Acta Cryst. (1997). A53, 376-387

Errors in Maximum-Entropy Charge-Density Distributions Obtained from Diffraction Data

BO BRUMMERSTEDT IVERSEN,[a]*† JENS LEDET JENSEN[b] AND JACOB DANIELSEN[a]

[a]*Department of Chemistry, and* [b]*Department of Mathematics, Aarhus University, DK-8000 Aarhus C, Denmark. E-mail: bo@kemi.aau.dk*

(Received 29 July 1996; accepted 20 January 1997)

Abstract

A scheme to estimate errors in maximum-entropy-method (MEM) charge-density distributions *via* Monte Carlo simulations is presented. Knowledge of the errors in the density allows evaluation of the reliability of fine features in MEM electron densities. As a test example, the errors in the MEM electron-density distribution of metallic beryllium are examined based on calculations using both uniform and non-uniform prior distributions. The study shows that the MEM introduces systematic bias in the density and that the bias is closely related to the fact that after a MEM optimization most of the χ^2 value is carried by a few low-order reflections. An iterative procedure to estimate the bias is presented and this allows a correction of the MEM density to be performed. The systematic bias is in some regions an order of magnitude larger than the random error in the density. The bias-corrected MEM density has a more uniform residual distribution than the uncorrected density. The topological features of the electron-density distribution in metallic beryllium are discussed based on the bias-corrected MEM densities.

1. Introduction

In recent years, the maximum-entropy method (MEM) has been used in a number of studies of charge distributions in crystals. When X-ray diffraction data are used, the MEM yields the electron-density distribution (EDD) without assuming any model (Collins, 1982; Sakata & Sato, 1990), whereas neutron diffraction data allows the direct-space nuclear probability density function to be determined (Sakata, Uno, Takata & Howard, 1993). From limited numbers of accurate X-ray diffraction data, EDD's have been determined by the MEM in a number of systems, such as Si, Be and YC_{82} (Sakata & Sato, 1990; Takata, Kubota & Sakata, 1993; Takata, Umeda, Nishibori, Sakata, Saito, Ohno & Shinoshara, 1995). Maps that qualitatively reveal bonding features have been obtained in these and many other studies. Although this is of interest in itself, quantification and detailed analysis of the derived MEM charge densities is highly desirable to make the method generally useful. Quantification of MEM densities has been introduced for nuclear density distributions through parametrization and least-squares fitting of the MEM nuclear density to a function based on the one-particle potential model (Takata, Sakata, Kumasawa, Larsen & Iversen, 1994; Iversen, Nielsen & Larsen, 1995: Kumazawa, Takata & Sakata, 1995). In the case of electron densities, Iversen, Larsen, Souhassou & Takata (1995) have employed the theory of atoms in molecules (Bader, 1991) to perform a detailed topological analysis of the MEM density of metallic beryllium. In that study, it was possible to reveal subtle but nevertheless physically important features such as non-nuclear maxima and bond critical points in the density. However, interpretation of fine features in the density raises the question of the reliability of the MEM density. Several authors have pointed out that unphysical features can appear in MEM densities and, depending on the quality and the completeness of the data, fine features in the density may be artifacts of the density reconstruction (Jauch & Palmer, 1993; Iversen, Nielsen & Larsen, 1995; Papoular, Vechter & Coppens, 1996; Takata & Sakata, 1996). It has furthermore been pointed out that use of an entropy term as a regularizing function in the reconstruction inevitably will introduce systematic bias into the result (Jauch, 1994; Donoho, Johnstone, Hoch & Stern, 1992).

In the conventional empirical modeling methods widely used in accurate charge-density reconstruction from diffraction data (Stewart, 1976; Hirshfeld, 1977; Hansen & Coppens, 1978; Figgis, Reynolds & Williams, 1980), estimates of errors in the density and in derived properties can be calculated within the framework of the least-squares method. Such estimates rely on several assumptions, among which are the adequacy of the refined model. Several studies (Figgis, Iversen, Larsen & Reynolds, 1993; Chandler, Figgis, Reynolds & Wolff, 1994; Iversen, Larsen, Figgis & Reynolds, 1995) have shown that even the very sophisticated models currently used in empirical EDD modeling are inadequate to describe very fine density features present in the data and, in general, least-squares estimates of EDD's will therefore contain systematic bias owing to the model. However, the least-squares error estimates allow to some extent assessment of the reliabilty of conclusions drawn from the model densities. An error analysis has so far not been developed in the case of MEM charge densities.

† Present address: Department of Chemistry, University of California at Santa Barbara, Santa Barbara, CA 93106, USA.

It is clear that if the MEM is to become more generally applicable in chemistry and physics some kind of error assessment is necessary. This is especially imperative for very accurate charge-density studies.

The present paper presents an attempt to derive estimates of the error in MEM densities through Monte Carlo simulations. As a test case, the beryllium MEM density derived from the 58 structure factors measured by Larsen & Hansen (1984) has been investigated. Metallic beryllium is particularly well suited for testing MEM error estimates because it is expected that systematic errors in the data are few. This is due to a number of facts. First, beryllium has a very high Debye temperature (1340 K), which means that the mean-square amplitude of vibration is small. Thus, even at room temperature, beryllium scatters well to high values of $\sin\theta/\lambda$. The hardness of beryllium also effectively reduces contributions of thermal diffuse scattering (TDS) to the Bragg intensities. Furthermore, since elastic constants are known, the limited TDS effects present in the high-order intensities can adequately be corrected for. For the beryllium data set, the maximum TDS correction amounted to 2.6% of the intensity. The beryllium crystals have excellent diffraction properties and this allows use of small single crystals. The small size of the sample crystal ensures a good beam homogeneity of the part of the X-ray beam hitting the crystal. As a further benefit, extinction effects are minimal. Owing to the low atomic number of beryllium as well as the small size of the sample crystal, absorption effects are also effectively reduced. For the present data set, transmission factors varied between 0.990 and 0.995. Finally, beryllium has a very high symmetry, which means that a large number of equivalent reflections can be measured. This allows for removal of outliers and the calculation of unusually precise standard uncertainties on the structure factors. In conclusion, the beryllium data contain comparatively few systematic errors and provide a good case for calculation of errors in the MEM density. If systematic errors are detected in the MEM result, they are likely to be due to the MEM algorithm and not the data.

2. Monte Carlo simulation of errors in MEM densities

The data that are available form a complete set of 58 structure factors ($\sin\theta/\lambda < 1.21$ Å^{-1}) with standard uncertainties (s.u.'s) derived from the scatter of equivalent reflections (Larsen & Hansen, 1984). In the case of beryllium, the s.u.'s on the structure factors are typically, owing to the high symmetry of the h.c.p. structure, estimated from approximately 16 measurements of reflections related by symmetry. If we assume that systematic errors in the data are negligible and furthermore assume that the error distributions around the true structure factors are Gaussian, then the error distributions around the individual true structure factors are defined from the estimates of the standard uncertainties. From the set of observed structure factors and their standard uncertainties [F^{obs}, $\sigma(F^{obs})$], we can calculate a MEM electron-density distribution (EDD), ρ^0, using for instance the *MEED* algorithm (Kumazawa, Kubota, Takata, Sakata & Ishibashi, 1993). For a discussion of the *MEED* algorithm and the MEM itself, readers are referred to the references given in the *Introduction*. In Fig. 1, the MEM density based on 58 structure factors is shown for the (110) plane of the beryllium hexagonal close-packed structure, Fig. 2.

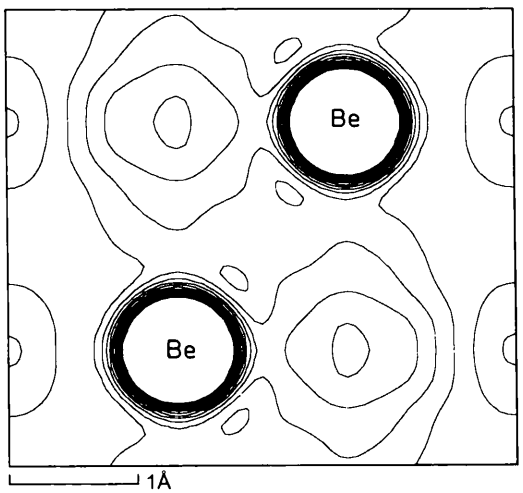


Fig. 1. Contour plot in the (110) plane of the MEM electron-density distribution of beryllium based on X-ray single-crystal room-temperature diffraction data. The plot is on a linear scale with 0.05 e Å^{-3} intervals. The contour plotting was truncated at 1.0 e Å^{-3} to put emphasis on the lower-density regions.

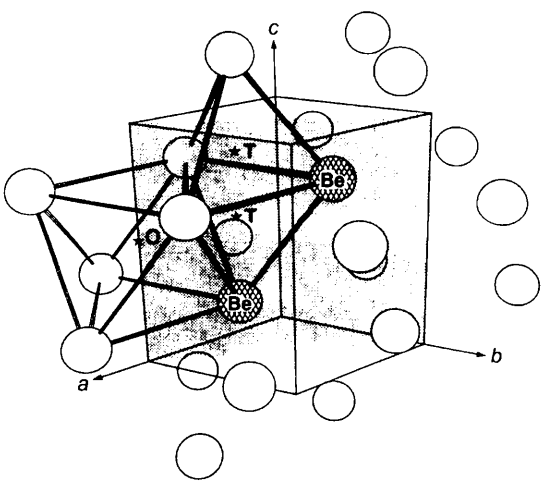


Fig. 2. The hexagonal-close-packed beryllium structure. The space group is $P6_3/mmc$ (no. 194) with beryllium at (1/3, 2/3, 1/4). The unit-cell dimensions derived from X-ray data are $a = 2.2853$ (3) and $c = 3.5842$ (2) Å. The shaded area within the outlined unit cell is the (110) plane, which goes through the tetrahedral holes, marked *T*, and through the octahedral holes, marked *O*.

The Monte Carlo method for estimation of errors that we have used closely follows the presentation of Press, Teukolsky, Vetterling & Flannery (1992). For an in-depth discussion on the Monte Carlo method, readers are referred to this text and references therein. We will briefly outline our adaptation of their Monte Carlo scheme as well as the main points of the procedure. The MEM density, ρ^0, will have a corresponding set of calculated structure factors, F^0. We can construct synthetic data sets by adding random noise to F^0 according to the known error distribution around the true values. Methods to add random noise that follows a given distribution are well described in standard text books on numerical methods (see Press, Teukolsky, Vetterling & Flannery, 1992). In this way, synthetic data sets, $F^{i,\mathrm{syn}}$, can be generated, and these synthetic data sets can be used as input to a series of MEM calculations. The result will be a series of Monte Carlo MEM densities, $\rho^{i,\mathrm{MC}}$, and the scatter of these densities can be used to give an estimate of the error in the original MEM density, ρ^0, which was obtained by the MEM from the observed structure factors, F^{obs}. The basic idea of the Monte Carlo simulation is that we are not assuming F^0 to be equal to the true structure factors. We merely assume that the relationship between F^0 and the synthetic data sets, $F^{i,\mathrm{syn}}$, is the same as the relationship between the true structure factors, F^{true}, and the set we happened to measure, F^{obs}. In Fig. 3, a schematic representation of the Monte Carlo simulations is shown.

3. Calculations

As MEM calculations even for a single data set are computationally quite demanding, obviously Monte Carlo simulations require substantial computing power. The calculations presented in this paper were carried out on a Digital DEC3000 work station and as an example the CPU time for one MEM optimization with approximately 3000 cycles was of the order of 30 min. However, as the problem is ideal for parallel computing, the speed of calculation can be much improved. In all calculations, the beryllium h.c.p. unit cell was divided into a 60 × 60 × 60 grid. In the study of Iversen, Larsen, Souhassou & Takata (1995), a 120 × 120 × 120 grid was used in the final calculations of the topological features. However, in that study it was shown that a 60 × 60 × 60 density contains almost identical features to a 120 × 120 × 120 pixel density and, for practical reasons (a factor of eight in computation time), the present study therefore employs a 60 × 60 × 60 grid. For practical reasons, we have also limited the present study to 200 Monte Carlo data sets. The conclusions presented below do not depend critically on this number. In future work, the Monte Carlo procedure may be optimized to save calculation time but the main purpose of this paper is to present the general idea of the calculations. In all Monte Carlo calculations, the iterations were stopped when the constraint $\chi^2 = 1/N\sum(F_{\mathbf{H}}^o - F_{\mathbf{H}}^c)^2/\sigma_{\mathbf{H}}^2$ had reached a value of 1 and all calculations employed the same value of the Lagragian multiplier ($\lambda = 0.000015$).

4. Systematic bias in the MEM

Once 200 Monte Carlo MEM densities are available, the standard uncertainty in each pixel, x, can be calculated by

$$\sigma(\rho_x) = \left[\sum_{i=1}^{200}(\rho_x^i - \overline{\rho_x})^2/199\right]^{1/2}. \qquad (1)$$

If no systematic features are introduced by the MEM algorithm, we will expect that

$$\overline{\rho_x} = \rho_x^{\mathrm{ave}} = (1/200)\sum_{i=1}^{200}\rho_x^i = \rho_x^0. \qquad (2)$$

In (2), ρ_x^0 is the density in pixel number x of the MEM density obtained from the observed structure factors. However, the average value of the Monte Carlo densities, ρ_x^{ave}, turns out to be systematically different from ρ_x^0. This is illustrated in Fig. 4, where a plot of

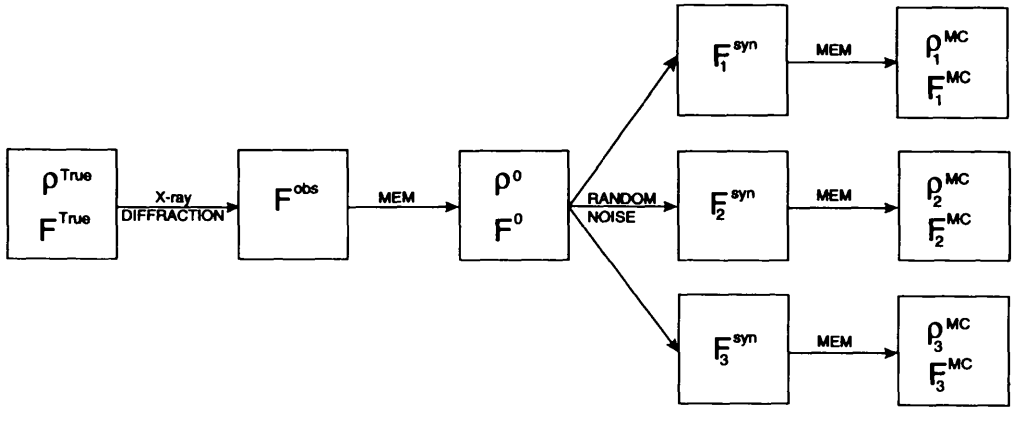


Fig. 3. Schematic representation of a Monte Carlo simulation of the error in the MEM density.

$b(\rho^0) = \rho^{ave} - \rho^0$ is shown for the (110) plane. The map indicates where systematic bias, $b(\rho^0)$, is introduced in the density by the MEM algorithm. If a similar systematic bias was introduced in the calculation of ρ^0 from the observed structure factors, F^{obs}, then the map also suggests where ρ^0 may be systematically different from the true electron density of beryllium. It is instructive to analyze the distribution of the Monte Carlo MEM densities for individual pixels in order to get a better understanding of the features in Fig. 4. In Fig. 5, such distributions are shown for two different pixels. The first pixel is the beryllium maximum whereas the second pixel is a randomly chosen pixel. As can be seen, the distributions are not centered on the ρ^0 values. In other words, the means of the distributions have been systematically shifted. The large atomic maximum is the most affected distribution and for this pixel the density has been shifted -0.74 e Å^{-3}. The random pixel, Fig. 5(*b*), has a relatively smaller shift of the mean of the distribution (0.01 e Å^{-3}).

In order to better understand the systematic bias in specific regions of the density, we have also analyzed the individual structure-factor values obtained after the MEM. Jauch & Palmer (1993) pointed out that after a MEM optimization the distribution of structure-factor residuals, $\Delta F_j/\sigma(F_j)$, is very non-uniform. The MEM tends to have the residual concentrated in a few low-order structure factors whereas the rest of the reflections fit almost perfectly.* For beryllium, it is the 101 reflection that carries most of the error after the optimization, $\Delta F(101)/\sigma[F(101)] = -6.9$, as compared to the second largest residual, $\Delta F(112)/\sigma[F(112)] = -1.6$. It is of interest to see whether the 101 reflection also carries most of the error in the Monte Carlo densities or whether it has been 'pulled back' towards the observed value. In Fig. 6, distributions of the structure-factor values corresponding to the Monte Carlo densities, F_j^{MC}, are shown for the 101 reflection and for comparison also for a typical low-order reflection, 100. As can be seen, the 101 reflection is shifted away from the original MEM value, whereas the 100 reflection is nicely distributed around it. The 101 reflection is in other words shifted further away from the observed value in

* A table of structure-factor residuals for various MEM densities have been deposited with the IUCr (Reference: SH0090). Copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

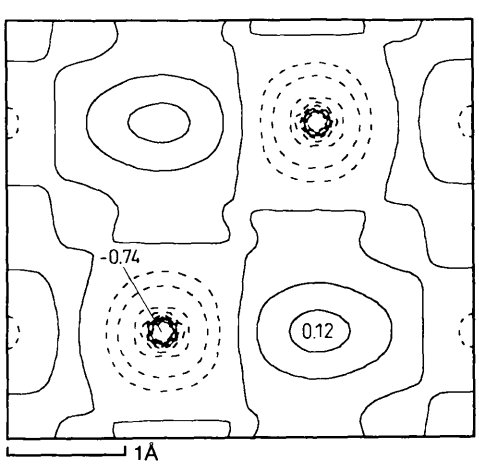


Fig. 4. Contour plot in the (110) plane of the bias distribution, $b(\rho^0) = \rho^{ave} - \rho^0$, where ρ^{ave} is given in formula (2). The plot is on a linear scale with 0.05 e Å^{-3} intervals. The maximum value in the map is 0.115 e Å^{-3} and the minimum -0.736 e Å^{-3}. The broken lines are negative contours and the solid lines are positive contours.

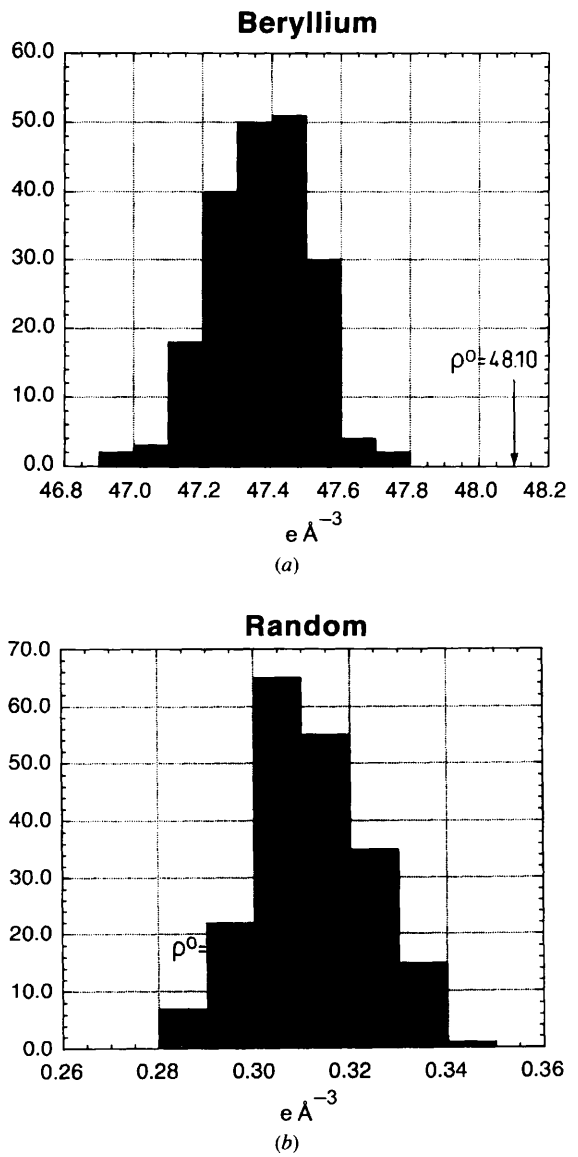


Fig. 5. Plots of the distributions of the Monte Carlo densities, $\rho^{j,MC}$. (*a*) Beryllium, (*b*) random pixel. The values of the average Monte Carlo density, ρ^{ave}, and the original MEM density, ρ^0, are marked on the plots.

the Monte Carlo calculation. The two reflections have similar distributions of the Monte Carlo structure factors but only the 101 reflection has a large systematic shift of the mean.

de Vries, Briels & Feil (1994) noted that the larger the amplitude of a structure factor the larger the entropy gain when such a structure factor is made to deviate from the observed value. For beryllium, it is not necessarily the large-amplitude structure factors that carry most of the error. The strongest reflection for beryllium is the 002 reflection and this reflection only has the sixth-largest deviation after the MEM optimization. There is also no obvious connection between the length of a scattering vector, $|\mathbf{H}|$, and the residual, although the poorly fitting reflections always are low-order reflections.

An interesting question is why it is the 101 reflection that is 'chosen' by the MEM to carry most of the error. We have calculated the entropy of the MEM density after convergence of single-structure-factor MEM calculations for the 58 structure factors. The single-structure-factor MEM calculations are analogous to performing a Fourier 'summation' with just one structure factor. The entropy of the 101 reflection is considerably smaller than for the other reflections. Thus, the optimization of the entropy term makes the 101 reflection carry most of the error. If we remove the 101 reflection and perform the MEM optimization, it is the 112 and the 103 reflections that have most of the error. These reflections are next on the list of minimal entropy. It seems that the entropy of the individual reflections is related to which reflections contribute most to the χ^2 deviation after the optimization.

In Fig. 7, a plot of the MEM density obtained from the 101 single-structure-factor MEM calculations is shown. The 101 reflection builds up density at the beryllium position but, contrary to the other strong reflections, it introduces a minimum in the bipyramidal space of tetrahedral holes. These are exactly the regions with the largest systematic bias and in general the bias is found in highly symmetric positions. Clearly, the systematic bias in the MEM density is closely related to the fact that the scatter of the residuals, $\Delta F_j/\sigma(F_j)$, is very nonuniform after a MEM calculation.

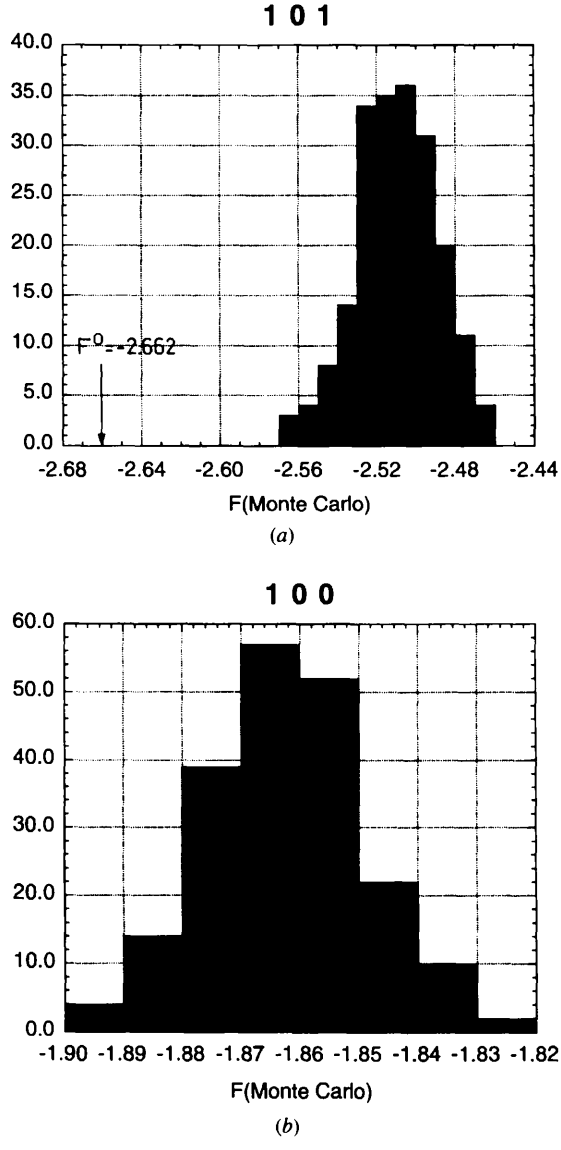


Fig. 6. Distributions of Monte Carlo structure factors, $F^{i,MC}$, for (*a*) the 101 and (*b*) the 100 reflections, respectively. The original MEM values, F^0, as well as the average values, F^{ave}, are marked on the plots.

5. Direct-space correction for systematic bias

Since the average value of the Monte Carlo densities in a given pixel is different from ρ^0, the MEM density

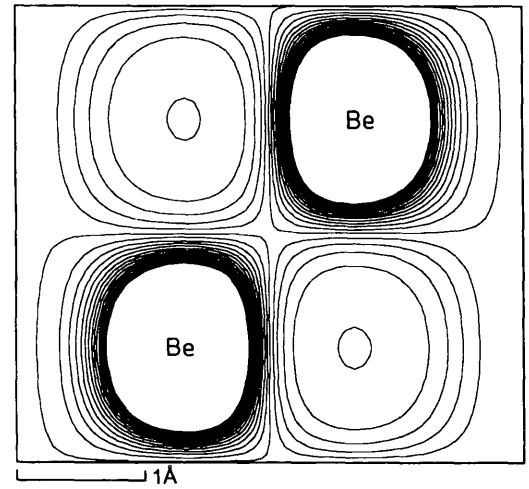


Fig. 7. Contour plot of the 101 single structure-factor density in the (110) plane. Contours as in Fig. 1.

must contain systematic bias relative to the true density, ρ^{true}. We do not know the MEM bias on ρ^{true}, but from the Monte Carlo simulations we know the bias on ρ^0. We can use an iterative procedure to estimate the bias on ρ^{true}, $b(\rho^{\text{true}})$, which is defined by ($E[\rho^0]$ denotes the expected value of ρ^0)

$$\rho^{\text{true}} = E[\rho^0] - (E[\rho^0] - \rho^{\text{true}}) = E[\rho^0] - b(\rho^{\text{true}}). \quad (3)$$

Similarly, the bias on the MEM density, ρ^0, is

$$\rho^0 = \rho^{\text{ave}} - (\rho^{\text{ave}} - \rho^0) = \rho^{\text{ave}} - b(\rho^0). \quad (4)$$

If ρ^0 is not very different from ρ^{true}, we can hope that $b(\rho^{\text{true}}) \approx b(\rho^0)$ and use the known bias on ρ^0, $b(\rho^0)$, to correct ρ^0 for systematic bias. If the assumption is valid, the correction will give us a better estimate of the true density, $\hat{\rho}^1$,

$$\hat{\rho}^1 = \rho^0 - b(\rho^0). \quad (5)$$

Because $\hat{\rho}^1$ is a better estimate of ρ^{true}, we then carry out the Monte Carlo simulations using $\hat{\rho}^1$ and the corresponding structure factors, $\hat{F}^1$, in place of ρ^0 and F^0. This results in an estimate of the bias in $\hat{\rho}^1$ that is an improved estimate of the bias in ρ^{true},

$$b(\hat{\rho}^1) = \hat{\rho}^{1,\text{ave}} - \hat{\rho}^1. \quad (6)$$

Knowing $b(\hat{\rho}^1)$, we can calculate an improved estimate of ρ^{true} as

$$\hat{\rho}^2 = \rho^0 - b(\hat{\rho}^1). \quad (7)$$

The iterative procedure can be stopped when the difference in the bias estimates between consecutive iterations is small. The result after the ith iteration is an estimate of the bias on ρ^{true} and this can be used to correct the original MEM density,

$$\rho^0_{\text{corrected}} = \rho^0 - b(\hat{\rho}^i). \quad (8)$$

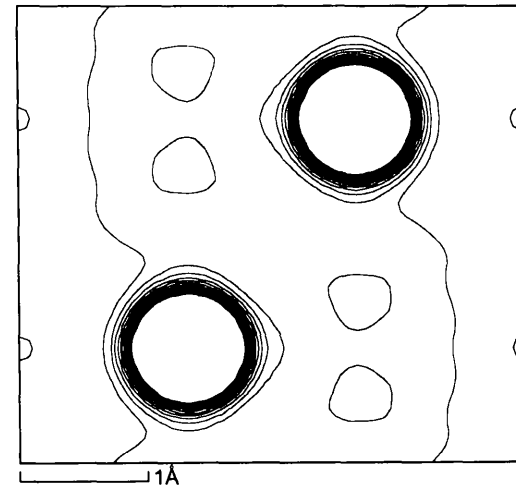
In Fig. 8, plots of the bias-corrected MEM density after one and two iterations are shown for the (110) plane. It is notable that the region around the non-nuclear maximum (NNM) in the basal plane has become flatter although it still contains density accumulation relative to the rest of the valence region. The importance of this will be discussed below.

If the scheme presented above to correct for the bias is valid, then the scatter of the residuals should be more uniform for $\rho^0_{\text{corrected}}$ than for ρ^0. If no bias is introduced by the optimization algorithm, the residual distribution of the structure factors should follow a normalized Gaussian distribution, $N(0,1)$, under the assumption of Gaussian error distributions around the true structure factors. In Fig. 9, the distribution of residuals in the structure factors before and after the correction is shown. However, since the value of χ^2 after the first correction is 0.13 (0.02 after the second iteration), the comparison is not straightforward. We have therefore in Fig. 9(*d*) included the distribution of residuals for the uncorrected density with the MEM iterations stopped at $\chi^2 = 0.13$. The scatter of residuals for the bias-corrected density is clearly more uniform than for the original MEM density.

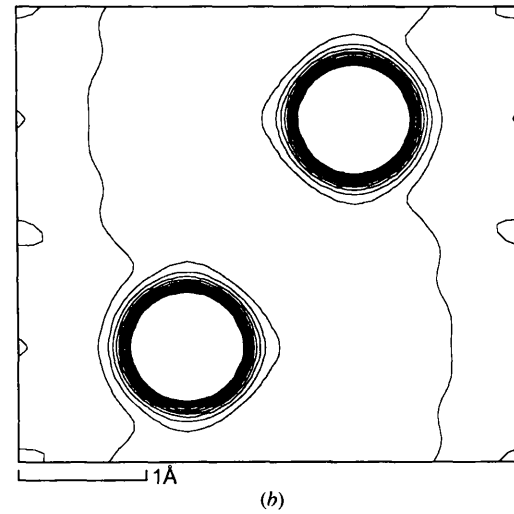
6. Error estimates

The mean-square error of the estimate is

$$E[\rho_x^{\text{true}} - \hat{\rho}_x^i]^2 = E[\hat{\rho}_x^i - E[\hat{\rho}_x^i]]^2 + [E[\hat{\rho}_x^i] - \rho_x^{\text{true}}]^2. \quad (9)$$



(*a*)



(*b*)

Fig. 8. Contour plot in the (110) plane of the bias-corrected MEM density. In (*a*), the first-order corrected density is plotted, while (*b*) shows the second-order corrected density. Contours as in Fig. 1.

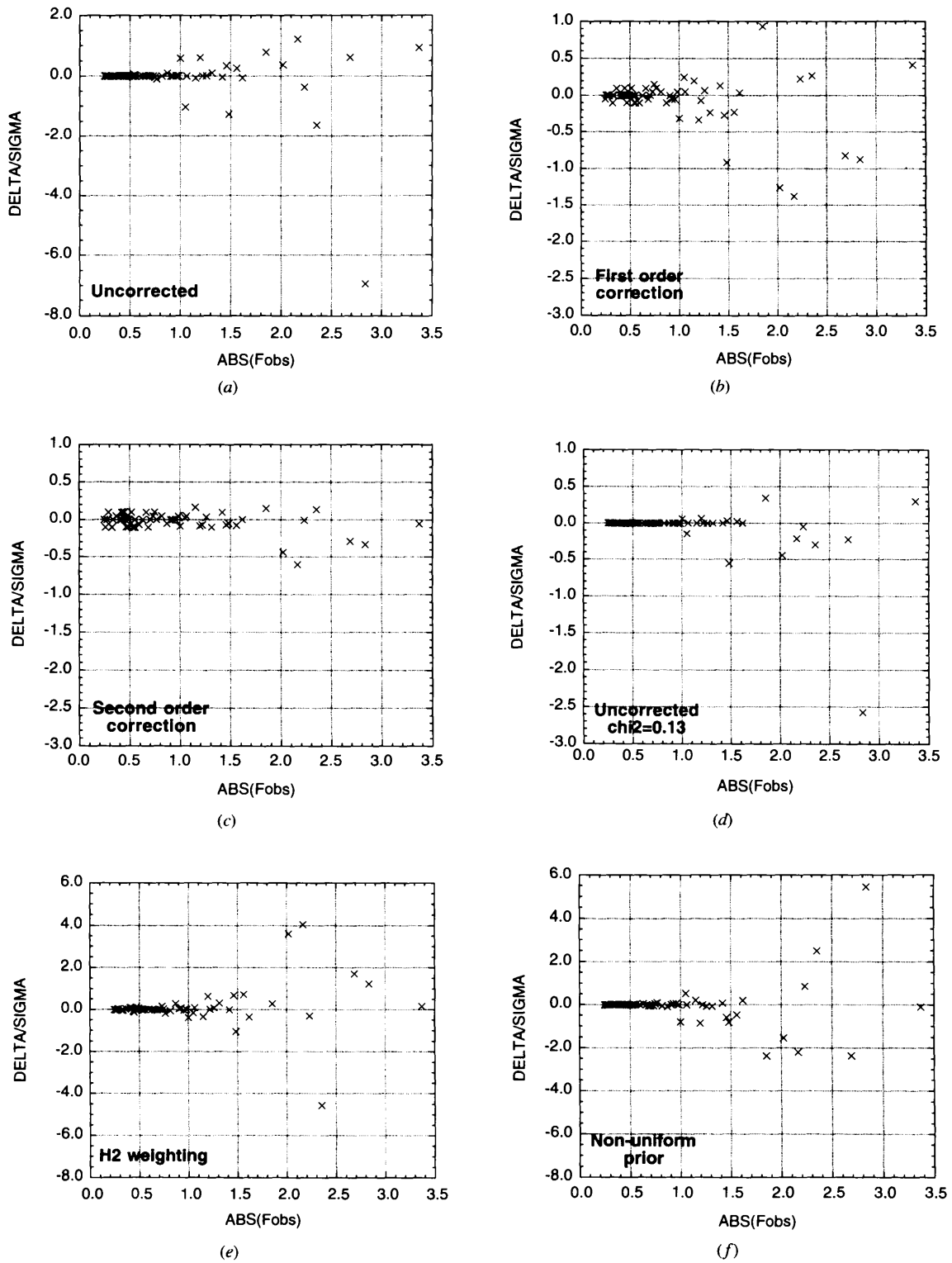


Fig. 9. Distributions of residuals, $\delta F/\sigma(F)$, *versus* the absolute values of the structure factors for various MEM densities. (*a*) Uncorrected density, (*b*) first-order bias-corrected density, (*c*) second-order bias-corrected density, (*d*) MEM density with iterations stopped at $\chi^2 = 0.13$, (*e*) MEM density obtained with H^2 weighting and (*f*) MEM density with non-uniform prior. Note that the plots have different scales on the abcissa in order to illustrate both the outliers in plots (*a*), (*e*) and (*f*), and the fine features in plots (*b*), (*c*) and (*d*).

The first term is the variance and it can be estimated by

$$\tau_x^2 = [1/(N-1)] \sum_j (\hat{\rho}_x^{i-1,j} - \hat{\rho}_x^{i-1,\mathrm{ave}})^2. \qquad (10)$$

In (10), i denotes the iteration number, j the number of the Monte Carlo density, and x the pixel number. The random error τ_x in $\hat{\rho}_x^i$ is plotted in Fig. 10. Generally, τ_x is very small in the valence regions with a slightly larger value in high-symmetry positions. The largest value, 0.14 e Å^{-3}, is at the beryllium position. For comparison, we show in Fig. 11 the estimated random error distribution of the deformation density, $\Delta\rho = (1/V)\sum_H (F^{\mathrm{obs}}/k) - F^{\mathrm{free\ atom}}$, obtained by Larsen & Hansen (1984) using the method developed by Rees (1978). The error on $\Delta\rho$ contains contributions from both the structure-factor errors and from the least-square errors on the refined structural parameters. However, for beryllium, the errors on the structural parameters are small. The $\Delta\rho$ error estimate is primarily affected by the error in the scale factor, and the effect is almost exclusively seen at the Be nuclear position. In the valence regions, the error on $\Delta\rho$ is probably comparable to estimates of the error in ρ itself. The random error in the MEM density is as expected slightly smaller than for $\Delta\rho$, but in magnitude in very good agreement with the previous error estimate. The values of the random error should be compared to the range -0.74 to 0.12 e Å^{-3} for $b(\rho^0)$, *i.e.* the bias of ρ_x^0 is at some positions five to ten times bigger than the standard error τ_x. In Fig. 12, $b(\hat{\rho}^1)_x - b(\rho^0)_x = b[\rho^0 - b(\rho^0)]_x - b(\rho^0)_x$ is plotted. The range of $b(\hat{\rho}^1)_x - b(\rho^0)_x$ is -0.014 to 0.057 e Å^{-3}, and we may therefore expect that $b(\rho^0 + \delta)_x \simeq b(\rho^0)_x$ when δ is of the order τ. We then have

$$\begin{aligned} E[\hat{\rho}_x^1] &= E[\rho_x^0 - b(\rho^0)_x] \\ &\simeq E[\rho_x^0] - b(\rho^0)_x \\ &= \rho_x^{\mathrm{true}} + b(\rho^{\mathrm{true}})_x - b(\rho^0)_x. \end{aligned} \qquad (11)$$

The second term in (9) can therefore be approximated by $[b(\rho^{\mathrm{true}})_x - b(\rho^0)_x]^2$. We still do not know this quantity, but as a rough guide we can use $[b(\hat{\rho}^1)_x - b(\rho^0)_x]^2$ and as mentioned above this is of the same order of magnitude as τ_x^2.

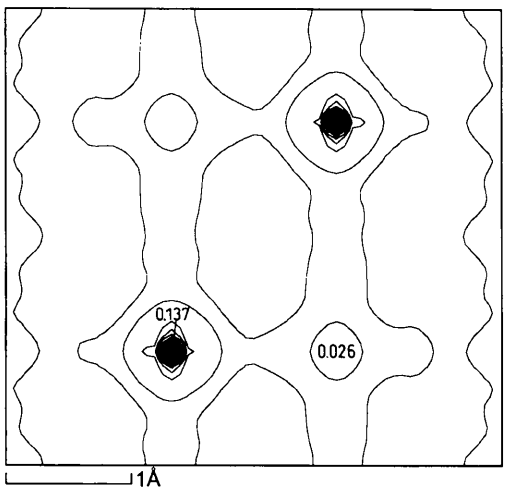


Fig. 10. Contour plot in the (110) plane of the first error term in equation (9). The contours are plotted on a linear scale with 0.0 1e Å^{-3} intervals. The maximum value in the map is 0.137 e Å^{-3} and the minimum is 0.007 e Å^{-3}.

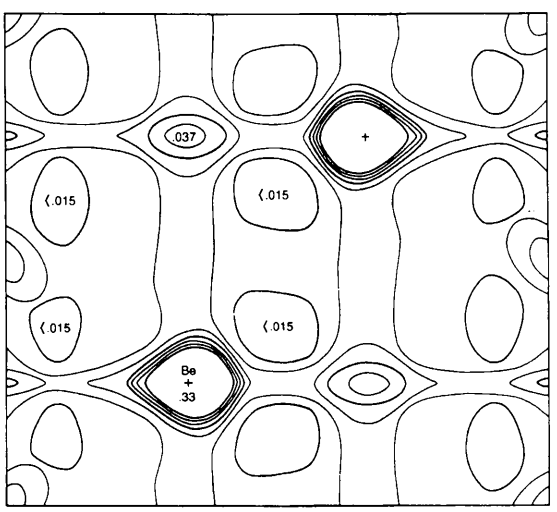


Fig. 11. Estimated error distribution of the deformation density, $\Delta\rho$, in the (110) plane as obtained by Larsen & Hansen (1984) using the method of Rees (1978). Contours are plotted at 0.005 e Å^{-3} intervals starting at 0.015 e Å^{-3}. Contours above 0.045 e Å^{-3} have been omitted.

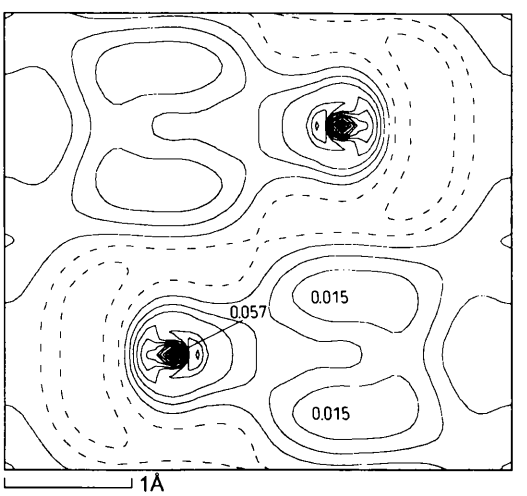


Fig. 12. Contour plot in the (110) plane of the second error term in equation (9). The contours are plotted on a linear scale with 0.005 e Å^{-3} intervals. The maximum value in the map is 0.057 e Å^{-3} and the minimum is -0.014 e Å^{-3}. The broken lines are negative contours and the solid lines are positive contours.

7. Weighting schemes

de Vries *et al.* (1994) as well as Iversen, Nielsen & Larsen (1995) showed that introduction of weighting schemes in the MEM calculations may improve the quality of the derived MEM densities. It was also noted that the distribution of structure-factor residuals of the MEM procedure becomes more uniform when larger weights are given to the low-order reflections during the optimization. de Vries *et al.* (1994) found that using $\sigma_{\rm MEM} = H^2\sigma_F$, H being the length of the scattering vector, gives the most satisfactory results. Iversen, Nielsen & Larsen (1995) used three different physical criteria to judge the densities obtained with different weighting schemes. In that study, use of $\sigma_{\rm MEM} = H\sigma_F$ was found to improve high-temperature Mg nuclear densities, whereas low-temperature densities degraded upon introduction of weights. In the low-temperature densities, spurious features, which were not present in the unweighted densities, were introduced when using the H weighting. The weighting schemes are therefore not an inherent improvement of the MEM but merely an *ad hoc* way to remove certain features in the map. However, the introduction of weights do improve the residual distribution also in the case of the beryllium MEM EDD. In Fig. 13, the MEM density obtained with H^2 weighting is shown. The corresponding residual distribution is shown in Fig. 9(*e*). Fig. 9(*e*) should be compared with Fig. 9(*a*) showing the residual distribution for the unweighted MEM calculation. Because large-amplitude structure factors also tend to be low-order reflections, these reflections carry relatively more weight when $\sigma_{\rm MEM} = H^2\sigma_F$ is used, and the poorly fitting reflections are therefore forced to agree better. However, one could just as well have used a weighting scheme based on the structure-factor amplitude. In some sense, the task is to introduce a weighting scheme that in a precise way counterbalances the MEM bias. A more attractive route than the use of weighting schemes is to introduce additional well defined constraints in the optimization. In the present MEM algorithm, only the variance of the residual distribution is constrained. A much stronger constraint would be to demand that the optimization is unbiased by enforcing the residual distribution to resemble an $N(0,1)$ distribution. However, this will require an extensive rewriting of for instance the *MEED* algorithm.

8. MEM calculations with non-uniform prior distributions

The use of non-uniform priors in MEM charge-density calculations to express *a priori* knowledge/prejudice from other sources than the actual experiment has been discussed by Jauch & Palmer (1993), Jauch (1994) and Iversen, Larsen, Souhassou & Takata (1995). A natural non-uniform prior in the case of beryllium metal is a distribution consisting of two thermally smeared Be atoms placed at their unit-cell positions. The non-uniform prior distribution we have used was calculated using wavefunctions from Clementi & Roetti (1974) and neutron diffraction thermal parameters by Larsen, Brown, Lehman & Merisalo (1982). In Fig. 14, the MEM density obtained with use of this prior is shown. The structure-factor residuals corresponding to this density are plotted in Fig. 9(*f*). The residual distribution is also for this calculation quite non-uniform. The reflection carrying most of the error is again the 101 reflection but, contrary to the calculations with a uniform prior, the MEM density this time predicts a much too large absolute value. It is notable that the density is less smooth and more flat in the valence regions than the density obtained with the uniform prior. The structure factors corresponding to the density shown in Fig. 14 were used as F^0 input to a Monte Carlo calculation. The resulting MEM bias is shown in Fig. 15 and the bias-corrected density is shown in Fig. 16. The bias features outside the Be atom are considerably smaller than the bias obtained with a uniform prior (Fig. 4), which suggests that the non-uniform prior is preferable to the uniform prior. The various features in the density will be discussed below.

9. Discussion

With a knowledge of the errors, it is possible to assess the reliability of various topological features in the MEM density of metallic beryllium. The bias-corrected MEM density obtained with the uniform prior is quite different from the original MEM density especially in the low-density regions. The most notable difference is that the clear non-nuclear maximum in the basal plane (NNM1) has disappeared. The whole area around the bipyramidal

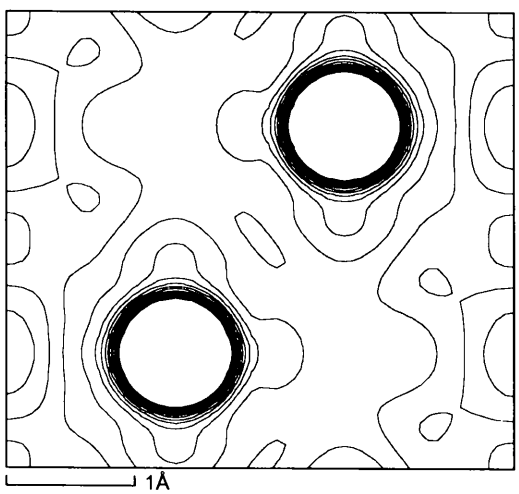


Fig. 13. Contour plot in the (110) plane of the MEM density obtained using H^2 weighting ($\sigma_{\rm MEM} = H^2\sigma_F$). Contours as in Fig. 1.

region of tetrahedral holes is now very flat although still of considerably higher density than the rest of the unit cell outside the Be atomic region. A flattening of NNM1 is also observed if the MEM iterations are allowed to progress to lower values of χ^2. Iversen, Larsen, Souhassou & Takata (1995) noted that at $\chi^2 = 0.5$ NNM1 splits into two new maxima below and above the basal plane. The bias-corrected density is more noisy than the uncorrected density and a topological analysis with the *CPGRID* program (Souhassou, 1993) reveals a very large number of critical points without physical meaning. The topological analysis can therefore only be discussed on a qualitative level. It is notable that the bias-corrected density still contains a region of minimal density around $(0, 0, 1/4)$, which is also the minimum in ρ^0. The region around the octahedral hole, $(0, 0, 0)$, comes out as a weak secondary maximum (NNM2), but it is very flat and of lower density than the bipyramidal region. Based on the bias-corrected density, we must conclude that the topological network proposed by Iversen, Larsen, Souhassou & Takata (1995) is uncertain although the bias-corrected density resembles a flatter version of the uncorrected density. Since the random error in the MEM density is very small, it will probably not influence these conclusions.

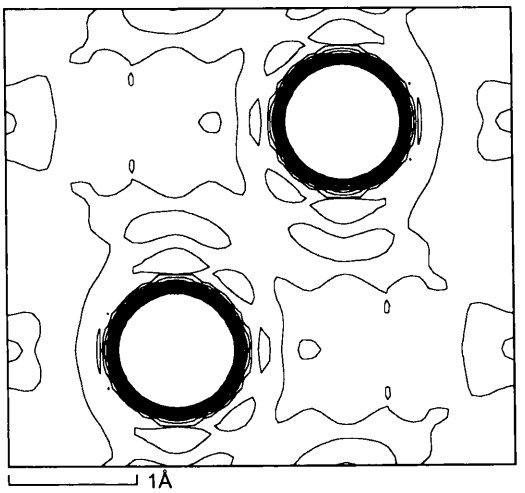


Fig. 14. Contour plot in the (110) plane of MEM density obtained using a non-uniform prior distribution. Contours as in Fig. 1.

The bias-corrected density obtained using a non-uniform prior is even flatter than the bias-corrected density obtained with a uniform prior. In this density, the valence region is almost without any structure and there is less than $0.04\,\mathrm{e\,\AA^{-3}}$ difference between the extremes in the valence regions. As explained above, the MEM density with non-uniform prior also has the largest error carried by the 101 reflection but with an opposite sign compared to the calculation with uniform prior. The 101 reflection is primarily responsible for removing the density from the bipyramidal region (see Fig. 7) and the large error on the 101 reflection results in the MEM density with non-uniform prior instead having an underestimated density in this region. In Table 1, values of the MEM density are compared for some of the high-symmetry points in the structure. It may be that the atomic prejudice in the calculations with a non-uniform prior is too strong to move electrons away from the atom and into the valence regions. Iversen, Larsen, Souhassou & Takata (1995) showed that if the low-order data get increased weight in the MEM calculation with non-uniform prior then the corresponding valence density gets more structure and the topology resembles what

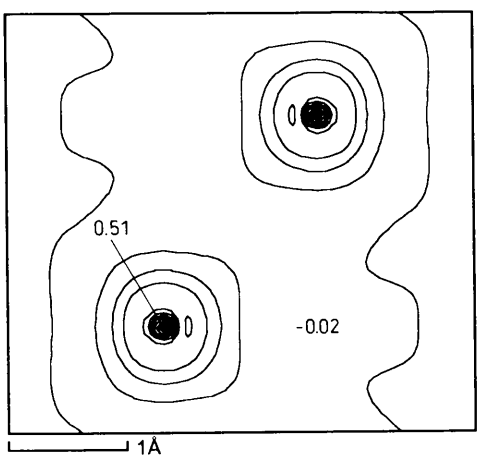


Fig. 15. Contour plot in the (110) plane of the bias distribution, $b(\rho^0) = \rho^{\mathrm{ave}} - \rho^0$, for the MEM density obtained with a non-uniform prior. The contours are plotted on a linear scale with $0.05\,\mathrm{e\,\AA^{-3}}$ intervals. The maximum value in the map is $0.506\,\mathrm{e\,\AA^{-3}}$ and the minimum is $-0.035\,\mathrm{e\,\AA^{-3}}$. The density contains regions with slight negative bias (< -0.05), which are not seen in the plot.

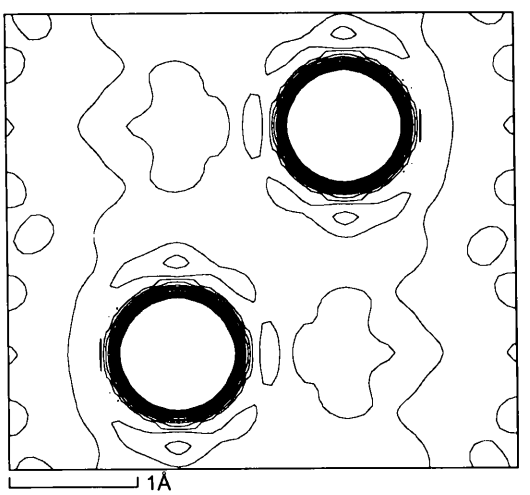


Fig. 16. Contour plot in the (110) plane of the bias-corrected MEM density obtained using a non-uniform prior distribution. Contours as in Fig. 1.

Table 1. *Values of various MEM densities in high-symmetry positions*

The values are given in e Å^{-3}. For each point, the corresponding pixel values in the 60 × 60 × 60 MEM grid are listed. Numbers in parentheses are the standard uncertainties.

	Be (20, 40, 15)	NNM1 (40, 20, 15)	NNM2 (0, 0, 0)	MIN (0, 0, 15)
Uniform prior	48.10	0.41	0.22	0.14
Uniform prior, bias corrected	48.84 (0.14)	0.29 (0.03)	0.20 (0.02)	0.20 (0.01)
Non-uniform prior	53.49	0.20	0.22	0.24
Non-uniform prior, bias corrected	52.98 (0.09)	0.23 (0.01)	0.20 (0.01)	0.20 (0.01)
Uniform prior, H^2 weighting	48.35	0.28	0.27	0.11
Non-uniform prior, H^2 weighting	53.86	0.31	0.20	0.23

is found with a uniform prior. In this respect, it should be mentioned that Iversen, Larsen, Souhassou & Takata (1995) also carried out conventional multipole modeling in order to test the results. The density obtained from a multipole model contains both NNM1 and the cage critical point at (0, 0, 1/4). Overall, it therefore seems probable that there is charge build up in the bipyramidal region but whether the maximum is in the basal plane or for instance in the tetrahedral holes is uncertain. A number of theoretical studies also indicate a charge build up in this region of the structure (Dovesi, Pisani, Ricca & Roetti, 1982; Chou, Lam & Cohen, 1983; Ross, Ermler, Kern & Pitzler, 1992).

de Vries, Briels & Feil (1996) have recently repeated the MEM calculations on Be with a non-uniform prior density previously reported by Iversen, Larsen, Souhassou & Takata (1995). These calculations indicate a minimum at (2/3, 1/3, 1/4). This is probably, as discussed above, because of too strong atomic prejudice in the prior. In conclusion, a network containing NNM1 either in the basal plane or in the tetrahedral hole, NNM2 in the octahedral hole and the cage critical point in (0, 0, 1/4) seems, based on the present analysis, the most convincing possibility.

10. Conclusions

A Monte Carlo type scheme to estimate the mean-square error in MEM charge densities has been presented. The Monte Carlo calculations show that the MEM densities are affected by systematic bias that is intimately related to the non-uniform distribution of the structure-factor residuals, $\Delta F_j/\sigma(F_j)$, after a MEM optimization. A scheme to estimate the systematic bias and correct the density has been presented. The bias-corrected density has a more uniform residual distribution than the uncorrected density. The study shows that the systematic bias in some regions is an order of magnitude larger than the random error in the MEM density. Correction for systematic bias may therefore be very important when interpreting fine details in MEM charge densities. The value of χ^2 is lower than 1 after the bias correction and the scheme is therefore only partially successful in correcting for the systematic bias of the MEM. It would be desirable if the systematic bias of the MEM was counterbalanced during the optimization process. This may be possible by introducing a new constraint that constrains the residual distribution to an $N(0,1)$ distribution. This will, however, demand an extensive rewriting of the *MEED* algorithm. Previously, various weighting schemes have been used in MEM calculations to reduce the bias of the method. For metallic beryllium, the use of $\sigma_{\rm MEM} = H^2\sigma_F$, H being the length of the scattering vector, reduces the bias slightly but in general the use of weighting schemes seems to be an *ad hoc* way to interfere with the MEM optimization. The weighting schemes are not optimal in counterbalancing the systematic bias of the MEM procedure.

The MEM density obtained with a non-uniform prior also contains systematic bias although less than the MEM density with uniform prior. This indicates that calculations with non-uniform priors lead to physically sounder results and this aspect of the MEM deserves further attention. In this way, more optimal non-uniform priors may be developed.

The critical-point network of metallic beryllium was discussed based on the bias-corrected MEM densities and the corresponding error estimates. It seems that the valence features are flatter than originally proposed by Iversen, Larsen, Souhassou & Takata (1995) but overall the present analysis qualitatively supports the network established in that study. The flatness of the valence features corroborates the fact that beryllium is an excellent conductor.

Finally, it should be reiterated that, because the present MEM algorithm does not contain a model, it cannot filter out inconsistencies in the data stemming from systematic errors. The MEM densities may therefore in the general case, where the data are of lesser quality than for Be, contain unphysical features, not only because of systematic bias in the calculation but also because of systematic errors in the data.

BBI gratefully acknowledges support for this work from the Danish Natural Science Research Council and the Carlsberg Foundation. Lektor Finn Krebs Larsen is thanked for encouragement, for fruitful discussions and for supplying Figs. 2 and 11. Mr Arne Lindahl is thanked for his kind assistance with plots and drawings.

References

Bader, R. F. W. (1991). *Atoms in Molecules. A Quantum Theory.* Oxford University Press.

Chandler, G. S., Figgis, B. N., Reynolds, P. A. & Wolff, S. K. (1994). *Chem. Phys. Lett.* **225**, 421–426.

Chou, M. Y., Lam, P. K. & Cohen, M. L. (1983). *Phys. Rev. B*, **32**, 1356–1358.
Clementi, E. & Roetti, C. (1974). *At. Data Nucl. Data Tables*, **14**, 177–478.
Collins, D. M. (1982). *Nature (London)*, **298**, 49–51.
Donoho, D. L., Johnstone, I. M., Hoch, J. C. & Stern, A. S. (1992). *J. R. Stat. Soc.* B**54**, 41–81.
Dovesi, R., Pisani, C., Ricca, F. & Roetti, C. (1982). *Phys. Rev. B*, **25**, 3731–3739.
Figgis, B. N., Iversen, B. B., Larsen, F. K. & Reynolds, P. A. (1993). *Acta Cryst.* B**49**, 794–806.
Figgis, B. N., Reynolds, P. A. & Williams, G. A. (1980). *J. Chem. Soc. Dalton Trans.* pp. 2339–2347.
Hansen, N. K. & Coppens, P. (1978). *Acta Cryst.* A**34**, 909–921.
Hirshfeld, F. L. (1977). *Isr. J. Chem.* **16**, 226–229.
Iversen, B. B., Larsen, F. K., Figgis, B. N. & Reynolds, P. A. (1995). *Trans. Am. Crystallogr. Assoc.* **31**, 1–10.
Iversen, B. B., Larsen, F. K., Souhassou, M. & Takata, M. (1995). *Acta Cryst.* B**51**, 580–592.
Iversen, B. B., Nielsen, S. K. & Larsen, F. K. (1995). *Philos. Mag.* A**72**, 1357–1380.
Jauch, W. (1994). *Acta Cryst.* A**50**, 650–652.
Jauch, W. & Palmer, A. (1993). *Acta Cryst.* A**49**, 590–591.
Kumazawa, S., Kubota, Y., Takata, M., Sakata, M. & Ishibashi, Y. (1993). *J. Appl. Cryst.* **26**, 453–457.
Kumazawa, S., Takata, M. & Sakata, M. (1995). *Acta Cryst.* A**51**, 453–457.
Larsen, F. K., Brown, P. J., Lehman, M. S. & Merisalo, M. (1982). *Philos. Mag.* B**45**, 31–50.
Larsen, F. K. & Hansen, N. K. (1984). *Acta Cryst.* B**40**, 169–179.
Papoular, R. J., Vechter, Y. & Coppens, P. (1996). *Acta Cryst.* A**52**, 397–408.
Press, W. H., Teukolsky, S. A., Vetterling, N. T. & Flannery, B. P. (1992). *Numerical Recipes*. Cambridge University Press.
Rees, B. (1978). *Acta Cryst.* A**34**, 254–256.
Ross, R. B., Ermler, W. C., Kern, C. W. & Pitzler, R. M. (1992). *Int. J. Quantum Chem.* **41**, 733–747.
Sakata, M. & Sato, M. (1990). *Acta Cryst.* A**46**, 263–270.
Sakata, M., Uno, T., Takata, M. & Howard, C. J. (1993). *J. Appl. Cryst.* **26**, 159–165.
Souhassou, M. (1993). Program *CPGRID*. Laboratoire de Minéralogie et Cristallographie, Université de Nancy I, BP 239, 54506 Vandoeuvre CEDEX, France.
Stewart, R. F. (1976). *Acta Cryst.* A**32**, 565–574.
Takata, M., Kubota, Y. & Sakata, M. (1993). *Z. Naturforsch. Teil A*, **48**, 75–80.
Takata, M. & Sakata, M. (1996). *Acta Cryst.* A**52**, 287–291.
Takata, M., Sakata, M., Kumazawa, S., Larsen, F. K. & Iversen, B. B. (1994). *Acta Cryst.* A**50**, 330–337.
Takata, M., Umeda, B., Nishibori, E., Sakata, M., Saito, Y., Ohno, M. & Shinoshara, H. (1995). *Nature (London)*, **377**, 46–49.
Vries, R. Y. de, Briels, W. J. & Feil, D. (1994). *Acta Cryst.* A**50**, 383–391.
Vries, R. Y. de, Briels, W. J. & Feil, D. (1996). *Phys. Rev. Lett.* **7**, 1719–1722.