An introduction to a cryo-EM simulator: InSilicoTEM

for group meeting at ISS, AS

Chun Lu Huang for [1]

Motivation

Overview of recent simulators

Results in InSilicoTEM

Conclusion

References

Motivation

In order to estimate contrast transfer functions, CTFs, in the cryo-EM images, a simulator for cryo-EM is needed. The simulator should produce the cryo-EM images that reflect the given physical conditions (i.e. input parameters). Then those produced cryo-EM images can be observed as a benchmark for the estimation of contrast transfer function.

Overview of recent simulators

Recently, some cryo-EM simulators which take into account the physical situations have been developed. The simulators are:

- InSilicoTEM [1] which is to simulate cryo-EM images of specimens embedded in vitreous ice, and is implemented in MATLAB with DIPimage, released on their github
- cisTEM, which refers to another simulator. See their tutorials

Overview of InSilicoTEM

- A particle or wave can be a probe to an object. This idea is taken by InSilicoTEM and its simulation is firstly to model interaction between particle/wave (electrons) and object (potential of specimen)
- Besides, the microscope aberrations introduce phase shift into the electron wave left from specimen
- the final image is captured by detector. The conversion from electron wave to a digital signal involves various detector properties affecting the final images.

Electric potential from a point charge

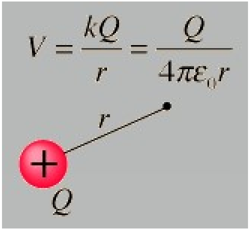
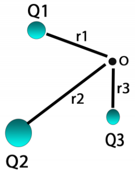


Figure: The electric potential, V, at a distance, r, produced by a point charge, Q. Figure from http://hyperphysics.phy-astr.gsu.edu/hbase/electric/potpoi.html

Electric potential from multiple charges



$$V = \frac{kQ1}{r1} + \frac{kQ2}{r2} + \frac{kQ3}{r3}$$

Figure: The electric potential, V, at a distance, o, produced by three point charges, Q_1, Q_2 and Q_3 in distances, r_1, r_2, r_3, respectively. Figure from https://byjus.com/physics/electric-potential-point-charge/

Interaction potential

The image formation model incorporated in InSilicoTEM models the three physics considerations:

- Interaction potential modeled as a sum of two interaction potentials:

$$V^{\mathrm{int}}(\mathbf{r}) = V^{\mathrm{atom}}(\mathbf{r}) + V^{\mathrm{bond}}(\mathbf{r}) \tag{1}$$

V^{atom}: superposition of atomic potentials (each atom has no correlation) in specimen.
V^{bond}: changes in charge density due to electrostatic interaction between atoms in specimen and due to influence of solvent and ions interactions (about 5% [2] compared to V^{atom})

Isolated atom superposition approximation, IASA

We can consider the specimen as a set of isolated atoms, and the total potential of system is given by:

$$V_{\mathrm{atom}}(\mathbf{r}) = \sum_{j=1}^{m} V_{Z_j}(\mathbf{r} - \mathbf{R}_j)$$

where V_Z is the potential of an isolated atom centered at R_j and Z is atomic number. With some physics approximation, such potential can be written as:

$$V(\mathbf{r}) = \frac{16\pi\hbar^2}{me} \int f_Z^{(e)}(\xi) e^{4\pi i \xi r} d^3\xi \tag{2}$$

where $f_Z^{(e)}$ is the scattering factor of the atom, and ξ is related to the spatial frequency. A scattering factor is an observable observing the scattering amplitude vs ξ (spatial frequency)

Double-slit experiment with bullets

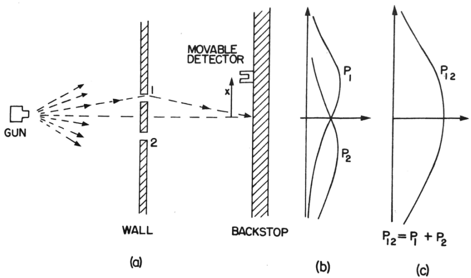
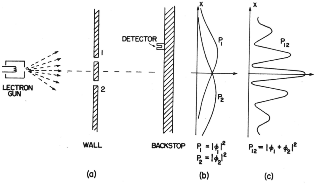


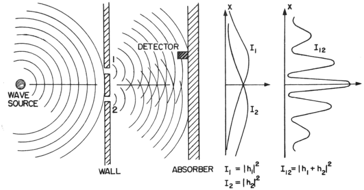
Figure: A double-slit experiment with indestructable bullets. Figure from http://www-thphys.physics.ox.ac.uk/people/FabianEssler/QM/QM2019.pdf

The detector at position, x detects bullets. Probability, $P_1(x)$ and $P_2(x)$, of measuring bullets with closing slits, 2 and 1, respectively are obtained. $P_{12}(x)$ is sum of P_1 and P_2.

Double-slit experiment with electrons



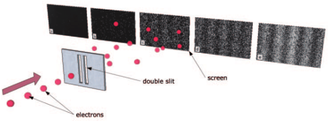
(a) A double-slit experiment with electrons.



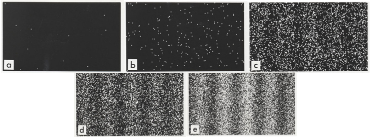
(b) A double-slit experiment with waves.

In quantum mechanics, $P_{12}(x)$ in (a) looks like the interference pattern that is observed in the experiment with waves in (b). So, $P_{12}(x)$ is now $P_{12} = |\phi_1 + \phi_2|^2$ instead of $P_{12}(x) \neq P_1(x) + P_2(x)$. The above figure is from http://www-thphys.physics.ox.ac.uk/people/FabianEssler/QM/QM2019.pdf

Image obtained from double-slit experiment with electrons



(a) A double-slit experiment with electrons.



(b) A double-slit experiment with waves.

Plates (a)-(e) are the buildup of interference pattern at electron level over time.

Wave function and probability

- $\psi(x)$ is probability amplitude at x
- $P = \int_{-\infty}^{\infty} \psi^*(x)\psi(x)dx$ is probability of finding a particle at x

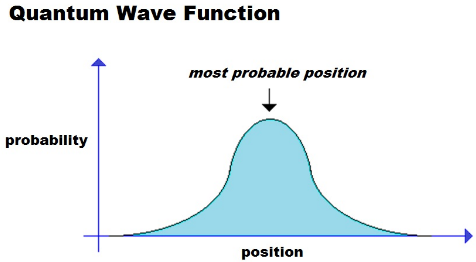


Figure: Probability distribution. Figure from https://themeaningolife.wordpress.com/tag/quantum-theory/

Wave propagation in InSilicoTEM

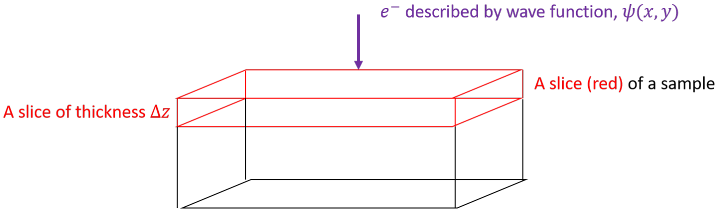


Figure: Wave propagation through slices of a sample.

Wave propagation in InSilicoTEM

- When electron wave propagates though specimen, its wave function varies slowly with z, the optical axis. To solve wave function, multislice method is utilized. Within the method, Δz is the thickness of a slice of specimen and $V_z(x, y, z)$ is prjected potential inside slice, then wave function at top of n+1 slice of specimen is given by

$$\psi_{n+1}(x, y) = \mathcal{F}^{-1}\{P_n(q_x, q_y, \Delta z_n)\mathcal{F}[\exp(i\sigma V_z(x, y, z))\psi_n(x, y)]\} \tag{3}$$

where $P_n(q, \Delta z) = \exp(-i\pi\lambda\Delta z q^2)$ is Fresnel propagator over a slice of thickness Δz and it is depends on spatial frequency (q_x, q_y)

Eq. 3 shows that wave propagation through specimen can be seen as a recursive transmission and propagation finishes until wave leaves specimen, then the wave becomes $\psi_{\mathrm{exit}}(x, y)$.

Mean free path in gas

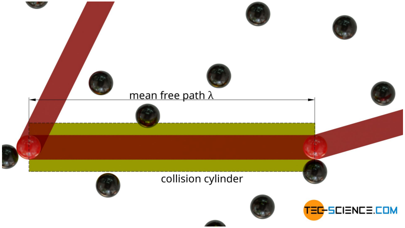


Figure: Illustration of the mean free path, λ, in gas consisting of one type molecule. A particle in red travels along the red cylinder path and collides with two other particles in black. The distance between the two collisions is λ. Figure from https://www.tec-science.com/thermodynamics/kinetic-theory-of-gases/mean-free-path-collision-frequency/

Inelastic interactions

The transmitted electrons, the main contribution to the cryo-EM image, can either lost energy (inelastic scattering) or lost no energy (elastic scattering) when leaving the sample. The effects of inelastic scattering are modeled as imaginary part of total potential:

$$V_{\mathrm{tot}}^{\mathrm{int}} = V_{\mathrm{ph}} + iV_{\mathrm{ab}} \tag{4}$$

- V_{ph} is interaction potential shown in Eq. 1
- V_{ab} influencing the amplitude contrast

The dominant contribution to inelastic interactions is plasmons. Plasmons can be described by mean free path Λ_{in} which is related to V_{ab}:

$$V_{\mathrm{ab}}(x, y, z) = 1/(2\sigma\Lambda_{\mathrm{in}})$$

Interaction constant $\sigma = 2\pi m|e|\lambda h^2$. λ, e, m and h are wavelength, charge, mass of electron and Planck's constant, respectively.

Electro-optical effects in cryo-EM

The cryo-EM image is formed mainly due to phase contrast. The wave function, $\psi_{\mathrm{bf}}(q,\alpha)$, in back focal plane is written as:

$$\psi_{\mathrm{bf}}(q,\alpha) = \mathcal{F}\left[\psi_{\mathrm{exit}}(x,y)\right] T(q,\alpha)$$

where $T(q,\alpha)$ is contrast transfer function taking into account the following properties:

- temporal incoherence (modeled as $K_c(q)$)
- spatial incoherence (modeled as $K_s(q,\alpha)$)
- aperture function (modeled as $A_p(q)$ including phase plate)

and is given by $T(q,\alpha) = K_s(q,\alpha)K_c(q)A_p(q)e^{-i\chi(q,\alpha)}$. Note that $\chi(q,\alpha)$ is aberration function related to spherical aberration, defocus, α and q.

Image sensor: Charge coupled device CCD

The CCD working principle is

- liberate electrons by exposing to photons
- collect charges at pixel levels
- read out (transfer) sequentially charges (from analog to digitized signals)

The read out signal, $S^{\mathrm{out}}_{x',y'}$, can be written as:

$$S^{\mathrm{out}}_{x',y'} = \{s_{x'=1,y'=1}, s_{x'=2,y'=1}, \ldots, s_{x'=x,y'=1}, s_{x'=1,y'=2}, \ldots, s_{x'=x,y'=2}, s_{x'=x,y'=y}\}$$

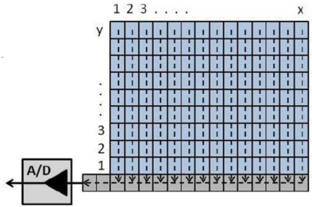


Figure: A CCD sensor array with pixel size $x \times y$ Figure from http://particle.astro.ru.nl/ps/astropract1-1112-hk5.pdf

Introduction to some terminologies for detector

We want a detector which can have high signal, low noise, and high resolution, and which doesn't loss information while transferring intensity into its captured image. We can define quantities to see a detector performance.

- **MTF** is image contrast as a function of spatial frequency in a detector, meaning that MTF observes the detector capability to transfer the input signal at a given spatial frequency to its output.
- **NTF** is noise as a function of spatial frequency. It can see how much noise is arose in the output with the input signal at a given spatial frequency.
- **DQE** refers to the efficiency of a detector in converting incident intensity into an image signal. The ideal DQE is one. High DQE indicate that less exposure is needed to achieve identical image quality.

Detector response

- Detected image is formed through conversion from electron wave to digital signal in detector. This process involves detector effects characterized by, for example, conversion factor (CF), detective quantum efficiency (DQE), shot noises... etc.

Detected image, I (number of recorded photons/electrons), can be written as:

$$I(x, y) = I_{\mathrm{rn}} + I_{\mathrm{dn}} + CF\mathcal{F}^{-1}\left\{\mathcal{F}\left[P_{\mathrm{poisson}}\left(\phi_e \cdot \mathcal{F}^{-1}(\tilde{I}_0(q)\sqrt{\mathrm{DQE}(q)})\right)\right]\mathrm{NTF}(q)\right\} \quad (5)$$

- I_{rn} readout noise generated from amplifier
- I_{dn}, dark noise, which is arose from dark current building up on sensor caused by thermal energy
- conversion factor CF [ADU/e^-]
- P_{poisson} describes shot noise following Poisson distribution
- $\mathrm{DQE} = (\frac{\mathrm{MTF}^2}{\mathrm{NTF}^2})$ is detective quantum efficiency and MTF is modulation transfer function depending on spatial frequency q
- NTF ($\mathrm{NTF}^2 = \frac{\mathrm{NPS}}{CF^2\phi_e}$) is noise transfer function. NPS noise power spectrum and ϕ_e incident electron flux
- $\tilde{I}_0$ is noise-free signal in fourier domain

Results in InSilicoTEM: bond contributions

A comparison is constructed to see the effect of interaction potential modeled in Eq. 1. Four particles, proteasome, hemoglobin in 50mM, 1M, 3M NH_4Ac, are used to perform the comparison (see figure below).

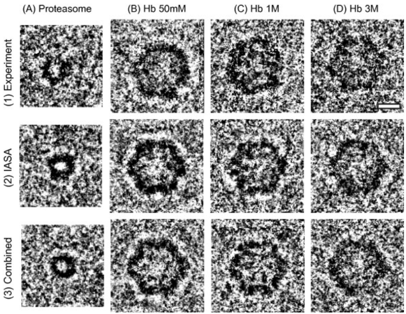


Figure: (1) experimental images (2) simulated images where the potential is V_{atom}. (3) simulated images where the potential is $V_{\mathrm{atom}} + V_{\mathrm{bond}}$. The scale bar is 10 nm. From (3), the contribution of V_{bond} is weak. Figure from [1]

Results in InSilicoTEM: defocus series

Various defocus series, ranging from 500 nm to 4000 nm, are set to observe the effect in the simulated images. In Figure below, 20S proteasome is used with the conditions of 80 kV for 0.5 s, 1 s exposure time.

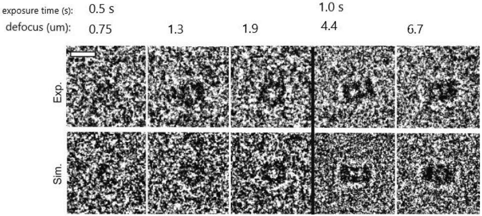


Figure: Experimental and simulated images of 20S proteasome are shown. The simulations predict the changes in the images with increasing defocus. At low defocus the contrast is too small to distinguished from noise, while at large defocus the white fringes can be recognized. The scale bar is 10 nm. Figure from [1]

Results in InSilicoTEM: Detector's DQE

The DQE is necessary in modeling the detector.

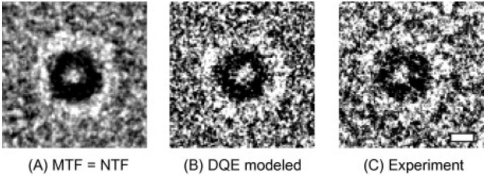


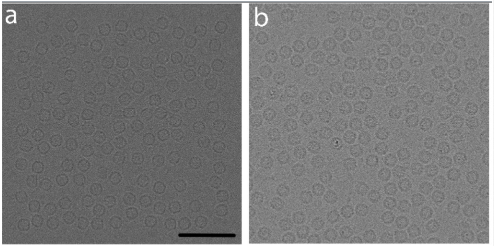
Figure: A comparison is shown with and without DQE in the detector simulation. (A) simulated image assuming MTF=NTF. (B) simulated image taking into account the measured DQE. (C) experimental image for exposure time of 1 s, ice thickness, d = 92 nm, Δf = 6713 nm, and motion factor, $\sigma_{mot} = 6$ $\AA$. The scale bar is 5 nm. Figure from [1]

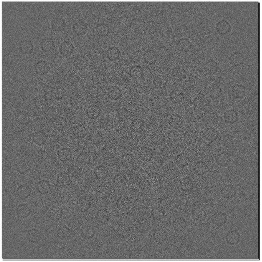
Tests within InSilicoTEM

- reproduction of the example figure in the paper [3]
- comparison between simulated and experimental micrographs–study A
- production of the simulated micrographs with various defocus values in order to see the cability concerning defocus–study B

Reproduce micrograph in the paper

The paper [3] uses InSilicoTEM to study the structure and there is an figure (see below)





Apo-ferritin example from paper [3] in (a) simulated micrograph at 1220 nm; and in (b) experimental micrograph at 1201 nm.

The right panel shows a simulated micrograph at 1246 nm which was produced from my side.

Setup for study A

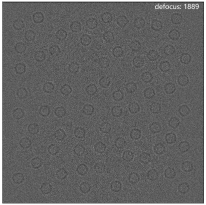
Data, EMPIAR-10216 (apo-ferritin), is used. This data have 59 frames in a micrograph and are also used in [3] which utilizes InSilicoTEM as well.

Experimental conditions:

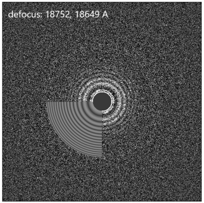
- operating voltage: 300 kV
- Spherical aberrations: 2.7 mm
- dose: 50 $e/\AA^2$
- detector: Falcon III in counting mode
- image size: 4k

Procedure for processing data: take one micrograph → apply CTFFind4 to the micrograph
Procedure for processing simulated production: set the same input conditions → execute the simulation to produce a micrograph → apply CTFFind4 to the simulated micrograph.

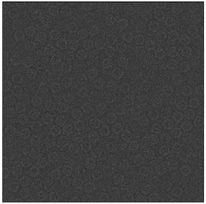
Comparison between the micrographs



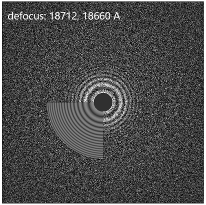
(a) simulation at 1889 nm



(b) ctf estimation corresponding to (a)



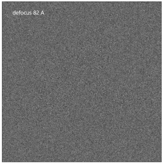
(c) data



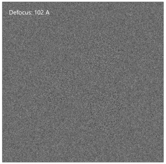
(d) ctf estimation corresponding to (c)

Defocus variations

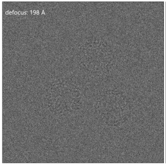
Defocus variations from [0,50] to [150, 200], and [1900,2000] (nm) are given. Image size is 1024×1024.



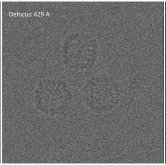
(a) Defocus: 82 nm



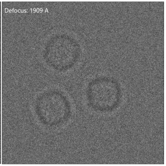
(b) Defocus: 102 nm



(c) Defocus: 198 nm



(d) Defocus: 629 nm



(e) Defocus: 1909 nm

Figures (a)-(e) shows simulated micrographs with increasing defocus. The defocus range [0,50] is a not possible range for InSilicoTEM to produce micrographs.

Comments on InSilicoTEM

- InSilicoTEM take throughout physics mechanisms, such as interaction between specimen and electron beams, electro-optics effects, detector response, in cryo-EM and simulates the ideal micrographs;
- InSilicoTEM cannot produce movie data, however, a low dose micrograph can be simulated.

References

[1] Miloš Vulović et al. "Image formation modeling in cryo-electron microscopy". In: *Journal of Structural Biology* 183.1 (2013), pp. 19–32. DOI: https://doi.org/10.1016/j.jsb.2013.05.008.

[2] Peng LM et al. *High Energy Electron Diffraction and Microscopy*. Vol. 61. 2004. DOI: https://doi.org/10.1016/j.jsb.2013.05.008.

[3] Y. Zhang et al. "Could Egg White Lysozyme be Solved by Single Particle Cryo-EM?" In: *J. Chem. Inf. Model.* 60.5 (2020), pp. 2605–2613.