

# Documentation of C Inversion Library

Paul O'Brien (paul.obrien@aero.org)

March 2, 2009

## Contents

### 1 Introduction

This document describes the functions provided by the inversion library “invlib.so” or “invlib.dll” for determination of a best-fit energy spectrum from a multi-channel measurement.

### 2 Changes

- 3/2/2009 fixed bug in specinv that caused PLE to only do PL.
- 2/17/2009 Created `ana_spec_inv_multi`
- 2/13/2009 Changed `void *outFilePtr` to `char *outFile`, expanded meaning of verbose input code to allow output to a file (codes 4 and 5), and added a new output error value (-502) for unable to open the output file. (Thanks for Reiner Friedel for the suggestion)
- 9/2/2008 Added double relativistic Maxwellian (RM2) and Appendix ??.
- 8/11/2008 Added power law with exponential tail (PLE) analytical spectrum to `ana_spec_inv`. Note: now `real_params` must have length 3 to use PLE.
- 4/25/2008 Library and documentation created and released.

### 3 Library Functions `ana_spec_inv` and `ana_spec_inv_multi`

The library provides a spectral inversion function, `ana_spec_inv`, with the following C prototype (found in “invlib.h”):

```
int ana_spec_inv(const double *y, const double *dy,
                const double *Egrid, const double *H, const double *b,
                const long int *int_params, const double *real_params,
                char *outFile, double *Eout,
                double *flux, double *dlogflux);
```

This function takes a set of counts `y`, estimated relative error on those counts `dy`, and measurement matrix `H` defined on energy grid `Egrid`, with estimated background counts `b`, and returns `flux` on output energy grid `Eout`. The user can select up to 3 different analytical functions that will be individually fit to the data, and then a “weighted expected value” method will be used to combine multiple analytical fits. The user can also control which nonlinear minimization solver is used, and whether and where progress messages are printed. The function returns a code that indicates success or one of several failure modes. There is a multi-case version `ana_spec_inv_multi` described at the end of this section.

### 3.1 Calculations Performed

In continuous form, the function `ana_spec_inv` solves

$$\vec{y} \approx \vec{\lambda} = T \int_0^\infty \vec{G}(E) f(E) dE + \vec{b}, \quad (1)$$

where  $\vec{y}$  is a vector of observed counts,  $\vec{\lambda}$  is a vector of expected counts,  $T$  is the integration time,  $\vec{G}$  is a vector of energy-geometric factors (response functions), and  $f(E)$  is the differential particle flux at energy  $E$ , and  $\vec{b}$  is a vector of expected background counts. In this equation, all vectors have length  $N_y$ , the number of energy channels. *Throughout this subsection, we use the one-based linear algebra convention for vector and matrix subscripts.*

#### 3.1.1 Numerical Problem to be Solved

The first step toward a numerical solution to (??) is to discretize the integral:

$$\vec{y} \approx \vec{\lambda} = \underline{\underline{H}} \vec{f} + \vec{b}, \quad (2)$$

$$H_{ij} \approx T G_i(E_j) \Delta E_j, \quad (3)$$

$$f_j = f(E_j). \quad (4)$$

Where  $\vec{f}$  is now defined on a grid in energy, with  $N_E$  points, and  $\underline{\underline{H}}$  has size  $N_y \times N_E$ . For a sufficiently fine grid in  $\Delta E_j$ , the approximation for  $H_{ij}$  above should work well. However, slightly better choices are available if the grid is not coarse or unevenly spaced. It is usually worthwhile to use at least a trapezoidal integral, regardless of the grid:

$$\Delta E_j = \begin{cases} E_j = (E_{j+1} - E_j)/2 & j = 1 \\ E_j = (E_j - E_{j-1})/2 & j = N_E \\ E_j = (E_{j+1} - E_{j-1})/2 & \text{otherwise} \end{cases} \quad (5)$$

#### 3.1.2 Penalty Functions

The next step is to define a “penalty” function based on the likelihood of observing counts  $\vec{y}$  given expected counts  $\vec{\lambda}$ . The observations differ from the expected counts due to a variety of possible error processes. The most common two errors are Poisson error and calibration error.

**3.1.2.1 Poisson Error Penalty Function.** The Poisson error arises from the finite number of counts observed, and is often referred to as “counting statistics” or “counting error.” The likelihood function, or the probability of observing  $y$  given  $\lambda$  from a Poisson counting process, is given by:

$$p^{(\text{Poisson})}(y|\lambda) = \frac{\lambda^y e^{-\lambda}}{y!}. \quad (6)$$

The penalty function is given by the negative natural log of the likelihood function, with terms that do not depend on  $\lambda$  removed:

$$\ell^{(\text{Poisson})}(\lambda) = -\ln p(y|\lambda)^{(\text{Poisson})} = \lambda - y \ln \lambda + \text{constants}. \quad (7)$$

We remove the terms that do not depend on  $\lambda$  because they will not affect our solution (which, varies  $\lambda$  by varying  $\vec{f}$  to minimize the penalty function).

We will eventually also need the first and second derivatives with respect to  $\lambda$ :

$$\frac{d\ell^{(\text{Poisson})}}{d\lambda} = 1 - y/\lambda, \quad (8)$$

$$\frac{d^2\ell^{(\text{Poisson})}}{d\lambda^2} = y/\lambda^2. \quad (9)$$

**3.1.2.2 Relative Error Penalty Function.** The relative error process arises from incomplete knowledge of the instrument calibration, either due to incomplete preflight testing, or changes in the instrument that occur during launch or on orbit, or due to assumptions made about the angular response of a wide-angle or omnidirectional sensor relative to the local angular flux distribution.

Relative errors are assumed to have a Gaussian shape in the log of the counts:

$$p(y|\lambda)^{(\text{Relative})} = \frac{\exp[-((\ln y - \ln \lambda)/\delta y)^2/2]}{\sqrt{2\pi}\delta y}. \quad (10)$$

The relative error penalty function is:

$$\ell^{(\text{Relative})}(\lambda) = -\ln p(y|\lambda)^{(\text{Relative})} = ((\ln y - \ln \lambda)/\delta y)^2/2 + \text{constants}. \quad (11)$$

The derivatives with  $\lambda$  are:

$$\frac{d\ell^{(\text{Relative})}}{d\lambda} = (\ln \lambda - \ln y)/(\delta y)^2/\lambda, \quad (12)$$

$$\frac{d^2\ell^{(\text{Relative})}}{d\lambda^2} = (1 + \ln y - \ln \lambda)/(\lambda\delta y)^2. \quad (13)$$

**3.1.2.3 Selecting a Penalty Function for Each  $y$ .** In order to keep things simple, we select either the Poisson or relative error penalty function for each  $y$ , whenever the relative Poisson counting error ( $1/\sqrt{y}$ ) is larger than the Gaussian relative error ( $\delta y$ ).

$$\ell_i = \begin{cases} \ell^{(\text{Poisson})} & y < (\delta y)^{-2} \\ \ell^{(\text{Relative})} & \text{otherwise} \end{cases} \quad (14)$$

Thus, the final function to be minimized is:

$$\ell(\vec{\lambda}) = \sum_i \ell_i(\lambda_i). \quad (15)$$

This equation assumes that the error processes in the  $N_y$  channels are independent; i.e., the deviations  $y_i - \lambda_i$  are not correlated with each other.

The derivatives are given by:

$$\frac{\partial \ell}{\partial \lambda_i} = \frac{\partial \ell_i}{\partial \lambda_i}, \quad (16)$$

$$\frac{\partial^2 \ell}{\partial \lambda_i \partial \lambda_k} = \delta_{ik} \frac{\partial \ell_i}{\partial \lambda_i}. \quad (17)$$

### 3.1.3 Analytical Spectral Functions

Next, we must define the analytical (parametric) spectral functions that we wish to fit.

**3.1.3.1 Power Law Spectrum.** The power law (PL) spectrum is given by:

$$f^{(\text{PL})}(E) = \exp(q_1 - q_2 \ln E), \quad (18)$$

where we have chosen this parameterization so that there is no positivity constraint on  $q_1$  and to reduce numerical effects of large free parameters, e.g., if we'd defined  $q_1$  as a prefactor.

Its derivatives are:

$$\frac{\partial f^{(\text{PL})}}{\partial q_1} = f(E), \quad (19)$$

$$\frac{\partial f^{(\text{PL})}}{\partial q_2} = -\ln E f(E), \quad (20)$$

$$\frac{\partial^2 f^{(\text{PL})}}{\partial q_1^2} = f(E), \quad (21)$$

$$\frac{\partial^2 f^{(\text{PL})}}{\partial q_1 \partial q_2} = -\ln E f(E) = \frac{\partial^2 f^{(\text{PL})}}{\partial q_2 \partial q_1}, \quad (22)$$

$$\frac{\partial^2 f^{(\text{PL})}}{\partial q_2^2} = (\ln E)^2 f(E). \quad (23)$$

**3.1.3.2 Exponential Spectrum.** The exponential (EXP) spectrum is given by:

$$f^{(\text{EXP})}(E) = \exp(q_1 + q_2 E), \quad (24)$$

where, again, we have chosen this parameterization so that there is no positivity constraint on  $q_1$  and to reduce numerical effects of large free parameters. Its derivatives are:

$$\frac{\partial f^{(\text{EXP})}}{\partial q_1} = f(E), \quad (25)$$

$$\frac{\partial f^{(\text{EXP})}}{\partial q_2} = E f(E), \quad (26)$$

$$\frac{\partial^2 f^{(\text{EXP})}}{\partial q_1^2} = f(E), \quad (27)$$

$$\frac{\partial^2 f^{(\text{EXP})}}{\partial q_1 \partial q_2} = E f(E) = \frac{\partial^2 f^{(\text{EXP})}}{\partial q_2 \partial q_1}, \quad (28)$$

$$\frac{\partial^2 f^{(\text{EXP})}}{\partial q_2^2} = E^2 f(E). \quad (29)$$

**3.1.3.3 Relativistic Maxwellian Spectrum.** The relativistic Maxwellian (RM) spectrum is given by:

$$f^{(\text{RM})}(E) = E(1 + E/E_0/2) \exp(q_1 + q_2 E), \quad (30)$$

where, again, we have chosen this parameterization so that there is no positivity constraint on  $q_1$  and to reduce numerical effects of large free parameters. The constant  $E_0$  is the rest energy of the particle species (511 keV for electrons, 938 MeV for protons). Its derivatives are given by the same expressions as for the exponential spectrum (25)-(29).

**3.1.3.4 Double Relativistic Maxwellian Spectrum.** The double relativistic Maxwellian (RM2) spectrum is given by:

$$f^{(\text{RM2})}(E) = E(1 + E/E_0/2) [\exp(q_1 + q_2 E) + \exp(q_3 + q_4 E)], \quad (31)$$

where, again, we have chosen this parameterization so that there is no positivity constraint on  $q_1$  or  $q_3$  and to reduce numerical effects of large free parameters. The constant  $E_0$  is the rest energy of the particle species (511 keV for electrons, 938 MeV for protons). Its derivatives are given by the straightforward extension of (25)-(29).

**3.1.3.5 Power Law Spectrum with Exponential Tail.** The power law spectrum with exponential tail (PLE) is given by:

$$f^{(\text{PLE})}(E) = \begin{cases} \exp(q_1 - q_2 \ln E) & E \leq E_{\text{break}}, \\ \exp(q_1 - q_2 \ln E_{\text{break}} - (E - E_0)) & E > E_{\text{break}}. \end{cases} \quad (32)$$

where we have chosen this parameterization so that there is no positivity constraint on  $q_1$  and to reduce numerical effects of large free parameters, e.g., if we'd defined  $q_1$  as a prefactor. Nominal values for inner zone protons are  $E_0 = 345$  MeV,  $E_{\text{break}} = 100$  MeV.

Its derivatives are:

$$\theta = \begin{cases} -\ln E & E \leq E_{\text{break}}, \\ -\ln E_{\text{break}} & E > E_{\text{break}}. \end{cases} \quad (33)$$

$$\frac{\partial f^{(\text{PLE})}}{\partial q_1} = f(E), \quad (34)$$

$$\frac{\partial f^{(\text{PLE})}}{\partial q_2} = \theta f(E), \quad (35)$$

$$\frac{\partial^2 f^{(\text{PLE})}}{\partial q_1^2} = f(E), \quad (36)$$

$$\frac{\partial^2 f^{(\text{PLE})}}{\partial q_1 \partial q_2} = \theta f(E) = \frac{\partial^2 f^{(\text{PLE})}}{\partial q_2 \partial q_1}, \quad (37)$$

$$\frac{\partial^2 f^{(\text{PLE})}}{\partial q_2^2} = \theta^2 f(E). \quad (38)$$

### 3.1.4 Fit Errors and Combination of Multiple Spectral Fits

For a given spectral function  $f^{(k)}(E)$ , the fit is the minimization of  $\ell^{(k)}(\vec{\lambda})$  with respect to  $\vec{q}^{(k)}$ , which obtains the maximum likelihood estimate  $\hat{q}^{(k)}$ . Each fit is performed with multivariate minimization routines provided by the Gnu Scientific Library. Some of these minimizations require gradients of  $\ell^{(k)}$  with respect to  $\vec{q}^{(k)}$ , which are given by (without the  $(k)$  superscripts):

$$\begin{aligned}\frac{\partial \ell}{\partial q_m} &= \sum_i \frac{\partial \ell_i}{\partial \lambda_i} \sum_j \frac{\partial \lambda_i}{\partial f_j} \frac{\partial f_j}{\partial q_m}, \\ &= \sum_i \frac{\partial \ell_i}{\partial \lambda_i} \sum_j H_{ij} \frac{\partial f_j}{\partial q_m}.\end{aligned}\quad (39)$$

The Hessian, used for computing the error bar, is given by:

$$\frac{\partial^2 \ell}{\partial q_m \partial q_{m'}} = \sum_i \frac{\partial^2 \ell_i}{\partial \lambda_i^2} \sum_j H_{ij} \frac{\partial f_j}{\partial q_m} \sum_{j'} H_{ij'} \frac{\partial f_{j'}}{\partial q_{m'}} + \sum_i \frac{\partial \ell_i}{\partial \lambda_i} \sum_j H_{ij} \frac{\partial^2 f_j}{\partial q_m \partial q_{m'}}. \quad (40)$$

We treat the error on the resulting flux as having a log-normal distribution with standard deviation (i.e., standard error) given by the energy-dependent expression:

$$\begin{aligned}\sigma_{\ln f^{(k)}(E)} &= \sqrt{\sum_m \sum_{m'} \frac{\partial \ln f^{(k)}}{\partial q_m^{(k)}} \text{cov}(q_m^{(k)}, q_{m'}^{(k)}) \frac{\partial \ln f^{(k)}}{\partial q_{m'}^{(k)}}}, \\ &= \sqrt{\sum_m \sum_{m'} \frac{1}{f^{(k)}} \frac{\partial f^{(k)}}{\partial q_m^{(k)}} \text{cov}(q_m^{(k)}, q_{m'}^{(k)}) \frac{1}{f^{(k)}} \frac{\partial f^{(k)}}{\partial q_{m'}^{(k)}}},\end{aligned}\quad (41)$$

$$\text{cov}(q_m^{(k)}, q_{m'}^{(k)}) = \left( \begin{array}{ccc} \vdots & & \\ \cdots & \frac{\partial^2 \ell^{(k)}}{\partial q_m^{(k)} \partial q_{m'}^{(k)}} \Big|_{\hat{q}^{(k)}} & \cdots \\ \vdots & & \end{array} \right)^{-1}, \quad (42)$$

$$p(\ln f^{(k)}(E)) = \frac{1}{\sqrt{2\pi} \sigma_{\ln f^{(k)}(E)}} \exp \left[ -\frac{1}{2} \left( \frac{\ln f(E) - \ln f^{(k)}(E)}{\sigma_{\ln f^{(k)}(E)}} \right)^2 \right] \quad (43)$$

The approximation of an error covariance by the inverse of the Hessian of the penalty function is equivalent to approximating the curvature of the penalty function near the maximum likelihood value with the curvature of a Gaussian, i.e., a second-order Taylor series expansion of the penalty function.

We choose to combine the analytical fits according to  $\ell^{(k)}$ :

$$w_k = \frac{\exp(-\ell^{(k)} - N_q^{(k)})}{\sum_k \exp(-\ell^{(k)} - N_q^{(k)})}, \quad (44)$$

where  $N_q^{(k)}$  is the length of  $\vec{q}^{(k)}$  (always 2 in the spectral functions used so far). When evaluating the exp functions above, it is best first to subtract of the least  $\ell^{(k)}$  from all the  $\ell^{(k)}$  to avoid floating point overflow. This combination process is equivalent to assuming (reasonably) there are multiple possibilities for each  $f_j$ , with probability given by  $\exp(-\ell^{(k)})$ .

We then have a probability distribution that combines the log-normal probability distributions for the individual fits:

$$p(\ln f^{\text{combined}}(E)) = \sum_k w_k p(f^{(k)}(E)), \quad (45)$$

$$\ln \hat{f}(E) = \langle \ln f^{\text{combined}}(E) \rangle = \sum_k w_k \ln f^{(k)}(E), \quad (46)$$

$$\begin{aligned}\delta \ln \hat{f}(E) &= \sqrt{\text{var} \ln f^{\text{combined}}(E)} \\ &= \sqrt{\sum_k w_k \left( \sigma_{\ln f^{(k)}(E)}^2 + \ln^2 f^{(k)}(E) \right) - \langle \ln f^{\text{combined}}(E) \rangle^2}.\end{aligned}\quad (47)$$

When the function returns,  $\text{flux}[j] = \hat{f}(E_j)$ , and  $\text{dlogflux}[j] = \delta \ln \hat{f}(E_j)$ .

### 3.2 Function Arguments

- `y` (input) `y[NY*t+i]` = counts in channel `i` in case `t`.
- `dy` (input) `dy[i]` = relative error for channel `i` (dimensionless).
- `Egrid` (input) `Egrid[j]` = nominal energy of  $j^{\text{th}}$  grid point (e.g., in keV).
- `H` (input) `H[NY*j+i]` = response of channel `i` to flux at energy `j` (e.g., in keV cm<sup>2</sup> sr s).
- `dt` (input) `dt[t]` = integration time for case `t`.
- `b` (input) `b[NY*t+i]` = expected background counts in channel `i` for case `t`. Best fit when  $y[i] \approx b[NY*t+i] + \sum_j H[NY*j+i] \text{flux}[j]$ .
- `int_params` (input) integer parameters, length = 7 (TBR).
  - `int_params[0]` Number of energy channels, `NY`.
  - `int_params[1]` Number of energy grid points, `NE`.
  - `int_params[2]` Number of energy output points `NEout`.
  - `int_params[3]` Spectral Function Bit Mask (combined via bitwise OR):
    - [1] Power law (PL),
    - [2] Exponential (EXP),
    - [4] Relativistic Maxwellian (RM),
    - [8] Power law with exponential tail (PLE).
    - [16] Double Relativistic Maxwellian (RM2),
  - `int_params[4]` Choice of minimizer (choose one):
    - [0] Broyden-Fletcher-Goldfarb-Shanno, BFGS (recommended),
    - [1] Conjugate Fletcher-Reeves, Conjugate FR,
    - [2] Conjugate Polak-Ribiere, Conjugate PR,
    - [3] Nelder-Mead Simplex.
  - `int_params[5]` Maximum number of iterations by minimizer (recommend 10,000).
  - `int_params[6]` Verbose setting (choose one):
    - [0] no text output,
    - [1] text output to standard output stream,
    - [2] text output to standard error stream,
    - [3] text output to outFile (assumes outFile is actually a FILE \*).
    - [4] text output to outFile, overwrite existing file
    - [5] text output to outFile, append to existing file
  - `int_params[7]` `NT` Number of cases to invert.
- `real_params` (input) real parameters, length 3 (TBR).
  - `real_params[0]` = rest energy of particle species.
  - `real_params[1]` =  $E_{\text{break}}$  used by PLE.
  - `real_params[2]` =  $E_0$  used by PLE.
- `outFile` (input) provides filename or FILE \* for verbose output (see verbose setting above). Null terminated string.
- `flux` (output) `flux[NE*t+j]` = inverted flux at  $j^{\text{th}}$  energy grid point (e.g., in #/keV/cm<sup>2</sup>/sr/s), case `t`.
- `dlogflux` (output) `dlogflux[NE*t+j]` = standard error of natural log of `flux[NE*t+j]` (dimensionless).

Notes:

1. Array and matrix indices above are zero-based, following C convention rather than one-based linear algebra convention.
2. `ana_spec_inv` operates on *counts* not *count rate*.
3. Missing counts can be replaced with NaN, and the channel will be ignored.
4. If `dy` is unknown, use  $\ln(2)/2$  for factor of 2 95% confidence bounds, or 0 for Poisson error only.
5. Energy units must be consistent throughout: e.g., fluxes in `/keV`, energy grid points in `keV`, and rest energy in `keV`.
6. `H` is stored in row-major format, i.e., a sequence of contiguous rows of `H`.
7.  $H_{ij} \approx TG_i(E_j)\Delta E_j$ , where  $T$  is the integration time,  $G_i(E_j)$  is the geometric factor for channel  $i$  at energy  $j$ , and  $\Delta E_j$  is the energy bandwidth of the  $j^{\text{th}}$  grid point. Alternatives are possible: for example, the first and last columns of `H` can be divided by 2 to approximate trapezoidal integration if the energy grid is uniform.
8. It is best not to fit both an exponential and a relativistic Maxwellian in combination with a power law (or any other spectral function we add later) spectrum. Because exponential and relativistic Maxwellians are so similar, they'll combine to overrule the power law, inadvertently giving two votes to roughly the same spectrum.
9. `outFile` is ignored for verbose settings 0, 1, and 2.
10. The 95% confidence interval on `flux[j]` is given by `flux[j] × exp(±1.9600 × dlogflux[j])`.

### 3.3 Return Codes

- 1 Success! No Error.
- 0 Unknown error (this only happens if there's a bug).
- 101 NULL passed where pointer expected.
- 102 One or fewer valid data points; e.g., `(NY ≤ 1)`, or `NE ≤ 1` or `NEout < 1`.
- 103 one or more invalid (NaN or infinite) in input arrays, or some `y[i] < 0`, `dy[i] < 0`, `Egrid[j] ≤ 0`, or `Egrid[j] ≤ 0`.
- 104 no counts in any channel.
- 201 No functions selected in function bit map.
- 202 Invalid function selected in function bit map.
- 301 Relativistic Maxwellian (RM) requested with NULL `real_params` or `real_params[0] ≤ 0`.
- 302 Power law with exponential tail (PLE) requested with NULL or negative `real_params`.
- 401 Invalid minimizer selected.
- 402 Invalid iterations requested (zero or negative).
- 501 Invalid value of verbose provided or NULL value provided for user-requested stream (`outFile`).
- 502 User-requested verbose output file (`outFile`) could not be opened.

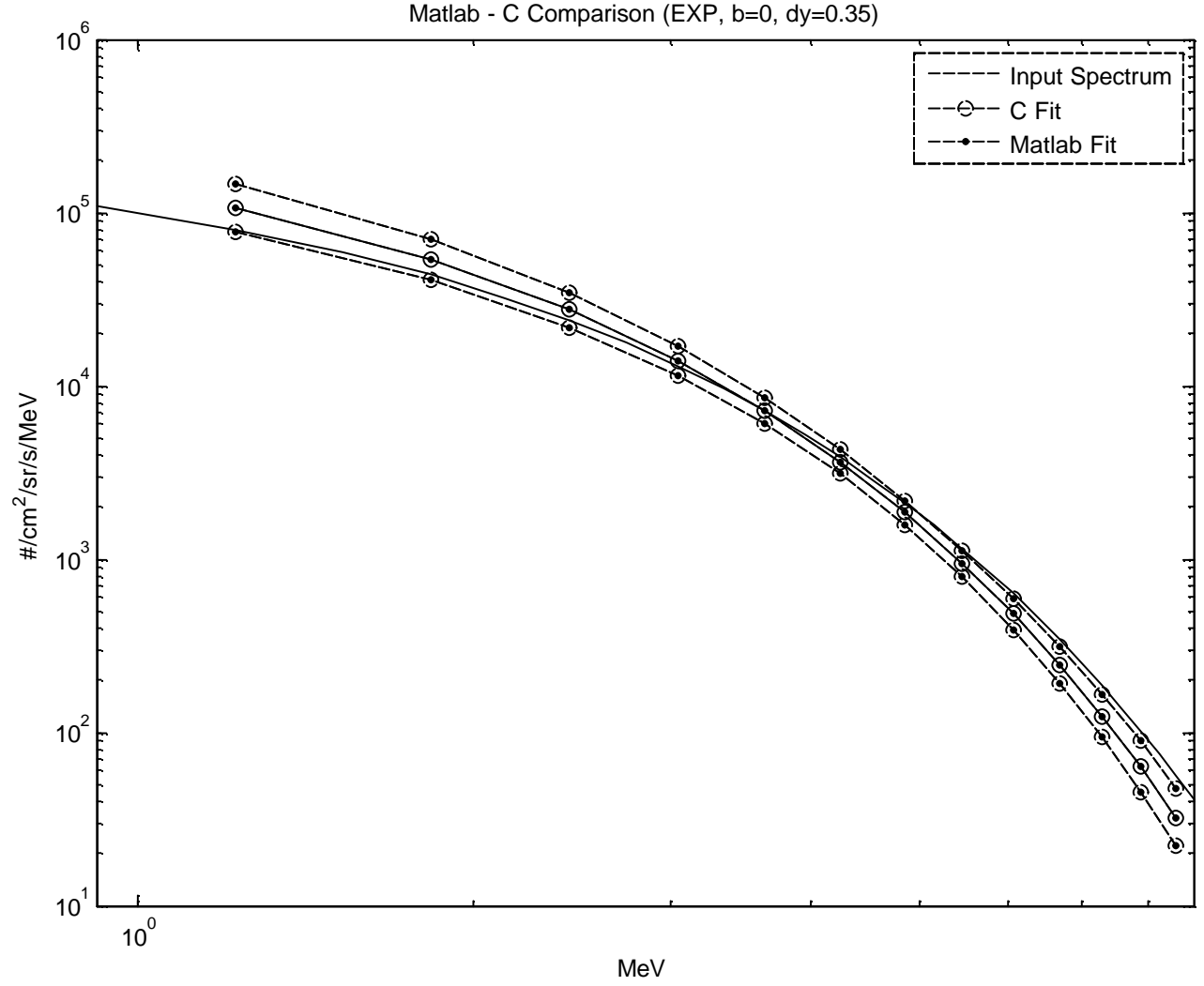


Figure 1: An example of fitting power law (PL) and exponential (EXP) spectra to the counts in 5 ICO electron channels to a simulated exponential with Poisson noise. Dotted lines indicate 1 standard error above and below the fits.



### 3.4 Validation

The routine `ana_spec_inv` and many of its dependencies have been validated against a Matlab counterpart “EnergySpectralInversionAnalytical.m”. The C test code for `ana_spec_inv` is “specinv\_test.c” and the Matlab test code is “specinv\_test.m”. The program “specinv\_test.exe” reads the text file “specinv\_test.in1” and produces the text file “specinv\_test.out1”, which is then compared to the same spectral inversion solved in Matlab. The results are shown in Figure ??.

Validation has consisted of validation of individual “private” routines found in the various ‘.c’ files but not shared via “invlib.h”, as well as the shared `ana_spec_inv` routine. The C and Matlab codes agree to within 5 or 6 decimal places for inputs of noisy power-law and exponential distributions fit to combinations of power-law, exponential and relativistic Maxwellian spectra.

### 3.5 Library Function `ana_spec_inv_multi`

The library also provides a mult-case function, `ana_spec_inv_multi`, with the following C prototype (found in “invlib.h”):

```
int ana_spec_inv_multi(const long int Ntimes,
    const double *y, const double *dy,
    const double *Egrid, const double *H0,
    const double *dt, const double *b,
    const long int *int_params, const double *real_params,
    char *outFile, double *Eout,
    double *flux, double *dlogflux, int *result_codes);
```

This function calls `ana_spec_inv` `Ntimes`, in each case advancing the `y` and `b` pointers by `NY` (i.e., moving one row forward, if `y` and `b` are interpreted as `Ntimes×NY` row-major matrices), and, similarly, advancing `flux` and `dlogflux` by `NE`. `result_codes` provides the functional return values. The primary difference between `ana_spec_inv` and `ana_spec_inv_multi` is that the latter requires the aggregation time (`dt`) be passed in separately from the instrument response function. For each call to `ana_spec_inv`, a temporary `H` will be set to `H0×dt[t]`, to reflect a possibly changing integration time. The other inputs are passed directly to `ana_spec_inv`, without alteration.

The return code of `ana_spec_inv_multi` is typically the first non-success result code returned by the multiple calls `ana_spec_inv`. Of course, if all calls to `ana_spec_inv` succeed, the return code from `ana_spec_inv_multi` will indicate success. The only additional exception is if `Ntimes < 1`, in which case the empty input array error code will be returned (-102).

## 4 Library Functions `omni2uni` and `wide2uni`

This section describes planned angular inversion functions for single-channel omnidirectional and wide-angle measurements.

The following C prototypes will be added to “invlib.h”:

```
int omni2uni(const double *omniflux, const double *dlogomniflux,
    const long int *int_params,
    const double *real_params,
    char *outFile,
    double *uniflux, double *dloguniflux);

int wide2uni(const double *wideflux, const double *dlogwideflux,
    const double *PAGrid, const double *H,
    const long int *int_params,
    const double *real_params,
    char *outFile,
    const long int *ialpha0,
    double *uniflux, double *dloguniflux);
```

Returns estimated flux at `PAGrid[ialpha0]`, where `H` describes angular dependence on grid `PAGrid`.

## 4.1 Calculations Performed - TEM-1 Method

The omnidirectional or wide-angle flux is given by:

$$f_{\text{iso}} = \int_0^{2\pi} \int_0^\pi h(\alpha, \phi) f_{\text{uni}}(\alpha) \sin \alpha d\alpha d\phi. \quad (48)$$

We denote the wide-angle or omnidirectional flux  $f_{\text{iso}}$  under the assumption that **wideflux** or **omniflux** is given in the form of a isotropic unidirectional flux (e.g., #/cm<sup>2</sup>/s/sr/keV). For the wide-angle approximation, one has a bit more flexibility: it is only necessary to provide **H** and **wideflux** in appropriate units so that the units of **uniflux** will be the units of **wideflux** divided by the units of **H**.

The omnidirectional flux is a special case of the wide-angle flux problem:

$$h(\alpha, \phi) = \frac{1}{4\pi}. \quad (49)$$

First, we discretize this equation:

$$f_{\text{iso}} \approx \sum_i H_i f_i, \quad (50)$$

$$H_i \approx \int_0^{2\pi} h(\alpha_i, \phi) d\phi \sin \alpha_i \Delta \alpha_i, \quad (51)$$

$$f_i = f_{\text{uni}}(\alpha_i). \quad (52)$$

TEM-1 follows Vette's AE-8 atmospheric cutoff:

$$\frac{B_m}{B_0} = \begin{cases} 0.6572 L_m^{3.452} & L_m < 2.4 \\ 0.196 L_m^{4.878} & 2.4 \leq L_m < 3.0 \\ 1.4567 L_m^{3.050} & L_m > 3.0 \end{cases} \quad (53)$$

Any flux outside (i.e., with  $B_m/B_0$  beyond the limit) is treated as implicitly zero by removing it from the  $\vec{f}$ . By implication, there is no electron flux at any pitch angle for  $L_m < 1.1293$ . Requesting a unidirectional flux at a local pitch angle that mirrors beyond the cutoff will result in an error code (see Return Codes, section ??).

Next, we set the problem up as a constrained maximum likelihood equation in the log fluxes:

$$x_i = \ln f_i, \quad (54)$$

$$y = \text{wideflux}, \quad (55)$$

$$\delta y = \text{dlogwideflux}, \quad (56)$$

$$\lambda = \sum_i H_i f_i, \quad (57)$$

$$\ell = \Lambda(\ln \lambda - \ln y) + \frac{1}{2}(\vec{x} - \vec{\mu})^T \underline{\Sigma}^{-1}(\vec{x} - \vec{\mu}). \quad (58)$$

The LaGrange multiplier  $\Lambda$  ensures that the local minimum in  $\ell$  satisfies  $y = \lambda$ , i.e., the inverted angular distribution exactly reconstructs the observed wide-angle flux.

Note: I also tried using a typical log-normal penalty function on  $y$  based on  $\delta y$ , but that approach causes the solution  $\vec{x}$  to be systematically biased toward  $\vec{\mu}$ . This is a typical problem with maximum likelihood. I believe we ought to treat separately the uncertainty in the wide-angle measurement and the uncertainty in the conversion to unidirectional flux. This LaGrange multiplier approach does just that.

The second term in (??) is the likelihood of a particular local pitch angle distribution at the specified energy and location according to TEM1. TEM1 provides the median  $m_{50}(E, \alpha_{eq}, L_m)$ , 95<sup>th</sup> percentile  $m_{95}(E, \alpha_{eq}, L_m)$ , and spatial correlation coefficient  $\rho(\alpha_{eq,i}, \alpha_{eq,i'} | E, L_m)$ , and we can convert between  $\alpha$  and  $\alpha_{eq}$  using the  $B/B_0$  parameter.

$$\sin \alpha_{eq,i} = \frac{\sin \alpha_i}{\sqrt{B/B_0}}, \quad (59)$$

$$\mu_i = \ln m_{50}(E, \alpha_{eq,i}, L_m), \quad (60)$$

$$\sigma_i = (\ln m_{95}(E, \alpha_{eq,i}, L_m) - \mu) / \Phi^{-1}(0.95), \quad (61)$$

$$\Phi^{-1}(0.95) = 1.6448536270, \quad (62)$$

$$\Sigma_{ii'} = \sigma_i \sigma_{i'} \rho(\alpha_{eq,i}, \alpha_{eq,i'} | E, L_m). \quad (63)$$

To solve for  $\ell$  will require the usual gradients and derivatives:

$$\frac{\partial \lambda}{\partial x_i} = H_i f_i \quad (64)$$

$$\frac{\partial \ell}{\partial x_i} = (\Lambda/\lambda) H_i f_i + \sum_{i'} (\underline{\Sigma}^{-1})_{ii'} (x_{i'} - \mu_{i'}), \quad (65)$$

$$\frac{\partial \ell}{\partial \Lambda} = \ln \lambda - \ln y, \quad (66)$$

$$\frac{\partial^2 \ell}{\partial x_i \partial x_{i'}} = (\Lambda/\lambda) H_i f_i \delta_{ii'} - (\Lambda/\lambda^2) H_i f_i H_{i'} f_{i'} + (\underline{\Sigma}^{-1})_{ii'}, \quad (67)$$

$$\frac{\partial^2 \ell}{\partial x_i \partial \Lambda} = H_i f_i / \lambda, \quad (68)$$

$$\frac{\partial^2 \ell}{\partial \Lambda^2} = 0. \quad (69)$$

The initial guess is  $y/(\vec{H}^T \exp(\vec{\mu}))$ , i.e., the median fluxes are rescaled to reconstruct  $y$ . The fit for  $\vec{x}$  is performed with multivariate root finding routines provided by the Gnu Scientific Library, producing the maximum likelihood value  $(\hat{\vec{x}}^T, \hat{\Lambda})$ , which finds simultaneous zeros of (??) and (??).

Thus the estimated flux is:

$$\mathbf{uniflux} = \exp(\hat{x}_{i_{\alpha_0}}). \quad (70)$$

We discard all the other components of  $\hat{\vec{x}}$  because we want to conserve the number of observations (one cannot get something for nothing) for subsequent analysis. Doing otherwise would risk over-weighting the “typical” pitch angle distribution shape in subsequent analysis of the inverted flux.

The uncertainty on the log flux is given by:

$$\mathbf{dloguniflux} = \sqrt{(\delta y)^2 + (\delta \ln f_{i_{\alpha_0}})^2}, \quad (71)$$

$$\delta \ln f_{i_{\alpha_0}} = \delta x_{i_{\alpha_0}} = \sqrt{Q_{i_{\alpha_0}, i_{\alpha_0}}}, \quad (72)$$

$$\underline{\underline{Q}} = \begin{pmatrix} \cdots & \frac{\partial^2 \ell}{\partial x_i \partial x_{i'}} \Big|_{(\hat{\vec{x}}^T, \hat{\Lambda})} & \cdots & \frac{\partial \ell}{\partial \Lambda} \Big|_{(\hat{\vec{x}}^T, \hat{\Lambda})} \\ \vdots & & & \\ \frac{\partial \ell}{\partial \Lambda} \Big|_{(\hat{\vec{x}}^T, \hat{\Lambda})} & & & 0 \end{pmatrix}^{-1}. \quad (73)$$

Note that  $\underline{\underline{Q}}$  is the inverse of the Hessian of  $\ell$  in the  $(\vec{x}^T, \Lambda)$  super-space, and  $\mathbf{dloguniflux}$  is a combination of the angular inversion error and the measurement error  $\mathbf{dlogwideflux}$ , as if they were independent (not a bad assumption).

## 4.2 Function Arguments

### 4.2.1 `omni2uni`

- `omniflux` (input) Estimated isotropic unidirectional flux (for TEM-1 method, use  $\#/\text{cm}^2/\text{sr}/\text{s}/\text{keV}$ ).
- `dlogomniflux` (input) relative error for `omniflux` (dimensionless).
- `int_params` (input) integer parameters, length = 5 (TBR).
  - `int_params[0]` NA = number of grid points for angular integral.
  - `int_params[1]` angular inversion method:
    - `[-1]` - TEM-1.
  - `int_params[2]` Verbose setting (choose one):
  - `int_params[3]` Choice of root finder (choose one):
    - `[0]` Powell’s Hybrid method, scaled (recommended),

- [1] Powell's Hybrid method
  - [2] Newton's method
  - [3] Pseudo-global Newton's method
  - [4] Powell's Hybrid method w/ numerical Hessian, scaled
  - [5] Powell's Hybrid method w/ numerical Hessian
  - [6] Discrete Newton's method (numerical Hessian)
  - [7] Broyden's algorithm (numerical Hessian)
- int\_params**[4] Maximum number of iterations by minimizer (recommend 10,000).
- [0] no text output,
  - [1] text output to standard output stream,
  - [2] text output to standard error stream,
  - [3] text output to outFile (assumes outFile is actually a FILE \*).
  - [4] text output to outFile, overwrite existing file
  - [5] text output to outFile, append to existing file
- **real\_params** (input) real parameters, length 3 (TBR).
    - real\_params**[0] = Energy of electron flux energy, keV
    - real\_params**[1] =  $B/B_0$  of spacecraft location, dimensionless.
    - real\_params**[2] =  $L_m$ , McIlwain L for locally mirroring electron, in Olson-Pfitzer Quiet.
  - **outFile** (input) provides filename or FILE \* for verbose output (see verbose setting above). Null terminated string.
  - **uniflux** (output) Locally-mirroring unidirectional flux, (e.g., in  $\#/\text{keV}/\text{cm}^2/\text{sr}/\text{s}$ ).
  - **dloguniflux** (output) Standard error of natural log of **uniflux** (dimensionless).

#### 4.2.2 wide2uni

- **wideflux** (input) Estimated isotropic unidirectional flux (for TEM-1 method, use  $\#/\text{cm}^2/\text{sr}/\text{s}/\text{keV}$ ).
- **dlogwideflux** (input) relative error for **wideflux** (dimensionless).
- **PAGrid** (input) **PAGrid**[*i*] local pitch angle at grid point *i*, degrees, must be monotonically increasing grid.
- **H** (input) **H**[*i*] Angular weighting for unidirectional flux at grid point *i*, dimensionless.
- **int\_params** (input) integer parameters, length = 5 (TBR), same as **omni2uni** except:
  - [0] = NA number of pitch angles in grid.
- **real\_params** (input) same as **omni2uni**.
- **outFile** (input) same as **omni2uni**.
- **ialpha0** (input) zero-based index of pitch angle grid point for uniflux output (e.g., grid point nearest sensor bore sight).
- **uniflux** (output) same as **omni2uni**.
- **dloguniflux** (output) same as **omni2uni**.

Notes:

1. Pointers are used throughout for ease of access from FORTRAN, which can only pass arguments by reference (i.e., pointers).
2. Array and matrix indices above are zero-based, following C convention rather than one-based linear algebra convention.

3. `omni2uni` and `wide2uni` operate on estimated isotropic unidirectional flux.
4. Alternatively, `H` and `wideflux` can be defined in terms of omnidirectional flux, so long as the units of `wideflux/H` give units of `uniflux`.
5. If `dlogomniflux` or `dlogwideflux` is unknown, use  $\ln(2)/2$  for factor of 2 95% confidence bounds.
6. `omniflux`, `wideflux`, `dlogomniflux`, or `dlogwideflux` negative or zero will result in an error (see Return Codes, section ??).
7. `outFile` is ignored for verbose settings 0, 1, and 2.
8. The 95% confidence interval on `uniflux` is given by  $\text{uniflux} \times \exp(\pm 1.9600 \times \text{dloguniflux})$ .

### 4.3 Return Codes

- 1 Success! No Error.
- 0 Unknown error (this only happens if there's a bug).
- 101 NULL passed where pointer expected.
- 102 One or fewer valid data points ( $NY \leq 1$ ), or  $NE \leq 1$  or  $NE_{out} < 1$ .
- 103 One or more invalid (NaN or infinite) in input arrays, or some  $y[i] < 0$ ,  $dy[i] < 0$ ,  $E_{grid}[j] \leq 0$ , or  $E_{grid}[j] \leq 0$ .
- 104 Negative or zero input `omniflux`, `wideflux`, `dlogomniflux`, `dlogwideflux`.
- 401 Invalid minimizer selected.
- 402 Invalid iterations requested (zero or negative).
- 501 Invalid value of verbose provided or user-requested file stream NULL (`outFile`).
- 502 User-requested verbose output file (`outFile`) could not be opened.
- 601 Output pitch angle requested out of range (index out of range or  $B_m$  beyond cutoff).
- 602 Invalid angular inversion method requested.

### 4.4 Validation

The routine `omni2uni` has been validated for the TEM-1 method by running it for various cases and examining the output for sanity (not very sophisticated validation). The C test code is `“omni2uni_test.c”`, which passes a single test case to `omni2uni` and prints its output.

Here is a transcript of `omni2uni_test.exe`:

```
[PROMPT]> ./omni2uni_test.exe
omni2uni_test:inputs = omni=1.00000e+04 (dlog=3.46574e-01) @ [300.0 keV, B/B0=400.00, Lm=6.60]
fzero Invoked with solver hybridsj, MaxIter=1000
fzero: 1/1000: 0.313671:
fzero: 2/1000: 0.0140756:
fzero: 3/1000: 1.27369e-06:
fzero: 4/1000: 3.37084e-09:
fzero: completed after 4/1000: success
omni2uni_test:inputs = omni=1.00000e+04 (dlog=3.46574e-01) @ [300.0 keV, B/B0=400.00, Lm=6.60]
omni2uni_test:result = 1, uni=2.73852e+04 (dlog=3.46656e-01)
[PROMPT]>
```

In this case, because  $B/B_0 \gg 1$ , there is a substantial loss cone, thus creating a significant pitch-angle anisotropy, which results in a much larger locally-mirroring flux than would be assumed from isotropy alone (by a factor of 2.7...).

## 5 Dependencies and Compiling

The inversion library routines rely on the GNU Scientific Library, available for free download at <http://www.gnu.org/software/gsl>. The GSL base library and the GSL CBLAS library are required (both come with a typical GSL install).

The library is built for the gcc compiler suite. A makefile (“Makefile”) is provided. The `make` command alone will display the helps required to build the `invlib` shared object (DLL).

The compilation may generate some `warning` and `Info` messages:

```
warning: comparison between signed and unsigned
warning: assignment discards qualifiers from pointer target type
Info: resolving _gsl_multimin_fdfminimizer_conjugate_fr by linking to
__imp__gsl_multimin_fdfminimizer_conjugate_fr (auto-import)
Info: resolving _gsl_multimin_fdfminimizer_conjugate_pr by linking to
__imp__gsl_multimin_fdfminimizer_conjugate_pr (auto-import)
Info: resolving _gsl_multimin_fminimizer_nmsimplex by linking to
__imp__gsl_multimin_fminimizer_nmsimplex (auto-import)
Info: resolving _gsl_multimin_fdfminimizer_vector_bfgs2 by linking to
__imp__gsl_multimin_fdfminimizer_vector_bfgs2 (auto-import)
```

I have not figured out how to make a stand-alone DLL in cygwin, I think because GSL libraries depend on “cygwin1.dll”. However, I have created a DLL with MSYS/MINGW, and the makefile has commands to do this.

## A Adding a new spectral form

The following modifications must be made to add a new spectral form:

In `specinv.h`,

- add a new `ASI_FXN_<NEW>` constant, the next power of two.
- increment `ASI_FXN_POW2` constant,
- insert `ASI_FXN_<NEW>` in bitwise or of all allowed functions (`ASI_FXN_ALL`)

In `invlib_const.h`, add any new error codes associated with improper options/settings for new analytical spectrum.

In `specinv.c`,

- create new `flux_<new>` function, following template from existing analytical flux functions.
- add input checks to `ana_spec_inv`, e.g., checking for existence and positivity of rest mass parameter in `params`.
- add case for `ASI_FXN_<NEW>` to `switch(fxn_bit)` in `ana_spec_inv`.

In `invlib.tex` (this file), add appropriate documentation:

- Note changes in section ??, “Changes”.
- Add definition and derivatives of new spectral function to section ??.
- Update argument list for `ana_spec_inv`, section ??.

Test compilation with `make` commands. Recompile this file with `pdflatex`.