

Information theory based crystallographic symmetry classifications of a noisy 2D periodic scanning tunneling microscope image

Peter Moeck¹, Andrew Dempsey¹, and Connor Shu¹

¹ Nano-Crystallography Group, Department of Physics, Portland State University, Portland, OR 97207-0751

Modern scanning tunneling microscopes (STMs) are often employed in the imaging of crystalline surfaces at atomic resolution, Fig. 1. In recent years, we adapted crystallographic image processing (CIP) techniques that originated in the field of electron crystallography some 50 years ago to the imaging with scanning probe microscopes [1-3]. In cases of images from STMs and atomic force microscopes, our CIP method allows for the removal of the effects of a blunt scanning probe tip [2,3].

Recent improvements to our method allow for fully objective, i.e. completely researcher independent, classifications of noisy 2D periodic images into Bravais lattice types [3,4], Laue classes [4], and plane symmetry groups [4]. The preconditions that need to be fulfilled for the application of our new information theory based method are a sufficiently long-range ordered crystalline material, a sufficient number of repeats of the unit cell in the image, and a relatively modest amount of generalized Gaussian noise [4,5]. Because discrete Fourier transforms are involved, noise in the images reduces to “noise per unit cell” with the square root of the number of unit cells that are processed [4]. Generalized noise arises from multiple sources in all imaging and image processing steps as well as from structural defects in the crystalline sample itself. None of these sources is allowed to dominate so that the central limit theorem applies and the sum total of all errors possesses approximately a Gaussian distribution [4,5]. The type of the microscope is not important; it is only important that its operation is reasonably stable and uncorrected systematic errors are small compared to the sum total of generalized random errors.

We applied our method to a STM image from a graphite sample that is openly available (in the on-line supporting material of [6]), Fig. 1. Utilizing sets of ratios of geometric-bias corrected squared residuals for the involved symmetry hierarchy branches, see right hand side of Fig. 1, the plane symmetry group of the selected region of the STM image was determined to be $h31m$ with confidence levels of approximately 26 % over $c11m$ (on the basis of the primitive sub-units of oc lattices [8] averaged over three settings) as well as 95 % over $h3$, respectively. There are, thus, no serious doubts that the prevailing translation symmetry of the STM-imaged graphite sample is that of the 2D hexagonal Bravais lattice type (as also obtained in [6]). Note in passing that the $h31m$ result is indicative of a rhombohedral layer stacking that may be due to the sample preparation. Note also that $h31m$ on a triple hexagonal unit cell features the same plane symmetries as $p3m1$ on a three times smaller cell [5], since the former is a minimal non-isomorphic supergroup of index 3 of the latter. Our classification enabling geometric-bias corrected residuals are in the form of geometric Akaike information criteria [9] as derived in [4]. Our confidence levels are based on the information content equation in [9] as derived in [3] and recently generalized for Laue class and plane symmetry group classifications.

Note also that the image in Fig. 1 was recently classified with respect to its Bravais lattice type by means of a machine learning (artificial intelligence, AI) system [6]. Since that AI system performed often better than chance for oblique lattices with translational pseudo-symmetry around a metric specialization [8] but misclassified a hexagonal lattice as being oblique (see, e.g. supplemental figure 5 in [6]), our information theory based classifications show that human intelligence and analytic ingenuity [3-5] combined with sound geometric statistics [9] represent a superior approach to crystallographic symmetry classifications at the present time. It is well known that the higher symmetric 2D Bravais lattice types are compatible with several plane symmetry groups and Laue classes so that our crystallographic symmetry classifications are in any case much more diagnostic than the AI results of [6].

Deep Convolutional Neural Network (DCNN) applications that deal with crystallographic symmetry classifications in two and three dimensions were recently discussed in [5] in conjunction with a review of the state of the art in classifications of crystallographic symmetries by a wide range of analytical approaches. The existence of pseudo-symmetries in noisy 2D periodic images present serious challenges to DCNNs as they are currently not trained on such images [7]. Our information theory based approach to crystallographic symmetry classifications [3-5] delivers, on the other hand, noise-level dependent classifications for pseudo-symmetric images [10] that can in turn be used to label such images for future training rounds of DCNNs.

References

- [1] P. Moeck, Proc. 11th IEEE Intern. Conf. on Nanotechnology, Portland, August 15-18, 2011, pp. 520–525.
- [2] J. C. Straton *et al*, Advanced Structural and Chemical Imaging **1** (2015) paper 14 (12 pages).
- [3] P. Moeck, In: Microscopy Book Series No. 7, pp. 503–514, 2017, <http://www.microscopy7.org/book/503-514.pdf>.
- [4] P. Moeck, Symmetry **10** (2018) paper 133 (46 pages), open access.
- [5] P. Moeck, arXiv:1902.04155, April 2019 (23 pages).
- [6] R. K. Vasudevan *et al*, npj Computational Materials **30** (2018) 1–9.
- [7] R. K. Vasudevan, private communication, the STM image is freely available as support material of [6].
- [8] P. Moeck and P. DeStefano, Advanced Structural and Chemical Imaging **4** (2018), paper 5 (33 pages).
- [9] K. Kanatani, Intern. J. Computer Vision **26** (1998) 171–189.
- [10] P. Moeck and A. Dempsey, this conference proceedings.

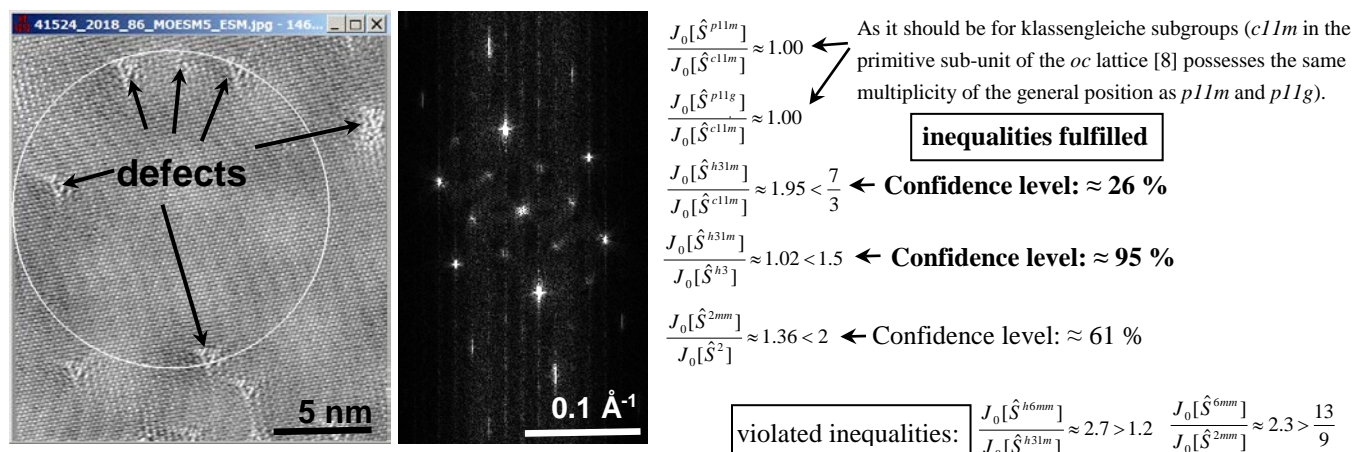


Figure 1. Left: Atomic resolution STM image of graphite [7], the top layer is somewhat rotated with respect to the layer below. The selected sample region consists of more than a thousand unit cells, each containing three very well resolved carbon atoms. This image is openly accessible under a creative commons license [6]. **Middle:** Central part of the amplitude map (power spectrum) of a discrete Fourier transform (dFT) that was calculated from the selected area to the left with the electron crystallography program CRISP [1,5,8]. Note the relative high amount of “1/spatial frequency noise” in this map, which is typical for raw STM images. The selected (slightly elliptical) image region was classified with respect to its plane symmetry group and Laue class. The fulfillment of a set of defining inequalities to the **right** of the dFT amplitude map identifies plane symmetry group $h31m$ as the one that minimize the expected Kullback-Leibler information loss [4]. Image symmetrization in this particular group by CIP in Fourier space provides, therefore, the best possible separation of structural information from generalized noise (at the derived six atoms unit cell and the prevailing generalized noise level of the selected image region). That particular symmetrization revealed indeed weak contrasts corresponding to three more carbon atoms per “surface-region” unit cell. Reference [6] shows this particular image as its figure 5d, but proclaims incorrectly [7] that it was taken from graphene.