UiT The Arctic University of Norway

Faculty of Science and Technology • Department of Physics and Technology

Quantitatively reinterpreting atomic force microscopy via the data science paradigm

Chia-Yun Lai

A dissertation for the degree of Philosophiae Doctor, November 2019

Acknowledgment

This thesis is a collective effort of many people. In my journey of pursuing this doctorate, I have been helped and encouraged. In the past few years, I have gained not only professional knowledge, but also the ability to practice critical thinking which could be arguably said as the most important skill in this information explosion era, especially when the media is no longer the agent impartially reporting events. I wanted to thank my supervisor Matteo Chiesa who has brought me to Norway to broaden my horizon, to see very different scenery, to experience another environment, and gives me advice whenever is needed. I also wanted to thank Sergio Santos who has taught me that people with knowledge should not be restricted within academic questions, and to think what the world needs. In addition, I wanted to thank my labmate for all the discussions and laugh we had, and this has made the time memorable. Of course, all of the achievements I got are built on the selfless love of my family. It is because I know they are always by my side so that I can do what I want to do without worrying anything else, and go further.

Finally, I want to thank the Ph. D defence committee for spending their valuable time reading this thesis.

Chia-Yun Lai

Tromsø, September 2019

Abstract

From the invention of the atomic force microscope (AFM) in 1986, tremendous efforts have been put into developing this tool. The AFM has long been considered as one of the top choices to probe the nanoscale world with the ability to achieve nanoscale resolution imaging of surfaces under different environments. Advances in instrumentation combined with the exploitation of sophisticated data analysis methodologies are set to meet the demand for higher resolution forms of microscopy that allow for direct visualization, identification, nanometric or atomic defects and structure, and material phases. This combination is not arbitrary but responds to the necessity of employing algorithms to decouple and interpret the complex signals and contrast channels that result from both standard instrumentation and the extra complexity added by the instrumentation designed to increase throughput and enhance resolution and quantification. Starting with interpreting AFM data using single mode force spectroscopy method to explicating multiple channels acquired with advanced multifrequency methods, it has reached a point that resorting to big data approaches might provide broader understanding toward surface properties in the material science community. Finally, this thesis shows that it is possible to submit the data capturing complex physical phenomena like the tip-surface interaction in AFM to a specific question and obtain the answer regardless of the complexity and/or unknown factors of the phenomena.

List of Papers

This thesis is based on the following papers:

- **Chia Y. Lai*,** Sergio Santos, Matteo Chiesa. **Machine Learning Assisted Quantification of Graphitic Surfaces Exposure to Defined Environments.** *Appl. Phys. Lett*, 2019.
- **Chia Y. Lai*,** Karen Sloyan*, Jin Y. Lu, Boulos Alfakes, Saeed Al Hassan, Ibraheem Almansouri, Marcus S. Dahlem, Matteo Chiesa. **Discerning the Contribution of Morphology and Chemistry in Wettability Studies.** *J. Phys. Chem. A*, 2018, 122.
- **Chia Y. Lai**, Sergio Santos, Carlo A Amadei, Karim R Gadelrab, Tzu-Chieh Tang, Albert Verdaguer, Victor Barcons, Josep Font, Jaime Colchero, Matteo Chiesa. **The Mendeleev-Meyer force project.** *Nanoscale*, 2016, 8.
- **Chia Y. Lai**, Saveri Perri, Sergio Santos, Ricardo Garcia, Matteo Chiesa. **Rapid quantitative chemical mapping of surfaces with sub-2nm resolution**. *Nanoscale*, 2016, 8.
- **Chia Y. Lai**, Sergio Santos, Matteo Chiesa. **Systematic multidimensional quantification of nanoscale systems from bimodal AFM data.** *ACS Nano, 2016.*
- **Chia Y. Lai**, Sergio Santos, Matteo Chiesa. **Reconstruction of height of sub-nm steps with bimodal atomic force microscopy.** *Nanotechnology*, 2016, 27.
- **Chia Y. Lai***, Tuza Olukan*, Sergio Santos, Amal Al Ghaferi, Matteo Chiesa. **The power laws of nanoscale forces in ambient conditions.** *Chem Comm*, 2015, 51.
- **Chia Y. Lai,** Monica Cozzolino, Maria Vittoria Diamanti, Matteo Chiesa. **Underling mechanism of time dependent surface properties of calcite ($CaCO_3$): a baseline for investigations of reservoirs wettability.** *J. Phys. Chem. C*, 2015, 119 (52).
- Matteo Chiesa, **Chia Y. Lai**. **AFM-Based Methodology for the Investigation of Surface Evolution**. *Phys Chem Chem Phys*, 2018, 29.
- Sergio Santos, **Chia Y. Lai,** Tuza Olukan, Matteo Chiesa. **Multifrequency AFM: from origins to convergence.** *Nanoscale*, 2017, 9.
- Corrado Garlisi, Gabriele Scandura, Giovanni Palmisano, Matteo Chiesa, **Chia Y. Lai**. **Integrated Nano-and Macroscale Investigation of Photoinduced Hydrophilicity in TiO2 Thin Films.** *Langmuir*, 2016, 45.
- Yun H. Chang*, Tuza Olukan*, **Chia Y. Lai**, Sergio Santos, Tze Y. Lin, Harry Apostoleris, Josep Font, Victor Barcons, Matteo Chiesa. **Establishing nanoscale heterogeneity with nanoscale force measurements**. *J. Phys. Chem. C*30T, 2015, 119 (32).

List of Tables

List of Figures

Contents

Chapter 1

Overview

This thesis is arranged in the following manner. Chapter 2 provides the background introduction to the importance of surfaces and interfaces, the reason why to investigate materials surfaces properties, and some common surface characterization techniques. In particular, a brief development of atomic force microscope (AFM) is presented. In Chapter 3, the methodologies used in this thesis are reported. These methodologies include the force reconstruction and Hamaker coefficient determination. Within the force reconstruction section, two factors, *i.e.* the assumption of the power law, and the statistics applied for data analysis, that could affect the validation of this method are discussed. Then, in Chapter 4, 3 case studies to test the methodologies are presented in the structure of motivation, sample preparation, and results and discussions sections. Next, a discussion of applying computer science assisted approach for analyzing AFM data is showed in Chapter 5. An example for this application is shown in Chapter 6. Finally, Chapter 7 provides the possible future works and the summary of this thesis.

This thesis is based on the publications listed in the List of Papers section.

Chapter 2

Background

2.1 Surfaces and interfaces

Solid surfaces, being discontinuities from one material to another, exhibit unique properties different from the bulk[1]. As a new surface is created from a solid material, either surface relaxation occurs that atoms closed to the surface exhibit expanded interlayer separation compared to the bulk, or surface reconstruction arises in which the lateral distances between surface atoms change. The surface of a material is the part that interacting with the surroundings. All interactions (both physical or chemical interactions) take places on the surfaces, such as all chemical reactions, catalysis, corrosion, adhesion, *etc*. Furthermore, as exposing to ambient conditions, a solid surface would be covered with (either physically or chemically adsorbed) contaminants such as gases, water vapor, and hydrocarbons, which exist on every surface of any solid matter. This affects the way of how we perceive a material.

Since almost all the industrial processes like oxidation, chemical activity, deformation and fracture, bonding, friction, lubrication, *etc*., involve or depend on the surface properties of the material, it is necessary and essential to tell whether a surface possess the desired properties. In the field of surface

analysis, surface characterization techniques encompass physical, chemical, mechanical, structural, thermal, optical, and electrical aspects, making it a cross-disciplinary area.

2.2 Characterization techniques

Most of surface characterization techniques involve a probe interacting with the material. This probe can be electrons, photons (light), x-rays, neutral species, ion beams or physical cantilevers. In some techniques, a probe beam strikes the material, and the information extracted either from the changes of the material or from the probes are used for analysis. As for mechanical techniques, mechanical contacts between a probe and a material surface are required. In this thesis, atomic force microscopy (AFM) is focused and methodologies are developed based on this instrument to study the material surface properties. A brief introduction and the reasons behind using AFM is present in the following session.

- Atomic Force Microscopy (AFM)

The AFM was presented to the community in 1986[2], and the same year in which a Nobel Prize went to the invention of its sister instrument, the Scanning Tunnelling Microscope. In 1986, the general idea of AFM was to employ some form of physical phenomena that would allow imaging surfaces in different environments[2]. To achieving this, a micro-cantilever with a sharp probe at its end is brought into proximity with a sample's surface and the emerging atomic forces between the tip and the sample are monitored. Dealing with this sample perturbations and with the wide range of acting forces at the tip-sample junction has recurrently led the development of the instrument.

An example of an AFM tip imaging a molecule adsorbed onto a mica's surface in the static or quasi-static mode is illustrated in Figure 2.1. In this mode, by exploiting Hooke's law, the constant applied force is applied. As the tip makes mechanical contact with the surface and raster scans over

the sample, lateral and normal forces can deform, cut and/or displace the molecule. In the dynamic modes, deformation and lateral perturbations can be minimized[3] (Figure 2.2), while improving resolution, sensitivity and quantification of the different atomic forces and phenomena.

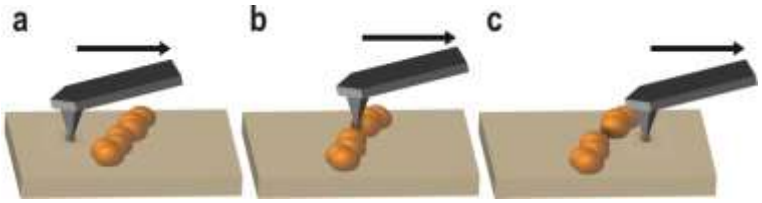

Figure 2.1 Illustration of static AFM imaging. An AFM tip, operated in static or quasi-static mode, imaging a molecule before (a), during (b) and after (c) contacting the sample.

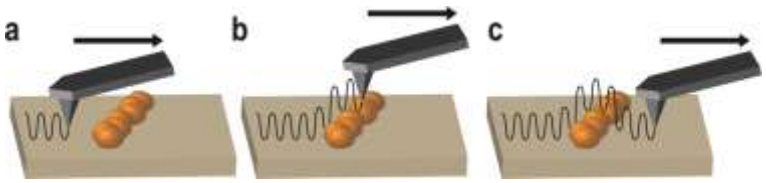

Figure 2.2 Illustration of dynamic AFM imaging. An AFM tip, operated in the dynamic mode, imaging a molecule before (a), during (b) and after (c) contacting the sample.

Later, two schools of dynamic AFM had differentiated by feedback mechanism: the frequency modulation FM AFM[4], and the amplitude modulation AM AFM[3]. Controlling a single frequency was manageable and quantification of dissipative and conservative interactions was possible in both AM[5] and FM[6] in the late nineties and early two thousand. At around this time, it became increasingly obvious[7-8] that ignoring higher frequencies implied ignoring information about the atomic processes. From an energy point of view, it could be said that energy from the frequency at which we excite disperses to higher frequencies because of the non-linear impact. An illustration of this phenomenon is shown in Figure 2.3. Such relationships were also rapidly recognized[8] that one had to "hammer"[9] the sample in order to increase the signal to noise ratio of the higher frequency components. This turn of events seemed worrisome, since one of the big goals consisted in decreasing the interaction in order to image soft matter[10-12]. Garcia *et al*. provided a solution to this dilemma in 2004 with what

is considered the origin of multifrequency AFM in the AM mode. Their solution implied directly and externally exciting higher frequencies before the interaction even occurred (Figure 2.3). They found that these higher frequencies were now sensitive and accessible without even mechanically contacting samples and were therefore suitable for gentle imaging[13]. In an important contribution to the theory, the FM community proposed that the higher driven frequency was influenced by the average derivative of the tip-sample force[14-15] (Figure 2.3). The proposed expression was rapidly adopted by some of the original proponents of multifrequency AFM[16] to directly map material properties such as the effective Young modulus of proteins[16]. This work further led to close form solutions in FM AFM for the effective Young Modulus, sample-deformation and viscosity[17].

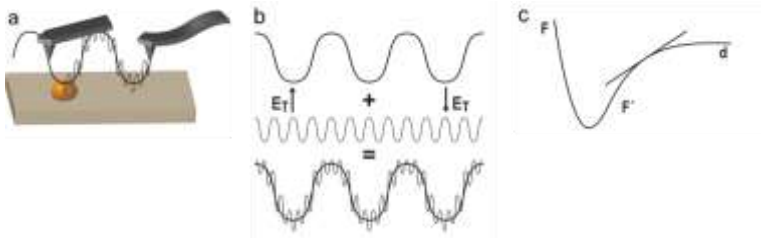

Figure 2.2 Illustration of the contrast mechanism behind multifrequency AFM. a, Imaging with a single drive frequency (standard or monomodal dynamic AFM) and with large oscillation amplitudes excites higher frequency signals. b, Externally exciting two frequencies (multifrequency AFM). c, The higher driven frequency in multifrequency AFM is influenced by the derivative of the tip-sample interaction F'.

Meanwhile extended forms of multifrequency[18-19] and methods[20-22] to exploit higher harmonic responses have been proposed by others. The overall result could be summarized as the availability of multiple observables, signals, expressions, and contrast channels in a myriad of driving scenarios that currently allow imaging with small and large amplitudes in the liquid, air and vacuum environments. The outcome is somehow reminiscent of the so-called zoo of particles that emerged in particle physics. Only, in this case, it was the zoo of contrast maps that emerged. It is almost like the community was driving the field into the big data era, something that had otherwise not escaped the notice of some pioneers[23].

2.3 State of the art machine learning in AFM

Exploiting computer science techniques to assist the data analysis involved in other fields, such as biology and the social sciences, has been routine for several years now. In materials science this phenomenon is much more recent and is only starting to materialise. Optimized computational methods are particularly relevant when dealing with data sufficiently complex that the simpler statistical or analytical methods become human "unreadable". With this we mean that there are now standard machine learning techniques, particularly well-suited to imaging and image interpretation, and that these may facilitate human understanding when multiple parameters, or higher dimensional analysis, makes direct interpretation complex. Sometimes, analytic expression, for which fundamental laws or principles are unavailable or overly complex. More and more research groups starting to adopt the approach of incorporating machine learning methods in the design of the project. Huang et al.[24] employed a support vector machine algorithm in AFM images for pattern recognition, feature identification without human interference. Sokolov et al.[25] used three different machine learning algorithms to identify cancer cells from normal cells with images obtained from height and adhesion channels. Mencattini et al.[26] studied cell-nanoparticle interactions with the help of two types of classification algorithms: support vector machine and linear discriminant analysis. Most of the applications are image-based methods, and very few attempts have been made to use forces reconstructed from AFM measurements. One of the reasons being that imaging is an easier technique for users to get hold of, yet there are advantages to develop force-based method. While recording forces with AFM measurements correctly could be very laborious, it provides availability of more parameters/features that could be used into machine learning techniques provided the phenomena is richer than the simpler models used in imaging. Second, with the many attempts by the community to ever increasing the number of functionally independent number of experimental observables, let that be via enhancing instrumentation sensitivity or extracting information via probing the force through different interaction regimes[27], the interpretation of the imaging channels can be assisted by the increasingly advancing imaging extraction and interpretation techniques in machine learning.

Chapter 3

Parameters and methodology

3.1 Force

In addition to using AFM to scan over samples for topographical investigation, it's also exploited as a force spectroscopy wherein on reconstructs the nanoscale force profile from experimental observables to recover the force as a function of tip-sample separation distance (d)[28-30]. While in the contact mode of operation, one could say it is the simplest way to obtain force profiles, yet the noise or jump to contact phenomenon might leave the results meaningless[31-32]. In dynamic AFM, the integral form of the equation of motion (first derived in 1997 by exploiting the Hamilton Jacobi formalism[33]) is typically derived by assuming Euler–Bernoulli beam theory[34]. It can be reduced to a standard driven harmonic oscillator with damping and the addition of the tip-sample force that introduces the non-linearities. Several authors have inverted the integral equation of motion by exploiting transforms including the Laplace transform leading to the modified Bessel function of the first kind and allowing

reconstruction via a Pade approximant representation[35] (with error below 5%), the Legendre transform[36] and the Fourier transform[37].

Here, the method used to reconstruct the force-distance curves (FDCs) is the Sader-Jarvis-Katan formalism[35, 38-39], which is a derivation of the results obtained via the Laplace transform and show that solutions are equivalent in AM and FM[38]. The FDCs are reconstructed (both conservative and dissipative) by considering variations in cantilever amplitude (A) and phase (P) as a function of variations in separation distance *d*. The conservative tip-sample force F can be written as:

$$F = 2k\int_{u=d}^{u=\infty}\left[\left(1+\frac{A^{\frac{1}{2}}(u)}{8\sqrt{\pi(u-d)}}\right)\Omega(u)-\frac{A^{\frac{3}{2}}(u)}{\sqrt{2(u-d)}}\frac{d\Omega(u)}{du}\right]du \qquad \textit{Eq. 1}$$

where u is the variable of separation distance, A is the amplitude of tip oscillation, k is the spring constant, and the phase Ω is the normalized frequency shift. Ω is expressed by[40-41]:

$$\Omega(d) = \left[1+\frac{A_0}{QA}cos(\Phi(d))\right]^{1/2} - 1 \qquad \textit{Eq. 2}$$

where A_0 is the free amplitude of tip oscillation, Q is the quality factor, and Φ is the phase lag relative to the driving force. Noted that A_0 is a key parameter to achieve a smooth transition to the repulsive regime, i.e., avoiding bistability and discontinuity in the amplitude-phase-distance curves[42-43]. Usually, A_0 is set to 3 times higher than the critical amplitude A_c value[44-46], which is the minimum free amplitude A_0 required to reach the repulsive regime. Eq. 1 and Eq. 2 has been implemented in Matlab with standard functions and the resulting raw force F has been smoothened with the standard rloess method (moving average filter of 30) from Matlab prior to calculating F_{AD}. The speed of acquisition was limited by the AFM, *i.e*. one amplitude and phase curve ~1 second.

The absolute value of minimum force in the FDCs is extracted from the experimental data and defined as adhesion force (F_{AD}) as illustrated in Figure 3.1. The magnitude of F_{AD} depends on the AFM tip radius and sample surface chemistry, and F_{AD} can be expressed as:

$$F_{AD} = -\frac{R_{tip}H}{6a_0^2} \qquad \textit{Eq. 3}$$

where R_{tip} is the tip radius, H is the Hamaker coefficient which will be explained in the later section and a_0 is the intermolecular distance. Noted that all the forces reconstructed in this thesis were acquired with monomodal AM AFM.

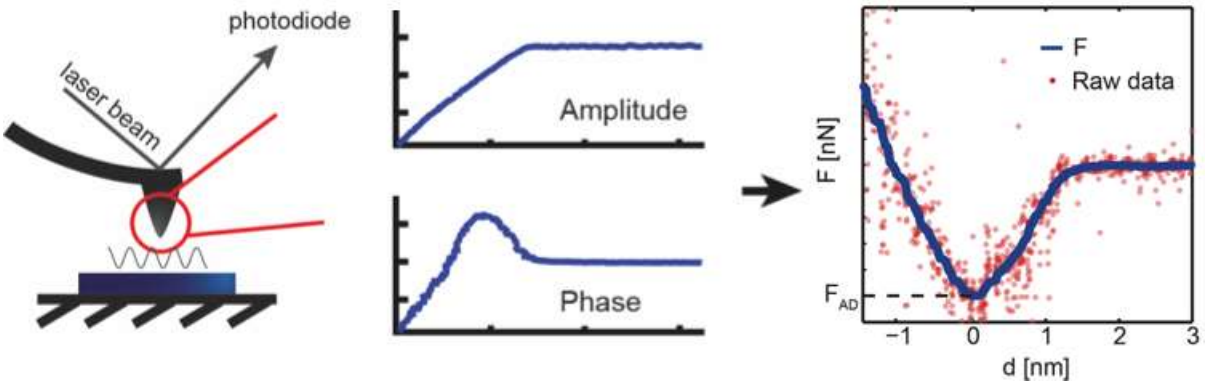

Figure 3.1 Schematic of reconstructing FDC from the amplitude and phase-distance curve.

Furthermore, by using the sphere-flat plane model, F_{AD} between an AFM probe and the sample surface under investigation can be expressed as:

$$|F_{AD}| = 4\pi R_{tip}\gamma \qquad \textit{Eq. 4}$$

where γ is the surface energy.

As it could be seen in Eq. 5, fundamental forces are typically written in terms of power laws[47-48] with inverse-square laws being central in fundamental physics[47]. In this respect, one of the questions in nanosciences is to verify whether physical laws are dependent on the size of the interacting bodies[48]. In the context of van der Waals forces, using Hamaker[49] and Lifshitz[50-51] theories provide fundamental expressions for the interaction between a sphere and a plane so that these can be tested against AFM experimental results. For the interaction between a sphere of radius R and a plane (Figure 3.2), the theories agree in predicting inverse-square laws at fractions of a nm or several nm of separation, in ambient conditions and in a vacuum. In ambient conditions, there is a lack of availability of experimental

data to test the ubiquitous inverse-square law. Therefore, in the following section, we tested the inverse square law for attractive forces in the proximity of the surface, and the dependence of such law on the size of the AFM tip. The methodology is based on the ideal sphere-plane interaction for simplicity and for the reasonable possibility of modelling an AFM tip as a sphere and a sample's surface as a plane.

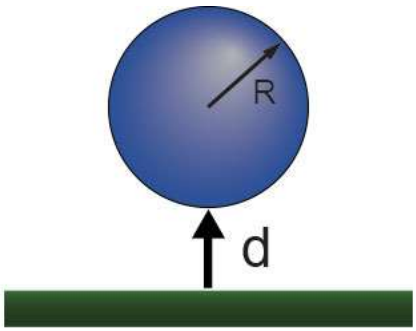

Figure 3.2 Sphere-plane model for an AFM tip-surface interaction.

Power laws of the force

The AFM data was collected in ambient conditions. In the experiments, R_{tip} was monitored *in-situ* with the A_C method providing a functional relationship in dynamic AFM between the minimum free amplitude A_0 required to reach the repulsive regime and the effective radius R_{tip}, *i.e.* $R_{tip}=4.75(A_c)^{1.1}$. The A_c method allows computing the effective radius in seconds and was used here to monitor R_{tip}. Data sets for 5 different cantilevers were collected. We start by writing a general power law[52-53] with power n as

$$F = -\frac{\alpha}{d^n} \qquad \textit{Eq. 5}$$

where F is the force at a distance d and the parameter α might contain the geometrical and/or chemical properties of the interacting system. In the Hamaker and Lifshitz formalisms[50], α is written in terms of a constant H (chemistry) and the radius of the sphere R (geometry) according to

$$\alpha = \frac{RH}{6} \qquad \textit{Eq. 6}$$

We then start by transforming Eq. 3 and assuming the following power law applies

$$F = -\frac{RH}{6d^n} \qquad \textit{Eq. 7}$$

The experimental force profile acquired on a graphite surface and shown in Figure 3.3 is employed for illustrative purposes. First, we identify F_{AD} as the adhesion force[50]. Second, we identify the intermolecular distance $d=a_0$ at which $F=F_{AD}$. This means that by further decreasing d below a_0, atoms on the sphere will repel atoms on the surface[54] and vice versa and imposes a (lower-bound) limit in the applicability of the power law[52]. Furthermore, it allows us to define absolute experimental distances d that agree with standard definitions[43, 55]. Third, a one to one relationship between d and F is assumed.

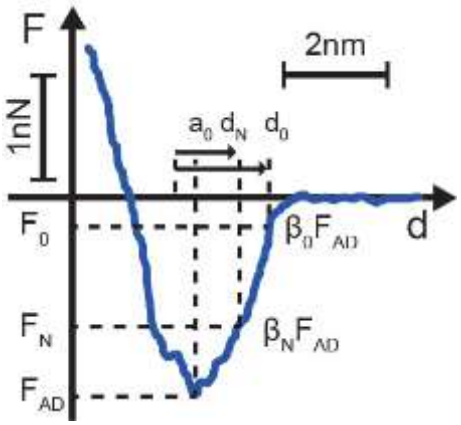

Figure 3.3 Experimental FDC parameterized via F_{AD} and β

To test the range of applicability of Eq. 7, we parameterize the FDCs using F_{AD} by considering the force F at a distance d such that $F=\beta F_{AD}$. By varying β from 0 to 1, the FDCs can be fully parameterized and quantified[56] (see Figure 3.3). Then, we consider a reference value of β to define a force ($F_0= \beta_0 F_{AD}$)-distance (d_0) pair as follows

$$F_0 = \beta_0 F_{AD} \equiv -\frac{\alpha}{d_0^n} \qquad \textit{Eq. 8}$$

All other force-distance pairs (F_i, d_i) can also be parameterized

$$F_i = \beta_i F_{AD} \equiv -\frac{\alpha}{d_i^n} \qquad \textit{Eq. 9}$$

Combining (4) and (5) results in the following expression

$$\left[\frac{\beta_i}{\beta_0}\right]^{1/n} = \frac{d_0}{d_i} \qquad \textit{Eq. 10}$$

Noted that n=2 according to standard Hamaker[49] and Lifshitz[51] formalisms for van der Waals-Cassimir[57] forces with a sphere-plane geometry[50] and when d<2-3 nm[38]. Eq. 10 makes the ratio d_0/d_i being easily computed from experimental FDCs and thus allows to test Eq. 7 for the full range of distances of interest. Eq. 10 predicts the ratio $(\beta_i/\beta_0)^{1/n}$ is independent of R. For this purpose, experimental FDCs were acquired on a graphite sample as a function of R. The value of a_0 for graphite can be estimated by employing the standard expression of surface sciences[50, 55, 58]

$$a_0 = \sqrt{\frac{H}{24\pi\gamma}} \qquad \textit{Eq. 11}$$

The value of H for graphite-silicon dioxide systems interacting in vacuum-air[59] has been reported to be H≈1.35×10^{-19} J while[60] $\gamma \approx 55$ mJ/m^2 resulting in $a_0 \approx 0.18$ nm[50, 61]. Values of R ranging from ≈4-36 nm was obtained on the graphite sample with 5 different cantilevers. In order to test the predictions of Eq. 10 on the data, a reference value at β_0=0.15 was chosen and other values of β_i were β_i= 0.25, 0.35, 0.45, 0.55, 0.65, 0.75 and 0.85 resulting in β_i/β_0= 1.7, 2.3, 3.0, 3.7, 4.3, 5.0 and 5.7. The respective experimental ratios d_0/d_i were computed for the 5 different cantilevers[62].

In Figure 3.4, the results obtained for d_0/d_1 (β_1/β_0) and d_0/d_7 (β_7/β_0) have been plotted versus tip radii R. The respective β ratios for ratios predicted for a power law of n=2 have also been plotted in the figure for comparison (dashed lines). Fig. 3.4 indicates that there is a possible functional relationship between d_0/d_i and R that can be expressed as (R in nm)

$$\frac{d_0}{d_i} = \theta_{1i} R + \theta_{0i} \qquad \textit{Eq. 12}$$

The results from linear regression (obtained with the standard lm function in R and using the model in Eq. 12) are reported in Table 1. Normality is assumed for the predicted intervals at 95% confidence. The goodness of fit is obtained via the standard coefficient of determination RR.

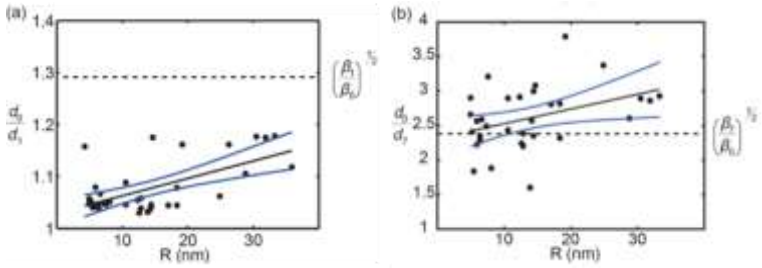

Figure 3.4 Experimental values (circles) as a function of R of (a) d_0/d_1 and (b) d_0/d_7 versus the predictions (dashed lines) of $(\beta_1/\beta_0)^{1/2}=1.3$ and $(\beta_7/\beta_0)^{1/2}=2.4$ and lines of best fit (blue lines) and respective prediction intervals (black lines) according to Eq. 12.

By combining Eq. 10 and Eq. 12, the power n can be further written as

$$\left[\frac{\beta_i}{\beta_0}\right]^{1/n} = \theta_{1i}R + \theta_{0i} \qquad \textit{Eq. 13}$$

Then, the predicted power n at a given β_i/β_0 ratio follows at once from Eq. 13:

$$n_i(R) = \frac{log(\beta_i/\beta_0)}{log(\theta_{1i}R+\theta_{0i})} \qquad \textit{Eq. 14}$$

To get the physical implication, the power law in Eq. 7 obtained from the experimental data can be written as

$$F = -\frac{RH}{6}d^{-\left[\frac{log(\beta_i/\beta_0)}{log(\theta_{1i}R+\theta_{0i})}\right]} \qquad \textit{Eq. 15}$$

The numerical values of the regressors θ_{1i} and θ_{0i} can be found in Table 1. The predicated powers n at β_1/β_0 (continuous lines) and β_7/β_0 (dashed lines) computed from Eq. 14 are shown in Figure 3.5 with corresponding predicted 95% confidence intervals. The non-zero value of the regressors θ_{1i} in Table 1 gives the direct physical implication that the effective power n depends on the tip radius R and presumably also on the distance d in the FDC at which n is computed relative to β_0. In particular,

Figure 3.5 shows that for β_1/β_0, the power n might be an order of magnitude ($n \approx 20$) larger than that predicted by standard vdW models ($n \approx 2$) provided the tip is sharp, *i.e.* R < 5-20 nm. Then, n gets closer to $n \approx 2$ as R increases.

Table 1 Numerical results from the regression analysis carried out with R as a single variable for Eq. 12.

β_i/β_0	θ_1	θ_1 at CI 95%	θ_0	θ_0 at CI 95%	$(\beta_i/\beta_0)^{1/2}$	RR
1.7	0.003	0.002-0.005	1.029	1.004-1.055	1.291	0.611
2.3	0.006	0.003-0.009	1.084	1.031-1.138	1.528	0.543
3.0	0.011	0.006-0.016	1.133	1.056-1.210	1.732	0.626
3.7	0.016	0.009-0.023	1.222	1.116-1.328	1.915	0.649
4.3	0.021	0.011-0.031	1.396	1.237-1.555	2.082	0.601
5.0	0.021	0.004-0.038	1.772	1.507-2.038	2.236	0.392
5.7	0.021	0.002-0.040	2.317	2.029-2.605	2.380	0.388

The dependence of n on β_i/β_0 can be exploited by writing a more general model that includes the ratio β_i/β_0 and three regressors λ_2, λ_1 and λ_0:

$$\frac{d_0}{d_i} = \lambda_2 R + \lambda_1 \beta^* + \lambda_0 \qquad \textit{Eq. 16}$$

where the variable β_i/β_0 has been written as β^* for simplicity. Eq. 16 is the equation of a plane and the best fit (plane) with the full set of raw data acquired (circles), is shown in Figure 3.6. Table 2 shows the numerical values of the regressors, the predicted 95% confidence intervals and the goodness of fit (RR).

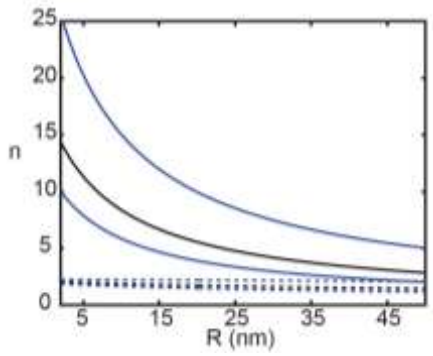

Figure 3.5 Prediction of n from Eq. (10) for β_1/β_0 (continuous lines) and β_7/β_0 (dashed lines).

The prediction for the effective power *n* can now be written as a single expression at any distance *d* as a function of R:

$$n(R,\beta^*) = \frac{log(\beta^*)}{log(\lambda_2 R + \lambda_1 \beta^* + \lambda_0)} \qquad \textit{Eq. 17}$$

Again, by combining Eq. 7 and Eq. 17, the power law results in

$$F(R,\beta^*) = -\frac{RH}{6} d^{-\left[\frac{log(\beta^*)}{log(0.014R+0.361\beta^*+0.99)}\right]} \qquad \textit{Eq. 18}$$

The predictions of Eq. 17 are shown as a function of $\beta^* = \beta_i/\beta_0$ and *R* in Figure 3.7. The vertical axes correspond to the predicted power *n* and the *x* and *y* axes correspond to the β^* and *R*. For *R*=5, 20 and 35, Eq. 17 predicts means of n=15, 2.6 and 2.0 respectively.

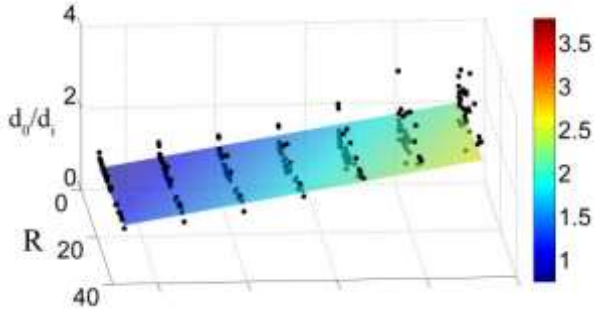

Figure 3.6 Experimental values for d_0/d_i as a function of R and β and regression results according to Eq. 16.*

Table 2 Numerical results from the regression analysis carried out with R (tip radius) and β as variables for Eq. 16.

λ_2	λ_2 **at** CI 95%	λ_1	λ_1 **at CI 95%**	λ_0	λ_0 **at CI 95%**	**RR**
0.014	0.009-0.018	0.361	0.332-0.390	0.099	-0.029-0.227	0.72

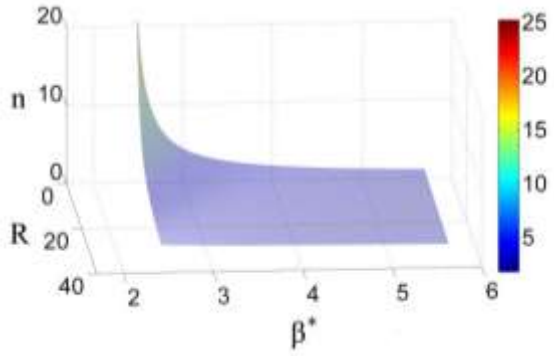

Figure 3.7 Respective predictions of n according to Eq. 17.

In summary, from experimental AFM data, the effective power *n* has been shown to depend on the tip radius *R* and the distance *d* in the FDC. The dependence of the power *n* on *R* implies that there is a nanoscale to (loosely) mesoscale transition in the power law. This transition is very sharp since n can be as large as 10-30 for very sharp tips, *i.e.* $R < 5$ nm. A physical implication is that very sharp tips might be rapidly trapped onto a surface once they get sufficiently close to it. The relationship of the power with distance physically implies that a single power n is not sufficient to completely characterize the FDC, probably because of the combination and presence of forces of different nature. On the other hand, provided the size of the tip is large enough, i.e. $R > 20$-30 nm, the universal inverse-square law seems to be reasonably matching our results. The results here apply only for a graphite-silicon dioxide tip, but the methodologies employed here can be easily extended to any other tip-sample system. This provides a means to find a universal nanoscale power law.

- AFM cantilever details

The AFM cantilevers employed in this section were OLYMPUS AC160TS with k ≈ 40N/m, f_0 ≈ 300kHz, and Q factor ≈ 500.

Statistics

A major motivation behind mapping material properties with nanoscale resolution relates to understanding the relationships between dimensions[48, 63-64] and properties since these might allow fine tuning these properties. The AFM is commonly employed to map nanoscale heterogeneity[65-67]. From FDCs, several material dependent features can be recorded. Being one of the important prerequisites to analysis the acquired data, exploring the experimental conditions and requirements to establish the presence or absence of nanoscale compositional heterogeneity by considering experimental errors in the context of accuracy and precision as function of the samples' size *N* is indispensable. The results show that it is possible to improve precision, *i.e.* decrease the interval or margin of error, while maintaining accuracy, *i.e.* repeatedly including in the given interval or within a margin of error the true mean of the parameter being measured, by sufficiently increasing *N*. However, this is not achievable by directly assuming normally distributed distribution. Applying standard theory of inference concepts with the normal distribution assumptions (or Student's *t*-distribution) leads to very large errors and finally to erroneous or inconsistent conclusions. In this session, the results show that averaging over at least 200-300 points might be required to obtain a normally distributed distribution[68]. In order to obtain sample sizes of 10-30 data points, at least 2000-9000 data points per experiment are acquired and hence the associated time-cost would be considerable[17, 22, 69-71]. A set of metrics to deal with accuracy and precision of force measurements and a protocol for measurements are employed, and a set of standards to compare between sample compositional heterogeneity with nanoscale force measurements is defined. The fundamental

principle behind the investigation deals with the very definition of reproducibility and repeatability in terms of accuracy and precision. Therefore, general criteria to ensure that these hold without restricting assumptions is established.

Sapphire has been employed as a model system to establish accuracy and precision in the measurements and the convergence to a value as the sample's size increases. Two parameters that can lead to systematic errors have been carefully considered. First, during a given experiment the tip radius *R* might vary due to wear – typically increasing *R*. In order to avoid such errors in our experiments, the tip's effective radius was monitored *in situ* with the A_c method[45] and remained constant which lies within the error of the A_c method. Second, variations in the position onto which the laser beam is reflected from the cantilever's surface might lead to slight variations in the volt to meter conversion. To avoid this systematic error, the laser was aligned and adjusted for at least 30 minutes prior to acquiring data.

FDCs were collected on a sapphire surface at a constant rate of 0.5 Hz, *i.e.* 1 force profile every 2 seconds. Sets of data of 5000 points or more were acquired continuously for hours. As an example, two force profiles have been plotted in Figure 3.8. One of them is n=100 (blue line), and the other one belongs to the same set of data (5000 points) for n=4000 (purple line), *i.e.* data points are separated in time by at least two hours. The fact that $F_{AD} \approx$ -1 nN for these 2 data points provides that the tip radius *R* remained constant throughout the measurements since F_{AD} should rapidly increase with *R* according to Eq. 4.

A full set of raw data (light blue circles) collected continuously for ~5000 data points is shown in Figure 3.9. The data have been smoothened with the standard rlowess function of Matlab[72] with a smoothing coefficient of 0.03 (continuous blue line). The mean of the 5000 data points is shown in the dashed black line. At this point, the metrics Accuracy Ratio (AR) and Interval of Error (IE) were defined; IE refers to a given radius of an interval throughout.

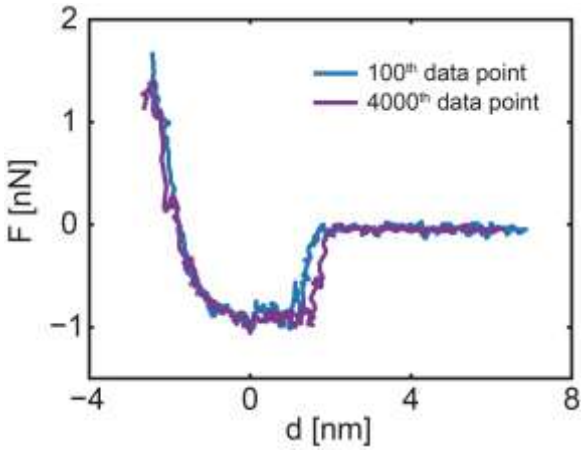

Figure 3.8 Two FDCs acquired on a sapphire surface. Blue line: n=100. Purple line: n=4000.

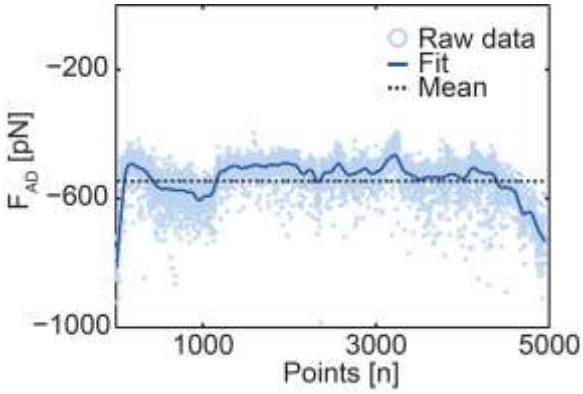

Figure 3.9 A set of experimental raw data with 5000 data points. Light blue circles are raw data, black dashed line is the mean of this data set and blue line is the smoothened result using Matlab.

First, accuracy is defined with the concept of AR to compute the confidence that the estimated mean, accounting for the error, *i.e.* an IE, will include the true mean within a given experimental set-up. More thoroughly, if an AFM experiment was set up without readjusting or recalibrating cantilever-photodiode parameters, the constraint AR<0.05 means that any IE produced only will include the true mean at least 95% of the times the measurement performed. This concept is illustrated in Figure 3.10 where three intervals IE are shown. The first two on the left do not include the true mean (dashed lines) as indicated by crosses while the other one on the right does (tick). The AR metric produces AR=2/3≈0.66 indicating lack of accuracy or a confidence of 33%.

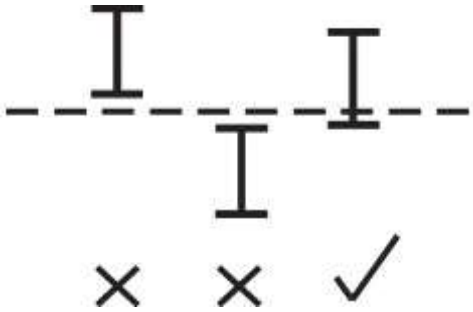

Figure 3.10 Illustration exemplifying the method to numerically compute accuracy of the intervals in a measurement.

The focus here is on the detection or identification of the relative contrast necessary to establish heterogeneity in the context of repeatability within an experiment. Thus, the mean referred here is which would be measured if the measurement was conducted many times. Assumed that 1000-5000 data points are enough to conclude that the true mean coincides with the estimation of the mean. Thus, the use of the term true mean should not lead to ambiguity. The IE provides the precision of the measurement. That is, the maximum resolution with which two properties can be distinguished.

First, let's assume that a Student's *t*-distribution can be employed to define IE via confidence interval CI. These intervals IE(CI) are IE determined from a CI. From this, F_{AD} can be written as the mean of the N data points $\langle F_{AD} \rangle$ with a given uncertainty computed as

$$F_{AD} = \langle F_{AD} \rangle \pm t\alpha_{/2} \frac{\sigma}{\sqrt{N}} \qquad \textit{Eq. 19}$$

where the term in brackets is the sample's mean $\langle F_{AD} \rangle$, σ is the estimate of the standard deviation, N is the sample's size and $t_{\alpha/2}$ comes from a Student's *t*-distribution for a given α (here 0.02). The error or precision is then defined by

$$IE(CI, N) = t\alpha_{/2} \frac{\sigma}{\sqrt{N}} \qquad \textit{Eq. 20}$$

and coincides with the product between the 0.98 quantile of the Student's *t*-distribution and the Standard Error (N -1 degrees of freedom). Eq. 20 is used to compute IE(CI) of data acquired on a sapphire's surface for 4 different data sets (5000 data points each) as shown in Figure 3.11. The vertical axis corresponds to the IE(CI) at α=0.02 and the horizontal axis is the sample's size *N*.

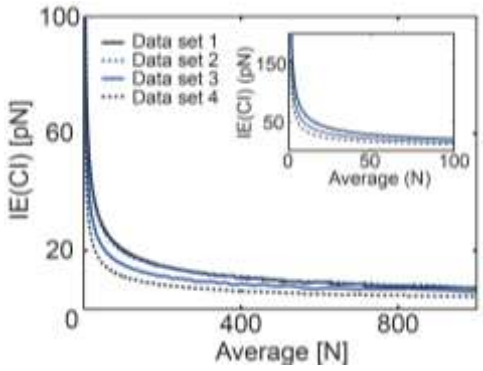

Figure 3.11 IE(CI) computed with four different data sets (5000 points each).

Reproducibility has been tested by acquiring data in two ways. First, by not readjusting the system, the data was obtained with the first cantilever (continuous black lines), and then the data in dashed blue lines were obtained after stopping the data collection for several minutes. As shown in the figure, differences in IE(CI) with increasing N for these two data sets are minimal. Later, the data in continuous blue lines were obtained by using a different cantilever. With a third cantilever, the data in dashed black lines were obtained. A general outcome is that the IE(CI) consistently and monotonically decreases with increasing *N*.

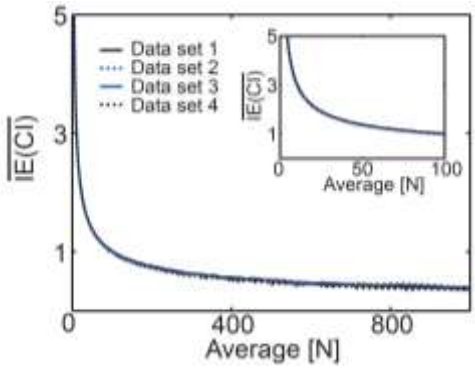

Figure 3.12 Normalized IE(CI) of four data sets (at N=100).

Practically, Eq. 19 and Eq. 20 imply that it is possible to increase precision by increasing *N*. Furthermore, the largest variations in IE occurs from changing cantilevers and readjusting the photodiode. Variations in σ as a function of *N* however were independent of experiment as shown in Figure 3.12 by normalizing IE(CI) at N=100 as

$$\overline{IE(CI)} = \frac{t\alpha/_2 \frac{\sigma(N)}{\sqrt{N}}}{t\alpha/_2 \frac{\sigma(100)}{\sqrt{N}}} \qquad \textit{Eq. 21}$$

Figure 3.1.12 shows strong support that the assumption of 5000 data points suffices to reach high precision and are representative of the system under study.

The IE(CI) metric has been put to test by employing a sapphire's surface as a model sample. The data from Figure 3.1.11 has been grouped into sets of *N*=30 data points (Figure 3.13) and means (black dashed lines) and IE(CI)s (continuous blue lines) have been computed via Eq. 19 and Eq. 20.

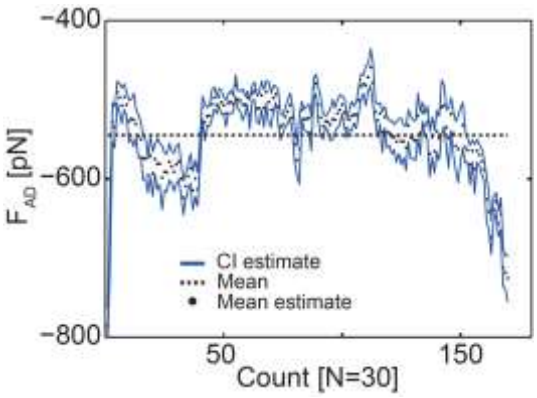

Figure 3.13 Calculated IE(CI)s with N=30. Black dots are means for N=30, black dashed line is the mean for N=5000, and blue lines are IE(IC).

By inspection, most of the IE(CI) do not include the mean calculated with the 5000 data points. This situation does not improve by increasing the number of data points (Figure 3.14). The immediate practical implication from this would be erroneously conclusion that the sapphire surface presents nanoscale heterogeneity. A direct consequence from this conclusion is that the F_{AD} depends on the number of points that the user takes. The objective of this study is to establish consistency

between measurements on a given location of a sample, so that comparisons between different locations of a sample, or different samples, can be carried out. Thus, this inconsistency needs to be solved.

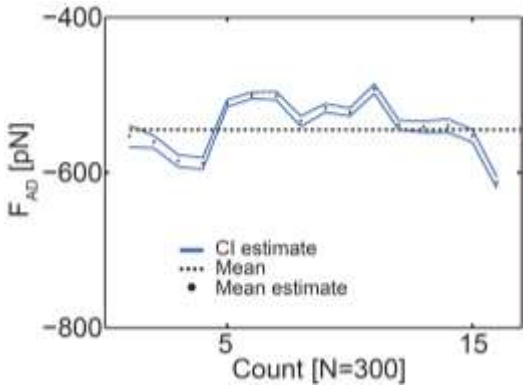

Figure 3.14 Calculated IE(CI)s with N=300. Black dots are means for N=300, black dashed line is the mean for N=5000, and blue lines are IE(IC)s.

The normality of the data sets was established with the use of both the standard jbtest and lillietest normality tests from Matlab. This implies that a normal distribution could be obtained with means obtained from *N*=200-300 data[68], and the data sets did not pass the test when averaging over smaller values of *N*. Now, the concept of AR is defined from the IE(CI) metric computed from Eq. 19 and Eq. 20

$$AR(CI,N) = \frac{Excluded\ (IE(CI))}{Total} \qquad Eq.\ 22$$

where *N* is the sample's size employed to compute an IE(CI), excluded (IE(CI)) is the number of IE(CI)s that exclude the 5000 data points mean, and total refers to the total number of CIs (here 5000/N). The AR(CI, N) metric addresses how well IE(CI)s do in including the mean of 5000 data points as a function of *N* (Figure 3.15). The IE(CI)s do best when including less than 30 data points, *i.e*. *N*<30. When averaging the data over *N*=30, AR(CI, 30)= 0.70 (70% of the intervals exclude the mean) and when averaging over *N*=300, AR(CI, 300)= 0.75 (75% of the intervals exclude the mean).

In summary, the AR(CI, N) metric is too large when employing IE computed with the use of Eq. 19 and Eq. 20.

A metric that provides 1) an error interval that becomes smaller, *i.e*. higher precision, as N increases while 2) also providing accuracy, *i.e*. the intervals should include the true mean needs to be established. The method described earlier fails in the second requirement. Therefore, we turn to the mean and standard deviation estimates from the sample's populations and construct a theory related to Chebyshev's inequality[73].

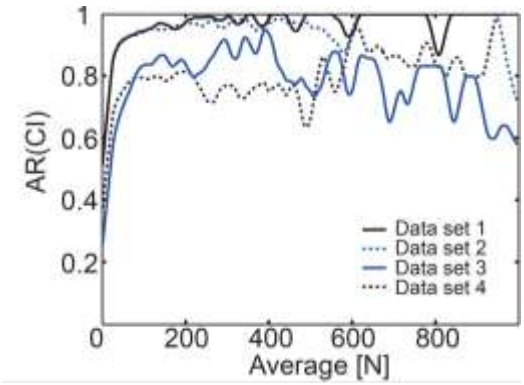

Figure 3.15 The behavior of AR(CI)s of four data sets with respect to N.

The same data from Figure 3.11 has been employed to plot the behavior of the standard deviation σ (N-1 degrees of freedom) as a function of N (Figure 3.16). σ increases by≈20% from N=2 to N≈10, ≈10% from N≈10 to N≈100 (Figure 3.17). Considering σ as a function of N, F_{AD} with errors from the estimate of σ can be written as

$$F_{AD} = \langle F_{AD} \rangle \pm \lambda\sigma \qquad \textit{Eq. 23}$$

where λ (λ>0) is a factor for selecting a given width for the error and can be related to the parameter k in Chebyshev's inequality. The precision of the measurement can be written as

$$IE(\lambda) = \lambda\sigma \qquad \textit{Eq. 24}$$

The use of Eq. 23 and Eq. 24 now reduces to deducing λ that is consistent with the measurements. Therefore, an accuracy ratio AR(λ) can be written as

$$AR(\lambda) = \frac{Excluded\ (IE(\lambda))}{Total} \qquad Eq.\ 25$$

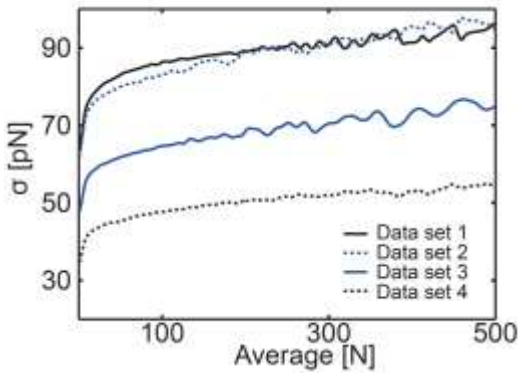

Figure 3.16 σ calculated from 4 data sets.

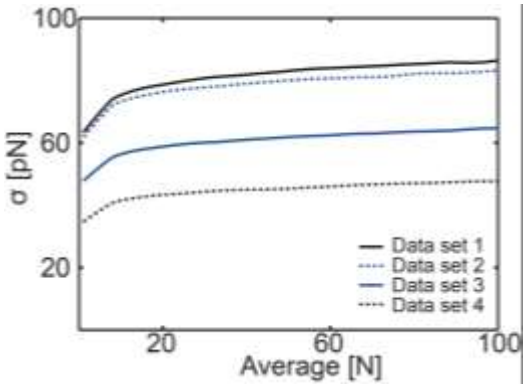

Figure 3.17 σ of four data sets increases slightly as N increase.

The above expression quantifies the accuracy of the measurement and can be employed to test the validity of Eq. 23. In particular, AR(λ)< 0.05 with a meaning similar to that of a CI of 95% would be desired. The dependence of AR(λ) on *N* and σ is shown in Figure 3.18. The vertical axis is AR(λ) and the horizontal axis stands for *N*. The values of λ are 0.5 (dashed black lines), 1 (dashed blue lines), 2 (continuous black lines) and 3 (continuous blue lines). AR(λ) monotonically decreases

with increasing N independently of λ. Yet, larger λ is required for the condition AR(λ) < 0.05 to apply with a small number of data points, *i.e. N*~10-100. This implies that higher precision requires more data points.

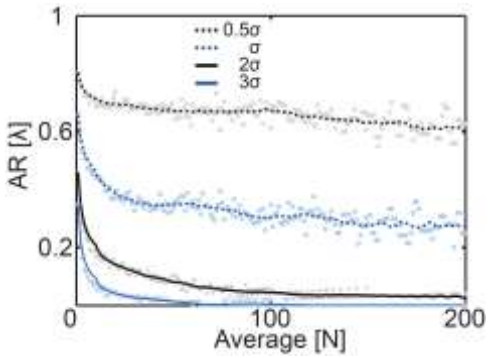

Figure 3.18 AR(λ)'s behavior with respect to N.

The actual values of AR in Figures 3.19 and Figure 3.20 are: AR(2)= 0.12 (N=30) and AR(2)= 0.04 (N=100) where λ=2 throughout. The accuracy of the measurement is determined by the minimum number of data points required to reach a given accuracy in terms of AR(λ). Thus, if accuracy of 95% is required, a minimum value of N needs to be found such that AR(2)<0.05. The pair N=30 and λ=2 should be excluded since AR(2)>0.05 (N=30). The pair N=100 and λ=2 is sufficient for this experiment since AR(2)<0.05 (N=100). Since λ=2, the estimated error interval is IE(2)=2σ(N=100). That is, σ is ≈48 pN and the total uncertainty is ≈96 pN. Thus, heterogeneity could be established if the means in F_{AD} of two materials were at least 96 pN apart. If more precision was required (smaller λ), larger N would be needed. The behavior of IE(λ) or precision with increasing λ is shown in Figure 3.21. The accuracy of the measurement increases with decreasing λ; while precision decreases with increasing λ. Here the small conclusion could be drawn: the accuracy increases with increasing λ and N while the precision increases with decreasing λ.

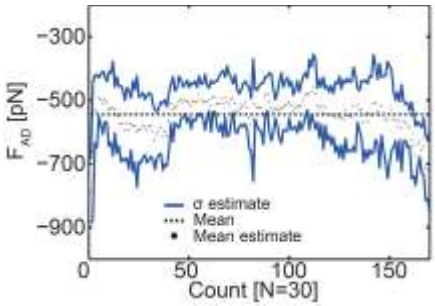

Figure 3.19 Calculated AR(2) with N=30.

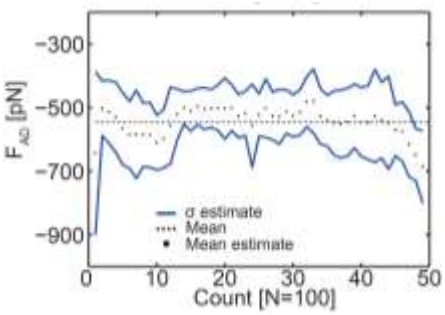

Figure 3.20 Calculated AR(2) with N=300.

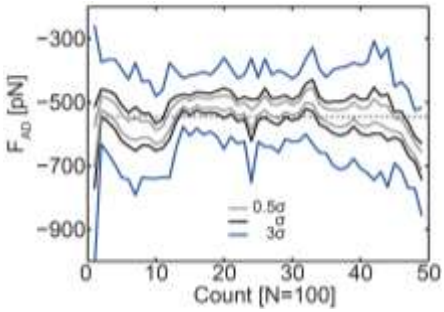

Figure 3.21 The behavior of IE(λ)s calculated with N=100 under different precision criteria.

Next, a block copolymer polystyrene-b-polymethyl methacrylate (PS-b-PMMA) thin film was employed to exemplify the methodology explained above. The two phases (PS and PMMA) were identified (cross for phase 1 and triangle for phase 2). The characteristic cylinders of the PMMA[17] can be seen in Figure 3.22.

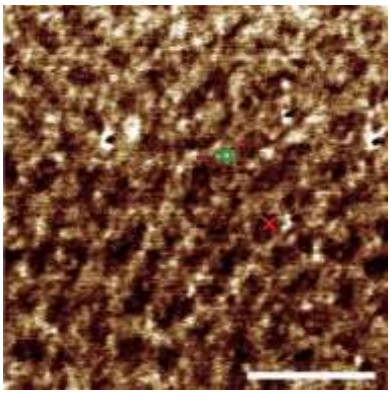

Figure 3.22 AFM scanning operated in the repulsive regime. Cross indicates phase 1 and triangle represents for phase 2. Scale bar of 100 nm.

The aim is to establish the minimum number of points N necessary to accept the hypothesis that the two phases are different in F_{AD}, and provide the IE (radii) with a given confidence level, *i.e.* 95%. The minimum difference D_m is defined, accounting for errors, and computed using the metrics Eq. 23-Eq. 25. First, the difference between estimated means in the two phases is

$$\Delta\mu = \|\langle F_{AD}(1) - F_{AD}(2)\rangle\| \qquad \textit{Eq. 26}$$

where (1) and (2) stand for the two phases under comparison and the parameters are identified with the estimates of the means μ_1 and μ_2 respectively, *i.e.* $\Delta\mu=\mu_2-\mu_1$. If the IEs are computed from Eq. 24, then D_m can be written as

$$D_m = \Delta\mu - 2\lambda\sigma^* \qquad \textit{Eq. 27}$$

where σ^* is the mean of the two standard deviations. While σ monotonically increases with N, the rate of change is so small that σ can be considered a constant in the context of Eq. 24 and Eq. 27. A constraint to determine nanoscale heterogeneity can now be written as

$$D_m > 0 \qquad \textit{Eq. 28}$$

or

$$\frac{D_m}{\sigma^*} > 0 \qquad \textit{Eq. 29}$$

The advantage of Eq. 29 is that D_m is given in terms of σ. From Eq. 26 and Eq. 27, the minimum value of λ to establish nanoscale heterogeneity can also be estimated:

$$\frac{\Delta\mu}{2\sigma^*} > \lambda \qquad \textit{Eq. 30}$$

from which the critical value of λ_c can be written as

$$\lambda_c = \frac{\Delta\mu}{2\sigma^*} \qquad \textit{Eq. 31}$$

λ controls precision via Eq. 24 and this value cannot be selected arbitrarily for a given *N* since accuracy also depends on it from Eq. 25. The constraint AR(λ)<0.05 will ensure that estimates in F_{AD} for a given material will be consistent in the repeatability and reproducibility of experiments while λ will control D_m that can be detected, *i.e*. precision.

The values of Eq. 28 and Eq. 29 obtained for the PS-b-PMMA copolymer are given in Table 3 as a function of λ and *N* while ensuring AR(λ)<0.05. A minimum of N=250 data points are required for an accuracy of AR<0.05 when λ=0.5. From the table, a minimum N≈80 (80 points per phase) is needed to ensure that D_m>0 with λ=1. Replicates (Rep) are given in Table 3 for λ= 0.5, 1, 2 and 3. The results imply that with N<100, differences in the order of 10-10^2 pN can be detected.

In summary, we have shown that reporting the sample's mean, standard errors and standard deviations only might lead to inconsistent conclusions**.** Then, a set of metrics have been introduced in terms of accuracy and precision in the measurements that have been shown to ensure reproducibility and repeatability in experiments. These metrics have been employed to conclusively establish the presence or absence of compositional heterogeneity via a given parameter derived from force measurements with a given number of data points N, and with a given margin of error while ensuring that the results are repeatable and reproducible. Finally, this work should aid to produce robust comparisons between data sets originating from nanoscale force measurements and will assist to produce repeatable and reproducible outcomes in the field.

Table 3 Numerical values for the minimum number of data points N to establish nanoscale heterogeneity from F_{AD} (on PS-b-PMMA). Positive D_m/σ values are required to establish nanoscale heterogeneity and the minimum number N follows for a maximum value of λ, i.e. N≈80 and λ≈1. Expression (17) predicts a maximum $\lambda_c \approx 1.5$.

minimum N	**λ**	**AR**	**mean difference**	**Rep.1**	**Rep.2**	**Rep.3**	**Rep.4**	**Rep.5**
250	0.5	<0.05	D_m [pN]	158	113			
			D_m/σ	2.4	1.9			
80	1	<0.05	D_m [pN]	141	117	80	38	85
			D_m/σ	1.6	2.3	1.3	0.7	1.6
12	2	<0.05	D_m [pN]	-17	0	5	170	78
			D_m/σ	-0.5	0	0.1	3.6	1.5
4	3	<0.05	D_m [pN]	-37	-138	-90	-78	-71
			D_m/σ	-1.2	-3.6	-2.1	-1.8	-1.8

- AFM cantilever details

The AFM cantilevers employed in this section were OLYMPUS AC160TS with k ≈ 40N/m, f_0 ≈ 300kHz, and Q factor ≈ 500.

3.2 Hamaker coefficient

- Introduction

Rapid chemical mapping of substances with nanoscale resolution has been a target of nanotechnologists[17, 74-75]. The broader community relies on probing and identifying chemical substances via standard spectrometry methods that exploit electromagnetic radiation generating footprints associated to a wavelength of the electromagnetic spectrum at which resonance is observed[76]. The preference for such methods is based on robust and reproducible quantification and parameterization in measurements achieved by standard spectroscopy methodologies and the possibility to directly map a physically relevant parameter to chemical substances. To advance nanotechnology or nanosciences, higher lateral resolution is often mandatory[77-78]. While AFM

methods offer the possibility to enhance lateral resolution to sub-nm levels, quantification commonly requires very specialized equipment[18], special environmental conditions such as ultra-high vacuum[74-75, 79] or the use of atomically flat surfaces[74]. Here, we set to map a parameter related to the sample's chemical composition, *i.e*. the Hamaker coefficient H, directly from the standard observables in bimodal AFM via the non-invasive non-contact mode of operation whereby mechanical contact with the sample is avoided.

- Sample preparation

Highly Ordered Pyrolytic Graphite (HOPG) was cleaved with traditional scotch tape method and aged in air for more than 48 hours for the surfaces to reach thermodynamic equilibrium with the ambient air (Temperature at 23±2°C and relative humidity (RH) ~55±5%). Calcite samples were cleaved along $(10\bar{1}4)$ plane and aged in ambient conditions for more than 48 hours for heterogeneity to form on the surfaces and be visible in the AFM data acquired here. (1H,1H,2H,2H perfluorodecyl) acrylate (PFDA) samples were provided by Prof. Gleason. The PFDA was deposited on Si wafer with the use of the Initiated Chemical Vapor Deposition (iCVD) method.

- Bimodal AFM operation

The AFM was operated in standard bimodal AFM by keeping the perturbed amplitude, A_1 or A_{sp}, of the first mode constant while imaging and allowing the phase shifts of the first and second modes and the amplitude of the second mode respond freely to the tip-surface force. AFM cantilevers were excited at the first 2 modal resonance frequencies. The 2 resonance frequencies were determined by thermal analysis of the cantilevers near the sample surface (~50 nm). The effective tip radii were calculated through the A_c method[45]. The A_c value was kept monitoring throughout at intervals of 10 to 30 minutes in between experiments. If no reasonable variations in the A_c value were observed, we assumed the tip remained constant at R≈10 nm as reported by the manufacturer.

- AFM cantilever details

The AFM cantilevers employed in this section were OLYMPUS AC160TS with k ≈ 40N/m, f_0 ≈ 300kHz, and Q factor ≈ 500 for the force reconstruction method, and OLYMPUS AC240TS with k ≈ 2N/m, f_0 ≈ 70kHz, and Q factor ≈ 100 for the bimodal operation.

- Results and discussions

The objective here is to routinely produce maps that is recognized by the broader scientific community, *i.e*. the H coefficient, with nm or sub-nm resolution. Thus, let's recall Eq. 3[43, 49-50] and rewrite it as

$$F \approx -\frac{RH}{6d^2} \qquad \textit{Eq. 32}$$

where *d* is the tip-surface distance. The parameter *H* has many advantages. For example, it can be tabulated and associated to a given chemical substance[59], it can be derived from fundamental and well-known properties of a material, such as the dielectric constant and the refractive index, and can be derived from fundamental quantum field theory[50, 80]. Physically, the Lifshitz theory shows that an effective H value can be found to apply in Eq. 32 and account for London dispersion (frequency dependent) forces, permanent (Keesom) and permanent-induced (Debye) dipoles (zero-frequency contributions)[50]. Here, the experimental H is an effective H that accounts for such broad range of phenomena is assumed. Then, provided R is known, the peak force in a single oscillation of the cantilever in dynamic AFM can be written as

$$F\,(d_{min}, H) \approx -\frac{RH}{6d_{min}^2} \qquad \textit{Eq. 33}$$

where d_{min} is typically referred as the minimum distance of approach[81]. In bimodal AFM the virial V expressions, forming the basis of the bimodal theory[82-84], can be written as

$$V_m(d_{min}) \equiv \frac{1}{T}\oint F z_m dt = -\frac{A_m(d_{min})k_m A_{0m}}{2Q_m} cos\,\phi_m(d_{min}) \qquad \textit{Eq. 34}$$

where *m* stands for mode number, *i.e*. *m*=1 and 2 respectively in bimodal AFM, *F* is an arbitrary tip-sample force, A_m is the oscillation amplitude, A_{0m} is the free or unperturbed oscillation amplitude, Q_m is the quality factor- for simplicity, $A_{01} \equiv A_0$, $A_1 \equiv A_{sp}$ and as customary in dynamic AFM (Figure 3.23). Then, by combining Eq. 33 and Eq. 34

$$d_{min} + b d_{min}^{2/3} + c = 0 \qquad \textit{Eq. 35}$$

where d_{min} is the only unknown since *b* and *c* can be written in terms of the known parameters. The *H* coefficient then follows as[85]

$$H = -\frac{3\pi k_2 A_{02} \cos\phi_2}{0.83 R Q_2 A_2} \sqrt{d_{min}^5 A_{sp}} \qquad \textit{Eq. 36}$$

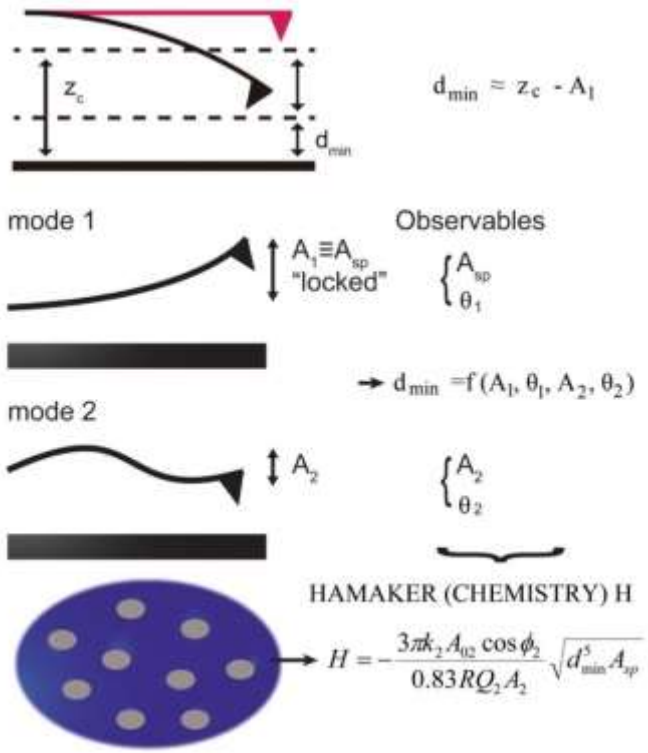

Figure 3.23 Scheme of the geometrical and operational parameters and excitation in bimodal AFM. From the four experimental observables the H is computed explicitly for each pixel in the image.

Bimodal AFM images were obtained in standard bimodal[13] AFM operated in the AM mode for the HOPG sample, *i.e*. constant first mode amplitude where A_{sp} is "locked" and employed to recover

standard topography[13, 17, 83]. Eq. 32 applies only at *d* large enough with no mechanical contact[43, 50] and should be controlled in the experiments. In dynamic AFM, the tip-sample proximity *d* is known to be controlled via the operational parameters A_0 and set-point A_{sp}[86-88]. In short, when above the critical amplitude A_c, mechanical contact occurs during the interaction[45]. Here, $A_0^C = A_0/A_c$ and a set point ratio $A_r = A_{sp}/A_0$ are defined, and we work at $A_0^C < 2/3$. The experimental results obtained in 54 bimodal images on an HOPG sample with 5 different cantilevers and 3 different AFM users are shown in Figure 3.24 for H vs A_r, in Figure 3.25 for H vs A_0^C and in Figure 3.26 for H vs A_r and A_0^C. The black circles account for the mean of each experimental image (256x256 pixels per image) and the best fits (obtained with standard libraries in the R language) are also shown. The coefficients of determination were 0.13, 0.68 and 0.77 for H vs A_r, A_0^C and A_r -A_0^C implying that the main controlling parameter is A_0^C. Here we focus on the range at which the values of *H* obtained in the bimodal images approximately match the values obtained by fitting Eq. 33 to the FDCs. Thus, the data indicates that provided $A_0^C \approx 0.5\text{-}0.6$ (Figure 3.26) the effective H obtained via the bimodal images approximately matches the results from FDCs.

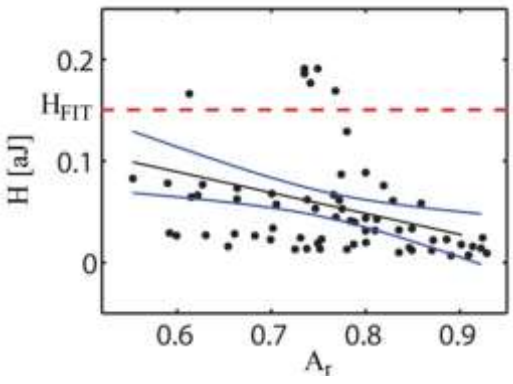

Figure 3.24 Experimental results (circles) for H as a function of A_r.

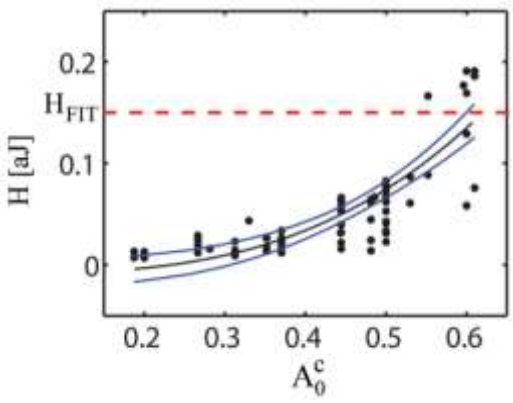

Figure 3.25 Experimental results (circles) for H as a function of A_0^C.

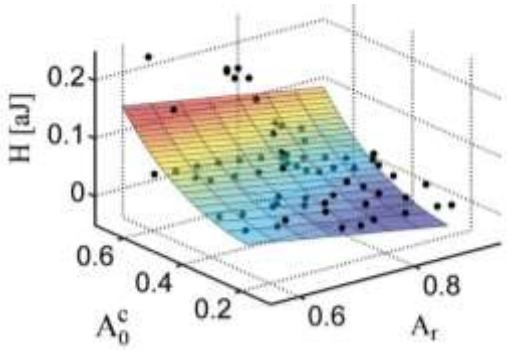

Figure 3.26 Experimental results (circles) for H as a function of A_r -A_0^C.

Table 4 Table presenting experimental parameters, H values obtained from FDCs, bimodal images, the Lifshitz theory, and errors for HOPG, mica, calcite and PFDA.

	A_{sp} [nm]	A_0 [nm]	A_r	A_o ^{C}	H_{IMG} [aJ]	H_{LT} [aJ]	ΔH_{LT} %	H_{FIT}	ΔH_{FIT} %
HOPG 1	3.12	4.77	0.65	0.28	0.016	0.132	-737	0.150	-851
HOPG 2	12.41	15.90	0.78	0.60	0.129	0.132	-2	0.150	-16
mica 1	3.93	4.34	0.90	0.22	0.013	0.080	-528	0.100	-686
mica 2	9.97	12.19	0.82	0.56	0.128	0.080	37	0.100	22
calcite 1	3.21	4.13	0.78	0.31	0.019	0.081	-326	0.031	-63
calcite 2	5.83	7.04	0.83	0.53	0.030	0.081	-170	0.031	-3
PFDA 1	3.74	4.95	0.75	0.58	0.010			0.014	-41
PFDA 2	3.33	5.30	0.63	0.63	0.017			0.014	17

A summary of the results is provided in Table 4 for images obtained at $A_0^C \approx 1/4$ and $A_0^C \approx 0.6$ for the HOPG and mica samples. The set of data provided in Table 4 consists of; A_{sp}, A_0, A_r, A_0^C, H_{IMG} (mean H obtained from the bimodal image), H_{LT} (H predicted by the Lifshitz theory), H_{FIT} (H obtained from the fit from FDCs), the error in H from the image relative to the prediction by the Lifshitz theory (ΔH_{LT}) and the error relative to the fit (ΔH_{FIT}) – errors are given in %. Figure 3.27 shows an example of H_{FIT} from experimental reconstructed FDCs.

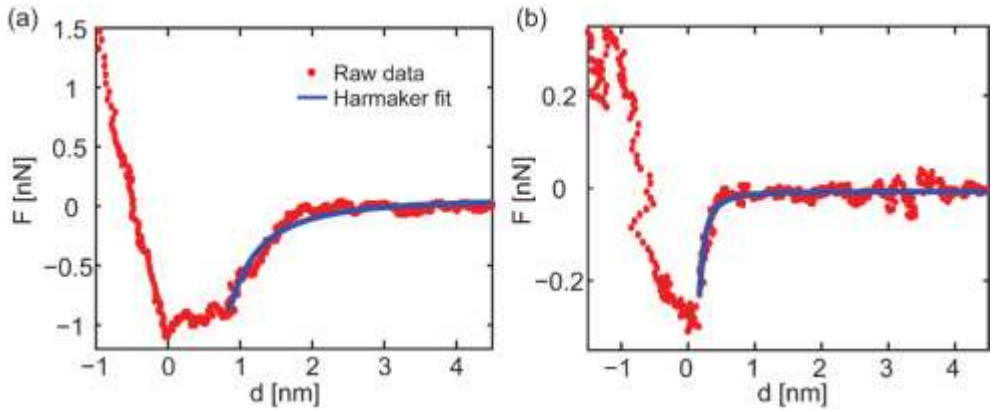

Figure 3.27 Raw experimental (red dots) HOPG and b) PFDA FDCs and (blue lines) best fits obtained by employing standard linear regression on the raw data.

Data is also provided in Table 4 for a PFDA sample[89] (Figure 3.27) and a calcite sample (Figure 3.28). The PFDA sample presents chemical heterogeneity in the 1-2 nm range[90] which complicates manually selecting a given homogeneous position on the surface to probe chemistry as opposed to the HOPG and mica samples. Thus, the H_{FIT} values obtained from FRCs in Table 4 for PFDA are a mean average of the heterogeneous surface. Bimodal images however can provide information in the 1-2 nm range[14, 16] while simultaneously providing a quantitative value for H with such resolution. The calcite sample displays two distinct regions when the sample is exposed to the ambient air. This is the heterogeneity that we observed on the calcite surface for which FDCs and images were acquired. An H value can be computed from the two regions directly from the FDCs. The values that resulted from force profiles were $H \approx 0.047$ and 0.015 aJ for the two phases respectively.

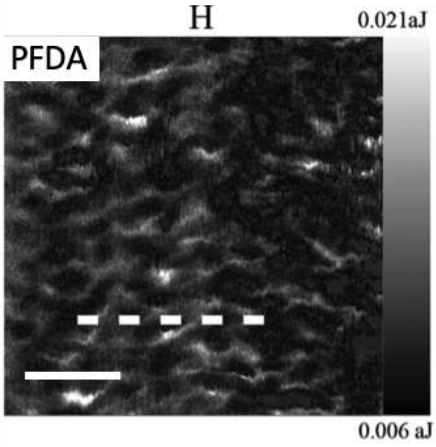

Figure 3.28 H map of PFDA obtained in bimodal AFM. Scale bar: 10 nm.

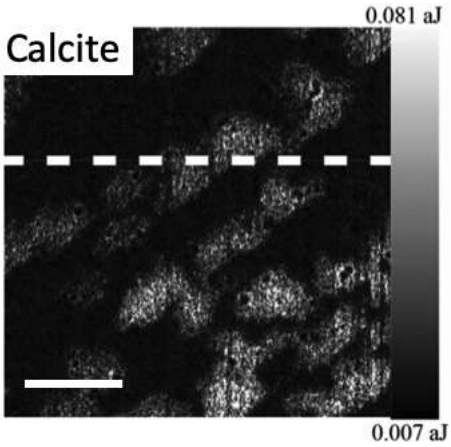

Figure 3.29 H map of calcite obtained in bimodal AFM. Scale bar: 75 nm.

The H map in Figure 3.27 agrees with that the PDFA sample presents chemical heterogeneity in the 1-2 nm range[89-90]. Cross sections corresponding to the H images are shown for PFDA (Figure 3.29) and calcite (Figure 3.30). Crests and troughs of the cross-sections of H are found in the sub 2nm range – as measured at half height (Figure 3.29). The raw data in presented as blue dots per pixel and a smoothened fit (blue lines) obtained with the use of the smoothing rloess function in Matlab (10% span).

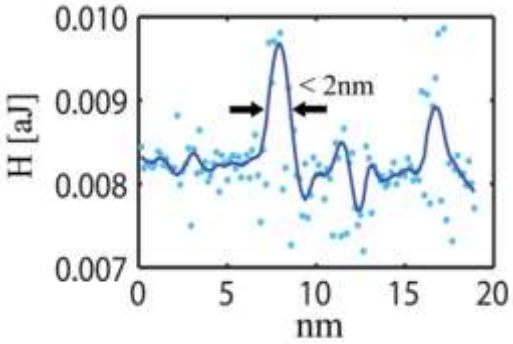

Figure 3.30 Cross sections of the H map of PFDA corresponding to the dashed lines in Figure 3.27.

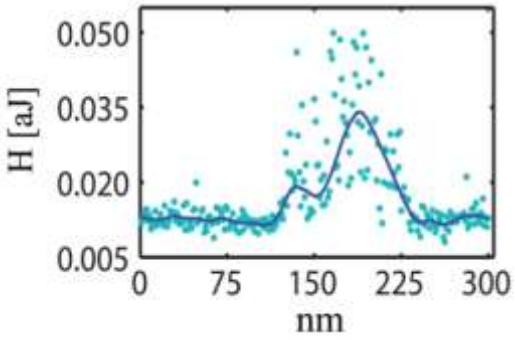

Figure 3.31 Cross sections of the H map of calcite corresponding to the dashed lines in Figure 3.28.

- Conclusion

In short, a theory of bimodal AFM has been presented that can be employed to recover an effective value of the Hamaker coefficient H directly from experimental observables while imaging in standard bimodal AFM in the non–invasive attractive mode of operation. The values of H obtained from such computation are in good agreement with the standard Lifshitz theory and with the fits resulting from experimental FDCs. In addition, a simple methodology has been further presented to optimize the range of operational parameters for which H is in closest agreement with the standard Lifshitz theory.

Chapter 4

Validation of Methodology

In this Chapter, validation of previously described methodology is presented. Various materials were prepared and investigated including $CaCO_3$, TiO_2, and Si. Results are arranged as case study style that put together motivation, sample preparation, discussion and conclusion as a comprehensive section for each material.

4.1 $CaCO_3$

- Motivation

Waterflooding is the dominating secondary recovery strategy used in the oil industry for its efficiency and economic feasibility[91]. One of the most important factors affecting the efficiency of waterflooding processes is the wettability of porous surfaces within reservoir rocks[92], which carry the inherent complexity of the physical morphology and the chemical composition. While the former includes permeability, pore connectivity, and pore size distribution in the reservoir, the latter links with the molecular interactions between the different phases crude oil, brine and rock (CBR). Studies relating to this have been extensively performed[91-96], but a conclusive understanding of the parameters including the intrinsic properties of the material composing the reservoir and the extrinsic conditions that the reservoir has been subjected to has not yet been completely obtained.

It has been shown that oil recovery of the carbonate reservoirs can be improved by altering the reservoir wettability[97-99] to slightly water-wet. However, the fundamental understanding of reasons behind this improvement remains elusive[91]. Due to the lack of spatial resolution, macroscopic techniques fail to provide information for the wetting behavior within the pore, hinting higher spatial resolution measurements, *i.e.*, in submicrometer to micrometer range, are needed.

Calcite ($CaCO_3$) with rhombohedric crystallographic structure is the most common carbonate mineral used to represent the rock formation[100] and the most stable polymorph of calcium carbonate. In addition to the intrinsic properties of the material, *i.e.*, cleavage planes, the wetting history, *i.e.*, the calcite being wetted first by water or oil, is also expected to play a role in affecting the wettability of calcite. Furthermore, upon exposure to the environment, surface wettability changes due to adsorption of atmospheric contaminants[29, 101].

- AFM cantilever details

The AFM cantilevers employed in this section were OLYMPUS AC160TS with k ≈ 40N/m, f_0 ≈ 300kHz, and Q factor ≈ 500.

- Sample preparation

Calcite Iceland spar was mechanically cleaved along the $(10\bar{1}4)$ cleavage plane with a hammer with gloves and handled carefully to avoid contamination from the environment due to aerosol adsorption. Calcite cleaves along the unit rhombohedron labeled as plane 1 as the perfect cleavage plane and was confirmed with x-ray diffraction XRD with rotating sample stage to reduce the preferred-orientation effect as shown in Figure 4.1.

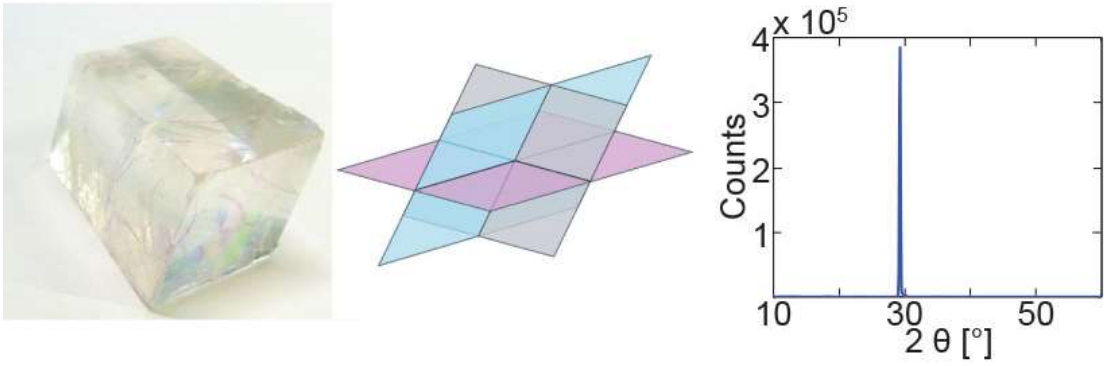

Figure 4.1 Calcite Iceland spar used in this study and the XRD spectrum (peak at 29.4°) showing the plane of $(10\bar{1}4)$.

- Results and Discussions

We examine the wettability of the $(10\bar{1}4)$ calcite cleavage plane upon exposure to the atmosphere. We aimed to show the nanoscale measurements based on the AFM techniques described in the previous chapter are able to describe macroscale wettability indicating it is possible to assess the pores wettability with improved spatial resolution approach.

First, macroscopic measurements, *i.e.*, water contact angle measurements, were carried out to evaluate the wettability of calcite when exposed to ambient conditions. As shown in Figure 4.2, freshly cleaved $(10\bar{1}4)$ plane exhibited superhydrophilic property with contact angles (CA) smaller than 5°. Here, when calcite was freshly cleaved, we defined t=0h. Upon aging, CAs increased and reached ~73.8±8.5° after 120 hours of cleaving. The AFM root-mean-square (RMS) roughness examination of the surface difference between as cleaved and after 24h exposure is 50 pm, showing the morphological variation is negligible and cannot be responsible for such wettability alteration. Yet, as shown in Figure 4.3, AFM phase imaging - which provides information on the surface chemical composition[102] - presented contrasts, implying the formation of chemistry heterogeneity as a consequence of exposure to ambient conditions.

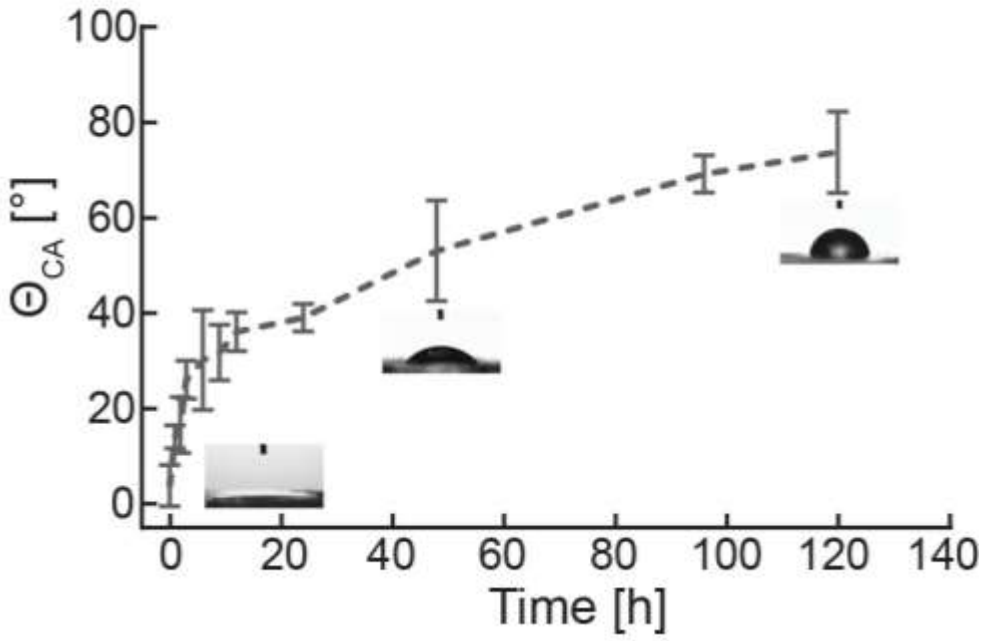

Figure 4.2 Time dependent contact angles for calcite $(10\bar{1}4)$ plane.

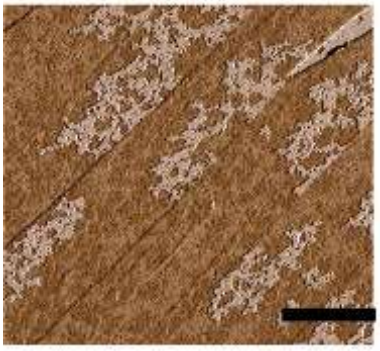

Figure 4.3 AFM phase image for calcite ($10\bar{1}4$) plane. Scale bar: 500 nm.

With this finding, we recorded the phase contrasts of the AFM scan with aging time, *i.e.* from t=0h to t=10h, to study the development of chemistry heterogeneity of calcite surfaces. As shown in Figure 4.4 (a), no contrast was present in phase images at t=0h: we called this state of pure calcite 1st phase. With time elapsed, 2nd phase emerged in the form of lighter color. We can see from the figure that the 2nd phase established from the surface steps edges and formed continuous patches over the scanned area. As shown in Figure 4.4 (b), the increase in the 2nd phase percentage is linear with $R^2 > 0.97$. The presence of surface heterogeneity strengthens the need for micro to nanoscale understanding of the surface properties.

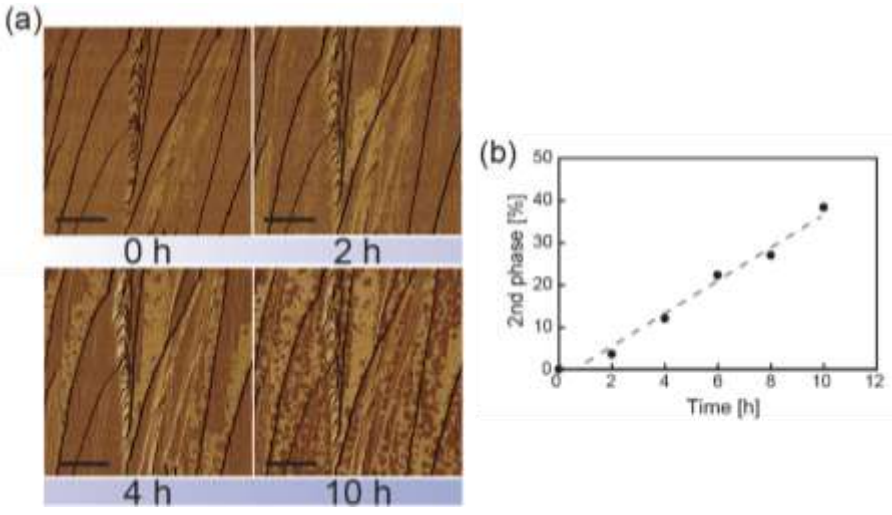

Figure 4.4 (a) Time sequence AFM phase images for (10$\bar{1}$4) plane. Scale bar: 500 nm. (b) show the percentage of 2nd phase growing with time.

AFM force spectroscopy technique was also performed to quantify how the calcite surface properties changed when the 2nd phase appeared. Bimodal distribution[103] in the adhesion force F_{AD} can be clearly seen in Figure 4.5. Since R_{tip} was monitored throughout the experiments ensuring no change occurred, F_{AD} here is then solely affected by sample surface properties, indicating that the two phases present on the calcite surfaces are chemically different. In addition, both AFM force measurements and AFM phase images showed that chemistry heterogeneity develops as freshly cleaved calcite is exposed to the ambient conditions, and consequently changes the surface properties.

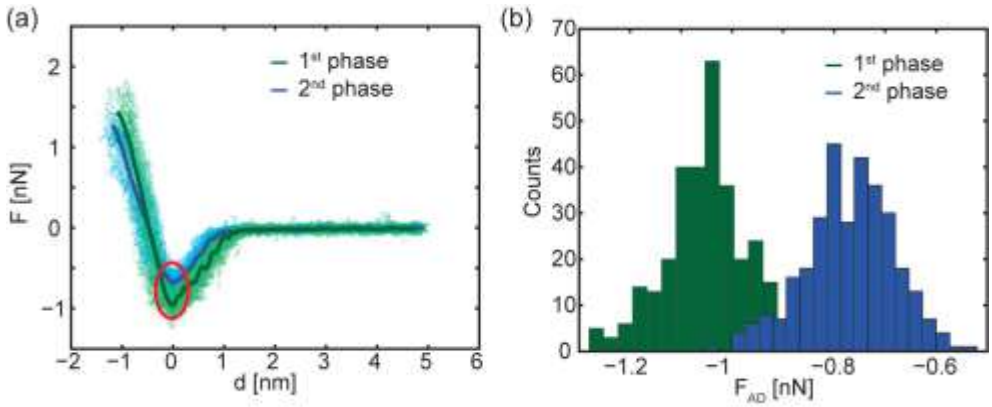

Figure 4.5(a) Force profiles for both phases on (10$\bar{1}$4) plane. (b) Histogram of F_{AD} for the 1st and 2nd phase. Blue and green dots represent for experimental data while the continuous lines stand for averaged force curves.

Furthermore, by invoking the relationship between F_{AD} and the surface energy γ (Eq. 4)

, together with the 2nd phase growth rate presented in Figure 4.4, we could estimate the change in effective adhesion force F_{AD} of a certain area and normalize the change with respect to the F_{AD} value at t=0h to disregard the uncertainty in evaluating the true value of R_{tip}. Therefore, we could have:

$$F_i^* = \frac{|F_{AD,i}|}{|F_{AD,0}|} = \frac{\gamma_i}{\gamma_0} \qquad \textit{Eq. 37}$$

where F^*_i stands for normalized effective F_{AD} at time i. For macroscopic measurements, by recalling Young-Dupré relation, we have:

$$W = \gamma_L(1 + cos\theta) \qquad \textit{Eq. 38}$$

where W is the work of adhesion, γ_L is the surface energy of water and θ is the contact angle. By normalizing the change at t=0h, we could compare the results in macro- and nanoscale. That is, normalized macroscale parameter W can be written as:

$$W_i^* = \frac{W_i}{W_0} = \frac{(1+cos\theta_i)}{(1+cos\theta_0)} \qquad \textit{Eq. 39}$$

where W^*_i represents normalized work of adhesion of time i. In Figure 4.6, by plotting F^*_i together with W^*_i, the macroscale and nanoscale measurements show good agreement as linear regression test showing that the fitting slopes had no significance difference. W^*_i determines the strength of the water contact macroscopically and F^*_i captures at the nanoscale adhesion forces variation. Nevertheless, while like W^*_I, F^*_i is capable of describing the surface wettability, F^*_i holds additional advantages that it studies the calcite surface with higher spatial resolution.

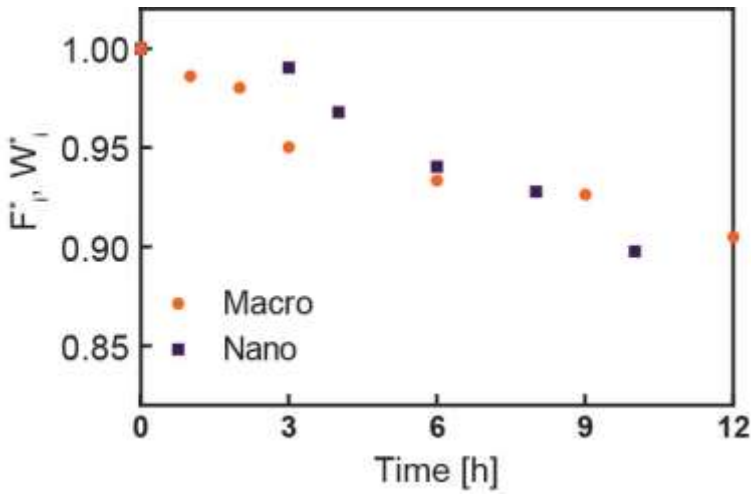

Figure 4.6 Macroscopic and nanoscopic measurements trend on $(10\bar{1}4)$ plane.

FTIR experiments were then performed to investigate the possible composition of 2nd phase. A newly cleaved calcite $(10\bar{1}4)$ plane was put under to mid-IR beam and then left exposed to the ambient air. IR spectrum at t=0h was taken as the baseline signal and compared with the spectra after 24 hours of exposure: the increase in spectra peaks intensity will indicate the surface composition variation. Figure 4.6 shows the absorbance spectra of $(10\bar{1}4)$ plane. In the figure, the peak at 2400 cm^{-1} belongs to the instrument artifact signal which fluctuated in every measurement independent of users and samples. The only variance in the spectrum falls in 3200-3600 cm^{-1} which belongs to O-H stretching signal, while the C-O bending and stretching vibrations of calcite characteristic peak[104] in 1420 cm^{-1} show no difference. The increase in O-H bonding signal could be the result of loosely bound water on the surface as reported earlier. In order to reach thermodynamic equilibrium, freshly cleaved calcite adsorbs water from the humid air forming hydrated[105] $CaCO_3$ to minimize the surface energy.

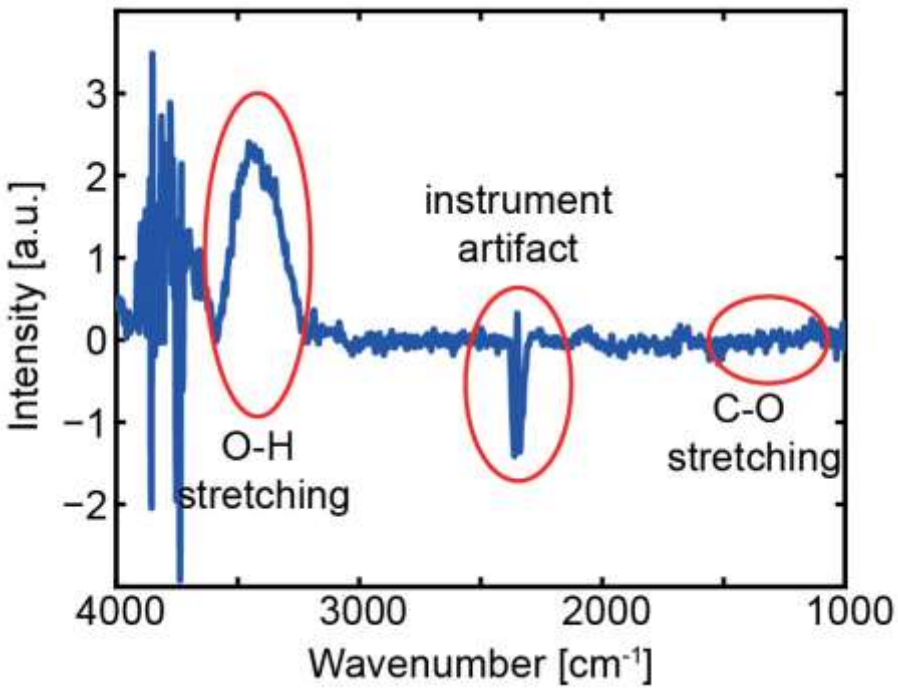

Figure 4.7 FTIR spectrum shows the change in peaks intensity of aged calcite $(10\bar{1}4)$ plane

- Conclusion

We have shown that freshly cleaved calcite undergoes the wettability transition from superhydrophilic/hydrophilic to hydrophobic after exposure to ambient conditions for 120 h by adsorbing water from the environment and forming hydrated $CaCO_3$. The presence of hydrated $CaCO_3$ film is supported by FTIR spectra. Corroborating nanoscale measurements allow us to study calcite surface properties with finer spatial resolution.

4.2 TiO_2

- Motivation

Titanium dioxide (TiO_2) attracts a lot of research interests for its suitable for many applications[106-109]. It has been commercially exploited as it has the ability to combine the effect of photocatalysis and photoinduced, specifically UV-induced, hydrophilicity. UV light could induce a large number of surface oxygen vacancies and dangling bonds generation and facilitate the molecular water dissociative adsorption[110-111].

Water contact angle measurements have been widely used to study TiO_2 macroscopic wettability during and after UV exposure. However, the nanoscale mechanisms of the photoinduced wettability and the changes in surface chemistry are difficult to assess. Previous studies of TiO_2 wettability[112-113] have shown that it is difficult to decouple the effects of surface chemistry, *i.e.*, crystal arrangement and hydroxylation, from those of morphology.

Better understanding of the wettability properties of TiO_2 films could lead to improving TiO_2-based self-cleaning coatings. However, for indoor applications where UV light is limited, such coatings lose their superhydrophilic properties. Therefore, developing a means of inducing permanent hydrophilicity in the TiO_2 film that is not dependent on UV illumination would be highly desirable. High-temperature annealing treatments are found to be one possibility, which is already

widely used in thin film fabrication since it reduces the strain generated during the fabrication process and modifies the film crystallinity, morphology, and adhesion to the substrate. Therefore, annealing may induce desirable changes in the TiO_2 surfaces.

- AFM cantilever details

The AFM cantilevers employed in this section were OLYMPUS AC240TS with k ≈ 2N/m, $f_0 \approx 70$kHz, Q factor ≈ 100, k ≈ 80 N/m, $f_2 \approx 420$ kHz, and $Q_2 \approx 400$.

- Sample preparation

Deposition of TiO_2 films were carried out on 25 × 75 mm^2 soda-lime glass substrates (Sigma-Aldrich) with e-beam evaporation. Acetone and isopropanol (10 min for each) were used to clean the substrates with an ultrasonic bath. 99.9% pure TiO_2 (Plasmaterials) pellets (1−3 mm) were used as source materials in the Temescal BJD-2000 e-beam evaporation system. The deposition chamber was vacuum pumped to 3.0×10^{-6} Torr, and the substrates were rotated at 40 rpm during the deposition. The electron gun voltage and the deposition rate were 10 kV and 1 Å s^{-1} and obtained a final thickness of 250 nm. After deposition, films were annealed in air using the following program: first heating to 475 °C (ramp rate: 10 °C/min), 5 min at 475 °C, heating to 500 °C (ramp rate: 2.5 °C/min), and 4 h at 500 °C. Apart from 500°C, another batch of films was annealed at 350 °C following a similar temperature program. We labeled the samples without thermal treatment "as-deposited TiO_2", whereas "350-TiO_2" and "500-TiO_2" refer to the samples that were annealed at 350 and 500 °C, respectively.

- Results and Discussions

In this study, the AFM techniques described in the previous chapter is incorporated with macroscopic measurements, *i.e.*, the water static contact angle (SCA), to develop a more comprehensive understanding of TiO_2 thin film surface modifications.

TiO_2 thin film surfaces are first checked with AFM imaging. Figure 4.8 shows the AFM topography and phase images for as-deposited and annealed samples. As-deposited TiO_2 and 350-TiO_2 present similar morphology, while 500-TiO_2 surface displayed some structures that were further analyzed with XRD measurements. Figure 4.9 shows the XRD patterns of the TiO_2 films for as-deposited and annealed samples showing that the film becomes crystalline after annealing at 500 °C with the peaks characteristic of the anatase phase[114] located at 2θ = 25.33, 37.82, 48.08, and 55.12°, which can be indexed as the (101), (004), (200), and (211) planes.

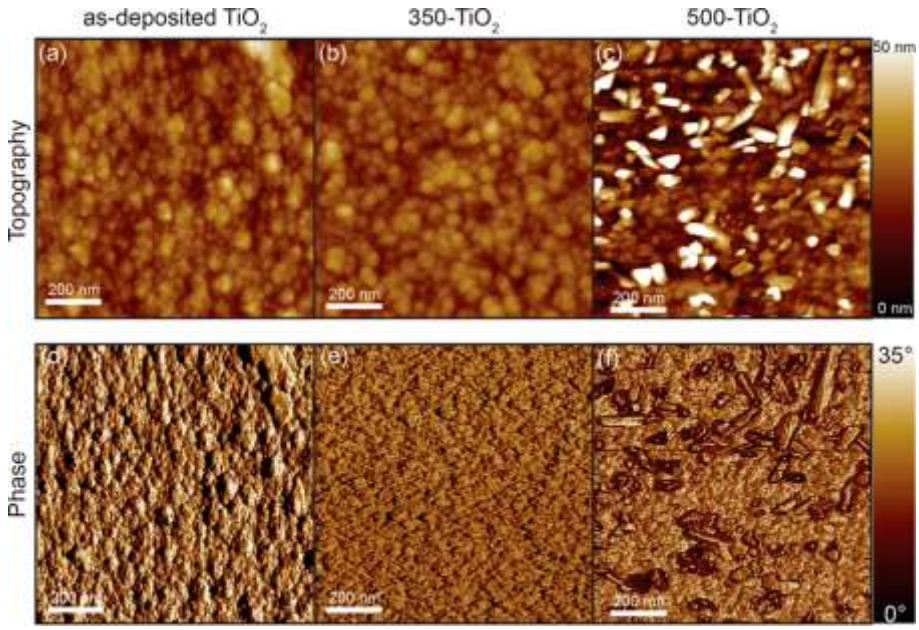

Figure 4.8 AFM topography image of as-deposited TiO2 (a), 350-TiO2 (b), and 500-TiO2 (c) and AFM phase image of as-deposited-TiO2 (d), 350-TiO2 (e), and 500-TiO2 (f).

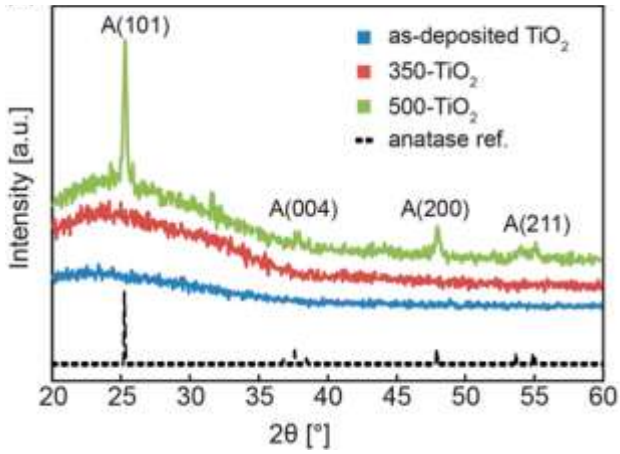

Figure 4.9 XRD diffractogram of as-deposited TiO2, 350-TiO_2, and 500-TiO_2, in which the dashed black diffractogram is a reference XRD of pure anatase.

Figure 4.10 shows that the CA of the as-deposited films before UV irradiation is 96 ± 10°, exhibiting hydrophobic characteristics. On the other hand, the 350-TiO_2 sample showed intermediate behavior (CA = 45 ± 8°), while the 500-TiO_2 sample is moderately hydrophilic with CA = 20 ± 7°. After a 150 min exposure period to UV light, the water droplet spreads out on the film surface almost completely, yielding small CAs smaller than 10° for all the samples. Similar trends were observed with TiO_2 films prepared by different deposition techniques like MOCVD[115] (metal organic chemical vapor deposition), the sol−gel process[116], and radio-frequency magnetron sputtering[117]. However, the temporary range needed for those samples to switch from hydrophilic back to hydrophobic after discontinuing UV irradiation ranged from a few hours to a maximum of 3 days. Yet, the samples presented in this study have much more stable hydrophilicity characteristics after being exposed to UV irradiation for 150 min. This is confirmed by the CA measurements that equal 18 ± 5, 35 ± 5, and 60 ± 6° for 500 TiO_2, 350-TiO_2, and as-deposited TiO_2, respectively, after 3 months of storage under dark conditions. Thus, the result obtained with 500-TiO_2 is improved in terms of long-lasting hydrophilicity.

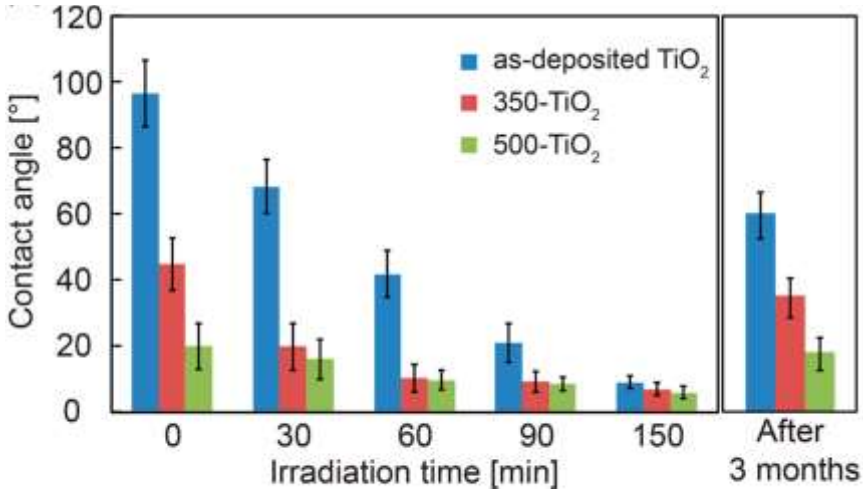

Figure 4.10 Time evolution under UV irradiation of the CA for as-deposited TiO_2, 350-TiO_2, and 500-TiO_2, along with CA after 3 months of storage in the dark.

As material wettability is affected by morphology and surface chemistry[118-121], the surface RMS roughness of TiO_2 thin film samples has been examined. The RMS roughness before and after UV irradiation are 3.7 ± 1.1 and 4.1 ± 0.6 nm for as-deposited TiO_2, 4.8 ± 3.8 and 4.1 ± 3.2 nm for 350-TiO_2, and 13.6 ± 1.5 and 11.3 ± 1.8 nm for 500-TiO_2. This small difference in the RMS roughness values before and after UV treatment for these samples implying the wettability alteration of the TiO_2 thin film should not be attributed to morphological changes. However, since 500-TiO_2 presents a higher roughness than as-deposited and 350-TiO_2, to fully assess the effect of roughness on wettability, the following Wenzel equation has been employed to estimate contact angle for the smooth surface[122-123]:

$$\cos\theta_m = r\cos\theta_Y \qquad \textit{Eq. 40}$$

where θ_m is the experimentally measured contact angle, r is the roughness factor, and θ_Y is the calculated contact angle for a perfectly smooth surface. We obtained calculated θ_Y values of 96.3, 45.1, and 26.5° for as-deposited, 350-TiO_2, and 500-TiO_2 samples before UV treatment and 11.3, 9.3, and 18.5° after UV treatment, respectively. The detailed CAs are recorded in Table 5. The calculated contact angle θ_Y for as-deposited TiO_2 and 350-TiO_2 before and after UV treatment are within the CA experimental measurement error, *i.e.*, $\theta_Y \in [\theta_m - \text{error}, \theta_m + \text{error}]$, yet this is not the case for 500-TiO_2. This indicates that the RMS roughness of the 500-TiO_2 film plays a role in

affecting the hydrophilic properties of the surface. That is, the more hydrophilic character of 500-TiO_2 may be the result of roughness of the surface. However, the wettability variation resulting from UV exposure is dependent on a significant change in surface chemistry. Additionally, as mentioned before, the hydrophilic properties of TiO_2 thin films are preserved for a longer time in the absence of UV when the film has been treated at 500 °C.

Table 5 Detailed θ_m and θ_Y for As-Deposited TiO_2, 350-TiO_2, and 500-TiO_2 before and after UV Treatment.

	Before UV		**After UV**	
	θ_m	θ_Y	θ_m	θ_Y
as-deposited TiO_2	96 ± 10°	96.3°	9 ± 2°	11.3°
350-TiO_2	45 ± 8°	45.1°	7 ± 2°	9.3°
500-TiO_2	20 ± 7°	26.5°	6 ± 2°	18.5°

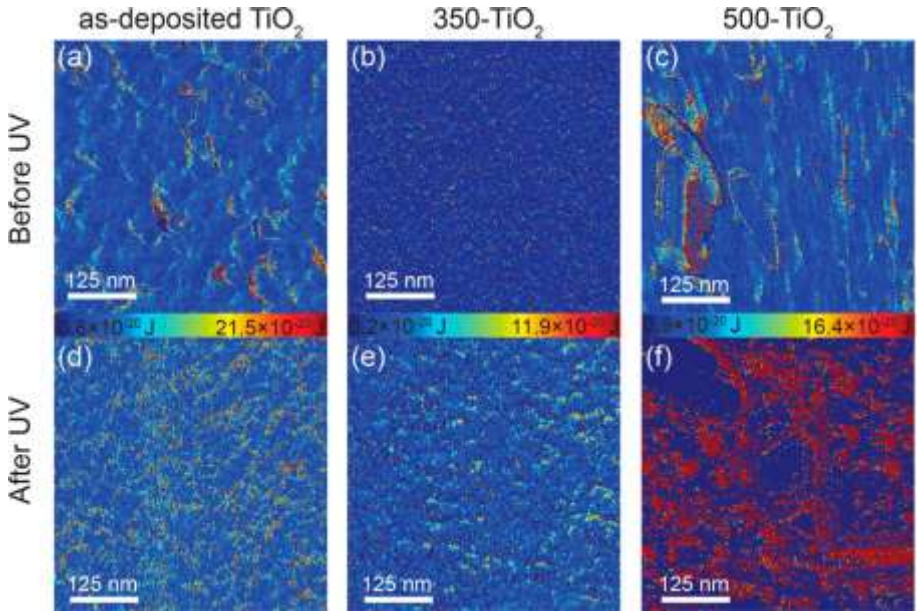

Figure 4.11 Hamaker mapping of as-deposited TiO_2 (a), 350-TiO_2 (b), and 500-TiO_2 (c) before UV irradiation and as-deposited-TiO_2 (d), 350-TiO_2 (e), and 500-TiO_2 (f) after UV irradiation.

The variation in surface chemistry was examined by comparing the Hamaker coefficient values for all three samples before and after UV exposure. To avoid complications in this work, we normalized Hamaker coefficients[85] by the value before UV exposure so that we can evaluate the change is an easier way. As presented in Figure 4.11, we could see increased Hamaker coefficient values after UV exposure of 94.9, 155.6, and 52.7% for as-deposited TiO_2, 350-TiO_2, and 500-TiO_2.

To further look into the reasons behind how annealed TiO_2 could possess such properties, the amplitude versus distance curves recorded via AFM have been analyzed as shown in Figure 4.12. The six cantilever oscillation amplitude (A_1) versus distance (Z_c) curves represent as-deposited and 350-TiO_2 and 500-TiO_2. A negative slope at small Z_c curves after UV exposure can be seen in the figure. As it has been explained in the previous chapter that this feature indicates the presence of a nanoscale water layer on the sample surface[56, 124] and thereby demonstrates the nanoscale hydrophilicity of the UV-treated surfaces.

Table 6 Adhesion Force for As-Deposited TiO_2, 350-TiO_2, and 500-TiO_2 before and after UV Treatment.

	Before UV		**After UV**	
	F_{AD} (nN)	Standard error	F_{AD} (nN)	Standard error
as-deposited TiO_2	0.76	0.25	0.84	0.34
350-TiO_2	0.62	0.28	1.71	3.21
500-TiO_2	2.82	1.64	6.84	2.54

In addition, when reconstructing the tip-sample interaction force profiles, the F_{AD} extracted from the profiles is 0.76 ± 0.25 nN for as-deposited TiO_2, 0.62 ± 0.28 nN for 350-TiO_2, and 2.82 ± 1.64 nN for 500-TiO_2. By using Eq. 4, we know that a larger F_{AD} yields a larger surface energy, and hence a smaller contact angle is expected. The results of 500-TiO_2 samples having a higher F_{AD} are in consistent with the macroscopic CA measurements. The detailed adhesion force for all of the samples before and after UV treatment is recorded in Table 6.

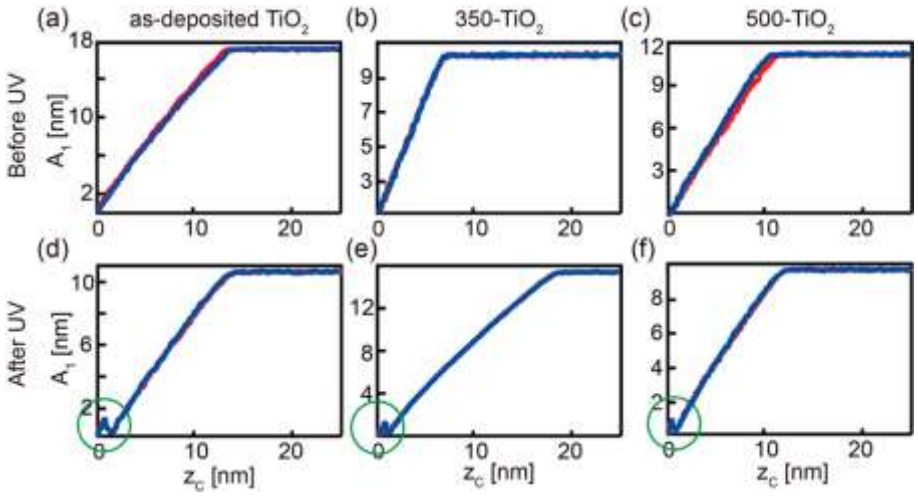

Figure 4.12 AFM probe A_1 versus Z_c curves for as-deposited TiO_2 (a), 350-TiO_2 (b), and 500-TiO_2 (c) before UV irradiation and for as-deposited TiO_2 (d), 350-TiO_2 (e), and 500-TiO_2 (f) after UV irradiation.

With the force reconstruction method, we obtained information on a single point of the studied sample. To get F_{AD} for a larger region, we employed Bimodal-SASS method[124] to acquire data for calculating the tip−sample interaction force while scanning. As a prerequisite of this technique that the AFM probe oscillating in the negative slope region highlighted in Figure 4.12 with green circles, and it could not be performed in the absence of UV irradiation. Figure 4.13 shows average force maps for all of the samples after UV irradiation. All three samples exhibited heterogeneity on the nanometer scale. Furthermore, for the sample annealed at 500 °C and highlighted by red circles, it shows larger TiO_2 crystals, and are characterized by a higher adhesion force: both factors are thought to be responsible for the lower CAs and the higher F_{AD} for the 500-TiO_2 samples.

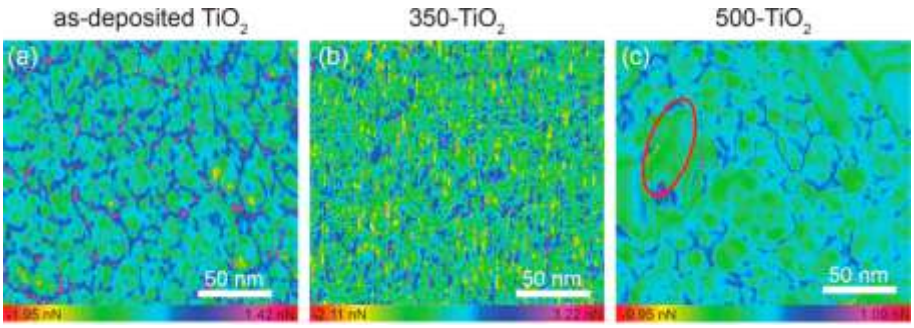

Figure 4.13 Average force maps for as-deposited-TiO_2 (a), 350-TiO_2 (b), and 500-TiO_2 (c) after UV irradiation

- Conclusion

The wettability alteration of e-beam-evaporated TiO_2 thin films with modification of annealing at 350 and 500 °C and UV exposure have been studied with nano- and macroscale characterization techniques. It has been shown that the TiO_2 films exhibit crystallization in the anatase polymorph after being annealed at 500 °C. 500-TiO_2 (CA ∼20°) is found to be substantially more hydrophilic than 350-TiO_2 (CA ∼45°) and as-deposited TiO_2 (CA ∼96°) without UV irradiation. The reasons causing this difference in wettability can be attributed to the RMS roughness of the samples and F_{AD}, which is smaller for as-deposited and 350-TiO_2 samples than for 500-TiO_2. The higher F_{AD} of 500-TiO_2 indicates the higher surface energy than the other two samples, resulting enhanced wetting properties. AFM observations provided indisputable evidence of an adsorbed water layer on all the UV-exposed samples. Force mapping techniques showed that the crystalline part of the 500-TiO_2 sample exhibited an increased F_{AD}. On the basis of this, it was proposed that the hydrophilicity of the 500-TiO_2 film is a result of this partially crystallization and the consequent surface chemistry modification instead of being a mainly morphology-dependent effect. In addition, the annealing process and the annealing temperature prolong the TiO_2 film hydrophilic properties as CA measurements of 500-TiO_2, 350-TiO_2, and as-deposited TiO_2 were ∼18, ∼ 35, and ∼60°, respectively, after 3 months of aging in the dark. With these results, the nanoscale mechanisms that determine the wetting properties of TiO_2 films have been shown to improve hydrophilicity performance. The effectiveness of annealing TiO_2 thin films above 350 °C to induce UV-independent hydrophilicity may suggest an alternative pathway to designing multipurpose indoor/outdoor coatings.

4.3 Si

- Motivation

CA measurements are often involved in wettability studies as the primary approach - a method to assess the degree of wetting putting a droplet of liquid in contact with a solid surface. When considering an ideal solid surface, which is flat, rigid, chemical homogeneous, nonreactive, and insoluble, the equilibrium contact angle captures the minimal Gibbs free energy of the solid/drop system. Yet when encountering rough surfaces, the Gibbs free energy of the system would encompass an extra variable *f*, in addition to the apparent contact angle, θ. For this rough surfaces case, Wenzel[123], Cassie and Baxter[125] established the basis for studying equilibrium rough surfaces wetting many years ago by thermodynamically modeling the entire system and providing equations that give the relationship between an apparent CA describing the wetting state and the Young contact angle. Both Wenzel and Cassie−Baxter equations are applicable in macroscopic scale when the droplet is sufficiently large, in comparison to the roughness scale. However, in microscale wettability where chemical heterogeneity plays an important role, these equations could not provide an accurate model. The lack of spatial resolution in traditional methods gives out the message that it is necessary to discern the role of chemistry and morphology to study microscale wettability.

There is a good example illustrating the importance of well-characterization of microscale wettability from the oil industry. In waterflooding-based enhanced oil recovery (EOR), especially for sandstone-like reservoirs, the most important factor affecting the efficiency of waterflooding processes[92] is the wetting properties of the reservoirs. Hence, the influence of wettability on the efficiency of this process has been extensively studied, but still lacks a satisfactory understanding. The key of this problem lies in the physical morphology and chemical composition of the core since these factors influence the behavior of a reservoir and determine the relative permeability and distribution of fluids inside of it.

- AFM cantilever details

The AFM cantilevers employed in this section were OLYMPUS AC160TS with k ≈ 30N/m, f_0 ≈ 280kHz, and Q factor ≈ 400.

- Sample preparation

Periodically staggered structures on a Si substrate were created with E-beam lithography (Raith e-LINE). Among the structures, the characteristic diameter of the pores is ~180 nm with the depth ~120 nm and lattice constant of 510 and 270 nm (Figure 4.14). A new Si wafer was diced and sonicated with acetone and IPA (each step ~5 min), then immersed in DI water ~1 min to remove any trace of solvent. Then, the wafer was annealed at 200 °C for 30 min. PMMA A2 (950K) was used as an ebeam positive resist. After spin coating with HMDS and PMMA A2 (3000 rpm for 60 s) and prebaking (softbaking) at 180 °C for 90 s, the resist thickness was measured with as a Filmetrics F40-UV reflectometer and with reading of 72 nm. The pattern was imprinted on a 1.4 × 1.4 mm^2 area. An acceleration voltage (EHT) of 25 kV and an aperture size of 20 μm were chosen to obtain the necessary resolution, and to reduce the time of the large-area patterning. The dose was 140 $\mu C/cm^2$ and element step size for patterning was 10 nm. After ~2 h of patterning, the resist was developed using MIBK/IPA (1:3) for 30 s and then rinsed by IPA and DI water for 35 and 30 s. After this, the structure was etched by SAMCO RIE-200iP fluorine with SF_6 (10 sccm) and CF_4 (100 sccm) at 1 Pa for 55 s, using the RF powers of BIAS: 15 W and ICP: 100 W. The hard mask was removed with a flow of O_2 at 50 sccm for 2 min. Finally, native oxides or any kind of residuals from earlier processes was removed by HF etching (49% for 2 min).

For SiO_2-Coated structures, SiO_2 was deposited on the structure with an Oxford FlexAL ALD tool at 150 °C for 200 cycles. Precursors were BTBAS-*t*-butylaminosilane and O_2, and the purging gas was Ar. A detailed deposition cycle was 3 s pulse of BTBAS precursor (80 Torr), 3 s Ar purge, a 3 s pulse of 60 sccm of O_2 at a 250 W plasma power (15 Torr), and a 2 s Ar purge. A J.A. Woollam variable angle ellipsometer was used to measure SiO_2 layer thickness, and fitted with a Cauchy model, giving a thickness ~30 nm and refractive index of 1.44. “Silanization” was done by

adsorption of $C_{19}H_{42}O_3Si$ in a fume hood. Five drops of $C_{19}H_{42}O_3Si$ were placed in an aluminum foil cap and the wafer was placed on top as a seal for the cap. After 2.5 h, the wafer was placed on a hot plate (150 °C) for 10 min to cure and evaporate the excessive silane.

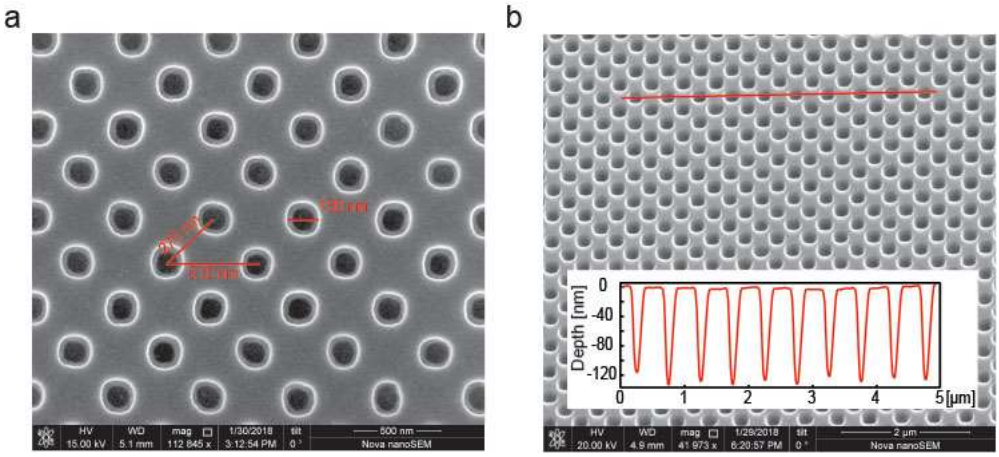

Figure 4.14 (a) Two-dimensional square lattice of pores on a Si substrate. (b) AFM scans show a pore depth ~120 nm.

- Results and Discussions

The aims of this project are to investigate the wetting properties of well-defined macroscopic surfaces by combining CA measurements and the AFM technique described in the previous chapter. The submicron pores of reservoir rocks were portrayed in our controlled laboratory method by fabricating idealized periodically staggered structures on Si using e-beam lithography. The process consistently creates self-similar, repeatable pore dimension structures. We modify the chemistry of this predefined Si pattern by coating it with a 30 nm SiO_2 layer and silane functionalization. The SiO_2 layer was deposited with atomic layer deposition (ALD) to assure uniform coverage between and within the pores.

For the CA experiments on structures with different chemistry, a sufficiently small deionized (DI) water droplet volume, *i.e.*, 1 µL, was used so that any spreading is fully contained within the 1.4 × 1.4 mm^2 area of the structure when investigating the patterned structure (see Figure 4.15). In order to minimize evaporation of the water droplet and ensure experimental repeatability, the

experiment was carried out at conditions close to those of saturated vapor. Despite the small size, the drop is large enough compared with the characteristic length scales of the periodic structures.

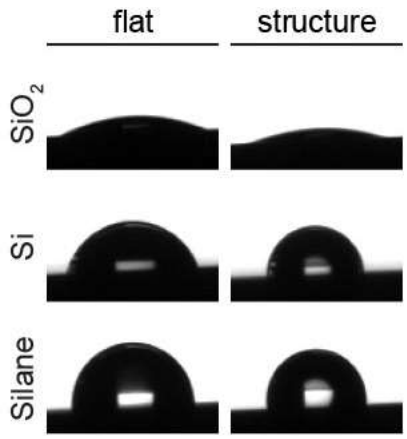

Figure 4.15 Photograph of the 1 μL of DI water placed on flat and periodically staggered (structure) SiO_2-, Si-, and silane-functionalized substrates.

Figure 4.16 shows the CAs collected on flat and patterned surfaces of SiO_2-, Si-, and silane-functionalized substrates. The CAs on the flat surfaces are in consistent with the data reported previously[126-127], which gives 24.1 ± 3.0, 76.7 ± 0.9, and 91.7 ± 1.5° for SiO_2-, Si-, and silane-functionalized Si substrates, respectively. As for the CAs on the structure surface, 14.1 ± 1.1, 85.5 ± 2.3, and 98.9 ± 0.2° were obtained for SiO_2-, Si-, and silane-functionalized Si substrates.

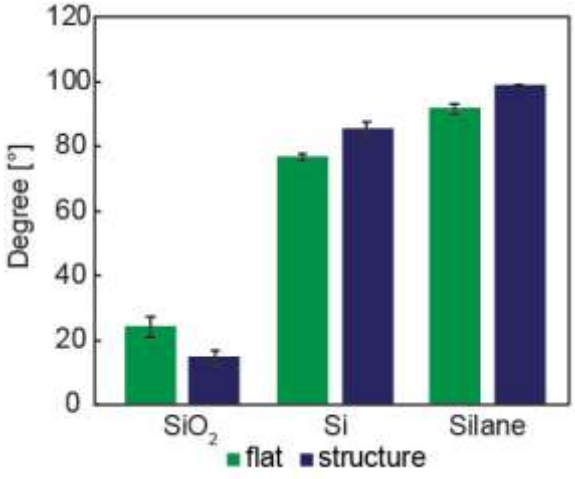

Figure 4.16 Average values for CAs of a 1 μL of DI water droplet.

All of the numbers reported in this work are averages of 30 droplets over 5 different samples, that is 6 droplets for each sample, to show the repeatability of CA measurements. The differences between CAs on the flat and the structures are examined with Wenzel or Cassie−Baxter models. The well-known Wenzel[123] and Cassie−Baxter[125] equations give:

$$\text{Wenzel: } \cos\theta_m = r\cos\theta_Y \qquad \textit{Eq. 40}$$

$$\text{Cassie-Baxter: } \cos\theta_m = r_w f\cos\theta_Y + f - 1 \qquad \textit{Eq. 41}$$

where θ_m is the measured contact angle, θ_Y is the contact angle on an ideal surface, r is the roughness ratio, r_w is the roughness ratio of the wetted area, and f is the fraction of the wetted area. Here we used the values of CA on the flat surfaces for θ_Y, 1.078 for r, 1.003 for r_w, and 0.868 for f. r, r_w, and f were obtained from AFM imaging analysis. Applying the Wenzel model for the SiO_2 sample and the Cassie−Baxter model to the Si- and silane-functionalized Si substrates, the calculated θ_m values for SiO_2-, Si-, and silane-functionalized Si substrates are 10.2, 86.1, and 99.0°, which are very similar to the experimentally obtained CAs on the periodically staggered structure. All of the CA values are summarized in Table 7.

Table 7 CAs on Flat and Structured SiO_2-, Si-, and Silane-Functionalized Si Substrates along with CAs Calculated from the Cassie−Baxter and Wenzel Models

	SiO_2 (deg)	**Si (deg)**	**silane (deg)**
flat	24.1 ± 3.0	76.7 ± 0.9	91.7 ± 1.5
periodically staggered	14.1 ± 1.1	85.5 ± 2.3	98.9 ± 0.2
Cassie−Baxter		86.1	99.0
Wenzel	10.2		

To decouple the roles of chemistry and morphology, we exploit the AFM technique to reconstruct the force field exerted by the sample surface. Force profiles on flat and structures on SiO_2-, Si-, and silane-functionalized Si substrates are reported in Figure 4.17. Each presented curve is an average of 200 measurements taken at 5 different spots on each sample. Student's t test showed

that there is no significant difference between flat and structured regions within each sample and significant difference across different samples. By comparing the shape of the profile, force profiles on flat and structured area are identical. As the AFM probe tip radius (~10 nm) is at least 1 order of magnitude smaller than the scales of the structure (~200 nm), this makes the probe unaffected by the morphology of the surface. Furthermore, this outcome also indicates that by disregarding the morphology factor of the surface the AFM probe senses the same chemistry on the flat and structured surfaces. This is a conclusion that cannot be derived with macroscopic measurements due to spatial limitation.

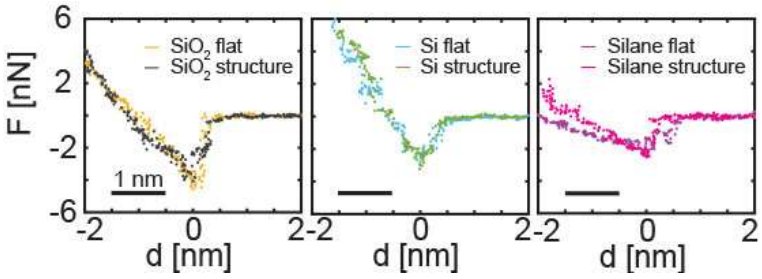

Figure 4.17 Force profiles on flat and structures on SiO_2-, Si-, and silane-functionalized substrates.

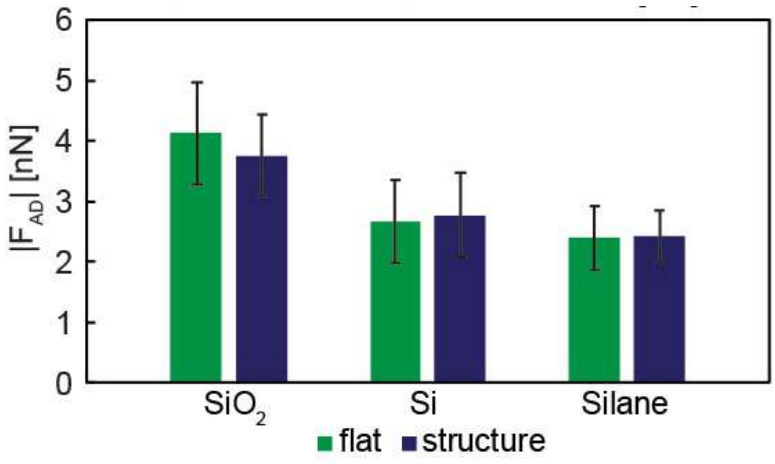

Figure 4.18 The $|F_{AD}|$ for the flat and structures for all the samples is within the standard deviation of the experiment.

Next, by using Eq. 4 and ensuring R constant throughout the experiment, we could compare F_{AD} across the samples and refer to a direct proportionality between F_{AD} and γ. As the F_{AD} values are identical for the flat and structures with no statistical significance (Figure 4.18), we confirm that the AFM tip with the spatial resolution of ~10 nm could determine the sole effect of surface chemistry while the macroscopic CA measurements would be affected by the surface roughness.

- Conclusions

In this study, we prepared flat and periodically staggered Si surfaces with SiO_2 coating or silane functionalization. CA and AFM measurements were both performed on these different structures (flat and staggered pores) and surface chemistries. This study showed that the CA measurement is constrained by its intrinsic spatial resolution limit and hence when characterizing the wettability of surfaces with microscopic roughness, it's not able to decouple the effects of morphology and chemistry. This method gave a ~10° difference in CA on the flat and periodically staggered surfaces. However, with AFM measurements, it can disregard the surface morphology, sensing only the chemistry of the rough solid surface. Our methodology has shown that with a higher spatial resolution-characterizing technique, we could overcome the morphology−chemistry coupling issue that complicates wettability studies and can be potentially employed to probe the wettability alteration due to surface chemistry functionalization and rough surfaces caused by nanoscale patterning.

Chapter 5

Machine learning in AFM

It could be argued that the AFM community has reached a point reminiscent to what physicists lived when facing the emerging zoo of particles in particle physics. At this point, we face a growing set of contrast maps emerging from multiple observables, signals, expressions, and contrast channels, that currently allow imaging with small and large amplitudes in liquid, air and vacuum environments. In this chapter, we propose a radical form of data analysis where AFM data is directly transformed into abstract machine learning features. The concept is encapsulated in the Mendeleev-Meyer Force Project (TMMFP) where data should be tabulated in a manner reminiscent of the construction of the periodic table. The goal is to group and tabulate substances using nanoscale force footprints rather than atomic number or electronic configuration as in the periodic table. The process is divided into: 1) acquiring force data from materials, 2) parameterizing the raw data into standardized input features to generate a library, 3) feeding the standardized library into an algorithm to generate or exploit a model to identify a material or property. We propose producing databases mimicking the Materials Genome Initiative, the Medical Literature Analysis and Retrieval System Online (MEDLARS) or the PRoteomics IDEntifications database (PRIDE) and making these searchable online via search engines resembling Pubmed or the PRIDE web interface.

Background

The starting point for the standardization or tabulation of materials consists of FDCs acquired with AFM. These curves contain the nanoscale footprint of the substance or material and are typically acquired with an AFM. The force between the tip of an AFM and a surface is monitored as a function of separation or distance. The sensitivity of the AFM should provide information from all the relevant nanoscale force footprints or force contributions between materials[2]. On the other hand, the quest to identify and recognize atoms or materials from atomic footprints or FDC data has remained an active research field[79, 128] which remains challenging when considering the generalization and standardization of measurements and procedures[74, 129]. In addition, experiments are typically sophisticated and are reported by carrying out extensive analysis from complex models or fundamental theory[79] rather than via automated processes. About a decade ago, a significant advance was reported by invoking a particular form of normalization of the raw FDC data[74] and single atoms were identified via specific atomic footprints. Similar approach was more recently employed to identify more complex heterogeneous systems[79]. Other forms of sample recognition and identification consist of modelling and parameterizing the FDC with physically relevant parameters such as stiffness[17], adhesion, viscoelasticity[130] or other parametric models[131], or model free parameters[71, 132-133]. Parameterization typically involves an intermediate step after acquiring the raw data which consists of quantification and comparison. In this way, detected differences are exploited as parameter contrast maps that could be used to discriminate between materials[130]. Standardization and tabulation are lagging far behind[134] other research fields such as proteomics, metabolomics and genomics that are heavily assisted by computer science, large databases, powerful search engines and submission protocols[135] that allow rapid access to the databases. Yet, Kalinin *et al*. have been early proponents of the exploitation of computer science techniques in probe microscopy.[23, 133] Here we propose to integrate force spectroscopy and advanced computer science techniques. The objective is to parameterize the FDC raw data into features with the abstract meaning typically given to the features employed in machine learning algorithms. In this way, no restriction us imposed to the number of input features to identify a given material or family of materials or substances. Features are then employed to construct feature libraries for groups of families

or specific families. Finally, an algorithm is exploited to generate a model from a given feature library that groups materials according to similarity. Also, the concept of classification from standard machine learning where the output of the algorithm is zero when a non-match is predicted and one when the algorithm predicts a match is employed. In the prototype, a multilayer neural network is trained with the backpropagation method in Matlab[72]. F-score is used as a figure of merit to quantify Precision and Recall for the models. Precision and Recall are defined as in machine learning where Precision is the ratio between true positives and predicted positives and Recall is the ratio between true positives and actual positives. The F-score parameter combines Precision and Recall as

$$F - score = 2 \times \frac{Precision \times Recall}{Precisioni + Recall} \qquad Eq.\ 42$$

The advantage of employing the F-score rather than Precision or Recall alone is that high values in F-score will be obtained if and only if both Recall and Precision are high simultaneously. In a more intuitive note, Precision could be defined as specificity and Recall as sensitivity implying that high F-score values include both high specificity and sensitivity. More detail on these figures of merit is given when discussing a practical example below.

Methods

- AFM cantilever details

The AFM cantilevers employed in this section were OLYMPUS AC160TS with k ≈ 30N/m, f_0 ≈ 280kHz, and Q factor ≈ 400, and OLYMPUS AC240TS with k ≈ 2N/m, f_0 ≈ 70kHz, and Q factor ≈ 100.

- Raw data acquisition:

The initial step for parametrization and tabulation involves acquiring the nanoscale force footprint in the form of an FDC. This force arises from the atomic interactions between the atoms

on the tip and the atoms on the sample. It is also typical to associate F_{AD} with mechanical contact between the AFM tip and the surface[43]. Only points in the force curves satisfying F<0 nN are considered since these provide enough information to classify materials.

- Parametrizing raw data and transformation into input features:

The second step consists of parametrizing the raw FDC. Measuring the distances in the well of the FDC in a similar fashion to that recently proposed elsewhere[29, 62] is chosen. The steps are as follows:

1) Take the F_{AD} as the force reference for a given curve. This reference allows considering all other force-distance pairs with the use of a factor β as $F=\beta F_{AD}$.
2) Vary β from 0 to 1. Any arbitrary force curve can be fully parameterized and quantified[62] for F<0 nN.
3) Limit β to 0.85, 0.75… 0.05 and normalize the distances in the well of the curves with the reference β=0.85.
4) Compute the absolute distances $dF_\beta=dF_{0.85}$, …, $dF_{0.05}$ where β=0.85, …, 0.05 as illustrated in Figure 5.1 (ii). This produces 9 distances as input features for each single curve.
5) Normalize the distances dF_β by computing the ratios $dF_i=dF_\beta/dF_{0.85}$ where i=1 to 8 resulting in 8 normalized distances as illustrated in Figure 5.1 (iii).
6) The distances dF_1 to dF_8 can be now employed as a table of input features for a machine learning algorithm to generate a model. To remove noise, the distances for a given substance or family of substances is averaged over 40-100 samples.

An example of tabulation of input features to generate a feature library is shown in Table 8. In Table 8, polyethylene high-density (PEHD), Polycaprolactone (PCL), glass, and silicon have been employed to generate three sets of input features. The three sets for each family form a feature

library for polymers and silica respectively and concluding the second step of the procedure as shown as an illustration in Figure 5.1(iv).

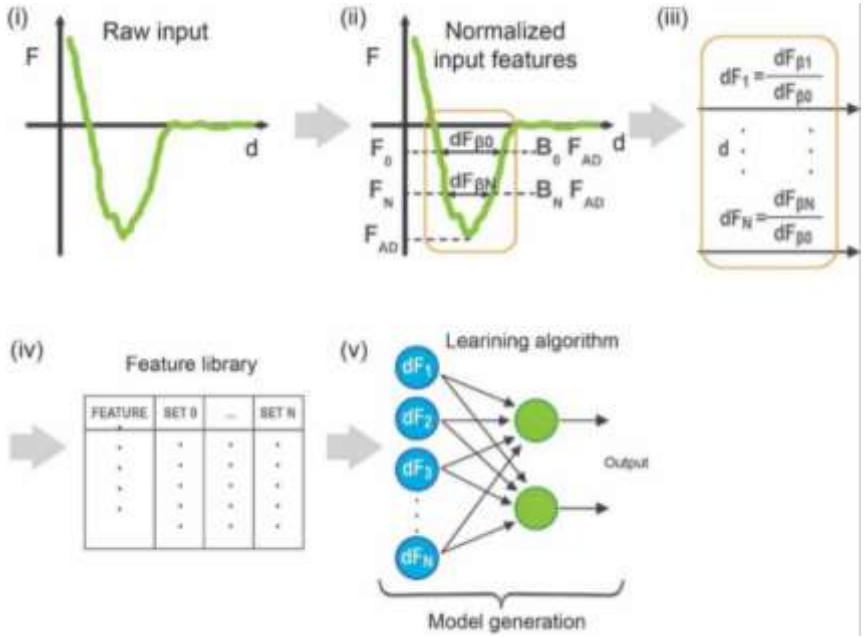

Figure 5.1 Illustrations of i) raw data, ii) input features, iii) normalized input features, to iv) generation of a feature library and v) model generation from a feature library set.

- Feeding the standardized input feature library into a learning algorithm:

The third step consists of generating a model from a feature library as shown in Figure 5.1(v). In the case of Table 8, this model should be able to identify or detect whether input features belong to the polymer family or the silica family. In order to generate models, a standard multilayer artificial neural network in Matlab[72] that included a regularization term λ to avoid overfitting is implemented. The steps are as follows:

1) Inputting an input feature library, as shown in Table 8 and in Figure 5.1(iv), into a machine learning algorithm as illustrated in Figure 5.1(v). An artificial neural network composed of U units per layer L is chosen. Units stand for unit cells or neurons and each unit is modelled with a sigmoid function where the inputs are processed by the function and the output fed into the units U in the next layer L as illustrated in Figure 5.2. The very last layer of the system will

produce the predicted outputs. In the case of Table 8, the two unit cells in the last layer L will produce the predictive outcomes for the polymer (one unit) and silica (another unit) families.

2) The model is first trained with a set of input features from a given library where the output is known. In the case of Table 8, the last unit cell for polymers should produce ones if and only if data from polymers is fed into the system and similarly for the unit cell of the silica family.
3) Then the model is tested by inputting data into the model generated from the training data and comparing the output to the known values for the output. This is typical from supervised algorithms where the algorithm learns from inputting data for which the outcome is known by the user in advance. Errors in the outcomes are quantified via Recall and Precision and together via the F-score parameter as discussed above. In the experimental section we report errors on testing sets of data via Precision, Recall and F-score.

This concludes the procedure of training and testing a model from feature libraries for substance identification. An illustration of the full process is shown in Figure 5.2.

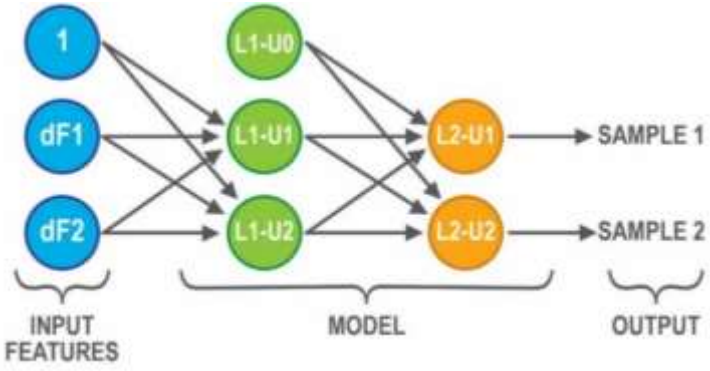

Figure 5.2 Scheme of the hierarchy and ordering of the number of layers L and number of unit cells U in the artificial neural networks.

To test the performance of the models, raw data obtained by different users and with different cantilever-sample systems are acquired and fed into trained neural networks model produced from the feature libraries. Silicon is the first testing sample. Approximately 1000 data points were collected and then fed into the trained model. The performance of the models was calculated by computing Precision,

Recall and F-score. We defined true positives=TP, false positives=FP and false negatives=FN. Then Precision=TP/(TP+FP), Recall=TP/(TP+FN) and F-score as described above in Eq. 42. The results of the test are shown in Table 8 for a 2 L 2 U model, a 3 L 3 U model and a 4 L 4 U model. Values of F-score above 0.5 indicates significant predictive power of the model to successfully identify and discriminate between the silica and polymer families. Values of 0.5 or below imply that the models lack sufficient predictive power. In Table 8, it can be seen that arbitrarily increasing L or U might not result in a better model. That is, relatively simple models, consisting of a few numbers of layers L and units U, might produce models with enough predicting power and might not be improved by increasing the complexity of the model arbitrarily[136]. Finally, zero values in the figures of merit indicate overfitting, *i.e.* the output from the unit cells in the last layer of the model is always 0.5 independently of the input.

Table 8 Example of libraries employed as the input data to generate models for two groups of materials: polymers family (Pol) and silica (Silica) family. Figures of merit for 2L-2U, 3L-2U and 4L-4U models are also shown.

β	**Set 1 Pol**	**Set 2 Pol**	**Set 3 Pol**	**Set 1 Silica**	**Set 2 Silica**	**Set 3 Silica**
0.75	0.95	0.95	0.94	0.89	0.89	0.9
0.65	0.88	0.88	0.88	0.8	0.78	0.81
0.55	0.8	0.8	0.81	0.71	0.69	0.71
0.45	0.7	0.72	0.73	0.61	0.6	0.61
0.35	0.59	0.62	0.6	0.52	0.5	0.5
0.25	0.45	0.48	0.48	0.4	0.38	0.37
0.15	0.3	0.31	0.31	0.27	0.24	0.26
0.05	0.1	0.1	0.13	0.11	0.09	0.1
	Precision		**Recall**		**F1 Score**	
2 L 2 U	0.8		0.66		0.72	
3 L 2 U	0.78		0.71		0.74	
4 L 4 U	0		0		0	

Table 9 shows a training feature library produced to discriminate between the PEHD and PCL polymers. The data was employed to train and produce models and the models were then tested. Again, we see that a single layer and two unit cells suffice to produce a model with enough predicting power to discriminate between the two samples, *i.e.* F-score > 0.5. These two examples illustrate how by using

multiple processes or steps, specific samples can be identified. That is, a given model can be employed to discriminate between families, as in the case of the silica and polymer families described above (Table 8). Then, a more specific model can be exploited to discriminate between samples in the same family (Table 9). A scheme of such flow is illustrated in Figure 5.3.

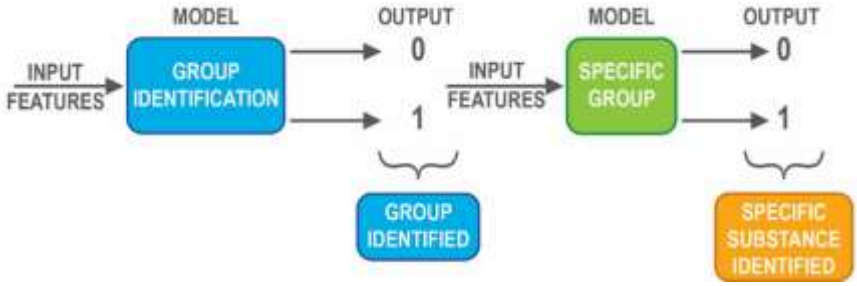

Figure 5.3 Process of identifying and classifying data first into families or groups and then into specific substances or samples.

Table 9 Example of libraries employed as input data to generate models for PCL and PEHD samples. Figures of merit for 1L-2U, a 2L-2U and a 3L-4U models obtained when feeding a test sample (PEHD) into the trained model are given

β	**Set 1 PCL**	**Set 2 PCL**	**Set 3 PCL**	**Set 1 PEHD**	**Set 2 PEHD**	**Set 3 PEHD**
0.75	0.949	0.95	0.947	0.929	0.927	0.936
0.65	0.9	0.889	0.893	0.855	0.865	0.862
0.55	0.839	0.823	0.819	0.772	0.785	0.783
0.45	0.752	0.724	0.736	0.689	0.689	0.692
0.35	0.655	0.606	0.624	0.572	0.573	0.585
0.25	0.523	0.467	0.487	0.45	0.455	0.462
0.15	0.349	0.322	0.317	0.292	0.309	0.31
0.05	0.124	0.133	0.111	0.11	0.11	0.113
	Precision		Recall		F-Score	
1 L 2 U	0.71		0.59		0.64	
2 L 2 U	0.7		0.64		0.67	
3 L 4 U	0		0		0	

Next we employ a feature library capable to differentiate between calcite and CaF_2. The surface of calcite can further be divided into two different substances or phases. The two phases are termed calcite P1 and calcite P2 and these are shown as a standard phase contrast image in Figure 5.4. The third substance of the feature library consists of CaF_2. A model with 4L and 6U that gave an F-score=1 when

tested is constructed. Then, FDCs on an 80 nm^2 area of the calcite sample were acquired. Then the model consisting of calcite P1, calcite P2 and CaF_2 was tested against the 1024 data points collected.

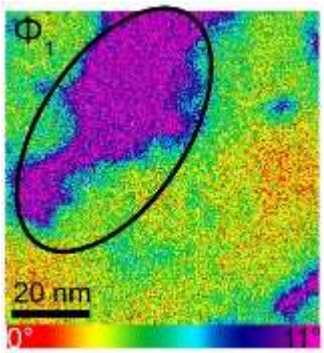

Figure 5.4 Two phases of calcite P2 (pink-purple) and calcite P1 (rest of the image) acquired as a standard phase image in dynamic AFM.

The results are shown in Figure 5.5. The black pixels imply that the model could not assign or identify (107 out of 1024 pixels) any of the three samples to that pixel. The blue pixels stand for positive identification of calcite P1 (402 out of 1024 pixels), the green stands for positive identification of calcite P2 (459 out of 1024 pixels), and the red pixels stand for positive identification of CaF_2 (56 out of 1024 pixels). There was thermally-induced drift when collecting the force data. The patch of calcite P2 in Figure 5.4 has been circled in order to relate it to the displaced patch predicted by the model in Figure 5.5. A feature library for calcite P1 and calcite P2 was then employed to generate models to discriminate between these two substances only. Figure 5.6 and Figure 5.7 show for a model of 1L-2U and 2L-3U where black pixels stand for ambiguity, blue for calcite P1 and green for calcite P2. As Figure 5.6 and Figure 5.7 look almost identical, this implies that increasing the complexity of the model from 1L-2U to 2L-3U did not significantly improve predicting power. The 1L-2U model is superior since it is equally predictive and computationally much simpler.

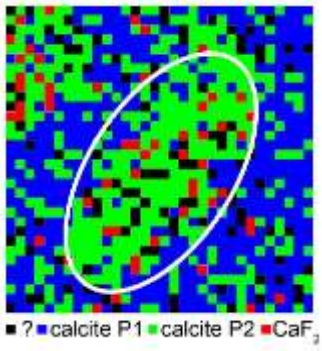

Figure 5.5 Prediction of the model produced from a feature library consisting of calcite P1 (blue), calcite P2 (green) and CaF_2 (red). The black pixels refer to pixels where the model could not predict any output unambiguously.

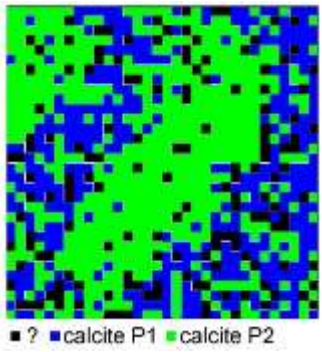

Figure 5.6 Predictions of the model (1L-2U) produced from a feature library consisting of calcite P1 (blue), and calcite P2 (green). The black pixels refer to pixels where the model couldn't guess any output unambiguously.

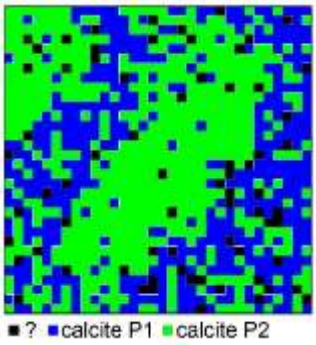

Figure 5.7 Predictions of the model (2L-3U) produced from a feature library consisting of calcite P1 (blue), and calcite P2 (green). The black pixels refer to pixels where the model couldn't guess any output unambiguously.

The generality of the model produced to discriminate between calcite P1 and calcite P2 above was tested with a second set of calcite data collected using a different AFM tip and a different calcite

sample. The results are shown in Figure 5.8. The thermal drift was much smaller in this case. A standard phase contrast image is shown in Figure 5.8a and the prediction of the model against force data for the same spot in Figure 5.8b.

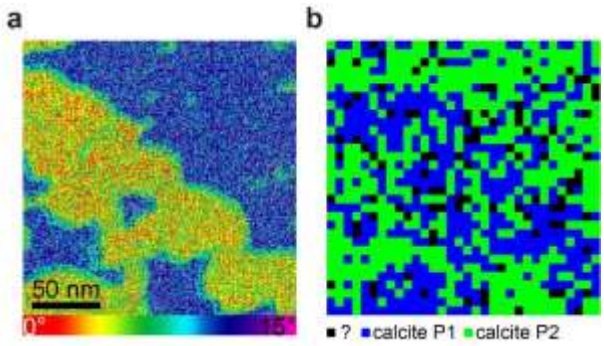

Figure 5.8 ***a****, Two phases of calcite acquired with different tip and sample phase image in dynamic AFM.* ***b****, Guess of the model produced from another data set. The blue pixels refer to calcite P1, green pixels refer to calcite P2 and black pixels refer to pixels where the model could not predict any output unambiguously.*

Conclusion

The use of libraries and models generated from libraries can be exploited to identify substances from force data alone. The generation of models should not be restricted to artificial neural networks either, but could be enhanced, or even replaced, by other methodologies. By implementing well known methodologies in machine learning such as support vector machines or Bayesian networks and exploiting them in parallel, it is possible to improve predictive power. These methods are standard in the machine learning field and packages can be found in Matlab, python and the R languages. Finally, the ordering or fabrication of libraries does not need to be limited to the air environment, but can be expanded to liquid and vacuum environments and to the use of probes other than silicon. Furthermore, the classification into families does not have to be restricted to material properties but can be enhanced to, for example, identifying the presence or absence of atomic irregularities or dislocations or identifying biological patterns or behavior of systems for which distinct features might be produced[67, 133]. The

number and type of input features might be further increased into linear or non-linear combinations of the features presented here or any other input feature, such as temperature, relative humidity, tip radius, geometry or chemistry, power law (as presented in the earlier chapter) that might enhance identification or recognition. Arguably the intuition of researchers working in a particular field will suggest the number and type of features that will make a given feature library preferable to produce a given model. In summary, the application of models and the massive testing of data should ultimately tell us what the limit of the proposed method of standardization is in the field of force spectroscopy.

Chapter 6

Application of Machine Learning Assisted Analysis

Motivation

There is arguably a divide in the approach to finding answers to questions in the general sciences. The first one has two parts and consists in either looking into phenomena and finding expressions corresponding to fundamental laws, *i.e.* the strictest theoretical side of the sciences, or explaining phenomena probed experimentally directly from these first principles as mechanisms that provide the physical insight into our findings, *i.e.* the practical science that exploits the link between experiment and the strictest theory. The second consists in asking a question to a phenomenon and finding a correlation, *i.e.* an association in general, between knowns, *i.e.* the data that we have access to, and the unknowns, *i.e.* the answer to the question we are asking. We could argue that the task of the second method to finding answers to questions was typically dealt with by statistics for a long time.

In the recent years, computer systems have considerably enabled and extended the second form of enquiry by allowing computing over extremely large sets of data, *i.e.* big data, rapidly and efficiently, *i.e.* via artificial neural networks, vector supporting machines *etc.*, and, importantly, via the exploitation of these "model free" algorithms that build a model without being explicitly programmed to, in order to

answer a specific question based on data alone[137]. Arguably, this last point defines the field of machine learning. It is now possible to find standard tool-sets in any standard programming language, *i.e.* Matlab, python or R, that include powerful machine learning algorithms to solve problems in this way[138]. It is not clear however where the two methods that we pointed out earlier meet or will meet, but we claim that it is at least advantageous to be able to answer specific questions of practical relevance from sets of data extracted from complex physical systems for which fundamental laws, principles, or expressions are either too complex or unavailable[139].

In the field of AFM, a sharp tip of nanometric radius held at the end of a microscopic cantilever is made to interact with a surface. In the interaction, complex conservative and dissipative phenomena[40] affect the dynamics of the cantilever[41]. A main paradigm of the field has been to extract information from the dynamics in relation to the properties of the surfaces[140,141]. On the other hand, the AFM field is expanding via two main fronts. Namely, 1) improving the instrumentation so the data itself contains more information of this tip-surface interaction[22,18], and 2) improving the data processing, modelling and overall understanding of the phenomena in order to understand and interpret the available data[142]. It is not clear whether more information can be extracted from the phenomena, that is, whether enhancing the resolution of the system would provide more information about the nanoscale properties of the surface, or whether we have reached a point where data processing is the main bottleneck in terms of advancing in the field[143,144]. On the other hand, we have recently shown that simple power laws[62] are not enough to explain the rich phenomena that we can probe in ambient conditions already, where air contaminants, such as CO_2 and water, start adhering to the surface almost immediately after surfaces are produced[145]. This surface phenomena under ambient conditions, which we term here "surface aging", is a dynamic process with many unknowns and possibly related to very complex physical processes such as ion exchange, surface energy dynamics and surface-tension thermodynamic equilibrium[146,147], thus providing us with a suitable candidate question in AFM. Namely, "how long has a surface been exposed to ambient conditions?". In principle, we expect that the tip-surface force will change with time and that these changes can be associated to our question via machine learning algorithms.

Application

We employ HOPG as a model sample because it is easy to cleave with standard scotch tape. In this way, HOPG is cleaved creating a pristine surface and labelled 0h. After cleaving, the sample is exposed to ambient air (temperature T≈ 22 ± 2ºC and relative humidity (RH)≈ 50% ± 5%) for 24h at which point, and in principle, phenomena, physically and known or unknown by us as researchers, will act on the surface inducing the "surface aging" phenomena. In any case, we stand by our simple question "how old is this surface in hours?" and attend simply to the outcome of the machine learning algorithm.

As input data, we employ FDCs that are standard in AFM. A complete set of experiments consists of force data taken in 6-time steps: 0h, 1h, 3h, 6h, 12h, and 24h. The data was collected with standard OLYMPUS AC160TS cantilevers with k (spring constant) ≈ 30N/m, f_0 (natural frequency) ≈ 300kHz, and Q ≈ 400. Thermal analysis was employed to calibrate the f_0 of the cantilever with the free cantilevers at ~50nm above the sample surface[148].

Since it is well known that the tip radius R strongly affects the tip-sample interaction, the effective radius R was monitored *in-situ* with the use of the critical amplitude A_C method[45] and found to be 8nm ± 2nm. It is worth noting that R could also be employed as an input feature in the machine learning model. On the whole, this capacity to easily add input features that are believed to influence the outcome or to better characterize the model is a further advantage of these methods.

After acquiring the FDCs, the profiles were parameterized into standardized input features for generating a feature or model library to answer our specific question. The way to create input features followed the same manner as described in the previous chapter. In short, F_{AD} was taken as the force reference for a given force profile and only part of the well was considered, *i.e.* F < 0 nN (net attractive). Next, a factor β was introduced for parametrizing the force curves by varying β from 0 to 1, so as F = βF_{AD}. In this work, we limited β to 0.05, 0.15, …, 0.85 and normalized these distances in the well with β = 0.85. That is, 9 absolute distances $disF_\beta$ were first computed for each single curve, and then the normalization of input features was achieved by calculating the ratios $disF_i = disF_\beta/disF_{0.85}$ where i = 1 to 8 giving 8 normalized distances. In order to remove noise, the final input features of $disF_i$ were the

average of 50 samples for each time step. Table 10 shows an example of tabulation of 2 sets final input features for each time step. A total of 56 sets input features formed the library for HOPG aging in the ambient conditions for 24h (the HOPG aging library).

Figure 6.1 shows some of the force data and an illustration of the changes on the surface that might have produce the difference in F. It is worth nothing that while we represent water-adhesion in the illustration, knowledge of this physical phenomenon is, in principle, not required in order to answer our question. At this point, machine learning allows for 1) unsupervised machine learning methods to extract candidate input features for our supervised model, *i.e*. the generation of the library, or 2) our physical insight and inspection of the data to select input feature candidates. As explained, we decided to exploit the normalized distance in the well. We could have also reduced the data to nm units rather than go for pure normalization as we have done. This might have enhanced the explanatory power of our candidate features and might be something to explore in the future. On the other hand, in order to generate what is considered standard in machine learning, *i.e*. purely normalized features, here, we did not follow this possibility through. Instead, the features were directly fed into the Neural Network with L layers and U units per layer and an optimum model, *i.e*. the optimum number of layers and units, was sought.

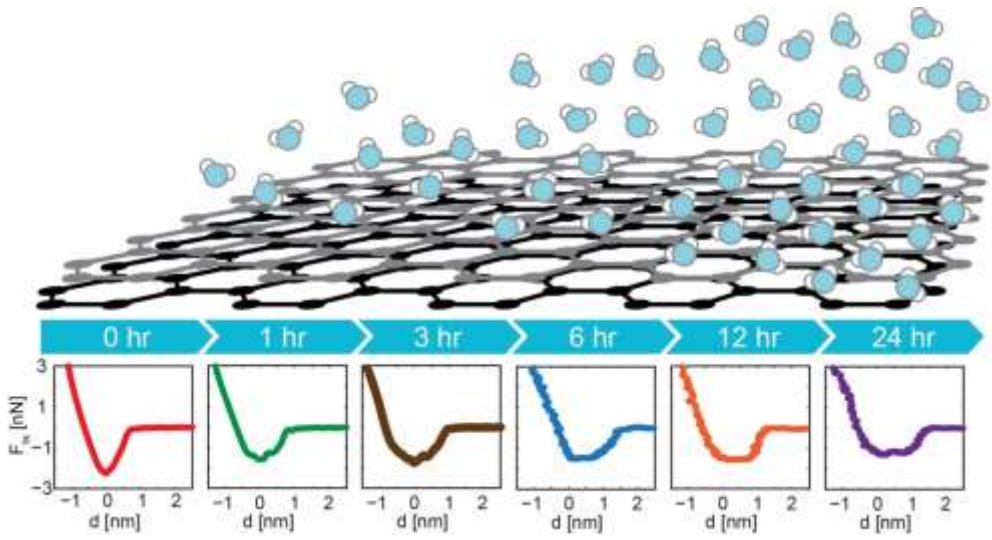

Figure 6.1 Evolution of force profiles as a function of d at each time step in hours h, as the HOPG surface ages by exposure to ambient conditions.

Table 10 Example of 2 sets of input features for each time step used to generate models for HOPG aging in the ambient conditions.

β	**0 h**	**0 h**	**1 h**	**1 h**	**3 h**	**3 h**	**6 h**	**6 h**	**12 h**	**12 h**	**24 h**	**24 h**
	Set 1	**Set 2**	**Set 1**	**Set 2**	**Set 1**	**Set 2**	**Set 1**	**Set 2**	**Set 1**	**Set 2**	**Set 1**	**Set 2**
0.75	0.91	0.94	0.92	0.87	0.93	0.92	0.92	0.93	0.93	0.95	0.93	0.92
0.65	0.83	0.84	0.85	0.81	0.88	0.87	0.85	0.85	0.88	0.89	0.85	0.84
0.55	0.75	0.77	0.79	0.75	0.82	0.81	0.79	0.78	0.83	0.83	0.76	0.76
0.45	0.67	0.69	0.68	0.68	0.74	0.73	0.72	0.72	0.76	0.76	0.67	0.66
0.35	0.57	0.60	0.61	0.57	0.63	0.63	0.65	0.65	0.67	0.67	0.57	0.57
0.25	0.44	0.47	0.50	0.48	0.50	0.47	0.55	0.55	0.53	0.54	0.46	0.48
0.15	0.30	0.31	0.34	0.32	0.29	0.29	0.39	0.38	0.33	0.37	0.32	0.32
0.05	0.12	0.12	0.13	0.12	0.11	0.12	0.15	0.15	0.15	0.13	0.14	0.13

The performance of the models was carried out by computing Precision, Recall and F_1 score to obtain validation sets as typically carried out in machine learning. In standard supervised machine learning, the last step consists in submitting the model to predict from data that the model has never been exposed to. This provides final validation of the predicting power of the model. We took the model with the highest $F_1 \approx 0.78$ score reported in Table 11 (3L and 3U). FDCs were acquired with a different tip that otherwise had similar R and the data was labelled as before in order to confirm the viability of the model. A total about 1800 data points were collected, processed and averaged over 45 samples for 1h, 6h, and 24h time steps. Then we tested the 3L-3U model against the testing datasets. We took 80%, *i.e*. 0.8 where the range is 0 to 1 in our hypothesis function h, as the trigger for the hypothesis value. The outcome of the predictions is shown in Figure 6.2. The open circles stand for either false positives or false negatives while the solid circles indicate true positives or true negatives, *i.e*. correct predictions. At 1h, the model produced only 7 false positive predictions (out of 16 averaged data points) with most of them falling into the 0h range. On the other hand, there were 10 true positives and 6 false negatives for the 1h set. The results are 5 and 1 false positive predictions for the 6h (8 false negatives and 9 true positives) and 24h (3 false negatives and 13 true positives) datasets respectively (out of 17 averaged data points per set).

Table 11 Performance of 2L 2U, 2L 3U, 3L 2U, 3L 3U, 4L 2U and 4L 4U models are evaluated by Precision, Recall and F_1 score.

	Precision	**Recall**	**F_1 score**
2 L 2 U	0.79	0.63	0.70
2 L 3 U	0.79	0.63	0.70
3 L 2 U	0.93	0.58	0.72
3 L 3 U	0.82	0.75	0.78
4 L 2 U	1.00	0.00	0.00
4 L 3 U	0.75	0.75	0.75

Conclusion

In summary, we have shown that a standard Neural Network can be employed to predict the time of exposure to ambient conditions of a graphite surface directly from AFM data. No knowledge of the underlying phenomena is required in order to predict the time of exposure. The complexity of the data and difficulty to reduce it to well-known first principles might in fact act as a positive in terms of enhancing the predictive power of the model since Neural Networks, and other Machine Learning methods, benefit from complexity and detail. We propose that model libraries can be employed to predict phenomena that answers to specific questions in AFM and that libraries could be customized and standardized in a similar way to that presented here.

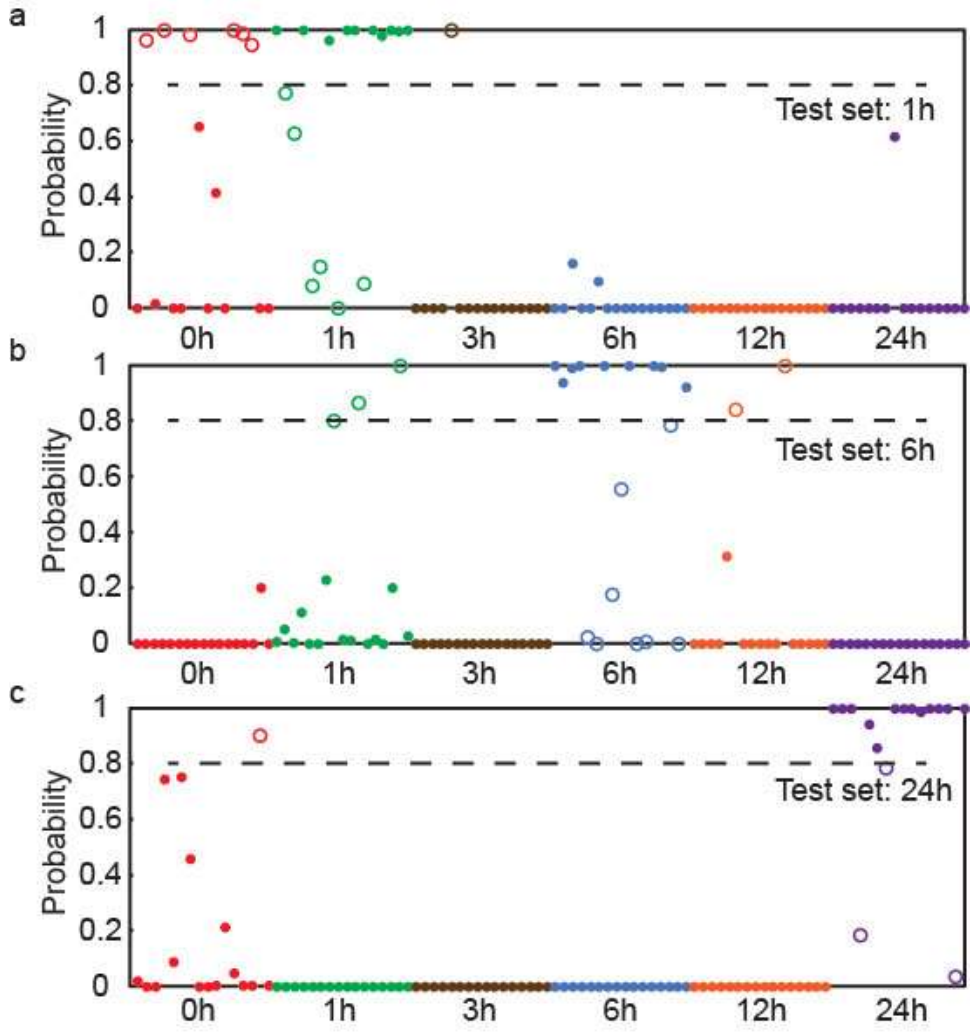

Figure 6.2 Model exploiting a 3L3U model against input data sets at 1h (a), 6h (b), 24h (c). The open circles stand for incorrect predictions while the solid circles indicate correct predictions. Color codes are used to refer to the specific time steps: 0 h (red), 1 h (green), 3 h (brown), 6 h (blue), 12 h (orange), and 24 h (purple).

Chapter 7

Wrapping up

Future work

This thesis has provided methodologies to investigate and characterize material surface properties. Furthermore, machine learning assisted data analysis approach is applied and shown to be a promising direction to advance. There are possible extensions to the current work:

- The methodologies developed in this thesis are based on the experiments carried out in the ambient environment. This experimental condition may pose great limitation to the investigation of soft or biological samples that are viable normally in liquid conditions. Therefore, extending the methodologies applicability to liquid conditions could be one of the extensions.
- As mentioned earlier, the machine learning algorithm applied in this thesis is artificial neuron networks. There are other algorithms like support vector machines or Bayesian networks that could be used as alternatives or exploited in parallel, and this could possibly improve the predictive power. In addition, it has only been explored the use of one parameter as features to train the machine learning model. Combining more parameters such as tip radius or Hamaker

coefficient, *etc*. as input features could be another extension that worth investigating to advance the method.

Summary

In this thesis, we have explored two ambits of current interest in AFM. First, that of force reconstruction under ambient conditions. Second, that of image generation in multifrequency AFM. Arguably, the tip-surface force profile contains all the phenomena to be explained in AFM. In particular, the force profile can, and has typically been, explained by models where the motivation typically is 1) to show that a given model, derived or known, matches the experimental force profile to then 2) extract material properties from it by either providing as much accuracy as possible, increasing throughput, or both.

The first part of this thesis has been dedicated to inspecting the force profile as such without necessarily concerning ourselves with throughput. Our approach has shown that:

1) the "attractive" part of the force profile, typically explained as the short-range van der Waals force, does not necessarily follow an inverse square law (2015). In fact, we have shown that power laws of 3-20 might be more suitable to describe the tip-surface interaction. This is particularly the case for the sharpest tips, i.e. R<20-30 nm. Above this range the power law can be said to be 2 as expected. This point puts restrictions to high resolution imaging since the shape of the force becomes tip size dependent. We could argue that it might be possible to modify the force expression to make it independent of the tip size by we have not shown it in our work.

2) That the procedure of recovering the force profile does not necessarily meet the conditions required by the central limit theorem which in turn allows assuming a normal distribution (2015). The main consequence of this second point is that acquiring ~30 data points per pixel does not justify

providing a mean and a standard deviation of a given parameter in order to account for errors. That is, the true mean is not correctly explained in this way implying that reproducibility is at stake.

3) This first part of the thesis finishes with a machine learning approach to parametrizing the force (2016, 2019). With this approach the underlying phenomena can manifest via any complexity, i.e. power law, or any possible distance dependent function, without restricting accuracy, overfitting or bias. A disadvantage of this method is the increase of complexity in the computational side. In the thesis, this part is actually presented last.

The second part of the thesis is dedicated to recover the parameter that is typically used to explain the van der Waals force interaction in ambient AFM, i.e. the Hamaker or H, while imaging. We assumed a power law of 2 (2015) and analytically derived an expression to recover H by exploiting the multifrequency method. Here two expressions are experimentally available and describe the attractive part of the force in terms of two unknowns, i.e. the minimum tip-surface distance and H. We have shown that an analytical solution is available assuming certain standard experimental conditions. This approach allows for higher throughput and quasi-instantaneous recovery of the parameter while performing standard multifrequency imaging.

Bibliography

1. Miyoshi, K., Surface characterization techniques: An overview. In *Mechanical Tribology*, CRC Press: 2004; pp 45-68.
2. Binnig, G.; Quate, C. F.; Gerber, C., Atomic Force Microscope. *Phys. Rev. Lett.* **1986,** *56*, 930-933.
3. Zhong, Q.; Inniss, D.; Kjoller, K.; Elings, V. B., Fractured polymer/silica fiber surface studied by tapping mode atomic force microscopy. *Surface Science* **1993,** *290* (1–2), L688-L692.
4. Albrecht, T. R.; Grutter, P.; Horne, D.; Rugar, D., Frequency modulation detection using high-Q cantilevers for enhanced force microscope sensitivity. *Journal of Applied Physics* **1991,** *69* (2), 668-673.
5. Paulo, A. S.; Garcia, R., Tip-surface, amplitude, and energy dissipation in amplitude-modulation (tapping mode) force microscopy. *Physical Review B* **2001,** *64*, 193411-193414.
6. Gauthier, M.; Pérez, R.; Arai, T.; Tomitori, M.; Tsukada, M., Interplay between Nonlinearity, Scan Speed, Damping, and Electronics in Frequency Modulation Atomic-Force Microscopy. *Phys. Rev. Lett.* **2002,** *89*, 146104.
7. Dürig, U., Interaction sensing in dynamic force microscopy. *NJPh* **2000,** *2*, 5-5.
8. Stark, R. W.; Heckl, W. M., Fourier transformed atomic force microscopy: tapping mode atomic force microscopy beyond the Hookian approximation. *Surface Science* **2000,** *457* (1), 219-228.
9. Stark, R. W., Spectroscopy of higher harmonics in dynamic atomic force microscopy. *Nanotechnology* **2003,** *15* (3), 347-351.
10. Radmacher, M.; Fritz, M.; Hansma, P. K., Imaging soft samples with the atomic force microscope: gelatin in water and propanol. *Biophysical journal* **1995,** *69* (1), 264-270.
11. García, R.; Magerle, R.; Perez, R., Nanoscale compositional mapping with gentle forces. *Nature materials* **2007,** *6*, 405.
12. Jalili, N.; Laxminarayana, K., A review of atomic force microscopy imaging systems: application to molecular metrology and biological sciences. *Mechatronics* **2004,** *14* (8), 907-945.
13. Rodriguez, T.; Garcia, R., Compositional mapping of surfaces in atomic force microscopy by excitation of the second normal mode of the microcantilever. *Appl. Phys. Lett.* **2004,** *84* (3), 449-551.
14. Kawai, S.; Glatzel, T.; Koch, S.; Such, B.; Baratoff, A.; Meyer, E., Systematic Achievement of Improved Atomic-Scale Contrast via Bimodal Dynamic Force Microscopy. *Phys. Rev. Lett.* **2009,** *103* (22), 220801.
15. Aksoy, M. D.; Atalar, A., Force spectroscopy using bimodal frequency modulation atomic force microscopy. *Physical Review B* **2011,** *83* (7), 075416.
16. Martinez-Martin, D.; Herruzo, E. T.; Dietz, C.; Gomez-Herrero, J.; Garcia, R., Noninvasive Protein Structural Flexibility Mapping by Bimodal Dynamic Force Microscopy. *Phys. Rev. Lett.* **2011,** *106*, 198101-198104.

17. Herruzo, E. T.; Perrino, A. P.; Garcia, R., Fast nanomechanical spectroscopy of soft matter. *Nat Commun* **2014,** *5*, 3126.
18. Sahin, O.; Magonov, S.; Su, C.; Quate, C. F.; Solgaard, O., An atomic force microscope tip designed to measure time-varying nanomechanical forces. *Nature nanotechnology* **2007,** *2*, 507.
19. Solares, S. D.; An, S.; Long, C. J. Multi-frequency tapping-mode atomic force microscopy beyond three eigenmodes in ambient air *Beilstein journal of nanotechnology* [Online], 2014, p. 1637-1648. PubMed. http://europepmc.org/abstract/MED/25383276

http://europepmc.org/articles/PMC4222484?pdf=render

http://europepmc.org/articles/PMC4222484

https://doi.org/10.3762/bjnano.5.175 (accessed 2014).
20. Raman, A.; Trigueros, S.; Cartagena, A.; Stevenson, A. P. Z.; Susilo, M.; Nauman, E.; Contera, S. A., Mapping nanomechanical properties of live cells using multi-harmonic atomic force microscopy. *Nature nanotechnology* **2011,** *6*, 809.
21. Platz, D.; Tholén, E. A.; Pesen, D.; Haviland, D. B., Intermodulation atomic force microscopy. *Appl. Phys. Lett.* **2008,** *92* (15), 153106.
22. Jesse, S.; Vasudevan, R. K.; Collins, L.; Strelcov, E.; Okatan, M. B.; Belianinov, A.; Baddorf, A. P.; Proksch, R.; Kalinin, S. V., Band excitation in scanning probe microscopy: recognition and functional imaging. *Annu. Rev. Phys. Chem.* **2014,** *65* (1), 519-36.
23. Kalinin, S. V.; Strelcov, E.; Belianinov, A.; Somnath, S.; Vasudevan, R. K.; Lingerfelt, E. J.; Archibald, R. K.; Chen, C.; Proksch, R.; Laanait, N.; Jesse, S., Big, Deep, and Smart Data in Scanning Probe Microscopy. *ACS nano* **2016,** *10* (10), 9068-9086.
24. Huang, B.; Li, Z.; Li, J., An artificial intelligence atomic force microscope enabled by machine learning. *Nanoscale* **2018,** *10* (45), 21320-21326.
25. Sokolov, I.; Dokukin, M. E.; Kalaparthi, V.; Miljkovic, M.; Wang, A.; Seigne, J. D.; Grivas, P.; Demidenko, E., Noninvasive diagnostic imaging using machine-learning analysis of nanoresolution images of cell surfaces: Detection of bladder cancer. *Proceedings of the National Academy of Sciences* **2018,** *115* (51), 12920-12925.
26. Mencattini, A.; Casti, P.; Fazio, G.; Ghibelli, L.; Luce, M.; Cricenti, A.; Martinelli, E.; Natale, C. D., Uncertainty Evaluation of a VBM System for AFM Study of Cell-Cerium Oxide Nanoparticles Interactions. *IEEE Trans. Instrum. Meas.* **2018,** *67* (7), 1564-1572.
27. Eslami, B.; Damircheli, M., Biharmonic versus bimodal AFM: Numerical and experimental study on soft matter. *Journal of Applied Physics* **2019,** *126* (9), 095301.
28. Santos, S.; Barcons, V.; Christenson, H. K.; Billingsley, D. J.; Bonass, W. A.; Font, J.; Thomson, N. H., Stability, resolution, and ultra-low wear amplitude modulation atomic force microscopy of DNA: Small amplitude small set-point imaging. *Appl. Phys. Lett.* **2013,** *103* (6), 063702.
29. Lai, C.-Y.; Tang, T.-C.; Amadei, C. A.; Marsden, A. J.; Verdaguer, A.; Wilson, N.; Chiesa, M., A nanoscopic approach to studying evolution in graphene wettability. *Carbon* **2014,** *80* (0), 784-792.
30. Amir, F. P.; Daniel, M.-J.; Ricardo, G., Force reconstruction from tapping mode force microscopy experiments. *Nanotechnology* **2015,** *26* (18), 185706.
31. Giessibl, F. J., Advances in atomic force microscopy. *Reviews of modern physics* **2003,** *75* (3), 949.

32. Katan, A. J.; Oosterkamp, T. H., Measuring Hydrophobic Interactions with Three-Dimensional Nanometer Resolution. *The Journal of Physical Chemistry C* **2008,** *112* (26), 9769-9776.
33. Giessibl, F. J., Forces and frequency shifts in atomic-resolution dynamic-force microscopy. *Physical Review B* **1997,** *56* (24), 16010-16015.
34. Steidel, R., *An introduction to mechanical vibrations*. 3 ed.; John Wiley & Sons: 1989.
35. Sader, J. E.; Jarvis, S. P., Accurate formulas for interaction force and energy in frequency modulation force spectroscopy. *Appl. Phys. Lett.* **2004,** *84* (10), 1801-1803.
36. Lee, M.; Jhe, W., General Theory of Amplitude-Modulation Atomic Force Microscopy. *Phys. Rev. Lett.* **2006,** *97* (3), 036104.
37. Hu, S.; Raman, A., Inverting amplitude and phase to reconstruct tip–sample interaction forces in tapping mode atomic force microscopy. *Nanotechnology* **2008,** *19* (37), 375704.
38. Katan, A. J.; van Es, M. H.; Oosterkamp, T. H., Quantitative force versus distance measurements in amplitude modulation AFM: a novel force inversion technique. *Nanotechnology* **2009,** *20* (16), 165703.
39. Tzu-Chieh Tang; Carlo A. Amadei; Neil H. Thomson; Chiesa, M., Ion-Exchange and DNA Molecular Dip-Sticks: Studying the Nanoscale Surface Wetting of Muscovite Mica. *Journal of Physical Chemistry C* **2014,** *Accepted*.
40. Cleveland, J. P.; Anczykowski, B.; Schmid, A. E.; Elings, V. B., Energy dissipation in tapping-mode atomic force microscopy. *Appl. Phys. Lett.* **1998,** *72* (20), 2613-2615.
41. Tamayo, J.; Garcia, R., Relationship between phase shift and energy dissipation in tapping-mode scanning force microscopy. *Appl. Phys. Lett.* **1998,** *73* (20), 2926-2928.
42. Garcia, R.; San Paulo, A., Amplitude curves and operating regimes in dynamic atomic force microscopy. *Ultramicroscopy* **2000,** *82* (1-4), 79-83.
43. Garcia, R.; San Paulo, A., Attractive and repulsive tip-sample interaction regimes in tapping-mode atomic force microscopy. *Physical Review B* **1999,** *60* (7), 4961.
44. Ramos, J., Tip radius preservation for high resolution imaging in amplitude modulation atomic force microscopy. *Appl. Phys. Lett.* **2014,** *105* (4), 043111-043114.
45. Santos, S.; Guang, L.; Souier, T.; Gadelrab, K.; Chiesa, M.; Thomson, N. H., A method to provide rapid in situ determination of tip radius in dynamic atomic force microscopy. *Review of Scientific Instruments* **2012,** *83* (4), 043707.
46. Maragliano, C.; Glia, A.; Stefancich, M.; Chiesa, M., Effective AFM cantilever tip size: methods for in-situ determination. *Meas. Sci. Technol.* **2014,** *26* (1), 015002.
47. Goldstein, H.; Poole, C. P.; Safko, J. L., *Classical Mechanics* 3rd ed.; Addison-Wesley: 2001.
48. R. M. Brydson; C. Hammond; D. Mowbray; M. R. J. Gibbs; I. Todd; M. Grell; I. W. Hamley; M. Geoghegan; Jones, R. A. L.; G. J. Leggett, *Nanoscale Science and Technology*. Wiley: Chichester, 2005.
49. Hamaker, H. C., The London—van der Waals attraction between spherical particles. *Physica* **1937,** *4* (10), 1058-1072.
50. Israelachvili, J., *Intermolecular & Surface Forces*. 2 ed.; Academic Press: New York, 1991.
51. Visser, J., On Hamaker constants: A comparison between Hamaker constants and Lifshitz-van der Waals constants. *Advances in Colloid and Interface Science* **1972,** *3* (4), 331-363.

52. Clauset, A.; Shalizi, C. R.; Newman, M. E. J., Power-Law Distributions in Empirical Data. *SIAMR* **2009,** *51* (4), 661-703.
53. Goldstein, M. L.; Morris, S. A.; Yen, G. G., Problems with fitting to the power-law distribution. *Eur. Phys. J. B* **2004,** *41* (2), 255-258.
54. García, R.; Pérez, R., Dynamic atomic force microscopy methods. *Surface Science Reports* **2002,** *47* (6), 197-301.
55. Zitzler, L.; Herminghaus, S.; Mugele, F., Capillary forces in tapping mode atomic force microscopy. *Physical Review B* **2002,** *66* (15), 155436.
56. Lo Iacono, F.; Bologna, N.; Diamanti, M. V.; Chang, Y.-H.; Santos, S.; Chiesa, M., General Parametrization of Persisting Long-Range Nanoscale Phenomena in Force Measurements Emerging under Ambient Conditions. *The Journal of Physical Chemistry C* **2015,** *119* (23), 13062-13067.
57. Casimir, H. B. G.; Polder, D., The Influence of Retardation on the London-van der Waals Forces. *Phys. Rev.* **1948,** *73* (4), 360-372.
58. Derjaguin, B.; Muller, V.; Toporov, Y. P., Effect of contact deformations on the adhesion of particles. *J. Colloid Interface Sci.* **1975,** *53* (2), 314-326.
59. Krajina, B. A.; Kocherlakota, L. S.; Overney, R. M., Direct determination of the local Hamaker constant of inorganic surfaces based on scanning force microscopy. *The Journal of chemical physics* **2014,** *141* (16), 164707.
60. Wang, S.; Zhang, Y.; Abidi, N.; Cabrales, L., Wettability and Surface Free Energy of Graphene Films. *Langmuir : the ACS journal of surfaces and colloids* **2009,** *25* (18), 11078-11081.
61. Hammer, M. U.; Anderson, T. H.; Chaimovich, A.; Scott Shell, M.; Israelachvili, J., The search for the hydrophobic force law. *Faraday Discuss.* **2010,** *146*, 299-401.
62. Lai, C.-Y.; Olukan, T.; Santos, S.; Al Ghaferi, A.; Chiesa, M., The power laws of nanoscale forces under ambient conditions. *Chem. Commun.* **2015,** *51* (99), 17619-17622.
63. Souier, T.; Santos, S.; Al Ghaferi, A.; Stefancich, M.; Chiesa, M., Enhanced electrical properties of vertically aligned carbon nanotube-epoxy nanocomposites with high packing density. *Nanoscale research letters* **2012,** *7* (1), 630.
64. Masuda, T.; Matsuki, Y.; Shimoda, T., Spectral parameters and Hamaker constants of silicon hydride compounds and organic solvents. *J. Colloid Interface Sci.* **2009,** *340* (2), 298-305.
65. Liu, Y. H.; Wang, D.; Nakajima, K.; Zhang, W.; Hirata, A.; Nishi, T.; Inoue, A.; Chen, M. W., Characterization of Nanoscale Mechanical Heterogeneity in a Metallic Glass by Dynamic Force Microscopy. *Phys. Rev. Lett.* **2011,** *106* (12), 125504.
66. Tai, K.; Dao, M.; Suresh, S.; Palazoglu, A.; Ortiz, C., Nanoscale heterogeneity promotes energy dissipation in bone. *Nature materials* **2007,** *6*, 454.
67. Dokukin, M. E.; Guz, N. V.; Gaikwad, R. M.; Woodworth, C. D.; Sokolov, I., Cell Surface as a Fractal: Normal and Cancerous Cervical Cells Demonstrate Different Fractal Behavior of Surface Adhesion Maps at the Nanoscale. *Phys. Rev. Lett.* **2011,** *107* (2), 028101.
68. Walpole, R. E.; Myers, R. H.; Myers, S. L., *Probability and statistics for engineers and scientists*. 6th ed.; Prentice Hall: Upper Saddle River, N.J., 1998; p xii, 739 p.
69. Cartagena, A.; Raman, A., Local Viscoelastic Properties of Live Cells Investigated Using Dynamic and Quasi-Static Atomic Force Microscopy Methods. *Biophysical Journal* **2014,** *106* (5), 1033-1043.

70. Voss, A.; Dietz, C.; Stocker, A.; Stark, R., Quantitative measurement of the mechanical properties of human antibodies with sub-10-nm resolution in a liquid environment. *Nano research* **2015,** *8* (6), 1987-1996.
71. Forchheimer, D.; Forchheimer, R.; Haviland, D. B., Improving image contrast and material discrimination with nonlinear response in bimodal atomic force microscopy. *Nat Commun* **2015,** *6*.
72. Matlab and Simulink, R. a., The MathWorks, Inc. Natick *Matlab and Simulink, Release 2010a*, The MathWorks, Inc. Natick 2010.
73. Miller, I., Probability, Random Variables, and Stochastic Processes. *Technometrics* **1966,** *8* (2), 378-380.
74. Sugimoto, Y.; Pou, P.; Abe, M.; Jelinek, P.; Pérez, R.; Morita, S.; Custance, O., Chemical identification of individual surface atoms by atomic force microscopy. *Nature* **2007,** *446* (7131), 64-67.
75. Gross, L.; Mohn, F.; Moll, N.; Liljeroth, P.; Meyer, G., The Chemical Structure of a Molecule Resolved by Atomic Force Microscopy. *Science* **2009,** *325* (5944), 1110-1114.
76. Banwell, C. N.; McCash, E. M., *Fundamentals of molecular spectroscopy*. McGraw-Hill New York: 1994; Vol. 851.
77. Gross, L.; Mohn, F.; Moll, N.; Meyer, G.; Ebel, R.; Abdel-Mageed, W. M.; Jaspars, M., Organic structure determination using atomic-resolution scanning probe microscopy. *Nat. Chem.* **2010,** *2*, 821.
78. Hadlington, S. Nanostripe controversy in new twist. http://www.rsc.org/chemistryworld/2014/11/nanostripe-stripeynanoparticle-controversy-new-twist.
79. Setvín, M.; Mutombo, P.; Ondráček, M.; Majzik, Z.; Švec, M.; Cháb, V.; Ošťádal, I.; Sobotík, P.; Jelínek, P., Chemical Identification of Single Atoms in Heterogeneous III–IV Chains on Si(100) Surface by Means of nc-AFM and DFT Calculations. *ACS nano* **2012,** *6* (8), 6969-6976.
80. Lifshitz, E., The theory of molecular attractive forces between solids. *Sov.Phys.JETP* **1956,** *2*, 73-83.
81. John, E. S.; Takayuki, U.; Michael, J. H.; Alan, F.; Yoshikazu, N.; Suzanne, P. J., Quantitative force measurements using frequency modulation atomic force microscopy—theoretical foundations. *Nanotechnology* **2005,** *16* (3), S94.
82. Lozano, J. R.; Garcia, R., Theory of Multifrequency Atomic Force Microscopy. *Phys. Rev. Lett.* **2008,** *100* (7), 076102.
83. Santos, S., Phase contrast and operation regimes in multifrequency atomic force microscopy. *Appl. Phys. Lett.* **2014,** *104* (14), 143109.
84. Lozano, J. R.; Garcia, R., Theory of phase spectroscopy in bimodal atomic force microscopy. *Physical Review B* **2009,** *79* (1), 014110.
85. Lai, C.-Y.; Perri, S.; Santos, S.; Garcia, R.; Chiesa, M., Rapid quantitative chemical mapping of surfaces with sub-2nm resolution. *Nanoscale* **2016**.
86. Santos, S.; Barcons, V.; Font, J.; Thomson, N. H., Bi-stability of amplitude modulation AFM in air: deterministic and stochastic outcomes for imaging biomolecular systems. *Nanotechnology* **2010,** *21* (22), 225710.
87. Tamayo, J.; Garcia, R., Deformation, contact time, and phase contrast in tapping mode scanning force microscopy. *Langmuir : the ACS journal of surfaces and colloids* **1996,** *12* (18), 4430-4435.

88. Garcia, R.; Gómez, C. J.; Martinez, N. F.; Patil, S.; Dietz, C.; Magerle, R., Identification of Nanoscale Dissipation Processes by Dynamic Atomic Force Microscopy. *Phys. Rev. Lett.* **2006,** *97* (1), 016103.
89. Baxamusa, S. H.; Gleason, K. K., Random Copolymer Films with Molecular-Scale Compositional Heterogeneities that Interfere with Protein Adsorption. *Adv. Funct. Mater.* **2009,** *19* (21), 3489-3496.
90. Amadei, C. A.; Yang, R.; Chiesa, M.; Gleason, K. K.; Santos, S., Revealing amphiphilic nanodomains of anti-biofouling polymer coatings. *ACS applied materials & interfaces* **2014,** *6* (7), 4705-4712.
91. Sheng, J., *Enhanced oil recovery field case studies*. Gulf Professional Publishing: 2013.
92. Anderson, W. G., Wettability Literature Survey-Part 6: The Effects of Wettability on Waterflooding. *Journal of Petroleum Technology* **1987,** *39* (12), 1605-1622.
93. Craig, F. F., *The reservoir engineering aspects of waterflooding*. Richardson, TX: Henry L. Doherty Memorial Fund of AIME, Society of Petroleum Engineers: 1993.
94. Morrow, N. R.; Lim, H. T.; Ward, J. S., Effect of Crude-Oil-Induced Wettability Changes on Oil Recovery. *SPE Formation Evaluation* **1986,** *1* (01), 89-103.
95. du Petrole, F.; Malmaison, F. R., Evaluation of reservoir wettability and its effect on oil recovery. In *Interfacial Phenomena in Petroleum Recovery*, CRC Press Boca Raton, FL: 1990; p 319.
96. Karabakal, U.; Bagci, S., Determination of Wettability and Its Effect on Waterflood Performance in Limestone Medium. *Energy Fuels* **2004,** *18* (2), 438-449.
97. Al-Mjeni, R.; Arora, S.; Cherukupalli, P.; Van Wunnik, J.; Edwards, J.; Felber, B. J.; Gurpinar, O.; Hirasaki, G. J.; Miller, C. A.; Jackson, C., Has the Time Come for EOR? *Oilfield review* **2010,** *22* (4), 16-35.
98. Fathi, S. J.; Austad, T.; Strand, S., Effect of Water-Extractable Carboxylic Acids in Crude Oil on Wettability in Carbonates. *Energy Fuels* **2011,** *25* (6), 2587-2592.
99. Buckley, J. S.; Liu, Y., Some mechanisms of crude oil/brine/solid interactions. *Journal of Petroleum Science and Engineering* **1998,** *20* (3–4), 155-160.
100. Greenberg, M. L.; Castagna, J. P., SHEAR-WAVE VELOCITY ESTIMATION IN POROUS ROCKS: THEORETICAL FORMULATION, PRELIMINARY VERIFICATION AND APPLICATIONS1. *GeopP* **1992,** *40* (2), 195-209.
101. Zhiting Li; Yongjin Wang; Andrew Kozbial; Ganesh Shenoy; Feng Zhou; Rebecca McGinley; Patrick Ireland; Brittni Morganstein; Alyssa Kunkel; Sumedh P. Surwade; Li, L.; Liu, H., Effect of airborne contaminants on the wettability of supported graphene and graphite. *Nature Material* **2013,** *12* (10), 925-031.
102. Nicolás, F. M.; Ricardo, G., Measuring phase shifts and energy dissipation with amplitude modulation atomic force microscopy. *Nanotechnology* **2006,** *17* (7), S167.
103. Freeman, J.; Dale, R., Assessing bimodality to detect the presence of a dual cognitive process. *Behav. Res. Methods* **2013,** *45* (1), 83-97.
104. Chu, D. H.; Vinoba, M.; Bhagiyalakshmi, M.; Hyun Baek, II; Nam, S. C.; Yoon, Y.; Kim, S. H.; Jeong, S. K., CO2 mineralization into different polymorphs of CaCO3 using an aqueous-CO2 system. *RSC Advances* **2013,** *3* (44), 21722-21729.
105. Kendall, T. A.; Martin, S. T., Water-Induced Reconstruction that Affects Mobile Ions on the Surface of Calcite. *The Journal of Physical Chemistry A* **2007,** *111* (3), 505-514.
106. Nakata, K.; Fujishima, A., TiO2 photocatalysis: Design and applications. *Journal of Photochemistry and Photobiology C: Photochemistry Reviews* **2012,** *13* (3), 169-189.

107. Galstyan, V.; Comini, E.; Faglia, G.; Sberveglieri, G., TiO2 nanotubes: recent advances in synthesis and gas sensing properties. *Sensors* **2013,** *13* (11), 14813-14838.
108. Carp, O.; Huisman, C. L.; Reller, A., Photoinduced reactivity of titanium dioxide. *Prog. Solid State Chem.* **2004,** *32* (1), 33-177.
109. Kenanakis, G.; Vernardou, D.; Dalamagkas, A.; Katsarakis, N., Photocatalytic and electrooxidation properties of TiO2 thin films deposited by sol–gel. *Catal. Today* **2015,** *240*, 146-152.
110. Seki, K.; Tachiya, M., Kinetics of Photoinduced Hydrophilic Conversion Processes of TiO2 Surfaces. *The Journal of Physical Chemistry B* **2004,** *108* (15), 4806-4810.
111. Li, Y.; Sasaki, T.; Shimizu, Y.; Koshizaki, N., A Hierarchically Ordered TiO2 Hemispherical Particle Array with Hexagonal-Non-Close-Packed Tops: Synthesis and Stable Superhydrophilicity Without UV Irradiation. *Small* **2008,** *4* (12), 2286-2291.
112. Watanabe, T.; Nakajima, A.; Wang, R.; Minabe, M.; Koizumi, S.; Fujishima, A.; Hashimoto, K., Photocatalytic activity and photoinduced hydrophilicity of titanium dioxide coated glass. *Thin Solid Films* **1999,** *351* (1), 260-263.
113. Diamanti, M. V.; Gadelrab, K. R.; Pedeferri, M. P.; Stefancich, M.; Pehkonen, S. O.; Chiesa, M., Nanoscale Investigation of Photoinduced Hydrophilicity Variations in Anatase and Rutile Nanopowders. *Langmuir : the ACS journal of surfaces and colloids* **2013,** *29* (47), 14512-14518.
114. http://rruff.info/repository/sample_child_record_powder/by_minerals/Anatase__R070582-9__Powder__Xray_Data_XY_RAW__9462.txt.
115. Miyauchi, M.; Kieda, N.; Hishita, S.; Mitsuhashi, T.; Nakajima, A.; Watanabe, T.; Hashimoto, K., Reversible wettability control of TiO2 surface by light irradiation. *Surface Science* **2002,** *511* (1), 401-407.
116. Watanabe, T.; Fukayama, S.; Miyauchi, M.; Fujishima, A.; Hashimoto, K., Photocatalytic Activity and Photo-Induced Wettability Conversion of TiO2 Thin Film Prepared by Sol-Gel Process on a Soda-Lime Glass. *J. Sol-Gel Sci. Technol.* **2000,** *19* (1), 71-76.
117. Zeman, P.; Takabayashi, S., Self-cleaning and antifogging effects of TiO2 films prepared by radio frequency magnetron sputtering. *Journal of Vacuum Science & Technology A* **2002,** *20* (2), 388-393.
118. Vernardou, D.; Kalogerakis, G.; Stratakis, E.; Kenanakis, G.; Koudoumas, E.; Katsarakis, N., Photoinduced hydrophilic and photocatalytic response of hydrothermally grown TiO2 nanostructured thin films. *Solid State Sciences* **2009,** *11* (8), 1499-1502.
119. Suchea, M.; Christoulakis, S.; Tudose, I. V.; Vernardou, D.; Lygeraki, M. I.; Anastasiadis, S. H.; Kitsopoulos, T.; Kiriakidis, G., Pure and Nb2O5-doped TiO2 amorphous thin films grown by dc magnetron sputtering at room temperature: Surface and photo-induced hydrophilic conversion studies. *Materials Science and Engineering: B* **2007,** *144* (1), 54-59.
120. Meng, F.; Sun, Z.; Song, X., Influence of Annealing and UV Irradiation on Hydrophilicity of Ag-TiO. *Journal of Nanomaterials* **2012,** *2012*, 7.
121. Xin, B.; Hao, J., Reversibly switchable wettability. *Chem. Soc. Rev.* **2010,** *39* (2), 769-782.
122. Marmur, A., Soft contact: measurement and interpretation of contact angles. *Soft Matter* **2006,** *2* (1), 12-17.
123. Wenzel, R. N., RESISTANCE OF SOLID SURFACES TO WETTING BY WATER. *Industrial & Engineering Chemistry* **1936,** *28* (8), 988-994.

124. Lai, C.-Y.; Santos, S.; Chiesa, M., Systematic Multidimensional Quantification of Nanoscale Systems From Bimodal Atomic Force Microscopy Data. *ACS nano* **2016,** *10* (6), 6265-6272.
125. Cassie, A. B. D.; Baxter, S., Wettability of porous surfaces. *Transactions of the Faraday Society* **1944,** *40* (0), 546-551.
126. Egatz-Gomez, A.; Majithia, R.; Levert, C.; Meissner, K. E., Super-wetting, wafer-sized silicon nanowire surfaces with hierarchical roughness and low defects. *RSC Advances* **2012,** *2* (30), 11472-11480.
127. Shahsavan, H.; Quinn, J.; d'Eon, J.; Zhao, B., Surface modification of polydimethylsiloxane elastomer for stable hydrophilicity, optical transparency and film lubrication. *Colloids Surf. Physicochem. Eng. Aspects* **2015,** *482*, 267-275.
128. Schuler, B.; Meyer, G.; Peña, D.; Mullins, O. C.; Gross, L., Unraveling the Molecular Structures of Asphaltenes by Atomic Force Microscopy. *J. Am. Chem. Soc.* **2015,** *137* (31), 9870-9876.
129. Kalinin, S. V.; Sumpter, B. G.; Archibald, R. K., Big-deep-smart data in imaging for guiding materials design. *Nature materials* **2015,** *14* (10), 973-980.
130. Haviland, D. B.; van Eysden, C. A.; Forchheimer, D.; Platz, D.; Kassa, H. G.; Leclere, P., Probing viscoelastic response of soft material surfaces at the nanoscale. *Soft Matter* **2016,** *12* (2), 619-624.
131. Soussen, C.; Brie, D.; Gaboriaud, F.; Kessler, C. In *Modeling of force-volume images in atomic force microscopy*, 2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 14-17 May 2008; 2008; pp 1605-1608.
132. Nikiforov, M. P.; Reukov, V. V.; Thompson, G. L.; Vertegel, A. A.; Guo, S.; Kalinin, S. V.; Jesse, S., Functional recognition imaging using artificial neural networks: applications to rapid cellular identification via broadband electromechanical response. *Nanotechnology* **2009,** *20* (40), 405708.
133. Kumar, A.; Ovchinnikov, O.; Guo, S.; Griggio, F.; Jesse, S.; Trolier-McKinstry, S.; Kalinin, S. V., Spatially resolved mapping of disorder type and distribution in random systems using artificial neural network recognition. *Physical Review B* **2011,** *84* (2), 024203.
134. Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. A., Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater.* **2013,** *1* (1), 011002.
135. Griss, J.; Jones, A. R.; Sachsenberg, T.; Walzer, M.; Gatto, L.; Hartler, J.; Thallinger, G. G.; Salek, R. M.; Steinbeck, C.; Neuhauser, N.; Cox, J.; Neumann, S.; Fan, J.; Reisinger, F.; Xu, Q.-W.; del Toro, N.; Pérez-Riverol, Y.; Ghali, F.; Bandeira, N.; Xenarios, I.; Kohlbacher, O.; Vizcaíno, J. A.; Hermjakob, H., The mzTab Data Exchange Format: Communicating Mass-spectrometry-based Proteomics and Metabolomics Experimental Results to a Wider Audience. *Mol. Cell. Proteomics* **2014,** *13* (10), 2765-2775.
136. Boschetti, A.; Massaron, L., *Python data science essentials*. Packt Publishing Ltd: 2015.
137. Bottou, L.; Curtis, F. E.; Nocedal, J., Optimization Methods for Large-Scale Machine Learning. *SIAMR* **2018,** *60* (2), 223-311.
138. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V., Scikit-learn: Machine learning in Python. *Journal of machine learning research* **2011,** *12* (Oct), 2825-2830.

139. Borodinov, N.; Neumayer, S.; Kalinin, S. V.; Ovchinnikova, O. S.; Vasudevan, R. K.; Jesse, S., Deep neural networks for understanding noisy data applied to physical property extraction in scanning probe microscopy. *npj Computational Materials* **2019,** *5* (1), 25.
140. Santos, S.; Lai, C.-Y.; Olukan, T.; Chiesa, M., Multifrequency AFM: from origins to convergence. *Nanoscale* **2017,** *9* (16), 5038-5043.
141. Santos, S.; Lai, C.-Y.; Amadei, C. A.; Gadelrab, K. R.; Tang, T.-C.; Verdaguer, A.; Barcons, V.; Font, J.; Colchero, J.; Chiesa, M., The Mendeleev–Meyer force project. *Nanoscale* **2016,** *8* (40), 17400-17406.
142. Garcia, R.; Herruzo, E. T., The emergence of multifrequency force microscopy. *Nat Nano* **2012,** *7* (4), 217-226.
143. Benaglia, S.; Gisbert, V. G.; Perrino, A. P.; Amo, C. A.; Garcia, R., Fast and high-resolution mapping of elastic properties of biomolecules and polymers with bimodal AFM. *Nat. Protoc.* **2018,** *13* (12), 2890-2907.
144. Efremov, Y. M.; Cartagena-Rivera, A. X.; Athamneh, A. I. M.; Suter, D. M.; Raman, A., Mapping heterogeneity of cellular mechanics by multi-harmonic atomic force microscopy. *Nat. Protoc.* **2018,** *13* (10), 2200-2216.
145. Santos, S.; Verdaguer, A., Imaging Water Thin Films in Ambient Conditions Using Atomic Force Microscopy. *Materials* **2016,** *9* (3), 182.
146. Balmer, T. E.; Christenson, H. K.; Spencer, N. D.; Heuberger, M., The Effect of Surface Ions on Water Adsorption to Mica. *Langmuir : the ACS journal of surfaces and colloids* **2007,** *24* (4), 1566-1569.
147. Christenson, H. K.; Thomson, N. H., The nature of the air-cleaved mica surface. *Surface Science Reports* **2016,** *71* (2), 367-390.
148. Hutter, J. L.; Bechhoefer, J., Calibration of atomic-force microscope tips. *Review of Scientific Instruments* **1993,** *64* (7), 1868-1873.