A resolution criterion for electron tomography based on cross-validation

Giovanni Cardone a, Kay Grünwald b, Alasdair C. Steven a, *

a Laboratory of Structural Biology, National Institute of Arthritis, Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD 20892, USA
b Department of Molecular Structural Biology, Max Planck Institute of Biochemistry, 82152, Martinsried, Germany

Received 18 March 2005, and in revised form 27 April 2005
Available online 31 May 2005

Abstract

Despite much progress in electron tomography, quantitative assessment of resolution has remained a problematic issue. The criteria that are used in single particle analysis, based on gauging the consistency between density maps calculated from half data sets, are not directly applicable because of the uniqueness of a tomographic volume. Here, we propose two criteria based on a cross-validation approach. One, called \( FSC_{eo} \), is based on a Fourier shell correlation comparison between tomograms calculated from the even and odd members of a tilt series. The other, called noise-compensated leave-one-out (NLOO), is based on Fourier ring correlation comparisons between an original projection and the corresponding reprojection of the tomogram calculated from all the other projections, taking into account the differing noise statistics. Plotted as a function of tilt angle, they allow assessment of the angular dependence of resolution and quality control over the series of projections. Integrated over all projections, the results give a global figure for resolution. Tests on simulated tomograms established consistency between these criteria and the \( FSC_{ref} \), a correlation coefficient calculated between a known reference structure and the corresponding portion of a tomogram containing that structure. The two criteria—\( FSC_{eo} \) and NLOO—are mutually consistent when residual noise is the major resolution-limiting factor. When the size of the tilt increment becomes a significant factor, NLOO provides a more reliable criterion, as expected, although it is computationally intensive. Applicable to entire tomograms or selected structures, NLOO has also been tested on experimental tomographic data. Published by Elsevier Inc.

Keywords: Fourier shell correlation coefficient; Anisotropic resolution; Signal-to-noise ratio

1. Introduction

Electron tomography (ET) is a rapidly developing three-dimensional imaging technique that bridges the resolution gap between cryo-EM/single-particle analysis (SPA) and optical microscopy (Grünewald et al., 2003a; McEwen and Koster, 2002). Its most notable feature is its general applicability: ET may be used to study the structural organization of complex molecular objects in situ (e.g., Medalia et al., 2002; Perkins et al., 2001) or of individual isolated complexes (e.g., Nitsch et al., 1998). On the other hand, the capability to visualize single specimens imposes limitations for vitrified specimens, in terms of the maximum tolerable electron dose and other technical issues.

In this context, quantitative assessment of resolution remains a key issue. Standard SPA criteria cannot be applied to electron tomograms in a straightforward way. In the Fourier shell correlation (FSC) (Harauz and van Heel, 1986) and the differential phase residual (DPR) (Frank et al., 1981)—two measures commonly used in SPA—the input projections are randomly divided into half-sets, and resolution is measured in terms of the consistency between the resulting density maps. The efficacy of these measures relies on the (usually high)...
redundancy of the input data: when the number of projections is limited, as in tomography, they may be expected to underestimate the resolution obtainable from the full data set.

If a tomogram happens to contain a recognizable component whose structure is already known (from structural analysis by SPA or crystallography) to higher resolution than is likely for the tomogram, it may be used to estimate the resolution of the tomogram by calculating a FSC correlation coefficient (Grünewald et al., 2003b). (Similar bootstrapping is done in resolution estimation for SPA in situations in which a reconstructed complex contains a component for which a high resolution crystal structure exists e.g., Ludtke et al., 2001; San Martin et al., 2001.) This approach is intuitively appealing but is limited to tomograms that contain a suitable reference structure, which will not in general be the case.

To date, there have been few formal measurements of resolution in electron tomographic studies and investigators have tended to be guided by visual clues concerning the visibility or otherwise of certain features in the tomogram. Recently, some approaches to resolution assessment have been proposed (Penczek, 2002; Taylor et al., 1997; Unser et al., 2005—see Section 4). However, no such measure has yet entered widespread use.

In statistics, the term cross-validation (Efron, 1979) covers a wide class of model evaluation methods which have in common the separation of the input data into two sets—one for performing the analysis, and the other for assessing its goodness-of-fit. This approach has been pursued in many fields. In crystallography, it underlies the R-factor, $R_{\text{free}}$ (Brünger, 1992), which gives an unbiased measure of the consistency between atomic models and diffraction data. In EM, cross-validation has been used to quantify the bias introduced by a reference model in projection-matching (Shaikh et al., 2003). These authors suggested a procedure to derive a Free-FSC curve for SPA, i.e., a curve theoretically free from reference bias. Here, we present a method for assessing the resolution of tomograms based on cross-validation in the projection domain.

2. Theory

2.1. Definitions

A tomographic tilt series consists of $N_p$ projection images. We define $X^{(i)}_{m,n}$ as component $(m,n)$ of the Fourier transform of projection $i$. The corresponding component for the reprojection calculated from the tomogram is $\tilde{X}^{(i)}_{m,n}$, and $\tilde{X}^{(-i)}_{m,n}$ for the reprojection of the tomogram from which projection $i$ was excluded. The Fourier ring correlation (Saxton and Baumeister, 1982) between transforms $F^{(i)}$ and $G^{(i)}$ is given by

\[
\text{FRC}_{FG}^{(i)}(k) = \frac{\sum_{m,n < R(k)} \text{Re}\left\{F^{(i)}_{m,n}G^{(i)}_{m,n}\right\}}{\left(\sum_{m,n < R(k)} |F^{(i)}_{m,n}|^2\right)^{1/2} \left(\sum_{m,n < R(k)} |G^{(i)}_{m,n}|^2\right)^{1/2}},
\]

where the asterisk denotes the complex conjugate, $k$ is the radial frequency, and $R(k)$ refers to an annular zone in Fourier space of mean radius $k$. Its width is given by the reciprocal of the minimum dimension of the volume under analysis. The Fourier shell correlation for three-dimensional data is analogous to Eq. (1), except that in this case $R(k)$ is a shell and the frequency components of a transform have three indices $(m,n,p)$.

2.2. The ideal experiment

We first need to define a correlation-based measure of resolution that is applicable in tomography. Operationally, resolution is limited by noise as well as by resolving power. In this context, the Fourier shell correlation (FSC) between two maps of the same volume, calculated from independent but statistically equivalent data sets, is related to the signal-to-noise ratio (SNR) of the maps by (Frank and Al-Ali, 1975; Penczek, 2002)

\[
\text{FSC}(k) = \frac{\text{SNR}(k)}{\text{SNR}(k) + 1}.
\]

We define resolution in terms of the FSC comparing two such maps. This definition is idealized in that it requires two reconstructions from independent ‘full’ data sets, which is not usually possible. However, we show below that approximations to this measure of resolution may be obtained by defining proper criteria.

2.3. The even/odd Fourier shell correlation

In SPA, the FSC compares two density maps calculated from half data sets of single projections of many nominally identical particles. Since, in the general case, a tomogram represents a unique volume, there will not be two such maps to compare. However, independent tomographic representations of the same volume may be calculated from subsets of projections in a tilt series. A natural division is to select the even and odd projections, respectively. This strategy guarantees the most uniform coverage in the Fourier domain and retains the property of equally spaced projections. If the number of projections (angular increment) is the resolution-limiting factor, the resolution of the subtomograms should be lower by a factor of two than that of the full tomogram, according to the formula

\[
r_d \approx d\theta \cdot D,
\]

where $D$ is the diameter of the particle and $d\theta$ is the tilt increment in radians (Bracewell and Riddle, 1967;
Crowther et al., 1970). This formula assumes a noise-free situation. In practice, noise may place a more stringent limit on the resolution of tomograms, and the limited range of accessible tilt-angles imposes a further constraint (Frank, 1992).

In SPA, maps from half data sets are necessarily noisier than the map obtained from the full data set: accordingly, the FSC under-estimates the resolution of the latter map. To compensate for this effect, considerations of signal-to-noise (Rosenthal and Henderson, 2003) allow one to derive a relationship between the FSC(k) calculated in this way and FSC(k), a measure that approximates the ideal experiment defined above, i.e.,

\[
\text{FSC}'(k) = \frac{2\text{FSC}(k)}{\text{FSC}(k) + 1}.
\]

This equation assumes that the SNR for each map from a half data set is half that of the map from the full data set. This assumption is reasonable in SPA because of the expected redundancy of input projections for each orientation, while it is less likely to be satisfied in electron tomography. Nevertheless, we empirically adopt this correction in defining the even/odd FSC, i.e., first we calculate a correlation coefficient FSC*(k) from the two half tilt series tomograms, and then correct it as follows:

\[
\text{FSC}_{e/o}(k) = \frac{2\text{FSC}'(k)}{\text{FSC}'(k) + 1}.
\]

Since each subtomogram has some information content that is missing from the other, according to the central section theorem (Kak and Slaney, 1988), it follows that the accuracy of FSC_{e/o}(k) in estimating the resolution of the full tomogram depends on the degree of overlap in the Fourier domain between adjacent projections in the tilt series: the larger the tilt increment or the diameter of the structure under analysis, the greater is the resolution underestimate given by FSC_{e/o}(k). However, we have observed that the estimate provided by FSC_{e/o} is in agreement with that given by other criteria that are not limited in this way to the same extent (see below), provided that the resolution obtained, r, is inferior to the critical resolution, according to

\[
r > r_{\text{crit}} = 2 \cdot r_3,
\]

where the limit r_3 is posed by the angular increment of the full tilt series Eq. (3).

2.4. The leave-one-out method

Since the number of projections is likely to limit resolution in ET, at least for moderately thick biological specimens, halving this number is a drawback, despite compensatory approximations (above). An alternative approach to assessing self-consistency is to compare a single projection with the corresponding reprojection of a tomogram calculated from all the other projections except the projection in question, which is omitted to avoid bias. Since only one projection is missing, the resolution of this tomogram should be very close to that of the complete tomogram. This approach, which we refer to as leave-one-out (LOO), represents another example of cross-validation. It is, in fact, an extreme case of the many possible ways in which the full tilt series could be partitioned and analyzed.

Since there are N_p input projections, an equal number of tomograms should, in principle, be calculated. However, an adequate assessment of resolution may be achieved with a subset of LOO calculations (see below). Although this approach is computationally intensive, it has compensations. Plotting resolution against tilt angle allows one to assess the dependence of resolution upon direction and to appraise the tilt series.

2.5. 2D noise-compensated leave-one-out estimation

Since projections and reprojections have differing noise statistics, one must allow for this distinction in using them to define a measure of resolution. We assume that the noise in the input projections is additive. The reprojections also have a noise component, but its variance is reduced in proportion to the number of projections used to calculate the tomogram. Furthermore, the tomogram can be calculated only for a partial thickness of the specimen, or we may be interested only in a specific structure confined to a subvolume. In the latter cases, reprojections from the selected subvolume contain just the signal of interest, while the corresponding input projections also have contributions from adjacent features. In the following, contributions to the input projections from any object that is not part of the (sub)volume selected for resolution analysis is treated as noise.

On this basis, we estimate the noise as the difference between the input projection and the corresponding reprojection from the tomogram generated from all the input projections. The discrepancy between an input projection and the corresponding reprojection from a tomogram reconstructed without that projection can be quantified by means of a normalized square difference. Thus, exploiting the relationship between correlation and normalized square difference (Appendix A), we define the noise-compensated leave-one-out in two-dimensions (NLOO-2D) measure as

\[
\text{NLOO-2D}^{(i)}(k) = \frac{\text{FRC}_{XX}^{(i)}(k)}{\text{FRC}_{XX}^{(i)}(k)},
\]

where FRC_{XX}^{(i)}(k) and FRC_{XX}^{(i)}(k) are defined in Eq. (1). Here, the denominator is a noise compensation term. This measure, NLOO-2D, closely approximates that which would be obtained in the ideal experiment, i.e., by comparing the entire tomogram, via FSC, to a
hypothetical additional tomogram calculated from an independent but statistically equivalent tilt series (Appendix B).

Since the 2D Fourier transform of a projection forms a slice of the 3D Fourier transform of that volume, each NLOO-2D\(^{(k)}\) curve can be assumed to provide a resolution for the corresponding slice of the reconstructed volume.

### 2.6. Noise-compensated leave-one-out estimation in 3D

We may generalize NLOO-2D\(^{(k)}\) to define a measure of resolution for the whole tomogram,

\[
\text{NLOO-3D}(k) = \frac{1}{\sqrt{N_p}} \frac{\sum_{i \in \mathcal{R}(k)} \text{Re}\{X_n^{(i)} Y_{-m,n}^{(i)}\}}{\sqrt{\sum_{i \in \mathcal{R}(k)} |X_n^{(i)}|^2 \sum_{i \in \mathcal{R}(k)} |Y_{-m,n}^{(i)}|^2}}
\]

(see Section 2.1 for definitions). This expression is obtained from (7) by simply performing further summations over both tilt series. (For the sake of clarity, redundant terms in (8) have not been eliminated). Thus, NLOO-3D may be interpreted as the ratio between two Fourier shell correlations.

### 2.7. Dependence of resolution on tilt-angle and location within the volume

The quality of a tomogram is usually not homogeneous. In particular, the limited angular range leaves an unsampled portion of Fourier space—the “missing wedge” for single-axis tilt series, or “missing pyramid” for dual-axis tilt series—that results in inherently anisotropic resolution (Frank, 1992). NLOO-2D, applied to each projection, provides a straightforward way to assess resolution as a function of the tilt-angle. The increasing thickness of the specimen at higher tilt-angles, as well as focal gradients and uncertainties in alignment parameters, also contributes to non-uniformity of resolution. To assess it, both criteria may be applied to selected subvolumes.

### 3. Results

To explore the performance of NLOO-3D and FSC\(_{c/o}\) as measures of resolution, we compared them with results given by a reference FSC. This was done in two ways. First, we tested them on synthetic tomograms calculated from tilt series that were derived computationally from a reference density map. In this experiment, the reference structure occupied the full tomographic volume. Then we applied the measures to a real cryo-ET data set, for which a reference map was available for one (minor) component of the reconstructed volume. In both cases, we calculated the reference FSC over the Fourier domain covered by the tilt series.

All resolution estimates were determined at a threshold of 0.3, unless otherwise specified. According to Eq. (2), this choice corresponds to a SNR\((k)\) close to 0.5 (0.429), i.e., the resolution is identified as the (highest) radial frequency at which the spectral density of the signal falls below half that of the noise. All reconstructions were performed using the \(R\)-weighted back-projection algorithm as implemented in IMOD (Kremer et al., 1996). Both NLOO-3D and FSC\(_{c/o}\) have been implemented in ELECTRA (ELECtron Tomography Resolution Assessment), a software package written in C that relies on Bsoft (Heymann, 2001). Running on many Unix-like platforms, it is freely available from the authors, on request.

#### 3.1. A model experiment

A density map of hepatitis B virus (HBV) capsid, a spherical particle of \(~310\) Å in diameter, was used to generate tilt series. This map has an isotropic resolution of \(9.1\) Å, according to the FSC with a threshold of 0.5 (Conway et al., 1997), and is sampled at \(2.63\) Å. Projections were generated around a single tilt axis, covering the range \(-68\)° to \(+68\)°. Three series were calculated, corresponding to angular steps of \(1\)°, \(2\)°, and \(4\)° (137, 69, and 35 images). White Gaussian noise was added to the projections at a SNR of 1 (e.g., Fig. 1). A SNR of 1 is unrealistically high for many cryo-ET applications, but the effect of this parameter is considered further below. Tomograms were calculated from these data and their resolutions measured by all three criteria.

This experiment is idealized in the sense that real tomograms are also compromised by other factors, including imperfect alignment of the projections, variations in defocus, phase contrast effects, radiation damage, and the modulation transfer function of the camera. Nevertheless, it allows one to calibrate the new measures relative to a reference FSC. Because we
introduced some noise in our simulation, the FSC between the generated tomograms and the reference map, which is essentially noise-free, does not satisfy Eq. (2). A reference measure $FSC_{ref}$ that is consistent with our definition of the ideal experiment (see Section 2.2), and hence satisfies Eq. (2), is given by

$$FSC_{ref}(k) = FSC^2(k),$$

where $FSC$ is directly calculated between the tomogram and the reference map (see the Appendix of Stewart and Grigorieff, 2004).

The results are shown in Fig. 2 where, for each angular increment, $FSC_{ref}$, NLOO-3D, and $FSC_{elo}$ are compared. With an angular step of 1° (Fig. 2A), the three measures provide almost coincident curves. The resolution is given as 11.7 Å in each case. With an angular step of 2° (Fig. 2B), there is still good agreement between NLOO-3D (14.2 Å) and $FSC_{ref}$ (13.9 Å), while $FSC_{elo}$ falls off at higher frequencies, giving a resolution of 14.7 Å. Finally, with an angular step of 4° (Fig. 2C), NLOO-3D and $FSC_{elo}$ both underestimate the resolution as given by $FSC_{ref}$, more severely so for $FSC_{elo}$. The corresponding figures are 22.6 Å ($FSC_{elo}$), 17.5 Å (NLOO-3D), and 16.2 Å ($FSC_{ref}$).

We interpret these results as follows. At 1° tilt increment, the major factor limiting the resolution of the simulated tomogram is residual noise. Both NLOO-3D and $FSC_{elo}$ give values close to $FSC_{ref}$, the highest achievable resolution, and sampling-limited resolution is not a factor ($r_0 \sim 5.4$ Å). At 2° and 4°, NLOO-3D remains close to $FSC_{ref}$, although there is more noise in the tomograms since they are based on fewer projections: in contrast, the $FSC_{elo}$ estimate of resolution is systematically too low as sparse angular sampling becomes important ($r_0 \sim 21.6$ Å at 4°). According to the a posteriori criterion (6), the resolution given by $FSC_{elo}$ is reliable only for the 1° tilt increment. Thus the expected improvement of NLOO-3D over $FSC_{elo}$ under these conditions (see Section 2.4) is realized.

We performed variations of this experiment with different SNRs but otherwise maintaining the same conditions (4° angular step)—Fig. 3. A consistent trend is observed: NLOO-3D gives better results, i.e., ones that are closer to $FSC_{ref}$, although the distinction between the two measures is less pronounced at low values of the SNR, where noise becomes the dominant resolution-limiting factor.

### 3.2. Results on a cryo-electron tomographic data set

We tested the new measures on experimental data for a field of isolated virions of herpes simplex virus type 1, HSV-1 (Grünewald et al., 2003b). The virion has a nucleocapsid of $\sim 1250$ Å diameter, surrounded by a tegument and envelope, giving a full diameter of $\sim 2200$ Å. The 150-Å thick capsid shell is icosahedrally symmetric but neither the DNA that it contains nor the tegument or envelope share this symmetry. The capsid, which accounts for $\sim 20\%$ of the virion mass (Lampert et al., 1969; Newcomb et al., 1989), has been reconstructed by cryo-EM/SPA in numerous studies, at resolutions from $\sim 25$ to 8.5 Å (e.g., Cheng et al., 2002; Newcomb et al., 1993; Zhou et al., 1994, 2000). Fig. 4 shows the zero-tilt projection of one series collect-
ed at 1.5° increments over the range −63° to +63° (Grünewald et al., 2003b). Projections and tomograms were sampled at 7.4 Å.

The resolution given by NLOO-3D for the entire tomogram (Fig. 5A) is 77 Å, while FSC_{e/o} gives a value of 92 Å. However, resolution is expected to vary in different parts of the volume and, in particular, to be lower in regions far from the tilt axis (see Section 3.1) or at the periphery, where some features may not be present in all projections. To assess resolution locally, we selected an intact virion near the center of the field (Fig. 4). The region of interest was defined as a sphere of 1240 Å radius. Its resolution according to NLOO-3D is ~71 Å, whereas
FSC\(_{e/o}\) gave 76 Å. These resolutions are indeed higher than the (average) resolutions obtained for the entire tomogram, and there is less discrepancy between the two measures—Fig. 5B. Taking the virion to be \(\sim2200\) Å thick, \(r_0 \sim 58\) Å.

Previously, these tomograms were assessed by comparing capsids contained in them with a reference density map from SPA/cryo-EM (Gruenwald et al., 2003b). In this study, we followed a similar procedure. Basically, we extracted the capsid from a virion by selecting a spherical shell between radii of 520 and 640 Å, which we then aligned, translationally and orientationally, with the reference. We expected considerable discrepancy between the two capsid representations, since the data were recorded at very different values of defocus (8 vs. 1–2 \(\mu\)m). For this reason and to maintain comparability with the earlier measurements (Gruenwald et al., 2003b), we used a straightforward FSC\(_{ref}\) in this comparison, i.e., without invoking Eq. (9).

The curves obtained by NLOO-3D and FSC\(_{e/o}\) show similar behavior (Fig. 5C). Both give the resolution to be around 68 Å, slightly lower than the critical resolution, \(r_{crit}\) (65 Å; see Eq. (6)), where \(r_{crit}\) was obtained using a diameter \(D\) of 1280 Å. These numbers are closely consistent with FSC\(_{ref}\), which gave a value of 69 Å, although we note that the latter curve shows systematically lower correlation at low frequencies, reflecting the afore-mentioned disparity in defocus. Finally, resolution analysis for capsids in other virions in the same field (see Fig. 4) gave very similar results (67.5–69 Å).

### 3.3. Using NLOO-2D to measure the variation of resolution with tilt angle and to appraise a tilt series

The foregoing numbers represent averages over all viewing directions and resolution should be higher in-plane because of the missing wedge effect. The resulting anisotropy has been considered on theoretical grounds (e.g., Gruenwald et al., 2003b; Unger et al., 1999): however, in practice, resolution may fall off even faster at high tilt-angles, on account of greater projected thickness (hence, multiple scattering) and focal gradients. NLOO-2D allows one to measure resolution empirically as a function of tilt angle. This dependence is plotted in Fig. 6A for the whole cryo-tomogram containing HSV-1 virions described above (results given in Fig. 5A). The best resolution, at low tilt angles, is 70–72 Å, but falls off more rapidly at tilt-angles above 50° than geometrical considerations would anticipate (Gruenwald et al., 2003b). Interestingly, there were several projections towards the end of the series, at 50°–55°, that were noticeably poorer. In retrospect, this represented an inconsistency in focusing that was corrected again by the time the tilt series reached 58°–60°. Such curves allow one to check for radiation damage towards the end of a cryogenic tilt series (high positive tilt angles, in this case). Thus, resolution is slightly but systematically worse around +40° than it is around −40° which probably represents incipient radiation damage, although we would expect this effect to be more apparent in studies that extend to higher resolutions (Conway et al., 1993).
We also used NLOO-2D to evaluate the resolution for a single virion near the center of the field (Fig. 6B, corresponding to results in Fig. 5B). Here the best resolution, at low tilt-angles, was \( \sim 65 \, \text{Å} \), indicating that the figure of \( 72 \, \text{Å} \) for the whole tomogram was compromised by averaging in lower resolution data from peripheral regions. The poorer resolution of the 50°–55° projections is again apparent.

Finally, we measured resolution by NLOO-2D for a single capsid (Fig. 6C, corresponding to results given in Fig. 5C). Resolution varies from \( \sim 55 \, \text{Å} \) for near-zero tilts to \( \sim 100 \, \text{Å} \) for the highest tilts (cf. 68 Å for the overall resolution from NLOO-3D). Thus, the best resolution, in-plane and near the center of the field, is 55 Å, and the deterioration in the perpendicular dimension is more pronounced than previously estimated, which was calculated to be \( \sim 90 \, \text{Å} \) at 90° (Grünewald et al., 2003b).

4. Discussion

In this paper, we address the issue of resolution measurement in the context of electron tomography and derive two criteria, NLOO and FSC\(_{\text{el/lo}}\), based on principles of cross-validation. These criteria were evaluated in model experiments and with real cryo-ET data, whereby the resolutions that they gave were compared with the results of reference-FSC calculations. In the intuitively appealing reference-FSC approach, the calculated structure (or a component thereof) is correlated with a known reference structure. The new criteria observed close consistency with FSC\(_{\text{ref}}\).

4.1. Advantages and limitations of NLOO-3D and FSC\(_{\text{el/lo}}\)

In calculating FSC\(_{\text{el/lo}}\), the projections are partitioned into odd and even half-sets. The formula (5) used in defining this measure is quite well known in SPA (Gri gorieff, 2000; Rosenthal and Henderson, 2003). NLOO-3D represents a further step in the same direction. Here, the underlying idea is to estimate the quality of a tomogram in terms of its ability to predict a missing projection. In theory, this measure provides a good approximation to the ideal but unattainable resolution of an FSC calculated between two tomograms derived from full but independent, statistically equivalent, tilt series of the same volume (Appendix B). A priori, we expected NLOO-3D to give more accurate, and indeed higher, resolutions than FSC\(_{\text{el/lo}}\) because more information goes into the tomograms that are evaluated. Such turned out to be the case in the validation experiments (see Section 3), although the two measures are quite consistent if the resolution is lower than \( r_{\text{crit}} \). In such a situation, which reflects relatively high noise levels (from any of several sources), the sparser sampling of Fourier space in FSC\(_{\text{el/lo}}\) is of minimal importance. Finally, appraisal resolution over a tilt series of projections by NLOO-2D allows for a posteriori quality control of the data, e.g., screening out seriously deficient projections, in addition to providing an assessment of the variation of resolution with viewing angle. This measure also provides an estimate of the tomogram resolution in the most favorable situation, which will usually be in-plane (near-zero tilt-angles).

4.2. Computational considerations in practice

FSC\(_{\text{el/lo}}\) is less demanding than NLOO-3D in terms of computational cost, since it only requires the calculation of two tomograms. Accordingly, it may be economical in practice to use FSC\(_{\text{el/lo}}\) first. If the resulting resolution is lower than \( r_{\text{crit}} \), it is likely that NLOO-3D will yield a very similar number, and little is to be gained by making the lengthier NLOO-3D calculation. Nevertheless, a short-cut can be achieved in the NLOO-3D calculation by overall reducing the number of projections analyzed or by only reducing the number of tomograms to be generated. In the first case, the estimate is restricted to comparisons between every 2nd (3rd, 4th...) projection and the corresponding reprojection from the tomogram calculated without that projection. In the second case a reduced number of tomograms is generated by excluding from each of them multiple suitably spaced projection, instead of only one, and then all the reprojections corresponding to the excluded tilt angles are calculated. In both cases, we expect a little loss in accuracy while obtaining a computational speed-up that is proportional to the reduction factor.

NLOO-2D may be used to obtain information about the anisotropy of resolution, which is an important consideration for thick specimens and likely to be underestimated by simply geometrical considerations.

There are reasons (see above) to suppose that resolution may vary somewhat within a tomographic volume, although the uniformity of resolutions that we obtained for different HSV capsids suggests that such variations may be encountered mainly around the edges (we note that this result tends to validate the FSC\(_{\text{ref}}\) approach). However, if there should be a region that is particularly important to the investigation in hand, it may be worthwhile to calculate the resolution of this subvolume. ELECTRA is configured to effect calculations of all these kinds.

4.3. Other measures of tomographic resolution

To our knowledge, only a few approaches to resolution assessment have been proposed in electron tomography. A leave-one-out technique (not so named) was one of several criteria tried by Taylor et al. (1997). Specifically, the resolution achieved with a tomographic
reconstruction method suited for paracrystalline objects was assessed by three criteria applied in the projection domain (FSC, DPR, and unweighted phase residuals). The reference image for each input projection was obtained from a tomogram calculated from the whole tilt series, as well as from one calculated omitting the projection in question (i.e., leave-one-out). Non-exclusion greatly affects the outcome of this calculation (above, and unpublished results), as was reflected in the results obtained.

In this study, we have shown that an appropriate combination of these quantities (Eqs. (7) and (8)) gives a measure that is more consistent with those used in SPA. More recently, application of the spectral signal to noise ratio (Unser et al., 1987) to tomographic reconstructions has been investigated in separate ways by Penczek (2002) and Unser et al. (2005). Penczek (2002) derived and tested a 3D formulation, but only for a particular class of Fourier space interpolation-based reconstruction methods. Unser et al. (2005) have proposed an interesting generalization of this approach but it has yet to be applied in practice.

4.4. Local variations in resolution

To complement measurements of the variation of resolution with tilt-angle (see Section 3.3), we can also define a measure describing the 3D distribution of NLOO, in the Fourier domain as well as in real space. In the Fourier domain it would be similar to what has been proposed for the SSNR (Penczek, 2002; Unser et al., 2005), and would be useful for assessing the distribution of the signal and to design anisotropic low-pass filters optimized for the tomogram. In our case, this measure would be obtained by omitting the summation over $R(k)$ in Eq. (8), then evaluating this modified expression of NLOO-3D for each frequency voxel. An adequate mapping of the frequency 2D coordinates of the single projections to the three-dimensional Fourier domain would be required, which could be derived via the Central Section Theorem. However, this kind of measure is expected to give noisy results, because of the low number of projections involved in the estimation, thus requiring a proper regularization. As an alternative, volumetric variations of the resolution in the real space domain may be assessed by calculating the resolutions of sizeable subvolumes that, once combined, cover all the tomogram, or just selected parts of it (e.g., at the center or the periphery of the field).

4.5. Density of angular sampling: implications

The resolution of a tomogram is ultimately limited by the angular increment between projections. Interestingly, results obtained in the model experiment in Section 3.1 showed that tomographic resolution, as assessed by NLOO-3D and FSC_ref, may exceed the limit, $r_0$, given by Eq. (3). To explain this discrepancy, we note that, as derived (Crowther et al., 1970), the relationship was not as a strict equality, but was based on an approximation—that the Bessel function $J_n(x)$ should be $\sim 0$ for $0 < x < n - 2$. This consideration placed a limit on the highest order, $N$, of Bessel function that could be solved at a given radial frequency limit for a particle of specified diameter. In fact, higher order Bessel functions have small but non-zero values that intrude into this domain. We infer that their contributions, in effect, underlie the detection of signal to the observed resolution limits. We note that degree of consistency between $r_0$ and other measures should also depend, to some extent, on the threshold adopted for the latter measures. Notwithstanding the observed discrepancy, this relationship Eq. (2) provides a useful rule-of-thumb for resolution as constrained by the size of the angular increment.

4.6. Application to dual-tilt series-based tomograms

Currently, there is growing interest in the use of dual-axis tilt series (Penczek et al., 1995) to reduce the missing wedge to a missing pyramid. In this context, it is desirable to have a method for quantifying the improvement in resolution that is achieved. The final tomogram is produced by calculating and then merging two tomograms from the separate tilt series. Before merging, the two tomograms are processed by applying a transformation to align them and to compensate for potential radiation-induced distortions of the volumes between the recording of the two tilt series (Mastronarde, 1997). Here, the NLOO approach is readily applicable, provided that the experimental projections are compared with reprojections in the same frame of reference, i.e., without any relative transformation applied. In principle, the two sub-tomograms could be compared in an FSC calculation although this approach is likely to suggest a lower resolution than would be obtained with a projection partition that follows the order of recording because, particularly with cryo-specimens, the second tilt series depicts a specimen that has been subjected to more radiation.

4.7. Resolution in tomograms of plastic sections

Although progress is being made with vitrified sections (Al-Amoudi et al., 2004; Hsieh et al., 2002; Zhang et al., 2004), it is likely that most cellular tomography will continue, for some time to come, to be done with plastic sections of freeze-substituted or conventionally embedded material. Such tomograms may also be evaluated by the present criteria. In preliminary estimates for tomograms of a $\sim 110$ nm plastic section of a virus-infected cell (G.C. et al., work in progress), NLOO-3D
gave an overall resolution of 55 Å, whereas analysis by NLOO-2D gave 40 Å in-plane, increasing to 140 Å at 60°. With respect to the notably high in-plane resolution, we note that this figure refers to internal consistency of the representation given by the stained plastic section. If it were possible to calculate a FSC ref between such a tomogram and the corresponding native hydrated structure, it would surely yield a considerably lower resolution. The notably low resolution at 60° represents another example of its falling off faster than geometrical considerations would predict.

**Acknowledgments**

We thank David Belnap and Bernard Heymann for fruitful discussions, James Conway for useful suggestions and making available the HBV density map, Benes Trus for making available the HSV1 capsid density map, and Eric Freed, Vlad Speransky, Dennis Winkler for collaboration on the tomography project on HIV budding.

**Appendix A. Derivation of the NLOO-2D formula**

**A.1. Relationship between Fourier ring correlation and normalized square difference**

Given two images \( F^{(i)} \) and \( G^{(i)} \) in the Fourier domain, we define the normalized square difference (NSD) for a region \( R(k) \) centered around radial frequency \( k \), as

\[
\text{NSD}_{\text{NSD}}^{(i)}(k) = \frac{\sum_{m,n \in R(k)} \left| F_{m,n}^{(i)} - G_{m,n}^{(i)} \right|^2}{\sum_{m,n \in R(k)} \left| F_{m,n}^{(i)} \right|^2 + \sum_{m,n \in R(k)} \left| G_{m,n}^{(i)} \right|^2}. \tag{A.1}
\]

where the frequency components of the images are designated by the index pair \((m,n)\).

The NSD can assume values between 0 and 2, which is evident after manipulating the expression \( \text{(A.1)} \), and rewriting it as

\[
\text{NSD}_{\text{NSD}}^{(i)}(k) = 1 - \frac{2\sum_{m,n \in R(k)} \text{Re} \left\{ F_{m,n}^{(i)} G_{m,n}^{(i)*} \right\}}{\sum_{m,n \in R(k)} \left| F_{m,n}^{(i)} \right|^2 + \sum_{m,n \in R(k)} \left| G_{m,n}^{(i)} \right|^2}. \tag{A.2}
\]

If we normalize the images in the region \( R(k) \), to obtain \( \hat{F}^{(i)} \) and \( \hat{G}^{(i)} \) such that

\[
\hat{F}_{m,n}^{(i)} = \frac{F_{m,n}^{(i)}}{\sqrt{\sum_{m,n \in R(k)} \left| F_{m,n}^{(i)} \right|^2}} \quad \text{and} \quad \hat{G}_{m,n}^{(i)} = \frac{G_{m,n}^{(i)}}{\sqrt{\sum_{m,n \in R(k)} \left| G_{m,n}^{(i)} \right|^2}}, \tag{A.3}
\]

the NSD between these two quantities is

\[
\text{NSD}_{\text{NSD}}^{(i)}(k) = 1 - \frac{2\sum_{m,n \in R(k)} \text{Re} \left\{ \hat{F}_{m,n}^{(i)} \hat{G}_{m,n}^{(i)*} \right\}}{\sum_{m,n \in R(k)} \left| \hat{F}_{m,n}^{(i)} \right|^2 + \sum_{m,n \in R(k)} \left| \hat{G}_{m,n}^{(i)} \right|^2}. \tag{A.4}
\]

Since both terms in the denominator are equal to 1, NSD and FRC are related by

\[
\text{NSD}_{\text{FRC}}^{(i)}(k) = \frac{\sum_{m,n \in R(k)} \left| \hat{F}_{m,n}^{(i)} \right|^2 + \sum_{m,n \in R(k)} \left| \hat{G}_{m,n}^{(i)} \right|^2}{\sum_{m,n \in R(k)} \left| \hat{F}_{m,n}^{(i)} \right|^2 + \sum_{m,n \in R(k)} \left| \hat{G}_{m,n}^{(i)} \right|^2} = 1 - \text{FRC}_{\text{FRC}}^{(i)}(k). \tag{A.5}
\]

**A.2. Derivation of the noise-compensated leave-one-out from the normalized square difference formulation**

The NSD Eq. \( \text{(A.1)} \) between the input projection \( X^{(i)} \) and the reprojection \( \hat{X}^{(i)} \) is affected by the different noise statistics of the two images. Furthermore, the input projection may contain some additional signal that is unwanted, since it is not present in the reprojection. We assume that a proper noise term \( N^{(i)} \) can take into account these two effects.

We define the noise-compensated normalized square difference (NNSD) as

\[
\text{NNSD}_{\text{NSD}}^{(i)}(X^{(i)} - \hat{X}^{(i)})(k) = \frac{\sum_{m,n \in R(k)} \left| X_{m,n}^{(i)} - \hat{X}_{m,n}^{(i)} \right|^2 - \sum_{m,n \in R(k)} \left| N_{m,n}^{(i)} \right|^2}{\sum_{m,n \in R(k)} \left| X_{m,n}^{(i)} \right|^2 + \sum_{m,n \in R(k)} \left| \hat{X}_{m,n}^{(i)} \right|^2 - \sum_{m,n \in R(k)} \left| N_{m,n}^{(i)} \right|^2}, \tag{A.5}
\]

where the noise term is introduced both at numerator and denominator. The object of this modified formulation is to provide a new normalized measure by removing from the squared difference any variance due either to the differing noise statistics or to additional signals that may be present in the input projections. We then approximate the frequency components of the noise as \( \left| N_{m,n}^{(i)} \right| \approx \left| X_{m,n}^{(i)} - \hat{X}_{m,n}^{(i)} \right| \). Assuming that the discrepancy between an input projection and the corresponding reprojection from the tomogram generated from all the input projections is always less than the discrepancy between the same input projection and the reprojection from a tomogram reconstructed without that projection, this choice for \( N_{m,n}^{(i)} \) guarantees that the NNSD can not become negative. It follows that
The evaluation of the NNSD on the projections/reprojections, after they are normalized in the region \( R(k) \), gives

\[
\text{NNSD}^{(g)}_{\hat{X}_m,\hat{X}_n}(k) = 1 - \frac{2 \sum_{m,n \in R(k)} |\text{Re}\{\hat{X}_m^{(g)} \hat{X}_n^{(-g)}\}|^2}{\sum_{m,n \in R(k)} |\hat{X}_m^{(g)}|^2 - 2 \sum_{m,n \in R(k)} |\hat{X}_m^{(g)}|^2 + 2 \sum_{m,n \in R(k)} |\text{Re}\{\hat{X}_m^{(g)} \hat{X}_n^{(-g)}\}|^2}.
\]  

(A.6)

By analogy with (A.4), the ratio between the two Fourier ring correlations in (A.7) represents a correlation measure, which we define as 2D noise-compensated leave-one-out resolution estimate.

**Appendix B. An approximated analysis of the NLOO-2D formula**

In the following we demonstrate how the NLOO-2D approximates the ideal experiment (see Section 2.2) in assessing the resolution. Precisely we show that the NLOO-2D gives a relationship near to the equivalent of Eq. (2) for the FRC. This result can then be extended to the NLOO-3D by exploiting the Fourier Slice Theorem.

We consider the case when the reconstruction is performed by nearest-neighbor interpolation in the Fourier domain, an approach already followed in (Penczek, 2002) with regard to the SSNR. Furthermore, in the analysis we do not consider the presence in the input projections of signal components that can instead be missing in the corresponding reprojections.

We assume that each input projection \( X^{(g)}_m \) is corrupted by additive random noise, uncorrelated between the projections, such that each frequency component is

\[
X^{(g)}_{m,n} = S_{m,n} + N^{(g)}_{m,n},
\]  

(B.1)

where \( S_{m,n} \) is the signal to reconstruct, while \( N^{(g)}_{m,n} \) is a zero-mean Gaussian noise component. The signal \( S_{m,n} \) describes a component in a voxel of the three-dimensional Fourier space that is uniquely defined by the values of \( i, m, \) and \( n \). Since we are focused only on the signal in a defined slice, in the following we omit the index \( i \).

Assuming that the noise is independent both between frequencies and projections, we have

\[
E \left[ N^{(g)}_{m,n} N^{(-g)}_{p,q} \right] = \delta_{m,p} \delta_{n,q} \delta_{i,j} \sigma^2_{m,n},
\]  

(B.2)

where \( E[ \cdot ] \) indicates the expectation value, \( \delta_{i,j} \) is the Kronecker delta, and \( \sigma^2_{m,n} \) is the variance of the noise component, which is assumed to be the same for each projection.

Each reprojection \( \hat{X}^{(i)}_{m,n} \) from the reconstructed volume similarly represents, in the Fourier domain, a slice from the reconstructed object. When the reconstruction is performed by nearest-neighbor interpolation, each component \( \hat{X}^{(i)}_{m,n} \) is obtained as an average from \( n_{m,n} \) adjacent projections

\[
\hat{X}^{(i)}_{m,n} = \frac{1}{n_{m,n}} \sum_{j=1}^{n_{m,n}} X^{(j)}_{m,n}.
\]  

(B.3)

In this formulation the index \( j \) in the sum does not indicate the position of the input projection in the tilt series, but refers to an arbitrary sequence which includes only the input projections that contain the signal \( S_{m,n} \) to be reconstructed. Moreover, we assume that in the new sequence the index \( i \) corresponds to \( j' \). We remark that the number \( n_{m,n} \) of projections involved is different for each spatial frequency, and it typically decreases with higher frequency values.

Under these assumptions, from (B.1) and (B.3) it follows that the signal-to-noise ratio of the \( i \)-th slice of the reconstruction, measured for a radial frequency \( k \), is

\[
\text{SNR}^{(g)}(k) = \frac{\sum_{m,n \in R(k)} |S_{m,n}|^2}{\sum_{m,n \in R(k)} \delta_{m,n} \sigma^2_{m,n}}.
\]  

(B.4)

Moreover, the frequency components of the reprojection \( \hat{X}^{(-j)}_{m,n} \), since obtained by excluding the corresponding input projection from the reconstruction process, are given by

\[
\hat{X}^{(-j)}_{m,n} = \frac{1}{n_{m,n} - 1} \sum_{j=1, \neq j'}^{n_{m,n}} X^{(j)}_{m,n}.
\]  

(B.5)

Under these conditions, we can then evaluate the expectation value of the NLOO-2D as

\[
E[\text{NLOO-2D}^{(g)}(k)] \simeq \frac{E \left[ \text{FRC}^{(g)}_{\hat{X}^{(-j)}_{m,n}}(k) \right]}{E \left[ \text{FRC}^{(g)}_{\hat{X}^{(i)}_{m,n}}(k) \right]}.
\]  

(B.6)

By making use of (1), (B.1), (B.2), and (B.5), the expectation value of the FRC at numerator is

\[
E \left[ \text{FRC}^{(g)}_{\hat{X}^{(-j)}_{m,n}}(k) \right] \simeq \frac{\sum_{m,n \in R(k)} |S_{m,n}|^2}{\sqrt{\sum_{m,n \in R(k)} |S_{m,n}|^2 + \sum_{m,n \in R(k)} \delta_{m,n} \sigma^2_{m,n} + \sum_{m,n \in R(k)} \delta_{m,n} \sigma^2_{m,n-1}}}
\]  

(B.7)

The expectation value of the FRC at the denominator is obtained in a similar way, just using (B.3) instead of (B.5).
\[ E \left[ \text{FRC}^{(i)}_{\text{XX}}(k) \right] \]
\[ \approx \frac{\sum_{m,n\in R(k)} |S_{m,n}|^2}{\sqrt{\sum_{m,n\in R(k)} |S_{m,n}|^2 + \sum_{m,n\in R(k)} \frac{\sigma_{m,n}^2}{\sigma_{m,n}^2}}} \]
\[ (B.8) \]
Substitution of (B.7) and (B.8) in (B.6) gives
\[ E \left[ \text{NLOO-2D}^{(i)}(k) \right] \]
\[ \approx \frac{\sum_{m,n\in R(k)} |S_{m,n}|^2}{\sqrt{\sum_{m,n\in R(k)} |S_{m,n}|^2 + \sum_{m,n\in R(k)} \frac{\sigma_{m,n}^2}{\sigma_{m,n}^2}}} \]
\[ (B.9) \]
which, according to (B.4), can also be written as
\[ E \left[ \text{NLOO-2D}^{(i)}(k) \right] \]
\[ \approx \frac{\text{SNR}^{(i)}(k)}{\text{SNR}^{(i)}(k) + 1} \left( \sum_{m,n\in R(k)} |S_{m,n}|^2 + \sum_{m,n\in R(k)} \frac{\sigma_{m,n}^2}{\sigma_{m,n}^2} \right)^{1/2} \]
\[ (B.10) \]
We observe that the first ratio term on the right side of (B.10) gives the required relationship between correlation measure and signal-to-noise ratio. The second ratio term under square root represents the underestimation factor, which is generally close to one.

References