BMC Bioinformatics

METHODOLOGY

Open Access

Compressed sensing improved iterative reconstruction-reprojection algorithm for electron tomography

Lun Li[1,2], Renmin Han[3*], Zhaotian Zhang[4], Tiande Guo[2], Zhiyong Liu[1] and Fa Zhang[1*]

*Correspondence:
hanrenmin@gmail.com;
zhangfa@ict.ac.cn
[1]High Performance Computer
Research Center, Institute Of
Computing Technology, Chinese
Academy of Sciences, 100190
Beijing, China
[3]Research Center for Mathematics
and Interdisciplinary Sciences,
Shandong University, 266237
Qingdao, People's Republic of China
Full list of author information is
available at the end of the article

Abstract

Background: Electron tomography (ET) is an important technique for the study of complex biological structures and their functions. Electron tomography reconstructs the interior of a three-dimensional object from its projections at different orientations. However, due to the instrument limitation, the angular tilt range of the projections is limited within $+70°$ to $-70°$. The missing angle range is known as the missing wedge and will cause artifacts.

Results: In this paper, we proposed a novel algorithm, compressed sensing improved iterative reconstruction-reprojection (CSIIRR), which follows the schedule of improved iterative reconstruction-reprojection but further considers the sparsity of the biological ultra-structural content in specimen. The proposed algorithm keeps both the merits of the improved iterative reconstruction-reprojection (IIRR) and compressed sensing, resulting in an estimation of the electron tomography with faster execution speed and better reconstruction result. A comprehensive experiment has been carried out, in which CSIIRR was challenged on both simulated and real-world datasets as well as compared with a number of classical methods. The experimental results prove the effectiveness and efficiency of CSIIRR, and further show its advantages over the other methods.

Conclusions: The proposed algorithm has an obvious advance in the suppression of missing wedge effects and the restoration of missing information, which provides an option to the structural biologist for clear and accurate tomographic reconstruction.

Keywords: Electron tomography, Compressed sensing, Matching pursuit, Iterative reconstruction-reprojection

BMC

Li *et al. BMC Bioinformatics* 2020, **21**(Suppl 6):202

Page 2 of 19

Background

Electron tomography (ET) is an important technique for the study of complex biological structures and their functions, which reconstructs the interior of a three-dimensional object from its projections at different orientations [1]. However, due to the instrument limitation, the angular tilt range of the projections is limited within $+70°$ to $-70°$. Such a series of the projections compose an angle-limited projection dataset, in which the unsampled high tilt angles are called the missing wedge and will cause artifacts. Therefore, the reconstruction of the tomogram in ET is an optimization problem with incomplete information. Additionally, the optimization problem becomes more difficult due to the extreme noise caused by the low electron doses in the collection of Electron Microscopy (EM) images.

Traditional 3D reconstruction methods, such as Weighted back-projection (WBP) and Filtered back-projection (FBP), suffer from the incomplete data and extremely noisy projections, resulting in the undesired artifacts. Many algorithms have been proposed to deal with the incompleteness caused by the limited projections and strong noise. By combining the algebra reconstruction technique (ART) and the nonlinear diffusion (ND) filter technique, filtered iterative reconstruction technique (FIRT) [2] could partially restore some information at the non-sampled angular region. Iterative reconstruction-reprojection (IRR) [3–5] and its variations, such as discrete IRR algorithm and the finite impulse response (FIR) lowpass filter (DIRRLF) [6], iterative reconstruction-reprojection (IRR) algorithm with total variation (TV) constraint (IRR-TV) [7], iterative reconstruction-reprojection in projection space (IRRP) algorithm [8] and algebraic iterative reconstruction-reprojection (AIRR) [9], estimate the missing projects by an iterative optimization procedure between the projection and image domain.

Improved iterative reconstruction-reprojection (IIRR) [10, 11] is one of the variations of the IRR algorithm, whose convergence under extremely noisy condition has been proved [12]. However, without prior information, IIRR is not efficient in suppressing the effect of the missing wedge and restoring the missing information for ET data, because of the low correlation between known projections and unknown projections in the missing wedge. On the other hand, other algorithms reduce artifacts by introducing prior information. Iterative Compressed-sensing Optimized Non-uniform fast Fourier transform reconstruction (ICON) [13–15] combines compressed sensing (CS) with non-uniform fast Fourier transform (NUFFT) in order to restore the missing information. There are two main steps in ICON. One is the fidelity preservation step which minimizes the difference between the reconstructed volume and the projections. Another one is prior sparsity restriction step which increases the sparsity of the reconstructed volume. Nonetheless, a drawback of ICON is its slow computational speed.

In this paper, we proposed a novel algorithm which follows the schedule of IIRR but further considers the sparsity of the biological ultra-structural content in specimen, resulting in an estimation of the electron tomography with faster execution speed and better reconstruction result. A comprehensive experiment on both the simulated and real-world dataset is carried out to estimate the performance of our algorithm and compare the new algorithm with other state-of-the-art methods. The experimental results indicate that the proposed algorithm has the ability to suppress the effect of missing wedge and restore the missing information within the missing wedge. Specifically, compared with the classical methods, our algorithm can achieve better reconstruction with faster convergence.

Li *et al. BMC Bioinformatics* 2020, **21**(Suppl 6):202

Page 3 of 19

Preliminaries

Without loss of generality, our following discussion is defined in two-dimensional space. Nevertheless, the algorithm we proposed can be extended to three-dimensional space easily.

From filtered backprojection (FBP) to iterative reconstruction-reprojection (IRR)

The filtered back-projection (FBP) algorithm is one of the oldest algorithms in computing tomography. Its mathematical model could be concluded as following:

Given a function $f(x)$ ($x \in \mathbb{R}^2$), the projections of function f is defined as a linear integral along a straight line $L = \{x \in \mathbb{R}^2 | (\cos \alpha) x_1 + (\sin \alpha) x_2 = s\}$

$$
\begin{aligned}
p(s, \alpha) &= \mathscr{R}f \\
&= \int_L f(x)\, dx \\
&= \int_0^\pi f(t \sin \alpha + s \cos \alpha, -t \cos \alpha + s \sin \alpha)\, dt,
\end{aligned}
\tag{1}
$$

where $\mathscr{R}$ is the Radon transform operator, $s \in \mathbb{R}$ is the distance of line L from the origin, α is the angle between normal vector to L and the x_1 axis. Suppose the value of p on $\Omega \subset \{(s, \alpha) | \alpha \in (0, \pi], s \in \mathbb{R}\}$ is known in prior. Let $\chi_\Omega(s, \alpha) = \begin{cases} 1, (s, \alpha) \in \Omega \\ 0, (s, \alpha) \notin \Omega \end{cases}$ be the characteristic function of set Ω, $p_{unknown}(s, \alpha) = (1 - \chi_\Omega(s, \alpha))p(s, \alpha)$ be the unknown projections and $p_{known}(s, \alpha) = \chi_\Omega(s, \alpha)p(s, \alpha)$ be the known projections. FBP will first apply a filter on the projections and then backproject these filtered projections to get an estimation of the original signal f:

1) Convolution (Filtering)

$$
\begin{aligned}
p_c(s, \alpha) &= p_{known}(s, \alpha) * \mathscr{F}_1^{-1}|\omega| \\
&= \int p_{known}(t, \alpha) \cdot (\mathscr{F}_1^{-1}|\omega|)(s - t)dt
\end{aligned}
\tag{2}
$$

where * denotes convolution, $\mathscr{F}_1^{-1}$ denotes the one dimensional Fourier transform.

2) Backprojection

$$
\begin{aligned}
\widetilde{f}(x_1, x_2) &= \mathscr{B}p_c \\
&= \int p_c(x_1 \cos \alpha + x_2 \sin \alpha, \alpha)d\alpha,
\end{aligned}
\tag{3}
$$

where $\widetilde{f}$ is the estimation of f, x_i is the ith element of x, $\mathscr{B}$ is the backprojection operator. It should be noted that, in FBP, the unknown projections $p_{unknown}$ have been set to 0.

Meanwhile, the iterative reconstruction-reprojection algorithm (IRR) is an iterative reconstruction method which is able to suppress the effect of insufficient tilt angles by estimating the unknown projections. At the beginning, all the unknown projections in the system $p_{unknown}^1$ are set to 0 as in FBP. During the computational process of IRR, for the ith iteration, the ith estimation of f will be reconstructed from both the known projections p_{known} and the ith estimated unknown projections $p_{unknown}^i$:

$$
f^i(x_1, x_2) = \mathscr{B}((p_{known} + p_{unknown}^i) * \mathscr{F}_1^{-1}|\omega|).
\tag{4}
$$

Li *et al. BMC Bioinformatics* 2020, **21**(Suppl 6):202

Page 4 of 19

However, indeed, we do not know the unknown projections $p^i_{unknown}$. In IRR, the (i)th unknown projections are estimated from f in the previous iteration:

$$p^i_{unknown}(s, \alpha) = (1 - \chi_\Omega)\mathscr{R}f^{i-1}. \tag{5}$$

By estimating the unknown projections rather than setting them to 0, IRR suffers less from the 'missing wedge' problem than FBP.

The improved algorithms based on IRR

Improved iterative reconstruction-reprojection (IIRR)

A drawback of IRR is that the convergence of IRR under extremely noisy condition is not guaranteed. By introducing a relaxed parameter λ into IRR, the convergence under extremely noisy condition has been proved. More precisely, IIRR changes Eq. (4) into

$$f^i(x_1, x_2) = \mathscr{B}((p_{known} + \lambda p^i_{unknown}) * \mathscr{F}_1^{-1}|\omega|), \tag{6}$$

where $0 < \lambda < 1$ is a relaxed parameter. The smaller value of λ considers more in the known projections while larger value of λ considers more about the unknown projections. IIRR can restrain the artifact caused by the incomplete data in Computerized Tomography (CT).

Iterative reconstruction-reprojection with total variation (IRR-TV)

IRR-TV is an reconstruction method which is developed for few-view projections, in which the angular tilt range of the projections is not limited but the number of projections is limited. IRR-TV poses the reconstruction problem as an optimization problem with the form as:

$$\begin{aligned} &\min_{f} \ TV(f) \\ &s.t. \ Af = b, \end{aligned} \tag{7}$$

where $Af = b$ is the fidelity term,

$$TV(f) = \sum_{(x_1, x_2, x_3) \in \mathbb{Z}^3} ((f_{x_1, x_2, x_3} - f_{x_1-1, x_2, x_3})^2 + (f_{x_1, x_2, x_3} - f_{x_1, x_2-1, x_3})^2 + (f_{x_1, x_2, x_3} - f_{x_1, x_2, x_3-1})^2)^{\frac{1}{2}} \tag{8}$$

is the total variation norm, f is the unknown 3D reconstruction volume, $(x_1, x_2, x_3) \in \mathbb{Z}^3$ is the coordinate of voxels in unknown volume, A is the projection matrix, and b is the observations or projections. By introducing TV into IRR, IRR-TV achieves an accelerated convergence rate compared with the classic IRR algorithm.

Specifically, in each iteration before the estimation of unknown projections, IRR-TV minimizes the TV norm by deepest gradient descent:

$$f^i = f^i - d \cdot \frac{\partial TV(f^i)}{\partial f^i}, \tag{9}$$

where d is step size, f^i is the ith estimation of f, and $TV(\cdot)$ is the TV norm.

Matching pursuit

Matching pursuit (MP) is an algorithm which represents a signal in a redundant basis. Let $\mathscr{D} = \{g_\gamma\}_{\gamma \in \Gamma}$ be a family of vectors in Hilbert space and f be a signal. f can be

Li *et al. BMC Bioinformatics* 2020, **21**(Suppl 6):202

Page 5 of 19

represented as a linear combination of N vectors selected from $\mathscr{D}$

$$f \approx \sum_{n=1}^{N} a_{\gamma_n} g_{\gamma_n}, \tag{10}$$

where a_{γ_n} is the weight of g_{γ_n}, g_{γ_n} is the vectors selected from $\mathscr{D}$, N is the total number of vectors selected from $\mathscr{D}$. The basic idea of MP is choosing vectors from $\mathscr{D}$ greedily. More precisely, let $R_1 = f$ be the initial approximation error. For $i = 1, 2, 3, ...$, MP finds g_{γ_i} with maximum absolute value of inner product $| < R_i, g_{\gamma_i} > |$, where $R_i = f - \sum_{n=1}^{i} a_{\gamma_n} g_{\gamma_n}$ is the ith approximation error. Then we set $a_i = < R_i, g_{\gamma_i} >$. The process continues until the approximation error is small enough.

Methods

Problem definition

Despite the success of IIRR in Computed Tomography (CT), IIRR is not efficient in suppressing the effect of the missing wedge and restore the missing information for ET data. Here, an algorithm based on the schedule of IIRR with the consideration of sparse information is proposed to achieve better ET reconstruction.

According to the previous study [13], the biological ultra-structural content is relatively sparse with respect to the surrounding solvent. Such sparsity is used as the prior information in our proposed method. Formally, let $f(x)$ be the space function ($x \in \mathbb{R}^2$) representing the sample and suppose that f is supported on set T which is a subset of $\mathbb{Z}^2$. The consequent optimization problem could be defined as

$$\begin{aligned} &\min_{f} \ \|f\|_0 \\ &s.t. \ \ \chi_\Omega \mathscr{R} f = p_{known}, \end{aligned} \tag{11}$$

where $\|f\|_0 = \#\{x \in \mathbb{Z}^2 \cap T | f(x) \neq 0\}$ is the l_0 norm of f, x is a 2 dimensional vector representing the coordinate of pixels in picture f, $\mathscr{R}$ is the Radon transform, p_{known} is the known projections, χ_Ω is the characteristic function of set Ω, and Ω is the subset of projection space corresponding to the known projections.

Compressed sensing improved iterative reconstruction-reprojection (CSIIRR)

We designed a new algorithm to solve the optimization problem (11) by combining IIRR and CS. The proposed algorithm is elaborated as follows.

- Step 0: Preprocess
- Step 1: Estimate the unknown projections
- Step 2: Reconstruct via modified matching pursuit (MMP), return to step 1

Step 0: preprocess

Alignment is needed for the raw projections before reconstruction for ET data [16, 17]. Then we need an initial guess of specimen f which can be the reconstruction result via any reconstruction method or even a zero function. We denote the initial guess of f as f^0.

Step 1: estimate the unknown projections

Let f^{k-1} be the reconstruction result in the $(k-1)$th step, $k = 1, 2, ...$ We estimate the unknown projections by reprojecting f^{k-1} at the corresponding tilt angles. We denote the estimated projections as

$$p_{unknown}^{k-1} = (1 - \chi_\Omega) \mathscr{R} f^{k-1}, \tag{12}$$

Li *et al. BMC Bioinformatics* 2020, **21**(Suppl 6):202

Page 6 of 19

where the notation $\mathscr{R}$ denotes the Radon transform, Ω is the subset of projection space corresponding to the known projections, and χ_Ω denotes the characteristic function of Ω.

Step 2: reconstruct via modified MP (MMP)

In this step, we want to reconstruct specimen f^k from the known projections p_{known} and the estimated projections $p_{unknown}^{k-1}$. We seek for the sparsest f^k such that $\mathscr{R}f^{k-1} = p_{known} + \lambda p_{unknown}^{k-1}$, where $0 < \lambda < 1$ is a relaxed parameter. A small λ is suggested for extremely noisy data because the reliability of estimated unknown projections is low with high noise levels. So we introduced compressed sensing to this step by a modified matching pursuit method. The detailed process of our modified matching pursuit method is described in Algorithm 1.

Algorithm 1 Modified matching pursuit

Input: Known projections p_{known} and estimated projections $p_{unknown}^{k-1}$.

Output: Reconstruction result f^k

    function $\mathscr{B}_{mmp}(p_{known}, p_{unknown}^{k-1})$

        $R_0 \leftarrow p_{known} + \lambda p_{unknown}^{k-1}$

        $f_{mmp}^0 \leftarrow 0$

        $l \leftarrow 1$

        repeat

            Find the largest M absolute values of pixel value in $\chi_T \mathscr{B}(R_{l-1} * \mathscr{F}^{-1}|\omega|)$

            Restore the location of the M pixels in $x_1, x_2, ..., x_M$

$$f_{mmp}^l \leftarrow f_{mmp}^{l-1} + (\sum_{i=1}^{M} \chi_{x_i})\mathscr{B}(R_{l-1} * \mathscr{F}^{-1}|\omega|) \tag{13}$$

$$R_l \leftarrow p_{known} + \lambda p_{unknown}^{k-1} - \mathscr{R}f_{mmp}^l \tag{14}$$

            $l \leftarrow l + 1$

        until $\|R\| < \epsilon$

        $f^k \leftarrow f_{mmp}^{l-1}$

    end function

As shown in Algorithm 1, we first initialize the approximation of reconstruction result f_{mmp}^0 to zero vector, and the residual R_0 to $p_{known} + \lambda p_{unknown}^{k-1}$. In the l-th iteration, we want to find the largest M absolute values of pixel value in $\chi_T \mathscr{B}(R_{l-1} * \mathscr{F}^{-1}|\omega|)$ and restore the location of the M voxels in $x_1, x_2, ..., x_M$. The parameter M affects the sparsity of the reconstruction result. A small M is suggested if the biological ultra-structural content is sparse in the specimen. Next, we use the restored information to update the approximation f_{mmp}^l by Eq. (13) and the consequent residual R_l by Eq. (14). The process continues until $\|R\| < \epsilon$. After obtaining f^k, the algorithm returns to step 1 in order to optimize the estimated projections $p_{unknown}^k$.

The main advantage of combining CS with IIRR is to introduce sparsity into IIRR. There are two main differences between the Algorithm 1 and MP methods. First, Algorithm 1 estimates the tomogram not only from the known projections p_{known} but also from the estimated projections $p_{unknown}^{k-1}$, which makes the proposed algorithm more stable than

Li *et al. BMC Bioinformatics* 2020, **21**(Suppl 6):202

Page 7 of 19

MP. Second, instead of choosing only one atom, Algorithm 1 chooses multiple pixels in each iteration. By doing so, Algorithm 1 can achieve a quick convergence compared with the classic MP. Moreover, there are also some differences between IRR-TV and CSI-IRR. First, the two methods choose different regularization. IRR-TV minimizes the TV norm while CSIIRR minimizes the l_0 norm. Second, IRR-TV adds a 'deepest gradient descent' step after the back projection, which is time-consuming. Third, IRR-TV focuses on the few-view problem in CT while CSIIRR focuses on the angle-limited problem in ET. Especially, CSIIRR is able to handle extremely noisy data on which IRR-TV usually fails.

Results

The CSIIRR is challenged by both the simulated data and real-world data. Expect for CSIIRR, the filtered back-projection (FBP), IIRR, weighted back-projection (WBP), Iterative Compressed-sensing Optimized Non-uniform fast Fourier transform reconstruction (ICON) [13, 14] and total variation (TV) are also tested in our experiment as a comparison. (Here, the algorithms of FBP, IIRR and CSIIRR are implemented with MATLAB, the implementation of ICON is download from http://ear.ict.ac.cn and the implementation of WBP comes from the 'tilt' module in IMOD.) In the experiment, the relaxation factor λ is set to 0.99 for both IIRR and the CSIIRR.

Performance on simulated data

Overall evaluation

First, CSIIRR is challenged on the noise-free simulated data. The ribosome structure downloaded from EMDB-3489 [18] is used as the ground-truth, whose volume size of EMDB-3489 is 400^3 voxels. A set of the projections with the range of the tilt angle from $60°$ to $-60°$ is generated as the measured observations, with an angular increment of $1°$.

To reconstruct the tomogram, the number of iterations is set to 10 in IIRR. For the sake of fairness, the number of outer-loop iterations in our proposed algorithm is also set to 10. To make reconstruction results converge, the number of iterations for modified MP in the proposed algorithm is set to 50. To trade off the convergence rate of CSIIRR and the sparsity of the output, the number of atoms selected in each iteration (M in Algorithm 1) is set to 335103 in modified MP.

Figure 1 demonstrates the reconstruction results obtained by FBP, IIRR, CSIIRR, ICON and TV. As shown in the first row, the central slice of the reconstruction produced by CSIIRR suffers the least artifacts caused by the incompleteness of the sampling angle compared with the ones of other methods. The log-scaled power spectrums on the second row of Fig. 1 shows that CSIIRR has filled some missing data in the missing wedge. Furthermore, the phase difference of the tomograms reconstructed by different methods has been presented in the third row of Fig. 1 (The point-to-point phase difference [13] Δ of a and b is defined as $\Delta = |\frac{|a-b|}{\pi} - 1|$.), which shows that CSIIRR has restored some correct information in the missing wedge. Moreover, the histograms demonstrated in the fourth row of Fig. 1 further support the conclusion made from the log-scaled power spectrum, in which the histogram shows that the reconstruction result by CSIIRR has a much more similar voxel distribution with the ground-truth, compared with the ones obtained from other methods. Additionally, the histogram accurately presents the sparseness of the reconstruction generated by CSIIRR.

Li *et al. BMC Bioinformatics* 2020, **21**(Suppl 6):202

Page 8 of 19

Fig. 1 Missing wedge analysis. Comparison of central slices from the ground truth and the tomograms reconstructed by FBP, IIRR, CSIIRR, ICON, TV, of the log-scaled power spectrum, phase difference, and histogram
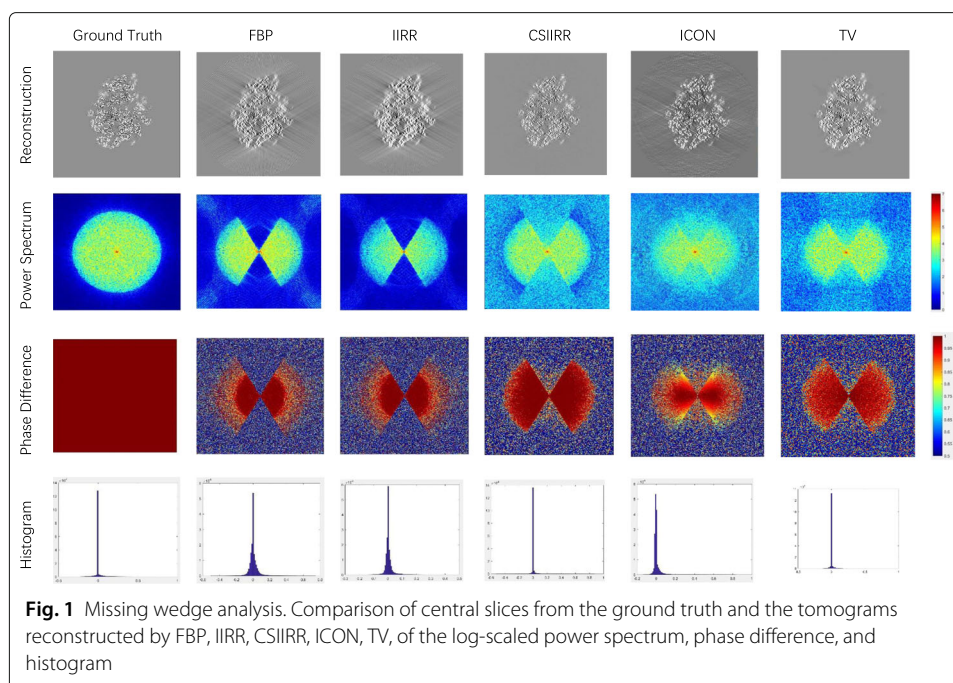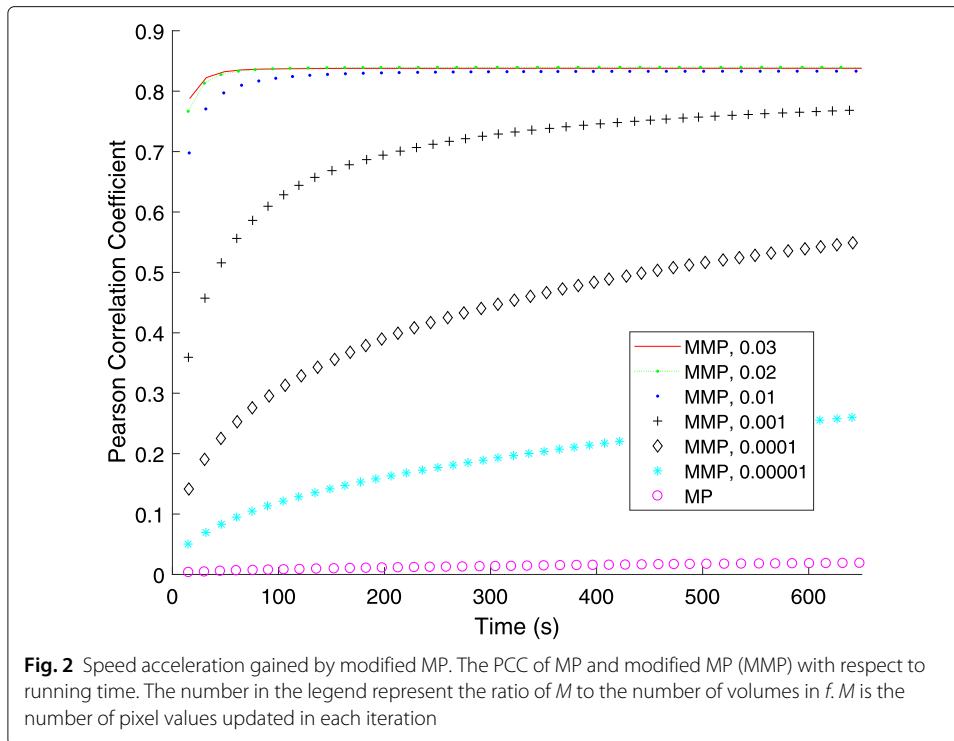
Convergence analysis

Except for the reconstruction accuracy, the convergence and running speed are also the important issues for a novel iterative method. Here, we first investigated the convergence of the modified MP algorithm (MMP) and the classic matching pursuit algorithm (MP) under different parameter configuration. And then compared the runtime of the proposed CSIIRR with the other classical methods (e.g. WBP, IIRR and ICON). The simulated data shown in the previous section is used as the benchmark dataset. When comparing the proposed modified MP and classic MP, we omitted the estimation of unknown projection in MMP (i.e., we set $p_{unknown}^{k-1} = 0$ in Algorithm 1) for fair comparison. One point should be noted is that the classic MP is a special case of MMP if setting the parameter M to 1. The Pearson correlation coefficient (PCC) between the ground-truth and each turn's reconstruction result is measured to indicate the improvement of reconstruction quality during the iteration process. Figure 2 demonstrates the change of PCC value for the two methods. It can be found that the MMP algorithm has a much faster convergence rate compared with the classic MP algorithm, no matter how we set the parameters. Furthermore, in the early iterations, the larger the M is, the faster the convergence speed will be.

When comparing the run-time between the whole CSIIRR framework and the classic WBP, IIRR, and ICON methods, the M used in modified MP is set to 335103 (i.e., with 0.01 ratio). We use structural similarity (SSIM) which is a measurement of image quality to evaluate CSIIRR. Figure 3 shows the experimental result, in which the SSIM curves of CSIIRR, ICON, WBP and IIRR with respect to the running time are demonstrated. It can be observed that CSIIRR converged to the highest SSIM value with the fastest convergence rate.
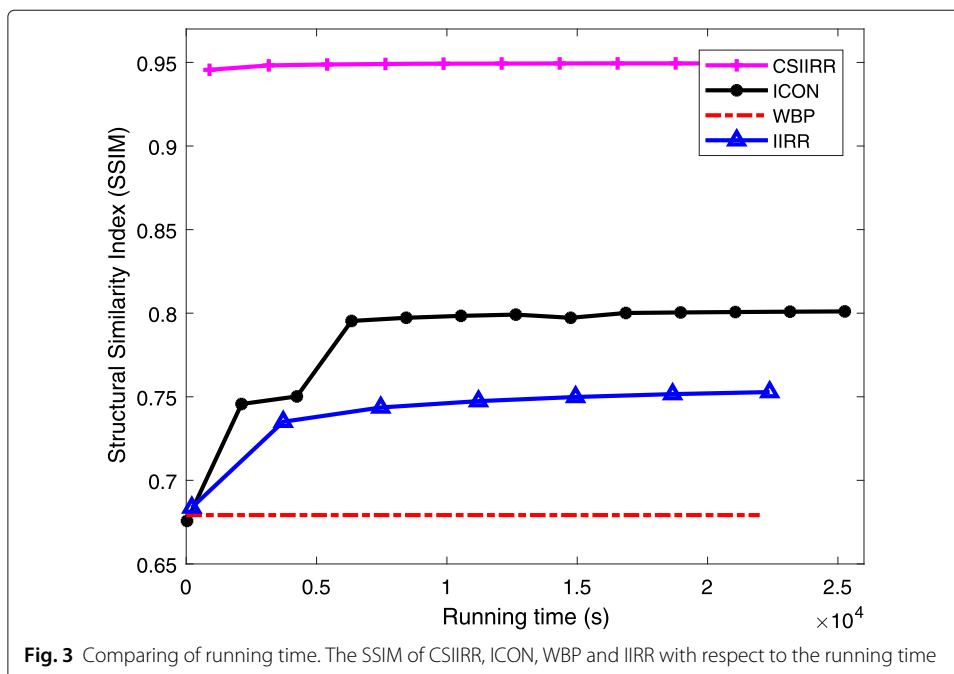
Effect of tilt angle change

Next, we tested the robustness of CSIIRR on different tilt angle ranges and intervals.
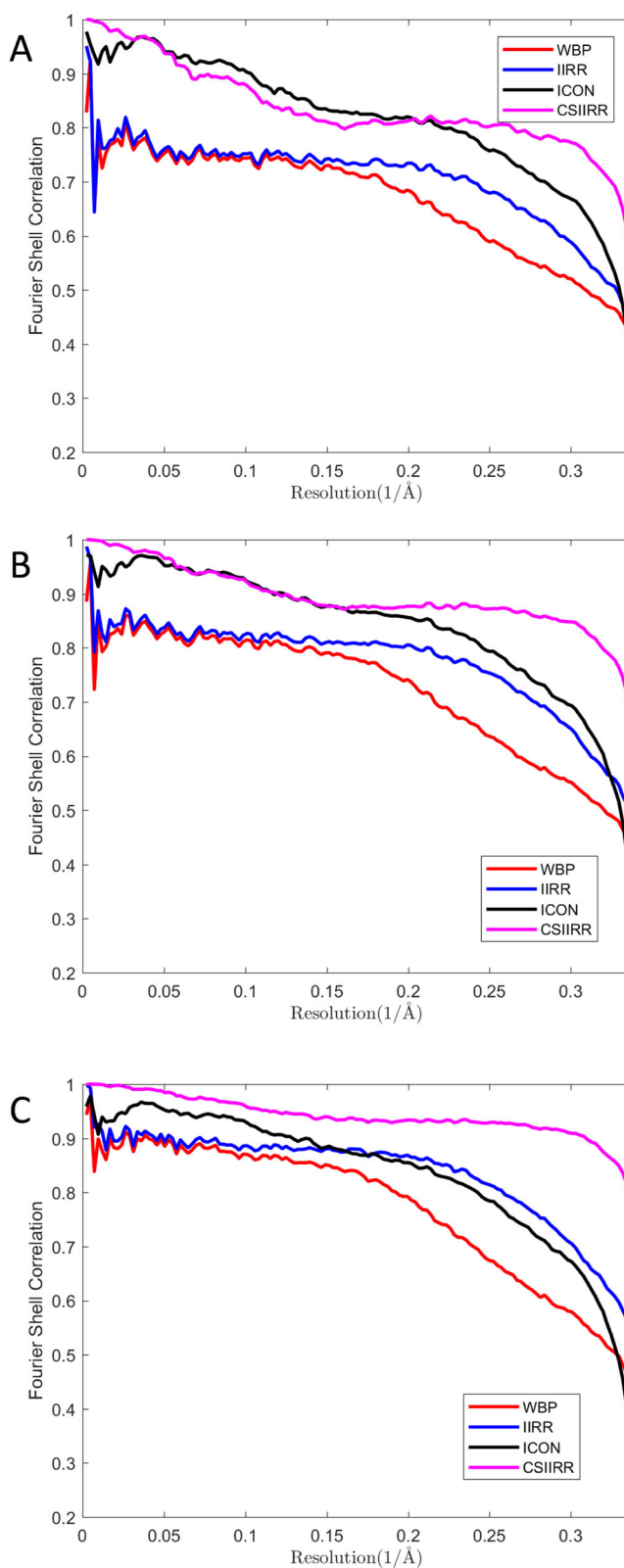
Li *et al. BMC Bioinformatics* 2020, **21**(Suppl 6):202

Page 9 of 19

Fig. 2 Speed acceleration gained by modified MP. The PCC of MP and modified MP (MMP) with respect to running time. The number in the legend represent the ratio of *M* to the number of volumes in *f*. *M* is the number of pixel values updated in each iteration

CSIIRR was tested on simulated data with angular range set to $-50°$ to $50°$, $-60°$ to $60°$ and $-70°$ to $70°$ (Fig. 4). As shown in Fig. 4, CSIIRR's FSC curve is significantly higher than the ones of other methods.

Furthermore, CSIIRR was tested on simulated data with angular intervals $1°$, $2°$ and $3°$ (Fig. 5). As shown in Fig. 5, CSIIRR's FSC curve has significantly higher coefficients than that of other methods, especially for high frequency.

Fig. 3 Comparing of running time. The SSIM of CSIIRR, ICON, WBP and IIRR with respect to the running time

Li *et al. BMC Bioinformatics* 2020, **21**(Suppl 6):202

Page 10 of 19

Fig. 4 Effect of angle range. **a**, **b**, **c** is the FSC curves of tomograms reconstructed from simulated data with angular range set to −50° to 50°, −60° to 60° and −70° to 70° respectively

Li *et al. BMC Bioinformatics* 2020, **21**(Suppl 6):202

Page 11 of 19
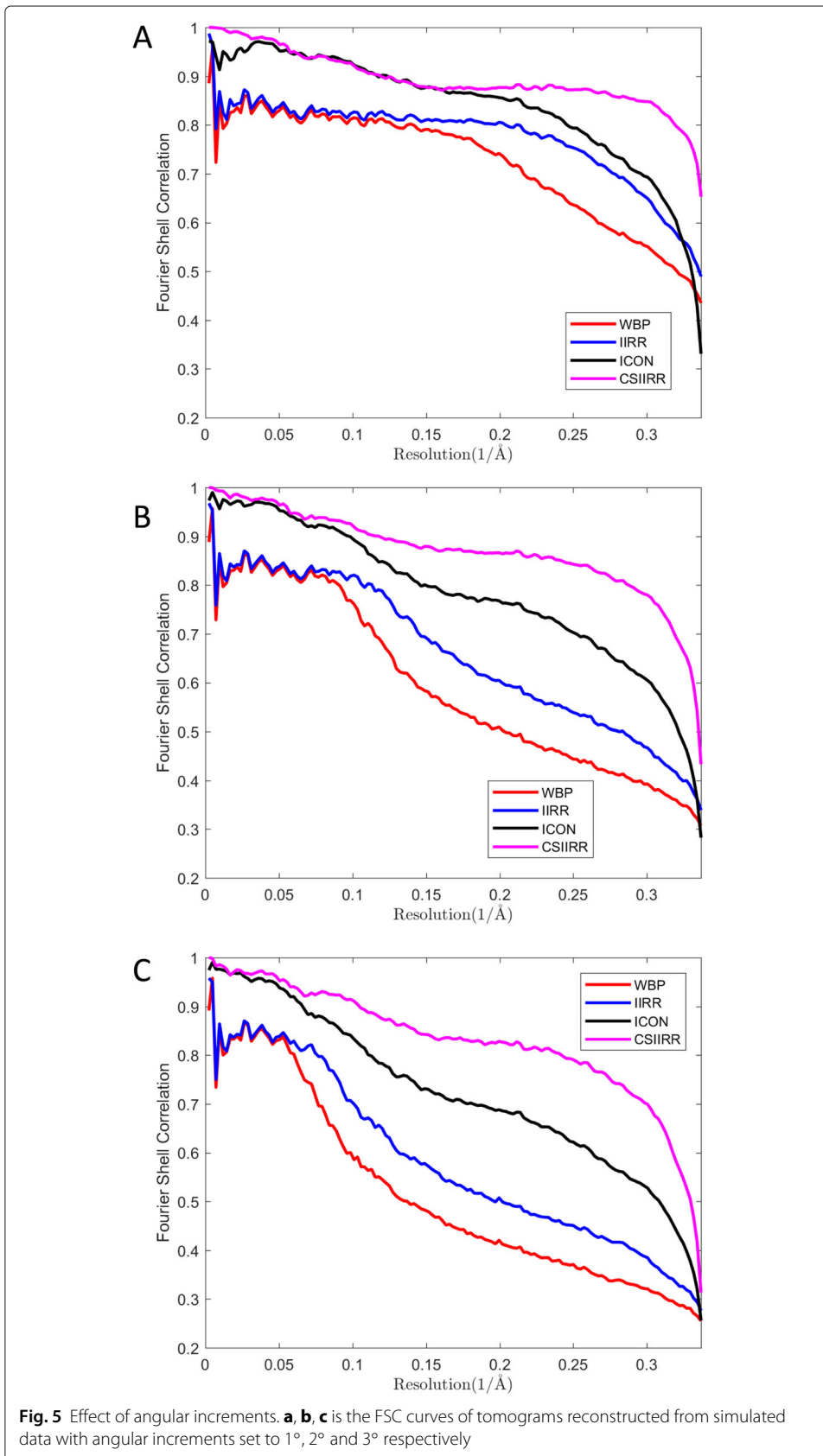
Fig. 5 Effect of angular increments. **a**, **b**, **c** is the FSC curves of tomograms reconstructed from simulated data with angular increments set to 1°, 2° and 3° respectively

Li *et al. BMC Bioinformatics* 2020, **21**(Suppl 6):202

Page 12 of 19

Effect of noise

The effect of noise on the behavior CSIIRR has been investigated. The noise with signal-to-noise ratio $SNR = 0.5, 1, 2$ is added to the projections and the tomogram is reconstructed by WBP, IIRR, ICON and CSIIRR. The Fourier shell correlation (FSC) between the reconstructed tomogram and the ground-truth is calculated to show the reconstruction quality of each method. Figure 6 shows corresponding FSC curve. For the data with $SNR = 0.5$ and 1(Fig. 6a, b), the result obtained by CSIIRR achieves the highest correlation value for each Fourier frequency compared with the ones of WBP, IIRR and ICON. For the data with $SNR = 2$ (Fig. 6c), the result obtained by CSIIRR achieves the highest correlation value for high frequency compared with the ones of WBP, IIRR and ICON.

Performance on real data

The proposed CSIIRR was further challenged by real-world datasets. We first made a whole tilt series reconstruction to see the behavior differences between the proposed CSI-IRR and the other classical methods (e.g., IIRR and ICON). Then, a leave-one-out analysis was carried out to further make a quantitative analysis.

The leave-one-out [19] validation can show how precise the missing information is restored by a reconstruction method. In detail, we exclude the projection with the minimum tilt angle from the tilt series and then reconstruct the tomogram without considering the excluded projection. After the reconstruction, we reproject the tomogram according to the tilt angle value of the excluded projection and calculate the similarity between the reprojection and the exact excluded projection by Fourier ring correlation.

Experiment on centriole

The Centriole dataset is a tilt series of plastic embedded cell section around a centriole region, which was taken on a FEI TF30 microscope (operated at 300 kV) with a Gatan Camera. This dataset is downloaded from IMOD's website [20]. The tilt angles of the projections range from $+65.0°$ to $-61.0°$ at $2°$ intervals. The size of each projection is 1024×1024 with a pixel size of 1.01 nm. The tilt series is aligned in advance by IMOD [20].
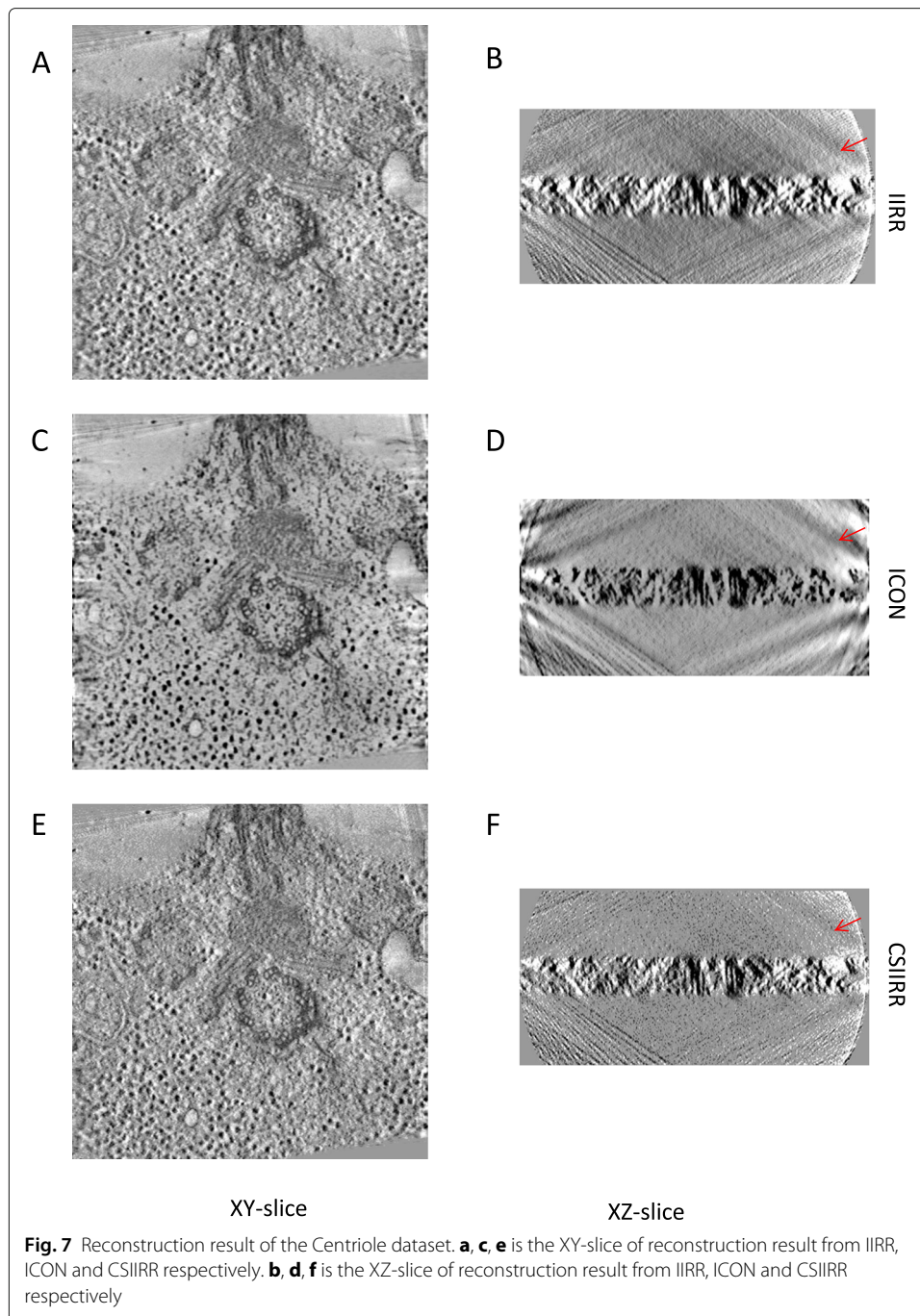
To reconstruct the tomogram, the number of iterations is set to 15 in IIRR. For the sake of fairness, the number of outer-loop iterations in our proposed algorithm is also set to 15. The number of iterations for modified MP in CSIIRR is set to 3, and the number of atoms selected in each iterations M is set to 19,191,398 in modified MP. The iteration is set to '30,120,300' in ICON.

Figure 7 shows the XY-slice and XZ-slice of the reconstructed tomograms by different methods. From Figs. 7a, b and e, f, it can be found that the tomogram reconstructed by CSIIRR has a clearer background comparing with the ones reconstructed by IIRR. Additionally, according to Fig. 7c, d, stripe noise appears around the boundary of the XY-slice of tomogram from ICON while this kind of stripe noise does not appear around the boundary of the XY-slice of tomogram from CSIIRR.

Figure 8 demonstrates the result of reprojection analysis. It can be found that the reprojection produced by IIRR is blurred while the reprojection produced by CSIIRR and ICON are very clear and keep high similarity with the exact excluded groundtruth. The Fourier
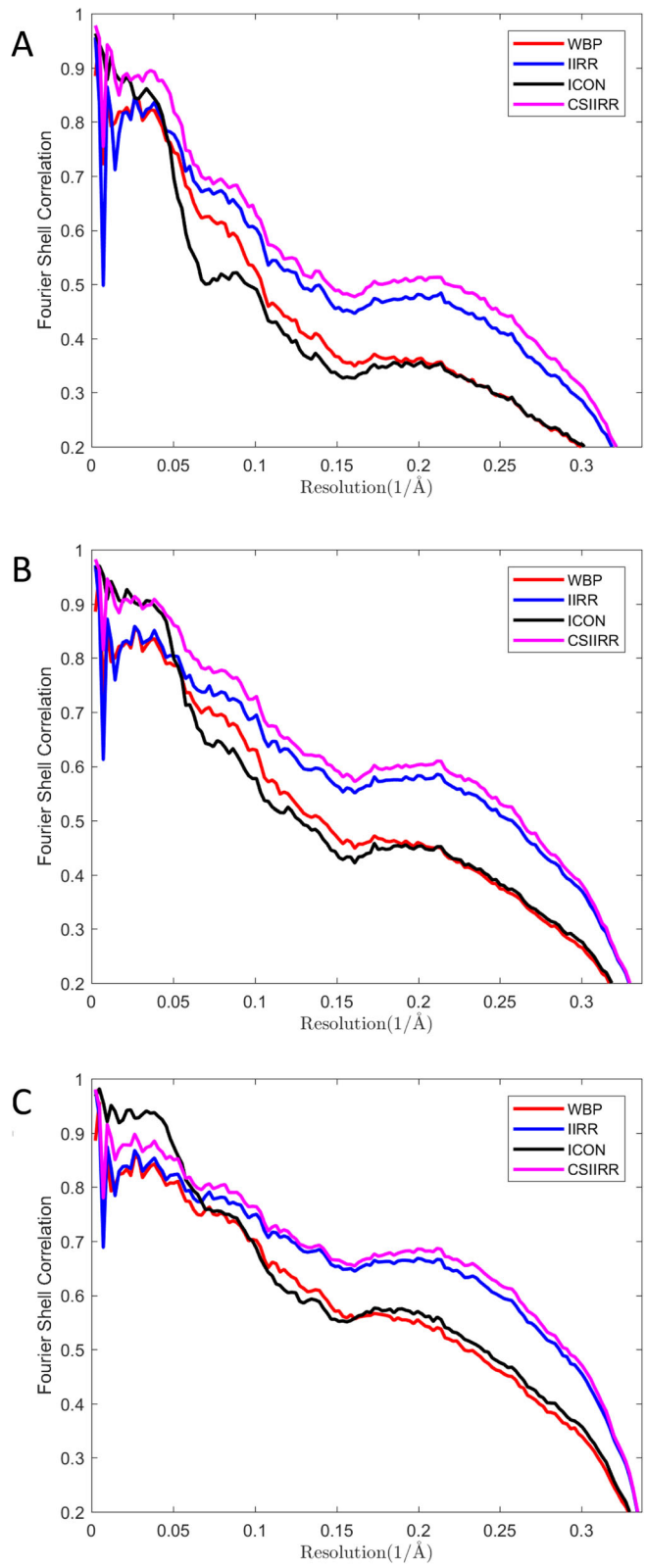
Fig. 6 Effect of noise. **a**, **b**, **c** is the FSC curves of tomograms reconstructed from simulated data with SNR set to 0.5, 1 and 2 respectively

Li *et al. BMC Bioinformatics* 2020, **21**(Suppl 6):202

Page 14 of 19

Fig. 7 Reconstruction result of the Centriole dataset. **a**, **c**, **e** is the XY-slice of reconstruction result from IIRR, ICON and CSIIRR respectively. **b**, **d**, **f** is the XZ-slice of reconstruction result from IIRR, ICON and CSIIRR respectively

ring correlation (FRC) curve (Fig. 8b) further indicates that CSIIRR could estimate the missing information more precisely compared with ICON.

Experiment on mitochondria

The Mitochondria dataset is a tilt series of mitochondria from mouse hepatic taken with a FEI Tecnai 20, with voltage at 200 kV. cells without fiducial markers. The tilt angles of the projections range from −52.0° to +59.0° at 1° intervals. The tilt series is aligned in advance by the marker-free alignment module in AuTom [16, 21].

Li *et al. BMC Bioinformatics* 2020, **21**(Suppl 6):202

Page 15 of 19

Fig. 8 Reprojection result of the Centriole dataset. **a** The excluded projection (groundtruth) and the reprojections at the minimum tilt angle produced by IIRR, ICON and CSIIRR. **b** The FRC curve of IIRR, ICON and CSIIRR, by comparison with the groundtruth
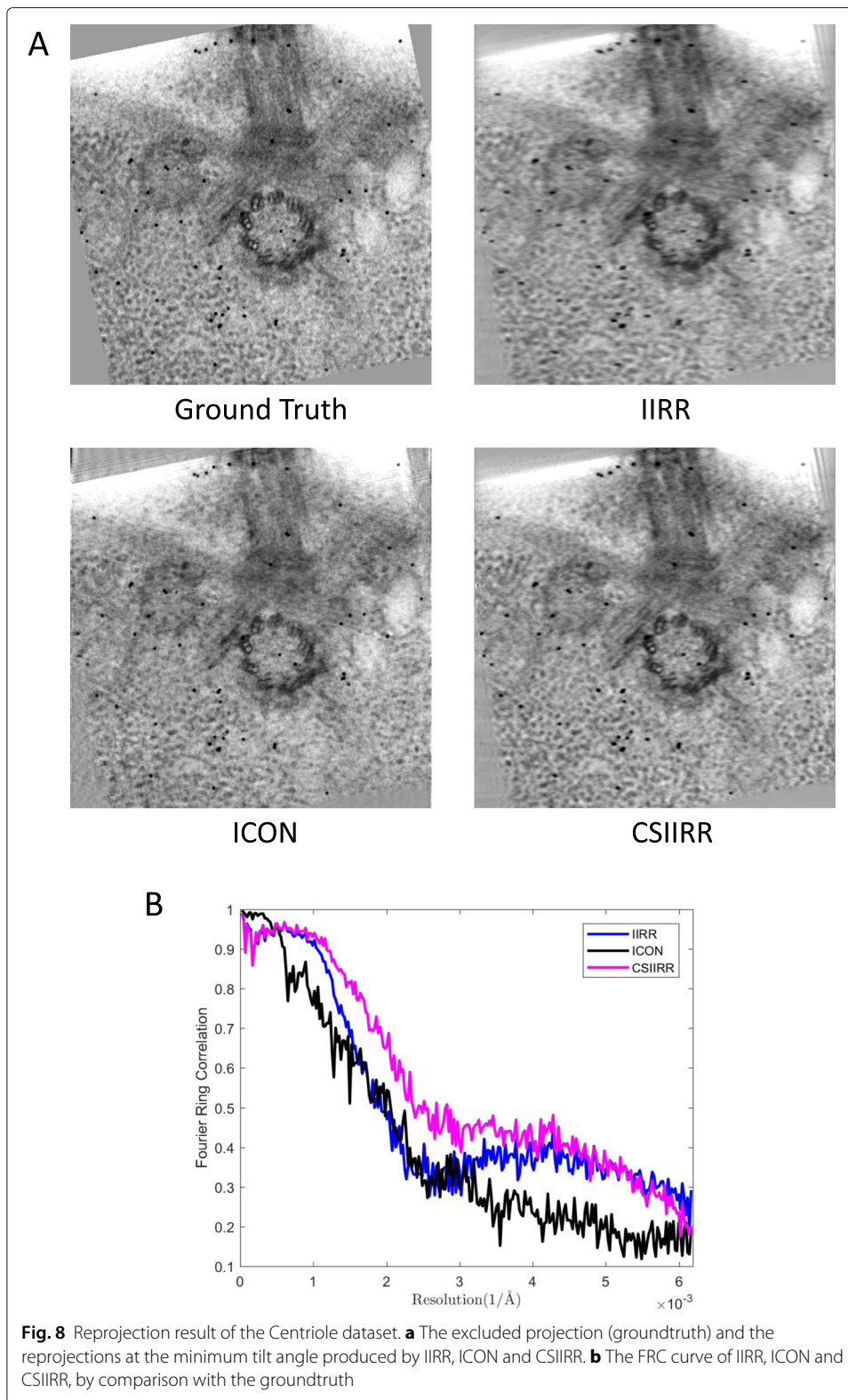
Figure 9 shows the XY-slice and XZ-slice of the reconstructed tomograms. As shown in Fig. 9a, b and e, f, the tomogram reconstructed by CSIIRR has a better contrast comparing with the ones reconstructed by IIRR. According to Fig. 9c, d, stripe noise appears around
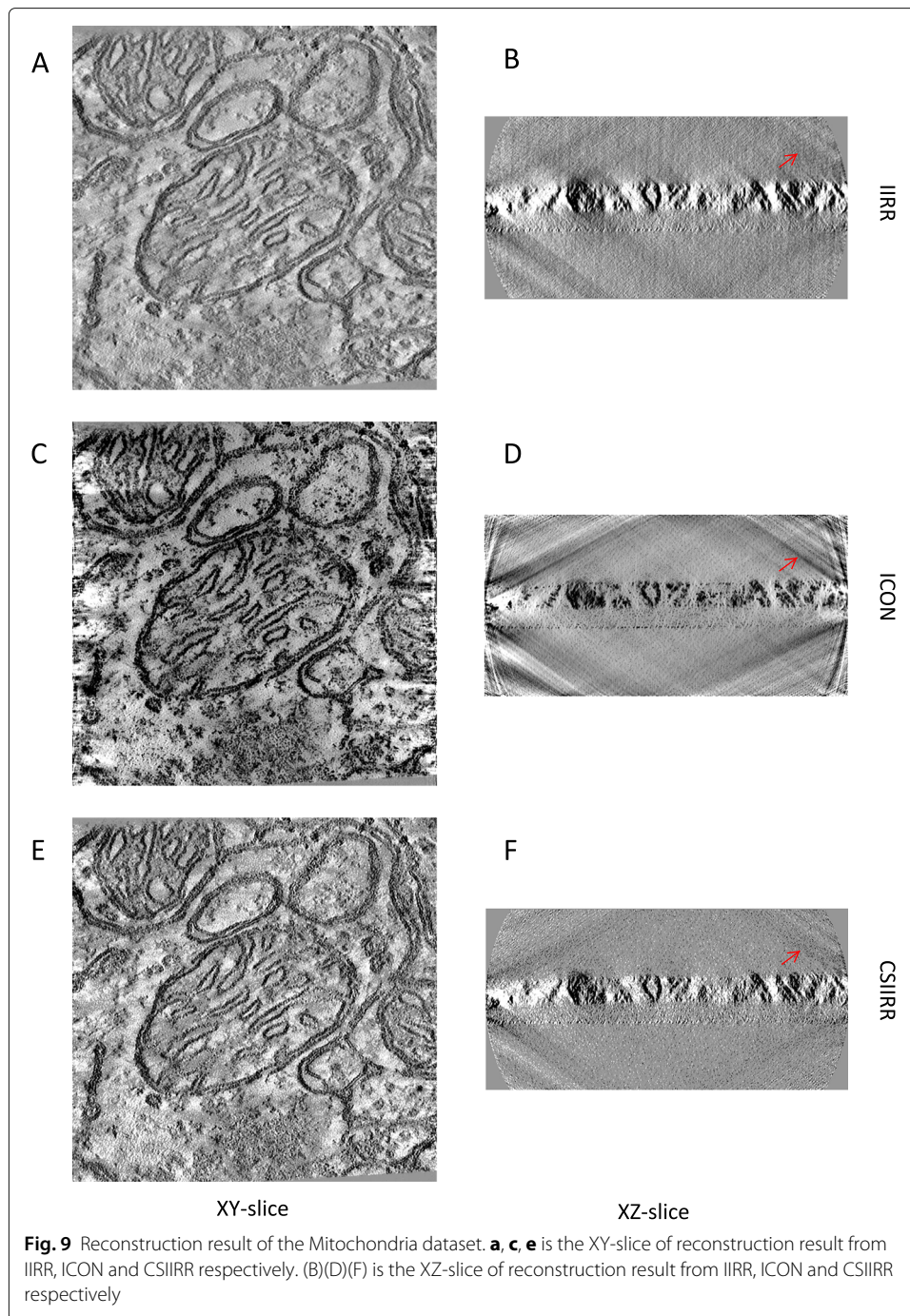
Li *et al. BMC Bioinformatics* 2020, **21**(Suppl 6):202

Page 16 of 19

Fig. 9 Reconstruction result of the Mitochondria dataset. **a**, **c**, **e** is the XY-slice of reconstruction result from IIRR, ICON and CSIIRR respectively. (B)(D)(F) is the XZ-slice of reconstruction result from IIRR, ICON and CSIIRR respectively

the boundary of the XY-slice of tomogram from ICON while this kind of stripe noise does not appear around the boundary of the XY-slice of tomogram from CSIIRR.

Figure 10 demonstrates the experiment result. As shown in Fig. 10a, the reprojection produced by IIRR has been blurred compared with the groundtruth. The reprojection produced by ICON looks better than IIRR. Nevertheless, strip artifacts still exist around the boundary of the projection produced by ICON. The reprojection produced by CSIIRR

Fig. 10 Reprojection result of the Mitochondria dataset. **a** The excluded projection (groundtruth) and the reprojections at the minimum tilt angle produced by IIRR, ICON, and CSIIRR. **b** The FRC curve of IIRR, ICON and CSIIRR, by comparison with the groundtruth
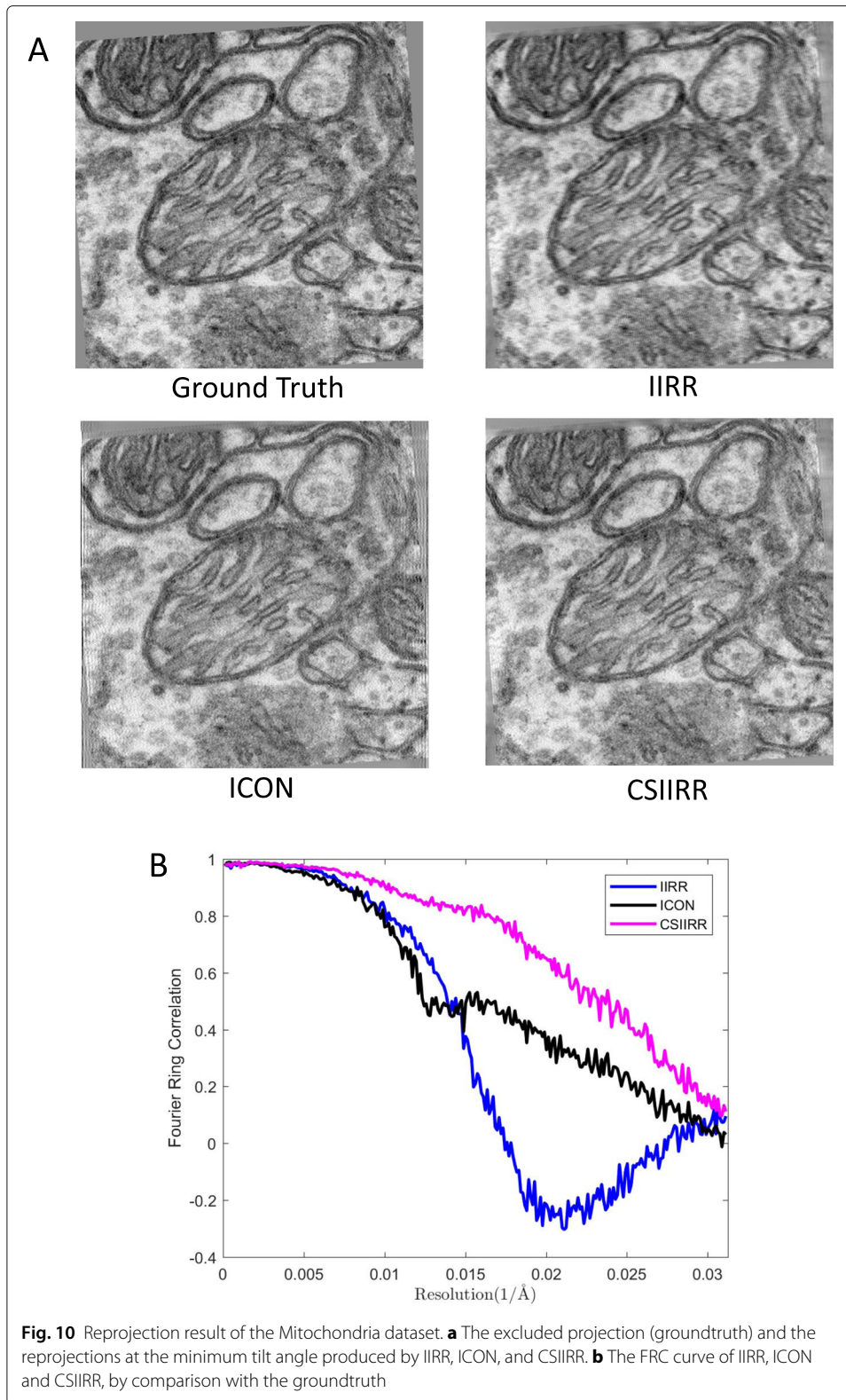
has the most visual details and is the most similar to the original excluded projection. The FRC curves shown in Fig. 10b further support our conclusion.

Li *et al. BMC Bioinformatics* 2020, **21**(Suppl 6):202

Page 18 of 19

Conclusion

In this work, we proposed a novel algorithm for the reconstruction of electron tomography. By combining the concept of compressed sensing (CS) and improved iterative reconstruction-reprojection (IIRR) together, this algorithm is able to suppress the effect of the missing wedge and restore the missing information.

Setting a proper value to the parameters in CSIIRR is really important to the performance of this method. First, the factor λ should be set carefully to ensure the convergence of the method. The author proposing the IIRR suggested that λ should be close to 1, so we set $\lambda = 0.99$ in the experiment. Next, the number of atoms M should be set carefully to trade off the convergence rate and the sparsity of the reconstructed volume.

The proposed algorithm was challenged by both the simulated data and real-world data. The results show that the proposed algorithm has an obvious advance in the suppression of missing wedge effects and the restoration of missing information.

Author details

[1] High Performance Computer Research Center, Institute Of Computing Technology, Chinese Academy of Sciences, 100190 Beijing, China. [2] University of Chinese Academy of Sciences, Beijing, China. [3] Research Center for Mathematics and

Li *et al. BMC Bioinformatics* 2020, **21**(Suppl 6):202

Page 19 of 19

Interdisciplinary Sciences, Shandong University, 266237 Qingdao, People's Republic of China. [4]National Natural Science Foundation of China, 100085 Beijing, China.

References

1. Frank J. Electron Tomography: Methods for Three-dimensional Visualization of Structures in the Cell. New York: Springer; 2008.
2. Chen Y, Zhang Y, Zhang K, Deng Y, Wang S, Zhang F, Sun F. FIRT: Filtered iterative reconstruction technique with information restoration. J Struct Biol. 2016;195(1):49–61.
3. Nassi M, Brody WR, Medoff BP, Macovski A. Iterative reconstruction-reprojection: an algorithm for limited data cardiac-computed tomography. IEEE Trans Biomed Eng. 1982;5:333–41.
4. Medoff BP, Brody WR, Nassi M, Macovski A. Iterative convolution backprojection algorithms for image reconstruction from limited data. JOSA. 1983;73(11):1493–500.
5. Kim J, Kwak K, Park S, Cho Z. Projection space iteration reconstruction-reprojection. IEEE Trans Med Imaging. 1985;4(3):139–43.
6. Wan X, He X, Zou W, Chen X, Cai M. A reconstruction algorithm with iterative reconstruction-reprojection and an fir filter. In: Information Optics and Photonics Technology, vol. 5642. International Society for Optics and Photonics; 2005. p. 364–72. https://doi.org/10.1117/12.568775.
7. Duan X, Zhang L, Xing Y, Chen Z, Cheng J. Few-view projection reconstruction with an iterative reconstruction-reprojection algorithm and tv constraint. IEEE Trans Nucl Sci. 2009;56(3):1377–82.
8. Kim J, Kwak K, Park S. Iterative reconstruction-reprojection in projection space. Proc IEEE. 1985;73(6):1140–1.
9. Yazdanpanah AP, Regentova EE, Bebis G. Algebraic iterative reconstruction-reprojection (airr) method for high performance sparse-view ct reconstruction. Appl Math. 2016;10(6):1–8.
10. Zhang Z, Xiong X, Yang W. An iterative image reconstruction algorithm with improved convergence. In: Proc. of 4th World Congress on Industrial Process Tomography. United Kingdom: VCIPT; 2005. p. 708–13.
11. Zhang Z, Zhang P. Improved iterative re-projection algorithm for tomographic image reconstruction. In: The 3rd International Symposium on Process Tomography. Poland; 2004. p. 183–6.
12. Zhang Z. Research on incomplete data tomography theory and application. PhD thesis. 2005.
13. Deng Y, Chen Y, Zhang Y, Wang S, Zhang F, Sun F. Icon: 3d reconstruction with 'missing-information'restoration in biological electron tomography. J Struct Biol. 2016;195(1):100–12.
14. Chen Y, Wang Z, Zhang J, Li L, Wan X, Sun F, Zhang F. Accelerating electron tomography reconstruction algorithm icon with gpu. Biophys Rep. 2017;3(1-3):36–42.
15. Wang Z, Chen Y, Zhang J, Li L, Wan X, Liu Z, Sun F, Zhang F. Icon-mic: Implementing a cpu/mic collaboration parallel framework for icon on tianhe-2 supercomputer. J Comput Biol. 2018;25(3):270–81.
16. Han R, Zhang F, Wan X, Fernández J-J, Sun F, Liu Z. A marker-free automatic alignment method based on scale-invariant features. J Struct Biol. 2014;186(1):167–80.
17. Han R, Wang L, Liu Z, Sun F, Zhang F. A novel fully automatic scheme for fiducial marker-based alignment in electron tomography. J Struct Biol. 2015;192(3):403–17.
18. James NR, Brown A, Gordiyenko Y, Ramakrishnan V. Translational termination without a stop codon. Science. 2016;354(6318):1437–40.
19. Cardone G, Grünewald K, Steven AC. A resolution criterion for electron tomography based on cross-validation. J Struct Biol. 2005;151(2):117–29.
20. Kremer JR, Mastronarde DN, McIntosh JR. Computer visualization of three-dimensional image data using imod. J Struct Biol. 1996;116(1):71–6.
21. Han R, Wan X, Wang Z, Hao Y, Zhang J, Chen Y, Gao X, Liu Z, Ren F, Sun F, et al. Autom: A novel automatic platform for electron tomography reconstruction. J Struct Biol. 2017;199(3):196–208.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.