

ARTICLE OPEN



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 Faes³, Ece Arslan Irmak^{1,2}, Peter D. Nellist⁴ and Lewys Jones^{5,6}

We introduce a Bayesian genetic algorithm for reconstructing atomic models of monotype crystalline nanoparticles from a single projection using Z-contrast imaging. The number of atoms in a projected atomic column obtained from annular dark field scanning transmission electron microscopy images serves as an input for the initial three-dimensional model. The algorithm minimizes the energy of the structure while utilizing a priori information about the finite precision of the atom-counting results and neighbor-mass relations. The results show promising prospects for obtaining reliable reconstructions of beam-sensitive nanoparticles during dynamical processes from images acquired with sufficiently low incident electron doses.

npj Computational Materials (2022)8:216; <https://doi.org/10.1038/s41524-022-00900-w>

INTRODUCTION

It is commonly accepted that the three-dimensional (3D) atomic structure of metallic nanoparticles determines their catalytic properties^{1–5}. Indeed, the presence of highly undercoordinated atoms or stepped facets at the surface governs many catalytic reactions. A quantitative characterization of the atomic configuration at the surface is therefore essential to reveal the active sites of the nanoparticle where reactant molecules are preferentially adsorbed, to understand the mechanisms of the catalytic behavior, and to improve the performance of these systems. Atomic-resolution annular dark field scanning transmission electron microscopy (ADF STEM) has become an invaluable tool for imaging metallic nanostructures^{6–10}. In this context, electron tomography has been used to provide insights into the 3D shape of nanostructures^{11–13}, but this technique requires a significant electron dose for the multiple projections. Consequently, this approach is not feasible when investigating small beam-sensitive catalysts or dynamical processes. Therefore, alternative methods have been developed where the 3D atomic structure is reconstructed from a single ADF STEM projection^{8,14–17}. For this purpose, the number of atoms contained in an atomic column along the third dimension is retrieved from an ADF STEM image. These atom counts are combined with prior knowledge about a material's crystal structure and used to create an initial atomic model. This serves as an input for energy minimization to obtain a relaxed 3D reconstruction of a nanostructure. The validity of this atom-counting/energy minimization method has qualitatively been verified using electron tomography^{10,16} and is applied to study several systems^{5,8,10,16,18}.

Nowadays, different possible approaches are available for energy minimization that incorporates experimental observations. Indeed, the aim is not to find the global energy minimum from a purely computational study but to reconstruct also the metastable configurations in which the nanoparticles often reside. Yu et al. refined the atomic structure by minimizing the particle energy together with matching forward linear modeling and the experimental ADF STEM images¹⁵. In this study, an approximate model for the STEM image simulation was required limiting the

accuracy of the reconstructed models. Other methodologies use atom-counting results as an input, and mainly, two different strategies can be distinguished here. Using the first approach, the energy is minimized by shifting the atomic columns up and down while keeping the number of atoms in a column fixed to the outcome of the atom-counting procedure^{8,18}. The second approach consists of a full molecular dynamics simulation to relax the particle's structure^{5,10,16}. The first method is potentially too restrictive by ignoring the finite atom-counting precision, especially at lower doses. The expected inevitable counting imprecision in this case⁹, will likely result in slightly more roughness at the reconstructed atomic surface in the direction parallel to the beam direction^{8,16}. On the other hand, the second method runs the risk of ending up in a global energy minimum and violating the physical constraints of the experimental observation. Both approaches hamper a reliable 3D reconstruction of the atomic configuration at the surface, especially for images acquired at lower doses.

Here, we propose a Bayesian genetic algorithm that includes the finite atom-counting precision via a Bayesian inference scheme to improve the 3D atomic models for small nanostructures. Bayesian methods are powerful tools in which a priori information is rationally combined with the observed data and have been successful in other fields of science including business, computer science, economics, educational research, environmental science, epidemiology, genetics, geography, imaging, law, medicine, political science, psychometrics, public policy, sociology, and sports^{19–22}. Next to the finite atom-counting precision, the incorporation of additional prior knowledge from neighbor-mass relations will be beneficial when reconstructing atomic models from extremely low-dose ADF STEM images. This prior knowledge is fused into the genetic algorithm which uses atom-counting results as an input for reconstructing the 3D atomic structure. Genetic algorithms are typically used as efficient tools for solving large optimization problems where finding a direct solution is not possible. These evolutionary methods are widely applied for structure prediction of systems with increasing complexity, such as atomic clusters,

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

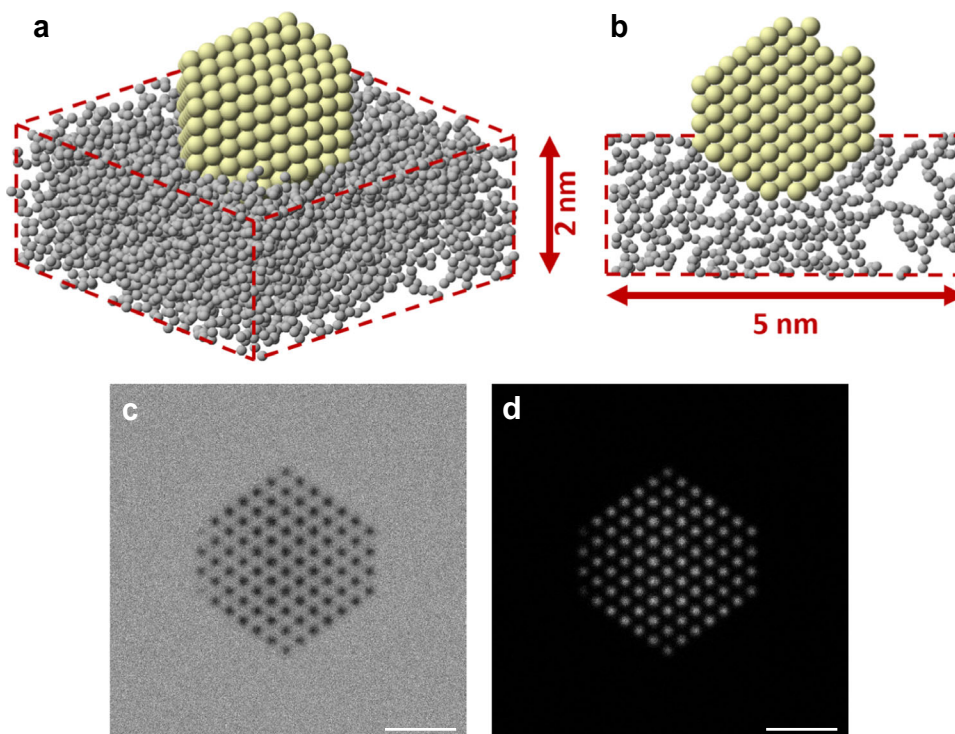


Fig. 1 Simulated Pt nanoparticle partially embedded in an amorphous carbon support. **a** 3D model and **b** cross-section, **c** simulated ABF STEM image, and **d** simulated ADF STEM image at $10^4 e^- \text{Å}^{-2}$. The scale bars in **c–d** are 1 nm.

crystals, interfaces, surfaces, thin films, nanoparticles, and defect structures^{15,21,23–29}.

In this study, we focus on the 3D characterization of small, single-crystal, mono-metallic catalyst nanoparticles. The technologically relevant size of the catalytic particles considered here, for which the catalytic activity is peaking³⁰, is around 3 nm in size, but the method will be applicable to structures up to 10 nm in diameter. The method will also be suitable for the different types of single-element crystalline nanostructures for which the model-based atom-counting method is designed^{31,32}. This covers atomic-resolution ADF STEM images of nanostructures with a varying thickness, acquired along a main zone axis, which are defect-free along the beam direction. For more complex shapes, e.g., concave surfaces, it will depend on the direction of the surface with respect to the incident electron beam whether the specific feature can be reconstructed reliably from the atom-counting results of a single projection. In this work, the quality of the obtained reconstructions with our advanced Bayesian genetic algorithm is quantitatively evaluated via an extensive simulation study, in terms of the reliability with which the surface atoms can be reconstructed in 3D. In the last part, the algorithm is applied to retrieve 3D atomic models from an experimental time series of a Pt nanoparticle.

RESULTS AND DISCUSSION

Pt nanoparticle model and atom-counting

To count the number of atoms, a very well-established approach is utilized^{8,11,16,31,32}. In the past years, the high accuracy and precision of this method have been demonstrated for nanocrystals of arbitrary shape, size, and atom type. For atom-counting, so-called scattering cross-sections (SCSs), corresponding to the total intensity of electrons scattered toward the ADF detector for every atomic column, have been introduced in ADF STEM. These SCSs can be measured using statistical parameter estimation theory^{33,34} or by integrating intensities over the probe positions in the vicinity of a single column of atoms³⁵. For our simulation study, SCSs are

determined from noise realizations at different doses of a simulated ADF STEM image of a Pt nanoparticle partially embedded in a carbon support, illustrated in Fig. 1a, b. The created particle largely resembles the Wulff construction solid for Pt₅₈₆ (four atoms along each edge) and was modified to include several diagnostic features of interest commonly observed for catalytic metallic nanoparticles including a surface adatom, a surface vacancy, a terrace edge, a small area of {110} face, and rounded corners. With these modifications, the particle model contains 587 atoms. The slab of amorphous carbon measures 5 nm × 5 nm × 2 nm and the geometry follows the work of reference³⁶. Image simulations were performed for the Pt particle viewed along the [110] zone axis. The details of the simulations are described in the Methods section. Figure 1c shows the annular bright field (ABF) STEM image, illustrating the presence of the carbon support and Fig. 1d shows the ADF STEM image using an electron dose of $10^4 e^- \text{Å}^{-2}$. In Fig. 2a, b simulated ADF STEM images are shown using lower incident electron doses corresponding to 10^3 and $10^2 e^- \text{Å}^{-2}$, respectively. The simulated ADF STEM images can be modeled as a superposition of Gaussian functions using the StatSTEM software¹⁶. The refined models for the simulated ADF STEM images are shown in Fig. 2c, d. From the estimated model parameters, the SCSs are determined for each atomic column and are represented in a histogram in Fig. 2g, h. In a subsequent analysis, the distribution of the SCSs of all atomic columns is decomposed into overlapping normal distributions, i.e., a Gaussian mixture model, as illustrated in Fig. 2g, h^{31,32}. The locations of the normal distributions are matched to the expected SCS values from simulations for a column containing g atoms³⁷ and their widths are determined. It should be noted that this procedure only depends on the values of the estimated SCSs and is independent of the subjective choice of bins in the histogram for visualization. The number of atoms in each projected atomic column in Fig. 2e, f is then obtained by assigning the SCS to the component that generates this SCS value with the highest probability. More details on the atom-counting methodology are

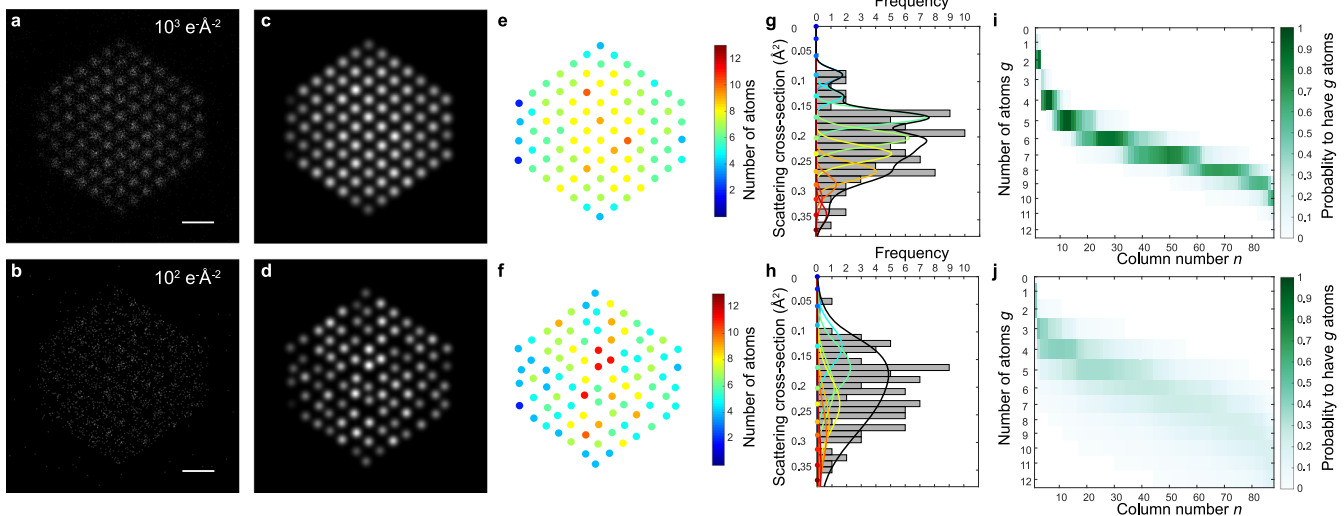


Fig. 2 Statistical atom-counting analysis and construction of the probability matrix. **a, b** Simulated ADF STEM images of the Pt nanoparticle embedded in the carbon support along the [110] zone axis for an incident electron dose of 10^3 and $10^2 \text{ e}^- \text{Å}^{-2}$. **c, d** Refined parametric models of the images shown in **a, b**. **e, f** Estimated number of atoms in each atomic column for the images shown in **a, b**. **g, h** Histograms of SCSs for the images shown in **a, b**. The decomposition into overlapping Gaussian components is shown in color, corresponding to the number of atoms. The black curve shows the estimated mixture model. **i, j** Probability matrices indicating the probability for a column n to contain a specific number of atoms g . The scale bars in **a** and **b** are 5 Å .

provided in the Methods section. The width of the normal distributions reflects the finite precision of the atom-counting results and will be used as prior knowledge for reconstructing the 3D atomic structure.

Bayesian inference: finite atom-counting precision and neighbor-mass relations

As an input we need the probability that an atomic column contains a specific number of atoms which can be defined from the decomposition into normal distributions, illustrated in Fig. 2g, h. Indeed, the normal distribution functions describe the probability that component g generates the n th SCS, i.e., $p(\text{SCS}_n|g)$, and by using Bayes' theorem, the probability that the n th SCS has g atoms, i.e., $p(g|\text{SCS}_n)$, can be computed:

$$p(g|\text{SCS}_n) = \frac{p(g)p(\text{SCS}_n|g)}{p(\text{SCS}_n)} = \frac{p(g)p(\text{SCS}_n|g)}{\sum_g p(g)p(\text{SCS}_n|g)}. \quad (1)$$

Equal probabilities are assigned to the probability of having g atoms in a column $p(g)$. The value of $p(g)$ depends on the maximum number of g that is considered. However, for the computation of the probability $p(g|\text{SCS}_n)$, it is irrelevant since a constant $p(g)$ cancels in Eq. (1). The probability $p(g|\text{SCS}_n)$ is visualized by a probability matrix in Fig. 2i, j. From the representation of the Gaussian mixture model on top of the histogram, the relation between the probability matrix and the width of the normal distributions is clear.

For improving the quality of the reconstructions further at lower doses, we can include even more relevant prior knowledge in the form of neighbor-mass relations. A neighbor-mass matrix helps to predict the column mass based on the average mass of the neighboring columns. For the small spherical convex nanoparticles in this study, abrupt discontinuities are highly non-physical³⁰. Therefore, we propose here a diagonal neighbor-mass matrix. The matrix is visualized in Supplementary Fig. S2. The probability profile is Gaussian and the width is chosen such that the interval ± 1 atom contains 80% of the probability. One should keep in mind that for particles with significantly different surfaces, the neighbor-mass matrix should be designed differently. The normalized neighbor-mass probability, $p(g|\text{NB}_n)$ with NB_n indicating the average neighbor-mass, is combined with the probability matrix

accounting for the atom-counting reliability $p(g|\text{SCS}_n)$, in order to take the two types of prior knowledge into account:

$$p(g|\text{SCS}_n \cap \text{NB}_n) = \frac{p(g|\text{SCS}_n)p(g|\text{NB}_n)}{\sum_g p(g|\text{SCS}_n)p(g|\text{NB}_n)}. \quad (2)$$

Bayesian genetic algorithm for 3D atomic modeling

The introduced probability matrices are used as input of the prior knowledge for the genetic algorithm that we will use to reconstruct the 3D atomic structure of the nanoparticle, hence the name Bayesian genetic algorithm. A genetic algorithm is an iterative process where first a population of randomly generated individuals is created. In our algorithm, this initial population is generated by randomly modifying the number of atoms and the height offset of the atomic columns of a 3D starting configuration within a certain mutation range. This 3D starting configuration is obtained by positioning the atoms (i.e., the outcome of the atom-counting procedure) in each atomic column parallel to the beam direction and symmetrically around a central plane. Based on prior knowledge about the crystal structure, the atoms are separated at a fixed distance $a/\sqrt{2}$ with $a = 3.92 \text{ Å}$ corresponding to the [110] zone orientation. For this orientation, the atomic columns at the even planes are shifted along the beam direction by $a/(2\sqrt{2})$ with respect to the odd planes. A population size of 500 is used in all calculations, the count mutation range equals 1, and the height mutation range for the offset of a column equals a lattice step, i.e., the interatomic spacing between the Pt atoms along the [110] direction. In each iteration, i.e., a generation, the fitness of every individual in the population is evaluated by the cost-function of the optimization problem. The individuals with the best cost-function values are selected from the current population, and a new complete population of candidate solutions is formed by recombining and mutating the selected individuals. The fraction of the population that is used for the recombination step equals 50%. For each recombined member, two parents are randomly chosen from the selected individuals, and cross-over is performed by randomly selecting columns from both parents. A mutation density of 2% is included to avoid ending up immediately in a local minimum by randomly modifying the number of atoms and

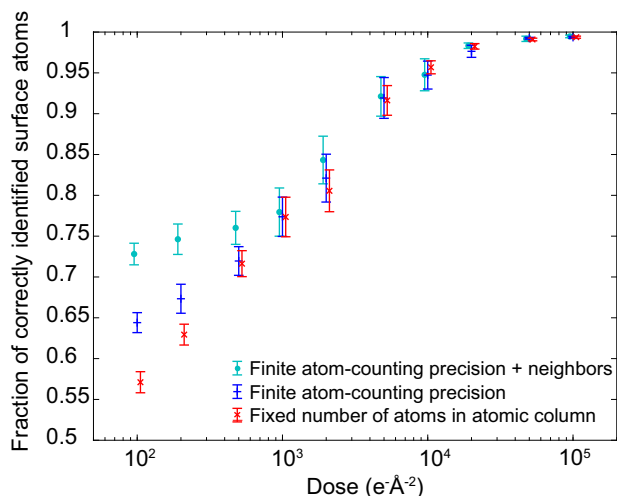


Fig. 3 Quantification evaluation of the reconstructed 3D models. Fraction of the surface atoms that are correctly reconstructed in 3D when including the finite atom-counting precision and neighbor-mass relations as prior knowledge. As a reference, the results when using a fixed number of atoms in a column are also displayed. The error bars indicate 95% confidence intervals.

height offset for 2% of the atomic columns in each new member. In addition to the usual iterations over many breeding generations, in this work we introduce a second loop to provide for multiple unique starting initializations, specifically to reduce the risk of finding only local-minima solutions. It should be noted that in the population of candidate solutions the atoms are not relaxed in x -, y -, and z -direction in order to search a broad population with fast convergence. A geometry relaxation, performed using LAMMPS³⁸, would not affect the ranking of the models and thus not influence the reliability of the reconstructed models as illustrated in Supplementary Table S2. The geometry relaxation only reduces the cost-function by a consistent amount as shown in Supplementary Fig. S3. More details of the genetic algorithm are provided in the Methods section. The cost-function χ that we use to evaluate the candidate solutions within the different generations of the genetic algorithm is given by:

$$\chi = \frac{\sum E_a}{\sum_n g_n} \cdot \left(1 + \sqrt[n]{\prod_n p(g_n | \text{prior knowledge})} \right), \quad (3)$$

where $\sum E_a$ is the sum of the energies per atom given by the EAM potential^{39,40} and $\sum_n g_n$ is the total particle mass. This cost-function consists of two factors where the first represents the average energy per atom that we wish to minimize. The second factor allows us to penalize the energy per atom with the probability of the candidate solution based on the prior knowledge (Eq. (1) or Eq. (2)). This form of the cost-function weights the penalty term with the average energy per atom. The model probability itself is based on the geometric mean of the probabilities of each individual column and needs to be maximized. Since the average energy per atom is negative, this cost-function is minimized.

Simulation results

In order to evaluate the quality of the reconstructions using the Bayesian genetic algorithm, an extensive simulation study is carried out. For this purpose, the electron dose is varied between 10^2 and 10^5 e-Å⁻² and 30 noise realizations are generated at each electron dose. Supplementary Fig. S4 summarizes the atom-counting results obtained following the methodology illustrated in Fig. 2 and which are used as an input for our Bayesian genetic algorithm. The results of the reconstructions are quantitatively summarized by comparing the reconstructed 3D models with the

ground truth model. As a criterion to evaluate the 3D atomic models, we used the fraction of surface atom positions, i.e., with coordination number less than 12, that are correctly defined in 3D. These are the atoms that are of interest for catalysis. Figure 3 shows this fraction for the reconstructed atomic models. As a reference, we also included the fraction following the approach where the number of atoms is fixed to the outcome of the atom-counting procedure and where during the reconstruction the atomic columns are only shifted up and down. A significant, targeted improvement for the reconstructed surface atoms is observed for the lower incident electron doses when including the finite atom-counting precision and the neighbor-mass relations. For the lowest incident electron dose of 10^2 e-Å⁻², we see an improvement from 57% to 73%. It is important to notice that only Poisson noise is included in the simulation study. In experiments, environmental distortions will introduce additional uncertainties in the atom-counting results⁴¹. Therefore, we expect that the Bayesian genetic algorithm will also improve the reliability of the reconstructed 3D models at higher incident electron doses for experiments. If the probability matrix of the finite atom-counting precision (Fig. 2i, j) shows very high probabilities for all the atomic columns, i.e., when there is no ambiguity in the assignment of the number of atoms, the less computational demanding method, where the number of atoms in the atomic columns is fixed during the reconstruction, might be preferred as it will result in a reliable 3D atomic model in a more fast manner. In Supplementary Fig. S5a, the results obtained when using the neighbor-mass relations only are also included, next to the results of Fig. 3. The quality of the neighbors-only reconstructions is dose-independent and significantly lower at the higher incident dose values. In this case, there are no constraints set by the experimental observations and the reconstruction is more determined by a lower energy solution. The high performance at low doses for the neighbors-only reconstructions is partially owing to the choice of a relatively well-faceted, low-energy particle. These results illustrate that prior knowledge about the finite atom-counting precision is essential in the Bayesian genetic algorithm and that it is the combination of both types of prior knowledge, that leads to improved performance as illustrated in Fig. 3. Furthermore, the results in this figure demonstrate that the different contributions in the cost-function, i.e., the average energy per atom, the finite atom-counting precision, and the neighbor-mass relations, are properly balanced and that the reconstructed 3D model is not fully determined by one of the contributions.

In order to evaluate the reconstructed 3D atomic models in a bit more detail, the 4th worst and 4th best reconstructions of the 30 noise realizations at each dose can be visualized as an 80% prediction interval for the reconstructions. The 4th worst and 4th best reconstructions are selected based on the percentage of correctly reconstructed surface atoms. These intervals are shown in Fig. 4 and Supplementary Fig. S5b. The colors of the atoms correspond to the coordination number and the reconstruction in the box in Fig. 4c corresponds to the ground truth model. The coordination number serves as a powerful predictor for surface adsorption strength on Pt nanoparticles, and hence as a predictor of chemical activity^{3,42–44}. Even for the lower doses in Fig. 4c, the shape is very well reconstructed and a vast improvement is observed when including relevant prior knowledge resulting in less roughness from the finite atom-counting precision at the surface as compared to the low-dose reconstructions in Fig. 4a, where the number of atoms in an atomic column has been kept fixed. Some features will be more easily reconstructed unambiguously than others. Surface facets are typically well reconstructed. As expected, smaller details are more challenging. For example, an adatom will be easier to detect if it is located at the side of the nanoparticle in projection. However, in this case, it is more difficult to locate this atom correctly along the depth direction. On the other hand, an adatom on top of the nanoparticle, will be more

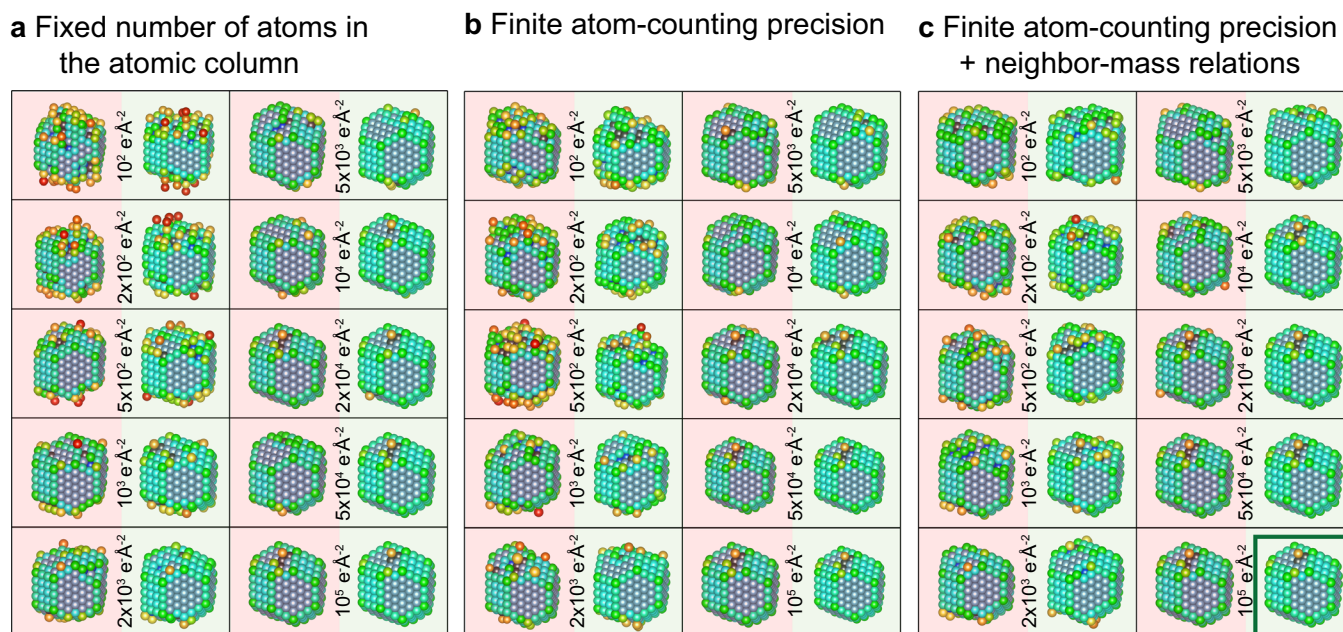


Fig. 4 Visualization of the reconstructed 3D atomic models. A lower bound (red background) and upper bound (green background) of an 80% prediction interval are represented for the reconstructions when using **a** a fixed number of atoms in the atomic column during the reconstruction, **b** the finite atom-counting precision, and **c** the finite atom-counting precision and the neighbor-mass relations. The reconstruction in the box in the lower right corner corresponds to the ground truth model. The coloring of the atoms corresponds to the nearest-neighbor coordination, from 1 in red to 12 in dark blue.

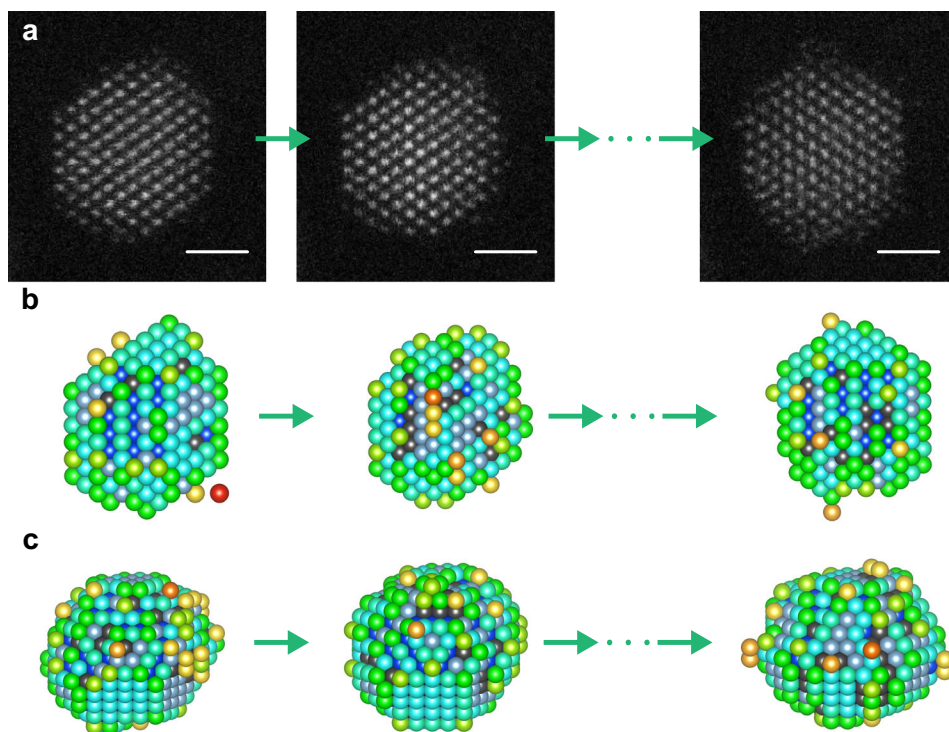


Fig. 5 Analysis of an experimental time series of a catalyst Pt nanoparticle. **a** ADF STEM time series. **b** Corresponding reconstructed 3D atomic models for the time sequence viewed along the beam direction. **c** Rotated models to show the dominant surface facets. The coloring of the atoms corresponds to the nearest-neighbor coordination, from 1 in red to 12 in dark blue. The scale bars in **a** are 10 Å.

difficult to detect if the atom-counting precision is limited, but once detected, it can more easily be located along the depth direction. In general, if the direction of the incident electron beam is not immediately visible based on the reconstructed 3D atomic models, a trustworthy reconstruction is obtained.

Experimental results

As a last part of this work, we apply the Bayesian genetic algorithm to 25 frames of an experimental time series of a catalyst Pt nanoparticle³⁰. The experimental details and corrections for scan noise and tilt are described in the Methods section.

The incident electron dose is $1.38 \cdot 10^4 \text{ e}^{-\text{\AA}^{-2}}$ per frame, i.e., higher than the doses where we observe an improved reconstruction from Fig. 3. However, as already mentioned, during experiments, environmental distortions introduce additional uncertainties and this can be observed from the estimated width of the overlapping normal distributions of the Gaussian mixture model. This estimated width σ_{GMM} reaches values up to 0.0228, and can be compared to the values shown in Supplementary Fig. S1. This experimental value corresponds to the width for a dose lower than $10^3 \text{ e}^{-\text{\AA}^{-2}}$, when considering Poisson noise only. Furthermore, the probability matrices for the finite atom-counting results from the experiment and the simulation study can be compared. Supplementary Fig. S6 illustrates that the experimental probability matrix corresponds more to the probability matrix of $5 \times 10^2 \text{ e}^{-\text{\AA}^{-2}}$ for which the Bayesian genetic algorithm clearly benefits.

To reliably count the number of atoms from the time series of images, we used a hidden Markov model which explicitly describes the possibility of structural changes over time^{45,46}. The atom-counting results from every single frame have been used as an input for our Bayesian genetic algorithm in which we utilize the finite atom-counting precision and neighbor-mass relations. The reconstructed models are schematically represented in Fig. 5. The ADF STEM images and corresponding reconstructed models for all frames are shown in Supplementary Figs. S7 and S8. This approach enables a reliable 3D quantification of the structural changes of the Pt nanoparticle under the electron beam. From the evaluation of the coordination numbers (Supplementary Fig. S9), we can conclude that although each image has unique noise and that the structure is moving under the electron beam, the number of atoms with the same coordination number is consistent throughout time. Since these coordination numbers are very important to relate to the catalytic properties, it is important to point out here that the small changes clearly do not change the overall catalysis-relevant information that we can extract.

To summarize, we introduced a powerful alternative to the initially developed atom-counting/energy minimization method for the 3D reconstruction of nanoparticles from a single viewing projection. This Bayesian genetic algorithm takes advantage of the finite atom-counting precision and neighbor-mass relations during the reconstruction. This results in more reliable reconstructions of the 3D atomic structure, especially at lower incident electron doses below $10^4 \text{ e}^{-\text{\AA}^{-2}}$. The increased quality of the 3D atomic models has been validated by a quantitative evaluation of the reconstructed surface atoms. This result shows great promise to use these reconstructions to predict the adsorption properties of catalytic nanoparticles.

METHODS

STEM simulations

Image simulations were performed using the MULTEM package^{47,48}. An acceleration voltage of 200 kV, a semi-convergence angle of 22.48 mrad, and a pixel size of 0.124 Å were chosen, and averaging over 30 unique phonon configurations was performed. A source size having a Gaussian profile with an FWHM of 1.0 Å was added to further reflect experiments recorded under the same conditions. The full set of simulation settings is listed in Supplementary Table S1.

Atom-counting methodology

Using the StatSTEM software³⁴, a parametric imaging model is fitted to the simulated ADF STEM images. This model consists of a superposition of N Gaussian functions and describes the intensity

at the pixel (k, l) at the position (x_k, y_l) of the ADF STEM image:

$$f_{kl}(\boldsymbol{\theta}) = \zeta + \sum_{n=1}^N \eta_n \exp\left(-\frac{(x_k - \beta_{x_n})^2 + (y_l - \beta_{y_n})^2}{2\rho^2}\right) \quad (4)$$

with ζ a constant background accounting for the amorphous carbon support, ρ the width of the Gaussian peak, η_n the column intensity of the n th Gaussian peak, and β_{x_n} and β_{y_n} the x - and y -coordinate of the n th atomic column, respectively. The unknown parameters of the imaging model are given by the parameter vector:

$$\boldsymbol{\theta} = (\beta_{x_1}, \dots, \beta_{x_N}, \beta_{y_1}, \dots, \beta_{y_N}, \rho, \eta_1, \dots, \eta_N, \zeta) \quad (5)$$

and are estimated in the least square sense. From the obtained estimated parameters $\hat{\boldsymbol{\theta}}$, the estimated SCSs can be calculated from the volumes under the Gaussian peaks above the background^{33,34}:

$$\text{SCS}_n = 2\pi\hat{\eta}_n\hat{\rho}_n^2. \quad (6)$$

The estimated SCSs are visualized in a histogram in Fig. 2g, h and are regarded as independent statistical draws from a so-called Gaussian mixture model. In a sense, the assumption of independent statistical draws implies that cross-talk between neighboring atomic columns is not significantly contributing^{49–52}. The model is defined as a superposition of Gaussian components and describes the probability that a specific SCS value is observed. The probability density function of a mixture model with G components can parametrically be written as:

$$f_{\text{mix}}(\text{SCS}_n; \boldsymbol{\Psi}_G) = \sum_{g=1}^G \pi_g \frac{1}{\sqrt{2\pi}\sigma_{\text{GMM}}} \exp\left(-\frac{(\text{SCS}_n - \mu_g)^2}{2\sigma_{\text{GMM}}^2}\right). \quad (7)$$

The locations μ_g of the normal distributions are matched to the expected SCS values for a column containing g atoms obtained from image simulations performed for a Pt crystal in [110] orientation up to 30 atoms thickness. The settings for the frozen lattice multislice simulation are the same as for the simulation of the Pt nanoparticle embedded in the carbon support (listed in Supplementary Table S1). The symbol $\boldsymbol{\Psi}_G$ in Eq. (7) represents the vector containing all unknown parameters in the mixture model with G components:

$$\boldsymbol{\Psi}_G = (\pi_1, \dots, \pi_{G-1}, \sigma_{\text{GMM}}). \quad (8)$$

The parameters π_g and σ_{GMM} denote the mixing proportion of the g th component and the width of the components respectively. The mixing proportions π_g sum up to unity, therefore the G th mixing proportion is omitted in the parameter vector $\boldsymbol{\Psi}_G$. The parameters $\boldsymbol{\Psi}_G$ are estimated from the measured SCSs using the maximum likelihood estimator for a given number of components G . Here, G equals the maximum thickness of the image simulations of the Pt crystal, i.e., 30 atoms, since mixing proportions of components that exceed the maximum thickness of the Pt nanoparticle are estimated zero.

In principle, the width σ_{GMM} reflects the finite atom-counting precision. However, it is known that the width of the Gaussian distributions might be underestimated for lower electron doses³⁷. In order to counterbalance this underestimation, an effective width σ_{eff} will be used to describe the width of the normal components in the Gaussian mixture model. For this purpose, σ_{GMM} is evaluated with respect to the expected dose-dependent width σ_{dose} as explained in more detail in the Supplementary methods.

Genetic algorithm

The parameters of the genetic algorithm are specified for the reconstruction of small nanoparticles and balanced with the

Table 1. Genetic algorithm parameters.

Parameter	Value
Population size	500
Recombination fraction	50%
Mutation density	2%
Height mutation range	±1 lattice step
Count mutation range	±1 atom

available RAM/CPU resources (desktop computer) (Table 1). A larger population size will capture a more diverse spread of solutions and result in a longer computation time. For small particles, a smaller population size can be used. Here, a population size of 500 particle configurations is used which is significantly larger than the number of atomic columns in the particle (≈ 100 atomic columns). A fraction of this population, i.e., the recombination fraction or cross-breeding fraction will survive from one generation to the next and will be used as “parents” for recombining or cross-breeding new solutions. A too small fraction would reduce the diversity in the solution. A too large fraction, on the other hand, reduces the amount of new children in each generation. Therefore, a 50% fraction is chosen which means that every solution with a quality above average will be preserved and selected for breeding the next generation. If only existing parents are used in the recombination step, there is a risk of ending up in a local minimum. For this reason, after each breeding operation, a small percentage of mutants for 2% of the atomic columns can be introduced to inject some randomness into the atomic models. If these traits are an improvement, they will persist, otherwise they will automatically disappear. A very large degree of mutation is undesirable. In such a case, the benefit from the history of the evolution might be lost. The degree of mutation is expressed in terms of the height mutation range and count mutation range. If these values are too large, very non-physical solutions are proposed. Here we consider a mutation range of ± 1 atom in mass for a given atomic column and ± 1 position height shift or lattice step. This choice is also justified by both the finite atom-counting precision and the expected finite surface roughness. Important to notice here is that if a mutation is beneficial, then a cumulative evolution is possible. In this manner, a solution of ± 2 or more atoms and/or ± 2 or more height changes can be obtained over different generations. Next to the iterations throughout the different generations, a second loop feeds multiple unique starting initializations to the algorithm to reduce the risk of ending with local-minima solutions. For the simulation study, for each reconstruction, 25 unique random starting populations are used in the algorithm. It should be noted that a structure with a better cost-function might be found when increasing this number. For the experimental time series, we used 100 unique populations initializations throughout the reconstruction procedure.

Experimental Pt series

The ADF STEM images were recorded on a JEOL ARM200CF fitted with a probe-aberration corrector using an acceleration voltage of 200 kV, a probe convergence angle of 22.48 mrad, an annular detector ranging from 52 to 248 mrad, a dwell time of 4 μ s and an incident electron dose of $1.38 \cdot 10^4 \text{ e}^- \text{ \AA}^{-2}$ per frame. The images of the time series were corrected for drift and other distortions using non-rigid registration⁵³. During the time series, the Pt nanoparticle tilts slightly away from zone axis orientation and back, which affects the SCSs⁵⁴. Therefore, the SCSs of the individual frames need to be compensated for tilt as an input for the hidden Markov model⁴⁵. This is done by applying a linear scaling (up to 10%) to the SCSs of the individual frames^{37,55}, assuming that the total number of atoms in the nanoparticle remains constant throughout the time series.

This assumption is valid since the threshold energy for sputtering Pt atoms from a convex surface with step sites is 379 keV⁵⁶, well above the incident electron energy of 200 keV. We therefore do not expect sputtering of atoms from the surface, only surface diffusion⁴¹. This is also further confirmed directly from the experimental time series since no single Pt atoms are observed on the carbon around the Pt particle.

DATA AVAILABILITY

The datasets that support the findings of this study are available from the corresponding authors upon reasonable request.

CODE AVAILABILITY

The codes that support the findings of this study are available from the corresponding authors upon reasonable request.

Received: 24 March 2022; Accepted: 27 September 2022;

Published online: 12 October 2022

REFERENCES

- Narayanan, R. & El-Sayed, M. A. Shape-dependent catalytic activity of platinum nanoparticles in colloidal solution. *Nano Lett.* **4**, 1343–1348 (2004).
- Tao, F. et al. Break-up of stepped platinum catalyst surfaces by high CO coverage. *Science* **327**, 850–853 (2010).
- Calle-Vallejo, F., Sautet, P. & Loffreda, D. Understanding adsorption-induced effect on platinum nanoparticles: an energy-decomposition analysis. *J. Phys. Chem. Lett.* **5**, 3120–3124 (2014).
- Barron, H. & Barnard, A. S. Using structural diversity to tune the catalytic performance of Pt nanoparticle ensembles. *Catal. Sci. Technol.* **5**, 2848 (2015).
- Liu, P. et al. Three-dimensional atomic structure of supported Au nanoparticles at high temperature. *Nanoscale* **13**, 1770–1776 (2021).
- Fujita, T. et al. Atomic origins of the high catalytic activity of nanoporous gold. *Nat. Mater.* **11**, 775–780 (2012).
- Yankovich, A. B. et al. Picometre-precision analysis of scanning transmission electron microscopy images of platinum nanocatalysts. *Nat. Commun.* **5**, 4155 (2014).
- Jones, L., MacArthur, K. E., Fauske, V. T., van Helvoort, A. T. J. & Nellist, P. D. Rapid estimation of catalyst nanoparticle morphology and atomic-coordination by high-resolution Z-contrast electron microscopy. *Nano Lett.* **14**, 6336–6341 (2014).
- De Backer, A. et al. Dose limited reliability of quantitative annular dark field scanning transmission electron microscopy for nano-particle atom-counting. *Ultramicroscopy* **151**, 56–61 (2015).
- Altantzis, T. et al. Three-dimensional quantification of the facet evolution of Pt nanoparticles in a variable gaseous environment. *Nano Lett.* **19**, 477–481 (2019).
- Van Aert, S., Batenburg, K. J., Rossell, M. D., Erni, R. & Van Tendeloo, G. Three-dimensional atomic imaging of crystalline nanoparticles. *Nature* **470**, 374–377 (2011).
- Bals, S., Goris, B., Liz-Marzán, L. M. & Van Tendeloo, G. Three-dimensional characterization of noble-metal nanoparticles and their assemblies by electron tomography. *Angew. Chem.* **53**, 10600–10610 (2014).
- Miao, J., Ercius, P. & Billinge, J. L. Atomic electron tomography: 3D structures without crystals. *Science* **353**, 1380 (2016).
- Bals, S. et al. Atomic scale dynamics of ultrasmall germanium clusters. *Nat. Commun.* **3**, 897 (2012).
- Yu, M., Yankovich, A. B., Kaczmarowski, A., Morgan, D. & Voyles, P. M. Integrated computational and experimental structure refinement for nanoparticles. *ACS Nano* **10**, 4031–4038 (2016).
- De Backer, A. et al. Three-dimensional atomic models from a single projection using Z-contrast imaging: verification by electron tomography and opportunities. *Nanoscale* **9**, 8791–8798 (2017).
- Arslan Irmak, E., Liu, P., Bals, S. & Van Aert, S. 3D atomic structure of supported metallic nanoparticles estimated from 2D ADF STEM images: a combination of atom-counting and a local minima search algorithm. *Small Methods* **5**, 2101150 (2021).
- Geuchies, J. J. et al. In situ study of the formation mechanism of two-dimensional superlattices from PbSe nanocrystals. *Nat. Mater.* **15**, 1248–1254 (2016).
- Gelman, A. et al. *Bayesian Data Analysis* 3rd edn (Chapman & Hall/CRC Texts in Statistical Science, 2014).
- Lawson, A. B. & Lesaffre, E. *Bayesian Biostatistics* (John Wiley & Sons, 2012).
- Fernandez, M. & Caballero, J. Chapter 4 – genetic algorithm optimization of Bayesian-regularized artificial neural networks in drug design. In *Artificial Neural*

- Network for Drug Design, Delivery and Disposition (eds. Puri, M., Pathak, Y., Sutariya, V. K., Tipparaju, S. & Moreno, W.) 83–102 (Academic Press, 2016).
22. Yamashita, T. et al. Crystal structure prediction accelerated by Bayesian optimization. *Phys. Rev. Mater.* **2**, 013803 (2018).
 23. Goldberg, D. E. *Genetic Algorithms in Search, Optimization & Machine Learning* (Addison-Wesley Longman Publishing Co., Inc., 1989).
 24. Mitchell, M. *An Introduction to Genetic Algorithms* (A Bradford Book The MIT Press, 1996).
 25. Johnston, R. L. Evolving better nanoparticles: genetic algorithms for optimising cluster geometries. *Dalton Trans.* **22**, 4193–4207 (2003).
 26. Dugan, N. & Erkoç, S. Genetic algorithms in application to the geometry optimization of nanoparticles. *Algorithms* **2**, 410–428 (2009).
 27. Rossi, G. & Ferrando, R. Searching for low-energy structures of nanoparticles: a comparison of different methods and algorithms. *J. Phys. Condens. Matter* **21**, 084208 (2009).
 28. Wu, X. & Wu, G. An adaptive immune optimization algorithm with dynamic lattice searching operation for fast optimization of atomic clusters. *Chem. Phys.* **440**, 94–98 (2014).
 29. Kaczmarowski, A., Yang, S., Szlufarska, I. & Morgan, D. Genetic algorithm optimization of defect clusters in crystalline materials. *Comput. Mater. Sci.* **98**, 234–244 (2015).
 30. Aarons, J. et al. Predicting the oxygen binding properties of platinum nanoparticle ensembles by combining high-precision electron microscopy & DFT. *Nano Lett.* **17**, 4003–4012 (2017).
 31. Van Aert, S. et al. Procedure to count atoms with trustworthy single-atom sensitivity. *Phys. Rev. B* **87**, 064107 (2013).
 32. 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).
 33. Van Aert, S. et al. Quantitative atomic resolution mapping using high-angle annular dark field scanning transmission electron microscopy. *Ultramicroscopy* **109**, 1236–1244 (2009).
 34. De Backer, A., van den Bos, K. H. W., Van den Broek, W., Sijbers, J. & Van Aert, S. StatSTEM: an efficient approach for accurate and precise model-based quantification of atomic resolution electron microscopy images. *Ultramicroscopy* **171**, 104–116 (2016).
 35. E, H. et al. Probe integrated scattering cross sections in the analysis of atomic resolution HAADF STEM images. *Ultramicroscopy* **133**, 109–119 (2013).
 36. Ricolleau, C., Le Bouar, Y., Amara, H., Landon-Cardinal, O. & Alloyeau, D. Random vs realistic amorphous carbon models for high resolution microscopy and electron diffraction. *J. Appl. Phys.* **114**, 213504 (2013).
 37. 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).
 38. LAMMPS Molecular Dynamics Simulator. <https://lammps.org>. Accessed Nov 2019.
 39. Foiles, S. M., Baskes, M. I. & Daw, M. S. Embedded-atom-method functions for the FCC metals Cu, Ag, Au, Ni, Pd, Pt, and their alloys. *Phys. Rev. B* **33**, 7983 (1986).
 40. Lee, B.-J., Shim, J.-H. & Baskes, M. I. Semi-empirical atomic potentials for the FCC metals Cu, Ag, Au, Ni, Pd, Pt, Al, and Pb based on first and second nearest-neighbor modified embedded atom method. *Phys. Rev. B* **68**, 144112 (2003).
 41. 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).
 42. Calle-Vallejo, F., Martinez, J. I., García-Lastra, J. M., Sautet, P. & Loffreda, D. Fast prediction of adsorption properties for platinum nanocatalysts with generalized coordination numbers. *Angew. Chem. Int. Ed.* **53**, 8316–8319 (2014).
 43. Calle-Vallejo, F., Loffreda, D., Koper, M. T. & Sautet, P. Introducing structural sensitivity into adsorption-scaling relations by means of coordination numbers. *Nat. Chem.* **7**, 403–410 (2015).
 44. Calle-Vallejo, F. et al. Finding optimal surface sites on heterogeneous catalysts by counting nearest neighbors. *Science* **350**, 185–189 (2015).
 45. De wael, A. et al. Measuring dynamic structural changes of nanoparticles at the atomic scale using scanning transmission electron microscopy. *Phys. Rev. Lett.* **124**, 106105 (2020).
 46. De wael, A., De Backer, A. & Van Aert, S. Hidden Markov model for atom-counting from sequential ADF STEM images: methodology, possibilities and limitations. *Ultramicroscopy* **219**, 113131 (2020).
 47. Lobato, I. & Van Dyck, D. MULTEM: a new multislice program to perform accurate and fast electron diffraction and imaging simulation using graphics processing units with CUDA. *Ultramicroscopy* **156**, 9–17 (2015).
 48. Lobato, I., Van Aert, S. & Verbeeck, J. Progress and new advances in simulating electron microscopy datasets using MULTEM. *Ultramicroscopy* **168**, 17–27 (2016).
 49. Allen, L. J., Findlay, S. D., Oxley, M. P. & Rossouw, C. J. Lattice-resolution contrast from a focus coherent electron probe. Part I. *Ultramicroscopy* **96**, 47–63 (2003).
 50. Fertig, J. & Rose, H. Resolution and contrast of crystalline objects in high-resolution scanning transmission electron microscopy. *Optik* **59**, 407–429 (1981).
 51. Nellist, P. D. & Pennycook, S. J. Incoherent imaging using dynamically scattered coherent electrons. *Ultramicroscopy* **78**, 111–124 (1999).
 52. Martinez, G. T., van den Bos, K. H. W., Alania, M., Nellist, P. D. & Van Aert, S. Thickness dependence of scattering cross-sections in quantitative scanning transmission electron microscopy. *Ultramicroscopy* **187**, 84–92 (2018).
 53. Jones, L. et al. Smart align – a new tool for robust non-rigid registration of scanning microscope data. *Adv. Struct. Chem. Imaging* **1**, 8 (2015).
 54. MacArthur, K. E., D'Alfonso, A. J., Ozkaya, D., Allen, L. J. & Nellist, P. D. Optimal ADF STEM imaging parameters for tilt-robust image quantification. *Ultramicroscopy* **156**, 1–8 (2015).
 55. De wael, A., De Backer, A., Lobato, I. & Van Aert, S. Modelling ADFSTEM images using elliptical Gaussian peaks and its effects on the quantification of structure parameters in the presence of sample tilt. *Ultramicroscopy* **230**, 113391 (2021).
 56. Egerton, R. F., McLeod, R., Wang, F. & Malac, M. Basic questions related to electron-induced sputtering in TEM. *Ultramicroscopy* **110**, 991–997 (2010).

ACKNOWLEDGEMENTS

This work was supported by the European Research Council (Grant 770887 PICOMETRICS to S.V.A. and Grant 823717 ESTEEM3). The authors acknowledge financial support from the Research Foundation Flanders (FWO, Belgium) through project fundings (G.0267.18N, G.0502.18N, G.0346.21N) and a postdoctoral grant to A.D.B. L.J. acknowledges Science Foundation Ireland (SFI – grant number URF/RI/191637), the Royal Society, and the AMBER Centre. The authors acknowledge Aakash Varambha for his assistance and expertise with the experimental recording and use of characterization facilities within the David Cockayne Centre for Electron Microscopy, Department of Materials, University of Oxford, and in particular the EPSRC (EP/K040375/1 South of England Analytical Electron Microscope).

AUTHOR CONTRIBUTIONS

A.D.B., S.V.A., P.D.N., and L.J. conceived the idea. L.J. implemented the genetic algorithm. A.D.B. implemented the Bayesian aspect in the genetic algorithm and performed the analysis. C.F. assisted with the formulation of the Bayesian cost-function. E.A.I. performed the geometry relaxations using LAMMPS. A.D.B. wrote the manuscript. All authors reviewed and commented on the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41524-022-00900-w>.

Correspondence and requests for materials should be addressed to Sandra Van Aert or Lewys Jones.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022