(5)] was found to provide only marginal, if any, improvement in the B estimate, and to have a number of drawbacks. The concept of Debye-curve inflexion points is introduced and a straightforward and relatively robust method for improving the least-squares process, based on predictable features of a Debye curve, is described. Values of B estimated by the inflexion-point method are, on average, 10% better than those calculated by conventional methods.

The authors wish to acknowledge the assistance of the Australian Research Grants Committee (Grant: C7915302) during the tenure of this work.

References

DAVIS, C. L., MASLEN, E. N. & VARGHESE, J. N. (1978). *Acta Cryst.* A**34**, 371–377.

DEBYE, P. (1915). *Ann. Phys. (Leipzig)*, **46**, 809.

DECLERCQ, J. P., GERMAIN, G. & VAN MEERSSCHE, M. (1972). *Cryst. Struct. Commun.* **1**, 13–15.

FRENCH, S. & WILSON, K. (1978). *Acta Cryst.* A**34**, 517–525.

HALL, S. R. (1972). XRAY72 system of crystallographic programs : program *NORMSF* edited by J. M. STEWART *et al.* Tech. Rep. TR192, Computer Science Center, Univ. of Maryland, College Park, Maryland.

HALL, S. R. (1978). *Acta Cryst.* A**34**, S348.

HALL, S. R. (1981). *GENEV:* Generation of E values. *The XTAL system of crystallographic programs.* Tech. Rep. TR873. Computer Science Center, Univ. of Maryland, College Park, Maryland.

HALL, S. R., RASTON, C. L. & WHITE, A. H. (1978). *Aust. J. Chem.* **31**, 685–688.

HALL, S. R., STEWART, M. J. & MUNN, R. J. (1980). *Acta Cryst.* A**36**, 979–989.

HENDRICKSON, W. A. (1980). Personal communication.

ISAACS, N. W. & AGARWAL, R. C. (1978). *Acta Cryst.* A**34**, 782–791.

LADD, M. F. C. (1978). *Z. Kristallogr.* **147**, 279–296.

MAIN, P. (1976). *Crystallographic Computing Techniques*, edited by F. R. AHMED, pp. 97–105. Copenhagen; Munksgaard.

MAIN, P., FISKE, S. J., HULL, S. E., LESSINGER, L., GERMAIN, G., DECLERCQ, J. P. & WOOLFSON, M. M. (1980). *MULTAN*80. Univ. of York.

SKELTON, B. W. & WHITE, A. H. (1981). Personal communication.

SUBRAMANIAN, V. & HALL, S. R. (1982). *Acta Cryst.* A**38**, 577–590.

TEETER, M. M. & HENDRICKSON, W. A. (1979). *J. Mol. Biol.* **127**, 219–223.

WATENPAUGH, K. D., SIEKER, L. C., HERRIOTT, J. R. & JENSEN, L. H. (1973). *Acta Cryst.* B**29**, 943–956.

WILSON, A. J. C. (1942). *Nature (London)*, **150**, 151.

Acta Cryst. (1982). A**38**, 598–608

Normalized Structure Factors. III. Estimation of Errors

BY S. R. HALL AND V. SUBRAMANIAN†

Crystallography Centre, University of Western Australia, Nedlands 6009, *Australia*

(*Received* 6 *May* 1981; *accepted* 20 *January* 1982)

Abstract

A method for calculating the expected errors in $|E_h|$ values is outlined. It is based on the precision of the measured data and the Wilson-plot parameters; and allows for errors arising from the use of the profile scaling function and/or the index rescaling procedure in the normalization scheme. Six refined structures are used to test the estimated errors in $|E_h|$ against values deduced from a comparison with the 'true' normalized structure factor $|\mathscr{E}_h|$.

Introduction

One of the most serious obstacles to structure solution by statistical invariant methods is the sensitivity of all phasing procedures to errors in the initial phase relationships. The generation of a single incorrect phase in the early stages of a phasing procedure can often result in the failure of the entire process. For this reason computer programs place a strong emphasis on the choice of initial starting phases and on the order in which the invariants are processed.

There are a number of different approaches to the selection of starting phases but all of them depend on one fundamental quantity, namely, the magnitude of

† Deceased 27 December 1981.

0567-7394/82/050598-11$01.00

the normalized structure factor $|E_h|$. It is therefore of particular concern that the average precision of estimated $|E_h|$ values from existing computer programs is low (Hall & Subramanian, 1982). Nevertheless, the success ratio of direct-methods procedures using these $|E_h|$ values is relatively high and this seems to suggest that the precision of $|E_h|$ values is not an important factor in the application of structure-invariant relationships. This conclusion, however, is questionable because the sensitivity of phase initialization procedures to relatively small changes in $|E_h|$ values is well known. Changes in $|E_h|$ values of a few percent can cause the structure-invariant relationships to be processed in an entirely different order; or, rather, the phasing process to follow a different 'phase path'. Fortunately, more than one phase path can lead to a correct solution. In fact, the success of the 'multi-solution' approach employed in many phasing procedures relies on the variations in the phase path caused by a permutation of phases. A multi-solution procedure repeats the phase generation process for each permutation of the starting phases, and this has the effect of buffering the phasing process against the need to identify the most reliable phase path. It should be remembered, however, that much of this repetition is also necessary because of the fragility of the initial phase selection process itself, and a consequent need to increase the number of starting phases.

The reliability of structure-invariant relationships decreases as a function of the atomic content of the unit cell. For structures with molecular weights in excess of about 750 daltons, the average reliability of phase relationships decreases to the point where the overall effectiveness and the practicality of existing multi-solution procedures become seriously limited. In these cases it is particularly important that the Monte Carlo aspects of the multi-solution approach are minimized, and that a more rigorous treatment of phase probabilities is pursued. Subramanian & Hall (1982) have shown that improved estimates of $|E_h|$ provide commensurate improvements in the reliability of phase relationships. It follows that if a measure of the precision of each $|E_h|$ value in terms of its expected error can be obtained, this should lead to a more correct estimate of the phase probability. At the very least it will provide a better statistical foundation for the crucial phase selection process.

Measured errors in $|E_h|$ estimates

The study of normalization scaling functions by Ladd (1978) and Subramanian & Hall (1982) highlight the importance of the calculated 'true' normalized structure factor $|\mathscr{E}_h|$ as a measure of the precision of estimated $|E_h|$ values. These studies show that $|\mathscr{E}_h|$ values provide a measure for assessing the errors in the different estimates of $|E_h|$, and their relative dependence on parameters such as s^2, $|\mathscr{E}_h|$, and $|F_h|$. In this study the values of $|\mathscr{E}_h|$ and $|E_h|$ are compared for six refined structures (see Table 1), and are used to estimate the errors inherent in each estimate of $|E_h|$.

Table 1. *Test structures*

$R = \sum||F_o| - |F_c||/\sum|F_o|$.
$\bar{B}$ is the overall temperature factor.

	Formula	Space group	R value	$\bar{B}$ (Å^2)	s^2_{max}	Reference
BEKA4	$C_{58}H_{90}N_2O_6$	$P\bar{1}$	0·055	4·5	0·24	(*a*)
CANON2	$C_{18}H_{18}O_5$	$P2_1/n$	0·058	3·8	0·24	(*b*)
ANTH1	$C_{34}H_{26}O_4$	$P2_1/c$	0·034	4·5	0·36	(*a*)
CORT	$C_{21}H_{28}O_5$	$P2_12_12_1$	0·058	3·3	0·32	(*c*)
K22BR	$C_{35}H_{48}O_6$	$Iba2$	0·049	4·8	0·22	(*a*)
KCPP	$C_{16}H_{19}KO_{11}$	$Pcab$	0·042	3·2	0·36	(*a*)

References: (*a*) Skelton & White (1981); (*b*) Hall, Raston & White (1978); (*c*) Declercq, Germain & Van Meerssche (1972).

Fig. 1 displays the variation of $|\mathscr{E}_h|$ and $|E_h|_n$ with s^2 for the test structures. In these plots the data are averaged into 41 ranges between 0·0 and s^2_{max} and have been overlapped to ensure that a minimum of 200 reflections contributed to the mid-range. The average difference between $|\mathscr{E}_h|$ and $|E_h|_n$, ΔE_n, is shown as a function of s^2, $|F_h|$, and $|\mathscr{E}_h|$ in Figs. 2(*a*), 3(*a*) and 4(*a*), respectively. All plots in this study were prepared using the program *ESCAN* (Hall & Subramanian, 1980).

The estimated $|E_h|$ values contain error contributions from at least three main sources; (1) the systematic difference between the mean $|\mathscr{E}_i|$ and the mean $|E_i|$; (2) the mean random error of $|\mathscr{E}_h|$ about the mean $|\mathscr{E}_i|$; and (3) the mean random error of $|E_h|$ about the mean $|E_i|$. The mean error for the ith range, ΔE_i, may be expressed as

$$\Delta E_i = \big||\mathscr{E}_i| - |E_i|\big| + (\sigma^2|\mathscr{E}_i| + \sigma^2|E_i|)^{1/2}. \qquad (1)$$

The variation of the systematic error difference $\big||\mathscr{E}_i| - |E_i|\big|$ with s^2 may be gauged from Fig. 1. This shows relatively small differences between $|\mathscr{E}_h|$ and the estimates derived from the exponential scale $\mathbf{k}\exp(Bs^2)$ with overall and index rescaling ($|E_h|_1$ and $|E_h|_3$). The differences for the E values calculated using the profile scale ($|E_h|_2$ and $|E_h|_4$) are, however, a significant fraction of ΔE_i, as a function of both s^2 and $|\mathscr{E}_h|$.

It is evident from Figs. 1 and 2 that the principal contributors to the ΔE values, however, are the random errors associated with the distribution of $|E_h|$ about E_i and $|\mathscr{E}_h|$ about $|\mathscr{E}_i|$. Figs. 2(*a*) and 4(*a*) indicate that the magnitude of the random errors increases both as a function of s^2 and as a function of $|\mathscr{E}_h|$ itself. Figs. 5(*a*) and 6(*a*) show the variation of $\Delta E_n/|E_n|$ with s^2 and $|\mathscr{E}_h|$ and serve to illustrate the variation of the errors as fractional differences.

Estimated errors of $|E_h|$

The estimation of errors in a normalized structure factor must take into account both the errors in the parameters used in the calculation of $|E_h|$ and the uncertainty in the Wilson-plot process itself. The normalized structure factor is calculated from the expression

$$|E_h| = \frac{|F_h^m| \mathbf{k} \exp(Bs^2)}{\langle |F_h^2| \rangle^{1/2}}. \tag{2}$$

Assuming that the expectation value $\langle |F_h^2| \rangle$ is derived directly from the random-atom scattering term $\varepsilon \sum f^2$, then the variance of $|E_h|$ may be expressed as

$$\sigma^2|E_h| = |E_h^2| \left\{ \frac{\sigma^2|F_h|}{|F_h^2|} + \frac{\sigma^2 \exp(Bs^2)}{\exp(2Bs^2)} + \frac{2\sigma^2[\mathbf{k}, \exp(Bs^2)]}{\mathbf{k} \exp(Bs^2)} \right\}, \tag{3}$$

where the $\sigma^2[\mathbf{k} \exp(Bs^2)]$ is the covariance of $\mathbf{k}$ and $\exp(Bs^2)$. The third term in (3) may be expanded as

$$\sigma^2 \exp(Bs^2) = s^2 \exp(Bs^2)\, \sigma^2 B. \tag{4}$$

Similarly, the covariance term is expanded in terms of the correlation coefficient as

$$\sigma^2(P, Q) = r(P, Q)\, \sigma P \sigma Q. \tag{5}$$

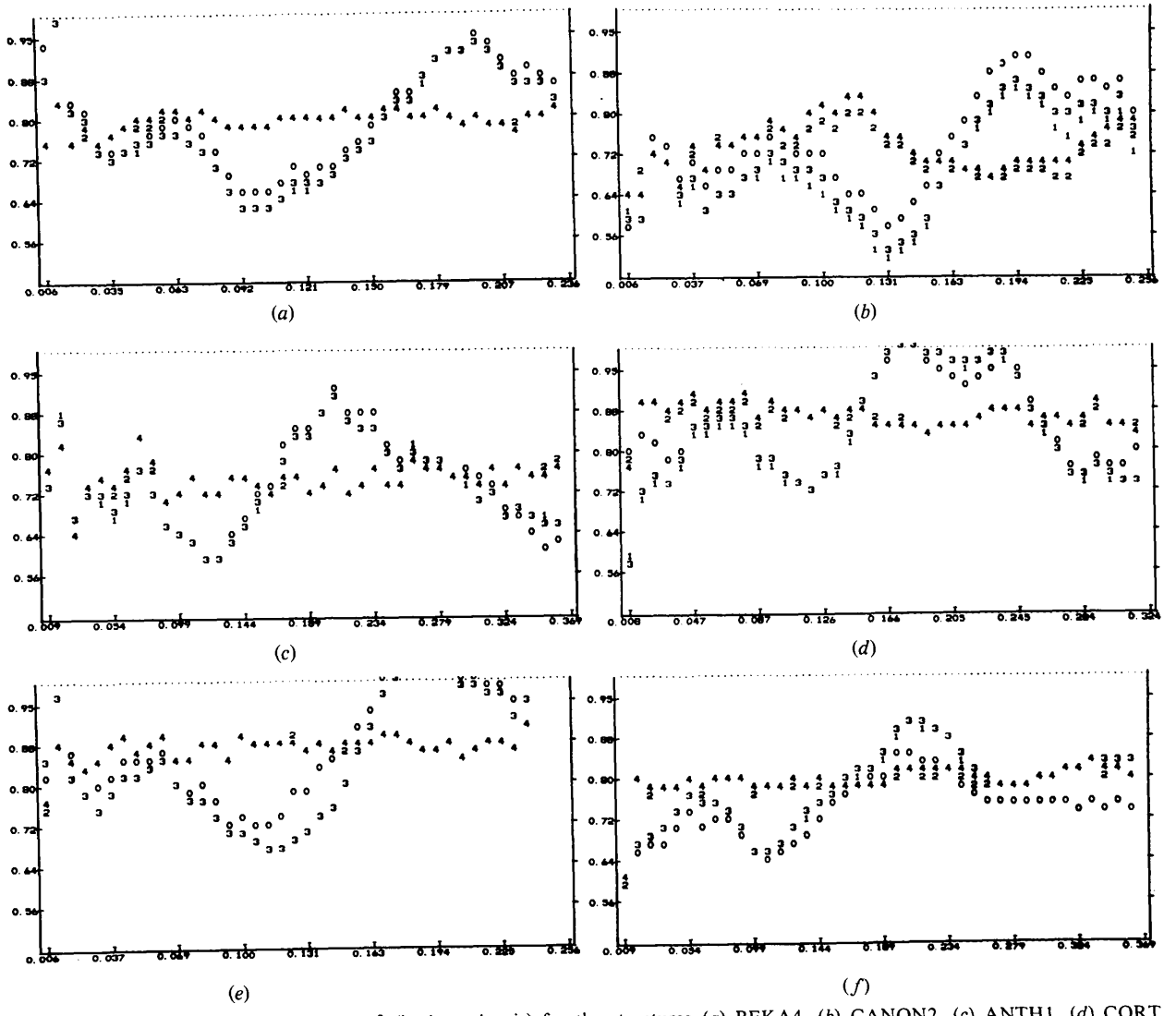


Fig. 1. Plots of $|\mathscr{E}_h|$ and $|E_h|_n$ *versus* s^2 (horizontal axis) for the structures (*a*) BEKA4, (*b*) CANON2, (*c*) ANTH1, (*d*) CORT, (*e*) K22BR and (*f*) KCPP. The complete description of $|\mathscr{E}_h|$ (plotted as 0) and the four $|E_h|_n$ values (plotted as 1, 2, 3, 4) is given by Subramanian & Hall (1982).

Equations (4) and (5) having been substituted in (3), the variance becomes

$$\sigma^2|E_h| = |E_h^2|\left\{\frac{\sigma^2|F_h|}{|F_h^2|} + \frac{\sigma^2\,\mathbf{k}}{\mathbf{k}^2} + s^4\,\sigma^2 B + \frac{2s^2}{\mathbf{k}}\,r(\mathbf{k},B)\,\sigma\mathbf{k}\sigma B\right\}. \quad (6)$$

The variance equation (6) contains the parameters s, $|F_h|$ and σF_h derived from the measurement process; $\mathbf{k}$, B and $|E_h|$ derived from the Wilson-plot calculation, and $\sigma\mathbf{k}$, σB and $\sigma(\mathbf{k},B)$ which can be obtained from the Wilson-plot least-squares procedure (*vide infra*).

The essential relationships on which the Wilson-plot process is based may be written as

$$\ln\frac{|F_h^2|}{\langle|F_h^2|\rangle} = -2\ln\mathbf{k} - 2Bs^2. \quad (7)$$

The plot of $\ln(|F_h^2|/\langle|F_h^2|\rangle)$ *versus* s^2 is a straight line with $-2B$ as the slope and $-2\ln(\mathbf{k})$ as the intercept at $s^2 = 0$. In practice the slope and intercept are obtained by the application of linear least squares to the χ^2 term,

$$\chi^2 = \sum_i^n w_i(R_i - C - Ss_i^2)^2, \quad (8)$$

so that it is a minimum for the slope S and the intercept C. R_i is the logarithm of the ratio of $|F_h^2|$ and $\langle|F_h^2|\rangle$, n is the number of data points in the summation and w_i is the least-squares weight of each point.

Estimates for the variances σ^2C, σ^2S, $\sigma^2(C,S)$ and the correlation coefficient $r(C,S)$ are available directly from the least-squares process (see Appendix for details). Their relationships to the variances in (6) are as follows:

$$\sigma^2 C = \sigma^2(2\ln\mathbf{k}) = 2\sigma^2\mathbf{k}/\mathbf{k}^2 \quad (9)$$

$$\sigma^2 S = \sigma^2(2B) = 2\sigma^2 B \quad (10)$$

and

$$\sigma^2(C,S) = r(C,S)\,\sigma C\sigma S = r(C,S)\frac{2\sigma\mathbf{k}}{\mathbf{k}}\sigma B. \quad (11)$$

Substituting (9), (10) and (11) into (4), and assuming that the linear correlation factor $r(C,S)$ is equivalent to $r[\mathbf{k},\exp(Bs^2)]$, one obtains

$$\sigma^2|E_h| = |E_h^2|\left\{\frac{\sigma^2|F_h|}{|F_h^2|} + \frac{\sigma^2 C}{2} + \frac{s^4\,\sigma^2 S}{2} + s^2\,\sigma^2(C,S)\right\}. \quad (12)$$

Profile scale error contribution

The variance $\sigma^2|E_h|$ expressed by (12) excludes the non-random errors of the type discussed in connection with the measured errors derived from (1). The derivation of the variance $\sigma^2|E_h|$ ignores any *systematic* differences between the estimated $|E_h|$ value and the true normalized structure factor. This is valid only if the mean $|E_h|$ is the same as the mean $|\mathscr{E}_h|$ for all values of s^2, $|F_h|$ and $|\mathscr{E}_h|$. The plots in Fig. 1 indicate that this is a reasonable approach for the estimate calculated from the exponential scaling function and overall rescale, $|E_h|_1$; and may also be legitimate for the estimate based on the exponential scale and index rescale, $|E_h|_3$, despite the presence of non-random errors due to the index rescale (*vide infra*). Fig. 1 also clearly shows that there is a significant difference between the mean $|\mathscr{E}_h|$ and the estimates based on the profile scale, $|E_h|_2$ and $|E_h|_4$. These systematic differences must be included in any estimation of errors in $|E_h|$. The error contribution due to the profile scaling function $\mathbf{K}(s)$ can be evaluated from its difference from the linear scale as

$$\Delta^2\,\mathbf{K}(s) = [\mathbf{k}\exp(Bs^2) - \mathbf{K}(s)]^2. \quad (13)$$

Index rescale error contribution

The analysis of overall agreement between the estimated $|E_h|_3$ values and calculated $|\mathscr{E}_h|$ values indicates that the index rescaling is also a potential source of systematic error. This error will be smaller than the profile scale contribution except in special cases. The non-random errors due to index rescaling will also tend to be less conspicuous than errors from other sources due to its dependency on combinations of h, k, l rather than parameters such as s^2, $|F_h|$ or $|\mathscr{E}_h|$.

A suitable approximation of the systematic errors due to the index rescale value $\mathbf{K}(hkl)$ may be calculated from its difference with the overall rescale value $\mathbf{k}$,

$$\Delta^2\,\mathbf{K}(hkl) = [\mathbf{k} - \mathbf{K}(hkl)]^2. \quad (14)$$

Comparison of measured and estimated errors

The difference between the calculated normalized structure factor $|\mathscr{E}_h|$ and the estimated normalized structure factor $|E_h|$ provides information on both the random and systematic errors in the various estimates of $|E_h|$. The use of (1) for this purpose presupposes that $|\mathscr{E}_h|$ is a reasonable measure of the 'true' normalized structure factor and that the random errors associated with $|\mathscr{E}_h|$ are relatively small. A study of Subramanian & Hall (1982) on the reliability of $|\mathscr{E}_h|$ and $|E_h|$ in phasing procedures supports this and it is reasonable to assume that the contribution of $\sigma^2|\mathscr{E}_h|$ to (1) is relatively small at the level of refinement reached for the test structures (see Table 1). The plots of the

ΔE_i against s^2, $|F_h|$ and $|\mathscr{E}_h|$ for the test structures shown in Figs. 2(*a*), 3(*a*) and 4(*a*),* are expected therefore to be a good representation of the errors in the different $|E_h|$ estimates. The fractional differences $\Delta E_h/|\mathscr{E}_h|$ provide further information about the expected error distribution (see Figs. 5*a* and 6*a*).

These estimated errors in $|E_h|$ values have been calculated from the standard deviations of measured data and from the Wilson-plot parameters [see (12)]. The systematic errors due to the profile scale and index

* Plots for ANTH1 and K22BR similar to those in Figs. 2, 3, 4 and 6 have been deposited with the British Library Lending Division as Supplementary Publication No. SUP 36644 (5 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

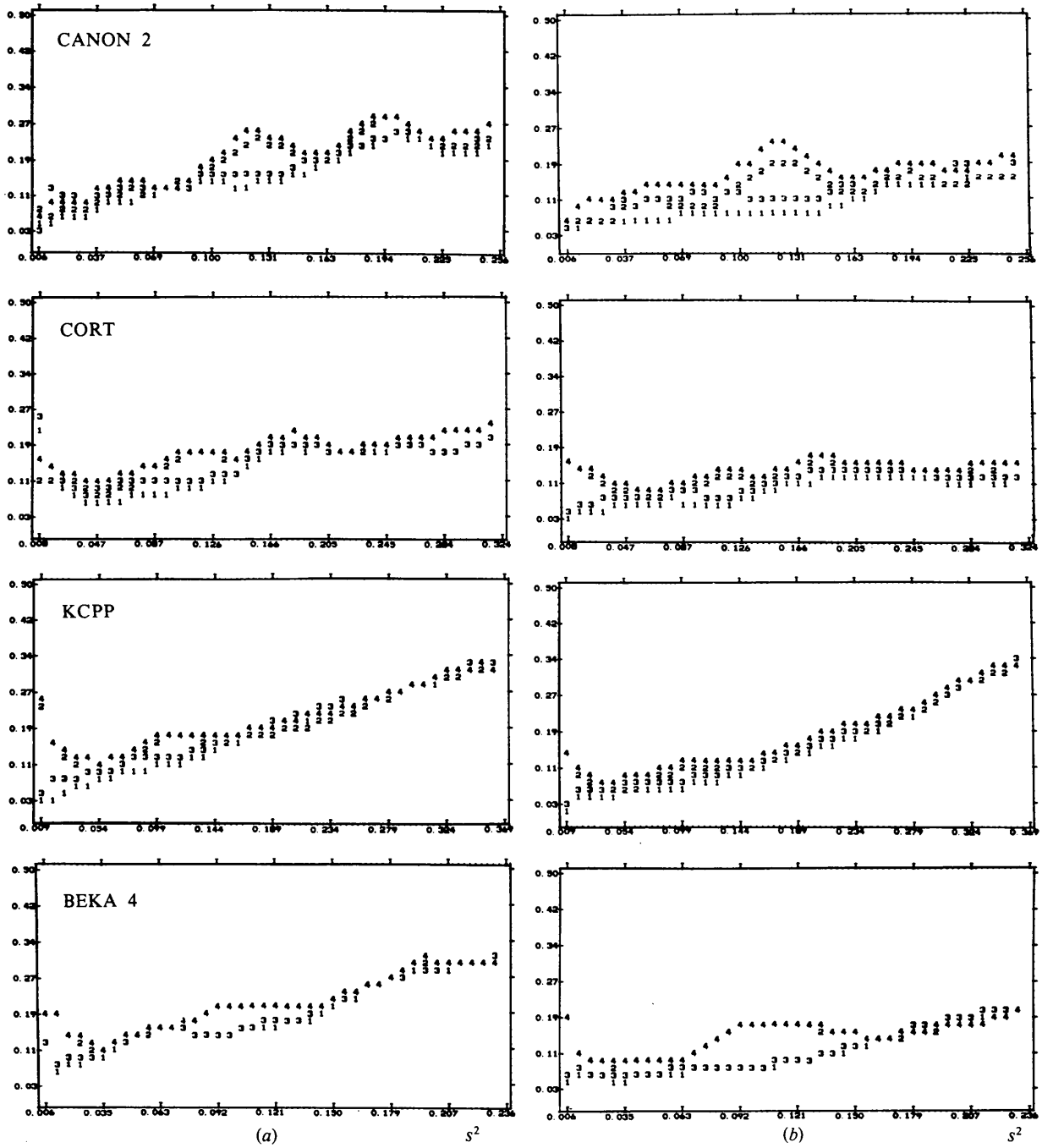


Fig. 2. Plots of (*a*) the measured ΔE_n and (*b*) the estimated ΔE_n *versus* s^2 for four of the test structures. (Plots for ANTH1 and K22BR have been deposited.) The measured ΔE_n is $||\mathscr{E}_h| - |E_h|_n|$ and estimated ΔE_n is calculated using (13), (14) and (15).

rescale terms have been added when appropriate. Plots of the estimated errors are shown in Figs. 2(*b*), 3(*b*) and 4(*b*) as a function of s^2, $|F_h|$ and $|\mathscr{E}_h|$, respectively. Plots of estimated fractional errors are shown in Figs. 5(*b*) and 6(*b*).

The variations of the mean measured and estimated errors with s^2 are shown in Fig. 2. For all test structures the measured error tends to be larger because of the contributions from $\sigma^2|\mathscr{E}_h|$. Apart from this the main features of these plots match closely, and the relative variations among different estimates in the two plots are also similar.

In Fig. 3 the distribution of measured and estimated errors is plotted against $|F_h|$. Once again their general

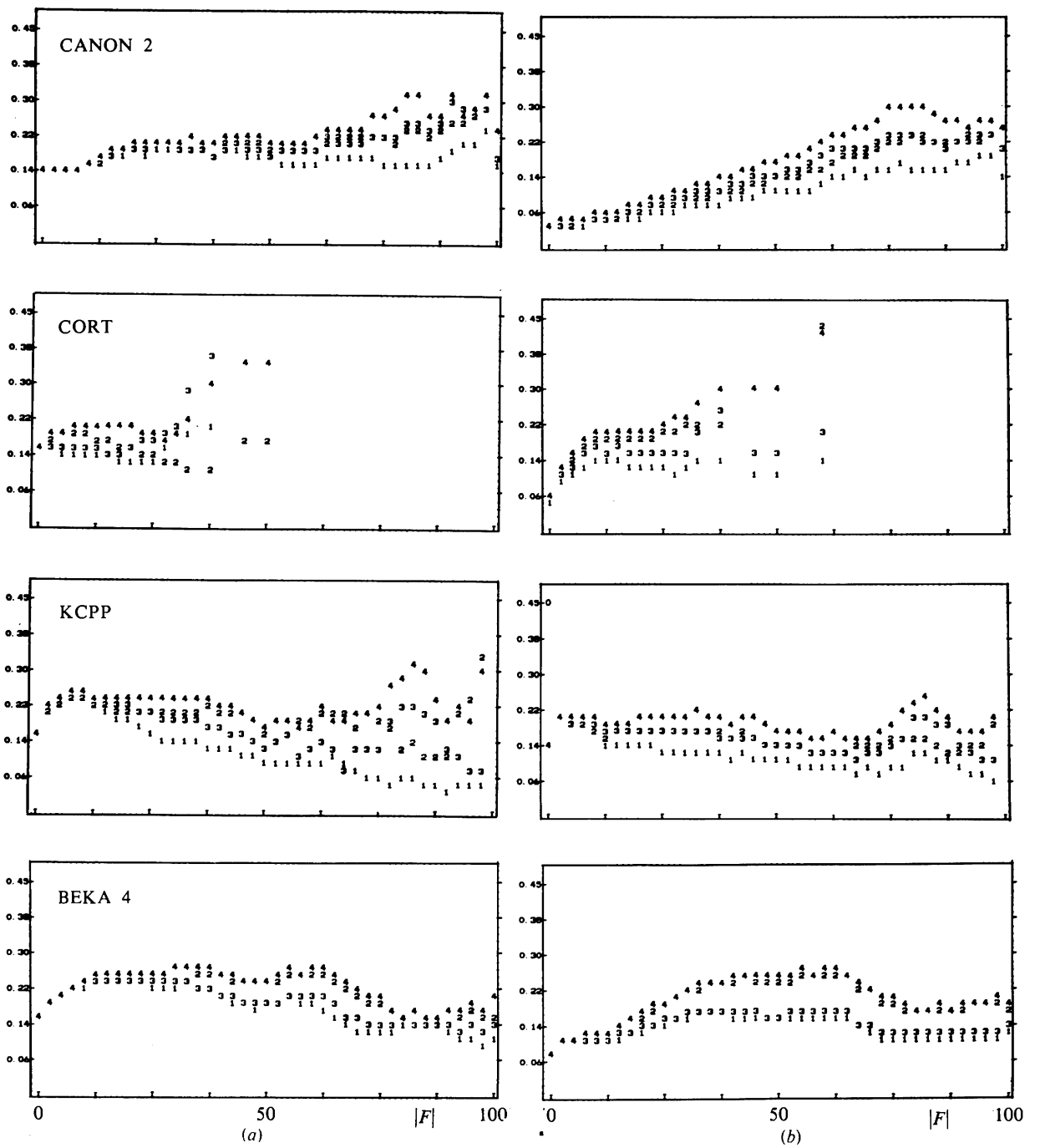


Fig. 3. Plots of (*a*) the measured ΔE_n and (*b*) the estimated ΔE_n *versus* $|F|$ for four of the test structures. (Plots for ANTH1 and K22BR have been deposited.)

similarity is obvious. However, there is a lack of correspondence in the plots of CANON2, especially at low $|F_h|$ values. This is due to the incorrectly low values of σI which give rise to poor values of $\sigma|F_h|/|F_h|$ even after the application of limited Bayesian statistics (Hall & Subramanian, 1982). A similar difference between measured and estimated errors occurs for the CORT data, where no $\sigma|F_h|$ values were available for inclusion in (12).

The plot of measured and estimated errors as a function of $|\mathscr{E}_h|$ is shown in Fig. 4. The good agreement between the plots is evident, with some minor dif-

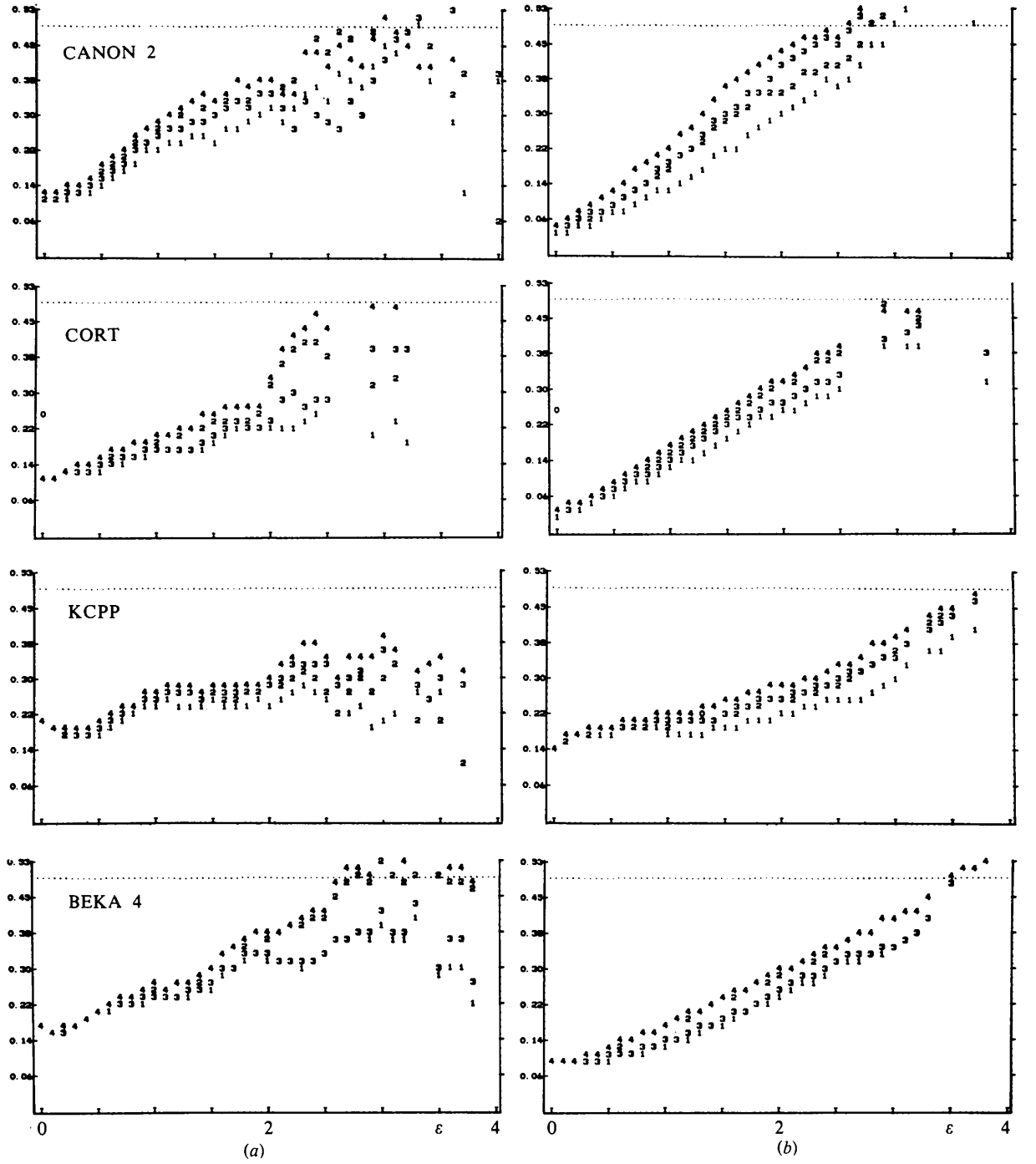
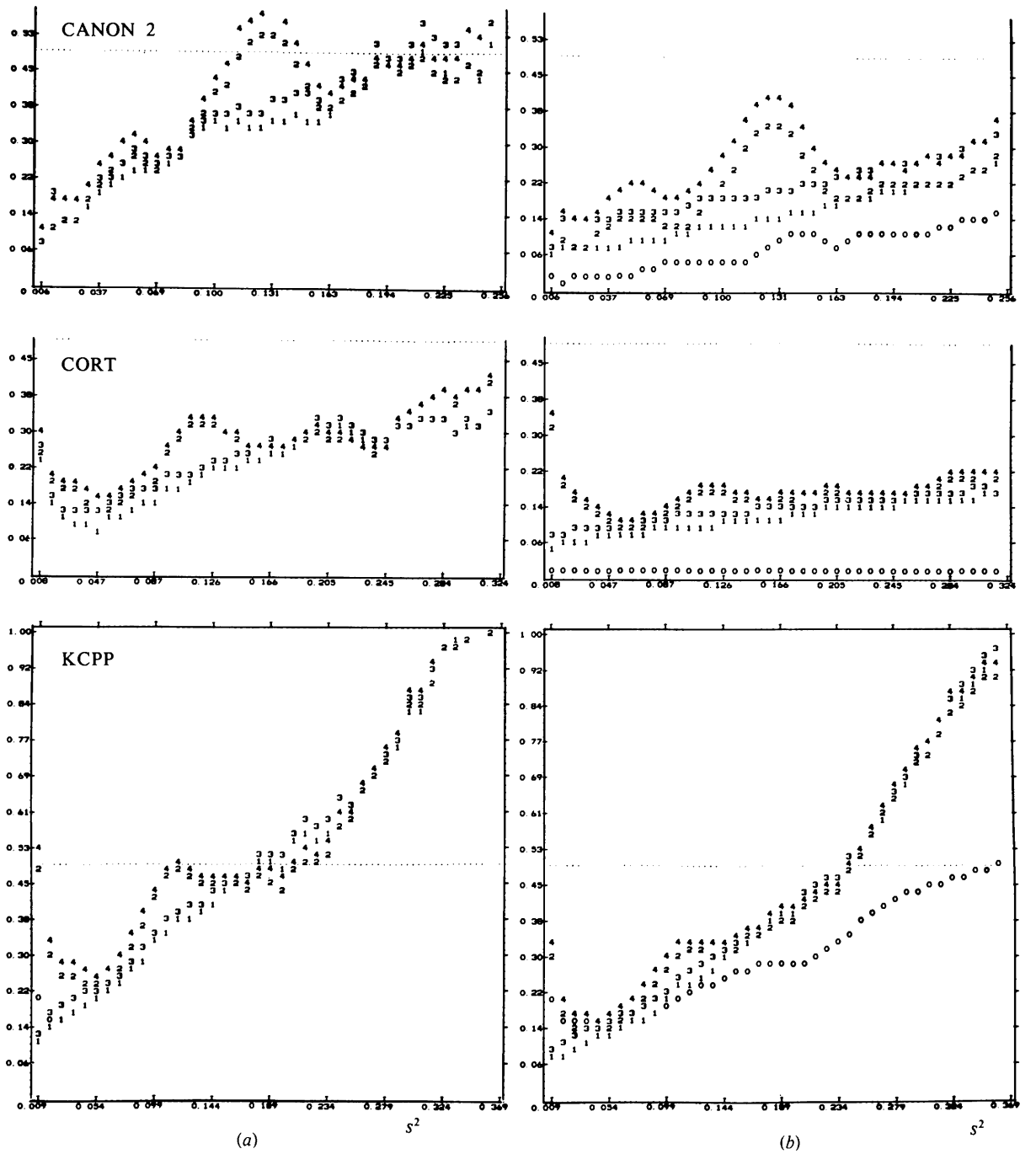


Fig. 4. Plots of (*a*) the measured ΔE_n and (*b*) the estimated ΔE_n *versus* $|\mathscr{E}|$ for four of the test structures. (Plots for ANTH1 and K22BR have been deposited.)

ferences occurring at low $|\mathscr{E}_h|$ values. It is probably due to the contribution of $\sigma^2|\mathscr{E}_h|$ to the measured errors.

In Figs. 5 and 6 the measured and estimated *fractional* errors are compared in terms of s^2 and $|\mathscr{E}_h|$, respectively. Only the plots for the test structures CANON2, CORT and KCPP are illustrated in Fig. 5 because of space considerations. These were selected to emphasize the dependence of the estimated errors on the precision of the $\sigma|F_h|/|F_h|$ values (shown in Fig. 5*b* as 0). For CANON2 and CORT the $\sigma|F_h|/|F_h|$ values are underestimated and unavailable, respectively, and this has a predictable effect on the estimated $\sigma|E_h|$ in both cases. For KCPP and the other three data sets, the $\sigma|F_h|/|F_h|$ values are more reliable and the



(*a*) (*b*)

Fig. 5. Plots of (*a*) the measured $\Delta E_n/|\mathscr{E}|$ and (*b*) the estimated $\Delta E_n/|\mathscr{E}|$ *versus* s^2 for the structures CANON2, CORT and KCPP. $\sigma|F_h|/|F_h|$ is shown as 0 in (*b*).

agreement between measured and estimated errors, as a function of s^2, is excellent. It is also worthy of note that despite the lack of the reliable values of $\sigma|F_h|/|F_h|$ for CANON2 and CORT the errors arising from other sources still produce the correct form of the error distribution. There is no reason to doubt that should the precise values of $\sigma|F_h|/|F_h|$ be available for these two data sets the agreement would be equally good.

The same general observations may be made about the fractional errors with respect to $|\mathscr{E}_h|$ (see Fig. 6).

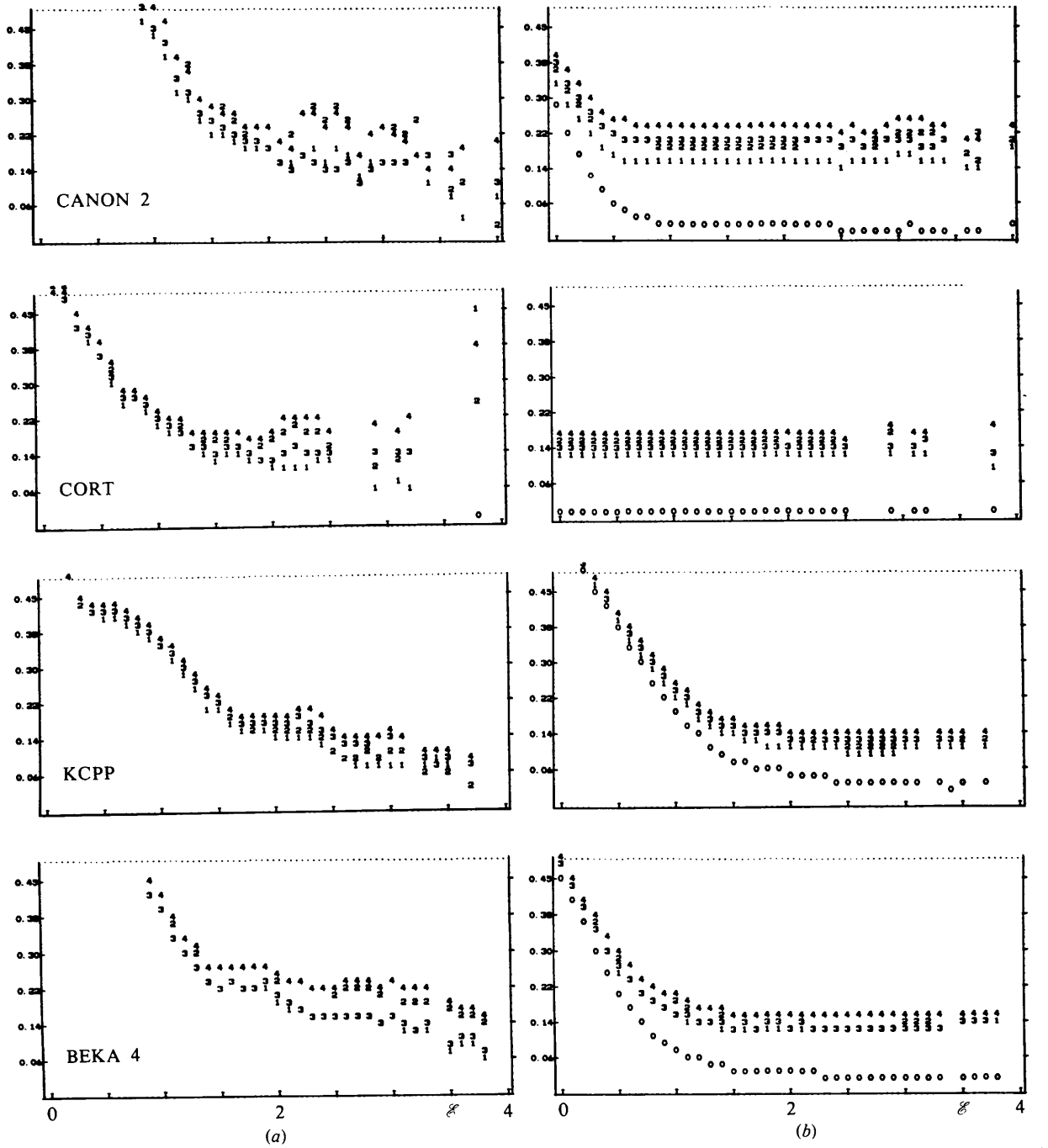


Fig. 6. Plots of (*a*) the measured $\Delta E_n/|\mathscr{E}|$ and (*b*) the estimated $\Delta E_n/|\mathscr{E}|$ *versus* $|\mathscr{E}|$ for four of the test structures. (Plots for ANTH1 and K22BR have been deposited.) $\sigma|F_h|/|F_h|$ is shown as 0 in (*b*).

Data sets with reliable estimates of $\sigma|F_h|/|F_h|$ [shown as 0 in (*b*)] show close agreement between measured and estimated errors, especially for high $|\mathscr{E}_h|$ values. This is important because it is the data with $|\mathscr{E}_h|$ values above the threshold of about 1·5 that are used in phasing procedures. Note that above 1·5 the *average* fractional errors remain relatively constant – though it is clear from the other plots that errors in *individual* $|E_h|$ values may vary significantly with s^2 and $|F_h|$.

Conclusions

A method for calculating the expected errors for different $|E_h|$ values is proposed through the application of (12), (13) and (14). The random error given by (12) relies on experimental estimates of $\sigma|F_h|/|F_h|$ and on the least-squares parameters from the Wilson plot, while the estimates of systematic errors given by (13) and (14) depend on deviations of profile and index scales from the appropriate mean scale factors. Errors estimated in this way are in good agreement with observed errors derived from the difference between $|\mathscr{E}_h|$ and $|E_h|$. This has been verified by comparing the estimated and the measured errors as functions of s^2, $|F_h|$ and $|E_h|$. These comparisons also provide information on the variation of errors for the different methods of calculating $|E_h|$ values. This study also confirms the observations made by Subramanian & Hall (1982) concerning the relative reliabilities of different $|E_h|$ values. It is expected that the availability of estimated errors for $|E_h|$ values prior to structure solution will have important implications for structure-invariant phasing procedures. The application of these estimated errors to phasing procedures, and the minimization of associated noise propagation during phase extension process, are being studied.

The authors wish to acknowledge the assistance of the Australian Research Grants Committee (Grant: C7915302) during the tenure of this work. We are indebted to Dr E. N. Maslen of this laboratory for valuable suggestions on this work.

APPENDIX

In linear least squares the χ^2 summation [see (8)] is minimized with respect to the slope S and the intercept C at $s = 0$ so that

$$\frac{\partial\chi^2}{\partial C} = \sum_i^n w_i (R_i - C - Ss_i^2) = 0, \qquad (A1)$$

$$\frac{\partial\chi^2}{\partial S} = \sum_i^n w_i s_i^2 (R_i - C - Ss_i^2) = 0. \qquad (A2)$$

Equations (*A*1) and (*A*2) can be rewritten as the simultaneous equations

$$\sum w_i R_i = C\sum w_i + S\sum w_i s_i^2 \qquad (A3)$$

$$\sum w_i R_i S_i^2 = C\sum w_i s_i^2 + S\sum w_i s_i^4. \qquad (A4)$$

(*A*3) and (*A*4) may be expressed in vector notation as

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} a_{11}\, a_{12} \\ a_{21}\, a_{22} \end{pmatrix} \begin{pmatrix} C \\ S \end{pmatrix} \qquad (A5)$$

or simply

$$\mathbf{v} = \mathsf{A}\mathbf{p}. \qquad (A6)$$

It follows that

$$\mathbf{p} = \mathsf{B}\mathbf{v} \qquad (A7)$$

where

$$\mathsf{B} = \begin{pmatrix} b_{11}\, b_{12} \\ b_{21}\, b_{22} \end{pmatrix} \qquad (A8)$$

$$= \begin{pmatrix} a_{22} - a_{12} \\ -a_{21}\, a_{11} \end{pmatrix} / |\mathsf{A}| \qquad (A9)$$

and

$$|\mathsf{A}| = a_{11}\, a_{22} - a_{12}\, a_{21}. \qquad (A10)$$

From (*A*7), the intercept C is

$$C = (a_{22}\, v_1 - a_{12}\, v_2/|\mathsf{A}|, \qquad (A11)$$

and the slope S is

$$S = (a_{11}\, v_2 - a_{21}\, v_1)/|\mathsf{A}|. \qquad (A12)$$

The variance of each parameter may also be obtained from the least-squares matrix B as

$$\sigma^2 p_i = \frac{b_{ii}\,\chi^2}{n-2}. \qquad (A13)$$

The variance of the intercept and the slope are therefore

$$\sigma^2 C = \frac{b_{11}\,\chi^2}{n-2} = \frac{a_{22}\,\chi^2}{|\mathsf{A}|\,(n-2)} \qquad (A14)$$

$$\sigma^2 S = \frac{b_{22}\,\chi^2}{n-2} = \frac{a_{11}\,\chi^2}{|\mathsf{A}|\,(n-2)}, \qquad (A15)$$

and

$$\sigma^2(C,S) = \frac{b_{12}\,\chi^2}{n-2} = \frac{a_{12}\,\chi^2}{|\mathsf{A}|\,(n-2)}. \qquad (A16)$$

The value of the χ^2 summation may be calculated from the expression

$$\chi^2 = C^2 a_{11} + 2CSa_{12} + S^2 a_{22} - 2Cv_1 - 2Sv_2 + v_3, \quad (A17)$$

where

$$v_3 = \sum^{n} w_i R_i^2. \quad (A18)$$

The correlation coefficient between C and S is

$$r(C, S) = b_{12}/(b_{11} b_{22})^{1/2}. \quad (A19)$$

References

DECLERCQ, J. P., GERMAIN, G. & VAN MEERSSCHE, M. (1972). *Cryst. Struct. Commun.* **1**, 13–15.

HALL, S. R., RASTON, C. L. & WHITE, A. H. (1978). *Aust. J. Chem.* **31**, 685–688.

HALL, S. R. & SUBRAMANIAN, V. (1980). XRAY76 system of crystallographic programs: program *ESCAN* for the comparison of E values. Univ. of Western Australia.

HALL, S. R. & SUBRAMANIAN, V. (1982). *Acta Cryst.* A**38**, 590–598.

LADD, M. F. C. (1978). *Z. Kristallogr.* **147**, 279–296.

SKELTON, B. W. & WHITE, A. H. (1981). Personal communication.

SUBRAMANIAN, V. & HALL, S. R. (1982). *Acta Cryst.* A**38**, 577–590.

Acta Cryst. (1982). A**38**, 608–611

Use of a Statistical Distribution of Electron Density to Describe the Bragg Scattering from a Linearly Disordered Crystal

BY RAYMOND P. SCARINGE, LAUREL J. PACE AND JAMES A. IBERS

Department of Chemistry and Materials Research Center, Northwestern University, Evanston, Illinois 60201, *USA*

(*Received* 13 *November* 1981; *accepted* 9 *March* 1982)

Abstract

The modulated non-zero electron density distribution that results from linear disorder of iodine chains in systems of stacked planar organic molecules or metallomacrocycles partially oxidized by iodine is modeled by an integrable statistical distribution function. The contributions to the Bragg scattering of the iodine disorder are fit in an excellent manner with the use of at most two extra variables.

Introduction

Several highly conducting one-dimensional systems have been prepared by partial oxidation with iodine of planar organic molecules or metallomacrocycles (for a review see Hoffman, Martinsen, Pace & Ibers, 1982). These systems typically contain stacks of the oxidized species surrounded by linear chains of polyiodide anions that are disordered. The form of the iodine (*e.g.* I_2, I_3^-, I_5^-) can be elucidated spectroscopically either by resonance Raman or Mössbauer methods (Marks, 1978), and in several cases structural information on the iodine species has been obtained by analysis of the diffuse X-ray scattering that results from the disorder (Endres, Keller, Mégnamisi-Bélombé, Moroni, Pritzkow, Weiss & Comès, 1976; Scaringe & Ibers, 1979; Schramm, Scaringe, Stojakovic, Hoffman, Ibers & Marks, 1980). Here, for several systems containing I_3^- we shall consider the effect of this disorder on the Bragg scattering.

In the systems Ni(Pc)I* (Schramm *et al.*, 1980), Ni(tbp)I (Martinsen, Pace, Phillips, Hoffman & Ibers, 1982) and $M(bqd)_2I_{0.5}$ (Endres, Keller & Weiss, 1975; Brown, Kalina, McClure, Schultz, Ruby, Ibers, Kannewurf & Marks, 1979) the disorder of the I_3^- anions is not severe, perhaps because the spacing between macrocycles in the stack is short, less than 3·24 Å. All of the diffuse X-ray lines in these systems can be indexed on the basis of a superlattice spacing that is

* Abbreviations used: bqd, 1,2-benzoquinonedioximato; Pc, phthalocyaninato; tbp, tetrabenzporphyrinato; omtbp, 1,4,5,8,9,12,13,16-octamethyltetrabenzporphyrinato; tmp, 5,10,15,20-tetramethylporphyrinato; φ_4DTP, tetraphenyldithiapyranylidene; TTT, tetrathiatetracene.

0567-7394/82/050608-04$01.00