Big Data and Machine Learning in Microscopy

MLM24

17-19 September 2024
Kanazawa, Japan

Programme

Click on a speaker's name to see the associated abstract and you can download a pdf with all the abstracts ***here***.

Tuesday 17.9

9:30 tutorial session (**Priante/Federici**)
12:00 *lunch/discussion*
Chair: Hofmann
13:00 ***Li***
14:00 ***Abe***
14:30 ***Kolev***
15:00 *coffee*
Chair: Chan
15:30 ***Huang***
16:00 ***Silveira***
16:30 ***Gao***
17:30 *end day 1*

Wednesday 18.9

Chair: Abe
9:00 ***Liu***
10:00 ***Sullivan***
10:30 *coffee*
Chair: Li
11:00 ***Hofmann***
11:30 ***Chan***
12:30 *lunch/discussion*
Chair: Gao
14:00 ***Mizoguchi***
15:00 ***Wu***
15:30 ***Flechsig***
16:00 *coffee*
Chair: Flechsig
16:30 ***Heath***
17:30 ***Qiu***
18:00 *end day 2*
18:45 *bus pickup for dinner*
19:00 *workshop dinner*
~22:00 *bus return*

Thursday 19.9

Chair: Tsuda
9:00 ***Ghiringhelli***
10:00 ***Kurki***
10:30 *coffee*
Chair: Ghiringhelli
11:00 ***Ren***
11:30 ***Tsuda***
12:30 *lunch/discussion*
Chair: Silveira
14:00 ***Wagner***
15:00 ***Stock***
15:30 *meeting end*

Enhancing Microscopy with Machine Learning: A Hands-on Tutorial

Fabio Priante [1] and **Filippo Federici**[2,3]

[1]Chemistry, Aalto University, Espoo, Finland

[2]Nanolayers Research Computing LTD

[3]Applied Physics, Aalto University, Espoo, Finland

This session offers a basic introduction to machine learning (ML) techniques for microscopy image analysis. First, essential ML concepts and model training procedures are outlined using PyTorch. The tutorial then moves to practical implementations, showcasing the use of a Convolutional Neural Network (CNN) for classification tasks, with high-resolution AFM images as an illustrative example. Subsequently, the common challenge of image segmentation is addressed. Considering the problem of identifying cell nuclei in optical microscopy images, a traditional computer vision approach is first attempted, and then compared with modern deep learning methods, particularly with the popular U-Net CNN architecture. The tutorial aims to equip participants with practical knowledge, that could be readily applied to their own microscopy problems.

An introduction to computational algorithms in Microscopy: From images to graphs

Li Jial[1], Su Ji[1], Lu Jion[1]

[1]National University of Singapore, Chemistry Department

In the rapidly evolving field of materials science, Scanning Tunneling Microscopy (STM) and Atomic Force Microscopy (AFM) generate vast amounts of data that require advanced computational techniques for effective analysis, including automation, interpretation, and prediction. This tutorial will introduce the latest computational algorithms designed to process and analyze microscopy data. We will start with image-tailored algorithms such as convolutional neural networks (CNNs), which are adept at handling continuous image data to extract meaningful features. Following this, we will explore graph neural networks (GNNs), which are suited for analyzing non-Euclidean graph data, such as molecules and crystals, enabling the extraction of atomic and bonding information at the microscopic level. This session will include practical demonstrations and hands-on examples to illustrate the implementation and benefits of these algorithms in real-world microscopy applications. By the end of the tutorial, attendees will gain a solid understanding of how to apply CNNs and GNNs (e.g., Chemprop) to enhance their analysis of STM and AFM data, ultimately advancing their research in materials science. This tutorial is intended for both novice and intermediate researchers who aim to leverage machine learning techniques in their microscopy studies.

AI-equipped Scanning Probe Microscopy for Self-driving Measurements at the Atomic Scale

Masayuki Abe[1], Zhuo Diao[1], Keiichi Ueda[1,2], Linfeng Hou[1], Fengxuan Li[1], Hayato Yamashita[1]

[1]Osaka University
[2]Tokyo Metropolitan Industrial Technology, Research Institute

We present an AI-equipped scanning probe microscopy system (AI-SPM), which is designed to autonomously conduct measurements at the atomic scale. This advanced system can autonomously target specific atomic positions and perform tasks with atomic precision such as the acquisition of spectroscopic data and atom manipulations. One of the key features of our AI-SPM is its inherent ability to identify and circumvent (or target) areas with surface imperfections. In addition, the AI-SPM system has been designed to effectively compensate for common problems such as drifting from the target position and atomic fluctuations at the tip apex. These implementations are critical to maintaining the fidelity of site-specific surface measurements and analysis. We have tested our AI-SPM system in the harsh conditions imposed by room temperature experiments, in which thermal drift and atomic fluctuations at the tip apex are ubiquitous. In room-temperature experiments on the Si(111)-(7x7) surface, the system was able to autonomously distinguish areas free of adsorbates and defects and then automatically acquire thousands of current-voltage (I-V) spectroscopy measurements over the four different adatom sites, while compensating the thermal drift and monitoring the probe apex condition in an autonomous experiment. Such experiments yield big data with the statistical significance necessary for the reliable characterization of materials, which confirms the capabilities of our AI-SPM implementation for enhanced data collection. The incorporation of AI into the SPM field paves the way for more efficient, accurate, and reliable surface analyses at the atomic level, potentially transforming our approach to material characterization.

Unsupervised and few-shot learning applied to scanning tunnelling microscope image segmentation

Nikola Kolev[1,2], Taylor Stock[1,2], David Gao[3, 4, 1], Filippo Federici[3,5], Emily Hoffman[1,2], Steven Schofield[1,2], Max Trouton[1,2], Geoff Thornton[1,2]

[1]University College London
[2]London Centre for Nanotechnology
[3]Nanolayers Research Computing LTD
[4]Norwegian Institute of Science and Technology
[5]Aalto University

Machine learning (ML) has demonstrated exceptionally high accuracies across various tasks, including image recognition, natural language processing, and gameplay. The field of material science has increasingly adopted these methods. However, a significant challenge remains: the most popular ML techniques are often supervised, requiring vast amounts of labelled data to train accurate models. This poses a difficulty when working with novel data, as is frequently the case in academia. One way to circumvent this is to use ML frameworks that require less data. In this work, we explore the utility of unsupervised learning and few-shot learning (FSL) for segmenting scanning tunnelling microscopy (STM) images to identify and classify both inherent surface defects, and adsorbates formed from surface chemical reactions. Unsupervised learning is employed to generate training data for a UNet, which identifies defect and adsorbate locations on surfaces. Subsequently, the FSL algorithm classifies these features. After initial training, FSL can classify previously unseen classes using only a few new labelled data points. FSL has been widely used in the ML community for various data types, such as images, audio, and radar, and has gained some attention in astronomy for differentiating galaxy types. However, its application in microscopy has been limited. We investigate its applicability by testing five different FSL algorithms on three different surfaces (Si(001):H, Ge(001), and TiO2(110)). This method demonstrates the potential for greater flexibility of ML algorithms in microscopy and a significant reduction in the required training data.

Enhancing AFM image analysis through machine learning with style translation and data augmentation

Jie Huang[1], Niko Oinonen[1, 2], Lauri Kurki[1], Adam S. Foster[1, 3]

[1]Department of Applied Physics, Aalto University, Finland
[2]Nanolayers Research Computing Ltd., London N12 0HL, United Kingdom
[3]WPI Nano Life Science Institute (WPI-NanoLSI), Kanazawa University, Japan

Atomic Force Microscopy (AFM) is critical for atomic-scale nanostructure characterization. Simulations, especially using Particle Probe AFM (PPAFM), provide a cost-effective means for rapid image generation. Leveraging state-of-the-art machine learning models and substantial PPAFM-generated datasets, properties like molecular structures, electrostatic force potential, and molecular graphs can be accurately predicted using AFM images from simulations or experiments. However, transferring model performance from PPAFM to real AFM images poses challenges due to the subtle variations in real experimental data compared to the seemingly flawless nature of simulations. Our study explores Cycle GANs for style translation to augment data and improve the predictive accuracy of machine learning models in surface property analysis. Focused on mitigating the gap between simulated PPAFM and authentic AFM images, we optimize hyperparameters, showcasing the method's effectiveness through paired data comparisons. This research promises valuable insights, providing a novel approach to enhance machine learning model efficiency in the absence of abundant experimental data.

Image interpretation methods for high-resolution tip-enhanced Raman spectroscopy images

Orlando J Silveira[1], Markus Junttila[1], Lauri Kurki[1], Shawulienu Kezilebieke[2], Peter Liljeroth[1], Adam S Foster[1,3]

[1]Department of Applied Physics, Aalto University, Aalto, Helsinki, Finland
[2]Department of Physics, Department of Chemistry and Nanoscience Center, University of Jyväskylä, Jyväskylä, Finland
[3]WPI Nano Life Science Institute (WPI-NanoLSI), Kanazawa University, Kanazawa, Japan

Tip-enhanced Raman spectroscopy (TERS) is a method capable of mapping vibrational modes of molecules in a unique way with Å-resolution, capturing not only the physical structure of a sample but also its chemical bonds [1]. Both effects, in combination with the Raman fingerprints of different chemical groups, can fully define the structural arrangement and the constituents of a single molecule. As TERS is promising for being a widely adopted method in materials science, a full understanding of the produced images and the physics behind the different contrasts is necessary. In this work, we propose image interpretation tools to extract physical information from TERS images using machine learning. Image analysis methods using machine learning have been already proven effective in scanning tunnelling microscopy (STM) [2] and atomic force microscopy (AFM) [3]. In here, we bring similar concepts to demonstrate the extraction of structural and chemical information directly from simulated TERS images. We also analysed different possibilities of simulating the TERS signal starting from density functional theory (DFT) calculation and investigated how they correspond to experimental images.

[1] Zhang, R., Zhang, Y., Dong, Z. et al. Chemical mapping of a single molecule by plasmon-enhanced Raman scattering. Nature 498, 82–86 (2013)

[2] Lauri Kurki, Niko Oinonen, and Adam S. Foster. ACS Nano 2024 18 (17), 11130-11138

[3] Oinonen, N., Kurki, L., Ilin, A., Foster, Adam S. MRS Bulletin 47, 895–905 (2022)

LabCore: Research Digitalization and Automation for Scanning Probe Microscopy

David Gao[1,2,3], Filippo Federici Canova[1,4], Nikola Kolev[3,5], Taylor Stock[3,5], Neil Curson[3,5]

[1]Nanolayers Research Computing LTD
[2]Norwegian Institute of Science and Technology
[3]University College London
[4]Aalto University
[5]London Centre for Nanotechnology

Modern research techniques have massively increased our ability to generate large quantities of both experimental and simulated data. This data can be leveraged with data science, machine learning, and automation techniques to create significant impact in industry. Scanning probe lithography for example is a promising manufacturing technique for microelectronic and quantum devices that require atomic scale precision. However, typical laboratory techniques are not suitable for industrial scale manufacturing and require a specialist to perform repetitive, tedious, and prohibitively time-consuming tasks. The LabCore digital research platform and ALANN graphical user interface are being developed to bridge this gap by combining data management and machine learning tools with instrument control and automation capabilities. Our software automation plug-in connects directly to the scanner and allows the user to control all relevant imaging parameters, acquire scans, and automatically save them in a convenient format. At the heart of the software lies integrated advanced image filtering routines that remove slope, spikes and creep from the data, compensate for small tip changes and ultimately construct a binary map of the detected atomic step edges. Defects in atomic scale images are automatically detected and identified using machine-learning methods and a custom correlation algorithm realigns the SPM images as they are acquired by stitching their step maps together to eliminate drift and aid in navigation. ALANN constructs a map of the entire sample that can be used for navigation, instrument control to write user specified patterns, and autonomously alignment of consecutive lithography stages to ensure precise results.

Navigating the Microscopic World: Synergy of Human Expertise, ML efficiency, and API Enabled Automation

Yongtao Liu[1]

[1]Oak Ridge National Laboratory

Microscopy has significantly advanced our understanding of structure-function relationships at the nanoscale, becoming a staple in characterization laboratories. However, traditional microscopy methods have largely been constrained by manual operations centered around human intervention. Therefore, we present the integration of application program interface (API) with machine learning to address these limitations. We developed AEcroscoPy, a cross-platform Python API designed to automate microscopy experiments, and showcase the combined power of human expertise, ML efficiency, and API-driven automation for accelerating scientific discovery. Our development of automated and autonomous experiment (AE) in scanning probe microscopy (SPM) facilitates the exploration of material functionalities and mechanisms. Using AE-SPM, we discovered coexistence and interplay of two ferroelectric subsystems in wurtzite ferroelectric thin films. Employing ML-driven approaches, we have probed ferroelectric materials to study phenomena such as domain wall dynamics and switching mechanisms, as well as the interactions between domain structures and local properties. By incorporating physical hypotheses in active learning model, our approach has enabled the microscope to autonomously discover the physical laws influencing domain switching. Although these methodologies were applied to specific materials, they possess broad potential to revolutionize various characterization techniques, including the assessment of stiffness and adhesion through force-distance curves in SPM.

Toward an Automation Framework for Scanning Probe Microscopy

Nicholas Sullivan[1], Peter Grutter[2], Kirk H. Bevan[1]

[1]Department of Mining and Materials Engineering, McGill University,
[2]Department of Physics, McGill University,

Scanning probe microscopy (SPM) is a valuable technique by which one can investigate the physical characteristics of the surfaces of materials. However, its ability to gain broader statistical information is hampered by the time-consuming nature of running an experiment, the significant domain knowledge necessary to do so, and the decision-making of the individual researcher guiding the device. Recent studies have shown the value of machine learning-based automation in assisting these, from guiding tasks that require constant user vigilance to redefining the experimental workflow around more statistics-driven decision-making. A remaining limitation is reusability: the experiment is often designed in a device-specific manner, making it non-trivial to integrate created automation into other researchers' experiments. We present an automation framework for SPM (afspm) with the aim of decoupling device specifics from the experiment and any automation components developed. This design limits device-specifics in the software to a single component, the 'microscope translator', which translates between device-specific and afspm-generic terminology. By further defining responsibilities between the components of an experiment, we hope to help researchers develop reusable components that can be easily shared within the community (e.g., 'tip detector/corrector'). The goal is to make it easier to develop and share automation components among the SPM community. We test the framework on two different SPM devices (a softdB/gxsm home-built controller, and an Asylum MFP-3D system), validating the reusability of individual components on devices with different interfaces and programming language requirements.

Automated structure discovery for scanning tunnelling microscopy

Lauri Kurki[1], Niko Oinonen[1,2], Adam S. Foster[1,3]

[1]Aalto University, Espoo, Finland
[2]Nanolayers Research Computing Ltd., London, UK
[3]WPI Nano Life Science Institute, Kanazawa, Japan

Scanning tunnelling microscopy (STM) and atomic force microscopy (AFM) functionalized with a CO molecule on the probe apex capture sub-molecular level detail of the electronic and physical structures of a sample from different perspectives [1]. However, the produced images are often difficult to interpret with respect to both physical and chemical structure. To accelerate the analysis, we propose automated machine learning image interpretation tools to extract sample properties directly from bond-resolved STM images [2]. In recent years, there has been rapid development in image analysis methods using machine learning with particular impact in medical imaging. These concepts have been proven effective also in SPM in general and especially for extracting sample properties from AFM images [3,4,5]. We build upon these models utilising convolutional neural networks for image analysis, and show that we can extract atomic positions directly from STM images. Finally, we establish the limits of the approach in an experimental context by predicting atomic structures from STM images of various small organic molecules. We also test the chemical sensitivity of the method by predicting chemical compositions of some organic molecules.

[1] Cai, S., Kurki, L., Xu, C., Foster, A. S., Liljeroth, P., J. Am. Chem. Soc. 144, 20227-20231 (2022).

[2] Kurki, L., Oinonen, N., Foster, A.S., ACS Nano, 18, 17, 11130–11138 (2024).

[3] Alldritt, B., Hapala, P., Oinonen, N., Urtev, F., Krejci, O., Canova, F. F., Kannala, J., Schulz, F., Liljeroth, P., Foster, A. S., Sci. Adv. 6, eaay6913 (2020).

[4] Carracedo-Cosme, J., Perez, R., npj Comput. Mater. 10, 19 (2024).

[5] Oinonen, N., Kurki, L., Ilin, A., Foster, A.S., MRS Bulletin 47, 895-905 (2022).

AI/ML and Data Infrastructure for Microscopy

Maria K Y Chan[1]

[1]Argonne National Laboratory

The combination of high throughput computational modeling and in situ/operando microscopy and spectroscopy experiments has given rise to significant opportunities to study an ever-increasing number of materials in ever-increasing detail. However, the integration of computational modeling with experimental measurements still often happens in a post-hoc and inefficient way. Data science techniques such as machine learning (ML), artificial intelligence (AI), and computer vision may be able to accelerate this integration. In this talk, we will discuss how we use AI/ML in conjunction with theory-based modeling to extract information in an automated fashion from microscopy data. We will describe our FANTASTX (Fully Automated Nanoscale To Atomistic Structures from Theory and eXperiment) [1] and Ingrained [2] codes which allow such integration, as well as EXSCLAIM! [3] and Plot2Spectra [4] codes for extracting microscopy and spectroscopy data from scientific literature. We will discuss the improvements to EXSCLAIM afforded by recent developments in large language models, as well as how data extracted from EXSCLAIM is used to aide tagging, data infrastructure, and retrieval of microscopy images. The use of these codes to discover the structure of novel materials [5] as well as understand technologically important energy materials [6] will also be discussed.

[1] D. Unruh, V. S. C. Kolluru, A. Baskaran, Y. Chen, and M. K. Y. Chan “Theory+AI/ML for Microscopy and Spectroscopy – Challenges and Opportunities”, MRS Bulletin 47, 1024–1035 (2022).

[2] E. Schwenker, V. S. Chaitanya Kolluru, J. Guo, X. Hu, Q. Li, M. C. Hersam, V. P. Dravid, R. F. Klie, J. R. Guest, M. K. Y. Chan, “Ingrained: an automated framework for fusing atomic-scale image simulations into experiments,” Small 18, 2102960 (2022).

[3] E. Schwenker, W. Jiang, T. Spreadbury, N. Ferrier, O. Cossairt, M. K. Y. Chan, “EXSCLAIM! -- Harnessing materials science literature for labeled microscopy datasets,” Patterns 4, 100843 (2023).

[4] W. Jiang, K. Li, T. Spreadbury, E. Schwenker, O. Cossiart, M. K. Y. Chan, “Plot2Spectra: an Automatic Spectra Extraction Tool,” Digital Discovery 1, 719-731 (2022).

[5] Q. Li, V. S. C. Kolluru, M. S. Rahn, E. Schwenker, S. Li, R. G. Hennig, P. Darancet, M. K. Y. Chan, M. C. Hersam, “Synthesis of borophane polymorphs through hydrogenation of borophene,” Science, 371(6534), 1143-1148 (2021).

[6] X. Liu, G.-L. Xu, K. V. S. Kolluru, et al, “Origin and regulation of oxygen redox instability in high-voltage battery cathodes,” Nature Energy 7, 808–817 (2022).

Unlocking the Potential of EELS and XAFS with Machine Learning

Teruyasu Mizoguchi[1]

[1]Institute of Industrial Science, The Unviersity of Tokyo

Electron Energy Loss Spectroscopy (EELS) and X-ray Absorption Spectroscopy (XAFS) observed using STEM and synchrotron facilities are now indispensable tools for characterizing functional materials due to their superior spatial and time resolutions, high sensitivity, and abundant information. In particular, the near-edge structure in the EELS/XAFS spectrum (ELNES/XANES) reflects the partial density of states of the conduction band at the excited state. Recently, the application of machine learning to EELS/XAFS has been reported. Our work aims to transcend the traditional physics of spectrum generation through machine learning. First, we constructed a spectral database comprising more than 100,000 Carbon-K edges of organic molecules [1]. Using this database, we achieved the prediction of extensive properties, such as molecular weight and internal energy, which are typically considered unrelated to ELNES features, via machine learning [2] Additionally, we have attempted to extract valence band information, similar to that obtained from XPS, from the ELNES/XANES features [3,4]. Furthermore, the same information as EXAFS was also obtained from ELNES/XANES features [5]. In my presentation, I will discuss the applications of machine learning in unlocking the potential of these core-loss spectroscopies.

[1] Shibata, K., Kikumasa, K., Kiyohara, S., & Mizoguchi, T. (2022). Scientific Data, 9, 214.

[2] Kikumasa, K., Kiyohara, S., Shibata, K., & Mizoguchi, T. (2022). Advanced Intelligent Systems, 4, 2100103.

[3] Chen, P.Y., et al. (2023). Journal of Physical Chemistry Letters, 14, 4858.

[4] Takahara, I., Shibata, K., & Mizoguchi, T. submitted.

[5] Kiyohara, S., and Mizoguchi, T., (2020), J. Phys. Soc. Jpn (Letter), 89, 103001.

Autonomous chemical reactions in scanning tunneling microscopy

Nian Wu[1], Markus Aapro[1], Alexander Ilin[2], Robert Drost[1], Joakim Jestilä[1], Nan Cao[1], Roman Fasel[3, 4], Peter Lijeroth[1] and Adam S. Foster[1,6]

[1]Applied Physics, Aalto University, Espoo, Finland
[2]Computer Science, Aalto University, Espoo, Finland
[3]Nanotech@surfaces Laboratory, Empa-Swiss Federal Laboratories for Materials Science and Technology, 8600 Dübendorf, Switzerland
[4]Department of Chemistry, Biochemistry and Pharmaceutical Sciences, University of Bern, 3012 Bern, Switzerland
[5]Departamento de Química Orgánica, Universidad Autónoma de Madrid, Campus de Cantoblanco, 28049 Ma-drid, Spain
[6]WPI Nano Life Science Institute, Kanazawa University, Kanazawa, Japan

Several breakthrough studies have harnessed scanning probe microscopy (SPM) manipulations to control chemical reactions in on-surface molecular synthesis. In general, for scanning tunnelling microscope (STM) manipulations, they are predominantly controlled via parameters of the tip position, pulse voltages and tunneling conductance. However, the selection of proper parameters requires extensive domain knowledge, which is time consuming and not necessarily transferable to new systems. Recent research has allowed the automation of a wide range of challenges in SPM, including image quality assessment, lateral and vertical manipulation. However, the automation for breaking or forming covalent bonds, which is an indispensable step during chemical synthesis is, as yet, unexplored. To address this problem, we developed our deep reinforcement learning (DRL) approach to automate bromine removal from 5,15-bis(4-bromo-2,6-methyl-phenyl)porphyrin (Br2Me4DPP) through learning manipulation parameters in STM. Meanwhile, DFT calculations were implemented to explore the reaction mechanism in combination with STM results and DRL results.

Machine learning applications towards automatized analysis and interpretation of HS-AFM imaging data

Holger Flechsig[1], Romain Amyot[1]

[1]WPI Nano Life Science Institute, Kanazawa University, Japan

While high-speed atomic force microscopy (HS-AFM) experiments play a leading role to observe functional dynamics of proteins, the interpretation of observations is generally difficult, mainly for two reasons: HS-AFM images have no atomistic resolution, and, automatized analysis of large imaging data sets acquired from experiments is practically absent. Both limitations can be overcome only by the methods provided by computational science. Our previous work including simulation AFM, development of fitting methods, and integrative modelling has significantly improved interpretation of HS-AFM observation. However, large-scale automatized analysis to fully exploit the explanatory power of HS-AFM observations requires application of machine learning methods. In this situation, synthetic AFM data computed from the enormous amount of high-resolution structural data of proteins (available from the Protein Data Bank and AlphaFold predictions) and atomistic modelling of functional dynamics is essential to infer atomistic information from resolution-limited imaging. In this talk we present our recently developed algorithms for GPU-accelerated calculation of simulated AFM imaging, which allows to efficiently generate a massive database of simulated AFM topographic images which can be correlated with measured HS-AFM images. We then present our attempts to employ synthetic AFM data of proteins in machine learning methods and demonstrate first applications to experimental HS-AFM data.

Towards seamless automated AFM movie analysis

George Heath[1], Emily Micklethwaite[1], Tabitha Storer[1]

[1]University of Leeds

Here we present open-source software developed for high-throughput automatic processing of Atomic Force Microscopy (AFM) image and movie data. AFM, High-Speed AFM, simulation AFM and Localization AFM (LAFM) enable the study of surfaces with increasingly higher spatiotemporal resolution. However, efficient and rapid analysis of the images and movies produced by these techniques can be challenging, often requiring the use of multiple image processing software applications and scripts. The combined co-localization and development of these tools will enable new analysis opportunities and improved resolution. Our software, NanoLocz, has been developed as an AFM and high-speed AFM analysis platform that facilitates import of raw data, automatic pre-processing and various single-particle analysis workflows designed to quickly leverage the rich data AFM has to offer. Workflows include but are not limited to: single-particle tracking, single-particle topographic feature analysis, single-molecule LAFM, time-resolved LAFM, and simulation LAFM.

HS-AFM single-molecule structural biology reveals the foundation of nuclear pore kinetics

Yujia Qiu, Shinnosuke Narimatsu, Elma Sakinatus Sajidah, Keesiang Lim, Richard W Wong

[1]Division of Nano Life Science, Graduate School of Frontier Science Initiative, Kanazawa University, Kanazawa 920-1192, Japan
[2]WPI Nano Life Science Institute (WPI-NanoLSI), Kanazawa University, Kanazawa 920-1192, Japan
[3]Cell-Bionomics Research Unit, Innovative Integrated Bio-Research Core, Institute for Frontier Science Initiative, Kanazawa University, Kanazawa 920-1192, Japan

Nuclear pore complexes (NPCs) on the nuclear membrane surface govern the flow of small and macromolecules between the cell nucleus and cytoplasm through their complicated core channel, which resembles a spiderweb. Few nanoscale technologies exist to reliably evaluate nuclear pore dynamics on nuclear membranes. Traditional optical imaging cannot resolve a variety of organelles and proteins in cells owing to diffraction. By allowing nanoscale subcellular inspection, super-resolution methods have overcome this restriction. However, fixed samples are generally needed for these approaches. This also raises the question of how well a static picture portrays intracellular dynamics. The unique technology of high-speed atomic force microscopy (HS-AFM) allows dynamic structural biology researchers to investigate molecules in motion near to their natural states. We use HS-AFM, which provides simultaneous structural and temporal resolution, to study protein kinetic mode switching. We created Y-shaped protein reconstitution for subcomplex and molecular analysis. With detection of local maxima, height extraction, merging of height modules from the localization atomic force microscopy (LAFM) and our novel JAVA coding, we could link structural states to a functional chronology and solve structures from a single molecule. Our novel modification approaches enable the development of more sophisticated time-resolved dynamic single-molecule structural methods for discovering structures using AFM.

Machine-learning autonomous molecular manipulations

Bernhard Ramsauer[1], Grant J Simpson[2], Leonhard Grill[2], **Oliver T Hofmann**[1]

[1]Institute of Solid State Physics, TU Graz, Austria
[2]Institute of Chemistry, Uni Graz, Austria

Finding the optimal parameters to manipulate individual molecules with STM is challenging and time consuming, even for human experts. This is particularly the case for larger molecules, where there molecule cannot be moved with arbitrary precision and the final location after a manipulation attempt is given by a probability distribution. Here, we present a combination of a reinforcement-leaning based exploration strategy with Bayesian statistics to find efficient protocols that allow bring molecules to arbitrary positions on the surface and to assemble them into functional nanostructures. Exemplarily, we build a “tic-tac-toe” board and a quantum corral. In both cases, the automated assembly achieves the assembly much faster than a human operator. Still, the remaining time bottleneck is the necessity to (re-)locate the molecule after each manipulation. We discuss options to overcome this via automated image recognition, neural networks to analyze the tunneling current during manipulation, and other search strategies.

Atomically Precise Fabrication using Machine Learning Assisted Scanning Tunneling Microscopy

Taylor Stock[1]

[1]University College London

Scanning Tunneling Microscopy Hydrogen Resist Lithography (STM-HRL) is a fabrication technique capable of producing electronic devices in silicon with true atomic precision. In this technique, the probe of an STM is used to pattern a single atomic layer of hydrogen on an atomically perfect silicon surface. This forms a chemically resistant mask for the atomically precise positioning of substitutional dopant atoms, delivered via gas phase molecular precursors. STM-HRL is the most precise semiconductor device fabrication technique available, and in principle, could be used to fabricate a solid-state universal quantum computer. The potential applications of STM-HRL thus hold great promise to deliver transformative technologies in the future, however the pathway to realization is challenging. Currently, device fabrication relies on skilled STM operators who handcraft individual devices, one at a time, in a labor-intensive process. By introducing machine learning assistance or replacement of the STM operator for tasks including complex real-time image analysis and dynamic instrument and process optimization, it should be possible to significantly increase the complexity, yield, and throughput of atomically precise device fabrication. In this work, we introduce the full process of STM-HRL fabrication and explore opportunities for the introduction of machine learning assistance. In particular, we explore machine learning image recognition as a means to identify and map distributions of dopant atoms and atomic-scale defects and to align these features within STM-HRL fabricated device structures.

Machine-driven Scanning Tunneling Microscopy Analysis and Rapid Surface Structure Construction

Pengju Ren[1,2], Xueqian Pan[2,3], Xiao-Dong Wen[1,2]

[1]Synfuels China Co. Ltd.
[2]Institute of coal chemistry, Chinese Academy of Sciences
[3]University of Chinese Academy of Sciences

A prerequisite for understanding the physical and chemical properties of the surface is detailed knowledge of the atomic structure of surfaces. Scanning tunneling microscopy (STM) is a high-resolution imaging technique for surfaces, which can obtain atomic level images of surfaces and assembly structures of molecules on substrates. Using complex theoretical models combined with electronic-structure calculations, increasingly efforts attempt to determine the surface atomic structure in experimental STM images. However, STM image analysis heavily rely on scientists' expertise and experience. Furthermore, high computational cost on solving electronic structures limits the extensive exploration of surface compositional and conformational space. Therefore, accurate and efficient analysis of experimental STM images remains an urgent challenge. Our work proposes a machine-driven approach for deciphering surface atomic structures from STM images. This approach can efficiently analyze experimental STM spectra and automatically recommend the most probable atomic structure models. Through the machine learning method driven by chemical knowledge and theoretical data, the extensive exploration of surface conformational space is accelerated. By effectively integrating multiple machine-driven filtering processes, the reliability and efficiency of analysis results are greatly improved. Taking complex iron carbides system as model system, this method deciphers the surface atomic structures which are most consistent with the experimental observation. This approach will facilitate the studies of the materials field and catalysis field related to surface structure.

Designing Materials with Generative AI

Koji Tsuda[1]

[1]University of Tokyo

It is increasingly common that machine learning is used together with experiments and simulations to discover new materials. Target materials include inorganic materials such as metals, ceramics and nanoparticles and organic molecules for drugs and photovoltaics. In this talk, I will introduce case studies of materials design for thermal emitters, fluorescent molecules, fluorescent proteins and antimicrobial peptides. I will also discuss about interpretable issues in materials design and automated experiments using robots.

Toward nanofabrication with molecular building blocks

Christian Wagner[1]

[1] Peter Grünberg Institut (PGI-3), Forschungszentrum Jülich, 52425 Jülich, Germany

The control of single-molecule junctions with a scanning probe microscope (SPM) not only directly benefits molecular electronics and molecular machines, but is also a stepping stone to the more ambitious goal of (supra)molecular nanofabrication. We use the versatile concept of two-contact manipulation, where the SPM tip actuates a molecule via a single chemical bond, while the surface provides a second (weaker) fixation. The biggest challenge of this approach is the lack of information about the atomic configuration of the molecule during manipulation. I present the components of an SPM-based single-molecule manipulation setup, specifically (1) a simulation with a machine-learned model of the tip-molecule-surface junction based on the message-passing neural network PaiNN [1] trained on DFT data, (2) a probabilistic search that compares experimental and simulated force gradient data to find the best molecular configuration estimate [2], and (3) an immersive virtual reality interface. A complementary approach to molecular nanofabrication is the use of an autonomous agent to replace the decision making of a human experimenter. I will describe such an agent, which is based on reinforcement learning principles and performs a nanofabrication task even in the face of large uncertainties, sparse feedback, and without further knowledge of single-molecule mechanics [3].

[1] K. Schütt et al., Proceedings of the 38th International Conference on Machine Learning, 9377 (2021) [2] J. Scheidt et al., J. Phys. Chem. C 127, 13817 (2023) [3] P. Leinen et al., Sci. Adv. 6, eabb6987 (2020)