

## The Estimation of Standard Deviations in Powder Diffraction Rietveld Refinements

BY H. G. SCOTT

*CSIRO Division of Materials Science, Advanced Materials Laboratory, PO Box 4331, Melbourne, Victoria 3001, Australia*

(Received 8 June 1982; accepted 12 August 1982)

### Abstract

It is shown that when the observed and calculated intensities in a powder diffraction pattern differ because of model errors which affect the calculated integrated intensities, in addition to the random counting errors, Rietveld refinement with weights determined solely by counting statistics yields estimated standard deviations for the crystal structure parameters which tend to be measures of the precision rather than the accuracy of these parameters. Under these conditions, e.s.d.'s calculated by Rietveld refinement may be significantly smaller than those obtained by integrated intensity refinement of the same data set. A method of adjusting Rietveld e.s.d.'s, to provide comparability with integrated intensity refinement, is proposed.

### Introduction

The technique of crystal structure refinement by profile fitting, using powder diffraction data, pioneered by Rietveld (1969) is now widely used. It enables considerably more information to be obtained from overlapping peaks than does the conventional integrated intensity method which largely ignores the information contained in the peak shapes. In essence, the method involves refining parameters describing the peak shapes, unit-cell dimensions and crystal structure to obtain the best fit between the observed and calculated diffraction profiles. In the original version the background was removed graphically from the experimental data, but some current programs include the background with adjustable parameters in the calculated profile which significantly improves the accuracy of the background estimates (Prince, 1981; Toraya & Marumo, 1980).

It seems generally accepted that Rietveld refinement yields unbiased estimates for the crystal structure parameters. However, there is some ambiguity as to the physical significance of the estimated standard deviations calculated by current Rietveld refinement programs. For example, Haywood & Shirley (1977) found that it was necessary to multiply the e.s.d.'s for the crystal structure parameters by factors of 2 to 3 to

reconcile the refinements of three sets of profile data for tetraiodoethylene. Sakata & Cooper (1979) have examined the problem analytically and concluded that the correlation of errors in adjacent observations could lead to the e.s.d.'s being underestimated in Rietveld refinement: on the basis of some experimental results they suggested that a factor of at least 2 could be involved.

Subsequently, Hewat & Sabine (1981) and Prince (1981) have challenged the analysis of Sakata & Cooper and claimed that the usual Rietveld refinement programs do provide correct e.s.d.'s. However, they have demonstrated this only for the case in which counting statistics are the sole source of error, which probably arises seldom.

Pawley (1980) has asserted that 'it is probable that no published powder refinement has been limited by counting statistics'. In principle, if the weights of the observations are known, a means of testing this assertion is provided by the goodness-of-fit index,  $S^2$ , which is equal to the function minimized divided by the number of degrees of freedom in the refinement, or by the standard deviation of an observation of unit weight,  $S$ , which is equal to  $R(\text{weighted})/R(\text{expected})$ . If the weights are put equal to the reciprocals of the counting variances, and the discrepancies between the observed and calculated intensities arise solely from counting statistics,  $S$  has an expected value of  $(1 - 1/4\nu)$  and a standard deviation of  $(2\nu)^{-1/2}$ , where  $\nu$  is the number of degrees of freedom. A survey of papers published in *Acta Crystallographica B* and *Journal of Solid State Chemistry* in the last three years revealed 28 papers reporting profile refinements of crystal structures:  $S$  was reported or implied in only nine of these papers, and only for one structure was it consistent with the assumption that counting statistics were the only source of error. It seems certain therefore that in most cases the effect of other sources of error must be considered.

This paper examines the numerical values and physical significance of the e.s.d.'s of the crystal structure parameters when the structure model yields integrated intensities which are not, in the statistical sense, a good description of the profile data, as would be indicated by a value of  $S$  outside the limits implied above.

### Model errors and goodness of fit

The estimation of standard deviations in Rietveld refinement is discussed on the basis of the following assumptions.

(1) The data contains random errors arising from counting statistics.

(2) There are discrepancies between the data and the calculated intensities due to inadequacies in the model from which the intensities are calculated.

(3) For simplicity, the defects in the model affect only the integrated intensities: the background and peak shapes are well described.

(4) The theoretically resolvable Bragg peaks are fully resolved. This assumption also is made for simplicity (*cf.* Hewat & Sabine, 1981; Prince, 1981; Sakata & Cooper, 1979).

Neither of the last two assumptions is likely to be satisfied in practice, and some consequences of this will be discussed later.

On this basis, the  $j$ th ordinate of the  $k$ th peak in a profile may be written as

$$Y_{jk}(\text{obs}) = G_{jk}I_k + b_{jk} + e_{jk}, \quad (1)$$

where  $G_{jk}$  is the shape function, with  $\sum_j G_{jk} = 1$ ,  $I_k$  is the integrated intensity of the  $k$ th peak,  $b_{jk}$  is the background, and  $e_{jk}$  is the counting error, with  $E\{e_{jk}\} = 0$  and  $E\{e_{jk}^2\} \simeq Y_{jk}$ . The corresponding residual may be written as

$$\Delta Y_{jk} = Y_{jk}(\text{obs}) - Y_{jk}(\text{calc}) \simeq e_{jk} + G_{jk}\Delta I_k, \quad (2)$$

where  $\Delta I_k$  is the model-determined error in the calculated integrated intensity of the  $k$ th peak.

In Rietveld refinement the function minimized,  $M_p$ , which enters into the calculation of the e.s.d.'s is given by

$$M_p = \sum_k \sum_j w_{jk} (\Delta Y_{jk})^2 \quad (3)$$

and the observation weights,  $w_{jk}$ , are normally set equal to  $1/Y_{jk}(\text{obs})$ . As Hewat & Sabine (1981) have pointed out, there is a formal error in using  $Y(\text{obs})$  rather than  $Y(\text{expected})$  in the weighting scheme, but the effect of this approximation is usually negligible.

Combining (2) and (3), the expected value of  $M_p$  in a Rietveld refinement may be expressed as

$$E\{M_p\} = E\left\{ \sum_k \sum_j [w_{jk}e_{jk}^2 + w_{jk}G_{jk}^2(\Delta I_k)^2] \right\}. \quad (4)$$

The first group of terms in the summation is the contribution from counting statistics, and is approximately equal to  $N - P_p$ , where  $N$  is the number of profile points and  $P_p$  is the number of parameters describing the profile, *i.e.* all those parameters in the model which do not directly affect the integrated intensities. The second group of terms arises from the

defects in the model for the integrated intensities. There are no cross terms in the summation since  $e_{jk}$  and  $\Delta I_k$  are independent. There is a formal similarity between this equation and Sakata & Cooper's (1979) equation (42).

Provided that the intensity in the background is small compared to that in the peaks

$$\sum_j w_{jk}G_{jk}^2 \simeq (1 - B_k/I_k)/I_k, \quad (5)$$

where  $B_k$  is the integrated background under the  $k$ th peak, since  $w_{jk} = 1/Y_{jk}(\text{obs})$ ,  $\sum_j G_{jk} = 1$  and  $G_{jk} = [Y_{jk}(\text{obs}) - (b_{jk} + e_{jk})]/I_k$  from (1). For small backgrounds, the right-hand side of (5) may be written as  $1/(I_k + B_k)$  to terms of first order in  $B_k/I_k$  (this restriction to small backgrounds may be unnecessarily severe, since it is easily shown that for the rather artificial case of a rectangular peak superimposed on a uniform background the approximation becomes almost exact, regardless of the ratio of background to peak area), so that

$$E\{M_p\} = N - P_p + \sum_k (\Delta I_k)^2 / (I_k + B_k). \quad (6)$$

Superficially, the summation over  $k$  is just  $M_I$ , the function minimized in an integrated intensity refinement, since  $1/(I_k + B_k)$  is the observation weight in this case if the background is estimated with negligible variance [*cf.* Sakata & Cooper (1979), equation (16)]. However, because of the way in which  $\Delta I_k$  has been defined here, it is actually the limit of that function in the absence of counting errors. When the contribution from counting errors is included the expected value of  $M_I$  for the data set is given by

$$E\{M_I\} = N_B - P_C + \sum_k (\Delta I_k)^2 / (I_k + B_k), \quad (7)$$

where  $N_B$  is the number of Bragg peaks and  $P_C$  is the number of crystal structure parameters.

At this stage it is informative to consider plausible numerical values for the goodness-of-fit index  $S^2$ . For example, consider an integrated intensity refinement with  $N_B - P_C = 50$  for which  $S^2 = 3.8$ : such a value for  $S^2$  is not atypical and would be unlikely to cause concern, even though it clearly indicates that the model is not a good fit to the data in the statistical sense. If the same data were analysed by Rietveld refinement with  $N - P = 2000$  (say), the expected value for  $S^2$  would be about 1.07: this value is still outside the acceptable range for  $S^2$  for a good model, but much less strikingly so than in the integrated intensity case. In short, if the number of observations is much larger than the number of peaks, as is invariably the case in Rietveld refinement, the goodness-of-fit index may be rather insensitive to errors in the model for the intensities.

### Calculation of estimated standard deviations

In a least-squares refinement, the e.s.d. of the  $i$ th parameter is calculated from the expression

$$\sigma^2(p_i) = A_{ii}^{-1} M / (N - P) = A_{ii}^{-1} S^2, \quad (8)$$

where  $A_{ii}^{-1}$  is the appropriate diagonal element of the inverse of the matrix of the normal equations. For this direct matrix

$$A_{mn} = \sum_k \sum_j w_{jk} \frac{\partial Y_{jk}}{\partial p_m} \frac{\partial Y_{jk}}{\partial p_n}. \quad (9)$$

Hewat & Sabine (1981) have shown that, in the absence of background, the matrix elements which relate to the crystal structure parameters are the same for Rietveld and for integrated intensity refinements: consequently the ratio of the e.s.d.'s determined by the two methods is equal to the square root of the ratio of the goodness-of-fit indices.

For a profile refinement,  $w_{jk} \propto 1/t$ ,  $Y_{jk} \propto t$  (where  $t$  is the counting time for a single profile point) and there are  $N$  terms in the summations for the matrix elements. Consequently, if the profile points are reasonably closely spaced in  $2\theta$ ,  $A_{mn}$  is proportional to  $T$  (the total counting time for the entire profile), but does not depend on  $N$  as such if  $T$  is unchanged. Thus

$$A_{ii}^{-1}(T) = [T_0/T] A_{ii}^{-1}(T_0) \quad (10)$$

and similarly

$$\sum_k (\Delta I_k)^2 / (I_k + B_k) = [T/T_0] Q, \quad (11)$$

where  $Q$  is the value of the summation for some arbitrary time,  $T_0$ , and is a measure of the non-statistical discrepancy between the data and the model intensities. Since  $N$  is large compared to  $P$ , (6), (8), (10) and (11) may be combined to yield

$$E\{\sigma^2(p_i)\} = A_{ii}^{-1}(T_0) [T_0/T + Q/N]. \quad (12)$$

Clearly then, by sufficiently increasing  $N$  and  $T$ , the standard deviations in a Rietveld refinement may be made arbitrarily small, regardless of the value of  $Q$ , *i.e.* even if the model provides a poor fit to the data. This apparent absurdity arises because the e.s.d.'s calculated by Rietveld refinement in this limit are measures of the precision of the parameters, which in the presence of model errors give little information about the accuracy of the parameters.

The physically unsatisfactory nature of the e.s.d.'s calculated in a Rietveld refinement may also be seen as follows. From a data set collected at steps of  $0.05^\circ$  (say) in  $2\theta$ , a second set may be derived by adding the original data points in pairs, so that it contains only half the number of observations. As far as integrated intensities are concerned, the contracted data set is essentially as good as the original set, and should yield the same e.s.d.'s for the crystal structure parameters if

Table 1. Variation of some figures of merit as a function of the number of observations in a Rietveld refinement, for constant total counting time

$\beta$ -PbO<sub>2</sub> sample: neutron diffraction data.

	Initial data set	Contracted data set	Ratio of contracted/initial
$N$	2641	1321	0.5
$N_B$	65	65	1.0
$P_p$	10	10	1.0
$P_c$	5	5	1.0
$R_e^*$	17.80	12.67	0.712
$R_{wp}^\dagger$	23.98	20.35	0.849
$S_p^\ddagger$	1.347	1.606	1.192
$\sigma[B(\text{Pb})]_\S$	0.0355	0.0424	1.194

$$*R_e = 100[(N - P_p - P_c) / \sum_i w_i (Y_i - b_i)^2]^{1/2}.$$

$$^\dagger R_{wp} = 100[\sum_i w_i (\Delta Y_i)^2 / \sum_i w_i (Y_i - b_i)^2]^{1/2}.$$

$$^\ddagger S_p = R_{wp}/R_e.$$

$\S$ The other e.s.d.'s could not usefully be compared because of limitations imposed by the program output format.

these are to be physically meaningful. However, (12) predicts that the e.s.d.'s calculated from the contracted data will be increased by a factor of between 1 and  $\sqrt{2}$ , depending on the value of  $Q$ . This prediction has been verified experimentally by Hill (1982) and his results are presented in Table 1.

Bacon & Lisher (1980) have shown that deleting every alternate point from a profile data set collected at steps of  $0.05^\circ$  increased the Rietveld e.s.d.'s by almost exactly  $\sqrt{2}$ . This is just the increase predicted by (12), and indicates that there is no significant loss of information attributable to the greater step width as such.

### Discussion

In considering the reliability of a physical measurement, it is important to distinguish between precision and accuracy. Precision refers to the statistical variation likely to occur in the measurement, and this can be determined by repeated measurements. Accuracy, on the other hand, refers to the difference between the measured value of a quantity and its 'true' value, assuming that a true value exists. The accuracy can never be greater than the precision, and may well be much less if the measurements are subject to uncorrected systematic errors. A similar distinction applies to parameters estimated by the method of least squares, and only if the model is good and complete (which implies *inter alia* that any systematic errors in the observations are represented by appropriate parameters in the model) are the statistical precision and physical accuracy the same.

When the model used in a least-squares refinement does contain errors, the calculated e.s.d.'s for the parameters depend on both the statistical and the model errors, but the relative importance accorded to

the two types of error may depend on the way in which the data is handled. Rietveld refinement places more emphasis on the statistical errors than does integrated intensity refinement, as may be seen by comparison of (6) and (7). Consequently, it yields e.s.d.'s which reflect the precision rather than the accuracy of the model parameters. Furthermore, when the number of profile points is large, the goodness-of-fit index ceases to be an efficient indicator of model errors.

Nevertheless, it must be emphasized that this uncertainty about the exact significance of the e.s.d.'s does not detract seriously from the fundamental merit of the Rietveld refinement technique. The parameter estimates themselves are unbiased, and almost certainly more accurate than those that would be obtained by the integrated intensity method, since profile methods make use of information contained in the peak shapes. Moreover, in many cases it is the values of the structure parameters which are of primary interest, with the estimates of accuracy being of lesser interest.

However, there are cases in which the accuracy of the structure parameters is significant, and in these cases some modification of current profile techniques may be required. One approach is to analyse profile data in two stages (Cooper, Rouse & Sakata, 1981; Pawley, 1981; Will, 1981): in the first stage numerical values for the integrated intensities are obtained from the profile data, without reference to the crystal structure parameters, while the second stage is a conventional structure refinement using the integrated intensities obtained in the first stage. Alternatively, *ad hoc* adjustments to the e.s.d.'s may be made. Pawley (1980) suggests that in calculating e.s.d.'s by the Rietveld method the actual number of observations should be divided by  $\Delta$ , the number of observations in the average full width at half height of well resolved peaks. This has the effect of multiplying all the e.s.d.'s by  $\sqrt{\Delta}$ , which is typically in the range 2 to 3. While this may be reasonable in practice, it is theoretically unsound to apply such a correction without regard to how well the model fits the data.

A different method of adjustment is to convert Rietveld e.s.d.'s into equivalent integrated intensity e.s.d.'s, on the grounds that these better reflect the effect of model errors and are widely accepted by crystallographers as measures of accuracy, even if the statistical basis for this assumption is dubious. Equations (6), (7) and (8) suggest a means of making the conversion, which requires  $S_p^2 = M_p/(N - P)$  to be replaced by  $S_I^2 = M_I/(N_B - P_C) = 1 + (M_p - N + P_p)/(N_B - P_C)$  in calculating the profile e.s.d.'s of the crystal structure parameters. Since this adjustment is based on the assumption that only the integrated intensities are subject to model errors it should not be applied to the lattice, background or peak-shape parameters. The adjustment can be applied to published structure data provided that the goodness-of-fit index or equivalent

Table 2. *Adjustment factors for structure-parameter e.s.d.'s for some published Rietveld refinements*

	LaTaO <sub>4</sub> <sup>a</sup>	CeTaO <sub>4</sub> <sup>b</sup>
$N$	3816	1277
$N_B$	167*	290
$P_p$	17	8
$P_C$	19	25
$R_e$	4.49	8.08
$R_{wp}$	8.56	10.04
whence		
$S_p$	1.906	1.234
$S_I$	8.256	1.860
and		
$S_I/S_p$	4.332	1.497

References: (a) Cava & Roth (1981); (b) Santoro, Marezio, Roth & Minor (1980).

\* Calculated from the lattice parameters and scan range of the profile.

information is available: the other figures required are usually given, or can be inferred with fair accuracy.

When the suggested procedure is applied to the data in Table 1 it yields adjustment factors,  $S_I/S_p$ , of 4.488 and 3.699 for the full and contracted data sets respectively: the adjusted values of the e.s.d. for the isotropic temperature factor of the lead atom,  $\sigma[B(\text{Pb})]$ , are 0.159 and 0.157, so that the adjustment is at least internally consistent. Other examples of the procedure are illustrated in Table 2.

The adjustments calculated above are in broad agreement with the *ad hoc* adjustments suggested previously, but take specific account of the goodness-of-fit index for the structure refinement being considered. However, they are still only approximations, since in deriving the form of the adjustment it has been assumed that the model for the peak shape is good. This is seldom true, particularly for X-ray data, so that  $S_p^2$  should be reduced by an uncertain amount corresponding to the peak-shape error. On the other hand, it has been assumed that the Bragg peaks are completely resolved, which is also seldom true. Consequently,  $N_B$  should be replaced by the smaller number of actually resolved peaks, though this is ill-defined when partial overlap of peaks occurs. Notwithstanding these criticisms, it is considered that the procedure above imposes reasonable restraints on the uncritical use of the e.s.d.'s generated by Rietveld refinements as measures of the accuracy of the refined structure parameters.

While the general conclusions above also apply when profile data is collected photographically, the suggested adjustments can seldom be applied. This is a consequence of the difficulty in determining absolute statistical weights when the observations are optical densities, rather than numbers of counts obeying the Poisson distribution. Without these absolute weights, the goodness-of-fit index is as much a scale factor to be

applied to the weights as it is a test for the adequacy of the model.

The author wishes to thank Drs A. W. Hewat, R. J. Hill, J. K. Mackenzie and E. Prince for helpful correspondence and discussions.

#### References

- BACON, G. E. & LISHER, E. J. (1980). *Acta Cryst.* **B36**, 1908–1916.
- CAVA, R. J. & ROTH, R. S. (1981). *J. Solid State Chem.* **36**, 139–147.
- COOPER, M. J., ROUSE, K. D. & SAKATA, M. (1981). *Z. Kristallogr.* **157**, 101–117.
- HAYWOOD, B. C. & SHIRLEY, R. (1977). *Acta Cryst.* **B33**, 1765–1773.
- HEWAT, A. W. & SABINE, T. M. (1981). NSW Institute of Technology, Internal Report NSWIT/PMN 24.
- HILL, R. J. (1982). Private communication.
- PAWLEY, G. S. (1980). *J. Appl. Cryst.* **13**, 630–633.
- PAWLEY, G. S. (1981). *J. Appl. Cryst.* **14**, 357–361.
- PRINCE, E. (1981). *J. Appl. Cryst.* **14**, 157–159.
- RIETVELD, H. M. (1969). *J. Appl. Cryst.* **2**, 65–71.
- SAKATA, M. & COOPER, M. J. (1979). *J. Appl. Cryst.* **12**, 554–563.
- SANTORO, A., MAREZIO, M., ROTH, R. S. & MINOR, D. (1980). *J. Solid State Chem.* **35**, 167–175.
- TORAYA, H. & MARUMO, F. (1980). Report of the Research Laboratory of Engineering Materials, Tokyo Institute of Technology, Vol. 5, pp. 55–64.
- WILL, G. (1981). *Acta Cryst.* **A37**, C274.