








RESEARCH ARTICLE | DECEMBER 05 2023

# Machine learning assisted multifrequency AFM: Force model prediction

Lamiaa Elsherbiny ; Sergio Santos ; Karim Gadelrab ; Tuza Olukan ; Josep Font ; Victor Barcons ; Matteo Chiesa 

 Check for updates

*Appl. Phys. Lett.* 123, 231603 (2023)

<https://doi.org/10.1063/5.0176688>

  
View  
Online

  
Export  
Citation

CrossMark



**APL Quantum**  
Bridging fundamental quantum research with technological applications

**Now Open for Submissions**  
No Article Processing Charges (APCs) through 2024

**Submit Today**

 AIP  
Publishing

# Machine learning assisted multifrequency AFM: Force model prediction

Cite as: Appl. Phys. Lett. **123**, 231603 (2023); doi: 10.1063/5.0176688

Submitted: 15 September 2023 · Accepted: 20 November 2023 ·

Published Online: 5 December 2023



View Online



Export Citation



CrossMark

Lamiaa Elsherbiny,<sup>1</sup> Sergio Santos,<sup>2,a)</sup> Karim Gadelrab,<sup>3</sup> Tuza Olukan,<sup>1</sup> Josep Font,<sup>4</sup> Victor Barcons,<sup>4</sup> and Matteo Chiesa<sup>1,2,a)</sup>

## AFFILIATIONS

<sup>1</sup>Laboratory for Energy and NanoScience (LENS), Khalifa University of Science and Technology, Masdar Institute Campus, 127788 Abu Dhabi, United Arab Emirates

<sup>2</sup>Department of Physics and Technology, UiT-The Arctic University of Norway, 9037 Tromsø, Norway

<sup>3</sup>Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

<sup>4</sup>Departament d'Enginyeria Minera, Industrial i TIC, UPC BarcelonaTech, 08242 Manresa, Spain

<sup>a)</sup>Authors to whom correspondence should be addressed: [ssantos78h@gmail.com](mailto:ssantos78h@gmail.com) and [matteo.chiesa@ku.ac.ae](mailto:matteo.chiesa@ku.ac.ae)

## ABSTRACT

Multifrequency atomic force microscopy (AFM) enhances resolving power, provides extra contrast channels, and is equipped with a formalism to quantify material properties pixel by pixel. On the other hand, multifrequency AFM lacks the ability to extract and examine the profile to validate a given force model while scanning. We propose exploiting data-driven algorithms, i.e., machine learning packages, to predict the optimum force model from the observables of multifrequency AFM pixel by pixel. This approach allows distinguishing between different phenomena and selecting a suitable force model directly from observables. We generate predictive models using simulation data. Finally, the formalism of multifrequency AFM can be employed to analytically recover material properties by inputting the right force model.

© 2023 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0176688>

A large body of physics is dedicated to understanding the behavior of forces.<sup>1–3</sup> The atomic force microscope (AFM) is employed to understand and quantify force in the nanoscale<sup>4,5</sup> with nanometric,<sup>6</sup> atomic,<sup>7–9</sup> and sub-atomic<sup>10</sup> resolution. Forces provide information about samples and are employed to quantify and characterize properties, structure and functions of cells,<sup>5,11–13</sup> biomolecular processes,<sup>14</sup> thin and ultrathin films,<sup>15</sup> and a range of other interesting phenomena<sup>16</sup> including magnetic<sup>17</sup> and hydration forces.<sup>18</sup> Nanoscale forces might vary in terms of the shape of their profile and/or their magnitude or strength.<sup>2</sup> The particular phenomena involved in the interaction is responsible for the specific shape of the force profile, whereas the magnitude is a characteristic of the strength of the interaction.<sup>2</sup> In dynamic AFM, the force profile can be reconstructed by acquiring curves where observable parameters, such as phase, amplitude,<sup>19,20</sup> or frequency shift,<sup>21,22</sup> are obtained in terms of cantilever-sample distance. Experimentally monitored data are then transformed into force distance curves.<sup>19,21,23</sup> In order to quantify material properties, a second step is required. This involves considering a force model that reasonably fits the experimental force profile and simultaneously expresses the phenomena in terms of physically meaningful

parameters or properties.<sup>24–26</sup> Conversely, force models are also exploited to understand the dynamics of the cantilever.<sup>2,24,27</sup> A method consists of numerically integrating the equation of motion so the dynamics of the system can be interpreted in terms of the sample's properties, i.e., the Young's modulus or the coefficient of viscosity,<sup>28–30</sup> rather than the observables of the microscope, i.e., amplitude, frequency, or phase, that are relatively meaningless or complex to interpret for the broader community. Different levels of complexity for the models might be assumed,<sup>12,31</sup> and experimentally, errors might follow where *a priori* assumptions do not meet the requirements to fit the experimental data.<sup>32</sup>

Here, we focus on the extraction of material properties in multifrequency AFM,<sup>9,33,34</sup> an advanced method that allows simultaneously scanning surfaces and quantifying material properties.<sup>9,22,25,26,35</sup> With this method, the full force profile is not directly recoverable, rather a force model must be first inputted into the formalism.<sup>25,36</sup> The current upgrade<sup>37</sup> to 26 force models of a simulator aiming to assist data interpretation in monomodal and bimodal AFM illustrates the complexity and diversity of forces that might be encountered while scanning. We propose to employ machine learning methods to first predict a suitable

force profile for each pixel and then exploit the analytic expressions from the multifrequency formalism to extract material properties. This approach effectively combines a data-driven machine learning approach<sup>38</sup> with the physical modeling developed over the past decades in AFM.<sup>9,26,28,39</sup> For simplicity and to showcase the method, we focus on amplitude modulation (AM) multifrequency only and provide simulation data only. We illustrate the problem theoretically by focusing on inverse power laws in the long range of interaction where there is no mechanical contact between the tip and the sample.

The standard multifrequency formalism is equipped with two independent expressions, one for the fundamental mode, i.e., lower frequencies, and other one for the higher modes, i.e., higher frequencies.<sup>39,40</sup> Material properties can be quantified by exploiting these two expressions.<sup>25,36</sup> Material properties can be recovered with the multifrequency formalism from conservative force models.<sup>39–41</sup> In short, once a force model is selected, it is possible to recover material properties by exploiting the spectroscopy formalism that multifrequency is equipped with.<sup>9,25,36</sup> Figure 1 schematically illustrates the process. We first discuss the suitability of machine learning to predict a suitable force model and then illustrate the process with an example based on inverse power laws. Machine learning approaches are suited to extract patterns from observable data and are increasingly being exploited in this direction.<sup>42</sup> The shape of the force profile is a suitable candidate for such classification problem<sup>43</sup> in multifrequency AFM since a discrete number of force models can be taken as the “labels.” For every pixel, the observable data consist of the full set of observables. In principle, it would be possible to employ experimental data<sup>44</sup> to generate a suitably labeled dataset (see Fig. 1). On the other hand, some<sup>42</sup> have recently recommended testing the performance of machine learning algorithms by exploiting simulations and physical models when dealing with complex systems. In our case, the physical model is well known in the AFM community,<sup>8,26,45,46</sup> i.e., a set of differential equations that are standard in multifrequency AFM. We present results employing this second approach. Some have further emphasized<sup>42</sup> that data-driven machine learning will not replace physical modeling but strongly complement and enrich it. In our approach, the machine learning results are the first step toward material properties quantification since they are employed to predict the force model or shape of the force profile. Once the force profile is known, the standard analytical spectroscopy formalism in multifrequency AFM can be exploited to recover material properties (see schematic in Fig. 1).

The problem to be solved via machine learning can be stated as follows. Since a set of observables are obtained at every pixel in multifrequency AFM while scanning, i.e., amplitudes and phases, and several other parameters can be arbitrary set in the experiment, i.e., drive amplitudes and drive frequencies, the objective is to use these “known” parameters as input features in a machine learning algorithm to predict the shape of the force profile. The force profiles are the outputs or labels and form a discrete set. Finally, since cantilever parameters, i.e., the spring constants  $k_i$ , quality factors  $Q_i$ , and natural angular frequencies  $\omega_{oi}$  of each mode  $i$ , can also vary from experiment to experiment, these parameters could be added to the set, i.e., these are potential input features. Nevertheless, we simplify the problem here by selecting relatively standard cantilever values in multifrequency AFM as discussed in the example below.

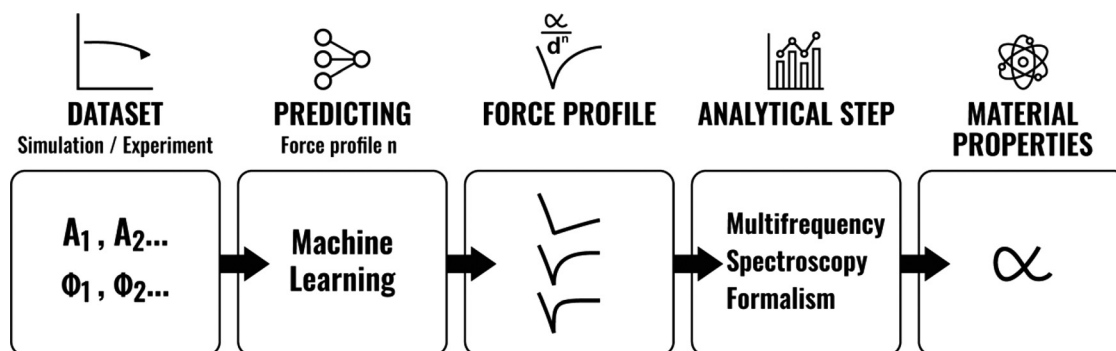
First, a set of possible force models must be considered. In order to illustrate the problem with an example, we focus on a general expression for inverse power laws, where the force  $F_{ts}$  can be written as follows:

$$F_{ts} = -\frac{\alpha}{d^n}, \quad \text{where } d > a_0, \quad (1)$$

where  $d$  is the tip-sample distance,<sup>47</sup>  $\alpha$  dictates the magnitude, or the strength, of the force,<sup>29,40</sup> and  $a_0$  is an intermolecular distance introduced to avoid the divergence and indicates that matter interpretability is forbidden.<sup>47</sup> The power  $n$  controls the shape or “profile” of the force and therefore determines the behavior of the phenomena being probed. For this reason,  $n$  is the parameter to be predicted, where  $n=2$  the force can be identified with the ubiquitous van der Waals (vdW) force provided the tip can be assumed to be a sphere of radius  $R$  and the surface can be assumed to be an infinite plane at a distance  $d$ .<sup>48</sup> The expression was first derived by Hamaker<sup>48</sup> and assumes atoms on the tip and sample interact via a pair potential inversely proportional to the sixth power of the distance. It follows that the tip-sample force can be written in terms of an inverse power law with power  $n=2$ ,

$$F_{ts} = -\frac{\alpha}{d^2} \equiv \frac{RH}{6d^2}, \quad \text{vdW force for a sphere-plane interaction,} \quad (2)$$

where  $H$  is the so-called Hamaker constant.  $H$  provides chemical<sup>36</sup> information about atoms and atomic packing on the sample’s surface.<sup>2,40</sup> The situation in (2) is compatible with experiments, where the



**FIG. 1.** Schematic of the full process that combines machine learning to predict the shape or profile of the force directly from observable data and physical modeling to extract material properties analytically once the force profile is known and a force model selected.

tip is sufficiently sharp that the tip interacts as if it were a sphere of radius  $R$  and the surface were a plane. Here, the range in tip-sample distance  $d$  lies in the order of  $\sim 1$ – $10$  nm, as in standard multifrequency AFM, a tip of radius  $R \sim 1$ – $10$  nm reasonably satisfies the condition in (2). For larger tip radii, i.e.,  $R \gg 1$ – $10$  nm, and similar distances, i.e.,  $d \sim 1$ – $10$  nm, the geometry better approximates that of two planes.<sup>2,48</sup> Hamaker worked out that for these conditions, the vdW force is inversely proportional to the third power of the distance,

$$F_{ts} = -\frac{\alpha}{d^3} \equiv \frac{H}{6\pi d^3}, \quad \text{vdW force for a plane-plane interaction.} \quad (3)$$

Equations (2) and (3) demonstrate the relevance of the geometry in the interaction's force profile and magnitude. Equations (2) and (3) also demonstrate that while the tip's radius  $R$  can be assumed to be approximately constant when scanning, by varying  $R$  from experiment to experiment, and for a given sample and force, the interaction can be made to display different inverse power laws, i.e.,  $n = 2$  and  $n = 3$  for the above-mentioned cases. The interaction profile can also vary because of the type of force emerging from a given phenomenon.<sup>2</sup> For example,<sup>41</sup> where the surface of the sample has a zero net charge, but where it can be assumed to be a matrix of polarizable electric dipoles, every pixel can be assumed to be an effective polarizable dipole with moment  $\vec{p}$  interacting with a polarized tip with effective moment  $\vec{p}'$ . Standard silicon oxide tips meet this requirement since they are polarized under the influence of an electric field. Then, the force can be modeled as a power law inversely proportional to the fourth power, i.e.,  $n = 4$ , as follows:

$$F_{ts} = \frac{B}{d^4} \vec{p} \vec{p}', \quad \text{dipolar electric interaction for} \\ \text{randomly oriented dipoles,} \quad (4)$$

where  $B$  is a coefficient. When the effective dipole on the tip and the effective dipole of the pixel are oriented in the same direction, the magnitude of the force in the vertical axis is

$$F_{ts} = \pm \frac{\alpha}{d^4} \equiv \pm \frac{3}{2\pi\epsilon_0 d^4} |\vec{p} \vec{p}'|, \\ \text{dipolar electric interaction (same orientation),} \quad (5)$$

where  $\epsilon_0$  is the permittivity, and the sign of the force depends on the relative orientation of the dipoles. The magnetic interaction has the same form as that in (5) when the tip sample distance  $d \gg R$ . For completeness, we write the expression of the magnitude of the magnetic force in the vertical axis for sufficiently large distances, i.e.,  $d \gg R$ , as

$$F_{ts} = \frac{\alpha}{d^4} \equiv \frac{3\mu_0}{2\pi d^4} |\vec{\mu} \vec{\mu}'|, \\ \text{dipolar electric interaction (dipoles in the same direction),} \quad (6)$$

where  $\vec{\mu}$  and  $\vec{\mu}'$  are the effective magnetic moments of the pixel and tip, respectively, and  $\mu_0$  is the permeability in vacuum. While not exhaustive, the above-mentioned discussion shows that models for inverse power laws [Eq. (1)] are experimentally relevant. Next, we discuss the problem of predicting the force model for inverse power laws with powers  $n = 2$  to 4 by exploiting machine learning algorithms.

The input data, i.e., features, for training the machine learning models have been produced by numerical integration of the equations of motion. We reduce the dynamics of the cantilever-tip-sample system to three modes, i.e., trimodal AFM,<sup>39,49</sup> as follows:

$$m\ddot{z}_i = -k_i z_i - \frac{m\omega_{0i}}{Q_i} \dot{z}_i + \sum_{i=1}^{i=M} F_{0i} \cos \omega_i t + F_{ts}(z), \quad (7)$$

where  $m$  is the effective mass,  $z$  is the tip's displacement from its rest position,  $M$  is the number of modes employed to model the system,  $F_{0i}$  and  $\omega_i$  are the driving force and the driving angular frequency at or near the resonance of each mode, and  $F_{ts}(z)$  is the tip-sample force acting at  $z$ . The position  $z$  can be expressed in terms of the frequency components coinciding with the drive frequencies  $\omega_i$ ,

$$z(t) = \sum_{i=1}^{i=M} z_i + O(\epsilon) \approx \sum_{i=1}^{i=M} A_i \cos(\omega_i t - \phi_i). \quad (8)$$

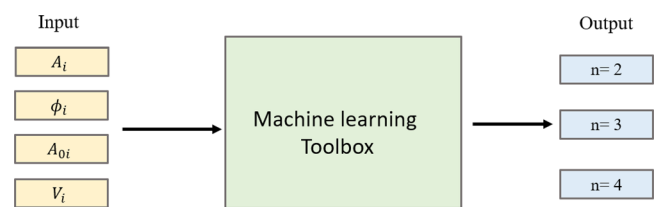
$O(\epsilon)$  carries the contributions of higher harmonics and higher modes. The driving forces  $F_{0i}$  near resonance can be written in terms of the experimental value of the free or unperturbed amplitude  $A_{0i}$ . Then,

$$F_{0i} \approx \frac{A_{0i} k_i}{Q_i}. \quad (9)$$

As input features, we employed the perturbed amplitudes  $A_i$ , phases  $\phi_i$ , free amplitudes  $A_{0i}$ , and virials of interaction  $V_i$  (Fig. 2). We added the terms  $V_i$  to the set of input features because these are known<sup>22,45,50</sup> to control conservative forces and therefore act as additional information by enforcing the physical laws<sup>43</sup> for training purposes. The virials can be expressed in terms of known parameters as follows:

$$V_i = \langle F_{ts} z_i \rangle = \frac{1}{T} \int_0^T F_{ts} z_i dt \approx -\frac{1}{2} F_{0i} A_i \cos \phi_i. \quad (10)$$

We investigated the predictive power according to the setup in Fig. 2 in monomodal, bimodal, and trimodal AFM. Exploiting machine learning to predict  $n$  from observables in monomodal, bimodal, and trimodal AFM allowed us to understand the predictive power available in each mode of operation. In monomodal, only the first mode is excited. In bimodal, we excited either modes 1 and 2 or modes 1 and 3, and in trimodal, we simultaneously excited modes 1, 2, and 3. For the simulations, the constant parameters were  $k_1 = 2$  N/m,  $k_2 = 80$  N/m,  $k_3 = 600$  N/m,  $Q_1 = 100$ ,  $Q_2 = 600$ ,  $Q_3 = 1800$ ,  $f_1 = 70$  kHz,  $f_2 = 420$  kHz,  $f_3 = 1190$  kHz,  $R = 20$  nm, and  $a_0 = 0.165$  nm. These are



**FIG. 2.** Schematic showing how the input features are fed into machine learning algorithms to generate models to predict the power  $n$ . The full MATLAB toolbox was employed to test the several machine learning models offered in the machine learning package. A schematic of the workings of support vector machine (SVM) algorithms is shown in the figure.



relatively standard cantilever parameters in multifrequency AM AFM. The coefficient  $\alpha$  in Eq. (1), together with its units, depends on the power  $n$ . Nevertheless, it is possible<sup>2</sup> to relate minima in force at  $d = a_0$  to the surface energy  $\gamma$  by considering the force of adhesion  $F_{AD}$  and relating it to  $\gamma$ ,

$$F_{AD} = -4\pi R\gamma \equiv -\frac{\alpha}{a_0^n}. \quad (11)$$

Combining Eqs. (1) and (11),

$$\alpha = 4\pi R\gamma a_0^n. \quad (12)$$

Equation (12) allows writing  $\alpha$  in terms of  $\gamma$  for any power law  $n$  in Eq. (1). This is useful because it allows establishing the magnitude of the force in terms of the common parameter  $\gamma$  for any power  $n$ . In the simulations, the magnitude of the force was varied by varying the values of  $\gamma$  and data were produced for power laws with  $n = 2, 3$ , and 4 according to Eq. (1). The variations in  $\gamma$  were discrete, and sets of data were produced for  $\gamma = 20, 40, 60$ , and 80 mJ. Note that since  $R$  and  $\gamma$  are inversely proportional [Eq. (12)], variations in  $R$  are equivalent to variations in  $\gamma$  so only one or the other has to be modified to account for variations in  $\alpha$ . The free amplitudes of the first mode  $A_{01}$  were 0.8, 1.8, and 2.8 nm, and for  $A_{02}$  and  $A_{03}$ , they were 80, 180, and 280 pm. The above are standard values in multifrequency AFM<sup>33</sup> where higher mode amplitudes are approximately 10% that of the fundamental. This choice of free amplitudes limits the use of the ML models since the ML models and our predictions will be valid provided users stay within this range of values.

Approximately 650 points of simulation data were recorded in monomodal, bimodal modes 1 and 2, bimodal modes 1 and 3, and trimodal with modes 1, 2, and 3. The data were fed into the MATLAB Statistics and Machine Learning Toolbox<sup>51</sup> (refer to the supplementary material for raw data, codes, and video explanation on the use and application of the MATLAB Classification Learner Toolbox). We exploited the classification and regression learner apps to programmatically build the set of predictive models available in the toolbox (Classification Learner App<sup>52</sup>). This includes support vector machine (SVM), deep neural network (DNN), and shallow neural network (SNN) classifiers among other (see supplementary material Sec. 2.2.2). The results of Table I were produced by employing the amplitudes  $A_i$ , phases  $\phi_i$ , and virials  $V_i$  as input features to predict the parameter  $n$  in Eq. (1) (all the raw data are available as the supplementary material). In Table I, we provide the best performing algorithm and the performance of the algorithms, or classification errors, in terms of accuracy, i.e., the number of positively identified examples divided by the total

number of examples. These results are those resulting after cross-validating with training data to minimize overfitting and submitting the models to a test (see supplementary material Sec. 2.2.4). The confusion matrices for all results in the tables can be found in supplementary material Sec. 2.2.5. The MATLAB toolbox is further equipped with feature selection methods to identify variables with the best predictive power. These results are also presented in Table I. The higher the number, the better the predictive power. The results of Table I provide several insights.

First, the two main observables in monomodal AFM, amplitude  $A_1$ , and phase  $\phi_1$ , together with the virial of interaction  $V_1$  that includes [Eqs. (8) and (9)] information about the spring constant  $k_1$ , the free amplitude  $A_{01}$ , and the quality factor  $Q_1$ , are sufficient to predict the power law  $n$  in Eq. (1) with an accuracy of  $\approx 80\%$  with this setup, i.e., classifications for  $n = 2$  to 4 with  $A_i$ ,  $\phi_i$ , and  $V_i$  as input features. The predictive power is lowest for  $A_1$  and highest for  $\phi_1$ , i.e., phase is the best predictor. This behavior is reproduced for the first mode in bimodal and trimodal AFM throughout. In bimodal AFM, both when driving with the first and second modes and with the first and third modes, the accuracy increases to 98.4% and 98.5%, respectively. The implication is that by driving with two modes, the system already provides most of the information required to predict the force profile, i.e., close to 100% with the input features of choice  $A_1$ ,  $V_1$ , and  $\phi_1$ . With the higher modes, however, the predictive power of the features changes. In bimodal AFM, the best predictor from the higher modes is still the phase, but the worst predictor is the virial. In trimodal AFM ( $m1$ ,  $m2$ , and  $m3$ ), the accuracy slightly increases in relation to bimodal AFM  $\approx 99\%$ . The predictive power of the features reverses with  $A_1$  performing worst and  $\phi_1$  performing best as in monomodal. The virial  $V_i$  is a parameter that includes information about the phase, the amplitude, and the free amplitude [Eq. (10)]. The virial is, thus, compounded, while the free amplitude is both arbitrary set by the user and not compounded. In Table II, we present the results of including the free amplitudes  $A_{0i}$  as input features.

The results of Table II show the important result that, when also employing the free amplitudes as input features, monomodal AFM is sufficient to obtain accuracies  $>98\%$ . Furthermore, the free amplitudes provide the best predictive power for mode 1 in all cases. The virials  $V_i$  become relatively irrelevant in comparison (see that their predictive power is 0 or 1 throughout). The phases are no longer the best features to predict  $n$  for the higher modes. Overall, these results show that the performance of the algorithms is very sensitive to the choice of input features, i.e., standard observables in multifrequency AFM like amplitudes  $A_i$ , phase  $\phi_i$ , or free amplitudes  $A_{0i}$ . As a final step to extract material properties, the algorithms generated and employed to

**TABLE I.** Prediction of  $n$  ( $n = 2$  to 4) in monomodal, bimodal, and trimodal AFM in terms of accuracy and best performing algorithm in each case. The predictive power of features in each case is also presented, where amplitudes  $A_i$ , phases  $\phi_i$ , and virials  $V_i$  are employed as input features.

	Accuracy (%)	Best algorithm	Feature predictive power								
			$A_1$	$\phi_1$	$V_1$	$A_2$	$\phi_2$	$V_2$	$A_3$	$\phi_3$	$V_3$
Monomodal m1	79	Trilayered neural network	70	80	76	x	x	x	x	x	x
Bimodal, m1, m2	98.4	Narrow neural network	24	57	42	32	52	30	x	x	x
Bimodal, m1, m3	98.5	Medium neural network	30	69	47	x	x	x	46	62	29
Trimodal, m1, m2, m3	99.2	Narrow neural network	46	93	70	34	77	47	25	83	50

**TABLE II.** Prediction of  $n$  ( $n = 2$  to  $4$ ) in monomodal, bimodal, and trimodal AFM in terms of accuracy and best performing algorithm in each case. The predictive power of features in each case is also presented, where amplitudes  $A_i$ , phases  $\phi_i$ , virials  $V_i$ , and free amplitudes  $A_{0i}$  are employed as input features.

	Accuracy (%)	Best algorithm	Feature predictive power											
			$A_1$	$\phi_1$	$V_1$	$A_{01}$	$A_2$	$\phi_2$	$V_2$	$A_{02}$	$A_3$	$\phi_3$	$V_3$	$A_{03}$
Monomodal m1	98.4	Trilayered neural network	52	58	0	72	x	x	x	x	x	x	x	x
Bimodal, m1, m2	98.4	Narrow neural network	25	30	0	57	49	24	0	41	x	x	x	x
Bimodal, m1, m3	100	Medium neural network	20	46	1	68	x	x	x	x	52	35	1	108
Trimodal, m1, m2, m3	100	Narrow neural network	53	26	0	91	64	73	0	37	66	46	0	45

produce the data in Tables I and II should be used while scanning to predict the force profile for every pixel. This is relatively straightforward since the generated model takes an input vector, i.e., amplitudes, phases, and free amplitudes, for each pixel and directly predicts the output as a standard function (see supplementary material Sec. 3.2 for an example on using our models experimentally). From here, material properties can be recovered while scanning by exploiting the analytic formalism in multifrequency AFM. This second step (see schematic in Fig. 1) has been developed during the last decade, as reported elsewhere,<sup>4,25,26,33,49,53</sup> and requires solving a set of integrals and driving with at least two modes.<sup>25,36,39,54</sup>

In summary, we have shown that the spectroscopy formalism of multifrequency AFM can benefit from data-driven predictions obtained by exploiting machine learning algorithms. While one dimensional force curves provide the full force profile and allow inspecting the full curve to assess the validity of a given force mode, multifrequency AFM requires assuming a force model without accessing the full force profile. Exploiting machine learning and the set of available observables in multifrequency AFM allows distinguishing between different phenomena for each pixel and selecting a suitable force model. We have illustrated this method theoretically by employing simulation data to generate the models and restricted the data to a set of cantilever parameters, i.e.,  $k$ ,  $\omega_0$ , and  $Q$ , a set of drive amplitudes, i.e.,  $A_{01} \approx 1$ – $3$  nm, and a set of models involving inverse power laws in the long range with powers  $n = 2, 3$ , and  $4$ . Clearly, this presentation is limited to the choice of models in Eq. (1) with  $n = 2, 3$ , and  $4$  and cantilever/user parameters. The approach, however, can be easily expanded to account for other force models and cantilever parameters in the future if required.

See the supplementary material for the following: MATLAB codes, raw data from simulations (training/testing datasets), models imported from MATLAB, and video explaining the use of the raw data to produce the models and Tables I and II with the MATLAB Classification Learner App.

## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

### Author Contributions

Lamiaa Elsherbiny, Sergio Santos, and Matteo Chiesa contributed equally to this work.

**Lamiaa Elsherbiny:** Conceptualization (equal); Formal analysis (equal); Investigation (lead); Methodology (equal); Software (lead);

Supervision (equal); Writing – original draft (equal); Writing – review & editing (equal). **Sergio Santos:** Conceptualization (lead); Formal analysis (equal); Methodology (equal); Validation (equal); Visualization (equal); Writing – original draft (lead); Writing – review & editing (equal). **Karim Gadelrab:** Conceptualization (equal); Formal analysis (supporting); Methodology (supporting); Supervision (equal); Writing – original draft (equal); Writing – review & editing (supporting). **Tuza Adeyemi Olukan:** Conceptualization (equal); Formal analysis (equal); Investigation (supporting); Methodology (supporting); Supervision (supporting); Writing – original draft (equal); Writing – review & editing (supporting). **Josep Font:** Conceptualization (equal); Formal analysis (equal); Methodology (equal); Supervision (equal); Writing – original draft (equal); Writing – review & editing (supporting). **Victor Barcons:** Conceptualization (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Software (lead); Writing – original draft (equal); Writing – review & editing (supporting). **Matteo Chiesa:** Conceptualization (lead); Formal analysis (equal); Investigation (equal); Methodology (equal); Supervision (lead); Writing – original draft (equal); Writing – review & editing (lead).

## DATA AVAILABILITY

The data that support the findings of this study are available within the article and its supplementary material.

## REFERENCES

- <sup>1</sup>H. Goldstein, C. Poole, and J. L. Safko, *Classical Mechanics* (Pearson, 2001).
- <sup>2</sup>J. N. Israelachvili, *Intermolecular and Surface Forces* (Elsevier Academic Press, London, 2005).
- <sup>3</sup>R. M. Brydson, C. Hammond, D. Mowbray, M. R. J. Gibbs, I. Todd, M. Grell, I. W. Hamley, M. Geoghegan, R. A. L. Jones, and G. J. Leggett, *Nanoscale Science and Technology* (Wiley, Chichester, 2005).
- <sup>4</sup>A. L. Eichhorn, M. Hoffer, and C. Dietz, *Carbon* **200**, 124 (2022).
- <sup>5</sup>K. Walter, J. Bourquin, A. Amiri, N. Scheer, M. Dehnert, A. L. Eichhorn, and C. Dietz, *Soft Matter* **19**(25), 4772–4779 (2023).
- <sup>6</sup>F. Gramazio, M. Lorenzoni, F. Pérez-Murano, L. Evangelio, and J. Fraxedas, *Ultramicroscopy* **187**, 20 (2018); A. Raman, S. Trigueros, A. Cartagena, A. P. Z. Stevenson, M. Susilo, E. Nauman, and S. Antoranz Contera, *Nat. Nanotechnol.* **6**(12), 809 (2011).
- <sup>7</sup>A. J. Weymouth, O. Gretz, E. Riegel, and F. J. Giessibl, *Jpn. J. Appl. Phys., Part 1* **61**, SL0801 (2022).
- <sup>8</sup>A. L. Eichhorn and C. Dietz, *Sci. Rep.* **12**(1), 8981 (2022).
- <sup>9</sup>S. Kawai, T. Glatzel, S. Koch, B. Such, A. Baratoff, and E. Meyer, *Phys. Rev. Lett.* **103**, 220801 (2009).
- <sup>10</sup>A. Liebig, C. Setescak, A. Weindl, and F. J. Giessibl, *J. Phys. Chem. C* **126**(51), 21716 (2022).

- <sup>11</sup>R. D. Turner, J. Kirkham, D. Devine, and N. H. Thomson, *Appl. Phys. Lett.* **94**(4), 043901 (2009).
- <sup>12</sup>Y. M. Efremov, W.-H. Wang, S. D. Hardy, R. L. Geahlen, and A. Raman, *Sci. Rep.* **7**(1), 1541 (2017).
- <sup>13</sup>A. Stylianou, M. Lekka, and T. Stylianopoulos, *Nanoscale* **10**(45), 20930 (2018).
- <sup>14</sup>T. Ando, T. Uchihashi, and S. Scheuring, *Chem. Rev.* **114**(6), 3120 (2014); A. Engel and D. J. Müller, *Nat. Struct. Biol.* **7**(9), 715 (2000); A. Lostao, K. Lim, M. C. Pallarés, A. Ptak, and C. Marcuello, *Int. J. Biol. Macromol.* **238**, 124089 (2023); N. H. Thomson, *Ultramicroscopy* **105**, 103 (2005).
- <sup>15</sup>V. G. Gisbert and R. Garcia, *ACS Nano* **15**(12), 20574 (2021); F. Güzelçimen, B. Tanören, Ç. Çetinkaya, M. D. Kaya, H. İbrahim Efker, Y. Özen, D. Bingöl, M. Sirkeci, B. Kinaci, M. Burçin Ünlü, and S. Özçelik, *Vacuum* **182**, 109766 (2020); M. Swierczewski, A. Chenneviere, L.-T. Lee, P. Maroni, and T. Bürgi, *J. Colloid Interface Sci.* **630**, 28 (2023).
- <sup>16</sup>J. E. Sader, B. D. Hughes, F. Huber, and F. J. Giessibl, *Nat. Nanotechnol.* **13**(12), 1088 (2018).
- <sup>17</sup>A. Hrabec, J. Sampaio, M. Belméguenai, I. Gross, R. Weil, S. M. Chérif, A. Stashkevich, V. Jacques, A. Thiaville, and S. Rohart, *Nat. Commun.* **8**(1), 15765 (2017); Y. Feng, P. Mirzadeh Vaghefi, S. Vranjkovic, M. Penedo, P. Kappenberger, J. Schwenk, X. Zhao, A. O. Mandru, and H. J. Hug, *J. Magn. Mater.* **551**, 169073 (2022); A. Klaassen, F. Liu, F. Mugele, and I. Siretanu, *Langmuir* **38**(3), 914 (2022).
- <sup>18</sup>M. R. Uhlig, S. Benaglia, R. Thakkar, J. Comer, and R. Garcia, *Nanoscale* **13**(10), 5275 (2021).
- <sup>19</sup>S. Hu and A. Raman, *Nanotechnology* **19**(37), 375704 (2008); A. J. Katan, M. H. van Es, and T. H. Oosterkamp, *Nanotechnology* **20**(16), 165703 (2009).
- <sup>20</sup>J. R. Lozano and R. Garcia, *Phys. Rev. B* **79**(1), 014110 (2009); H. Holscher, *Appl. Phys. Lett.* **89**(12), 123109 (2006).
- <sup>21</sup>J. E. Sader and S. P. Jarvis, *Appl. Phys. Lett.* **84**(10), 1801 (2004).
- <sup>22</sup>F. J. Giessibl, *Phys. Rev. B* **56**(24), 16010 (1997).
- <sup>23</sup>F. Payam Amir, D. Martin-Jimenez, and R. Garcia, *Nanotechnology* **26**(18), 185706 (2015).
- <sup>24</sup>D. Forchheimer, D. Platz, E. A. Tholén, and D. B. Haviland, *Phys. Rev. B* **85**(19), 195449 (2012).
- <sup>25</sup>E. T. Herruzo, A. P. Perrino, and R. Garcia, *Nat. Commun.* **5**(1), 3126 (2014).
- <sup>26</sup>B. Rajabifar, A. Bajaj, R. Reifengerger, R. Proksch, and A. Raman, *Nanoscale* **13**(41), 17428 (2021).
- <sup>27</sup>R. Garcia and A. San Paulo, *Phys. Rev. B* **61**, R13381 (2000); R. Garcia and R. Pérez, *Surf. Sci. Rep.* **47**(6), 197 (2002).
- <sup>28</sup>M. Damircheli, U. Jung, and R. Wagner, *Phys. Scr.* **98**(3), 035708 (2023); S. Benaglia, C. A. Amo, and R. Garcia, *Nanoscale* **11**(32), 15289 (2019).
- <sup>29</sup>R. Garcia, C. J. Gómez, N. F. Martínez, S. Patil, C. Dietz, and R. Magerle, *Phys. Rev. Lett.* **97**, 016103 (2006).
- <sup>30</sup>J. Tamayo and R. Garcia, *Langmuir* **12**(18), 4430 (1996).
- <sup>31</sup>S. Abuhattum, D. Mokbel, P. Müller, D. Soteriou, J. Guck, and S. Aland, *iScience* **25**(4), 104016 (2022).
- <sup>32</sup>M. R. McCraw, B. Uluutku, H. D. Solomon, M. S. Anderson, K. Sarkar, and S. D. Solares, *Soft Matter* **19**(3), 451 (2023).
- <sup>33</sup>R. Garcia and E. T. Herruzo, *Nat. Nanotechnol.* **7**(4), 217 (2012).
- <sup>34</sup>R. Garcia and R. Proksch, *Eur. Polym. J.* **49**(8), 1897 (2013); A. Labuda, M. Kocun, W. Meinhold, D. Walters, and R. Proksch, *Beilstein J. Nanotechnol.* **7**, 970 (2016); M. Kocun, A. Labuda, W. Meinhold, I. Revenko, and R. Proksch, *ACS Nano* **11**(10), 10097 (2017); R. Proksch, *Appl. Phys. Lett.* **89**, 113121 (2006).
- <sup>35</sup>J. E. Sader, T. Uchihashi, M. J. Higgins, A. Farrell, Y. Nakayama, and S. P. Jarvis, *Nanotechnology* **16**(3), S94 (2005).
- <sup>36</sup>C.-Y. Lai, S. Perri, S. Santos, R. Garcia, and M. Chiesa, *Nanoscale* **8**(18), 9688 (2016).
- <sup>37</sup>V. G. Gisbert and R. Garcia, *Soft Matter* **19**(31), 5857–5868 (2023).
- <sup>38</sup>A. Ghadami and B. I. Epureanu, *Philos. Trans. R. Soc. A* **380**(2229), 20210213 (2022).
- <sup>39</sup>S. Santos, K. Gadelrab, L. Elsherbiny, X. Drexler, T. Olukan, J. Font, V. Barcons, and M. Chiesa, *J. Chem. Phys.* **158**(20), 204703 (2023).
- <sup>40</sup>S. Santos, K. Gadelrab, T. Olukan, J. Font, V. Barcons, and M. Chiesa, *Appl. Phys. Lett.* **122**(7), 071603 (2023).
- <sup>41</sup>C. A. Amo, “Microscopía de Fuerzas Bimodal y no Resonante para Medir Propiedades Físicas y Químicas a Escala Nanométrica,” Ph.D. thesis (Instituto de Ciencia de Materiales de Madrid, 2019); available at <https://repositorio.uam.es/handle/10486/686790>.
- <sup>42</sup>M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, and Prabhat, *Nature* **566**(7743), 195 (2019).
- <sup>43</sup>G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, *Nat. Rev. Phys.* **3**(6), 422 (2021).
- <sup>44</sup>S. Santos, C.-Y. Lai, C. A. Amadei, K. R. Gadelrab, T.-C. Tang, A. Verdaguer, V. Barcons, J. Font, J. Colchero, and M. Chiesa, *Nanoscale* **8**(40), 17400 (2016); C.-Y. Lai, S. Santos, and M. Chiesa, *Appl. Phys. Lett.* **114**, 241601 (2019).
- <sup>45</sup>J. R. Lozano and R. Garcia, *Phys. Rev. Lett.* **100**, 076102 (2008).
- <sup>46</sup>S. D. Solares and G. Chawla, *J. Appl. Phys.* **108**, 054901 (2010).
- <sup>47</sup>R. Garcia and A. San Paulo, *Phys. Rev. B* **60**(7), 4961 (1999).
- <sup>48</sup>H. C. Hamaker, *Physica* **4**(10), 1058 (1937).
- <sup>49</sup>S. D. Solares and G. Chawla, *Meas. Sci. Technol.* **21**(12), 125502 (2010).
- <sup>50</sup>Á. S. Paulo and R. Garcia, *Phys. Rev. B* **64**(19), 193411 (2001).
- <sup>51</sup>The MathWorks, Inc., *MATLAB and Simulink, Release R2022b* (The MathWorks, Inc., Natick, 2022).
- <sup>52</sup>MathWorks, see <https://ch.mathworks.com/help/stats/classificationlearner-app.html> for a description provided by Matlab of the Classification Learner app to classify data.
- <sup>53</sup>R. Garcia, *Chem. Soc. Rev.* **49**(16), 5850 (2020).
- <sup>54</sup>D. Martínez-Martin, E. T. Herruzo, C. Dietz, J. Gomez-Herrero, and R. Garcia, *Phys. Rev. Lett.* **106**(19), 198101 (2011).