

WISCONSIN

CONF-960543--26

UNIVERSITY OF WISCONSIN • MADISON, WISCONSIN

PLASMA PHYSICS

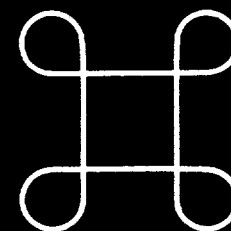
MAXIMUM-LIKEHOOD FITTING OF DATA DOMINATED BY
POISSON STATISTICAL UNCERTAINTIES

45
7-11-96 JS @

M.R. Stoneking
D.J. Den Hartog

DOE/ER/53198-274

June 1996



WISCONSIN

NOTICE

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for any third party's use or the results of such use of any information, apparatus, product or process disclosed in this report, or represents that its use by such third party would not infringe privately owned rights.

Printed in the United States of America
Available from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road
Springfield, VA 22161

NTIS Price codes

Printed copy: A02

Microfiche copy: A01

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

Maximum-Likelihood Fitting of Data Dominated by Poisson Statistical Uncertainties

M. R. Stoneking and D. J. Den Hartog

University of Wisconsin - Madison, Department of Physics, 1150 University Avenue, Madison, WI 53706

Abstract

The fitting of data by χ^2 -minimization is valid only when the uncertainties in the data are normally distributed. When analyzing spectroscopic or particle counting data at very low signal level (e.g. a Thomson scattering diagnostic), the uncertainties are distributed with a Poisson distribution. We have developed a maximum-likelihood method for fitting data that correctly treats the Poisson statistical character of the uncertainties. This method maximizes the total probability that the observed data are drawn from the assumed fit function using the Poisson probability function to determine the probability for each data point. The algorithm also returns uncertainty estimates for the fit parameters. We compare this method with a χ^2 -minimization routine applied to both simulated and real data. Differences in the returned fits are greater at low signal level (less than ~ 20 counts per measurement). The maximum-likelihood method is found to be more accurate and robust, returning a narrower distribution of values for the fit parameters with fewer outliers.

I. Introduction

When experimental data is fit to a function that is known to represent the phenomenon under investigation, the standard strategy is to minimize χ^2 with respect to changes in the parameters of the "fit" function. Such a strategy relies on the assumption that the uncertainties in the data are normally distributed. For non-normal uncertainty distributions such as the Poisson distribution, minimizing χ^2 does not maximize the *likelihood* that the fitted parameters reflect the data. The maximum-likelihood method is described in the standard texts (1,2). Direct application of the method using Poisson distributed uncertainties was made previously for a specific X-ray photon counting application (3). However, the general treatment in the case of Poisson distributed uncertainties for an arbitrary fit function (Eq. 2) and its application to Thomson scattering data from (nonrelativistic) high temperature plasma (Eqs. 4) is new. Published work dealing with Thomson scattering data (4,5) have pointed out the need for correctly evaluating the measurement uncertainty in the case of low signal photon counting, but those uncertainties were then used in least squares (or χ^2 -minimization) fitting routines, rather than the method described in this paper. Our motivation to develop this method arose from the need to accurately and correctly fit Thomson scattering data with very low signal levels (<20 counts per measurement).

II. Description of the Maximum-Likelihood Method

In order to fit data, the function that represents the phenomenon being investigated must be known. Let the function be written as $u(x, a_1 \dots a_M)$, where x is the independent variable and $a_1 \dots a_M$ are adjustable parameters to be determined by the fit. For a given set of parameters one can calculate the probability of measuring the data points, $y_i(x_j)$. The maximum-likelihood fitting method is founded on the assumption that the probability of measuring a set of data points, $y_1 \dots y_N$, given a set of adjustable parameters, $a_1 \dots a_M$, is a measure of the *likelihood* of the parameters given

the data, and that maximizing the former also maximizes the latter(2). For counting measurements, the probability of measuring a single value, y_i , is given by the Poisson probability function, $P_i = u^{y_i} \exp(-u)/y_i!$. The total probability of measuring the N values of y_i is the product of the individual probabilities,

$$P = \prod_{i=1}^N \frac{\exp(-u)u^{y_i}}{y_i!}. \quad (1)$$

If the measurement uncertainties were *normally* distributed the product would be over the Gaussian probabilities with standard deviation, σ , given by the measurement uncertainties. Maximizing the product of Gaussian probabilities is equivalent to minimizing the χ^2 merit function; χ^2 -minimization is therefore an application of the maximum-likelihood method for the case of normally distributed uncertainties.

The total Poisson probability is maximized by setting the partial derivatives (of the natural logarithm) of Eq. 1 with respect to the fit parameters to zero, $\partial/\partial a_j (\ln P) = 0$. The result is a set of M equations the roots of which yield the fit parameters that maximize the Poisson probability,

$$\frac{\partial}{\partial a_j} (\ln P) = \sum_{i=1}^N \frac{\partial u(x_i)}{\partial a_j} \left(1 - \frac{y_i}{u(x_i)} \right) = 0. \quad (2)$$

We now apply this method for the fit function applicable to Thomson scattering data reduction. For this diagnostic method, monochromatic radiation is Thomson scattered off the electrons in a high temperature plasma(6). The scattered spectrum of radiation has a Doppler spread determined by the electron velocity distribution. Scattered light is collected in wavelength bands close to the wavelength of the incident light, and fit to the spectrum expected from a Maxwellian electron distribution.

Relativistic effects can be neglected for temperatures less than ~ 1 keV(6), and in that case the fit function is a Gaussian with two parameters,

$$u(x_i) = a_1 \exp\left(\frac{-x_i^2}{2a_2^2}\right) \Delta x_i, \quad (3)$$

where x is the wavelength shift ($\Delta\lambda$) and Δx is the channel width (in wavelength units). The Gaussian width, a_2 , is related to the electron temperature by $a_2^2 = 2\lambda_o^2 T_e / mc^2$ (90° scattering). The electron density is proportional to the area under the Gaussian, $a_1 a_2 \propto n_e$. It is important to insert the channel width (and any channel to channel calibration factors) into the fitting function (Eq. 3) rather than modifying the raw data. If the raw photon count data is modified, the Poisson statistical character of the errors is distorted.

Applying the maximum-likelihood method to the fit function in Eq. 3 yields the equations,

$$\sum_{i=1}^N (u(x_i) - y_i) = 0, \quad \sum_{i=1}^N x_i^2 (u(x_i) - y_i) = 0. \quad (4)$$

We employ the Newton-Raphson method to find the roots of Eqs.4. Since there are only two fit parameters, the Jacobian matrix is 2X2, can be immediately inverted. For more than two fit parameters, a numerical matrix inversion is required.

An estimate of the uncertainty in the fit parameters is made by direct application of the error propagation equation for a set of N independent measurements,

$$\sigma_a^2 = \sum_{i=1}^N \sigma_{y_i}^2 \left(\frac{\partial a}{\partial y_i} \right)^2, \quad (5)$$

where $\sigma_{y_i} = \sqrt{y_i}$ are the Poisson measurement uncertainties for each data point. The derivative in Eq. 5 is evaluated using a 5-point numerical derivative. The fit parameters are determined for 4 "perturbed" values of y in addition to the measured value that determines the best fit.

III. Testing the Fitting Algorithm with Simulated Data

Simulated Thomson scattering data was generated for five wavelength channels by drawing random deviates that were Poisson distributed about a mean given by Eq. 3. The values of the fit parameters returned by the maximum-likelihood (Poisson) code are compared with the simulation input values and with those returned by the χ^2 -minimizing code. As expected, differences between the two codes, and between the code and the simulation inputs are smaller at high signal intensity (larger simulation values for a_1). Figure 1a shows the difference (%) in the returned values for the electron temperature from the Poisson and χ^2 codes. The range of the difference is less than 5% at values of a_1 greater than 20 (more than ~50 counts per channel). The difference range between the fit parameters and the actual value (350 eV) is larger, about 20% at high signal (Fig. 1b). This reflects the rapid convergence of the Poisson probability distribution toward the normal probability distribution with increased mean values. The less rapid convergence of the fits toward the simulation is understood, since the uncertainty (*i.e.* random fluctuation) in each data point is significant beyond the signal level where the two probability distributions converge.

The true test of the maximum-likelihood method with Poisson uncertainties comes at low signal intensities where the difference between the Poisson distribution and the normal distribution is greatest. Figure 2 is a scatter plot of values for the electron temperature returned by the two codes plotted against each other for approximately 3000 fits to simulated data. The amplitude parameter a_1 was held fixed at a value of 2.0 that simulates the low end of our experimentally realized signal range

(about 50 total counts in 5 channels), and the electron temperature was fixed at 350 eV. All of the scatter is a result of the random (Poisson) fluctuations; no additional noise sources are simulated. The scatter in the temperature values are not independent but tend to lie along the "line of agreement." However, there is greater scatter vertically than horizontally, indicating that the distribution of returned values for the maximum-likelihood code is narrower with fewer outliers than the distribution from the χ^2 -minimizing code.

The distribution of returned temperature values from the Poisson code is asymmetric having a wing on the high temperature side (Fig. 3) with a mean value 11% higher than the simulated temperature. The χ^2 code returns a mean temperature that is 17% higher than the simulated value with a similarly shaped distribution. At higher signal levels, the distribution becomes symmetric and the mean approaches the actual value (for both codes).

The other parameter of interest in Thomson scattering data analysis is the area under the Gaussian spectrum ($\text{area} = a_1 a_2 \sqrt{2\pi}$) which is proportional to electron density. For the same data set shown in Fig. 2, the distribution of returned values of area is symmetric, with mean values close to that of the simulation. The maximum-likelihood fit value for the area is larger than the simulated value by 1.5%, and the χ^2 value is 5% low.

The histogram in figure 4 was generated to assess the estimated uncertainty in the fit parameters determined using Eq. 5. It shows the number of returned fits that fell within $1\sigma, 2\sigma$, etc., of the simulated value for the temperature, where the σ estimated from the code was used for each fit. If the distribution is Gaussian and σ is the Gaussian width then 68% of the fits should fall within 1σ of the simulated value, 95% within 2σ , etc. The distribution of values is not Gaussian (Fig. 3), and the statistical analysis of the uncertainties shows that the algorithm (Eq. 5) underestimates the uncertainty as defined for a Gaussian distribution. 40-50% of the returned fits fall

within 1σ of the simulated value. The uncertainty estimate should be more closely identified with the probable error, defined such that 50% of the values fall within the error bounds. If a similar histogram is generated for the area parameter which has a nearly Gaussian distribution of values, the results are unchanged; Eq. 5 returns an estimate of the probable error rather than the Gaussian width, σ .

IV. Testing the Fitting Algorithm with Real Data

Thomson scattering data was acquired in the Madison Symmetric Torus (MST), a reversed field pinch plasma (7,8). Three plasma density values (~30 spectra for each case) were obtained to get Thomson scattering signal variation of ~3X. The plasma current was held fixed at ~340 kA.

The trends observed in the simulations are also seen in the fitting of real data. At low plasma density (hence low Thomson scattering signal), the maximum-likelihood method is more robust returning a narrower distribution of values for the electron temperature with fewer outliers. Since the parent distribution is unknown in the experiment and there may be significant discharge to discharge fluctuations in the parameters of the parent distribution as well as additional sources of signal contamination (especially due to *bremstrahlung* radiation at higher plasma densities), it is more difficult to assess the differences between the fitting routines with real data than it is with simulated data, hence our reliance on simulations for the critical tests of the fitting algorithms.

To correctly include background plasma light, the fit function (Eq. 3) should be modified to include *bremstrahlung* counts. Such a modification could be incorporated into the maximum-likelihood fitting method. We have not implemented such a modification because the *bremstrahlung* intensity at low plasma density is negligible (a few nanoWatts, compared to tens of nanoWatts of Thomson scattered signal). *Bremstrahlung* contamination is a more serious problem at higher density as

the continuum emission increases with n_e^2 while the TS signal increases linearly with density. At higher density (higher signal), however the choice of fitting algorithm is inconsequential as seen in simulations and with real data. The *bremsstrahlung* counts are sampled before and after the laser pulse and the average is subtracted from the signal counts before the fitting method is applied (8).

To estimate the values for the parent distribution at each density, we summed the Thomson scattering data for all the discharges and fit the accumulated spectrum. The fit parameters from the two fitting routines for this "ensemble" spectrum are in close agreement. Table 1 is a summary of the comparison of the two fitting algorithms using real data. Column 3 gives the mean of the approximately 30 temperature values returned by the maximum-likelihood method and the standard deviation. Column 4 presents the mean and standard deviation for fits returned by the χ^2 routine. At the lowest density, the maximum-likelihood fitting routine returns a distribution of temperature values with mean closer to the ensemble value and with a smaller standard deviation than the χ^2 routine. At high density, the two codes return temperature values with very similar mean and standard deviation. The trends agree with the simulation results and confirm the conclusion that for low signal level the maximum-likelihood fitting routine is more robust, returning fewer outliers, a narrower distribution of fit parameter values and a mean closer to the parent distribution value. At higher signal level, the two routines are indistinguishable and equally robust.

The estimated uncertainties for the fit parameters calculated using Eq. 5 can be compared with the mean values returned in the same manner that the was used for simulated data. The results are consistent with Fig. 4. About 50% of returned fits lie within 1σ of the mean.

V. Acknowledgments:

The authors would like to acknowledge the contributions of David Ruppert to the development of this analysis technique. This work was supported by the U.S. Department of Energy.

References:

- (1) P.R. Bevington and D.K. Robinson, Data Reduction and Error Analysis for the Physics Sciences, 2nd ed. (McGraw Hill, New York, 1992).
- (2) W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, Numerical Recipes in FORTRAN, The Art of Scientific Computing, 2nd ed. (Cambridge University Press, 1992).
- (3) J. H. M. M. Schmitt and T. Maccacaro, *Astrophys. J.* 310, 334 (1986).
- (4) G.W. Hart, F.M. Levinton, and D.H. McNeill, *Rev. Sci. Instrum.* 57, 2218 (1986).
- (5) M. Mattioli and R. Papoular, *Plasma Phys.* 17, 165 (1975).
- (6) J. Sheffield, Plasma Scattering of Electromagnetic Radiation, (Academic Press, New York, 1975).
- (7) R.N. Dexter, D.W. Kerst, T.W. Lovell, S.C. Prager, and J.C. Sprott, *Fusion Tech.* 19, 131 (1991).
- (8) D.J. Den Hartog and M. Cekic, *Meas. Sci. Technol.* 5, 1115 (1994).

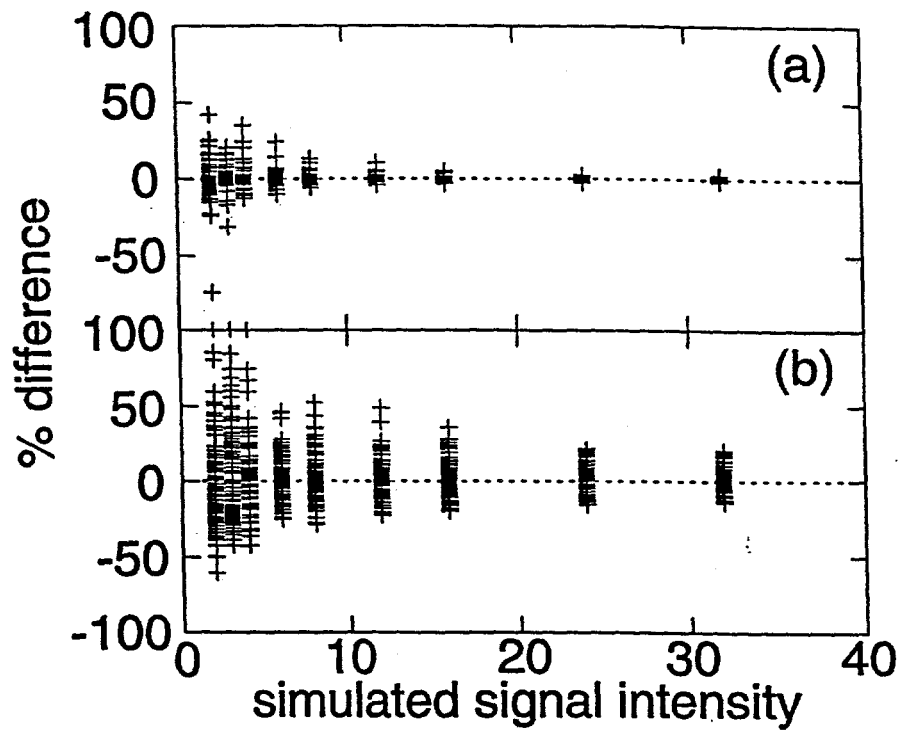


Figure 1: (a) The difference between the temperature determined by the maximum-likelihood method and χ^2 -minimization method for simulated data, (b) the difference between maximum-likelihood method and the input value for temperature (350 eV).

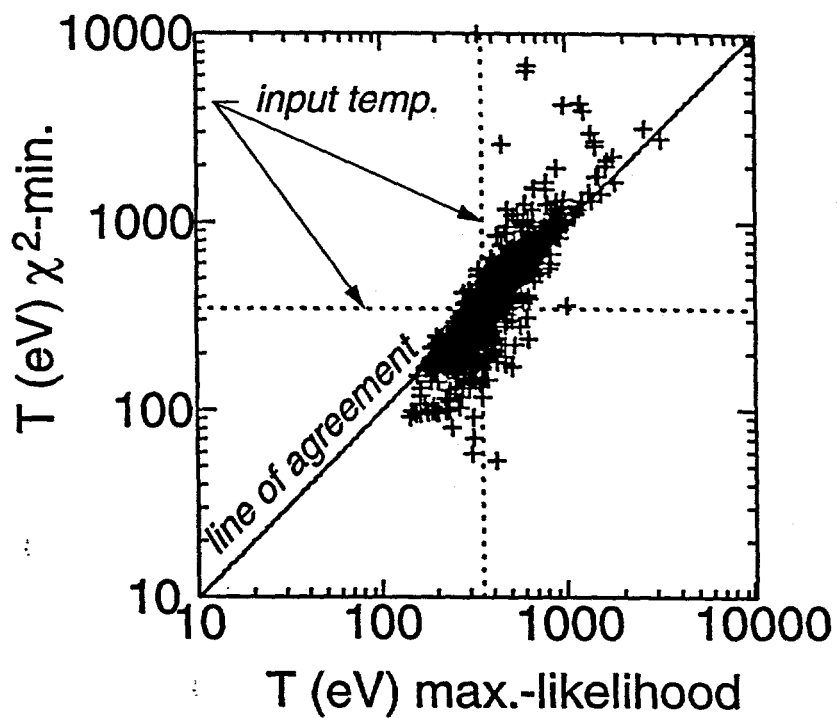


Figure 2: Fit values for temperature from maximum-likelihood and χ^2 -minimization methods plotted against each other (simulation input value is 350 eV).

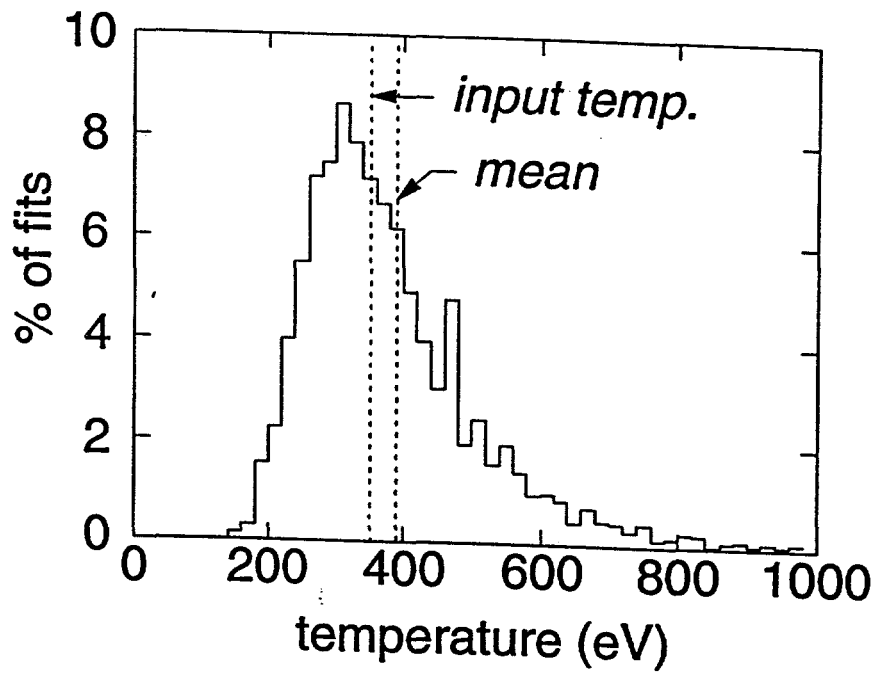


Figure 3: The distribution of values for the temperature returned by the maximum-likelihood fitting routine is asymmetric.

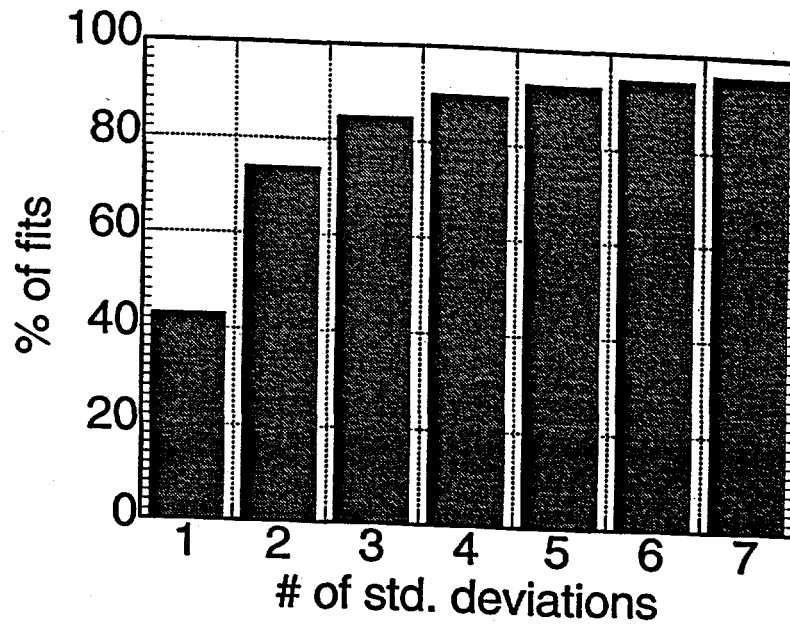


Figure 4: Histogram showing that the uncertainty estimate from Eq. 5 produces an uncertainty identified as the probable error (50% of values within one error unit).

Density (10^{19} m^{-3})	Ensemble Fit Temp. (eV)	Max.-Likelihood		χ^2 -minimization	
		mean	std. dev.	mean	std. dev.
0.5	379	405	164	428	205
1.3	289	307	121	310	137
1.8	210	223	78	218	76

Table 1: Summary of the comparison of the maximum-likelihood and χ^2 -minimization methods using real Thomson scattering data from MST.

EXTERNAL DISTRIBUTION IN ADDITION TO UC-20

S.N. Rasband, Brigham Young University
R.A. Moyer, General Atomics
J.B. Taylor, Institute for Fusion Studies, The University of Texas at Austin
E. Uchimoto, University of Montana
F.W. Perkins, PPPL
O. Ishihara, Texas Technical University
M.A. Abdou, University of California, Los Angeles
R.W. Conn, University of California, Los Angeles
P.E. Vandenplas, Association Euratom-Etat Belge, Belgium
Centro Brasileiro de Pesquisas Fircas, Brazil
P. Sakanaka, Institute de Fisica-Unicamp, Brazil
Mme. Monique Bex, GANIL, France
J. Radet, CEN/CADARACHE, France
University of Ioannina, Greece
R. Andreani, Associazione EURATOM-ENEA sulla Fusione, Italy
Biblioteca, Istituto Gas Ionizzati, EURATON-ENEA-CNR Association, Italy
Plasma section, Energy Fundamentals Division Electrotechnical Laboratory, Japan
Y. Kondoh, Gunma University, Kiryu, Gunma, Japan
H. Toyama, University of Tokyo, Japan
Z. Yoshida, University of Tokyo, Japan
FOM-Instituut voor Plasmafysica "Rijnhuizen," The Netherlands
Z. Ning, Academia Sinica, Peoples Republic of China
P. Yang, Shandong University, Peoples Republic of China
S. Zhu, University of Science & Technology of China, People's Republic of China
I.N. Bogatu, Institute of Atomic Physics, Romania
M.J. Alport, University of Natal, Durban, South Africa
R. Storer, The Flinders University of South Australia, South Australia
B. Lehnert, Royal Institute of Technology, Sweden
Librarian, CRPP, Ecole Polytechnique Federale de Lausanne, Switzerland
B. Alper, Culham Laboratory, UK
A. Newton, UK

2 for Chicago Operations Office
5 for individuals in Washington Offices

INTERNAL DISTRIBUTION IN ADDITION TO UC-20
80 for local group and file