BRAGG-PEAK LOCATION EMPLOYING A MAXIMUM-ENTROPY FORMALISM

M. Lehmann, T. Robinson, S. Wilkins

To cite this version:


HAL Id: jpa-00225824
https://hal.archives-ouvertes.fr/jpa-00225824
Submitted on 1 Jan 1986

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Bragg-peak location employing a maximum-entropy formalism

M.S. Lehmann, T.E. Robinson and S.W. Wilkins

Institut Laue-Langevin, 156X, F-38042 Grenoble Cedex, France

Abstract - The maximum entropy method has been tried on simulated data from a small 2-dimensional position-sensitive detector. Constraints were introduced to account for smoothness and the fact that only one peak was found within the frame of the recording. Analysis of a large number of weak Bragg peaks with I/σ(I) < 9 and of different size and background showed the method to give virtually bias-free results. The computing time is sufficiently low to allow real time use on measurements of a single Bragg spot with a small detector. The reliability of the method, as measured by the value of I/σ(I) where the method begins to fail in around 50% of the test cases, was found to be about 4.

I - INTRODUCTION

In the recording of a 2-dimensional scattering pattern using position-sensitive detection techniques the detector is most commonly divided (physically or logically) into N pixels of equal size, and the result of the measurement is a set of numbers \{P_j\}, where P_j is the accumulated count in pixel j. In the case of diffraction data from a single crystal the object is then to integrate the Bragg peaks present in the recording, but to do this their locations have to be identified, often from the noisy data. The first goal is therefore to find the most likely position of the Bragg peak. For a total number of accumulated counts P in the detector a large number of possible sets of \{P_j\} exist, and to find the one most likely to occur we can use an approach first developed by Boltzmann /1/.

Imagine that we are randomly putting P particles into the N pixels, one after another, until a given set of \{P_j\} counts is obtained. This can be done in W ways, where W is given by the multinomial coefficient

\[ W = \frac{P!}{P_1! P_2! \ldots P_N!} \]

The most likely set is then taken to be the one that can be obtained in most ways, i.e. the one that maximizes W, subject to whatever information, in the form of mathematical constraints, can be imposed on the P_j. This information is the

Résumé - La méthode de l'entropie maximale a été appliquée aux données d'un multidetecteur à deux dimensions. Des contraintes mathématiques ont été introduites pour imposer que les spectres ne contiennent qu'une seule raie de Bragg et que le profil de raie finale ait une forme raisonnablement lisse. La méthode a été testée sur un large nombre de spectres synthétiques avec I/σ(I) < 9. Les résultats montrent que les erreurs systématiques sont très petites, et que le temps de calcul permet un traitement en temps réel. La valeur en I/σ(I) au-dessus de laquelle plus que 50% de raies de Bragg sont observées correctement, est 4.

JOURNAL DE PHYSIQUE
Colloque C5, supplément au n° 8, Tome 47, août 1986
recording of the map, together with information about peak shapes, detector efficiency, position of powder lines, etc. If \( P \) is reasonably large*, Stirling's approximation for factorials can be used to give

\[
\ln W \approx P \cdot S = -P \sum_{j} \ln(P_j/P)
\]

An approach to the integration of the Bragg peak would thus be to use the most likely map as obtained by maximizing \( S \) on the observed data, and when the peak boundary is found from this map, to use the raw data to get the intensity.

Following thermodynamic language \( S \) is called the entropy, and the method of determining the most likely count for each pixel by maximizing \( S \) is termed the Maximum Entropy Method (MEM). Originally, the use of MEM was developed in statistical mechanics /2/. More recently, it has been employed in image reconstruction and crystallography /3/ and numerical techniques are under constant development. In this paper we report a study of the application of MEM for the peak location on a small 2-dimensional area detector. Whereas most position sensitive detectors have been designed for the recording of many Bragg reflections, construction is under way at the ILL of small area detectors (16 x 16 or 32 x 32 detector elements) aimed at the study of one or a few Bragg reflections at a given time. These detectors would mainly be used for samples with very small unit cells, and should lead both to a faster data collection rate, and a more precise knowledge of the Bragg peak(s).

II - MANIPULATION OF THE MAP

In order to maximize the \( S \), subject to a series of constraints \( \{f_k\} \) and the condition that the total number of counts is constant, we search for the maximum of the function

\[
Q = -\sum_{ij} P_{ij} \ln p_{ij} - \lambda_0 \sum_{ij} P_{ij} - \sum_{k} \lambda_k f_k
\]

with respect to \( \{p_{ij}\} \), where \( p_{ij} \) is the MEM count in pixel \((ij)\) in a 2-dimensional PSD with \( N \) pixels. \( \{\lambda_k\} \) are Lagrange undetermined multipliers, and the first of these, \( \lambda_0 \), is used to ensure that the number of counts is kept constant, i.e.

\[
\sum_{ij} p_{ij} = \sum_{ij} o_{ij} = P
\]

where \( o_{ij} \) is the recorded count.

The constraints imposed depend on the nature of the problem. In the present list there are three constraints, namely:

**Constraint 1:**

\[
f_1 = \chi^2 = \frac{1}{N} \sum_{ij} w_{ij} (p_{ij} - o_{ij})^2 \leq 1
\]

This is the most essential constraint, and ensures that \( \{p_{ij}\} \) agree with \( \{o_{ij}\} \) within the counting statistical limits. For low counts it can be difficult to estimate the weight \( w_{ij} \) for individual pixels, so we approximate \( w_{ij} \) by the mean over the detector, i.e.

\[
w_{ij} = N/\sum_{ij} o_{ij}
\]

If \( f_1 \) is larger than 1 then we have not taken full advantage of the information contained in \( \{o_{ij}\} \), while \( f_1 \) less than 1 means that the noise has been allowed to impose structure on the final map.

* 32 counts distributed on 16 pixels, e.g., would be enough!
Constraint 2: $f_2 = 1/2N \sum_{ij} w_{ij} [(p_{ij} - p_{ij-1})^2 + (p_{ij} - p_{i-1j})^2] < \delta$

We do know that the resolution function is slowly varying, and the Bragg peak is therefore smooth. Although the MEM using only constraint 1 by its nature will lead to a smooth set of $\{p_{ij}\}$, there is no direct coupling between neighbours (except via the condition that the sum of $p_{ij}$ is constant). Reducing $f_2$ therefore serves to enforce local smoothness. The quantity $\delta$ is chosen so that the mean square difference between neighbouring pixels does not exceed a certain fraction of the variance of $p_{ij}$.

Constraint 3: $f_3 = I = \sum_{ij} p_{ij} r_{ij}^2 / \sum_{ij} p_{ij}$

The sum is over points with $p_{ij}$ larger than the mean count, and $r_{ij}$ is the distance from the centre of gravity of the peak. We have in our analysis assumed that all the Bragg intensity is found on one region of the detector surface. Under these conditions the influence of small, artificial peaks can be reduced by lowering the moment of inertia $I$ (the variance of the peak distribution function). This constraint is of limited importance, serving mainly cosmetic purposes.

III - ADJUSTMENT OF $\{\lambda_k\}$

The $\{p_{ij}\}$ are determined from the $N$ equations

$$\frac{\partial q}{\partial p_{ij}} = 0$$

which can be written as

$$p_{ij} = \exp(-\lambda_0 - \sum_{k=1,3} \lambda_k \frac{\partial f_k}{\partial p_{ij}})$$

$\lambda_0$ is related to $\{\lambda_k, k=1,3\}$ by $\sum_{ij} p_{ij} = P$.

The derivatives of $f_k$ are

$$\frac{\partial f_1}{\partial p_{ij}} = 2/N w_{ij} (p_{ij} - o_{ij})$$

$$\frac{\partial f_2}{\partial p_{ij}} = 4/N w_{ij} (p_{ij} - \bar{p})$$

where $\bar{p}$ is the mean over the four nearest neighbours to $p_{ij}$.

$$\frac{\partial f_3}{\partial p_{ij}} = r_{ij}^2 / \sum_{ij} p_{ij}$$

For a given $\{\lambda_k, k=1,3\}$, the equations are solved by an iterative procedure /4/ in $p_{ij}$ and $\lambda_0$ until $\Sigma p_{ij} = P$ is reached.

The variation of $\{\lambda_k, k=1,3\}$ is done for one $\lambda_k$ at a time. First $\lambda_1$ is adjusted to get $X^2 = 1$. This is done iteratively. Improved estimates of $\lambda_1$ are obtained as $\lambda_1' = \lambda_1 - \delta$, where $X^2_{old}$ is the value of $X^2$ for $\lambda_1'$. When $X^2$ is within given limits of 1 the process is stopped.

The quantity $f_2$ is then calculated for fixed $\lambda_1'$, and if it is larger than $\delta$, ...
\( \lambda_2 \) is increased until this is no longer the case while ensuring \( \chi^2 \) stays below an upper limit. New estimates of \( \lambda_2 \) are obtained in a way similar to that for \( \lambda_1 \).

Finally \( \lambda_3 \) is estimated in a single cycle. It is obvious from the expressions for \( p_{ij} \) and \( \delta s_j / \delta p_{ij} \) that \( p_{ij} \) is most sensitive to points far from the centre of gravity. The point giving largest \( r_{ij} \)^2 \( p_{ij} \) is found, and \( \lambda_3 \) is then determined from the ratio between \( p_{ij} \) and the mean value \( \Sigma_{ij} p_{ij}/N \). It was found that setting \( \lambda_3 = p_{ij}/(\Sigma_{ij} p_{ij}/N) \) gave satisfactory results for the reduction in size of these \( p_{ij} \).

IV - CONTOURING THE MAP

The peak can now be located from the smooth map \( \{p_{ij}\} \). In exceptional cases this is done by visual inspection of the data. In the majority of cases, however, this is done by a computer algorithm, which aims at contouring the map, setting a limit between the peak and the background. A main requirement of such an algorithm is speed, and if the critical contour level is known this can easily be achieved. Unfortunately, most often this is not so, and for weak and consequently very flat peaks even small changes in the level of the limiting contour lead to large changes in the size of the peak. The other approach is to use knowledge about the size of the peak, which is either obtained from an analysis of the distribution of \( \{p_{ij}\} \) or is based on prior observations. One interesting method is to study carefully the histogram of \( \{o_{ij}\}/5 \), but again for weak peaks good results are difficult to obtain. In the present analysis we have therefore assumed that the number of points in the peak, \( M \), is known approximately, and report below some tests on how much error can be tolerated in the choice of \( M \).

The procedure is then computationally simple. The \( \{p_{ij}\} \) are sorted, and the \( M \) largest values define the peak. Some of the points located will not be connected with the main group of points. They are therefore removed using the condition that a point can only belong to the peak if two or more of its neighbours are peak points. Secondly, the edge of the peak can be somewhat ragged, and this is cured by including iteratively all points that have four or more neighbours belonging to the peak.

It is clear that a simple manipulation of peak and background points as described above is in itself a powerful method /5/, and it may be asked whether the prior smoothing using MEM adds anything to this. If we first consider constraint 1, then in our formulation it does not radically change the relative distribution of values in \( \{p_{ij}\} \). Large \( o_{ij} \) will lead to somewhat smaller \( p_{ij} \), and small \( o_{ij} \) give somewhat larger \( p_{ij} \), so whether the \( M \) largest values are obtained from \( \{o_{ij}\} \) or \( \{p_{ij}\} \) is much the same. Constraint 2, however, gives a penalty to singular points being too large, i.e. points outside the peak, but increases low values in a high count environment, and via constraint 3 any point above the mean lying far from the peak will effectively be removed from the group of the \( M \) largest points. We therefore have good reason to believe (confirmed by numerical tests), that the MEM does improve the starting values for the manipulation.

When the peak has been located, the integrated intensity \( I \) and the standard--deviation \( \sigma(I) \) are then calculated using the raw data \( \{o_{ij}\} \). By selecting at some stage the largest values in \( \{p_{ij}\} \) as being the peak, the integrated intensity does however risk having a positive bias. This bias is partially or totally removed by taking away points that obviously do not belong to the peak, and by adding points to the boundary of the peak as described above. How successful this approach is can best be determined from simulation calculations, and we have therefore undertaken the data reduction of synthetic peaks of different intensities and sizes.
Although eventually the strength of any data reduction method will be judged from its use on 'real' data, these experiences do not easily tell us how correct the intensity extraction is. We have therefore produced artificial data with known intensities and counting errors and used these to probe the method. Throughout this analysis a grid of 16 x 16 pixels was used, and the Bragg peaks were constructed as a sum of a constant background and a 2-dimensional Gaussian distribution function. The peak was assumed to cover the region defined by twice the standard deviations of the distribution. Peak intensities were chosen in the range from 0 to 180 counts, i.e. weak peaks, and the average background count in a pixel was from 2 to 4 in the case of the estimates of bias, and from 1 to 27 for the estimate of sensitivity. After the model peak had been produced the data was converted to "experimental" data using a routine by Antoniadis et al. /6/ which for a given Poisson distribution supplies individual samples. For later comparison the intensity of the "experimental" map was then calculated using the known peak location. Starting values for the calculation were the flat map, and the parameters in the data reduction program were:

\[ \chi^2 \] was initially required to fall between 0.95 and 1.05, and to stay below 1.2 when adjusting \( \lambda \).

\( \delta \) was set to \( 0.005 \times \bar{\Delta}_{ij} / N \), i.e. the mean difference between neighbouring points was approximately 7% of the standard deviation. \( \delta \) could be varied considerably without causing much change in the final results. It was found, however, that for peaks above or around 300 counts this condition was too strict, but as at present we are mainly considering weak peaks, the study of this was postponed.

To illustrate the procedure an example is shown in Fig. 1. The first plot (a) shows the model peak with background 1, intensity 150 and 69 points in the peak. The second plot (b) shows the experimental map. The third plot (c) shows \( \{ p_{ij} \} \). To obtain a better idea of \( \{ p_{ij} \} \), it was rescaled, and this is shown in the fourth plot (d). The results are encouraging.

Fig. 1. Example of Bragg peak

a. Model peak

b. Observed peak
First a test of the sensitivity was done. The procedures used were either to keep the peak count fixed and increase the background, or to keep the background fixed and vary the peak. For each case 50 frames were treated, the result was estimated by visual inspection and the percentages of success were determined. The peak determination was considered to fail if all peak points were not connected, if the peak location was wrong, or if the peak size was too large or too small. The results are plotted in Fig. 2 as a function of $I/\sigma(I)$. Open dots are for peaks with 45 points, dark points are for peaks with 69 points. It is obviously easier to determine sharp peaks (i.e. 45 points) than broad peaks, and the range of sensitivity (measured as the $I/\sigma(I)$ corresponding to 50% success) lies in the range from 3.5 to 4.5.

Fig. 2. Percentage of success in data reduction. Squares are for constant intensity (150 counts), and circles and triangles are for constant background (1 and 3 counts/pixel, respectively).
The bias was determined for peaks with sizes of 45 points and 69 points, and the results are shown in Fig. 3. Here the intensity obtained in the data reduction is plotted versus the intensity calculated with the correct boundary, and each point is the average of 100 synthetic Bragg-peaks. There is indication of a small positive bias for the larger peaks. It should be noted that \( \frac{I}{\sigma(I)} \) for the points plotted are in the range from 0 to 9, so we are dealing with low intensity peaks. In the worst case the bias is about 20% of the standard deviation \( \sigma(I) \).

Fig. 3. Bias in intensity determination.
\( \tilde{I}_{IN} \) is model intensity, \( \tilde{I}_{OUT} \) is derived intensity. Squares are for Gaussian peak with 45 points and circles are for Gaussian peak with 69 points. Open symbols are for background of 2 counts/pixel, while filled symbols are for 4 counts/pixel.

As the number of peak points, \( M \), is assumed to be known, the sensitivity of \( I \) with respect to this number was then tested by doing similar calculations varying \( M \) from 80% to 120% of its original value. The results are shown in Fig. 4. In this case we plot the least squares line going through the points for the different cases. As seen the sensitivity is small, so we have a large margin for choice of \( M \). This value would probably be estimated by visual inspection of stronger peaks.

Fig. 4 Correlation between choice of \( M \) and bias. Axis as in Fig. 3. The lines are the least squares lines through points for given value of \( M \).

VI - CONCLUSION

The maximum entropy method is a flexible tool for obtaining from a noisy measurement of a diffraction pattern a smooth map revealing the location of any diffraction peak present in the recording. Constraints based on knowledge of the nature of the Bragg peak can easily be introduced, but the computing might be heavy. In this paper we have studied the constraint of smoothness in the case where only one peak is present, and for simplicity we have only taken the case of 2-dimensional data. We find that the bias is very limited, and for the non-optimised program available at present the computing time on a VAX 750 is less than 2 sec. for a frame of 16 x 16.
data points. It should thus be possible to use the method in real time on data from small position sensitive detectors. These detectors are under construction and test, so within the next year we will be able to try the method in a routine fashion on real data.

REFERENCES