Supplementary information for Three-dimensional deconvolution for large-angle illumination annular dark-field scanning transmission electron microscopy depth sectioning

Taichi Kusumi[1], Hajime Ueda[1], Toshihiro Futazuka[2], Masatoshi Hanai[3], Shun Katakami[1], Kazuaki Kawahara[2], Ryo Ishikawa[2*], Naoya Shibata[2], Masato Okada[1*]

[1]Graduate School of Frontier Sciences, The University of Tokyo, 5-1-5, Kashiwa, 277-8561, Chiba, Japan.
[2]Institute of Engineering Innovation, School of Engineering, The University of Tokyo, 7-3-1, Bunkyo, 113-8656, Tokyo, Japan.
[3]Information Technology Center, The University of Tokyo, 7-3-1, Bunkyo, 113-8656, Tokyo, Japan.

*Corresponding author(s). E-mail(s): ishikawa@sigma.t.u-tokyo.ac.jp; okada@edu.k.u-tokyo.ac.jp;

Supplementary Note 1: Formalizing 3D deconvolution as linear regression

In this section, we formalize 3D deconvolution as linear regression. Here, we denote by $\otimes_i$ $(i = 1, 2, 3)$ the 1D, 2D, and 3D convolution operators, respectively. First, we consider the 1D convolution. Since the n-th entry of the convolution of $\boldsymbol{a} = [a_0, a_1, \cdots, a_{M-1}]^\top \in \mathbb{R}^M$ and $\boldsymbol{b} = [b_0, b_1, \cdots, b_{N-1}]^\top \in \mathbb{R}^N$ is given by

$$[\boldsymbol{a} \otimes_1 \boldsymbol{b}]_n = \sum_i a_{n-i} b_i, \tag{S1}$$

the convolution can be expressed as

$$\begin{aligned}
\boldsymbol{a} \otimes_1 \boldsymbol{b} &= \underbrace{\begin{bmatrix}
a_0 & 0 & 0 & \cdots & 0 \\
a_1 & a_0 & 0 & \cdots & 0 \\
a_2 & a_1 & a_0 & \ddots & \vdots \\
\vdots & \vdots & \vdots & \ddots & 0 \\
a_{M-1} & a_{M-2} & a_{M-3} & & a_0 \\
0 & a_{M-1} & a_{M-2} & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & a_{M-3} \\
0 & \cdots & 0 & a_{M-1} & a_{M-2} \\
0 & \cdots & 0 & 0 & a_{M-1}
\end{bmatrix}}_{\text{Toeplitz matrix}}
\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_{N-1} \end{bmatrix} \\
&=: \boldsymbol{T}(\boldsymbol{a})\boldsymbol{b}.
\end{aligned} \tag{S2}$$

Here, $\boldsymbol{T}(\boldsymbol{a}) \in \mathbb{R}^{(M+N-1)\times N}$ denotes the Toeplitz matrix.

Next, we consider the 2D convolution. The (m, n)-th entry of the convolution of $\boldsymbol{A} \in \mathbb{R}^{M_1 \times M_2}$ and $\boldsymbol{B} \in \mathbb{R}^{N_1 \times N_2}$ is given by

$$\begin{aligned}
[\boldsymbol{A} \otimes_2 \boldsymbol{B}]_{m,n} &= \sum_j \underbrace{\sum_i A_{m-i,n-j} B_{i,j}}_{m\text{-th entry of 1D convolution}} \\
&= \sum_j [\boldsymbol{a}_{n-j} \otimes_1 \boldsymbol{b}_j]_m \quad (\because \text{Eq. (S1)}) \\
&= \sum_j [\boldsymbol{T}(\boldsymbol{a}_{n-j})\boldsymbol{b}_j]_m \quad (\because \text{Eq. (S2)}) \\
&= \left[\sum_j \boldsymbol{T}(\boldsymbol{a}_{n-j})\boldsymbol{b}_j\right]_m,
\end{aligned} \tag{S3}$$

where $\boldsymbol{a}_{n-j} = [A_{0,n-j}, A_{1,n-j}, \cdots, A_{M_1-1,n-j}]^\top$ and $\boldsymbol{b}_i = [B_{0,j}, B_{1,j}, \cdots, B_{N_1-1,j}]^\top$ are the $(n-j)$- and j-th column of $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively. Thus, the 2D convolution can be expressed as

$$
\begin{aligned}
\mathrm{vec}(\boldsymbol{A} \otimes_2 \boldsymbol{B}) &= \begin{bmatrix} \sum_j \boldsymbol{T}(\boldsymbol{a}_{0-j})\boldsymbol{b}_j \\ \sum_j \boldsymbol{T}(\boldsymbol{a}_{1-j})\boldsymbol{b}_j \\ \vdots \\ \sum_j \boldsymbol{T}(\boldsymbol{a}_{M_2-j})\boldsymbol{b}_j \end{bmatrix} \\
&= \begin{bmatrix}
\boldsymbol{T}(\boldsymbol{a}_0) & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\
\boldsymbol{T}(\boldsymbol{a}_1) & \boldsymbol{T}(\boldsymbol{a}_0) & \mathbf{0} & \cdots & \mathbf{0} \\
\boldsymbol{T}(\boldsymbol{a}_2) & \boldsymbol{T}(\boldsymbol{a}_1) & \boldsymbol{T}(\boldsymbol{a}_0) & \ddots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \mathbf{0} \\
\boldsymbol{T}(\boldsymbol{a}_{M_2-1}) & \boldsymbol{T}(\boldsymbol{a}_{M_2-2}) & \boldsymbol{T}(\boldsymbol{a}_{M_2-3}) & & \boldsymbol{T}(\boldsymbol{a}_0) \\
\mathbf{0} & \boldsymbol{T}(\boldsymbol{a}_{M_2-1}) & \boldsymbol{T}(\boldsymbol{a}_{M_2-2}) & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \boldsymbol{T}(\boldsymbol{a}_{M_2-3}) \\
\mathbf{0} & \cdots & \mathbf{0} & \boldsymbol{T}(\boldsymbol{a}_{M_2-1}) & \boldsymbol{T}(\boldsymbol{a}_{M_2-2}) \\
\mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \boldsymbol{T}(\boldsymbol{a}_{M_2-1})
\end{bmatrix}
\begin{bmatrix} \boldsymbol{b}_0 \\ \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_{N_2-1} \end{bmatrix} \\
&=: \boldsymbol{T}_2(\boldsymbol{A})\mathrm{vec}(\boldsymbol{B}),
\end{aligned} \tag{S4}
$$

where $\boldsymbol{T}_2(\boldsymbol{A}) \in \mathbb{R}^{(M_1+N_1-1)(M_2+N_2-1)\times N_1 N_2}$ denotes the doubly block Toeplitz matrix and $\mathrm{vec}(\cdot)$ represents the vectorization. Hence, the 2D convolution can be expressed as the matrix multiplication.

Similarly, the vectorization of 3D convolution between $\mathcal{A} \in \mathbb{R}^{M_1\times M_2\times M_3}$ and $\mathcal{B} \in \mathbb{R}^{N_1\times N_2\times N_3}$ is expressed as

$$
\begin{aligned}
\mathrm{vec}(\mathcal{A} \otimes_3 \mathcal{B}) &= \begin{bmatrix}
\boldsymbol{T}_2(\boldsymbol{A}_0) & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\
\boldsymbol{T}_2(\boldsymbol{A}_1) & \boldsymbol{T}_2(\boldsymbol{A}_0) & \mathbf{0} & \cdots & \mathbf{0} \\
\boldsymbol{T}_2(\boldsymbol{A}_2) & \boldsymbol{T}_2(\boldsymbol{A}_1) & \boldsymbol{T}_2(\boldsymbol{A}_0) & \ddots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \mathbf{0} \\
\boldsymbol{T}_2(\boldsymbol{A}_{M_3-1}) & \boldsymbol{T}_2(\boldsymbol{A}_{M_3-2}) & \boldsymbol{T}_2(\boldsymbol{A}_{M_3-3}) & & \boldsymbol{T}_2(\boldsymbol{A}_0) \\
\mathbf{0} & \boldsymbol{T}_2(\boldsymbol{A}_{M_3-1}) & \boldsymbol{T}_2(\boldsymbol{A}_{M_3-2}) & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \boldsymbol{T}_2(\boldsymbol{A}_{M_3-3}) \\
\mathbf{0} & \cdots & \mathbf{0} & \boldsymbol{T}_2(\boldsymbol{A}_{M_3-1}) & \boldsymbol{T}_2(\boldsymbol{A}_{M_3-2}) \\
\mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \boldsymbol{T}_2(\boldsymbol{A}_{M_3-1})
\end{bmatrix}
\begin{bmatrix} \mathrm{vec}(\boldsymbol{B}_0) \\ \mathrm{vec}(\boldsymbol{B}_1) \\ \mathrm{vec}(\boldsymbol{B}_2) \\ \vdots \\ \mathrm{vec}(\boldsymbol{B}_{N_3-1}) \end{bmatrix} \\
&=: \boldsymbol{T}_3(\mathcal{A})\mathrm{vec}(\mathcal{B}),
\end{aligned} \tag{S5}
$$

where

$$\boldsymbol{A}_i = \begin{bmatrix} \mathcal{A}_{0,0,i} & \mathcal{A}_{0,1,i} & \cdots & \mathcal{A}_{0,M_2-1,i} \\ \mathcal{A}_{1,0,i} & \mathcal{A}_{1,1,i} & \cdots & \mathcal{A}_{1,M_2-1,i} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{A}_{M_1-1,0,i} & \mathcal{A}_{M_1-1,1,i} & \cdots & \mathcal{A}_{M_1-1,M_2-1,i} \end{bmatrix} \in \mathbb{R}^{M_1 \times M_2},$$
$$\boldsymbol{B}_i = \begin{bmatrix} \mathcal{B}_{0,0,i} & \mathcal{B}_{0,1,i} & \cdots & \mathcal{B}_{0,N_2-1,i} \\ \mathcal{B}_{1,0,i} & \mathcal{B}_{1,1,i} & \cdots & \mathcal{B}_{1,N_2-1,i} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{B}_{N_1-1,0,i} & \mathcal{B}_{N_1-1,1,i} & \cdots & \mathcal{B}_{N_1-1,N_2-1,i} \end{bmatrix} \in \mathbb{R}^{N_1 \times N_2}. \tag{S6}$$

Here, we model ADF-STEM depth sectioning as

$$\boldsymbol{I} = \boldsymbol{P} \otimes_3 \boldsymbol{O} + \boldsymbol{n}', \tag{S7}$$

where $\boldsymbol{I}, \boldsymbol{O}, \boldsymbol{P}$ represent the intensity of an ADF-STEM depth sectioning image, an 3D object function, and a 3D probe function, respectively, and $\boldsymbol{n} \sim \mathcal{N}(\boldsymbol{0}, \gamma^{-1}\boldsymbol{E})$ denotes the additive white Gaussian noise (AWGN). Using Eq. (S5), we obtain

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{n}, \tag{S8}$$

where $\boldsymbol{y} = \mathrm{vec}(\boldsymbol{I}), \boldsymbol{A} = \boldsymbol{T}_3(\boldsymbol{P}), \boldsymbol{x} = \mathrm{vec}(\boldsymbol{O}), \boldsymbol{n} = \mathrm{vec}(\boldsymbol{n}')$, respectively. This is equaivalent to Eq. (2) in the main text. Therefore, 3D deconvolution, which recovers the 3D object function $\boldsymbol{O}$ from the intensity of an ADF-STEM depth sectioning image $\boldsymbol{I}$, can be formalized as linear regression.

Table S1 Estimated effective source size of ADF-STEM depth sectioning image of Si at various eletron dose.

Electron dose ($e^-/\text{Å}^2$)	10^5	2×10^5	5×10^5	10^6	2×10^6	5×10^6	10^7	Infinite
Estimated effective source size (pm)	59.0	54.5	52.7	50.0	42.5	48.6	48.1	47.3