2013 American Control Conference (ACC)
Washington, DC, USA, June 17-19, 2013

Topography detection using innovations mismatch method for high speed and high density dynamic mode AFM

Sayan Ghosal[1], Govind Saraswat[2], Aditya Ramamoorthy[3] and Murti Salapaka[4]

***Abstract*— The atomic force microscope (AFM) is one of the major advances in recent science that has enabled imaging of samples at the nanometer scale. Over the years, different techniques have been developed to improve the speed, resolution and accuracy of imaging using AFM. As AFMs can scan and deform material with extremely high resolution, it has also been used as a data read-write system where the high or low topography of the sample surface are interpreted as a one or a zero bit. Data storage using this method can produce extremely high data storage density. In this paper a new method called the innovations mismatch method (IM) is developed that can be used for both imaging and data storage applications. The IM scheme utilizes the dynamic mode of operation and therefore is applicable to soft matter interrogation. In this work, IM method is shown to outperform previously developed techniques.**

I. INTRODUCTION

Since its invention [2], the atomic force microscope (AFM) has become the tool of choice for investigating materials with nanometer scale resolution. The basic operation of an AFM is described in Fig. 1. As AFMs can be used to deform material and measure topography on a material with nanometer scale resolution, researchers have developed it for ultra-high density data storage purposes [11]. AFM based data storage systems can achieve extremely high areal density (upto 3 Tb per square inch) [6]. In probe based data storage using static mode of operation [3], the cantilever tip always interacts with the sample. The tip movement parallel to the surface creates large lateral frictional forces resulting in wear and tear of the material especially when the material is soft. In the amplitude modulation AFM (AM-AFM) which is a dynamic mode of operation (see Fig. 1), the cantilever is forced sinusoidally at or near the first resonant frequency. Under the presence of the sample, the cantilever interacts with the sample intermittently every oscillation cycle that alters the amplitude and phase of cantilever oscillations. The changes in measured amplitude of the cantilever trajectory is used to interpret the sample topography. The dynamic mode reduces the large lateral frictional forces present in the contact mode scheme substantially due to the intermittent contact with sample. However, the high quality factors of the cantilever (low damping) creates large transients and slow down imaging. In [7], a faster dynamic mode imaging method called the transient force atomic force microscopy (TF-AFM) was introduced that achieved significantly higher speed than conventional AM-AFM. TF-AFM utilized the

[1]S. Ghosal, [2]G.Saraswat and [4] M.Salapaka are with Dept. of Electrical Engineering, University of Minnesota, Minneapolis, MN 55455, USA, (e-mail: {ghos0087,saras006,murtis}@umn.edu). [3] A. Ramamoorthy is with Dept. of Electrical and Computer Engg. at Iowa State University, Ames IA 50010, USA, (e-mail: adityar@iastate.edu)

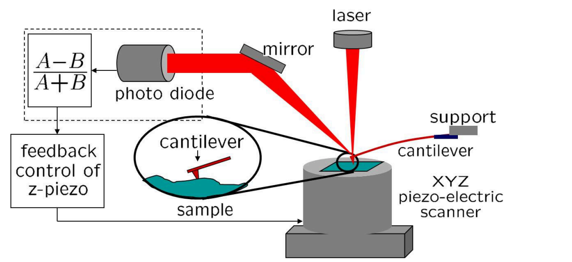

Fig. 1. Atomic force microscope (AFM). The main components are a micro-cantilever, a piezoelectric nanopositioner and a laser-photodiode based cantilever tip position detection system. In dynamic mode of operation, the cantilever is sinusoidally oscillated typically near the first resonance frequency of the cantilever. The tip-sample interactions modify the oscillation amplitude and phase of the cantilever tip and these modifications are utilized in infering sample topography.

linear time-invariant (LTI) behavior of the cantilever dynamics to realize an observer. Typically the cantilever amplitude is in the 27-200 nm range whereas the force on the cantilever due to the sample is effective only within 3-4 nm range from the sample surface. Exploiting the short duration of the interaction, TF-AFM method assumes that the cantilever interaction with the sample can be approximated as an impulsive hit in every oscillation cycle that abruptly changes the cantilever state. Then the problem is reduced to the detection of the cantilever state jumps. With small measurement noise, the observer's tracking bandwidth can be decoupled from the quality factor (damping) of the cantilever where the observer is capable of tracking the cantilever state even when it is in transient phase after an interaction with the sample. Thus, TF-AFM is a fast imaging method. However, the magnitude of the impulsive impacts strongly depends on the past cantilever trajectory. After an impulsive interaction with the sample, typically, the cantilever tip interaction with the sample becomes milder. This is evident in Fig. 2. The signal labeled sep in Fig. 2 is the separation between the cantilever base and the sample. High sep indicates free air cantilever oscillations. Low sep (or raised sample) implies cantilever interacts with the sample with interaction length of 2 nm. When the sample is raised at time 1 $\mu\,sec$, the cantilever experiences a large impulsive interaction with the sample, that leads to loss of tracking resulting in large error signal (labeled *innov*) between the actual cantilever deflection and the deflection estimated by the observer. However, with topography raised at the same level, amplitude of the cantilever oscillation reduces. Thus, the strength of the impulsive forces on the cantilever are reduced leading to a smaller *innov* signal. As a result, persistent high topography becomes difficult to detect. In data storage settings, different topographic profiles etched on the material encode bits 0 and 1. For example, a raised topographic profile can encode

978-1-4799-0178-4/$31.00 ©2013 AACC

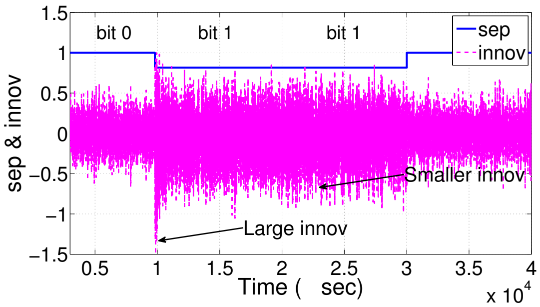

Fig. 2. Innovation from free space observer vs separation of cantilever base from sample. During the initial phase of interaction *innov* has high magnitude. However, with persistent high topography *innov* loses its initial strength.

bit 1 and a depressed topographic profile can encode bit 0. It is evident from the earlier discussion that TF-AFM suffers while detecting a consecutive sequence of one bits. To overcome this drawback, in [5], the cantilever system is modeled as a communication channel, and a maximum a posterior estimation methodology is adopted. This method involves learning how the *innov* signal changes for different bit sequences and then exploiting this information in a real-time detection application. This technique resulted in significant reduction in bit error rates. However, it involves a learning step where the exact topographic profile that encodes a one and a zero plays an important role. Such an assumption is valid in a data storage scenario where the topographic profile is designed. However, in imaging situation such an assumption is not valid as the topographic profile is not known a-priori.

In this article we substantially improve over the TF-AFM method, where we obtain, an LTI *equivalent* model, that characterizes the cantilever behavior under sample's influence using averaging techniques. This model is used in conjunction with the cantilever model to realize two observers; one matched to the sample free situation and another to the situation where sample is present. The effectiveness of this strategy is demonstrated first via simulations followed by experimental results. We show that the problem of low SNR of the TF-AFM method under persistent interaction is alleviated. A maximum likelihood sequence detection scheme is also developed that gains over bit by bit detection. This article is organized as follows. Section II explains the basic system and detection models that are developed in previous works. In section III, the concept of equivalent cantilever model is explained and the idea of innovations mismatch (IM) is developed. Section IV explains the sequence detection method applied on innovations mismatch signal. Section V verifies the benefit of IM with simulation results. An experimental result is presented in section VI that provides strong evidence for applicability of IM. Section VII summarizes this paper and concludes with future works.

II. Background and Significance

A. System model and test signal

The system model and the detection schemes used in the previous works are briefly explained in this section. The first mode approximation of the cantilever is given by the spring mass damper dynamics as below [9]:

$$\ddot{p} + \frac{\omega_0}{Q}\dot{p} + \omega_0^2 p = f(t) = \frac{1}{m}\left(\eta + g + h\right), \qquad (1)$$

$$h = \phi(p, \dot{p}),\; y = p + v, \qquad (2)$$

where p is tip deflection, $\dot{p}$ is the tip velocity, m is the cantilever mass, $f(t)$ is force per unit mass on the cantilever, η is thermal noise input, g is the dither forcing input and $h = \phi(p, \dot{p})$ is tip-media interaction force (ϕ is a nonlinear function of the cantilever position and velocity). Typically, g is a sinusoidal signal with frequency close or equal to the first resonant frequency ω_0. y is the measured deflection and v is the measurement noise. Q is the quality factor and ω_0 is the first resonant frequency given by $\omega_0 = \sqrt{\frac{k}{m}}$, $Q = \frac{\sqrt{km}}{c}$. c and k are the damping constant and the stiffness of the spring-mass-damper model of the cantilever (obtained by the first mode approximation). A state space representation is described by

$$\dot{x} = Ax + Bf,\; y = Cx + v. \qquad (3)$$

Here, x is the state vector given by $[p \;\; \dot{p}]^T$, where $\dot{p}$ is the cantilever tip velocity. The matrices A, B and C are

$$A = \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & -\frac{\omega_0}{Q} \end{bmatrix},\; B = \begin{bmatrix} 0 \\ 1 \end{bmatrix},\; C = \begin{bmatrix} 1 & 0 \end{bmatrix}. \qquad (4)$$

As explained in [7], the TF-AFM scheme uses a Kalman observer together with the original system to have an estimate of the tip deflection. A continuous time Kalman filter matched to the free space model is given by

$$\dot{\hat{x}} = A\hat{x} + B\frac{g}{m} + L(y - \hat{y}),\; \hat{x}(0) = \hat{x}_0,\; \hat{y} = C\hat{x}, \qquad (5)$$

where L is the Kalman gain, $\hat{x}$ is the estimated state vector$=$ $[\hat{p} \;\; \dot{\hat{p}}]^T$. Hence the error between the actual state and the estimated state is governed by the following dynamics

$$\dot{\tilde{x}} = [A - LC]\,\tilde{x} + \begin{bmatrix} B & -L & B \end{bmatrix}\begin{bmatrix} \eta/m \\ v \\ h/m \end{bmatrix}. \qquad (6)$$

$$e = (y - \hat{y}) = C\tilde{x} + \begin{bmatrix} 0 & 1 \end{bmatrix}\begin{bmatrix} \eta \\ v \end{bmatrix}. \qquad (7)$$

e (also termed as innovation) is the error between the measured cantilever deflection and the deflection estimated by the observer. For sinusoidal excitation g, cantilever oscillates sinusoidally in steady state with amplitude typically in the 25 nm to 200 nm range. Normally in dynamic mode operation, interaction with the sample extends withing a few nanometers which is a very small fraction of cantilever orbit. Hence, the sample force h can be modeled as an impulsive force that suddenly changes the cantilever state. With such an assumption the discretized model of the cantilever dynamics is given by [7]

$$x_{k+1} = Fx_k + G(g_k + \eta_k) + \delta_{\theta,k+1}\nu,\; y_k = Hx_k + v_k. \qquad (8)$$

Matrices F, G and H can be obtained from continuous time model by discretization. Here, $\delta_{\theta,k+1}\nu$ denotes the sudden change in the state at the time of impact. ν models the impact strength and θ models the time of impact ($\delta_{i,j} = 1$ when $i = j$, else 0). The discrete time Kalman observer model is

$$\begin{aligned}\hat{x}_{k+1} &= F\hat{x}_k + Gg_k + L_K\left(y_k - H\hat{x}_k\right), \\ \hat{y}_k &= H\hat{x}_k,\ e_k = \left(y_k - \hat{y}_k\right).\end{aligned} \tag{9}$$

B. Detection algorithm

From equations (8) - (9) it can be shown that once there is an impact at time θ, the discrete time innovation e_k can be written as [7]

$$e_k = \Gamma_{k;\theta}\nu + n_k;\ \ \Gamma_{k;\theta} = H\left(F - L_K H\right)^{k-\theta}, \tag{10}$$

where n_k is discrete time Gaussian noise with variance V. Hence, $\Gamma_{k;\theta}$ (which is the impulse response from an impulsive force input at time θ to the innovation signal) can be used to realize a matched filter to extract an impulsive force interaction from the sampled innovations. Let M be the sample window size that is used for detection. The binary hypotheses

$$H_0 : e_k = n_k,\ H_1 : e_k = \Gamma_{k,\theta}\nu + n_k, \tag{11}$$

represent absence and presence of sample; $k = 1, 2, \cdots, M$. H_0 is the hypothesis that e_k is white noise. H_1 is the hypothesis that the profile $\Gamma_{k;\theta}\nu$ (ν is the strength of the impulsive interaction) appears in e_k. Here, e_k and n_k are the sampled innovation and noise. In case of locally most powerful (LMP) technique the likelihood ratio is given by (see [5])

$$l_{lmp}(M) = \bar{e}^T V^{-1}\Gamma, \tag{12}$$

where $\Gamma = [H, H(F - L_K H), \cdots, H(F - L_K H)^{M-1}]^T$. $\bar{e} = [e_1, e_2, \cdots, e_M]^T$ is the observed innovations vector. We remark that l_{lmp} correlates the discrete innovation sequence with the profile $\Gamma_{k;1}$. M is chosen depending on the effective duration of $\Gamma_{k;\theta}$. Observed noise vector $\bar{n} = [n_1, n_2, \cdots, n_M]^T$ is zero mean with covariance matrix $VI_{M\times M}$. For imaging, l_{lmp} is compared to a threshold τ_1 $\left(l_{lmp} \lessgtr_{H_1}^{H_0} \tau_1\right)$ that is chosen to have a good balance between false alarm and detection probability. As pointed out in Section I, one of the main issues with TF-AFM is that it is a force detector and the current impulsive force ν depends on the past topography. Indeed, a persistent raised topography reduces the strength of ν. Thus, TF-AFM detection method generates high SNR at the introduction of a step whereas SNR is small near its steady state. In case of probe storage, suppose a single bit lasts for q oscillation cycles. Then l_{lmp} obtained in each cycle is observed and in a bit duration (say T), q LMP outputs $(l_{1,lmp}(M), l_{2,lmp}(M), \cdots, l_{q,lmp}(M))$ are stored. The presence or absence of a bit is decided by (see [5]) $\max\left(l_{1,lmp}(M), l_{2,lmp}(M), \cdots, l_{q,lmp}(M)\right) \lessgtr_{H_1}^{H_0} \tau_1$. Here again, τ_1 is chosen to tailor a desired balance between detection and false alarm probabilities. In [5], a communications model is developed that takes into account the past topography encountered. This method involves a maximum a-posteriori estimation of the topography using Viterbi detection and yields significant improvement in bit error rate (BER). However, it requires a statistical learning step that is effective when the nature of the topographic profile encoding a bit is known (this holds in a data storage application). In an imaging application, where the nature of the topography is less certain the effectiveness remains to be seen. This article introduces a new method that alleviates drawbacks of the TF-AFM method and the sequence estimation based approach. The details follow in the next section.

III. Use of on-sample innovations mismatch

In this section a new approach is proposed that is based on the *equivalent cantilever* perspective. When the cantilever interacts with the material, $h = \phi(p, \dot{p})$ in (1) is not zero. Since, $\phi(p, \dot{p})$ is a complicated nonlinear function (see [10]), it is difficult to obtain a direct solution of the differential equation (1). However, when the steady state is reached, interaction occurs in each cycle of oscillation. Hence the force h can be assumed to be periodic with the same period as g. The cantilever gain at resonant frequency is high. So the excitation g required is small. Again, most of the cantilever orbit is in free air, so the damping c is small. Finally, the periodic nonlinear interaction h for typical interaction lengths considered is very small too. With such small magnitude assumptions on c, g and $h = \phi(p, \dot{p})$, they can be replaced by $c = \epsilon c_1$, $g = \epsilon g_1$, $\phi = \epsilon\phi_1$. where ϵ is a small parameter. Ignoring the noise term, from (1) we can write $\ddot{p}+\omega^2 p = \frac{\epsilon}{m}\phi_1(p, \dot{p}) + \frac{\epsilon}{m}g_1(t) - \frac{\epsilon}{m}c_1\dot{p}$. Then, we can change the coordinates from $(p, \dot{p}) \mapsto (a, \theta)$ using $p = a\cos(\omega t+\theta)$, $\dot{p} = -a\omega\sin(\omega t + \theta)$. Differentiating the new coordinates with respect to time we can get dynamics of the changed coordinates as:

$$\begin{aligned}\dot{a} = &-\frac{\epsilon}{\omega m}[\phi_1(a\cos(\omega t + \theta), -a\omega\sin(\omega t + \theta)) \\ &+ c_1 a\omega\sin(\omega t + \theta) + g_1(t)]\sin(\omega t + \theta),\end{aligned} \tag{13}$$

$$\begin{aligned}\dot{\theta} = &-\frac{\epsilon}{\omega m}[\phi_1(a\cos(\omega t + \theta), -a\omega\sin(\omega t + \theta)) \\ &+ c_1 a\omega\sin(\omega t + \theta) + g_1(t)]\cos(\omega t + \theta).\end{aligned} \tag{14}$$

Assuming the forcing $g_1(t) = E\sin(\omega t)$ it can be seen that dynamics of (13) and (14) are periodic with period $\frac{2\pi}{\omega}$. Now we can apply the first order periodic averaging theory [8] that states that if we consider

$$\dot{x} = \epsilon f(t, x);\ \dot{x}_{av} = \epsilon f_{av}(x_{av}) \tag{15}$$

where $f(t, x)$ is T periodic in the variable t and $f_{av}(x) := \frac{1}{T}\int_0^T f(\tau, x)\,\mathrm{d}\tau$ with initial conditions $x(0) = x_{av}(0) = x_0$. Then with mild assumptions on f, there exist constraints L and M' such that

$$\sup_{t\in[0,L/\epsilon]}|x(t) - x_{av}(t)| \leq M'\epsilon. \tag{16}$$

Application of theorem (16) to the time varying dynamics (13) and (14) results in the averaged time invariant dynamics (using the notations a and θ for time averaged amplitude and

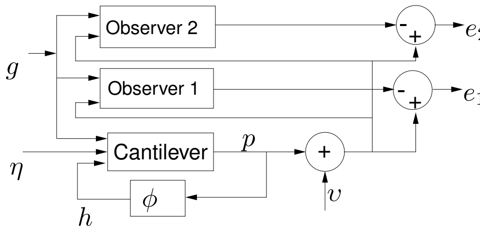

Fig. 3. Two observers architecture of IM technique

phase) as described by

$$\begin{aligned}\dot{a} &= -\frac{c_e(a)}{2m}a - \frac{\epsilon}{2m\omega}E\sin\theta,\\ \dot{\theta} &= \omega_e(a) - \omega - \frac{\epsilon}{2ma\omega}E\cos\theta,\end{aligned} \tag{17}$$

where $\omega_e^2(a) = w^2 - \frac{2}{am}\bar{\phi}_c$, $\frac{c_e(a)}{2m} = \frac{c_1}{m} + \frac{1}{am\omega}\bar{\phi}_d$. The parameters $\bar{\phi}_c$ and $\bar{\phi}_d$ are given by $\bar{\phi}_c = \frac{1}{2\pi}\int_0^{2\pi}\phi\left(a\cos\psi, -a\omega\sin\psi\right)\cos\psi\,\mathrm{d}\psi$, $\bar{\phi}_d = \frac{1}{2\pi}\int_0^{2\pi}\phi\left(a\cos\psi, -a\omega\sin\psi\right)\sin\psi\,\mathrm{d}\psi$. One can get back to the original coordinates $(p, \dot{p})$ using the relations $p = a\cos(\omega t + \theta)$, $\dot{p} = -a\omega\sin(\omega t + \theta)$. a and θ follow the averaged dynamics (17). It can be shown that p obtained this way satisfies the following equation

$$\ddot{p} + \frac{c_e}{m}\dot{p} + \omega_e^2 p = \frac{1}{m}g(t). \tag{18}$$

This dynamics describes an equivalent cantilever with changed resonant frequency ω_e and damping c_e. We remark that the equivalent damping and equivalent stiffness are functions of amplitude. However, the amplitude evolves on a slow time scale and thus can be considered constant over the time scale on which the sample presence/absence has to be inferred. Thus we have realized an equivalent LTI model of the cantilever-sample interconnection. Corresponding equivalent stiffness is $k_e = m\omega_e^2$ and quality factor $Q_e = \sqrt{k_e m}/c_e$. As found in [1], ω_e and Q_e normally reduce with increase in interaction length l_{int}. If for some value of l_{int}, one has the estimate of ω_e and Q_e, a new state space model can be obtained as shown in (4). Hence, one can design a new observer and obtain a new innovation signal similar to (5) and (6). Let the innovation signal obtained from the free space Kalman observer be e_1 and that from equivalent Kalman observer be e_2 (see figure 3). For free space oscillations, e_1 has small magnitude whereas it is large when cantilever interacts with material. e_2 on the other hand is expected to follow an opposite trend. TF-AFM uses only one innovation signal $e_1(t)$ to run detection algorithms. However, in IM, both the signals $e_1(t)$ and $e_2(t)$ as shown in Fig. 3 are utilized. A new signal called the innovations mismatch signal $i_m(t)$ is defined as $i_m(t) = |e_1(t)| - |e_2(t)|$. Since, $i_m(t)$ is formed from the difference of the instantaneous magnitudes of $e_1(t)$ and $e_2(t)$, it is expected to respond faster than either of the innovation signals when they are considered individually. Hence, whenever there is a change from interaction to no interaction (or the reverse), $i_m(t)$ responds faster than e_1 or e_2. Also, it will be seen in the next section that the envelope of the signal $i_m(t)$ conveys the topography information of the sample. A low pass filter with appropriate bandwidth can be used to extract the envelope of $i_m(t)$. The low pass filtered version of $i_m(t)$ is termed as $s_m(t)$. $s_m(t)$ can be compared with an appropriate threshold to infer presence or absence of sample. The cut-off frequency of the low pass filter is determined from the estimate of how fast the topography changes are encountered. In probe based data storage applications $s_m(t)$ can be further processed to detect bits. A square wave $\sqcap(T)$ signal can be used as the matched filter on $s_m(t)$ which provides the area under the curve corresponding to the bit interval. By comparing this area with appropriate threshold τ, whether the bit is 1 or 0 can be decided. In case of sequence detection, $s_m(t)$ can be used as the test signal which is explained in the next section.

IV. Sequence Detection

A gist of the main concepts of sequence detection is provided here. The sequence detector processes the entire sequence of samples from signal $s_m(t)$ to decide the maximum likely sequence of source bits. In a data storage setting let $\bar{a} = [a_0, a_1, \cdots, a_{N-1}]^T$ be the source bit sequence ($a_k \in \{0, 1\}$ for $k = 0, 1, \cdots, N-1$). This bit sequence is etched onto the sample. A raised topography denotes a 1 and a lowered topography denotes a 0. Let T denote the time needed by the cantilever to scan the topography representing a single bit. Let q be the number of cantilever oscillations per bit duration T. The sampled $s_m(t)$ is denoted by the vector $\bar{z} = [z_0, z_1, \cdots, z_{(Nq-1)}]^T$. Let $f(\bar{z}|\bar{a})$ be the conditional probability density function of the output vector $\bar{z}$ given a source bit sequence $\bar{a}$. The maximum likely bit sequence is given by $\hat{\bar{a}} = \arg\max_{\bar{a}\in\{0,1\}^N} f(\bar{z}|\bar{a}) = \arg\max_{\bar{a}\in\{0,1\}^N}\prod_{i=0}^{N-1} f(\bar{z}|\bar{a}, \bar{z}_0^{i-1})$. $\bar{z}_j^i$ is defined as the vector $\bar{z}_j^i = [z_{jq}, z_{jq+1}, \cdots, z_{iq-1}]^T$. We also define $\bar{z}_i = [z_{iq}, z_{iq+1}, \cdots, z_{i(q+1)-1}]^T$ as the sampled vector corresponding to i^{th} bit interval. The i^{th} state S_i is defined as $S_i = [a_{i-m-m_I+1}, a_{i-m-m_I+2}, \cdots, a_i]^T$ where the memory in the system is given by the past $m + m_I$ source symbols. Hence, the problem of identifying the maximum likely sequence of bits can be recast as the problem of identifying the maximum likely sequence of states. The conditional pdf $f(\bar{z}_i|\bar{a}, \bar{z}_0^{i-1})$ can be simplified by removing the history of the source bits $\bar{a}$ and observed outputs $\bar{z}$ beyond the assumed memories. In [5] it is assumed that $f(\bar{z}_{i-m_I}^i|S_i, S_{i-1})$ is Gaussian i.e. $f(\bar{z}_{i-m_I}^i|S_i, S_{i-1}) \sim N\left(\bar{\mathcal{Y}}(S_i, S_{i-1}), \mathcal{C}(S_i, S_{i-1})\right)$ with the mean vector $\bar{\mathcal{Y}}(S_i, S_{i-1})$ and the covariance matrix $\mathcal{C}(S_i, S_{i-1})$. $\bar{\mathcal{Y}}$ and $\mathcal{C}$ can be estimated by a statistical learning step before doing the actual bit detection. Once the statistics is available, the maximum likely sequence of states $\hat{\bar{S}}$ can be estimated from the observation vector $\bar{z}$ as below:

$$\begin{aligned}\hat{\bar{S}} = \arg\min_{\text{all}\bar{S}} &\sum_{i=0}^{N-1}\log\left(\frac{\mathcal{C}(S_i, S_{i-1})}{c(S_i, S_{i-1})}\right)\\ &+(\bar{z}_{i-m_I}^i - \bar{\mathcal{Y}}(S_i, S_{i-1}))^T\mathcal{C}(S_i, S_{i-1})^{-1}\\ &(\bar{z}_{i-m_I}^i - \bar{\mathcal{Y}}(S_i, S_{i-1}))\\ &-(\bar{z}_{i-m_I}^i - \bar{y}(S_i, S_{i-1}))^T c(S_i, S_{i-1})^{-1}\\ &(\bar{z}_{i-m_I}^i - \bar{y}(S_i, S_{i-1})),\end{aligned} \tag{19}$$

where $c(S_i, S_{i-1})$ is upper $m_I q \times m_I q$ principal minor of $\mathcal{C}(S_i, S_{i-1})$ and $\bar{y}(S_i, S_{i-1})$ is the first $m_I q$ elements of the mean vector $\bar{\mathcal{Y}}(S_i, S_{i-1})$. With the help of this cost function (19), the maximum likely sequence of source bits can be found out employing the Viterbi algorithm (see [4]). We denote Viterbi detection applied on $s_m(t)$ as Viterbi(new). Viterbi detection on a single innovation based only on the signal $e_1(t)$, as described in [5] is referred to as Viterbi(old). Detection performances of different techniques are compared in the next section.

V. Simulation results

In this section, superior performance of IM is verified with simulation results. Simulations are done with a Simulink model and MATLAB codes that capture the dynamic mode AFM setup fairly well. The cantilever is oscillated with its first resonance frequency $f_0 = 63.15\,\text{kHz}$ and the noisy measurement of deflection is taken. sep denotes the normalized separation between the cantiliver base and the sample. When sep is high, the cantilever is oscillating in free air. sep low indicates that cantilever is interacting with the material with interaction length of $l_{int} = 2$ nm. The equivalent plant model corresponding to l_{int} of 2 nm can be found using a recursive least squares method [1]. A second Kalman observer matched to the estimated model is designed. Using two Kalman observers, signals e_1 and e_2 as described in previous section can be generated. The time evolution of e_1 and e_2 corresponding to sep is shown in Fig. 4. Fig. 4 shows that for smaller sep (interaction), e_1 is stronger than e_2 whereas the reverse happens when sep is high (no interaction). $|e_2| - |e_1|$ corresponding to the same separation pattern is shown in Fig. 5. We also define $i'_m(t) = -i_m(t) = (|e_2| - |e_1|)$. Clearly from Fig. 5, the innovations mismatch signal $i'_m(t)$ responds faster than the individual signals $e_1(t)$ or $e_2(t)$. When there is a change in topography, envelope of $i'_m(t)$ follows the change rapidly. $s'_m(t)$ is generated by low pass filtering $i'_m(t)$ (see Fig. 6). Bit duration is $T = 200\,\mu$sec. Hence, the low pass filter band width is chosen to allow signal changing at rate $1/T$. $s'_m(t)$ is less noisy and conveys topography information. Decision about sep is high or low can be taken by thresholding $s'_m(t)$ with respect to a threshold τ. Behavior of $s'_m(t)$ corresponding to a few more bit intervals is shown in Fig. 7. It can be seen that even if the cantilever continues to interact with the sample for longer time, $s'_m(t)$ does not lose its steady value substantially. Hence, imaging does not suffer from not detecting a long sequence of 1's or 0's. In Fig. 8, a performance comparison of different bit detection methods is shown. Interaction length l_{int} is gradually changed from 1 nm to 2 nm. Locally most powerful (LMP) technique uses matched filtering on $e_1(t)$ as described in section II. Viterbi(old) uses sequence detection only on e_1 as explained in [5] using memories $m = 1$ and $m_I = 2$. For IM method, a square signal $\sqcap(T)$ is used as matched filter on $s'_m(t)$ and its output corresponding to each bit interval is thresholded to decide 1 or 0. Viterbi(new) uses $s'_m(t)$ for sequence detection with the same memory lengths. It can be seen that IM method provides much better

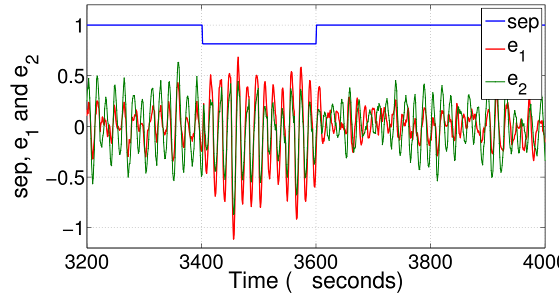

Fig. 4. Innovations from two observers vs separation

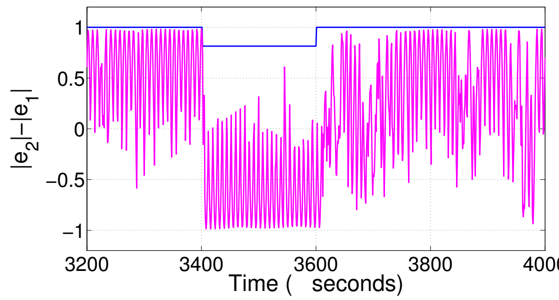

Fig. 5. Innovations mismatch signal $i'_m(t)$ vs separation

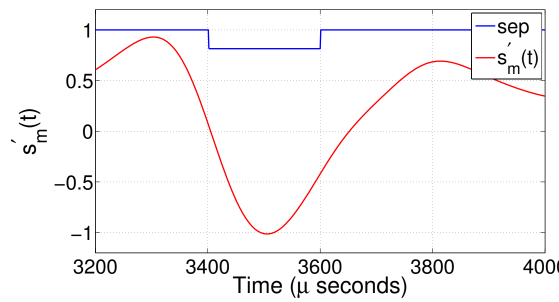

Fig. 6. Low pass filtered $i'_m(t)$ ($s'_m(t)$) vs separation

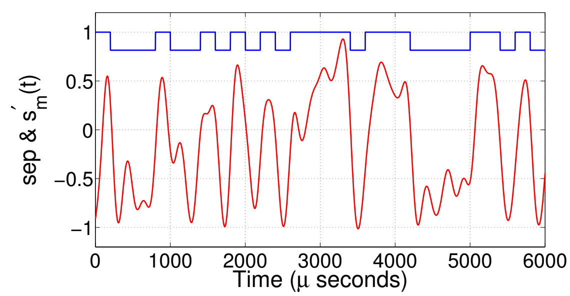

Fig. 7. Low pass filtered $i'_m(t)$ ($s'_m(t)$) vs separation

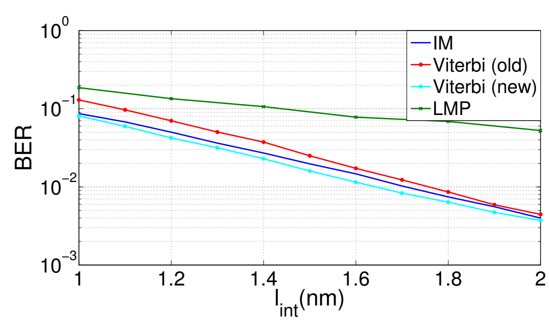

Fig. 8. Comparison of BER performances of different bit detection techniques

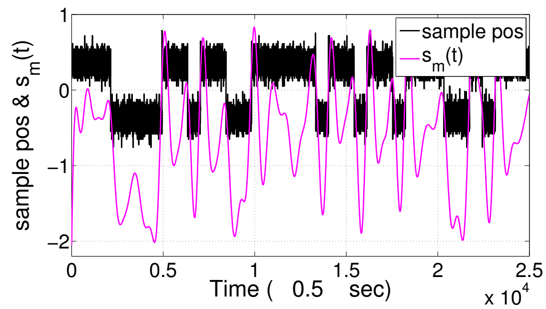

Fig. 9. $s_m(t)$ vs sample position obtained from experimental data

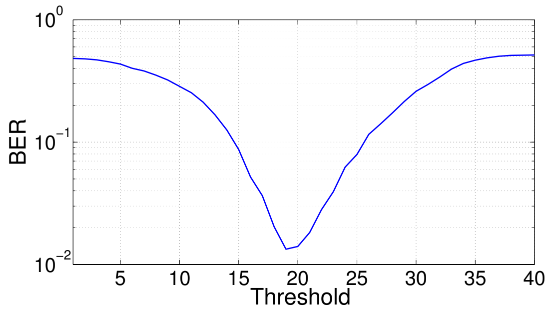

Fig. 10. BER corresponding to the choice of threshold from experimental data

bit error rate (BER) than LMP. Moreover, it works slightly better than Viterbi(old). Viterbi(new) is slightly better than IM. However, signal processing is more complicated. Even though l_{int} is varied from 1 nm to 2 nm, $e_2(t)$ is always obtained from a Kalman observer matched to the 2 nm equivalent cantilever model. Hence, it is not necessary to find equivalent models corresponding to each interaction lengths between 1 nm and 2 nm. Just one Kalman observer matched to a suitable equivalent model is necessary to synthesize the mismatch signal. Signal processing involved in generating $s_m(t)$ is extremely simple (requires a rectifier and a low pass filter). For the simplicity of IM and availability of analog circuits for rectification and low pass filtering, IM can be implemented for real time imaging.

VI. Experimental Evidence

A cantilever with $f_0 = 71.8$ kHz and $Q_0 = 127.3$ is used for experiments. A mica sheet is used as the sample material that can be moved up and down in the z direction using a piezo electric setup. A random bit sequence is generated and an FPGA based circuit is used to actuate the setup to place the mica sheet high or low corresponding to bit 1 or 0 respectively. For bit 0, the centilever oscillates in free air. Interaction occurs when bit is 1. Bit width is $T = 340\,\mu$ sec. One Kalman observer is matched to free space cantilever model and the other is matched to a model with $f_e = 70.8$ kHz (100 Hz less than free space model) and the same quality factor ($Q_e = Q_0$). $s_m(t)$ is generated from the measured cantilever deflection and shown in Fig. 9. In Fig. 9, high sample position indicates interaction length of 1 nm. It can be clearly seen that even small interaction such as 1 nm can be identified from $s_m(t)$. A square matched filter $\sqcap(T)$ is applied on $s_m(t)$ and its output in each bit interval is recorded. After observing the matched filter outputs for 10^5 bits, the maximum and the minimum values of it are chosen (say, L_{max} and L_{min}). The range $[L_{min}, L_{max}]$ is equally divided into 40 points and each point is chosen to be the decision threshold τ and corresponding BER is calculated. BER corresponding to the choice for τ is shown in the Fig. 10. From Fig. 10, it can be seen that the minimum BER that can be achieve is close to 10^{-2}. We remark that the experimental results are for bit by bit detection which has the considerable advantage of ease of implementation (no need for learning and Viterbi based decoding).

VII. Conclusions and Future Work

A new topography imaging method, called the IM technique is developed in this paper. IM works on the dynamic mode of AFM operation and hence suitable for soft material imaging. Older method for imaging and data detection using dynamic mode AFM called the TF-AFM is briefly explained and the associated problems are delineated. Simulation results show that IM method which employs two observers beats previous bit by bit detection scheme and even the sequence detection method on a single innovation signal on performance. A significant advantage of the IM method is that it is considerably simpler to implement than any sequence based detection schemes such as the one reported in [5]. An experiment is also performed validating the applicability of IM. Even though the advantages of using IM are clear, a detailed theoretical analysis of performance of IM is still pending and forms the future work of this paper.

References

[1] P. Agarwal and M. Salapaka, "Real time estimation of equivalent cantilever parameters in tapping mode atomic force microscopy," *Applied Physics Letters*, vol. 95, no. 8, pp. 083 113–083 113, 2009.

[2] G. Binnig, C. Quate, and C. Gerber, "Atomic force microscope," *Physical review letters*, vol. 56, no. 9, pp. 930–933, 1986.

[3] E. Eleftheriou, T. Antonakopoulos, G. Binnig, G. Cherubini, M. Despont, A. Dholakia, U. Durig, M. Lantz, H. Pozidis, H. Rothuizen, *et al.*, "Millipede-a mems-based scanning-probe data-storage system," *Magnetics, IEEE Transactions on*, vol. 39, no. 2, pp. 938–945, 2003.

[4] A. Kavcic and J. Moura, "The viterbi algorithm and markov noise memory," *Information Theory, IEEE Transactions on*, vol. 46, no. 1, pp. 291–301, 2000.

[5] N. Kumar, P. Agarwal, A. Ramamoorthy, and M. Salapaka, "Maximum-likelihood sequence detector for dynamic mode high density probe storage," *Communications, IEEE Transactions on*, vol. 58, no. 6, pp. 1686–1694, 2010.

[6] H. Mamin, R. Ried, B. Terris, and D. Rugar, "High-density data storage based on the atomic force microscope," *Proceedings of the IEEE*, vol. 87, no. 6, pp. 1014–1027, 1999.

[7] D. Sahoo, P. Agarwal, and M. Salapaka, "Transient force atomic force microscopy: A new nano-interrogation method," in *American Control Conference, 2007. ACC'07.* IEEE, 2007, pp. 2135–2140.

[8] J. Sanders and F. Verhulst, *Averaging methods in nonlinear dynamical systems.* Springer, 1985, vol. 59.

[9] A. Sebastian, A. Gannepalli, and M. Salapaka, "The amplitude phase dynamics and fixed points in tapping-mode atomic force microscopy," in *American Control Conference, 2004. Proceedings of the 2004*, vol. 3. IEEE, 2004, pp. 2499–2504.

[10] A. Sebastian, M. Salapaka, D. Chen, and J. Cleveland, "Harmonic and power balance tools for tapping-mode atomic force microscope," *Journal of Applied Physics*, vol. 89, p. 6473, 2001.

[11] P. Vettiger, G. Cross, M. Despont, U. Drechsler, U. Durig, B. Gotsmann, W. Haberle, M. Lantz, H. Rothuizen, R. Stutz, *et al.*, "The millipede-nanotechnology entering data storage," *Nanotechnology, IEEE Transactions on*, vol. 1, no. 1, pp. 39–55, 2002.