CAMBRIDGE
UNIVERSITY PRESS

RESEARCH ARTICLE

GoldDigger and Checkers, computational developments in cryo-scanning transmission electron tomography to improve the quality of reconstructed volumes

Genevieve Buckley, Georg Ramm and Sylvain Trépout

Ramaciotti Centre for Cryo-EM, Monash University, Clayton, VIC, Australia
Corresponding author: Sylvain Trépout; Email: sylvain.trepout@monash.edu

Abstract

In this work, we present a pair of tools to improve the fiducial tracking and reconstruction quality of cryo-scanning transmission electron tomography (STET) datasets. We then demonstrate the effectiveness of these two tools on experimental cryo-STET data. The first tool, GoldDigger, improves the tracking of fiducials in cryo-STET by accommodating the changed appearance of highly defocussed fiducial markers. Since defocus effects are much stronger in scanning transmission electron microscopy than in conventional transmission electron microscopy, existing alignment tools do not perform well without manual intervention. The second tool, Checkers, combines image inpainting and unsupervised deep learning for denoising tomograms. Existing tools for denoising cryo-tomography often rely on paired noisy image frames, which are unavailable in cryo-STET datasets, necessitating a new approach. Finally, we make the two software tools freely available for the cryo-STET community.

Impact Statement

During the fiducial alignment of cryoSTET tilt-series, out-of-focus fiducials can be missed by the alignment software, leading to poor alignment quality. GoldDigger is a new software that improves the detection of out-of-focus fiducials in cryoSTET images. 3D cryoSTET reconstructions often contain an important amount of noise. Checkers is the first deep-learning-based denoising tool to significantly reduce the noise in cryoSTET reconstructions. Together, GoldDigger and Checkers help better interpret the structural information of cellular samples in cryoSTET.

---

1. Introduction

Cryo-scanning transmission electron tomography (STET) is a 3D imaging method that is used to observe the structure of biological samples cryo-fixed in their native state. After collecting projection images of a unique sample tilted at different angles inside the electron microscope, usually between $-60°$ and $+60°$, the 3D structure of the object of interest is mathematically reconstructed using the tilted projections. Cryo-STET is complementary to the more broadly known cryo-electron tomography (ET) method, as it allows imaging of much thicker samples up to ~1.5 μm,[1,2] whereas cryo-ET is limited to the study of samples thinner than ~300 nm.[3] Biological samples can be observed as is using cryo-STET if their thickness does not exceed 1.5 μm, which significantly reduces and simplifies the sample preparation workflow. Briefly, samples can be

deposited (or grown) directly on electron microscopy grids, plunge-frozen in a cryogenic fluid, and then imaged at the electron microscope without any additional thinning or milling step which is usually required in cryo-ET. Cryo-STET allowed the first observation of several different types of samples (e.g., large bacteria, unicellular eukaryotic parasites, mammalian cells) without cutting or milling them.[2,4,5]

First developed in 2014,[6] cryo-STET is a scanning transmission electron microscopy (STEM) method that can be operated in bright-field (BF) and/or dark-field (DF) modes and works by scanning a focused electron beam across the sample, measuring the amplitude contrast at each point to produce an image.[7,8] Cryo-STET is better performed in BF mode compared to DF mode as multiple scattering (beam broadening) occurring in thick samples is detected on the DF detector and generates noisy images.[9,10] Cryo-STET resolutions in BF mode are lower than that of cryo-ET (a few nanometers vs sub-nanometric, respectively). However, STEM can also be performed in integrated differential phase contrast (iDPC) mode, and recent developments have already been made to perform cryo-STET using this more recent imaging modality.[11−13] Work performed in 2D has shown that iDPC single particle analysis can reach resolutions of 3.5 Å on protein samples.[12] This demonstrates the high potential for higher-resolution studies in cryo-STET. Given the advantages of its imaging capabilities, cryo-STET is the tool of choice to study thick cryo-fixed biological specimens such as thick cellular samples at the nanometer level.

A critical initial step of tomography is the fine alignment of the collected projection images to accurately reconstruct the observed sample without reconstruction artifacts. In tomography, this is typically performed by adding gold beads to the sample which are high contrast features in the images and can be used as fiducial markers to accurately measure the image movements throughout the whole tilt-series.[14] Originally performed by hand, fiducial picking and tracking is now fully automatic, enabling bulk processing of the ever-increasing amount of collected data.[15−17] Fiducial-based tilt-series alignment is particularly well suited for cryo-ET. However, cryo-STET has several fundamental differences from conventional cryo-ET. First, the samples are much thicker. The thin slab geometry generally accepted in cryo-ET is no longer respected because of the thickness and the non-slab geometry of the sample. Consequently, the contrast and noise in cryo-STET images are challenging. Also challenging is the fact that because the samples are thicker, the gold bead position in Z is much more scattered and less likely to fall exactly in the plane of focus. This can create differences in the appearance of gold beads, such as blurring. This can cause problems identifying and tracking the gold beads, necessary for image alignment. All these negative effects become worse at higher tilt angles, where the tilt angle causes the electron beam path to pass through a greater thickness of the sample.

These differences mean that tools and algorithms developed specifically for cryo-ET data may not be fully compatible with cryo-STET data. Because cryo-STET is an emerging method, it lacks the immense variety of post-acquisition tools available in cryo-ET, which hampers the downstream processing of collected cryo-STET data. One existing tool developed to improve the alignment of thick cryo-samples, is ClusterAlign, an alternative fiducial-based tilt-series alignment software.[18] Instead of tracking individual fiducial markers separately, ClusterAlign identifies clusters of fiducial markers and tracks the vectors defining the clusters throughout the tilt-series. This strategy is more robust for thick samples such as those observed in cryo-STET. Nevertheless, there is still a present need for more tools to support data processing for the growing cryo-STET community.

In cellular tomography samples are usually relatively thick (up to ~300 nm in cryo-ET and up to ~1.5 μm in cryo-STET), generating noisy images which makes the analysis difficult, particularly because intracellular spaces are intricate with packed and intertwined cytosolic molecules and membranes. Denoising tools can be used to improve the quality of reconstructed tomograms, increasing interpretability. CryoCARE[19] is an unsupervised deep-learning denoising algorithm which uses the Noise2Noise strategy[20] and is the current gold standard tool for cryo-ET denoising. In the Noise2Noise strategy, two images of the same sample with independent noise are used to separate the signal from the noise. Because, direct electron detector cameras used in cryo-ET acquire multiple movie frames which can be accurately aligned,[21] the aligned movie frames can be split into two groups (i.e., even vs odd numbered frames) and independently reconstructed. Finally, the twin independently noisy volumes are used as input for the Noise2Noise strategy. CryoCARE enhances the contrast of the 3D reconstructions, significantly

improving the quality of details in the observed structures. This better facilitates the identification of proteins inside their native environment or the discrimination of adjacent membranes in intricate cellular compartments. Denoised 3D reconstructions also improve the effectiveness of segmentation tools, which are used to better interpret and present structures within the data.

However, another important difference is the collection time of a cryo-STET image which is much longer than for conventional cryo-ET (several tens of seconds vs. typically a few seconds). This means that some drift can happen during the first collected pixel and the last one, generating image distortion. In material sciences, this type of image distortion can be corrected using the revolving STEM method which consists of the collection of two orthogonal images to measure and determine a potential drift, but this strategy is not easily applicable to biological samples, mainly because two images must be collected generating additional beam damage.[22] Moreover, in cryo-STET, there are no such things as movie frames because the acquisition time per frame is too slow. This prevents the use of motion correction algorithms[21] or denoising as performed in CryoCARE[19] which are instrumental in the cryo-ET data analysis pipeline.

In this work, we present two new tools, targeting two crucial steps of the cryo-STET pipeline, (i) fiducial tracking to align cryo-STET datasets when visibility of gold beads is limited (e.g., out-of-focus) and (ii) denoising of cryo-STET reconstructions using deep-learning. GoldDigger is a fiducial picking and tracking tool: it uses a novel strategy combining existing tools and can track out-of-focus gold beads or beads of varying diameters. Checkers is a denoising tool: it uses a combined strategy to transform a single image collected in cryo-STET into paired noisy images, then enhanced with unsupervised deep-learning-based denoising algorithms such as Cryo-CARE.[19]

2. Wet lab methods

2.1. Bacteria cell culture, sample preparation, and data collection

E. coli MG1655 was grown until the mid-stationary phase, up to an OD_{600} 1.5. Bacteria were plunge-frozen on holey carbon-coated gold Quantifoil Finder grids (Quantifoil Micro Tools, Großlöbichau, Germany) R2/2 (reference: Au G200F1) previously glow-discharged using a PELCO EasiGlow (Ted Pella, Inc., Redding, CA, USA). To allow tilt-series fiducial alignment, gold beads (15 nm from Aurion) were deposited and dried onto the grids prior to cell addition. Grids were manually blotted with a Whatman filter paper using a Leica EM-CPC (Leica Microsystems, Wetzlar, Germany). More details about the cells can be found in the literature as cells from the same culture batch, prepared using the same setup, have previously been investigated using cryo-soft X-ray tomography.[23] Cryo-plunged Quantifoil EM grids were mounted on a Gatan 914 high-tilt cryo-holder (Gatan, Pleasanton, CA, USA).

Cryo-STET datasets were collected on a JEOL 2200FS 200 kV FEG electron microscope (JEOL, Tokyo, Japan). 3 k by 3 k images were collected in BF mode using a 20 μm condenser aperture (corresponding convergence semi-angle about 6 mrad) and an on-axis JEOL STEM detector (40 cm camera length, corresponding collection semi-angle about 10 mrad). The dwell time was set between 2 and 4 μs/pixel and the magnifications used were 40,000× (corresponding pixel size 1.5 nm). Images and tilt-series acquisition were performed in Digital Micrograph using the Digiscan II. Fully automatic cryo-STET tilt-series collection was performed using a lab-made script developed in Digital Micrograph.[24] When possible, tilt-series were collected up to ±75° using 3° tilt increments (50 tilts overall), however, some tilt-series could not be collected using this important angular range as their position on the EM grid did not allow it. The estimated total electron dose received by the sample was about 80 e⁻/Å².

2.2. Protozoan cell culture, sample preparation, and data collection

The *Trypanosoma brucei* cryo-STET data used for the development and testing of the GoldDigger comes from the original publication of the structural analysis of the *T. brucei* flagellum attachment zone.[2] Information about sample preparation can be found in the original publication. Briefly, *T. brucei* AnTat 1.1E bloodstream forms were used and fixed when they reached the exponential growth phase. Chemically fixed cells were deposited on Quantifoil 200 mesh R2/2 grids (Quantifoil, Großlöbichau, Germany)

pre-coated with a special gold bead solution composed of commercial 15 nm gold beads (Aurion) and lab-made gold nanorods of various dimensions (synthesized at Li's laboratory, Ecole Normale Supérieure Chimie Paris- Tech, Paris, France). Grids were manually plunge-frozen using a Leica EM-CPC equipment (Leica, Wetzlar, Germany).

Cryo-STET datasets were collected on JEOL 2200FS microscope (JEOL, Tokyo, Japan) in bright-field mode using an on-axis JEOL STEM detector (60 cm camera length) and a 40 µm condenser aperture (beam convergence and collection semi-angles were 9.3 and 6.6 mrad, respectively). The dwell time was set between 1 and 3 µs/pixel and the magnifications used ranged between 30,000× and 50,000× (corresponding pixel sizes ranged between 2 and 1.3 nm, respectively). Tilt series were collected in Digital Micrograph using a Digiscan II with a lab-made script. Tilt-series were collected between $-70°$ and $+70°$ using $2°$ tilt increments. The total electron dose received by the sample ranged between 40 and 80 e$^-$/Å^2.

3. Software methods

3.1. *GoldDigger, an extensive fiducial detection and tracking tool for cryo-STET datasets*

The data used for designing and testing GoldDigger consisted of 7 cryo-STET tilt-series of *E. coli* bacteria (named TS Ec001 to TS Ec007 throughout the text) and 12 cryo-STET tilt-series of *T. brucei* (named TS Tb001 to TS Tb012 throughout the text) collected on a JEOL 2200FS, a side-entry electron microscope.

3.1.1. *Rationale and software overview*

GoldDigger was created due to the fact that existing cryo-ET tilt-series alignment software perform poorly on cryo-STET datasets. At the time the development of GoldDigger was initiated, the software Cluster-Align was not published yet. We identified two main issues hampering alignment or causing misalignment of cryo-STET tilt-series. The number of detected fiducials is limited especially at high tilt-angles and the fiducials are often incorrectly tracked. This can lead to having to discard entire tilt-series. Also, high throughput data cannot be automatically aligned with tilt-series alignment requiring manual correction which is time-consuming.

GoldDigger solves this by more accurately detecting out-of-focus gold beads, improving the fiducial tracking used to align the tilt-series. Because cryo-STET is used to observe thick specimens, many gold beads can be out-of-focus, especially at high tilts and when high convergence semi-angles are used. High convergence semi-angles allow small probe sizes (sub-nm) for high-resolution studies but have limited depth of focus (few tens of nm).[25] In cryo-STEM, out-of-focus gold beads appear as a blurry spot with a diameter larger than the nominal size of the gold bead.

Moreover, initial raw alignment of cryo-STEM tilt-series often fails especially at high-tilts which makes the tracking of the gold beads from low-tilts to high-tilts difficult. This might be due to the contrast difference between the thick sample and the thin carbon films which support it. At high-tilts, the specimen contours are no longer clearly resolved because of the strong darkening in the image caused by the low number of electrons reaching the STEM detector. This issue becomes even more important on side-entry electron microscopes because of the strong inherent stage instability compared to recent electron microscopes equipped with an autoloader (e.g., TFS Krios and JEOL Cryo ARM). GoldDigger was developed and tested on cryo-STET datasets collected on a side-entry electron microscope ensuring that fiducial tracking does not fail even when the raw tilt-series suffers strong shifts, as it often happens on side-entry cryo-electron microscopes.

The advantage of GoldDigger is how it can detect more fiducials by tracking out-of-focus beads. First, GoldDigger searches for gold beads of different sizes, before merging the different tracks of the detected gold beads. To analyze most of the collected data and tackle the pre-alignment difficulties, GoldDigger proposes an optional pre-alignment step using StackReg,[26] which showed better and more consistent results than any other alignment software we tried, and allowed better subsequent detection of gold beads using the main core of GoldDigger. The different computing steps performed in GoldDigger workflow are presented in a diagram (Figure 1). GoldDigger was developed on Linux, it uses a main bash script which
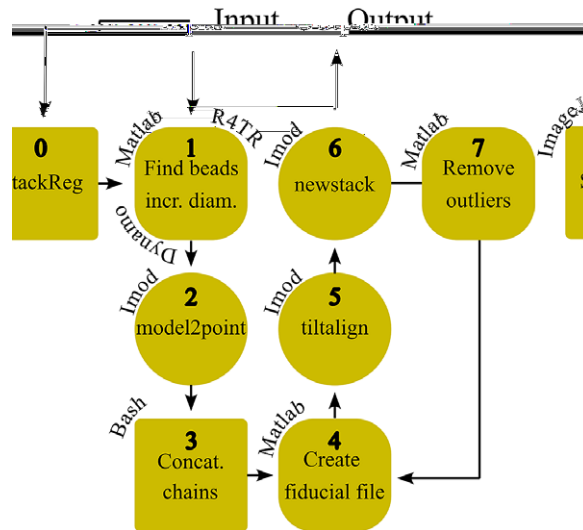
Figure 1. *Diagram of GoldDigger workflow. The whole workflow is performed by a bash script which calls several other software (ImageJ, MATLAB, R4TR, Dynamo, IMOD). The main processing loop is located at the center of the diagram (step 1: detection of the beads in MATLAB using R4TR and Dynamo; step 3: concatenation of the fiducial chains in bash; step 4: merging of the fiducial chains and creation of the final fiducial file in MATLAB; step 5: alignment in IMOD; and step 6: aligned tilt-series computation in IMOD). Two optional steps are displayed on the left-hand side (step 0: pre-alignment step of the raw tilt-series using StackReg in ImageJ) and the right-hand side of the diagram (step 7: post-alignment removal of outlying fiducials in MATLAB).*

calls different software: MATLAB (The MathWorks Inc., Natick, MA, USA), Dynamo,[16] IMOD[14] and ImageJ[27] as well as the relion4_tomo_robot (shortened to R4TR in this work).[17] To use GoldDigger, the installation of all auxiliary software is required. GoldDigger is executed by running a bash script which contains all the user-adjustable parameters (among others: gold bead size, raw data path, software paths, enabling/disabling of pre- and post-alignment step) and runs through all processing steps (Figure 1). Details about how to run GoldDigger can be found at https://github.com/CryoSTEM-tools/GoldDigger.

3.1.2. Detection and merging of different-sized gold beads

The detection of gold beads is performed using R4TR. This software uses Dynamo under MATLAB to robustly find gold beads of a given diameter. The gold beads are detected by means of cross-correlation using the gold bead diameter as an input parameter. The highest peaks of the cross-correlation matrix are selected as they are supposed to indicate the position of the gold beads. The peaks are then indexed to generate fiducial chains using several parameters to discard or retain the peaks (e.g., maximal shift allowed, minimum number of markers per tilt). It outputs an IMOD-compatible coordinate file containing the positions of the gold beads identified on each image of the tilt-series. This coordinate file can be inputted to IMOD to generate a fiducial model which is subsequently used to compute the alignment matrix. The GoldDigger strategy consists of looping the R4TR detection of the gold beads, increasing the given diameter at each step (Figure 1, step 1). This allows cross-correlating the tilt-series with templates of increasing diameters, which finds different sets of gold beads at each round. The multiple IMOD-type coordinate files are then transformed into text files using the *model2point* command from IMOD (Figure 1, step 2) and subsequently concatenated into a single text file (Figure 1, step 3). Because the detection of the gold beads is looped in the first step of GoldDigger, some gold beads can be detected multiple times. In MATLAB, the coordinates of all detected gold beads are compared to identify and

delete the duplicated ones (Figure 1, step 4). Step 4 of GoldDigger is also used to merge the gold bead chains which have been detected during each loop of R4TR. The purpose is to generate longer gold bead chains which are usually associated with better alignment quality. The output of step 4 is a new gold bead coordinate file, exempt of duplicates and containing longer chains, which can be inputted to *tiltalign* (IMOD command computing the transform matrix to align a tilt-series using the tilt-angles, the orientation of the tilt-axis and the gold bead coordinates) to compute the alignment of the tilt-series (Figure 1, step 5). Finally, a *newstack* command (IMOD command used to make modifications to a tilt-series, in our case, the modification consists of transforming the tilt-series images using a transformation matrix previously computed with *tiltalign*) is run to generate the aligned tilt-series (Figure 1, step 6). The final fiducial coordinate file can also be used to align the tilt-series using different software such as TomoAlign which can use the position of the gold beads throughout the tilt-series to measure and compensate the deformation of the sample during data collection.[28] Because motion deformation measurements in TomoAlign must be performed using a maximum of 100 fiducials, the number of fiducials in the final fiducial coordinate file outputted by GoldDigger is by default limited to 100. This is achieved by deleting the fiducials with the shortest fiducial chains as they are more prone to introduce bad fiducial positions during the filling of the fiducial model (the nogaps model generated by *tiltalign*).

3.1.3. Optional pre- and post-gold bead detection alignment steps

The preliminary optional step of our data processing pipeline is pre-alignment of the tilt-series with StackReg[26] (Figure 1, step 0). The additional pre-alignment step is most useful when the raw tilt series contains large spatial shifts between images, such that R4TR may fail to accurately track the detected gold beads. Pre-alignment of the tilt-series reduces the chance that R4TR will confuse one gold bead with a different bead in the next tilt, resulting in a poor overall alignment. We tried several registration methods (IMOD,[14] AreTomo,[29] and StackReg[26]) and found StackReg gave the best and more consistent results on the test data. The good performance of StackReg for cryo-STET datasets might come from the fact that StackReg uses a coarse-to-fine image registration sequence which is less prone to be trapped into a local optimum. The StackReg alignment should be performed in translation mode since rotations do not need to be computed at this stage. Because the StackReg image registration translates the image frames in relation to one another, the geometric transformation requires access to data beyond the image frame, and the final aligned result will include blank pixels at the border edges. The StackReg result is converted into an IMOD-type transform file. This pre-alignment step must be performed only when fiducial tracking fails on the raw tilt-series as it has a limitation. The pre-aligned tilt-series generated with *newstack* contains empty border areas corresponding to the applied shifts at the edge of the images. These empty border areas generate NaN errors in R4TR when Dynamo performs the cross-correlation between the tilt-series and the gold bead template. The pre-aligned tilt-series has to be cropped to discard the empty border areas, which reduces the area where gold beads can be detected and tracked. The cropping has to be performed manually. The final aligned tilt-series is generated using the original raw tilt-series as input, the StackReg and GoldDigger alignments being fused together using the *xfproduct* command (IMOD command merging two transformation matrices into a single one) to avoid introducing multiple interpolations of the data. Using this workflow, the aligned tilt-series has the same size as the raw tilt-series. The limitation of this strategy is then only present during the detection of the beads. This optional pre-alignment is indicated in the diagram (Figure 1, step 0).

The first step of GoldDigger can track a significant amount of gold beads and other features because of the wide diameter search range. On some rare occasions, gold beads or features detected by R4TR were wrongly tracked throughout the tilt-series generating large residuals in the *tiltalign* step (Figure 1, step 5). The large residuals can be ignored using the robust fitting option in IMOD. However, in TomoAlign the measurement of the sample deformation is based on the position of the gold beads. Because badly tracked fiducials will generate wrong deformations in TomoAlign, we added two optional polishing steps to get rid of the large residual fiducials (Figure 1, step 7). Each polishing step consists in removing the gold beads which residuals are not contained within a certain range around the median value of all residuals.
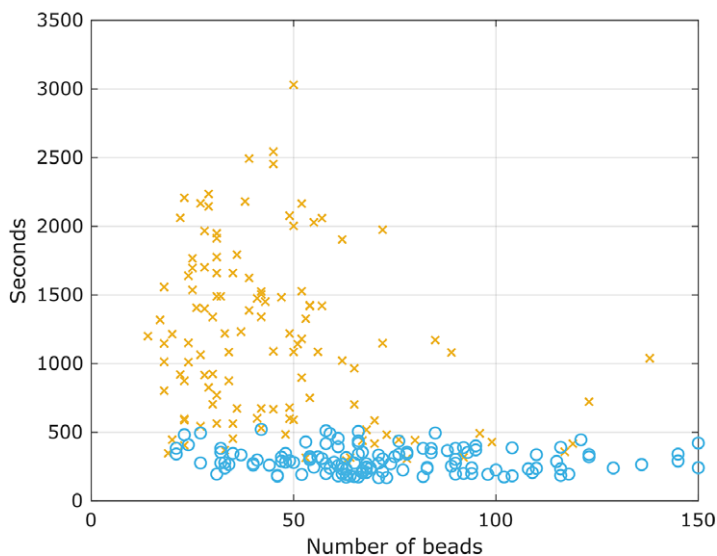
Figure 2. *GoldDigger (step 1) computation time. Time to complete step 1 of GoldDigger as a function of the number of detected gold beads. GoldDigger step 1 consists of 10 gold bead detection loops using R4TR. This plot shows the time taken by GoldDigger step 1 to detect and track gold beads using the default (orange crosses) and fast (blue circles) modes of the R4TR software.*

The range corresponds to a multiple of the standard deviation of all residuals and is user-adjustable to fit special requirements. The exclusion of big residuals can be performed in a single step if their residuals are not big, or in two steps if very large residuals are present. The purpose of the two filtering steps is to first exclude the fiducials with a very large residual and then discard the remaining fiducials with more classic residuals (Supplementary Figure S1). To allow this, the range of both filtering steps can be set separately using user-adjustable parameters. The filtering steps are optional and can be skipped. If only one pass of big residual removal is required, then the second filtering step can be skipped. All these parameters need to be set in the bash script.

3.1.4. Results on TS Ec tilt-series

To test the performance of GoldDigger, we used cryo-STET datasets of *E. coli* cells. We found that GoldDigger improved the tracking of fiducials on cryo-STET datasets, tracking more fiducials (about 50% more) while targeting the gold bead center more precisely, compared to R4TR designed for cryo-ET datasets.

Because GoldDigger step 1 performs several loops of gold bead detection, we first measured the computation time of GoldDigger. To measure the computation time, GoldDigger was run on the cryo-STET datasets using both *default* and *fast* R4TR settings (Figure 2). Surprisingly, the *default* settings tracked fewer gold beads than the *fast* settings. On average the *default* settings (Figure 2, orange crosses) tracked 45 gold beads per tilt-series (range of 14–138) and the *fast* settings (Figure 2, blue circles) tracked 72 gold beads per tilt-series (range of 21–150). The average computation time of the *default* settings to track the gold beads was 1175 s (295 s to 3030 s) and for the *fast* settings 296 s (169 s to 520 s). These timing measurements were performed on a server with 13 CPUs (Intel Xeon Gold 6230R @ 2.10 GHz), 225 GB DDR4 memory, and two Tesla T4 GPUs with 16 GB of memory each. One observation is that the computation time is not proportional to the number of detected and tracked gold beads for both *default* and *fast* settings. A possible explanation is that the computation time is dependent on how difficult it is to detect and track gold beads, because of low image contrast or thick areas.

Table 1. *The number of fiducials tracked on different tilt-series (TS Ec) using R4TR and GoldDigger methods*

|	R4TR *default*	GoldDigger *default* (step 3)	GoldDigger *default* (step 4)	R4TR *fast*	GoldDigger *fast* (step 3)	GoldDigger *fast* (step 4)
TS Ec001	95	839	181	122	1147	212
TS Ec002	38	407	108	93	833	151
TS Ec003	51	293	85	57	408	104
TS Ec004	17	307	143	107	737	259
TS Ec005	41	418	170	71	582	206
TS Ec006	18	427	179	120	779	293
TS Ec007	30	408	168	92	591	228

Note: The columns R4TR *default* and R4TR *fast* present the number of fiducials detected using each mode of R4TR. The columns GoldDigger *default/ fast* (step 3) and GoldDigger *default/fast* (step 4) present the number of fiducials detected using GoldDigger before and after merging the chains, respectively.

The combination of steps 1–4 of the GoldDigger workflow is important as they represent the novelty compared to previous fiducial-based alignment software. To characterize them, the number of tracked fiducials was measured (Table 1). On average, R4TR *default* tracked 41 (±27) gold beads, R4TR *fast* tracked 95 (±24) gold beads, GoldDigger *default* tracked 442 (±183) gold beads, and GoldDigger *fast* tracked 725 (±235) gold beads (Table 1, R4TR and GoldDigger step 3). Using the *default* settings, GoldDigger tracked more gold beads than R4TR, which was expected since GoldDigger computes 10 loops of R4TR. These values are valid until step 3 of the GoldDigger workflow as the gold bead chains are not merged yet. After merging, fewer gold beads remained (Table 1, GoldDigger step 4), as on average GoldDigger *default* tracked 148 (±38) gold beads and GoldDigger *fast* tracked 204 (±69) gold beads. Fewer gold beads were expected as longer chains were generated. It must be noted that after step 4, the fiducial file outputted by GoldDigger does not contain more than 100 fiducials as this is the limit of TomoAlign.

A final comparison has been performed, including now ClusterAlign,[18] which as previously mentioned was not released yet when the development of GoldDigger was initiated. In this comparison, the number of tracked fiducials per tilt-angle is plotted for each tilt-series (Supplementary Figure S2). As can be observed, the number of tracked fiducials per tilt-angle varies importantly when using R4TR (Supplementary Figure S2, left-hand side panel), as on some tilt-images only a few gold beads have been tracked, especially for the high-tilt ones but also for some images collected at low tilt. Using GoldDigger (Supplementary Figure S2, center panel), much more beads are tracked, more consistently and even the high-tilt images possess several gold beads tracked. Using GoldDigger, there is almost always more than 40 beads tracked per tilt-image. Using ClusterAlign (Supplementary Figure S2, right-hand side panel), the number of beads tracked is quite constant throughout the tilt-series, which must be a characteristic of using gold bead clusters instead of individual gold beads. However, ClusterAlign does not track many gold beads after ±60°. Compared to the developers of ClusterAlign, the tilt schemes of our data collection always go beyond ±60° and often up to ±70°. Our hypothesis explaining why ClusterAlign has difficulties tracking gold bead clusters at such high tilt-angles is that the component of the cluster vectors in the direction perpendicular to the tilt-axis become very small at such high angles, which prevents accurate tracking and identification of the clusters.

The lengths of the fiducial chains were measured to better characterize step 4 of the GoldDigger workflow (Figure 3). Fiducial chain lengths of all datasets are available in Supplementary Figure S3. As can be observed by the shift to the right of the data distribution curve and of the mean line for each dataset, GoldDigger generated longer fiducial chains after merging of the fiducial chains compared to the initial chains tracked by R4TR (Figure 3 and Supplementary Figure S3). GoldDigger fiducial chains were on average 54 and 53% longer than initial R4TR ones (Table 2, *default* and *fast* settings, respectively). The fiducial initially detected with R4TR were often tracked on less than half the images constituting the tilt-
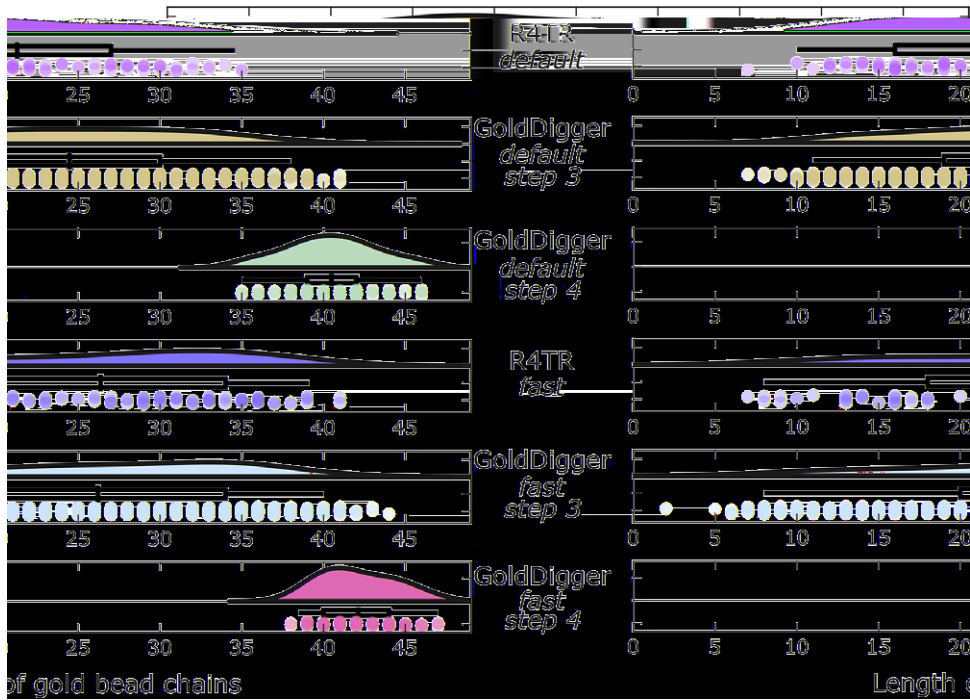
Figure 3. *Characterization of gold beads chains. Comparison of the gold bead chain lengths on tilt-series TS 001 using the different fiducial tracking methods: R4TR, GoldDigger step 3 (before merging of the fiducial chains) and step 4 (after merging of the fiducial chains). For each method, the default and fast modes are presented. Each plot consists of three rows. In the first row, the colored curve represents the distribution of the data. In the second row, the boxplot ranges from the position of the 25% quantile to that of the 75% quantile and the thick vertical line corresponds to the mean value of the data. Finally, the third row contains colored dots representing each data point, using the same color as the data distribution in the first row.*

series (Table 2, R4TR columns, red cells). As a reminder, each tilt-series was composed of between 41 and 45 images. This clearly demonstrates the difficulty in tracking gold beads in cryo-STET datasets and shows the amount of manual intervention one might need to perform to complete the fiducial tracking. Using the R4TR *fast* mode allowed to track fiducials on slightly more images than the R4TR *default* mode. After merging the fiducial chains, GoldDigger allowed to generate much longer chains which were often longer than 2/3 of the images of the tilt-series (Table 2, GoldDigger step 4 columns, green cells). These results show that the combination of GoldDigger steps 1, 3, and 4 allowed to tracked more fiducials, which once combined, were tracked on most images of the tilt-series. This allows alignment of cryo-STET tilt-series which otherwise would have needed manual intervention to pick additional fiducials, as it is the case for tilt-series TS Ec002, TS Ec005, and TS Ec007 for which very low number of fiducials were detected and tracked (Table 2 and Supplementary Figure S3). The following part of the alignment is carried out using standard methods (GoldDigger step 5, *tiltalign*).

To verify that merging the fiducial chains effectively performed better detection and tracking, the coordinates of the tracked fiducials were compared to the actual position of the corresponding gold beads. Only fiducials which were tracked by all methods were considered for the analysis to allow a side-by-side comparison of the methods. To present the tracking accuracy of each method, 30 fiducials were randomly selected per tilt-series (Supplementary Figure S4). The first eight fiducials of the cryo-STET dataset TS Ec001 are displayed below (Figure 4). From this small subset, it is visible that, even though not perfect, GoldDigger *default* and *fast* tracking methods (Figure 4, second and fourth rows, purple and green circles,
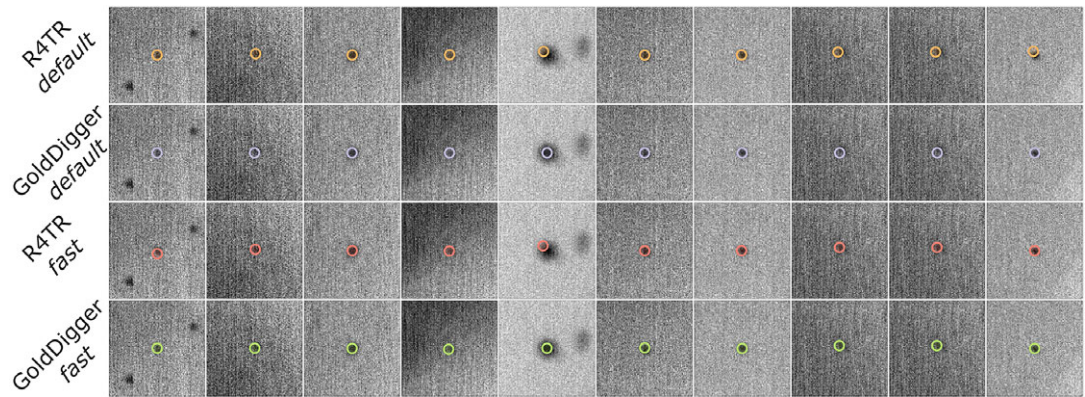
Figure 4. *Comparison of the coordinates of eight fiducials of TS 001 as tracked by R4TR and GoldDigger, and the actual position of the underlying gold bead. Visual comparison of the accuracy of the different fiducial tracking methods (from top to bottom): R4TR default (orange circles), GoldDigger default (purple circles), RATR fast (red circles), and GoldDigger fast (green circles). The columns represent different gold beads.*

**Table 2.** *Length of fiducial chains tracked using R4TR and GoldDigger methods*

	R4TR *default*	GoldDigger *default* (step 3)	GoldDigger *default* (step 4)	R4TR *fast*	GoldDigger *fast* (step 3)	GoldDigger *fast* (step 4)
TS Ec001 (49)	21	24	41	26	26	42
TS Ec002 (46)	16	19	29	21	21	35
TS Ec003 (45)	27	21	31	23	24	32
TS Ec004 (47)	25	23	33	33	29	43
TS Ec005 (43)	15	18	27	17	17	29
TS Ec006 (46)	22	20	32	22	21	36
TS Ec007 (47)	17	19	30	17	19	31

Note: The columns R4TR *default* and R4TR *fast* present the average length of fiducial chains detected using each mode of R4TR. The columns GoldDigger default/fast (before merge) and GoldDigger default/fast (after merge) present the average length of fiducial chains detected using GoldDigger at step 3 and step 4 of the workflow, respectively. The tilt-series number is indicated in the left-most column and the numbers between brackets represent the total number of tilt-angles. Cells are marked in red when on average the fiducials are tracked on less than half the images of the tilt-series. Cells are marked in green when on average the fiducials are tracked on more than 2/3 the images of the tilt-series.

respectively) are more accurate than the classic R4TR *default* and *fast* methods (Figure 4, first and third rows, orange and red circles, respectively), as the position of the fiducial more often corresponds to the actual center of the gold bead. The same assessment can be made after analyzing the positions of the 30 fiducials extracted from each tilt-series compared to the position of the underlying gold beads (Supplementary Figure S4).

3.1.5. Additional tests on more challenging data

The *E. coli* samples used so far were about 600 to 700 nm-thick cells, for which GoldDigger proved to be efficient, generating accurate picking of the gold beads. To test the robustness of GoldDigger even more, we applied it on *T. brucei* cryo-STET datasets which were collected for a previous study[2] (named TS Tb001 to TS Tb012 hereafter). These datasets previously had to be fully manually aligned because the alignment tools available at that time could not accurately pick and track the gold beads in such thick samples (up to 1.6 μm). Because of the cell thickness and the high luminosity and contrasts changes, these
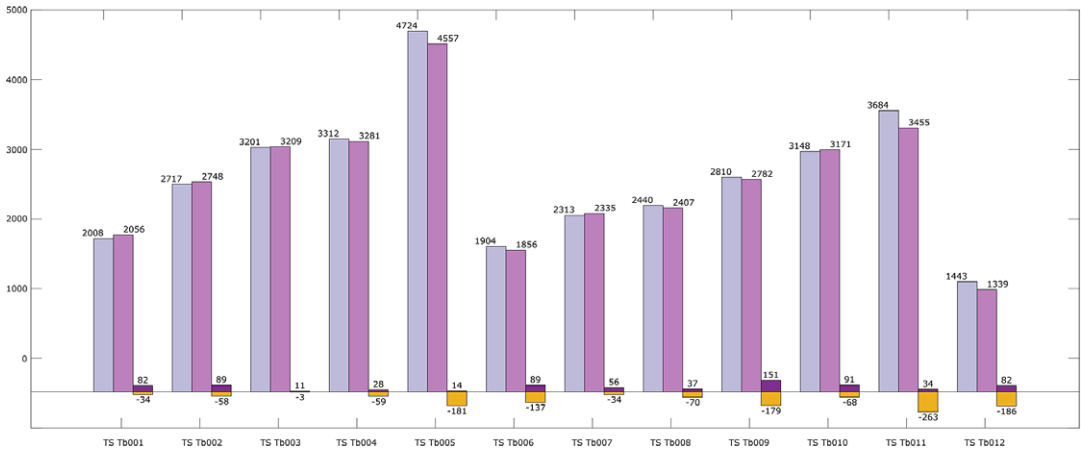
Figure 5. *Manual editing of the GoldDigger-detected gold bead coordinates on TS Tb tilt-series. For each TS Tb tilt-series, three histograms are plotted: (i) the number of gold beads picked by GoldDigger (gray), (ii) the number of gold beads after manual editing (pink), and (iii) a double histogram showing the number of added (purple) and deleted (yellow) gold beads. The number of added gold beads is reported in positive values whereas the number of deleted gold beads is displayed in negative values. The values next to each histogram represent the number of corresponding gold beads. Each tilt-series consists of about 70 images, meaning that 50 fully tracked gold beads would represent a number of 3500 picked gold beads.*

datasets are particularly challenging. In these tests, GoldDigger optional step 7 was systematically used as the set of gold beads detected in step 4 and used for alignment in step 5 contained too many outliers which compromised the alignment quality. Two rounds of removing outliers were then processed, which led to much better alignment. However, the alignment quality was still perfectible, as shown by the distorted shape of some gold beads in the 3D reconstruction (Supplementary Figure S5, dotted circles). For these datasets, we do not compare R4TR and GoldDigger as performed in the previous chapters as R4TR failed to pick enough gold beads to compute an alignment (data not shown).

After manual refinement of the gold bead coordinate file generated after step 7 of GoldDigger, we measured how many gold beads were manually deleted or added (Figure 5). For each TS Tb tilt-series, the number of gold beads picked in the GoldDigger-generated file, the number of gold beads after manual modification, and the number of added and deleted gold beads are reported. Each TS Tb tilt-series consists of about 70 images, meaning that 50 fully tracked gold beads would represent a number of 3500 picked gold beads. Values displayed in the first two columns can relate to that, considering that gold beads are rarely tracked on all images. The third column shows the number of added gold beads in positive values (Figure 5, purple histograms) and the number of deleted gold beads in negative values (Figure 5, yellow histograms). The number of added gold beads is often low, representing the equivalent of tracking one single gold bead throughout the whole tilt-series, except for TS Tb010 for which 151 new gold beads were added. Gold beads were added in particular at high tilts which often contained too few gold beads to accurately align the images. The number of deleted gold beads is noticeably higher than that of the added ones, however, deleting gold beads can be very quick as an entire gold bead chain can be deleted in a single click. Gold beads were deleted when their position was off, and entire gold bead tracks were deleted when the track path was obviously wrong. Compared to the total number of gold beads picked, the amount of manual work required to add and delete gold beads is significantly lower. This means that even on challenging tilt-series, GoldDigger can be used to generate a gold bead coordinate file which can be improved with minimal manual work, taking away the lengthy fully manual gold bead picking.

To better demonstrate why manual refinement was necessary, 3D reconstruction slices were extracted from a couple of TS Tb datasets (Supplementary Figure S5). The strong reconstruction artifacts present in the 3D reconstructions computed with the raw GoldDigger fiducial coordinates (Supplementary Figure S5A,C,

dotted circles) are better observed using a combination of XY, XZ, and YZ slices. After manual correction of the fiducial coordinates, the reconstruction artifacts are much less present if not totally absent (Supplementary Figure S5B,D).

### 3.2. Checkers, an image-splitting tool to enable Noise2Noise denoising methods in cryo-STET

3.2.1. Rationale and software overview

The design for a software such as Checkers comes from the lack of denoising tools for cryo-STET 3D reconstructions such as the deep-learning denoising software found for cryo-ET. Basic filtering can be performed using median or Gaussian 3D filters, Fourier band-pass filters, or more advanced non-anisotropic diffusion and edge-enhancing denoising filters,[30] but they do not achieve the same denoising quality as what can be achieved with CryoCARE.[19] The main obstacle preventing from using Cryo-CARE on cryo-STET data is the absence of image pairs as available in cryo-ET. Generating the image pairs in cryo-STET is far from being trivial. First, because the data collection of a STEM image takes much longer than that of a transmission electron microscopy (TEM) image, the eventuality of a sample drift during the collection of the image pair would generate artifacts when processed with CryoCARE because the images of the pair would represent slightly different fields of view of the sample. The image pair could be aligned as it is performed with MotionCor2 in cryo-EM but this has some limitations as collecting two cryo-STEM images with reduced electron dose to prevent beam damage would generate low signal-to-noise ratio images, which will be deleterious for the whole downstream processing of the data. Last, an image pair might be created by splitting the dataset by odd-even tilt angles before reconstruction. However, this method has been shown to produce less fine detail and more reconstruction artifacts, both when compared to splitting by movie frames in cryo-ET (Tim-Oliver Buchholtz, Thesis, Figure 2.7)[31] and when compared to splitting by pixels in cryo-STET.[32] Because state-of-the-art denoising software do not exist for cryo-STET data, it creates an important bottleneck in the processing pipeline, reducing the quality and interpretability of 3D reconstructions, and preventing automatic segmentations, which in turn limits our understanding of the structural data.

The novelty of Checkers relies on the creation of an image pair from a single unique image. This is performed by splitting the image into two halves (even/odd pixel split) and then reconstructing the missing pixels in each half using inpainting. This strategy has been designed based on previous expertise with the collection of sparse images which are incomplete and can be reconstructed using different types of algorithms.[24] The computing steps performed in Checkers have been summarized in a diagram (Figure 6). As for GoldDigger, Checkers was developed on Linux and it uses a main bash script which calls four software (MATLAB, IMOD,[14] Tomo3D,[33] and CryoCARE[19]). To use Checkers, the installation of auxiliary software is compulsory. Checkers is executed by running a bash script (in the future it will also be present as a Python/Notebook script) and only requires a few input parameters from user parameters (final size of the 3D reconstruction, number of inpainting iterations, data path). The internal parameters of Checkers have been set to provide efficient denoising as tested on different data during development. Details about how to run Checkers can be found at https://github.com/CryoSTEM-tools/Checkers.

3.2.2. From image to dual-frame movie

The input data of Checkers is an aligned tilt-series. Even though it has not been tested what the effect would be between using Checkers on unaligned data compared to aligned data, using aligned data makes the use of Checkers smoother. It allows the user to use the alignment software of their choice prior to using Checkers. The only requirement is to provide the input data in either *mrc* or *tif* format. Any other data format should first be converted into *mrc* or *tif*. To better keep track of the different computing steps performed in Checkers, a supplementary figure has been assembled to show the image content during the first two steps, which correspond to the novelty introduced with this software (Supplementary Figure S6). In MATLAB, each individual image of the tilt-series (Supplementary Figure S6, input) is split into two, pixel-wise (Figure 6, step 1). The pixels with even coordinates (2–2, 2–4, …,4–2, 4–4, etc.) are used to
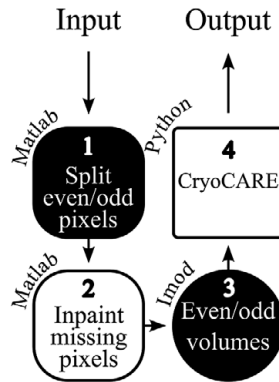
Figure 6. *Diagram of Checkers workflow. The whole workflow is performed by a bash script which calls several other software, indicated by their name in the diagram. The input data is the aligned tilt-series. The even/odd pixel split is performed in MATLAB and generates a tilt-series pair. Subsequent inpainting of the paired datasets is performed in MATLAB. Each inpainted dataset is then 3D reconstructed creating a pair of volumes. The volume pair is then inputted in CryoCARE which will extract the training dataset, train the network, and then denoise the data. Checkers output is a single volume with enhanced contrast and reduced noise.*

create a first new image, while the pixels with odd coordinates (1–1, 1–3, …, 3–1, 3–3, etc.) are used to generate a second new image (Supplementary Figure S6, split even/odd). Note that when splitting the pixels, only half of the pixels are used as one-fourth of the pixels are kept per even or odd image. During testing of the software, using all the pixels, i.e. keeping half of them per even or odd image produced inferior denoising of the data. The reason for this might be that using all the pixels ends up in generating even/odd pair images with such minor differences that efficient denoising with CryoCARE is not achieved. However, we did not investigate this effect further. After splitting the data, inpainting using discrete cosine transforms[34,35] is applied to reconstruct the missing pixels (Figure 6, step 2). The gaps introduced by the pixel splitting are filled after inpainting (Supplementary Figure S6, inpainting even/odd). The inpainting procedure performs smoothing of the data, the smoothing factor being iteratively refined when values are missing, as in sparse images. After processing all images of the tilt-series, a pair of new even/odd tilt-series is created and saved as *tif* files. Two 3D reconstructions are reconstructed using weighted back-projection (WBP) and CTF deconvolution in Tomo3D[33,36,37] (Figure 6, step 3). Alternatively, 3D reconstructions can be computed using IMOD, but doing so might result in different denoising results. After this step, the reconstructions are cropped to the dimensions defined by the user. This step is important as the reconstructions should not contain too much bright areas (e.g., empty Quantifoil hole) or too dark areas (e.g., electron-opaque structures) as this will prevent Cryo-CARE from performing accurate denoising. The final step consists of running the whole CryoCARE classic workflow (Figure 6, step 4). The CryoCARE parameters used in Checkers are the same as those found in version 0.3 of CryoCARE. The output of Checkers is the denoised volume with enhanced contrast.

3.2.3. Results

To assess the denoising and contrast improvement of Checkers, it has been tested on a *T. brucei* cryo-STET dataset TS Tb002. Information about data collection parameters and sub-cellular structures observed by cryo-STET in entire *T. brucei* cells have been detailed in a previous article.[2] In this previous article, the 3D reconstructions were computed using WBP and then filtered using edge-enhancing noise-reduction anisotropic diffusion filter (EED).[30] In the present work, several reconstruction and filtering methods are compared, including the one used in the original cryo-STET work[2] (Figure 7). Wider fields
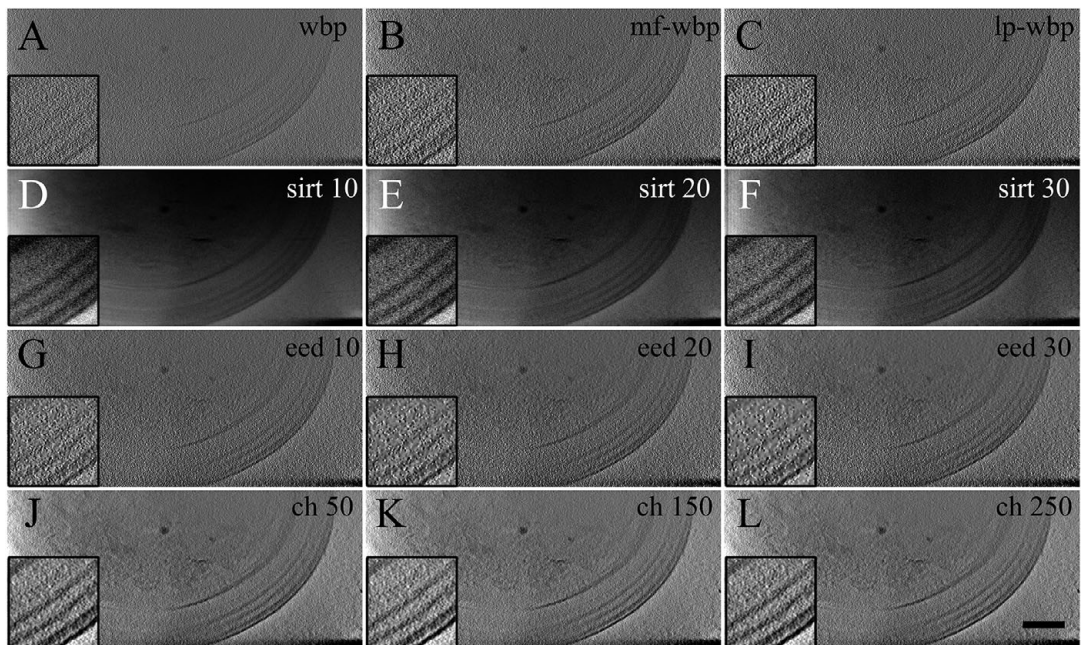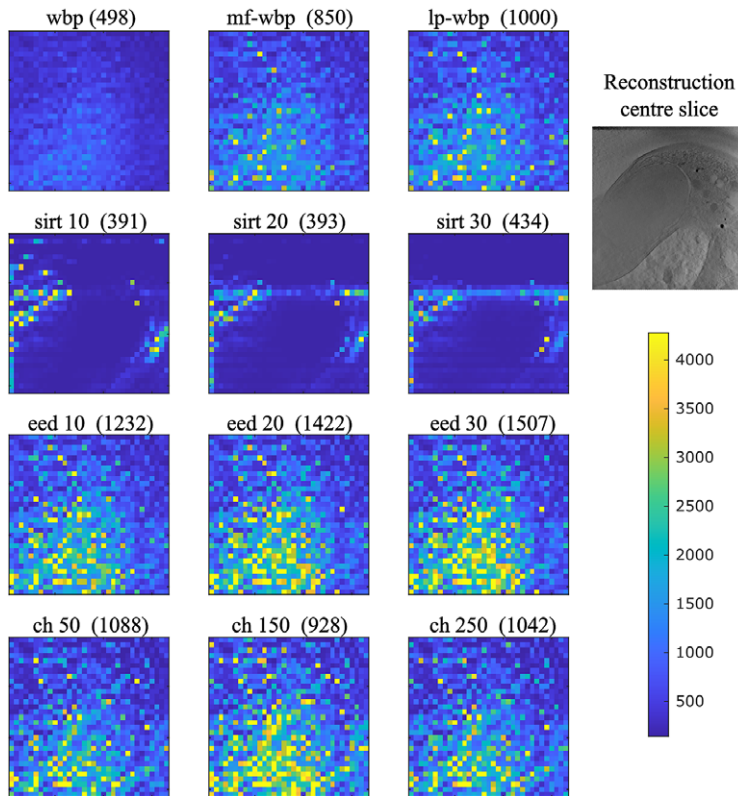
Figure 7. Comparing the quality of data computed with various 3D reconstruction and filtering methods. The images show a single Z-slice (2 nm-thick) extracted from the 3D reconstructions (TS Tb002). More information about this cryo-STET dataset of a whole T. brucei cell can be found in a previous work.[2] (a) WBP reconstruction. (b) Median-filtered (3px wide, at the level of the tilt-series) WBP reconstruction. (c) Low-pass filtered (radius 0.15, sigma 0.05, at the level of the tilt-series) WBP reconstruction. (d) 10 SIRT iterations reconstruction. (e) 20 SIRT iterations reconstruction. (f) 30 SIRT iterations reconstruction. (g) WBP reconstruction filtered with 10 EED iterations. (h) WBP reconstruction filtered with 20 EED iterations. (i) WBP reconstruction filtered with 30 EED iterations. (j) Checkers reconstruction with 50 inpainting iterations. (k) Checkers reconstruction with 150 inpainting iterations. (l) Checkers reconstruction with 250 inpainting iterations. Scale bar is 400 nm. Note that the images presented in this figure are flipped compared to images of the same dataset presented in Figure 8 because of software having different image coordinate systems (top-left corner VS bottom-left corner).

of view of the reconstructions are available in Supplementary Figure S7. All reconstructions were computed in Tomo3D (v2.2).

The reconstructions show different visual contrast and texture: some structures are barely visible in some reconstructions, and strikingly visible in others. During WBP, CTF deconvolution was applied in Tomo3D using default parameters. The classic WBP is as expected the reconstruction which looks the noisiest (Figure 7a). The WBP volume was further filtered using median filtering (IMOD *clip* command 3 px) and low-pass filtering (IMOD *mtfilter* command radius 0.15, sigma 0.05). These two filters slightly removed the noise of the WBP reconstruction, increasing the contrast and allowing better visualization of microtubules (Figure 7b,c).

Reconstructions using the simultaneous iterative reconstruction technique (SIRT) were also computed, using 10, 20, and 30 iterations (Figure 7d–f). In the SIRT reconstructions, the noise is less visible, but the pixels inside the reconstructions have a significant intensity ramp, which is dependent on the sample thickness. At the center of the *T. brucei* cell, the voxels are particularly dark, whereas around the cell where the ice is thinner, the voxels are brighter. This is a classic problem in SIRT reconstructions when sample thickness varies significantly. This ramp hampers the analysis of the 3D reconstructions,

Figure 8. *Comparing the signal-to-noise ratio (SNR) of data computed with various 3D reconstruction and filtering methods. The heat-map images show the SNR computed on single Z-slices (2 nm-thick) located near the center of the 3D reconstructions. The color-coded SNR value is displayed for each 50 × 50-pixel patch constituting the full image. Next to each reconstruction method, the value given in between brackets corresponds to the average SNR value of all 50 × 50-pixel patches of the entire volume. From left to right and top to bottom the various methods are WBP, median-filtered WBP (3px wide, at the level of the tilt-series), low-pass filtered WBP (radius 0.15, sigma 0.05, at the level of the tilt-series), 10 SIRT iterations, 20 SIRT iterations, 30 SIRT iterations, WBP filtered with 10 EED iterations, WBP filtered with 20 EED iterations, WBP filtered with 30 EED iterations, Checkers with 50 inpainting iterations, Checkers with 150 inpainting iterations and Checkers with 250 inpainting iterations. The top-right gray-scale image is the central Z-slice of a Checkers reconstruction and shows the area on which SNR values were computed.*

especially of the intracellular elements. When more SIRT iterations are computed the densities are less blurry, as can be seen by the increased visibility of the microtubule lumen (Figure 7d–f, inserts).

Three EED reconstructions were also computed using 10, 20, and 30 iterations (other parameters were set to default) for comparison purposes as they represent what was best achieved in 2020[2] (Figure 7g–i, respectively). The EED reconstructions are less noisy than the median or low-pass filtered WBP ones and they do not possess the ramp of the SIRT reconstructions. When more EED iterations were computed, the cellular structures became darker but bright pixels were also generated (Figure 7i, insert).

Finally, three Checkers reconstructions were computed using 50, 150, and 250 iterations of inpainting procedure (Figure 7j–l, respectively). Varying the number of inpainting iterations was to test the effect of the inpainting quality on that of the final 3D reconstruction, but little effect was observed. The Checkers reconstructions outperform any other type of reconstruction. They have much less noise, the cellular structures are much better defined and there is no ramp effect throughout the reconstruction.

In addition to analyzing the visual quality of the different reconstructions, numerical image quality descriptors were also computed to quantify the quality of the reconstructions. In the present study, there was no ground-truth reconstruction which could be used as a reference, thus image quality descriptors such as structure similarity index measurement or peak signal-to-noise ratio (SNR) could not be used. Instead, the SNR was computed on single images/reconstructions using a method[38] previously used in the analysis of electron microscopy images. Instead of computing a single SNR value for the whole image as it is usually performed, it was measured on 50 × 50-pixel patches to better visualize the effect of each reconstruction and filtering method on the cellular structure themselves (Figure 8). As previously mentioned, the classical WBP reconstruction is noisy and this is confirmed by the low SNR values (Figure 8, first row). Median and low-pass filtering help to slightly increase the SNR values, but they still remain quite low. The SIRT reconstructions have strong contrast at the edge of the *T. brucei* cell, but inside the cell the contrast is very low (Figure 8, second row). This matches the dark areas previously observed inside the cell where cellular structures were barely visible (Figure 7d–f). The three EED reconstructions have high SNRs which increase with the number of computed iterations (Figure 8, third row). These high SNRs probably originate from the aberrant bright pixels generated during filtering, which presence was increased by the number of EED iterations (Figure 7g–i). The three Checkers reconstructions also have high SNR (Figure 8, fourth row), in agreement with the visual high quality of the images (Figure 7j–l). Varying the number of inpainting iterations had an effect on the SNR values, with the reconstruction computed with 150 iterations being the one with the highest SNR values. This analysis confirms that denoising with Checkers allows to substantially increase the SNR of 3D cryo-STET reconstructions and outperforming other existing methods, and greatly improving the visibility of the different cellular structures at the edge of the cells as well as deep inside them where the sample is the thickest.

4. Discussion and conclusion

In this work, we present two novel tools to process cryo-STET datasets and demonstrate their performance on experimental data. These tools, GoldDigger and Checkers, tackle two crucial steps of the tomography workflow. GoldDigger addresses the specific issue of insufficient tracked fiducials before tilt-series alignment. GoldDigger identifies more fiducial chains of greater length compared to algorithms initially developed for the analysis of cryo-ET data. This allows us to align cryo-STET datasets which would otherwise have been discarded, or painstakingly manually aligned. This means GoldDigger is an important tool enabling the automated bulk data processing pipelines required for high-throughput cryo-STEM tomography.

Checkers is a denoising tool which combines a data-splitting and inpainting strategy with state-of-the-art neural networks for unsupervised image denoising. Existing state-of-the-art unsupervised denoising for conventional cryo-ET cannot be applied directly to cryo-STET datasets, because cryo-STET does not easily produce independent noisy image pairs. Checkers uses a pixel-based splitting and inpainting strategy to generate paired noisy images which are then suitable for unsupervised Noise2Noise denoising. The denoising significantly improves the visibility of cellular components, interpretability of the 3D reconstructions, and understanding of the sample structure. We expect that high-quality 3D reconstructions will facilitate and extend the use of automatic segmentation algorithms on cryo-STET data in future work.

References

1. Rez P, Larsen T and Elbaum M (2016) Exploring the theoretical basis and limitations of cryo-STEM tomography for thick biological specimens. *Journal of Structural Biology* **196**, 466–478.
2. Trépout S (2020) In situ structural analysis of the flagellum attachment zone in Trypanosoma brucei using cryo-scanning transmission electron tomography. *Journal of Structural Biology: X* **4**, 100033.
3. Aoyama K, Takagi T, Hirase A and Miyazawa A (2008) STEM tomography for thick biological specimens. *Ultramicroscopy* **109**, 70–80.
4. Wolf SG, Mutsafi Y, Dadosh T, Ilani T, Lansky Z, Horowitz B, Rubin S, Elbaum M and Fass D (2017) 3D visualization of mitochondrial solid-phase calcium stores in whole cells. *elife* **6**, e29929.
5. Kirchenbuechler D, Mutsafi Y, Horowitz B, Levin-Zaidman S, Fass D, Wolf SG and Elbaum M (2015) Cryo-STEM tomography of intact vitrified fibroblasts. *AIMS Biophysics* **2**, 259–273.
6. Wolf SG, Houben L and Elbaum M (2014) Cryo-scanning transmission electron tomography of vitrified cells. *Nature Methods* **11**, 423–428.
7. Pennycook SJ (2011) A scan through the history of STEM. In Pennycook SJ and Nellist PD (eds.), *Scanning Transmission Electron Microscopy*. New York: Springer, pp. 1–90. http://doi.org/10.1007/978-1-4419-7200-2_1.
8. Nellist PD (2011) The principles of STEM imaging. In Pennycook SJ and Nellist PD (eds.), *Scanning Transmission Electron Microscopy*. New York: Springer, pp. 91–115. http://doi.org/10.1007/978-1-4419-7200-2_2.
9. Hohmann-Marriott MF, Sousa AA, Azari AA, Glushakova S, Zhang G, Zimmerberg J and Leapman RD (2009) Nanoscale 3D cellular imaging by axial scanning transmission electron tomography. *Nature Methods* **6**, 729–731.
10. Walther P, Bauer A, Wenske N, Catanese A, Garrido D and Schneider M (2018) STEM tomography of high-pressure frozen and freeze-substituted cells: A comparison of image stacks obtained at 200 kV or 300 kV. *Histochemistry and Cell Biology* **150**, 545–556.
11. Seifer S, Houben L and Elbaum M (2021) Flexible STEM with simultaneous phase and depth contrast. *Microscopy and Microanalysis* **27**, 1476–1487.
12. Lazić I, Wirix M, Leidl ML, de Haas F, Mann D, Beckers M, Pechnikova EV, Müller-Caspary K, Egoavil R, Bosch EGT and Sachse C (2022) Single-particle cryo-EM structures from iDPC–STEM at near-atomic resolution. *Nature Methods* **19**, 1126. http://doi.org/10.1038/s41592-022-01586-0.
13. Li X, Lazić I, Huang X, Wirix M, Wang L, Deng Y, Niu T, Wu D, Yu L and Sun F (2022) Imaging biological samples by integrated differential phase contrast (iDPC) STEM technique. *Journal of Structural Biology* **214**, 107837.
14. Kremer JR, Mastronarde DN and McIntosh JR (1996) Computer visualization of three-dimensional image data using IMOD. *Journal of Structural Biology* **116**, 71–76.
15. Mastronarde DN and Held SR (2017) Automated tilt series alignment and tomographic reconstruction in IMOD. *Journal of Structural Biology* **197**, 102–113.
16. Castaño-Díez D, Kudryashev M, Arheit M and Stahlberg H (2012) Dynamo: A flexible, user-friendly development tool for subtomogram averaging of cryo-EM data in high-performance computing environments. *Journal of Structural Biology* **178**, 139–151.
17. Pyle E (2023) relion4_tomo_robot.
18. Seifer S and Elbaum M (2022) ClusterAlign: A fiducial tracking and tilt series alignment tool for thick sample tomography. *Biological Imaging* **2**, e7.
19. Buchholz T-O, Jordan M, Pigino G and Jug F Cryo-CARE: Content-aware image restoration for cryo-transmission electron microscopy data, 5.
20. Lehtinen J, Munkberg J, Hasselgren J, Laine S, Karras T, Aittala M and Aila T (2018) Noise2Noise: Learning image restoration without clean data. In *Proceedings of the 35th International Conference on Machine Learning*. Stockholm, Sweden: PMLR, pp. 2965–2974.
21. Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng Y and Agard DA (2017) MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods* **14**, 331–332.
22. Sang X and LeBeau JM (2014) Revolving scanning transmission electron microscopy: Correcting sample drift distortion without prior knowledge. *Ultramicroscopy* **138**, 28–35.
23. Cossa A, Trépout S, Wien F, Groen J, Le Brun E, Turbant F, Besse L, Pereiro E and Arluison V (2022) Cryo soft X-ray tomography to explore Escherichia coli nucleoid remodeling by Hfq master regulator. *Journal of Structural Biology* **214**, 107912.

24. Trépout S (2019) Tomographic collection of block-based sparse STEM images: Practical implementation and impact on the quality of the 3D reconstructed volume. *Materials* **12**, 2281.

25. Biskupek J, Leschner J, Walther P and Kaiser U (2010) Optimization of STEM tomography acquisition — A comparison of convergent beam and parallel beam STEM tomography. *Ultramicroscopy* **110**, 1231–1237.

26. Thévenaz P, Ruttimann UE and Unser M (1998) A pyramid approach to subpixel registration based on intensity. *IEEE Transactions on Image Processing* **7**, 27–41.

27. Schneider CA, Rasband WS and Eliceiri KW (2012) NIH image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671–675.

28. Fernandez J-J, Li S, Bharat TAM and Agard DA (2018) Cryo-tomography tilt-series alignment with consideration of the beam-induced sample motion. *Journal of Structural Biology* **202**, 200–209.

29. Zheng S, Wolff G, Greenan G, Chen Z, Faas FGA, Bárcena M, Koster AJ, Cheng Y and Agard DA (2022) AreTomo: An integrated software package for automated marker-free, motion-corrected cryo-electron tomographic alignment and reconstruction. *Journal of Structural Biology: X* **6**, 100068.

30. Moreno JJ, Martínez-Sánchez A, Martínez JA, Garzón EM and Fernández JJ (2018) TomoEED: Fast edge-enhancing denoising of tomographic volumes. *Bioinformatics* **34**, 3776–3778.

31. Buchholz T-O (2022) Content-aware image restoration techniques without ground truth and novel ideas to image reconstruction.

32. Donati L, Nilchian M, Trépout S, Messaoudi C, Marco S, Unser M (2017) Compressed sensing for STEM tomography. *Ultramicroscopy* **179**, 47–56.

33. Agulleiro J-I and Fernandez J-J (2015) Tomo3D 2.0 – Exploitation of advanced vector eXtensions (AVX) for 3D reconstruction. *Journal of Structural Biology* **189**, 147–152.

34. Garcia D (2010) Robust smoothing of gridded data in one and higher dimensions with missing values. *Computational Statistics and Data Analysis* **54**, 1167–1178.

35. Wang G, Garcia D, Liu Y, de Jeu R and Johannes Dolman A (2012) A three-dimensional gap filling method for large geophysical datasets: Application to global satellite soil moisture observations. *Environmental Modelling & Software* **30**, 139–142.

36. Agulleiro JI and Fernandez JJ (2011) Fast tomographic reconstruction on multicore computers. *Bioinformatics* **27**, 582–583.

37. Tegunov D and Cramer P (2019) Real-time cryo-electron microscopy data preprocessing with warp. *Nature Methods* **16**, 1146–1152.

38. Thong JTL, Sim KS and Phang JCH (2006) Single-image signal-to-noise ratio estimation. *Scanning* **23**, 328–336.

39. Goscinski WJ, McIntosh P, Felzmann U, Maksimenko A, Hall CJ, Gureyev T, Thompson D, Janke A, Galloway G, Killeen NE, Raniga P, Kaluza O, Ng A, Poudel G, Barnes DG, Nguyen T, Bonnington P, Egan GF (2014) The multi-modal Australian ScienceS imaging and visualization environment (MASSIVE) high performance computing infrastructure: Applications in neuroscience and neuroinformatics research. *Frontiers in Neuroinformatics* **8**, 30.