CRYOWAVEGS: WAVE-OPTICAL GAUSSIAN SPLATTING FOR CRYO-ELECTRON TOMOGRAPHY

Anonymous authors
Paper under double-blind review

ABSTRACT

Gaussian Splatting has recently been adapted to tomographic reconstruction, but existing formulations assume a ray model that discards the depth at which scattering occurs. This approximation fails past a computable switch point that the resolution of current cryo-ET reconstructions does not reach, and that atomic-resolution series do. We introduce CryoWaveGS, a Gaussian Splatting reconstruction with a wave-optical forward model. Because Fresnel propagation maps a Gaussian to a Gaussian with complex covariance, depth-resolved diffraction can be rendered in a single rasterization pass at a cost similar of the ray-based renderer, and within $1.3\times$ end to end on our densest run. We validate the renderer against independent multislice simulation and map where the ray approximation is adequate and where it is not. We evaluate on measured cryo-ET and atomic-resolution tilt series against conventional, implicit-field and matched ray-based baselines under a single scoring pipeline.

1 INTRODUCTION

Cryo-electron tomography (cryo-ET) reconstructs a three-dimensional specimen from a *tilt series*, a set of transmission electron micrographs recorded while the specimen stage is tilted, typically over $\pm 60^\circ$ in steps of 2–3°. Two facts dominate the inverse problem. The electron dose is severely limited, so the signal-to-noise ratio of each micrograph is far below one. The limited tilt range also leaves a *missing wedge* of unmeasured directions in Fourier space, which makes the resolution anisotropic and worst along the beam axis. Resolution is therefore best reported per direction and not as a single average over the volume.

Classical reconstruction treats each micrograph as a *line integral* of a scalar scattering potential along the beam. Weighted back-projection (WBP) (Radermacher, 2006) filters each tilt image and smears it back through the volume along its viewing direction. The simultaneous iterative and algebraic reconstruction techniques SIRT (Gilbert, 1972) and SART (Andersen & Kak, 1984) solve the same linear system iteratively, by repeatedly back-projecting the difference between measured and predicted projections. All three are fast, linear and blind to depth. An electron scattered near the entrance surface of the specimen contributes to the image in the same way as one scattered near the exit surface. Real image formation is governed by wave optics. Electrons scatter, propagate and interfere, and the contrast transfer function (CTF) of the microscope mixes amplitude and phase depending on the defocus, which varies with depth inside a tilted specimen. Vulović et al. (2014) derive validity conditions for the line-integral (projection) model and for the weak-phase object approximation. Both hold for tomograms coarser than about 30 Å but begin to fail near 5 Å for typical macromolecules (1 Å $= 0.1$ nm). As cryo-ET moves toward finer resolution, the ray model approaches the limit of its validity.

The multislice method (Kirkland, 2010) computes the exit wave accurately, and PhaseT3M (Lee, 2026) already inverts it for cryo-ET. Multislice, however, propagates the wave through a dense stack of planes one after another, so its cost grows with the number of planes and not with the complexity of the specimen. We instead carry the propagation inside each primitive of the representation, which for a compact set of primitives costs as little as the ray renderer.

3D Gaussian Splatting (Kerbl et al., 2023) represents a scene by anisotropic Gaussians that are rasterized in one differentiable pass. R^2-Gaussian (Zha et al., 2024) adapted it to X-ray CT and CryoETGS (Zhang, 2026) to cryo-ET. Gaussians suit tomography because integrating a 3D Gaussian

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

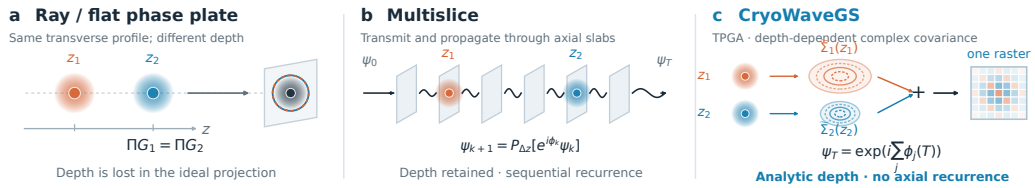

Figure 1: Wave optics relevance. *Left:* the ray model collapses two Gaussians at different z into the same ideal projection. *Middle:* multislice recovers the difference, but requires a sequential recurrence through many axial slabs. *Right:* CryoWaveGS assigns each primitive its own depth-dependent complex covariance and rasterizes once.

along the beam yields a 2D Gaussian in closed form, so the renderer never samples points along a ray. Our starting observation is that wave propagation also preserves Gaussians. Propagating a wave over a distance z convolves it with the kernel $\exp(i\pi\|\mathbf{r}\|^2/\lambda z)$, where $\mathbf{r}$ is the lateral position and λ the electron wavelength. This kernel is a Gaussian with an imaginary variance, and the convolution of two Gaussians is a Gaussian. A propagated Gaussian therefore stays a Gaussian, and z enters only as an *imaginary term added to its covariance*. Depth then no longer needs to be discretized into planes and becomes a parameter of each primitive (Figure 1).

We present **CryoWaveGS**, a cryo-ET reconstruction method that keeps the explicit, adaptive Gaussian representation but replaces the line-integral forward model with a wave-optical one. Each Gaussian is rendered with the *thick-phase-grating approximation* (TPGA) of Vulović et al. (2014), which gives it a projected covariance that is complex and depends on its depth. Summing all contributions yields the exit wave of an entire tilt in one rasterization pass, after which the transfer function of the microscope and the detector model are applied as after a multislice simulation. Setting the wavelength inside the specimen to zero turns the renderer into a ray renderer, so comparing ray-based and wave-based reconstructions of the same data shows when depth-resolved physics changes the result. Our contributions are:

1. **Wave-optical Gaussian splatting:** a TPGA forward model in which each Gaussian carries a complex covariance encoding the defocus of its own depth, so that depth-resolved diffraction is rendered in one rasterization pass, with no axial planes and no multislice recursion.
2. **A cryo-ET reconstruction technique with two image formation models:** differentiable Gaussian reconstruction with joint refinement of alignment, defocus and an analytic ice model, rendering either by line integral or by depth-aware wave propagation. The line integral is the $\lambda \to 0$ limit of the same renderer, so the two differ only in propagation.
3. **Validation against a gold standard and matched ray-versus-wave reconstructions:** against converged multislice from an independent code, the wave renderer keeps an image correlation of about 0.9 down to 1.6 Å, while the ray renderer fails below about 5 Å. In reconstruction the two agree to the third decimal on cryo-ET data and separate on an atomic-resolution series, with a Fourier correlation of 0.85 versus 0.42 to the PhaseT3M reference at 2 Å.

2 Background

Let $V(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^3$, be the scattering potential of the specimen and $R_\theta \in \mathrm{SO}(3)$ the rotation that tilts it by θ. With detector coordinates (u, v) and the coordinate z along the beam, called depth, the projection assumption states that each micrograph records the line integral

$$P_\theta(u, v) \;=\; \int V\big(R_\theta^{-1}[u, v, z]^\top\big)\, dz. \tag{1}$$

The integration removes the depth at which a contribution originated, so WBP, SIRT and SART, which all invert equation 1, cannot distinguish scatterers at different depths.

A micrograph is not fundamentally a measurement of equation 1. The incident electron is a coherent wave, and biological specimens are nearly transparent to it. They absorb few electrons and instead

shift the phase of the wave in proportion to the electrostatic potential along its path. The detector records intensity and is blind to phase, so contrast arises only because propagation and a deliberately defocused objective lens convert phase variations into intensity variations. Both are interference effects that depend on distance. Let $\mathbf{q} \in \mathbb{R}^2$ be the spatial frequency in the detector plane, λ the relativistic electron wavelength (1.97 pm at 300 kV) and $\mathcal{F}\{f\}(\mathbf{q}) = \int f(\mathbf{r})\, e^{-2\pi i\, \mathbf{q}\cdot\mathbf{r}}\, d\mathbf{r}$ the Fourier transform used throughout. Paraxial propagation over a distance Δz multiplies each spatial frequency of the wave by the Fresnel propagator $P_{\Delta z}(\mathbf{q}) = \exp(-i\pi\lambda\Delta z\|\mathbf{q}\|^2)$. Its phase grows with distance, so identical scatterers near the entrance and the exit surface reach the detector with different phases. The objective lens then multiplies the spectrum of the *exit wave* ψ_{exit} leaving the specimen by the CTF $H_\theta(\mathbf{q}) = E_\theta(\mathbf{q}) \exp[-i\chi_\theta(\mathbf{q})]$, where the phase χ_θ collects defocus and lens aberrations and the envelope E_θ the damping by partial coherence and the detector. The detector records

$$I_\theta = \left| \mathcal{F}^{-1}\left[H_\theta \odot \mathcal{F}\{\psi_{\text{exit}}\} \right] \right|^2, \tag{2}$$

where $\odot$ denotes pointwise multiplication. The squared modulus makes image formation nonlinear in the potential, and because the phase of a scatterer depends on its depth, the depth that equation 1 integrates away is measurable. The multislice method (Kirkland, 2010) simulates this process by alternating the phase shifts of thin slabs with Fresnel propagation between them, and it is the reference for our approximation.

How thick may a specimen be before depth matters? A scatterer at depth Δz from the plane in which the image is formed acquires, relative to its projection, the Fresnel phase $\pi\lambda\Delta z\|\mathbf{q}\|^2$. The projection assumption holds while this phase is small, $\lambda\Delta z\|\mathbf{q}\|^2 \ll 1$, for every depth in the specimen and every frequency to be resolved (Vulović et al., 2014). For a target period $d = 1/\|\mathbf{q}\|$ this bounds the thickness by the depth of field

$$t_{\text{crit}}(d) \;=\; d^2/\lambda. \tag{3}$$

A specimen thinner than $t_{\text{crit}}(d)$ can be reconstructed at period d with the ray model, while a thicker one requires wave optics. Since t_{crit} grows with d^2, coarse reconstructions tolerate thick specimens and fine ones do not. Stricter forms such as $\|\mathbf{q}\| \ll (4\lambda\Delta z)^{-1/2}$ (Vulović et al., 2014) differ by a constant factor. We use equation 3 because the switch points in Section 5.3 follow it. Binning a tilt series by k multiplies the pixel size by k, and the finest period is twice the pixel size (bin 1 is native sampling). At 300 kV and a native pixel of 1.5 Å, t_{crit} is about 740 nm at bin 4 ($d = 12.1$ Å), 180 nm at bin 2 (6.0 Å) and 46 nm at bin 1 (3.0 Å). For lamellae of 100–300 nm the ray model is thus adequate at bin 4, borderline at bin 2 and violated at native sampling, at least on paper. In practice what counts is the finest period the data resolve (Section 5.2). CryoETGS (Zhang, 2026) reconstructs at bin 4 and 8, well inside the ray regime.

The criterion therefore decides whether a given specimen and target resolution need wave optics. CryoWaveGS provides a ray renderer for the regime it admits and a wave renderer beyond, and the two share every other component (Section 4). The separate weak-phase object approximation (WPOA) requires the phase shift σV_z to be small, where σ is the interaction constant of the electron and V_z the potential projected along the beam, with $\sigma V_z < 0.36$ a common worst-case bound. A weakly scattering specimen can still violate the projection assumption, because diffraction accumulates during propagation.

The depth-dependent phase also has a simple picture in Fourier space. By the Fourier-slice theorem, the 2D Fourier transform of a projection equals a flat central slice through the 3D Fourier transform of the specimen. At finite wavelength, the waves that reach the detector sample the 3D transform on a curved surface instead, the Ewald sphere, which near the beam axis is a paraboloid that departs from the flat slice by $q_z = -\lambda\|\mathbf{q}\|^2/2$ along the beam. Sampling on this paraboloid is equivalent to imaging every scatterer at a defocus given by its depth. NovaCTF (Turoňová et al., 2017) corrects for this depth-dependent defocus in conventional reconstruction by dividing the tomogram into slabs and correcting each with the CTF of its mean depth, which yielded a 3.4 Å subtomogram average. CryoWaveGS builds the effect into the forward model, where every scatterer is imaged at the defocus of its own continuous depth.

3 Related Work

Weighted Backprojection (WBP), Simultaneous Iterative Reconstruction Technique (SIRT) and Simultaneous Algebraic Reconstruction Technique (SART) remain the standard reconstruction algo-

rithms (Kremer et al., 1996; Zheng et al., 2022). As baselines they separate gains due to a new representation or forward model from gains due to iterative fitting alone. NovaCTF (Turoňová et al., 2017) and related 3D-CTF methods account for depth-dependent defocus as a correction within a linear reconstruction, not as part of a wave-optical forward model.

Learned methods change the representation or add priors while keeping the ray model. ICE-TIDE (Debarnot et al., 2025) represents the specimen as a coordinate-based implicit neural field and jointly estimates the alignment, which makes it the learned method closest to our setting. IsoNet (Liu et al., 2022) and DeepDeWedge (Wiedemann & Heckel, 2024) learn to fill in the missing wedge of an existing reconstruction and leave the forward model unchanged.

R^2-Gaussian (Zha et al., 2024) adapted Gaussian Splatting to tomographic reconstruction. It supports parallel-beam geometry and removed the integration bias that arises when a graphics splatting renderer is reused for line-integral imaging. CryoETGS (Zhang, 2026) adapts it to cryo-ET. Exact-GS (Yang, 2026) removes the affine approximation that splatting makes under cone-beam projection. For the parallel beam of cryo-ET this approximation is already exact, so Exact-GS leaves cryo-ET reconstructions unchanged. All of these methods use a ray forward model.

PhaseT3M (Lee, 2026) is the most closely related work. It embeds multislice in iterative electron tomography with Bayesian alignment and aberration correction, and it optimizes a stack of slices through the sequential multislice recursion. CryoWaveGS optimizes an adaptive set of Gaussians with a renderer that needs no recursion over depth. Outside electron microscopy, Gaussian Wave Splatting (Choi, 2025) and Complex-Valued Holographic Radiance Fields (Zhan, 2026) carry wave optics inside Gaussian rasterization to render holograms of known scenes. Moško et al. (2025) reconstruct X-ray phase-contrast tomograms with Gaussians but propagate the exit wave of the whole specimen in a single Fresnel step, which treats the specimen as thin. CryoWaveGS propagates each Gaussian from its own depth and so models diffraction inside a thick specimen.

4 CryoWaveGS

We represent the scattering potential by N anisotropic Gaussians,

$$V(\mathbf{x}) \;=\; \sum_{k=1}^{N} m_k \,\mathcal{N}(\mathbf{x};\boldsymbol{\mu}_k,\Sigma_k), \qquad \Sigma_k = R_k\,\mathrm{diag}(s_{k,x}^2, s_{k,y}^2, s_{k,z}^2)\,R_k^\top, \tag{4}$$

with center $\boldsymbol{\mu}_k \in \mathbb{R}^3$, rotation $R_k \in \mathrm{SO}(3)$, scales $s_{k,x}, s_{k,y}, s_{k,z} > 0$ and mass $m_k > 0$. Since $\mathcal{N}$ integrates to one, m_k is the integral of primitive k. A primitive can therefore be widened, or split into two whose masses sum to m_k, without changing the integrated potential $\int V\,d\mathbf{x} = \sum_k m_k$.

For tilt θ we rotate each primitive into beam coordinates $\mathbf{x} = (u, v, z)$, replacing $\boldsymbol{\mu}_k$ by $R_\theta\boldsymbol{\mu}_k$ and Σ_k by $R_\theta\Sigma_k R_\theta^\top$ while keeping the symbols, and partition $\boldsymbol{\mu}_k = [\boldsymbol{\mu}_{k,\perp}^\top,\, \mu_{k,z}]^\top$ and $\Sigma_k = \left[\begin{smallmatrix} A_k & \mathbf{b}_k \\ \mathbf{b}_k^\top & c_k \end{smallmatrix}\right]$. Here $\boldsymbol{\mu}_{k,\perp} \in \mathbb{R}^2$ is the lateral position, $\mu_{k,z}$ the depth, A_k the 2×2 lateral covariance that sets the footprint on the detector, $\mathbf{b}_k \in \mathbb{R}^2$ the covariances between the lateral coordinates and depth, which are nonzero when the primitive is tilted relative to the beam, and c_k the variance along the beam. Integrating a primitive along the beam as in equation 1 yields its marginal over depth, the 2D Gaussian $m_k\mathcal{N}(\mathbf{u};\boldsymbol{\mu}_{k,\perp}, A_k)$ with $\mathbf{u} = (u, v)$ (Zha et al., 2024), and the ray renderer sums these footprints. They do not depend on $\mu_{k,z}$, $\mathbf{b}_k$ or c_k, so two primitives that differ only in depth produce identical images. The wave renderer removes this invariance.

Depth-resolved rendering with a complex covariance. TPGA (Vulović et al., 2014) writes the exit wave as $\psi_{\mathrm{exit}} = \exp(i\sigma U_\theta)$ with the *propagated potential*

$$U_\theta = \mathcal{F}^{-1}\big[\,\widehat{V}\big(\mathbf{q}, q_z(\mathbf{q})\big)\big], \qquad q_z(\mathbf{q}) = -\tfrac{\lambda}{2}\|\mathbf{q}\|^2, \tag{5}$$

where a hat denotes a Fourier transform, $\widehat{V}$ is the 3D transform of the rotated specimen and $\mathcal{F}^{-1}$ the inverse 2D transform. TPGA thus samples $\widehat{V}$ on the Ewald paraboloid instead of the flat central slice. For $\lambda \to 0$ the paraboloid flattens and, by the Fourier-slice theorem, U_θ becomes the projection P_θ of equation 1, so the ray model is a limit of TPGA. Vulović et al. (2014) show that TPGA reproduces multislice exit waves for specimens much thicker than t_{crit}, where the projection assumption fails, and that it breaks down only under strong multiple scattering.

Since equation 5 is linear in V, each Gaussian contributes separately. The 3D Fourier transform of a Gaussian is again a Gaussian, $\widehat{V}_k(\boldsymbol{\xi}) = m_k \exp(-2\pi^2 \boldsymbol{\xi}^\top \Sigma_k \boldsymbol{\xi}) \exp(-2\pi i\, \boldsymbol{\xi}^\top \boldsymbol{\mu}_k)$ for $\boldsymbol{\xi} \in \mathbb{R}^3$. Evaluating it on the paraboloid, $\boldsymbol{\xi} = (\mathbf{q}, q_z)$, and splitting both exponents with the partition above gives the contribution of primitive k to $\widehat{U}_\theta$,

$$\widehat{G}_k(\mathbf{q}) = m_k \exp[-2\pi^2(\mathbf{q}^\top A_k \mathbf{q} + 2q_z \mathbf{b}_k^\top \mathbf{q} + c_k q_z^2)] \exp(-2\pi i\, \mathbf{q}^\top \boldsymbol{\mu}_{k,\perp}) \exp(-2\pi i\, q_z \mu_{k,z}). \quad (6)$$

The mass $m_k = \widehat{G}_k(\mathbf{0})$, the integral of the contribution, sets the strength of the primitive. The first exponential is a real envelope that sets its shape. The first term of its exponent is the spectrum of the lateral footprint, as in the ray renderer. The other two exist only because the paraboloid leaves the flat slice ($q_z \neq 0$). Parts of the primitive at different depths carry different defocus, which damps fine detail in proportion to the axial variance c_k, and a tilt relative to the beam ($\mathbf{b}_k$) makes the defocus vary across the footprint. The second and third exponentials have unit modulus and change only the phase. By the Fourier shift theorem, the second places the primitive at its lateral position $\boldsymbol{\mu}_{k,\perp}$ and the third shifts it along the beam to its depth $\mu_{k,z}$. On the flat slice $q_z = 0$, the third exponential is one and depth is invisible, as in the ray model. On the paraboloid, $q_z = -\lambda\|\mathbf{q}\|^2/2$ turns it into

$$\exp(-2\pi i\, q_z \mu_{k,z}) = \exp(i\pi\lambda\mu_{k,z}\|\mathbf{q}\|^2) = P_{-\mu_{k,z}}(\mathbf{q}),$$

the Fresnel propagator of Section 2, which carries the wave of the primitive from its depth back to $z = 0$, a defocus by its depth. The ray renderer discards this *depth-dependent defocus phase*.

The rasterizer can only draw 2D Gaussians, whose Fourier transforms are exponentials of quadratic forms in $\mathbf{q}$. Two terms of equation 6 are not quadratic, because q_z itself grows with $\|\mathbf{q}\|^2$. These are the cross term $2q_z\mathbf{b}_k^\top\mathbf{q}$, set by the tilt of the primitive, and the axial term $c_k q_z^2$, set by its extent along the beam. We drop both, so depth enters only through the center $\mu_{k,z}$ and each Gaussian is imaged as a flat object at that depth. This is accurate when the primitive is thin compared with the depth of field, just as the projection model is for a thin specimen. Our primitives are far thinner, 0.4 Å on the Co_3O_4 series against $t_{\mathrm{crit}} \approx 21$ nm at 2 Å.

Propagating all contributions to the reference plane $z = T$, where the CTF is applied, multiplies them by P_T and turns the depth phase into $P_{T-\mu_{k,z}}$, the propagation from the depth of the primitive to T. The two retained factors then combine into one Gaussian with a complex covariance,

$$\begin{aligned} \exp\big(-2\pi^2\, \mathbf{q}^\top A_k \mathbf{q}\big)\ \exp\big(-i\pi\lambda(T-\mu_{k,z})\|\mathbf{q}\|^2\big) &= \exp\big(-2\pi^2\, \mathbf{q}^\top A_k^{\mathbb{C}}\, \mathbf{q}\big), \\ A_k^{\mathbb{C}} &= A_k + i\,\tfrac{\lambda(T-\mu_{k,z})}{2\pi}\, I_2, \end{aligned} \quad (7)$$

with I_2 the 2×2 identity. The real part of $A_k^{\mathbb{C}}$ is the footprint and the imaginary part is the defocus. In real space this is a broadened blob with a phase that varies quadratically across it, which is how a small Gaussian object looks after propagating over the distance $T - \mu_{k,z}$. The rasterizer draws one such Gaussian per primitive and sums them into $U_\theta = \sum_k G_k$ in one pass.

From exit wave to reconstruction. Writing $U_\theta = U'_\theta + iU''_\theta$, the exit wave $\exp(i\sigma U_\theta) = \exp(i\sigma U'_\theta)\exp(-\sigma U''_\theta)$ has its phase set by U'_θ and its amplitude set by U''_θ. This amplitude modulation is not absorption. It is produced by propagation alone, is fixed by the masses and depths of the primitives and vanishes for $\lambda \to 0$. The image follows from equation 2, with the measured modulation transfer function of the detector applied to the intensity. Under the WPOA, $\exp(i\sigma U_\theta) \approx 1 + i\sigma U_\theta$ and the model reduces to the linear CTF model of conventional cryo-ET.

We place the reference plane T through the rotation center of the tilt series. The rotation center stays fixed while the stage tilts, and the defocus estimated from the Thon rings of a micrograph describes approximately the middle of the specimen. The per-tilt defocus thus keeps its meaning at every tilt, and each primitive adds the defocus of its depth relative to T through $A_k^{\mathbb{C}}$. The ray renderer is the same model with the depth phase set to zero.

As in PhaseT3M, we minimize an amplitude loss between the predicted images I_θ^{pred} of equation 2 and the measured images I_θ^{obs},

$$\mathcal{L} = \sum_\theta \big\| W_\theta \big[(I_\theta^{\mathrm{pred}})^{1/2} - (I_\theta^{\mathrm{obs}})^{1/2} \big] \big\|_2^2, \quad (8)$$

summed over all tilts. The square root makes the counting noise approximately equally strong in bright and dark pixels. The fixed linear filter W_θ weights the spatial frequencies of the residual.

Where raw movie frames exist (EMPIAR-10164), each frequency is weighted by its spectral signal-to-noise ratio (SSNR), measured once from the odd and even frames, so that noisy frequencies count less. Without frames, as for the Co_3O_4 series, W_θ is the identity. Adam optimizes the Gaussian parameters together with per-tilt rotation and translation corrections δ_{rot} and δ_{trans} to the nominal pose of each tilt. On the cryo-ET series it also refines the defocus of each tilt and an analytic slab that models the vitreous ice, so that the Gaussians represent only the specimen. Positive masses, a support sized by the ice thickness and a lower bound on the widths act as regularizers. Where the fit needs more detail, primitives are cloned or split with their masses divided, which conserves the total mass. On the cryo-ET series we fit from coarse to fine, from bin 8 down to native sampling, each stage initializing the next. Coarse stages are cheap and have a high signal-to-noise ratio per pixel, so they fix the alignment and large structures first.

5 EXPERIMENTS

As established in section 2, the electrostatic potential of the specimen scatters electrons as they propagate through the slab. Scattering near the entry of the slab result carry over longer distances than scatterings from interactions near the slab exit. Ray-based rendering cannot account for this, as it ignores the depth along the projection. As we move to higher resolution detections ($\rightarrow$ bin1), this effect becomes more pronounced and given a fixed slab thickness, there is a threshold resolution below which the effect is undetectable (Equation 3). In this regime, ray-based approximations don't exhibit a measurable error, but once we cross this threshold, accounting for the wave interactions becomes crucial. Therefore, we distinguish two types of experiments, depending on which side of this threshold the available data is on. The thin-specimen regime covers experiments on datasets with t_{crit} *below* the threshold, i.e. where ray-based methods are still reasonable approximations. The thick-specimen regime covers experiments on data whose t_{crit} *exceeds* this threshold, i.e. where wave-based optics are required to achieve accurate reconstructions.

5.1 EXPERIMENTAL SETTINGS

Datasets. We evaluate on three public datasets (Table 5). All are measured tilt series, and together they cover the two regimes of equation 3. EMPIAR-10164 (immature HIV-1 VLPs) and EMPIAR-10643 lie in the thin regime (the specimen is thinner than t_{crit} at the resolution the data support), where propagation should not matter. EMPIAR-10164 is used by both NovaCTF-related work and PhaseT3M and ships raw frames that permit frame-half reconstructions at identical tilt geometry, and EMPIAR-10643, the dataset used by CryoETGS, provides an even/odd-tilt split. For the thick regime (thicker than t_{crit} at the lattice periods), the Co_3O_4 nanoparticle focal series deposited with PhaseT3M (Lee, 2026) samples past t_{crit} and requires wave optics.

All methods reconstruct from the same tilt series with the same poses, physical crop and declared normalization. In the thin regime the baseline is WBP from IMOD (Kremer et al., 1996) on the same half-stacks. We also run ICE-TIDE (Debarnot et al., 2025), the implicit-field method closest to our setting, with its released code on the same series and half splits. It uses its own alignment procedure, so it is compared with a WBP of matching input stacks. In the thick regime the reconstruction deposited with PhaseT3M (Lee, 2026) serves as the reference.

Evaluation metrics. The experimental cryo-ET tilt series EMPIAR-10164 and EMPIAR-10643 come with no ground-truth tomogram, so resolution is read from experimental Fourier shell correlation (FSC) between reconstructions of independent halves of the data, at two thresholds. Half-map FSC between two fitted fields, or between two WBP volumes, is read at 0.5 (van Heel & Schatz, 2005). At bin1 we instead report the cross-FSC of one half's fitted field against the prior-free WBP of the other half, read at 0.143 following Rosenthal & Henderson, with the WBP–WBP curve as the floor (§5.2). Frame halves are preferred to even/odd-tilt splits, which change noise and angular sampling at once (Cardone et al., 2005), and are used wherever frames exist. On the Co_3O_4 series the deposited reconstruction serves as the reference, and we adopt the deposited analysis: the $\{400\}$ shell score, the Fourier correlation to the deposit at 2 Å and the image R-factor, plus the amplitude error and Fourier ring correlation (FRC) on a 0° tilt held out from training (§5.3). Appendix G explains why projection PSNR and SSIM are not reported.

Table 1: **Thin specimen regime.** (a) CryoWaveGS and ICE-TIDE (Debarnot et al., 2025), each against WBP. We report half-map FSC at 0.5 through bin 2 and cross-FSC at 0.143 at bin 1 (10643 in both directions). 10164 is bead-masked, and its bin 8/4 entries are omitted because the mask is coarser than a bead there. ICE-TIDE uses the same halves through its own pipeline, with its own alignment and box and no mask, against a WBP of its own input. Its numbers are not directly comparable to the CryoWaveGS group. † marks pairs flagged by its data-consistency control (Appendix D). (b) Wave (TPGA) and ray with the same seed and schedule. The unmasked 10164 bin 2 pair compares the two renderers, not the two pipeline groups in (a).

(a) Reconstruction comparisons

		CryoWaveGS pipeline		ICE-TIDE pipeline	
Dataset (halves)	Bin	WBP	CryoWaveGS	WBP	ICE-TIDE
10643 (tilts)	8	167	**134**	**191**	208
	4	170	**135**	**193**	231
	2	173	**102**	**210**	249
	1	81	**51 / 59**	–	–
10164 (frames)	8	–	–	80	**78**†
	4	–	–	**80**	82
	2	97	**74**	81	**38**†
	1	31.4	**31.7**	–	–

(b) Wave (TPGA) vs. Ray

Dataset	Metric	Wave	Ray
	Unmasked		
10164	bin 2, FSC 0.5	36.12 Å	36.15 Å
	bin 1 cross-FSC at frequency		
10164	$1/100$ Å^{-1}	0.782	0.782
	$1/60$ Å^{-1}	0.436	0.436
	$1/40$ Å^{-1}	0.292	0.292
	bin 1 cross-FSC at frequency		
10643	0.143	60.75 Å	60.76 Å

Implementation details. CryoWaveGS extends the CUDA rasterizer of Kerbl et al. (2023) to accumulate complex amplitudes with the depth phase of equation 7. Hardware, primitive counts, run times and memory for every run are in Appendix F.

5.2 Thin-specimen regime: cryo-ET

No ground truth tomogram exists for EMPIAR-10164 or EMPIAR-10643, so every reported number is based on comparing reconstructions from the even and odd subsets of the tilt series. Each half is initialized, calibrated and fitted on its own. The WBP baseline (IMOD) uses the same half-stacks, and ICE-TIDE is fitted to the same halves using its own alignment and box and compared with a WBP of its own input (§5.1). All other comparisons are made inside one ≈ 250 nm ROI per dataset over the full ice thickness (see Fig. 2a), fitted using a binning curriculum ($8 \to 4 \to 2 \to 1$) with bin1 on a ≈ 130 nm box inside it. The fiducials in the 10164 ROI are masked identically in both methods. We report the half-map FSC and the cross-FSC against prior-free WBP, where any excess over WBP–WBP is signal recovered from the data.

Table 1 shows the FSC results of this experiments and Figure 2 visualizes the qualitative results and FSC plots (see Appendix A for additional results). CryoWaveGS achieves better reconstructions than WBP at every binning on both datasets. At full resolution (bin 1, the camera's own pixel size), each half's CryoWaveGS field is compared with the WBP of the other half. WBP uses no prior, so any agreement between the field and the WBP comes from the data alone, not from assumptions shared by two fitted halves. All methods reach similar values at bin 1, limited by the noise in the data rather than by the reconstruction, as neither series holds signal beyond ≈ 30 Å. The bin 1 curves in Fig. 2d lie below the bin 2 curves for both CryoWaveGS and WBP because the bin 1 box is smaller and contains less of the specimen. ICE-TIDE and IsoNet2 ran through their own pipelines (middle block), so each is compared only with the WBP of its own input. ICE-TIDE does not improve on that WBP at the 0.5 threshold while IsoNet2 improves on it more than CryoWaveGS does on 10643 at bin 2. On 10164 the gold fiducials dominate any unmasked FSC, so the top block masks the fiducials in every method. The ICE-TIDE pipeline does not mask the fiducials, and the ICE-TIDE values on 10164 are not comparable with the top block.

This regime is the control for §5.3. A scatterer at depth z differs from its projection by the Fresnel phase $\pi\lambda z q^2$, which at 300 kV and $z = t/2 = 81$ nm is 0.06 rad at $1/30$ Å and reaches 1 rad only near $1/7$ Å; by equation 3 the specimens are thin by more than an order of magnitude at the resolution they support. The wave model must therefore reduce to the ray model here, and it does: with the same seed, schedule and Gaussian count, the two arms differ only in λ and agree to the third decimal in every FSC band on both datasets at the same cost to within 1% (Table 1, last row; Appendix F). The gains above are a property of the representation and its joint refinement of field, shifts, defocus and ice, not of wave optics; the wave term becomes decisive only where §5.3 places it.

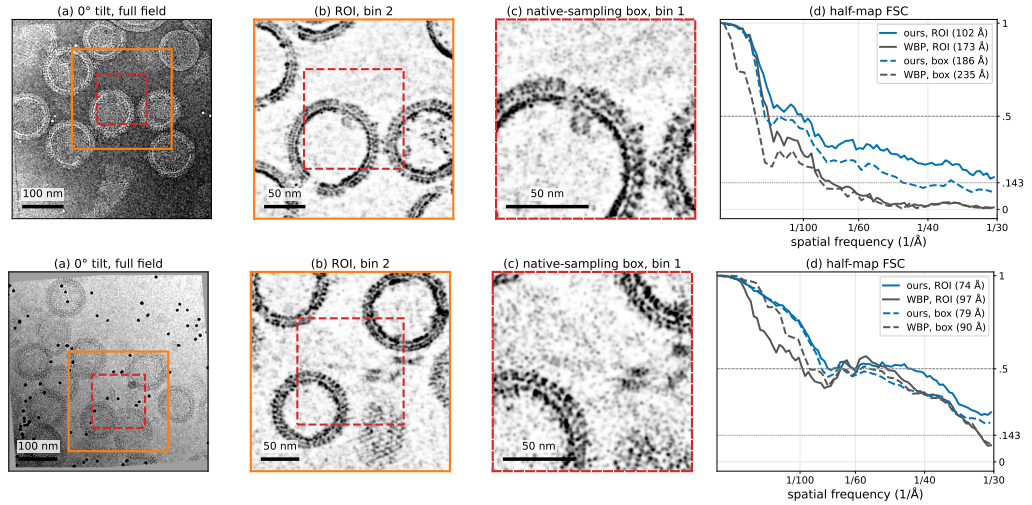

Figure 2: **CryoWaveGS recovers finer consistent detail than WBP from the same data, with continuous membranes down to native sampling**. (a) The 0° tilt of the full field with the ROI (orange) used at every binning and the native-sampling box (red, dashed). (b) The ROI reconstructed by CryoWaveGS at bin 2, mean of the two half fields, 80 Å slab. (c) The native-sampling box at 1.18 / 1.35 Å/px. (d) FSC between the two halves inside each box, CryoWaveGS (solid) and WBP of the same half-stacks (dashed), 0.5 crossings in the legend; on 10164 the fiducials inside the boxes are masked in both methods.

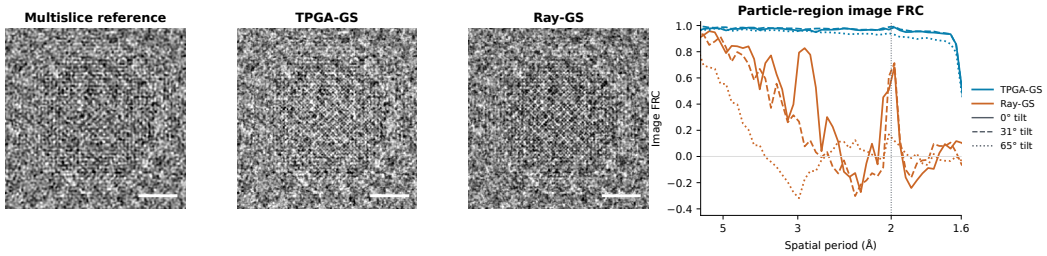

Figure 3: **Renderer validation on a frozen Gaussian field. Left:** abTEM multislice reference, TPGA-GS and Ray-GS at 0° tilt and the middle defocus, scale bar 2 nm. **Right:** particle-region image FRC against the reference for all three tilts, pooled over the three defoci.

5.3 Thick-specimen regime: atomic-resolution nanoparticle

The Co_3O_4 series images a $\sim$7 nm cobalt oxide nanoparticle, a crystal with the spinel lattice, partially embedded in an amorphous carbon support of about 30 nm (Lee, 2026). The particle carries the lattice that sets the target periods, the carbon carries no lattice but sets the propagation depth, and the deposited reconstruction spans 312 Å along the beam (600 slices at 0.52 Å), beyond the depth of field of equation 3 for every period below 2.5 Å, so the ray model is predicted to fail at the periods that carry the lattice.

We first test the renderers with no fit at all. One fixed field of 200,000 Gaussians, fitted to a simulated copy of the particle on extended carbon, is imaged with the deposited optics by each renderer and by multislice, with no optimization (Figure 3). The reference is abTEM (Madsen & Susi, 2021), the standard open-source multislice code for atomic-resolution TEM, which shares no code with our renderers; an independent multislice written for this work reproduces its images of the same potential to a relative error below 2×10^{-5}, so the reference is not an artefact of one implementation.

Both renderers match multislice down to about 5 Å. Below that, the ray model loses correlation while the complex-covariance renderer holds at about 0.9 down to 1.6 Å, and the period at which the ray model fails moves with tilt exactly as the depth-of-field criterion equation 3 predicts (Figure 3).

TPGA-GS and Ray-GS reconstruct the measured series under the PhaseT3M settings: published optics held fixed, positivity, and the PhaseT3M amplitude loss. The particle starts as 1.6 M fixed-width Gaussians on a 0.5 Å grid inside an image-derived mask, with uniform mass; the carbon outside the mask is a jointly fitted voxel field. For 100 passes over the 131 tilts only the masses train, so the particle is a voxel-like field in a Gaussian basis and the two renderers differ only in

Table 2: Co_3O_4 reconstructions scored against the PhaseT3M deposit (§5.3). Particle and particle + carbon: Fourier correlation at 2 Å, over all directions and inside the missing wedge. R: image R-factor over all 393 images, lower is better. Grid $\sigma = 0.4$ Å; center steps capped at 0.02 Å per update and 2 Å in total.

				Particle		Particle + carbon			
Method	**Pass**	**Gaussians**	$\{400\}$	**all**	**wedge**	**all**	**wedge**	R	**T(min)**
PhaseT3M (deposited)	–	–	1.78	1	1	1	1	–	–
TPGA-GS	100	1,598,858	1.04	**0.93**	**0.82**	**0.85**	**0.80**	–	125
TPGA-GS	150	1,420,862	**1.09**	0.88	0.63	0.82	0.76	**0.161**	188
Ray-GS	100	1,598,858	0.88	0.74	0.67	0.42	0.31	–	97
Ray-GS	150	1,468,761	0.81	0.65	0.40	0.40	0.29	0.168	146

propagation. Gaussians below 5% of the median mass are then removed and the centers refine for 50 passes, which a voxel field cannot do.

We score against the PhaseT3M reconstruction with the analysis published alongside it. The finest reflection that analysis reports is $\{400\}$ of the Co_3O_4 spinel lattice at 2.02 Å, scored as the fraction of the 2.02 Å shell's power on the lattice positions divided by the same fraction for a phase-randomized copy with identical Fourier magnitudes: 1 means no concentration beyond chance, below 1 means the power sits off the lattice positions, and PhaseT3M scores 1.78. To this we add the Fourier correlation to PhaseT3M at 2 Å, over all directions and inside the missing wedge, for the particle alone and for the full field with carbon, and the image R-factor with the same formula.

The wave (TPGA-GS) and ray (Ray-GS) reconstructions recover the lattice alike (Figure 4, Appendix B) but separate at the fine periods where the forward check predicts the ray model fails (Table 2). At pass 100, the end of the fixed-position stage, the wave reconstruction correlates better with PhaseT3M at 2 Å: 0.85 against 0.42 over the full field, with the same ordering inside the missing wedge and on the particle alone. The wave power spectrum follows PhaseT3M to 2 Å, where the ray spectrum carries about half the power. The $\{400\}$ score separates the two reconstructions least, as expected for a lattice spot, and neither reaches the PhaseT3M value of 1.78. Center refinement lowers the image R-factor of both but reduces agreement with the deposit, so we quote pass 100. Each entry is one run.

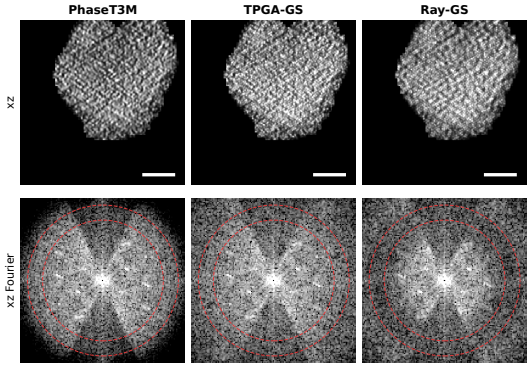

Figure 4: **Reconstruction of Co_3O_4 tilt series.** *Top:* 2.08 Å *xz* slab through the particle, the plane along the beam in which missing-wedge elongation and depth errors appear. *Bottom:* 2D Fourier power of the *xz* slab, dashed rings at 2 and 1.6 Å. Scale bars 2 nm.

6 Discussion and Conclusion

CryoWaveGS moves the Fresnel phase into a complex primitive covariance, so wave optics renders in one rasterization pass rather than a propagated voxel stack. The depth of field $t_{\mathrm{crit}} = d^2/\lambda$ equation 3 states when that matters, and the two regimes behave as it predicts. On cryo-ET tilt series at current binnings the specimens are thin by more than an order of magnitude at the resolution the data support, and the wave and ray arms agree to the third decimal in every FSC band (Table 1); the gain over WBP there belongs to the representation and its joint refinement, not to wave optics. On the atomic-resolution Co_3O_4 series, which lies past the switch point, the same two arms separate where the forward check against multislice predicts: at 2 Å the wave arm reaches a Fourier correlation of 0.85 to the PhaseT3M reference against 0.42 for the ray arm, within 1.3× its cost (Table 2).

Limitations. TPGA and its complex-covariance renderer are approximations, so multislice remains the reference. The method does not fill the missing wedge. All stated resolutions come from frequency metrics with declared scope. On the thick-specimen series the reference is itself a reconstruction, so agreement with PhaseT3M is not agreement with the specimen, and neither renderer predicts a held-out 0° view beyond about 7 Å. Neither renderer reaches the 2 Å lattice contrast of that reference whilst refining the Gaussian centers after the grid stage did not improve on the grid, so the compact atom-like representation that Gaussians promise is not yet shown on measured data.

AI USE STATEMENT

We used generative AI tools to draft and edit text, to search for related work beyond our manual literature review, to gain feedback for the drafts of the paper, and, as coding agents, to implement, debug and run parts of the software, experiments and analysis. All generated text, code, analyses and suggested references were checked by the authors, who designed the study, verified every reported result and take full responsibility for the content of the paper.

ETHICS STATEMENT

This work uses only publicly deposited datasets (EMPIAR-10164, EMPIAR-10643 and the Co_3O_4 series deposited with PhaseT3M) released under open licences, and simulated data generated by the authors. It involves no human subjects, no personally identifying information, and no foreseeable dual-use concern.

REPRODUCIBILITY STATEMENT

The forward model is specified in §4 with its exact and approximate components listed explicitly; the datasets, splits and metrics are specified in §5.1; the wave and ray implementations share initialization, nuisance parameters, budget and seeds as stated in §5.2 and §5.3.

REFERENCES

Anders H. Andersen and Avinash C. Kak. Simultaneous algebraic reconstruction technique (SART): A superior implementation of the ART algorithm. *Ultrasonic Imaging*, 6(1):81–94, 1984.

Giovanni Cardone, Kay Grünewald, and Alasdair C. Steven. A resolution criterion for electron tomography based on cross-validation. *Journal of Structural Biology*, 151(2):117–129, 2005.

et al. Choi. Gaussian wave splatting for computer-generated holography. *ACM Transactions on Graphics*, 44, 2025.

Valentin Debarnot et al. ICE-TIDE: Implicit cryo-ET imaging and deformation estimation. *IEEE Transactions on Computational Imaging*, 2025.

EMPIAR-10164. EMPIAR-10164: Cryo-electron tomography of immature HIV-1 virus-like particles. Electron Microscopy Public Image Archive.

Peter Gilbert. Iterative methods for the three-dimensional reconstruction of an object from projections. *Journal of Theoretical Biology*, 36(1):105–117, 1972.

Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3D Gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 42(4), 2023.

Earl J. Kirkland. *Advanced Computing in Electron Microscopy*. Springer, 2nd edition, 2010.

James R. Kremer, David N. Mastronarde, and J. Richard McIntosh. Computer visualization of three-dimensional image data using IMOD. *Journal of Structural Biology*, 116(1):71–76, 1996.

et al. Lee. PhaseT3M: 3D imaging at 1.6 å resolution via electron cryo-tomography with nonlinear phase retrieval. *Nature Communications*, 17:690, 2026.

Yun-Tao Liu, Heng Zhang, Hui Wang, Chang-Lu Tao, Guo-Qiang Bi, and Z. Hong Zhou. Isotropic reconstruction for electron tomography with deep learning. *Nature Communications*, 13:6482, 2022.

Jacob Madsen and Toma Susi. The abTEM code: transmission electron microscopy from first principles. *Open Research Europe*, 1:24, 2021. doi: 10.12688/openreseurope.13015.2.

D. Moško, P. Szeles, I. Draci, P. Vagovič, and J. Uličný. Differentiable Gaussian splatting with Fresnel propagation for X-ray phase contrast imaging. *Research Square preprint rs-10201845*, 2025.

Michael Radermacher. Weighted back-projection methods. In Joachim Frank (ed.), *Electron Tomography: Methods for Three-Dimensional Visualization of Structures in the Cell*, pp. 245–273. Springer, 2006.

Peter B. Rosenthal and Richard Henderson. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *Journal of Molecular Biology*, 333(4):721–745, 2003.

Beata Turoňová, Florian K. M. Schur, William Wan, and John A. G. Briggs. Efficient 3D-CTF correction for cryo-electron tomography using NovaCTF improves subtomogram averaging resolution to 3.4 å. *Journal of Structural Biology*, 199(3):187–195, 2017.

Marin van Heel and Michael Schatz. Fourier shell correlation threshold criteria. *Journal of Structural Biology*, 151(3):250–262, 2005.

Miloš Vulović, Lenard M. Voortman, Lucas J. van Vliet, and Bernd Rieger. When to use the projection assumption and the weak-phase object approximation in phase contrast cryo-EM. *Ultramicroscopy*, 136:61–66, 2014.

Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4): 600–612, 2004.

Simon Wiedemann and Reinhard Heckel. A deep learning method for simultaneous denoising and missing wedge reconstruction in cryogenic electron tomography. *Nature Communications*, 2024.

et al. Yang. Exact-GS: Mathematically rigorous and accurate 3D Gaussian splatting for 3D X-ray reconstruction. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2026.

Ruyi Zha, Tao Lin, Yuanhao Cai, Jiwen Cao, Yanhao Zhang, and Hongdong Li. R^2-Gaussian: Rectifying radiative Gaussian splatting for tomographic reconstruction. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2024. arXiv:2405.20693.

et al. Zhan. Complex-valued holographic radiance fields. *ACM Transactions on Graphics*, 45, 2026.

et al. Zhang. Adaptive Gaussian representation for differentiable cryo-electron tomography reconstruction. *Journal of Structural Biology*, 218:108281, 2026.

Shawn Zheng, Georg Wolff, Garrett Greenan, et al. AreTomo: An integrated software package for automated marker-free, motion-corrected cryo-electron tomographic alignment and reconstruction. *Journal of Structural Biology: X*, 6:100068, 2022.

APPENDIX

A ADDITIONAL RESULTS: THIN-SPECIMEN REGIME

These figures support §5.2. Slabs are 80 Å thick and show the mean of the two half fields without filtering, and FSC is computed between the two halves inside the ROI of Fig. 2. Figure 5 shows the wave and ray fields behind the bottom block of Table 1. Figures 6 and 7 follow one region through the curriculum, and Fig. 8 shows the fiducial mask behind the 10164 values. Figures 9 and 10 compare with CryoETGS and ICE-TIDE. Figures 11, 12 and 13 document the support, the primitive size floor and the density left in empty ice. Figure 14 adds a third specimen, EMPIAR-10453, in 272 nm of ice.

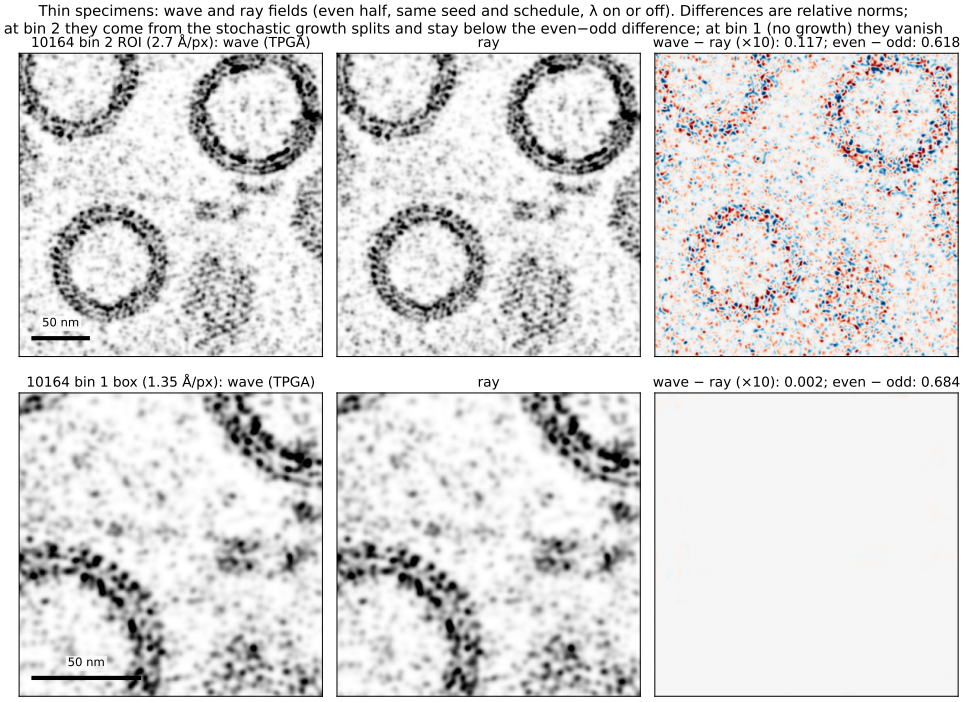

Figure 5: **Wave (TPGA) and ray fields on EMPIAR-10164, even half, same seed and schedule, λ on or off**. Right: the difference at ten times the display contrast, with the relative norms of wave−ray and, for scale, even−odd. At bin 2 the difference, 0.117, comes from the stochastic split offsets of the growth schedule and is five times smaller than the half-to-half difference, 0.618; at bin 1, without growth, the two fields agree to 0.2%.

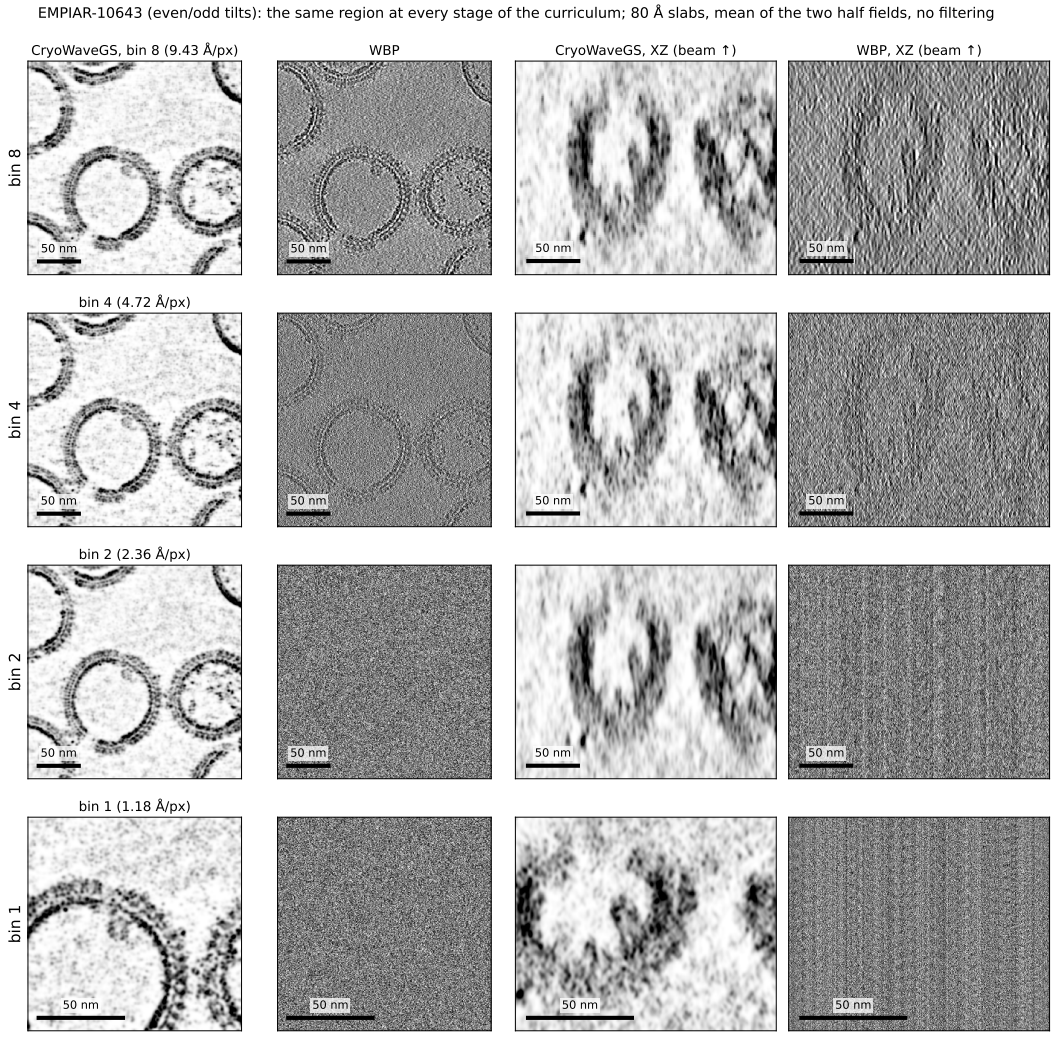

Figure 6: **EMPIAR-10643 (even/odd tilts): the same region at every stage of the curriculum**. Columns: CryoWaveGS and WBP, 80 Å XY slab, then 80 Å XZ slab with the beam vertical; mean of the two half fields, no filtering in either method. The bin 1 row shows the smaller native-sampling box. The CryoWaveGS VLP membranes stay continuous and keep their double-layer profile from bin 8 to bin 1, while WBP loses the specimen into noise at bin 2 and 1; the XZ panels show the missing-wedge elongation both methods share.

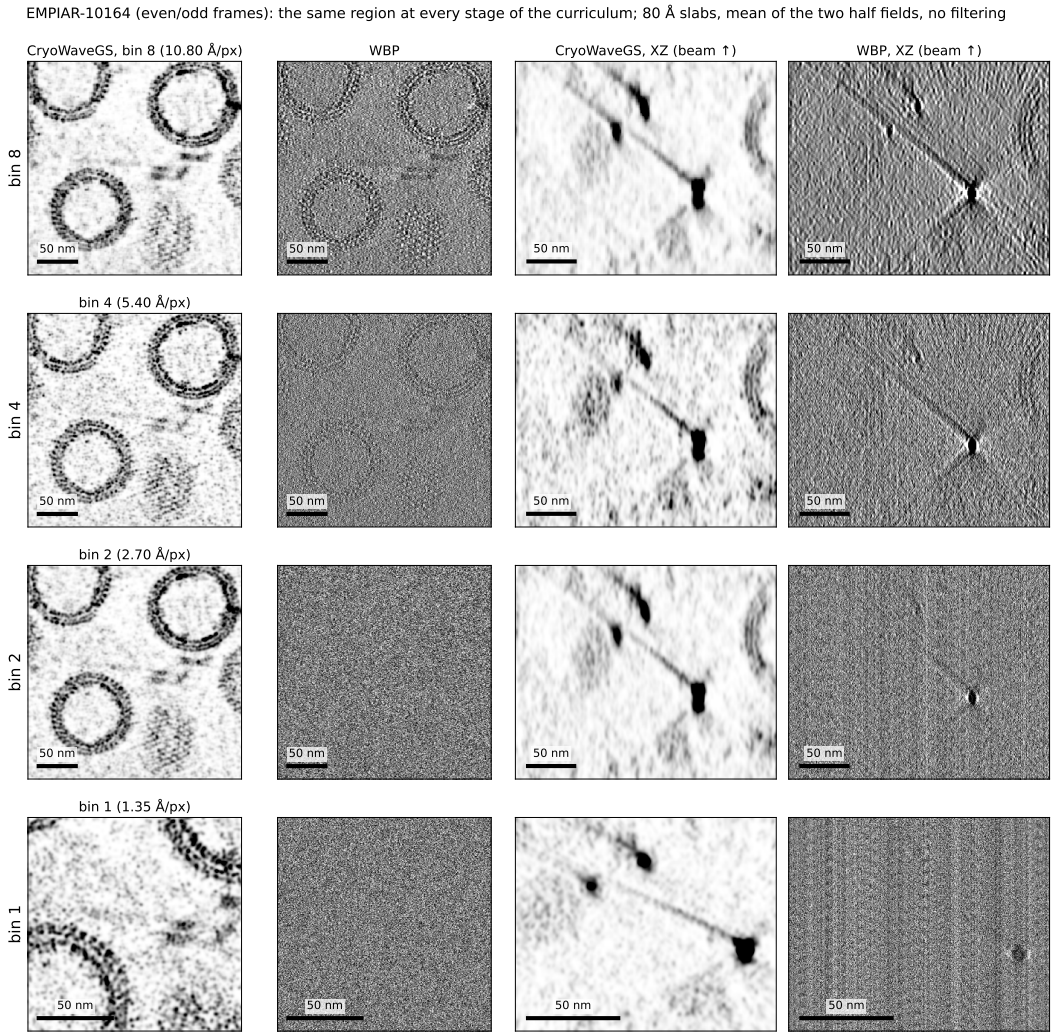

Figure 7: **EMPIAR-10164 (even/odd movie-frame halves), as Fig. 6**. The immature Gag lattice is resolved on the lower-right particle at bin 8 to 2. The XZ panels contain three gold fiducials and the carbon edge: WBP shows their tilt-streak artifacts, and CryoWaveGS models them as compact objects. Fiducials are masked before every FSC (Fig. 8).

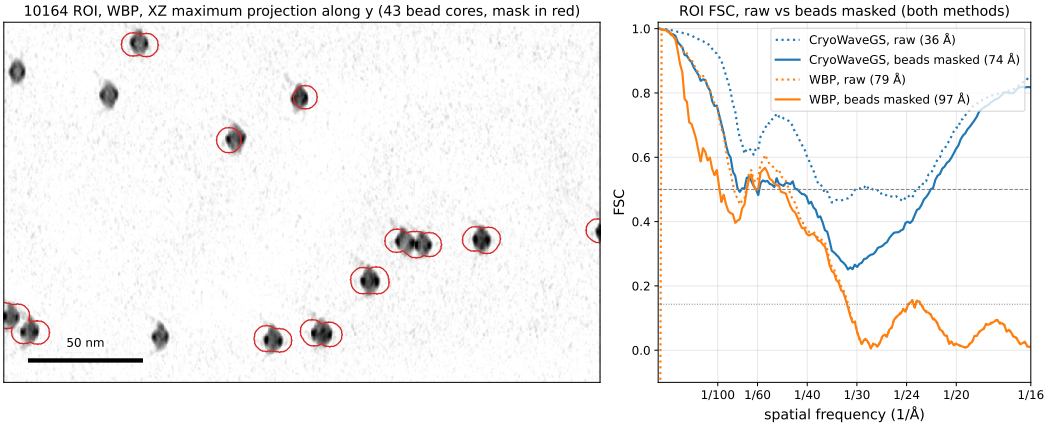

Figure 8: **Gold fiducials inside the EMPIAR-10164 ROI**. Left: WBP maximum projection along y with the bead mask (red), one mask for all methods. Right: ROI FSC before (dotted) and after (solid) masking. The beads carry most of the correlation of the raw CryoWaveGS curve, 36 Å raw against 74 Å masked at 0.5, and move WBP from 79 to 97 Å; the masked values are the ones reported in Table 1.

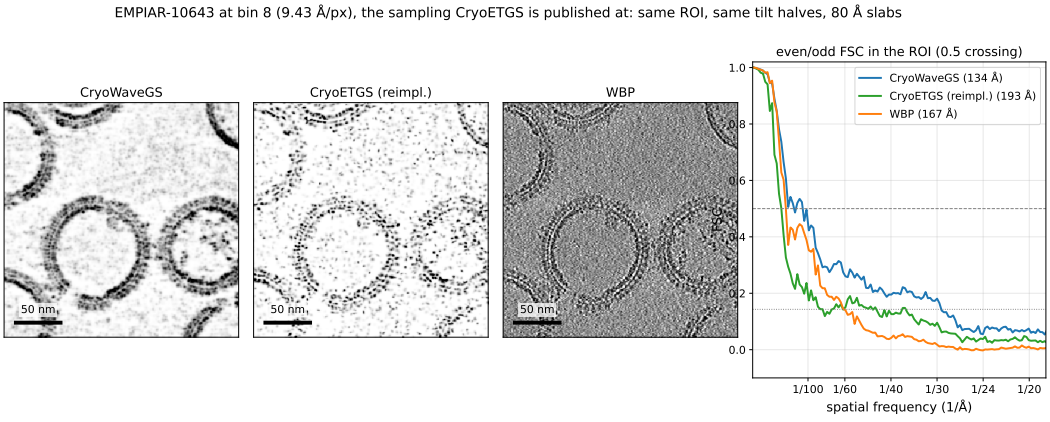

Figure 9: **EMPIAR-10643 at bin 8, the sampling at which CryoETGS is published**: CryoWaveGS, CryoETGS (our reimplementation with the authors' per-dataset parameters, 9.4 M Gaussians) and WBP in the same ROI from the same tilt halves, 80 Å slabs. Right: even/odd FSC in the ROI; at 0.5 CryoWaveGS crosses at 134 Å, WBP at 167 Å and CryoETGS at 193 Å.

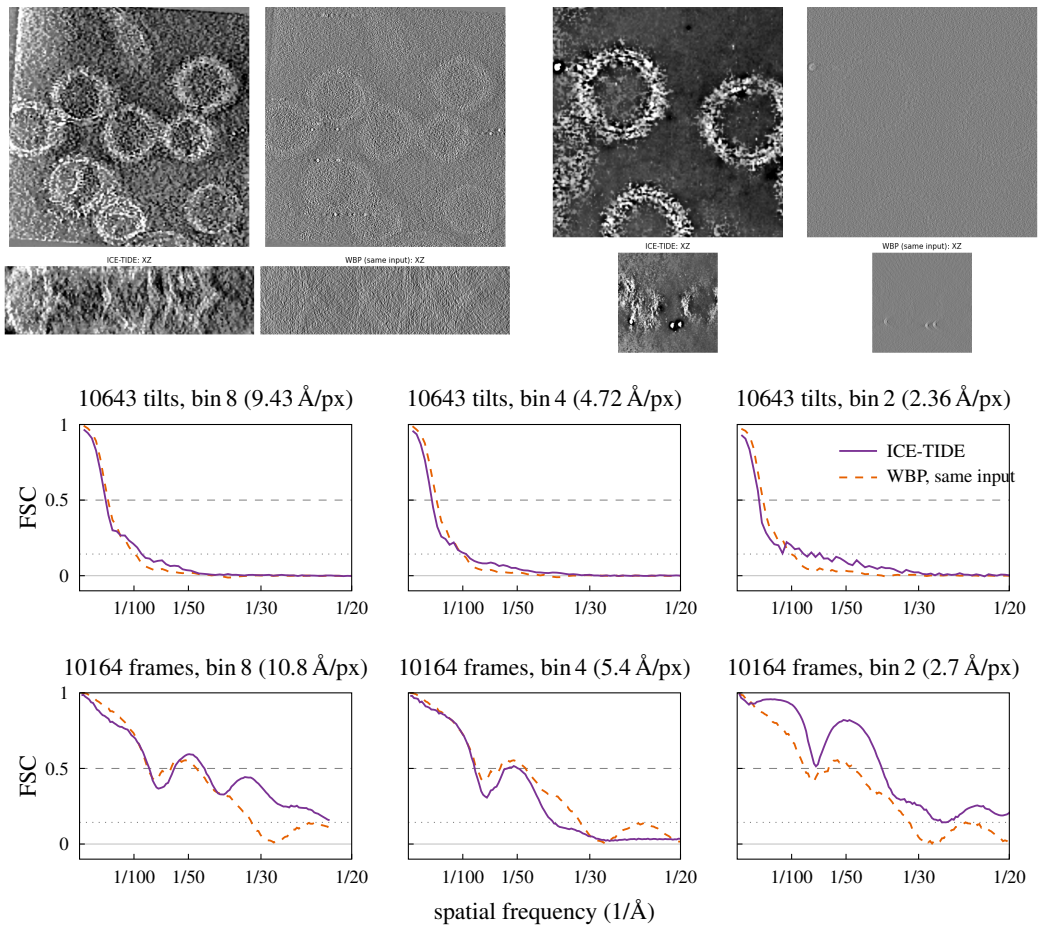

Figure 10: **ICE-TIDE and a WBP of the same input stacks (Appendix D)**. Top: central slices of the even half, EMPIAR-10643 at bin 8 (left) and EMPIAR-10164 at bin 2 (right); in each, ICE-TIDE on the left and WBP on the right, XY above XZ. Bottom: half-map FSC of the 10643 tilt halves and the 10164 frame halves at bin 8, 4 and 2, with 0.5 (dashed) and 0.143 (dotted); the crossings are in Table 4. The 10164 curves include the fiducials, which ICE-TIDE renders as bright compact blobs (top right).

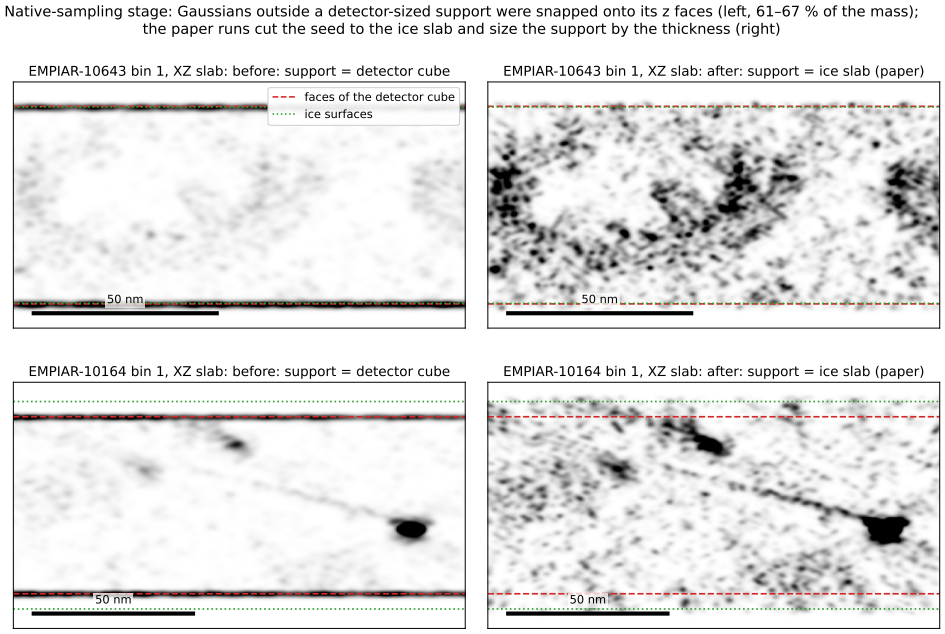

Figure 11: **The native-sampling stage before and after sizing the support by the ice thickness (XZ slabs through the bin 1 box, even half).** Left: with the support equal to the detector cube (red), Gaussians outside it were moved onto its two z faces, which then held 61–67% of the mass and dominated the half-map FSC; on 10164 the cube is thinner than the ice (green). Right: the reported runs cut the seed to the ice slab and size the support by the thickness.

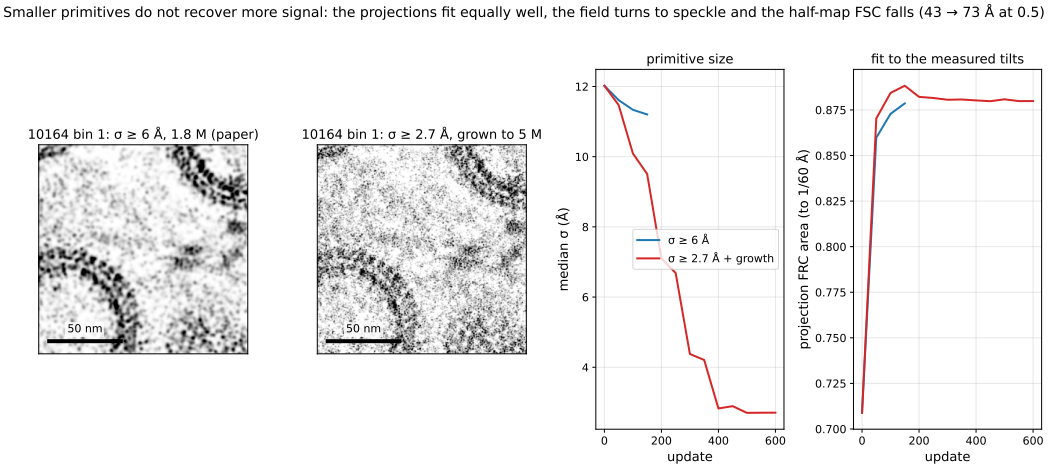

Figure 12: **Ablation of the primitive size floor on EMPIAR-10164 at bin 1**. Lowering the floor from 6 to 2.7 Å and growing the field to 5 M Gaussians drives the median σ to the floor while the fit to the measured tilts stays the same; the field turns to speckle and the half-map FSC 0.5 coarsens from 43 to 73 Å. The information limit is the data, not the primitive size, which is why the reported runs keep the 6 Å floor.

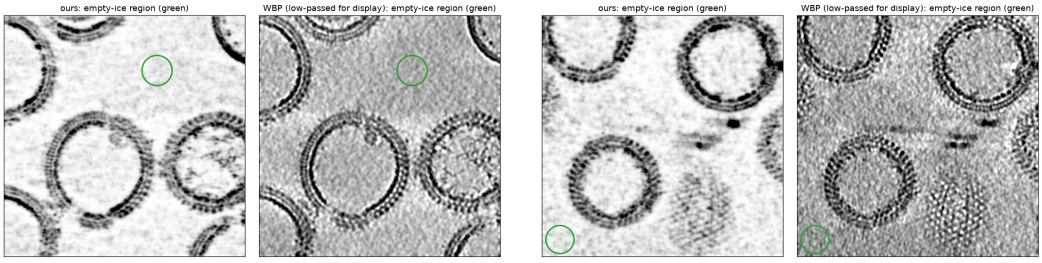

Figure 13: Empty-ice probe (green circle): a 15 nm-radius cylinder placed at the least-variance region of the WBP slab, shown on CryoWaveGS and on WBP low-passed for display, for EMPIAR-10643 (left pair) and EMPIAR-10164 (right pair). The mean CryoWaveGS density inside the probe is 9% (10643) and 3.5% (10164) of the membrane mean.

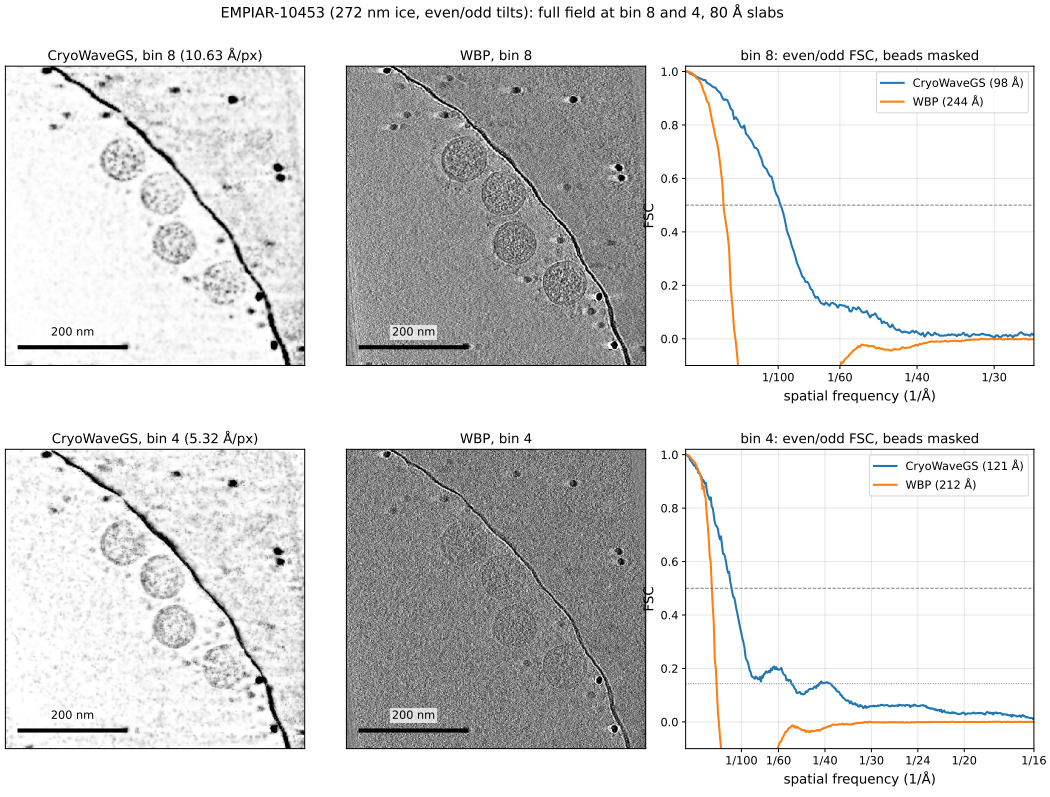

Figure 14: EMPIAR-10453, SARS-CoV-2 virions in 272 nm of ice (even/odd tilts): full field at bin 8 and 4, 80 Å slabs, and the bead-masked half-map FSC. At 0.5 CryoWaveGS crosses at 98 Å against 244 Å for WBP at bin 8, and at 121 against 212 Å at bin 4.

B ADDITIONAL RESULTS: THICK-SPECIMEN REGIME

These figures support §5.3. Figures 15 and 16 extend the forward check of Fig. 3 to all three tilts. Figures 17 to 20 show the reconstructions behind Fig. 4 and Table 2: the reconstruction figures show pass 150, the state after center refinement, as in Fig. 4, while the text quotes pass 100. Figure 21 shows the image fit behind the R-factor, Fig. 22 the held-out test, Fig. 23 both runs over the 150 passes, and Fig. 24 the refined Gaussian centers.

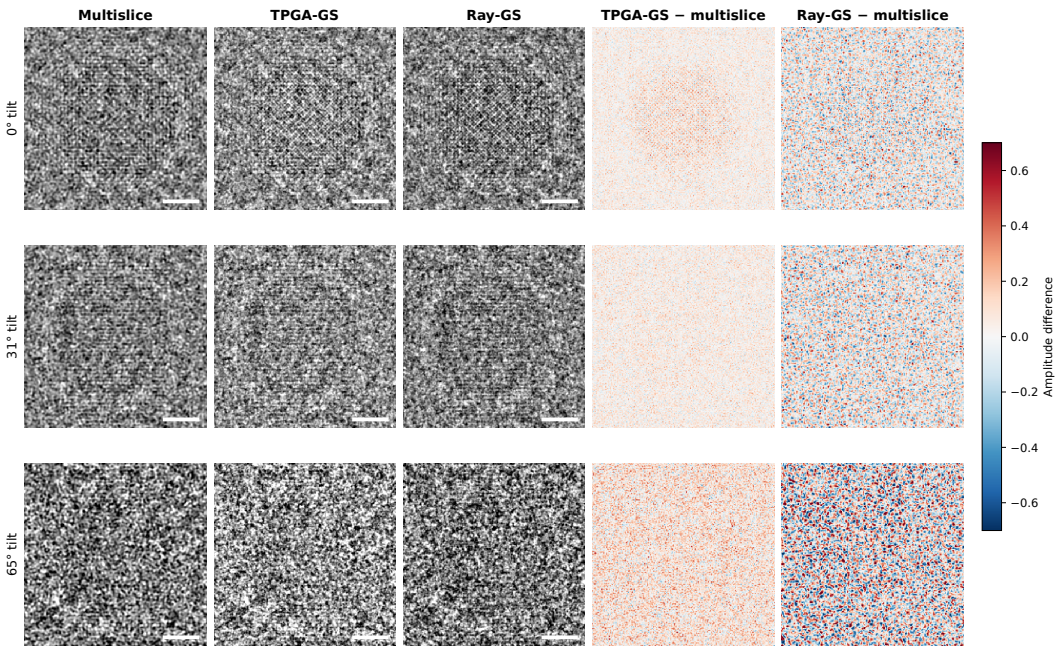

Figure 15: Forward check of Fig. 3 at all three tilts: multislice reference, TPGA-GS and Ray-GS images of the same frozen field, and each renderer's difference from multislice on a common scale. The TPGA-GS difference stays small at every tilt, while the Ray-GS difference grows with tilt as the depth along the beam increases. Forward predictions, not reconstructions.

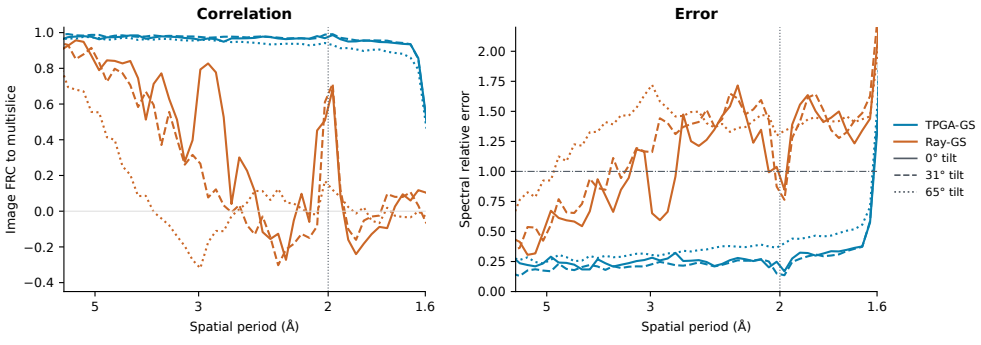

Figure 16: Spectral view of Fig. 15, pooled over the three defoci: image FRC to multislice (left) and spectral relative error (right) for each tilt. TPGA-GS keeps the correlation near or above 0.9 and the relative error near or below 0.5 to about 1.7 Å at every tilt. Ray-GS loses correlation between 5 and 3 Å, earliest at 65°, and at finer periods its relative error exceeds 1, worse than predicting zero; its peak near 2 Å coincides with the {400} reflection at 2.02 Å.

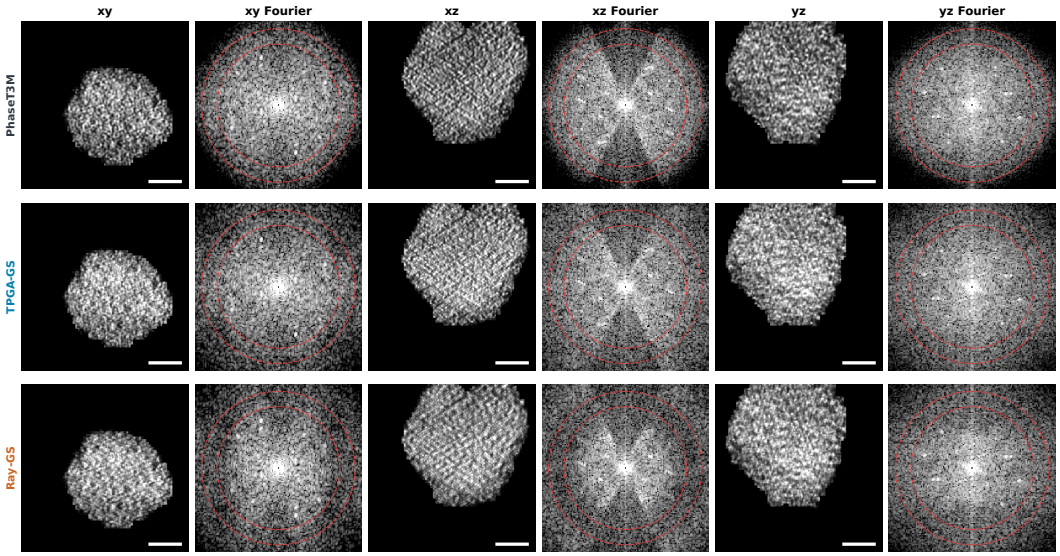

Figure 17: Orthogonal slabs through the particle (xy, xz, yz) and their 2D Fourier power for the PhaseT3M deposit, TPGA-GS and Ray-GS at pass 150, extending Fig. 4 to all three planes; slab thickness, rings and scale bars as in Fig. 4. The missing wedge appears as the empty double cone in the xz spectra of all three volumes.

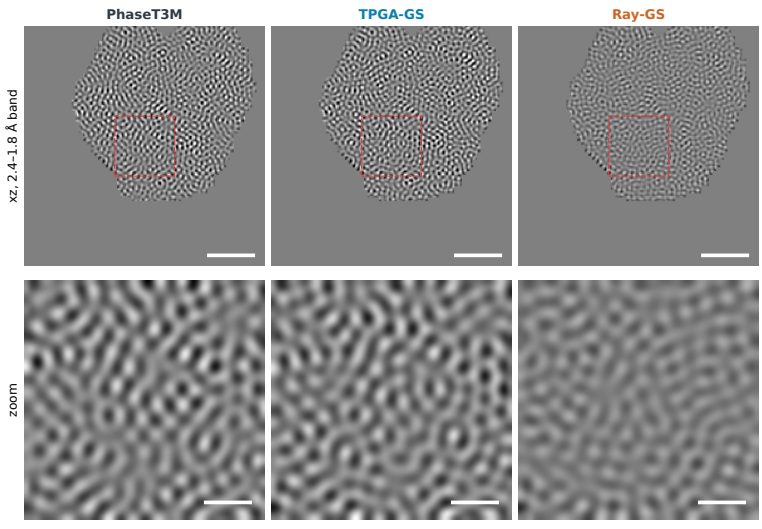

Figure 18: The xz slab of Fig. 4 (pass 150) band-passed to 2.4–1.8 Å, the band around the $\{400\}$ reflection at 2.02 Å, with a zoom on the red box (bottom). TPGA-GS reproduces the fringe pattern and contrast of the deposit; the Ray-GS fringes are weaker.

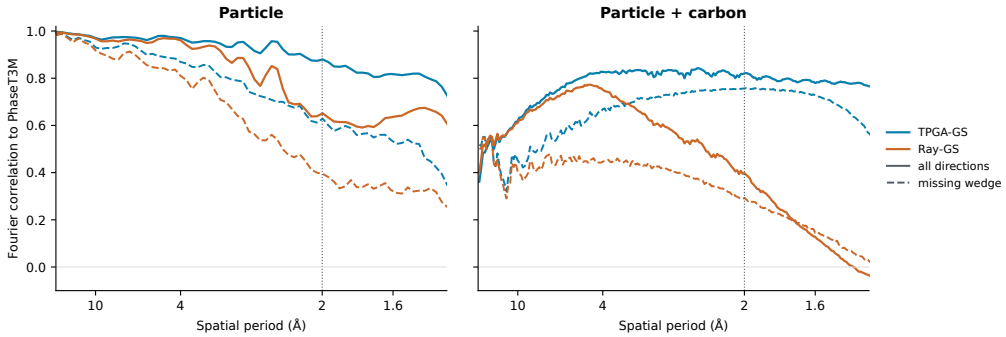

Figure 19: Fourier correlation to the PhaseT3M deposit against spatial period, over all directions (solid) and inside the missing wedge (dashed), for the particle (left) and particle + carbon (right) at pass 150; the values at 2 Å (dotted line) are the pass-150 rows of Table 2. Over particle + carbon, TPGA-GS stays near 0.8 from 4 to 1.6 Å while Ray-GS falls steadily below 4 Å.

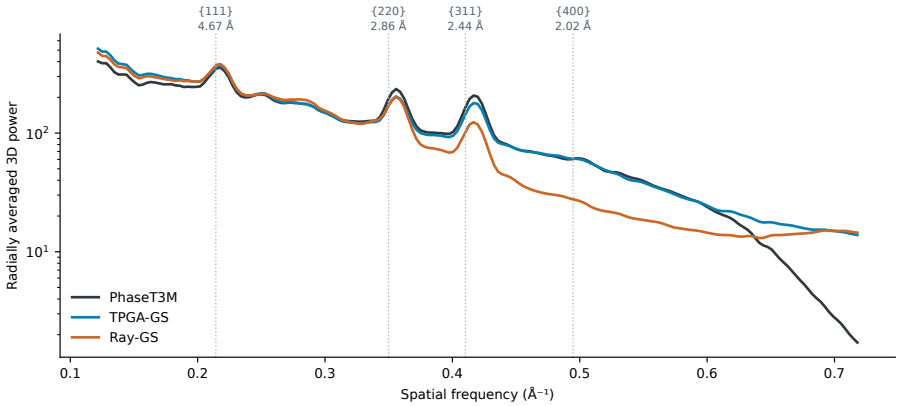

Figure 20: Radially averaged 3D power spectrum of the particle with the spinel reflections marked. TPGA-GS follows the deposit through the {111}, {220}, {311} and {400} reflections and on to about 1.6 Å, where the deposit's power falls off; Ray-GS falls below both from the {220} reflection on and carries about half the power beyond 2.4 Å.

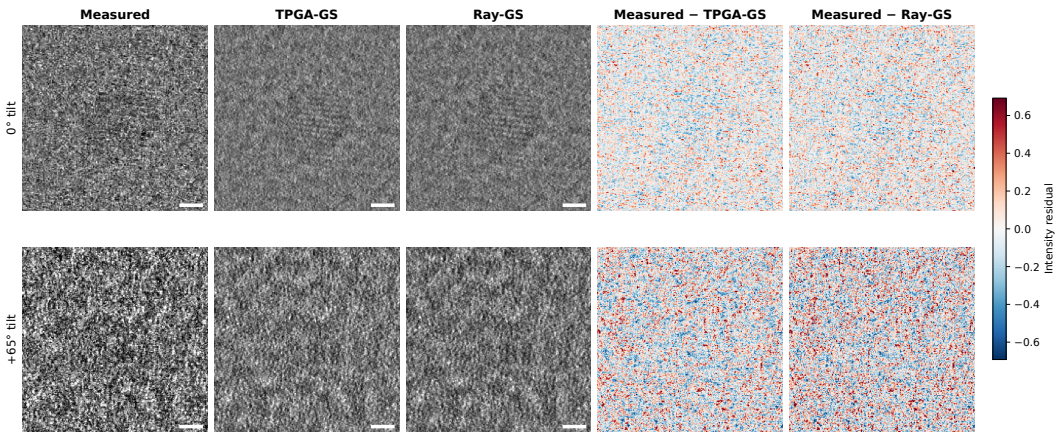

Figure 21: Measured images at 0° and +65°, the predictions of both renderers and the residuals on a common scale. Both residuals are dominated by noise, and at +65° both retain low-frequency structure. Over all 393 images the R-factor is 0.161 for TPGA-GS and 0.168 for Ray-GS at pass 150 (Table 2).

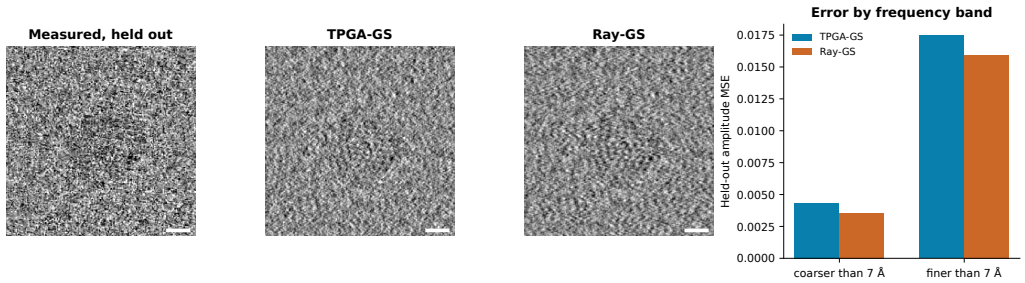

Figure 22: Leave-one-out test with the 0° defocus triplet held out from training. Left: a measured held-out image and the two predictions. Right: held-out amplitude error split at 7 Å. Ray-GS has the lower error in both bands, 0.0195 against 0.0218 in total, and neither prediction correlates with the held-out images beyond about 7 Å.

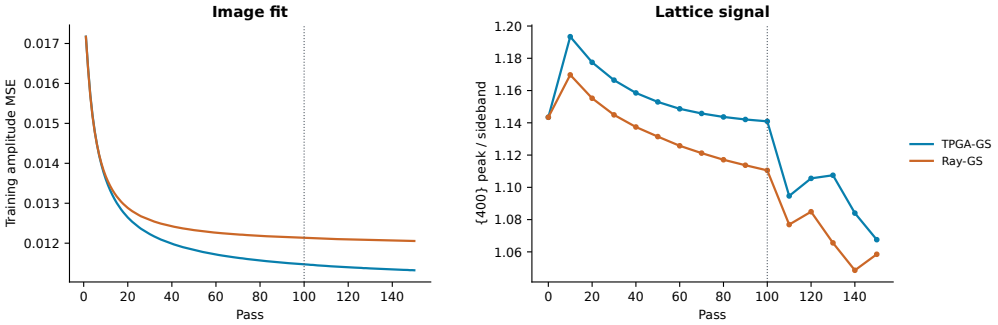

Figure 23: Both Co_3O_4 runs over the 150 passes. Left: training amplitude error; TPGA-GS fits the measured images better throughout. Right: a training-time lattice proxy, the $\{400\}$ peak-to-sideband ratio, which rises in the first passes, declines slowly during the fixed-position stage and drops at the prune after pass 100 (dotted line). The proxy is a different quantity from the shell score of Table 2.

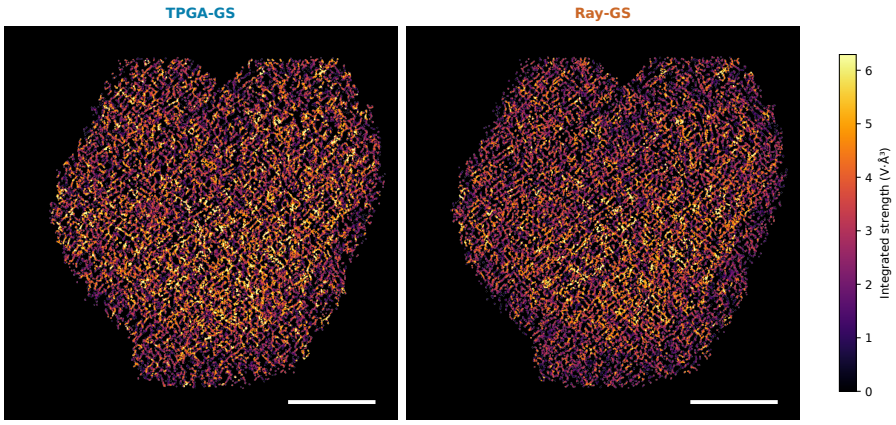

Figure 24: Gaussian centers of the particle after pruning and center refinement (pass 150), viewed in the xz plane and colored by integrated strength. In both fields the centers form a dense connected network rather than isolated atomic sites, so center refinement does not yet yield the compact atom-like representation (§6, Limitations).

C REPRODUCTION SETTINGS FOR THE THICK-SPECIMEN EXPERIMENTS

Table 3 lists every setting of the Co_3O_4 reconstructions of §5.3. TPGA-GS and Ray-GS share all of them and differ only in the propagation row. The leave-one-out pair repeats the same settings with the 0° triplet (original index 65) removed before preprocessing, leaving 130 training triplets. No densification or splitting is used at any stage.

Setting	Value
Data and optics	
Series	131 tilts over $\pm 65^\circ$, 3 defocus images per tilt, 600^2 pixels at 0.52 Å, 300 kV ($\lambda = 0.019687$ Å)
Optics	deposited per-image aberrations (angular order ≤ 2, radial order ≤ 4, 8 coefficients) and 2 image shifts, held fixed
Volume	600^3 voxels at 0.52 Å, 312 Å deep; exit plane at $z = 155.74$ Å
Interaction constant	$\sigma_e = 6.5262 \times 10^{-4}$ rad V^{-1} Å^{-1}
Loss and schedule	
Loss	PhaseT3M amplitude loss $\| \lvert\psi\rvert - \sqrt{I} \|^2$, no fitted gain
Update unit	one tilt (3 defoci) per update, order reshuffled each pass with seed 0; 131 updates per pass
Passes	100 (stage 1) + 50 (stage 2) = 150; 19,650 updates
Precision	FP32, TF32 disabled
Particle (Gaussians)	
Support	image-derived SIRT mask: 256-bin Otsu threshold, largest 26-connected component, holes filled, dilated 3 Å; 200,257 Å^3
Initialization	one Gaussian per 0.5 Å grid site in the mask, 1,598,858 in total; uniform mass 0.1915 V Å^3 (one 0.0005-rad object step)
Width	$\sigma = 0.4$ Å, isotropic, fixed throughout
Mass optimizer	SGD, no momentum, lr $= 0.0005/(\sigma_e/p^2)^2 = 73.37$ with pitch $p = 0.5$ Å; masses clamped to $[0, 10^6]$
Stage 1	centres frozen, masses trained, passes 1–100
Prune (after pass 100)	remove masses ≤ 0 and $< 0.05\times$ median positive mass; applied once
Stage 2	centres trained with Adam, lr 0.003 Å, $\beta = (0.9, 0.999)$; step cap 0.02 Å per update; total cap 2 Å from the post-prune position; centres kept within ± 155.74 Å
Carbon film (voxels, outside the mask)	
Initialization	zero
Optimizer	gradient descent, step 0.0005 (PhaseT3M default)
Constraints	positivity after every update and after every pass; Butterworth low-pass at 0.8 Å^{-1}, order 10, every pass
Rotation	cubic interpolation (PhaseT3M)
Propagation	
TPGA-GS	particle by complex-covariance splat to the exit plane; carbon slices depth-propagated in Fourier space and summed; one exponential
Ray-GS	particle and carbon projected to $z = 0$, one exponential, Fresnel propagation over 155.74 Å
Compute (per renderer, one RTX 4090, 24 GB)	
Time	TPGA-GS 75 s/pass, 188 min; Ray-GS 58 s/pass, 146 min
Memory	TPGA-GS 12.7 GiB, Ray-GS 9.5 GiB

Table 3: Settings of the Co_3O_4 reconstructions. All values are fixed before training; no setting is tuned on the PhaseT3M reference or on the held-out view.

The frozen field of Figure 3 has $200{,}000$ Gaussians with widths 0.34–0.60 Å (median 0.45 Å). The multislice reference is abTEM 1.0.10 on an 800^2 grid cropped to 600^2 after the lens, 0.5 Å slices of the exact erf-integrated Gaussian potential plus the atomic carbon, with abTEM's $2/3$-Nyquist antialiasing and $C_s = 2.3$ mm. An independent multislice reproduces the abTEM intensities to a relative error of 7.6×10^{-6}.

The $\{400\}$ score is computed on a common 192^3 region in a $\pm 2.2\%$ shell around $1/2.02$ Å^{-1}: the concentration of Fourier power in the shell, its 99.95th percentile over its mean, divided by the same concentration after the shell phases are replaced by a fixed random field, so a lattice spot raises the score above 1 and diffuse power does not. Fourier correlations are signed, without alignment or rescaling, after masking and removing the mean; the missing-wedge subset is the Fourier cone outside the sampled $\pm 65^\circ$.

D ICE-TIDE BASELINE

Protocol. ICE-TIDE (Debarnot et al., 2025) was run with its released code on EMPIAR-10164 (TS_01) and EMPIAR-10643 (b2tilt40). The tilt series were motion-corrected (10164, with an even/odd frame split) and globally aligned with AreTomo3 (shifts and tilt axis, no local patches, since ICE-TIDE estimates local deformations itself), rotated to put the tilt axis along y, Fourier-cropped to bin 8, 4 and 2, and split into even/odd frames (10164) and even/odd tilts (both series). One network is trained per half for 1000 epochs with the published loss, optimizer and regularization weights. Four settings differ from the released defaults: the recorded tilt angles replace a uniform grid; the box thickness is the measured specimen thickness plus 40 nm; the finest hash-grid level matches the largest volume dimension; and each half has its own random seed and a random coordinate offset, because halves that share the hash-grid lattice correlate at fine scale through the lattice rather than through the data. The WBP baseline is a ram-lak filtered back-projection (ASTRA) of exactly the same input stacks, without CTF correction, since ICE-TIDE has none. FSC is computed in the central box of at most 1024 px per axis with a 10 px cosine taper and no fiducial mask, and the ICE-TIDE halves are registered within ± 5 px before FSC. Two controls accompany every ICE-TIDE pair: deliberately misaligned halves, whose FSC vanishes in every case, and each half's field against the WBP of the same half, which measures what the field shares with the data. Figure 10 shows central slices and the FSC curves behind Table 4.

Table 4: ICE-TIDE and a WBP of the same input stacks: half-map FSC crossings in Å at 0.5 and 0.143, with ICE-TIDE training time per half and peak GPU memory on one Quadro RTX 8000. *: the curve does not fall below 0.143 before Nyquist (21.6 Å). †: the data-consistency control finds half-map correlation in a frequency band that the field does not share with the data. 10164 values include the fiducials. Tilt halves double the angular step, which penalizes WBP more than a fitted field.

			ICE-TIDE		WBP			
Dataset	**Halves**	**Bin**	**0.5**	**0.143**	**0.5**	**0.143**	**min/half**	**GiB**
EMPIAR-10643	tilts	8	208	88	191	99	7.5	5.2
		4	231	94	193	101	13.2	8.8
		2	249	82	210	100	23.2	15.1
EMPIAR-10164	frames	8	78†	–*†	80	32	16.8	7.0
		4	82	38	80	32	28.6	12.7
		2	38†	27†	81	32	54.3	23.6
	tilts	8	86	62	213	154	8.4	7.0
		4	85	41	217	156	14.3	12.6
		2	75	29	226	164	28.5	23.3

Reading. On 10643, where Table 1 uses the same tilt split, ICE-TIDE crosses 0.5 coarser than its WBP at every binning and 0.143 finer, and the controls pass at every binning. CryoWaveGS crosses 0.5 finer than its own WBP at every binning (Table 1); its 0.143 half-map crossing is not a data statement (§5.2), so 0.5 is the only threshold at which the two fitted methods can be compared. On 10164 the frame-half curves of both ICE-TIDE and WBP carry the fiducials, whose signal is removed by the mask in Table 1; the unmasked values here are therefore not comparable with the masked ones there.

E DATASETS

EMPIAR series. EMPIAR-10164 (immature HIV-1 VLPs) is the experimental bridge: it ships unaligned multi-frame movies, permitting frame-half reconstructions at identical tilt geometry, and it has been used by both NovaCTF-related work and PhaseT3M (EMPIAR-10164; Lee, 2026). EMPIAR-10643 (in vitro HIV-1 Gag assemblies) is the dataset used by CryoETGS (Zhang, 2026); it ships no frames, so its halves are even and odd tilts. We note that all shipped CryoETGS configurations run at $\leq \sim 960$ px, so no published Gaussian-splatting cryo-ET baseline exists near native sampling. EMPIAR-10453 (SARS-CoV-2 virions, tilt series TS_082, 272 nm of ice) is reconstructed

Table 5: Datasets and reconstruction settings. Pixel sizes are given at every binning level used. The Co_3O_4 series has three defoci per tilt; EMPIAR-10643 contains in vitro assemblies. EMPIAR-10453 appears only in Appendix A.

Dataset	Specimen	Tilts / range	Bin	Pixel (Å)
EMPIAR-10164	Immature HIV-1 VLPs	41, $\pm 60^\circ/3^\circ$	8/4/2/1	10.8/5.4/2.7/1.35
EMPIAR-10643	HIV-1 Gag assemblies	41, $\pm 60^\circ/3^\circ$	8/4/2/1	9.44/4.72/2.36/1.18
EMPIAR-10453	SARS-CoV-2 virions	37, -60° to $+48^\circ/3^\circ$	8/4	10.63/5.32
Co_3O_4 (PhaseT3M)	Nanoparticle on carbon	131, $\pm 65^\circ$	1	0.52

only for Fig. 14. The Co_3O_4 series is the focal tilt series deposited with PhaseT3M, with per-image optics and a deposited reconstruction (§5.3).

F COMPUTE

Cryo-ET series. Each half runs on one NVIDIA RTX PRO 6000 Blackwell Max-Q (96 GB), and the two halves run in parallel on two GPUs, so one dataset takes about 2 h of wall time (Table 6). One rendered view costs 0.55–1.9 s including the backward pass, on a $2\times$ padded grid. The wave and ray arms cost the same to within 1%: at 10164 bin 2 they take 34.5 and 34.2 min with 9.7 GB each, and in the matched bin 1 pair 50.7 and 50.4 min with 7.8 and 7.7 GB.

Table 6: Compute per half on the cryo-ET series, one NVIDIA RTX PRO 6000 Blackwell Max-Q (96 GB) per half. Ranges span the two datasets.

Stage	Gaussians	Updates $\times$ views	s/update	min/half	Peak GPU
bin 8, full field	12 M	500×2	1.2	10	8 GB
bin 4, full field	12 M	1000×2	1.7–1.8	29–30	9 GB
bin 2, ROI	10 M	800×2	2.6–3.9	34–54	9.7 GB
bin 1, box	1.3–1.8 M	150×20	11–14	28–35	5 GB
Total per dataset				108 (10164), 122 (10643)	$\leq$ 10 GB

Co_3O_4 series. Each renderer runs on one NVIDIA RTX 4090 (24 GB) of a host with a 32-core Threadripper PRO 3975WX and 251 GB of RAM (Table 8). TPGA-GS is $1.29\times$ slower than Ray-GS and uses $1.3\times$ more memory. On voxels, TPGA is faster than multislice: on the same GPU one voxel TPGA update (0.136 s) is $1.8\times$ faster than one multislice update (0.245 s) in a paired benchmark, and 100 passes of voxel TPGA take 29 min against 122 min for a PhaseT3M run, which also refines the optics, so that ratio is indicative only. End to end, the flat-grid run at 1.6 M Gaussians takes 75 s per pass against 17.5 s for voxel TPGA (12.7 vs. 6.8 GiB), which puts TPGA-GS at about the speed of PhaseT3M (188 min for 150 passes against about 183 min extrapolated). The renderer accounts for little of that gap: it costs 67 ms of the roughly 440 ms per update by which the flat-grid run exceeds voxel TPGA, and the rest is pipeline overhead, the voxel carbon field and memory clearing between steps, not yet profiled.

Operator cost. Table 7 isolates the renderers: one training update of a Gaussian particle 312 Å deep, against multislice of the same field with 600 slices. Multislice pays 600 sequential FFT pairs, about 0.24 s whatever the number of Gaussians, while TPGA grows with the number of primitives, about 42 ns each. At about 20 k Gaussians, one per atom of this particle, TPGA costs the same as the ray renderer and $139\times$ less than multislice; at the 1.6 M of the flat grid the saving falls to $3.7\times$. The TPGA-to-ray ratio grows with the count, up to $5.8\times$, because the depth phase widens each footprint about $2\times$ and so covers about $4\times$ more pixels. A cost advantage in practice therefore requires the compact representation that §5.3 does not yet reach on measured data.

Table 7: One training update (forward and backward) of a Gaussian particle, 312 Å deep, on one NVIDIA RTX 4090; multislice uses 600 slices. The benchmark favors multislice: each Gaussian enters only one slice, while a real multislice spreads it over about five.

Gaussians	TPGA-GS	Ray-GS	Multislice	Multislice / TPGA
20 k	1.7 ms	1.7 ms	236 ms	139×
100 k	4.4 ms	1.5 ms	237 ms	54×
400 k	15.8 ms	3.4 ms	243 ms	15×
1.6 M	66.7 ms	11.6 ms	250 ms	3.7×

Table 8: Compute per renderer on the Co_3O_4 series, one NVIDIA RTX 4090 (24 GB) each. GPU memory is the CuPy pool plus the PyTorch peak.

	TPGA-GS	Ray-GS
Gaussians, pass 1–100 / after prune	1,598,858 / 1,420,862	1,598,858 / 1,468,761
Trainable parameters, particle	1.6 M masses, then + 4.3 M center coordinates	
Carbon voxel field	$600^3 = 216$ M voxels	
Updates	19,650 (131 tilts × 150 passes, 3 defocus images each)	
Time per pass / per tilt update	75 s / 0.57 s	58 s / 0.44 s
Cumulative time, pass 100 / 150	125 / 188 min	97 / 146 min
GPU memory	$9.3 + 3.4 \approx 12.7$ GiB	$7.7 + 1.8 \approx 9.5$ GiB
Checkpoint size	0.89 GB	0.89 GB
Image scoring, 393 images	39 s	32 s

G CHOICE OF EVALUATION METRICS

Each entry states the question it answers and the caveat that bounds it.

Even/odd-tilt FSC and frame-half FSC (experimental). *Is the reconstruction reproducible under independent noise?* We prefer frame-half splits, which separate noise realizations at identical tilt geometry. Even/odd *tilt* splits share the specimen, alignment and common missing wedge and are correlated through common alignment errors, so they measure reproducibility rather than resolution (Cardone et al., 2005); frame-half splits separate noise but not model error. Half-map FSC between two fitted fields is read at 0.5 (van Heel & Schatz, 2005), because the shared Gaussian size floor keeps two fitted fields correlated above 0.143 up to the band limit. The 0.143 criterion (Rosenthal & Henderson, 2003) is calibrated for independent halves and is used only for the cross-FSC of one half's field against the prior-free WBP of the other half (§5.2).

SSNR per tilt from odd/even frame splits. Used both as the loss weighting W_θ of §4 and as an evaluation quantity.

Projection PSNR and SSIM are widely reported in the Gaussian-splatting tomography literature. We do not report them here, for three reasons specific to this setting. First, per-tilt SNR in cryo-ET is far below one — the SHREC 2020 simulation uses 0.004 — so a metric computed against a noisy reference rewards fitting the noise as well as the signal. Second, both metrics are computed in projection space, and projection fidelity and reconstructed-volume fidelity are not equivalent for an unbounded primitive representation. The clearest evidence comes from within this literature: the R^2-Gaussian ablation reports that rectifying the integration bias is worth $+3.15$ dB in 2D rendering PSNR but $+17.77$ dB in 3D reconstruction PSNR (Zha et al., 2024), so a small 2D difference can correspond to a large 3D one. Third, PSNR is dominated by global scale and offset and by the low-frequency ice background, and SSIM's local luminance, contrast and structure decomposition (Wang et al., 2004) assumes spatially stationary statistics and a meaningful local mean, neither of which holds at cryo-ET SNRs. The frequency-based metrics of §5.1 are chosen because each answers a question that a single scalar in projection space cannot.

The same reasoning applies to the R-factor, which is an in-sample reprojection measure. We report it for comparability with PhaseT3M and read it as a data-consistency check, not as a resolution claim.

H Extended Background: Electron Image Formation

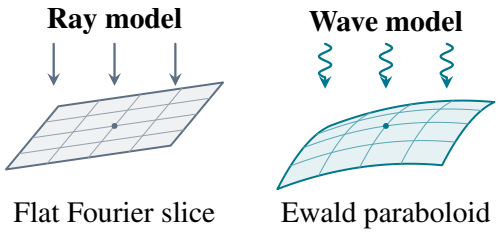

Figure 25: Fourier sampling at one tilt: flat slice versus Ewald paraboloid (curvature exaggerated).

Propagation and transfer. Paraxial free-space propagation over a distance Δz multiplies a spatial frequency $\mathbf{q}$ by the Fresnel transfer function $P_{\Delta z}(\mathbf{q}) = \exp[-i\pi\lambda\Delta z\|\mathbf{q}\|^2]$, where λ is the relativistic electron wavelength ($\lambda \approx 1.97$ pm at 300 kV). After the specimen, the objective lens applies $H_\theta(\mathbf{q}) = E_\theta(\mathbf{q})\exp[-i\chi_\theta(\mathbf{q})]$, where χ collects defocus, astigmatism, spherical aberration, and higher aberration terms, and E collects the spatial/temporal coherence and detector envelopes. The detector image follows from equation 2, with the measured MTF applied to the intensity.

Multislice. The object is divided into axial slabs. Slab k with projected potential $V_k(x, y)$ has transmission function $t_k(x, y) = \exp[i\sigma V_k(x, y)]$, where σ is the interaction constant, and the wave advances as

$$\psi_{k+1} = \mathcal{F}^{-1}[P_{\Delta z}\,\mathcal{F}\{t_k\psi_k\}]\,. \tag{9}$$

The wave entering slab $k + 1$ depends on every preceding slab, so the recurrence cannot be parallelized over depth. This is the structural cost that the primitive formulation of §4 avoids, and it is also why PhaseT3M's own discussion places multislice inversion at roughly $100\times$ conventional reconstruction compute.

Validity boundaries. Vulović et al. (2014) derive quantitative conditions for four models — the projection assumption (PA), PA+WPOA, WPOA alone, and TPGA. PA and WPOA are satisfied in cryo-EM tomograms coarser than ∼30 Å but fail near ∼5 Å for typical macromolecules, while TPGA produces exit waves comparable to full multislice under far weaker conditions. TPGA is therefore our forward-model target: multislice-accurate where it matters, closed-form where possible. Table 9 tabulates the depth of field $t_{\mathrm{crit}} = d^2/\lambda$ of equation 3 against the binning levels used by published Gaussian-splatting cryo-ET work and against the Co_3O_4 series.

Table 9: Depth of field $t_{\mathrm{crit}} = d^2/\lambda$ at 300 kV ($\lambda = 1.97$ pm), every row from the same λ. The sampling d is twice the pixel size, for a native pixel of about 1.5 Å. Cryo-ET lamellae are typically 100–300 nm thick; the Co_3O_4 particle and carbon span 31 nm.

Level	**Sampling d**	t_{crit}	**Ray model on paper**
bin6	18–24 Å	1.6–2.9 μm	Adequate
bin4	12.1 Å	∼740 nm	Adequate
bin2	6.0 Å	∼180 nm	Borderline for lamellae
bin1	3.0 Å	∼46 nm	Violated for lamellae
Co_3O_4, {400}	2.02 Å	∼21 nm	Violated (31 nm specimen)

Weak-phase condition. A commonly used worst-case form is $\sigma V_z < 0.36$. This concerns phase strength and multiple scattering and must not be confused with the projection assumption: a specimen can satisfy the weak-phase condition while still violating the projection assumption, because diffraction accumulates during propagation.

I Positioning Along the Three Axes

Table 10: Positioning of the closest prior work along the three axes.

Method	Representation	Forward physics	Depth	Raw tilts	Align.
WBP / SIRT / SART	Voxels	Ray	Discarded	Yes	Ext.
NovaCTF	Voxels	Ray + 3D-CTF	Quantized slabs	Yes	Ext.
IsoNet / DeepDeWedge	Voxels / CNN	Post-reconstruction	—	No	No
ICE-TIDE	Implicit field	Ray	Discarded	Yes	Joint
R^2-Gaussian	Gaussians	Ray	Discarded	Yes	—
CryoETGS	Gaussians	Ray	Discarded	Yes	Joint
PhaseT3M	Voxel / slices	Multislice	Discretized	Yes	Joint
CryoWaveGS	**Gaussians**	**TPGA / wave**	**Continuous**	**Yes**	**Joint**