

Artificial Intelligence End-to-End Workflow for Transmission Electron Microscopy: From Data Analysis Automation to Materials Knowledge Unveiling

Marc Botifoll^{1,*},†, Ivan Pinto-Huguet^{1,†}, Enzo Rotunno^{2,*}, Thomas Galvani¹, Catalina Coll¹, Payam Habibzadeh Kavkani^{2,3}, Maria Chiara Spadaro^{4,5}, Yann-Michel Niquet⁶, Martin Børstad Eriksen⁷, Sara Martí-Sánchez¹, Georgios Katsaros⁸, Giordano Scappucci⁹, Peter Krogstrup¹⁰, Giovanni Isella¹¹, Andreu Cabot^{12,13}, Gonzalo Merino⁷, Pablo Ordejón¹, Stephan Roche^{1,13}, Vincenzo Grillo², and Jordi Arbiol^{1,13,*}

¹Catalan Institute of Nanoscience and Nanotechnology (ICN2), CSIC and BIST, Campus UAB, Bellaterra, 08193, Barcelona, Catalonia, Spain

²CNR Istituto Nanoscienze, Via Campi 213/A, 41125 Modena, Italy

³University of Modena and Reggio Emilia, 42122, Reggio Emilia RE, Italia

⁴Department of Physics and Astronomy “Ettore Majorana”, University of Catania, via S. Sofia 64, 95123 Catania, Italy

⁵CNR-IMM, via S. Sofia 64, 95123 Catania, Italy

⁶Univ. Grenoble Alpes, CEA, IRIG-MEM-L Sim, Grenoble, France

⁷PIC, IFAE, Campus UAB, Bellaterra, 08193, Barcelona, Catalonia, Spain

⁸Institute of Science and Technology Austria, Klosterneuburg, Austria

⁹QuTech and Kavli Institute of Nanoscience, Delft University of Technology, PO Box 5046, 2600 GA, Delft, The Netherlands

¹⁰NNF Quantum Computing Programme, Niels Bohr Institute, University of Copenhagen, Denmark

¹¹Laboratory for Epitaxial Nanostructures on Silicon and Spintronics, Physics Department, Politecnico di Milano, Como, Italy

¹²Catalonia Institute for Energy Research – IREC, Sant Adrià de Besòs, Barcelona, 08930 Spain

¹³ICREA, Passeig Lluís Companys 23, 08010, Barcelona, Catalonia, Spain

*Corresponding authors: marc.botifoll@icn2.cat, enzo.rotunno@nano.cnr.it, arbiol@icrea.cat; † Equal contribution

Abstract

This article introduces a groundbreaking analytical workflow designed for the holistic characterisation, modelling and physical simulation of de-

vice heterostructures. Our innovative workflow autonomously, comprehensively and locally characterises the crystallographic information and 3D orientation of the crystal phases, the elemental composition, and the strain maps of devices from (scanning) transmission electron microscopy data. It converts a manual characterisation process that traditionally takes days into an automatic routine completed in minutes. This is achieved through a physics-guided artificial intelligence model that combines unsupervised and supervised machine learning in a modular way to provide a representative 3D description of the devices, materials structures, or samples under analysis. To culminate the process, we integrate the extracted knowledge to automate the generation of both 3D finite element and atomic models of millions of atoms acting as digital twins, enabling simulations that yield essential physical and chemical insights crucial for understanding the device’s behaviour in practical applications. We prove this end-to-end workflow with a state-of-the-art materials platform based on SiGe planar heterostructures for hosting coherent and scalable spin qubits. Our workflow connects representative digital twins of the experimental devices with their theoretical properties to reveal the true impact that every atom in the structure has on their electronic properties, and eventually, into their functional quantum performance. Notably, the versatility of our workflow is demonstrated through its successful application to a wide array of materials systems, device configurations and sample morphologies.

Keywords— artificial intelligence, transmission electron microscopy, physical modelling, end-to-end automation, quantum materials and devices, machine learning

1 Introduction

In an era marked by profound digital transformations, semiconductor heterostructures within a chip have emerged as crucial and widespread assets, driving major industrial value chains. They support advancements in both novel sectors (such as automated vehicles, cloud computing, Internet of Things, space exploration, supercomputing, and quantum technologies) and traditional ones (including computing and communications, industrial automation, entertainment, and healthcare). [1, 2] Miniaturisation, now reaching the nanoscale and approaching the atomic limit, stands out as a primary driver of progress. This trend enhances device capabilities, lowers costs, and reduces energy consumption. [3]

The tackling of these novel technologies and their required extreme miniaturisation signifies a paradigm shift in device design, where “every atom matters”. In contrast to traditional electronic devices, the diminishing sizes introduce significant nuances, such as fluctuations in dopant concentration, interdiffusion at interfaces, and local strain fields at the nanoscale, profoundly impacting device function and performance. [4, 5] Devices for quantum computing exemplify the extreme case. Any structural deviation from the ideal conceptual design will be fatal for the quantum performance, even though the exact causes and cor-

relations between these structural features and the final functional properties are still unknown in most systems. This ubiquitously occurs at different scales in material systems in which miniaturisation is key towards property improvement, the precise characterisation of every single atom will be key in understanding its properties when embedded in a full device (i.e., ferroelectrics and their interatomic distances, catalytic particles and their progression into single-atom catalysts, batteries and the atom-wise *in-situ* analysis of their degradation mechanisms, and others). [6, 7, 8, 9, 10, 11]

Beyond its characterisation, the exploration and development of novel materials and devices, as well as the optimisation of existing ones for various applications, constitute a multifaceted process involving necessity identification, literature review, material proposal, device engineering, characterisation, and application testing. [12] This iterative cycle, driven by multidisciplinary and collaborative efforts, forms the backbone of scientific progress. However, it is often hindered by its time-consuming and expensive nature, particularly when atomic scale precision is required for understanding the functionality of materials and heterostructured devices. The gold standard for achieving this atomic characterisation is (Scanning) Transmission Electron Microscopy ((S)TEM). However, this technique is currently constrained by the aforementioned challenges, which restrict the number of experimental repetitions and diminish the statistical significance of the results. [13, 14, 15, 16]

While substantial strides have been made in automating (S)TEM, especially in industrial settings for metrology and process characterisation, a substantial challenge persists in extracting meaningful physical insights from the vast amount of data (raw images and spectra) generated during experimental analysis. [17, 18] This challenge hinders a deep understanding of embedded material heterostructures in devices at the atomic level. Conventional fab and lab metrology tools fall short in providing a comprehensive and efficient analysis of these intricate device architectures, leading to a lack of statistical sampling for understanding performance variability among individual devices. As a result, the demand for high-throughput analyses that provide statistical significance and link structural characterisation with functional properties is more justified than ever.

In the present work, we address the inherent challenges in the traditional (S)TEM-based materials exploration process and introduce a revolutionary analytical paradigm that links it to physical modelling facilitated by recent breakthroughs in data analysis. [19, 20, 21] Machine learning (ML), deep learning (DL), computer vision (CV), and artificial intelligence (AI), have transformed nearly every facet of our daily lives, and materials science is not an exception, enabling levels of accuracy, precision, and noise tolerance previously considered unachievable in (S)TEM-related analyses. [16, 17, 22, 23] However, since its introduction in electron microscopy for materials science, the challenge of generalising its methods has been regarded as its greatest limitation.

The early stages of AI-driven methodologies associated with (S)TEM data analysis, characterised by the utilisation of relatively straightforward unsupervised unmixing algorithms to decompose hyperspectral signals like Energy

Dispersive X-Ray Spectroscopy (EDX) or Electron Energy Loss Spectroscopy (EELS) spectra, have evolved into the adoption of advanced models such as convolutional neural networks, autoencoders, or reinforcement learning. [16, 24, 25, 26, 27, 28] These advanced models are employed to unveil and learn features from images and high-dimensionality signals such as spectral images or 4D-STEM data. [29, 30] Notably, these advancements have paved the way towards the automation of both experiments and data analysis.[31, 32, 33, 34, 35] These innovations result in unparalleled insights from AI-based data analysis while still constrained by case-specific routines and limited statistical significance. [36, 37, 38, 39, 40]

In the present manuscript, we explore the integration of these advanced techniques into a comprehensive and automated characterisation workflow, aiming to overcome the traditionally slow and tedious aspects of materials research. Not only does our solution provide a new paradigm of automation in STEM characterisation but also an easy way to generate new knowledge from representative 3D models of the experimental devices which would otherwise imply an impractical manual crafting (atom-by-atom in atomic models, and contour-by-contour in finite element models). The proposed workflow starts by automating the data analysis process, traditionally considered a bottleneck, turning it into a solution for rapid and reproducible knowledge retrieval. The manuscript outlines a step-by-step approach inspired by the logical progression of human microscopists, beginning with low-magnification segmentation to gauge device morphology and culminating in phase and orientation-sensitive Geometrical Phase Analysis (GPA), for detailed structural insights on local atomic displacements, strain and defects.[41, 42]. Importantly, the workflow extends beyond mere data analysis and incorporates the automated generation of representative 3D atomic (3DAMs) and Finite Element (FEMs) models, utilizing the experimental data collected. Notably, these models comprehensively capture all the experimental information obtained through the preceding automated steps. As a result, the automated workflow can simulate a device that closely matches the originally designed, engineered, and grown device (i.e., digital twin). We refer to this process as an “experimental simulation”, as the models are created entirely from experimental data and parameters. For example, the finite elements of the FEMs are derived from the contours outlined by the segmentation of low-magnification images, while the atomic positions in the 3DAMs can be determined by the displacements identified through GPA on atomic resolution STEM images. This capability facilitates an efficient workflow and enables unprecedented exploration of atomistic models (digital twins) comprising millions of atoms. The structural relaxation properties of these models are represented by finite element relaxations and Keating models, while their electronic structures are derived from precise tight-binding Hamiltonians and computed using linear scaling algorithms. Ultimately, this establishes a direct link between realistic atomistic representations of as-grown materials and the variations in both local and global physical properties of the associated quantum devices.

The manuscript demonstrates the flexibility and adaptability of the proposed workflow, emphasizing its modular nature. This modularity allows seamless in-

tegration of cutting-edge research and open-source tools from diverse sources, positioning the workflow at the forefront of AI-driven analysis of electron microscopy datasets. The potential applications of this automated approach extend to diverse fields, including energy and environmental research, classical communications, quantum technologies, mechanical engineering, and fundamental chemical research.

In subsequent sections, we delve into the details of each step within the proposed workflow, elucidating the interconnectedness and modularity that make this method a promising avenue for accelerating materials science research. The manuscript also discusses the broader implications of this AI-driven methodology and its potential to reshape the landscape of materials exploration and characterisation. [17, 43]

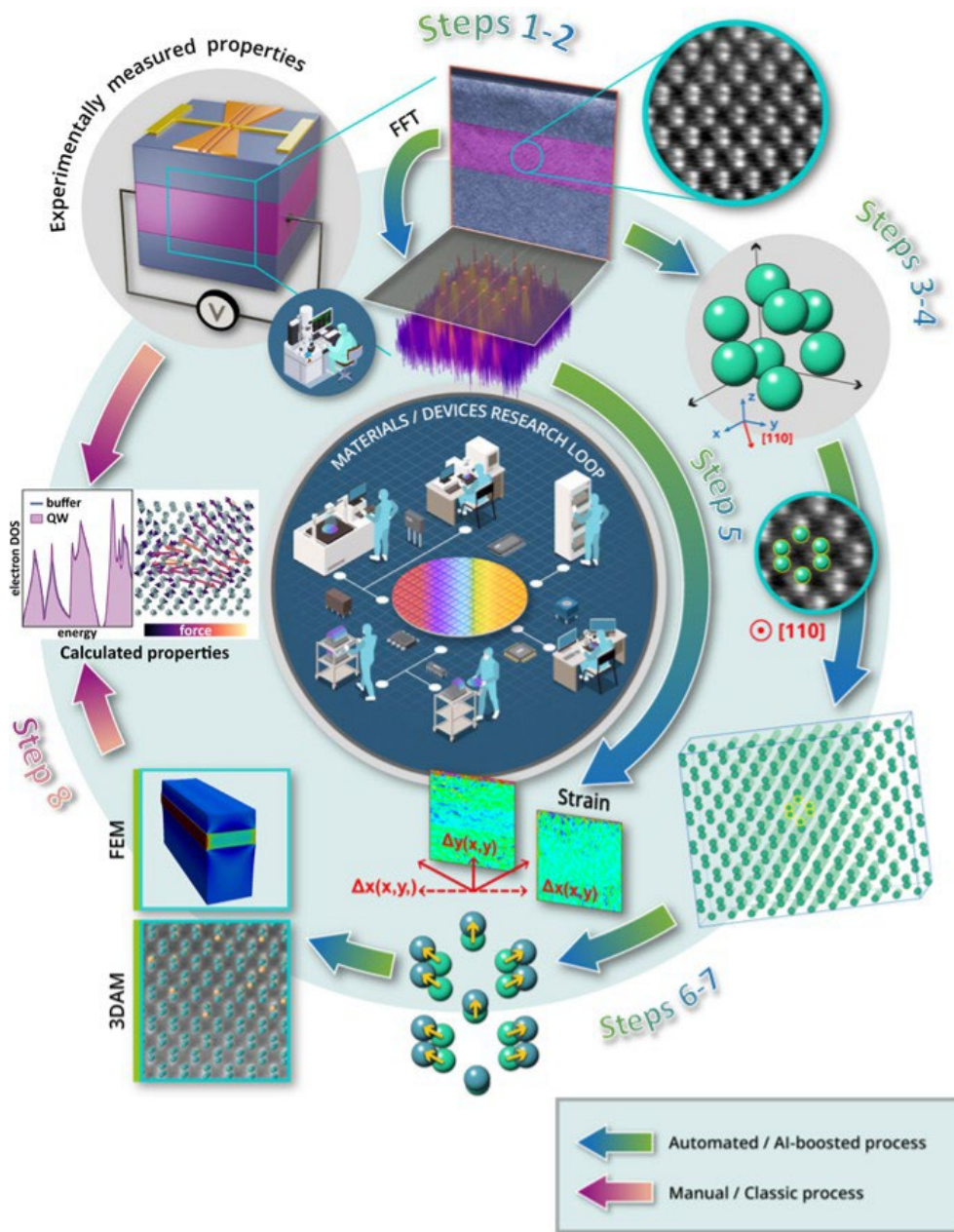


Figure 1: Scheme of the proposed workflow: Step 1: Initially, images are segmented, breaking them down into distinct regions. Steps 2-4: The micrograph analysis in Fourier space characterises the crystallography within each segmented region. Step 5: This comprehensive local crystallographic characterisation enables to map strain and to get finer structural details, such as interfaces or defects. Steps 6-7: The compiled information is then transferred to create Finite Element and 3D Atomic Models (FEM/3DAM). Step 8: These models, in turn, serve as the basis for computing functional properties like the electronic and vibrational states, or atomic forces and strain relaxations. This sequential process ensures a comprehensive analysis and understanding of the material's functional characteristics and behaviour.

2 Workflow

The workflow connects the data that is obtained directly from the electron microscope with the properties of the imaged heterostructure or device, in an intuitive and automated manner. The focus is on single images or stacks of them, as well as spatially correlated stacks of spectra, known as spectrum or spectral images. This paper concentrates on automating image analysis rather than spectroscopy, although compositional information is incorporated when necessary to enhance accuracy.

The choice between a single micrograph and a stack of images depends on the device's or heterostructure's size, imaged features, and microscope sampling. A single micrograph is suitable when it captures both the morphology and the atomically-resolved structural information. Conversely, a stack of images is used for larger devices or structures, requiring a progressive decrease in magnification to transition from the morphological overview to atomic and structural details. Both situations are handled differently, with the latter requiring an additional processing step to link information gathered at each magnification.

The modular workflow is represented in figure 1. Firstly, the autonomous modules can independently produce results on their own. Secondly, the modules are autonomously interconnected to link their outputs into a final result of additive complexity. The full cycle essentially consists of eight main modules which sequentially 1) segment the micrographs, 2) peak-find the Fast Fourier Transform (FFT), identify the crystal phase from 3) a single image or 4) a stack of them, 5) compute strain, build representative 3D models of the experimental sample or device, either 6) FEMs or 7) 3DAMs, and finally 8) calculate functional properties based on the model representation of the experimental device (Fig. 1).

Every independent module, its contribution to the state of the art, and its additive complexity towards our proposed “experimental simulations” are described next.

2.1 Segmentation

Initially, our goal is to reveal the morphology of devices, encompassing their size, identification of key regions, their spatial distribution and their interfaces and contours. These regions typically contain different materials or diverse crystal phases or orientations. The concept is that any features within the image distinguishable by significant differences in pixel intensity can be separated as segmented units, which is particularly challenging given the study's aim for a general solution applicable to a broad range of devices and samples. [44, 45] This targeted diversity and the consequent absence of labelled data has required an unsupervised machine learning approach. [46, 47, 48, 49] After evaluating multiple algorithms, we chose the Canny edge detection algorithm, which we optimized and automated, and generalised for our target data type. [49] The details on the optimisation and automation are included in the supplementary information, section 1.1 entitled “Segmentation”.

In addition, the model is combined with a state-of-the-art general-purpose segmentation neural network, the Segment Anything Model (SAM). [50] Canny edge model stands as a comparably precise but faster solution compared to SAM. Specifically, up to two orders of magnitude faster in demanding samples where SAM underestimates the contours between segmented regions. On the other hand, SAM allowed us to expand the domains of our targeted systems to parallel beam TEM data whose diffraction contrast hinders an intuitive naked-eye segmentation. [50, 51] The optimised Canny edge detection model is tailored to detect typical (S)TEM edges (in high angle annular dark field (HAADF), bright field (BF) or integrated differential phase contrast (iDPC) STEM, BF TEM and high-resolution TEM (HRTEM) imaging modes) comprising dozens of pixels while SAM’s general digital image processing focuses on few-pixel intensity gradients. With this double model, the workflow offers the selection of the segmentation model based on the targeted device, highlighting the benefits of the human-in-the-loop approach to add control layers to a fully automated process, if wanted. Comprehensive details of the model performance, training, metrics, labelling processes, and more, are available in the supplementary information, section 1.1.3 entitled “Final segmentation model proposed”.

2.2 Fast Fourier Transform peak finding

After segmentation, automated structural characterisation begins by transforming real space images into reciprocal space using Fast Fourier Transform (FFT).[52, 53, 54] The key is using the segmented information to spatially filter the information represented in the FFT, either by masking or by cropping from the segmentation (supplementary information, section 1.2 “Peak finding in the Fourier spectra”). The goal in either case is to find the reciprocal space coordinates of the frequency peaks in the FFT representing crystallographic planes, as they encode the information of the imaged local crystallographic phases. The workflow has been ultimately designed to be universal and to detect amorphous phases, single crystals, and polycrystalline materials. Therefore, we designed our peak-finding method to be robust enough to maximise the crystallographic planes detected (recall: % of detected real planes from all real planes in the image) without noise (precision: % of real planes from all detected planes). Three or more detected planes per crystallographic phase are enough to instil confidence in model-based phase identification. Therefore, we prioritise maximising recall over precision to minimise the possibility of missing local crystallographic phases represented by a reduced number of planes in the FFT.

The existing peak-finding algorithms are abundant. [55, 56, 57, 58] Nevertheless, they all imply manual fine-tunings and sample-dependent considerations. As a result, comparing them to our method would overlook the key aspect of automation that distinguishes our approach, making it, to the best of our knowledge, the most accurate and robust autonomous FFT peak-finding method available in the literature. Our peak finding model relies on the successive application and evaluation of up to three distinct methods: 1) experimental ML-based 2D Gaussian fitting, 2) a trained U-Net model on synthetic data, and

3) a 1D profile scanning. [44, 56] The first and second methods lie within supervised ML/DL, while the third can be regarded as a computer vision algorithm. First, the ML-based 2D Gaussian fitting automatically optimises the parameters that would classically demand manual tuning in a 2D Gaussian fitting. Second, a custom model based on noisy kinematical diffraction patterns is deployed to train a U-Net model peak identifier. [33, 44] The third and complementary 1D profile scanning method parallelly scans the FFT vertically and horizontally, merging both outputs to eliminate misleading cross-shape artefacts. These three approaches avoided manual labelling approaches to rely on model-based labelling.

The individual performance of all three methods could not equilibrate the recall-precision balance required to grant the posterior successful crystal phase identification step in the wide range of tested materials systems and devices. Their performance is detailed in the supplementary information, section 1.2.5 “Peak finding performance metrics”. Therefore, we combined the three through a pipeline capable of, first, detecting if the crystal phase is amorphous or crystalline, and second, maximising the recall to deal with crystalline samples with multiple identifiable spots. While not flawless, it achieves the optimal balance between true positives and false positives and demonstrates adaptability across a wide range of materials and corresponding image types. [44, 59]

By testing this global model with 1000 manually labelled experimental FFTs from multiple materials, geometries and crystalline configurations, its precision of 69.78%, recall of 70.89%, and F1 score of 61.87%, make it, to the best of our knowledge, the most robust automated peak finding model available (details on metrics available in the supplementary information, section 1.2.5 “Peak finding performance metrics”). The recall and precision are high enough to ensure that the initial order Laue zones are well-identified, entailing a successful crystal phase identification, which is the eventual goal of the peak finding. Moreover, it detects neighbouring spots typically accounting for mismatched heterostructures or defects, which will be accounted for in the following steps. The nature and details of the peak-finding algorithm are carefully presented in the supplementary information, section 1.2.4 “Global combined peak finding model”. However, the refinement, further implications and added potential of the proposed peak finding model are out of the scope of the present text and will be disclosed in the future.

2.3 Phase identification

Once the crystallographic planes are located in the FFT, we can evaluate their goodness of fit with a database of candidate unit cells to match a crystallographic phase and its 3D orientation. To do so, a physics-aware model is convenient. Thus, this process involves ranking each potential crystal phase and assigning a score based on how well the experimental diffraction pattern encoded in the FFT aligns with their theoretical kinematical diffraction. [60, 61] The candidate unit cells are extracted from a crystallographic database and filtered by prior knowledge (i.e., involved chemical species) to optimise the phase identification.

[62, 63, 64, 65] The key lies in the iterative and combinatorial comparison of every pair of crystal planes detected in the experimental FFT with their best 3D fit within the candidate phases' theoretical pair of diffracted planes. By treating the problem with spot pairs as the validating unit, we make the phase identification robust and sensitive to structural defects that introduce additional planes in the FFT like twin boundaries or stacking faults. The output is an automated plane-wise indexation of the FFT and the consequent 3D picture of the most likely crystallographic phases considered from the database. Importantly, the workflow incorporates an interactive graphical interface for visualising the raw FFT indexation and the sorted list of likely phases. The nature of the user interfaces can be found in section 2 "Proofs of the automated phase identification" of the supplementary information, as well as in the supplementary audiovisual material (videos showing the workflow in real-time).

The sequential segmentation and peak-finding enable the distinctive identification of peaks that would be too close together in a single global FFT of combined regions, allowing for separate frequency and phase identification. Consequently, the algorithm is resistant to cumulative spot position shifts caused by drift, or sensitive to subtle lattice mismatches in heterostructures. The algorithm's tolerance between experimental and theoretical differences is set to 5%, providing the optimal balance between accuracy in phase identification and robustness to deviations from drift or calibration shifts. Steps 1-4 (Fig. 1) present the sequential process till the indexation of the crystallographic phase and its 3D information. Figure 2 proves this versatility: it shows three examples of paradigmatic nanomaterials and heterostructure configurations in which the workflow univocally identifies the crystal phases and their 3D arrangement. The workflow is capable of detecting highly mismatched interfaces such as cubic defective InSb grown on InP and detecting their relative orientation (Fig. 2.a). [66, 67] It can also identify the closest pure unit cell in a binary compound, like the $\text{Si}_{0.5}\text{Ge}_{0.7}/\text{Ge}/\text{Si}_{0.5}\text{Ge}_{0.7}$ quantum well in Figure 2,b to set the ground of the stoichiometrical refinement that will be addressed later in the workflow's pipeline. [68] Showcasing the model's versatility and robustness, in a different materials science domain, low-contrast TEM micrographs of CuTe nanoparticles are indexed and Bragg-filtered. Our phase identification can successfully point at the correct $Pm\bar{3}n$ phase among the up to 12 different candidate CuTe crystal phases (Fig. 2.c). [69]

As discussed, the phase-finding module is robust to false positives and false negatives in peak detection, thanks to the use of crystal plane-pair correlations when accepting a possible detected phase. This is demonstrated in Figure 2 and more extensively proved for additional heterostructures, devices, crystal phase types and spatial groups, morphologies, and orientations, in the supplementary information, sections 1.5 "Phase identification" and 2 "Proofs of the automated phase identification". The automated phase identification proposed in this work, grounded in a model-experiment comparison, achieves remarkable accuracy and robustness. Furthermore, its added value lies in its integration into a comprehensive analytical workflow that exploits its output to access further structural and functional insights of devices.

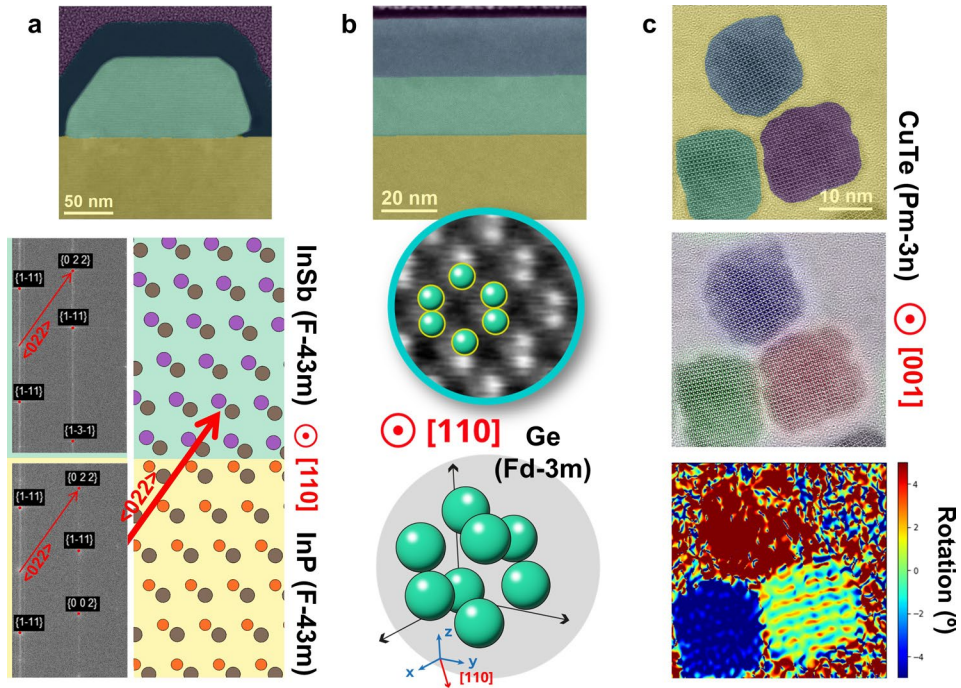


Figure 2: Outputs from the automated workflow for the comprehensive structural characterisation of devices. a) Scanning Transmission Electron Microscopy (STEM) image of an InSb nanowire cross-section on an InP substrate oriented along $[011]$ axis. The contour-based segmentation accurately separates the four distinguishable regions of the heterostructure within the lamella. The automated phase retrieval indexes the Fast Fourier Transform (FFT) per region and finds the spatial orientation of both constituting phases. [66] b) STEM micrograph of a $\text{Si}_{0.3}\text{Ge}_{0.7}/\text{Ge}/\text{Si}_{0.3}\text{Ge}_{0.7}$ quantum well. [68] The segmentation and phase identification conclude the device is entirely formed by Ge with diamond structure, along $[011]$. Not surprisingly, the stoichiometry of $\text{Si}_{0.3}\text{Ge}_{0.7}$ from the buffer layers is not retrieved, as the candidate unit cells are pure cubic Si and Ge. c) High-resolution parallel beam TEM micrograph of three CuTe square particles. SAM perfectly segments the particle-like format and highlights how the workflow is general and applicable to multiple device morphologies. From the segmentation and the $Pm-3n$ phase identification, we can Bragg filter (middle panel) the particles based on their in-(image)-plane relative rotation, as well as computing their local strain, validating their relative rotation (bottom panel). [69]

2.4 Low-to-high magnification correlation

The workflow will diverge if dealing with an entire image dataset of the device instead of with a single image. Step 4 applies to image stacks and circumvents the insufficiency of a single image to capture every morphological and structural detail, which is common in large devices of several μm . The process stacks micrographs of the same device, but with varying experimental parameters like magnification, focus, and sample orientation. These parameters are manually adjusted and defined during the acquisition process. The core idea involves sorting these images by field of view (FOV) and matching them in a chain of template and query images. The mathematical core, multiscale template matching, ensures the adequate pixel size-based scaling of the template-query pairs to maximise their matching based on cross-correlation. The low magnification images are segmented to reveal device morphology taking advantage of the reduced number of pixels per contour. Next, the automated reciprocal space analysis is performed on higher-magnification images containing structural details. This enables mapping crystal phases from atomically-resolved images, but linking them to the lowest magnification images, providing structural information in FOVs where no atomic resolution can be achieved. For instance, we have achieved crystal phase mapping (i.e., identification, 3D orientation, indexation) in FOVs of up to 1197 nm. This particular FOV would demand an image of around 20000 pixels per side to be within the Nyquist regime of atomic resolution. [70] We then replicate the advantages of 4D-STEM acquisitions without venturing into its big data, while also anticipating the future application of this workflow in 4D-STEM. The multiscale matching process is detailed in the supplementary information, section 1.4 “Low-to-high magnification correlation”.

2.5 Strain analysis

Back in the single-image scenario, having a complete description of reciprocal space is powerful. The combined knowledge retrieved so far (local FFT indexation, crystallographic phase and device morphology) reveals a global picture of material arrangement. This information facilitates mapping epitaxial relations and identifying heterojunctions or single crystalline blocks. We correlate the mapped reciprocal space through Geometrical Phase Analysis (GPA) in an automated fashion by leveraging knowledge obtained from real and reciprocal space. The method involves selecting a crystalline segmented region as the reference while identifying optimal planes (g vectors) from which to compute the geometrical phase and the strain. [41, 42, 60, 61] For instance, in the examples showcased in Figure 2, different epitaxial relations are automatically found. In figure 2, a and b, the workflow gauges a perfect epitaxy with high 10.3% (Fig. 2.a) and no (Fig. 2.b) mismatch, while in Fig. 2.c a partially misaligned epitaxy equivalent to a polycrystalline nature is retrieved. Further practical details and the resulting automated strain maps from these and other samples can be found in the supplementary information, section 1.6 “Strain analysis: Geometrical Phase Analysis automation”.

The crystallographic frequencies from the segmented regions are compared to the chosen g vectors, and a mask is drawn around them to fit neighbouring indexed crystal planes and to balance spatial resolution and noise. Still with examples from Figure 2, in Fig. 2.a the mask grabs the two planes representing the heteroepitaxy (Fig. S23), while in Fig. 2.b just the main one indicating the homogenous cell parameter (Fig. S24). In Fig. 2.c the mask opens to allocate three planes, one representative of each of the three nanoparticles (Fig. S30), to calculate their relative in-(image)-plane orientation within the same zone axis (further details in supplementary information, section 1.6.2 “Selection of the optimal g vectors pair and mask resolution”).

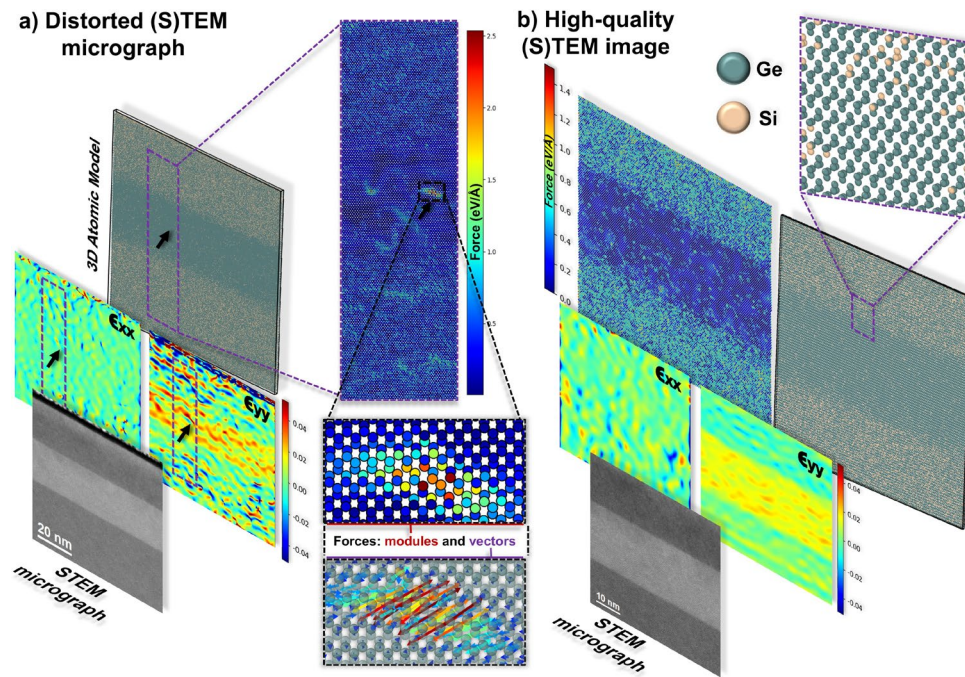


Figure 3: Transference of atomic-resolution information and relative strain mapping into 3D atomic models (3DAMs). From high-resolution (scanning) transmission electron microscopy ((S)TEM) images, the ϵ_{xx} and ϵ_{yy} relative strain maps are computed and its contained information is transferred to 3D atomic models to represent the experimental strain fields in an automated way. The fidelity of the eventual 3DAMs depends on the quality of the experimental (S)TEM micrographs from which they originate: a) Process involving a STEM micrograph of a $\text{Si}_{0.3}\text{Ge}_{0.7}/\text{Ge}/\text{Si}_{0.3}\text{Ge}_{0.7}$ quantum well with scanning artefacts and heterogeneous resolution of the atomic lattice. [68] The effects, turning into uneven strain maps, are transferred into the 3DAM unrealistically distorting the atomic lattice and causing strong atomic force accumulations (black arrow in the strain and atomic forces maps). b) Process involving a high-quality STEM micrograph of the same device or heterostructure with a smooth and even representation of the atomic lattice and consequently, of the strain. Its transfer to atomic models ends up with a realistic 3DAM where slight heterogeneities in the atomic forces arise due to the placement of the individual Si and Ge atoms in a strained lattice based on local quantitative electron energy loss spectroscopy. The atomic force maps are computed with a Keating model of Si and Ge alloys.

This methodology provides insights into the spatial dependencies of the elastic strain tensor components, the junctions between materials, their structural interaction or epitaxial relations, and the presence of dislocations in their in-

interface. Assessing the relative orientation of crystals and their epitaxial relations is valuable for visually mapping planar defects within the crystal phase, particularly with the automated Bragg filtering in the workflow. In fact, the heteroepitaxial relations automatically found are also used to adjust the Bragg filtering of crystal planes linked to each phase, revealing their spatial distribution. [71, 72] Moreover, our automated GPA routine highlights subtle structural features within the different interfaces of the device as abrupt local variations of the measured strain fields, going beyond the capabilities of the previous crystal phase identification step. This is particularly interesting for the global semiconductor field, including its powerful industry. Overall, it detects defects, dislocations, and alterations from the expected perfect crystal lattice, offering detailed insights that are not easily discernible to the naked eye. For example, the automated rotation map (supporting information, Fig. S23) in the system in Fig. 2.a highlights the stack of dislocations and the presence of a stacking fault. [66, 67] These elements were invisible to the workflow until this step. Furthermore, the quality and smoothness of the GPA maps depend on the quality of the original (S)TEM data, as with the original manual GPA routine. Figure 3 compares the quality of the retrieved strain maps of a Ge quantum well based on image quality and demonstrates how the smoothness of the resulting maps is closely linked to it. This concept is crucial for understanding how this knowledge is transferred to the eventual atomic model, as detailed later (Fig. 3). Similar examples prove the versatility of the method in the supplementary information, section 1.6.5 “Proofs of the automated GPA”.

Interpreting GPA’s output remains a manual task for the trained materials scientist. Indeed, the interpretation of the measured strain with this workflow is equivalent to a traditional manual GPA calculation. Therefore, the retrieved strain maps represent strain relative to the chosen reference region. [41, 42] From the resulting relative strain maps, the translation from relative dilatation to absolute in-plane strain components can be immediately computed, as the relaxed cell parameters and their local changes in each segmented region belonging to each phase have been automatically retrieved in the previous steps. In fact, it is the combination of these outputs that allows the posterior generation of atomic models that accurately represent the retrieved structural details, as discussed below.

2.6 Finite Element Model (FEM) building

At this stage, the workflow has gathered sufficient information to craft 3D models that are representative of the device or sample. However, it is important to note that this model generation is meaningful when we can assume translational invariance along, at least, one axis of the device.

FEMs creation: Our first approach is the automated creation of Finite Element Models (FEM), which describe intricate physical systems using discrete geometric elements. This enables numerical calculations that would otherwise be impractical when considering the entire global system at once. [73, 74, 75]

We transform the segmentation representing the morphology of the device into sequences of contour vectors forming a boundary elements model. We encode it into a Graphic Data Stream (.gds) file, which contains contour information, materials, and identified crystal phases. The level of morphological detail of the contour model is easily adjusted by manipulating the number of total nodes defining the contour. Thus, we trade-off between smoothing curves to reduce aliasing when segmenting downscaled images, and capturing subtler details by keeping more boundary elements. Additionally, we automatically align the 3D indexation of the identified crystal phases with the coordinates system of the FEM calculation to incorporate crystal phase-specific or orientation-dependent properties in the simulation, such as anisotropic strain tensors.

The workflow-based model generation outperforms traditional approaches to FEM work. It is automated, fast, and does not require manual input for system properties or tedious building of morphology with simple polygons. Moreover, as it is based on experimental data, it is more accurate and representative of the actual heterostructure or device being simulated. However, the introduction of simulation properties (e.g., mesh size and resolution, material properties database, boundary conditions) still requires manual setup in the simulation software of choice. [76, 77, 78]

FEM calculation example and discussion: We subjected the workflow to scrutiny by importing the InSb-InP nanowire cross-section presented in Figure 2 in COMSOL to evaluate its relaxation in physical scenarios of interest represented by different boundary conditions. [66, 67, 76] The results are illustrated in Figure 4.a, demonstrating the straightforward adaptability of simulations based on the proposed workflow. The device’s high mismatch of 10.3% experimentally forces the creation of an array of dislocations in the InSb/InP interface. This effect is observed in the automatically computed strain maps (Fig. 4.a, “Experimental GPA”, and supplementary information, “Proofs of the automated GPA”). As dislocations are an atomic effect, we manually fine-tune the automatic output from the workflow (Fig. S61) to consider them in our continuous simulation system ((Fig. 4.a, “FEM strain relaxation”). We also manually determine the calculation setup: we first impose a thermal strain (i.e., thermal expansion coefficient) of 10% to mimic the mismatch and the elastic component of the relaxation. Second, we manually create an array of circumferences spatially coinciding with the dislocation positions. We impose that a displacement of one epitaxial plane is fixed in each, representing the dislocation. The workflow then proves to be versatile enough to output a base FEM model that can intuitively be manually modified *ad-hoc* based on the particular needs of the system to represent.

We meshed our geometry with a varying mesh density that is maximised in the dislocations. We applied the boundary conditions sets detailed in the supplementary information, section 4.2 “FEM simulations” to present a common yet fully unresolved issue within materials science and TEM: the effect of lamellae thinning in strain mapping. We compare the unthinned device con-

sidering infinite translational invariance with a thinned TEM lamella of 40 nm. This respectively involves applying or not applying a boundary condition of null displacement in the transversal facets of the device. This approach allows us to study the expected difference in the strain relaxation between our measured thinned lamellae and the real device in the wafer. The complete description of the problem is detailed in the supplementary information (section 4.2 “FEM simulations”), but for simplicity, we present the representative component, ϵ_x , in Figure 4.a to compare both scenarios. The first and most obvious observation is the overall larger component in the unthinned system. Specifically, when considering the infinitely thick “lamella”, we obtained average values of 11-12 % dilatation in the nanowire. However, in the thinned version, the dilatation only reaches around 9 % in both components. This implies that the effect of the lamella thinning reduces the components by approximately 22 % ($1 - \frac{9\%}{11.5\%}$). In other words, thinning a 40 nm lamella in the present system releases 22 % of the expected transversal strain through longitudinal relaxation as observed by the lamellae expansion (figure 4.a and supplementary information, section 4.2 “FEM simulations”). Consequently, our experimental TEM measurements would reflect 22 % less than the unmodified value we would ideally measure.

It is also interesting to observe from the 3D views of the components, how the shape of the NW, especially at the kinks on both edges of the interface, unevenly modulates the strain. This highlights the importance of having the actual experimental morphology of the sample rather than just a schematic simplification of the ideal system. However, the comparison with the experimental strain measured with GPA is still discordant with the simulations, which release more strain than the limit of the lattice mismatch. This discrepancy arises because the simulations only consider a stack of dislocations along the transversal dimension and not longitudinally. The x and y components compensate for the remaining unrelaxed longitudinal strain. Notably, the y -component (Fig. 4.a) weighs a larger percentage of this relaxation, given that the epitaxial conditions tightly constrain the x -component. In summary, with the present workflow, we can intuitively tune our experimental FEM simulation to improve the match of our experimental system. As a consequence, a sample-wise analysis of the studied effect is unlocked, as well as the exploration of the implications of varying lamella thicknesses, among endless possibilities. Further details and discussion about the automated FEM process are available in the supplementary information section 4.2 “FEM simulations”.

2.7 3D Atomic Model (3DAM) building

We can go beyond FEMs with 3DAMs and provide a more precise description of the actual device and its structural features. [79, 80] While the atomic description allows for higher theoretical levels in simulations, its building and the actual calculations come at a greater cost in computing time and resources than the FEM. [60, 61, 81, 82, 83, 84] The fundamental concept here is to generate a representative atomic model: a file containing the three spatial coordinates of all atoms constituting the entire device or a specific region of interest. To achieve

this, we require the combination of the output from every previous module of the workflow.

The gathered knowledge from the workflow is first used to populate the volume occupied by the device (defined by the segmentation) with atoms based on the found conventional unit cells and their symmetry operations. [85] These atomic positions are subsequently linked with GPA analysis, particularly with displacement maps. The central concept is the correlation of the displacement maps with the atomic positions to calculate their local atom-wise displacement to the position where they are experimentally found. The key is to build the initial atomic model based on what we refer to as “virtual unit cell”, which is the unit cell whose cell parameters match those of the reference area of our automated strain analysis. In other words, since displacements are relative to our reference, we need to build everything based on this reference to later apply the modifications (i.e., displacements) relative to it. We use the two crystal planes (g vectors) used for GPA to obtain the experimental plane spacing, which we refine at sub-pixel level from the FFT, to compute the resulting virtual unit cell (details in section 1.9.2 “Virtual unit cell calculation” of supplementary information).

The automated strain analysis module is aware of epitaxial relations. Consequently, the construction of the 3DAM will be guided by this knowledge. Specifically, epitaxy detection groups two or more epitaxed regions into a replicated single virtual unit cell, from which the atoms are then displaced. For instance, in the $\text{Si}_{0.3}\text{Ge}_{0.7}/\text{Ge}/\text{Si}_{0.3}\text{Ge}_{0.7}$ quantum well displayed in Figure 3, a single virtual pure Ge building block is used for the three regions of the device: the quantum well and the upper and lower buffer layers (details in supplementary information, section 1.9.4 “Compositional information: Spectroscopic mapping and quantification”). [68] Nevertheless, in this case, the resulting atomic model built of only Ge atoms does not consider the local binary stoichiometry of the SiGe alloy yet, as this cannot be inferred from reciprocal space-based phase identification. We can circumvent this by adding quantitative compositional maps that spatially showcase the local Si:Ge ratio. Figure 3 displays the results after refining the 3DAM with electron energy loss spectroscopy (EELS) Si/Ge quantitative maps. The compositional refinement is based on a multi-scale template matching procedure similar to the one used in correlating image stacks to assign an element type to every atom (step 4). In this case though, the templates are the quantitative maps that correlate spatial coordinates with elemental composition. More examples and details can be found in the supplementary information, section 1.9.4 “Compositional information: Spectroscopic mapping and quantification”.

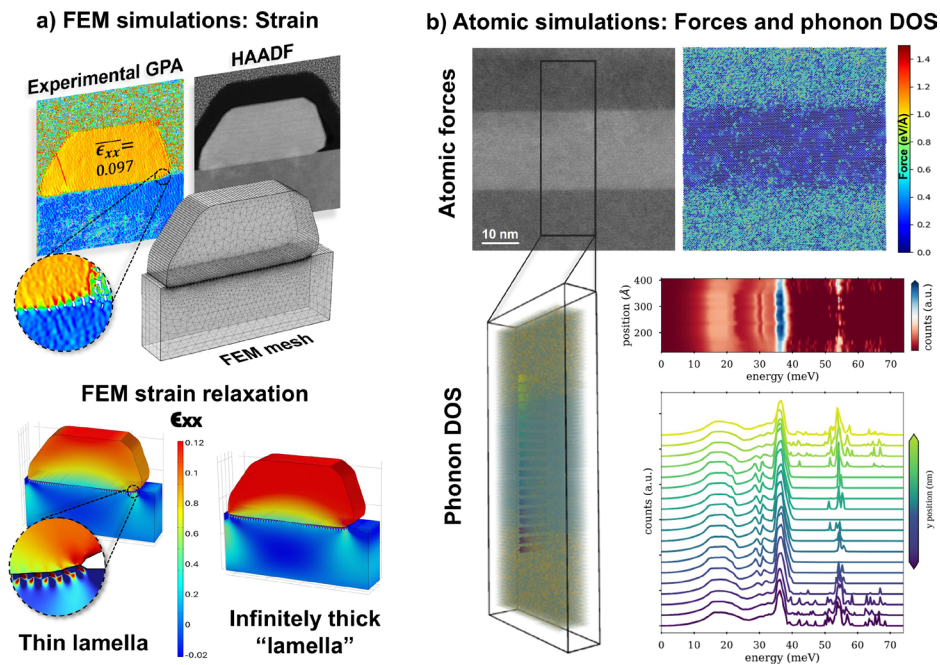


Figure 4: Multiscale computational materials science: Finite Element Model (FEM) and atomic simulations from the workflow’s outputs. a) FEM simulations of an InSb nanowire grown on InP. [66] The segmentation extracts its morphology to make a 3D geometrical FEM mesh of the experimental structure. Among endless possibilities, it unlocks computing strain relaxations and gauging the effect of lamellae preparation even in the presence of fine anomalies like dislocations (zoom-in) and subtle morphological features. b) Atomic simulations from the 3D Atomic Model (3DAM) of a $\text{Si}_{0.3}\text{Ge}_{0.7}/\text{Ge}/\text{Si}_{0.3}\text{Ge}_{0.7}$ quantum well. [68] The Keating model is used to compute atomic forces, which are larger at the top and bottom buffer layers due to the inclusion of a high percentage (around 30%) of Si atoms in a lattice closer to pure Ge. Furthermore, the Keating model maps the theoretical phonon density of states vertically along the heterostructure, highlighting the modulation of the optical mode of Si at 55 meV when going through the variations of the local Si content.

The situation slightly differs if the epitaxy arises from highly mismatched device regions, separated by the segmentation model at the beginning of the workflow. In these cases, the crystal planes of both regions appear separately in the FFT, which enables their separate crystal phase identification (see details in supplementary information, section 1.9.5 “Compositional information: Segmentation and symmetry equivalences”). Multiple strain calculations are sequentially computed by varying the reference position in every material par-

icipating in the epitaxy. Every virtual atomic block is displaced independently based on its own virtual reference unit cell. Eventually, the blocks are merged into the 3DAM representing the full device. This pertains to the InSb-InP structure displayed in figure 2.a and 4.a, the discussion and 3DAMs of which can be found in the supplementary information, section 3 “Proofs of the full workflow”. [66, 67]

The automated method described above is capable of modeling dislocations, stacking faults and local strain modulations. For instance, the formation mechanisms of stacking faults are captured for the InSb nanowire of Figures 2.a and 4.a, as detailed in the supplementary information, section 3 “Proofs of the full workflow”. In this case, the defect begins to form by placing the atoms with the expected mirroring of the dumbbells from the undistorted lattice, as fully described and detailed in Figure S59.

Achieving a perfect 3DAM is sensitive to the micrograph acquisition process. Importantly, the quality of the reconstruction is tightly linked to the micrograph quality and its consequent assessment in the automated strain maps retrieved, as showcased in Figure 3. Thus, as in classic GPA, scanning artefacts, noise or undersampling affect the smoothness of the resulting strain maps, and cumulatively, these will be showcased in the calculated 3DAM as model artefacts (Fig. 3.a). Therefore, although the phase identification is robust even with flawed data, the original data from which 3DAM are to be built needs to be high-quality data (Fig. 3.b). Intuitively, providing atomically precise outputs requires an also atomically precise flawless starting point that validates the fidelity of the output. Therefore, the automatic strain maps serve as a first validation stage to determine if the starting quality is enough to build a reliable 3DAM and digital twin, as they visually highlight imaging artefacts. Figure 3 showcases how uneven strain maps (ϵ_{xx} , ϵ_{yy}) with strong local modulations (black arrow) from a distorted STEM micrograph end up with defective 3DAMs characterised by local nonphysical accumulations of out-of-equilibrium atomic forces (see calculation details below) (Fig. 3.a). On the other hand, high-quality data, represented by smooth GPA maps, produces representative 3DAM with smooth atomic force maps (Fig. 3.b). This demonstrates that the calculation of atomic forces is a complementary verification of the quality of the 3DAM reconstruction. In this case, local peaks in the atom-wise forces field point inform us of imaging artefacts when a double-check with the local image quality and the strain maps themselves validates them. In the next section, we discuss additional approaches to ensure quantitative data quality for artefact-free 3DAM building.

Equivalently, these effects caused by imaging artefacts can be understood as the result of applying a custom strain field to the actual device to test its impact on the properties of a future device candidate (i.e., reverse engineering). For instance, a distortion such as the pointed by the black arrow in Figure 3 hypothetically induced by a mechanical indentation in the device. This approach is rooted in the experimental simulations generated by the proposed workflow and the concept of creating a reliable digital twin of the device under study.

Overall, these cases illustrate the workflow’s unprecedented capabilities to generate 3DAMs whose construction would be impractical manually due to the local atom-wise displacements. This is detailed with further examples supporting these claims in the supplementary information, section 2 “Proofs of the automated phase identification”.

Beyond providing a complete structural understanding of the heterostructure functionality, these 3DAMs open up the ground to infinite possibilities of unveiling new and previously inaccessible physics of the materials and devices.

2.8 Simulation of functional properties

Once the atomic model represents the experimental elastic displacements of the device, it opens up endless possibilities. We can simulate functional properties with a precision that is only achievable when considering the discrete atomic nature. We can validate the structures and perform quantitative (S)TEM analyses through immediate linear or multislice (S)TEM simulations of the resulting 3DAM (Fig. S6o). For instance, we present the results of a linear STEM image simulation of a Ge quantum well with varying lamella thicknesses, demonstrating the ease of performing quantitative checks such as focal series or varying depth of focus STEM analyses or finer evaluations of the effect of zero-point vibrations in micrographs (supplementary information, section 4.1 “STEM simulation details”).

The 3DAMs can also be used as inputs for simulation methods such as molecular dynamics or *ab initio* calculations that enable the computation of functional properties like electrical and thermal conductivity, electronic and phononic DOS and band gaps, dispersion relations, atomic forces or any property relevant to understanding intricate devices like quantum wells, hybrid quantum nanowires, catalytic particles, among many others. This is precisely what we have done with the particularly interesting $\text{Si}_{0.3}\text{Ge}_{0.7}/\text{Ge}/\text{Si}_{0.3}\text{Ge}_{0.7}$ quantum well reported throughout the article (Fig. 4.b and Fig. 5). [68] Its outstanding properties as a platform for allocating singlet-triplet spin qubits make it perfect for evaluating and envisioning the potential of having its representative 3DAMs digital twins to delve further into its physics.

We have first calculated the forces on each of the atoms of our 3DAMs. This can be done using a number of different methods, ranging from first-principles calculations such as Density Functional Theory (DFT) to empirical classical force fields. [86, 87, 88] DFT does not contain empirical parameters, and therefore does not require fitting to previous experimental or theoretical data, has a high predictive power, and provides very accurate forces. However, its computational cost is quite large, specially for systems with very large numbers of atoms like the 3DAMs considered here, which becomes unpractical for the purposes of this work. Nevertheless, reduced scaling DFT algorithms [89, 90] and access to massively parallel computers make these calculations feasible nowadays, and will be considered in future evolutions of our workflows. In the other extreme are empirical classical force fields, which have relatively simple functional forms as a function of the atomic positions, but are fitted to previously known data

and work well only for systems sufficiently similar to those where the fitting was done. The simple functional form makes them very inexpensive and capable of being applied to systems with very many atoms. Here we use a simple Keating model,[91] specifically developed for mixed SiGe systems [92] which provides a good balance between accuracy and computational cost. The computation of the atomic forces using the Keating force field model for systems of the order of 10⁵ atoms (such as as the ones presented here) takes only seconds in a desktop computer.

Fig. 3.a shows atomic forces computed using this Keating model, demonstrating its capability to identify artifacts in the construction of the 3DAM from the experimental STEM data. The unexpectedly large forces in the region highlighted with the black arrow originate from a scanning artifact in the image that is hard to identify in the 3DAM with the naked eye. Mapping its position back to the strain map image reveals local imperfections in the imaged atomic columns, such as a slightly off-axis orientation of the sample and a damaged lamella. Thus, the force calculations serve as a fast and inexpensive way to screen the models and to assess their quality: large forces likely indicate artifacts and faulty models.

In Figures 3.b and 5.b we show the forces obtained for a 3DAM obtained from high quality STEM data, showing that the forces on all the atoms are small and validating the model. Interesting information can be extracted from these results. In particular, very small forces are obtained for the atoms in the Ge-rich region of the quantum well, whereas the forces are larger in upper and lower Si-containing regions. This is due to the assumptions done in the construction of the 3DAM, where pure Ge is taken as a reference. When the model is built placing actual Si atoms in the Ge lattice to match the experimental position-dependent concentration, pairs of different species as first neighbors are expected to have different interatomic distances from the nominal Ge-Ge one, but this is not reflected in the model, thus producing forces in those regions where Si-Si or Ge-Si pairs occur. This seldom happens in the Ge-rich quantum well, as there are very few Si atoms, but is much more frequent in the buffer layers containing Si, therefore increasing the average atomic forces in those regions.

The computed atomic forces could serve to further refine the structural 3DAM model, producing more realistic interatomic distances in the regions which deviate from the pure Ge stoichiometry. This can be easily done by moving the atoms according to the forces until the total energy is minimized, either through a minimization algorithm such as conjugate gradients, or through molecular dynamics with force quenching. [93, 94] We have done this for the model shown in Figure 5b), producing a structure very similar to the original 3DAM, but with more realistic interatomic distances for each Si-Si, Si-Ge and Ge-Ge pairs (see Supplementary Information, section 4.3 “Atomic forces and relaxations - KEating model”). We have not used this refined model further in the calculations of other physical properties shown below, for simplicity, although more realistic structures would lead to more accurate properties.

The Keating model can further be used to obtain more information about the system. We can calculate the force constant matrix (second derivatives of

the energy with respect to the atomic positions), from which dynamical matrix and the harmonic vibrational frequencies and modes can be obtained. The computation of the force constants is marginally more costly than that of the forces, and can also be performed in very little time in desktop computers. Computing the vibrational modes from the force constant matrix, though, requires significantly more computational work, as it involves the diagonalisation of the dynamical matrix, which has a very large size for systems with many atoms as those considered here. However, we have developed an alternative algorithm based on Greens functions (described in detail in the supplementary information) which allows to obtain the vibrational Local Density of States (vLDOS) in selected regions of our 3DAMs with very little computational effort. Figure 4.b shows the vLDOS for the high-quality 3DAM described above, for different regions of the device. In particular, we calculate the vLDOS for columns of atoms at different locations along the variant axis (see (supplementary information, section 4.4 “Phonon DOS - Keating model”). Figure 4.b shows how the local chemistry modulates the acoustic and optical bands of the alloy. The vLDOS map shows a quantum well displaying close to bulk Ge modes (e.g., main optical peak around 37 meV, with very little signal of higher frequency modes from the presence of a few Si atoms in that region). In contrast, noticeable differences are observed in the acoustic and optical modes at the more Si-rich buffer zones, where clear Si-Ge and Si-Si stretching optical bands are observed at higher frequencies (around 55 and 65 meV, respectively). This knowledge can be very valuable to assess relevant physical properties of the device, such as conductivity and transport properties of quantum devices, or the figure of merit in thermoelectric devices. In addition, the spatial mapping of the vLDOS would enable a direct comparison with experimental STEM vibrational spectroscopy.

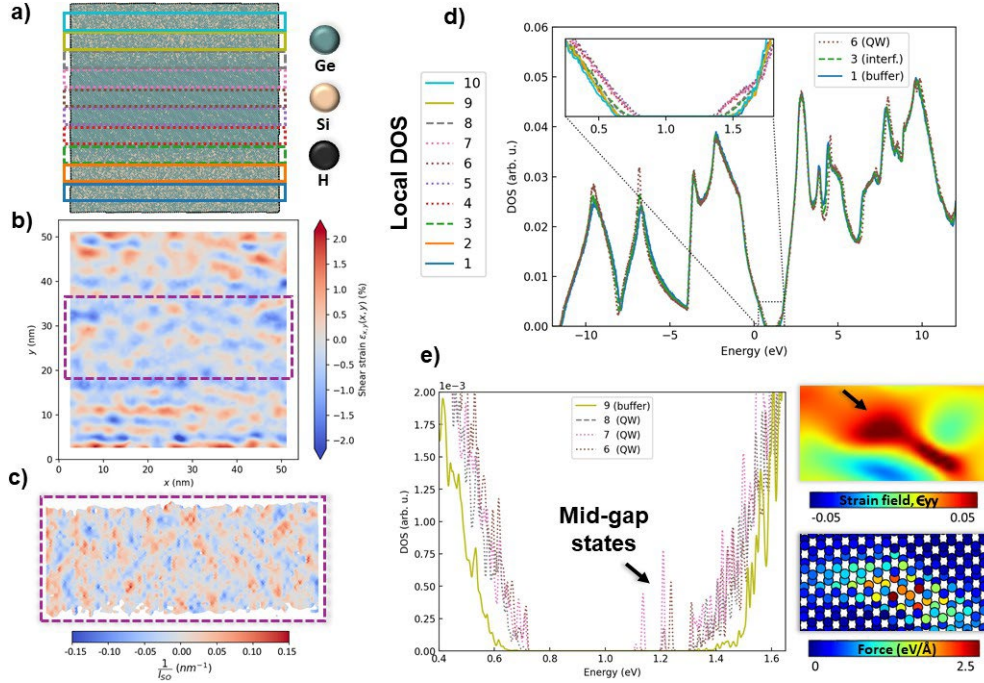


Figure 5: Local electronic density of states (DOS) of a $\text{Si}_{0.3}\text{Ge}_{0.7}/\text{Ge}/\text{Si}_{0.3}\text{Ge}_{0.7}$ through tight-binding atomic simulations. [68] a) Built upon the proposed workflow, the 3D atomic model (3DAM) used for the simulation is divided into boxes where the local DOS is computed. The 3DAM is passivated with H atoms to complete atomic bonding at its edges. b) The automation to compute the experimental shear strain permits us to map (c) crucial local properties for the spin qubits such as the inverse spin-orbit length at high spatial resolutions. d) Local electronic DOS profiling of the quantum well as depicted in panel a). The zoom-in showcases a clean band gap, displaying a valence band edge which is higher within the well and lower in the buffer layers, as expected. e) Zoomed-in near band gap electronic DOS for an atomic distortion centred in the quantum well. The distortion can be interpreted as the application of a custom strain accumulation indicated by the black arrow to the 3DAM digital twin of the Ge quantum well device (Fig. 3). The atomic distortion altering the regular lattice in the pm range raises mid-gap states potentially harmful to the transport and quantum properties of the device.

We now turn to exploring the electronic properties of our 3DAM. Again, for that purpose, methods ranging from first-principles DFT to empirical, simplified electronic Hamiltonians, can be used. As for the vibrational properties, here we chose to work with simplified and relatively inexpensive models instead of accurate and predictive but highly expensive DFT approaches. In particular,

we adopt an established sp^3ds^* tight-binding model for SiGe alloys, including the effects of strain and spin-orbit coupling. [95] This tight-binding model takes as an input the atomic positions, as specified by the 3DAM, and yields a Hamiltonian which describes the electronic properties of the system. Once this Hamiltonian is found, a variety of atomistic theoretical methods can be brought to bear to extract sample properties: sparse diagonalisation, non-equilibrium Green's functions (NEGF) for device simulation, or kernel polynomial methods (KPM). [96] The computational cost of the KPM scales linearly with the number of atoms in the system, making it tractable for very large sample sizes. We demonstrate this workflow capability here by a DOS calculation using the KPM and a study of spin-orbit coupling features related to the deformation fields experimentally obtained from the workflow.

For DOS calculations, we consider two different model cells, obtained from different 3DAMs. The first one is obtained from the purple dotted square in Fig. 3.a (which presents imaging artefacts) and contains about 800.000 atoms; while the second is extracted from the data of Fig. 3.b (high quality image) and contains about four million atoms (about 80 million spin-orbitals). In both cases, we compute averaged local densities of states (LDOS) in spatial regions profiling the 3DAM along the growth direction, to explore the corresponding change in electronic properties. For the large cell, these regions are shown in Fig. 5.a (for the smaller cell regions, see supporting information 4.5.2 “Results of the experimental/tight-binding/Keating correlation” and figures S67 and S68 therein).

In the smallest model cell (Fig. 3.a, dotted purple rectangle), we observe sharp DOS peaks within the electronic gap (Fig. 5.e), originating from regions inside the quantum well. These peaks stem from the imaging artefacts discussed above. Indeed, using sparse diagonalisation techniques, one can extract electronic wavefunctions at energies of interest. We do so for the in-gap peak at ~ 1.13 eV and find the resulting wavefunction to be strongly localized on the aforementioned imaging artefact, consistently with atomic forces calculations (Fig. 3.a, Fig. 5.e, and supporting information, section 4.5.2 “Results of the experimental/tight-binding/Keating correlation”). In other words, if a Ge quantum well experiences the depicted local strain peak or distortion (black arrow) in its centre, it may exhibit such localised mid-gap states.

The largest model cell is presented in Figure 5.a, showing the disordered distribution of Ge and Si atoms across the qubit device. We plot the LDOS over the whole spectral range of the system as well as a zoom-in near the band gap (Fig. 5.d and inset) showing the variability of the DOS along the 10 regions vertically profiling the 3DAM. Of particular interest to hole spin qubit physics is the behavior of the valence band edge, which can be seen in the inset. Three types of regions can be distinguished: the buffer regions (1, 2, 8, 10 – solid lines, $\text{Si}_{0.5}\text{Ge}_{0.7}$), the interface regions (3, 8 – dashed lines) and the quantum well regions (4, 5, 6, 7 – dotted lines, $\text{Si}_{0.05}\text{Ge}_{0.97}$), both in the atomic model and in their DOS signals. [68] It can be observed that the valence band edge is found at a higher energy inside the quantum well compared to the buffer region, while intermediate energies are likely to be found along the interface. This is

consistent with the expected band alignment and the quantum well’s purpose to provide a potential well for holes in the Ge portion. The conduction band edge is not directly relevant to the operation of this hole spin qubit, but is discussed in the supplementary material, section 4.5 “Electronic DOS - Tight-binding simulations”.

We should also highlight here the compatibility of the workflow with continuum methods, such as k-p methodologies, which also provide a powerful investigation tool for nanostructures. While lacking atomic accuracy, they are well developed and may offer an advantage in terms of shortening computation times and sometimes in the interpretability of the results. As a matter of illustration, it has recently been pointed out that inhomogeneities in strain and composition across the interfaces in planar SiGe quantum wells could influence which spin-orbit mechanisms are active in the system, and therefore available to drive the qubit. Better understanding and quantitative estimation of these inhomogeneities are therefore of crucial interest to the optimisation of qubit systems. [97, 98].

Figure 5.b provides the experimental map of absolute shear strain across the sample, which is seen here to be strongly inhomogeneous. Abadillo-Uriel and co-authors have shown that such inhomogeneities promote linear-in momentum spin-orbit interactions. They provide explicit corrections to the minimal k-p Luttinger-Kohn Hamiltonian for all strain-induced spin-orbit interactions as a function of Si and Ge Luttinger parameters, deformation potentials (which are tabulated), and derivatives of the strains, which can be evaluated for experimental structures using the proposed workflow.[97] By way of illustration, we display in Figure 5.c the inverse spin-orbit length which provides a metric exemplifying the strength of such effects in the quantum well part of the structure. Further details on the performed simulations and their conclusions can be found in the supplementary information, section 4 “Experimental simulation of key functional properties”.

2.9 Full workflow wrap up

The previous sections traced the workflow’s path from single electron micrographs to simulated properties. We have utilized a SiGe-based quantum device known for its exceptional performance in spin qubit computing to validate and demonstrate our findings, particularly due to its promising future potential for improvement and device integration. [68] The intermediate results of the workflow schematised in Figure 1 are presented for this device through figure 2.b (steps 1-4), Figure 3 (steps 5-7) and Figures 4.b and 5 (steps 7-8). They jointly offer a full perspective of the expected local and global outputs of the workflow. Furthermore, we have validated the workflow with additional device types and materials configurations representing a broad range of scenarios within materials science, including variations of SiGe quantum planar devices, quantum nanowires, and even nanoparticles for plasmonics. [66, 69, 99, 100, 101] This detailed description and discussion can be found in the supplementary information, section 3 “Proofs of the full workflow”.

A central part of the conducted research is about making the workflow as general as possible. This results in, to the best of our knowledge, the most general and comprehensive data analysis workflow available in the field. However, the room for improvement is huge. For instance, our unique 3DAMs are robust and representative when elastic strain mechanisms are present. However, they currently demand a manual adjustment in capturing strong plastic deformations such as stacking faults, which are only partially represented. As a result, the improvement can happen both in terms of the individual modules themselves, but especially in expanding the workflow by complementing it with further analytical modules that converge into a larger growing unit that progressively embraces more materials science situations, which grants it endless continuity and attention in time.

The workflow's immeasurable potential and its unveiling of physical knowledge in a theoretical exploratory manner go beyond the improvement of the workflow itself. It unlocks the optimisation of devices with promising perspectives beyond the state-of-the-art, such as the described SiGe heterostructures for spin qubits. We have demonstrated unprecedented experimental fidelity in 3D devices and heterostructures and the comprehensive understanding it can achieve in accessing functional properties affected by the experimental nuances of the actual built devices. For instance, the distortion depicted in Figures 3.a and 5.e can be regarded as applying a custom strain field to the device's digital twin to test its effects on the band structure. The answer is that a local lattice distortion in the pm range would be fatal for the transport and quantum properties of the device, which must guide the specifications of the next generation of devices. This pushes the state-of-the-art predicting power of digital twins unlocking hypothesis testing such as custom strain fields, compositional profiles or isotopic configurations for new device candidates. These steps towards risk reduction in the materials research cycle open a path towards reverse engineering and cost-effective device optimisation even when extremely subtle physics such as quantum phenomena are involved.

3 Conclusions

We have introduced an innovative end-to-end workflow that autonomously translates electron micrographs into 3D models suitable for theoretical analysis. This novel approach provides a rapid, accurate, and comprehensive structural description of imaged heterostructures and devices. Furthermore, it utilises this structural information to generate realistic models, either finite elements or atomic, that empower theoretical simulations aimed at extracting functional properties (stress fields and strain relaxation, forces, phonon and electronic DOS) in the final device configuration (digital twin). Our work represents a significant achievement with demonstrated reproducibility and validity over the examples and scenarios showcased throughout the main text and supplementary information. For instance, complete (S)TEM analysis which could take days of expert time can now be automatically done in a matter of minutes. The data

analysis modules take a few minutes to output results. On the other hand, the construction through local displacements of the 3DAMs is currently more time-consuming, taking a few hours, although it produces invaluable information.

Crucially, we believe that this workflow establishes a new paradigm in automated data analysis for experimental techniques. The underlying concept can be extended to various methods, with scanning probe and atomic force microscopies being prominent examples, especially for 2D materials. [8, 40, 102, 103, 104, 105, 106] We envision its potential widespread adoption within the microscopy community, both in academia and in industry. For instance, the semiconductor industry would thoroughly benefit from the whole robust workflow and its independent and flexible modules. This tool's automated, human-bias-free and user-friendly nature is crucial in overcoming the long-standing limitation of TEM: its difficulty in achieving statistical significance. That is, measuring the same properties in multiple devices or samples to minimise its statistical uncertainty, which requires the high-throughput unlocked by the workflow. Moreover, its prospects for combining it with ongoing developments in (S)TEM data analysis and automation for data acquisition are pivotal in transforming TEM into a high-throughput analytical technique, thus accelerating scientific discoveries. [43, 107]

The room for improvement, refinement and integrability is massive. Pre-cleaning steps like denoising, drift correction, or scanning artefacts correction could be easily added to enhance the robustness even further. [35, 108, 109, 110] Additionally, there is ample room to expand into other acquisition modes like 4D-STEM, or spectroscopy-related functionalities in both low and high-loss regimes. [111, 112, 113, 114, 115] Even extending these considerations into fully automating the last modules currently involving manual processing is feasible. [83, 84, 116, 117, 118] For instance, as an initial step, by grouping condition-specific simulation profiles to expedite the process. While these considerations extend beyond the present scope, they represent fascinating avenues for future research.

We want to end with a final concern about the human-AI interaction. We envision a general generative model capable of automatically retrieving anything from experimental devices: from acquisition to data analysis. We have set the first major cornerstone for it, and although this eventual model might be far off, we believe that humans should never fully abandon the intermediate steps of the workflow. Thus, the *a priori* advantage of automation could be limited in purpose to require a human cross-check. Its benefits would be twofold. On the one hand, certain steps such as the post-processing of the segmentation or the peak identification could be validated by a trained human to enrich the analysis. On the other hand, we raise a wall to the AI in its decision tree. In its current state, this concern is overkill, as the implied AI is not generative or online learning. However, reaching the envisioned point in a future version of the workflow could potentially entail these ethical issues. As a result, establishing strict control points or security checks from the very beginning is beneficial and does not hinder the original aim of fully revolutionising the electron microscopy analysis as we know it to date.

4 Acknowledgements

ICN2 acknowledges funding from Generalitat de Catalunya 2021SGR00457. This study is part of the Advanced Materials programme and was supported by MCIN with funding from European Union NextGenerationEU (PRTR-C17.I1) and by Generalitat de Catalunya. We acknowledge support from CSIC Interdisciplinary Thematic Platform (PTI+) on Quantum Technologies (PTI-QTEP+). This research work has been funded by the European Commission – NextGenerationEU (Regulation EU 2020/2094), through CSIC’s Quantum Technologies Platform (QTEP). ICN2 is supported by the Severo Ochoa program from Spanish MCIN / AEI (Grant No.: CEX2021-001214-S) and is funded by the CERCA Programme / Generalitat de Catalunya. Part of the present work has been performed in the framework of Universitat Autònoma de Barcelona Materials Science PhD program. IPH acknowledges funding from AGAUR-FI scholarship (2023FI-00268) Joan Oró of the Secretariat of Universities of the Generalitat of Catalonia and the European SocialPlus Fund. M.B. acknowledges support from SUR Generalitat de Catalunya and the EU Social Fund; project ref. 2020 FI 00103. This study was supported by EU HORIZON INFRA TECH 2022 project IMPRESS (Ref.: 101094299). Authors acknowledge the use of instrumentation as well as the technical advice provided by the Joint Electron Microscopy Center at ALBA (JEMCA). ICN2 acknowledges funding from Grant IU16-014206 (METCAM-FIB) funded by the European Union through the European Regional Development Fund (ERDF), with the support of the Ministry of Research and Universities, Generalitat de Catalunya. ICN2 is a founding member of e-DREAM.[119] SR is also supported by MICIN with European funds NextGenerationEU (PRTRC17.I1) funded by Generalitat de Catalunya and by 2021 SGR 00997. We thank the Catalan Quantum Academy for support. The authors acknowledge Dámaso Torres for his support in designing the graphical material.

5 Supplementary Information

Supplementary information is available under request to the corresponding authors.

6 Open data

The data and code used in the present work can be found in the following repository (a private GitHub repository available, to open to the public as soon as the manuscript is accepted).

7 Conflicts of interest

No conflicts of interest to declare.

8 Authors contributions

M.B. and J.A. conceived the idea and wrote the manuscript. M.B. led the conception of the workflow, all modules and their interconnection. M.B. and I.P-H developed the segmentation and peak finding modules. I.P-H developed the GUI. M.B., E.R. and V.G. developed the phase identification, GPA automation and 3DAM building. T.G. and S.R. led the tight-binding simulations. C.C. and P.O. led the Keating model-based simulations. P.H.K. and V.G. led the FEM simulations. M.C.S and E.R. led the STEM simulations. Y-M.N assisted with the tight-binding calculations. M.B.E. and G.M. assisted in the peak finding deep learning models. J.A, S.M.S and M.B. acquired the experimental (S)TEM data. G.I, A. C., G.K., G.S. and P.K. provided the materials platforms under study and the discussion on their functional properties and applications. All authors contributed to the manuscript and overall discussion.

References

- [1] R. Ferreira. “Introduction to Semiconductor Heterostructures”. In: *Semiconductor Modeling Techniques*. Ed. by Naci Balkan and Marie Xavier. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 1–17. isbn: 978-3-642-27512-8. doi: 10.1007/978-3-642-27512-8_1. url: https://doi.org/10.1007/978-3-642-27512-8_1.
- [2] Sandip Tiwari. *Semiconductor Physics: Principles, Theory and Nanoscale*. Oxford University Press, Sept. 2020. isbn: 9780198759867. doi: 10.1093/oso/9780198759867.001.0001. url: <https://doi.org/10.1093/oso/9780198759867.001.0001>.
- [3] G. Scappucci et al. “Crystalline materials for quantum computing: Semiconductor heterostructures and topological insulators exemplars”. In: *MRS Bulletin* 46.7 (July 2021), pp. 596–606. issn: 1938-1425. doi: 10.1557/s43577-021-00147-8. url: <https://doi.org/10.1557/s43577-021-00147-8>.
- [4] Eduardo J. S. Brás and Pedro Carlos de Barros Fernandes. “Miniaturization and microfluidic devices: an overview of basic concepts, fabrication techniques, and applications”. In: *Physical Sciences Reviews* 9.5 (2024), pp. 2009–2036. doi: doi:10.1515/psr-2022-0102. url: <https://doi.org/10.1515/psr-2022-0102>.
- [5] Paolo Di Barba and Slawomir Wiak. “Device Miniaturization Principles”. In: *MEMS: Field Models and Optimal Design*. Cham: Springer International Publishing, 2020, pp. 65–67. isbn: 978-3-030-21496-8. doi: 10.1007/978-3-030-21496-8_7. url: https://doi.org/10.1007/978-3-030-21496-8_7.
- [6] Henry H. Radamson et al. “Miniaturization of CMOS”. In: *Micromachines* 10.5 (2019). issn: 2072-666X. doi: 10.3390/mi10050293. url: <https://www.mdpi.com/2072-666X/10/5/293>.

-
- [7] Stefanos Mourdikoudis, Roger M. Pallares, and Nguyen T. K. Thanh. “Characterization techniques for nanoparticles: comparison and complementarity upon studying nanoparticle properties”. In: *Nanoscale* 10 (27 2018), pp. 12871–12934. doi: 10.1039/C8NR02278J. url: <http://dx.doi.org/10.1039/C8NR02278J>.
- [8] Rama K. Vasudevan et al. “Autonomous Experiments in Scanning Probe Microscopy and Spectroscopy: Choosing Where to Explore Polarization Dynamics in Ferroelectrics”. In: *ACS Nano* 15.7 (2021). PMID: 34228427, pp. 11253–11262. doi: 10.1021/acsnano.0c10239. eprint: <https://doi.org/10.1021/acsnano.0c10239>. url: <https://doi.org/10.1021/acsnano.0c10239>.
- [9] Cheng Chang et al. “Surface Functionalization of Surfactant-Free Particles: A Strategy to Tailor the Properties of Nanocomposites for Enhanced Thermoelectric Performance”. In: *Angewandte Chemie International Edition* 61.35 (2022), e202207002. doi: <https://doi.org/10.1002/anie.202207002>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/anie.202207002>. url: <https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.202207002>.
- [10] Ai Qin Wang, Jun Li, and Tao Zhang. “Heterogeneous single-atom catalysis”. In: *Nature Reviews Chemistry* 2.6 (June 2018), pp. 65–81. issn: 2397-3358. doi: 10.1038/s41570-018-0010-1. url: <https://doi.org/10.1038/s41570-018-0010-1>.
- [11] Maria Ibáñez et al. “Tuning Transport Properties in Thermoelectric Nanocomposites through Inorganic Ligands and Heterostructured Building Blocks”. In: *ACS Nano* 13.6 (2019). PMID: 31185159, pp. 6572–6580. doi: 10.1021/acsnano.9b00346. eprint: <https://doi.org/10.1021/acsnano.9b00346>. url: <https://doi.org/10.1021/acsnano.9b00346>.
- [12] Chen-Fu Chien, Hsuan-An Kuo, and Yun-Siang Lin. “Smart semiconductor manufacturing for pricing, demand planning, capacity portfolio and cost for sustainable supply chain management”. In: *International Journal of Logistics Research and Applications* 27.1 (2024), pp. 193–216. doi: 10.1080/13675567.2022.2076818. eprint: <https://doi.org/10.1080/13675567.2022.2076818>. url: <https://doi.org/10.1080/13675567.2022.2076818>.
- [13] Roser Fernández-Climent et al. “Highly Durable Nanoporous Cu₂-xS Films for Efficient Hydrogen Evolution Electrocatalysis under Mild pH Conditions”. In: *ACS Catalysis* 13.15 (2023), pp. 10457–10467. doi: 10.1021/acscatal.3c01673. eprint: <https://doi.org/10.1021/acscatal.3c01673>. url: <https://doi.org/10.1021/acscatal.3c01673>.
- [14] Sabbir A. Khan et al. “Epitaxially Driven Phase Selectivity of Sn in Hybrid Quantum Nanowires”. In: *ACS Nano* 17.12 (2023). PMID: 37317984, pp. 11794–11804. doi: 10.1021/acsnano.3c02733. eprint: <https://doi.org/10.1021/acsnano.3c02733>. url: <https://doi.org/10.1021/acsnano.3c02733>.

-
- [15] Marco Valentini et al. “Majorana-like Coulomb spectroscopy in the absence of zero-bias peaks”. In: *Nature* 612 (2022), pp. 442–447. doi: 10.1038 / s41586 - 022 - 05382 - w. eprint: <https://doi.org/10.1038/s41586-022-05382-w>.
- [16] Sara Martínez Sánchez et al. “Sub-nanometer mapping of strain-induced band structure variations in planar nanowire core-shell heterostructures”. In: *Nature Communications* 13 (2022), p. 4089. doi: /10.1038/s41467-022-31778-3. eprint: <https://doi.org/10.1038/s41467-022-31778-3>. url: <https://doi.org/10.1038/s41467-022-31778-3>.
- [17] Marc Botifoll, Ivan Pinto-Huguet, and Jordi Arbiol. “Machine learning in electron microscopy for advanced nanocharacterization: current developments, available tools and future outlook”. In: *Nanoscale Horiz.* 7 (12 2022), pp. 1427–1477. doi: 10.1039 / D2NH00377E. url: <http://dx.doi.org/10.1039/D2NH00377E>.
- [18] Sergei V. Kalinin et al. “Machine learning for automated experimentation in scanning transmission electron microscopy”. In: *npj Computational Materials* 9.1 (Dec. 2023), p. 227. issn: 2057-3960. doi: 10.1038/s41524-023-01142-0. url: <https://doi.org/10.1038/s41524-023-01142-0>.
- [19] Peter D. Nellist Stephen J. Pennycook. *Scanning Transmission Electron Microscopy*. Springer New York, 2011. isbn: 978-1-4419-7199-9. doi: /10.1007/978-1-4419-7200-2. eprint: <https://doi.org/10.1007/978-1-4419-7200-2>. url: <https://doi.org/10.1007/978-1-4419-7200-2>.
- [20] David B. Williams and C. Barry Carter. *Transmission Electron Microscopy*. 2nd ed. Springer New York, 2009. isbn: 978-0-387-76502-0.
- [21] David A. Muller. “Structure and bonding at the atomic scale by scanning transmission electron microscopy”. In: *Nature Materials* 8 (2009), pp. 263–270. doi: /10.1038/nmat2380. eprint: <https://doi.org/10.1038/nmat2380>. url: <https://doi.org/10.1038/nmat2380>.
- [22] Rama K. Vasudevan et al. “Materials science in the artificial intelligence age: High-throughput library generation, machine learning, and a pathway from correlations to the underpinning physics”. In: *MRS Communications* 9 (3 2019), pp. 821–838. doi: 10.1557/mrc.2019.95. eprint: <https://doi.org/10.1557/mrc.2019.95>. url: <https://doi.org/10.1557/mrc.2019.95>.
- [23] Sergei V. Kalinin et al. “Machine learning in scanning transmission electron microscopy”. In: *Nature Review Methods Primers* 2 (2022), p. 11. doi: /10.1038/s43586-022-00095-w. eprint: <https://doi.org/10.1038/s43586-022-00095-w>. url: <https://doi.org/10.1038/s43586-022-00095-w>.

-
- [24] A. Gloter et al. “Improving energy resolution of EELS spectra: an alternative to the monochromator solution”. In: *Ultramicroscopy* 96.3 (2003). Proceedings of the International Workshop on Strategies and Advances in Atomic Level Spectroscopy and Analysis, pp. 385–400. issn: 0304-3991. doi: [https://doi.org/10.1016/S0304-3991\(03\)00103-7](https://doi.org/10.1016/S0304-3991(03)00103-7). url: <https://www.sciencedirect.com/science/article/pii/S0304399103001037>.
- [25] Stephen Jesse and Sergei V Kalinin. “Principal component and spatial correlation analysis of spectroscopic-imaging data in scanning probe microscopy”. In: *Nanotechnology* 20.8 (Feb. 2009), p. 085714. doi: 10.1088/0957-4484/20/8/085714. url: <https://dx.doi.org/10.1088/0957-4484/20/8/085714>.
- [26] Pavel Potapov, Paolo Longo, and Eiji Okunishi. “Enhancement of noisy EDX HRSTEM spectrum-images by combination of filtering and PCA”. In: *Micron* 96 (2017), pp. 29–37. issn: 0968-4328. doi: <https://doi.org/10.1016/j.micron.2017.02.001>. url: <https://www.sciencedirect.com/science/article/pii/S0968432816303821>.
- [27] Lewys Jones et al. “Improving the SNR of Atomic Resolution STEM EELS and EDX Mapping while Reducing Beam-damage by using Non-rigid Spectrum-image Averaging”. In: *Microscopy and Microanalysis* 21.S3 (Sept. 2015), pp. 1215–1216. issn: 1431-9276. doi: 10.1017/S1431927615006868. eprint: <https://academic.oup.com/mam/article-pdf/21/S3/1215/48334791/mam1215.pdf>. url: <https://doi.org/10.1017/S1431927615006868>.
- [28] J.M.P. Nascimento and J.M.B. Dias. “Vertex component analysis: a fast algorithm to unmix hyperspectral data”. In: *IEEE Transactions on Geoscience and Remote Sensing* 43.4 (2005), pp. 898–910. doi: 10.1109/TGRS.2005.844293.
- [29] M. et al. Cao. “Machine Learning for Phase Retrieval from 4D-STEM Data”. In: *Microsc. Microanal.* 26 (2020), pp. 2020–2021.
- [30] C. Shi et al. “Rapid and Semi-Automated Analysis of 4D-STEM data via Unsupervised Learning”. In: *Microsc. Microanal.* 27 (2021), pp. 58–59.
- [31] Jacob Madsen et al. “A Deep Learning Approach to Identify Local Structures in Atomic-Resolution Transmission Electron Microscopy Images”. In: *Advanced Theory and Simulations* 1.8 (2018), p. 1800037. doi: <https://doi.org/10.1002/adts.201800037>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/adts.201800037>. url: <https://onlinelibrary.wiley.com/doi/abs/10.1002/adts.201800037>.
- [32] Kohei Aso et al. “Detection of picometer-order atomic displacements in drift-compensated HAADF-STEM images of gold nanorods”. In: *Microscopy* 65.5 (Sept. 2016), pp. 391–399. issn: 2050-5698. doi: 10.1093/jmicro/dfw018. eprint: <https://academic.oup.com/jmicro/article->

pdf/65/5/391/8038086/dfw018.pdf. url: <https://doi.org/10.1093/jmicro/dfw018>.

- [33] Maxim Ziatdinov et al. “Deep Learning of Atomically Resolved Scanning Transmission Electron Microscopy Images: Chemical Identification and Tracking Local Transformations”. In: *ACS Nano* 11.12 (2017). PMID: 29215876, pp. 12742–12752. doi: 10.1021/acsnano.7b07504. eprint: <https://doi.org/10.1021/acsnano.7b07504>. url: <https://doi.org/10.1021/acsnano.7b07504>.
- [34] Robert S. Pennington et al. “Neural-network-based depth-resolved multiscale structural optimization using density functional theory and electron diffraction data”. In: *Phys. Rev. B* 97 (2 Jan. 2018), p. 024112. doi: 10.1103/PhysRevB.97.024112. url: <https://link.aps.org/doi/10.1103/PhysRevB.97.024112>.
- [35] Feng Wang et al. “Noise2Atom: unsupervised denoising for scanning transmission electron microscopy images”. In: *Applied Microscopy* 50 (2020), p. 23. doi: /10.1186/s42649-020-00041-8. eprint: <https://doi.org/10.1186/s42649-020-00041-8>. url: <https://doi.org/10.1186/s42649-020-00041-8>.
- [36] Yuncheng Jin et al. “Machine learning guided rapid focusing with sensorless aberration corrections”. In: *Opt. Express* 26.23 (Nov. 2018), pp. 30162–30171. doi: 10.1364/OE.26.030162. url: <https://opg.optica.org/oe/abstract.cfm?URI=oe-26-23-30162>.
- [37] E. Rotunno et al. “Alignment of electron optical beam shaping elements using a convolutional neural network”. In: *Ultramicroscopy* 228 (2021), p. 113338. issn: 0304-3991. doi: <https://doi.org/10.1016/j.ultramic.2021.113338>. url: <https://www.sciencedirect.com/science/article/pii/S0304399121001224>.
- [38] Enzo Rotunno et al. “Convolutional neural network as a tool for automatic alignment of electron optical beam shaping devices”. In: *Microscopy and Microanalysis* 27.S1 (Aug. 2021), pp. 822–824. issn: 1431-9276. doi: 10.1017/S143192762100324X. eprint: <https://academic.oup.com/mam/article-pdf/27/S1/822/47965307/mam0822.pdf>. url: <https://doi.org/10.1017/S143192762100324X>.
- [39] Benjamin P. Cumming and Min Gu. “Direct determination of aberration functions in microscopy by an artificial neural network”. In: *Opt. Express* 28.10 (May 2020), pp. 14511–14521. doi: 10.1364/OE.390856. url: <https://opg.optica.org/oe/abstract.cfm?URI=oe-28-10-14511>.
- [40] Sergei V. Kalinin et al. “Automated and Autonomous Experiments in Electron and Scanning Probe Microscopy”. In: *ACS Nano* 15.8 (2021). PMID: 34269558, pp. 12604–12627. doi: 10.1021/acsnano.1c02104. eprint: <https://doi.org/10.1021/acsnano.1c02104>. url: <https://doi.org/10.1021/acsnano.1c02104>.

-
- [41] M.J. Hýtch, E. Snoeck, and R. Kilaas. “Quantitative measurement of displacement and strain fields from HREM micrographs”. In: *Ultramicroscopy* 74.3 (1998), pp. 131–146. issn: 0304-3991. doi: [https://doi.org/10.1016/S0304-3991\(98\)00035-7](https://doi.org/10.1016/S0304-3991(98)00035-7). url: <https://www.sciencedirect.com/science/article/pii/S0304399198000357>.
- [42] Martin J. Hýtch, Jean-Luc Putaux, and Jean-Michel P’ enisson. “Measurement of the displacement field of dislocations to 0.03 Å by electron microscopy”. In: *Nature* 423 (2003), pp. 270–273. doi: [/10.1038/nature01638](https://doi.org/10.1038/nature01638). eprint: <https://doi.org/10.1038/nature01638>. url: <https://doi.org/10.1038/nature01638>.
- [43] Suhas Somnath et al. *USID and Pycroscopy – Open frameworks for storing and analyzing spectroscopic and imaging data*. 2019. arXiv: 1903.09515 [physics.data-an].
- [44] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. “U-Net: Convolutional Networks for Biomedical Image Segmentation”. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. by Nassir Navab et al. Cham: Springer International Publishing, 2015, pp. 234–241. isbn: 978-3-319-24574-4.
- [45] Robbie Sadre et al. “Deep Learning Segmentation of Complex Features in Atomic-Resolution Phase-Contrast Transmission Electron Microscopy Images”. In: *Microscopy and Microanalysis* 27.4 (2021), pp. 804–814. doi: [10.1017/S1431927621000167](https://doi.org/10.1017/S1431927621000167).
- [46] Lingfeng Wang and Chunhong Pan. “Robust level set image segmentation via a local correntropy-based K-means clustering”. In: *Pattern Recognition* 47.5 (2014), pp. 1917–1925. issn: 0031-3203. doi: <https://doi.org/10.1016/j.patcog.2013.11.014>. url: <https://www.sciencedirect.com/science/article/pii/S0031320313005001>.
- [47] Toufiq Parag et al. “A context-aware delayed agglomeration framework for electron microscopy segmentation”. en. In: *PLoS One* 10.5 (May 2015), e0125825.
- [48] Randy M Sterbentz, Kristine L Haley, and Joshua O Island. “Universal image segmentation for optical identification of 2D materials”. en. In: *Sci. Rep.* 11.1 (Mar. 2021), p. 5808.
- [49] John Canny. “A Computational Approach to Edge Detection”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-8.6 (1986), pp. 679–698. doi: [10.1109/TPAMI.1986.4767851](https://doi.org/10.1109/TPAMI.1986.4767851).
- [50] Alexander Kirillov et al. *Segment Anything*. 2023. arXiv: 2304.02643 [cs.CV].
- [51] Zoltán Sylvester. *GitHub - zsyvester/segmenteverygrain: A SAM-based model for instance segmentation of images of grains*. Accessed: 2024-07-16. 2024. url: <https://github.com/zsyvester/segmenteverygrain>.

-
- [52] Rama K. Vasudevan et al. “Phases and Interfaces from Real Space Atomically Resolved Data: Physics-Based Deep Data Image Analysis”. In: *Nano Letters* 16.9 (2016). PMID: 27517608, pp. 5574–5581. doi: 10.1021/acs.nanolett.6b02130. eprint: <https://doi.org/10.1021/acs.nanolett.6b02130>. url: <https://doi.org/10.1021/acs.nanolett.6b02130>.
- [53] Artem Maksov et al. “Deep learning analysis of defect and phase evolution during electron beam-induced transformations in WS₂”. In: *npj Computational Materials* 5 (2019), p. 12. doi: /10.1038/s41524-019-0152-9. eprint: <https://doi.org/10.1038/s41524-019-0152-9>. url: <https://doi.org/10.1038/s41524-019-0152-9>.
- [54] R. Kannan et al. “Deep data analysis via physically constrained linear unmixing: universal framework, domain examples, and a community-wide platform”. In: *npj Computational Materials* 4 (2018), p. 6. doi: /10.1186/s40679-018-0055-8. eprint: <https://doi.org/10.1186/s40679-018-0055-8>. url: <https://doi.org/10.1186/s40679-018-0055-8>.
- [55] Magnus Nord et al. “Atomap - Automated Analysis of Atomic Resolution STEM Images”. In: *Microscopy and Microanalysis* 23.S1 (2017), pp. 426–427. doi: 10.1017/S1431927617002811.
- [56] Debangshu Mukherjee et al. “mpfit: a robust method for fitting atomic resolution images with multiple Gaussian peaks”. In: *Advanced Structural Chemical Imaging* 6.1 (2020), pp. 1199–1209. doi: /10.1186/s40679-020-0068-y. url: <https://doi.org/10.1186/s40679-020-0068-y>.
- [57] G. Bradski. “The OpenCV Library”. In: *Dr. Dobb’s Journal of Software Tools* (2000).
- [58] A. De Backer et al. “StatSTEM: An efficient approach for accurate and precise model-based quantification of atomic resolution electron microscopy images”. In: *Ultramicroscopy* 171 (2016), pp. 104–116. issn: 0304-3991. doi: <https://doi.org/10.1016/j.ultramic.2016.08.018>. url: <https://www.sciencedirect.com/science/article/pii/S0304399116301747>.
- [59] Debangshu Mukherjee and Raymond Unocic. “STEMTool: An Open Source Python Toolkit for Analyzing Electron Microscopy Datasets”. In: *Microscopy and Microanalysis* 26.S2 (2020), pp. 2960–2962. doi: 10.1017/S143192762002334X.
- [60] Vincenzo Grillo and Enzo Rotunno. “STEMCELL: A software tool for electron microscopy: Part I—simulations”. In: *Ultramicroscopy* 125 (2013), pp. 97–111. issn: 0304-3991. doi: <https://doi.org/10.1016/j.ultramic.2012.10.016>. url: <https://www.sciencedirect.com/science/article/pii/S0304399112002732>.

-
- [61] Vincenzo Grillo and Francesca Rossi. “STEMCELL: A software tool for electron microscopy. Part 2 analysis of crystalline materials”. In: *Ultramicroscopy* 125 (2013), pp. 112–129. issn: 0304-3991. doi: <https://doi.org/10.1016/j.ultramic.2012.10.009>. url: <https://www.sciencedirect.com/science/article/pii/S0304399112002665>.
- [62] S. et al. Gražulis. “Crystallography Open Database - An open-access collection of crystal structures”. In: *J. Appl. Crystallogr.* 42 (2009), pp. 726–729.
- [63] S. et al. Gražulis. “Crystallography Open Database (COD): An open-access collection of crystal structures and platform for world-wide collaboration”. In: *Nucleic Acids Res.* 40 (2012), pp. 420–427.
- [64] M. Hellenbrandt. “The inorganic crystal structure database (ICSD) - Present and future”. In: *Crystallogr. Rev.* 10 (2004), pp. 17–22.
- [65] A. D. Mighell and V. L. Karen. “NIST crystallographic databases for research and analysis”. In: *J. Res. Natl. Inst. Stand. Technol.* 101 (1996), pp. 273–280.
- [66] Pavel Aseev et al. “Ballistic InSb Nanowires and Networks via Metal-Sown Selective Area Growth”. In: *Nano Letters* 19.12 (2019). PMID: 31730748, pp. 9102–9111. doi: 10.1021/acs.nanolett.9b04265. eprint: <https://doi.org/10.1021/acs.nanolett.9b04265>. url: <https://doi.org/10.1021/acs.nanolett.9b04265>.
- [67] Pavel Aseev et al. “Selectivity Map for Molecular Beam Epitaxy of Advanced III–V Quantum Nanowire Networks”. In: *Nano Letters* 19.1 (2019). PMID: 30521341, pp. 218–227. doi: 10.1021/acs.nanolett.8b03733. eprint: <https://doi.org/10.1021/acs.nanolett.8b03733>. url: <https://doi.org/10.1021/acs.nanolett.8b03733>.
- [68] Daniel Jirovec et al. “A singlet-triplet hole spin qubit in planar Ge”. In: *Nature Materials* 20 (2021), pp. 1106–1112. doi: /10.1038/s41563-021-01022-2. eprint: <https://doi.org/10.1038/s41563-021-01022-2>. url: <https://doi.org/10.1038/s41563-021-01022-2>.
- [69] Wenhua Li et al. “CuTe Nanocrystals: Shape and Size Control, Plasmonic Properties, and Use as SERS Probes and Photothermal Agents”. In: *Journal of the American Chemical Society* 135.19 (2013). PMID: 23647089, pp. 7098–7101. doi: 10.1021/ja401428e. eprint: <https://doi.org/10.1021/ja401428e>. url: <https://doi.org/10.1021/ja401428e>.
- [70] Alberto Tosato et al. “Hard superconducting gap in germanium”. In: *Communications Materials* 4 (2023), p. 23. doi: /10.1038/s43246-023-00351-w. eprint: <https://doi.org/10.1038/s43246-023-00351-w>. url: <https://doi.org/10.1038/s43246-023-00351-w>.

-
- [71] Yuanyuan Zhu et al. “Interface lattice displacement measurement to 1pm by geometric phase analysis on aberration-corrected HAADF STEM images”. In: *Acta Materialia* 61.15 (2013), pp. 5646–5663. issn: 1359-6454. doi: <https://doi.org/10.1016/j.actamat.2013.06.006>. url: <https://www.sciencedirect.com/science/article/pii/S135964541300428X>.
- [72] Shizhe Feng and Zhiping Xu. “Strain Characterization in Two-Dimensional Crystals”. In: *Materials* 14.16 (2021). issn: 1996-1944. doi: 10.3390/ma14164460. url: <https://www.mdpi.com/1996-1944/14/16/4460>.
- [73] Y. Chandra et al. “Advances in finite element modelling of graphene and associated nanostructures”. In: *Materials Science and Engineering: R: Reports* 140 (2020), p. 100544. issn: 0927-796X. doi: <https://doi.org/10.1016/j.mser.2020.100544>. url: <https://www.sciencedirect.com/science/article/pii/S0927796X20300024>.
- [74] Jan Pomplun et al. “Adaptive finite element method for simulation of optical nano structures”. In: *physica status solidi (b)* 244.10 (2007), pp. 3419–3434. doi: <https://doi.org/10.1002/pssb.200743192>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/pssb.200743192>. url: <https://onlinelibrary.wiley.com/doi/abs/10.1002/pssb.200743192>.
- [75] P. Havu et al. “Finite-element implementation for electron transport in nanostructures”. In: *The Journal of Chemical Physics* 124.5 (Feb. 2006), p. 054707. issn: 0021-9606. doi: 10.1063/1.2162900. eprint: https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.2162900/15379852/054707_1_online.pdf. url: <https://doi.org/10.1063/1.2162900>.
- [76] COMSOL and AB. *COMSOL Multiphysics® v. 6.2*. Stockholm, Sweden. url: www.comsol.com.
- [77] Stefan Birner et al. “nextnano: General Purpose 3-D Simulations”. In: *IEEE Transactions on Electron Devices* 54.9 (2007), pp. 2137–2142. doi: 10.1109/TED.2007.902871.
- [78] H.G. Weller et al. “A Tensorial Approach to Computational Continuum Mechanics Using Object Orientated Techniques”. In: *Computers in Physics* 12 (Nov. 1998), pp. 620–631. doi: 10.1063/1.168744.
- [79] Mar’ia de la Mata et al. “The Role of Polarity in Nonplanar Semiconductor Nanostructures”. In: *Nano Letters* 19.6 (2019). PMID: 31039314, pp. 3396–3408. doi: 10.1021/acs.nanolett.9b00459. eprint: <https://doi.org/10.1021/acs.nanolett.9b00459>. url: <https://doi.org/10.1021/acs.nanolett.9b00459>.

-
- [80] Raquel Nafria et al. “Growth of Au–Pd₂Sn Nanorods via Galvanic Replacement and Their Catalytic Performance on Hydrogenation and Sonogashira Coupling Reactions”. In: *Langmuir* 34.36 (2018). PMID: 30096238, pp. 10634–10643. doi: 10.1021/acs.langmuir.8b02023. eprint: <https://doi.org/10.1021/acs.langmuir.8b02023>. url: <https://doi.org/10.1021/acs.langmuir.8b02023>.
- [81] Jacob Madsen and Toma Susi. “abTEM: ab Initio Transmission Electron Microscopy Image Simulation”. In: *Microscopy and Microanalysis* 26.S2 (2020), pp. 448–450. doi: 10.1017/S1431927620014701.
- [82] Toma Susi et al. “Efficient first principles simulation of electron scattering factors for transmission electron microscopy”. In: *Ultramicroscopy* 197 (2019), pp. 16–22. issn: 0304-3991. doi: <https://doi.org/10.1016/j.ultramic.2018.11.002>. url: <https://www.sciencedirect.com/science/article/pii/S0304399118302481>.
- [83] J. S. Smith, O. Isayev, and A. E. Roitberg. “ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost”. In: *Chem. Sci.* 8 (4 2017), pp. 3192–3203. doi: 10.1039/C6SC05720A. url: <http://dx.doi.org/10.1039/C6SC05720A>.
- [84] Nongnuch Artrith and Alexander Urban. “An implementation of artificial neural-network potentials for atomistic materials simulations: Performance for TiO₂”. In: *Computational Materials Science* 114 (2016), pp. 135–150. issn: 0927-0256. doi: <https://doi.org/10.1016/j.commatsci.2015.11.047>. url: <https://www.sciencedirect.com/science/article/pii/S0927025615007806>.
- [85] Ask Hjorth Larsen et al. “The atomic simulation environment—a Python library for working with atoms”. In: *Journal of Physics: Condensed Matter* 29.27 (June 2017), p. 273002. doi: 10.1088/1361-648X/aa680e. url: <https://dx.doi.org/10.1088/1361-648X/aa680e>.
- [86] P Hohenberg and W. Kohn. “Inhomogeneous Electron Gas”. In: *Phys. Rev.* 136 (1964), B864–B871. doi: 10.1103/PhysRev.136.B864. url: <https://doi.org/10.1103/PhysRev.136.B864>.
- [87] W. Kohn and L. Sham. “Self-Consistent Equations Including Exchange and Correlation Effects”. In: *Phys. Rev.* 140 (1965), A1133–A1138. doi: 10.1103/PhysRev.140.A1133. url: <https://doi.org/10.1103/PhysRev.140.A1133>.
- [88] Richard M. Martin. *Electronic Structure Basic Theory and Practical Methods*. Cambridge University Press, 2004. isbn: 0521782856. doi: 10.1017/CB09780511805769. url: <https://doi.org/10.1017/CB09780511805769>.
- [89] J.M. Soler et al. “The SIESTA method for ab initio order-N materials simulation”. In: *Journal of Physics: Condensed Matter* 14 (2020), p. 28745. doi: 10.1088/0953-8984/14/11/302. url: <https://iopscience.iop.org/article/10.1088/0953-8984/14/11/302>.

-
- [90] Alberto García et al. “Siesta: Recent developments and applications”. In: *The Journal of Chemical Physics* 152 (May 2020), p. 204108. issn: 0021-9606. doi: 10.1063/5.0005077. eprint: https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/5.0005077/16745439/204108_1_online.pdf. url: <https://doi.org/10.1063/5.0005077>.
- [91] P. N. Keating. “Effect of Invariance Requirements on the Elastic Strain Energy of Crystals with Application to the Diamond Structure”. In: *Phys. Rev.* 145 (2 May 1966), pp. 637–645. doi: 10.1103/PhysRev.145.637. url: <https://link.aps.org/doi/10.1103/PhysRev.145.637>.
- [92] S. Froyen, D. M. Wood, and A. Zunger. “Structural and electronic properties of epitaxial thin-layer Si_nGe_n superlattices”. In: *Phys. Rev. B* 37 (12 1988), pp. 6893–6907. doi: 10.1103/PhysRevB.37.6893. url: <https://doi.org/10.1103/PhysRevB.37.6893>.
- [93] William H. Press et al. *Numerical Recipes in Fortran 90: the Art of Scientific Computing*. Cambridge University Press, 1996. isbn: ISBN 0-521-57439-0.
- [94] Erik Bitzek et al. “Structural Relaxation Made Simple”. In: *Phys. Rev. Lett.* 97 (2006), p. 170201. doi: 10.1103/PhysRevLett.97.170201. url: <https://journals.aps.org/prl/abstract/10.1103/PhysRevLett.97.170201>.
- [95] Y. M. Niquet et al. “Onsite matrix elements of the tight-binding Hamiltonian of a strained crystal: Application to silicon, germanium, and their alloys”. In: *Phys. Rev. B* 79 (24 June 2009), p. 245201. doi: 10.1103/PhysRevB.79.245201. url: <https://link.aps.org/doi/10.1103/PhysRevB.79.245201>.
- [96] Zheyong Fan et al. “Linear scaling quantum transport methodologies”. In: *Physics Reports* 903 (2021), pp. 1–69.
- [97] José Carlos Abadillo-Uriel et al. “Hole-Spin Driving by Strain-Induced Spin-Orbit Interactions”. In: *Phys. Rev. Lett.* 131 (9 Sept. 2023), p. 097002. doi: 10.1103/PhysRevLett.131.097002. url: <https://link.aps.org/doi/10.1103/PhysRevLett.131.097002>.
- [98] Esteban A. Rodríguez-Mena et al. “Linear-in-momentum spin orbit interactions in planar Ge/GeSi heterostructures and spin qubits”. In: *Phys. Rev. B* 108 (20 Nov. 2023), p. 205416. doi: 10.1103/PhysRevB.108.205416. url: <https://link.aps.org/doi/10.1103/PhysRevB.108.205416>.
- [99] Brian Paquelet Wuetz et al. “Reducing charge noise in quantum dots by using thin silicon quantum wells”. In: *Nature Communications* 14 (2023), p. 1385. doi: /10.1038/s41467-023-36951-w. eprint: <https://doi.org/10.1038/s41467-023-36951-w>. url: <https://doi.org/10.1038/s41467-023-36951-w>.
- [100] Davide Degli Esposti et al. “Low disorder and high valley splitting in silicon”. In: *npj Quantum Information* 10.1 (Mar. 2024), p. 32.

-
- [101] Tom Willhammar et al. “Structure and vacancy distribution in copper telluride nanoparticles influence plasmonic activity in the near-infrared”. In: *Nature Communications* 8.1 (Mar. 2017), p. 14925. issn: 2041-1723. doi: 10.1038/ncomms14925. url: <https://doi.org/10.1038/ncomms14925>.
- [102] Kyle P. Kelley et al. “Fast Scanning Probe Microscopy via Machine Learning: Non-Rectangular Scans with Compressed Sensing and Gaussian Process Optimization”. In: *Small* 16.37 (2020), p. 2002878. doi: <https://doi.org/10.1002/sml.202002878>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/sml.202002878>. url: <https://onlinelibrary.wiley.com/doi/abs/10.1002/sml.202002878>.
- [103] Alexey G. Okunev et al. “Nanoparticle Recognition on Scanning Probe Microscopy Images Using Computer Vision and Deep Learning”. In: *Nanomaterials* 10.7 (2020). issn: 2079-4991. doi: 10.3390/nano10071285. url: <https://www.mdpi.com/2079-4991/10/7/1285>.
- [104] Maxim Ziatdinov et al. “Quantifying the Dynamics of Protein Self-Organization Using Deep Learning Analysis of Atomic Force Microscopy Data”. In: *Nano Letters* 21.1 (2021). PMID: 33306401, pp. 158–165. doi: 10.1021/acs.nanolett.0c03447. eprint: <https://doi.org/10.1021/acs.nanolett.0c03447>. url: <https://doi.org/10.1021/acs.nanolett.0c03447>.
- [105] Yue Liu et al. “General Resolution Enhancement Method in Atomic Force Microscopy Using Deep Learning”. In: *Advanced Theory and Simulations* 2.2 (2019), p. 1800137. doi: <https://doi.org/10.1002/adts.201800137>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/adts.201800137>. url: <https://onlinelibrary.wiley.com/doi/abs/10.1002/adts.201800137>.
- [106] Benjamin Alldritt et al. “Automated structure discovery in atomic force microscopy”. In: *Science Advances* 6.9 (2020), eaay6913. doi: 10.1126/sciadv.aay6913. eprint: <https://www.science.org/doi/pdf/10.1126/sciadv.aay6913>. url: <https://www.science.org/doi/abs/10.1126/sciadv.aay6913>.
- [107] Francisco de la Pena et al. “Electron Microscopy (Big and Small) Data Analysis With the Open Source Software Package HyperSpy”. In: *Microscopy and Microanalysis* 23.S1 (2017), pp. 214–215. doi: 10.1017/S1431927617001751.
- [108] Lewys Jones et al. “Smart Align - a new tool for robust non-rigid registration of scanning microscope data”. In: *Advanced Structural Chemical Imaging* 1 (2015), p. 8. doi: 10.1186/s40679-015-0008-4. eprint: <https://doi.org/10.1186/s40679-015-0008-4>. url: <https://doi.org/10.1186/s40679-015-0008-4>.

-
- [109] A. De Backer et al. “Atom counting in HAADF STEM using a statistical model-based approach: Methodology, possibilities, and inherent limitations”. In: *Ultramicroscopy* 134 (2013). 65th Birthdays of W Owen Saxton, David J Smith and Dirk Van Dyck / PICO 2013 From Multislice to Big Bang, pp. 23–33. issn: 0304-3991. doi: <https://doi.org/10.1016/j.ultramic.2013.05.003>. url: <https://www.sciencedirect.com/science/article/pii/S0304399113001228>.
- [110] Lewys Jones and Peter D. Nellist. “Identifying and Correcting Scan Noise and Drift in the Scanning Transmission Electron Microscope”. In: *Microscopy and Microanalysis* 19.4 (2013), pp. 1050–1060. doi: 10.1017/S1431927613001402.
- [111] David Rossouw et al. “A New Method for Determining the Composition of Core–Shell Nanoparticles via Dual-EDX+EELS Spectrum Imaging”. In: *Particle & Particle Systems Characterization* 33.10 (2016), pp. 749–755. doi: <https://doi.org/10.1002/ppsc.201600096>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/ppsc.201600096>. url: <https://onlinelibrary.wiley.com/doi/abs/10.1002/ppsc.201600096>.
- [112] J. Verbeeck, S. Van Aert, and G. Bertoni. “Model-based quantification of EELS spectra: Including the fine structure”. In: *Ultramicroscopy* 106.11 (2006). Proceedings of the International Workshop on Enhanced Data Generated by Electrons, pp. 976–980. issn: 0304-3991. doi: <https://doi.org/10.1016/j.ultramic.2006.05.006>. url: <https://www.sciencedirect.com/science/article/pii/S0304399106000945>.
- [113] J. Verbeeck and S. Van Aert. “Model based quantification of EELS spectra”. In: *Ultramicroscopy* 101.2 (2004), pp. 207–224. issn: 0304-3991. doi: <https://doi.org/10.1016/j.ultramic.2004.06.004>. url: <https://www.sciencedirect.com/science/article/pii/S0304399104001433>.
- [114] Jo Verbeeck et al. “Advancing EELS into an Unsupervised Quantification Method”. In: *Microscopy and Microanalysis* 29.Supplement 1 (July 2023), pp. 397–398. issn: 1431-9276. doi: 10.1093/micmic/ozad067.186. url: <https://doi.org/10.1093/micmic/ozad067.186>.
- [115] Sang-Hyeok Yang et al. “Deep Learning-Assisted Quantification of Atomic Dopants and Defects in 2D Materials”. In: *Advanced Science* 8.16 (2021), p. 2101099. doi: <https://doi.org/10.1002/advs.202101099>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/advs.202101099>. url: <https://onlinelibrary.wiley.com/doi/abs/10.1002/advs.202101099>.
- [116] Anton V. Sinititskiy and Vijay S. Pande. *Deep Neural Network Computes Electron Densities and Energies of a Large Set of Organic Molecules Faster than Density Functional Theory (DFT)*. 2018. arXiv: 1809.02723 [physics.chem-ph].

-
- [117] Jörg Behler. “First Principles Neural Network Potentials for Reactive Simulations of Large Molecular and Condensed Systems”. In: *Angewandte Chemie International Edition* 56.42 (2017), pp. 12828–12840. doi: <https://doi.org/10.1002/anie.201703114>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/anie.201703114>. url: <https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.201703114>.
- [118] Nongnuch Artrith, Tobias Morawietz, and Jörg Behler. “High-dimensional neural-network potentials for multicomponent systems: Applications to zinc oxide”. In: *Phys. Rev. B* 83 (15 Apr. 2011), p. 153101. doi: 10.1103/PhysRevB.83.153101. url: <https://link.aps.org/doi/10.1103/PhysRevB.83.153101>.
- [119] Regina Ciancio et al. “e-DREAM: the European Distributed Research Infrastructure for Advanced Electron Microscopy”. In: *Microscopy and Microanalysis* 28.S1 (Aug. 2022), pp. 2900–2902. issn: 1431-9276. doi: 10.1017/S1431927622010893. eprint: <https://academic.oup.com/mam/article-pdf/28/S1/2900/48822552/mam2900.pdf>. url: <https://doi.org/10.1017/S1431927622010893>.