Supplementary Information

Experimental reconstructions of 3D atomic structures from electron microscopy images using a Bayesian genetic algorithm

Annick De Backer[1,2], Sandra Van Aert[1,2,*], Christel Faes[3], Ece Arslan Irmak[1,2], Peter D. Nellist[4], and Lewys Jones[5,6,*]

[1]EMAT, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerp, Belgium
[2]NANOlab Center of Excellence, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerp, Belgium
[3]I-BioStat, Data Science Institute, Hasselt University, Hasselt, Belgium
[4]Department of Materials, University of Oxford, Parks Road, OX1 3PH Oxford, United Kingdom
[5]Advanced Microscopy Laboratory, Centre for Research on Adaptive Nanostructures and Nanodevices (CRANN), Dublin 2, Ireland
[6]School of Physics, Trinity College Dublin, The University of Dublin, Dublin 2, Ireland
[*]*E-mail: sandra.vanaert@uantwerpen.be, lewys.jones@tcd.ie*

Supplementary Tables

Parameter	Value
Acceleration voltage	200 kV
Defocus C_1	0 nm
Spherical aberration C_s	0 nm
Convergence angle α	22.48 mrad
Inner detector angle ADF	52 mrad
Outer detector angle ADF	248 mrad
Inner detector angle ABF	9 mrad
Outer detector angle ABF	21 mrad
FWHM of the source image	1.0 Å
Debye-Waller factor	0.384 Å^2
Pixel size	0.124 Å
Zone axis	[110]
Number of phonon configurations	30

Supplementary Table 1 – **Simulation settings.** Parameters for the frozen lattice simulation of a the Pt nanoparticle containing 587 atoms and Pt crystal in [110] zone axis up to 30 atoms thickness.

	Unrelaxed (cost-function - a.u.)	Geometry relaxation (cost-function - a.u.)
Dose 10^5 e$^-$Å^{-2}		
Model 1	-10.1418	-10.2374
Model 2	-10.0249	-10.1209
Model 3	-8.3578	-8.4502
Model 4	-6.7461	-6.8097
Model 5	-6.7388	-6.8089
Dose 10^2 e$^-$Å^{-2}		
Model 1	-7.2053	-7.2982
Model 2	-7.2051	-7.2967
Model 3	-7.1141	-7.2169
Model 4	-5.9676	-6.0438
Model 5	-5.9503	-6.0409

Supplementary Table 2 – Cost-function for models without and with geometry relaxion [1]. Five models are selected, with a very similar and a more different value for the cost-function. The models are taken from the final population of 4th best noise realization for the highest (10^5 e$^-$Å^{-2}) and lowest (10^2 e$^-$Å^{-2}) electron dose considered in this study.

Supplementary Methods

The estimated width σ_{GMM} should in principle take into account all dose-dependent and dose-independent uncertainties in the SCSs, including the limited electron dose and fluctuations in the SCSs from other effects such as the carbon support, different vertical onset of columns of the same number of atoms, intensity transfer between columns, relaxed atom positions, and the influence of neighboring columns of different number of atoms [2]. At low doses σ_{GMM} is underestimated [3]. This can be easily observed by comparing the estimated width σ_{GMM} with the expected width from the dose-dependent uncertainty. The width σ_{dose} can be predicted by [4]:

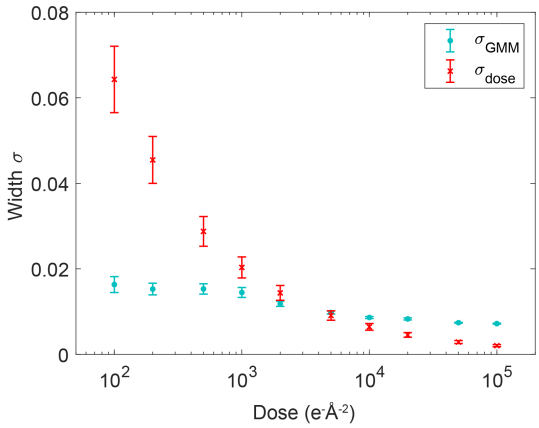
$$\sigma_{\mathrm{dose}} = \sqrt{\mu_g / d}, \tag{1}$$

where μ_g corresponds to the expected SCS for a column containing g atoms and d equals the incident electron dose. σ_{GMM}, based on the 30 noise realizations at each incident electron dose of the extensive simulation study presented in the main paper, and σ_{dose}, based on Equation (1), are shown in Figure 1 to illustrate the underestimation at lower incident electron doses. In order to have a better estimation of the finite atom-counting precision at lower incident doses, an effective width σ_{eff} is therefore introduced which is the maximum of the two contributions:

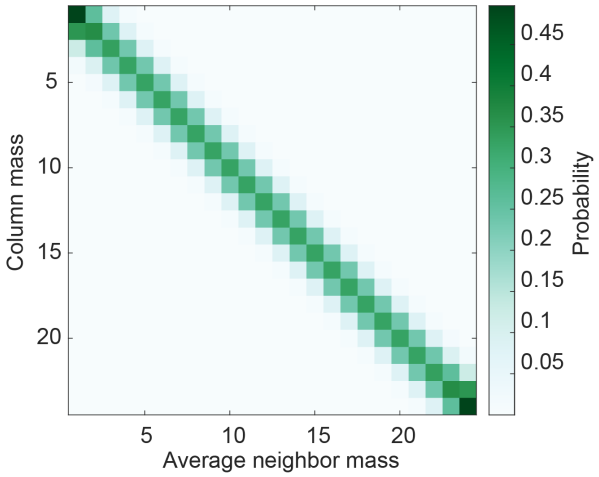
$$\sigma_{\mathrm{eff}} = max(\sigma_{\mathrm{dose}}, \sigma_{\mathrm{GMM}}). \tag{2}$$

This σ_{eff} is a very good approximation for the true width of the underlying Gaussian mixture model since σ_{dose} will be dominant at doses lower than $2 \times 10^3 e^- \text{Å}^{-2}$ as demonstrated by Van Aert *et al.* [4]. From Figure 1, it also clear that beyond this incident electron dose, the dose-independent width becomes more dominant.

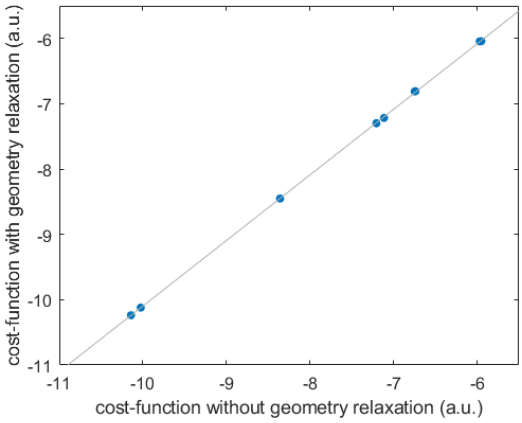
Supplementary Figures



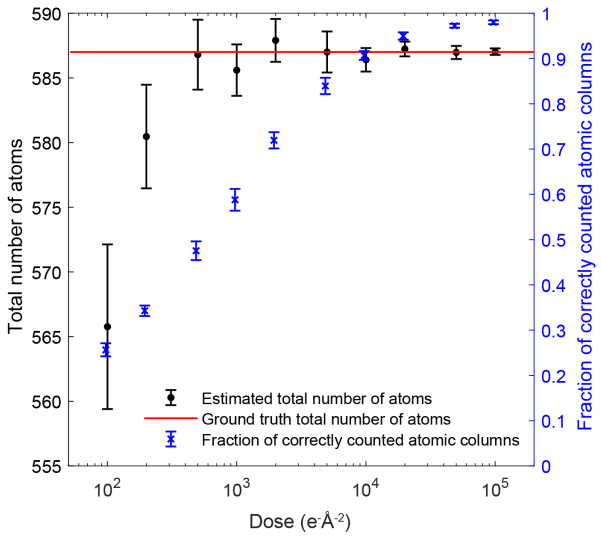
Supplementary Figure 1 – Illustration of the width of the Gaussian components. At lower incident electron doses the width σ_{GMM} is underestimated in comparison with the dose-dependent width σ_{dose}. The error bars indicate 95% confidence intervals. The 95% confidence intervals for the σ_{GMM} are computed using 30 noise realizations. For σ_{dose} the 95% confidence intervals indicate the spread due to thickness dependence of σ_{dose}.



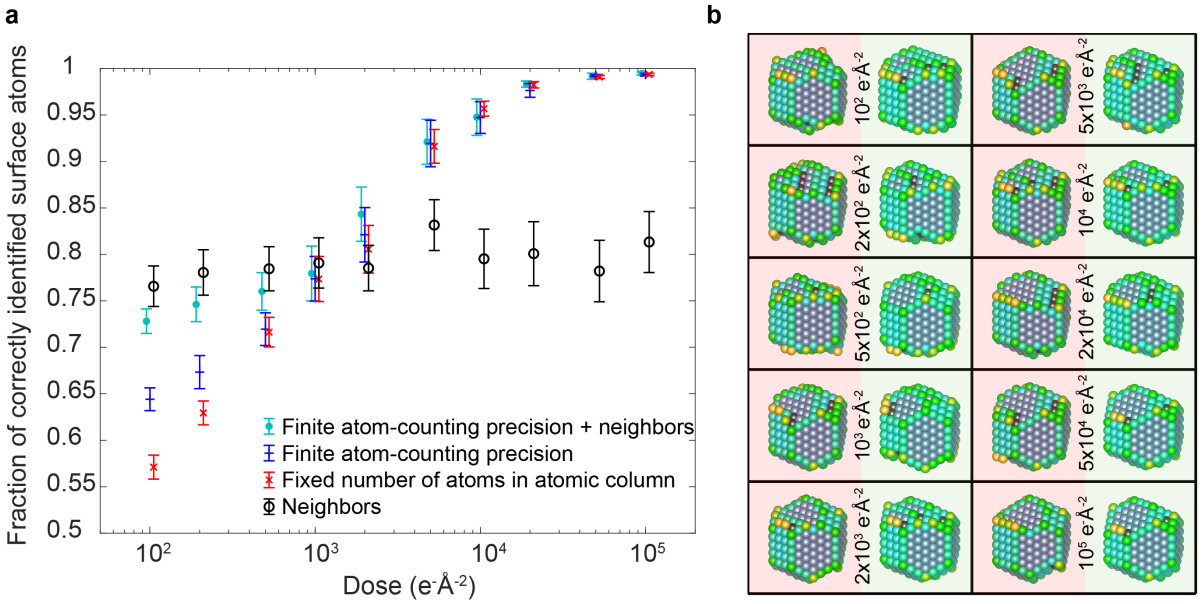
Supplementary Figure 2 – **Neighbor-mass probability matrix.** $p(g|\mathrm{NB}_n)$ The average mass of the neighboring columns is shown on the x-axis and the column mass on the y-axis. The colormap illustrates the probabilities.



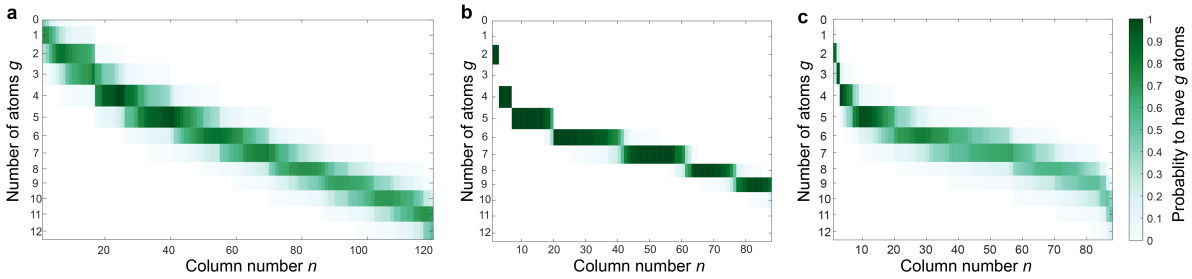
Supplementary Figure 3 – **Effect of geometry relaxation.** Cost-function values for the models with geometry relaxation as a function of the cost-function values for the unrelaxed models from Supplementary Table 2. The gray line is shown to highlight the linear relationship between the values of both approaches.



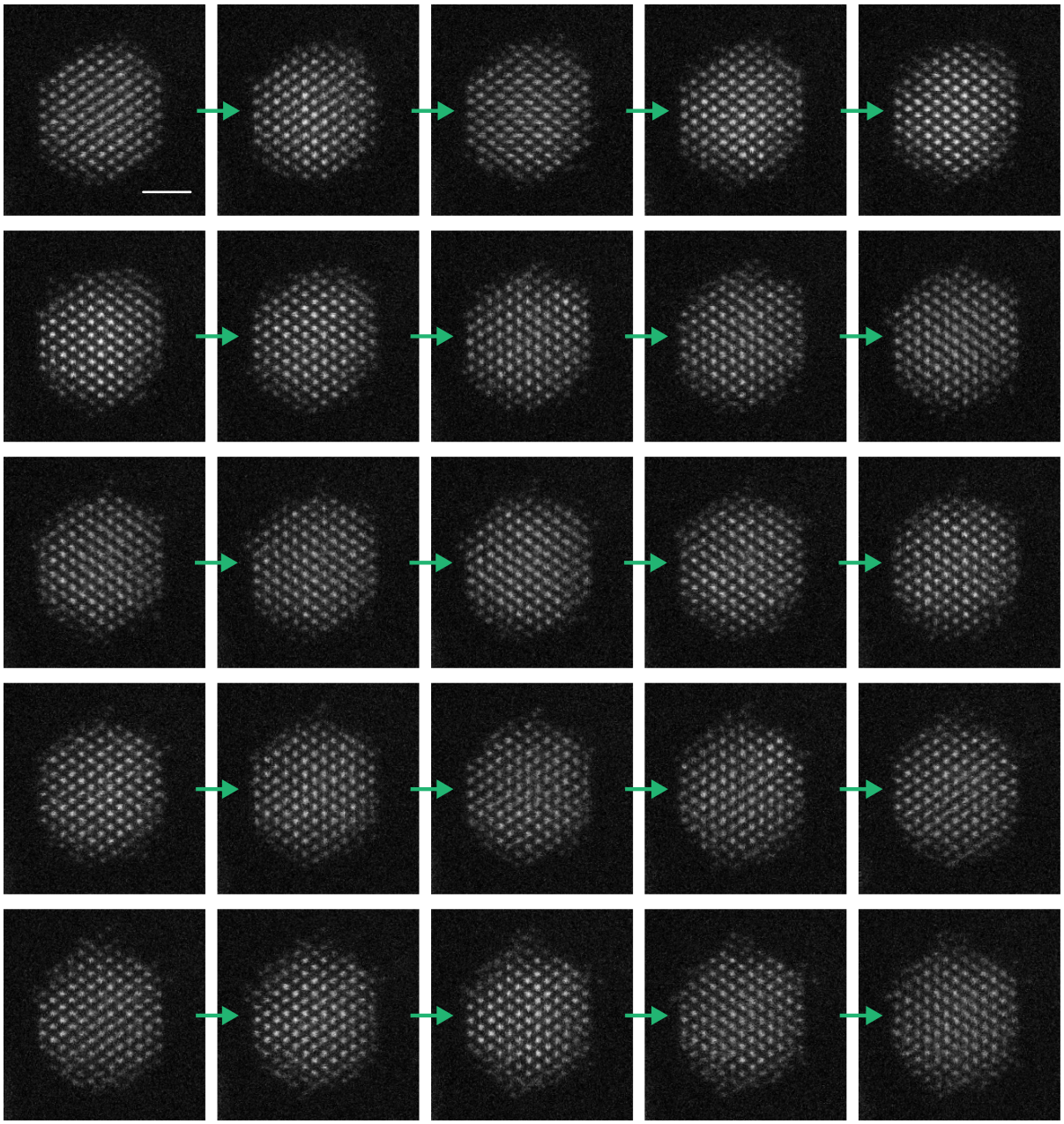
Supplementary Figure 4 – **Accuracy of the atom-counting results.** Estimated total number of atoms in the particle (left axis) and the fraction of atomic columns in which the number of atoms has been counted correctly (right blue axis). For the lowest doses, the total number of atoms is slightly underestimated and as expected the fraction of correctly counted atomic columns decreases when the incident electron dose decreases. These atom-counting results are used as an input for our Bayesian genetic algorithm. The error bars indicate 95% confidence intervals.



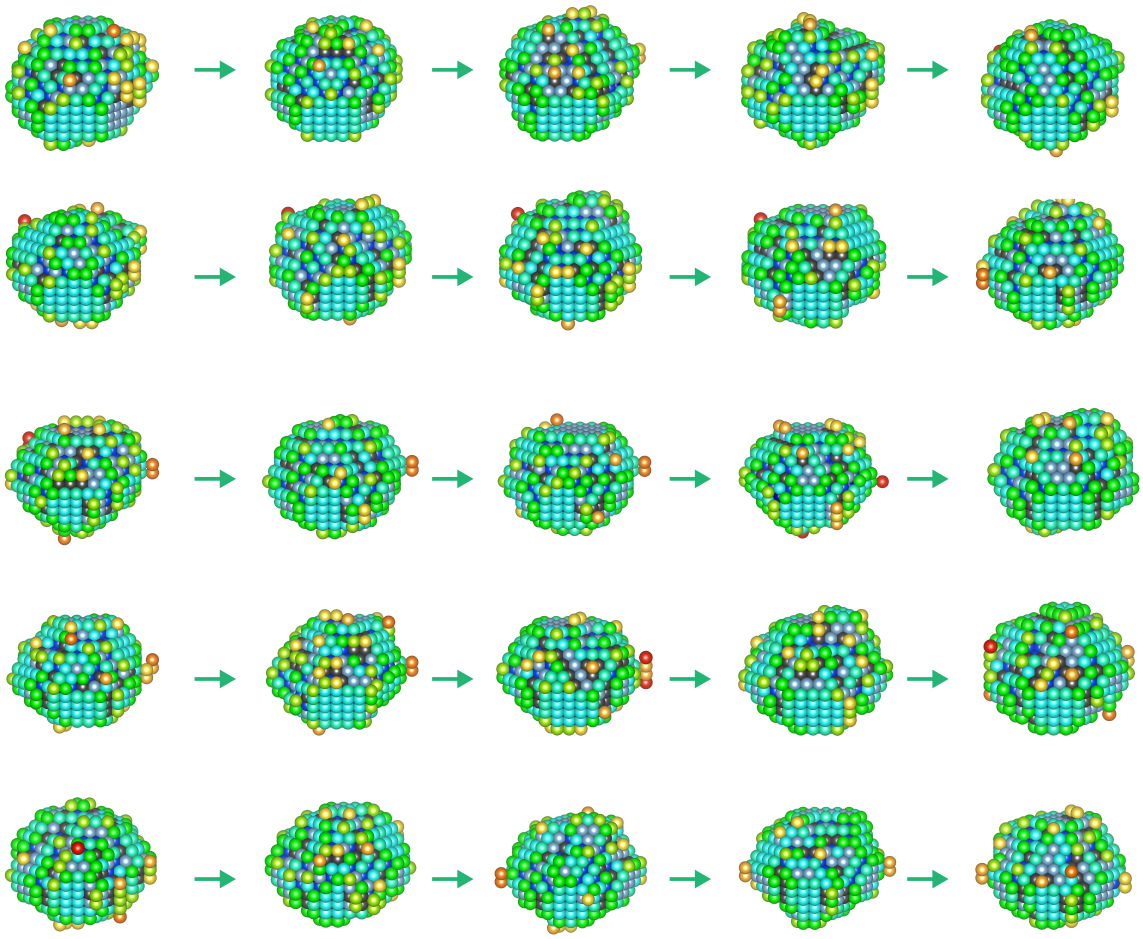
Supplementary Figure 5 – **Results for the reconstructions using the neighbor-mass relations only. a** Fraction of the surface atoms that are correctly defined in 3D when including both the finite atom-counting precision and neighbor-mass relations (light blue), the finite atom-counting precision only (dark blue), and the neighbor-mass relations only (black) as prior knowledge. As a reference the results when using a fixed number of atoms in a column are also displayed. The error bars indicate 95% confidence intervals. **b** Visualization of the reconstructed 3D atomic models represented by the lower bound (red background) and upper bound (green background) of a 80% prediction interval for the reconstructions when using the neighbor-mass relations only.



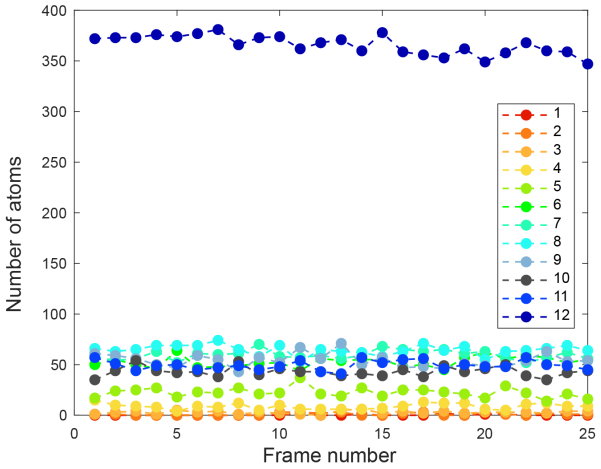
Supplementary Figure 6 – Comparison of probability matrices for finite atom-counting precision. Probability matrix for **a** the 4th frame of the experimental time series, **b** for a dose of 10^4 e$^-$Å^{-2} including only Poisson noise, and **c** for a dose of 5×10^2 e$^-$Å^{-2} including only Poisson noise.



Supplementary Figure 7 – Experimental ADF STEM time-series of the Pt nanoparticle. Time progresses along the rows. The scale bar is 10 Å.



Supplementary Figure 8 – **3D atomic models of the experimental Pt nanoparticle.** Reconstructions obtained by the Bayesian genetic algorithm including the finite atom-counting precision and neighbor-mass relations as prior knowledge for the time series from Supplementary Figure 7. The coloring of the atoms corresponds to the nearest-neighbor coordination, from 1 in red to 12 in dark blue.



Supplementary Figure 9 – **Atomic coordination number analysis.** The number of atoms with the a specific coordination number are shown for the time series of the Pt nanoparticle. The coloring of the data corresponds to the nearest-neighbor coordination, from 1 in red to 12 in dark blue.

References

[1] LAMMPS Molecular Dynamics Simulator. `https://lammps.org`. Accessed November 2019.

[2] De Backer, A., Martinez, G. T., Rosenauer, A. & Van Aert, S. Atom Counting in HAADF STEM Using a Statistical Model-Based Approach: Methodology, Possibilities, and Inherent Limitations. *Ultramicroscopy* **134**, 23–33 (2013).

[3] De wael, A., De Backer, A., Jones, L., Nellist, P. D. & Van Aert, S. Hybrid Statistics-Simulations Based Method for Atom-Counting from ADF STEM Images. *Ultramicroscopy* **177**, 69–77 (2017).

[4] Van Aert, S. *et al.* Control of Knock-On Damage for 3D Atomic Scale Quantification of Nanostructures: Making Every Electron Count in Scanning Transmission Electron Microscopy. *Phys. Rev. Lett.* **122**, 066101 (2019).