In equation (B1) and throughout this Appendix, repeated indices indicate summation, unless otherwise noted; the sum should be taken over the range of channels specified in § IV*a*. The propagated variance (14) can be written

$$\sigma_s^2 = a_i V_{ik} a_k\,, \tag{B3}$$

where $V_{ik}$ is the covariance matrix for the data. Equation (14) is recovered when $V_{ik}$ is diagonal and $V_{ii} = \sigma_{ni}^2$ (no sum).

Using equations (B1) and (B3), the weighted sum (eq. [15]) becomes the matrix equation,

$$w(E_j, \theta) = a_i W_{ik} a_k\,, \tag{B4}$$

where

$$W_{ik} \equiv S_{ik} \cos\theta + vV_{ik} \sin\theta \,. \tag{B5}$$

The normalization constraint (eq. [13]) can be similarly rewritten as

$$a_i u_i = 1 \,, \tag{B6}$$

where

$$u_i \equiv \int R_i(E) dE \,. \tag{B7}$$

The constant $v$ in equations (B5) and (15) determines how $\theta$ parametrizes the set of coefficients $a_i$. The exact value of $v$ is unimportant since all positive $v$ give the same tradeoff curve. The value of $v$ should make $vV_{ij}$ comparable to $S_{ij}$. We simply take $v$ equal to a fixed fraction of the ratio of the traces of the spread and covariance matrices, $v = (S_{ii}/15)/V_{ii}$.

The Backus-Gilbert inverse is obtained by choosing the $a_i$ which minimize equation (B4) subject to the constraint (B6). Using a Lagrange multiplier $\lambda$, this corresponds to choosing the $a_i$ and $\lambda$ that minimize

$$w'(E_j, \theta) = a_i W_{ik} a_k + \lambda a_i u_i - \lambda \,. \tag{B8}$$

The minimization equation is

$$\frac{\partial w'}{\partial a_j} = a_i W_{ij} + W_{jk} a_k + \lambda u_j = 0 \,. \tag{B9}$$

Minimizing $w'$ with respect to $\lambda$ reproduces equation (B6).

As both the spread and variance matrices are symmetric, equation, (B9) becomes

$$W_{jk} a_k = -\frac{\lambda}{2} u_j \,. \tag{B10}$$

The $a_i$ can now be determined by solving the simultaneous equations (B10) and (B6). In matrix notation, these equations read

$$\boldsymbol{W}\boldsymbol{a} = -\frac{\lambda}{2}\boldsymbol{u} \,, \tag{B11}$$

and

$$\boldsymbol{u} \cdot \boldsymbol{a} = 1 \,. \tag{B12}$$

Since the matrix $\boldsymbol{W}$ is positive-definite and symmetric, it has an inverse, and we can solve equation (B11) for $\boldsymbol{a}$, giving

$$\boldsymbol{a} = -\frac{\lambda}{2}\boldsymbol{W}^{-1}\boldsymbol{u} \,. \tag{B13}$$

The Lagrange multiplier may be found by dotting this with $\boldsymbol{u}$ and using equation (B12), which gives

$$\lambda = -2(\boldsymbol{u}\boldsymbol{W}^{-1}\boldsymbol{u})^{-1} \,. \tag{B14}$$

Thus the final solution is

$$\boldsymbol{a} = \frac{\boldsymbol{W}^{-1}\boldsymbol{u}}{\boldsymbol{u}\boldsymbol{W}^{-1}\boldsymbol{u}} \,, \tag{B15}$$

or, more explicitly,

$$a_i = \frac{(W^{-1})_{ij} u_j}{u_j (W^{-1})_{jk} u_k} \,. \tag{B16}$$

Using this, the spectral estimate is calculated from the data according to equation (11), and the spread and variance of this estimate can be calculated according to equations (B1) and (B3).

Note that the inverse coefficients $a_i$ depend on the data being inverted only through the covariance matrix $V_{ik}$. This has practical consequences when several inverted spectra of the same source must be calculated, because the spread matrix $S_{ik}$ need be calculated only once. Even greater computational economy can be realized if, in addition, a tradeoff parameter of zero is being used (as suggested in § IV*c*), for then the $a_i$ will be completely independent of the data and need be calculated only once. These considerations are relevant for time-resolved spectroscopy, because time binning affects the calculation of an inverse only through its possible affect on the variance of the data.

REFERENCES


Backus, G., and Gilbert, F. 1970, *Phil. Trans. Roy. Soc London, A*, **266**, 123.
Barat, C. 1983, in *Position-Electron Pairs in Astrophysics*, ed. M. L. Burns, A. K. Harding, and R. Ramaty (New York: AIP), p. 54.
Berger, M. J., and Seltzer, S. M. 1972, *Nuc. Instr. Meth.*, **104**, 317.
Chambless, D. A., and Broadway, J. A. 1981, *Nucl. Instr. Meth.*, **179**, 563.
Dennis, B. R., Frost, K. J., Kiplinger, A. L., Orwig, L. E., Desai, U., and Cline, T. L. 1981, in *Gamma-Ray Transients and Related Astrophysical Phenomena*, ed. R. E. Lingenfelter, H. S. Hudson, and D. M. Worrall (New York: AIP), p. 153.
Evans, W. D., *et al.* 1980, *Ap. J. (Letters)*, **237**, L7.
Fenimore, E. E. 1986, private communication.
Fenimore, E. E., Laros, J. G., Klebesadel, R. W., and Stockdale, R. E. 1982*a*, in *Gamma-Ray Transients and Related Astrophysical Phenomena*, ed. R. E. Lingenfelter, H. S. Hudson, and D. M. Worrall (New York: AIP), p. 201.
Fenimore, E. E., Klebesadel, R. W., and Laros, J. G. 1983, in *Gamma-Ray Astronomy in Perspective of Future Space Experiments* (London: Pergamon), p. 243.
Fenimore, E. E., Klebesadel, R. W., Laros, J. G., Stockdale, R. E., and Kane, S. R. 1982*b*, *Nature*, **297**, 665.
Frieden, B. R. 1975, in *Picture Processing and Digital Filtering*, ed. T. S. Huang (New York: Springer), p. 177.
———. 1976, in *Proc. Image Analysis and Evaluation*, ed. R. Shaw (Toronto: Society of Photographic Scientists and Engineers), p. 261.
———. 1983, *J. Opt. Soc. Am.*, **73**, 927.
———. 1984, in *Deconvolution with Applications to Spectroscopy*, ed. P. A. Jansson (Orlando: Academic), p. 327.
Frieden, B. R., and Swindell, W. 1976, *Science*, **191**, 1237.
Frieden, B. R., and Wells, D. C. 1978, *J. Opt. Soc. Am.*, **68**, 93.
Gilbert, F. 1971, *Geophys. J.R.A.S.*, **23**, 125.
Harding, A. K., Petrosian, V., and Teegarden, B. J. 1986, in *Gamma-Ray Bursts*, ed. E. P. Liang and V. Petrosian (New York: AIP), p. 75.
Heuter, G. J. 1984, in *High Energy Transients in Astrophysics*, ed. S. E. Woosley (New York: AIP), p. 373.
Hurley, K. 1986, unpublished.
Israel, H. I. Lier, D. W., and Storm, E. 1971, *Nucl. Instr. Meth.*, **91**, 141.
Jansson, P. A. 1984, in *Deconvolution with Applications to Spectroscopy*, ed. P. A. Jannsson (Orlando: Academic), p. 96.
Jeffrey, W., 1988, *Ap. J.*, **327**, 987.
Jeffrey, W., and Rosner, R. 1986, *Ap. J.*, **310**, 463.
Kahn, S. M., and Blissett, R. J. 1980, *Ap. J.*, **288**, 417.
Kahn, S. M., Charles, P. A., Bowyer, S., and Blissett, R. J. 1980, *Ap. J. (Letters)*, **242**, L19.
Klebesadel, R., Strong, I., and Olson, R. 1973, *Ap. J. (Letters)*, **182**, L85.
Lampton, M., Margon, B., and Bowyer, S. 1976, *Ap. J.*, **208**, 177.
Laros, J. G., Evans, W. D., Fenimore, E. E., and Klebesadel, R. W. 1983, *Ap. Space Sci*, **88**, 243.
Matz, S. M. 1986*a*, PhD. thesis, University of New Hampshire.
———. 1986*b*, private communication.
Mazets, S. V., Aptekar, R. L., Golenetskii, E. P., Guryan, Yu. A., Ilyinskii, V. L., and Panov, V. N. 1983, in *Positron-Electron Pairs in Astrophysics*, ed. M. L. Burns, A. K. Harding, and R. Ramaty (New York: AIP), p. 36.
Mazets, E. P., and Golenetskii, S. V. 1981, *Ap. Space Sci.*, **75**, 47.
Mazets, E. P., Golenetskii, S. V., Aptekar, R. L., Guryan, Yu. A., and Ilyinskii, V. L. 1981, *Nature*, **290**, 378.
Menke, W. 1984, *Geophysical Data Analysis: Discrete Inverse Theory* (Orlando: Academic).
Narayan, R., and Nityanada, R. 1986, *Ann. Rev. Astr. Ap.*, **24**, 127.
Nolan, P. L., Share, G. H., Chupp, E. L., Forrest, D. J., and Matz, S. M. 1984, *Nature* **311**, 360.
Parker, R. L. 1977, *Ann. Rev. Earth Planet. Sci.*, **5**, 35.
Saleh, B. E. A. 1974, *App. Optics*, **13**, 1833.
Sarkar, T. K., Weiner, D. D., and Jain, V. K. 1981, *IEEE Trans. Antennas and Propagation*, **AP-29**, 373.
Schwarz, V. J. 1979, in *Image Formation from Coherence Functions in Astronomy*, ed. C. van Schooneveld (Dordrecht: Reidel), p. 261.
Teegarden, B. J. 1984, in *High Energy Transients in Astrophysics*, ed. S. E. Woosley (New York: AIP), p. 352.
Teegarden, B. J., and Cline, T. 1980, *Ap. J. (Letters)*, **236**, L67.



RICHARD I. EPSTEIN: ESS-9, MS D436, Los Alamos National Laboratory, Los Alamos, NM 87545

THOMAS J. LOREDO: Astronomy and Astrophysics Center, University of Chicago, 5640 South Ellis Avenue, Chicago, IL 60637