Weighted simultaneous iterative reconstruction technique for single-axis tomography

D. Wolf*, A. Lubk, H. Lichte

Triebenberg Laboratory, Institute of Structure Physics, Technische Universität Dresden, 01062 Dresden, Germany

ABSTRACT

Tomographic techniques play a crucial role in imaging methods such as transmission electron microscopy (TEM) due to their unique capabilities to reconstruct three-dimensional object information. However, the accuracy of the two standard tomographic reconstruction techniques, the weighted back-projection (W-BP) and the simultaneous iterative reconstruction technique (SIRT) is reduced under common experimental restrictions, such as limited tilt range or noise. We demonstrate that the combination of W-BP and SIRT leads to an improved tomographic reconstruction technique: the weighted SIRT. Convergence, resolution and reconstruction error of the W-SIRT are analyzed by a detailed analytical, numerical, and experimental comparison with established methods. Our reconstruction technique is not restricted to TEM tomography but can be applied to all problems sharing single axis imaging geometry.

1. Introduction

Electron tomography (ET) provides an unique access to the three-dimensional (3D) structural, chemical or electrical properties of organic and inorganic materials with nanometer resolution [1–3]. For example, ET significantly contributes to the understanding of the prokaryotic ultrastructure [4,5] as well as complex catalysts [6], polymers [7], and semiconductor nanostructures [8].

ET basically includes three steps: First, the acquisition of a tilt series in the transmission electron microscope (TEM), i.e., a series of 2D electron micrographs (projections) while tilting the specimen under the electron beam typically within ca. ±70° at increments of 1–3°. Second, the alignment of the tilt series for corrections of residual displacements between the projections with respect to a common tilt axis; and third, the computerized 3D reconstruction of the tilt series by specific reconstruction techniques yielding finally the electron tomogram.

The most notable experimental limitation of the technique is an incomplete tilt range (for example ±70° instead of ±90°), which leads to a loss of information, visible as “missing wedge” in the Fourier transform (FT) of the tomogram. In real space, this corresponds to a reduced resolution in the tomogram in the direction of the missing wedge. Therefore, considerable effort is put into the development of adapted specimen geometries (e.g. needles [9]), holder designs (e.g. On-Axis Rotation Tomography Holder), and novel goniometers (e.g. TEAMstage [10]) facilitating at 180° tilt series acquisition. Accurate and stable rotation holders also reduce spurious drift of the sample which eases the requirements to the alignment procedure following the acquisition. Nonetheless, small alignment variations as well as detection noise and specimen damages cannot be completely avoided and impose a second important limitation, in particular when aiming for high-resolution tomograms. It is one important task of the reconstruction procedures to suppress the influence of these “non-projective” artifacts. This can be achieved by exploring additional information on the sample structure such as symmetries and by regularizing the reconstruction.

Nowadays, the tomographic reconstruction is usually performed either with the help of weighted back-projection (W-BP) methods [11–14] or iterative techniques, in particular the simultaneous iterative reconstruction technique (SIRT) [15,16]. Iterative methods are also often referred to as algebraic reconstruction techniques (ART) [17]. Numerous variations of these techniques have been developed in order to consider the above mentioned limitations: For example, the so-called Discrete ART (DART) [18] discretizes the range of the allowed reconstructions and can therefore robustly reconstruct samples that consist of only a few different materials (grey levels). More recently, another ART has been introduced which assumes a certain smoothness of the 3D data using compressive sensing [19].

In general, the mathematical structure of the projection–reconstruction corresponds to a matrix inversion, which usually only exists as a pseudoinverse [20]. This pseudoinverse is typically (mildly) ill conditioned (e.g. due to the missing wedge) and has to
be regularized in order to be robust against the “non-projective” errors such as noise [20]. Such a regularization can be achieved by adding auxiliary conditions weighted by a regularization parameter. Examples of such conditions are euclidean norm minimization (Tikhonov regularization, e.g. [21]), total variation minimization (TVM) [19] or basis function number minimization (compressive sensing) [22]. The iterative reconstruction techniques considered in the following utilize some sort of Tikhonov regularization where the regularization parameter is the number of iterations.

Here, we present a reconstruction algorithm that we refer to as Weighted SIRT (W-SIRT) because it is a combination of W-BP and SIRT. The combination of these two established methods (introduced in Section 2) and the advantage of a specific weighting filter (explained in Section 3) yield tomograms with higher signal fidelity and lateral resolution than those obtained from W-BP or SIRT. This is supported by comparing the results and the convergence behavior of W-SIRT or SIRT obtained from theoretical case studies (Section 4) and experimental examples (Section 5).

2. Two-dimensional weighted back-projection and SIRT

2.1. Radon transformation

If the tilt series is recorded at single axis geometry (one alternative is conical tilt geometry [23]), the description of the projection–reconstruction problem of a 3D (scalar) function \( f(x,y,z) \) can be reduced to 2D by a separate treatment of slices \( f(x,y=\text{const.},z) \) perpendicular to the tilt axis \( y \) (see Fig. 1). The process of projecting \( f(x,z) \) along lines \( l \) determined by a tilt angle \( \alpha \) and the distance to the origin \( I \), i.e.

\[
\hat{f}(I, \alpha) = \int f(x, z) ds,
\]

is referred to as Radon transformation \( \mathcal{R} \) [24], whose discrete result is also called sinogram. Thus, the tomographic reconstruction of a sinogram \((y\)-slice through the tilt series\) can be described mathematically as the inverse 2D Radon transformation \( \mathcal{R}^{-1} \). Therefore, algorithms for tomographic reconstruction optimally should be numerical realizations of \( \mathcal{R}^{-1} \).

2.2. Weighted back-projection

The W-BP is based on the simple or direct back-projection (S-BP) algorithm [4,16]. As the name back-projection suggests, each pixel of the sinogram is projected back into 2D space along the ray path, which contributed to the pixel during the projection process. The superposition of all back-projected paths yields the layergram (Fig. 1c), which reads in the continuous case [25]

\[
f_b(x,z) = (2\pi)^{-1} \int \hat{f}(n(\alpha) \cdot (x,z), \alpha) d\alpha
\]

with \( n(\alpha) = (\cos \alpha, \sin \alpha)^T \).

The layergram is only a blurred version of the desired object function \( f \). It can be abbreviated by

\[
f_b(x,z) = \mathcal{R}^T(\hat{f}(I, \alpha))
\]

with the transpose or adjoint Radon transformation \( \mathcal{R}^T \). However, the object function (tomogram) is computed by inverse Radon transformation by

\[
f(x,z) = \mathcal{R}^{-1}f_b(I,\alpha).
\]

The reason for the difference between \( f_b \) and \( f \), i.e. the blurring, can be understood from the projection-slice (or central slice) theorem, which states (for 2D) that the 1D projection of a 2D object corresponds in Fourier space to a 1D slice (line) through the origin (center) of the Fourier transform \( \mathcal{F} \) of the 2D object. Thus, the S-BP corresponds to a summation (integration) of central slices in Fourier space. This in turn means an inhomogeneous sampling decreasing from lower to higher spatial frequencies, which can be described by a transfer function (TF) for S-BP [14]. Consequently, the modulation of spatial frequencies \( g = (g_x, g_z) \) in the Fourier transform of the layergram \( f_b(g_x, g_z) = \mathcal{F}[f_b(x,z)] \) can be corrected by multiplication with a weighting function \( W(g) \), the inverse of the TF. This is the concept of the Weighted BP (W-BP) which finally retrieves the object function by

\[
f(x,z) = \mathcal{F}^{-1}[\mathcal{F}[f_b(x,z)] \cdot W(g_x, g_z)].
\]

In the continuous (analytical) case, i.e. infinitesimally small tilt increments and a tilt range of \( \pm 90^\circ \), the transfer function of the S-BP is the reciprocal modulus of the spatial frequency. Thus, the weighting function or the so-called analytical weighting filter (WF) is \( W_0(g_x, g_z) = |g| \).

Indeed the analytical WF is the 2D Jacobian for the Cartesian to polar coordinate transformation which is, however, missing if considering the S-BP (Eq. [2]) in Fourier space. However, in realistic cases, including possibly non-uniform tilt increments of about \( \pm 70^\circ \), the transfer function is neither axially symmetric nor its slope in radial direction is unity. To consider the limited and discrete number of projections better than the analytical WF, Harauz and van Heel introduced a so-called exact

![Fig. 1. Tomography in single-axis tilt geometry. (a) Schematic of the tilt series acquisition (projection) process. The tilt axis points in \( y \)-direction. The green arrow in the 2D projection corresponds to the 1D projection of the 2D object slice at a certain tilt angle \( \alpha \) as shown in (b). (c) The sinogram is composed of all available 1D projections through the 2D object slice under different tilt angles. Its back-projection leads to the layergram, a blurred version of the original 2D object slice. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)](image-url)
where \( \ell \) denotes the usual scalar product. Then, the following relation between subsequent projections errors hold

\[
\ell_{k+1} = \| \ell_{k} \| = \| (I - \omega R R^{\top}) \ell_{k} \|
\]

\[
= \sqrt{\ell_{k}^2 - 2 \omega \ell_{k} (R R^{\top} \ell_{k}) + \omega^2 (R R^{\top})^2 \ell_{k} \ell_{k}^{\top}}
\]

\[
\leq \sqrt{\ell_{k}^2 - 2 \omega \ell_{k} c_k + \omega^2 c_k^2}
\]

\[
\leq \sqrt{(1 - 2 \omega c_k + \omega^2 c_k^2) c_k}.
\]

Therefore, \( \omega \) has to be chosen such that the square root in the last line remains smaller than unity in order to decrease the error, i.e.

\[
0 < \sqrt{(1 - 2 \omega c_k + \omega^2 c_k^2) c_k} < 1.
\]

The left part of this inequality is always true because the parabola determined by the bracket under the square root never crosses zero from above. In other words, the zeros of the quadratic equation determined by

\[
\omega(\omega C_k - C_k) = 0
\]

are complex, which can be proven with the help of the Cauchy–Schwarz inequality and Eqs. (11) and (12) yielding

\[
(r_k, R R^\top r_k)^2 \leq \|r_k\|^2 \|R R^\top r_k\|^2
\]

\[
\Rightarrow c_k^2 \|r_k\|^4 \leq C_k \|r_k\|^4
\]

Consequently, the projection errors decrease and the reconstruction converges if the right hand side of the inequality (14) is fulfilled, i.e.

\[
2 \omega c_k - \omega^2 C_k \geq 0
\]

yielding the following condition for the damping parameter:

\[
\omega < \omega_{\max} \quad \text{with} \quad \omega_{\max} = 2 c_k / C_k
\]

in each iteration step. The projection error stops to decrease if \( r_k \in \ker(R R^\top) \). Indeed, to further minimize the error in projection space one has to replace the back-projection \( R^\top \) by a more suitable operator which reduces the size of \( \ker(R R^\top) \). The optimal operator is given by the Moore–Penrose inverse \( R^+ \) of the projection operator \( R \), defined by \( \min_{\ell} \| \ell - \ell \| = \| \ell - \ell - f \| \). If using this pseudoinverse, one iteration cycle is sufficient and no \( \omega \) is required. However, the computation of \( R^+ \) is very expensive, requires regularization and will be discussed in another publication. It is therefore convenient to use an easily computable approximation for \( R^+ \) which is automatically regularized by multiplying the damping factor \( \omega \) within the iterative scheme. Our approach uses the W-BP to approximate \( R^+ \) and combines it with a conventional SIRT to dampen the errors of the approximation. This Weighted SIRT (W-SIRT) leads to tomographic reconstructions with higher fidelity and enhanced resolution compared to SIRT and W-BP as it is demonstrated in the following.

3. Basic principle and implementation of W-SIRT

The major modification in the here presented weighted SIRT (W-SIRT) compared to a conventional SIRT is the use of the weighted back-projection (W-BP) at the beginning and at each iteration step of the algorithm instead of a simple back-projection (S-BP). The much more exact W-BP improves the convergence properties considerably as shown in the previous section. Moreover, an improved version of the so-called exact WF (Section 3.1) is used for the W-BP. However, the multiplication of such a “ramp-like” WF [see Eq. (6)] in Fourier space prior to the inverse Fourier transformation leads to artifacts in the W-BP tomogram. Since these artifacts appear especially at borders of the reconstructed image, we numerically generated the difference sinogram. This is back-projected again to obtain the final reconstruction.
region, the tomogram is multiplied by a window function (spherical or rectangular shape) that dampens the signal near the borders to zero before it is forward projected in the next iteration step.

We implemented the W-SIRT algorithm in Gatan Inc.’s Digital Micrograph (DM) scripting language, a widespread image processing software for electron microscopy. Moreover, to increase the processing speed, we implemented and compiled computationally expensive parts, i.e. forward- and back-projection, in C++ using the DM software developer’s kit (DMSDK). To exploit the facilities of state-of-the-art multi-core CPUs, the algorithm is parallelized by multithreaded programming.

Fig. 2 shows the user interface of the self-developed tomographic reconstruction software “Reconstruct 3D” [27], which provides a convenient way to adjust the parameters for the W-SIRT (e.g. number of iterations and filter type). The actual W-SIRT routine works for 2D only, however, as explained above, in single axis tilt geometry, the 3D reconstruction can be performed using subsequent reconstruction of 2D tomograms from sinograms perpendicular to the tilt axis.

3.1. Weighting filter

The concept of point spread function (PSF) and transfer function (TF) for tomographic reconstructions is described, e.g., in [14]. We can generate it also here, although its validity is limited in the discrete and finite sampling case as discussed previously in Section 2.2. To generate the PSF numerically, the layergram of a single central point (the central pixel is 1, whereas pixels are 0 elsewhere) is computed from its sinogram taking into account the same tilt directions as for the actual reconstruction later. The Fourier transform of the PSF leads to the transfer function (TF). To finally obtain the exact weighting filter, the reciprocal of the TF must be computed. However, as mentioned in Section 2.1, this is only meaningful above a certain threshold value to avoid division by zero and strong enhancement of noise.

The advantage of such an exact WF over an analytical one is demonstrated in Fig. 3, where the PSFs of three different tomographic reconstructions from a tilt range of –90° to +89.5° in steps of 0.5° are compared: The PSF of W-BP after using an analytical WF (a), the PSF of a SIRT after 30 iterations (b), and the PSF of W-BP with our exact WF (c). The PSFs (a) and (b) are extended over a few pixels, i.e. the resolution in a corresponding tomogram would be lower than one pixel, whereas the PSF in (c) is given by the central pixel only. This can be seen in detail in the 1D line profiles across the PSFs (Fig. 3d), which reveal the radial modulation of the PSF. It must be noted that only in the case of sufficiently small tilt increments, no missing wedge, and for the central pixel only, the W-BP with our exact WF provides a perfect reconstruction. As already mentioned above, this is generally not the case. The influence of pixel position and missing wedge is discussed in detail in Section 4.1.

In Fig. 3(e–g), the weighting filters which correspond to the PSFs (a–c) are shown. Since there is no WF in the conventional SIRT (only S-BP is used within each iteration), we computed an ‘effective’ WF for comparison (Fig. 3f) given by the ratio of the Fourier amplitudes of the final SIRT tomogram and the layergram. As depicted by the line profiles in Fig. 3(h), the ramp of the exact WF (black) in radial direction has a steeper slope than the analytical one (red). This indicates that higher spatial frequencies are amplified stronger by our exact WF leading to a higher lateral resolution in the tomogram. Interestingly, the slopes for SIRT (blue) and exact WF are almost identical at lower spatial frequencies, but at higher spatial frequencies the slope for SIRT falls off.

If the tilt range is limited to, e.g. ±70°, the Fourier space is undersampled by the central slices due to missing projections in real space. Thus, the generation of the WF from the reciprocal of the S-BP of the central point (transfer function) becomes critical below a certain threshold $t$ as mentioned in Section 2.2. A suitable choice for $t$ is

$$t = \frac{T_{WF}(0,0)}{N_t},$$

where $T_{WF}(0,0)$ is the value of the transfer function at the origin and $N_t$ the total number of tilts. The threshold restricts the

Fig. 2. Screen-shot of the self-developed tomographic reconstruction software “Reconstruct 3D” including the Weighted SIRT algorithm. The program runs in Gatan Inc.’s Digital Micrograph software. The parameters, e.g. number of iterations, filter type or size of tomogram, can conveniently be adjusted via a user interface.
application of WF such that only pixels are corrected, which are sampled by at least one central slice, i.e. the Fourier transform of a certain projection.

A WF for $70^\circ$ tilt range and $2^\circ$ step size ($N_\alpha = 71$) is shown in Fig. 4a. This WF does not change the mean value of the tomogram after its application (central pixel is 1), but amplifies higher spatial frequencies up to a maximum factor of $N_\alpha$. Finally, it is sometimes useful to multiply the threshold value by a factor $s$ depending on the quality (noise, total number of tilts) of the experimental data. This would correct spatial frequencies which are sampled by at least $s$ central slices.

4. W-SIRT vs. SIRT

In this section we compare the W-SIRT with a conventional SIRT with respect to signal fidelity, lateral resolution and convergence behavior. To this end, we choose a very simple and elementary object as shown in Fig. 5. The object consists of only three points (value 1): One in the center, one near the border and one in between. In contrast to the tomographic experiment the object is known, which facilitates studying the influence of the position on signal height and elongation of the reconstructed point (Section 4.1). Moreover, we demonstrate the capability of W-SIRT (incapability of SIRT) to reconstruct a sinusoidal modulation on a finite cylindrical object, although its modulation is originally in the direction of the missing wedge (Section 4.2).

The influence of noise is not addressed in detail here for the following reason: Noise is a non-projective error in the tilt series (sinogram), and its propagation into the reconstructed tomograms crucially depends upon the spatial structure of the former, e.g. high or low frequency, and the regularization parameter, e.g. number of iterations. A characterization of noise transfer would therefore require first to reproduce the correct projected data noise resulting from the experiment, and second to consider its influence under changing regularization on the reconstruction. This is beyond the scope of this paper. Nevertheless, we supplement (Suppl. S2) the reconstruction of a simulated object reproducing the experimental sample in Section 5 with realistic noise level. In this case the W-SIRT reconstruction was more accurate than the SIRT reconstruction in line with the following numerical results.

4.1. Case study 1: reconstruction of three points

The tomograms of the test pattern shown in Fig. 5 are reconstructed by SIRT and W-SIRT from two tilt ranges: from

Fig. 3. Tomographic reconstruction of the central pixel (value 1) from a tilt range of $-90^\circ$ to $+89.5^\circ$ in steps of $0.5^\circ$ using (a), W-BP with analytical weighting filter, (b) SIRT with 30 iterations, and (c) W-BP with an exact weighting filter. The line profiles (d) along the white arrows in (a–c) exhibit that only (c) represents a perfect reconstruction. (e–g) Weighting filters corresponding to (a–c), while an effective WF for the SIRT is computed by the ratio of the Fourier amplitudes of final the SIRT tomogram and the layergram (after S-BP). The line profiles (h) along the white arrows in (e–g) show the ramp shape of the weighting filters in radial direction. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

Fig. 4. (a) Weighting filter for $\pm 70^\circ$ tilt range and $2^\circ$ steps. The maximum value is given by the number of projections (here 71). (b) Point spread function of the weighted back-projection, i.e. inverse FT of (a).
90° to +89.5° in 0.5° steps, and from ±70° in 2° steps. In the W-SIRT tomograms, the three points are less smeared out than in the SIRT case which corresponds to a smaller PSF. Consequently, the spatial resolution is higher. Moreover, Fig. 5 shows the plotted maximum intensities of the three points as a function of iteration steps. The graphs reveal that the signal height converges faster and closer to the original value (unity) in the W-SIRT compared to the SIRT. Even at full tilt range and after 300 iterations, there is still a discrepancy in the SIRT tomogram for all three points; instead of the original value 1 one obtains 0.79 for point 1 (border), 0.76 for point 2 and 0.98 for point 3 (center). Also the W-SIRT does not provide a perfect reconstruction: only the central point 3 reaches the exact value immediately; for point 1 a rather high number of iterations (164) is needed to yield 0.99; and for point 2 one obtains a value of 0.93 after 300 iterations. Interestingly, although point 2 is closer to the center than point 3, its reconstructed value is smaller. We therefore conclude that no simple relationship between position and accuracy of the reconstructed point exists. In other words no general decay of resolution in radial direction from the center to the border could be observed. In fact, the complicated position dependency is introduced by sampling real and Radon space, which is different for each point (depending on tilt range and step). This restricts the validity for describing the resolution in tomograms with the concept of the PSF, i.e. with a position independent convolution.

At a tilt range of ±70° the signal height will never attain the original value for both W-SIRT and SIRT: Using W-SIRT (SIRT) point 1 reaches only 0.86 (0.85), point 2 0.70 (0.63) and point 3 0.70 (0.68). This is mainly due to the limited resolution in direction of the missing wedge. If considering also the elongation of these points, i.e. adding their adjacent pixels to the signal height, one obtains for both algorithms virtually the original value unity.

\[ \Delta = \sqrt{\frac{1}{N} \sum_{n=1}^{N} (f_n^{\text{rec}} - f_n)^2}, \]  

a measure for the discrepancy between original distribution \( f \) and computed tomogram \( f^{\text{rec}} \) with pixel number \( N \) of the discrete data-set, as a function of iterations. The graph clearly exhibits that the W-SIRT (solid lines) converges faster and closer than the SIRT (dashed line) to the original distribution at both tilt ranges full (black) and limited (red). As expected, when using W-SIRT the discrepancy at full tilt range is much lower than at limited tilt range. Interestingly, in case of SIRT there is virtually no difference between limited and full tilt range for \( \Delta \) throughout the first 20 iterations. This can be explained with the fact that SIRT starts to correct for lower spatial frequencies which are less affected by the
missing wedge and then corrects the higher spatial frequencies which are more affected by the missing wedge.

In the following, we investigate in more detail how the three points in the tomograms are elongated due to the limited resolution in direction of the missing wedge at the tilt range of \( \pm 70^\circ \). An elongation factor \( e_{xz} \), i.e. the ratio between elongated \( z \)- and non elongated \( x \)-direction can be defined according to [28] by

\[
e_{xz} = \sqrt{\frac{\alpha_m + \sin(\alpha_m) \cos(\alpha_m)}{\alpha_m - \sin(\alpha_m) \cos(\alpha_m)}}
\]

where \( \alpha_m \) denotes the maximum tilt angle. We note that such a definition can only be an approximate measure of the real object spread, e.g. here the point spread is approximated by a parabola (see below and Suppl. S3 for details). To measure the elongation with sub-pixel accuracy, we applied the so-called zero padding in Fourier space. This means the FT of the tomogram with originally 64 by 64 pixels is extended to a 2048 by 2048 image by adding pixels of value zero around the original data-set, and the inverse FT is computed subsequently. This corresponds in real space to a sub-pixel interpolation by convolution with a sinus cardinalis (sinc) function.

Therefore, a single pixel is spread to a sinc function whose main lobe has a full width half maximum (FWHM) of 1.2 px.

Fig. 7a shows the sub-pixel interpolated tomograms of the test object. The profiles below correspond to the line scans (marked by the arrows) across point 1 for S-BP and W-SIRT tomogram. The elongation of the point in \( z \)- with respect to \( x \)-direction is very high at S-BP, but is greatly reduced in the W-SIRT reconstruction. Because the expression of the elongation factor in Eq. (21) relies on approximating the PSF with a parabola, we plotted also parabolic functions as dashed lines, which intersect the maximum and the two FWHM positions of the profiles across point 1. They show that the parabolic approximation is reasonable only at the inner part of the PSF. The diagram in Fig. 7b depicts elongation factors \( e_{xz} \) of all three points after reconstructing them with S-BP, W-BP, SIRT (300 iterations) and W-SIRT (300 iterations), respectively. The elongation factors are determined by the ratio of the FWHM between \( z \)- and \( x \)-directions. The values for \( e_{xz} \) lie in a range from 1.17 to 1.52, where the W-SIRT always delivers the smallest elongation. However, regardless which reconstruction algorithm is used, there is a relatively large spread in elongation between the different points, for example, in the W-BP tomogram one obtains for point 2..

Fig. 7. Elongation due to missing wedge. (a) Sub-pixel interpolated tomograms of the test object (Fig. 5 reconstructed from a tilt range of \( \pm 70^\circ \) in steps of 2\(^\circ\) with intensity profiles in horizontal (\( x \)) and vertical (\( z \)) direction through point 1 (solid line)). The comparison of the profiles reveals an elongation due to the missing wedge along \( z \)-direction. The dashed lines represent a parabolic approximation of the line profiles. (b) Elongation factors using the FWHM ratios between \( z \)- and \( x \)-directions.
\(e_{zx} = 1.39\), but for point 3 only \(e_{zx} = 1.17\). This again illustrates how significantly the resolution depends on position. We can also compare these values with the elongation factor calculated by Eq. (21) yielding \(e_{zx} = 1.31\) at \(\alpha_m = 70^\circ\), which is in the range of the numerically obtained values.

4.2. Case study 2: reconstruction of sinusoidal object

A second test object, a sphere containing a sinusoidal modulation (Fig. 8), has been generated to demonstrate that it is possible to gain more information with W-SIRT than with conventional SIRT even when the sinusoidal modulation is in the direction of the missing wedge. In this particular case, a tilt range of \(\pm 74^\circ\) at a tilt increment of \(2^\circ\) has been chosen. Because the spatial frequency of this modulation is low and the size of the object is small enough, a part of the information about the corresponding reflection is convoluted out of the missing wedge region, from where it can be reconstructed. As depicted in Fig. 8a the reconstruction of the original object succeeds with the W-SIRT, whereas it fails with the SIRT. The supplementary information (Suppl. S1) shows snapshots after a certain number of iterations that visualize how the sinusoidal modulation propagates iteration by iteration from the horizontal border to the center of the object for both algorithms at different speed. The better quality of the W-SIRT reconstruction becomes also evident in the line profiles (Fig. 8b), and the RMS error vs. the number of iterations (Fig. 8c). This example shows that certain limits of conventional SIRT (and W-BP) can be overcome by using W-SIRT.

5. Experimental example

Recently, the W-SIRT algorithm has been successfully applied in the field of electron tomography, such as dark field TEM (DFTEM) tomography [29] and electron holographic tomography (EHT) [8]. Here, we show the 3D reconstruction of the electrostatic potential of a latex sphere on carbon foil shadowed with gold particles using EHT. A tilt series of off-axis electron holograms (\(\pm 66^\circ\) tilt range, \(2^\circ\) tilt steps) was recorded with a Philips CM200FEG ST/LL TEM in Lorentz mode using the Tomographic Holographic Microscope Acquisition Software package THOMAS that provides a largely automated acquisition procedure [3]. The Lorentz mode provides only medium resolution (ca. 5 nm), but a relatively large field of view (ca. 500 nm) for off-axis electron holography. The tilt series of object holograms and corresponding empty (object free) holograms has been reconstructed to obtain amplitude and phase of the object exit wave using the procedure described in detail by e.g. Lehmann and Lichte [30]. Empty holograms are necessary to

---

Fig. 8. (a) Test object containing a sinusoidal modulation in missing wedge direction and its reconstructions by W-SIRT and SIRT. (b) The line profiles in vertical direction (indicated by arrows in (a)) show that it is possible to reconstruct the original distribution with W-SIRT, but not with SIRT. (c) RMS error \(\Delta\) (Eq. (20)) to the original distribution as a function of iteration number.
Thus, dividing the phase tilt series by \( C_\ell \) leads to a tilt series of projected potentials, i.e., the desired data-set for the tomographic reconstruction. However, before starting the reconstruction the experimental tilt series has to be aligned, i.e. the residual displacements between the projections have to be corrected. Here, we used the gold clusters on the specimen as fiducial markers to apply the alignment routine implemented in the IMOD software package [31]. Subsequently, the tomographic reconstruction of the aligned tilt series is performed yielding the 3D distribution of the electrostatic potential.

The 3D potential of the latex sphere shadowed with gold particles is displayed in Fig. 9. It was reconstructed by W-SIRT using 100 iterations. Note that the colors in the volume rendering (Fig. 9b) are associated with quantitative potential values, i.e. the mean inner potential (MIP). This becomes more evident in the line profiles through the 2D slice shown in Fig. 9c,d: The peak on the left of the red line profile corresponds to the MIP of gold \( V_{\text{Au}}^{\text{MIP}} \), whereas the plateau corresponds to the MIP of polystyrene latex \( V_{\text{Lat}}^{\text{MIP}} \). The MIP values obtained from W-SIRT are with \( V_{\text{Au}}^{\text{MIP}} = 30.2 \text{ V} \) and \( V_{\text{Lat}}^{\text{MIP}} = 8.4 \text{ V} \) in excellent agreement with literature values demonstrating the quantitative character of EHT: Schowalter et al. reported an MIP for gold \( V_{\text{Au}}^{\text{MIP}} = 29.8 \text{ V} \) computed by density functional theory (DFT) [32]. Wang et al. determined a MIP of \( V_{\text{Lat}}^{\text{MIP}} = 8.5 \text{ V} \) by a recursive \( \chi^2 \) minimization routine at holographically reconstructed phase images (2D) [33].

A more detailed analysis about the convergence behavior of SIRT and W-SIRT for this experimental dataset is depicted in Fig. 10.

### 6. Summary and conclusion

We have developed a novel tomographic reconstruction scheme, referred to as W-SIRT, which employs an exact weighting filter in a sequential iterative reconstruction technique (SIRT) with an adaptive damping factor. Extensive numerical and experimental tests show a remarkable improvement with respect to the reconstruction error compared to standard SIRT and W-BP algorithms. Particularly in the experimentally often unavoidable case of a limited tilt range

Resolution and signal (potential) in the tomograms increase with iteration number for both techniques but slower in the case of SIRT. However, with higher iteration number also the noise is enhanced. For example, it is apparent in the tomograms after 400 iterations that the gold signal at the edge of the latex sphere begins to spread into the center hence adulterating the potential within the sphere. This also influences the determination of the MIP of the latex sphere when interpreting the mean value of the region which is indicated by the red dashed boxes in Fig. 10a as MIP. The latter is plotted as a function of iteration number in Fig. 10b and shows that this mean value is for W-SIRT at each iteration higher than for SIRT. The better convergence properties of the W-SIRT reconstruction are additionally supported by the lower RMS error in projection space at each iteration (Fig. 10c). This is consistent with our theoretical considerations in Section 2.3, where we derived that the kernel of the operators for backward and forward projection which determines the RMS error in projection space is smaller for W-SIRT than for SIRT.

Finally, the convergence behavior is analyzed in more detail in the supplementary material (Suppl. S2). To this end, we generated a model of the latex sphere, projected and reconstructed it by taking into account the same tilt range as in the experiment. As a result, the MIP values obtained from the W-SIRT tomogram are again closer to the original value even when we simulate a realistic noise level on the projected model data.

![Fig. 9. 3D reconstruction of a latex sphere on carbon foil shadowed with gold particles. The electrostatic 3D potential (b) was reconstructed from a phase image tilt series (a) covering a \( \pm 66^\circ \) tilt range in 2° steps using W-SIRT (100 iterations). The 1D line profiles (d) through the 2D slice (c) provide quantitative insight into the potential structure: The peak on the left in the red line profile corresponds to the MIP of gold, whereas the plateau corresponds to the MIP of polystyrene latex. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)](image)
W-SIRT tomograms show a significantly reduced elongated point spread in direction of the missing tilt angles, i.e., an improved reconstruction of spatial frequencies in the vicinity of the missing wedge. These improvements facilitate the simultaneous reconstruction of both fine details and larger features (in our case the 3D electrostatic potential of Au nanoparticles and latex spheres) in one tomogram at nanometer resolution, thus providing more detailed insight in the 3D nanostructure of emerging materials. Additionally, we have analyzed the position dependency of the reconstruction quality (resolution) and we provide analytic formulas for the adaptive damping factor. We also implemented the W-SIRT algorithm including a graphical user interface in Gatan Inc.’s Digital Micrograph, a widespread image acquisition and processing software for electron microscopy. Finally, we emphasize that the reconstruction algorithm is not restricted to the reconstruction of projected potentials in electron holography but can be applied to all tomographic problems sharing the same projection geometry.

Acknowledgments

We thank all members of the Triebenberg group for enlightening discussions, in particular Dr. F. Börnert for critical reading the manuscript. D.W. appreciates the great help from Dr. T. Niermann from TU Berlin on multithreaded programming in C++. The research leading to these results has received funding from the European Union Seventh Framework Programme under Grant Agreement 312483 – ESTEEM2 (Integrated Infrastructure Initiative – 13).

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.ultramic.2013.07.016.

References


Fig. 10. Convergence of SIRT and W-SIRT. (a) SIRT and W-SIRT reconstructions of the latex sphere from Fig. 9 at different numbers of iterations. (b) MIP values of polystyrene latex obtained from the mean value of the box indicated in (a) as a function of iterations. The error bars represent the standard deviations. (c) Root mean square of the difference between original and computed projection from the tomogram as a function of iterations. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)


[26] G. Harauz, M. van Heel, Exact filters for general geometry three dimensional reconstruction, Optik 73 (1986) 146–156.

[27] The W-SIRT algorithm implemented in Digital Micrograph may be obtained upon request to the authors. If you publish work using the W-SIRT algorithm, please reference us by citation on this paper.


[31] IMOD Software Package for Image Processing, Modeling and Display, Boulder Laboratory, University of Colorado.
