

Supplementary Material

Measuring dynamic structural changes of nanoparticles at the atomic scale using electron microscopy

Annelies De wael,¹ Annick De Backer,¹ Lewys Jones,^{2,3,4} Aakash Varambhia,^{2,†} Peter D. Nellist,² and Sandra Van Aert^{1,*}

¹*EMAT, NANOLab Center of Excellence, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerp, Belgium*

²*Department of Materials, University of Oxford, Parks Road, OX1 3PH Oxford, United Kingdom*

³*Advanced Microscopy Laboratory, Centre for Research on Adaptive Nanostructures and Nanodevices (CRANN), Dublin 2, Ireland*

⁴*School of Physics, Trinity College Dublin, The University of Dublin, Dublin 2, Ireland*

1. PARAMETER ESTIMATION

Table I summarizes the notation used for the factorial hidden Markov model for atom counting.

1.1. Expectation-Maximisation algorithm

The expectation-maximisation algorithm for the factorial hidden Markov model for atom counting, also called the Baum-Welch algorithm, updates the parameter estimates in order to maximize the complete data likelihood expressed by equations (1)-(4) [1].

$$p(\mathbf{H}, \mathbf{O}; \boldsymbol{\Omega}) = p(\mathbf{h}_1; \mathbf{I}) \prod_{t=2}^T p(\mathbf{h}_t | \mathbf{h}_{t-1}; \mathbf{A}) \prod_{t'=1}^T p(\mathbf{o}_{t'} | \mathbf{h}_{t'}; \boldsymbol{\mu}, \sigma), \quad (1)$$

with

$$p(\mathbf{h}_1; \mathbf{I}) = \prod_{n=1}^N \prod_{g=0}^G \iota_g^{h_{1g}^{(n)}}, \quad (2)$$

$$p(\mathbf{h}_t | \mathbf{h}_{t-1}; \mathbf{A}) = \prod_{n=1}^N \prod_{g_1=0}^G \prod_{g_2=0}^G A_{g_1 g_2}^{h_{t-1, g_1}^{(n)} h_{t, g_2}^{(n)}}, \quad (3)$$

$$p(\mathbf{o}_t | \mathbf{h}_t; \boldsymbol{\mu}, \sigma) = \prod_{n=1}^N \prod_{g=0}^G \mathcal{N}(o_t^{(n)} | \mu_g, \sigma)^{h_{tg}^{(n)}}, \quad (4)$$

where $\mu_g = a\mathcal{M}_g$, with a a linear scaling parameter and \mathcal{M}_g the simulated scattering cross section for a column with g atoms (also called library value).

In the M-step, the parameters are updated to maximize the likelihood, using the following update formulas:

$$\iota_g = \frac{\sum_{n=1}^N \mathbb{E}[h_{1g}^{(n)}]}{\sum_{n=1}^N \sum_{j=0}^G \mathbb{E}[h_{1j}^{(n)}]}, \quad (5)$$

$$A_{jg} = \frac{\sum_{t=2}^T \sum_{n=1}^N \mathbb{E}[h_{tg}^{(n)} h_{t-1, j}^{(n)}]}{\sum_{t=2}^T \sum_{n=1}^N \sum_{g'=0}^G \mathbb{E}[h_{tg'}^{(n)} h_{t-1, j}^{(n)}]}, \quad (6)$$

$$a = \frac{\sum_{t=1}^T \sum_{n=1}^N \sum_{g=0}^G \mathbb{E}[h_{tg}^{(n)}] \mathbf{o}_t^{(n)} \mathcal{M}_g}{\sum_{t=1}^T \sum_{n=1}^N \sum_{g=0}^G \mathbb{E}[h_{tg}^{(n)}] \mathcal{M}_g^2}, \quad (7)$$

$$\sigma = \sqrt{\frac{\sum_{t=1}^T \sum_{n=1}^N \sum_{g=0}^G \mathbb{E}[h_{tg}^{(n)}] (\mathbf{o}_t^{(n)} - a\mathcal{M}_g)^2}{\sum_{t=1}^T \sum_{n=1}^N \sum_{g=0}^G \mathbb{E}[h_{tg}^{(n)}]}}. \quad (8)$$

During the E-step of the algorithm, the quantities $\mathbb{E}[h_{tg}^{(n)}]$ and $\mathbb{E}[h_{tg}^{(n)} h_{t-1, j}^{(n)}]$ are determined, and the likelihood is calculated. The iterative algorithm keeps updating the parameter estimates until the value of the likelihood has converged. As such we obtain maximum likelihood estimates of the model parameters of the factorial hidden Markov model for atom counting.

1.2. Viterbi algorithm

Next, the hidden state sequence is retrieved based on the estimated parameters using a path backtracking algorithm, called the Viterbi algorithm [2, 3]. The goal of the Viterbi algorithm for hidden Markov models is to determine the most likely hidden state sequence. We could determine the individually most likely state of column n at time t , $\mathbf{h}_t^{(n)}$, as the state that maximizes the $\mathbb{E}[h_{tg}^{(n)}]$, but the transition probability from $t-1$ to this state at time t may be zero, causing this state to be invalid. Therefore, the Viterbi algorithm considers the entire state sequence.

[†] Present address: Johnson Matthey Technology Centre, Blounts Court Road, Sonning Common, Reading RG4 9NH, United Kingdom

* sandra.vanaert@uantwerpen.be

Symbol	Explanation
T	Number of frames in the time series
N	Number of atomic columns in each frame
G	Maximum number of atoms in a column during the time series
\mathbf{H}	$T \times N \times (G + 1)$ tensor with hidden states: $\mathbf{H} = \{\mathbf{h}_t\}$
\mathbf{h}_t	$N \times (G + 1)$ binary state matrix at time t : $\mathbf{h}_t = \{h_{tg}^{(n)}\}$
$h_{tg}^{(n)}$	Binary state variable: $h_{tg}^{(n)} = 1$ if column n has g atoms at time t
\mathbf{O}	$T \times N$ matrix with observed states: $\mathbf{O} = \{\mathbf{o}_t\}$
\mathbf{o}_t	$1 \times N$ observed vector at time t : $\mathbf{o}_t = \{o_t^{(n)}\}$
$o_t^{(n)}$	Scattering cross section of column n at time t
\mathbf{I}	$(G + 1) \times 1$ initial probability vector: $\mathbf{I} = \{\iota_g\}$
ι_g	Initial probability for a column to have g atoms in frame 1
\mathbf{A}	$(G + 1) \times (G + 1)$ transition matrix: $\mathbf{A} = \{A_{jg}\}$
A_{jg}	Transition probability from j to g atoms between two frames
$\boldsymbol{\mu}$	$G \times 1$ vector with elements μ_g
μ_g	Average scattering cross section corresponding to g atoms in a column
\mathcal{M}_g	Library value for a column with g atoms
a	Scaling parameter relating the average scattering cross section to the library: $\mu_g = a\mathcal{M}_g$
σ	Width of the Gaussian distribution around the average scattering cross section

Table I: Notation overview

In the following, we will write the actual number of atoms in column n at time t as $q_t^{(n)}$. This quantity equals the number of atoms $q_t^{(n)} = g$ for which the previously introduced binary state variable $h_{tg}^{(n)} = 1$. In order to retrieve the realized state sequence, given the observations, we define the following quantity:

$$\delta(t, n, g) = \max_{q_1^{(n)}, \dots, q_{t-1}^{(n)}} \left[P \left(q_1^{(n)} \dots q_t^{(n)} = g, o_1^{(n)} \dots o_t^{(n)} \right) \right]. \quad (9)$$

This expresses the score, i.e. the highest probability, along a single path, at time t for column n , which accounts for the first t observations of that column, and ends with g atoms in column n at time t . In the next frame, $t + 1$, we obtain the score by induction:

$$\delta(t + 1, n, g) = \max_{0 \leq j \leq G} [\delta(t, n, j) A_{jg}] \cdot \mathcal{N} \left(o_{t+1}^{(n)} | \mu_g, \sigma \right). \quad (10)$$

This expression leads to the complete data likelihood from equation (1) at time T :

$$P(\mathbf{H} = \mathbf{Q}, \mathbf{O}) = \prod_{n=1}^N \left(\max_{0 \leq g \leq G} [\delta(T, n, g)] \right), \quad (11)$$

with the hidden state sequence \mathbf{H} equal to the most likely path \mathbf{Q} retrieved by the Viterbi algorithm.

The algorithm calculates the score δ and keeps track of the number of atoms that maximizes this probability in an argument array ϕ . First, these quantities are initialized as follows:

$$\delta(1, n, g) = \iota_g \mathcal{N} \left(o_1^{(n)} | \mu_g, \sigma \right), \quad (12)$$

$$\phi(1, n, g) = 0, \quad (13)$$

with $\delta(1, n, g)$ the score for state g in column n at time t and ϕ the argument array.

The state with the best score for column n at a time t , is used to determine the scores for all possible states of that column n at the next time $t + 1$, expressed by

$$\delta(t + 1, n, g) = \max_{0 \leq j \leq G} [\delta(t, n, j) A_{jg}] \cdot \mathcal{N} \left(o_{t+1}^{(n)} | \mu_g, \sigma \right), \quad (14)$$

$$\phi(t + 1, n, g) = \arg \max_{0 \leq j \leq G} [\delta(t, n, j) A_{jg}], \quad (15)$$

where the number of atoms maximising the score is saved in the argument array.

In the final image of the time series, at time T , the number of atoms that maximizes the score for column n is equal to:

$$q_T^{(n)} = \arg \max_{0 \leq g \leq G} [\delta(T, n, g)]. \quad (16)$$

Finally, by path backtracking, the most likely hidden state sequence is retrieved:

$$q_t^{(n)} = \phi \left(t + 1, n, q_{t+1}^{(n)} \right), \quad (17)$$

yielding the the number of atoms in each column at each time.

2. DETAILS FOR THE SIMULATIONS OF THE HYPOTHETIC ADF STEM SERIES

We simulated scattering cross sections corresponding to hypothetical ADF STEM time series with 40 frames of a changing Pt nanoparticle with 215 atomic columns,

and a thickness up to 15 atoms. The number of atoms in a column was allowed to change by ± 1 from frame to frame throughout the time series, with a probability of 10%. The scattering cross sections of each atomic column are generated from a Pt library, using Poisson random draws in order to replicate the uncertainty from the finite electron dose. Additionally, scan distortion is included in the simulation of the scattering cross sections. Based on a previous study, the effect of scan distortion has been modeled as a normal distribution, with an experimentally determined variance of $4.5 \cdot 10^{-4} \mu_g$, with μ_g the average scattering cross section of an atomic column with g atoms [4]. For each electron dose, 200 different noise realizations of such time series were constructed. Using these data, we compare the atom counts obtained by the hidden Markov model and the collective hybrid

method with the ground truth.

3. PT TIME SERIES

In Fig. S1, we show all frames of the time series of the catalyst Pt nanoparticle analyzed in this paper, after image registration. From this series we can see that small amounts of sample tilt are present in some of the images. Furthermore, we see that the particle remains stable during acquisition. The counting results obtained by the hidden Markov model of the tilt-compensated time series are shown in Fig. S2. The experimental time series and corresponding counting results are also shown in the Supplementary Animations.

-
- [1] L. R. Rabiner, “A tutorial on hidden Markov models and selected applications in speech recognition,” *Proceedings of the IEEE* **77**, 257–286 (1989).
 - [2] A. J. Viterbi, “Error bounds for convolutional codes and an asymptotically optimum decoding algorithm,” *IEEE Transactions on Information Theory* **13**, 260–269 (1967).
 - [3] G. D. Forney, “The Viterbi algorithm,” *IEEE Transactions on Information Theory* **61**, 268–278 (1973).
 - [4] S. Van Aert, A. De Backer, L. Jones, G. T. Martinez, A. Béché, and P. D. Nellist, “Control of Knock-On Damage for 3D Atomic Scale Quantification of Nanostructures: Making Every Electron Count in Scanning Transmission Electron Microscopy,” *Physical Review Letters* **122**, 066101 (2019).

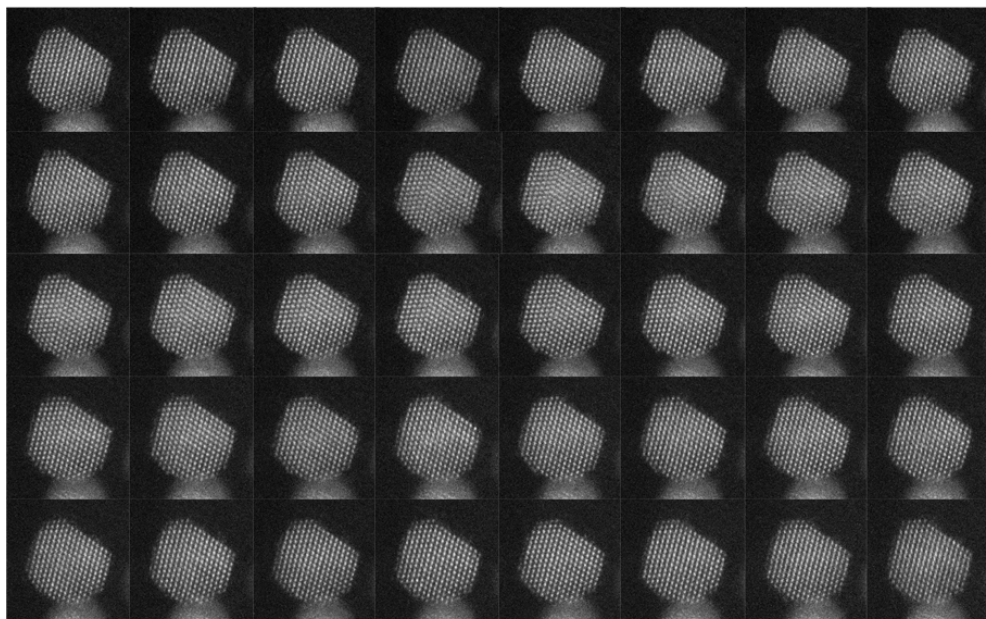


Figure S1: Experimental ADF STEM time series recorded of a Pt nanoparticle. Time progresses along the rows.

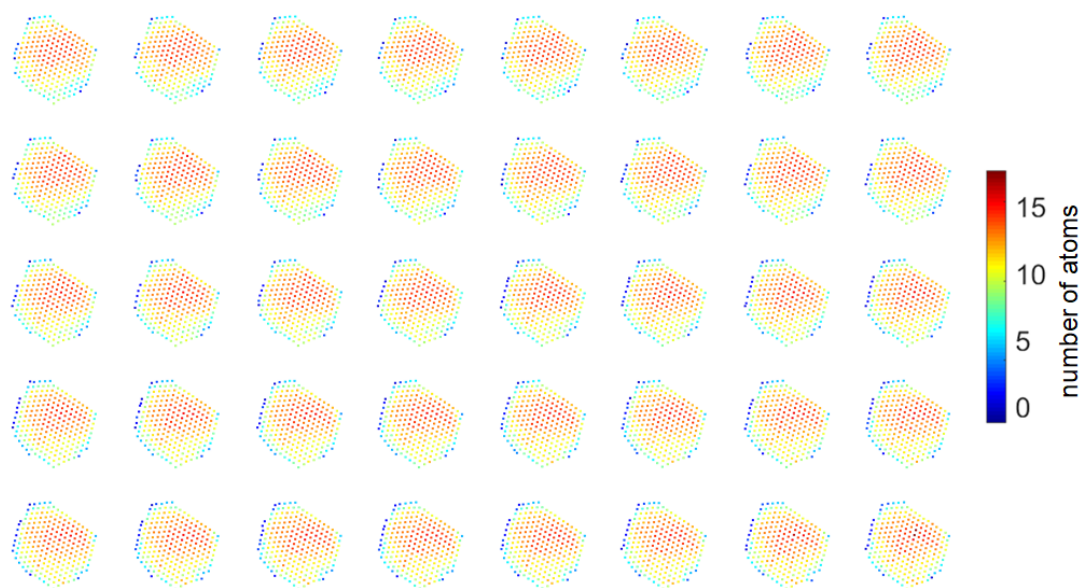


Figure S2: Counting results obtained by the hidden Markov model for the time series from Fig. S1.