

GPU Accelerated 3D Tomographic Reconstruction and Visualization from Noisy Electron Microscopy Tilt-Series

Julio Rey Ramirez, Peter Rautek, Ciril Bohak, Ondřej Strnad, Zheyuan Zhang, Sai Li, Ivan Viola, and Wolfgang Heidrich, *Fellow, IEEE*,

Abstract—We present a novel framework for 3D tomographic reconstruction and visualization of tomograms from noisy electron microscopy tilt-series. Our technique takes as an input aligned tilt-series from cryogenic electron microscopy and creates denoised 3D tomograms using a proximal jointly-optimized approach that iteratively performs reconstruction and denoising, relieving the users of the need to select appropriate denoising algorithms in the pre-reconstruction or post-reconstruction steps. The whole process is accelerated by exploiting parallelism on modern GPUs, and the results can be visualized immediately after the reconstruction using volume rendering tools incorporated in the framework. We show that our technique can be used with multiple combinations of reconstruction algorithms and regularizers, thanks to the flexibility provided by proximal algorithms. Additionally, the reconstruction framework is open-source and can be easily extended with additional reconstruction and denoising methods. Furthermore, our approach enables visualization of reconstruction error throughout the iterative process within the reconstructed tomogram and on projection planes of the input tilt-series. We evaluate our approach in comparison with state-of-the-art approaches and additionally show how our error visualization can be used for reconstruction evaluation.

Index Terms—tomographic reconstruction, electron tomography, tilt-series, visualization, cryo-ET, GPU acceleration

1 INTRODUCTION

TOGRAPHIC reconstruction is a process of transforming a series of tilted projection images (tilt-series) into a 3D volumetric representation (tomogram), where the information is stored in voxels organized in a regular grid. The problem was first addressed in the mathematical field by Johann Radon over a century ago. It was later put into practice in the 1930s as focal plane tomography using X-ray imaging in medicine, and many other modalities such as magnetic resonance imaging (MRI) [1], positron emission tomography [2], [3], functional MRI [4] and others within the medical domain and broader. The modalities mentioned above usually produce a series of uniformly distributed projections with small amounts of noise thanks to the characteristics of the imaging techniques and the objects under reconstruction. However, this is not true for specific domains where limitations are imposed in the imaging technique to avoid destroying the sample, as it is the case for high-energy electron microscopy (EM) techniques applied to fragile biological samples. For instance, cryo-electron tomography (cryo-ET) is capable of revealing the viral and cellular structures at molecular, or even atomic details [5], [6]. To avoid damaging the sample during the image acquisition

in these techniques, the energy of the electron beam and the number of projections is limited. Due to these limitations, projection images created in electron microscopy have a low signal-to-noise ratio, a limited number of projections, and a limited maximum tilt angle.

The tomographic reconstruction problem is addressed using a variety of algorithms, such as Fourier-domain reconstruction [7], back-projection reconstruction [8], fan-beam reconstruction [9], algebraic reconstruction techniques [10], [11], [12], and recently deep learning [13], [14]. While the first three approaches usually work well when we have a large number of projections with relatively low noise levels, they produce poor results when there is a limited number of projections or under high levels of noise, as is the case for electron microscopy. Algebraic reconstruction techniques provide better results under these conditions, but they require several iterations making them computationally more expensive.

There have been distinct approaches to improve the reconstruction quality for limited angle projection data, including deep-learning-based reconstruction methods [13] or traditional reconstruction techniques (e.g. back-projection reconstruction) coupled with a deep learning denoising approach. Here, denoising is applied either on the tilt-series before the reconstruction [15], [16], [17], or after the reconstruction on tomograms created with traditional techniques [18], [19]. One difficulty with such complex image processing pipelines is that the final reconstructions cannot easily be traced back to the individual raw input data to obtain an estimate of the reconstruction error or uncertainty. This makes it difficult to assess how reliable and trustworthy the reconstructions are, given that cryo-ET operates close

- Julio Rey Ramirez, Peter Rautek, Ciril Bohak, Ondřej Strnad, Ivan Viola, and Wolfgang Heidrich are with the Visual Computing Center at King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Kingdom of Saudi Arabia. E-mails: {julio.reyramirez | peter.rautek | ciril.bohak | ondrej.strnad | ivan.viola | wolfgang.heidrich}@kaust.edu.sa.
- Zheyuan Zhang and Sai Li are with the School of Life Sciences at Tsinghua University, Haidian District, Beijing, 100084, P.R. China. Emails: {zheyuanzhang | sai}@tsinghua.edu.cn.

Manuscript received April 19, 2005; revised August 26, 2015.

to the physical limits and there are usually no alternative imaging methods to validate the results.

In this paper, we present a GPU accelerated iterative tomographic reconstruction approach that integrates denoising in the reconstruction process by alternating between minimizing the re-projection error with algebraic reconstruction techniques and decreasing the noise using various regularization and denoising methods. The reconstructed tomograms are much cleaner than one could obtain using the existing approaches, removing the need to denoise the tilt-series prior to the reconstruction or denoise the reconstructed volumes. More importantly, the achieved reconstructions are denoised and, at the same time, directly tied to the original projection data in a way that enables a direct consistency check between the reconstruction and the raw input projection images. Once reconstructed, the tomograms can be visualized in 3D with volume rendering tools included in the framework. Moreover, we can visualize the error distribution and magnitude during the reconstruction process on individual projection planes or within the reconstructed tomogram.

To summarize, the main contributions of this paper are:

- a novel GPU accelerated technique for cryogenic electron tomography which iteratively reconstructs and denoises the tomograms, achieving excellent denoising results and sharp details in a matter of minutes;
- our approach is based on a flexible and robust optimization framework that integrates reconstruction and denoising, removing the necessity of pre-reconstruction or post-reconstruction denoising and maintaining the relationship between the original tilt-series and the reconstructed tomogram;
- a method for visualizing the reconstruction error throughout the reconstruction, either on individual projections or within the reconstructed volume.

In Section 2 we present how our approach relates to the existing research. In Section 3 we present our method and implementation details followed by the results and evaluation in Section 4. In Section 5 we present the conclusions and possible future extensions and adaptations of the presented method.

2 RELATED WORK

There are two well-known families of techniques for tomographic reconstruction. The first, *direct methods*, encompasses methods that exploit the Fourier slice theorem [20]. An overview of these methods can be found in a book by Hsieh [21]. The second, *iterative methods*, is a family of iterative algorithms that aim to solve tomographic reconstruction as a system of linear equations.

One of the first techniques from the direct methods is Weighted Back-projection (WBP) [22], with later adaptations [23], [24], where the reconstruction includes two main steps. First, the projections are filtered (typically in the Fourier domain) to avoid over-smoothed reconstructions due to the imbalance in the available low and high-frequency information. Second, the filtered projections are

back-projected into the reconstruction using the adjoint operator of the Radon transform [21].

The iterative reconstruction methods begin with the Algebraic Reconstruction Technique (ART) presented by Gordon *et al.* [10] and later adaptations the Simultaneous Iterative Reconstruction Technique (SIRT) [11] and Simultaneous Algebraic Reconstruction Technique (SART) [12]. In ART methods, the object is reconstructed into a discrete grid by iteratively projecting the current volume estimate, creating corrections based on the difference between the original projections and re-projections of the current volume estimate, and then back-projecting the corrections into the volume.

It has been shown that iterative methods outperform the FBP methods in the case of limited projections [25]. Moreover, the noise is handled better by algebraic methods [8]. These are also the reasons we implement our method as an iterative method which improves the reconstruction results by minimizing the error term. For an overview of more 3D reconstruction methods for multiple EM modalities, see work by Sorzano *et al.* [26].

As Pyle and Zanetti [27] already show, there are many software packages and toolboxes available for tomographic reconstruction such as IMOD [28], Air II - a Matlab toolbox [29], ASTRA Toolbox [30], [31] and its Python wrapper TomoPy [32], and AuTom [33] which are integrated in many microscopy pipelines. IMOD implements the FBP and SIRT approaches, Air II, and AuTom implement ART approaches, and ASTRA Toolbox offers FBP, SIRT, SART, and other reconstruction approaches. Most of these packages offer CPU as well as GPU implementations. Our approach solves tomographic reconstruction as an optimization problem, using SART or SIRT together with various regularizers to integrate denoising in the reconstruction process, and it is implemented for execution on GPUs.

In the specific case of cryo-ET on biological structures, the measured projections are very noisy. The reason for this is the destructive nature of the image acquisition process, where the specimen gets destroyed if exposed to the electron beam with higher energy. To avoid the sample damage due to the electron radiation, the image acquisition is performed using lower electron doses, resulting in a lower signal-to-noise ratio in the output tilt-series. A good overview of the current cryo-ET reconstruction approaches is presented by Donati [34] and the current strategies on data processing for cryo-ET as well as for subtomogram averaging are presented by Pyle and Zanetti [27]. The authors give a partial but extensive list of software tools used in individual steps of the cryo-ET pipeline before and after the tomographic reconstruction and identify their functionalities.

Additionally, due to the tilt limitation in cryo-ET, there is the *missing wedge* problem due to the lack of projections at high tilt-angle views [35]. While the effect of the missing information can be minimized already during the reconstruction step, this problem is usually addressed separately, since adding the missing information into the data is not part of reconstruction when not considering a single particle approach where the same particle is present in many parts of the sample. This is true for works presented by Trampert *et al.* [36], Yan *et al.* [37], and Ding *et al.* [38].

In order to handle angle-limited projection data, Würfl

et al. [13] proposed a deep-learning-based reconstruction method for limited angle tomography. Another common attempt is to couple a reconstruction technique, e.g., back-projection reconstruction, with a deep learning denoising approach. Here, denoising is applied either on the tilt-series before reconstruction [15], [16], [17], or after the reconstruction on tomograms created with traditional techniques [18], [19]. Alternatively, reconstruction methods including regularization terms have been proposed to handle this ill-posed problem, but only for 2D synthetic limited angle data [39], [40]. An example of the current reconstruction pipeline for obtaining state-of-the-art results in cryo-electron tomography (cryo-ET) is presented by Yao *et al.* [41]. It consists of downsampling the reconstructed tomograms using a binning approach and then applying subtomogram averaging to manually selected subtomograms. The tomograms are reconstructed using IMOD's weighted back-projection [28], and then binned 4 \times for further processing. Hundreds of tomograms are reconstructed and binned, and thousands of subtomograms are manually selected to apply subtomogram averaging.

Due to the low signal-to-noise ratio in cryo-ET tilt-series and the reconstructed tomograms, there is a need to denoise the data either before or after the tomographic reconstruction or even at both times. A recent overview on using the denoising in the cryo-ET reconstruction by Frangakis [42] presents what kind of methods are typically used prior or after the tomographic reconstruction ranging from simple linear, non-linear, bilateral, and trilateral filters in the spatial domain, filters in Fourier and Wavelet domains and denoising approaches using neural networks [16], [18]. The denoising using neural networks requires further studies to be well understood. There are still some concerns on whether such methods used in the pre-reconstruction step can introduce additional artifacts and/or biases. In our approach, we use a well-understood approach to denoising during the reconstruction step, which to the best of our knowledge does not introduce any biases into the data.

Multiple model-based iterative reconstructions approaches have been proposed for cryo-ET. Goris *et al.* [43] present a regularized algorithm that combines SIRT with Total Variation (TV), [44] for cryo-ET of nano-structured materials. These samples are not as susceptible as biological samples which allows using higher electron doses, obtaining projections in a higher angular range (-70° to 78°) and a higher number of projections (each 2° , 75 projections) with significantly less noise. Böhning *et al.* [45] present a cryo-ET reconstruction method based on compressed sensing. They propose using 3D second-order TV for tomographic reconstruction, casting the reconstruction problem as a regularized optimization problem, solved using a primal-dual hybrid gradient method implemented in python using ASTRA Toolbox [30], [31] for the data term. The virus particles in the tilt-series are separated and centered on the tilt axis for individual reconstruction. The reconstructed individual particles are then cropped and used for sub-tomogram averaging. The pipeline requires manual intervention between the steps, and relies on several software tools for the different steps. Yan *et al.* [37] applied the Model-Based Iterative Reconstruction (MBIR) [46] algorithm for cryo-ET reconstruction. The MBIR framework uses a probabilistic model for the

measurements (data term) and a probabilistic model for the object (regularization). Our framework is based on proximal algorithms [47], which allows for seamless combination of different data term solvers and denoisers. We experimented with SART and SIRT for the data term, and TV, Huber loss [48] and Non-Local means (NLM) [49] for regularization. Parting from aligned tilt-series, our framework is a self-contained software that allows 3D visualization of tomograms created through integrated reconstruction and denoising.

It is expected that the reconstruction processes are not perfect, which is especially true when done with insufficient input data in the first place. Uncertainty visualization is a central topic in visualization both for 2D as well as in 3D data as is reflected by multiple surveys [50], [51], [52], [53], [54]. There are only a few methods that address uncertainty in volume visualization, and it is not common for the current cryo-ET reconstruction pipelines to provide this uncertainty to the user, which is an option in our case. We provide an *error tilt-series*, which includes a re-projection error image for each original projection image. The error tilt-series is reconstructed into an *error volume* that shows the uncertainty in the reconstructed volume. The work by Kniss *et al.* [55] allows users to explore the uncertainty of surface boundaries in volumetric data interactively. Lundström *et al.* [56] presents how probabilistic animation can be used for uncertainty visualization in medical volume rendering, which was later also applied to concurrent visualization of real-time fMRI data presented by Nguyen *et al.* [57]. Lindholm *et al.* [58] propose the use of spatially conditioned transfer functions using local material distributions, which can also be applied to uncertainty visualization. This results in using separate transfer functions for individual modality, e.g., one being the volume of reconstructed values and the other volume of uncertainty values, and blending them in the rendering process. We experimented blending the error volume and the reconstructed volume in the rendering process, but found that visualizing the error and the reconstruction individually allowed for an easier interpretation of the error. Dinesha *et al.* [59] explore an option of using HDR volumetric rendering where regular transfer function is used for mapping the intensity values into HDR color space, and the luminance component of the color is exploited for capturing uncertainty.

3 METHOD

We propose a reconstruction framework that creates denoised 3D tomograms from previously aligned noisy 2D tilt-series and allows direct visualization of the reconstructed tomograms. We present tomographic reconstruction as a regularized optimization problem, which is solved using a proximal joint-optimization technique that integrates denoising and reconstruction, and we describe the proximal operators that are required in the framework. A diagram of our method's pipeline is depicted in Figure 1.

3.1 Tomography as an optimization problem

The projection images coming from cryogenic electron tomography represent the transmission images of parallel rays

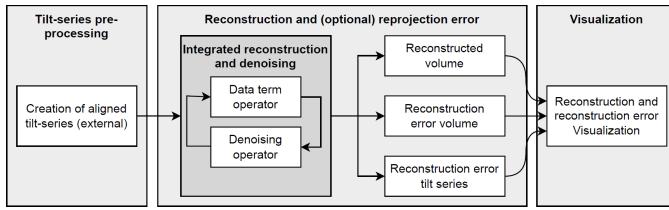


Fig. 1. Pipeline of our framework. The input is aligned tilt-series, which are used to create denoised tomograms by alternating between reconstruction and denoising instead of denoising the tilt-series or the resulting tomogram. The re-projection error and a corresponding error volume can be generated optionally.

passing through a sample. The image formation model is derived from the Beer-Lambert Law [60]. The measured intensity for a particular pixel p_i is

$$\hat{I}(p_i) = \hat{I}_0(p_i) \cdot \exp\left(-\int \sigma(r_{p_i}(t)) dt\right), \quad (1)$$

where $r_{p_i}(t)$ is a ray associated with the pixel p_i , σ is the attenuation cross-section, commonly known as density and $\hat{I}_0(p_i)$ is the intensity of the electron beam. The goal is to reconstruct a discrete volume $\sigma(x, y, z)$ from the available ray integrals, usually in log space as

$$I(p_i) = -\log\left(\frac{\hat{I}(p_i)}{\hat{I}_0(p_i)}\right) = \int \sigma(r_{p_i}(t)) dt. \quad (2)$$

In log space, the relationship between the projections and the 3D volume becomes linear and is simply the line integral of the volume densities $\sigma(x, y, z)$ along the ray $r_{p_i}(t)$. With this, the discretized image formation model can be represented by a *weight* matrix \mathbf{W} that contains the contribution of each volume element for each ray integral as in algebraic reconstruction techniques, where tomographic reconstruction is presented as a linear algebra problem $\mathbf{W}\mathbf{v} = \mathbf{p}$, where $\mathbf{v} \in \mathbb{R}^N$ is the unknown volume that we want to reconstruct, $\mathbf{p} \in \mathbb{R}^M$ is the set of known ray integrals (tilt-series), and the weight matrix $\mathbf{W} \in \mathbb{R}^{M \times N}$ represents the contribution of each voxel to each ray integral. We formulate tomographic reconstruction as the following optimization problem:

$$\arg \min_{\mathbf{v}} \underbrace{\frac{1}{2} \|\mathbf{W}\mathbf{v} - \mathbf{p}\|_2^2}_{\text{data fidelity}} + \underbrace{\Gamma(\mathbf{v})}_{\text{regularization}}. \quad (3)$$

The first term in Equation 3 is a data fitting term that measures how good of a fit the current volume estimate is to the original projection data. The second term in Equation 3 is a non-linear regularization term that constrains the space of acceptable solutions, which is required to deal with the noisy and angle-limited tilt-series from cryogenic electron tomography. To apply known non-linear solvers, we rewrite the data term as $F(\mathbf{v}) = \frac{1}{2} \|\mathbf{W}\mathbf{v} - \mathbf{p}\|_2^2$ and the regularization term as $G(\mathbf{K}\mathbf{v}) = \Gamma(\mathbf{v})$, where G is a non-linear function and \mathbf{K} is a matrix. Then, we rewrite Equation 3 as the following constrained optimization problem:

$$\arg \min_{\mathbf{v}, \mathbf{z}} F(\mathbf{v}) + G(\mathbf{z}), \text{ subject to } \mathbf{K}\mathbf{v} = \mathbf{z}. \quad (4)$$

3.2 Proximal Framework for Cryo-ET

Equation 4 is a standard problem in numerical optimization, and there are many non-linear solvers with guaranteed convergence when F and G are convex. There is a family of such solvers called *proximal algorithms*, which are powerful and flexible optimization algorithms that can be used to solve non-smooth, constrained, distributed or large-scale optimization problems. Proximal algorithms have a modular nature, since the optimization problems are split into smaller sub-problems called *proximal operators*, and each sub-problem can be solved independently. In our case, we require proximal operators for F and G . Let $H(\cdot)$ be a function that we want to minimize (i.e. either F or G). The proximal operator of $H(\cdot)$ can be interpreted as getting closer to the minimum of $H(\cdot)$ while staying close to a reference point \mathbf{u} , and is defined as [47]:

$$\text{prox}_{\lambda H}(\mathbf{u}) = \arg \min_{\mathbf{v}} H(\mathbf{v}) + \frac{1}{2\lambda} \|\mathbf{v} - \mathbf{u}\|_2^2, \quad (5)$$

where $\mathbf{v}, \mathbf{u} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ is a regularization parameter for controlling the relevance of the proximity term $\|\mathbf{v} - \mathbf{u}\|_2^2$. The proximal operators of many functions are easy to compute and have a closed form solution. Having the proximal operator for a given set of functions creates the possibility of solving complex optimization problems using proximal algorithms that combine these proximal operators, in our case, the proximal operators of F and G .

3.2.1 Linearized Alternating Direction Method of Multipliers

The Linearized Alternating Direction Method of Multipliers (LADMM) [47] (Algorithm 1) is a proximal algorithm that can be used to solve optimization problems of the form in Equation 4.

Algorithm 1 Linearized ADMM

```

Require:  $\mathbf{K} \in \mathbb{R}^{d \times N}$ ,  $\rho, \mu \in \mathbb{R}$  such that  $\frac{\mu}{\rho} \|\mathbf{K}\|^2 \leq 1$ , initial values  $\mathbf{v}^{(0)} \in \mathbb{R}^N$  and  $\mathbf{z}^{(0)} \in \mathbb{R}^d$ 
Initialize  $\mathbf{y}^{(0)} = \mathbf{0}_d$ 
for  $t = 1 \dots T_1$  do
     $\mathbf{v}^{(t+1)} = \text{prox}_{\mu F}(\mathbf{v}^{(t)} - \frac{\mu}{\rho} \mathbf{K}^T (\mathbf{K}\mathbf{v}^{(t)} - \mathbf{z}^{(t)} + \mathbf{y}^{(t)}))$ 
     $\mathbf{z}^{(t+1)} = \text{prox}_{\rho G}(\mathbf{K}\mathbf{v}^{(t+1)} + \mathbf{y}^{(t)})$ 
     $\mathbf{y}^{(t+1)} = \mathbf{y}^{(t)} + \mathbf{K}\mathbf{v}^{(t+1)} - \mathbf{z}^{(t+1)}$ 
end for
return  $\mathbf{v}^{(T_1)}$ 

```

In Algorithm 1, \mathbf{K} is a matrix related to the regularization term, ρ, μ arise when manipulating ADMM to solve problems of the form $\arg \min_{\mathbf{v}} F(\mathbf{v}) + G(\mathbf{K}\mathbf{v})$ (see [47] section 4.4.2), \mathbf{z} and \mathbf{y} are additional variables. T_1 is the number of iterations of the proximal algorithm.

3.3 Proximal operators

To apply Algorithm 1, we need the proximal operators for our data fidelity term and regularization terms.

3.3.1 Data term proximal operator

We want a proximal operator to solve the data fidelity term in Equation 3. By definition:

$$\text{prox}_{\mu F}(\mathbf{u}) = \arg \min_{\mathbf{v}} \frac{1}{2} \|\mathbf{W}\mathbf{v} - \mathbf{p}\|_2^2 + \frac{1}{2\mu} \|\mathbf{v} - \mathbf{u}\|_2^2 \quad (6)$$

That is, given a current volume \mathbf{u} , find a new volume \mathbf{v} close to the current estimate \mathbf{u} that minimizes $\frac{1}{2} \|\mathbf{W}\mathbf{v} - \mathbf{p}\|_2^2$. This is simply a least-squares problem, whose analytical solution is [47]:

$$(\mu\mathbf{W}^T\mathbf{W} + \mathbf{I})^{-1}(\mu\mathbf{W}^T\mathbf{p} + \mathbf{u}). \quad (7)$$

Any linear solver can be used to solve Equation 7. In our implementation, we use SART [12], which is a proven algorithm for tomographic reconstruction with a known proximal operator [61]. With SART, we can represent the projection matrix \mathbf{W} and its transpose in a procedural way, avoiding the explicit generation of the system matrix, which would otherwise incur an infeasible memory cost. Note that even if \mathbf{W} is not full rank, the inverse in the proximal operator is well-conditioned thanks to the addition of \mathbf{I} . The proximal operator for SART corresponds to Algorithm 2, where T_2 is the total number of SART iterations when calling the proximal operator (independent from T_1), α is the relaxation coefficient, μ is the regularization parameter of the proximal operator, \mathbf{p} is the original projection data and \mathbf{u} is an initial estimate of the reconstruction.

Algorithm 2 SART proximal operator

Require: $\mathbf{W} \in \mathbb{R}^{M \times N}$, $\alpha, \mu \in \mathbb{R}$, $\mathbf{p} \in \mathbb{R}^M$, $\mathbf{u} \in \mathbb{R}^N$

Initialize:

$$\begin{aligned} \mathbf{v}^{(0)} &= \mathbf{u} \\ \mathbf{y}^{(0)} &= \mathbf{0} \end{aligned}$$

for $t = 1 \dots T_2$ **do**

for all projection images P_θ **do**

$$\begin{aligned} y_j^{(t+1)} &= y_j^{(t)} + \alpha \sum_{i \in P_\theta} \frac{\sqrt{2\mu} p_i - \sqrt{2\mu} \sum_k w_{ik} v_k^{(t)} - y_i^{(t)}}{1 + \sqrt{2\mu} \sum_k w_{ik}} \\ v_j^{(t+1)} &= v_j^{(t)} + \alpha \frac{\sum_{i \in P_\theta} \frac{\sqrt{2\mu} p_i - \sqrt{2\mu} \sum_k w_{ik} v_k^{(t)} - y_i^{(t)}}{1 + \sqrt{2\mu} \sum_k w_{ik}}}{\sqrt{2\mu} \sum_{i \in P_\theta} w_{ij}} \sqrt{2\mu} w_{ij} \end{aligned}$$

$$v^{(t+1)} = \max(0, v^{(t+1)})$$

end for

end for

return $\mathbf{v}^{(T_2)}$

3.3.2 Regularization operators

Anisotropic Total Variation Proximal Operator: This regularizer was proposed by Rudin, Osher, and Fatemi and is known as the ROF model [44]. It was conceived under the idea that images with excessive (and probably noisy) details have high total variation (the integral of the absolute image gradient is high). In other words, images are assumed to be sparse in the gradient domain. Anisotropic Total Variation (ATV) is defined as the ℓ_1 norm of the image gradient. In our framework, we define $G_{ATV}(\cdot) = \|\cdot\|_1$ and $\mathbf{K} \in \mathbb{R}^{3N \times N}$ as the forward difference matrix that produces the discrete volume gradient. Since the proximal operator of the ℓ_1 norm is the *soft thresholding function* [47], the proximal operator of G_{ATV} is the soft thresholding function applied to each element of the volume gradient.

$$\text{prox}_{\rho_1 G_{ATV}}(\mathbf{u}_{n_i}) = \text{sign}(\mathbf{u}_{n_i}) \times \max(0, |\mathbf{u}_{n_i}| - \rho_1), \quad (8)$$

where \mathbf{u}_{n_i} is the i^{th} component of the volume gradient $\mathbf{u} = \mathbf{K}\mathbf{v}$ at voxel n and ρ_1 is a regularization parameter that controls how fast the volume gradient shrinks.

Huber Penalty Proximal Operator: This regularizer was proposed by Peter J. Huber [48]. The Huber penalty can be applied on the volume gradient to reduce the total variation, avoiding the sparse gradients produced by the ℓ_1 norm. It is defined as a piecewise function composed of ℓ_2 and an ℓ_1 segments:

$$G_{\text{Huber}} = \begin{cases} \frac{1}{2}x^2, & |x| \leq \delta, \\ \delta(|x| - \frac{1}{2}\delta), & |x| > \delta. \end{cases} \quad (9)$$

The proximal operator is again a soft-thresholding operation applied to each element of the volume gradient:

$$\text{prox}_{\rho_2 G_{\text{Huber}}}(\mathbf{u}_{n_i}) = \begin{cases} \frac{1}{1+\rho_2} u_{n_i}, & |u_{n_i}| \leq \delta + \delta\rho_2, \\ u_{n_i} + \rho_2\delta, & u_{n_i} < -(\delta + \delta\rho_2), \\ u_{n_i} - \rho_2\delta, & u_{n_i} > \delta + \delta\rho_2, \end{cases} \quad (10)$$

where ρ_2 controls how fast the gradient shrinks and δ controls the transition point between the linear and quadratic segments in G_{Huber} .

Non-local means as a proximal operator: Any Gaussian denoiser (e.g., Non-Local Means [49]) can be interpreted as a proximal operator with a fixed regularization parameter as the assumed variance of the noise, as shown in the works by Venkatakrishnan *et al.* [62] and Heide *et al.* [63]. We will use the Non-Local Means algorithm as a proximal operator in our framework since its operation principle resembles subtomogram averaging, and it is a common denoiser in cryo-ET pipelines (e.g., Thermo Fisher's Amira Software [64]). For each pixel in an image, Non-Local Means performs denoising by searching pixels with neighborhoods (square-shaped patches centered at the pixel) similar to that of the pixel under denoising in a square-shaped search zone centered at the pixel. The denoised value is computed as a weighted average of the pixel values in the search zone. The weight of each pixel in the search zone is proportional to the distance to the center pixel and the similarity of the neighborhoods of each pixel and the center pixel.

Assuming some prior distribution $G(\mathbf{v})$, the Non-Local Means algorithm [49] can be used directly in the framework as a proximal operator (with $\mathbf{K} = \mathbf{I}$):

$$\mathbf{v}_{MAP} = \text{prox}_{\sigma_{nlm}^2 \tilde{G}}(\mathbf{b}), \quad (11)$$

where $\tilde{G}(\mathbf{v}) = -\log(G(\mathbf{v}))$. The regularization parameter is the assumed variance of the noise σ_{nlm}^2 . For details, see the supplementary material.

Using the regularization proximal operators: In our implementation, we use TV and Huber penalty interchangeably in combination with NLM. First, we use Algorithm 1 with TV or Huber by performing the z update with Equation 8 or Equation 10 for several iterations (typically 50–80). Then, we incorporate NLM to reduce the noise in the reconstructed volume \mathbf{v} , using TV+NLM or Huber+NLM for a few iterations. We found that two final iterations with NLM improve the results by further reducing the noise levels without blurring the details in the reconstructed structures.

3.4 Implementation Details

Figure 2 includes a 2D diagram of the reconstruction scene. We can think of this diagram as the reconstruction of a single

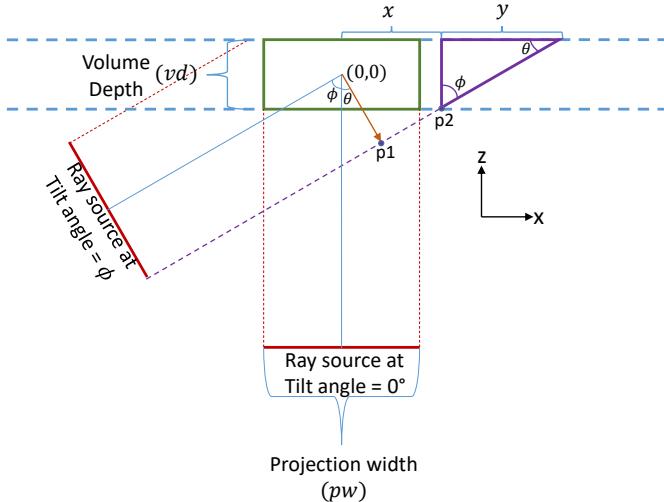


Fig. 2. Top view of the reconstruction scene. Horizontal padding is required for the extreme views.

horizontal slice. The volume (green rectangle) is centered at the origin. The rays travel from the source (red line) to the detector (not in the diagram) which is parallel to the ray source, located on the other side of the volume. In the domain of cryo-ET, the rays present parallel geometry and calculations can be easily parallelized. Our framework includes two implementations for the SART and SIRT proximal operators. The first is based on ray casting [65], and the second is based on voxel splatting [66]. In the voxel splatting implementation, the weights are read from pre-integrated lookup tables that are computed by line integrals of rays traversing radially symmetric filter kernels [67] at different distances. A 2D example of the ray-kernel interaction is shown in Figure 3.

3.4.1 Ray Casting Implementation of SART/SIRT

Forward Projection and Correction Term: From each pixel p_i in the ray source, a ray is cast across the volume towards the corresponding pixel in the detector. Samples are taken at equidistant points along the ray. At each sample point, a volume value is retrieved by tri-linear interpolation. The correction term is computed immediately after calculating the line integral as

$$\frac{I(p_i) - \text{sample_rate} \times \sum_{s=1}^S v_s^k}{\text{ray_length}}, \quad (12)$$

where $I(p_i)$ is the measured line integral value of the current ray in the log space, S corresponds to the number of samples, sample_rate is the distance between samples (in voxels), v_s^k is the current volume estimate at sample point s , and ray_length is the length in voxels of the ray-volume intersection.

Back-projection: For each voxel, we cast a ray parallel to the current view pointing to the detector. We compute the point at which the ray intersects the detector, and retrieve a correction value with bi-linear interpolation of the correction image. The voxel is updated with the interpolated correction value, times the relaxation factor. In this way, the forward projection and backward projection are not exact

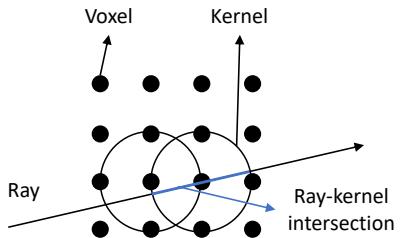


Fig. 3. 2D representation of ray-voxel interactions.

transposes of each other, but we avoid data races in both operations, and the result is practically the same as with the splat-based implementation, where the operators are transpose of each other.

3.4.2 Voxel Splatting Implementation of SART/SIRT

Forward Projection: Each voxel is splatted into the detector and updates all the pixels of a *ray-sum image* that are inside the splatted voxel's kernel support. A *weights image* is created accordingly. Data races can occur if two voxels try to update the same pixels at once, either in the ray-sum image or the weights image.

Correction Image: For each pixel in the correction image, the correction term is computed as

$$\alpha \times \frac{I(p_i) - \text{ray_sum_image}[i]}{\text{weights_image}[i]} \quad (13)$$

where α is the relaxation parameter, $I(p_i)$ is the measured line integral in the log space, $\text{ray_sum_image}[i]$ is the value of the pixel i of the computed projection, and $\text{weights_image}[i]$ is the sum of the weights for the pixel i .

Back-projection: For each voxel, we splat the voxel in the detector and retrieve a correction value by interpolation using the pre-computed weights. The forward projection and back-projection operators are transposes of each other. If we use linear interpolation, the result is practically the same as with the ray-based implementation. There are no data races in the back-projection operation.

3.4.3 CUDA implementation:

Our framework is implemented in CUDA. The most important kernels are described below.

- **Ray-casting Forward Projection:** Launches a thread for each pixel in the projected image and each thread computes a line integral and a correction term. It produces a correction image that is used in the back-projection.
- **Ray-casting Back-projection:** Launches a thread for each voxel. The corresponding voxel value is updated based on a correction term which is retrieved from the correction image.
- **Splatting Forward Projection:** Uses one thread per voxel. A synthetic projection and corresponding weights image are computed as described in Section 3.4.2.

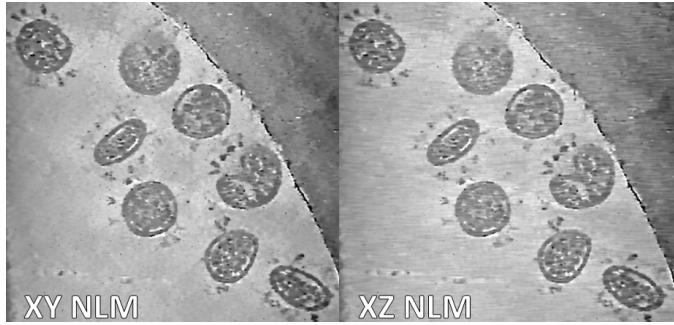


Fig. 4. Comparison between 2D XY NLM (left) and XZ NLM (right) using dataset 1. XZ NLM produces horizontal line-shaped artifacts.

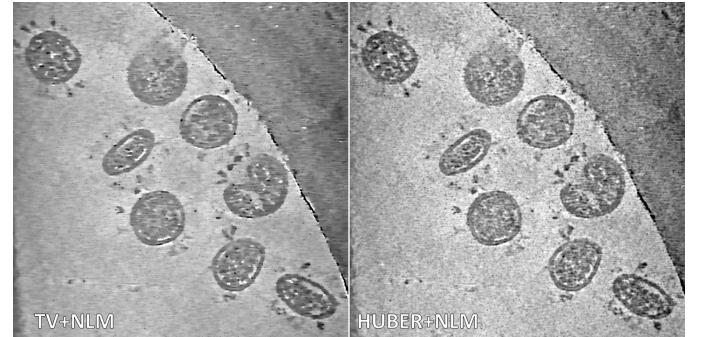


Fig. 5. Comparison between Huber+NLM vs. TV+NLM, dataset 1

- **Splatting Correction Image:** Using one thread per pixel, computes a correction image based on the previously computed synthetic projection, weight image and original projection.
- **TV Proximal Operator:** Executes one thread per voxel. Each thread updates the corresponding three components of the variable z .
- **NLM Proximal Operator:** Executes one thread per voxel. Each thread updates the corresponding voxel value based on similar patches in the surrounding neighborhood.

3.4.4 Horizontal Padding for Limited Angle Reconstructions

If the reconstructed volume has the same width as the projection images, the rays in the borders of the projection image are going to have a shorter intersection with the reconstruction volume as the tilt angle is increased. As seen in Figure 2, rays at tilt angle zero intersect the volume exactly at the boundaries, and all the rays intersect the volume for the same distance. Rays from views with maximum tilt angle ϕ will traverse the volume for different distances. Horizontal padding is required so that the distance that the rays travel inside the volume is the same for each view. For this, the required volume width vw including padding is

$$|pw \cdot \sin(\theta) + (pw \cdot \cos(\theta) - vd) \cdot \cot(\theta)| + \frac{2 \cdot vd}{\tan(\theta)} \quad (14)$$

Where pw is the width of the projection images, $\theta = 90 - \phi$, vd is the volume depth. The obtained value should be rounded up to the next even value. For example, for projection width = 1024, the padded volume width is 2568. Further details are included in the supplementary material.

3.5 3D Visualization of the Reconstructions

We use a simple adaptation of the regular Direct volume rendering ray marching approach [68] for visualizing the reconstructed tomograms. The rendering pipeline also includes soft shadows [69] and local ambient occlusion [70]. The visualizations include intermediate results (SART+TV) and final results (SART+TV+NLM).

3.6 Visualization of the Error

The re-projection error of the reconstruction can be optionally generated. An 'error tilt-series' is created as follows: For

each viewing angle included in the input tilt-series, we compute a synthetic projection with the estimated volume, and then a 're-projection error image' is created as the absolute value of the difference between the original and estimated projections. The error contribution in the individual error projections from the tilt-series is spread sparsely throughout the whole projections resulting in a very noisy error image. Using the error tilt-series, an 'error volume' is created by applying SART to the error tilt-series. Finally, to enhance the biggest error contributions for visualization purposes, we apply the following steps:

- 1) we first apply 3×3 Gaussian blur;
- 2) next we threshold the intensity values below $\frac{1}{8}$ -th of the max value;
- 3) finally we adjust the gamma value of the image to 2.0 to shift the mid-tones to brighter values.

This approach pronounces the regions in the error projection images with the highest error values and also pronounces structural features.

4 EVALUATION AND RESULTS

4.1 Datasets description

For the experimental results included in the paper, we used 4 different datasets:

- 1) SARS-CoV-2 specimen 1

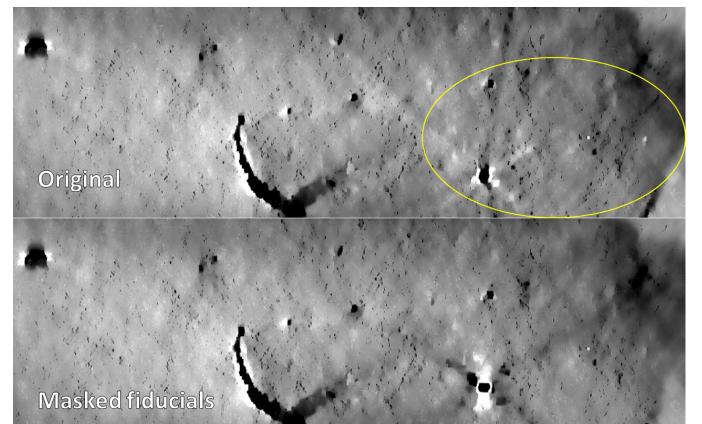


Fig. 6. Effect of masking the fiducials on a tilt-axis (XZ) slice on dataset 1.

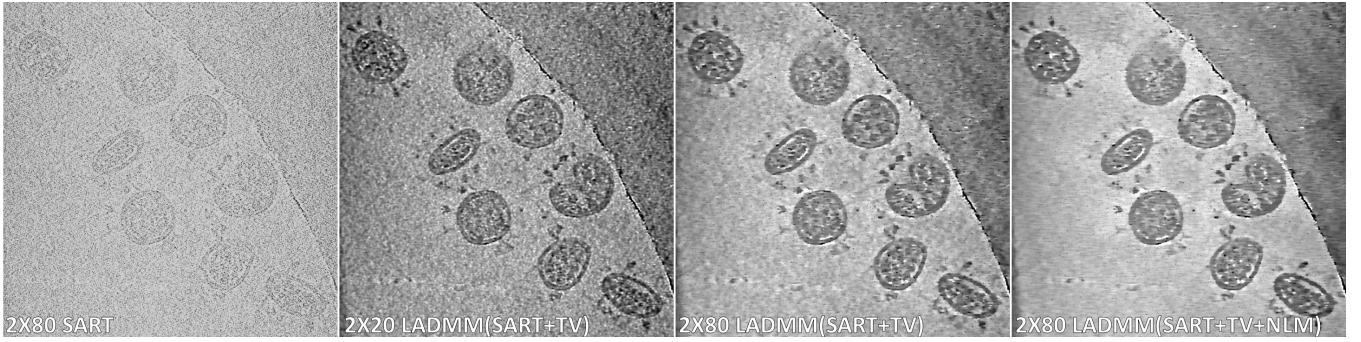


Fig. 7. Reconstructions from a SARS-CoV-2 tilt-series (dataset 1) using our framework with different settings. Without using regularization (first from the left), the reconstruction fits the noise in the projection data. By alternating between reconstruction (SART) and denoising (TV), the features are reconstructed and the noise is removed with the pass of iterations (second 2x20 iterations, third 2x80 iterations). Including NLM in the last two iterations further reduces the noise, preserving the reconstructed features.

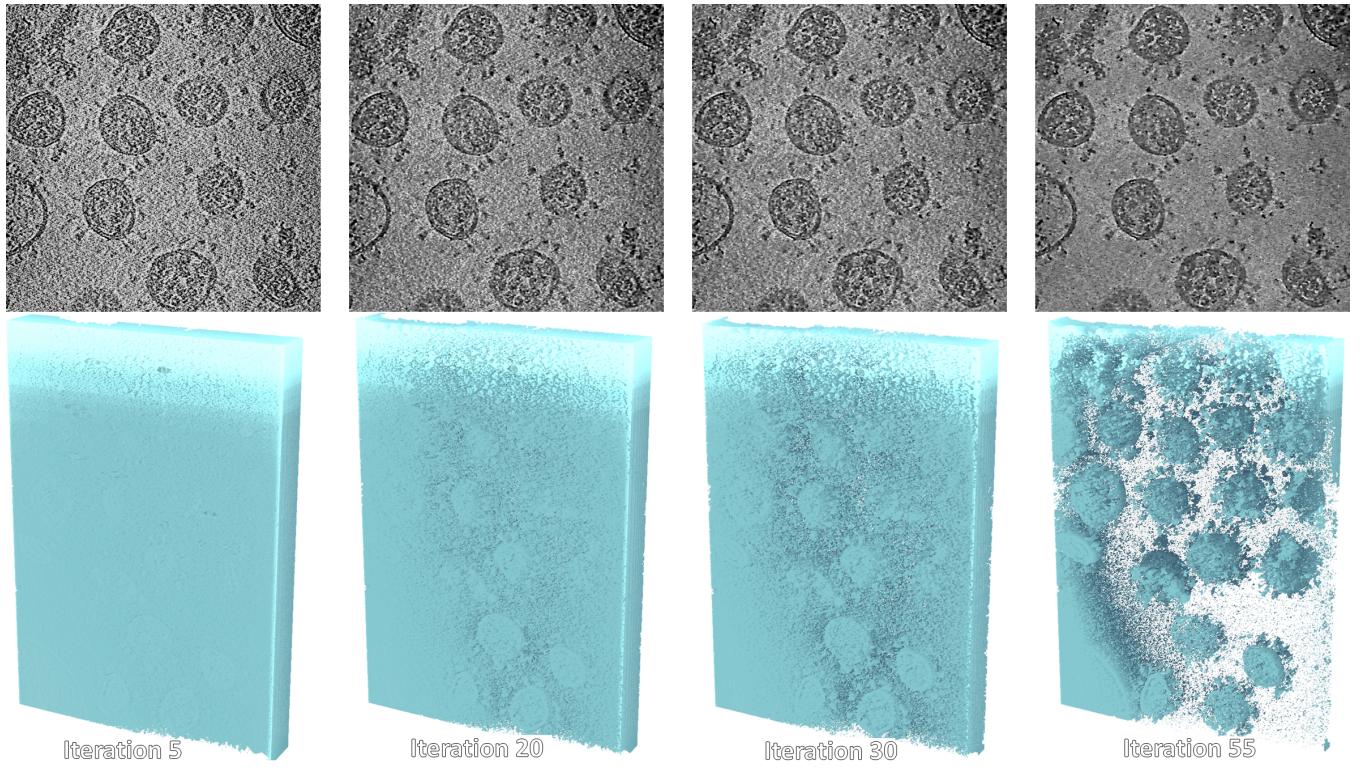


Fig. 8. Reconstruction progress from SARS-CoV-2 tilt-series (dataset 2) using our framework. The noise is reduced with the pass of iterations. The final iterations using NLM+TV reveal the virus particles in the reconstructed volume 3D visualization.

- 2) SARS-CoV-2 specimen 2
- 3) Influenza specimen
- 4) HIV-1 specimen (EMPIAR-10643 [71], [72])

The spatial information available in each dataset corresponds to Table 1.

4.2 Parameter settings

We denote the number of iterations as $T_2 \times T_1$, where T_1 is the number of iterations of the proximal algorithm, and T_2 is the number of SART iterations per iteration of the proximal algorithm. The last two iterations use TV+NLM, the rest use only TV. Unless otherwise specified, the parameters in our method are: $\alpha = 0.2$, $\rho_1 = 0.01$, $\|K\| = 11.99$, $\mu = \frac{0.99\rho}{\|K\|}$, $nlm_s = 21$, $nlm_w = 7$, $nlm_{skip} = 3$, $\sigma_{nlm} = 4 \times 10^{-8}$. A final iteration of the data term proximal operator is performed

to enforce consistency between the original projections and the reconstructed volume.

There is no set of optimal parameters for every dataset. A parameter search is necessary for every data source (different specimens, different acquisition devices, etc.), but we found that the set of optimal parameters is mostly unchanged for different datasets from the same source (e.g., datasets 1 and 2) and only individual parameters need to be adjusted. For example, the total number of iterations $T_2 \times T_1$ provides control over the ratio of data fidelity (T_2) and regularization (T_1), α controls the rate of convergence during SART iterations, ρ_1 controls the amount of TV regularization on a proximal iteration, the NLM parameters nlm_s , nlm_w , nlm_{skip} can be adjusted according to the size of the virions present in each dataset and σ_{nlm}^2 to control the

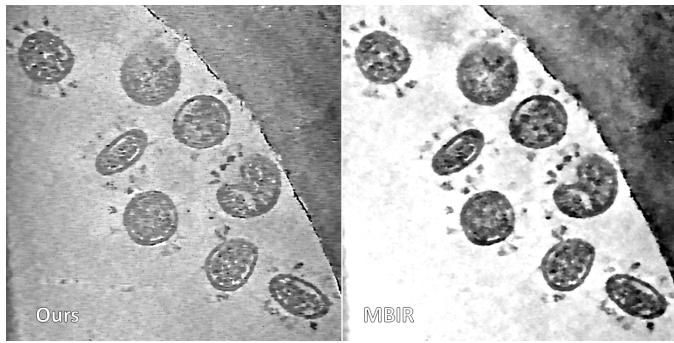


Fig. 9. Dataset 1 reconstructions. Left: Ours, 2×80 iterations of LADMM (SART+TV+NLM). Right: MBIR, 600 iterations.

regularization strength.

4.3 Computation hardware and time

A Volta V100 (32 GB) GPU was used for the reconstructions, which took about 30 minutes for the SARS-CoV-2 datasets and 1 hour for the HIV-1 dataset.

4.4 Experiments with Non-Local Means

We experimented with 3D and 2D NLM. The search region and the neighborhoods have a finite extent, are square-shaped in 2D and cube-shaped in 3D. The side of the search zone is controlled with nlm_s parameter, and is $2 \times nlm_s + 1$ (always odd). The side of the neighborhoods is controlled similarly, with parameter nlm_w . The size of the neighborhoods should be set accordingly to the size of sub-structures in the tomograms. We noticed that the reconstructions obtained with 3D and 2D NLM were very similar, and since the computational cost of the 2D version is much lower, we selected it for the rest of the experiments. However, we still found the 2D version computationally expensive, and we experimented with skipping some of the pixels in the search zone. The parameter nlm_{skip} controls the pixel skipping. There are nlm_{skip} ignored pixels between every two considered pixels on each axis of the search region. For the 2D versions, we found that the quality of the result was not affected significantly for $nlm_{skip} \leq 3$, and the computational cost was greatly reduced. Care should be taken to include the pixel in the center of the patch when skipping pixels since it is the one with the higher weight/contribution. We analyzed two different versions of 2D NLM:

- **XY NLM:** Applying 2D NLM on each XY slice of the volume

- **XZ NLM:** Applying 2D NLM on each XZ slice comparing neighborhoods on XY slices, with the idea of balancing the resolution between XY and XZ (tilt-axis) slices.

However, the XY slices produced by XZ NLM presented horizontal line-shaped artifacts, and there was not a noticeable improvement in the image quality on the XZ slices (Figure 4). Therefore, we selected 2D XY NLM as the default NLM operator for the framework. We experimented with applying NLM in the last, in the last two, and in the last three iterations only. We found that two iterations are enough to achieve good denoising, with $nlm_s = 21$, $nlm_w = 7$, $nlm_{skip} = 3$, and more iterations tend to blur the details.

4.5 TV vs. Huber penalty

A comparison between TV+NLM and Huber+NLM is presented in Figure 5. The sparse gradients enforced by TV achieve a cleaner, more homogeneous background. Huber achieves a slightly noisier reconstruction, but sharper details in the reconstructed virus particles. We refer the reader to the supplementary material for additional comparisons between TV+NLM and Huber+TV. The results in the following sections were obtained using TV+NLM.

4.6 Masking the fiducials

In order to minimize the artifacts coming from the fiducial markers, we created masks corresponding to the fiducial markers for each projection imaged for the two SARS-CoV-2 datasets. The masks were used to avoid back-projecting the pixels corresponding to the fiducials during the execution of Algorithm 2. An example of the artifacts removed by masking the fiducials on dataset 1 can be seen in Figure 6, which is an XZ slice. Additional images can be found in the supplementary material.

4.7 Ablation study

To show the contribution of the different regularizers in our framework, we created reconstructions of dataset 1 with different configurations. The results are shown in Figure 7. From left to right, first we have 2×80 iterations of proximal iteration (Section 4.1 of [47]) using the SART proximal operator (Algorithm 2). The second and third images correspond to LADMM (Algorithm 1) using SART+TV (Algorithm 2, Equation 8), after 2×20 and 2×80 iterations respectively. The rightmost image corresponds to 2×80 iterations of LADMM with Total Variation, using Non-Local Means instead of Total Variation in the last two iterations, plus a single final execution of the data term operator in the end. Without regularization, the reconstructions fit the noise in the projection data, leading to noisy reconstructions where it is hard to distinguish between virus particles and background. Regularization allows creating cleaner reconstructions by alternating between reconstruction and denoising, greatly reducing the noise and reconstructing rich details with the pass of iterations. Finally, Non-Local Means further reduces the noise in the empty regions while preserving the features of the virus particles.

Dataset	Cell size[Å ³]	Projection size [pixels ²]	Pixel size [Å ²]
1, 2	31334.4×22260.48×1.36	1024×1440 (bin-4)	5.44 ²
3	9794.4×10132.32×1.32	960×928 (bin-4)	5.28 ²
4	—	1366×1366 (bin-3)	3.54 ²

TABLE 1
Datasets spatial information



Fig. 10. Dataset 1 reconstructions. Left: Using Topaz Denoise to denoise the tilt-series and reconstructing it using SART. Right: Ours on the aligned noisy tilt-series, 2×80 iterations of LADMM (SART+TV+NLM)

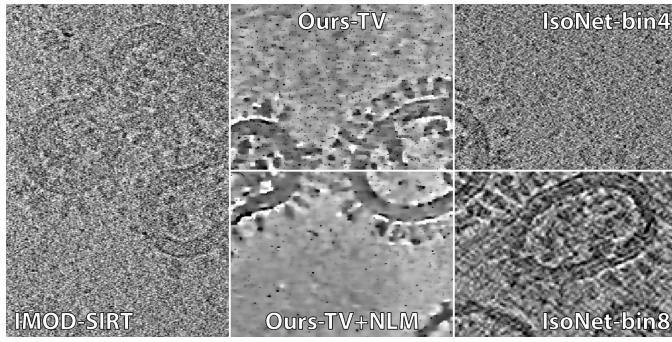


Fig. 11. Dataset 3 reconstructions. From left to right: IMOD, Our approach using TV (above) and TV+NLM (below), and IsoNet with bin-4 (above) and bin-8 (below).

4.8 Progress over time

We can see how the reconstruction progresses over time in Figure 8, which was created from a tilt-series of another SARS-CoV-2 specimen (dataset 2) with 3×55 iterations of LADMM. With the advance of the iterations, the features of the virus particles are gradually reconstructed and at the same time, the noise is diminished steadily. If we use NLM in the last two iterations, the noise in the empty regions is greatly reduced without significantly distorting the features of the virus particles gained through the reconstruction process. After the 55 iterations, a last execution of the data term operator is performed to ensure that the reconstructed volume fits the original projection data. The 3D visualizations at different stages show that the noise is gradually removed, and the final result allows visualization of the virus particles. Reconstructions with up to 3×100 iterations are available at the supplementary material.

4.9 Comparison with another model-based reconstruction method

We compare our results against those obtained with MBIR [37] on dataset 1. The reconstruction parameters for MBIR were: diffuseness 0.2, data offset 32768, smoothness 0.08, sigma X 1.52677×10^{-5} , 600 outer sub-iterations, 100 inner sub-iterations. The reconstruction time was 4 hours and 30 minutes (roughly 9 times slower than our method) using a processor Intel Xeon E5-2687W v4 (2x3 GHz processors). The results are very alike (Figure 9), which is

expected given the similarity between the two techniques. Additional results in orthogonal slices are included in the supplementary material.

4.10 Comparison with a pre-reconstruction denoising method

We processed dataset 1 using Topaz Denoise [16], developed for denoising cryo-ET tilt-series. The denoised tilt-series was reconstructed with 20 iterations of SART without regularization since the tilt-series was denoised and using additional regularization (e.g. LADMM with TV) produced over-regularized (blurry) reconstructions. We compare the result with the one from our framework using LADMM 2×80 iterations with TV, NLM instead of TV in the last 2 iterations plus a final data term iteration (Figure 10). Regardless of the pre-filtering applied to the tilt-series, the result using Topaz Denoise still contains noise in some sections of the reconstruction, and some details in the virus particles (e.g. the particle in the top-left corner) are blurred. Our result presents homogeneous denoising and sharper details.

4.11 Qualitative Comparison with Other Reconstruction and Denoising Methods

4.11.1 Influenza Tilt-Series

We created reconstructions from three tilt-series including the influenza virus particles (dataset 3) using IMOD, IsoNet, and our framework. First, we used IMOD's SIRT algorithm to create reconstructions with 10 and 20 iterations. We selected the result after 10 iterations since it was cleaner and presented better contrast. This is in accordance with what we found when running more than a hundred iterations with only SART in our ablation studies. We used IMOD's tomograms (bin-4 and bin-8) for training IsoNet (3 tomograms were used for training) for 20 iterations with the recommended configuration [19]. Finally, we created reconstructions with our framework, using LADMM for 2×80 iterations with and without NLM in the final two iterations. The results are displayed in Figure 11. IMOD's SIRT allows to reconstruct features in the virus particles, but there is a strong presence of noise in the reconstruction, which would only worsen with more iterations. IsoNet bin-4 results present slightly better contrast than IMOD's, but the noise level is very similar. IsoNet bin-8 achieves better contrast than bin-4 and a moderate level of denoising, but at the expense of image resolution. Our method manages to reconstruct the features of the virus particles and achieves excellent denoising results thanks to the iterative joint reconstruction and denoising nature of our technique.

4.11.2 HIV-1 (EMPIAR-10643)

We created reconstructions from tilt-series including 'HIV-1 GagdeltaMASP1T8I assemblies' (dataset 4) available at the Electron Microscopy Public Image Archive (EMPIAR-10643) [71], [72]. This dataset presents repetitive structures that are expected to work well for Non-Local Means but also any other denoising method based on self-similarity, including classical subtomogram averaging and IsoNet. We created the following reconstructions (volume depth = 380

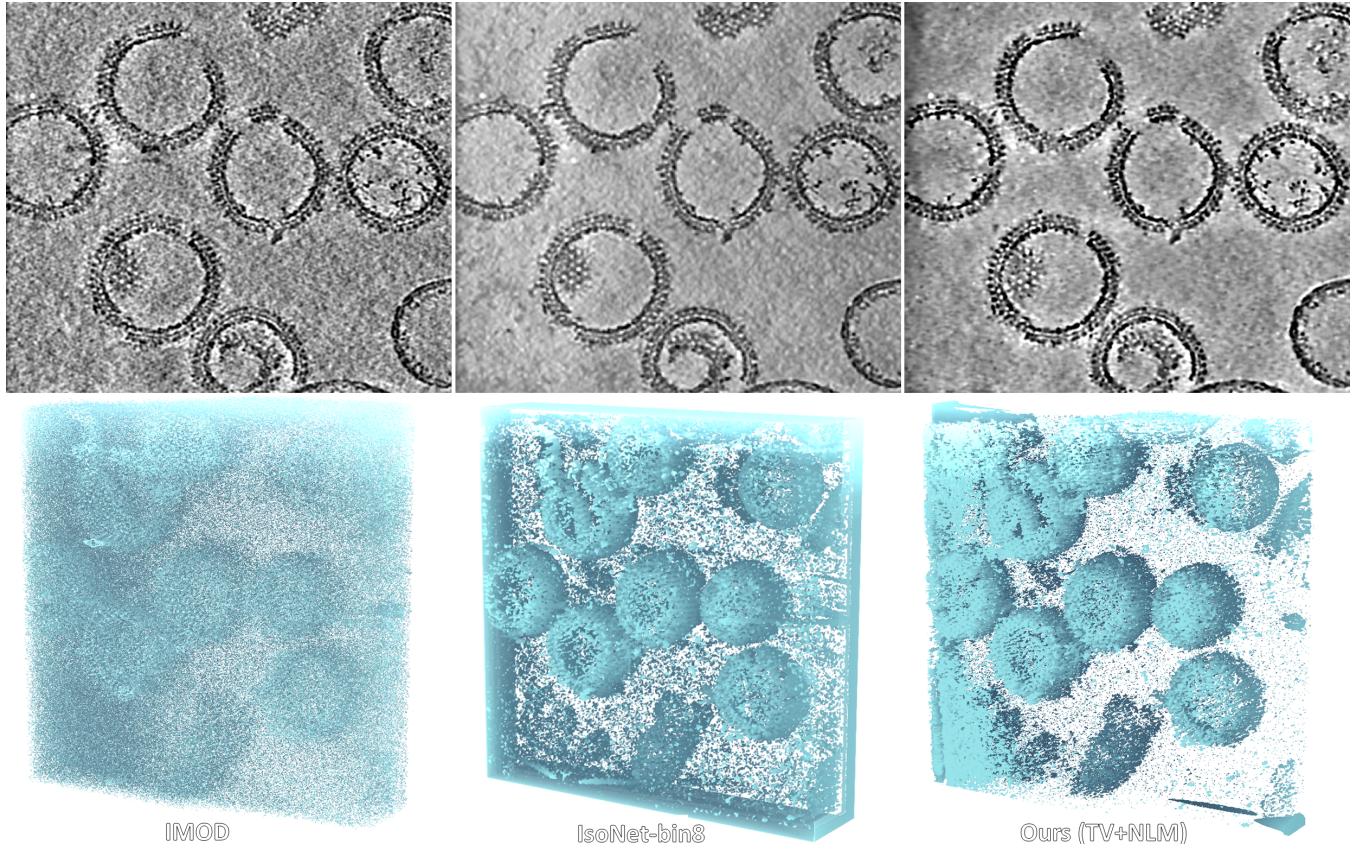


Fig. 12. 2D slices and 3D visualizations of reconstructions from HIV-1 tilt-series (dataset 4, EMPIAR-10643) using different methods. IMOD reconstructed the virus particles, but the 3D visualization is noisy. IsoNet reduces the noise from IMOD's reconstruction and allows for 3D visualization (at half the spatial resolution). Our method achieves cleaner results that produce cleaner 3D visualizations.

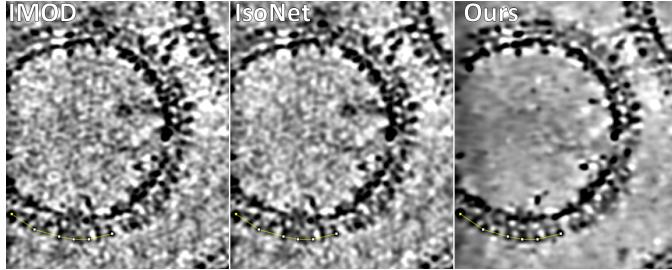


Fig. 13. Line for intensity profile comparison on dataset 4

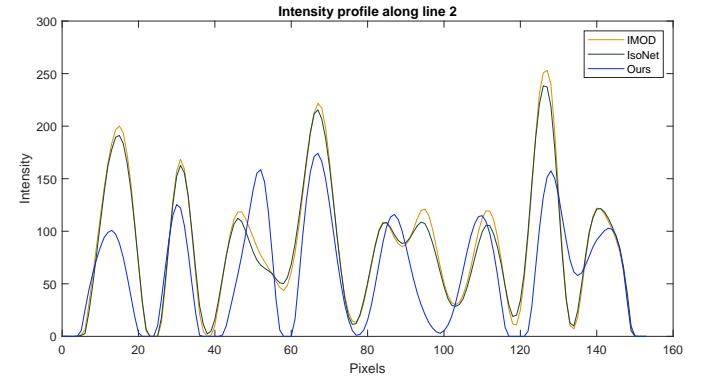


Fig. 14. Plot of the intensity profile along the line. The minimums (corresponding to spikes) are more pronounced in the profile from our reconstruction (blue).

for bin-3, 150 for bin-8): IMOD's SIRT 10 iterations, IsoNet bin-3 (pixel size 3.537 Å) and bin-8 (pixel size 9.432 Å), and ours using TV+NLM (NLM in the last 2 iterations). Additionally, the deconvolution step in IsoNet's framework was used for all the methods. 2D vertical slices and 3D visualizations of the results are depicted in Figure 12. IsoNet bin-3 achieves results very similar to IMOD's reconstruction (see supplementary material for IMOD vs IsoNet bin-3 vs IsoNet bin-8). IsoNet bin-8 achieves good denoising results, but the final result presents lower resolution than the other methods ($1366 \times 1366 \times 380$ vs $512 \times 512 \times 150$). Our technique achieves cleaner and sharper reconstructed features. To compare the sharpness and contrast of the results, we used a method suggested by a field expert. The idea is to plot the intensities along a line to create an intensity profile, and

we chose this dataset since it presents repetitive 'spikes' in the outer part of the virus particles, which is perfect for this method. The volume intensities were normalized from 0 to 255. We perform this analysis for IMOD's, IsoNet's bin-3, and our results on a line across the lower spikes of the leftmost particle in Figure 12. A close-up image included the line is shown in Figure 13, and the respective intensity profiles correspond to Figure 14. In the intensity profiles, the maximum peaks correspond to empty space between the spikes, and the minimum values correspond to the spikes.

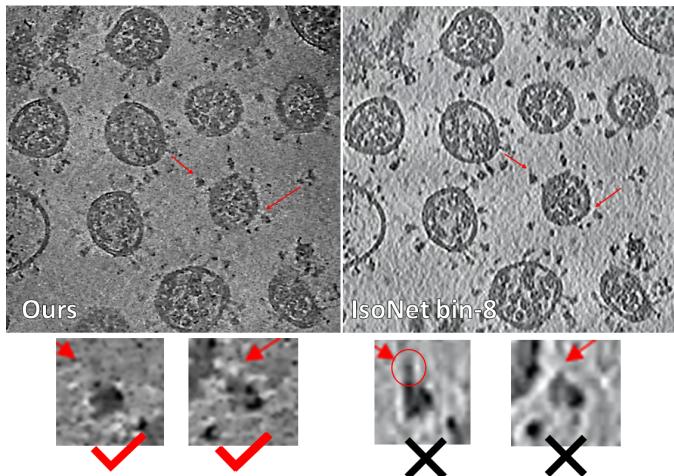


Fig. 15. Comparison between Ours (SART+TV+NLM, bin-4) and IsoNet (bin-8) on dataset 2. The virus inner structure is clearer after IsoNet processing. However, the spike's stem used to connect the membranes should have some density, and the spike's head should not have the encircled density.

For the majority of the spikes, our method presents smaller minimum values (darker spikes). In IMOD's and IsoNet's results there are several saturated blobs in the empty regions resulting from the deconvolution step due to the higher presence of noise, producing higher maximum values in the profile. Our method presents a more homogeneous background with significantly less blobs thanks to the achieved level of denoising. There is a clear visual difference in the reconstructed quality of the spikes between our method and the others, which can be appreciated in the regularity of Ours' intensity profile and at the spike at around 95-th pixel (the sixth minimum in our profile), since the spike is counted as two in the other methods.

4.12 Additional Comparison with IsoNet

We performed additional comparisons between our method (bin-4) and IsoNet (bin-8, half the spatial resolution) using dataset 2. The results are illustrated in Figure 15. IsoNet achieves better denoising results, which allows for a more clear view on the internal structure of the virus particles. However, the spike's stem used to connect the membranes should have some density, and the spike's head should not have the encircled density. Model based regularized approaches can provide more reliable results than learning-based reconstruction or refining methods.

4.13 Example of error visualization

Assuming that the reconstruction process was successful, the error should be uniformly distributed and not correlated with the reconstructed volume. We created two reconstructions of a tilt-series part of the EMPIAR-10643 dataset, one with good alignment, and one with a misaligned version of the same tilt-series. In Figure 16, we compare slices of the error volumes. The left image corresponds to the error volume of the reconstruction created with the aligned tilt-series, and the right image to the error volume of the reconstruction created with the misaligned tilt-series. The images

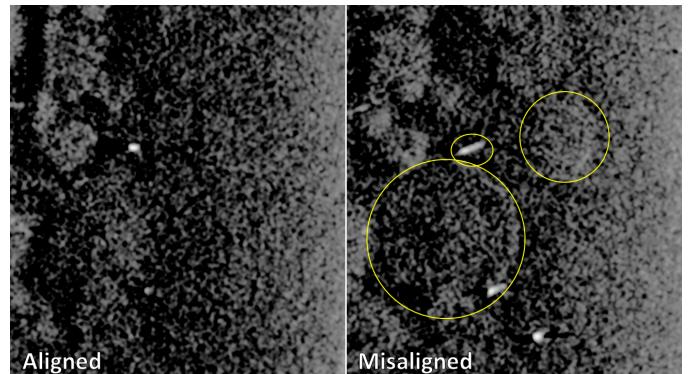


Fig. 16. Slice of the error volume, dataset 4. Left: Reconstructions from aligned tilt-series. Right: Reconstructions from misaligned tilt-series.

were processed as described in Section 3.6 to enhance the structures in the error. The error slice from the misaligned tilt-series contains visible structure in comparison with the aligned version, and there are white diagonal smudges corresponding to fiducials (both encircled in yellow).

We can observe the positive effect of regularization in Figure 17, which includes slices of the error volumes from a non-regularized reconstruction (SART only, left) and a regularized reconstruction (SART+TV+NLM, right). The error corresponding to the non-regularized reconstruction includes structures that clearly resemble the virus particles (encircled in yellow, see Figure 7 for reference). In contrast, the error from the regularized reconstruction presents practically no structure correlated to the volume, which indicates that the reconstructed virion's features obtained with regularization match the features available in the original projection data.

5 CONCLUSION

We have presented a new GPU accelerated framework for tomographic reconstruction of tilt-series from cryogenic electron microscopy. The framework is based on a flexible and robust optimization framework that integrates the reconstruction and denoising processes and achieves sharp denoised reconstructions, while preserving the relationship between final reconstruction and original tilt-series. The framework transforms noisy tilt-series into denoised visualizations in a short time and using a single piece of

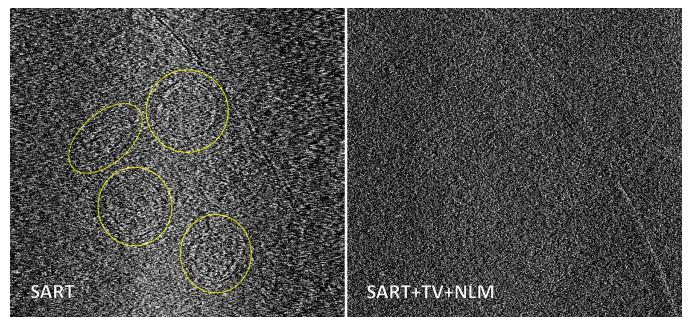


Fig. 17. Slice of the error volume, dataset 1. Left: SART+TV+NLM. Right: SART only. The error from the reconstruction without regularization presents structures correlated to the virus particles.

software. We demonstrate improvements in denoising and reconstruction quality compared to state-of-the-art methods. The framework's source code will become available upon publication. Given the flexibility of the proximal framework, a possible next step could be the use of self-supervised deep-learning-based denoising approaches as additional regularizers in the framework.

ACKNOWLEDGMENTS

This work was supported by the King Abdullah University of Science and Technology through a Competitive Research Grant (CRG), the VCC Center Competitive Funding (CCF), and award BAS/1/1680-01-01; by the Tsinghua University Spring Breeze Fund number 2021Z99CFZ004, and by the National Natural Science Foundation of China number 32171195. The authors would like to thank Rachid Sougrat, Director of Imaging and Characterization Core Lab at KAUST, for his constructive feedback on the evaluation approach used in the domain.

REFERENCES

- [1] P. Suetens, *Fundamentals of Medical Imaging*, 3rd ed. Cambridge University Press, 2017.
- [2] J. M. Ollinger and J. A. Fessler, "Positron-emission tomography," *IEEE signal processing magazine*, vol. 14, no. 1, pp. 43–55, 1997.
- [3] D. L. Bailey, M. N. Maisey, D. W. Townsend, and P. E. Valk, *Positron emission tomography*. Springer, 2005, vol. 2.
- [4] S. A. Huettel, A. W. Song, G. McCarthy *et al.*, *Functional magnetic resonance imaging - 3rd Edition*. Sinauer Associates Sunderland, MA, 2014, vol. 1.
- [5] S. Li, "Cryo-electron tomography of enveloped viruses," *Trends in Biochemical Sciences*, vol. 47, no. 2, pp. 173–186, 2022, special Issue: Pushing boundaries of cryo-EM. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0968000421001857>
- [6] L. Xue, S. Lenz, M. Zimmermann-Kogadeeva, D. Tegunov, P. Cramer, P. Bork, J. Rappoport, and J. Mahamid, "Visualizing translation dynamics at atomic detail inside a bacterial cell," *Nature*, vol. 610, pp. 1–7, 09 2022.
- [7] C. W. Stearns, D. A. Chesler, and G. L. Brownell, "Three dimensional image reconstruction in the fourier domain," *IEEE Transactions on Nuclear Science*, vol. 34, no. 1, pp. 374–378, 1987.
- [8] A. C. Kak and M. Slaney, *Principles of computerized tomographic imaging*. Society for Industrial and Applied Mathematics, 2001.
- [9] F. Noo, M. Defrise, R. Clackdoyle, and H. Kudo, "Image reconstruction from fan-beam projections on less than a short scan," *Physics in medicine & biology*, vol. 47, no. 14, p. 2525, 2002.
- [10] R. Gordon, R. Bender, and G. T. Herman, "Algebraic reconstruction techniques (art) for three-dimensional electron microscopy and x-ray photography," *Journal of Theoretical Biology*, vol. 29, no. 3, pp. 471–481, Dec. 1970.
- [11] P. Gilbert, "Iterative methods for the three-dimensional reconstruction of an object from projections," *Journal of Theoretical Biology*, vol. 36, no. 1, pp. 105–117, Jul. 1972.
- [12] A. H. Andersen and A. C. Kak, "Simultaneous algebraic reconstruction technique (sart): a superior implementation of the art algorithm," *Ultrasonic Imaging*, vol. 6, no. 1, pp. 81–94, Jan. 1984.
- [13] T. Würfl, M. Hoffmann, V. Christlein, K. Breininger, Y. Huang, M. Unberath, and A. K. Maier, "Deep learning computed tomography: Learning projection-domain weights from image domain in limited angle problems," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1454–1463, 2018.
- [14] G. Wang, J. C. Ye, K. Mueller, and J. A. Fessler, "Image reconstruction is a new frontier of machine learning," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1289–1296, 2018.
- [15] M. Su, H. Zhang, K. Schawinski, C. Zhang, and M. A. Cianfrocco, "Generative adversarial networks as a tool to recover structural information from cryo-electron microscopy data," *BioRxiv*, p. 256792, 2018.
- [16] T. Bepler, K. Kelley, A. J. Noble, and B. Bonnie, "Topaz-denoise: general deep denoising models for cryoem and cryoet," *Nature Communications*, vol. 11, no. 5208, p. 838920, Oct. 2020.
- [17] E. Palovcak, D. Asarnow, M. G. Campbell, Z. Yu, and Y. Cheng, "Enhancing snr and generating contrast for cryo-em images with convolutional neural networks," *bioRxiv*, 2020.
- [18] T. Wang, Y. Lei, Z. Tian, X. Dong, Y. Liu, X. Jiang, W. J. Curran, T. Liu, H.-K. Shu, and X. Yang, "Deep learning-based image quality improvement for low-dose computed tomography simulation in radiation therapy," *Journal of Medical Imaging*, vol. 6, no. 4, p. 043504, 2019.
- [19] Y.-T. Liu, H. Zhang, H. Wang, C.-L. Tao, G.-Q. Bi, and Z. H. Zhou, "Isotropic reconstruction for electron tomography with deep learning," *Nature communications*, 2022.
- [20] T. F. Budinger, G. T. Gullberg, and R. H. Huesman, "Emission computed tomography," in *Image reconstruction from projections. Implementaton and applications*. Springer-Verlag, 1979, pp. 147–246.
- [21] J. Hsieh, *Computed Tomography: Principles, Design, Artifacts, and Recent Advances*. SPIE PRESS, 2015.
- [22] R. N. Bracewell and A. C. Riddle, "Inversion of fan-beam scans in radio astronomy," *Astrophysical Journal*, vol. 150, p. 427, Nov. 1967.
- [23] L. A. Feldkamp, L. C. Davis, and J. W. Kress, "Practical cone-beam algorithm," *Josaa*, vol. 1, no. 6, pp. 612–619, 1984.
- [24] M. Radermacher, "Weighted back-projection methods," in *Electron tomography*. Springer, 1992, pp. 245–273.
- [25] A. Andersen, "Algebraic reconstruction in ct from limited views," *IEEE Transactions on Medical Imaging*, vol. 8, no. 1, pp. 50–55, Mar. 1989.
- [26] C. O. S. Sorzano, J. Vargas, J. Otón, J. de la Rosa-Trevín, J. Vilas, M. Kazemi, R. Melero, L. Del Caño, J. Cuenca, P. Conesa *et al.*, "A survey of the use of iterative reconstruction algorithms in electron microscopy," *BioMed research international*, vol. 2017, 2017.
- [27] E. Pyle and G. Zanetti, "Current data processing strategies for cryo-electron tomography and subtomogram averaging," *Biochemical Journal*, vol. 478, no. 10, pp. 1827–1845, May 2021.
- [28] J. R. Kremer, D. N. Mastronarde, and J. R. McIntosh, "Computer visualization of three-dimensional image data using imod," *Journal of Structural Biology*, vol. 116, no. 1, pp. 71–76, Jan. 1996.
- [29] P. C. Hansen and J. S. Jørgensen, "Air tools ii: algebraic iterative reconstruction methods, improved implementation," *Numerical Algorithms*, vol. 79, no. 1, pp. 107–137, Nov. 2018.
- [30] W. van Aarle, W. J. Palenstijn, J. De Beenhouwer, T. Altantzis, S. Bals, K. J. Batenburg, and J. Sijbers, "The astra toolbox: A platform for advanced algorithm development in electron tomography," *Ultramicroscopy*, vol. 157, pp. 35–47, 2015.
- [31] W. van Aarle, W. J. Palenstijn, J. Cant, E. Janssens, F. Bleichrodt, A. Dabrowski, J. D. Beenhouwer, K. J. Batenburg, and J. Sijbers, "Fast and flexible x-ray tomography using the astra toolbox," *Opt. Express*, vol. 24, no. 22, pp. 25129–25147, Oct 2016.
- [32] D. Gürsoy, F. De Carlo, X. Xiao, and C. Jacobsen, "Tomopy: a framework for the analysis of synchrotron tomographic data," *Journal of Synchrotron Radiation*, vol. 21, no. 5, pp. 1188–1193, Sep 2014.
- [33] R. Han, X. Wan, Z. Wang, Y. Hao, J. Zhang, Y. Chen, X. Gao, Z. Liu, F. Ren, F. Sun, and F. Zhang, "Autom: A novel automatic platform for electron tomography reconstruction," *Journal of Structural Biology*, vol. 199, no. 3, pp. 196–208, Sep. 2017.
- [34] L. Donati, "Reconstruction methods for cryo-electron microscopy: From model-based to data-driven," PhD Thesis, EPFL, Lausanne, 2020.
- [35] M. Turk and W. Baumeister, "The promise and the challenges of cryo-electron tomography," *FEBS Letters*, vol. 594, no. 20, pp. 3243–3261, 2020.
- [36] P. Trampert, W. Wang, D. Chen, R. B. Ravelli, T. Dahmen, P. J. Peters, C. Kübel, and P. Slusallek, "Exemplar-based inpainting as a solution to the missing wedge problem in electron tomography," *Ultramicroscopy*, vol. 191, pp. 1–10, 2018.
- [37] R. Yan, S. V. Venkatakrishnan, J. Liu, C. A. Bouman, and W. Jiang, "Mbir: A cryo-3d reconstruction method that effectively minimizes missing wedge artifacts and restores missing information," *Journal of structural biology*, vol. 206, no. 2, pp. 183–192, May 2019.
- [38] G. Ding, Y. Liu, R. Zhang, and H. L. Xin, "A joint deep learning model to recover information and reduce artifacts in missing-wedge sinograms for electron tomography and beyond," *Scientific reports*, vol. 9, no. 1, pp. 1–13, 2019.

- [39] J. Xu, Y. Zhao, H. Li, and P. Zhang, "An image reconstruction model regularized by edge-preserving diffusion and smoothing for limited-angle computed tomography," *Inverse Problems*, vol. 35, no. 8, jul 2019.
- [40] J. Frikel, "Sparse regularization in limited angle tomography," *Applied and Computational Harmonic Analysis*, vol. 34, no. 1, pp. 117–141, Jan. 2013. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1063520312000449>
- [41] H. Yao, Y. Song, Y. Chen, N. Wu, J. Xu, C. Sun, J. Zhang, T. Weng, Z. Zhang, Z. Wu, L. Cheng, D. Shi, X. Lu, J. Lei, M. Crispin, Y. Shi, L. Li, and S. Li, "Molecular architecture of the sars-cov-2 virus," *Cell*, vol. 183, no. 3, pp. 730–738, 2020. [Online]. Available: <https://www.ebi.ac.uk/emdb/EMD-33297>
- [42] A. S. Frangakis, "It's noisy out there! a review of denoising techniques in cryo-electron tomography," *Journal of Structural Biology*, vol. 213, no. 4, p. 107804, Dec. 2021.
- [43] B. Goris, T. Roelandts, K. Batenburg, H. Heidari Mezerji, and S. Bals, "Advanced reconstruction algorithms for electron tomography: From comparison to combination," *Ultramicroscopy*, vol. 127, pp. 40–47, 2013.
- [44] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, pp. 259–268, 1992. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/016727899290242F>
- [45] J. Böhning, T. A. Bharat, and S. M. Collins, "Compressed sensing for electron cryotomography and high-resolution subtomogram averaging of biological specimens," *Structure*, vol. 30, no. 3, pp. 408–417.e4, 2022.
- [46] S. V. Venkatakrishnan, L. F. Drummy, M. Jackson, M. De Graef, J. Simmons, and C. A. Bouman, "Model-based iterative reconstruction for bright-field electron tomography," *IEEE Transactions on Computational Imaging*, vol. 1, no. 1, pp. 1–15, 2015.
- [47] N. Parikh and S. Boyd, "Proximal algorithms," *Foundations and Trends in Optimization*, vol. 1, no. 3, pp. 127–239, Jan. 2014.
- [48] P. J. Huber, "Robust estimation of a location parameter," *The Annals of Mathematical Statistics*, vol. 35, no. 1, pp. 73–101, 1964.
- [49] A. Buades, B. Coll, and J.-M. Morel, "Non-local means denoising," *Image Processing On Line*, vol. 1, pp. 208–212, Sep. 2011.
- [50] C. R. Johnson and A. R. Sanderson, "A next step: Visualizing errors and uncertainty," *IEEE Computer Graphics and Applications*, vol. 23, no. 5, pp. 6–10, Oct. 2003.
- [51] K. Brodlie, R. Allendes Osorio, and A. Lopes, "A review of uncertainty in data visualization," *Expanding the frontiers of visual analytics and visualization*, pp. 81–109, 2012.
- [52] G.-P. Bonneau, H.-C. Hege, C. R. Johnson, M. M. Oliveira, K. Potter, P. Rheingans, and T. Schultz, "Overview and state-of-the-art of uncertainty visualization," in *Scientific Visualization*. Springer, 2014, pp. 3–27.
- [53] A. Kamal, P. Dhakal, A. Y. Javaid, V. K. Devabhaktuni, D. Kaur, J. Zaiantz, and R. Marinier, "Recent advances and challenges in uncertainty visualization: A survey," *Journal of Visualization*, vol. 24, no. 5, pp. 861–890, 2021.
- [54] C. Gillmann, D. Saur, T. Wischgoll, and G. Scheuermann, "Uncertainty-aware visualization in medical imaging - a survey," *Computer Graphics Forum*, vol. 40, no. 3, pp. 665–689, 2021.
- [55] J. M. Kniss, R. Van Uitert, A. Stephens, G.-S. Li, T. Tasdizen, and C. Hansen, "Statistically quantitative volume visualization," in *VIS '05. IEEE Visualization, 2005*. IEEE, 2005, pp. 287–294.
- [56] C. Lundström, P. Ljung, A. Persson, and A. Ynnerman, "Uncertainty visualization in medical volume rendering using probabilistic animation," *IEEE transactions on visualization and computer graphics*, vol. 13, no. 6, pp. 1648–1655, 2007.
- [57] T. K. Nguyen, H. Ohlsson, A. Eklund, F. Hernell, P. Ljung, C. Forsell, M. Andersson, H. Knutsson, and A. Ynnerman, "Concurrent volume visualization of real-time fmri," in *8th IEEE/EG International Symposium on Volume Graphics, Norrköping, Sweden, 2–3 May, 2010*. The Eurographics Association, 2010, pp. 53–60.
- [58] S. Lindholm, P. Ljung, C. Lundström, A. Persson, and A. Ynnerman, "Spatial conditioning of transfer functions using local material distributions," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 6, pp. 1301–1310, 2010.
- [59] V. Dinesha, N. Adabala, and V. Natarajan, "Uncertainty visualization using hdr volume rendering," *The Visual Computer*, vol. 28, no. 3, pp. 265–278, 2012.
- [60] A. C. Kak and M. Slaney, "Principles of computerized tomographic imaging," *SIAM*, pp. 37–45, 2001.
- [61] G. Zang, M. Aly, R. Idoughi, P. Wonka, and W. Heidrich, "Super-resolution and sparse view ct reconstruction," *European Conference on Computer Vision*, pp. 145–161, 2018.
- [62] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, "Plug-and-play priors for model based reconstruction," in *2013 IEEE Global Conference on Signal and Information Processing*, 2013, pp. 945–948.
- [63] F. Heide, M. Steinberger, Y. Tsai, M. Rouf, D. Pajak, D. Reddy, O. Gallo, J. Liu, W. Heidrich, K. Pajak, J. Kautz, and K. Pulli, "Flexisp: A flexible camera image processing framework," *ACM Transactions on Graphics (Proc. SIGGRAPH Asia)*, vol. 33, no. 6, pp. 1–13, 2014.
- [64] J. Kuba, J. Mitchels, M. Hovorka, P. Erdmann, L. Berka, R. Kirmse, J. König, J. De Bock, B. Goetze, and A. Rigort, "Advanced cryotomography workflow developments – correlative microscopy, milling automation and cryo-lift-out," *Journal of Microscopy*, vol. 281, no. 2, pp. 112–124, 2021.
- [65] A. Appel, "Some techniques for shading machine renderings of solids," *AFIPS Fall Joint Computing Conference*, 1968.
- [66] L. A. Westover, "Splatting: A parallel, feed-forward volume rendering algorithm," PhD Thesis, University of North Carolina at Chapel Hill, 1991.
- [67] B. Trifonov, D. Bradley, and W. Heidrich, "Tomographic reconstruction of transparent objects," in *Eurographics Symposium on Rendering*. The Eurographics Association, 2006.
- [68] N. Max, "Optical models for direct volume rendering," *IEEE Transactions on Visualization and Computer Graphics*, vol. 1, no. 2, pp. 99–108, 1995.
- [69] E. Veach, "Robust monte carlo methods for light transport simulation," PhD Thesis, Stanford University, Stanford, CA, USA, 1997.
- [70] F. Hernell, P. Ljung, and A. Ynnerman, "Local ambient occlusion in direct volume rendering," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 4, pp. 548–559, Jul. 2010.
- [71] A. Iudin, P. Korir, J. Salavert-Torres, G. Kleywegt, and P. A., "Empiar: A public archive for raw electron microscopy image data," *Nature Methods*, vol. 13, no. 5, pp. 387–388, 2016.
- [72] T. Ni, T. Frosio, L. Mendonça, Y. Sheng, D. Clare, B. A. Himes, and P. Zhang, "High-resolution in situ structure determination by cryo-electron tomography and subtomogram averaging using emclarify," *Nature Protocols*, vol. 17, pp. 421–444, 2022.



Julio Rey studied Electronics Engineering at the National University of Colombia (UNAL) from 2012 to 2018, where he worked on photometry and colorimetry of light sources at the Lighting Testing Lab (LABE - UNAL) from 2015 to 2018. He received his Masters degree in Computer Science at the King Abdullah University of Science and Technology (KAUST) in 2022. Since 2022, he is a Computer Science Ph.D student at KAUST, where he is part of the High-Performance Visualization Group (VCCVIS).



Peter Rautek received his MSc and PhD in computer science from Vienna University of Technology in 2005 and 2009, respectively. He worked as a research assistant and postdoctoral fellow at the same university between 2005 and 2011. Since 2011 until present, he is a research scientist at KAUST, Saudi Arabia, focused on interactive large-scale data visualization, virtual reality, domain-specific languages for visualization, and flow visualization.



Ciril Bohak is a research scientist at King Abdullah University of Science and Technology (KAUST), Saudi Arabia, and an assistant professor in the Faculty of Computer and Information Science, University of Ljubljana, Slovenia. He received his B.Sc., M.Sc., and Ph.D. from the University of Ljubljana. His research covers computer graphics, scientific visualization, and human-computer interaction.



Wolfgang Heidrich (Fellow, IEEE) received the Ph.D. degree from the University of Erlangen, Erlangen, Germany, in 1999. He was a Research Associate with the Computer Graphics Group of the Max-Planck-Institute for Computer Science in Saarbrücken, Germany, before joining UBC in 2000. He is currently a Professor of computer science with the King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia. He joined KAUST in 2014, after 13 years as a Faculty Member with the University of British Columbia, Vancouver, BC, Canada. His research interests include intersection of imaging, optics, computer vision, computer graphics, and inverse problems. His current research interests include computational imaging, focusing on hardware-software codesign of the next generation of imaging systems, with applications such as High-Dynamic Range imaging, compact computational cameras, hyperspectral cameras, to name just a few. He work on High-Dynamic Range Displays served as the basis for the technology behind Brightside Technologies, which was acquired by Dolby in 2007. He has chaired the papers program for both Siggraph Asia and the International Conference of Computational Photography among others. He was the recipient of a 2014 Humboldt Research Award.



Ondřej Strnad is a research scientist at King Abdullah University of Science and Technology (KAUST), Saudi Arabia. He received his Ph.D. from Masaryk University in Brno, Czech Republic, in 2014. His research interests span scientific visualization, geometry algorithms, and computer graphics. Recently, he joined the NANOVIS group at KAUST to work on technologies that deliver new visualizations and techniques for mesoscale biological models.



Zheyuan Zhang is a Research Assistant at Tsinghua University, China. He received a B.Eng. degree in Communication Engineering from Beijing Union University, China in 2020. His research interests include cryo-electron tomography and microscopic image processing.



Sai Li is an Associate Professor at Tsinghua University, China. He received a B.Sc. degree in applied physics from Wuhan University, China in 2006, and Ph.D. in biophysics from the University of Goettingen, Germany, in 2012. He worked as a postdoctoral and senior research scientist at the Oxford Particle Imaging Centre, University of Oxford, U.K. until 2018. His research interests include cryo-electron tomography, 3D reconstruction, and structural virology of pathogenic enveloped viruses.



Ivan Viola is a professor at King Abdullah University of Science and Technology (KAUST), Saudi Arabia. He graduated from TU Wien, Austria, in 2005 he took a postdoc position at the University of Bergen, Norway, where he was gradually promoted to the professor rank. In 2013 he received a WWTF grant to establish a research group at TU Wien. At KAUST, he continues developing new technologies that make visual, in-silico life at nanoscale possible. Viola co-founded the Nanographics startup to commercialize nanovisualization technologies.