Acta Cryst A
FOUNDATIONS
ADVANCES

ISSN 2053-2733

Image definition evaluation functions for X-ray crystallography: a new perspective on the phase problem

Hui Li,[a]* Meng He[b]* and Ze Zhang[c]

[a]Beijing University of Technology, Beijing, 100124, People's Republic of China, [b]CAS Key Laboratory of Nanosystem and Hierarchical Fabrication, National Center for Nanoscience and Technology, Beijing, 100190, People's Republic of China, and [c]Zhejiang University, Hangzhou, 310014, People's Republic of China. *Correspondence e-mail: huilicn@yahoo.com, mhe@nanoctr.cn

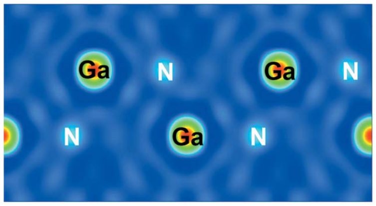
Received 5 March 2015
Accepted 24 June 2015

Edited by J. Miao, University of California, Los Angeles, USA

Keywords: phase problem; image definition evaluation functions; power spectrum entropy; iteration algorithms; charge-density maps.

Supporting information: this article has supporting information at journals.iucr.org/a

The core theme of X-ray crystallography is reconstructing the electron-density distribution of crystals under the constraints of observed diffraction data. Nevertheless, reconstruction of the electron-density distribution by straightforward Fourier synthesis is usually hindered due to the well known phase problem and the finite resolution of diffraction data. In analogy with optical imaging systems, the reconstructed electron-density map may be regarded as the image of the real electron-density distribution in crystals. Inspired by image definition evaluation functions applied in the auto-focusing process, two evaluation functions are proposed for the reconstructed electron-density images. One of them is based on the atomicity of the electron-density distribution and properties of Fourier synthesis. Tests were performed on synthetic data of known structures, and it was found that this evaluation function can distinguish the correctly reconstructed electron-density image from wrong ones when diffraction data of atomic resolution are available. An algorithm was established based on this evaluation function and applied in reconstructing the electron-density image from the synthetic data of known structures. The other evaluation function, which is based on the positivity of electron density and constrained power spectrum entropy maximization, was designed for cases where only diffraction data of rather limited resolution are available. Tests on the synthetic data indicate that this evaluation function may identify the correct phase set even for a data set with resolution as low as 3.5 Å. Though no algorithm for structure solution has been figured out based on the latter function, the results presented here provide a new perspective on the phase problem.



1. Introduction

It is interesting to compare X-ray crystallography with the optical imaging process (Miao *et al.*, 2015). In an optical imaging system, an object is illuminated by visible light, and the scattered light is then combined by the lenses to form an image which represents the object. In the case of X-ray crystallography, a crystal is illuminated by an X-ray beam and the diffraction data are then measured directly since no appropriate 'lens' is available for X-rays. One has to reconstruct the electron-density image of the crystal from the diffraction data, which has been the core theme of X-ray crystallography since its advent. Theoretically, the electron-density image of a crystal is just the inverse Fourier synthesis of its structure factors. Unfortunately, phases of structure factors are rarely recorded in diffraction experiments and, usually, only amplitudes of structure factors can be deduced from the diffraction data. Therefore, the electron-density distribution of crystals cannot be reconstructed by a straightforward Fourier synthesis. Even worse, the phases 'lost' in diffraction experiments are more important than amplitudes of structure factors in defining the electron-density distribution (Lattman & Derosier, 2008). This is the so-called 'phase problem' of X-ray crystallography.

Fourier synthesis can be made after assigning an arbitrary phase set to the observed moduli, but the resultant Fourier map generally will not represent the real electron-density distribution in the crystal. If all possible phase sets for the observed moduli can be tested in a limited time, the correct phase set and subsequently the electron-density image which represents the crystal must be included in these tests. Then the 'phase problem' is converted to the question of how to pick out the correct electron-density image among all the possible Fourier synthesis maps. Such a question is reminiscent of the auto-focusing process which is extensively performed in optical imaging devices. In the auto-focusing process, the image definition evaluation functions are calculated in real time and focusing is then tuned accordingly. The image definition evaluation functions describe the properties of photographs mathematically (Chen *et al.*, 2013) and the properly focused image is then picked out based on the evaluation functions. Inspired by this, we try to establish evaluation functions for an electron-density image reconstructed from diffraction data. Hopefully, the evaluation function can pick out the electron-density image which represents the crystal. For this purpose, the evaluation function should describe the general characteristics of the electron-density distribution in real crystals, and vary sensitively with phase sets assigned to the moduli to distinguish the faithfully reconstructed electron-density distribution from the artifacts.

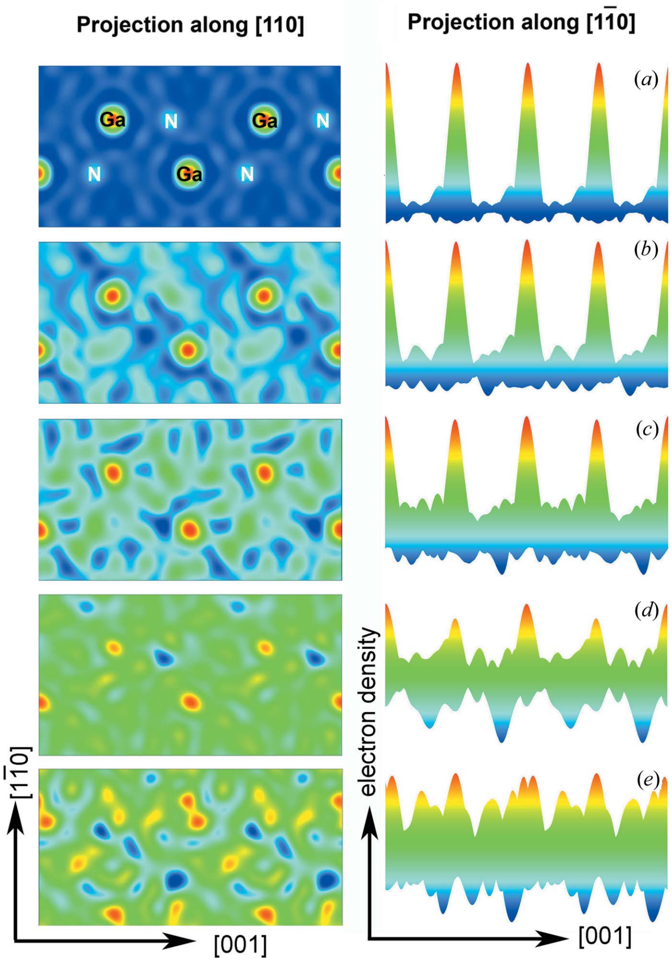


Figure 1
Electron-density distribution on the (110) plane of hexagonal GaN calculated by Fourier synthesis using synthetic structure-factor moduli and (*a*) correct phases, (*b*) 25% random phases, (*c*) 50% random phases, (*d*) 75% random phases and (*e*) 100% random phases. The resolution of the synthetic diffraction data is 0.8 Å.

The commonly recognized characteristics of electron density in crystals are atomicity and positivity. Actually, the atomicity and positivity have been extensively utilized in many (if not all) methods for structure solutions by X-ray diffraction. For example, the Patterson method (Patterson, 1934) and the recently popular charge-flipping algorithm (Oszlányi & Sütő, 2004, 2008) are based on atomicity. Atomicity and positivity are also prerequisites of the direct method (Woolfson, 1987; Woolfson & Fan, 1995), though apparently it works in reciprocal space. The success of the Patterson method, the direct method and the charge-flipping algorithm indicates that electron-density images can be faithfully reconstructed by combining the constraints of atomicity and/or positivity and observed structure-factor moduli, at least in some cases.

Information theory provides another constraint on the reconstructed electron-density image: the entropy of the reconstructed electron-density distribution should be maximized under the constraints of observed structure-factor moduli. An iterative procedure has been established based on the constrained entropy maximization for reconstructing an electron-density distribution from X-ray diffraction data (Collins, 1982; Sakata & Sato, 1990). This algorithm is usually referred to as the maximum entropy method (MEM) (Wu, 1997). The bound of positivity is usually maintained by the electron-density distribution reconstructed with the MEM.

In this study, we try to construct image definition evaluation functions for an electron-density image based on atomicity, positivity and constrained power spectrum entropy maximization, the validity of which in structure solutions has been confirmed in practice in past decades. Two evaluation functions are proposed. One of them is based on atomicity, applicable for cases where diffraction data of atomic resolution are available; the other combines the bound of positivity and constrained power spectrum entropy maximization, aiming for cases where only diffraction data of rather limited resolution are available.

2. Image definition of electron-density image and phases of structure factors

The concept of image definition has been accepted extensively but its accurate definition is still controversial. We prefer to define the image definition as a quantity which measures how well an image represents the object. The more information about the object is included in an image, the better the image definition. Many image definition evaluation functions have been proposed to assess the definition of an image quantitatively. These evaluation functions are calculated based on the image itself.

In comparison with the image definition of a photograph, the image definition of a reconstructed electron-density image seems to be more elusive, possibly because the real electron-density distribution of a crystal cannot be viewed directly by eye. The pictures shown in Fig. 1 may be helpful to understand the concept of image definition of a reconstructed electron-density image. The electron-density image of GaN reconstructed by combining the structure-factor moduli (up to 0.8 Å) and correct phases represents the structure well (Fig. 1*a*). With more and more phases of structure factors being replaced by random values, the information on the structure contained in the reconstructed electron-density images decreases gradually. In the electron-density image reconstructed with 25% random phases, the positions of N atoms can hardly be identified (Fig. 1*b*). When 50% or 75% of phases are replaced by random values, only the positions of Ga atoms can be discerned (Figs. 1*c* and 1*d*). In the electron-density image reconstructed with all random phases (Fig. 1*e*), no information on the structure is available. The information on the structure carried by the reconstructed electron-density image decreases with the increasing amount of phases being replaced by random values, resulting in worse image definition. The example shown in Fig. 1 not only illustrates the concept of image definition of electron-density images, but also demonstrates that the image definition varies with the phases of structure factors, and the best image definition is achieved when the correct phases are applied.

3. Image definition evaluation function for diffraction data of atomic resolution

3.1. Construction of image definition evaluation function

As mentioned above, many image definition evaluation functions have been established to assess the quality of photographs. An important kind of evaluation function is based on the idea that the more high-frequency information contained in the image, the better the image definition is. The energy of high-frequency information of an image is used to measure the image definition. However, this idea cannot be applied directly in the evaluation of the reconstructed electron-density images of crystals, because all the Fourier synthesis maps created by combining fixed structure-factor moduli with various phase sets will have the same energy of high-frequency information, which depends solely on the moduli of structure factors. Hence, evaluation functions for the reconstructed electron-density images have to be based on other intrinsic characteristics. As illustrated in Fig. 1, in a correctly reconstructed electron-density image, most electrons are concentrated at the positions of atoms while only a few electrons are distributed in the intermediate region between atoms. This is the well known atomicity of electron-density distribution. On the contrary, when random phases are assigned to the structure-factor moduli, the resultant electron-density image shows a much more even distribution of electrons. It implies that the extent to which the electrons are concentrated can be used to evaluate the image definition of the reconstructed electron-density image. The ratio of the sum of electron density at the atomic region to that in the intermediate region seems to be a good indicator to measure the extent to which the electrons are concentrated. Unfortunately, this ratio cannot be calculated before the structure is determined because one cannot discern the atomic and intermediate regions. An alternative way is to set a slightly positive threshold of electron density ρ_t, and use a quantity such as $\sum \rho_i^*$ (where $\rho_i^* > \rho_t$) to measure the extent to which the electrons are concentrated in the reconstructed electron-density images. Nevertheless, the threshold should be structure dependent, and it might be difficult to give a reasonable estimate of ρ_t.

Because of the truncation effect of Fourier synthesis, both positive and negative electron density will be observed in the reconstructed electron-density image. In a correctly reconstructed image, high positive electron density will appear at the position of atoms while in the intermediate region between atoms there are both small positive and negative electron densities. According to the property of the Fourier transform, for any electron-density image, the sum of positive density is equal to that of negative density when F_{000} is not included in the Fourier synthesis. This means that all negative density lies in the intermediate region while most positive density is concentrated at the atomic positions. In terms of Parseval's theorem, for any given group of structure-factor moduli we have

$$\sum_i \rho_+^2 + \sum_j \rho_-^2 = \text{constant}$$

where ρ_+ and ρ_- represent non-negative and negative density, respectively, at a certain grid of the electron-density image. Parseval's theorem and a brief derivation of the above corollary are presented in the supporting information. As discussed above, for a correctly reconstructed electron-density image, there is very little positive electron density in the intermediate region between atoms, so $\sum \rho_+^2$ is a good approximation of the sum of squared electron density at the atomic positions ($\sum \rho_{\rm atom}^2$), while $\sum \rho_-^2$ dominates the sum of squared electron density in the intermediate regions between atoms ($\sum \rho_{\rm inter}^2$). Hence, the ratio $\sum \rho_+^2 / \sum \rho_-^2$ is a good approximation for $\sum \rho_{\rm atom}^2 / \sum \rho_{\rm inter}^2$, which measures well the extent to which the electrons are concentrated. Thus, we establish the first image definition evaluation function for the reconstructed electron-density images:

$$\text{Tian1} \equiv \frac{\sum_i \rho_+^2}{\sum_j \rho_-^2}.$$

3.2. Verification and limitation

Tests are performed on the synthetic diffraction data of ten known structures to verify the validation of Tian1 and find its limitations. The moduli of structure factors calculated from the known structure data are taken as the synthetic diffraction data. A series of Fourier synthesis maps is then created by combining these structure-factor moduli with correct phases, all-zero phases and random phase sets. Evaluation function Tian1 is calculated for each Fourier synthesis map. Typical results of such tests are presented in Fig. 2. As revealed by Fig. 2, random phases will lead to a quite low Tian1 value, which is close to 1. This is reasonable because the random phases will lead to a random distribution of electron density, implying that statistically $\sum \rho_{+}^{2}$ is equal to $\sum \rho_{-}^{2}$. Both correct phases and all-zero phases result in high Tian1 values, indicating that electrons are well concentrated in the reconstructed electron-density images. When data of atomic resolution (1.0 Å) are available, the correct phases give a higher Tian1 value than all-zero phases do. When the resolution of the data is worse, decreased Tian1 values are observed for both the correct and all-zero phases. Unfortunately, the Tian1 value for the correct phase set decreases faster than that for all-zero phases. At the resolution of 1.5 Å, comparable Tian1 values are obtained for the correct phase set and all-zero phases. This implies the invalidation of Tian1 in identifying the correctly reconstructed electron-density image. The invalidation of Tian1 in cases where data of atomic resolution are not available can be well understood. In such cases, positive electron density will no longer be mainly concentrated at the positions of atoms in a correctly reconstructed electron-density image. Significant positive electron density will appear in the intermediate regions between atoms. On the contrary, all-zero phases always lead to the concentration of electron density at the origin of the unit cell. The evaluation function Tian1 measures the extent to which the electron density is concentrated, so it works well only when data of atomic resolution are available. The resolution limit for Tian1 might be structure dependent, but 1.2–1.8 Å seems to be a reasonable estimate.

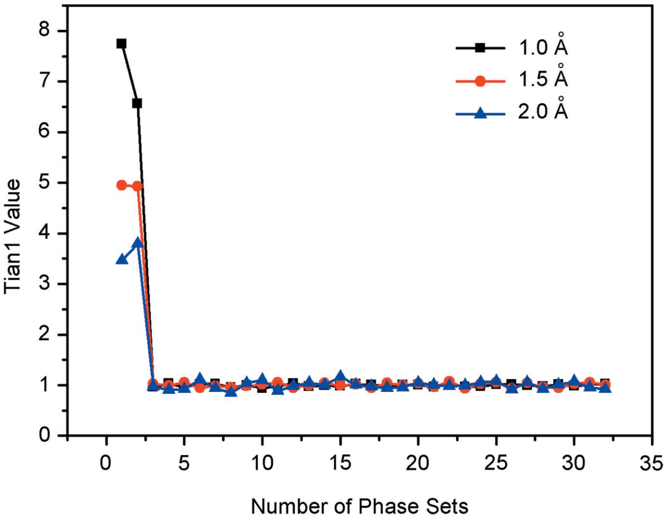
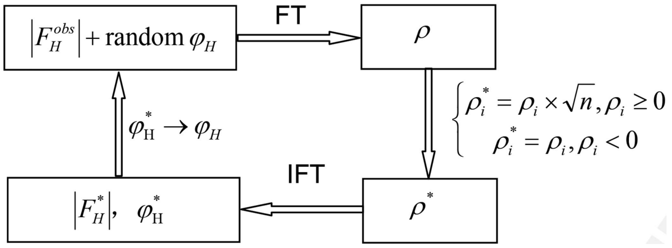


Figure 2
Evaluation function Tian1 of electron-density images created by Fourier synthesis using synthetic structure-factor moduli of $Al(IO_3)_3(H_2O)_8$ and correct phases, all-zero phases or random phases. The first phase set consists of correct phases, the second one is made up of all-zero phases, and all other phase sets are totally random. The unit cell is divided into 64 × 64 × 128 grids.

3.3. Algorithm for structure solution based on Tian1

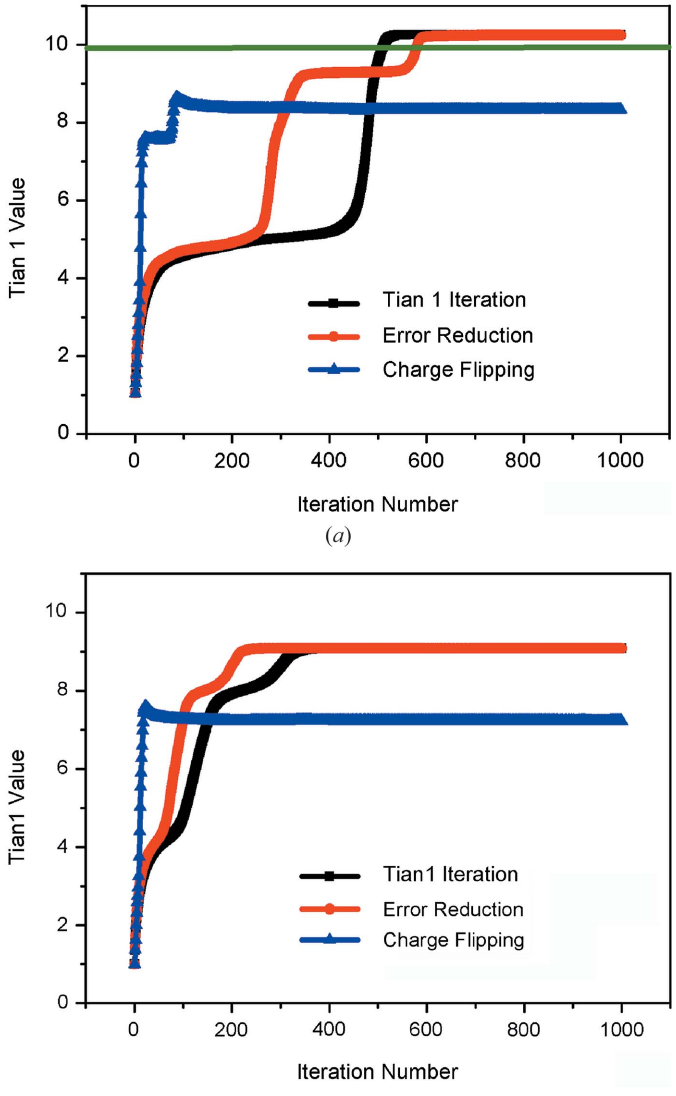
An algorithm for structure solution has been established based on the evaluation function Tian1. The flow chart of the algorithm is given as follows:



At the very beginning of the iterative process, a random phase is assigned to each observed structure-factor modulus, and then an electron-density image is calculated by Fourier synthesis. The original electron-density image ρ is subsequently modified by multiplying each non-negative electron density by $n^{1/2}$ (n is an arbitrary number greater than 1) while keeping all negative electron densities unchanged. The purpose of electron-density modification is to improve the Tian1 value of the original electron-density image by n times. The modified electron-density image ρ^* is then inversely Fourier transformed to generate new moduli and phases of structure factors. The latest phases are then combined with the observed structure-factor moduli to calculate a new electron-density image. The iterative process is repeated until a maximized Tian1 value has been achieved.

The algorithm is tested on synthetic diffraction data of known structures at a resolution of 0.8 Å. In the tests, symmetry was not taken into account and all structures were solved in the $P1$ space group. The resultant final electron-density images obtained by this algorithm were compared with the known crystal structure by superimposing the electron-density images upon the atomic structure model using the software *VESTA* (Momma & Izumi, 2008). Electron-density images are thought to be correct when they are consistent with the atomic structure model. It is worth noting that atomic coordinates in the structure model usually have to be translationally shifted and/or inverted in some cases to make a direct comparison with the electron-density images. The largest structure that has been tested with this algorithm is $K_{14}((H_2O)W_{19}P_2O_{69})(H_2O)_{24}$ (ICSD-66144), which contains 2064 non-hydrogen atoms in the unit cell including 112 K and 32 O undetermined. This structure was solved successfully in the $P1$ space group with the algorithm presented here, and all positions of the non-hydrogen atoms in the known structure model can be identified in the final reconstructed electron-density image.

This algorithm is somewhat similar to the auto-focusing process. In the auto-focusing process, the focusing is tuned to achieve the extremum of an image definition evaluation function. In the algorithm based on Tian1, the electron-density image is modified to maximize the Tian1 value. The algorithm is a dual-space method, and the similarity between this algorithm and the charge-flipping as well as the error-reduction algorithms (Fienup, 1982; Stark, 1987) is apparent: in all cases, the electron density is modified iteratively to converge to an image which represents the crystal. Computer programs are also developed based on the charge-flipping and error-reduction algorithms. All three algorithms are tested on the synthetic data of $Al(IO_3)_3(H_2O)_8$ (ICSD-200619) at a resolution of 0.8 Å. $Al(IO_3)_3(H_2O)_8$ crystallizes in the $P1$ space group, and the lattice parameters are $a = 7.473$, $b = 8.392$, $c = 13.574$ Å, $\alpha = 89.93$, $\beta = 87.62$, $\gamma = 63.58°$. There are 42 non-hydrogen atoms in the unit cell of $Al(IO_3)_3(H_2O)_8$ and all hydrogen atoms are undetermined in the structure model. The structure can be solved with all three algorithms. The Tian1 value is calculated in each iteration cycle for all three algorithms and presented in Fig. 3(*a*). For the algorithm proposed here, the Tian1 value increases monotonically in the iteration process as expected. It is interesting to note that this value increases monotonically also in the error-reduction iteration process. It is slightly different in the case of the charge-flipping algorithm. The value of the Tian1 function increases rapidly in the iteration process and soon an extremum is reached. After that this value decreases a little, and then is maintained at a level close to the extremum. The extremum achieved by the error-reduction and our algorithm is slightly larger than the value corresponding to the Fourier synthesis maps created using the calculated structure factors. This indicates that over-concentration of the electron density takes place in the iteration process. On the contrary, the extremum of the Tian1 function reached by the charge-flipping algorithm is significantly lower than the value corresponding to the theoretical Fourier synthesis map. This implies that electrons have not been sufficiently concentrated in the electron-density image obtained with the charge-flipping algorithm. Nevertheless, electron-density images created with all three algorithms are good enough to identify the positions of non-hydrogen atoms, as demonstrated by Fig. S1 in the supporting information.



(*a*)

(*b*)

Figure 3
The evolution of evaluation function Tian1 in the iterative cycles of structure solution of $Al(IO_3)_3(H_2O)_8$ from the simulated error-free data (*a*) and the simulated data with random uncertainties (*b*). The green horizontal line in (*a*) indicates the Tian1 value of the Fourier synthesis map created using calculated structure factors. The resolution of synthetic diffraction data is 0.8 Å.

Although the Tian1 value evolves similarly in the iterative process of all three algorithms, the algorithm based on the Tian1 value differs from the charge-flipping and error-reduction algorithms in that the goal of the iterative process is improving the Tian1 value for the former, while for the latter two the Tian1 value just increases unintentionally with the evolution of electron-density distribution. The variation of the Tian1 value observed in the iterative process of the charge-flipping and error-reduction algorithms further corroborates the idea that the image definition of the reconstructed electron-density images can be used to identify the correct phase assignment, and the Tian1 value is effective in evaluating the image definition of the electron-density images reconstructed from atomic resolution data.

A preliminary check has been performed to evaluate the effect of the uncertainties of the data on the applicability of the algorithm based on Tian1. We modified each calculated structure-factor modulus of $Al(IO_3)_3(H_2O)_8$ by multiplying it with $(1 + x \times 30\%)$ (where x is a random value in the range from -1 to 1) to simulate the experimental diffraction data with random errors. The structure can be retrieved from such a data set by the algorithm based on Tian1, the charge-flipping and error-reduction algorithms. The evolution of Tian1 in the iterative process is similar to that observed in the case of error-free data, as shown in Fig. 3(*b*). With the introduction of random errors in the simulated data, the maximum Tian1 values obtained in the iterative process decrease slightly in comparison with those resulting from error-free data for all three algorithms. Interestingly, the structure is retrieved with fewer iteration cycles from the data with random errors.

4. Image definition evaluation function for diffraction data of limited resolution

4.1. Construction of image definition evaluation function

As discussed above, the evaluation function Tian1 is valid only for a diffraction data set of atomic resolution. However, it is more desirable to establish an evaluation function which is applicable for the electron-density image reconstructed from diffraction data at rather limited resolution. Methods of structure solution from diffraction data at atomic resolution, such as the direct method, Patterson method and charge-flipping algorithm, have been well developed. Nevertheless, in many cases, such as in protein crystallography, diffraction data of atomic resolution are only rarely available. Usually, techniques of isomorphous replacement or anomalous dispersion have to be used to solve the phases. Anomalous dispersion requires the presence of heavy atoms in the structure, and the preparation of isomorphous heavy-atom derivative crystals is not trivial. Hence, the technique of reconstructing the electron-density image directly from a diffraction data set of low resolution is still strongly desired.

The positivity constraint on the reconstructed electron-density image is independent of the resolution. Under the constraint of positivity, a correctly reconstructed electron-density image should give the structure factors consistent with the observed data. For any phase set which is assigned to the observed moduli, the consistence between the calculated and 'observed' structure factors can be measured by the residual factor, which is defined as

$$R_{\rm mem} = \frac{1}{M}\sum \frac{(F^H_{\rm obs} - F^H_{\rm cal})^2}{(\sigma F^H_{\rm obs})^2}$$

where $F^H_{\rm obs}$ is the 'observed' structure factor, $F^H_{\rm cal}$ is the calculated structure factor, $\sigma F^H_{\rm obs}$ is the standard deviation of $F^H_{\rm obs}$ and M is the number of structure factors. The 'observed' structure factor is the combination of observed modulus and assigned phase. $F^H_{\rm cal}$ is generated by calculating the Fourier transform of the electron-density distribution.

In MEM calculations, the entropy of an electron-density image is defined as

$$H = -\sum_N \left(\frac{\rho_i}{Z}\right) \ln\left(\frac{\rho_i}{Z}\right)$$

where ρ_i is electron density at the ith grid of the unit cell, Z is the sum of ρ_i across the unit cell, and N is the number of grids in the unit cell. Reconstructing an electron-density image with the MEM means modifying the initial electron-density distribution iteratively to reach the lowest value of $R_{\rm mem}$ subject to maximization of the entropy.

It was found that the residual factor $R_{\rm mem}$, which combines the constraints of positivity and observed data on the electron-density image, is sufficient to identify the correct image when high-resolution diffraction data are available. When the resolution of diffraction data deteriorates, fewer constraints from the observed data are available on the reconstructed electron-density image. Then iteration of the MEM is more readily converged, yielding quite high entropy H and rather low $R_{\rm men}$. Unfortunately, in such cases, a low $R_{\rm mem}$ is no longer a guarantee of the correctness of the reconstructed electron-density image. An additional constraint has to be found to identify the correctly reconstructed image. The entropy H seems to be a good choice. Nevertheless, the entropies H of different electron-density images reconstructed with the MEM cannot be compared with each other directly since they are obtained under the constraint of various $R_{\rm mem}$. Here we introduce the power spectrum entropy S, which is defined as

$$S = -\sum_N \left(\frac{\rho_i^2}{C}\right) \ln\left(\frac{\rho_i^2}{C}\right)$$

where C is the sum of $\rho_i{}^2$ across the unit cell. Different from the entropy H which is calculated from a positive definite electron-density image, the power spectrum entropy S is based on a Fourier synthesis map. According to Parseval's theorem, for any given group of structure-factor moduli, C is a constant. The power spectrum entropy S of a Fourier synthesis map may be taken as the entropy H of an image of $\rho_i{}^2$. Then the power spectrum entropy S measures the extent to which the image of $\rho_i{}^2$ is close to a flat distribution. Because of the correlation between the image of $\rho_i{}^2$ and the Fourier synthesis map, we suppose that S also measures the flatness of the Fourier synthesis map approximately. We construct an evaluation function by combining S and $R_{\rm mem}$, which is defined as

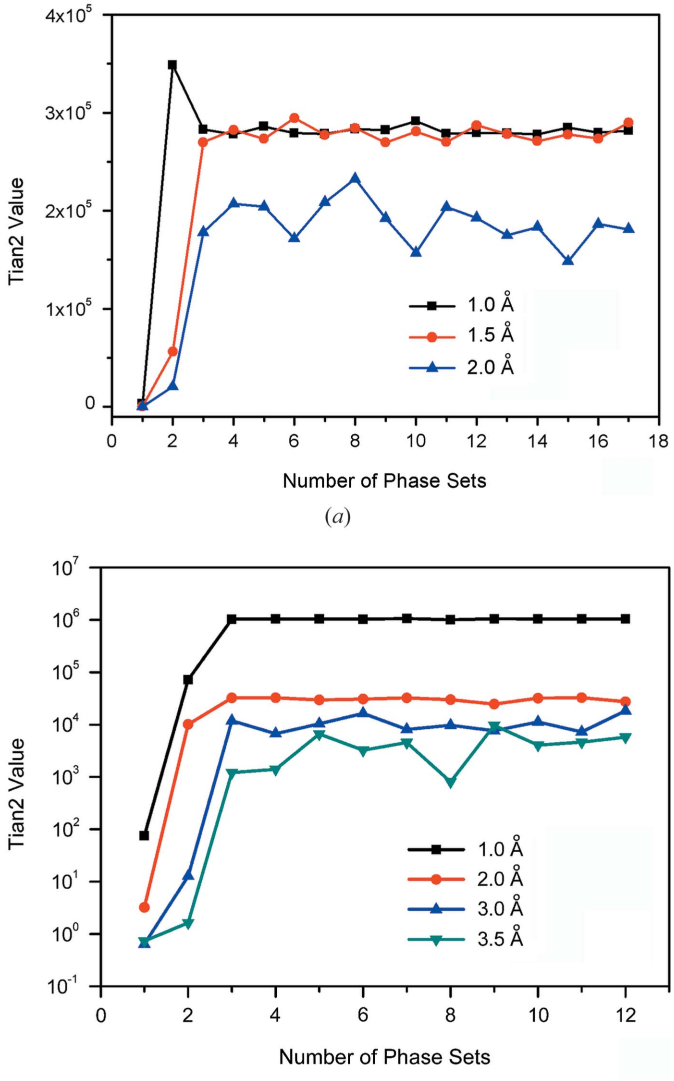
$$\text{Tian2} \equiv R_{\rm mem} \times \exp(S_{\rm ideal} - S)$$

where $S_{\rm ideal} = \ln N$, N is the number of grids in the unit cell. $S_{\rm ideal}$ is the power spectrum entropy of a completely flat electron-density image. This evaluation function is based on the idea that a correct phase set will result in a positive definite electron-density image consistent with the observed data and simultaneously a Fourier synthesis map which is as flat as possible.

4.2. Verification and limitation

Evaluation function Tian2 is tested on the synthetic diffraction data of two known structures. One is $Al(IO_3)_3(H_2O)_8$, the crystallographic data of which have been given in the previous section. The other is $C_{252}H_{326}O_{19}$ which is reported by Czugler *et al.* (2003). This compound crystallizes in a triclinic unit cell with lattice parameters a = 16.909, b = 18.772, c = 21.346 Å, α = 111.46, β = 103.38, γ = 107.74°. The space group of both structures is $P1$. The unit cells of $Al(IO_3)_3(H_2O)_8$ and $C_{252}H_{326}O_{19}$ are divided into 64 × 64 × 128 and 170 × 190 × 220 electron-density pixels, respectively. The moduli of structure factors calculated from the known structures are taken as synthetic diffraction data. For each synthetic data set, a phase set is assigned, and then a positive definite electron-density image is created with the MEM. The value of $R_{\rm men}$ is recorded when convergence of MEM iterations has been reached. A Fourier synthesis map is also calculated using the synthetic structure-factor moduli and the assigned phase set. Subsequently, the power spectrum entropy S of the Fourier synthesis map is calculated. The value of the evaluation function Tian2 is then calculated using $R_{\rm mem}$ and S.

For each synthetic data set of $Al(IO_3)_3(H_2O)_8$, the value of Tian2 is calculated for the correct phase set, all-zero phase set and 15 random phase sets. For $C_{252}H_{326}O_{19}$, the value of Tian 2 is calculated for ten random phase sets in addition to the correct and all-zero phase set. The results of tests are presented in Fig. 4. As revealed by Fig. 4, for the high-resolution data set, the value of Tian2 can distinguish the correct phase set from the all-zero and random phase sets readily. With the deterioration of data resolution, the difference among the values of Tian2 resulting from correct, all-zero and random phase sets decreases gradually. However, up to the resolution of 3.5 Å, the evaluation function Tian2 can still distinguish the correct phase set from the wrong ones for the structure of $C_{252}H_{326}O_{19}$. Based on the very limited tests, the evaluation function Tian2 is expected to be promising in cases where only a diffraction data set of rather limited resolution is available.



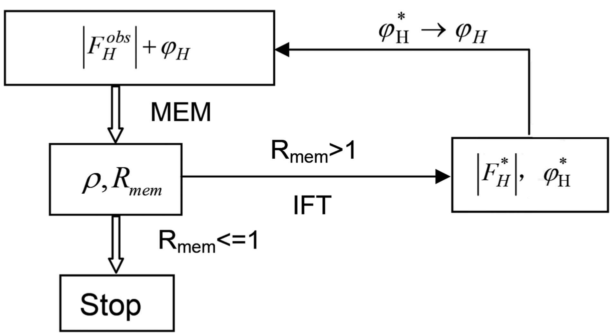
(*a*)

(*b*)

Figure 4
Evaluation function Tian2 of electron-density images reconstructed using synthetic structure-factor moduli and correct, all-zero and random phase sets. The first phase set consists of correct phases, the second one is made up of all-zero phases, and all other phase sets are totally random. (*a*) $Al(IO_3)_3(H_2O)_8$, (*b*) $C_{252}H_{326}O_{19}$. The unit cells of $Al(IO_3)_3(H_2O)_8$ and $C_{252}H_{326}O_{19}$ are divided into 64 × 64 × 128 and 170 × 190 × 220 grids, respectively.

4.3. Algorithm for structure solution based on $R_{\rm men}$

Algorithms have not been figured out yet to solve structures based on the evaluation function Tian2. Nevertheless, an algorithm based on $R_{\rm mem}$ is developed to solve structures from high-resolution data. The scheme of the algorithm is shown below:



The positive definite electron-density image ρ is first created by the MEM using the observed moduli of structure factors $|F_H^{\rm obs}|$ and the randomly assigned phases φ_H; $R_{\rm mem}$ is calculated simultaneously. If the iterative MEM process results in an $R_{\rm mem}$ which is less than 1, the electron-density image ρ is then considered to be correct. If the iterative MEM process results in an electron-density image ρ with $R_{\rm mem} > 1$, then ρ is inversely Fourier transformed to generate new moduli $|F_H^*|$ and phases φ_H^* of structure factors. φ_H is replaced with the latest φ_H^* and then the new iterative MEM process is started. The whole process is terminated when an $R_{\rm mem} < 1$ is achieved.

The algorithm is tested on several synthetic data sets at the resolution of 0.8 Å. It is demonstrated that the correct electron-density image can be reconstructed from high-resolution data under the constraint of $R_{\rm mem} < 1$.

5. Discussion and concluding remarks

In this contribution, the solution to the phase problem of X-ray crystallography is attempted from a new perspective of image definition evaluation functions. Two evaluation functions are established to assess the image definition of the electron-density images reconstructed from the X-ray diffraction data. The first function, Tian1, can be used to identify the image constructed with the correct phase set when atomic resolution data are available. The other function, Tian2, is still applicable when the resolution of the data set is as low as 3.5 Å. Therefore, it is expected that the function Tian2 may find applications in protein crystallography.

An iterative algorithm based on the function Tian1 has been established to solve structures from atomic resolution data. Algorithms of structure solution based on the function Tian2 have not been figured out yet. Global optimization methods are being considered in designing the algorithms based on Tian2. Introducing the direct method into the potential algorithms based on Tian2 may improve the efficiency of structure solution.

There are perhaps more appropriate functions to evaluate the image definition of the reconstructed electron-density image, and more efficient algorithms for structure solution from X-ray diffraction data. Nevertheless, the significance of this work lies in the fact that it provides a new perspective on the phase problem of X-ray crystallography.

Acknowledgements

The work was financially supported by the National Natural Science Foundation of China (grant Nos. 11474281 and 51272049). HL and MH are very grateful to Ms Guang-Lian Tian for her consistent support and encouragement.

References

Chen, L., Li, W. J., Chen, C., Qin, H. & Lai, J. L. (2013). *Comput. Eng. Appl.* **49**, 152–155.

Collins, D. M. (1982). *Nature (London)*, **298**, 49–51.

Czugler, M., Weber, E., Párkányi, L., Korkas, P. P. & Bombicz, P. (2003). *Chem. Eur. J.* **9**, 3741–3747.

Fienup, J. R. (1982). *Appl. Opt.* **21**, 2758–2769.

Lattman, E. & DeRosier, D. (2008). *Acta Cryst.* A**64**, 341–344.

Miao, J., Ishikawa, T., Robinson, I. K. & Murnane, M. M. (2015). *Science*, **348**, 530–535.

Momma, K. & Izumi, F. (2008). *J. Appl. Cryst.* **41**, 653–658.

Oszlányi, G. & Sütő, A. (2004). *Acta Cryst.* A**60**, 134–141.

Oszlányi, G. & Sütő, A. (2008). *Acta Cryst.* A**64**, 123–134.

Patterson, A. L. (1934). *Phys. Rev.* **46**, 372–376.

Sakata, M. & Sato, M. (1990). *Acta Cryst.* A**46**, 263–270.

Stark, H. (1987). Editor. *Image Recovery Theory and Application*. Orlando: Academic Press.

Woolfson, M. M. (1987). *Acta Cryst.* A**43**, 593–612.

Woolfson, M. M. & Fan, H. F. (1995). *Physical and Non-Physical Methods of Solving Crystal Structure.* Cambridge University Press.

Wu, N. (1997). *The Maximum Entropy Method*. Berlin: Springer.