THE ASTROPHYSICAL JOURNAL, **210**: 642–646, 1976 December 15


# ENERGY SPECTRA OF X-RAY CLUSTERS OF GALAXIES*

YORAM AVNI
Institute of Astronomy, Cambridge, England; and Center for Astrophysics, Cambridge, Massachusetts
*Received 1975 December 1; revised 1976 May 3*

## ABSTRACT

A procedure for estimating the ranges of parameters that describe the spectra of X-rays from clusters of galaxies is presented. The applicability of the method is proved by statistical simulations of cluster spectra; such a proof is necessary because of the nonlinearity of the spectral functions. Implications for the spectra of the Perseus, Coma, and Virgo clusters are discussed. The procedure can be applied in more general problems of parameter estimation.

*Subject headings:* galaxies: clusters of — X-rays: sources — X-rays: spectra

## I. INTRODUCTION

The problem of estimating the ranges of parameters that describe the X-ray spectra of clusters of galaxies has been a subject of extensive debate during the last year. The basic issue is whether or not these spectra have a low-energy cutoff (Kellogg 1973; Bahcall 1975 and references quoted therein). But the discussions in the literature have evolved into the general problem of finding confidence limits in nonlinear multiparameter fits (Kellogg, Baldwin, and Koch 1975; Margon *et al.* 1975; Kellogg 1975; Lampton, Margon, and Bowyer 1976). It has become evident that there is no general agreement on the value of the increment $\Delta\chi^2$ that has to be added to $\chi^2_{\min}$ in order to find confidence limits for the parameters when the "best values" are calculated by minimizing $\chi^2$.

In this paper we present the correct "minimum $\chi^2$" method to calculate confidence limits for any number of parameters. The elements of the method can be found in various references in statistics. (See the Appendix.) The value of $\Delta\chi^2$ depends on the number of parameters that are estimated simultaneously, and not on the total number of parameters in the fitting function. The procedure is exact for linear fits. For any particular nonlinear fitting problem (the degree of nonlinearity being determined both by the functional form and by the dispersion of the measurements), the applicability of the method can be tested by simulation, as we have done for the analysis of cluster spectra.

Based on the above-mentioned statistical analysis we find that *Uhuru* observations, when analyzed in terms of simple spectral functions, yield a formally significant low-energy cutoff for the Coma cluster, but not for the Virgo or Perseus clusters. We argue that no firm conclusion can be reached with the available *Uhuru* data since any such cutoff is outside the dynamic range in energy of the *Uhuru* detectors and since the spectral functions used are oversimplified.

We present the procedure for finding confidence limits in § II. The statistical simulations that we have performed for the problem of cluster spectra are described in § III. We apply the technique to *Uhuru* observations of cluster spectra in § IV. A summary of our results is given in § V. Some formal aspects of the estimation procedure are given in the Appendix, with references to the statistical literature.

* This paper was written while the author was visiting the Institute for Advanced Study, Princeton, New Jersey 08540, with partial support by the National Science Foundation, grant No. 40768X. The author was supported at the Institute of Astronomy by grants from the Israeli Academy and the SRC, and at the Center for Astrophysics by NASA grant NAS 5-20048 with the Smithsonian Astrophysical Observatory.

## II. FINDING CONFIDENCE LIMITS

Consider a set of observed quantities $O_i$, $i = 1, n$ (e.g., the counts of photons per channel in the $n = 7$ *Uhuru* energy bins), that are measured with an accuracy given by the rms dispersions $\sigma_i$. The observations are to be described by a functional form $C_i(\Theta) = C_i(\theta_1, \ldots, \theta_p)$ with $p$ parameters (e.g., a power-law spectrum with $p = 3$ parameters: low-energy cutoff, power-law index, and normalization). Suppose that for some specific application one needs to estimate simultaneously the values of the parameters $\theta_1, \ldots, \theta_q$, which we shall call the "interesting parameters," and is not concerned at all with the values of $\theta_{q+1}, \ldots, \theta_p$, which we shall call the "uninteresting parameters." In other words, we split the vector $\Theta$ into two components, $\Theta = (\Phi, \Psi)$, where $\Phi = (\theta_1, \ldots, \theta_q)$ and $\Psi = (\theta_{q+1}, \ldots, \theta_p)$, and we are interested in estimating $\Phi$ (e.g., we want to estimate the low-energy cutoff in a cluster spectrum and do not care about the power-law index or normalization).

The point-estimator ("best value") for $\Phi$ is obtained by the usual "minimum $\chi^2$" method. One minimizes the function

$$S(\Phi, \Psi) = \sum_{i=1}^{n} [O_i - C_i(\Phi, \Psi)]^2/\sigma_i^2 \qquad (1)$$

with respect to $\Phi$ and $\Psi$, and picks the value of $\Phi$ (which we denote by $\hat{\Phi}$) that yields $S_{\min}$.

The confidence limit for $\Phi$ that corresponds to a given confidence level $\alpha$ ($0 < \alpha < 1$) is obtained in the following way: first, calculate a number $\Delta(q, \alpha)$ so that

$$\text{Probability } (\chi^2(q \text{ degrees of freedom}) \leq \Delta) = \alpha\,; \quad (2)$$

i.e., $\Delta(q, \alpha)$ is the $\chi^2$ value that one finds from the table of a $\chi^2$ distribution with $q$ degrees of freedom. Then the confidence region $R_q^\alpha$ is the set of all the values of $\Phi$ such that

$$S(\Phi, \text{minimized over } \Psi) - S_{\min} \leq \Delta(q, \alpha)\,. \quad (3)$$

In other words, $\Delta(q, \alpha)$ is the required increment of "$\chi^2$," its value is determined by the number of interesting parameters, and "$\chi^2$" must be minimized with respect to the uninteresting parameters for each point in the $q$-dimensional subspace of the interesting parameters.

The region $R_q^\alpha$ is a region in the $q$-dimensional space of the interesting parameters. It is a statistic, namely, a result of the observed values $O_i$. It has the property that when the observation as a whole is repeated a large number of times, the region $R_q^\alpha$ will contain the true value $\Phi^0$ a fraction $\alpha$ of the times. For an outline of the proof of this property see the Appendix. This procedure is exact for linear fitting functions.

A practical numerical technique to find $R_q^\alpha$ is by using $\chi^2$-maps of the $p$-dimensional space of all parameters. One then finds the region $R$ such that inside it $S(\Theta) - S_{\min} \leq \Delta(q, \alpha)$, and then *projects* $R$ on the $q$-dimensional subspace of the vector $\Phi$.

A list of values $\Delta(q, \alpha)$ for some useful values of $q$ and $\alpha$ is given in Table 1. As one can see from this table, the appropriate increment for estimating one interesting parameter (e.g., a low-energy cutoff) at the 68 percent confidence level is 1, independent of the total number of parameters ($p$) in the fitting function.

### III. SIMULATIONS OF CLUSTER SPECTRA

We have used Monte Carlo simulations of X-ray spectra in order to test the validity of the general technique presented in § II for the specific case of cluster spectra. The results of the simulations are summarized in Table 2, and they show that the method works very well.

In the first simulation we mimicked the *Uhuru* observation of the Perseus cluster (Kellogg, Baldwin, and Koch 1975). For ease of computation we used a spectrum of the form

$$I(E) = aE^{-c} \exp(-b/E^3) \text{ photons keV}^{-1} \quad (4)$$

TABLE 1

CONSTANTS FOR CALCULATING CONFIDENCE REGIONS

| $\alpha$ (%) | $q$ (No. of Interesting Parameters) 1 | 2 | 3 |
|---|---|---|---|
| 68......... | 1.00 | 2.30 | 3.50 |
| 90......... | 2.71 | 4.61 | 6.25 |
| 99......... | 6.63 | 9.21 | 11.30 |

to describe the number of counts accumulated during an observation. The energy $E$ is measured in keV. The power-law index $c_0$ was taken as 2.1; the low-energy cutoff parameter $b_0$ was taken as 1.2 keV; the normalization constant $a_0$ was chosen as 15,000 in order to yield the same statistical accuracy as in the actual *Uhuru* experiment. We used seven energy bins with flat response functions. The central energies of the bins (2.09, 2.89, 3.94, 5.10, 6.40, 7.84, and 9.28 keV) and their widths (0.8, 0.85, 1.1, 1.2, 1.3, 1.4, and 1.4 keV) were chosen to correspond as closely as possible to the *Uhuru* bins. Using the input spectrum we calculated the expected counts in each of the seven bins. We then perturbed these numbers using Poisson statistics, and obtained a simulated observation. This observation was then analyzed using a minimum $\chi^2$ fitting program as described in the previous section, and the whole procedure was repeated 50 times.

We tested the procedure for constructing confidence regions for three different values of $q$, the number of interesting parameters. Let us associate $\theta_1 = b$, $\theta_2 = c$, $\theta_3 = a$. If one is interested in knowing simultaneously the values of all the three parameters, then $q = 3$ ($= p$ in this case). For each simulated spectrum we constructed the quantity

$$S_3 = S(b_0, c_0, a_0) - S_{\min}\,. \quad (5)$$

The correctness of the estimation procedure is equivalent to (and actually follows from) the fact that $S_3$ is distributed like $\chi^2$ with three degrees of freedom. (See Appendix. In the special case of $q = p$ this is the well-known decomposition theorem.) Thus one test is to compare the value of the average $\bar{S}_3$ that we find from the simulation to its expected value, and the agreement is good: the simulation gives a value of 2.88, whereas the expected value is $3 \pm 0.35$ ($1\,\sigma$). For a more direct test we chose two popular values of $\alpha$: 90 percent and 68 percent, and counted the number of times that the three-dimensional regions $R_3^\alpha$ contained the point ($b_0, c_0, a_0$), i.e., the "number of successes." This is the number of times that $S_3 \leq \Delta(3, \alpha)$ which we denote by $J_3(\alpha)$. From Table 2 one sees that the values found for $J_3(\alpha)$ agree well with those expected (90% confidence: 45 found, $45 \pm 2.1$ expected. 68% confidence: 35 found, $34 \pm 3.3$ expected).

If, however, one is interested in estimating simultaneously the values of $b_0$ and $c_0$, and does not care about the value of the normalization constant, then the proper value of $q$ is 2. For each simulated spectrum we constructed the quantity

$$S_2 = S(b_0, c_0, \text{minimized over } a_0) - S_{\min}\,. \quad (6)$$

The quantity $S_2$ should be distributed like $\chi^2$ with two degrees of freedom, and, as seen from Table 2, $\bar{S}_2$ which we find agrees with the expected value. The number of times that the two-dimensional region $R_2^\alpha$ contains the point ($b_0, c_0$), i.e., the "number of successes" for this estimation problem is denoted by $J_2(\alpha)$. Again the agreement is good.

Finally, if one just wants to estimate the value of the low-energy cutoff parameter $b_0$, and does not care

TABLE 2

RESULTS OF SIMULATIONS

| Parameters | Simulation No. 1 | Simulation No. 2 | Simulation No. 3 |
|---|---|---|---|
| $a_0$ | 15,000 | 15,000 | 50,000 |
| $b_0$ | 1.2 | 0 | 4.6 |
| $c_0$ | 2.1 | 2.1 | 1.5 |
| No. of Spectra Generated | 50 | 102 | 53 |
| $q = 3$: | | | |
| $\bar{S}_3$: | | | |
| Expected | 3 ± 0.35 | 3 ± 0.24 | 3 ± 0.34 |
| Found | 2.88 | 3.03 | 2.88 |
| $J_3$ (90%): | | | |
| Expected | 45 ± 2.1 | 91.8 ± 3.0 | 47.7 ± 2.2 |
| Found | 45 | 94 | 48 |
| $J_3$ (68%): | | | |
| Expected | 34 ± 3.3 | 69.4 ± 4.7 | 36.0 ± 3.4 |
| Found | 35 | 67 | 34 |
| $q = 2$: | | | |
| $\bar{S}_2$: | | | |
| Expected | 2 ± 0.28 | 2 ± 0.20 | 2 ± 0.27 |
| Found | 2.15 | 2.03 | 1.85 |
| $J_2$ (90%): | | | |
| Expected | 45 ± 2.1 | 91.8 ± 3.0 | 47.7 ± 2.2 |
| Found | 44 | 93 | 49 |
| $J_2$ (68%): | | | |
| Expected | 34 ± 3.3 | 69.4 ± 4.7 | 36.0 ± 3.4 |
| Found | 34 | 66 | 35 |
| $q = 1$: | | | |
| $\bar{S}_1$: | | | |
| Expected | 1 ± 0.20 | 1 ± 0.14 | 1 ± 0.19 |
| Found | 1.08 | 1.05 | 1.01 |
| $J_1$ (90%): | | | |
| Expected | 45 ± 2.1 | 91.8 ± 3.0 | 47.7 ± 2.2 |
| Found | 45 | 88 | 46 |
| $J_1$ (68%): | | | |
| Expected | 34 ± 3.3 | 69.4 ± 4.7 | 36.0 ± 3.4 |
| Found | 34 | 68 | 37 |

NOTE.—The ± statistical uncertainties correspond to 1 $\sigma$.

about the values of the other parameters, the correct value of $q$ to use is 1. For each simulated spectrum we calculated

$$S_1 = S(b_0, \text{minimized over } c_0 \text{ and } a_0) - S_{\min} . \quad (7)$$

$S_1$ is distributed like $\chi^2$ with one degree of freedom, and the comparison of $\bar{S}_1$ that we find from the simulation to the expected value is good. The number of times that the one-dimensional interval $R_1^\alpha$ contains the value $b_0$, which we denote by $J_1(\alpha)$, again shows good agreement between expectation and the results of the simulation.

The type of simulations described above test directly the basic property of confidence regions: the probability that in a given experiment the confidence contours enclose the true value of the parameter. Such simulations were conducted also by Lampton, Margon, and Bowyer (1976) [and also by Avni and Bahcall (1976) for a different problem].

The second set of simulations was identical to the first except that we used the value of $b_0 = 0.0$ for the low-energy cutoff parameter; i.e., an input spectrum without a low-energy cutoff. From Table 2 one can see that again in all the tests that we performed the agreement between the results of the simulation and expectation is very good. The purpose of this simulation was to show that the estimation procedure works equally well for input spectra without a low-energy cutoff as for input spectra with a low-energy cutoff. Therefore, the probability that a spectrum without a cutoff will show up in the results of an observation as if it did have a cutoff is correctly described by the confidence limits assigned by the estimation technique.

The third set of simulations is identical to simulation *A* of Lampton, Margon, and Bowyer (1976). We used their values of $a_0$, $b_0$, $c_0$, and also their energy bins (central energies: 1, 2, 3, 4, 5, 6, 7 keV; width: 1 keV). Again the results summarized in Table 2 show excellent agreement between expectation and simulation, for all values of $q$. Our results are, therefore, entirely consistent with those of Lampton, Margon, and Bowyer (1976) for the special case studied by them when all the parameters appearing in the fitting function are estimated simultaneously (i.e., $q = p = 3$).

We conclude that the technique discussed above is adequate to estimate any subset of parameters in the study of cluster spectra as measured by the *Uhuru* satellite. In particular, if the scientific issue is the value of the cutoff parameter, then $q = 1$, and the appropriate values of $\Delta\chi^2$ are 1 for 68 percent confidence and 2.7 for 90 percent confidence.

## IV. APPLICATIONS TO CLUSTER SPECTRA

We now apply these results to the *Uhuru* observations of the Perseus, Virgo, and Coma Clusters. By projecting the $\Delta\chi^2 = 1$ contours, given by Kellogg, Baldwin, and Koch (1975) for the Perseus cluster, on the low-energy cutoff axis we find that there is statistically significant evidence (more than 90%) for a cutoff if a power-law spectrum is used, but there is no evidence for a cutoff if a thermal spectrum is used. The overall quality of the two fits (as measured by $\chi^2_{\min}$) is similar. This results from the fact that any such cutoff is small, $\sim 1$ keV, and is therefore outside the range in energy of the *Uhuru* detectors. In fact, data at lower energies, as summarized by Kellogg, Baldwin, and Koch (1975), show a soft excess relative to the two spectral forms used (indicating that the actual spectrum is more complex) and definitely do not have a low-energy cutoff as large as 1 keV.

The situation for the Virgo cluster is similar. From the errors quoted by Kellogg, Baldwin, and Koch (1975) on the value of the low-energy cutoff when the power-law spectrum is used, and by projecting the $\Delta\chi^2$ contours given by Kellogg (1975) for the thermal fit onto the cutoff axis, we find significant evidence (more than 90%) for the existence of a cutoff in the first case; no evidence in the latter case. Observations at energies below 1 keV, summarized by Kellogg, Baldwin, and Koch (1975), show again indications of an excess soft flux, but the data are much more meager than for the Perseus cluster.

The *Uhuru* observations of the Coma cluster give a somewhat different picture. From the errors quoted by Kellogg, Baldwin, and Koch (1975) for the value of the low-energy cutoff when a power-law spectrum is used, one finds significant evidence for a cutoff (more than 90%), as in the previous two clusters. By projecting the $\Delta\chi^2$ contours given by Kellogg (1975) for the thermal fit onto the cutoff axis, we also find for this spectrum significant evidence (just above 90%: the 90% confidence interval for $b$ just fails to include zero) for a low-energy cutoff, unlike the other two clusters. Thus on the basis of *Uhuru* observations alone, and using the above simple spectral forms, one might conclude that there is a low-energy cutoff in the spectrum of the Coma cluster. We believe, however, that such a conclusion is premature because of the reasons discussed above: the low-energy cutoff being outside the dynamic range of the observation and the complexity of the actual spectrum. Observations below 1 keV, summarized by Kellogg, Baldwin, and Koch (1975), show a very slight indication for an excess soft flux, so the situation may perhaps be similar to that of the Perseus cluster. Surely, if the actual spectrum is more complex than the spectral form used in the fit, then even with much better counting statistics no firm conclusion can be made on the behavior of the spectrum outside the range where it was actually observed.

## V. SUMMARY

We have described a method, the elements of which can be found in statistical references (see Appendix), to calculate confidence limits for any number of parameters. The increment $\Delta\chi^2$ for a given confidence level depends on the number of parameters being estimated simultaneously, which in turn depends on the nature of the scientific question being asked, and not on the total number of parameters in the fitting function. We have shown by Monte Carlo simulations that this method (which is precise for linear fits) works well for the specific problem of studying the X-ray spectrum from clusters of galaxies. The scientific issue here is whether or not these spectra have a low-energy cutoff. Therefore the statistical problem is to estimate the value of the cutoff parameter; the appropriate values of $\Delta\chi^2$ are 1 for 68 percent confidence or 2.71 for 90 percent confidence. In using these increments $\Delta\chi^2$, one has to minimize $\chi^2$ with respect to the rest of the parameters (rather than keeping them fixed at the values which yield the overall $\chi^2_{\min}$).

By applying this technique to *Uhuru* observations of the Perseus, Virgo, and Coma clusters, we find marginal formal evidence for the existence of a small low-energy cutoff in the spectrum of Coma. As such cutoff is outside the range in energy of the detectors, no firm conclusion can be reached since the actual spectrum may be more complex than the spectral forms used in the fitting procedure.

It is a real pleasure to thank P. Joss for a stimulating discussion and persistent criticism. I wish also to thank E. Kellogg and B. Margon for sending me preprints of their work prior to publication and for useful comments.

## APPENDIX

In this Appendix we describe in outline form the formal aspects of the estimation technique. The minimum $\chi^2$ $p$-dimensional estimator $\hat{\Theta}$ in linear fits and for large samples has a multivariate normal distribution around the true value $\Theta^0$ with a $p \times p$ covariance matrix $V$ given asymptotically by (Martin 1971; Bevington 1969)

$$(V^{-1})_{j,k} = M_{j,k} = \frac{1}{2}\frac{\partial^2 S}{\partial\theta_j \partial\theta_k}\,, \quad 1 \leq j, k \leq p\,, \tag{A1}$$

where $S$ is given by equation (1), and the second derivative can be evaluated at $\Theta = \hat{\Theta}$ in the above limit. The distribution of the $q$-dimensional subvector $\Phi$ is therefore also a multivariate normal distribution, with a $q \times q$

covariance matrix $v$ which is obtained from $V$ by deleting the last $p - q$ rows and columns. Defining now $m = v^{-1}$ it follows that $T(\Phi^0)$, where $T(\Phi)$ is the bilinear form

$$T(\Phi) = (\Phi - \hat{\Phi})^T m(\Phi - \hat{\Phi}) , \tag{A2}$$

is distributed like $\chi^2$ with $q$ degrees of freedom. It is then obvious that a confidence region $R_q^\alpha$ can be defined as the set of all values of $\Phi$ for which $T(\Phi) \leq \Delta(\alpha, q)$, with $\Delta$ given by equation (2). It can now be shown using purely algebraic matrix manipulations (Arndt and MacGregor 1966) that

$$T(\Phi) = [(\Phi - \hat{\Phi}, \Psi - \hat{\Psi})^T M(\Phi - \hat{\Phi}, \Psi - \hat{\Psi})]_{\text{minimized over } \Psi} . \tag{A3}$$

Using now the relation (A1) between $M$ and $S$, and the fact that by definition $\partial S/\partial\Theta = 0$ at $\Theta = \hat{\Theta}$, we find

$$T(\Phi) = S(\Phi, \text{minimized over } \Psi) - S_{\min} , \tag{A4}$$

which yields the expression (3) for $R_q^\alpha$.

The same arguments should also apply for nonlinear fits if the sample size is large enough. For large samples the dispersion of $\hat{\Theta}$ around $\Theta^0$ becomes small, and therefore $C(\Theta)$ can be replaced by its linear approximation in the vicinity of $\Theta^0$. Under these conditions, $\hat{\Theta}$ is distributed normally, $S$ can be approximated by its expansion to second order, and all the rest of the arguments follow as before. The formal proof probably exists in the statistical literature. As these considerations apply only asymptotically, the applicability of the estimation procedure in any specific nonlinear problem with finite statistics should be checked by numerical simulations.

## REFERENCES


Arndt, R. A., and MacGregor, M. H. 1966, *Methods in Computational Physics*, **6**, 253.
Avni, Y., and Bahcall, N. A. 1976, *Ap. J.*, **209**, 16.
Bahcall, N. A. 1975, *Proc. Seventh Texas Symposium on Relativistic Astrophysics* (*Ann. N.Y. Acad. Sci.*).
Bevington, P. R. 1969, *Data Reduction and Error Analysis for the Physical Sciences* (New York: McGraw-Hill).
Kellogg, E. M. 1973, in *IAU Symposium No. 55, X- and Gamma-Ray Astronomy*, ed. H. Bradt and R. Giacconi (Dordrecht: Reidel).
Kellogg, E. M. 1975, preprint.
Kellogg, E. M., Baldwin, J. R., and Koch, D. 1975, *Ap. J.*, **199**, 299.
Lampton, M., Margon, B., and Bowyer, S. 1976, *Ap. J.*, **208**, 177.
Margon, B., Lampton, M., Bowyer, S., and Cruddace, R. 1975, *Ap. J.*, **197**, 25.
Martin, B. R. 1971, *Statistics for Physicists* (New York: Academic Press).


YORAM AVNI: Department of Physics, Weizmann Institute of Science, Rehovot, Israel