Acta Crystallographica Section A
**Foundations of Crystallography**
ISSN 0108-7673

Received 19 February 2004
Accepted 20 May 2004

# Phasing diffuse scattering. Application of the *SIR2002* algorithm to the non-crystallographic phase problem

**B. Carrozzini,[a] G. L Cascarano,[a] L. De Caro,[a] C. Giacovazzo,[a]* S. Marchesini,[b] H. Chapman,[b] H. He,[c] M. Howells,[c] J. S. Wu,[d] U. Weierstall[d] and J. C. H. Spence[d]**

[a]Istituto di Cristallografia, CNR, Via G. Amendola 122/o, 70126 Bari, Italy, [b]Lawrence Livermore Laboratory, 7000 East Avenue, Livermore, CA 94550-9234, USA, [c]Lawrence Berkeley Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA, and [d]Department of Physics and Astronomy, Arizona State University, Tempe, AZ 85287-1504, USA. Correspondence e-mail: carmelo.giacovazzo@ic.cnr.it

A new phasing algorithm has been used to determine the phases of diffuse elastic X-ray scattering from a non-periodic array of gold balls of 50 nm diameter. Two-dimensional real-space images, showing the charge-density distribution of the balls, have been reconstructed at 50 nm resolution from transmission diffraction patterns recorded at 550 eV energy. The reconstructed image fits well with a scanning-electron-microscope (SEM) image of the same sample. The algorithm, which uses only the density modification portion of the *SIR2002* program, is compared with the results obtained *via* the Gerchberg–Saxton–Fienup *HiO* algorithm. The new algorithm requires no knowledge of the object's boundary and proceeds from low to high resolution. In this way, the relationship between density modification in crystallography and the *HiO* algorithm used in signal and image processing is elucidated.



## 1. Introduction

The phase problem for non-periodic objects (Stark, 1987) has been studied by many different approaches, some of which have recently demonstrated striking success with experimental data:

(*a*) the Gerchberg–Saxton–Fienup *HiO* algorithm (Fienup, 1982, 1987); see also De Caro *et al.* (2002), who refer to it by using the concept of *confined structure*; this method is closely related to that used for the first successful experimental reconstruction of a non-periodic object from X-ray scattering by Miao *et al.* (1999);

(*b*) techniques based on analyticity and complex zeros (Liao *et al.*, 1997);

(*c*) the study of projections onto convex sets (Bautschke *et al.*, 2002);

(*d*) the transport of intensity equations (Paganin & Nugent, 1998);

(*e*) direct methods, for real and positive objects (Spence *et al.*, 2003).

Experimental implementation of the popular *HiO* algorithm in transmission geometry is complicated by loss of low spatial frequencies at the beam-stop. To overcome this problem, in the past the *HiO* oversampling method has been used in combination with an independent low-resolution image of the object (*i.e.* from SEM, X-ray zone plate or optical image) to provide the low spatial frequencies. More recently, the development of the *Shrinkwrap* (Marchesini *et al.*, 2003) and other variants of the *HiO* algorithm have successfully treated the missing data as adjustable parameters. Experimentally, the preparation of the isolated object needed for *HiO* can be difficult; only recently has the first atomic resolution image of a carbon nanotube been reconstructed by *HiO* from a coherent transmission electron diffraction pattern using nanotubes that span holes in a carbon grid (Zuo *et al.*, 2003).

In a recent paper (He *et al.*, 2003), images of randomly placed two-dimensional arrays of gold balls were reconstructed from their soft X-ray transmission diffraction patterns by means of the *HiO* algorithm. This algorithm is closely related to the electron-density modification procedure of X-ray crystallography, and is able to: (i) exploit prior knowledge of the geometrical form of a support, outside of which the electron density is known to be zero (the support is defined as the region in which the object density is known to be non-zero: it corresponds roughly to the boundary of the object); (ii) treat, *via* suitable constraints, the real and imaginary parts of the object wavefunction (Miao *et al.*, 1999). A necessary tool for success was the use of either a mask based on supplementary images provided by SEM techniques or a subjective estimate of the object boundary taken from the autocorrelation function. In the *Shrinkwrap* method (Marchesini *et al.*, 2003), a first estimate of the object support is obtained from the autocorrelation function of the object

(which is known to be twice as large as the object in any given direction) and this estimate is updated iteratively. If the structure consists of one isolated ball and two clusters (as in the case treated here), the autocorrelation function includes a faithful image of every cluster convoluted with the single ball and this may be used as a trial support function. The method is then similar to the heavy-atom method of X-ray crystallography or Fourier transform holography.

In a companion paper (Spence *et al.*, 2003), direct methods (DM) have been applied to a simulated diffraction pattern calculated from a structure of 28 gold balls. In that work, the phasing process was realized by the following steps. The simulated pattern was divided by the Fourier transform $f(\mathbf{h})$ of a single ball, so obtaining a simulated normalized pattern (SNP). The structure factors, in turn, were also normalized, according to

$$E_{\mathbf{h}} = \sum_{j=1}^{N} f_j(\mathbf{h}) \exp(2\pi i \mathbf{h} \cdot \mathbf{r}_j) \Big/ \left[ \sum_{j=1}^{N} f_j(\mathbf{h})^2 \right]^{1/2}, \qquad (1)$$

where $N$ is the total number of balls. Then a modified version of the *SIR2002* program was used to determine the positions of the balls from the SNP [Burla *et al.* (2003); *SIR2002* is a package originally dedicated to the crystal structure solution of small and large molecules]. The procedure consisted of a tangent procedure starting from random phases, followed by:

(*a*) the EDM (electron-density modification) step, where new phase sets $\{\varphi\}$ were estimated by iteratively applying electron-density-modification procedures, *i.e.* $\rho \rightarrow \{\varphi\} \rightarrow \rho$ cycles;

(*b*) the HAFR (heavy-atom reduced real-space Fourier refinement) step, consisting of $\rho \rightarrow \{\varphi\} \rightarrow \rho$ cycles, in which balls were associated with electron-density peaks;

(*c*) least-squares cycles alternating with $(2F_{\rm obs} - F_{\rm calc})$ map calculations to refine and complete the structural model.

Such an approach requires prior knowledge of the scattering factors of the balls and of their number $N$. However, it does not need prior knowledge of the geometry of the support. When applied to the experimental speckle pattern considered in this paper, such an algorithm did not succeed, probably owing to unavoidable errors in the measurements and in the model scattering function (the gold balls do not have exactly the same radius).

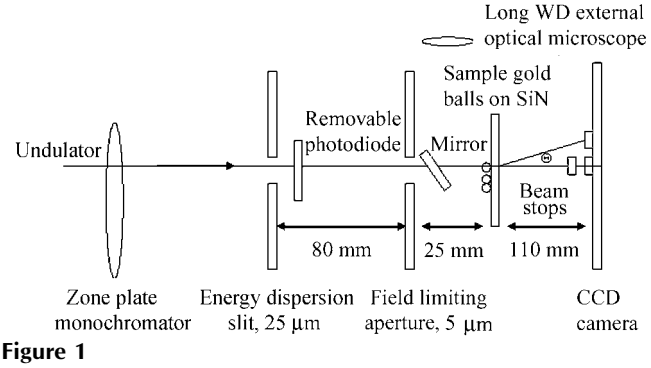


**Figure 1**
Set-up of the diffraction experiment.

This negative result has stimulated the search for a more general phasing procedure, described in this paper, based on the following general conditions:

(*a*) no prior knowledge of the scattering factor of the gold balls or of their number;

(*b*) no use of any support, such as masks based on the autocorrelation function or SEM images.

The algorithm described in this paper is restricted to real densities, such as a charge density illuminated by X-rays well away from any absorption edge or the electrostatic potential for electron diffraction in the absence of inelastic scattering. Significant spatial variation in X-ray absorption or multiple-scattering effects and inelastic scattering in electron or X-ray diffraction lead to a 'complex object', which cannot be treated by the methods of this paper. However, a spatially independent mean attenuation is acceptable. Without these inelastic effects, we may then assume that Friedel's law holds but we note that this assumption may hinder the precise reconstruction of the structure if they are actually present.

## 2. Experimental

An experimental soft X-ray transmission diffraction pattern from two clusters of gold balls of $50 \pm 5$ nm diameter, deposited on a silicon nitride window, was recorded at the Advanced Light Source at the Lawrence Berkeley Laboratory, using soft X-rays with a wavelength of 2.1 nm (He *et al.*, 2003). The experiments used the 'pink' 588 eV (2.11 nm) photon beam at beamline 9.0.1. Features of the beamline include: a 0.5 μm thick, 750 μm square Be window to separate the UHV beamline from the low-vacuum sample environment; a monochromator consisting of an off-axis segment of a zone plate and the diffraction experiment itself (Fig. 1). Samples were made by placing a droplet of solution containing 'gold conjugate' colloidal gold balls on a silicon nitride window (thickness 100 nm) and allowing it to dry. The gold balls formed several single-layered (two-dimensional) clusters on the SiN membrane, as shown by field-emission scanning-electron-microscope images.

The X-ray coherence length $l_c$ must be greater than the maximum path difference between any pair of interfering rays, *i.e.* $l_c > w\Theta_{\rm max}$, where $w$ is the width of the sample and $\Theta_{\rm max}$ is the maximum diffraction angle. For our geometry and wavelength, $\Theta_{\rm max} = 0.12$ rad and the resolution limit is 8.4 nm. For the 5 μm aperture (effectively the monochromator exit slit) shown in Fig. 1, the resolving power is about 500, the coherence length is then 1 μm and the maximum sample illumination area $8 \times 8$ μm. Similarly, the (spatial) coherence patch provided by the 5 μm aperture is $10 \times 10$ μm. Allowing for an empty (but still coherently illuminated) band around the sample, its allowed size is thus $<4 \times 4$ μm. Temporal and spatial coherence requirements for this coherent diffractive imaging are discussed in more detail by Spence & Howells (2004)

The periodic boundary conditions that are imposed by sampling are now considered. We first consider a one-dimensional case for simplicity. Shannon's theorem (see, for

example, Papoulis, 1962) gives the optimum sampling interval $\Delta x$ of a real-space band-limited function (which we take to be real), which will allow its reconstruction at all intermediate abscissa values. Shannon shows that this is possible if $\Delta x = 1/(2u_o)$, where $u_o$ is the band limit, *i.e.* the cut-off spatial frequency in the *complex* diffraction pattern. The theorem is shown in textbooks also to apply with the domains reversed, in which case a compact support of width $L$ in real space will play the role of a 'band limit'. However, the quantity detected in our experiments is the diffracted intensity, not the diffracted complex amplitude. The Fourier pair for this diffracted intensity is the autocorrelation of the object and we may now apply Shannon's band-limit sampling theorem to this Fourier pair. The autocorrelation function has width $2L$ and this now plays the role of a 'band limit' (but in real space); hence, by the theorem, the *optimum* sampling (not oversampling) of the diffracted intensity is $1/(2L)$ (half the Bragg angle for a crystal of period $L$). This is the sampling interval for the scattering intensity that extracts all useful information from it, including phase information. The extraction of this phase information requires an algorithm capable of solving a set of non-linear coupled equations, which the Fienup algorithm does. In this sense, this optimum sampling of the diffraction-plane intensity minimizes aliasing of the autocorrelation function. This Shannon interval for frequency-space sampling of the *intensity* is $1/(2w) = \Delta/\lambda z$, where $z$ is the sample-to-detector distance and $\Delta$ is the detector-plane increment (a 25 µm CCD pixel in our case). For our $\Delta$ and $z$ values, this leads to a sampling interval of 4 µm. Note that this would correspond to a twofold oversampling in each direction of the diffraction-plane complex amplitude. The choice of other sampling ratios in two and three dimensions is discussed in detail elsewhere (see Miao *et al.*, 1998; see also Sayre, 1952; Bates, 1982).

The semiangle $\Theta$ (see Fig. 1) subtended at the object by the first detector pixel beside the optic axis defines the lowest spatial frequency $1/D = \Theta/\lambda$ contributing to the reconstruction, and we treat this pixel as the first-order Bragg reflection, $(1, 0)$ Miller indices, for a square two-dimensional unit cell with symmetry $P1$. The use of *SIR2002* imposes periodic continuation of this cell. For the performed experiments, $D = 9.147$ µm. For our $1025 \times 1025$ CCD, the highest-order reflection is $(512, 512)$; the real-space resolution for pixel $(512, 0)$ at the boundary is $D/512 = 17.8$ nm, giving about four pixels per gold ball in the reconstructed image. The width of the region illuminated by the beam is about $B = 5$ µm, which is less than the coherence width of the beam. The diffraction pattern consists of Airy's disc-like pattern from a single ball, crossed by finer 'speckle' fringes due to interference between different balls. The first zero for a ball rather than a disc occurs at a semiangle $1.4\lambda/d$, which occurs at the reflection with pixel (Miller) indices $(256, 0)$. (The spacing $D1$ corresponding to the angle $1.4 \times \lambda/d$ is $D1 = d/1.4 = 50/1.4 = 35.7$ nm. As we have assumed $D = 9.147$ µm, the index is $D \times 1.4/50 = 256$.)

Fig. 2 shows a SEM image of a random set of gold balls of 50 nm diameter at 550 eV. Fig. 3 gives the corresponding diffraction pattern, showing fine speckle fringes modulating the pattern expected from a sphere. The first minimum of this pattern occurs at a value of $\sin\theta/\lambda = 1.4/d$ Å$^{-1}$, where $d$ is the diameter of the ball (the factor 1.4 is replaced by 1.2 for a disc).

The number of measured reflections (Friedel pairs included) is 992061, for data resolution up to 126 Å, about one quarter of the gold ball diameter; the number of measured unique reflections (Friedel opposite merged) is REFL = 522517; the number of unique reflections, not measured because of the beam stop, is UREFL = 1365.

**Figure 2**
SEM image of a random set of gold balls of 50 nm diameter at 550 eV.

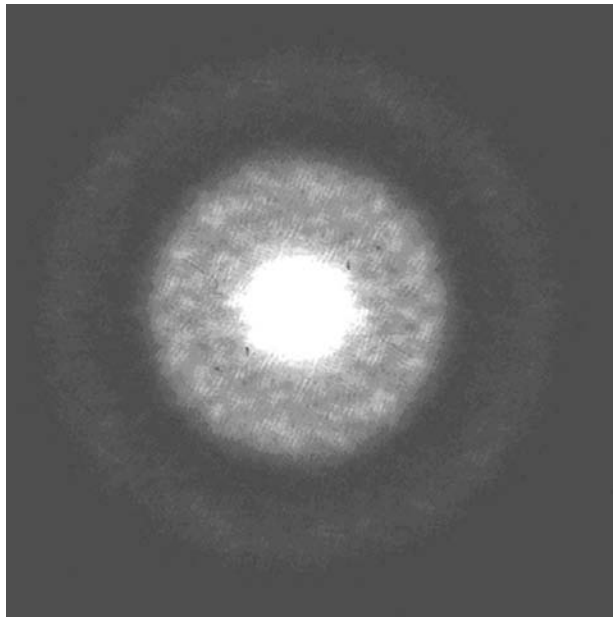

**Figure 3**
Experimental soft X-ray transmission diffraction pattern from the 50 nm diameter gold balls shown in Fig. 2. The X-ray wavelength is 2.5 nm.

## 3. The new phasing procedure

The new procedure uses only the electron-density-modification subroutine of the *SIR2002* software package. In real space, only a small portion of pixels with the highest values (about 0.25%) were selected as defining the region inside the support for the current estimate of the density function. In reciprocal space, the phasing proceeded gradually from low-order reflections outward. The modulus constraint was applied in each cycle. The initial assignment of phases is random. We assume that:

(*a*) the nature and the number of the scatterers are unknown (consequently, their scattering factors are also unknown);

(*b*) anomalous-dispersion effects can be neglected; accordingly, only the set of unique reflections will be considered in the calculations and thus the intensities of the Friedel opposites are merged.

The above assumptions exclude most of the steps of the standard *SIR2002* procedure, *e.g.* the normalization of the structure factors, the tangent process, the HAFR refinement cycles and the least-squares process (see §1). Accordingly, the new *SIR2002* procedure consists only of cycles of electron-density modification $\rho \rightarrow \{\varphi\} \rightarrow \rho$. The phasing process may be partitioned into two steps.

STEP 1. Because of their importance (see §4), only low-resolution reflections are used initially in the phasing process. A resolution threshold TRH is defined, equal to 0.25 of the experimental data resolution: let NREFTRH be the number of unique reflections satisfying the threshold resolution (in our case, NREFTRH = 52500). The program starts by assigning random phases to a very limited number of reflections: let us call NBEG this number, which corresponds to 5% of the NREFTRH reflections. Then 60 macrocycles, each constituted by 9 cycles $\rho \rightarrow \{\varphi\} \rightarrow \rho$ are performed. During each macrocycle, the number of phased reflections used in each hemicycle $\{\varphi\} \rightarrow \rho$ starts from NBEG and progressively increases up to NREFTRH: the selection of the reflections actively used is performed on the basis of the product $|F_{\rm obs}F_{\rm calc}|$. In each hemicycle $\rho \rightarrow \{\varphi\}$, the number of grid points of the electron-density map used for calculating phases is 0.0025 of their total number (we select the grid values with highest intensity).

At the end of the first macrocycle, the program estimates, in the hemicycle $\rho \rightarrow \{\varphi\}$, the intensities and the phases of the UREFL reflections. The intensities, calculated by inversion of $\rho$, are used as observed values in the next cycles $\rho \rightarrow \{\varphi\} \rightarrow \rho$ (in addition to the NREFTRH reflections) and are updated (together with the calculated phases) at the end of every macrocycle.

STEP 2. The phasing process is extended to the full set of experimentally available reflections. The phase values obtained in STEP 1 for the low-resolution reflections are used as pivots in the phasing process: their new values (obtained at the end of each macrocycle) are always combined (by the tangent formula) with the values obtained at the end of STEP 1. By contrast, the UREFL reflections are recalculated as in STEP 1.

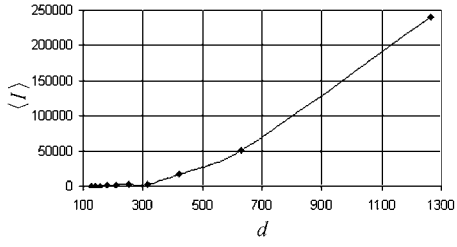


**Figure 4**
The average intensity $\langle I \rangle$ is plotted against the resolution $d$ (in Å).

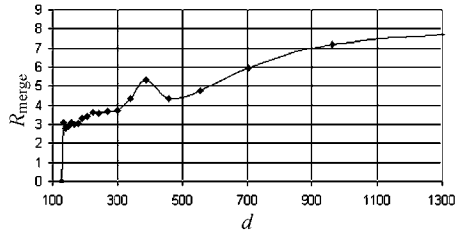


**Figure 5**
$R_{\rm merge}$ *versus* the resolution $d$ (in Å).

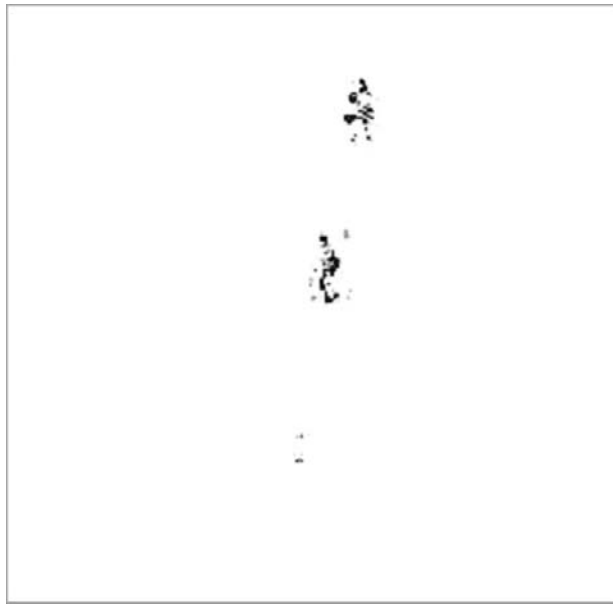

**Figure 6**
Recovered charge density from non-periodic array of 50 nm diameter gold balls using the modified *SIR2002* program.

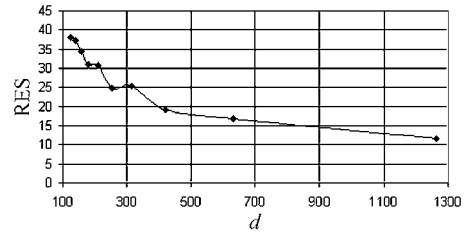


**Figure 7**
Value of RES for $n = 1$ *versus* $d$ (in Å).

These steps define the relationship between the new proposed algorithm and the classical *HiO* procedure.

(*a*) The new algorithm varies, from cycle to cycle, the number of structure factors used in the calculation of $\rho$.

(*b*) The new algorithm limits the number of grid points in the electron-density map, whose intensities are used to compute the phases. This practice is *not* suggested by the fact that the balls occupy a small part of the total area. Indeed, a similar procedure is used by the standard version of *SIR2002* for solving *ab initio* the crystal structure of proteins, and also of small molecules, which fill more densely the unit cell (only about 2% of the pixels are used for the electron-density inversion). In the case of oversampling, the percentage may be further reduced because the unit cell contains a large empty region.

(*c*) No information on the geometry of the support is used. This is an advantage when such information is unknown or imperfectly defined in the experiment.

## 4. Analysis of the diffraction data and of the results

Two features of the experimental diffraction pattern are evident. The first one regards the average intensity $\langle I \rangle$ as a function of the resolution $d$ (calculated in Å), plotted in Fig. 4. We see that the intensities of the low-order beams are very intense in comparison with the high orders. The phasing process, therefore, is expected to be dominated by the inner region of the pattern. This is the reason why our phasing procedure tries first: (*a*) to phase low-frequency reflections, and then uses them as pivots in a phase-extension process; (*b*) to estimate the number UREFL of unobserved low-frequency intensities. The subsequent estimates of these beam-stop reflections play a crucial role in the success of the procedure: without their use, the phasing process may fail. A similar observation was reported by Wu & Spence (2003) that the intensities of low-resolution fractional reflections calculated using a Hilbert transform formula was found to be important in performing phase extension.

The second feature that should be evidenced is that $I(h, k)$ is not perfectly symmetrical: significant deviations from Friedel's law were noted. This may be due to the following three reasons.

(*a*) Unavoidable parasitic scattering or absorption. For a well designed experiment, the effects of parasitic scattering or absorption (border effects, silicon nitride absorption *etc*.) should be minimized. The success of our phasing process indicates that this is the case for our experiment.

(*b*) Effects of anomalous scattering. If the balls are assumed to be identical, the scattering amplitudes $F_{h,k}$ and $F_{-h,-k}$ should be equal, however the corresponding phases are not constrained to reverse sign under index inversion if there is multiple scattering or if the phase shift along an optical path through a ball exceeds $\pi/2$. At large scattering angles also, all the intensities detected fall on the Ewald sphere, however reflection from one point on this sphere through the origin does not result in a new point on the sphere. This failure condition is discussed elsewhere (Spence *et al.*, 2003). Since the radii may not all be equal, and since the balls may be elastically deformed, we may expect that the equality $|F_{h,k}| = |F_{-h,-k}|$ is not strictly obeyed. The deviations, however, are expected to be small.

(*c*) Incorrect identification of the (0, 0) pixel in the detector (*zero pixel*).

Since the effects of (*a*) are expected to be small, we can identify the zero pixel as the one for which the value of

**Figure 8**
Recovered charge density from a non-periodic array of 50 nm diameter gold balls using the *HiO* algorithm.

**Figure 9**
Recovered charge density from a non-periodic array of 50 nm diameter gold balls using the modified *SIR2002* program at the end of STEP 1.

$$R_{\rm merge} = \frac{\sum_{h,k} (|F^+ - F_{\rm med}| + |F^- - F_{\rm med}|)}{\sum_{h,k} (F^+ + F^-)} \qquad (2)$$

is a minimum, where

$$F_{\rm med} = (F^+ + F^-)/2, \qquad (3)$$

with $F^+$ and $F^-$ representing $F_{h,k}$ and $F_{-h,-k}$, respectively.

The application of equation (2) to the experimental pattern shown in Fig. 3 did not suggest any origin shift. [We report in Appendix *A* the results obtained when we applied equation (2) to the pattern shown in Fig. 10.]

The deviation from Friedel's law decreases at increasing values of $\sin\theta/\lambda$ (see Fig. 5), where the value of $R_{\rm merge}$ is plotted against the resolution $d$ (Å). Such a trend does not comply with the expected behaviour of the anomalous-dispersion effects, which usually increase with $\sin\theta/\lambda$. This suggests that the deviations from Friedel's law are mainly due to unavoided parasitic scattering and to errors in measurements. (At high resolution, the diffraction intensities are particularly faint.)

We now consider the results of our phasing algorithm. The image resulting at the end of the phasing procedure (as seen on the computer screen) is shown in Fig. 6: it is in good agreement with the experimental SEM image shown in Fig. 2 (apart from a shift of origin).

We now require a quantitative measure of the efficiency of our phasing approach and of the accuracy of the calculated model. Since we have ignored the scattering factors of the balls, we cannot use standard crystallographic indicators as a measure of agreement between experiment and calculations (*i.e.* the crystallographic residual between the observed and scattered intensities). To overcome this problem, we treat the Fourier coefficients obtained by Fourier inversion of the final, suitably modified, electron-density map as the 'calculated' structure factors ($F_{\rm calc}$ in the following). The final map produced by *SIR2002* is modified (before the Fourier inversion) as follows:

$$\rho(x,y) = \begin{cases} \rho(x,y) & \text{if } \rho(x,y) > \langle\rho\rangle + n\sigma \\ \rho(x,y) = 0 & \text{otherwise.} \end{cases} \qquad (4)$$

$\langle\rho\rangle$ is the average value of the unmodified electron-density map, $\sigma$ is its standard deviation. We show in Table 1 the residual

$$\text{RES} = \frac{\sum_{h,k} (|F_{\rm obs} - |F_{\rm calc}||)}{\sum_{h,k} F_{\rm obs}} \qquad (5)$$

for various values of $n$. The following may be observed.

(*a*) The value of RES does not significantly change when $n$ varies from 1 to 7. For any value of $n$, RES is significantly larger than the figures usually obtained for small-molecule crystals at the end of the structure refinement process. Our previous observations suggest that this is probably due to experimental errors in the intensities rather than to some intrinsic lack of convergence of the phasing algorithm.

(*b*) The electron-density function is well separated into two domains: the domain where the scattering balls are confined and an empty region;

**Table 1**
The RES values for various $n$ values (see main text) are shown when the modified *SIR2002* and the *HiO* algorithms are used.

| *n* | *SIR2002* | *HiO* |
|---|---|---|
| 1 | 23 | 44 |
| 2 | 23 | 46 |
| 3 | 23 | 44 |
| 4 | 23 | 40 |
| 5 | 24 | 40 |
| 6 | 25 | 41 |
| 7 | 27 | 43 |

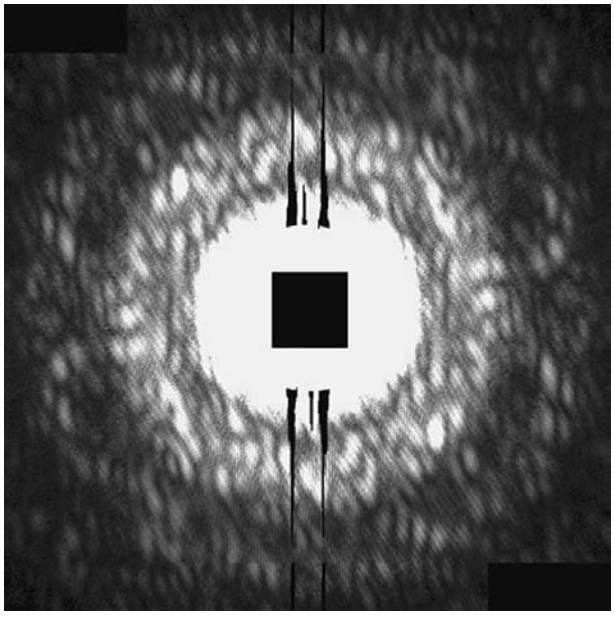

**Figure 10**
Experimental soft X-ray transmission diffraction pattern from the 50 nm diameter gold balls shown in Fig. 11. The X-ray wavelength is 2.5 nm.

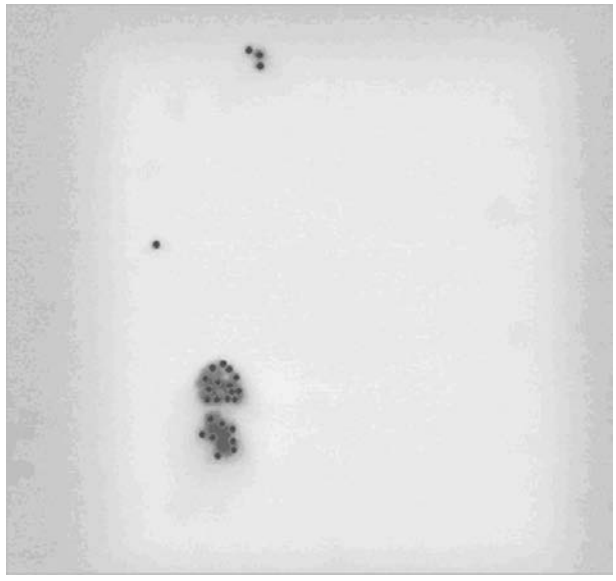

**Figure 11**
SEM image of a random set of 28 gold balls of 50 nm diameter at 550 eV.

(*c*) The refinement process converged to a satisfactory model.

We have calculated, for the case $n = 1$, the value of RES against resolution (a similar trend is obtained for other values of $n$). We observe (see Fig. 7) that RES substantially increases with $\sin\theta/\lambda$: this behaviour is partially expected since the high-angle intensities are weak but some systematic errors may remain. Luckily, the error is smaller for low-angle intensities: they drive the phasing process to convergence in spite of the large residual at high angles.

In Fig. 8, we show the map obtained by application of the *Shrinkwrap HiO* algorithm, in which a support is obtained from the autocorrelation function and then improved during *HiO* iterations. The map was calculated by assuming $|F_{h,k}| = (|F_{h,k}| + |F_{-h,-k}|)/2$ and, for the phases, we assume $\varphi_{-h,-k} = -\varphi_{h,k}$, ensuring a real object. The map correctly reveals the ball positions but shows a high background level. In Table 1, we show the values of RES for various $n$, obtained by using the *HiO* algorithm. The RES values are much higher than those obtained from the map provided by *SIR2002* because the *HiO* refinement process is unable to eliminate the large number of false low-intensity peaks.

Finally, it is useful to verify the efficiency of our algorithm when the data resolution is low. We therefore stopped the phasing procedure at the end of STEP 1, when we had phased NREFTRH = 51193 observed reflections and UREFL = 1365 beam-stop reflections. The data resolution was limited to 500 Å, comparable with the diameter of the balls. The *SIR2002* algorithm again succeeded and produced the map shown in Fig. 9. We note that:

(*a*) the structural model is essentially correct;

(*b*) the RES value is equal to 20, smaller than the value obtained at the end of STEP 2.

This last result is of particular interest: it shows that oversampling in reciprocal space provides such a high amount of information that the phase problem may be solved even in those cases in which the data resolution is insufficient for crystal structure solution.

This conclusion raises a further question: what degree of oversampling is optimal for solution of the phase problem? This issue has been discussed in detail elsewhere (Miao *et al.*, 1998).

In the absence of noise, Shannon's theorem indicates that the Nyquist rate allows reconstruction of the diffraction pattern at any angle (between samples) if the samples are independent. In the presence of noise, however, a higher degree of oversampling: (*a*) produces a high correlation between the intensities of neighbouring pixels and thus the total amount of information does not increases linearly with the amount of oversampling; (*b*) requires small detector pixels, variable working distance and/or large supports, but these conditions are not always easily satisfied: the first demands a high-quality detector and the second high spatial coherence in the incident beam; (*c*) generates a very large number of reflections (particularly for the three-dimensional case), not easily manageable using current computer programs.

## 5. Conclusions

A new phasing algorithm, based on the density-modification portion of the *SIR2002* direct-methods crystallographic program, has been applied to soft X-ray scattering from a non-periodic object. The algorithm successfully reconstructs phases from diffuse (continuous) measured scattering amplitudes. As a result, images of a cluster of 50 nm diameter gold balls have been reconstructed. The results provide an interesting comparison with the *HiO* Fienup–Gerchberg–Saxton algorithm, which operates on closely related principles. By contrast, however, no knowledge of the object boundary is assumed and the *SIR2002* phasing proceeds as a steadily increasing function of spatial frequency.

## APPENDIX *A*

We have calculated $R_{\rm merge}$ as a function of the index shift for the pattern shown in Fig. 10. This is obtained from a distribution of gold balls such as that shown in Fig. 11. The function $R_{\rm merge}$ is shown in Fig. 12. The minimum value of $R_{\rm merge}$ occurs by subtracting 3 from $h$ and adding 2 to $k$. After the origin shift, the average discrepancy between Friedel amplitudes decreases from 8.99% (obtained for unshifted data) to 6.33%. The shift for which the minimum is obtained is clearly defined.

This work was supported by ARO award DAAD190010500 (JCHS PI) and by the National Council of Research (Italy).

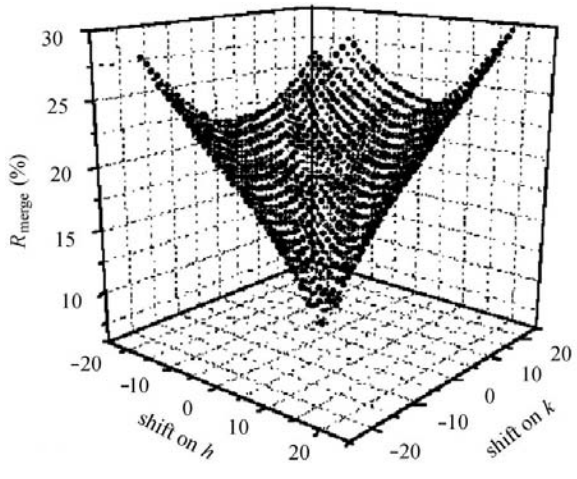


**Figure 12**
Behaviour of $R_{\rm merge}$ between Friedel opposites as a function of the index shift for the gold ball structure shown in Fig. 11.

## References

Bates, R. H. T. (1982). *Optik (Stuttgart)*, **61**, 247–262.
Bautschke, H., Combettes, P. L. & Luke, D. R. (2002). *J. Opt. Soc. Am.* **19**, 1344–1345.
Burla, M. C., Camalli, M., Carrozzini, B., Cascarano, G. L., Giacovazzo, C., Polidori, G. & Spagna, R. (2003). *J. Appl. Cryst.* **36**, 1103.

De Caro, L., Giacovazzo, C. & Siliqi, D. (2002). *Acta Cryst.* A**58**, 415–423.
Fienup, J. R. (1982). *Appl. Opt.* **21**, 2758–2769.
Fienup, J. R. (1987). *J. Opt. Soc. Am.* **4**, 118–123.
He, H., Marchesini, S., Howells, M., Weierstall, U., Hembree, G. & Spence, J. C. H. (2003). *Acta Cryst.* A**59**, 143–152.
Liao, C., Fiddy, M. & Byrne, C. (1997). *J. Opt. Soc. Am.* **14**, 3155–3161.
Marchesini, S., He, H., Chapman, H., Hau-Riege, S., Noy, A., Howells, M., Weierstall, U. & Spence, J. C. H. (2003). *Phys Rev.* **68**, 140101(R).
Miao, J., Charalambous, C., Kirz, J. & Sayre, D. (1999). *Nature (London)*, **400**, 342–344.
Miao, J., Sayre, D. & Chapman, H. N. (1998). *J. Opt. Soc. Am.* A**15**, 1662–1669.
Paganin, D. & Nugent, K. (1998). *Phys. Rev. Lett.* **80**, 2586–2589.
Papoulis, A. (1962). *The Fourier Integral and its Applications.* New York: McGraw-Hill.
Sayre, D. (1952). *Acta Cryst.* **5**, 843.
Spence, J. & Howells, M. (2004). *Ultramicroscopy*. In the press.
Spence, J. C. H., Wu, J. S., Giacovazzo, C., Carrozzini, B., Cascarano, G. L. & Padmore, H. A. (2003). *Acta Cryst.* A**59**, 255–261.
Stark, H. (1987). *Image Recovery: Theory and Applications.* New York: Academic Press.
Wu, J. S. & Spence, J. C. H. (2003). *Acta Cryst.* A**59**, 577–583.
Zuo, J. M., Vartanyants, I. A., Gao, M., Zhang, M. & Nagahara, L. A. (2003). *Science*, **300**, 1419.