INTERNATIONAL JOURNAL OF ROBUST AND NONLINEAR CONTROL
Int. J. Robust Nonlinear Control 2005; **15**:805–820
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/rnc.1025

Harnessing the transient signals in atomic force microscopy

Deepak R. Sahoo, Abu Sebastian and Murti V. Salapaka*,†

NanoDynamics Systems Lab, Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011, U.S.A.

SUMMARY

In the existing dynamic-mode operation of atomic force microscopes (AFMs) steady-state signals like amplitude and phase are used for detection and imaging of material. Due to the high quality factor of the cantilever probe the corresponding methods are inherently slow. In this paper, a novel methodology for fast interrogation of material that exploits the transient part of the cantilever motion is developed. This method effectively addresses the perceived fundamental limitation on bandwidth due to high quality factors. It is particularly suited for the detection of small time scale tip–sample interactions. Analysis and experiments show that the method results in significant increase in bandwidth and resolution as compared to the steady-state-based methods. This article demonstrates the effectiveness of a systems perspective to the field of imaging at the nano-scale and for the first time reports *realtime* experimental results and scanning applications of the transient method. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS: observers; Kalman filter; likelihood ratio test

1. INTRODUCTION

Desirable properties of manufactured products arise from the manner in which atoms are arranged in its material. Until recently, ways of manipulating and interrogating matter were limited to aggregate methods where the control and investigation of matter was achieved at scales much larger than atomic scales. The investigation (placement) of material to decipher (build) complex structures atom by atom was not possible, and thus optimal desired specificity of the material properties was not achievable.

Recent demonstrations of nanoscience provide ample evidence indicating the feasibility of rational *control*, *manipulation and interrogation* of matter at the atomic scale. There is a promise that material can be tailored at the atomic scale. The atomic force microscope (AFM) [1] is an instance of the impact of nanotechnology where a cantilever is utilized to image and manipulate

*Correspondence to: Murti V. Salapaka, Department of Electrical and Computer Engineering, Iowa State University, Ames, IA 50011, U.S.A.
†E-mail: murti@iastate.edu

Contract/grant sponsor: National Science Foundation Grant; contract/grant number: ECS-0330224

 Accepted 22 July 2005

sample properties (see Figure 1(I)). As stated in the National Nanotechnology Initiative Plan [2], '*These instruments, including scanning tunneling microscopes, atomic force microscopes, and near-field microscopes, provide the eyes and fingers required for nanostructure measurement and manipulation.*'

In a typical AFM (see Figure 1(I)), the deflection of the cantilever due to the sample is registered by the laser incident on the cantilever tip, which reflects into a split photo-diode. The piezoelectric scanner is used to position the sample. In static mode the cantilever deflection is solely due to the tip–sample interaction. The piezoelectric scanner is rastered in the lateral direction and the deflection of the tip is used to interpret sample properties. In the dynamic mode, the cantilever support is forced sinusoidally using a dither piezo. The changes in the oscillations introduced due to the sample are interpreted to infer its properties.

Cantilevers have been utilized in biological sciences to perform remarkable feats such as cutting DNA strands [3] and investigating the activity of RNA polymerase (a protein complex) [4]. On a similar note, there are impressive proposals on using cantilever-based nanoprobes to interrogate cell dynamics with significant impact on human health. Another intriguing application of the cantilever is in the detection of single electron spin based on magnetic detection principles [5–8]. Such research has significant ramifications for quantum computing technology. Yet another putative application of cantilever-based sensing is its use in detecting toxic chemical and biological agents with very high sensitivity, with obvious significance to security.

In spite of the underlying promise of cantilever-based technology, considerable challenges need to be overcome. Pivotal to harnessing its vast potential is ultra-fast *interrogation*. This is apparent as the interrogation, for example, of atoms or spins of electrons needs to be accomplished for material that has macroscopic dimensions. Most of the current techniques rely

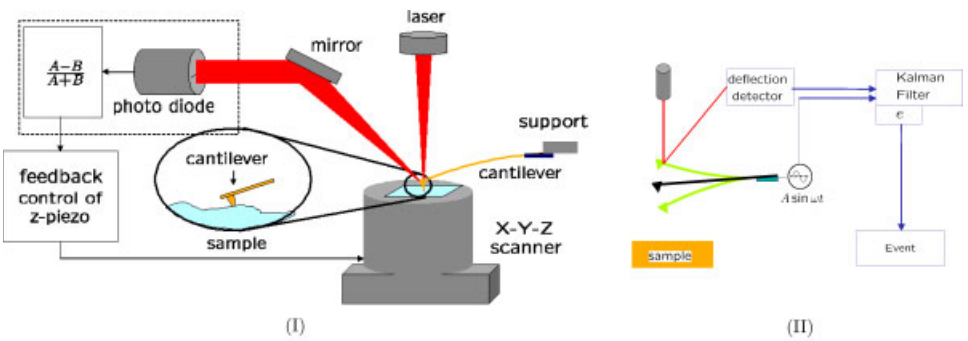

Figure 1. (I): In a typical AFM, the deflection of the cantilever due to the sample is registered by the laser incident on the cantilever tip, which reflects into a split photo-diode. The piezoelectric scanner is used to position the sample. In static mode the cantilever deflection is solely due to the tip–sample interaction. The piezoelectric scanner is rastered in the lateral direction and the deflection of the tip is used to interpret sample properties. In the dynamic mode, the cantilever support is forced sinusoidally using a dither piezo. The changes in the oscillations introduced due to the sample are interpreted to infer its properties. (II): In the new architecture an observer (Kalman filter) is designed based on a model of the cantilever that estimates the deflection of the cantilever from the photo-diode signal and the dither excitation signal. The error in estimation e arising due to the sample force is utilized to detect the tip–sample interactions as events.

on the steady-state data of the cantilever tip position. This results in very slow interrogation speeds due to the typical high quality factor of the cantilever that lead to high settling times. It is to be noted that high quality factors are a must for high resolution. Thus, in existing steady-state methods a trade-off has to be achieved between bandwidth and resolution.

In this paper, the hitherto unexplored idea of using the transient signals for increasing bandwidth and resolution is presented. A preliminary version of this paper has appeared as a letter in Reference [9]. Sensing based on transient signal has the potential to detect tip–sample interaction changes at extremely high rates. This method has the promise of detecting one event in 4 periods of cantilever oscillation, which translates to 25 million events per second for a 100 MHz cantilever. Such a technology has tremendous implications in numerous applications, for example, in cantilever-based retrieval of high-density data and detection of chemical and/or biological molecules on a surface. Researchers at IBM [10] have demonstrated areal densities of upto 3 $\mathrm{TB/in^2}$ (for data storage), however, the reading is performed in static mode (contact mode) which results in signal deterioration due to wear and, also needs to be corrected for thermal drift during extended periods of operation. Furthermore, the wear becomes more severe as the data rates are increased. Transient signal detection being a dynamic method is gentle and has virtually no wear. Moreover, as will be shown later the data rates depend only on the cantilever frequency and to a large extent independent of the quality factor Q. Thus, the high resolution needs can be effectively decoupled from the high bandwidth needs.

A systems viewpoint of the AFM dynamics and observer-based approach provide the basic analytical tools to investigate the transient signals. In Section 2, the state–space description of the cantilever dynamics is introduced. Section 3 describes the observer architecture that tracks the transient signals. Further assumptions on the character of tip–sample interactions can be used to analyse the resolution and bandwidth of the proposed scheme. Moreover, an efficient detection scheme is developed using tools from statistical signal processing. These results are presented in Section 4. Experimental results presented in Section 5 confirm the effectiveness of the new methodology.

In this article, apart from providing a detailed exposition of the transient method, we present the significant step of realizing the transient scheme in the realtime unlike the offline analysis of data in Reference [9]. Furthermore, for the first time scanning results utilizing transients are provided.

2. CANTILEVER MODEL

For many applications the cantilever is well modelled as a flexible structure. A multi-mode model accurately captures the cantilever dynamics (see Reference [11]). Typically, a first- or second-mode approximation is enough to describe the dynamics. The state–space representation of the cantilever dynamics is given by

$$\begin{aligned} \dot{\bar{x}} &= A\bar{x} + B(\eta + w) \\ y &= C\bar{x} + v \end{aligned} \tag{1}$$

where $\bar{x}$ is the state vector, A and B are matrices which are functions of the cantilever parameters, η is the thermal noise component, w describes all other external forces acting on the

cantilever, y is the photo-diode output that measures the deflection of the free end of the cantilever and v is the measurement (photo-diode) noise. A first-mode approximation of the cantilever dynamics given by

$$\ddot{p} + 2\xi\omega_0\dot{p} + \omega_0^2 p = \eta + w \tag{2}$$

that can be recast as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & -2\xi\omega_0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} (\eta + w)$$
$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + v \tag{3}$$

where state x_1 denotes the cantilever-tip position (p), state x_2 denotes the cantilever-tip velocity ($\dot{p}$), ω_0 and ξ denote the first resonant frequency of the cantilever and the damping factor in the operating medium, respectively. Note that $\xi = 1/2Q$ with Q being the quality factor of the cantilever. We will denote the equivalent dither forcing by g and the tip–sample interaction force by $\phi(p)$ that depends on the tip position p; ϕ typically has characteristics of long range attractive force and short range strong repulsive force [12]. In the above framework, we have $w = g + \phi(p)$ (see Figure 2(II)). The cantilever model described above can be identified precisely using the thermal noise response (see Reference [11]).

3. THE OBSERVER ARCHITECTURE

In the systems viewpoint the cantilever dynamics is separated as an independent system from the sample subsystem that affects the cantilever in a feedback manner. Such a perspective was first introduced into the AFM literature in References [13, 14] where it was utilized to identify the

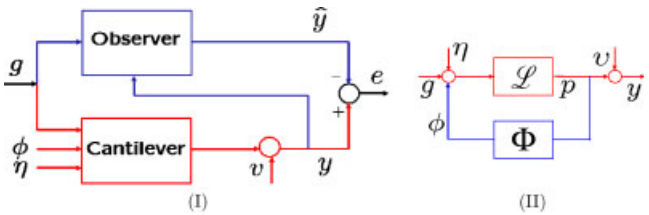

Figure 2. (I): The observer estimates the states to be $\hat{x}$. The actual state is $\bar{x}$. The error process is denoted by e which is the actual output minus the estimated output. The error process is a zero mean stationary stochastic process in the absence of the tip–sample interaction force ϕ. When the cantilever is subjected to the sample force the cantilever dynamics is altered whereas the observer dynamics remains the same. This is registered as a dynamic profile in the error process e. (II): In the systems perspective, the AFM dynamics is viewed as an interconnection of a linear cantilever system with the nonlinear tip–sample interaction forces in feedback. $\mathscr{L}$ models the cantilever dynamics that is well represented by a linear time-invariant system, Φ models the tip–sample interaction force that depends on the tip deflection p. g is the dither forcing, η is the thermal noise, v is the photo-diode (measurement) noise and y is the measured photo-diode signal.

sample interaction potentials. This systems perspective of the AFM (see Figure 2(II)) facilitates the design of an observer (see Figure 2(I)) that provides an estimate of the state x.

The observer dynamics is given by

$$\overbrace{\begin{aligned}\dot{\hat{x}} &= A\hat{x} + Bw + L(y - \hat{y});\ \ \hat{x}(0) = \hat{x}_0 \\ \hat{y} &= C\hat{x}\end{aligned}}^{\text{Observer}}$$

where $\hat{x}$ is the estimate of the state $\bar{x}$. The error in the estimate is given by $\tilde{\bar{x}} = \bar{x} - \hat{\bar{x}}$.

$$\overbrace{\begin{aligned}\dot{\tilde{\bar{x}}} &= A\bar{x} + B(w+\eta) - A\hat{\bar{x}} - Bw - L(y - \hat{y}) \\ &= (A - LC)\tilde{\bar{x}} + B\eta - Lv \\ \tilde{\bar{x}}(0) &= \bar{x}(0) - \hat{\bar{x}}(0)\end{aligned}}^{\text{State error dynamics}}$$

The error in the estimate of the output y is given by

$$e = y - \hat{y} = C\tilde{\bar{x}} + v \tag{4}$$

The error process is a good measure of the transients due to changes in tip–sample interactions. Note that only the cantilever-tip position is available as a measured quantity, not its velocity. The error between the observed state and the actual state of the cantilever, when no noise terms are present ($\eta = v = 0$) is only due to the mismatch in the initial conditions of the observer and the cantilever tip. It is evident that if the observer gain L is chosen so that the eigenvalues of the matrix $(A - LC)$ are in the strict left-half complex plane, the state error $\tilde{x}$ due to the initial condition mismatch $\tilde{x}(0)$ goes to zero with time. The system is observable and therefore the eigenvalues of $(A - LC)$ can be placed anywhere [15]. It can be shown that in the presence of the noise sources η and v, the error process e approaches a zero mean wide sense stationary stochastic process after the observer has tracked the state of the cantilever.

We now provide the intuition behind the new methodology. When the observer has tracked the cantilever state and the cantilever is subjected to a force due to the sample, the cantilever dynamics get altered. The observer sees the effects of the sample only through the measured cantilever position y which is the output of the photo-diode (see Figure 2(I)). This introduces a lag (that is dependent on the gain L) between the time when the cantilever sees the sample's influence and when the observer realizes the sample's influence. This is registered as a change in the signal e. Thus, the error signal e is a good measure of the transient signal. The magnitude of this error signal can be quite large even though the oscillation might not have changed significantly.

When the change in the tip–sample potential persists, the observer by utilizing the input y may eventually track the altered behaviour. Thus, the error signal e shows the signature at the initial part of the change but this deviation from stationarity might not persist even though the change in the tip–sample behaviour persists. This is in contrast to the steady-state methods where the information is available not in the initial part but after the cantilever has come to a steady state

(for example, in tapping-mode scheme this would mean an eventual lowered amplitude value when a step is encountered that persists). Thus, the transient signal-based scheme is a good *edge detector*.

There is considerable freedom on how fast the observer tracks the cantilever dynamics. Note that high quality factors are detrimental to high bandwidth in steady-state methods; however, required for high resolution. *By utilizing the observer-based architecture presented in this article a method for effectively isolating the high bandwidth needs from the high resolution needs is obtained.*

4. TIP–SAMPLE IMPACT MODEL

The error profile due to a tip–sample interaction change can be better characterized if a model of the effect of the tip–sample interaction change on the cantilever motion is available. We assume that the sample's influence on the cantilever tip is approximated by an impact condition where the position and velocity of the cantilever tip instantaneously assume a new value (equivalent to resetting to a different initial condition). This is satisfied in most typical operations because in the dynamic mode, the time spent by the tip under the sample's influence is negligible compared to the time it spends outside the sample's influence [16]. The assumption is also corroborated by experimental results provided later.

4.1. Bandwidth and noise analysis

For this part of the article the first-mode model for the cantilever is assumed (see Equation (3)). The dynamics of the signal e which is the difference between the photo-diode signal y and its estimated value ($\hat{y}$) is given by

$$e(s) = \frac{\eta(s) + \left(s^2 + \dfrac{\omega_0}{Q}s + \omega_0^2\right)v(s) + \left(s + \dfrac{\omega_0}{Q}\right)v_1 + v_2}{s^2 + \left(\dfrac{\omega_0}{Q} + l_1\right)s + \left(\omega_0^2 + l_2 + \dfrac{\omega_0}{Q}l_1\right)} \tag{5}$$

where $[v_1, v_2]^{\mathrm{T}}$ is the initial condition reset due to change in tip–sample interaction and $L = [l_1 \;\; l_2]^{\mathrm{T}}$ is the gain of the observer that must satisfy the stability criterion: $((\omega_0/Q) + l_1) > 0$ and $(\omega_0^2 + (\omega_0/Q)l_1 + l_2) > 0$. The signal e contains the dynamic profile due to tip–sample interaction change $[v_1, v_2]^{\mathrm{T}}$ which is corrupted by both thermal noise η and photo-diode noise v.

From Equation (5) it can be seen that the transfer function from the unknown state jump $[v_1, v_2]^{\mathrm{T}}$ to the signal e is a second-order transfer function. Hence the tracking bandwidth is characterized by,

$$B \propto \frac{\omega_0}{Q} + l_1 \tag{6}$$

Since the choice of the gain term l_1 is independent of the quality factor Q, the tracking bandwidth of the observer is effectively decoupled from Q.

The mean square contribution to e of the instantaneous state jump is given by

$$\langle p_v^2 \rangle = \frac{1}{\pi}\int_0^\infty \frac{\left(\frac{\omega_0}{Q}v_1 + v_2\right)^2 + \omega^2 v_1^2}{\left(\omega^2 - \omega_0^2 - l_2 - \frac{\omega_0 l_1}{Q}\right)^2 + \omega^2\left(\frac{\omega_0}{Q} + l_1\right)^2}\,\mathrm{d}\omega$$
$$= \frac{\left(\omega_0^2 + \frac{\omega_0 l_1}{Q} + l_2\right)v_1^2 + \left(\frac{\omega_0}{Q}v_1 + v_2\right)^2}{2\left(\omega_0^2 + l_2 + \frac{\omega_0 l_1}{Q}\right)\left(\frac{\omega_0}{Q} + l_1\right)} \tag{7}$$

$\langle p_v^2 \rangle$ characterizes the signal power.

The thermal noise acting on the cantilever is white and has mean equal to zero. Assuming the thermal noise power is equal to P, the mean square contribution to e of the thermal noise is given by,

$$\langle p_\eta^2 \rangle = \frac{P}{\pi}\int_0^\infty \frac{1}{\left(\omega^2 - \omega_0^2 - l_2 - \frac{\omega_0 l_1}{Q}\right)^2 + \omega^2\left(\frac{\omega_0}{Q} + l_1\right)^2}\,\mathrm{d}\omega$$
$$= \frac{P}{2\left(\omega_0^2 + l_2 + \frac{\omega_0 l_1}{Q}\right)\left(\frac{\omega_0}{Q} + l_1\right)} \tag{8}$$

Thus, the signal-to-noise ratio due to thermal noise per unity noise power is given by

$$\mathrm{SNR}_\eta = \sqrt{\frac{\langle p_v^2 \rangle}{\langle p_\eta^2 \rangle}} = \sqrt{\left(\omega_0^2 + \frac{\omega_0 l_1}{Q} + l_2\right)v_1^2 + \left(\frac{\omega_0}{Q}v_1 + v_2\right)^2} \tag{9}$$

It can be seen from Equation (6) and (9) that with increasing values of l_1 and l_2, SNR_η and bandwidth B increase.

The measurement noise in the photo-diode is assumed to be white and has mean equal to zero. Assuming the photo-diode noise power is equal to R, the power spectral density (psd) of the photo-diode noise v in signal e is given by

$$P_{vv}(\omega) = R\left[1 - \frac{\left(2\omega_0^2 + l_2 + \frac{\omega_0 l_1}{Q}\right)\left(\frac{\omega_0 l_1}{Q} + l_2\right) + \omega^2(l_1^2 - l_2)}{\left(\omega^2 - \omega_0^2 - l_2 - \frac{\omega_0 l_1}{Q}\right)^2 + \omega^2\left(\frac{\omega_0}{Q} + l_1\right)^2}\right]$$
$$= R[1 - P_{vv}^{\mathrm{lp}}(\omega)]$$

It can be assumed that the signal power due to state jump $[v_1, v_2]^{\mathrm{T}}$ in e is contained within a bandwidth $B = K((\omega_0/Q) + l_1)$ for $K >>> 0$. Note that the thermal noise contribution in e is also contained within bandwidth B. The signal e can be filtered out for $\omega > B$ and the signal-to-noise ratio in e (SNR_v) due to measurement noise can be analysed. Note that the noise power of the low-pass component $P_{vv}^{\mathrm{lp}}(\omega)$ of psd of photo-diode noise $P_{vv}(\omega)$ is contained within bandwidth B. The mean square contribution of the photo-diode noise to e (where e is filtered and effectively

contains the signal of interest) is approximately given by

$$\begin{aligned}\langle p_v^2\rangle &\approx \frac{1}{\pi}\int_0^B P_{vv}(\omega)\,\mathrm{d}\omega \\ &\approx R\left[K\left(\frac{\omega_0}{Q}+l_1\right)-\frac{1}{\pi}\int_0^\infty \frac{\left(2\omega_0^2+l_2+\dfrac{\omega_0 l_1}{Q}\right)\left(\dfrac{\omega_0 l_1}{Q}+l_2\right)+\omega^2(l_1^2-l_2)}{\left(\omega^2-\omega_0^2-l_2-\dfrac{\omega_0 l_1}{Q}\right)^2+\omega^2\left(\dfrac{\omega_0}{Q}+l_1\right)^2}\mathrm{d}\omega\right] \\ &= R\left[K\left(\frac{\omega_0}{Q}+l_1\right)-\frac{\left(\omega_0^2+l_2+\dfrac{\omega_0 l_1}{Q}\right)\left(\omega_0^2+\dfrac{\omega_0 l_1}{Q}+l_1^2\right)-\omega_0^4}{2\left(\omega_0^2+l_2+\dfrac{\omega_0 l_1}{Q}\right)\left(\dfrac{\omega_0}{Q}+l_1\right)}\right]\end{aligned} \tag{10}$$

The signal-to-noise ratio due to measurement noise per unity noise power is given by

$$\begin{aligned}\mathrm{SNR}_v &= \sqrt{\frac{\langle p_v^2\rangle}{\langle p_v^2\rangle}} \\ &\approx \sqrt{\frac{\left(\omega_0^2+l_2+\dfrac{\omega_0 l_1}{Q}\right)v_1^2+\left(\dfrac{\omega_0}{Q}v_1+v_2\right)^2}{\left(\omega_0^2+l_2+\dfrac{\omega_0 l_1}{Q}\right)\left((4K-1)\dfrac{\omega_0 l_1}{Q}+(2K-1)l_1^2+2K\left(\dfrac{\omega_0}{Q}\right)^2-\omega_0^2\right)+\omega_0^4}}\end{aligned} \tag{11}$$

It can be seen that SNR_v decreases with increasing values of l_1 and l_2. Therefore, the bandwidth constraint in the scheme to detection the signal due to state jump $[v_1, v_2]^{\mathrm{T}}$ is mainly imposed by the measurement noise. It is evident that a desired trade-off between signal-to-noise ratio and bandwidth can be obtained by an appropriate choice of l_1 and l_2 that is independent of Q. This provides considerable flexibility when compared to existing steady-state methods where Q determines the bandwidth. Note that due to the small measurement noise, the observer gain l_1 can be chosen large enough so that the cantilever state is tracked within a couple of cycles of the dither forcing. Therefore, the optimal bandwidth is primarily dictated by the resonant frequency ω_0 of the cantilever.

4.2. Detection

With the added assumption of the impact condition, the sample detection problem is formulated by considering a discretized model of the cantilever dynamics given by (1).

$$\begin{aligned}x_{k+1} &= Fx_k + G(g_k+\eta_k)+\delta_{\theta,k+1}v \\ y_k &= Hx_k + v_k, \quad k\geqslant 0\end{aligned} \tag{12}$$

where matrices F, G, and H are obtained from matrices A, B and C by discretizing the continuous time model of the cantilever described by Equation (1) and $\delta_{i,j}$ denotes the Dirac delta function defined as $\delta_{i,j}=1$ if $i=j$ and $\delta_{i,j}=0$ if $i\neq j$. θ denotes the time instant when the impact occurs and v signifies the magnitude of the impact. Essentially, the impact is modelled as an instantaneous change in the state by v at time instant θ. In this setting the time of impact and

the resulting change in the state are unknown quantities. The profile of the change in the mean of the error signal due to the sample can be pre-calculated and one can then employ detection and estimation methods to search for the presence of such a profile in the error sequence to not only detect the samples presence but also estimate the sample parameters.

We assume that the input noise and the output noise are white and uncorrelated. Note that the input and output noise power P and R can be measured experimentally. Given the following noise characteristics

$$E\left\{\begin{bmatrix}\eta_i\\ v_i\\ x_0\end{bmatrix}\begin{bmatrix}\eta_j\\ v_j\\ x_0\\ 1\end{bmatrix}^{\mathrm{T}}\right\}=\begin{bmatrix}P\delta_{ij} & 0 & 0 & 0\\ 0 & R\delta_{ij} & 0 & 0\\ 0 & 0 & \Pi_0 & 0\end{bmatrix}$$

the optimal transient observer is a Kalman filter [17]. Let the steady-state Kalman observer gain be given by $L = L_{\mathrm{K}}$. The error process is known as the innovation sequence when the optimal transient observer is employed. Moreover, when the tip–sample interaction is absent, the innovation process asymptotically approaches a zero mean white process.

With a Kalman observer and an impact model when the sample is encountered (given by the model in (12)), the innovation sequence e_k can be written as [18]

$$e_k = y_k - \hat{y}_k = \Upsilon_{k;\theta} v + \gamma_k \tag{13}$$

where $\{\Upsilon_{k;\theta} v\}$ is a known dynamic state profile with unknown arrival time θ defined by

$$\Upsilon_{k;\theta} = H(F - L_{\mathrm{K}} H)^{k-\theta} \tag{14}$$

and $\{\gamma_k\}$ is a zero mean white noise sequence which is the measurement residual had the jump not occurred. The statistics of γ is given by,

$$E\{\gamma_j \gamma_k^{\mathrm{T}}\} = V\delta_{ij}$$

where $V = HP_{\tilde{x}}H^{\mathrm{T}} + R$ and $P_{\tilde{x}}$ is the steady-state error covariance obtained from the Kalman filter. V is a function of P and R.

Thus, determining when the cantilever is 'hitting' the sample and when it is not, is equivalent to deciding whether the dynamic profile is present or not. A number of detection and estimation methods exist in statistical signal processing literature which address problems of similar nature. This problem can be formulated in the framework of binary hypothesis testing and a suboptimal version of the generalized likelihood ratio test [18] can be used to make the decision whether the profile is present or not. The innovation process is windowed into M samples for the analysis. The selection of the window size M is primarily determined by the length of the dynamic profile. This problem formulated in the framework of binary hypothesis testing is given by

$$\begin{aligned} &H_0 : e_k = \gamma_k, \quad k = 1, 2, \ldots, M\\ &\text{versus}\\ &H_1 : e_k = \Upsilon_{k;\theta} v + \gamma_k, \quad k = 1, 2, \ldots, M \end{aligned} \tag{15}$$

where γ_k is a zero mean white gaussian process, $e_k = \gamma_k$ is the observed innovation and $\{\Upsilon_{k;\theta} v\}$ for $k = 1, 2, \ldots, M$ is a known *dynamic profile* with unknown arrival time θ and unknown

magnitude of the state jump v. Note that if hypothesis H_0 is true, then the cantilever has not interacted with the sample within that window of M data points. Selection of hypothesis H_1 implies that the decision made is that the oscillating cantilever interacted with the sample producing the dynamic profile. To simplify the analysis, it is assumed that the impact or state jump occurs at the first sample of each time window of M samples (i.e. $\theta = 1$). This implies,

$$\bar{e} = \Upsilon v + \bar{\gamma} \tag{16}$$

where, $\bar{e} = [e_1, e_2, \ldots, e_M]^{\mathrm{T}}$, $\Upsilon = [H, H(F - L_{\mathrm{K}}H), H(F - L_{\mathrm{K}}H)^2, \ldots, H(F - L_{\mathrm{K}}H)^{M-1}]^{\mathrm{T}}$ and $\bar{\gamma} = [\gamma_1, \gamma_2, \ldots, \gamma_M]^{\mathrm{T}}$. Note that when the observer gain L is not equal to the Kalman observer gain L_{K}, the dynamic profile assumes a similar expression as in Equation (16) with observer gain L replacing L_{K}. However, in that case the noise $\bar{\gamma}$ will not necessarily be a white noise. When $\bar{\gamma}$ is not white and is a known coloured noise, a whitening filter may be introduced and the corresponding dynamic profile $\Upsilon^w v$ output from the whitening filter may be detected in corresponding white noise. When $\bar{\gamma}$ is white, the maximum likelihood estimate of the state jump, $\hat{v}$ is given by

$$\hat{v} = \left(\frac{\Upsilon^{\mathrm{T}}\Upsilon}{V}\right)^{-1}\frac{\Upsilon^{\mathrm{T}}\bar{e}}{V} \tag{17}$$

The dynamic profile if present in the selected M samples is given by

$$\bar{s} = \Upsilon\hat{v} = \Upsilon\left(\frac{\Upsilon^{\mathrm{T}}\Upsilon}{V}\right)^{-1}\frac{\Upsilon^{\mathrm{T}}\bar{e}}{V}$$

To select between the two hypotheses, the likelihood ratio is computed which is given by

$$l(M) = \frac{\bar{e}^{\mathrm{T}}\bar{s}}{V} = \frac{\bar{e}^{\mathrm{T}}\Upsilon}{V}\left(\frac{\Upsilon^{\mathrm{T}}\Upsilon}{V}\right)^{-1}\frac{\Upsilon^{\mathrm{T}}\bar{e}}{V} \tag{18}$$

The likelihood ratio is compared with a threshold value as $l(M) \gtrless_{H_0}^{H_1} \varepsilon$ to arrive at a decision whether the dynamic profile is present or not (equivalent to deciding whether a tip–sample interaction has occurred or not). The threshold ε is chosen to provide a suitable trade-off between the rate of false detection and the rate of missed detection [19]. The false alarm P_{F} and detection probability P_{D} are calculated as, $P_{\mathrm{F}} = P_0(\Upsilon) = \int_{\varepsilon}^{\infty} p(l = L|H_0)\,\mathrm{d}L$ and $P_{\mathrm{D}}(v) = P_1(\Upsilon) = \int_{\varepsilon}^{\infty} p(l = L|H_1, v)\,\mathrm{d}L$, respectively. $P_0(\Upsilon)$ and $P_0(\Upsilon)$ are the probability of detecting the dynamic profile Υ under hypothesis H_0 and H_1, respectively. $p(l = L|H_0)$ and $p(l = L|H_1, v)$ are the probability density functions of likelihood ratio l under hypothesis H_0 and H_1 with state jump v, respectively. Since $l = -\frac{1}{2}\sum_{i=1}^{M} e_i^{\mathrm{T}} V^{-1} e_i$ under hypothesis H_0 and e_i are independent identically distributed (iid) Gaussian random variables, $p(l = L|H_0)$ is Chi-squared (χ^2) density with M degrees of freedom. Similarly, $l = -\frac{1}{2}\sum_{i=1}^{M} e_i^{\mathrm{T}} V^{-1} e_i - \frac{1}{2}v\sum_{i=1}^{M} \Upsilon_i^{\mathrm{T}} V^{-1} e_i$ under hypothesis H_1, $p(l = L|H_1)$ is a non-central χ^2 density with non-centrality parameter $v^{\mathrm{T}}((\Upsilon^{\mathrm{T}}\Upsilon)/V)v$. Therefore, P_{D} is dependent upon values of v. For specified P_{F} or P_{D}, the threshold value ε can be computed from the tables in Reference [20]. Also given ε, the values P_{F} or $P_{\mathrm{D}}(v)$ can be computed similarly. To compute P_{D} we use v as the minimum jump that is required to be detected.

Since the dynamic profile Υ is a function of observer gain L, there is a possibility of further optimizing Equation (18) by making it a function of observer gain L.

5. EXPERIMENTAL RESULTS

Experiments were performed to ascertain the efficacy of the new transient signal-based approach. A *digital instruments* multi-mode AFM was used in the experiments. The piezodynamics was used to provide *data bits* like peaks in sample profile. The frequency response of the piezo was obtained using an HP 3563A control system analyser and a model was fit to the response. The model response is compared with that obtained experimentally in Figure 3(I). Figure 3(II) shows the experimental response of the piezo to a voltage pulse of amplitude 0.5 V, period 1000 μs and on time 500 μs. The piezodynamics results in the occurrence of four peaks separated by approximately 100 μs during the on time. The maximum width of each peak is approximately 35 μs.

The Kalman filter corresponding to the first-mode model of a cantilever at 60.025 kHz was implemented as shown in Figure 4(I). (Note that the experimental results presented in Reference [9] were derived from offline analysis of the data; here we report for the first time realtime experimental data.) The cantilever deflection signal y and the dither excitation signal g are inputs to the circuit and the estimated cantilever deflection signal $\hat{y}$ is output from the circuit. The cantilever deflection, the cantilever forcing and the estimated deflection signals are measured simultaneously at a sampling frequency of 5 MHz.

The innovation sequence in Figure 4(II;b) is the difference between the actual cantilever deflection signal and the estimated deflection signal from the analogue circuit. The innovation sequence loses its zero mean nature when the hits occur. The second-mode response of the cantilever is not included in the implemented analogue circuit for the observer. The effect of the second mode appears in the innovation sequence as the high frequency variations in the innovation signal after the initial spike due to a hit. The evidence that these effects seen in the innovation signal after a hit is due to the second mode is ascertained by computing one-mode observer-based innovation signal (see Figure 5(I;c,d)) and two-mode observer-based innovation signal (see Figure 5(II)) offline. The innovation sequence and likelihood ratio obtained from the

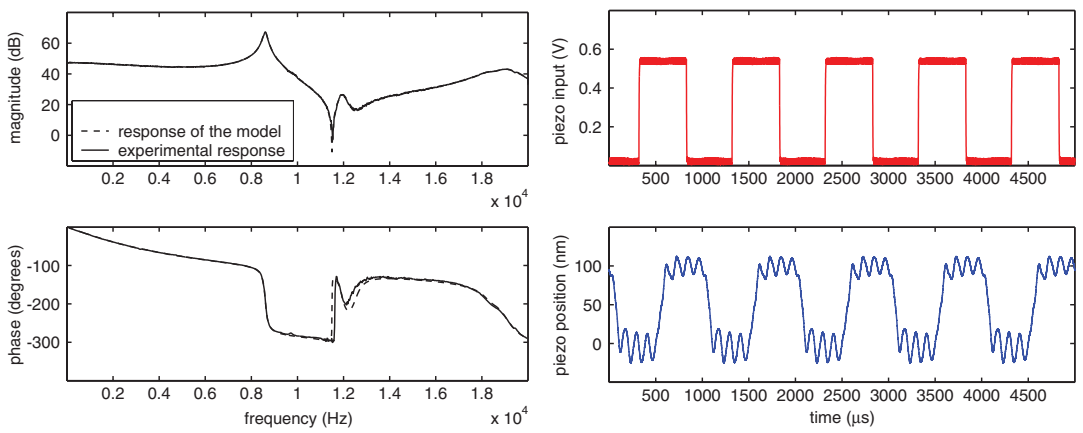

Figure 3. (I): The experimental frequency response of the z-piezo is compared with that of the model. (II): An input voltage pulse of amplitude 0.5 V, period 1000 μs and on time 500 μs results in the above piezo response. The four peaks during the on time are separated by approximately 100 μs.

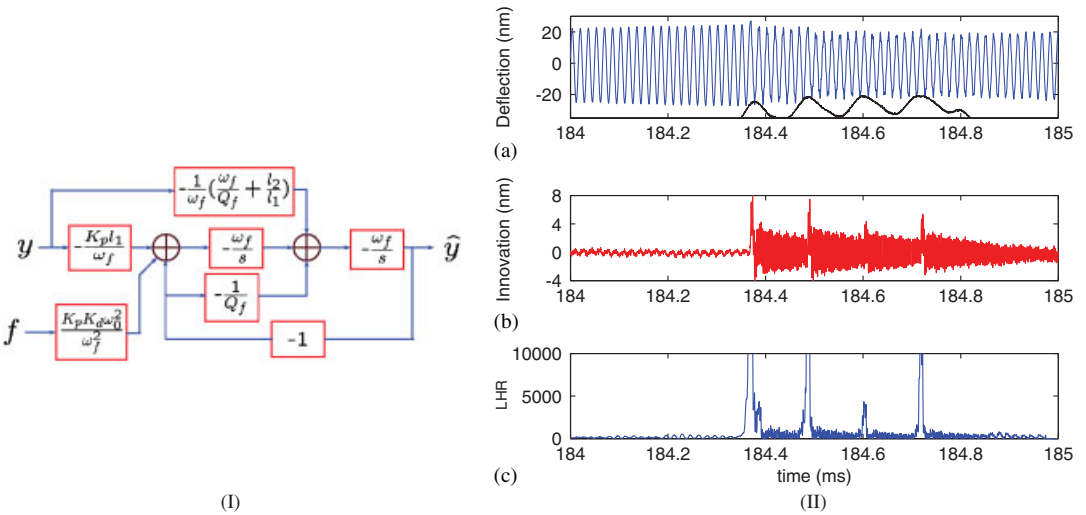

Figure 4. (I): The Kalman filter is implemented as shown in the block diagram. It is designed based on the first-mode model of the cantilever. y, f and $\hat{y}$ are the cantilever deflection, dither forcing and estimated deflection signals. K_d is the amplifier gain at the dither piezo, K_p is the photo-diode sensitivity. ω_f, Q_f and l_1, l_2 are the natural frequency and quality factor and gains of Kalman filter. (II): (a) The cantilever deflection signal is plotted against the generated sample profile; (b) every time the cantilever tip interacts with the sample the innovations loses its zero mean nature. The innovation sequence carries the signature profile and oscillations corresponding to the second mode of the cantilever; and (c) the likelihood ratio shoots up detecting the cantilever tip hitting the sample. It falls below threshold even when the second-mode oscillations persist in the innovation signal.

offline analysis and the realtime analogue circuit match well (see Figure 5(I)), confirming that the second-mode artifacts are being observed in realtime data. Note that eventhough the second mode is not included in the analogue observer (see Figure 4(I)), the likelihood ratio falls below threshold even when the second-mode artifacts of high frequency variations persist in the innovation signal. This is attributed to the fact that the dynamic profile with which the likelihood ratio is calculated is based on the first mode alone and it *filters* the second-mode effects.

Note that a bandwidth of 10 kHz is demonstrated by the experimental data. Simulation data indicate that a bandwidth of $(f_0/4)$ which equals approximately 12 kHz is possible. Further analysis of the data indicates that such a high rate is indeed possible.

The deflection and forcing signals measured above are analysed offline with a Kalman filter designed using the two-mode model (the second resonant frequency was determined experimentally to be at 346 kHz). The corresponding innovation sequence and likelihood ratio are shown in Figure 5(II). (The two-mode observer-based offline data analysis-based results was first reported in Reference [9].) It is evident that the high frequency variations in the innovation data are minimized. The two-mode-based Kalman filter is being implemented in an analogue circuit as part of ongoing work.

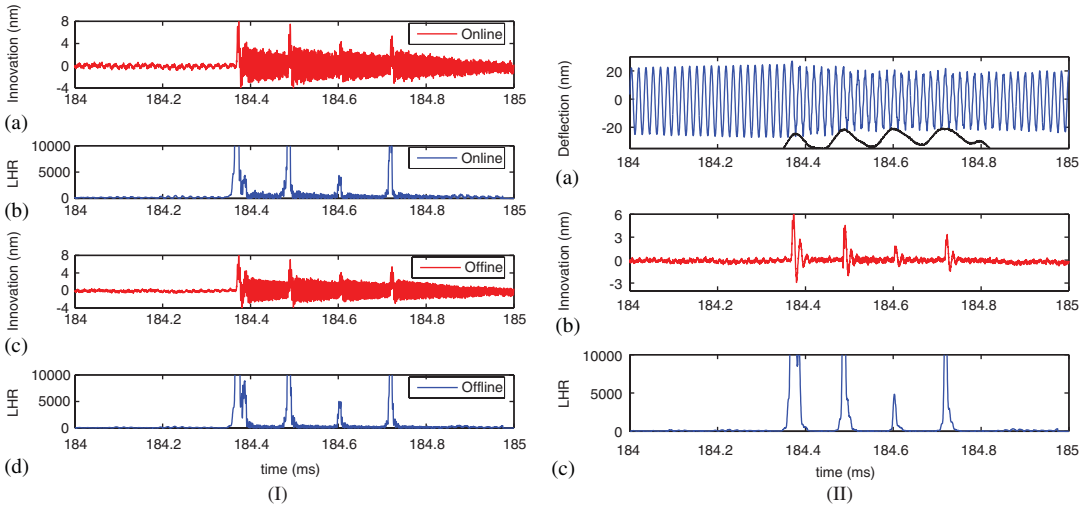

Figure 5. (I): (a) The innovation sequence obtained from realtime data is plotted; (b) the likelihood ratio computed from realtime data is plotted; (c) the innovation sequence calculated from offline Kalman filtering agrees with the realtime data in time and magnitude. Oscillations corresponding to second mode of cantilever are also present offline; and (d) the likelihood ratio computed from offline Kalman filtering data is of the same order of likelihood ratio computed from realtime Kalman filtering data. (II) Offline data processing when the two-mode model of the cantilever is incorporated in the observer: (a) the cantilever deflection signal is plotted against the pulse shape generated using the piezodynamics; (b) the innovation process is clean from the oscillations corresponding second-mode response of the cantilever. The dynamic profile is clearly visible when the hits occur; and (c) the likelihood ratio shoots up when the profile appear in the innovation sequence indicating the cantilever hitting the sample.

In another experiment the transient signal-based detection scheme was used on a calibration sample (*NanoDevices*) of pitch 2 μm and depth 20 nm. The objective was to detect the edges of the calibration sample as the oscillating cantilever is scanned over the sample surface. A scan size of 40 μm was chosen at a scan rate of 1.5 Hz. Since the scan size was large, the sample slope has a significant impact on the deflection signal if the scanning is performed in open-loop. To circumvent this problem, an amplitude feedback scheme (see Figure 6(I)) was employed where the amplitude of the oscillating tip is fed back to compensate for the slope. A proportional integral controller was employed for this purpose. The deflection signal measured while scanning in open-loop and closed loop are shown in Figure 6(II). Note that the open-loop scan results in non-uniform hits with the edges of the calibration sample in spite of a sample slope of less than 0.1°. With amplitude feedback, we are able to obtain uniform hits.

The error process is obtained as in the case of the previous experiment. In this experiment due to the amplitude feedback, the cantilever is constantly interacting with the sample surface. Since the amplitude set point is maintained close to the free oscillation amplitude, the impact is minimal except when the edges of the calibration sample are encountered. This is due to the slow response of the feedback controller that is based on the amplitude which is slowly varying.

In Figure 7(I), the deflection signal, error process and the likelihood ratio are shown for the scan size of 40 μm. Since the pitch of the calibration sample is 2 μm, there are 20 spikes in the

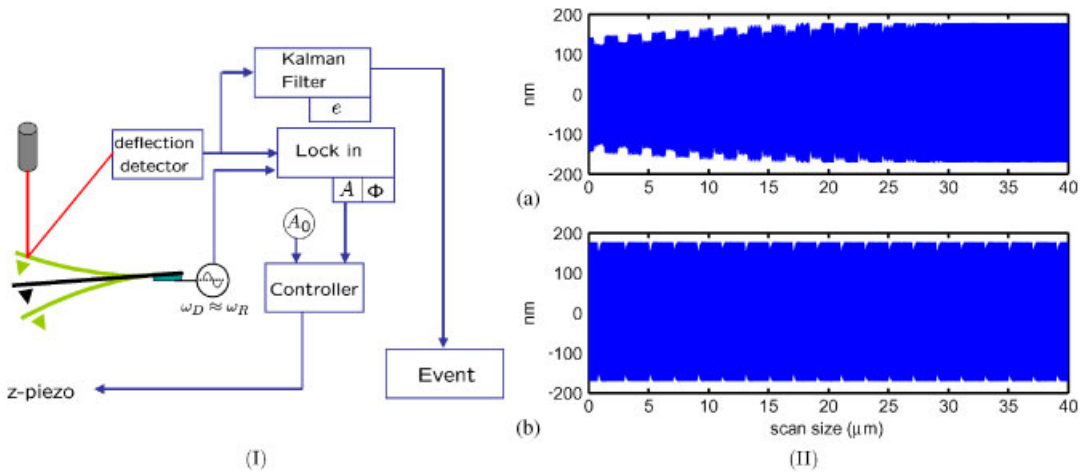

Figure 6. (I): The excitation frequency ω_D is set close to the resonant frequency ω_R of the cantilever. The amplitude A of the deflection signal of the cantilever is controlled at a set-point amplitude of A_0 by actuating the z-piezo in a feedback manner. The innovation signal from the Kalman filter is analysed by the transient detection method to infer variations in the sample profile. A_0 is kept close to the free oscillation amplitude so that the tip–sample interaction is mild. (II): (a) An oscillating cantilever is scanned over a calibration sample with pith 2 μm and depth 20 nm; and (b) the oscillating cantilever is scanned over the calibration sample while feeding back on the amplitude which corrects for the sample slope and low-frequency drifts seen in (a).

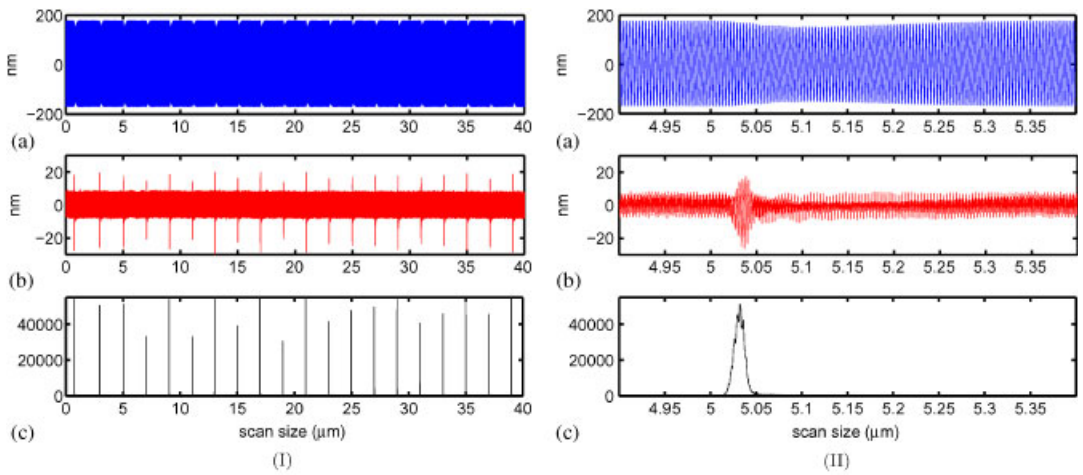

Figure 7. (I): (a) The deflection signal is sampled at 5 MHz; (b) the error process has significant higher harmonics due to the persistent interaction with the sample; and (c) the likelihood ratio could be used to detect all the edges with high resolution. (II): (a) The oscillating cantilever encounters an edge at 5.025 μm. The oscillations gradually decay to a smaller amplitude and recovers the set point amplitude at around 5.35 μm; (b) the error process shows the dynamic profile almost as soon as the encounter with the edge is made. The dynamic profile vanishes even when the cantilever is still in the transient state making it possible to detect another hit detectable during this period; and (c) the likelihood ratio distinctly shows the presence of the dynamic profile at high resolution.

innovations and the likelihood ratio. Figure 7(II) depicts a zoomed version of Figure 7(I). The error process starts showing the dynamic profile as soon as the impact with the edge of the calibration sample occurs. However, the deflection signal shows a gradual decay in amplitude. The error process, on the other hand, returns to the nearly white (albeit the higher harmonics are present due to the persistent interaction) nature even when the cantilever is still in the transient state. This experiment demonstrates that the pitch of the sample can be accurately determined very fast and has significant relevance to accurate calibration of sample pitch.

6. CONCLUSION

Note that all present schemes employ steady-state data. In this article a framework for ultra-fast interrogation of sample in atomic force microscopy is proposed which utilizes the tip-deflection data during the transient state of the cantilever probe. The systems perspective has facilitated the development of this methodology. The cantilever and its interaction with the sample is modelled. A first- and second-mode approximation model of the cantilever is considered and a Kalman filter is designed to estimate the dynamic states. The tip–sample interaction is modelled as an impulsive force applied to the cantilever in order to detect the presence of sample. The dynamics due to tip–sample interaction is calculated in the innovation sequence and a likelihood ratio test is performed to obtain the decision rule to infer the presence of sample. Experimental results tally with the simulation results verifying the proposed methodology and the sample-detection scheme.

Simulations show a bandwidth of $f_0/4$ (f_0 being the natural frequency of the micro-cantilever) in detecting small time scale tip–sample interactions. The result is corroborated in experiments and sample profiles appearing at 10 KHz were detected using a cantilever with resonant frequency equal to 70 KHz in realtime by implementing analogue observer. In this method high quality factor does facilitate high resolution but it does not limit the bandwidth as in steady-state data-based methods.

The transient signal-based sample-detection scheme is integrated with regular tapping-mode operation of AFM to detect high bandwidth content sample features at large scan rates in the case where the tapping-mode images failed to produce useful image. As part of ongoing work experiments to ascertain sub-nanometer scale features like atomic layers of graphite and the lattice structure of mica are being conducted at a scan rate an order higher than the regular tapping-mode operation with encouraging preliminary results.

ACKNOWLEDGEMENTS

This research is supported by National Science Foundation Grant ECS-0330224 to Prof. Murti V. Salapaka.

REFERENCES

1. Binnig G, Quate CF, Gerber C. Atomic force microscope. *Physical Review Letters* 1986; **56**(9):930–933.
2. National Nanotechnology Initiative: The Initiative and its Implementation Plan, http://www.nano.gov/nni2.pdf, July 2000, 2003.

3. Hoh JH, Lal R, John SA, Revel JP *et al*. Atomic force microscopy and dissection of gap junctions. *Science* 1991; **253**:1405–1408.
4. Kasas S, Thomson NH, Smith BL, Hansma HG *et al*. *Escherichia coli* RNA polymerase activity observed using atomic force microscopy. *Biochemistry* 1997; **36**(3):461–468.
5. Rugar D, Yannoni CS, Sidles JA. Mechanical detection of magnetic resonance. *Nature* 1992; **360**:563–566.
6. Sidles JA. Noninductive detection of single proton-magnetic resonance. *Applied Physics Letters* 1991; **58**(24): 2854–2856.
7. Sidles JA. Folded stern-gerlach experiment as a means for detecting nuclear magnetic resonance of individual nuclei. *Physical Review Letters* 1992; **68**:1124–1127.
8. Sidles JA, Garbini JL, Drobny GP. The theory of oscillator-coupled magnetic resonance with potential applications to molecular imaging. *Review of Science Instruments* 1992; **63**:3881–3899.
9. Sahoo DR, Sebastian A, Salapaka MV. Transient-signal-based sample-detection in atomic force microscopy. *Applied Physics Letters* 2003; **83**(26):5521–5523.
10. Mamin HJ, Ried RP, Terris BD, Rugar D. High-density data storage based on the atomic force microscope. *Proceedings of the IEEE* 1999; **87**(6):1014–1027.
11. Salapaka MV, Bergh HS, Lai J, Majumdar A, McFarland E. Multimode noise analysis of cantilevers for scanning probe microscopy. *Journal of Applied Physics* 1997; **81**(6):2480–2487.
12. Israelachvili JN. *Intermolecular and Surface Forces*. Academic Press: New York, 1985.
13. Sebastian A, Salapaka MV, Chen DJ, Cleveland JP. Harmonic analysis based modeling of tapping-mode afm. In *Proceedings of the American Control Conference*, California, San Diego, June 1999.
14. Sebastian A, Salapaka MV, Chen DJ, Cleveland JP. Harmonic and power balance tools for tapping-mode atomic force microscope. *Journal of Applied Physics* 2001; **89**(11):6473–6480.
15. Chen C-T. *Linear System Theory and Design*. Oxford University Press: Oxford, 1999.
16. Salapaka MV, Chen D, Cleveland JP. Linearity of amplitude and phase in tapping-mode atomic force microscopy. *Physical Review B* 2000; **61**(2):1106–1115.
17. Sayed AH, Kailath T, Hassibi B. *Linear Estimation*. Prentice Hall: NJ, 2000.
18. Willsky AS, Jones HL. A generalized likelihood ratio approach to the detection and estimation of jumps in linear systems. *IEEE Transactions on Automatic Control* 1976; **February**:108–112.
19. Vincent Poor H. *An Introduction to Signal Detection and Estimation* (2nd edn). Springer: Berlin, 1994.
20. Owen DB. *Handbook of Statistical Tables*. Addison Wesley: Reading, MA, 1922.

