



Statistical Properties of Maximum Likelihood Estimators of Power Law Spectra Information

L.W. Howell

Marshall Space Flight Center, Marshall Space Flight Center, Alabama

The NASA STI Program Office...in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the lead center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

- **TECHNICAL PUBLICATION.** Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA's counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.
- **TECHNICAL MEMORANDUM.** Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.
- **CONTRACTOR REPORT.** Scientific and technical findings by NASA-sponsored contractors and grantees.

- **CONFERENCE PUBLICATION.** Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or cosponsored by NASA.
- **SPECIAL PUBLICATION.** Scientific, technical, or historical information from NASA programs, projects, and mission, often concerned with subjects having substantial public interest.
- **TECHNICAL TRANSLATION.** English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized databases, organizing and publishing research results...even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI Program Home Page at <http://www.sti.nasa.gov>
- E-mail your question via the Internet to help@sti.nasa.gov
- Fax your question to the NASA Access Help Desk at (301) 621-0134
- Telephone the NASA Access Help Desk at (301) 621-0390
- Write to:
NASA Access Help Desk
NASA Center for AeroSpace Information
7121 Standard Drive
Hanover, MD 21076-1320



Statistical Properties of Maximum Likelihood Estimators of Power Law Spectra Information

L.W. Howell

Marshall Space Flight Center, Marshall Space Flight Center, Alabama

National Aeronautics and
Space Administration

Marshall Space Flight Center • MSFC, Alabama 35812

Acknowledgments

The author gives special appreciation to John Watts of the Cosmic-ray Group, Space Science Department, Marshall Space Flight Center, Huntsville, AL, for his many technical discussions and editorial remarks pertaining to this Technical Publication.

The author is grateful to Professor Herman J. Bierens, Professor of Economics, Pennsylvania State University, and part-time Professor of Econometrics, Tilburg University, the Netherlands, for his assistance in checking equation (39) in Appendix A—Closed-form Expression for the Cramer-Rao Bound of the Broken Power Law.

TRADEMARKS

Trade names and trademarks are used in this report for identification only. This usage does not constitute an official endorsement, either expressed or implied, by the National Aeronautics and Space Administration.

Available from:

NASA Center for AeroSpace Information
7121 Standard Drive
Hanover, MD 21076-1320
(301) 621-0390

National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
(703) 487-4650

TABLE OF CONTENTS

1. INTRODUCTION	1
1.1 Simple Power Law Energy Spectrum	2
1.2 Broken Power Law Energy Spectrum	10
1.3 Break-Size 0.3 Study	18
1.4 Analysis of Multiple Independent Data Sets	27
1.5 Example of Estimating α_1 of a Simple Power Law Using Two Data Sets	30
1.6 Using the CRB to Explore Instrument Design Parameters When Estimating $\theta=(\alpha_1, \alpha_2, E_k)$ of a Broken Power Law Using Three Data Sets	35
2. CONCLUSIONS	38
APPENDIX A—CLOSED-FORM EXPRESSION FOR THE CRAMER-RAO BOUND OF THE BROKEN POWER LAW	39
A.1 Derivation	39
A.2 Check on Equation (39) by Professor Bierens	43
APPENDIX B—CRAMER-RAO BOUND FOR MULTIPLE INDEPENDENT DATA SETS	44
REFERENCES	48

LIST OF FIGURES

1.	Mean estimates of α_{ML} as a function of the number of events, N , showing a bias for $N < 1,000$. Triangle marker at $N = 200$ and $52,000$ is for method of moments	5
2.	Standard deviation of α_{ML} and the CRB as a function of N	6
3.	MLE and method of moments as a function of instrument energy resolution for the proposed TSC with Gaussian response function	7
4.	Frequency histograms of α_{ML} for 10,000 simulated missions with (a) $N = 50$ and (b) $N = 52,000$ events, using the TSC with its 40-percent resolution Gaussian response function	7
5.	Frequency histogram of α_{ML} based on 10,000 simulated missions with $N = 50$ using an ideal detector. The smooth curve is the theoretical distribution of α_{ML} obtained from equation (13)	8
6.	Standard deviation of the ML estimator α_{ML} versus N for a 40-percent resolution Gaussian detector based on 10,000 simulated missions at each value of N and the CRB for the 40-percent Gaussian detector computed using equation (14) (—). The CRB for the ideal detector is included as a reference curve (— —)	9
7.	CRB as a function of N for detector energy resolutions in the range $0 \leq \rho \leq 0.40$. The proposed TSC instrument with its expected 52,000 events is indicated by the square, and the observing energy range was set to 20–5,500 TeV	10
8.	ML estimate of α_1 and α_2 as a function of detector collecting power	12
9.	ML estimate of E_k as a function of detector collecting power using a 40-percent resolution Gaussian response function (the TSC) and the ideal detector. The knee location E_k was set to 100 TeV in these simulations	13
10.	CRB of α_1 using TSC (— —) and ideal detector (—) obtained from equation (20) versus collecting power. Standard deviations of ML estimates from simulations for values of N in table 1 indicated by markers	14
11.	CRB of α_2 using TSC (— —) and ideal detector (—) obtained from equation (20) versus collecting power. Standard deviations of ML estimates from simulations for values of N in table 1 indicated by markers	15

LIST OF FIGURES (Continued)

12.	CRB of E_k using TSC (— —) and ideal detector (—) obtained from equation (20) versus collecting power. Standard deviations of ML estimates from simulations for values of N in table 1 indicated by markers	15
13.	Relative frequency histograms of the ML estimate of α_1 (leftmost two histograms) for $N = 11,439$ (broadest of the two) and $N = 114,390$ (narrow histogram). Rightmost two histograms similarly defined for α_2	16
14.	Relative frequency histograms of the ML estimates of E_k for $N = 11,439$ (broadest of the two and with bump in right-hand tail) and $N = 114,390$	17
15.	Relative frequency histograms of ML estimates of α_1 and α_2 for $N = 22,877$ and $N_2 = 1,000$	17
16.	ML estimate of α_1 and α_2 as a function of detector collecting power when the spectral break-size is 0.3 for the TSC and ideal detector	18
17.	ML estimate of E_k as a function of detector collecting power using the TSC and ideal detector when the break-size is 0.3. The actual concept TSC with its expected $N = 51,790$ events is indicated by the diamond and is based on 25,000 simulated missions (others are 5,000 missions each) and suggests the marker to its immediate right is probably a little on the high side	19
18.	CRB of α_1 using TSC (— —) and ideal detector (—) versus collecting power. Standard deviations of ML estimates from simulations for values of N in table 2 indicated by markers	20
19.	CRB of α_2 using TSC (— —) and ideal detector (—) versus collecting power. Standard deviations of ML estimates from simulations for values of N in table 2 indicated by markers	20
20.	CRB of E_k using TSC (— —) and ideal detector (—) versus collecting power. Standard deviations of ML estimates from simulations for values of N in table 2 indicated by markers	21
21.	Relative frequency histograms of ML estimates of α_1 and α_2 for the proposed TSC	21
22.	Relative frequency histograms of the ML estimates of E_k for the proposed TSC	22
23.	Objective function given by equation (18) for the ideal detector in the neighborhood of θ_{ML} for a simulated mission, keeping α_1 fixed at its ML estimate. There were 114,385 (4,819) events in this mission. The vertical axis has been scaled to 1	23

LIST OF FIGURES (Continued)

24.	Contour plot of the objective function for ideal detector in the neighborhood of θ_{ML} for a simulated mission, keeping α_1 fixed at its ML estimate. There were 114,385 (4,819) events in this mission	24
25.	Objective function for ideal detector in the neighborhood of θ_{ML} for a simulated mission, keeping α_1 fixed at its ML estimate. This mission consisted of 4,575 (207) events. The vertical axis has been scaled to 1	25
26.	CRB used to estimate where the distribution of α_2 begins to cross that of α_1 when (a) the spectral break-size is 0.5 and (b) the spectral break-size is 0.3	26
27a.	ML estimates of α_1 for 25 missions, with $E_1 = 5$ TeV (5,275 events for TRD and 593 for calorimeter)	31
27b.	ML estimates of α_1 for 25 missions, with $E_1 = 3$ TeV (11,660 events for TRD and 593 for calorimeter)	32
27c.	ML estimates of α_1 for 25 missions, with $E_1 = 1$ TeV (56,920 events for TRD and 593 for calorimeter). Note that α_{TRD} and α_{Both} are virtually indistinguishable	32
28.	Gamma, Gaussian, and broken Gaussian detector response functions to 40-TeV proton	36

LIST OF TABLES

1.	Number of events used in broken power law simulations for 0.5 break-size study	12
2.	Number of events used in broken power law simulations for 0.3 break-size study	18
3.	Summary statistics based on 25,000 simulated missions with $\theta = (2.8, 3.1, 100 \text{ TeV})$ and observing range 20–5,500 TeV, for the TSC having Gaussian response function	22
4.	Effect of lowering E_1 on the CRB for the TSC-sized detector with 0-, 20-, and 40-percent resolution Gaussian response function. The number of events N_2 above E_k is 2,255 for all values of E_1	27
5.	Summary statistics of α_{Cal} based on 1,000 simulated missions of the calorimeter	33
6.	Summary statistics of α_{TRD} based on 1,000 simulated missions for each value of E_1 of the TRD alone (all events between E_1 and 20 TeV)	33
7.	Summary statistics of α_{Both} based on 1,000 simulated missions for each value of E_1 of the TRD and calorimeter acting in combination	33
8.	Data sets and associated response functions, with CRB for all possible combinations of data sets	35
9.	Number of events from table 8 in data set C is increased by a factor of 10	36
10.	Number of events from table 8 in data set B is reduced by a factor of 2 and those in C increased by a factor of 10	37
11.	Number of events from table 8 in data set B is reduced by a factor of 2 and those in C increased by a factor of 10, resolution of A improved to 30 percent and C to 35 percent	37

LIST OF ACRONYMS

CRB	Cramer-Rao bound
EE	errant estimate
GCR	galactic cosmic ray
GEANT	GEometry ANd Tracking particle physics simulation program
ML	maximum likelihood
MLE	maximum likelihood estimation
rl	radiation length
TP	Technical Publication
TSC	thin sampling calorimeter

NOMENCLATURE

E	energy (units in TeV)
\mathbf{I}	information matrix
L	likelihood function
O	objective function
Z	standard normal random number

TECHNICAL PUBLICATION

STATISTICAL PROPERTIES OF MAXIMUM LIKELIHOOD ESTIMATORS OF POWER LAW SPECTRA INFORMATION

1. INTRODUCTION

A brief summary of the maximum likelihood estimation (MLE) procedure developed for estimating power law spectral parameters in earlier works^{1,2} begins with the probability density function of the astrophysics data set consisting of N detector responses y_i , e.g., energy deposit, as

$$g(y_i; \boldsymbol{\theta}) = \int_R g(y_i | E; \rho) \phi(E; \boldsymbol{\theta}) dE, \quad i = 1, \dots, N, \quad (1)$$

where $\boldsymbol{\theta}$ denotes the vector of spectral parameters of an assumed energy spectrum $\phi(E; \boldsymbol{\theta})$ to be estimated; N is the number of detected events from observing range, R , of the instrument having response function, g , and energy resolution, ρ . Then the corresponding likelihood function is

$$L(\boldsymbol{\theta}) = \prod_{i=1}^N \left[\int_R g(y_i | E; \rho) \phi(E; \boldsymbol{\theta}) dE \right] \quad (2)$$

and the ML estimate of $\boldsymbol{\theta}$, say $\boldsymbol{\theta}_{\text{ML}}$, is chosen so that for any admissible value of $\boldsymbol{\theta}$, $L(\boldsymbol{\theta}_{\text{ML}}) \geq L(\boldsymbol{\theta})$ or equivalently, $\log[L(\boldsymbol{\theta}_{\text{ML}})] \geq \log[L(\boldsymbol{\theta})]$.³ In practice, $\boldsymbol{\theta}_{\text{ML}}$ can be obtained from equation (3) using an optimization routine, such as the Nelder-Mead simplex search algorithm,⁴ to yield

$$\boldsymbol{\theta}_{\text{ML}} = \min_{\{\boldsymbol{\theta}\}} O(\boldsymbol{\theta}), \quad (3)$$

where the objective function, O , in equation (3) is defined as $O(\boldsymbol{\theta}) = -\log[L(\boldsymbol{\theta})]$ for this minimization algorithm so that minimizing $O(\boldsymbol{\theta})$ maximizes $\log[L(\boldsymbol{\theta})]$ as desired, and where the integral appearing in equation (2) and thus intrinsic to equation (3) can be very accurately evaluated by numerical methods such as the method of Gaussian quadratures.⁵ The Nelder-Mead algorithm does not require gradient information which is a vital consideration in selecting an optimization algorithm for this application because some energy spectra, such as the broken power law, are not differentiable everywhere.

The MLE theory generally leads to lower bounds on the statistical errors (standard deviations) of the spectra information and the existence of such a bound, called the Cramer-Rao bound (CRB), is the bound below which the variance of an unbiased estimator cannot fall. This implies that irrespective of the method used to quantify the parameters from the data, there is a lower bound on the precision that cannot be superseded.⁶ In the multiparameter case, if $\hat{\boldsymbol{\theta}}$ is any unbiased estimator of $\boldsymbol{\theta}$, then

$$\text{var}(\hat{\boldsymbol{\theta}}) \geq \left[N \left\langle \frac{\partial \log[g(y; \boldsymbol{\theta})]}{\partial \boldsymbol{\theta}} \times \frac{\partial \log[g(y; \boldsymbol{\theta})]}{\partial \boldsymbol{\theta}^T} \right\rangle \right]^{-1} \quad (4)$$

in the sense that the difference of these two matrices is positive semidefinite. The right-hand-side matrix in equation (4) is the CRB⁷ and notationally will be referred to as $\mathbf{I}^{-1}(\boldsymbol{\theta})$, where $\mathbf{I}(\boldsymbol{\theta})$ is frequently called the information matrix; the notation $\langle . \rangle$ denotes “expected value”; and the superscript T stands for vector transpose. Thus, the variance of one component of $\hat{\boldsymbol{\theta}}$, say $\hat{\theta}_i$, is bounded below as $\text{var}(\hat{\theta}_i) \geq \mathbf{I}_{ii}^{-1}(\boldsymbol{\theta})$, where $\mathbf{I}_{ii}^{-1}(\boldsymbol{\theta})$ is the i th diagonal element of $\mathbf{I}^{-1}(\boldsymbol{\theta})$. When $\boldsymbol{\theta}$ consists of a single spectral parameter, e.g., a simple power law energy spectrum is assumed, the CRB is the right-hand side of the inequality

$$\text{var}(\hat{\theta}) \geq \frac{1}{N \left\langle \left(\frac{\partial \log[g(y; \theta)]}{\partial \theta} \right)^2 \right\rangle} . \quad (5)$$

Additionally, ML estimators generally possess the favorable large sample properties of consistency (unbiased) (P1) and normality (P3).^{6,8,9} This Technical Publication (TP) investigates the conditions whereby these two properties, along with efficiency which is attainment of the CRB and is referred to as property P2 in this TP, are attained for an assumed simple power law energy distribution and a broken power law distribution, with emphasis on practical applications to instrument design and data analysis.

1.1 Simple Power Law Energy Spectrum

The simple power law suggests that the number of protons detected above an energy, E , is given by:

$$N_S(> E) = M_A \left(\frac{E}{E_A} \right)^{-\alpha_1 + 1} , \quad (6)$$

where E is in units TeV, α_1 is believed to be ≈ 2.8 , and M_A and E_A are numbers associated with the detector’s collecting power (combination of size and observing time). In statistical terms, N_S is assumed to represent an average number of events, while the actual number to be observed on any given mission would follow the Poisson probability distribution with mean number N_S . The probability density function for galactic cosmic ray (GCR) event energy, E , is then given by

$$\phi_S(E) = \frac{\alpha_1 - 1}{E_1^{1-\alpha_1} - E_2^{1-\alpha_1}} E^{-\alpha_1} \quad \text{for } E_1 \leq E \leq E_2 \quad (7)$$

over an energy range $[E_1, E_2]$ that does not depend on the parameter α_1 . Because the actual incident particle energies are never observed, but only a measure of their energy deposition from their passage through the detector, the random variable Y is introduced to represent the detector's response, e.g., energy deposition, of a GCR proton of incident energy, E , and its stochastic response function, g , with energy resolution, ρ , which may or may not be energy independent. For specificity, a response function, g , based on simulation studies of a thin sampling calorimeter (TSC) concept for the Advanced Cosmic-ray Composition Experiment for the Space Station (ACCESS) will be used. It has a planned 3-yr program life cycle and is composed of a carbon target and sampling calorimeter. The TSC area is 1 m² with a target ≈ 0.7 proton interaction lengths thick, sampled by X/Y pairs of square scintillating fibers. The fibers in the target are 2 mm thick and provide the approximate position of the interaction. The calorimeter consists of upper and lower parts totaling 25-radiation length- (rl-) thick lead and contains 28 X/Y pairs of 500- μ m square scintillating fibers. The upper 3-rl-thick calorimeter is sampled each 0.5 rl, and the lower part is sampled each 1.0 rl. The total weight of the target and calorimeter is $\approx 2,600$ kg, and the collecting power parameters, M_A and E_A , are estimated to be 160 and 500 TeV, respectively, implying that this TSC is expected to observe 160 proton events above 500 TeV over its expected 3-yr life cycle.

The TSC performance predictions are based on the geometry and tracking particle physics simulation program (GEANT) simulations of energy deposition for monoenergetic protons at specified energies at 0.1, 1, 10, 100, 1,000, and 5,000 TeV for this candidate detector. The Gaussian distribution was found to provide a reasonable description of the distribution of energy depositions at each of these incident energies.¹⁰ The mean detector response and the rms response were both found to be well approximated by a linear function of incident energy, E , in the range of interest for this study, which is typically in the range of 20 to 5,500 TeV. Thus, the mean energy deposition, Y , for a given incident energy, E , is defined to be $\mu_{Y|E} = (a + bE)$ and the rms response defined as $\sigma_{Y|E} = (c + dE)$, where the coefficients a , b , c , and d were estimated from the GEANT simulation results.

Before investigating the properties of MLE for the TSC instrument, it is instructive to consider the concept of a zero-resolution instrument or so-called ideal detector because it sets an upper bound on the expected performance of any real detector of equal collecting power. This measurement bound is determined by the CRB for the ideal detector which in turn establishes the limit in attainable precision with which unbiased spectra information can be obtained from a given science mission by any conceivable instrument with equal or less collecting power. Hence, it is useful in crafting realistic measurement goals for new science missions.

An ideal detector's energy resolution, ρ , is equal to zero, so the standard deviation $\sigma_{Y|E}$ is zero for all GCR event energies, E . Hence, the incident GCR energies are precisely known from the inverse mean response, so that for the TSC having linear mean response gives $E = (Y - a)/b$, and using equation (5) provides the CRB as the right-hand side of the inequality:

$$\text{var}(\hat{\alpha}) \geq \frac{1}{N \left(\frac{1}{(\alpha_1 - 1)^2} - \left[\frac{E_2^{1+\alpha_1} E_1^{1+\alpha_1} (\log[E_2] - \log[E_1])^2}{(E_2 E_1^{\alpha_1} - E_1 E_2^{\alpha_1})^2} \right] \right)}, \quad (8)$$

and is asymptotically attained by the ML estimator. A key question then arises, “For what values of N is this asymptotic property P2, as well as P1 and P3, achieved by MLE?” A battery of simulations was conducted to study this question consisting of 10,000 simulated missions for each of several values of N ranging from 50 to 52,000* events per mission and with GCR energies from the energy range of 20 to 5,500 TeV. The ML estimate α_{ML} was obtained for each mission by solving

$$\frac{\partial \log[L]}{\partial \alpha_1} = \frac{1}{\alpha_1 - 1} - \left[\frac{\log[E_2] E_2^{1-\alpha_1} - \log[E_1] E_1^{1-\alpha_1}}{E_1^{1-\alpha_1} - E_2^{1-\alpha_1}} \right] - \frac{1}{N} \sum_{i=1}^N \log[E_i] = 0 \quad (9)$$

in terms of α_1 for the ideal detector and then also for a 40-percent resolution Gaussian detector (the TSC) by application of equation (3). For comparison, the estimation technique referred to as the “method of moments” is included for the ideal detector ($\rho = 0$) for the case $N = 200$ events and $N = 52,000$ events, and consists of equating the sample mean \bar{E} to the population mean as

$$\bar{E} = \left(\frac{\hat{\alpha}_1 - 1}{\hat{\alpha}_1 - 2} \right) \frac{E_1^{2-\hat{\alpha}_1} - E_2^{2-\hat{\alpha}_1}}{E_1^{1-\hat{\alpha}_1} - E_2^{1-\hat{\alpha}_1}} \quad (10)$$

and then solving the nonlinear equation (10) in terms of $\hat{\alpha}_1$.² Figure 1 shows that MLE provides an unbiased estimate of α_1 when $N \geq 1,000$, but with an ever-increasing bias as the number of events diminishes. Note that even though α_{ML} is biased when $N = 200$ for the 40-percent resolution Gaussian TSC, its bias is significantly less than the bias of the method of moments estimator $\hat{\alpha}_1$ for the ideal detector having perfect energy resolution.

An analytical expression that allows one to compute the bias of α_{ML} for the ideal detector can be constructed by noting in equation (9) that α_{ML} is a function of the logarithm of the event energies by the term $\frac{1}{N} \sum_{i=1}^N \log[E_i]$. Thus, the random variable $W = \log[E]$ is introduced having probability density function

$$f(w) = \frac{(\alpha_1 - 1) E_1^{\alpha_1} E_2^{\alpha_1} e^{w(1-\alpha_1)}}{E_1 E_2^{\alpha_1} - E_1^{\alpha_1} E_2}, \quad \log[E_1] \leq w \leq \log[E_2] \quad (11)$$

*The TSC used in the ACCESS concept study would detect 52,000 events on average over the energy range of 20 to 5,500 TeV when the spectral parameter, α_1 , is assumed to be 2.8.

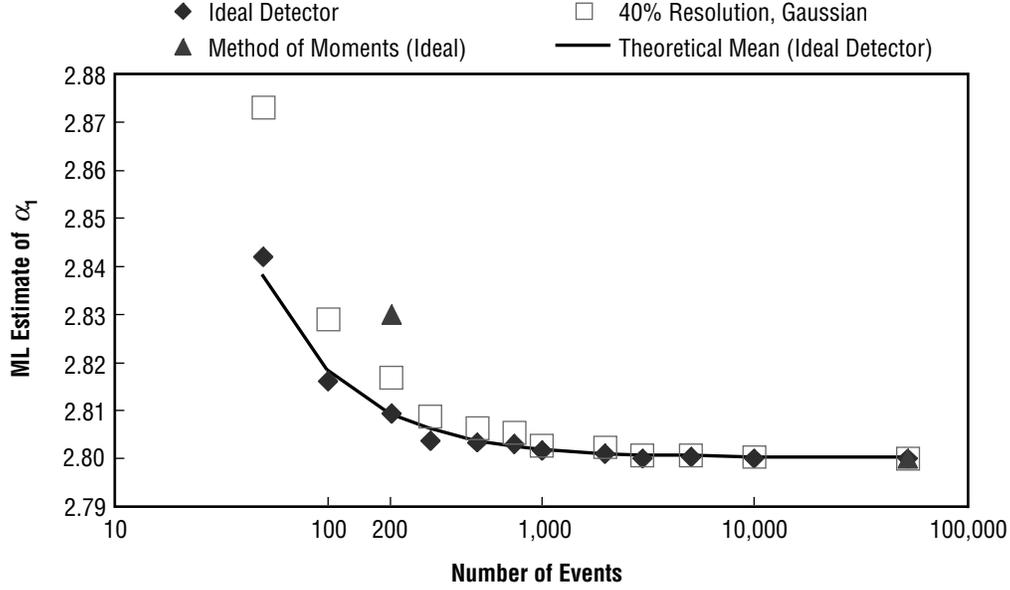


Figure 1. Mean estimates of α_{ML} as a function of the number of events, N , showing a bias for $N < 1,000$. Triangle marker at $N = 200$ and 52,000 is for method of moments.

and with mean and variance

$$\begin{cases} \mu_W = \log[E_1] + \frac{1}{-1 + \alpha_1} + \frac{E_1^{\alpha_1} E_2 \delta^2}{\omega} \\ \sigma_W^2 = \frac{1}{(-1 + \alpha_1)^2} - \frac{E_1^{\alpha_1} E_2 \delta^2}{\omega} - \frac{E_1^{2\alpha_1} E_2^2 \delta^2}{\omega^2} \end{cases} \quad (12)$$

where $\delta = \log[E_1] - \log[E_2]$, $\omega = -E_1^{\alpha_1} E_2 + E_1 E_2^{\alpha_1}$, and, by the central limit theorem, $\frac{1}{N} \sum_{i=1}^N w_i$ is normally distributed with mean μ_W and variance $\frac{1}{N} \sigma_W^2$. Consequently, the probability distribution of the ML estimator α_{ML} of α_1 using the ideal detector is obtained by solving the probability equation,

$$\Pr \left\{ \frac{\alpha_{ML} - 1 - \left[\frac{\log[E_2] E_2^{1-\alpha_{ML}} - \log[E_1] E_1^{1-\alpha_{ML}}}{E_1^{1-\alpha_{ML}} - E_2^{1-\alpha_{ML}}} \right] - \mu_W}{\frac{\sigma_W}{\sqrt{N}}} \leq Z \right\} = \int_{-\infty}^Z \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx, \quad (13)$$

in terms of α_{ML} for various values of Z . Letting Z vary from -5 to 5 , setting N to each of the number of events N used in figure 1, and then numerically evaluating the mean of α_{ML} from the probability density function constructed from equation (13) gives the solid curve shown in figure 1, indicating good agreement with the simulation results.

An interesting observation from figure 1 is that one would likely conclude, and incorrectly, that a significant difference between the slopes of two cosmic-ray elemental species exists if their respective number of events were significantly different from each other and at least one had fewer than 1,000 events. This is because for two given cosmic-ray elemental species, A and B, with simple power law parameters, α and β , the hypothesis $H_0: \alpha - \beta = 0$ (same “slopes”) versus $H_1: \alpha - \beta \neq 0$ uses the test statistic $\alpha_{ML} - \beta_{ML}$ that will inherit the bias(s) shown in figure 1 when N is $<1,000$. Thus, an interesting study would be to plot the estimate of the slope parameter for each of several elemental species as a function of N comprising their respective data sets to see if it resembles figure 1. It should also be understood that this bias as a function of N would be even worse had the method of moments estimation procedure been used to estimate the spectral parameters as previously noted in figure 1.

A comparison of the standard deviation of α_{ML} for the ideal detector with the CRB determined from equation (8) as a function of the number of events N is depicted in figure 2 and clearly shows that α_{ML} attains the CRB for $N > 100$. The standard deviation of the method of moments estimator $\hat{\alpha}_1$ for the ideal detector is also provided for comparison for $N = 200$ and $N = 52,000$, and note once again that MLE provides a superior estimator.

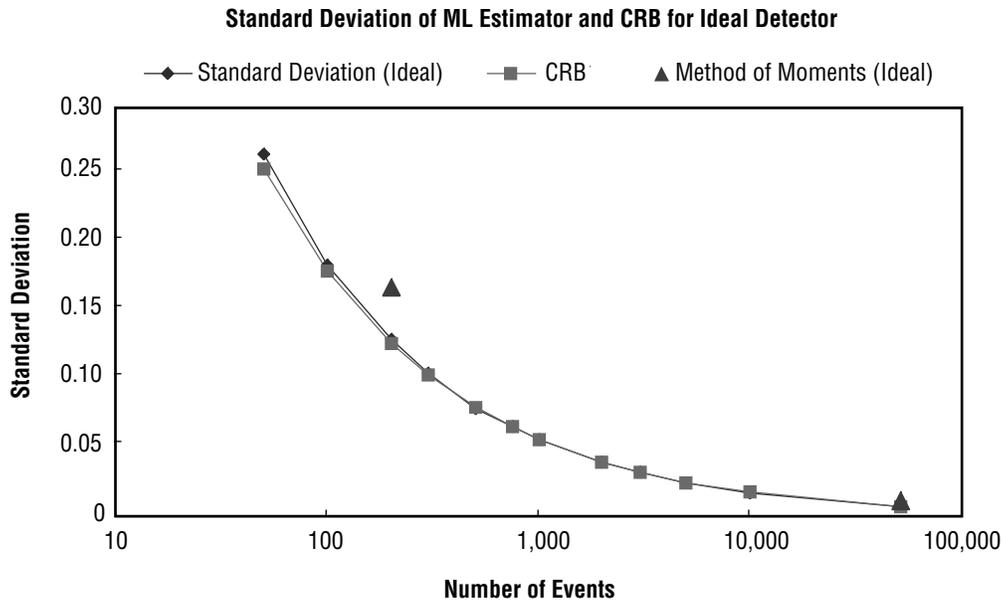


Figure 2. Standard deviation of α_{ML} and the CRB as a function of N .

The comparison for $N = 52,000$ (rightmost markers in fig. 2) is somewhat visually misleading because of the scale of the vertical axis. Their actual values are 0.008, 0.012, and 0.008 for α_{ML} , $\hat{\alpha}_1$, and the CRB, respectively, and the ratio 0.012 to 0.008 is 1.5 for the method of moments and 1.0 for MLE, which is a measure of the efficiency of the estimators and again shows MLE is far superior to the method of moments. Simply put, the improvement in measurement precision provided by MLE over the method of

moments can be roughly equated to doubling the collecting power of the instrument, because doubling the collecting power reduces the standard deviation by $1/\sqrt{2}$ when the CRB is attained. Furthermore, this ratio of ≈ 1.5 remained steady as the detector resolution varied from zero to 50 percent, as shown in figure 3. This fact, coupled with the better performance in achieving P1 as previously discussed, explains why MLE is superior to the method of moments when estimating power law spectra information.

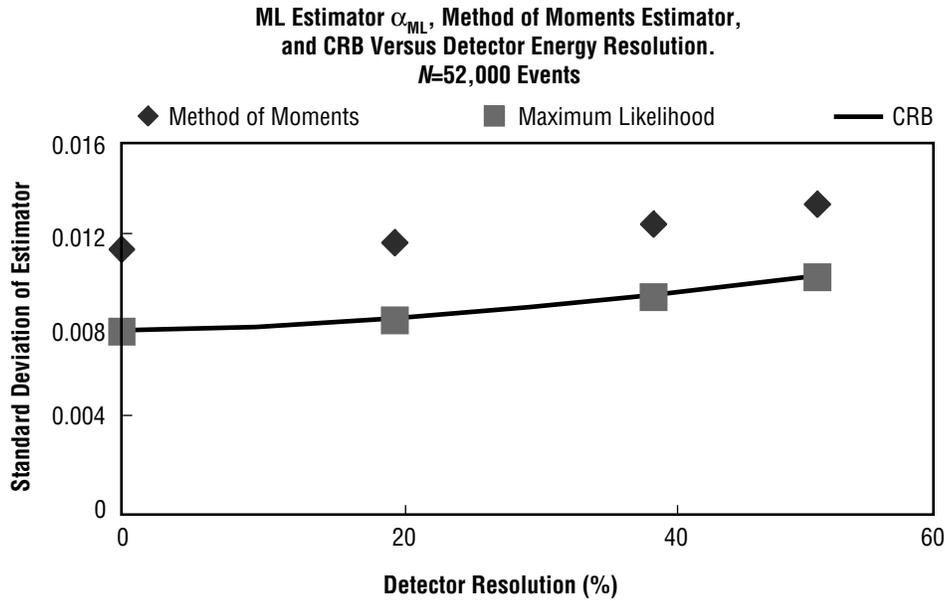


Figure 3. MLE and method of moments as a function of instrument energy resolution for the proposed TSC with Gaussian response function.

Property P3 is investigated using a frequency histogram of α_{ML} based on the 10,000 simulated missions when $N = 50$ events per mission and shows a significant right-hand skewness (fig. 4(a)), and thus, a clear departure from normality (Gaussian fit is illustrated as smooth curves in fig. 4(b)), while a similar comparison for the case $N = 52,000$ shows α_{ML} is very normally distributed. Visual studies of the intermediate values of N showed the frequency histograms to be normally distributed in appearance for $N \geq 1,000$ and is in concert with the bias study depicted in figure 1.

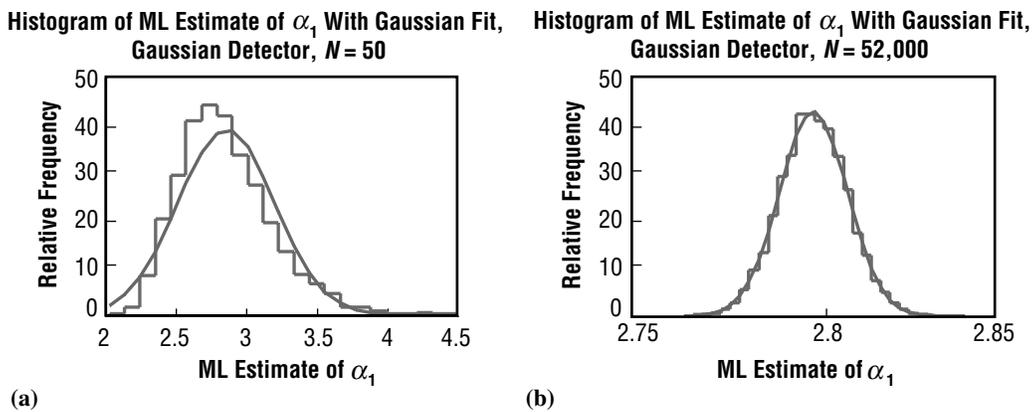


Figure 4. Frequency histograms of α_{ML} for 10,000 simulated missions with (a) $N = 50$ and (b) $N = 52,000$ events, using the TSC with its 40-percent resolution Gaussian response function.

A relative frequency histogram of α_{ML} based on 10,000 simulated missions with $N = 50$ per mission using an ideal detector having zero energy resolution is shown in figure 5. Also shown is the theoretical distribution of α_{ML} obtained from equation (13) with parameters set to $N = 50$, $E_1 = 20$ TeV, $E_2 = 5,500$ TeV, and $\alpha_1 = 2.8$ and illustrates the close agreement between simulation and theory.

Histogram of ML Estimate of α_1 With Theoretical Distribution When $N = 50$ for the Ideal Detector

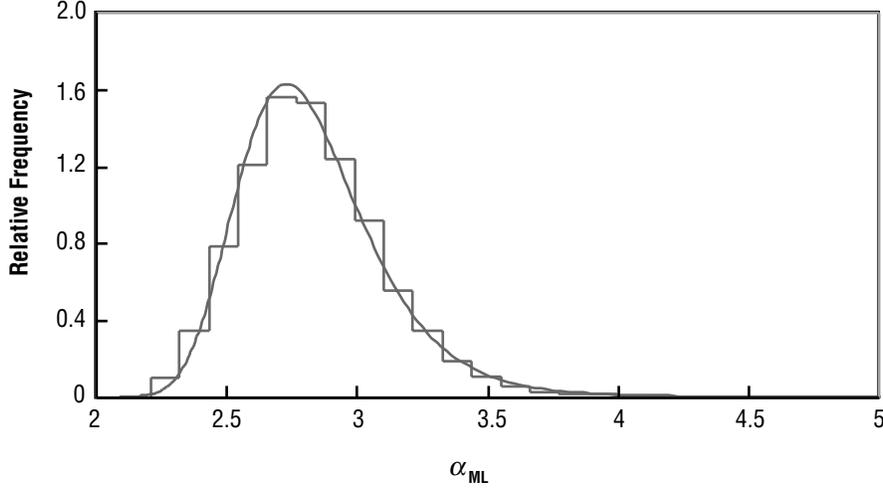


Figure 5. Frequency histogram of α_{ML} based on 10,000 simulated missions with $N = 50$ using an ideal detector. The smooth curve is the theoretical distribution of α_{ML} obtained from equation (13).

Solving equation (3) to obtain the ML estimate for the case where the events are measured by a real detector having nonzero energy resolution is straightforward, and checking consistency (P1) and normality (P3) is easily performed. However, checking efficiency (P2) can be quite formidable because of the term

$$\left\langle \left(\frac{\partial \log[g(y; \alpha_1)]}{\partial \alpha_1} \right)^2 \right\rangle = \int_0^{\infty} \left[\frac{\partial}{\partial \alpha_1} \log \left[\int_{E_1}^{E_2} g(y | E, \rho) \phi_S(E; \alpha_1) dE \right] \right]^2 \left[\int_{E_1}^{E_2} g(y | E, \rho) \phi_S(E; \alpha_1) dE \right] dy \quad (14)$$

required to compute the CRB, coupled with the fact that the detector response function, g , in equation (14) can be quite complicated. For example, g could be Gaussian with non-negativity constraint $y > 0$ and energy-dependent resolution function, $\rho(E)$, that in turn requires an energy-dependent normalizing coefficient. Fortunately, equation (14) can be numerically evaluated using the symmetrized form of the numerical derivative,¹¹

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h}, \quad (15)$$

to approximate the derivative in equation (14) and in conjunction with the method of Gaussian quadratures to calculate the definite integrals. The fact that the CRB in equation (8) for the ideal detector must match the CRB obtained from equation (14) when the detector resolution is zero ($\rho \rightarrow 0$) provides a means to tune the numerical differentiation parameter, h , and the integration parameters, e.g., the upper integration limit

for y as well as the number of partitions used in the numerical integration in both integration variables, E and y , in equation (14). For the TSC instrument with a data analysis range of 20 to 5,500 TeV, setting h to 0.0001 and the upper integration limit of y to 35,000 GeV in place of infinity in equation (14), and using 10-point Gaussian quadratures over subintervals over both integration ranges provides the somewhat surprising result of 13-decimal-place accuracy in the numerical evaluation of equation (14) when compared to the exact value obtained from equation (8) for the ideal detector. This accuracy in the numerical evaluation of equation (14) was independently confirmed using the numerical integration routine in *MATHEMATICA*[®].

Figure 6 illustrates the convergence of the standard deviation of the ML estimator α_{ML} to the CRB computed using equation (14) for a 40-percent resolution Gaussian detector as a function of N . The standard deviation of α_{ML} is based on a battery of 10,000 simulated missions for each value of N , where N ranges from 50 to 52,000 events per mission. The CRB for the ideal detector is included as a reference curve (---).

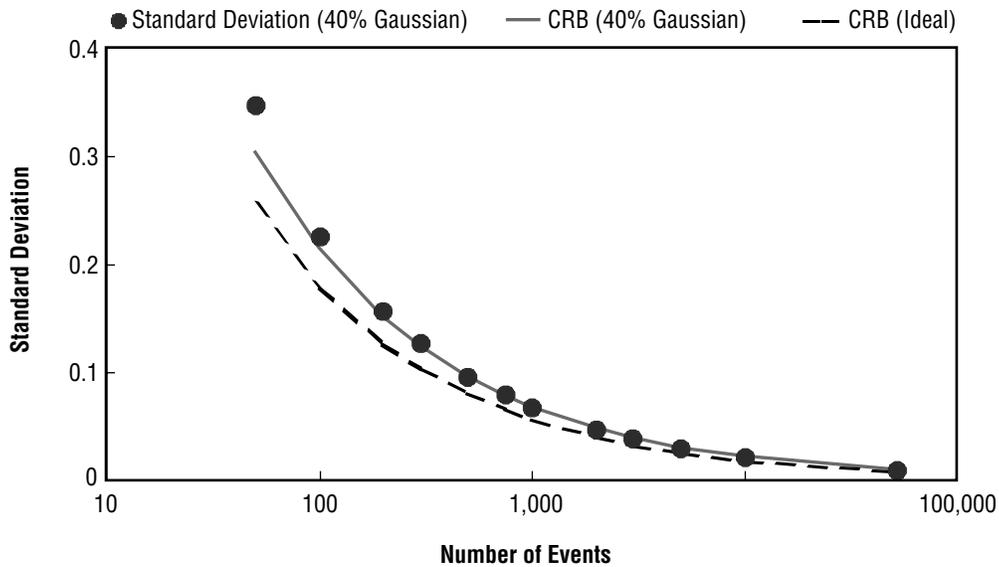


Figure 6. Standard deviation of the ML estimator α_{ML} versus N for a 40-percent resolution Gaussian detector based on 10,000 simulated missions at each value of N and the CRB for the 40-percent Gaussian detector computed using equation (14) (—). The CRB for the ideal detector is included as a reference curve (---).

When MLE is being used in the design phase of an instrument to estimate its expected performance and if the simulations indicate that MLE does in fact provide unbiased spectra information and approximate attainment of the CRB for the science mission under study, then equation (14) can be used to evaluate the relative merits of various instrument design parameters without performing additional simulations. This has tremendous practical value in design parameter trade studies because equation (14) can be evaluated in mere seconds, while the equivalent information from Monte Carlo simulations can take several days. For example, because we know (from Fig. 2) that MLE attains the CRB for $N \geq 100$ events, equation (14) can be used to compute the family of curves shown in figure 7 that relate the precision with which α_1 can be measured as a function of detector resolution and collecting power. This implies instrument designers should first attempt to maximize collecting power and then improve resolution, and in that order. The proposed TSC instrument, with its expected 52,000 events, is indicated by the square in figure 7. The reader is referred to figure 3 for a detailed view of the CRB calculated using equation (14) for the TSC as a function of detector energy resolution.

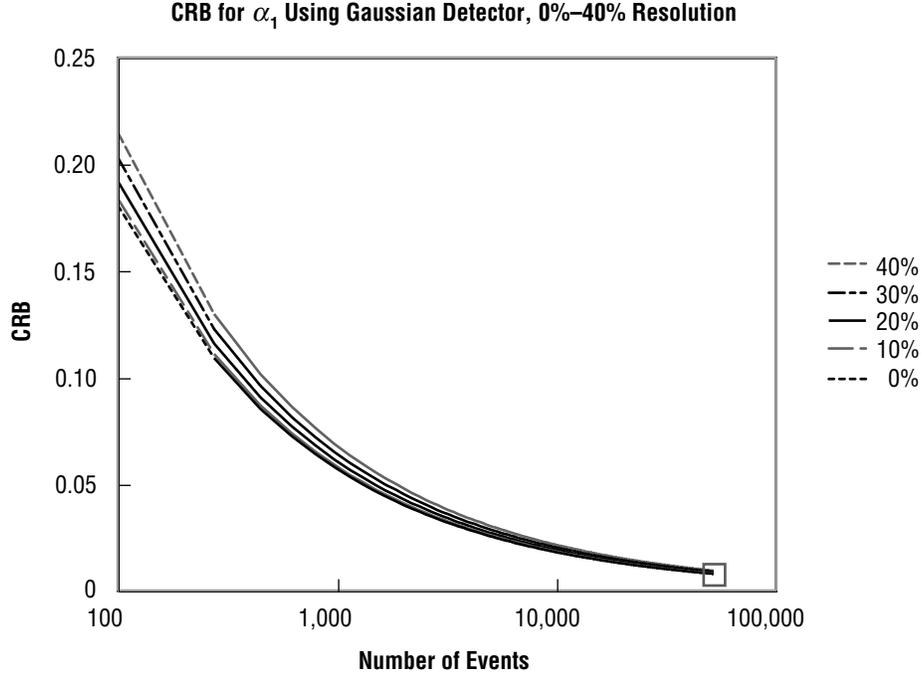


Figure 7. CRB as a function of N for detector energy resolutions in the range $0 \leq \rho \leq 0.40$. The proposed TSC instrument with its expected 52,000 events is indicated by the square, and the observing energy range was set to 20–5,500 TeV.

A detailed simulation study of the TSC-sized ideal detector, with its expected $N = 52,000$ events for the observing range 20–5,500 TeV and with $\alpha_1 = 2.8$, was conducted and α_{ML} obtained from equation (9) for each of 1 million missions (each mission detected 52,000 events), yielding a mean and standard deviation value of α_{ML} to be 2.80003 and 0.007905, respectively. Constructing the probability density function of α_{ML} from equation (13) and then numerically evaluating its mean and standard deviation gives 2.80003 and 0.007911, respectively, while the CRB calculated from equation (8) gives 0.00790998, illustrating the remarkable agreement between simulation and theory and attainment of P2. Note that α_{ML} is essentially unbiased too and thus P1 is approximately attained. Because 5.2×10^{10} random numbers were required for this 1 million simulated missions study, it is crucial to use a random number generator having a period longer than 5.2×10^{10} , such as the generator used in this study which has a period of $\approx 10^{18}$.

1.2 Broken Power Law Energy Spectrum

The broken power law energy spectrum suggests a transition from spectral index α_1 below the knee location at energy E_k to a steeper spectral index $\alpha_2 > \alpha_1$ above the knee.* The broken power law spectrum predicts that the number of protons detected above an energy, E , is given by

*The case $\alpha_2 < \alpha_1$ and where E_k is referred to as the ankle can also be handled by this MLE procedure.

$$N_B(> E) = \begin{cases} M_A \left(\frac{\alpha_1 - 1}{\alpha_2 - 1} \right) \left(\frac{E_k}{E_A} \right)^{-\alpha_1 + 1} \left(\frac{E}{E_k} \right)^{-\alpha_2 + 1} & \text{for } E \geq E_k \\ N_S(> E) - [N_S(> E_k) - N_B(> E_k)] & \text{for } E < E_k \end{cases}, \quad (16)$$

where E is in units TeV, M_A and E_A are 160 and 500 TeV, respectively, as before for the TSC instrument, $N_S(>E)$ is defined in equation (6), and currently available measurements suggest that α_1 is ≈ 2.8 , α_2 is thought to be somewhere between 3.1 and 3.3, and E_k is parameterized in the range of 100 to 300 TeV for this study. The broken power law probability density function ϕ_B is obtained by normalizing N_B over an observing range $[E_1, E_2]$ of interest and is defined in equation (21) of appendix A.

The likelihood function of a random sample of N GCR events from the broken power law spectrum detected by the ideal detector having perfect energy resolution, regarded as a function of the unknown vector of parameters $\boldsymbol{\theta} = (\alpha_1, \alpha_2, E_k)$, is

$$L(\boldsymbol{\theta}) = A(\boldsymbol{\theta})^N \left(\prod_{E_i < E_k} \frac{E_i}{E_k} \right)^{-\alpha_1} \left(\prod_{E_j \geq E_k} \frac{E_j}{E_k} \right)^{-\alpha_2}, \quad E_1 \leq E_i, E_j \leq E_2, \quad (17)$$

where the first product is over the event energies below the knee location E_k and the second product is over those event energies above E_k , and they total in number to N , and $A(\boldsymbol{\theta})$ is the normalizing coefficient given in equation (22).

The Nelder-Mead simplex method can then be used to obtain $\boldsymbol{\theta}_{\text{ML}}$ from equation (18), where the objective function $O(\boldsymbol{\theta})$ is defined as minus the log-likelihood function, so that

$$\boldsymbol{\theta}_{\text{ML}} = \min_{\{\boldsymbol{\theta}\}} \left[-N \log A(\boldsymbol{\theta}) + \alpha_1 \left(\sum_{E_i < E_k} \log \left[\frac{E_i}{E_k} \right] \right) + \alpha_2 \left(\sum_{E_j \geq E_k} \log \left[\frac{E_j}{E_k} \right] \right) \right] \quad (18)$$

for N events detected by the ideal detector, while equation (2) must be used to construct the likelihood function for a real detector having response function, g , and energy resolution, ρ , with N instrument responses y_i . Consequently, the ML estimate $\boldsymbol{\theta}_{\text{ML}}$ is

$$\boldsymbol{\theta}_{\text{ML}} = \min_{\{\boldsymbol{\theta}\}} \sum_{i=1}^N \left(-\log \left[\int_{E_1}^{E_k} g(y_i | E, \rho) A(\boldsymbol{\theta}) \left(\frac{E}{E_k} \right)^{-\alpha_1} dE + \int_{E_k}^{E_2} g(y_i | E, \rho) A(\boldsymbol{\theta}) \left(\frac{E}{E_k} \right)^{-\alpha_2} dE \right] \right), \quad (19)$$

where the range of integration must be split at E_k at each step in the simplex search for $\boldsymbol{\theta}_{\text{ML}} = (\alpha_1, \alpha_2, E_k)_{\text{ML}}$.

To numerically explore the properties of θ_{ML} for the broken power law distribution, the vector of spectral parameters is first set to $\theta = (2.8, 3.3, 100 \text{ TeV})$ and events simulated from the energy range 20 to 5,500 TeV for each of several values of N selected so as to provide an average of 500, 1,000, ..., 5,000 events above E_k as shown in table 1. The notation N_2 is introduced to denote $N_B(>E_k)$, N for $N_B(>E_1)$ so that $N_1=N-N_2$, and the notation $N(N_2)$ means “a total of N events, of which N_2 of them are above the spectral knee E_k .”

Table 1. Number of events used in broken power law simulations for 0.5 break-size study.

N_1	10,939	21,877	32,816	43,754	65,631	87,508	109,390
N_2	500	1,000	1,500	2,000	3,000	4,000	5,000
N	11,439	22,877	34,316	45,754	68,631	91,508	114,390

For each value of N in table 1, 10,000 missions were simulated and for each of these missions, θ_{ML} was obtained using equation (18) for an ideal detector and equation (19) for the TSC detector having Gaussian response function g and constant 40-percent energy resolution over the simulated energy range of 20 to 5,500 TeV.

Figure 8 depicts the mean of the ML estimates of α_1 and α_2 versus the number of events N used in the simulations and shows that when the collecting power of the detector provides $\geq 1,500$ events above E_k (corresponding to third set of markers from left), property P1 is essentially attained by the TSC instrument since the relative bias is < 3 percent for the 40-percent resolution Gaussian detector, and is even better for the ideal detector having zero energy resolution.

Similarly, figure 9 illustrates the bias (recall $E_k=100 \text{ TeV}$ in these simulations) of the ML estimate of E_k as a function of N for the TSC instrument and the ideal detector. Note that property P1 is again roughly attained by the TSC (relative bias is ≤ 2 percent) when there are $\geq 1,500$ events above E_k .

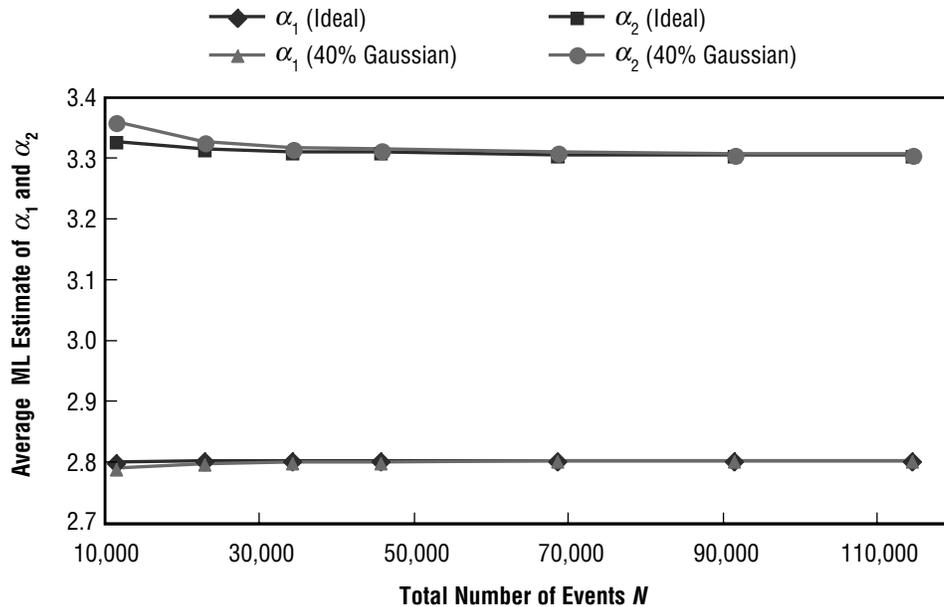


Figure 8. ML estimate of α_1 and α_2 as a function of detector collecting power.

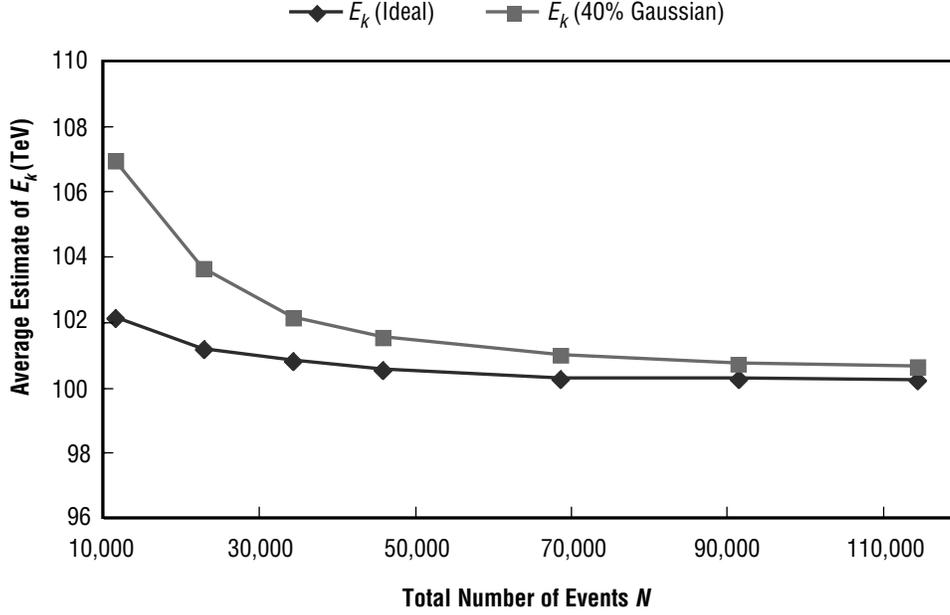


Figure 9. ML estimate of E_k as a function of detector collecting power using a 40-percent resolution Gaussian response function (the TSC) and the ideal detector. The knee location E_k was set to 100 TeV in these simulations.

Next, property P2 is investigated and requires the construction of the 3-by-3 information matrix $\mathbf{I}(\boldsymbol{\theta})$. Equation (32) of appendix A provides $\mathbf{I}(\boldsymbol{\theta})$ for the ideal detector, while for a real detector with response function, g , and energy resolution, ρ , the ij -element of $\mathbf{I}(\boldsymbol{\theta})$ is, by equation (4),

$$\begin{aligned}
 I_{ij}(\boldsymbol{\theta}) = N \int_0^{\infty} & \left[\frac{\partial}{\partial \theta_i} \log \left[\int_{E_1}^{E_2} g(y|E, \rho) \phi_B(E; \boldsymbol{\theta}) dE \right] \right] \left[\frac{\partial}{\partial \theta_j} \log \left[\int_{E_1}^{E_2} g(y|E, \rho) \phi_B(E; \boldsymbol{\theta}) dE \right] \right] \\
 & \times \left[\int_{E_1}^{E_2} g(y|E, \rho) \phi_B(E; \boldsymbol{\theta}) dE \right] dy
 \end{aligned} \tag{20}$$

and can be accurately computed using the numerical methods discussed in the simple power law section, and where the notation in equation (20) defines $\theta_1 \equiv \alpha_1$, $\theta_2 \equiv \alpha_2$, and $\theta_3 \equiv E_k$, and where the integration range $[E_1, E_2]$ must be split at E_k for the inner three integrals.

A comparison of the CRB obtained from equation (20) for α_1 using the TSC with its 40-percent Gaussian response function with the simulation results is shown in figure 10. Note the CRB is attained when the number of events above the knee location is $\geq 1,500$. The case 11,439 (500) had several simulated missions in which the MLE procedure gave an estimate $\boldsymbol{\theta}_{\text{ML}}$ of $\boldsymbol{\theta}$ that suggests a simple power law would probably be an adequate explanation of the simulated events. These errant estimates were characterized as EE1) E_k and α_2 are both very large relative to their assumed values of 100 TeV and 3.3 in the simulations,

and EE2) E_k and α_1 are both very small. While the condition $\alpha_1 \approx \alpha_2$ is normally associated with suggesting a simple power law adequately fits the data, these unlucky missions illustrate the beauty of the MLE procedure in finding two other conditions whereby a broken power law collapses into a simple power law. The first condition is a broken power law with E_k above the range of detected events and $\alpha_2 \rightarrow \infty$ in an effort to explain the absence of events above E_k , which is indeed just a simple power law over the range of detected events and implied by EE1. The second condition is a broken power law with E_k below the range of detected events and $\alpha_1 \rightarrow 0$ and implied by EE2. Eliminating these errant estimates of α_1 gives the trimmed standard deviation depicted at $N = 11,439$ (and $N_2 = 500$) in figure 10 and symbolized by a filled circle on the plot. The CRB for the ideal detector calculated from equation (42) is also shown in figure 10 with corresponding simulation results. Additionally, the difference between the covariance matrix of the ML estimates and $\mathbf{I}^{-1}(\boldsymbol{\theta})$ was noted to be positive definite as each of its three eigenvalues were positive, with two of them approximately zero for all values of N in table 1 used in the simulations.

Similar results are illustrated in figure 11 for α_2 and figure 12 for E_k . Trimmed estimates for these two make little difference because of their already larger variance relative to that of α_1 .

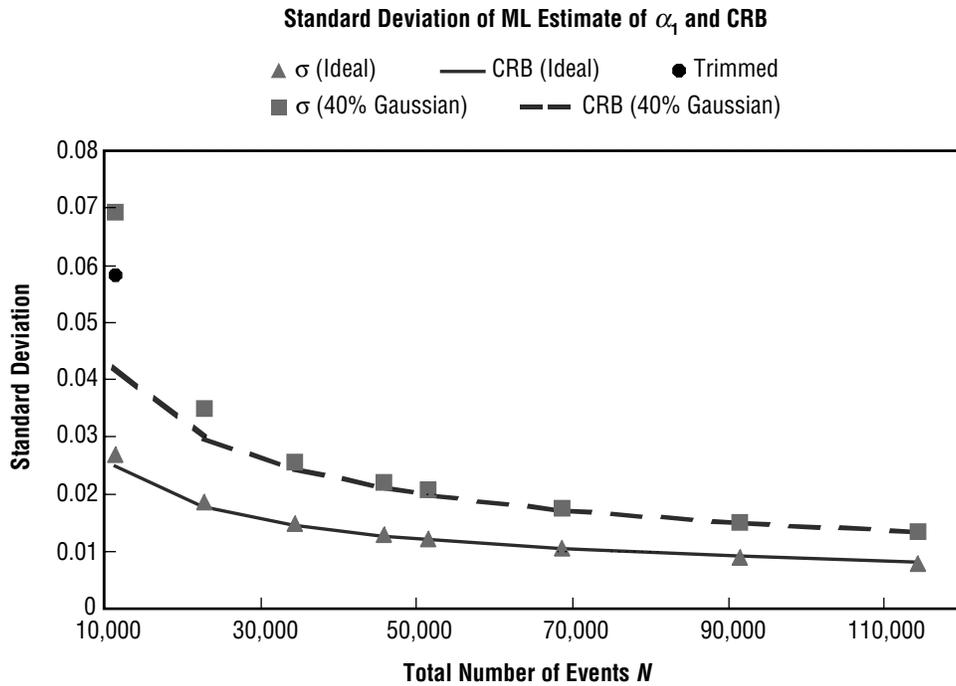


Figure 10. CRB of α_1 using TSC (— —) and ideal detector (—) obtained from equation (20) versus collecting power. Standard deviations of ML estimates from simulations for values of N in table 1 indicated by markers.

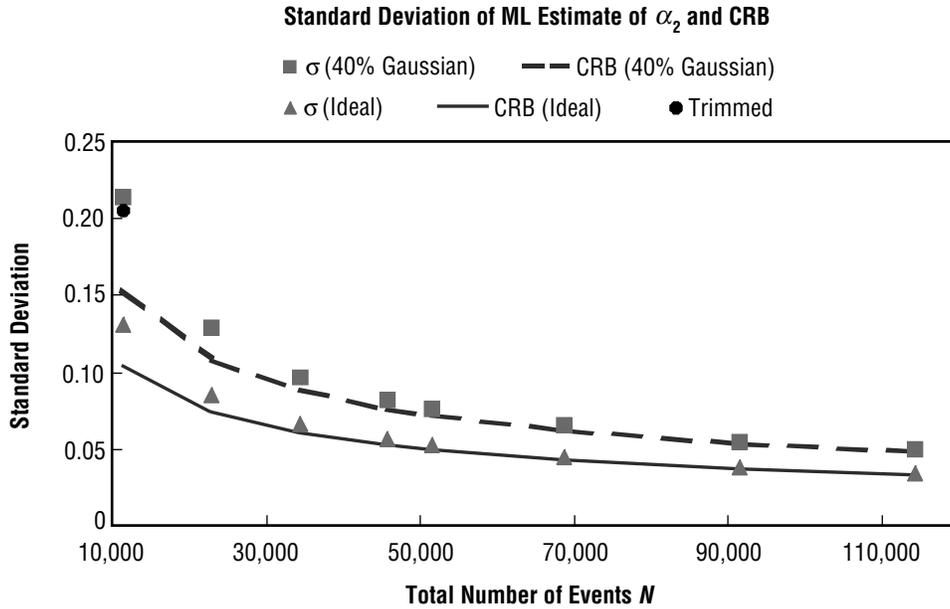


Figure 11. CRB of α_2 using TSC (— —) and ideal detector (—) obtained from equation (20) versus collecting power. Standard deviations of ML estimates from simulations for values of N in table 1 indicated by markers.

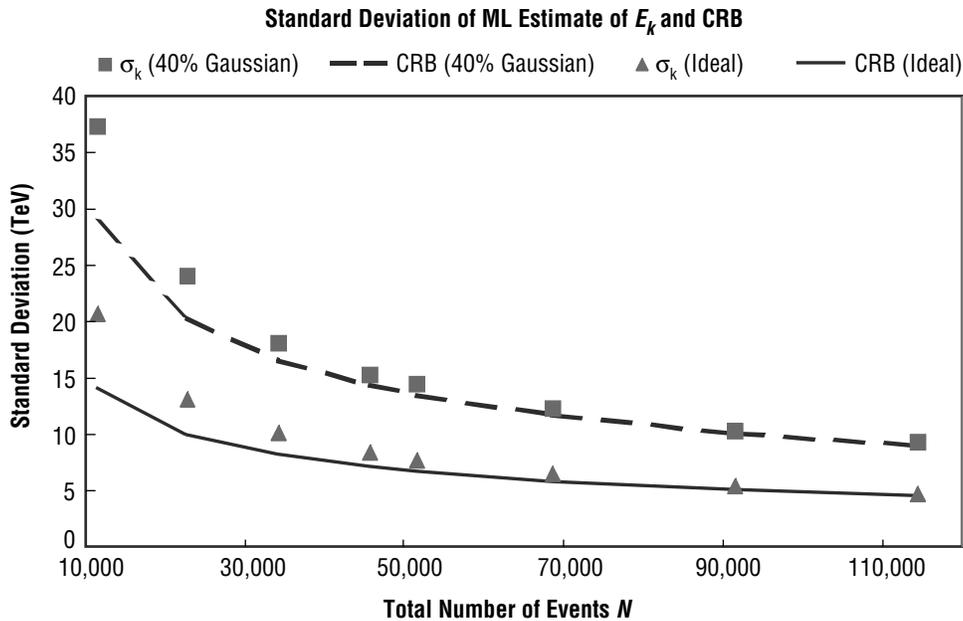


Figure 12. CRB of E_k using TSC (— —) and ideal detector (—) obtained from equation (20) versus collecting power. Standard deviations of ML estimates from simulations for values of N in table 1 indicated by markers.

The property of asymptotic normality (P3) of θ_{ML} is next investigated with the aid of relative frequency histograms of the components of θ_{ML} provided from the simulations. Figure 13 shows relative frequency histograms of the 10,000 ML estimates of α_1 and α_2 for the two collecting powers that provide 11,439 (500) events and also 114,390 (5,000) events and correspond to the first and last columns of table 1. As before, the detector here is the TSC with its Gaussian response function and 40-percent energy resolution and with its collecting power adjusted through the choice of N . Note that while the histograms corresponding to the larger collecting power are approximately normally distributed and well separated, those corresponding to the smaller detector are skewed and even slightly overlapping, indicating the onset of difficulties in detecting the broken power law parameters. Relative frequency histograms for the ideal detector (not shown) show no overlap for the $N = 11,439$ (500) case and suggest that this is the approximate boundary for fixed N where detector resolution can play a leveraging role for this set of parameters.

**Histograms of ML Estimate of α_1 and α_2 , for Gaussian Response Function With 40% Resolution.
 $\theta = (2.8, 3.3, 100 \text{ TeV})$, Energy Range 20–5,500 TeV. Average of 500 and 5,000 Events
 Above Knee, With 11,000 and 110,000 Below, Respectively.
 10,000 Simulated Missions**

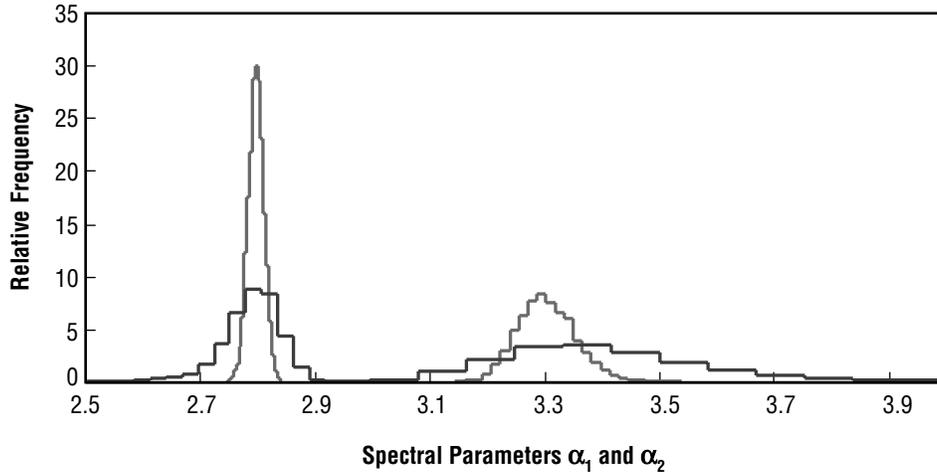


Figure 13. Relative frequency histograms of the ML estimate of α_1 (leftmost two histograms) for $N = 11,439$ (broadest of the two) and $N = 114,390$ (narrow histogram). Rightmost two histograms similarly defined for α_2 .

Frequency histograms of the ML estimates of E_k for these two cases of N are shown in figure 14 and once again, note that the larger sized detector has roughly attained P3 while the smaller sized detector has not, and in fact a “bump” in the tail of the broader distribution for the smaller detector is seen, suggesting a simple power law would likely be an adequate explanation of these particular mission results. This situation was previously discussed and referred to as EE1.

**Histograms of ML Estimate of E_k for Gaussian Response Function With 40% Resolution.
 $\theta = (2.8, 3.3, 100 \text{ TeV})$, Energy Range 20–5,500 TeV. Average of 500 and 5,000 Events
 Above Knee, With 11,000 and 110,000 Below, Respectively.
 10,000 Simulated Missions**

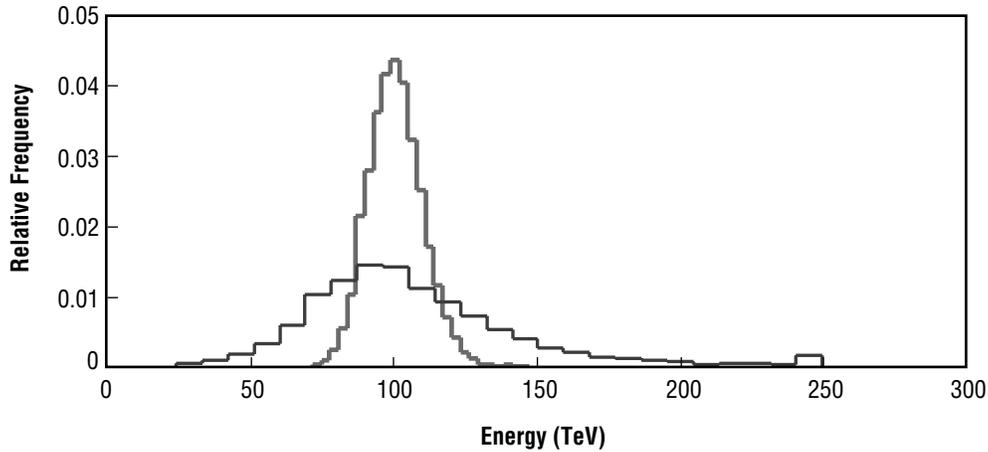


Figure 14. Relative frequency histograms of the ML estimates of E_k for $N = 11,439$ (broadest of the two and with bump in right-hand tail) and $N = 114,390$.

Next, figure 15 shows relative frequency histograms of the 10,000 ML estimates of α_1 and α_2 when $N = 22,877$, providing an average of $N_2 = 1,000$ events above E_k , and the two histograms are seen to be clearly separated. This suggests that a detector with this collecting power and a 40-percent resolution Gaussian response function could indeed measure the three broken power law spectral parameters when their true values are $\theta = (2.8, 3.3, 100 \text{ TeV})$. Because the concept TSC that was studied would detect 51,576 (2,255) events on average in the energy range 20–5,500 TeV when $\theta = (2.8, 3.3, 100 \text{ TeV})$, it is concluded that it could measure the three spectral parameters when $E_k \approx 100 \text{ TeV}$ and the break-size is ≈ 0.5 .

**Histogram of ML Estimate of α_1 and α_2 for Gaussian Response Function With 40% Resolution.
 $\theta = (2.8, 3.3, 100 \text{ TeV})$, Energy Range 20–5,500 TeV. Average of 1,000 Events
 Above Knee, 21,877 Below.
 10,000 Simulated Missions**

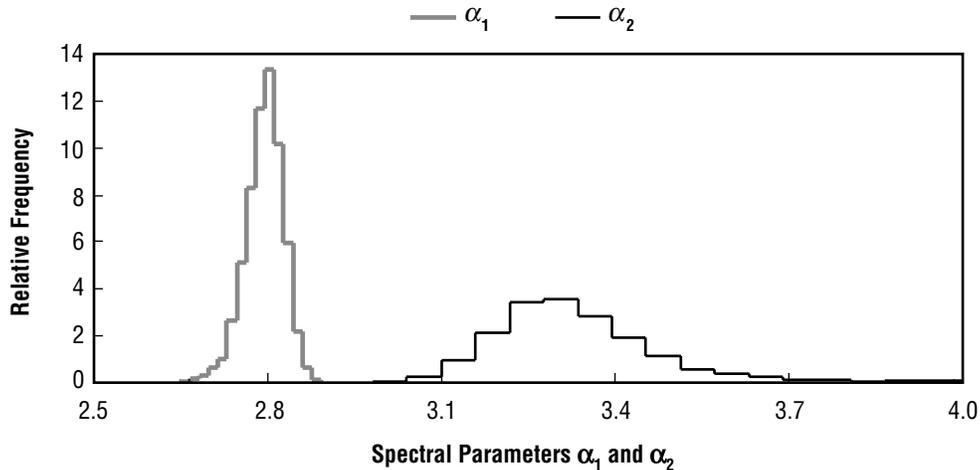


Figure 15. Relative frequency histograms of ML estimates of α_1 and α_2 for $N = 22,877$ and $N_2 = 1,000$.

1.3 Break-Size 0.3 Study

For this study the vector of spectral parameters is set to $\theta=(2.8, 3.1, 100 \text{ TeV})$ and events simulated from the energy range 20–5,500 TeV for each of several values of N selected so as to provide an average of 1,000, 2,000, ..., 5,000 events above E_k as shown in table 2. (Values of $N_2 < 1,000$ produced too many errant ML estimates θ_{ML} of θ to be useful.)

For each value of N in table 2, 5,000 missions were simulated and for each mission, θ_{ML} was obtained using equation (18) for an ideal detector, and equation (19) for the TSC detector having Gaussian response function and constant 40-percent energy resolution over the simulated energy range 20–5,500 TeV. Figure 16 shows that when the number of detected events above the knee is $\geq 2,000$, the ML estimate of α_1 and α_2 is essentially unbiased and property P1 is attained, while figure 17 indicates the ML estimate of E_k is still somewhat biased, even when $N_2 = 2,000$ (second markers from left) for the 40-percent resolution Gaussian detector, which is perhaps not surprising in light of the more difficult estimation task for this smaller break-size case.

Table 2. Number of events used in broken power law simulations for 0.3 break-size study.

N_1	19,977	39,954	59,931	79,909	99,886
N_2	1,000	2,000	3,000	4,000	5,000
N	20,977	41,954	62,931	83,909	104,886

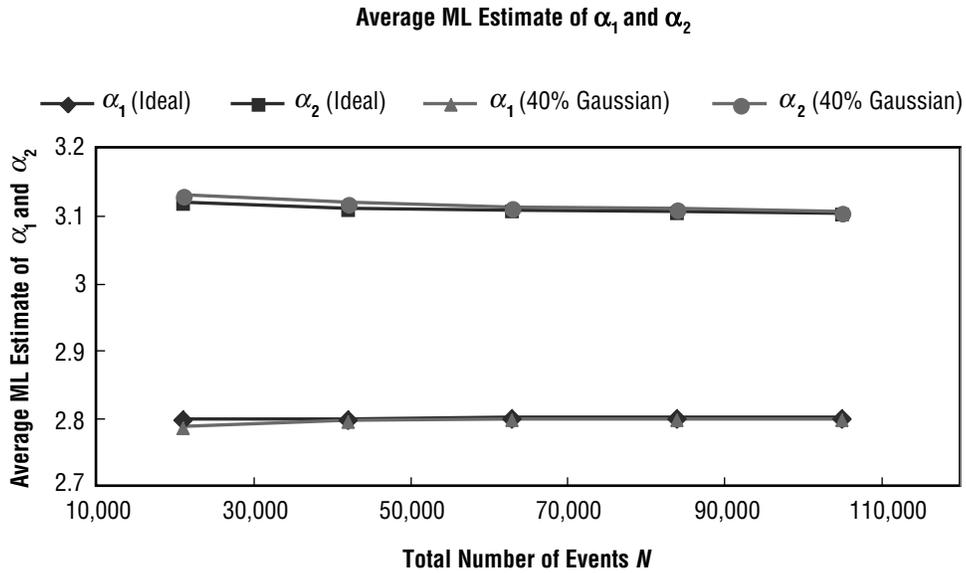


Figure 16. ML estimate of α_1 and α_2 as a function of detector collecting power when the spectral break-size is 0.3 for the TSC and the ideal detector.

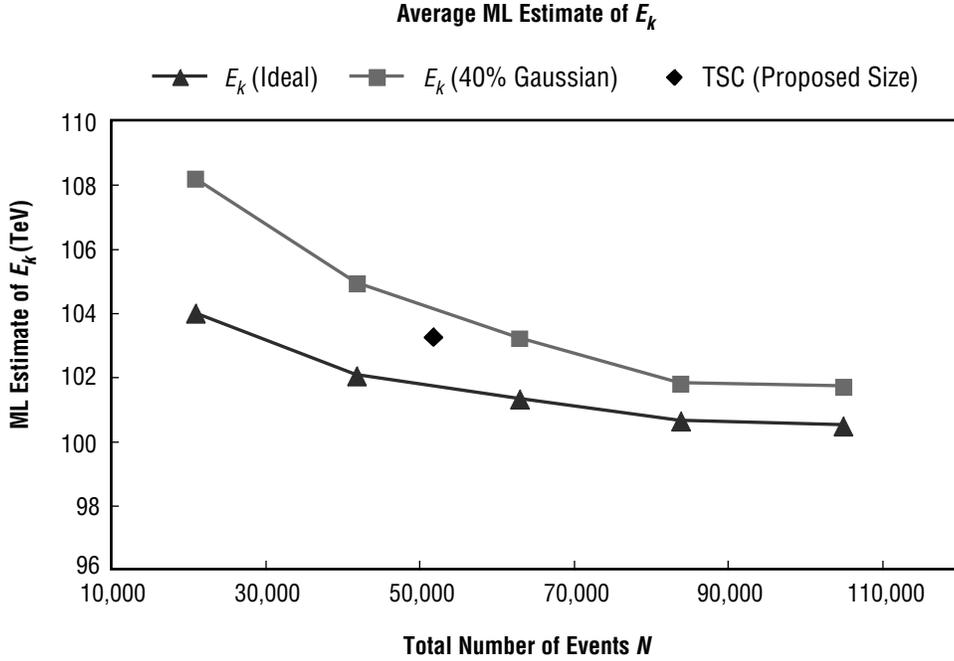


Figure 17. ML estimate of E_k as a function of detector collecting power using the TSC and ideal detector when the break-size is 0.3. The actual concept TSC with its expected $N = 51,790$ events is indicated by the diamond and is based on 25,000 simulated missions (others are 5,000 missions each) and suggests the marker to its immediate right is probably a little on the high side.

Figure 18 shows the CRB of α_1 using the TSC detector and also the ideal detector versus the number of detected events N . The markers represent the standard deviation of the 5,000 ML estimates of α_1 based on the simulations. Note that when $N_2 = 1,000$, MLE experienced several missions resulting in errant estimates of α_1 in its attempt to place the knee before the data and then drive $\alpha_1 \rightarrow 0$ (condition EE2), suggesting a simple power law might be a suitable fit for those simulated missions. Trimmed estimates are also provided in figure 18 corresponding to the cases where $N_2 = 1,000$ and $N_2 = 2,000$ and indicated by filled circles. Also note the ideal detector with its zero-percent resolution attains the CRB for all the values of N in table 2.

Similar comparisons between the CRB and the ML estimate of α_2 and E_k are shown in figures 19 and 20, respectively. Figure 20 indicates the CRB for E_k is particularly difficult to attain, even for the ideal detector. Trimmed estimates are indicated by filled circles for the first two values of N , corresponding to an average of 1,000 and 2,000 events above E_k , respectively.

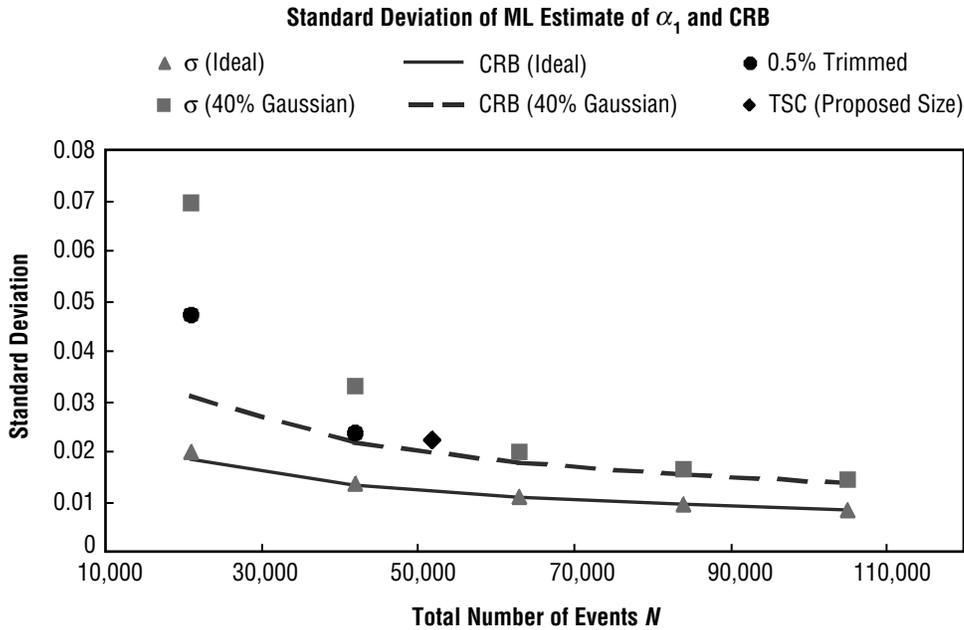


Figure 18. CRB of α_1 using TSC (— —) and ideal detector (—) versus collecting power. Standard deviations of ML estimates from simulations for values of N in table 2 indicated by markers.

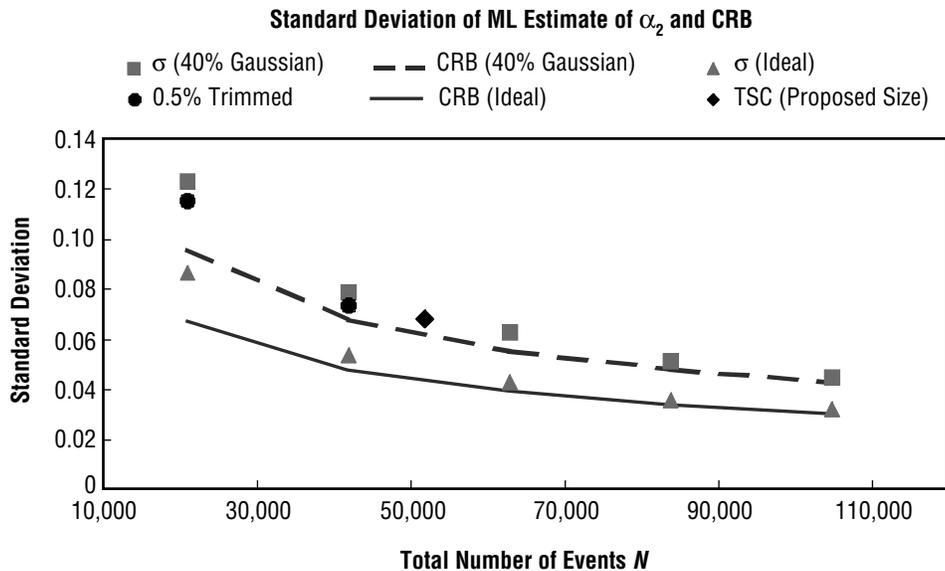


Figure 19. CRB of α_2 using TSC (— —) and ideal detector (—) versus collecting power. Standard deviations of ML estimates from simulations for values of N in table 2 indicated by markers.

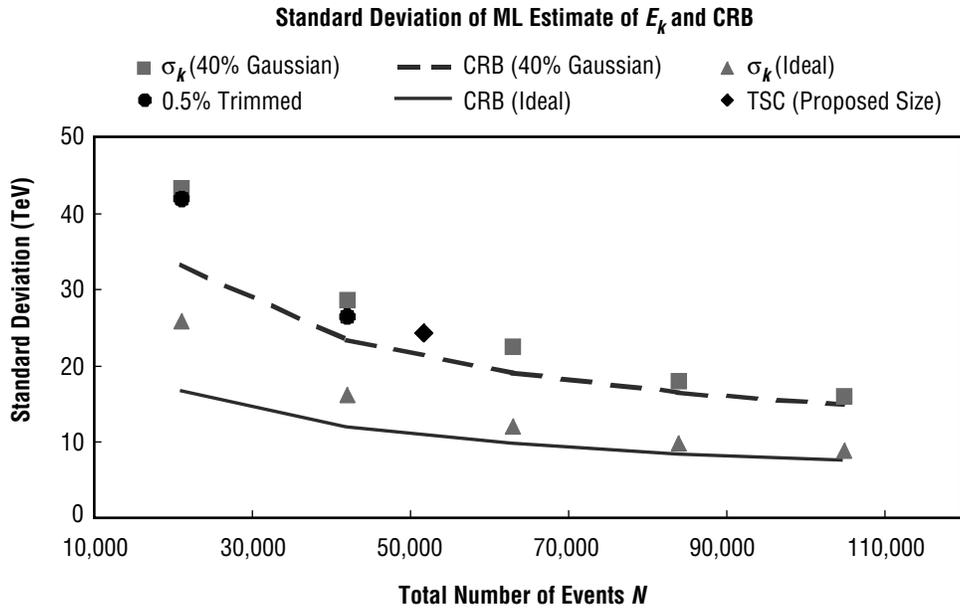


Figure 20. CRB of E_k using TSC (---) and ideal detector (—) versus collecting power. Standard deviations of ML estimates from simulations for values of N in table 2 indicated by markers.

To investigate the properties of θ_{ML} for the proposed TSC with its 40-percent resolution Gaussian response function, 25,000 missions were simulated with $\theta = (2.8, 3.1, 100 \text{ TeV})$, providing 51,790 (2,470) events on average from the observing range 20–5,500 TeV. Frequency histograms of the ML estimates of α_1 and α_2 are shown in figure 21 for the proposed TSC and a clear separation between the histograms is observed. A slight right-hand skewness in the ML estimates of α_2 is noted.

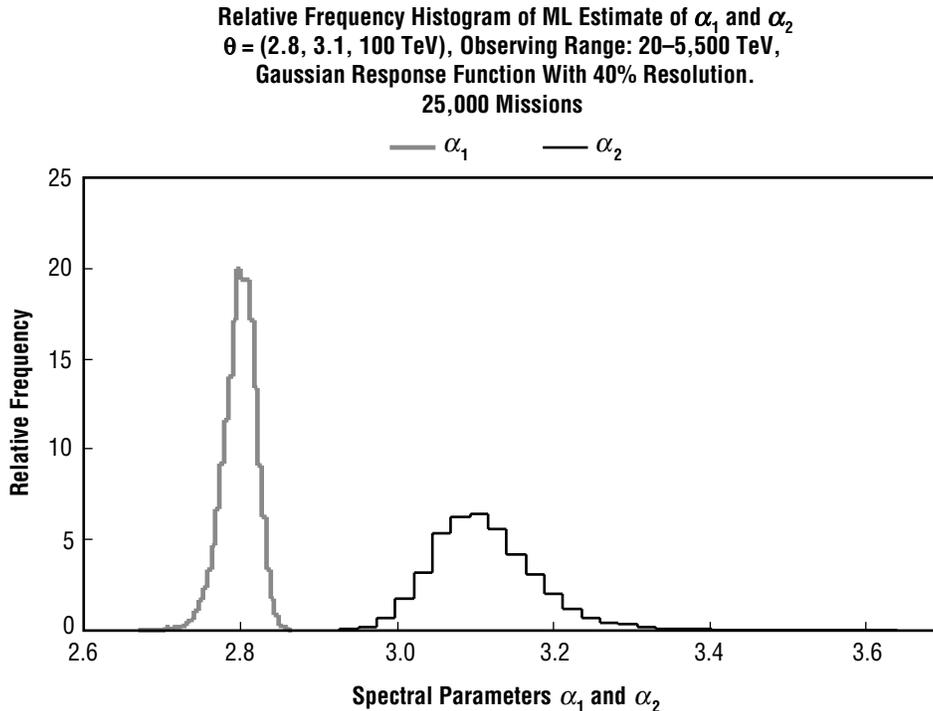


Figure 21. Relative frequency histograms of ML estimates of α_1 and α_2 for the proposed TSC.

Figure 22 shows the relative frequency histogram of the ML estimates of E_k using the proposed TSC and the long, right-hand tail suggests the possibility of a few missions that resulted in errant estimates of the form EE1. Also note the skewness and thus slight departure from normality (property P3).

Table 3 gives summary statistics of θ_{ML} for the simulated missions. The rows labeled “theoretical limits” under the Comments column provide the input parameters θ used in the simulations and the CRB obtained from equation (20), indicating that θ_{ML} is approximately unbiased, efficient, and normally distributed so that properties P1, P2, and P3 are roughly attained by the proposed TSC for this set of spectral parameters. Similar information for the zero-resolution ideal detector is also provided in table 3.

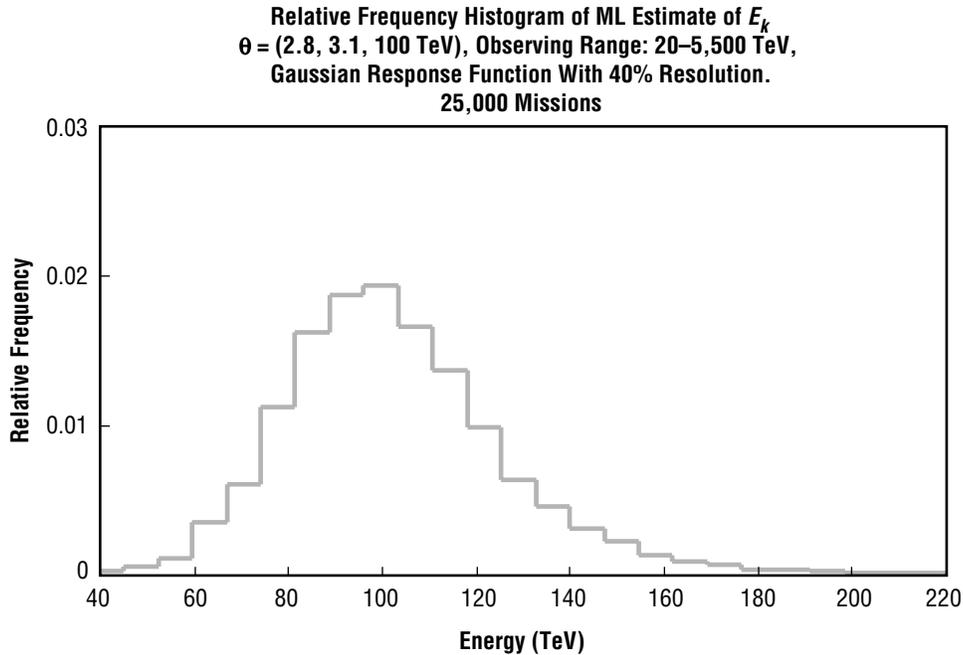


Figure 22. Relative frequency histograms of the ML estimates of E_k for the proposed TSC.

Table 3. Summary statistics based on 25,000 simulated missions with $\theta = (2.8, 3.1, 100 \text{ TeV})$ and observing range 20–5,500 TeV, for the TSC having Gaussian response function.

Resolution (%)	E_k (TeV) $N(N_2)$	Mean			Standard Deviation			Comments
		α_1	α_2	E_k (TeV)	α_1	α_2	E_k (TeV)	
0	100 51,790 (2470)	2.80	3.10	100	0.012	0.043	10.6	Theoretical limit
0		2.80	3.10	101	0.012	0.048	13.7	Simulation (25,000 missions)
40	100	2.80	3.10	100	0.020	0.061	21.2	Theoretical limit
40	51,790 (2470)	2.80	3.11	103	0.022	0.068	24.3	Simulation (25,000 missions)

The difficulty of the MLE task is only partly appreciated from the preceding study of the attainment of (or lack of) the three statistical properties—P1, P2, and P3. Figure 23 illustrates the objective function in the vicinity of θ_{ML} for the ideal detector for a particular simulation mission in which $\theta = (2.8, 3.3, 100 \text{ TeV})$ and there were 114,385 events from the energy range 20–5,500 TeV of which 4,819 were above the knee at 100 TeV, with equation (18) yielding $\theta_{ML} = (2.805, 3.319, 95.16 \text{ TeV})$. Figure 23 shows the objective function in the neighborhood of θ_{ML} , keeping α_1 fixed at 2.805 and letting α_2 and E_k vary in the region around θ_{ML} . Note the surface is well behaved and the minimum is easily found (and hence, θ_{ML}), and is representative of all the surfaces that were viewed when properties P1, P2, and P3 are approximately attained.

Objective Function Versus α_2 and E_k in Vicinity of θ_{ML} , With α_1 Fixed at its ML Estimate 2.805

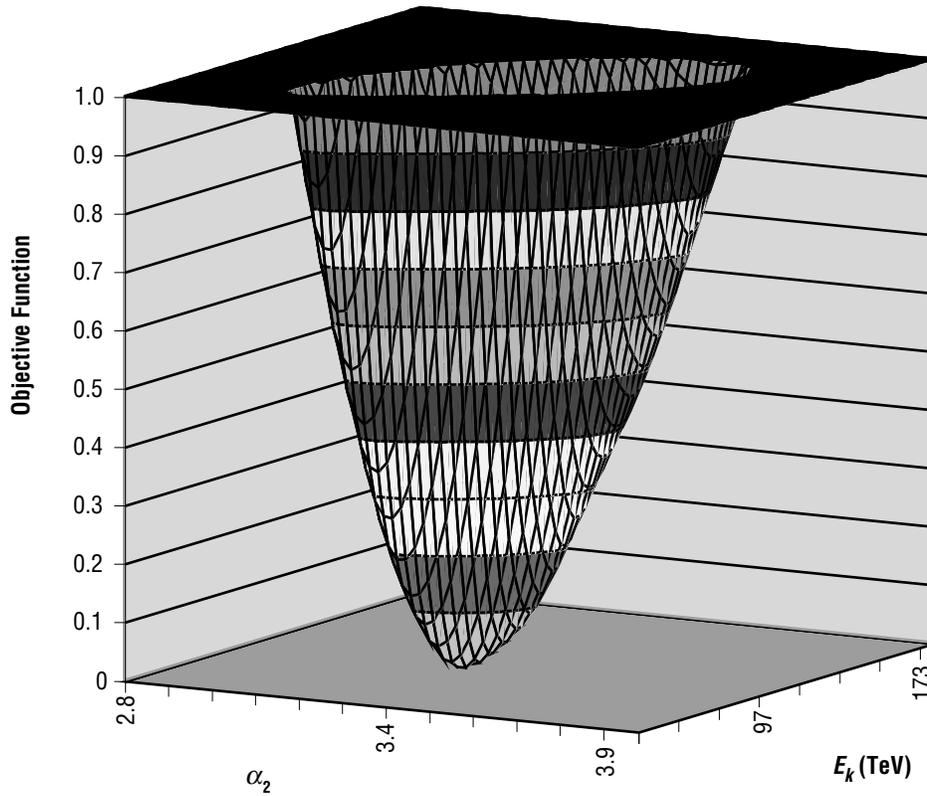


Figure 23. Objective function given by equation (18) for the ideal detector in the neighborhood of θ_{ML} for a simulated mission, keeping α_1 fixed at its ML estimate. There were 114,385 (4,819) events in this mission. The vertical axis has been scaled to 1.

The alignment or tilt of the surface in figure 23 is interesting and the contour plot in figure 24 illustrates the strong correlation between α_2 and E_k which is a direct consequence of the mathematical formulation of ϕ_B in equation (21) where the knee E_k acts as a “hinge” connecting the two distinct parts of the spectrum, and one can easily visualize a correlation between α_1 and E_k as well, whereas α_1 and α_2 appear to be only slightly correlated (not shown). The information matrix in equation (32) for the ideal detector can be used to approximate the correlations for this case and gives zero between α_1 and α_2 , 0.41 between α_1 and E_k , and 0.64 between α_2 and E_k .

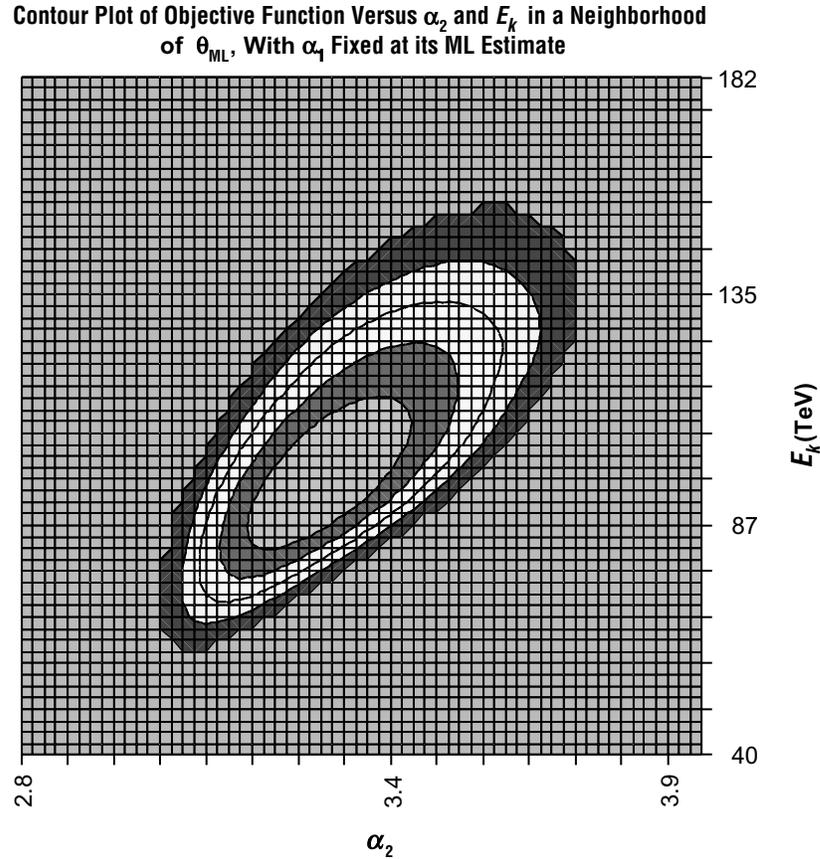


Figure 24. Contour plot of the objective function for ideal detector in the neighborhood of θ_{ML} for a simulated mission, keeping α_1 fixed at its ML estimate. There were 114,385 (4,819) events in this mission.

On the other hand, considering a much smaller ideal detector that detected 4,575 events of which 207 were above the knee on a particular simulated mission, equation (18) yields $\theta_{ML} = (2.856, 3.354, 156.57 \text{ TeV})$ and corresponds to the broader of the two minima shown in figure 25 which reveals a much more difficult estimation terrain in the neighborhood of θ_{ML} and in fact, suggests the possibility of multiple optimal solutions and confidence intervals for α_2 and E_k that would necessarily consist of disjoint subintervals!

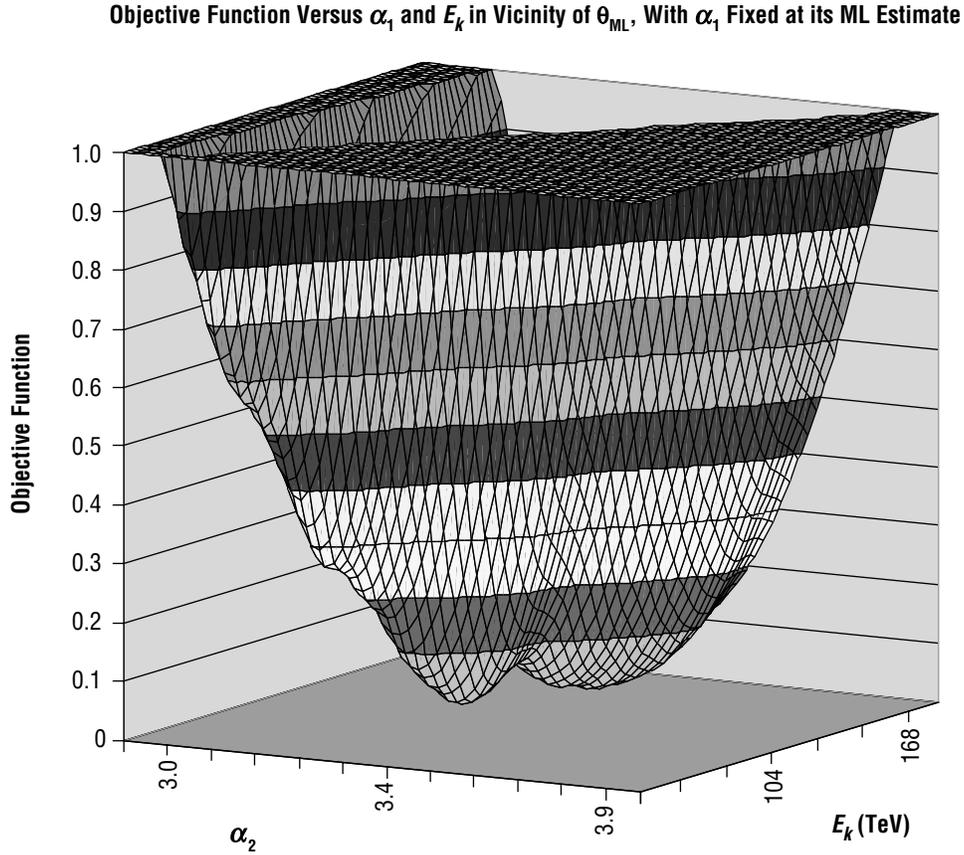


Figure 25. Objective function for ideal detector in the neighborhood of θ_{ML} for a simulated mission, keeping α_1 fixed at its ML estimate. This mission consisted of 4,575 (207) events. The vertical axis has been scaled to 1.

As observed in figures 13 and 15, when histograms of the ML estimates of α_1 and α_2 begin to overlap, it clearly signals the onset of difficulties in estimating the broken power law spectral parameters using MLE (and even more so for the method of moments since MLE was shown to be far superior in the simple power law section of this TP). Thus, an important scientific question is, “For what values of E_k will these distributions likely begin to overlap for a particular detector?”

If the concept TSC with its 40-percent Gaussian response function is considered and with an observing energy range of 20–5,500 TeV, then equation (20) provides the CRB for each of the three spectral parameters as a function of E_k . For example, for the case where $\theta = (2.8, 3.3, E_k)$ and calculating the CRB for $75 \leq E_k \leq 400$ TeV using equation (20), a 3σ curve describing the approximate width of the distribution (histogram) of the ML estimate of α_2 as a function of E_k can be constructed (3.3 minus three times the CRB of α_2) for values of E_k in the 75–400 TeV range. Sketching this curve versus E_k and noting where it begins to cross the line $\alpha_1 = 2.8$ suggests the value of E_k where the lower 3σ point of the distribution of the ML estimate of α_2 for this concept TSC will likely begin to overlap that of α_1 (fig. 26(a)). Also shown in figure 26(a) is the case when the resolution is set to 20-percent and also zero-percent (ideal detector), along with three additional dashed curves for the situation where the TSC’s collecting power is halved. Similar curves are provided in figure 26(b) for the case where $\alpha_2 = 3.1$, giving a spectral break-size of 0.3. Obviously these figures are no substitute for statistical hypothesis testing and furthermore do not consider the two errant estimation possibilities (EE1 and EE2) that also suggest a simple power law in favor of a broken power law, but nevertheless still provide important information.

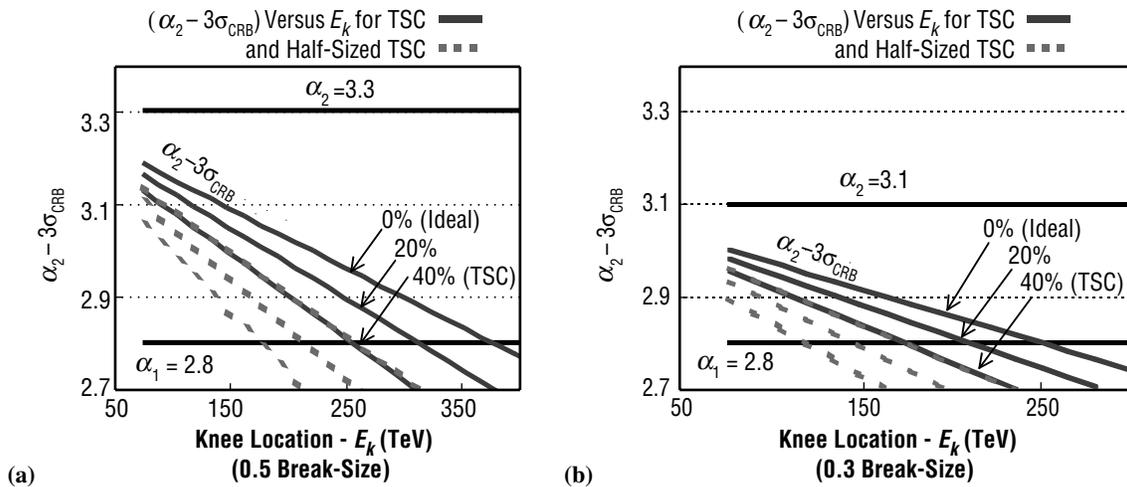


Figure 26. CRB used to estimate where the distribution of α_2 begins to cross that of α_1 when (a) the spectral break-size is 0.5 and (b) the spectral break-size is 0.3.

Note, too, that these figures represent a best-case scenario in that they are constructed using the CRB calculated from equation (20) which of course is not quite attained in practice, especially for the larger values of E_k and when the break-size is 0.3 as previously discussed. Thus, the actual overlap point would occur sooner (that is, for smaller values of E_k) because the variance of the ML estimator (or *any* other unbiased estimator of α_2) will be larger than the CRB.

Furthermore, figure 26 suggests that the 40-percent TSC is roughly equivalent to an ideal detector ($\rho=0$) of half its size in terms of measuring α_2 , while similar studies show this 40-percent TSC to be roughly equal to a 20-percent resolution detector of half its size in terms of estimating α_1 and E_k .² Consequently, instrument designers should consider first maximizing collecting power *and* then improving energy resolution, whenever possible.

It is important to realize that while raising E_2 to higher values in this analysis offers no benefit for this proposed TSC because of its previously stated collecting power, lowering E_1 does significantly benefit the measurement precision of α_1 and, because of their correlation but to a lesser degree, E_k . However, lowering E_1 offers no benefit in the measurement precision of α_2 when using the ideal detector and virtually none when using a real detector. These results are presented in table 4 for the case $\theta = (2.8, 3.3, 100 \text{ TeV})$ and in which E_1 is lowered incrementally from 20 to 1 TeV and once again illustrates the utility of the CRB determined by equation (20). Results for the case when the break-size is reduced to 0.3 are similar.

Table 4. Effect of lowering E_1 on the CRB for the TSC-sized detector with 0-, 20-, and 40-percent resolution Gaussian response function. The number of events N_2 above E_k is 2,255 for all values of E_1 .

E_1 (TeV)	N	CRB-Ideal Detector (0%)			CRB-20% Gaussian Detector			CRB-40% Gaussian Detector		
		α_1	α_2	E_k (TeV)	α_1	α_2	E_k (TeV)	α_1	α_2	E_k (TeV)
1	11,468,838	0.0005	0.0486	5.98	0.0007	0.0578	7.89	0.0009	0.0691	9.92
5	632,364	0.0024	0.0486	6.09	0.0028	0.0579	8.16	0.0034	0.0697	10.49
10	181,152	0.0050	0.0486	6.25	0.0058	0.0581	8.56	0.0073	0.0704	11.34
15	86,989	0.0081	0.0486	6.42	0.0095	0.0584	9.03	0.0126	0.0713	12.35
20	51,576	0.0117	0.0486	6.60	0.0144	0.0586	9.58	0.0199	0.0722	13.56

1.4 Analysis of Multiple Independent Data Sets

The ML theory required to estimate spectral parameters from an arbitrary number of data sets produced by science instruments having different observing ranges, different collecting powers, and different energy response functions is developed in this section. Application of this methodology will facilitate the interpretation of spectral information from existing data sets produced by astrophysics missions having different instrument characteristics and thereby permit the derivation of superior spectral information based on the combination of data sets. Furthermore, this procedure is of significant value to future astrophysics missions consisting of two or more detectors by allowing instrument developers to optimize each detector's design parameters through simulation studies in order to design and build complementary detectors that will maximize the precision with which the science objectives may be obtained.

This extension of the methodology developed in the previous sections to multiple data sets was motivated by such an application and is presented as an example in which two detectors, both assumed to have Gaussian response functions but different energy resolutions and observing ranges, were modeled separately and then in a collaborative effort to estimate the single parameter of a simple power law energy spectrum. A succinct comparison of the benefits from using the sets in concert is measured in terms of variance reduction of the estimator, as well as any biases resulting from poor statistics in one or both of the individual data sets that may be reduced when considered in combination.

The ML theory necessary for application to multiple astrophysics data sets is derived here for two independent data sets, A and B, produced by instruments having different (1) observing ranges, (2) collecting powers, (3) energy response functions, and (4) energy resolutions. These two data sets will be used together to estimate the energy spectra information and thereby benefit from the strengths of each detector, whereas, singly, they may be inadequate for achieving the science objectives. In practice, the data sets must be corrected for systematic errors in the energy response of the instruments in order to achieve the ultimate accuracy of the final spectra information based on the combination of astrophysics data sets. Generalization of this approach to more than two independent data sets then follows by induction.

To extend the ML theory to handle data sets A and B simultaneously, we begin with the probability density function for the data set of instrument responses $A = \{x_1, x_2, \dots, x_{N_A}\}$, given by

$$g_A(x_i; \boldsymbol{\theta}) = \int_{R_A} g_A(x_i | E, \rho_A) \phi(E; \boldsymbol{\theta}) dE, \quad i = 1, \dots, N_A, \quad (21)$$

so that the likelihood function is

$$L_A(\boldsymbol{\theta}) = \prod_{i=1}^{N_A} \left[\int_{R_A} g_A(x_i | E, \rho_A) \phi(E; \boldsymbol{\theta}) dE \right], \quad (22)$$

where $\boldsymbol{\theta}$ denotes the vector of spectral parameters of an arbitrary energy spectrum, $\phi(E; \boldsymbol{\theta})$, to be estimated; N_A is the number of detected events from observing range, R_A , of instrument A having response function, g_A , and energy resolution ρ_A , so that the corresponding objective function is

$$O_A(\boldsymbol{\theta}) = -\log [L_A(\boldsymbol{\theta})] \quad (23)$$

and the ML estimate $\boldsymbol{\theta}_A$, being that value of $\boldsymbol{\theta}$ for which $O_A(\boldsymbol{\theta})$ is a minimum, is obtained from equation (24) using the simplex search algorithm as

$$\boldsymbol{\theta}_A = \min_{\{\boldsymbol{\theta}\}} O_A(\boldsymbol{\theta}). \quad (24)$$

The likelihood function and objective function for data set B are similarly defined and because data sets A and B are assumed independent, the likelihood function for the two sets considered simultaneously is the product

$$L_{AB}(\boldsymbol{\theta}) = L_A(\boldsymbol{\theta})L_B(\boldsymbol{\theta}) , \quad (25)$$

so that upon taking the logarithm of each side gives the objective function as

$$O_{AB}(\boldsymbol{\theta}) = O_A(\boldsymbol{\theta}) + O_B(\boldsymbol{\theta}) \quad (26)$$

and the vector of ML estimates based on both data sets considered simultaneously is

$$\boldsymbol{\theta}_{AB} = \min_{\{\boldsymbol{\theta}\}} O_{AB}(\boldsymbol{\theta}) = \min_{\{\boldsymbol{\theta}\}} [O_A(\boldsymbol{\theta}) + O_B(\boldsymbol{\theta})] . \quad (27)$$

This procedure outlined above for two data sets is readily extended to k independent data sets so that any one of the $(2^k - 1)$ possible combinations of the data sets acting together provides the ML estimator $\boldsymbol{\theta}_S$ obtained by applying the Nelder-Mead algorithm to equation (28) as

$$\boldsymbol{\theta}_S = \min_{\{\boldsymbol{\theta}\}} \left\{ \sum_{j \in S} O_j(\boldsymbol{\theta}) \right\} , \quad (28)$$

where S can be any of the $(2^k - 1)$ nonvoid subsets of the set of integers $\{1, 2, \dots, k\}$.

The statistical properties of the derived ML estimator can then be studied for each simulated scenario and, in particular, those relating to (1) bias, (2) variance, and (3) normality can be rigorously investigated using graphical procedures and appropriate statistical techniques.^{*,12,13}

The CRB for the estimator of the vector of spectral parameters, $\boldsymbol{\theta}$, using two independent data sets, A and B, is derived in appendix B and shown to be the diagonal elements of the inverse of the covariance matrix, \mathbf{I} , whose I_{ij} elements are

$$I_{ij}(\boldsymbol{\theta}) = N_A \left\langle \frac{\partial \log [g_A(x; \boldsymbol{\theta})]}{\partial \theta_i} \times \frac{\partial \log [g_A(x; \boldsymbol{\theta})]}{\partial \theta_j} \right\rangle + N_B \left\langle \frac{\partial \log [g_B(y; \boldsymbol{\theta})]}{\partial \theta_i} \times \frac{\partial \log [g_B(y; \boldsymbol{\theta})]}{\partial \theta_j} \right\rangle \quad (29)$$

*Some statistical test procedures depend on the outcome of the normality test for estimators.

and where (1) the notation $\langle . \rangle$ denotes “expected value” and in practice can be accurately computed using Gaussian quadratures; (2) the partial derivatives are also numerically evaluated using equation (15); and (3) $g_A(x; \theta)$ is the probability density function for data set $A = \{x_i, i = 1, \dots, N_A\}$, given by equation (21) and similarly for data set $B = \{y_j, j = 1, \dots, N_B\}$. As noted in appendix B, the CRB given by equation (29) is readily extended to k independent data sets and provides a vital check on the performance of the derived ML estimation procedure. If the simulations show the ML estimator of the spectral information to be unbiased and also attains the CRB for a given spectrum-instrument combination, then this ML estimator will be the *best* (minimum variance) unbiased estimator possible from combining multiple data sets for that particular astrophysics mission scenario.

Furthermore, when this ML procedure is used in the design phase of an instrument and if the simulations show θ_{ML} is unbiased and attains the CRB for the science mission under consideration, then equation (29) can subsequently be used directly to evaluate the relative merits of various instrument design parameters (measured in terms of their impact on reducing the statistical error in measuring θ) without performing additional simulations. An example is given in section 1.6. This is of tremendous practical benefit because equation (29) can be evaluated in mere seconds while the equivalent information based on Monte Carlo simulations can take several days to obtain. Extending the ML procedure to multiple data sets, where in practice each may contain 10^5 to 10^6 events for a given science mission, is quite challenging. Practical studies conducted in Howell^{1,2} typically required $\geq 1,000$ simulated missions to obtain statistically meaningful inferences about a single detector design parameter under study. Such a simulation run would last ≥ 12 h when the broken power law spectrum was assumed (considerably less for the simple power law), largely because of the vast number of numerical integrations required in evaluating the objective functions and the many steps taken by the Nelder-Mead search for θ_{ML} .

1.5 Example of Estimating α_1 of a Simple Power Law Using Two Data Sets

The benefits of having the additional nuclei ($Z > 8$) from a proposed thin sampling calorimeter (TSC) for ACCESS at energies above the transition radiation detector (TRD) energy range was recently investigated¹⁴ using this methodology and serves to illustrate the application and utility of the approach. Specifically, the value of having these calorimeter-provided nuclei in addition to the expected number of TRD events was investigated, and the approach was to measure the benefit of including these additional events in the estimation of the single parameter, α_1 , of an assumed simple power law as compared to *not* including them.

In this example the TRD geometry factor is assumed to be $7.58 \text{ m}^2\text{-sr}$ with energy response saturating at 20 TeV/n and for an observing period of 1,000 d. Furthermore, it was decided to consider the species Ne-S with an assumed differential spectral index of 2.39, along with a calorimeter geometry factor of $5 \text{ m}^2\text{-sr}$, providing 593 events $> 20 \text{ TeV/n}$ on average for the calorimeter for a 1,000-d observing period.

Additionally, a Gaussian response function was assumed for both detectors, and the TRD was assumed to have the same linear mean response as the calorimeter but with a constant 35-percent energy resolution over its observing range E_1 to 20 TeV/n, while the TSC was assumed to have a constant resolution of 40-percent over its nonoverlapping observing range 20 to 2,000 TeV/n. This upper limit of 2,000 TeV was chosen for the calorimeter because the number of detected events $> 2,000 \text{ TeV}$ is negligible for a detector of this collecting power. Then, E_1 is taken to be different values ranging from 0.5 to 7 TeV as part

of a parametric study in this illustration. Of course, the upper energy boundary of 20 TeV/n could be extended to a higher value if the saturation limit of the TRD could be improved.

In the simulation, GCR event energies, E_i , are simulated from the TRD observing range, $R_{\text{TRD}} = [E_1, 20 \text{ TeV}]$, and where the number of these events, N_{TRD} , is a function of the observing range, geometry factor, and observing time. Then, for each of these simulated incident energies, E_i , a response, x_i , is simulated according to the assumed Gaussian response function and with 35-percent energy resolution, ρ_{TRD} . The calorimeter events are simulated in a similar manner using its defining parameters, N_{Cal} , R_{Cal} , and ρ_{Cal} , and the Gaussian response function.

Performing the simulation once defines a so-called astrophysics “mission” and provides a single ML estimate of α_1 for each instrument by solving equation (24) with $\theta = \alpha_1$ and all instrument characteristics appropriately modeled through the formulation of the detector response function, g , for each detector, and then solving equation (27) when the two data sets are used in combination in the ML estimation procedure. Notationally, α_{Cal} and α_{TRD} are the ML estimates of α_1 when the calorimeter and TRD are considered as stand-alone instruments, respectively, and α_{Both} is the ML estimate of α_1 when the two instruments are used in combination. The simulation is repeated for many missions, each time generating a new set of incident energies from the assumed simple power law spectrum for each instrument and their respective simulated energy deposits according to the detectors’ response functions, estimating α_1 using each data set separately and then in tandem. The statistical behavior of these estimates relative to plausible design variation of response function parameters is then studied.

Figure 27a shows the ML estimates α_{TRD} , α_{Cal} , and α_{Both} of α_1 for 25 simulated missions in which E_1 was first set to 5 TeV/n, thus providing 5,275 events on average for the TRD. Figure 27b shows the effect of lowering E_1 to 3 TeV/n, providing an average 11,660 events for the TRD, and then lowering E_1 to 1 TeV/n, providing an average of 56,920 events for the TRD for 25 missions (fig. 27c). We observe that as E_1 is lowered, thereby increasing the number of TRD events as well as extending its effective observing range, and hence “lever arm” effect in the estimation of α_1 , the calorimeter has an ever-diminishing role in its contribution to α_{Both} as illustrated by the α_{TRD} and α_{Both} estimates nearly coinciding in figure 27c.

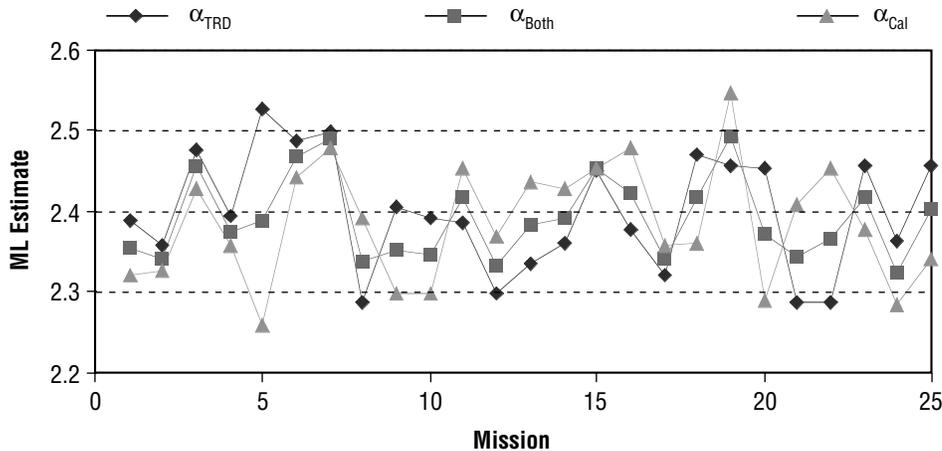


Figure 27a. ML estimates of α_1 for 25 missions, with $E_1 = 5 \text{ TeV}$ (5,275 events for TRD and 593 for calorimeter).

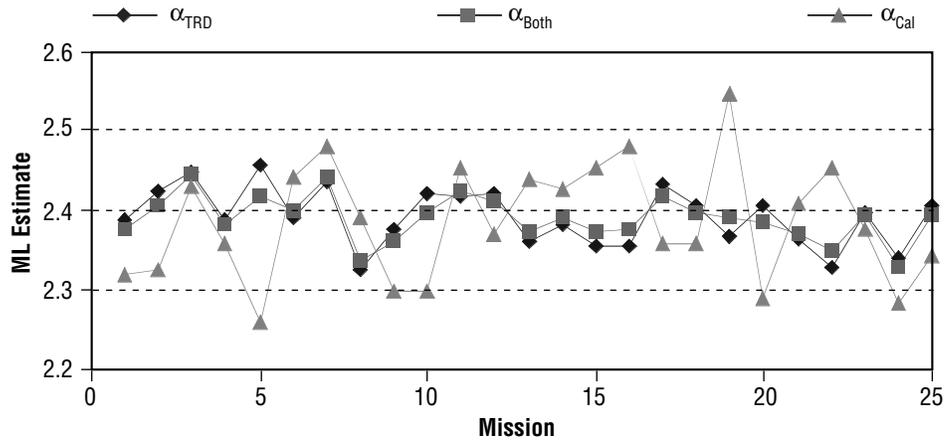


Figure 27b. ML estimates of α_1 for 25 missions, with $E_1 = 3$ TeV (11,660 events for TRD and 593 for calorimeter).

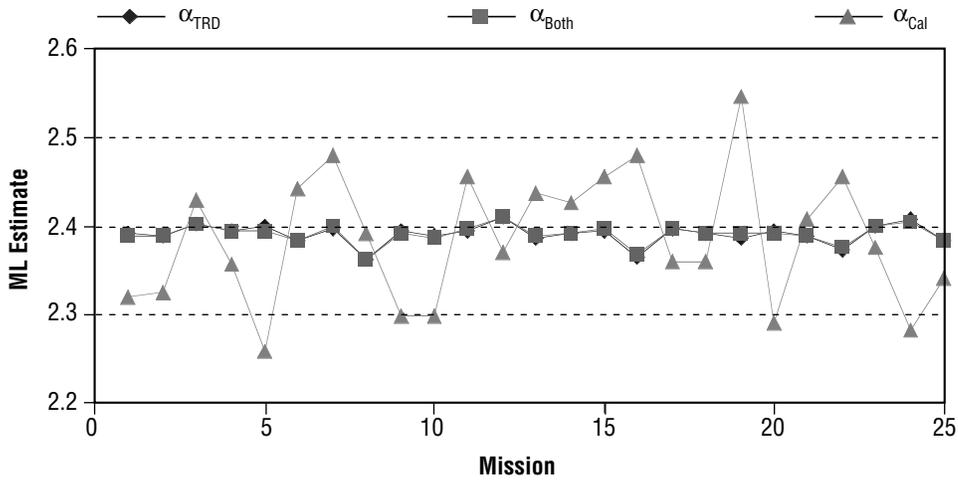


Figure 27c. ML estimates of α_1 for 25 missions, with $E_1 = 1$ TeV (56,920 events for TRD and 593 for calorimeter). Note that α_{TRD} and α_{Both} are virtually indistinguishable.

The simulation was run for 1,000 missions for each of several values of the E_1 parameter with summary statistics provided in table 5 for the calorimeter acting alone, table 6 for the TRD acting alone, and table 7 for the two data sets acting in concert to estimate α_1 . Tables 5 and 6 show that the calorimeter and TRD as stand-alone instruments each provide unbiased spectral information (recall $\alpha_1 = 2.39$ for this study) and attainment of the CRB.

Table 5. Summary statistics of α_{Cal} based on 1,000 simulated missions of the calorimeter.

E_1 (TeV)	E_2 (TeV)	Calorimeter Events $N(E > E_1)$	Mean	Standard Deviation σ_{Cal}	CRB
20	2,000	593	2.39	0.066 ^a	0.067 ^a

^aThe appearance that σ_{Cal} is less than the CRB is an artifact of the finite number of missions in the simulations.

Table 6. Summary statistics of α_{TRD} based on 1,000 simulated missions for each value of E_1 of the TRD alone (all events between E_1 and 20 TeV).

E_1 (TeV)	Saturation (TeV)	TRD Events	Mean	Standard Deviation σ_{TRD}	CRB
7.0	20	2,968	2.39	0.124	0.123
5.2	20	4,947	2.39	0.068	0.068
5.0	20	5,274	2.39	0.063	0.063
3.0	20	11,658	2.39	0.031	0.031
1.0	20	56,919	2.39	0.011	0.011
0.5	20	150,630	2.39	0.007	0.007

Table 7. Summary statistics of α_{Both} based on 1,000 simulated missions for each value of E_1 of the TRD and calorimeter acting in combination.

CRB for TRD and Calorimeter (Includes 593 Events >20 TeV)							
TRD		Calorimeter		Total Combined Events	Mean	Standard Deviation σ_{Both}	CRB
E_1 (TeV)	Saturation (TeV)	E_1 (TeV)	E_2 (TeV)				
7.0	20	20	2,000	3,560	2.39	0.058 ^a	0.059
5.2	20	20	2,000	5,539	2.39	0.048	0.048
5.0	20	20	2,000	5,866	2.39	0.045 ^a	0.046
3.0	20	20	2,000	12,249	2.39	0.027 ^a	0.028
1.0	20	20	2,000	57,511	2.39	0.011	0.011
0.5	20	20	2,000	151,220	2.39	0.007	0.007

^a The appearance that σ_{Cal} is slightly less than the CRB is an artifact of the finite number of missions in the simulations.

Table 7 shows that for all scenarios considered, the two data sets acting together likewise provide an unbiased estimate of the spectral parameter, α_1 , and attainment of the CRB computed from equation (29) with $\boldsymbol{\theta} = \alpha_1$ to give

$$\text{CRB}_{\text{Both}} = \frac{1}{N_{\text{TRD}} \left\langle \left(\frac{\partial \log [g_{\text{TRD}}(x; \alpha_1)]}{\partial \alpha_1} \right)^2 \right\rangle + N_{\text{Cal}} \left\langle \left(\frac{\partial \log [g_{\text{Cal}}(y; \alpha_1)]}{\partial \alpha_1} \right)^2 \right\rangle} \quad (30)$$

as the CRB for the TRD and calorimeter acting in combination. The CRB in table 5 for the calorimeter acting alone is obtained from equation (14) and likewise in table 6 for the TRD acting alone for the various values of E_1 . Numerical evaluation of equation (30) for the different scenarios of interest presented in table 7 indicates the two detectors acting in concert do indeed attain the CRB for all values of E_1 and achievement of this bound, coupled with the fact that the derived ML procedure provides an unbiased estimate of the spectral parameter, α_1 , implies that this procedure is the best (minimum variance) unbiased estimation technique possible.

Additionally, histograms of the 1,000-mission ML estimates for each instrument considered separately and also in tandem show the distribution of the ML estimator to be approximately normally distributed.

Comparing standard deviations of the different scenarios in tables 5–7 clearly shows the importance of not only the number of events or so-called “statistics” but also the observing range of the instruments; i.e., the “lever arm effect.” The very special case where $E_1 = 5.2$ TeV, found by trial and error, shows the calorimeter acting alone with only 593 events but with a much greater observing range (20 TeV/n to an average maximum event energy of 1,100 TeV/n) is effectively as good as the TRD in terms of estimating α_1 ($\sigma_{\text{Cal}} = 0.067$, and $\sigma_{\text{TRD}} = 0.068$ when $E_1 = 5.2$ TeV). For this case where $E_1 = 5.2$ TeV, the TRD had an average of 4,947 events but with a much shorter observing range. Furthermore, the TRD’s energy resolution was assumed to be 35 percent, whereas the calorimeter was assumed to be worse with a constant 40-percent energy resolution. We also note the fairly intuitive result in table 7 that $\sigma_{\text{Both}} = 0.048$ which is $\approx 0.068/\sqrt{2}$ when the calorimeter and TRD are effectively equal to each other and so σ_{Both} scales roughly by $\sqrt{2}$ for this case.

Nevertheless, as E_1 is successively lowered, we note the overpowering impact of the TRD’s increasing number of events and growing observing range, and in fact, for $E_1 \leq 1$ TeV, the calorimeter’s contribution to the variance reduction of the ML estimate of α_{Both} is virtually nil. Thus, we see that “how low” the TRD can observe plays a major role in assessing the value of these additional calorimeter events in the estimation of the spectral parameter, α_1 . It was concluded that since the proposed TRD would easily observe to ≤ 1 TeV/n, the contribution of the calorimeter’s additional events would be insignificant for measuring the spectra for $Z > 8$. Thus, this study resulted in an important design consideration for the charge particle identification detector for the calorimeter and underscores the utility of this ML theory for multiple independent data sets in the design phase of new science instruments.

1.6 Using the CRB to Explore Instrument Design Parameters When Estimating $\theta = (\alpha_1, \alpha_2, E_k)$ of a Broken Power Law Using Three Data Sets

The benefit of using three independent data sets in concert to measure the spectral parameters of an assumed broken power law proton spectrum will be investigated, and, for specificity, assume the spectral knee location is $E_k = 175$ TeV and the slope parameters α_1/α_2 below/above the knee are 2.8 and 3.2, respectively, to give a 0.4 spectral break-size.

Next, assume the three hypothetical data sets A, B, and C were collected by instruments having observing ranges, collecting powers, and stochastic response functions defined in table 8 so that data set A provides information regarding only the spectral parameter α_1 ; B about α_1, α_2 , and E_k ; and C, only about α_2 . It is further assumed that each data set was produced by a detector having the same linear mean response as the TSC previously introduced in section 1.1 so that their response functions may be visually compared in figure 28.

Table 8. Data sets and associated response functions, with CRB for all possible combinations of data sets.

Data Set Combinations and Instrument Response Function	Observing Range (TeV)	Average Number of Events	Cramer-Rao Bound for		
			α_1	α_2	E_k (TeV)
A (Gaussian)	[1, 20]	150,000	0.0095	–	–
B (Gamma)	[35, 5,000]	44,000 (2,000)*	0.0225	0.0780	34.24
C (Broken Gaussian)	[500, 10,000]	500	–	0.1050	–
AB			0.0087	0.0744	26.13
BC			0.0220	0.0626	30.23
ABC			0.0087	0.0607	23.05

*44,000 events above 35 TeV of which 2,000 are above $E_k=175$ TeV.

In this scenario, data set A was produced by an instrument having a Gaussian response function and with constant 40-percent energy resolution; B by an instrument having a gamma response function and thus capable of describing a wide variety of shapes with right-hand skewness (outer curve from the right in fig. 28); and C was measured by a detector having a “broken Gaussian” response function consisting of two blended normal distributions (middle curve from right in fig. 28) suggested by J. Ormes, 2000, private communication, for its closeness to the Gaussian response function but with a tail region as desired. These latter two response functions were introduced to address the possibility that some detector response functions may exhibit a right-hand skewness and might add to the difficulty of the ML estimation task. Note that while the gamma response function used here also has a constant energy resolution of 40 percent, the broken Gaussian actually has a 41-percent resolution because of the added skewness while keeping the rest of the distribution matching the Gaussian.

Furthermore, in the construction of these hypothetical data sets, the number of events in each data set was chosen so that each data set alone provides unbiased spectra information and approximate attainment of the CRB,^{1,2} as discussed in sections 1.1 and 1.2, so that any combination of the data sets

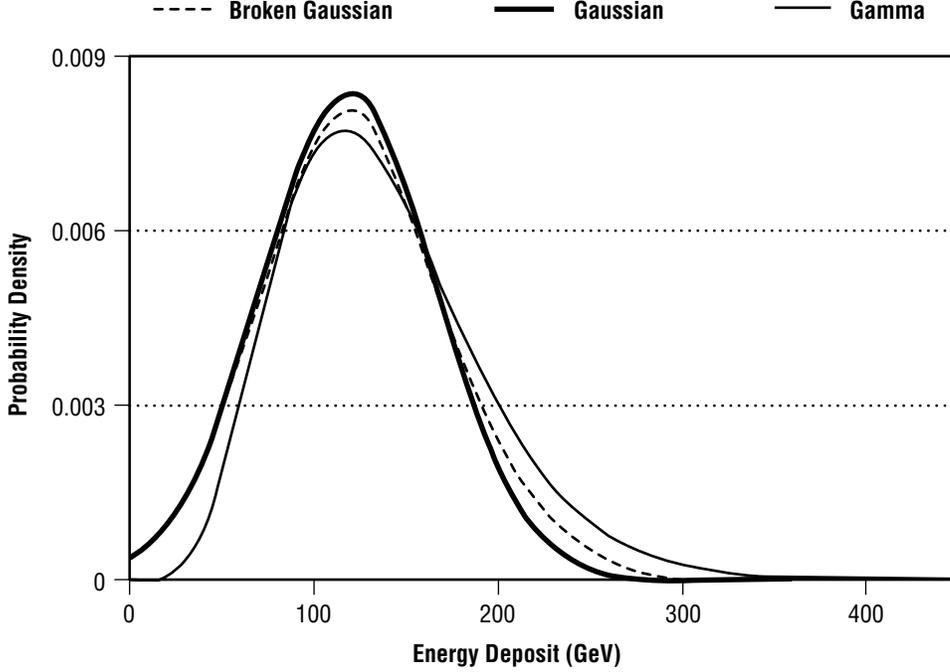


Figure 28. Gamma, Gaussian, and broken Gaussian detector response functions to 40-TeV proton.

should also attain the CRB. Consequently, the CRB can be calculated from equation (56) for the combination of data sets and used directly to investigate the relative merits of instrument design parameters without performing the simulations. Note that while A and C provide no information about the knee location E_k , they do provide an improvement in the measurement precision of E_k when they work together with data set B.

To illustrate the utility of the CRB for multiple independent data sets regarding design parameter studies, suppose the collecting power of instrument C is increased by a factor of 10 with the results given in table 9 which shows the merits of collecting power on the measurement precision of the spectral parameters for the combinations of data sets C, BC, and ABC. While C only contains information about α_2 , the measurement precision (standard deviation) of E_k reduces from 30.23 to 23.6 TeV for the BC combination and from 23.05 to 17.47 TeV for the ABC combination.

Table 9. Number of events from table 8 in data set C is increased by a factor of 10.

Instrument Combination	Observing Range (TeV)	Average Number of Events	Cramer-Rao Bound for		
			α_1	α_2	E_k (TeV)
A (Gaussian)	[1, 20]	150,000	0.0095	–	–
B (Gamma)	[35, 5,000]	44,000 (2,000)*	0.0225	0.0780	34.24
C (Broken Gaussian)	[500, 10,000]	5,000	–	0.0332	–
AB			0.0087	0.0744	26.13
BC			0.0214	0.0306	23.60
ABC			0.0086	0.0303	17.47

*44,000 events above 35 TeV of which 2,000 are above $E_k=175$ TeV.

Next, if the number of events in data set B is reduced by a factor of 2 and those in C increased by a factor of 10, we can once again explore the impact of various collecting power options shown in table 10. Finally, if we keep the collecting powers the same (B reduced by a factor of 2, C increased by a factor of 10), and additionally improve the resolution of the instruments for A and C to 30- and 35-percent resolution, respectively, we see the impact of detector energy resolution on measurement precision of the spectral parameters, as shown in table 11.

Table 10. Number of events from table 8 in data set B is reduced by a factor of 2 and those in C increased by a factor of 10.

Instrument Combination	Observing Range (TeV)	Average Number of Events	Cramer-Rao Bound for		
			α_1	α_2	E_k (TeV)
A (Gaussian)	[1, 20]	150,000	0.0095	–	–
B (Gamma)	[35, 5,000]	22,000 (1,000)*	0.0318	0.1104	48.42
C (Broken Gaussian)	[500, 10,000]	5,000	–	0.0332	–
AB			0.0091	0.1048	35.86
BC			0.0302	0.0318	31.81
ABC			0.0090	0.0317	22.36

*22,000 events above 35 TeV of which 2,000 are above $E_k=175$ TeV.

Table 11. Number of events from table 8 in data set B is reduced by a factor of 2 and those in C increased by a factor of 10, resolution of A improved to 30 percent and C to 35 percent.

Instrument Combination	Observing Range (TeV)	Average Number of Events	Cramer-Rao Bound for		
			α_1	α_2	E_k (TeV)
A (Gaussian, 30%)	[1, 20]	150,000	0.0083	–	–
B (Gamma, 40%)	[35, 5,000]	22,000 (1,000)*	0.0318	0.1104	48.42
C (Broken Gaussian, 35%)	[500, 10,000]	5,000	–	0.0321	–
AB			0.0080	0.1047	35.58
BC			0.0301	0.0309	31.70
ABC			0.0080	0.0307	22.02

*22,000 events above 35 TeV of which 2,000 are above $E_k=175$ TeV.

Obviously, the number of possible parametric studies are numerous, but the preceding sample investigation illustrates the utility of this procedure and the derived CRB for multiple independent data sets when considering the design of new, complementary detectors. Furthermore, the CRB values in tables 8–11 were computed using equation (56) in <1 min, whereas the equivalent information based on Monte Carlo simulations would take >1 wk to obtain. However, when the CRB is not attained in practice, simulations must be used to estimate the real performance benefits when using multiple independent data sets.

2. CONCLUSIONS

This TP investigates the statistical properties of the ML estimator of the single parameter of a simple power law energy spectrum and presents the conditions under which this estimator attains the three desirable properties: (P1) consistency, (P2) efficiency, and (P3) asymptotic normality. A comparison with the estimation procedure called the method of moments is also included and shows the ML estimator to be far superior in terms of these three desirable properties.

The properties of the ML estimator of the broken power law energy spectral parameters and the conditions under which P1, P2, and P3 are attained are also investigated under a wide range of parametric values. A crucial result of this research and necessity for investigating P2 is the derivation of a closed-form expression for the CRB of the broken power law distribution presented in appendix A. Another critical result is the calculation of the CRB using equation (20) for the broken power law distribution and equation (14) for the simple power law distribution when events are measured by a real detector having response function, g , and energy resolution, ρ . While this study considered an instrument having a Gaussian response function and with resolutions ranging from zero to 50 percent, any response function, g , can be used in equations (14) and (20) to calculate the CRB, such as various skewed distributions and others having nonconstant energy resolution as illustrated in section 1.6 and also discussed in references 1 and 2. Much insight into the estimation task of power law spectra information can be gleaned from the CRB as illustrated in this TP, as well as the fact that it provides a stopping rule in the search for the best (minimum variance) unbiased estimator of power law spectra information.

Simulations were conducted in parallel with these analytical results and the CRB played an unexpected but valuable oversight role by signaling errors in the simulation code when the simulations occasionally produced the impossible result of the ML estimator having a standard deviation smaller than the CRB. Furthermore, it is likely that these simulation errors, while small in practical terms, would have gone unnoticed if not for having the CRB available for comparison.

Additionally, several detector design parameter studies are included in this research and it is hoped that those designing instruments to measure power law spectra information will benefit from these studies. Additionally, this analysis should benefit those wishing to apply these techniques to the estimation of spectra information from existing data sets, which requires a modified likelihood function to handle the realistic situation in which the range of integration $[E_1, E_2]$ in the objective function is unknown, and is discussed in detail with examples in references 1 and 2.

The MLE procedure and companion analytical techniques to estimate spectra information from an arbitrary number of astrophysics data sets produced by vastly different science instruments is presented and demonstrates how complementary astrophysics missions can work in concert to achieve science goals. Additionally, the CRB for an unbiased estimator of spectra information based on these multiple independent data sets is derived in appendix B and provides a means of assessing the accuracy of different estimation techniques and, furthermore, provides a stopping rule in the search for estimation methodologies when the CRB is attained in practice. Several examples illustrating this ML method and utility of the CRB for multiple independent data sets are included.

APPENDIX A — CLOSED-FORM EXPRESSION FOR THE CRAMER-RAO BOUND OF THE BROKEN POWER LAW

A closed-form expression for the information matrix $\mathbf{I}(\boldsymbol{\theta})$ of the broken power law distribution is derived, and the inverse $\mathbf{I}^{-1}(\boldsymbol{\theta})$ is defined as the CRB. This bound corresponds to the so-called ideal cosmic-ray detector having perfect energy resolution and has tremendous utility because it sets the limit on the precision with which any conceivable detector of equal collecting power can measure the three broken power law spectral parameters. Furthermore, it provides a means to tune the integration and differentiation parameters in the numerical algorithm for evaluating equation (20) for real detectors because the CRB determined from equation (42) is exact and must equal equation (20) when the detector's energy resolution $\rho \rightarrow 0$.

A.1 Derivation

The broken power law probability density function for GCR event energy, E , is given by

$$\phi_B(E; \boldsymbol{\theta}) = \begin{cases} A(\boldsymbol{\theta}) \left(\frac{E}{E_k} \right)^{-\alpha_1} & \text{for } E_1 \leq E < E_k \\ A(\boldsymbol{\theta}) \left(\frac{E}{E_k} \right)^{-\alpha_2} & \text{for } E_k \leq E \leq E_2 \end{cases} \quad (31)$$

over an energy range $[E_1, E_2]$ that does not depend on the spectral parameters $\boldsymbol{\theta} = (\alpha_1, \alpha_2, E_k)$, and the normalizing coefficient $A(\boldsymbol{\theta})$ is

$$A(\boldsymbol{\theta}) = \frac{(\alpha_1 - 1)(\alpha_2 - 1)}{E_k \left[\alpha_1 - \alpha_2 + (\alpha_2 - 1) \left(\frac{E_1}{E_k} \right)^{1-\alpha_1} - (\alpha_1 - 1) \left(\frac{E_2}{E_k} \right)^{1-\alpha_2} \right]}. \quad (32)$$

Starting with

$$\begin{aligned} 1 &= \int_{E_1}^{E_2} \phi_B(E; \boldsymbol{\theta}) dE \\ &= \int_{E_1}^{E_k} e^{\log[\phi_B(E; \boldsymbol{\theta})]} dE + \int_{E_k}^{E_2} e^{\log[\phi_B(E; \boldsymbol{\theta})]} dE \end{aligned} \quad (33)$$

and writing $\log[\phi_B(E; \boldsymbol{\theta})]$ as

$$\log[\phi_B(E; \boldsymbol{\theta})] = \log[A(\boldsymbol{\theta})] - \alpha_1 \log\left[\frac{E}{E_k}\right] \delta_1 - \alpha_2 \log\left[\frac{E}{E_k}\right] \delta_2, \quad (34)$$

where δ_1 and δ_2 are indicator functions defined in equation (37), gives

$$\begin{aligned} 1 &= \int_{E_1}^{E_k} \exp\left[\log[A(\boldsymbol{\theta})] - \alpha_1 \log\left[\frac{E}{E_k}\right] \delta_1 - \alpha_2 \log\left[\frac{E}{E_k}\right] \delta_2\right] dE \\ &+ \int_{E_k}^{E_2} \exp\left[\log[A(\boldsymbol{\theta})] - \alpha_1 \log\left[\frac{E}{E_k}\right] \delta_1 - \alpha_2 \log\left[\frac{E}{E_k}\right] \delta_2\right] dE. \end{aligned} \quad (35)$$

Taking the derivative of equation (35) with respect to α_1 and α_2 is straightforward and gives

$$\begin{cases} 0 = \left\langle \frac{\partial \log[\phi_B(E; \boldsymbol{\theta})]}{\partial \alpha_1} \right\rangle = \langle A_1' - L_E \delta_1 \rangle \\ 0 = \left\langle \frac{\partial \log[\phi_B(E; \boldsymbol{\theta})]}{\partial \alpha_2} \right\rangle = \langle A_2' - L_E \delta_2 \rangle \end{cases}, \quad (36)$$

where $\langle . \rangle$ denotes mathematical expectation and the abbreviations

$$A_1' \equiv \frac{\partial \log[A(\boldsymbol{\theta})]}{\partial \alpha_1}, A_2' \equiv \frac{\partial \log[A(\boldsymbol{\theta})]}{\partial \alpha_2}, L_E \equiv \log\left[\frac{E}{E_k}\right], \delta_1 \equiv \begin{cases} 1, & E < E_k \\ 0, & E \geq E_k \end{cases}, \delta_2 \equiv \begin{cases} 0, & E < E_k \\ 1, & E \geq E_k \end{cases} \quad (37)$$

are introduced. Taking the derivative of equation (35) with respect to E_k and using Leibnitz's rule (see Professor Bierens' "check" of equation (39) in section A.2):

$$\frac{d}{da} \int_p^q f(x, a) dx = \int_p^q \frac{\partial}{\partial a} f(x, a) dx + f(q, a) \frac{dq}{da} - f(p, a) \frac{dp}{da} \quad (38)$$

and the continuity of ϕ_B at E_k gives

$$\begin{aligned}
0 &= \int_{E_1}^{E_k} \frac{\partial \log[\phi_B(E; \boldsymbol{\theta})]}{\partial E_k} \phi_B(E; \boldsymbol{\theta}) dE + \phi_B(E_k; \boldsymbol{\theta}) - 0 + \int_{E_k}^{E_2} \frac{\partial \log[\phi_B(E; \boldsymbol{\theta})]}{\partial E_k} \phi_B(E; \boldsymbol{\theta}) dE + 0 - \phi_B(E_k; \boldsymbol{\theta}) \\
&= \left\langle \frac{\partial \log[\phi_B(E; \boldsymbol{\theta})]}{\partial E_k} \right\rangle = \left\langle A'_k + \frac{\alpha_1}{E_k} \delta_1 + \frac{\alpha_2}{E_k} \delta_2 \right\rangle \text{ where } A'_k \equiv \frac{\partial \log[A(\boldsymbol{\theta})]}{\partial E_k}
\end{aligned} \tag{39}$$

and the three partial derivatives of $\log[A(\boldsymbol{\theta})]$ defined in equations (37) and (39) are

$$\left\{ \begin{aligned}
A'_1 &= \frac{(-1 + \alpha_2) \left(\frac{E_2}{E_k} \right)^{\alpha_2} \left(E_1 - \left(\frac{E_1}{E_k} \right)^{\alpha_1} E_k + (-1 + \alpha_1) E_1 \log \left[\frac{E_1}{E_k} \right] \right)}{(-1 + \alpha_1) \left((-1 + \alpha_2) E_1 \left(\frac{E_2}{E_k} \right)^{\alpha_2} + \left(\frac{E_1}{E_k} \right)^{\alpha_1} \left(E_2 - \alpha_1 E_2 + (\alpha_1 - \alpha_2) \left(\frac{E_2}{E_k} \right)^{\alpha_2} E_k \right) \right)} \\
A'_2 &= \frac{(-1 + \alpha_1) \left(\frac{E_1}{E_k} \right)^{\alpha_1} \left(-E_2 + \left(\frac{E_2}{E_k} \right)^{\alpha_2} E_k + (E_2 - \alpha_2 E_2) \log \left[\frac{E_2}{E_k} \right] \right)}{(-1 + \alpha_2) \left((-1 + \alpha_2) E_1 \left(\frac{E_2}{E_k} \right)^{\alpha_2} + \left(\frac{E_1}{E_k} \right)^{\alpha_1} \left(E_2 - \alpha_1 E_2 + (\alpha_1 - \alpha_2) \left(\frac{E_2}{E_k} \right)^{\alpha_2} E_k \right) \right)} \\
A'_k &= \frac{-\alpha_1 (-1 + \alpha_2) E_1 \left(\frac{E_2}{E_k} \right)^{\alpha_2} + \left(\frac{E_1}{E_k} \right)^{\alpha_1} \left((-1 + \alpha_1) \alpha_2 E_2 + (-\alpha_1 + \alpha_2) \left(\frac{E_2}{E_k} \right)^{\alpha_2} E_k \right)}{(-1 + \alpha_2) E_1 \left(\frac{E_2}{E_k} \right)^{\alpha_2} E_k + \left(\frac{E_1}{E_k} \right)^{\alpha_1} E_k \left(E_2 - \alpha_1 E_2 + (\alpha_1 - \alpha_2) \left(\frac{E_2}{E_k} \right)^{\alpha_2} E_k \right)}
\end{aligned} \right. \tag{40}$$

As a check, the expectation of the partial derivatives of ϕ_B with respect to the parameters α_1 , α_2 , and E_k defined in equations (36) and (39) are indeed seen to be zero when the derivatives A'_1 , A'_2 , and A'_k in equation (40) are used in conjunction with the expected value terms defined in equation (41), along with $\langle \delta_1 \rangle = \Pr\{E < E_k\}$ and $\langle \delta_2 \rangle = 1 - \Pr\{E < E_k\}$, and

$$\begin{cases} \langle L_E \delta_1 \rangle = -\frac{A(\boldsymbol{\theta})}{(-1+\alpha_1)^2} \left(\frac{E_1}{E_k}\right)^{-\alpha_1} \left(-E_1 + \left(\frac{E_1}{E_k}\right)^{\alpha_1} E_k + E_1 \log \left[\frac{E_1}{E_k} \right] - \alpha_1 E_1 \log \left[\frac{E_1}{E_k} \right] \right) \\ \langle L_E \delta_2 \rangle = -\frac{A(\boldsymbol{\theta})}{(-1+\alpha_2)^2} \left(\frac{E_2}{E_k}\right)^{-\alpha_2} \left(-E_2 + \left(\frac{E_2}{E_k}\right)^{\alpha_2} E_k + E_2 \log \left[\frac{E_2}{E_k} \right] - \alpha_2 E_2 \log \left[\frac{E_2}{E_k} \right] \right) \end{cases} \quad (41)$$

Next, the nine elements of the symmetric information matrix $\mathbf{I}(\boldsymbol{\theta})$ are constructed as follows:

$$\begin{cases} I_{11} = \left\langle \left(A_1' - L_E \delta_1 \right)^2 \right\rangle = \left\langle A_1'^2 - 2L_E A_1' \delta_1 + L_E^2 \delta_1^2 \right\rangle \\ I_{22} = \left\langle A_2'^2 - 2L_E A_2' \delta_2 + L_E^2 \delta_2^2 \right\rangle \\ I_{33} = \left\langle A_k'^2 + \left(\frac{\alpha_1}{E_k}\right)^2 \delta_1^2 + \left(\frac{\alpha_2}{E_k}\right)^2 \delta_2^2 + 2A_k' \left(\frac{\alpha_1}{E_k}\right) \delta_1 + 2A_k' \left(\frac{\alpha_2}{E_k}\right) \delta_2 + 2\frac{\alpha_1 \alpha_2}{E_k^2} \delta_1 \delta_2 \right\rangle \\ I_{12} = I_{21} = \left\langle A_1' A_2' + L_E^2 \delta_1 \delta_2 - L_E A_2' \delta_1 - L_E A_1' \delta_2 \right\rangle \\ I_{13} = I_{31} = \left\langle A_1' A_k' + A_1' \left(\frac{\alpha_1}{E_k}\right) \delta_1 + A_1' \left(\frac{\alpha_2}{E_k}\right) \delta_2 - L_E A_k' \delta_1 - L_E \left(\frac{\alpha_1}{E_k}\right) \delta_1 - L_E \left(\frac{\alpha_2}{E_k}\right) \delta_1 \delta_2 \right\rangle \\ I_{23} = I_{32} = \left\langle A_2' A_k' + A_2' \left(\frac{\alpha_1}{E_k}\right) \delta_1 + A_2' \left(\frac{\alpha_2}{E_k}\right) \delta_2 - L_E A_k' \delta_2 - L_E \left(\frac{\alpha_2}{E_k}\right) \delta_2 - L_E \left(\frac{\alpha_1}{E_k}\right) \delta_1 \delta_2 \right\rangle \end{cases} \quad (42)$$

where expectation terms in equation (42) not previously evaluated are $\langle \delta_1 \delta_2 \rangle = 0$ and

$$\begin{cases} \langle L_E^2 \delta_1 \rangle = \frac{A(\boldsymbol{\theta})}{(-1+\alpha_1)^3} \left(\frac{E_1}{E_k}\right)^{-\alpha_1} \left(2 \left(E_1 - \left(\frac{E_1}{E_k}\right)^{\alpha_1} E_k \right) + 2(-1+\alpha_1) E_1 \log \left[\frac{E_1}{E_k} \right] + (-1+\alpha_1)^2 E_1 \left(\log \left[\frac{E_1}{E_k} \right] \right)^2 \right) \\ \langle L_E^2 \delta_2 \rangle = \frac{A(\boldsymbol{\theta})}{(-1+\alpha_2)^3} \left(\frac{E_2}{E_k}\right)^{-\alpha_2} \left(2 \left(-E_2 + \left(\frac{E_2}{E_k}\right)^{\alpha_2} E_k \right) - 2(-1+\alpha_2) E_2 \log \left[\frac{E_2}{E_k} \right] - (-1+\alpha_2)^2 E_2 \left(\log \left[\frac{E_2}{E_k} \right] \right)^2 \right) \end{cases} \quad (43)$$

Finally, the matrix elements defined in equation (42) are multiplied by N to give the information matrix $\mathbf{I}(\boldsymbol{\theta})$. It is interesting to note that $\mathbf{I}(\boldsymbol{\theta})$ contains the intuitive terms $N\langle\delta_1\rangle$ and $N\langle\delta_2\rangle$, which are the expected number of events below and above the spectral knee E_k , respectively.

This completes the derivation of the closed-form expression of the information matrix $\mathbf{I}(\boldsymbol{\theta})$ corresponding to the so-called ideal detector having perfect energy resolution. Its inverse $\mathbf{I}^{-1}(\boldsymbol{\theta})$ gives the CRB, and the square root of the diagonal elements of $\mathbf{I}^{-1}(\boldsymbol{\theta})$ is the CRB for the three spectral parameters. In practice, one should verify that the difference of the covariance matrix \mathbf{C} and the CRB is positive definite by checking that the eigenvalues of the matrix $[\mathbf{C} - \mathbf{I}^{-1}(\boldsymbol{\theta})]$ are all positive. If the eigenvalues are zero, the CRB has been reached and the ML estimator has attained property P2 (efficiency) by achieving the CRB.

The computer program *MATHEMATICA*¹⁵ was used to provide equations (40)–(43) and performed the important check

$$\begin{cases} 0 = \langle A_1' - L_E \delta_1 \rangle \\ 0 = \langle A_2' - L_E \delta_2 \rangle \\ 0 = \langle A_k' + \frac{\alpha_1}{E_k} \delta_1 + \frac{\alpha_2}{E_k} \delta_2 \rangle \end{cases} \quad (44)$$

that confirms the correctness of the terms leading to the construction of the information matrix $\mathbf{I}(\boldsymbol{\theta})$. Furthermore, a format feature of *MATHEMATICA* called FortranForm was used to write these equations in FORTRAN source code for implementation in the overall simulation program and thereby eliminated the possibility of introducing human error in transferring the equations into the computer program.

A.2 Check on Equation (39) by Professor Bierens¹⁶

“Equation (39) is correct if (38) is correct, which in its turn is the case if (38) holds for fixed p and q . In the latter case (38) states that the derivative operator d/da can be moved inside the integral, which requires the application of the dominated convergence theorem. The conditions for the latter are:

1. For each x in $[p, q]$, $f(x, a)$ is differentiable in a , possibly except for x in a subset with Lebesgue measure zero.
2. There exists a function $b(x, a)$, say, with finite integral over $[p, q]$ such that $\sup\{n|f(x, a+1/n) - f(x, a)| < b(x, a)$ and $\sup\{n|f(x, a-1/n) - f(x, a)| < b(x, a)$, where the “sup” is taken over all $n > 0$.

In (39) you have applied (38) for the cases $q = a$ and p fixed, and q fixed and $p = a$, so that (39) is correct if the conditions 1 and 2 hold for $a = q$ and $a = p$. In the case $q = a$ and p fixed, the log-density $f(x, q)$ is differentiable in q for $x < q$, but not for $x = q$. Similarly, in the case q fixed and $p = a$, $f(x, p)$ is differentiable in p for $x > p$ but not for $x = p$. Thus condition 1 holds, with $\{p, q\}$ the subset with Lebesgue measure zero. Moreover, since the left and right partial derivatives of $f(x, a)$ to “ a ” are bounded, condition 2 holds as well. Consequently, (39) is correct.”

APPENDIX B—CRAMER-RAO BOUND FOR MULTIPLE INDEPENDENT DATA SETS

Derivation of the CRB using two independent data sets, A and B, considered simultaneously to estimate a single spectral parameter, θ , follow along the lines of the proof of the CR inequality in Hogg and Craig¹⁷ for a single data set. Starting with the probability density function for detector A's response (e.g., energy deposit) data set $A = \{x_i | i = 1, \dots, N_A\}$ as

$$g_A(x_i; \theta) = \int_{R_A} g_A(x_i | E, \rho_A) \phi(E; \theta) dE, \quad i = 1, \dots, N_A \quad (45)$$

and similarly for data set $B = \{y_j | j = 1, \dots, N_B\}$ as

$$g_B(y_j; \theta) = \int_{R_B} g_B(y_j | E, \rho_B) \phi(E; \theta) dE, \quad j = 1, \dots, N_B, \quad (46)$$

and because

$$1 = \int_{R_A} g_A(x_i; \theta) dx_i, \quad i = 1, \dots, N_A \quad (47)$$

and also for data set B, their product gives the joint probability density function of the two data sets considered simultaneously. Assuming the derivative operator $d/d\theta$ can be moved inside the integral, differentiating their product with respect to θ gives

$$0 = \int_{R_A} \frac{\partial \log [g_A(x_i; \theta)]}{\partial \theta} g_A(x_i; \theta) dx_i + \int_{R_B} \frac{\partial \log [g_B(y_j; \theta)]}{\partial \theta} g_B(y_j; \theta) dy_j. \quad (48)$$

Next, define the random variable, Z , as

$$Z = \sum_{i=1}^{N_A} \frac{\partial \log [g_A(x_i; \theta)]}{\partial \theta} + \sum_{j=1}^{N_B} \frac{\partial \log [g_B(y_j; \theta)]}{\partial \theta}. \quad (49)$$

It follows from equation (48) that the mean of Z is zero, or using expected value notation, $\langle Z \rangle = 0$. Moreover, Z is the sum of $(N_A + N_B)$ mutually stochastically independent random variables, each with mean zero and consequently with variance*

$$\sigma_Z^2 = N_A \left\langle \left(\frac{\partial \log [g_A(x; \theta)]}{\partial \theta} \right)^2 \right\rangle + N_B \left\langle \left(\frac{\partial \log [g_B(y; \theta)]}{\partial \theta} \right)^2 \right\rangle. \quad (50)$$

Next, assume there exists an unbiased statistic, U , of the parameter θ that is a function of the $(N_A + N_B)$ instrument responses comprising data sets A and B considered acting together, so that $U = u(x_1, \dots, x_{N_A}, y_1, \dots, y_{N_B})$ and hence

$$\begin{aligned} \theta = \int_{R_B} \cdots \int_{R_A} u(x_1, \dots, x_{N_A}, y_1, \dots, y_{N_B}) & g_A(x_1; \theta) \cdots g_A(x_{N_A}; \theta) g_B(y_1; \theta) \\ & \cdots g_B(y_{N_B}; \theta) dx_1, \dots, dx_{N_A}, dy_1, \dots, dy_{N_B} \end{aligned} \quad (51)$$

and differentiating both sides with respect to θ gives

$$\begin{aligned} 1 = \int_{R_B} \cdots \int_{R_A} u(x_1, \dots, x_{N_A}, y_1, \dots, y_{N_B}) & \left[\sum_{i=1}^{N_A} \frac{1}{g_A(x_i; \theta)} \frac{\partial g_A(x_i; \theta)}{\partial \theta} \right. \\ & \left. + \sum_{j=1}^{N_B} \frac{1}{g_B(y_j; \theta)} \frac{\partial g_B(y_j; \theta)}{\partial \theta} \right] \times g_A(x_1; \theta) \cdots g_A(x_{N_A}; \theta) g_B(y_1; \theta) \\ & \cdots g_B(y_{N_B}; \theta) dx_1 \cdots dx_{N_A} dy_1 \cdots dy_{N_B} \\ = \int_{R_B} \cdots \int_{R_A} u(x_1, \dots, x_{N_A}, y_1, \dots, y_{N_B}) & \left[\sum_{i=1}^{N_A} \frac{\partial \log [g_A(x_i; \theta)]}{\partial \theta} \right. \\ & \left. + \sum_{j=1}^{N_B} \frac{\partial \log [g_B(y_j; \theta)]}{\partial \theta} \right] \times g_A(x_1; \theta) \cdots g_A(x_{N_A}; \theta) g_B(y_1; \theta) \\ & \cdots g_B(y_{N_B}; \theta) dx_1 \cdots dx_{N_A} dy_1 \cdots dy_{N_B}. \end{aligned} \quad (52)$$

*The expectation of the cross-product terms is zero because of independence.

This shows that $\langle UZ \rangle = 1$, and, because $\langle UZ \rangle = \langle U \rangle \langle Z \rangle + \tau \sigma_U \sigma_Z$, where τ is the correlation coefficient of U and Z , we have

$$1 = \theta \times 0 + \tau \sigma_U \sigma_Z \quad \text{or} \quad \tau = \frac{1}{\sigma_U \sigma_Z} . \quad (53)$$

Since $\tau^2 \leq 1$, it follows that $\sigma_U^2 \geq 1/\sigma_Z^2$ and therefore

$$\sigma_U^2 \geq \frac{1}{N_A \left\langle \left(\frac{\partial \log [g_A(x; \theta)]}{\partial \theta} \right)^2 \right\rangle + N_B \left\langle \left(\frac{\partial \log [g_B(y; \theta)]}{\partial \theta} \right)^2 \right\rangle} , \quad (54)$$

which establishes the CRB for two independent data sets, A and B, for an unbiased estimator of the single spectral parameter of an energy spectrum. This bound readily extends to k independent data sets by induction and furthermore confirms that “it always pays to use additional data sets” because of the additional variance reduction and hence improvements in measurement precision, as long as the additional data sets provide unbiased information.

The above derivation generalizes to the case where $\boldsymbol{\theta}$ is a vector of spectral parameters by the additivity of information matrices¹⁸ so that the CRB for the individual parameters comprising $\boldsymbol{\theta}$ are the diagonal elements of the inverse of the information matrix, \mathbf{I} , having ij -elements:

$$I_{ij} = N_A \left\langle \frac{\partial \log [g_A(x; \boldsymbol{\theta})]}{\partial \theta_i} \frac{\partial \log [g_A(x; \boldsymbol{\theta})]}{\partial \theta_j} \right\rangle + N_B \left\langle \frac{\partial \log [g_B(y; \boldsymbol{\theta})]}{\partial \theta_i} \frac{\partial \log [g_B(y; \boldsymbol{\theta})]}{\partial \theta_j} \right\rangle \quad (55)$$

and, as in the single parameter case, equation (55) is readily extended to k independent data sets.

For the interesting case where a detector does not observe events in an energy range represented by one or more of the spectral parameters, the partial derivatives with respect to those parameters are zero. For example, suppose detector A observes only events below the spectral knee, detector C only observes events above the knee, and detector B observes both above and below the knee. Then the matrix to be inverted to obtain the CRB for an unbiased estimator of $\boldsymbol{\theta} = (\alpha_1, \alpha_2, E_k)$ is

$$\begin{aligned}
 I_{ij} = & N_A \begin{bmatrix} \left\langle \left(\frac{\partial \log [g_A(x; \alpha_1)]}{\partial \alpha_1} \right)^2 \right\rangle & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
 & + N_B \left\langle \frac{\partial \log [g_B(y; \boldsymbol{\theta})]}{\partial \theta_i} \frac{\partial \log [g_B(y; \boldsymbol{\theta})]}{\partial \theta_j} \right\rangle \\
 & + N_C \begin{bmatrix} 0 & 0 & 0 \\ 0 & \left\langle \left(\frac{\partial \log [g_C(x; \alpha_2)]}{\partial \alpha_2} \right)^2 \right\rangle & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{56}
 \end{aligned}$$

where the notation in the second term of the right-hand side of equation (56) defines $\theta_1 \equiv \alpha_1$, $\theta_2 \equiv \alpha_2$, and $\theta_3 \equiv E_k$ as before and where the integration range for this term must be split at E_k in its numerical evaluation. An example using equation (56) is provided in section 1.6 of this TP.

REFERENCES

1. Howell, L.W.: “Maximum Likelihood Estimation of the Broken Power Law Spectral Parameters With Detector Design Applications,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, In press, 2002.
2. Howell, L.W.: “A Recommended Procedure for Estimating the Cosmic-Ray Spectral Parameter of a Simple Power Law,” *Nuclear Instruments and Methods in Physics Research*, Vol. 480(2–3), pp. 739–751, 2002.
3. Kendall, M.G.; and Stuart, A.: *The Advanced Theory of Statistics*, Vol. 2, 2nd ed., Hafner Publishing Company, New York, NY, pp. 8–10 and pp. 35–67, 1963.
4. Press, W.H.; et al.: *Numerical Recipes in FORTRAN*, Cambridge University Press, New York, NY, pp. 402–405, 1992.
5. Press, W.H.; et al.: *Numerical Recipes in FORTRAN*, Cambridge University Press, New York, NY, pp. 140–155, 1992.
6. Kendall, M.G.; and Stuart, A.: *The Advanced Theory of Statistics*, Vol. 2, 2nd ed., Hafner Publishing Company, New York, NY, pp. 8–10 and pp. 55–60, 1963.
7. Bierens, H.J.: *Topics in Advanced Econometrics*, Cambridge University Press, New York, NY, pp. 60–67, 1994.
8. Lehmann, E.L.: *Theory of Point Estimation*, John Wiley & Sons, New York, NY, pp. 409–443, 1983.
9. Voinov, G.G.; and Nikulin, M.S.: *Unbiased Estimators and Their Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 90–104, 1993.
10. Lee, J.; Watts, J.; and Howell, L.: “Simulations of a Thin Sampling Calorimeter with GEANT/FLUKA,” *Nuclear Instruments and Methods in Physics Research*, Vol. 470, pp. 500–511, 2001.
11. Press, W.H.; et al.: *Numerical Recipes in FORTRAN*, Cambridge University Press, New York, NY, pp. 180–181, 1992.
12. Brownlee, K.A.: *Statistical Theory and Methodology in Science and Engineering*, John Wiley & Sons, Inc., New York, NY, 1960.
13. Conover, W.J.; Johnson, M.E.; and Johnson, M.M.: *Technometrics*, Vol. 23(4), p. 351, 1981.

14. Proceedings of the ACCESS Collaboration Meeting, J.H. Adams, ed., National Space Science Technology Center, Huntsville, AL, January 17–20, 2001.
15. Wolfram, S.: *MATHEMATICA*, Version 4, Wolfram Research, 1999.
16. Bierens, H.J.: Department of Economics, The Pennsylvania State University, University Park, PA, Private Communication, 2001.
17. Hogg, R.V.; and Craig, A.T.: *Introduction to Mathematical Statistics*, 3d ed., The Macmillan Company, New York, NY, 1972.
18. Lehmann, E.L.: *Theory of Point Estimation*, John Wiley & Sons, New York, NY, 1983.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operation and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503				
1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE December 2002	3. REPORT TYPE AND DATES COVERED Technical Publication		
4. TITLE AND SUBTITLE Statistical Properties of Maximum Likelihood Estimators of Power Law Spectra Information			5. FUNDING NUMBERS	
6. AUTHORS L.W. Howell				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) George C. Marshall Space Flight Center Marshall Space Flight Center, AL 35812			8. PERFORMING ORGANIZATION REPORT NUMBER M-1056	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Washington, DC 20546-0001			10. SPONSORING/MONITORING AGENCY REPORT NUMBER NASA/TP-2002-212020/REV1	
11. SUPPLEMENTARY NOTES Prepared by Space Science Department, Science Directorate				
12a. DISTRIBUTION/AVAILABILITY STATEMENT Unclassified-Unlimited Subject Category 90 Standard Distribution			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) A simple power law model consisting of a single spectral index, α_1 , is believed to be an adequate description of the galactic cosmic-ray (GCR) proton flux at energies below 10^{13} eV, with a transition at the knee energy, E_k , to a steeper spectral index $\alpha_2 > \alpha_1$ above E_k . The maximum likelihood (ML) procedure was developed for estimating the single parameter α_1 of a simple power law energy spectrum and generalized to estimate the three spectral parameters of the broken power law energy spectrum from simulated detector responses and real cosmic-ray data. The statistical properties of the ML estimator were investigated and shown to have the three desirable properties: (P1) consistency (asymptotically unbiased), (P2) efficiency (asymptotically attains the Cramer-Rao minimum variance bound), and (P3) asymptotically normally distributed, under a wide range of potential detector response functions. Attainment of these properties necessarily implies that the ML estimation procedure provides the best unbiased estimator possible. While simulation studies can easily determine if a given estimation procedure provides an unbiased estimate of the spectra information, and whether or not the estimator is approximately normally distributed, attainment of the Cramer-Rao bound (CRB) can only be ascertained by calculating the CRB for an assumed energy spectrum-detector response function combination, which can be quite formidable in practice. However, the effort in calculating the CRB is very worthwhile because it provides the necessary means to compare the efficiency of competing estimation techniques and, furthermore, provides a stopping rule in the search for the best unbiased estimator. Consequently, the CRB for both the simple and broken power law energy spectra are derived herein and the conditions under which they are attained in practice are investigated. The ML technique is then extended to estimate spectra information from an arbitrary number of astrophysics data sets produced by vastly different science instruments. This theory and its successful implementation will facilitate the interpretation of spectral information from multiple astrophysics missions and thereby permit the derivation of superior spectral parameter estimates based on the combination of data sets.				
14. SUBJECT TERMS cosmic ray, simple power law, broken power law, maximum likelihood, spectral information, Cramer-Rao bound			15. NUMBER OF PAGES 60	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited	