Model detection with application to probe based data storage ☆

Sayan Ghosal [a], Govind Saraswat [b], Murti Salapaka [a]

[a] *Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN, USA*
[b] *Oracle, Santa Clara, CA, USA*

ARTICLE INFO

ABSTRACT

Detection of a model from a set of models that best describes the behavior of a system is of primary importance in many applications. In this article two discriminating signals are derived from measurements for a plant that switches between two model behaviors, where the transfer functions from inputs to the two signals are identical when one model is effective while they are negative of each other when the other model is effective. The developed detection algorithm called the innovation squared mismatch is utilized for the read operation of probe based data storage. The innovations squared mismatch method offers better detection performance with significantly less computational complexity compared to prevalently used maximum a posteriori probability based methods in current data storage systems. The article further proposes employing maximum likelihood sequence detection based methods where the plant behavior can switch from one model to another at high rates and the transients from a previous behavior affect the current behavior causing inter-symbol interference. Exhaustive simulation and experimental results corroborate the efficiency of the proposed methods.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

In many applications, the system of concern can be modeled by a set of linear models where it switches from one model to another (Sun & Ge, 2005). In such systems it is often desired to identify the model that best describes the current behavior of the plant. Observer banks are widely employed to decipher plant models (Feng, JianQiang, & Keqiang, 2006; Gao, Li, Kum, & Zhang, 2015; Gao, Li, & Li, 2014; Han & Narendra, 2012; Huang, Wang, & Wang, 2015; Morse, 1996, 1997; Narendra & Balakrishnan, 1994, 1997) for closed loop control of switched systems. In principle, the observer which generates the minimum tracking error is considered closest to the active plant model (Morse, 1993). The switching index derived from tracking errors decides the closest plant. Various switching indices are proposed which include the index defined in Morse (1996, 1997) that integrates the square of tracking error with a tunable forgetting factor. In Narendra and Balakrishnan (1994, 1997), a switching index is developed

which uses a weighted combination of both the instantaneous and integrated value of square of tracking errors that improves the transient response of the closed-loop for uncertain plant. The switching index proposed in Narendra and Balakrishnan (1997) is further modified in Feng et al. (2006), Han and Narendra (2012), Gao et al. (2015, 2014) which simplifies the computation of switching index by removing past observation data (Datta, 2012), reduces number of candidate controllers (Gao et al., 2015) and allows derivation of analytical conditions for stability of a plant with large parametric uncertainties (Gao et al., 2015, 2014). One of the main emphases of these studies include analyzing and improving tracking performance of closed-loop switching systems while ensuring stability. This article, however, focuses on fast *detection* of a plant that switches between two *known* models. Such an assumption is particularly relevant for data storage application where plant characteristics when interacting with the bit 1 is different than interacting with bit 0. The maximum a posteriori probability (MAP) based methods are prevalent in bit patterned media recording and magnetic data storage (Cai et al., 2010; Carosino et al., 2013; Chen & Srinivasa, 2013; Khatami & Vasic, 2013; Krishnan et al., 2009) which are relatively matured technologies compared to probe based storage systems. Motivated by these examples, a recursive MAP based algorithm (see chapter 10, Anderson & Moore, 2012) is developed in this article and applied to online plant identification for probe based storage systems. In this article, a new method called the

innovation squared mismatch (ISM) is developed (introduced in Ghosal, Saraswat, Ramamoorthy, & Salapaka, 2013, for the first time). It is demonstrated that the newly synthesized signal offers considerable advantage over MAP while being considerably simpler to implement. For high bandwidth read operation, the issue of transients caused by fast transition between bits that leads to inter symbol interference has to be overcome. In this article we develop dynamic programming based sequence detection schemes. Remarkably we demonstrate that bit by bit ISM based detection performs better than sequence based detection with significantly higher complexity. This performance gain is achieved by the short time scales of the transients in ISM signals. We further demonstrate that sequence detection in conjunction with ISM can lead to further significant advances in read bandwidth. An important focus of this article is on enabling ultra-high density storage technology. The need for such technology is evident in the connected world of data and big data age.

With data storage density for memory devices gradually increasing, the data density for the hard disk drive technology will eventually be restricted by thermal limits on magnetism (super paramagnetic limit) (Campardo, Tiziani, & Iaculo, 2011; Richter, 2009). The minimum bit size for solid state drive memories is also limited by the resolution of the lithography processes (Lai, 2008) where new technology for higher resolution lithography is hugely expensive. The probe based surface manipulation and reading methods offer alternative route which can pack binary data at the nanometer or sub nanometer scale. Since the invention of atomic force microscope (Binnig, Quate, & Gerber, 1986), probe based investigation techniques have been employed in diverse applications (Bray, Cohen, & Lightbody, 2013) including probe based data storage systems (Campardo et al., 2011; Koelmans, Engelen, & Abelmann, 2015; Wright, Aziz, Shah, & Wang, 2011). Since the pioneering demonstrations of topographic probe based data storage (Cannara, Gotsmann, Knoll, & Dürig, 2008; Knoll et al., 2006; Mamin, Ried, Terris, & Rugar, 1999; Vettiger et al., 2002; Wiesmann et al., 2009) where an indentation of sample surface encodes the bit 1 and absence of indentation represents the bit 0, different physical parameters of the sample have been utilized to interpret binary data. Examples include phase-change (Bhaskaran, Sebastian, Drechsler, & Despont, 2009; Hamann, O'Boyle, Martin, Rooks, & Wickramasinghe, 2006), magnetic properties (Kappenberger et al., 2009; Kryder et al., 2008) or ferro-electric characteristics (Cho, Hashimoto, Odagawa, Tanaka, & Hiranaga, 2006; Tanaka & Cho, 2010) based memories. Even though extremely high data density in the order of 3–4 Tb/in^2 is reported in Wiesmann et al. (2009), Hamann et al. (2006), Tanaka and Cho (2010), a big challenge is the mechanical wear and tear of the probe as well as of the sample which severely cripples the lifetime well below the life span of present day data storage devices (Lantz, Wiesmann, & Gotsmann, 2009; Mamin et al., 1999). Thus attempts have been made to develop soft samples (Gotsmann et al., 2010) as well as to fabricate wear resistant cantilever tips (Bhaskaran et al., 2010; Lantz et al., 2012).

Most of the existing probe based data storage solutions utilize *contact mode* operation. Thus the tip of the flexure probe (see Fig. 1) is in contact with the storage media, causing wear and tear of the media as well as of the probe. The dynamic mode operation is an alternative to the contact mode method which drastically eliminates mechanical wear and tear during the reading process. Here the flexure probe is forced sinusoidally and the probe contact the media only intermittently. The information on the media encoded topographically modulates the flexure oscillations which is used to decipher information on the media. Due to slow read-out and complex data retrieval mechanism, the traditional dynamic mode method faces significant challenges for probe based data storage (Lantz et al., 2009). This article synthesizes new
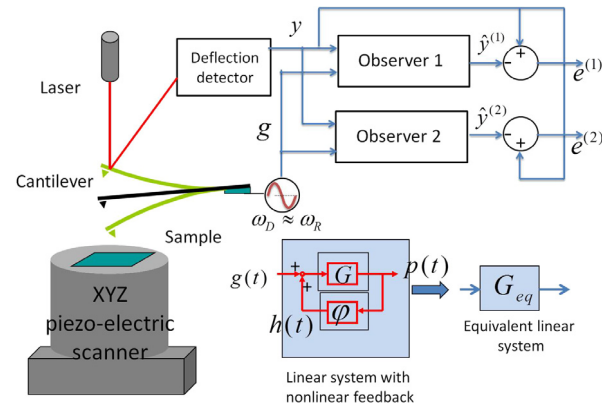

Fig. 1. The atomic force microscope (AFM). The sample sits on a positioning device which can move in all three spatial directions with respect to a flexible cantilever. The cantilever deflection which depends on inter atomic forces between the tip and the sample is recorded using a laser, photo-detector setup. In the dynamic mode, the cantilever is excited sinusoidally where its orbit depends on the topography of sample. Traditionally the oscillation amplitude A_{def} and phase ϕ_{def} of the cantilever deflection are utilized for dynamic mode imaging. In the proposed method, BCEWRLS algorithm is first utilized to obtain equivalent linear models when the cantilever oscillates in free air (model 1) and when it interacts with the sample (model 2). Observers generate tracking errors $e^{(1)}$ and $e^{(2)}$ that are processed by MAP and ISM techniques for real time interpretation of topography. MLSD, ISM-MLSD further improve the detection accuracy in presence of inter symbol interference.

methods for circumventing the limited speed of bit detection with dynamic mode operation. It is demonstrated that the oscillating cantilever while interacting with raised topography assumes one model behavior (model 1) which is different than model behavior assumed by the probe when traversing over lowered topography (model 0). Thus for topographic data storage where raised and lowered topography encode binary bits, the read task becomes detection of which model is active. These models can be identified prior to the reading operation with a bias compensated recursive least square (BCEWRLS) algorithm (Agarwal & Salapaka, 2009; Saraswat, Agarwal, Haugstad, & Salapaka, 2013) which remain valid for subsequent data reading. Thus it is a good assumption that the plant switches between two *known* models which need to be detected.

The conventional dynamic mode AFM (see Fig. 1) utilizes cantilever oscillation amplitude and phase to interpret sample topography. It is demonstrated that application of MAP and ISM offers significant detection bandwidth gains over conventional dynamic mode signals. ISM offers similar detection speed, better detection accuracy and substantial computational gain over MAP. Since the ISM and MAP algorithms offer improved but *finite* detection bandwidth, the issue of inter symbol interference (ISI) is addressed for fast bit detection using the maximum likelihood of an entire *sequence* of bits (Forney, 1972; Moon & Carley, 2012). In this article, sequence detection based on the ISM signal (the ISM-MLSD method) is developed where the time spent by the probe on single bit (the *dwell-time*) is known. ISM-MLSD consistently outperforms symbol-by-symbol ISM and MAP, however, at the expense of complex signal processing. The analytical and close to reality simulation evaluations are corroborated via exhaustive experimental results.

The main contributions of this paper are summarized as: (i) A new method for real time detection of plant behavior that switches between two models. (ii) A new maximum likelihood sequence detection method with superior bit error rates. (iii) A new method for real time detection as well as off-line sequence detection for probe based data storage with experimental validation. This article is organized as follows. Section 2 develops ISM. Modified symbol-by-symbol detection algorithms and ISM-MLSD are proposed

in Section 3. Section 4 provides a brief background for AFM and presents the simulation and experimental results. Section 5 concludes this work.

2. Real time plant detection with innovations squared mismatch (ISM) and maximum a posteriori probability (MAP) methods

Consider a continuous time plant which is linear and time varying described by

$$\dot{x}(t) = A(t)x(t) + B(t)g(t),$$
$$y(t) = C(t)x(t) + D(t)g(t) + \zeta(t), \tag{1}$$

where $\zeta(t)$ denotes measurement noise assumed to be a white and zero-mean Gaussian process with covariance $E[\zeta(t)\zeta^T(t_0)] = R_n\delta(t - t_0)$. It is assumed that $(A(t), B(t), C(t), D(t)) \in \{(A_i + \Delta A_i, B_i + \Delta B_i, C_i + \Delta C_i, D_i + \Delta D_i)\}$ with $i \in \{1, 2\}$. Thus the system can switch between two known models (A_i, B_i, C_i, D_i) with $(\Delta A_i, \Delta B_i, \Delta C_i, \Delta D_i)$ denoting uncertainties. Let $\hat{x}^{(i)}$ denote state of Kalman observer (Kailath, Sayed, & Hassibi, 2000) matched to model i, let $\hat{y}^{(i)}$ denote corresponding estimated output which are governed by

$$\dot{\hat{x}}^{(i)}(t) = A_i\hat{x}^{(i)}(t) + B_ig(t) + L_i(y(t) - \hat{y}^{(i)}(t)),$$
$$\hat{y}^{(i)}(t) = C_i\hat{x}^{(i)}(t) + D_ig(t). \tag{2}$$

Let $\tilde{x}^{(i)} = x - \hat{x}^{(i)}$ and $e^{(i)} = y - \hat{y}^{(i)}$ where $e^{(i)}$ is the output error (also referred in this article as innovation).

It follows from (1) and (2) that the error dynamics is described by

$$\begin{bmatrix} \dot{\tilde{x}}^{(1)}(t) \\ \dot{\tilde{x}}^{(2)}(t) \end{bmatrix} = \begin{bmatrix} (A(t) - A_1) - L_1(C(t) - C_1) \\ (A(t) - A_2) - L_2(C(t) - C_2) \end{bmatrix} x(t)$$
$$+ \begin{bmatrix} A_1 - L_1C_1 & 0 \\ 0 & A_2 - L_2C_2 \end{bmatrix} \begin{bmatrix} \tilde{x}^{(1)}(t) \\ \tilde{x}^{(2)}(t) \end{bmatrix}$$
$$+ \begin{bmatrix} (B(t) - B_1) + L_1(D_1 - D(t)) \\ (B(t) - B_2) + L_2(D_2 - D(t)) \end{bmatrix} g(t) + \begin{bmatrix} -L_1 \\ -L_2 \end{bmatrix} \zeta(t), \tag{3}$$

and

$$\begin{bmatrix} e^{(1)}(t) \\ e^{(2)}(t) \end{bmatrix} = \begin{bmatrix} C(t) - C_1 \\ C(t) - C_2 \end{bmatrix} x(t) + \begin{bmatrix} C_1 & 0 \\ 0 & C_2 \end{bmatrix} \begin{bmatrix} \tilde{x}^{(1)}(t) \\ \tilde{x}^{(2)}(t) \end{bmatrix}$$
$$+ \begin{bmatrix} D(t) - D_1 \\ D(t) - D_2 \end{bmatrix} g(t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \zeta(t). \tag{4}$$

Let $z_1(t) = (\tilde{x}^{(1)}(t) - \tilde{x}^{(2)}(t))$ and $z_2(t) = (\tilde{x}^{(1)}(t) + \tilde{x}^{(2)}(t))$. The output vectors e_{dif} and e_{sum} are defined by $e_{dif}(t) := e^{(1)}(t) - e^{(2)}(t)$ and $e_{sum}(t) := e^{(1)}(t) + e^{(2)}(t)$. Thus when the plant follows model i, the dynamics in the new coordinates are given by

$$\begin{bmatrix} \dot{z}_1(t) \\ \dot{z}_2(t) \end{bmatrix} = \begin{bmatrix} M_1 & M_2 \\ M_2 & M_1 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix}$$
$$+ \begin{bmatrix} A_{mi} + (L_2 - L_1)\Delta C_i \\ (-1)^i A_{mi} + 2\Delta A_i - (L_1 + L_2)\Delta C_i \end{bmatrix} x(t)$$
$$+ \begin{bmatrix} B_{mi} + (L_2 - L_1)\Delta D_i \\ (-1)^i B_{mi} - (L_1 + L_2)\Delta D_i \end{bmatrix} g(t)$$
$$+ \begin{bmatrix} L_2 - L_1 \\ -(L_1 + L_2) \end{bmatrix} \zeta(t), \tag{5}$$

and

$$\begin{bmatrix} e_{dif}(t) \\ e_{sum}(t) \end{bmatrix} = \begin{bmatrix} N_1 & N_2 \\ N_2 & N_1 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix}$$
$$+ \begin{bmatrix} C_{mi} \\ (-1)^i C_{mi} + 2\Delta C_i \end{bmatrix} x(t)$$
$$+ \begin{bmatrix} D_{mi} \\ (-1)^i D_{mi} + 2\Delta D_i \end{bmatrix} g(t) + \begin{bmatrix} 0 \\ 2 \end{bmatrix} \zeta(t), \tag{6}$$

with $M_1 := \frac{1}{2}[(A_1 + A_2) - (L_1C_1 + L_2C_2)]$, $M_2 := \frac{1}{2}[(A_1 - A_2) - (L_1C_1 - L_2C_2)]$, $N_1 := \frac{1}{2}(C_1 + C_2)$, $N_2 := \frac{1}{2}(C_1 - C_2)$, $A_{m1} := ((A_2 - A_1) + L_2(C_1 - C_2))$, $A_{m2} := ((A_2 - A_1) + L_1(C_1 - C_2))$, $B_{m1} := (B_2 - B_1 + L_2(D_1 - D_2))$, $B_{m2} := (B_2 - B_1 + L_1(D_1 - D_2))$, $C_{m1} = C_{m2} := (C_2 - C_1)$ and $D_{m1} = D_{m2} := (D_2 - D_1)$. The following theorem presents the main result of this section.

Theorem 1. Let $\mathcal{G}_i(s)$ denote the transfer function matrix from $[x^T(s), g^T(s)]^T$ to $[e_{dif}^T(s), e_{sum}^T(s)]^T$ when the plant is described by model i ($i \in \{1, 2\}$). Then

$$\mathcal{G}_i(s) = \begin{bmatrix} \mathcal{G}_{mi}^{(1)} + \Delta\mathcal{G}_{mi}^{(1)} & \mathcal{G}_{mi}^{(2)} + \Delta\mathcal{G}_{mi}^{(3)} \\ (-1)^i\mathcal{G}_{mi}^{(1)} + \Delta\mathcal{G}_{mi}^{(2)} & (-1)^i\mathcal{G}_{mi}^{(2)} + \Delta\mathcal{G}_{mi}^{(4)} \end{bmatrix}, \tag{7}$$

where

$$\mathcal{G}_{mi}^{(1)} = (N_1\mathcal{S}_1 + N_2\mathcal{S}_2 + N_1\mathcal{S}_2 + N_2\mathcal{S}_1)A_{mi} + C_{mi},$$
$$\mathcal{G}_{mi}^{(2)} = (N_1\mathcal{S}_1 + N_2\mathcal{S}_2 + N_1\mathcal{S}_2 + N_2\mathcal{S}_1)B_{mi} + D_{mi}, \tag{8}$$

with

$$\mathcal{S}_1 = ((sI - M_1)M_2^{-1}(sI - M_1) - M_2)^{-1}(sI - M_1)M_2^{-1},$$
$$\mathcal{S}_2 = ((sI - M_1)M_2^{-1}(sI - M_1) - M_2)^{-1}. \tag{9}$$

The perturbation matrices are defined by $\Delta\mathcal{G}_{mi}^{(1)} = (N_1\mathcal{S}_1 + N_2\mathcal{S}_2)(L_2 - L_1)\Delta C_i + (N_1\mathcal{S}_2 + N_2\mathcal{S}_1)(2\Delta A_i - (L_1 + L_2)\Delta C_i)$, $\Delta\mathcal{G}_{mi}^{(2)} = (N_1\mathcal{S}_2 + N_2\mathcal{S}_1)(L_2 - L_1)\Delta C_i + (N_1\mathcal{S}_1 + N_2\mathcal{S}_2)(2\Delta A_i - (L_1 + L_2)\Delta C_i) + 2\Delta C_i$, $\Delta\mathcal{G}_{mi}^{(3)} = (N_1\mathcal{S}_1 + N_2\mathcal{S}_2)(L_2 - L_1)\Delta D_i - (N_1\mathcal{S}_2 + N_2\mathcal{S}_1)(L_1 + L_2)\Delta D_i$, $\Delta\mathcal{G}_{mi}^{(4)} = (N_1\mathcal{S}_2 + N_2\mathcal{S}_1)(L_2 - L_1)\Delta D_i - (N_1\mathcal{S}_1 + N_2\mathcal{S}_2)(L_1 + L_2)\Delta D_i + 2\Delta D_i$.

Proof. The proof is provided in the Appendix (see Appendix A.1). □

It can also be shown that the transfer function matrix from the measurement noise $\zeta(s)$ to $[e_{dif}^T(s) \ e_{sum}^T(s)]^T$ is given by $[\mathcal{G}_{(dif,\zeta)}^T \ \mathcal{G}_{(sum,\zeta)}^T]^T$ where $\mathcal{G}_{(dif,\zeta)} = (N_1\mathcal{S}_1 + N_2\mathcal{S}_2)(L_2 - L_1) + (N_1\mathcal{S}_2 + N_2\mathcal{S}_1)(L_1 + L_2)$ and $\mathcal{G}_{(sum,\zeta)} = (N_1\mathcal{S}_2 + N_2\mathcal{S}_1)(L_2 - L_1) - (N_1\mathcal{S}_1 + N_2\mathcal{S}_2)(L_1 + L_2) + 2$. Thus in model i, it follows that

$$e_{dif}(s) = e_{mi}(s) + \Delta e_{(dif,i)}(s) + \zeta_{dif}(s), \text{ and}$$
$$e_{sum}(s) = (-1)^i e_{mi}(s) + \Delta e_{(sum,i)}(s) + \zeta_{sum}(s), \tag{10}$$

where $e_{mi}(s) := \mathcal{G}_{mi}^{(1)}x(s) + \mathcal{G}_{mi}^{(2)}g(s)$ represents the *signal* component. $\Delta e_{(dif,i)}(s) := \Delta\mathcal{G}_{mi}^{(1)}x(s) + \Delta\mathcal{G}_{mi}^{(3)}g(s)$, $\Delta e_{(sum,i)}(s) := \Delta\mathcal{G}_{mi}^{(2)}x(s) + \Delta\mathcal{G}_{mi}^{(4)}g(s)$ represent perturbation signals due to deviation of the plant model from the known reference model i. Further, $\zeta_{dif}(s) := \mathcal{G}_{(dif,\zeta)}(s)\zeta(s)$ and $\zeta_{sum}(s) := \mathcal{G}_{(sum,\zeta)}(s)\zeta(s)$ denote the noise components in e_{dif} and e_{sum} respectively. Notice that in (10), $e_{mi}(s)$ is multiplied with $(-1)^i$ in the expression for $e_{sum}(s)$. The next corollary follows immediately.

Corollary 2. For a plant switching between two stable models, under steady state the relative sign between the signal components present in $e_{dif}(t)$, $e_{sum}(t)$ are opposite when model 1 is active and the same when model 2 is active.

In (5) and (6), the decay rate of transients when the plant behavior switches depend on the eigenvalues of the matrix $\begin{bmatrix} M_1 & M_2 \\ M_2 & M_1 \end{bmatrix}$. The switching speed of the plant needs to be slower compared to the decay times of the transients in $e_{dif}(t)$ and $e_{sum}(t)$. This restricts the maximum switching speed of the plant that can be detected by the proposed method. We demonstrate later that the assumption holds quite well for practical reading speeds of probe based data storage. A particularly important case is when the external forcing is sinusoidal and the output is scalar.

Corollary 3. *Let* $g(t) = A\sin(\omega t)$. *For plant following model 1, in steady state, let* $e_{dif}(t) = \mathcal{A}_{m1}(t)\sin(\omega t + \phi_1(t)) + \zeta_{dif}(t)$ *where* $\mathcal{A}_{m1}(t)$ *and* $\phi_1(t)$ *denote the amplitude and phase of the signal component in* $e_{dif}(t)$ *respectively. Then* $e_{sum} = -\mathcal{A}_{m1}(t)\sin(\omega t + \phi_1(t)) + \zeta_{sum}(t)$. *For plant following model 2, in steady state, let* $e_{dif}(t) = \mathcal{A}_{m2}(t)\sin(\omega t + \phi_2(t)) + \zeta_{dif}(t)$. *Then* $e_{sum}(t) = \mathcal{A}_{m2}(t)\sin(\omega t + \phi_2(t)) + \zeta_{sum}(t)$. *Thus the innovation squared mismatch is given by*

$$(e^{(1)}(t))^2 - (e^{(2)}(t))^2 := e_{dif}(t)e_{sum}(t)$$

$$= \begin{cases} -\dfrac{\mathcal{A}_{m1}^2(t)}{2}(1 - \cos(2\omega t + 2\phi_1(t))) + \zeta_{ISM1}, & model\ 1 \\[3mm] \dfrac{\mathcal{A}_{m2}^2(t)}{2}(1 - \cos(2\omega t + 2\phi_2(t))) + \zeta_{ISM2}, & model\ 2 \end{cases} \tag{11}$$

where $\zeta_{ISMi}(t) = \zeta_{dif}(t)\zeta_{sum}(t) + A_{mi}(t)\sin(\omega t + \phi_i(t))(\zeta_{sum}(t) + (-1)^i\zeta_{dif}(t))$ *denotes the noise component present in the ISM test signal for model i.*

Remark. In steady state, the ISM signal $(e^{(1)}(t))^2 - (e^{(2)}(t))^2$ when processed by a low pass filter will result in the following:

$$\left((e^{(1)}(t))^2 - (e^{(2)}(t))^2\right)_{lpf} = \begin{cases} -\dfrac{\mathcal{A}_{m1}^2(t)}{2} + \zeta'_{ISM1}, & model\ 1 \\[3mm] +\dfrac{\mathcal{A}_{m2}^2(t)}{2} + \zeta'_{ISM2}, & model\ 2 \end{cases} \tag{12}$$

where ζ'_{ISMi} are low pass filtered components of ISM noise ζ_{ISMi}. The sign inversion in (12) can be used in real time to interpret the model that is being followed by the plant. It is also evident that the phase of the signal components in e_{dif} and e_{sum} differ by zero in model 2, whereas in model 1 the phase difference is π radians. The phase difference between e_{dif} and e_{sum} is the phase signal $\phi_m(t)$. We refer to $\left((e^{(1)}(t))^2 - (e^{(2)}(t))^2\right)_{lpf}$ as ISM_{lpf}.

The ISM methods described above apply for sinusoidal excitation signals. A simple correlation based implementation for determining which model is effective for a general input (including sinusoidal inputs) is to track the relative sign of $e_{dif}(t)$ and $e_{sum}(t)$ over a time window. Let the signals $e_{dif}[m]$ and $e_{sum}[m]$ denote corresponding digitally sampled signals where m is the discrete time index with sampling interval T_s. We define a *correlation based* test signal $T[m]$ as

$$T[m] = \frac{1}{M}\big(e_{dif}[m]e_{sum}[m] + e_{dif}[m-1]e_{sum}[m-1]$$
$$+ \cdots + e_{dif}[m-M+1]e_{sum}[m-M+1]\big), \tag{13}$$

where M is the window size. The windowing helps in mitigating the effect of noise. The decision policy is model 2 if $T[m] > 0$ else model 1. M can be adjusted according to user's choice for improving the signal to noise ratio in the test signal $T[m]$ where higher value of M typically tends to average out the effect of the noise process $\zeta_{ISM,i}$. However, increasing M also reduces the detection bandwidth of $T[m]$ when plant switching occurs.

2.1. MAP

In this part of the article we briefly provide description of a real time MAP algorithm which is motivated by the prevalently used MAP methods in data storage application. Consider a discrete time plant corresponding to sampling interval of T_s. Let i denote the index for the active model. The discrete time switching plant is given by

$$x[m+1] = A_{(d,i)}x[m] + B_{(d,i)}g[m],$$
$$y[m] = C_{(d,i)}x[m] + D_{(d,i)}g[m] + \zeta[m], \tag{14}$$

with $\zeta \sim N(0, R_{(d,i)})$. The subscript (d, i) indicates active model i for *discrete time* dynamics. A steady state discrete-time Kalman observer matched to model i with gain $L_{(d,i)}$ is governed by

$$\hat{x}^{(i)}[m+1|m] = A_{(d,i)}\hat{x}^{(i)}[m|m-1] + B_{(d,i)}g[m]$$
$$+ L_{(d,i)}\left(y[m] - \hat{y}^{(i)}[m]\right), \tag{15}$$
$$\hat{y}^{(i)}[m] = C_{(d,i)}\hat{x}^{(i)}[m|m-1] + D_{(d,i)}g[m].$$

Let θ_i denote model i and $\mathcal{Y}_m = \{y[0], y[1], \ldots, y[m]\}$. The a posteriori probability that the plant parameter is θ_i given the observation vector $\mathcal{Y}_m$ is denoted by $p(\theta_i|\mathcal{Y}_m)$. The decision θ_{MAP} is chosen from the candidates θ_i so that $p(\theta_i|\mathcal{Y}_m)$ is maximized. It can be shown that (see chapter 10, Anderson & Moore, 2012)

$$p(\theta_i|\mathcal{Y}_m) = \frac{p(y_m|\mathcal{Y}_{m-1}, \theta_i)p(\theta_i|\mathcal{Y}_{m-1})}{\sum_{j=1}^{2} p(y_m|\mathcal{Y}_{m-1}, \theta_j)p(\theta_j|\mathcal{Y}_{m-1})}. \tag{16}$$

From Anderson and Moore (2012), it also follows that $p(y_m|\mathcal{Y}_{m-1}, \theta_i)$ obeys a Gaussian distribution with mean $\hat{y}_i[m] \in \mathbb{R}^q$ and covariance $\Omega_{(m/i)}$. Hence

$$p(y_m|\mathcal{Y}_{m-1}, \theta_i) = \frac{1}{(2\pi)^{q/2}|\Omega_{(m/i)}|^{1/2}}$$
$$\exp\left\{-\frac{1}{2}\tilde{y}_{(m,i)}^T\Omega_{(m/i)}^{-1}\tilde{y}_{(m,i)}\right\}, \tag{17}$$

where $\tilde{y}_{(m,i)} = y - \hat{y}_{(m,i)}$. As a result, $p(\theta_i|\mathcal{Y}_m)$ can be updated from $p(\theta_i|\mathcal{Y}_{m-1})$ using

$$p(\theta_i|\mathcal{Y}_m) = c_m|\Omega_{(m/i)}|^{-\frac{1}{2}}$$
$$\exp\left\{-\frac{1}{2}\tilde{y}_{(m,i)}^T\Omega_{(m/i)}^{-1}\tilde{y}_{(m,i)}\right\} p(\theta_i|\mathcal{Y}_{m-1}), \tag{18}$$

where c_m is a normalization constant set to satisfy $\sum_{i=1}^{2} p(\theta_i|\mathcal{Y}_m) = 1$. The output covariance $\Omega_{(m/i)}$ in (18) and covariance of state estimation error $\Sigma_{(m+1/m,i)} = E\left[\tilde{x}_{(m+1,i)}\tilde{x}_{(m+1,i)}^T\right]$ with $\tilde{x}_{(m+1,i)} = (x[m+1] - \hat{x}^{(i)}[m+1|m])$ can be updated using

$$\Sigma_{(m+1/m,i)} = (A_{(d,i)} - L_{(d,i)}C_{(d,i)})\Sigma_{(m/m-1,i)}$$
$$(A_{(d,i)} - L_{(d,i)}C_{(d,i)})^T + L_{(d,i)}R_{(d,i)}L_{(d,i)}^T, \tag{19}$$
$$\Omega_{(m/i)} = C_{(d,i)}\Sigma_{(m/m-1,i)}C_{(d,i)}^T + R_{(d,i)}. \tag{20}$$

Notice that in (18), it is necessary to evaluate $\Omega_{(m/i)}$ at each stage which is accomplished using Eq. (20). Further, evaluation of r.h.s. of (20) requires availability of the state covariance matrix $\Sigma_{(m/m-1,i)}$ which in turn is determined at each recursive stage using (19). Finally, when the r.h.s. of (18) is calculated, parameter θ_i is decided to be effective if $p(\theta_i|\mathcal{Y}_m) > p(\theta_j|\mathcal{Y}_m)$.

2.2. Computational complexity of ISM and MAP

Assume the following dimensions of the matrices in (15), (19) and (20); $A_{d,i} \in \mathbb{R}^{m_A,m_A}$, $B_{d,i} \in \mathbb{R}^{m_A,n_B}$, $C_{d,i} \in \mathbb{R}^{m_C,m_A}$, $D_{d,i} \in \mathbb{R}^{m_C,n_B}$, $L_{d,i} \in \mathbb{R}^{m_A,m_C}$, $\Sigma_{m+1/m,i} \in \mathbb{R}^{m_A,m_A}$, $Q_{d,i} \in \mathbb{R}^{n_B,n_B}$, $R_{d,i} \in \mathbb{R}^{m_C,m_C}$, $\Omega_{m/i} \in \mathbb{R}^{m_C,m_C}$. Define a floating point multiplication, division, addition or subtraction operation as a flop. It is derived in Appendix A.2 that total number of flop counts required per update for (18)–(20) is $2(N_\Sigma + N_\Omega + N_p)$ where

$$
N_\Sigma =
\begin{cases}
\left(\frac{3}{2}m_A^4 + m_A^2 \left(\frac{3}{2}(m_C^2 + n_B^2) + 2m_C + 1 \right) \right. \\
\left. + m_A \left(\frac{3}{2}(m_C + n_B) - \frac{1}{2} \right) \right), \text{ for first update} \\
\frac{3}{2} \left(m_A^4 + m_A^2 \right), \text{ for subsequent updates}
\end{cases}
\tag{21}
$$

with $N_\Omega = \frac{3}{2}(m_C^2 m_A^2 + m_C m_A)$ and $N_p = m_C^3 + \frac{5}{2}m_C^2 + \frac{7}{2}m_C + 4$. The flop counts required for signal correlation based implementation of ISM is determined now. The test signal $T[m]$ defined in (13) required for ISM can be recursively updated using

$$
T[m+1] = T[m] + \frac{1}{M} \left(e_{dif}[m+1]^T e_{sum}[m+1] \right. \\
\left. - e_{dif}[m-M+1]^T e_{sum}[m-M+1] \right).
\tag{22}
$$

Here, $e_{dif}[m], e_{sum}[m] \in \mathbb{R}^{m_C}$ can be determined from $e^{(1)}$ and $e^{(2)}$ using m_C subtractions and m_C additions, thus $2m_C$ flops. Multiplication $e_{dif}[m]^T e_{sum}[m]$ needs m_C multiplications and $(m_C - 1)$ additions thus $(2m_C - 1)$ flops. Division by M, addition and subtraction operations in the r.h.s. of (22) require 1 flop each. Hence, update of $T[m+1]$ from $T[m]$ needs $2(2m_C - 1) + 2m_C + 3 = 6m_C + 1$ flops. The computational complexity for MAP depends on m_C and m_A by $O(m_C^3)$ and $O(m_A^4)$ respectively. Whereas, computational complexity of ISM is independent of m_A and depends on m_C by only $O(m_C)$. Thus the immense computational benefits achieved by ISM over MAP are evident.

3. Plant detection with known dwell interval; sequence detection

In many applications including the probe based data storage system each symbol is encoded in the media for an equivalent duration of T seconds on the read operation. The bit detection performance can be improved by averaging the measured interaction over this T seconds to reduce the effect of noise. Each of the schemes ISM and MAP has a *finite response time*; thus $T[m]$ in (13) and $p(\theta_i|\mathcal{Y}_m), i \in 1, 2$ from (18) take time to change to new steady values when the plant changes from model 1 to 2 and vice versa. The effect of contributions of past interactions on the present interaction causes inter symbol interference (ISI). ISI is more severe when the plant switches fast (dwell interval T is small). A significant gain over symbol-by-symbol schemes can be achieved by deciphering the *sequence* of plant models from the observation data. A widely accepted sequence detection technique by the communication systems and data storage community is the maximum likelihood sequence detection (see Forney, 1972) method which is based on dynamic programming. Such dynamic programming based approach is proven to be effective in different applications; one such interesting application is the detection of molecular steps (Aggarwal, Materassi, Davison, Hays, & Salapaka, 2012). As demonstrated here, maximum likelihood sequence detection (MLSD) can be applied in the context of binary plant model detection as well.

In this article MLSD is applied to process sequence of samples from the test signal $s_m(t)$ generated by two ways; firstly $s_m(t)$ is generated by suitably low pass filtered $(e^{(1)})^2$ which is termed
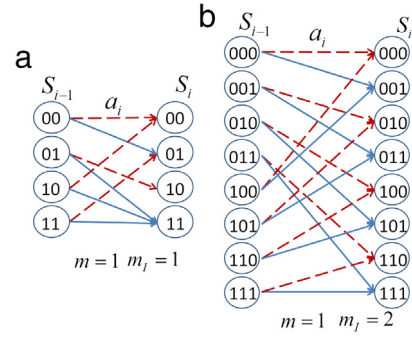

Fig. 2. (a) State transitions for $m = 1$, $m_I = 1$. At each stage i, S_i can assume four possible states. (b) Eight possible states S_i for $m = 1$, $m_I = 2$. Transition from stage S_{i-1} to S_i occurs when a new bit $a_i \in \{0, 1\}$ is input. State transitions for $a_i = 0$ ($a_i = 1$) are shown in broken (solid) lines.

as regular MLSD and secondly $s_m(t)$ is generated using low pass filtered $(e^{(1)})^2 - (e^{(2)})^2$ signal which is termed as ISM-MLSD. Let N be the total number of dwell intervals over which the sequence detection operation is performed. Let $\bar{a} = [a_0, a_1, \ldots, a_{N-1}]^T$ denote the bit sequence such that $a_k = 0$ when model 1 is active ($a_k = 1$ for model 2) during the time interval $t \in (kT, (k+1)T]$ with $k = 0, 1, \ldots, N - 1$. Let q be the number of samples of the test signal gathered in each dwell interval T. Since there are N such dwell intervals, the total number of samples acquired is Nq. The entire sampled sequence of $s_m(t)$ for MLSD detection is denoted by the vector $\bar{z} = [z_0, z_1, \ldots, z_{(Nq-1)}]^T$. Let $f(\bar{z}|\bar{a})$ be the conditional probability density function (pdf) of the output vector $\bar{z}$ given a source bit sequence $\bar{a}$. The maximum likelihood bit sequence is given by

$$
\hat{\bar{a}} = \arg\max_{\bar{a} \in \{0,1\}^N} \prod_{i=0}^{N-1} f(\bar{z}|\bar{a}, \bar{z}_0^{i-1}).
\tag{23}
$$

$\bar{z}_j^i$ is defined as the set of sampled data from the beginning of jth dwell interval till the end of $(i-1)$th dwell interval. Thus $\bar{z}_j^i = [z_{jq}, z_{jq+1}, z_{jq+2}, \ldots, z_{iq-1}]^T$. Further, define $\bar{z}_i$ as the set of samples acquired during ith dwell interval which is $[z_{iq}, z_{iq+1}, \ldots, z_{(i+1)q-1}]^T$. Define the ith state $S_i = [a_{i-m-m_I+1}, a_{i-m-m_I+2}, \ldots, a_i]^T$ where m is the length of system memory and m_I is length of the inter symbol interference memory (Kumar, Agarwal, Ramamoorthy, & Salapaka, 2010). Hence, the identification of maximum likelihood sequence of plant models can be recast as identifying the maximum likelihood sequence of states $\hat{S}_i$. Increasing the total memory $m + m_I$ improves the accuracy. However, the number of possible states S_i at each stage i increases by $2^{(m+m_I)}$ (see Fig. 2) which worsens the computational complexity of the algorithm. For practical purposes increasing the memory beyond a certain value does not improve the detection performance much. The conditional pdf $f(\bar{z}_i^i|\bar{a}, \bar{z}_0^{i-1})$ can be simplified by removing the history of the source bits $\bar{a}$ and observed outputs $\bar{z}$ beyond the assumed memories. In many data retrieving applications 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 $\mathcal{C}(S_i, S_{i-1})$. Even though the actual pdf $f(\bar{z}_{i-m_I}^i|S_i, S_{i-1})$ may not follow Gaussian distribution (Kumar et al., 2010), the Gaussian assumption allows us to write the cost function for the MLSD optimization problem in additive form. This additive cost function can be minimized using existing decoding algorithms. For each state transition shown in Fig. 2, $\bar{\mathcal{Y}}$ and $\mathcal{C}$ can be estimated by a statistical learning step before employing the detector for the read operation.

Lemma 4. *Let* $\hat{\bar{S}} = \arg\max\limits_{\bar{S}} f(\bar{z}|\bar{S})$. *Then*

$$
\hat{\bar{S}} = \arg\min_{\bar{S}} \sum_{i=0}^{N-1} \log\left(\frac{\mathcal{C}(S_i, S_{i-1})}{c(S_i, S_{i-1})}\right)
$$

$$
+ (\bar{z}_{i-m_I}^i - \bar{\mathcal{Y}}(S_i, S_{i-1}))^T \mathcal{C}(S_i, S_{i-1})^{-1}
$$

$$
(\bar{z}_{i-m_I}^i - \bar{\mathcal{Y}}(S_i, S_{i-1})) - (\bar{z}_{i-m_I}^{i-1} - \bar{y}(S_i, S_{i-1}))^T
$$

$$
c(S_i, S_{i-1})^{-1}(\bar{z}_{i-m_I}^{i-1} - \bar{y}(S_i, S_{i-1})), \tag{24}
$$

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* $\mathcal{Y}(S_i, S_{i-1})$.

Proof.

$$
\hat{\bar{S}} = \arg\max_{\bar{S}} f(\bar{z}|\bar{S}) = \arg\max_{\bar{S}} \prod_{i=0}^{N-1} f(\bar{z}_i|S_i, S_{i-1}, \bar{z}_{i-m_I}^{i-1})
$$

$$
= \arg\max_{\bar{S}} \prod_{i=0}^{N-1} \frac{f(\bar{z}_{i-m_I}^i|S_i, S_{i-1})}{f(\bar{z}_{i-m_I}^{i-1}|S_i, S_{i-1})}. \tag{25}
$$

Since $f(\bar{z}_{i-m_I}^i|S_i, S_{i-1}) \sim N\left(\bar{\mathcal{Y}}(S_i, S_{i-1}), \mathcal{C}(S_i, S_{i-1})\right)$, (24) follows immediately from (25). □

The cost function shown in the r.h.s. of (24) is additive function of states. Thus the maximum likelihood sequence of plant models can be found by employing the dynamic programming based Viterbi algorithm (Kavcic & Moura, 2000) which minimizes the cost function. The next section justifies the advantages of the proposed methods using close to reality simulation and exhaustive experimental results for dynamic mode probe based data storage application.

4. Application of ISM, ISM-MLSD to probe based imaging; benefits and comparison with MAP and regular MLSD based approach

Probe based data storage has the potential to revolutionize data storage where greater than 5Tb/in^2 densities can be achieved. We first describe the underlying principle. In the *dynamic* mode of operation of probe based data storage, the cantilever flexure is sinusoidally oscillated near its first resonant frequency. The flexure has a sharp tip near its end that intermittently interacts with the sample (San Paulo & Garcia, 2002). The tip–sample interaction occurring in every oscillation cycle alters the cantilever orbit. In conventional dynamic mode method the cantilever oscillation amplitude and phase of the fundamental harmonic (referred in this article as A_{def} and ϕ_{def} respectively) are utilized for imaging purposes. Research has demonstrated that the cantilever–sample system, which is nonlinear, can be mapped to an equivalent linear-time-invariant model where the model parameters depend on the height of the topography (Sebastian, Gannepalli, & Salapaka, 2007). These parameters if identified can be utilized to infer changes in the sample topography. In probe based data storage applications, high and low topography are used to encode different bits (Mamin et al., 1999; Wright et al., 2011). Binary topography detection technique, where the detection of bits is the main task, in the context of the dynamic mode probe based data storage is developed here.

The first mode model of the cantilever is specified by the parameters ω_0 (first resonant frequency) and Q_0 (quality factor). Here the cantilever flexure dynamics is governed by the following differential equation (Sebastian et al., 2007)

$$
\ddot{p} + \frac{\omega_0}{Q_0}\dot{p} + \omega_0^2 p = F(t) = g(t) + h(t), \tag{26}
$$

$$
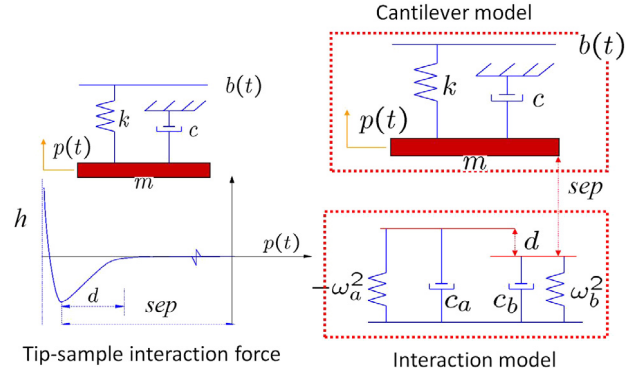h(t) = \varphi(p, \dot{p}), \quad y = p + \zeta,
$$

Fig. 3. Model for tip–sample interaction h used for simulation. $p(t)$ denotes tip trajectory of the cantilever of mass m. A model consisting of two spring–dampers mimics the interaction. The excitation signal $b(t) = g(t)/\omega_0^2$. Cantilever parameters k, m and c are measured experimentally from thermal measurements. Tip–sample interaction parameters ω_b, c_b, ω_a, c_a and d can be estimated using harmonic balance method described in Sebastian et al. (2001).

where p is the cantilever tip deflection and $\dot{p}$ is the tip velocity. The force per unit mass on the cantilever F consists of the excitation g and nonlinear tip–sample interaction force h. Output y is corrupted with the measurement noise ζ. Sinusoidal excitation $g(t)$ is normally chosen at or very near ω_0. The resulting system consists of the plant G specified by (ω_0, Q_0) and the nonlinear feedback term h as shown in Fig. 1. Using averaging theory (Sanders & Verhulst, 1985) techniques it can be shown that a cantilever steadily interacting with the sample in the dynamic mode can be modeled as an equivalent cantilever specified by ω_{eq} and Q_{eq} (Sebastian et al., 2007). The equivalent dynamics is described by

$$
\ddot{p} + \frac{\omega_{eq}}{Q_{eq}}\dot{p} + \omega_{eq}^2 p = g(t). \tag{27}
$$

ω_{eq} and Q_{eq} depend on the slowly varying cantilever oscillation amplitude a (see Sebastian, Gannepalli, & Salapaka, 2004) which can be identified using a bias compensated exponentially weighted recursive least square (BCEWRLS) algorithm reported in Agarwal and Salapaka (2009). The continuous time state-space model with matrices (A_1, B_1, C_1) describes cantilever dynamics when there is no sample interaction (that is $h = 0$ specified in (26)) is considered as model 1. The BCEWRLS method is used to estimate the parameters of the equivalent cantilever dynamics (27) by making the cantilever interact with the sample for a specified set point amplitude resulting in model 2 with state space matrices (A_2, B_2, C_2). It also holds true that the equivalent model differs from the free air model only in the A matrix. Thus, ISM detectors are developed for the models (A_i, B, C), $i \in \{1, 2\}$. Discrete time equivalent systems of (A_i, B, C) are used for MAP based detection as explained before.

4.1. Validation of equivalent model

A piecewise linear model for tip–sample interaction shown in Fig. 3 is utilized for simulations that works well for capturing interactions with hard substrates like mica in air (see Sebastian, Salapaka, Chen, & Cleveland, 2001). A cantilever with $f_0 = 63.15$ kHz, $Q_0 = 227.85$ is chosen for simulation. As explained before, the amplitude of sinusoidal cantilever oscillations is referred to as the *conventional amplitude* signal A_{def} whereas the phase of cantilever oscillation with respect to the cantilever excitation is referred to as the *conventional phase* signal ϕ_{def}. The cantilever with free air oscillation amplitude of 24 nm is gradually moved towards (approach) and away from the sample (retract) to increase the interaction length l_{int} followed by reducing it. l_{int} in this article is
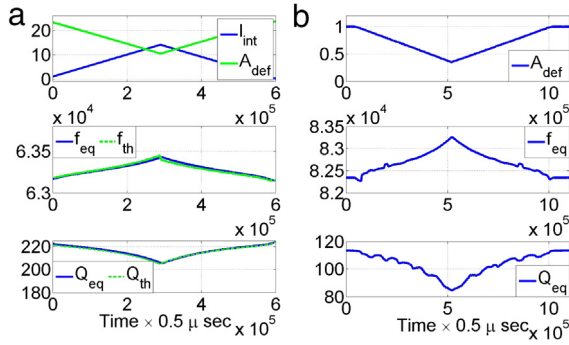
Fig. 4. (a) (Simulation result) Cantilever approach indicated by ascending l_{int} (nm) followed by descending l_{int} (cantilever retract) with cantilever oscillation amplitude A_{def} (nm). BCEWRLS estimates f_{eq} (Hz) and Q_{eq}. Using averaging theory equivalent parameters f_{th} and Q_{th} are calculated. (b) (Experimental data) f_{eq} (Hz) and Q_{eq} obtained by BCEWRLS method during an approach–retract cycle of an oscillating cantilever from mica surface. A_{def} is normalized with respect to free air oscillation amplitude.
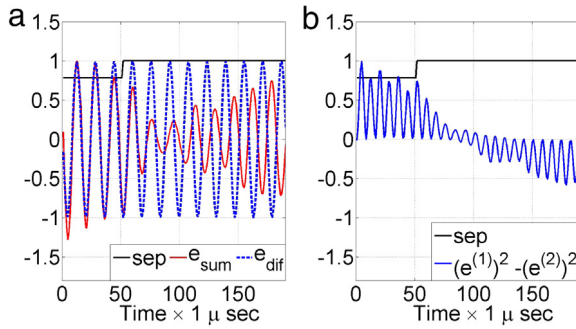

Fig. 5. (Simulation result) (a) e_{dif} and e_{sum} oscillate in phase (out of phase) when cantilever interacts with the sample (is in free air). (b) Envelope of $(e^{(1)})^2 - (e^{(2)})^2$ changes sign accordingly. *sep* high (low) implies free air oscillation (interaction). The original scales of *sep*, e_{dif} and e_{sum} are in nm. These signals are normalized and shifted appropriately for convenience of comparison.

defined as the difference between free air oscillation amplitude of the cantilever to the oscillation amplitude of the cantilever interacting with infinitely hard sample. Equivalent parameters (f_{eq} and Q_{eq}) are analytically evaluated using averaging theory (see Agarwal, 2010) and compared with estimated values using BCEWRLS. It is evident from Fig. 4 that BCEWRLS estimated values closely match with analytical results. Fig. 4(b) further validates BCEWRLS method with experimental results where trends for f_{eq} and Q_{eq} similar to the results obtained analytically in Fig. 4(a) are observed when a cantilever with $f_0 = 82.3$ kHz, $Q_0 = 115.68$ is made to interact with mica during an approach–retract cycle (see Fig. 4(b)).

4.2. Real time detection

The effectiveness of the model detection methods developed in this article are assessed below in the context of probe based data storage. In simulation studies the tip–sample interaction model shown in Fig. 3 is used where the separation between the cantilever base and sample (*sep*) is raised and lowered with a pseudo random bit pattern of bit duration 200 μs. High *sep* in Fig. 5 indicates model 1 (free air oscillations) and low *sep* indicates model 2 (interaction). The free air oscillation amplitude of cantilever is chosen to be 24 nm. The measurement noise variance is chosen to set the noise bed in cantilever deflection data within 100 pm. ISM detectors are designed by matching the free air cantilever model (model 1; (A_1, B, C)) and BCEWRLS estimated equivalent model (model 2; (A_2, B, C)) obtained using $l_{int} = 2$ nm. It is evident from Fig. 5(a), when *sep* is low (interaction), both e_{sum} and e_{dif} oscillate
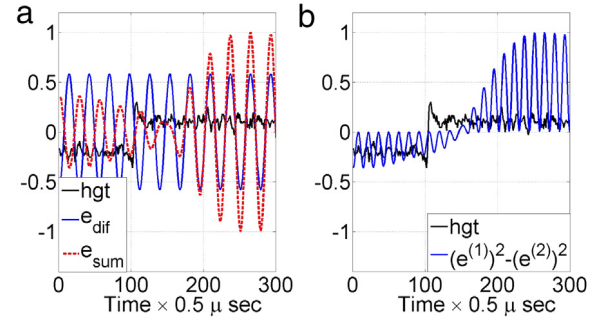

Fig. 6. (Experimental data) (a) *hgt* denotes measured height of sample. From 0–50 μs there is no interaction between the sample and the cantilever. e_{sum} and e_{dif} differ in phase by π radians in absence of interaction. After 50 μs the cantilever interacts with the sample. e_{sum} and e_{dif} signals align in phase within a couple of cantilever oscillations of the heightened topographic profile. (b) Envelope of $(e^{(1)})^2 - (e^{(2)})^2$ exhibits sign inversion within a few cycles of cantilever oscillation. The signals *hgt*, e_{dif} and e_{sum} are originally acquired in units of mV that are normalized and shifted appropriately for convenience of comparison.

in phase. However, high *sep* (free air oscillations) cause e_{dif} and e_{sum} to oscillate 180° out of phase confirming results stated in Corollary 3. In Fig. 5(b), the envelope of $(e^{(1)})^2 - (e^{(2)})^2$ follows the analytically predicted behavior in Section 2; the envelope is negative when the media follows depressed topography (the equivalent dynamics follows model 1) and is positive when the media follows a raised topography with an interaction length of 2 nm (the equivalent dynamics follows model 2). Note that the simulations incorporate the inevitable mismatch between the nonlinear time varying dynamics and the linear time-invariant dynamics of the equivalent model.

In the following the detection method is experimentally assessed. A field-programmable gate array (FPGA) based circuit actuates the piezoelectric nano-positioning system according to a pseudo random bit pattern of bit duration 350 μs. In Fig. 6(a) and (b), the signal *hgt* denotes the height of the sample. High *hgt* implies interaction and low *hgt* implies sample-free cantilever oscillations. The innovation signal $e^{(1)}$ is obtained from observer matched to the free air cantilever model ($f_0 = 71.73$ kHz, $Q_0 = 130.21$). The second observer is designed to match the equivalent cantilever model where f_{eq} is determined (using BCEWRLS method) to be 100 Hz less than f_0. Q_{eq} is determined to be same as Q_0. Behaviors of e_{sum}, e_{dif} and envelope of $(e^{(1)})^2 - (e^{(2)})^2$ resemble remarkably closely with theoretical predictions and simulation results (compare Figs. 5 and 6). The high bandwidth detection capability is also evident from experimental data where the sign inversion occurs within a few cycles of the cantilever oscillations (with time period $\approx 1/f_0$). Moreover the experimental results also support the fact that for probe based data storage application the deviations of the estimated models from the actual physical dynamics do not pose limitations on the ISM based method.

A comparison of ISM based amplitude and phase signals with conventional dynamic mode AFM signals (where excitation is sinusoidal) is presented below. A_{def} and ϕ_{def} in Fig. 7(a) and (b) denote the demodulated amplitude and phase of cantilever oscillations. A Butterworth low pass filter of bandwidth 5 kHz is used for extracting the ISM_{lpf} signal from $(e^{(1)})^2 - (e^{(2)})^2$. ISM_{lpf} and the phase difference between e_{dif}, e_{sum} (referred to as ϕ_m) signals are also shown in Fig. 7(a) and (b). In the time scale shown in Fig. 7(a) and (b), the conventional signals A_{def} and ϕ_{def} are always in their transient region. However, ISM based signals are observed to reach their steady state at a much shorter time scale. Further, images derived from conventional signals A_{def} and ϕ_{def} are dependent on history of the past topography encountered since a history of consistent high topography lowers the amplitude A_{def}. However, the ISM based signals ϕ_m and ISM_{lpf} follow a
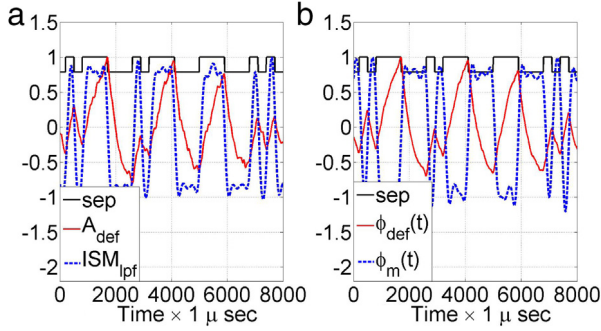
Fig. 7. (Simulation result) (a) Conventional dynamic mode AFM amplitude image (A_{def}) and ISM amplitude image (ISM_{lpf}) with cantilever–sample separation (*sep*). ISM_{lpf} is faster and exhibits significant independence of topographical history. (b) Regular phase image (ϕ_{def}) and ISM generated phase image (ϕ_m) from simulation. ϕ_m is more suitable for high speed binary deciphering of topography. The original units of *sep*, A_{def} and ISM_{lpf} are in nm, original units of ϕ_{def}, ϕ_m are in degrees. These signals are normalized and shifted accordingly for convenience of comparison.
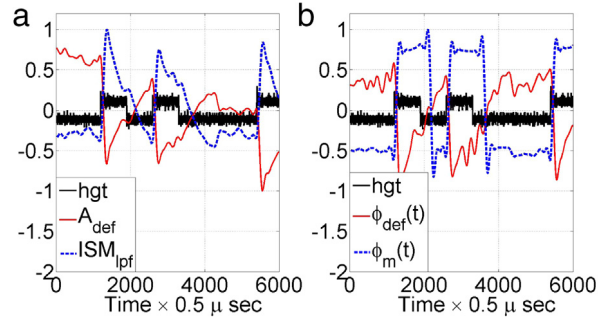

Fig. 8. (Experimental data) (a) Conventional amplitude A_{def} and ISM_{lpf} from experiment. A_{def} shows an overall drift from left to right. However, ISM_{lpf} has more consistent levels. (b) Regular phase ϕ_{def} and ISM based phase ϕ_m signals from experiment. ϕ_m has a very consistent binary behavior and better steady state magnitudes than ϕ_{def}. The original units of *hgt*, A_{def} and ISM_{lpf} are in mV. The original units of the phase signals ϕ_{def} and ϕ_m are in deg. These signals are normalized and shifted appropriately for convenience of comparison.

consistent binary behavior independent of history. Similar images are generated from experiments and compared (see Fig. 8). A low pass filter with cut off frequency of 3.3 kHz is utilized for extraction of ISM_{lpf} signal and also filtering out the high frequency noise present in ϕ_m. An interesting observation is that ϕ_m exhibits better steady behavior than ISM_{lpf} in experiment. In Fig. 8(a), value of A_{def} shows a downward drift from left to right. However, the magnitude levels of ISM_{lpf} are seen to be consistent implying that a threshold based detection is difficult from A_{def} but it is possible from ISM_{lpf}. ϕ_m is found to detect the topography changes at significantly higher bandwidth. When *hgt* changes, ϕ_m switches value almost seven times faster than ϕ_{def} (see Fig. 8(b)). Clearly ISM based detection outperforms conventional amplitude or phase images for binary detection of topography. Observation of Fig. 8(a) and (b) also suggest that the transition of test signals due to change of *hgt* from low to high is much faster than vice versa. A possible reason for this phenomenon is due to the fact that the oscillating cantilever trajectory evolves at slower rate towards its steady state when it goes from interaction to no-interaction state than the reverse case. The slower evolution of cantilever trajectory also influences the time scales of test signals which are generated by processing the cantilever trajectory itself. A detailed timing analysis for bit transitions for this application is an open problem which in this article is left for future studies.

A comparison of recursive MAP (18) and general signal correlation based implementation of ISM (13) is described here. For time varying systems, the MAP algorithm requires periodic refreshing of $p(\theta_1|\mathcal{Y}_m)$ and $p(\theta_2|\mathcal{Y}_m)$ (see chapter 10, Anderson &
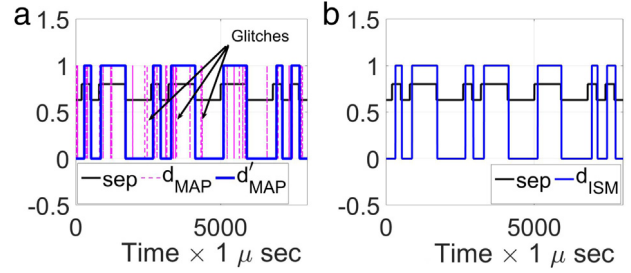

Fig. 9. (Simulation result) (a) MAP and (b) ISM based detection signals.
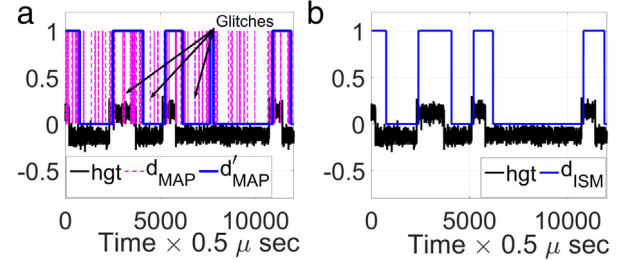

Fig. 10. (Experimental data) (a) MAP based decision image d_{MAP} is severely effected by decision glitches whereas d'_{MAP} has much better decision making capability. (b) d_{ISM} exhibits excellent detection performance.

Moore, 2012). A time window of 40 samples for periodic refreshing is found suitable for MAP algorithm. Such periodic refreshment is not necessary for ISM. In Fig. 9(a), the decision variable for MAP shown as d_{MAP} is set to 1 when $p(\theta_1|\mathcal{Y}_m) > p(\theta_2|\mathcal{Y}_m)$ otherwise 0. ISM signal $T[m]$ is gathered using a window size $M = 15$. d_{ISM} in Fig. 9(b) is set to 1 when $T[m] < 0$ otherwise 0. d_{MAP} as shown in Fig. 9(a) produces glitches in the decision even when *sep* is steady. A low pass filter with cut-off frequency of 3.3 kHz is further utilized for filtering $p(\theta_1|\mathcal{Y}_m)$ and $p(\theta_2|\mathcal{Y}_m)$ before decision making which can substantially eliminate the decision glitches in MAP algorithm and shown as d'_{MAP} in Fig. 10. d_{ISM} also shows excellent binary behavior without any errors in Fig. 9(b). Fig. 10(a) and (b) show a comparison of experimentally obtained decision images by recursive MAP and ISM. It is evident that the decision image from MAP without low pass filtering (d_{MAP}) results in considerable fraction of incorrect decisions (see Fig. 10(a)) which is mostly eliminated in d'_{MAP}. From Fig. 10(b) it is observed that ISM consistently yields high fidelity results. It should be noted that for the purpose of real time binary decision inferring, the plant detection approach using techniques MAP (d'_{MAP}) and ISM (ISM_{lpf}, ϕ_m and d_{ISM}) consistently offer superior bandwidth and accuracy of operation than traditional dynamic mode AFM amplitude and phase signals (observe Figs. 8 and 10). Moreover, ISM is computationally much less demanding than MAP. It should also be noted that observer based architectures have been studied before for topography detection with AFM (Sahoo, De Murti, & Salapaka, 2005; Sebastian, Sahoo, & Salapaka, 2003). However, it is already demonstrated in Ghosal et al. (2013) that ISM outperforms such previous arts for probe storage applications.

4.3. Symbol by symbol and sequence detection

A comparison of detection of plant models with finite dwell time as discussed in Section 3 is presented here. For simulations, *sep* is varied with a dwell time T (or equivalently bit duration) of 250 µs. Test signals for ISM in (13) and MAP in (18) are averaged over this time interval T before making decisions for symbol-by-symbol detection of plant models. In the previous work on MLSD for probe based data storage (see Kumar et al., 2010), the test signal $s_m(t)$ is chosen by applying a matched filter on the innovation
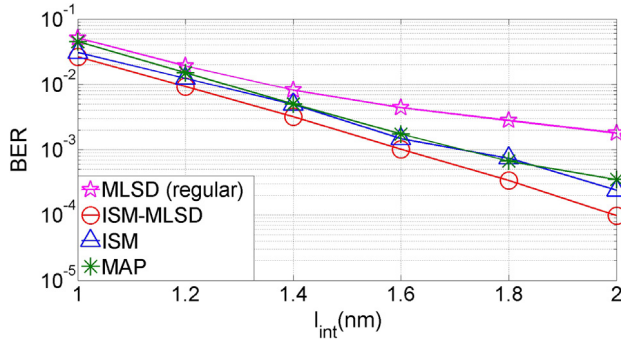
Fig. 11. (Simulation result) Bit error rate (BER) performance of symbol-by-symbol ISM and MAP with regular MLSD and ISM-MLSD as a function of interaction length l_{int}. Increasing l_{int} over the range 1–2 nm gradually improves BER performance of each scheme. Performance of ISM and MAP are close to each other. Regular MLSD lags slightly in error rate compared to other methods. However ISM-MLSD consistently outperforms all the other methods. Dwell time T is chosen as 250 μs.
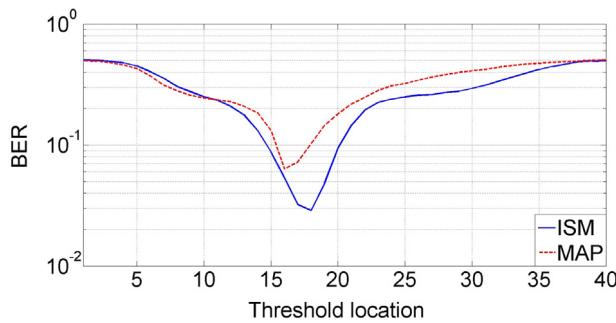

Fig. 12. (Experimental data) Bit error rate (BER) performance of symbol-by-symbol ISM and MAP with varying threshold location. Dwell time T is chosen as 350 μs.
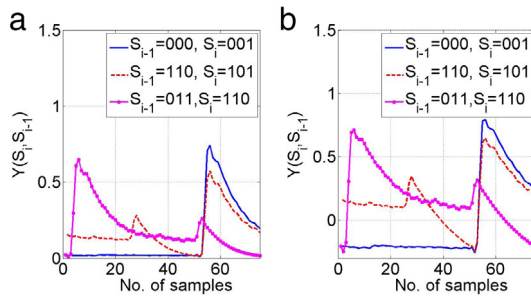

Fig. 13. (Experimental data) (a) Example mean observation vectors $\bar{\mathcal{Y}}$ for regular MLSD. (b) $\bar{\mathcal{Y}}$ for ISM-MLSD. The 3 state transitions shown here are $(S_{i-1}, S_i) \in \{(000, 001), (110, 101), (011, 110)\}$.

signal $e^{(1)}$. Apart from the assumption of known dwell time T, in Kumar et al. (2010) it is also assumed that the time when the oscillating cantilever tip impulsively interacts with the sample (q times during dwell interval T) is known and outputs of the matched filter are collected at their maximum value during each successive tip–sample interaction cycle. In Ghosal et al. (2013) it is shown that ISM performs as good as MLSD applied on test signal generated through processing $e^{(1)}$ as recommended in Kumar et al. (2010). However, from the perspective of implementation, having accurate knowledge of q instants where the cantilever interacts with the sample during each dwell time T is a difficult task. Hence in this article regular MLSD and ISM-MLSD are applied on test signals $s_m(t)$ generated by filtering $(e^{(1)})^2$ and $(e^{(1)})^2 - (e^{(2)})^2$ signals respectively with a low pass filter of cut-off frequency 3.3 kHz. This method is simpler to implement and the q samples during each interval T can be collected just by uniform sampling of the observation data without knowing the time instants of cantilever

tip–sample interaction. The decision error rates (also referred to as bit error rate; BER) of each method are observed where the bits are generated pseudo randomly. Average BERs are obtained for fixed on-sample interaction length l_{int} which is further varied from 1 nm to 2 nm. For all these cases, ISM signals are generated using the fixed equivalent model of $l_{int} = 2$ nm. The resulting BER vs l_{int} plots are compared in Fig. 11. Clearly, tapping the surface harder by increasing l_{int} improves the detection schemes. Even though the direct real time implementation of MAP algorithm suffers from decision glitches (see signal d_{MAP} in Fig. 10), when the MAP test signals in (18) are averaged over T, the resulting decision metric offers much better accuracy for symbol by symbol detection which is similar to the improvement obtained by utilizing low pass filters on real time MAP signals (see d'_{MAP} signal in Fig. 10). Thus BER performance of MAP follows closely to the performance by ISM in Fig. 11. In Fig. 11 ISM-MLSD performs better than other methods. However, it must be noted that for mild interactions of the order of $l_{int} \approx 2$ nm shown in Fig. 11, the slight BER improvement offered by ISM-MLSD compared to ISM is at the expense of much more computationally demanding signal processing. Similar simulations are performed by varying the dwell interval T and for all cases ISM-MLSD is observed to consistently outperform other methods.

Results from experiment are presented now. The sample height *hgt* is actuated with FPGA based actuator with a pseudo random bit pattern of bit duration $T = 350$ μs. Symbol-by-symbol test signals for ISM and MAP are obtained. The ranges of obtained values for ISM and MAP are divided into 40 grid points and they are used as decision thresholds for bit detection. Experimentally obtained BERs for ISM and MAP for various threshold locations are shown in Fig. 12. It is clear that the optimum threshold for detection (where BER is minimized) is slightly off from the mid-point of range of test values. The minimum BER for ISM and MAP for optimum choice of thresholds are obtained as 2.88% and 6.32% respectively. Hence more than 2-times improvement of accuracy is obtained by ISM over MAP.

Statistics for sequence detectors is gathered for known bit sequence with assumptions of system memory length $m = 1$ and inter symbol interference memory length $m_I = 2$. As shown in Fig. 2, there are 8 possible states S_i with 16 possible state transitions (S_{i-1}, S_i). For each possible state transition, average observation vector $\bar{\mathcal{Y}}$ and corresponding covariance matrix $\mathcal{C}$ are determined. Example of such experimentally gathered mean observation vectors for regular MLSD and ISM-MLSD corresponding to three possible state transitions $(S_{i-1}, S_i) \in \{(000, 001), (110, 101), (011, 110)\}$ are shown in Fig. 13. The set of statistics pairs $(\bar{\mathcal{Y}}, \mathcal{C})$ for 16 possible state transitions are utilized for Viterbi sequence detections using regular MLSD and ISM-MLSD. Interestingly, regular MLSD on the experimental data yields BER of 0.14% whereas ISM-MLSD yields only 0.07% BER. These experimental data suggest that MLSD and ISM-MLSD are capable of improvement of accuracy almost 20 times and 41 times than the best possible situation for symbol-by-symbol detection. The improvement of sequence detection over symbol by symbol method is achieved by correcting for the moderate ISI which is observed in the experimentally obtained test signals $(e^{(1)})^2$ and also $(e^{(1)})^2 - (e^{(2)})^2$. However, implementation of MLSD and ISM-MLSD are much more intricate compared to simple symbol-by-symbol methods and require a statistical learning step with known sequence of bits before detection of unknown bit sequence.

5. Conclusions

In this article a real time switching-model detection strategy ISM is developed for a system that switches between two models. The concept of utilizing the known plant models and sign-inverting test signals for ISM is developed. A MAP based detection algorithm

from existing literature is briefly outlined and compared with ISM. Detection schemes are modified for the situation when a dwell time with which the plant model switches is known a priori. Sequence detection method namely MLSD on innovation squared mismatch signal is proposed to further improve detection performance in the presence of inter symbol interference. In the application part, the dynamic mode AFM based fast topography detection system is cast as a binary switched system. Effectiveness and accuracy of the equivalent cantilever and switched plant model are discussed with theoretical predictions. The results indicate the superior speed of operation and the steady-state value advantages of plant model detection approach over the conventional AFM imaging. The simulation and experimental results also suggest that ISM is simpler to implement yet performs with similar detection speed and better accuracy compared to MAP for real time detection. Furthermore the robustness of ISM against the plant uncertainties, i.e., the situation when the observers are not accurately matched to the plant models and noise is present is borne out by thorough experimental data. Performance of sequence based detections are compared with ISM and MAP based symbol-by-symbol detections when the bit duration is known. It is verified with simulation and experiments that the proposed combination of ISM-MLSD offers significant gain in detection accuracy compared to other methods at the expense of computational complexity.

Appendix

A.1. Transfer function

From (5) and (6), the transfer function from $\left[x(s)^T, g(s)^T\right]^T$ to $\left[e_{dif}(s), e_{sum}(s)\right]^T$ in model i is given by

$$\mathcal{G}_i(s) = C_{big_i}(sI - A_{big_i})^{-1}B_{big_i} + D_{big_i}, \tag{A.1}$$

where A_{big_i} and B_{big_i} are given by

$$A_{big_i} = \begin{bmatrix} M_1 & M_2 \\ M_2 & M_1 \end{bmatrix},$$

$$B_{big_i}(:, 1) = \begin{bmatrix} A_{mi} + (L_2 - L_1)\Delta C_i \\ (-1)^i A_{mi} + 2\Delta A_i - (L_1 + L_2)\Delta C_i \end{bmatrix},$$

$$B_{big_i}(:, 2) = \begin{bmatrix} B_{mi} + (L_2 - L_1)\Delta D_i \\ (-1)^i B_{mi} - (L_1 + L_2)\Delta D_i \end{bmatrix},$$

where $(:, i)$ denotes the ith column of a matrix. Further, C_{big_i} and D_{big_i} are defined by

$$C_{big_i} = \begin{bmatrix} N_1 & N_2 \\ N_2 & N_1 \end{bmatrix}, D_{big_i}(:, 1) = \begin{bmatrix} C_{mi} \\ (-1)^i C_{mi} + 2\Delta C_i \end{bmatrix},$$

$$D_{big_i}(:, 2) = \begin{bmatrix} D_{mi} \\ (-1)^i D_{mi} + 2\Delta D_i \end{bmatrix}.$$

Using Woodbury's identity (Woodbury, 1950) it follows that

$$(sI - A_{big1})^{-1} = \begin{bmatrix} \mathcal{S}_1 & \mathcal{S}_2 \\ \mathcal{S}_2 & \mathcal{S}_1 \end{bmatrix}, \tag{A.2}$$

with $\mathcal{S}_1$ and $\mathcal{S}_2$ defined as

$$\mathcal{S}_1 = \left((sI - M_1)M_2^{-1}(sI - M_1) - M_2\right)^{-1}(sI - M_1)M_2^{-1},$$

$$\mathcal{S}_2 = \left((sI - M_1)M_2^{-1}(sI - M_1) - M_2\right)^{-1}. \tag{A.3}$$

From (A.1) it follows that

$$\mathcal{G}_i(s) = \begin{bmatrix} \mathcal{G}_{mi}^{(1)} + \Delta\mathcal{G}_{mi}^{(1)} & \mathcal{G}_{mi}^{(2)}(s) + \Delta\mathcal{G}_{mi}^{(3)} \\ (-1)^i\mathcal{G}_{mi}^{(1)} + \Delta\mathcal{G}_{mi}^{(2)} & (-1)^i\mathcal{G}_{mi}^{(2)} + \Delta\mathcal{G}_{mi}^{(4)} \end{bmatrix}, \tag{A.4}$$

where $\mathcal{G}_{mi}^{(1)}$ and $\mathcal{G}_{mi}^{(2)}$ are defined by

$$\mathcal{G}_{mi}^{(1)} = (N_1\mathcal{S}_1 + N_2\mathcal{S}_2 - (N_1\mathcal{S}_2 + N_2\mathcal{S}_1))A_{mi} + C_{mi},$$

$$\mathcal{G}_{mi}^{(2)} = (N_1\mathcal{S}_1 + N_2\mathcal{S}_2 - (N_1\mathcal{S}_2 + N_2\mathcal{S}_1))B_{mi} + D_{mi}. \tag{A.5}$$

Further, the transfer function matrices due to model perturbations can be found to be $\Delta\mathcal{G}_{mi}^{(1)} = (N_1\mathcal{S}_1 + N_2\mathcal{S}_2)(L_2 - L_1)\Delta C_i + (N_1\mathcal{S}_2 + N_2\mathcal{S}_1)(2\Delta A_i - (L_1 + L_2)\Delta C_i)$, $\Delta\mathcal{G}_{mi}^{(2)} = (N_1\mathcal{S}_2 + N_2\mathcal{S}_1)(L_2 - L_1)\Delta C_i + (N_1\mathcal{S}_1 + N_2\mathcal{S}_2)(2\Delta A_i - (L_1 + L_2)\Delta C_i) + 2\Delta C_i$, $\Delta\mathcal{G}_{mi}^{(3)} = (N_1\mathcal{S}_1 + N_2\mathcal{S}_2)(L_2 - L_1)\Delta D_i - (N_1\mathcal{S}_2 + N_2\mathcal{S}_1)(L_1 + L_2)\Delta D_i$, $\Delta\mathcal{G}_{mi}^{(4)} = (N_1\mathcal{S}_2 + N_2\mathcal{S}_1)(L_2 - L_1)\Delta D_i - (N_1\mathcal{S}_1 + N_2\mathcal{S}_2)(L_1 + L_2)\Delta D_i + 2\Delta D_i$.

A.2. Computational complexity of MAP

Consider the last term on the r.h.s. of (19) which is $(L_{d,i}R_{d,i}L_{d,i}^T)$. Write $L_{d,i}^T = [v_1, v_2, \ldots, v_{\mathfrak{m}_A}]$, v_j being jth column of $L_{d,i}^T$. (j, k)th element of $(L_{d,i}R_{d,i}L_{d,i}^T)$ is $(v_j^T R_{d,i}v_k)$. If $v_j \neq v_k$, $(v_j^T R_{d,i}v_k) = \sum_{m=1}^{\mathfrak{m}_C}\sum_{n=1}^{\mathfrak{m}_C} v_j(m)R_{d,i}(m, n)v_k(n)$. Clearly, $2\mathfrak{m}_C^2$ multiplications, $(\mathfrak{m}_C^2 - 1)$ additions with total flops $2\mathfrak{m}_C^2 + (\mathfrak{m}_C^2 - 1) = 3\mathfrak{m}_C^2 - 1$ are needed for calculating $(v_j^T R_{d,i}v_k)$ $(v_j \neq v_k)$. For symmetric $R_{d,i} \in \mathbb{R}^{\mathfrak{m}_C,\mathfrak{m}_C}$ and any $v_j \in \mathbb{R}^{\mathfrak{m}_C}$, computing $(v_j^T R_{d,i}v_j)$ requires $(\frac{3}{2}\mathfrak{m}_C^2 + \frac{3}{2}\mathfrak{m}_C - 1)$ flop counts (see Hunger, 2005). $(L_{d,i}R_{d,i}L_{d,i}^T)$ being symmetric, $\mathfrak{m}_A$ diagonal elements of the form $(v_j^T R_{d,i}v_j)$, $j = 1, 2, \ldots, \mathfrak{m}_A$ need to be computed each requiring $(\frac{3}{2}\mathfrak{m}_C^2 + \frac{3}{2}\mathfrak{m}_C - 1)$ flops. The $\mathfrak{m}_A(\mathfrak{m}_A - 1)/2$ off-diagonal terms need $(3\mathfrak{m}_C^2 - 1)$ flops. Thus computation of $(L_{d,i}R_{d,i}L_{d,i}^T)$ needs $(\mathfrak{m}_A(\frac{3}{2}\mathfrak{m}_C^2 + \frac{3}{2}\mathfrak{m}_C - 1) + (\mathfrak{m}_A(\mathfrak{m}_A - 1)/2)(3\mathfrak{m}_C^2 - 1)) = (\frac{3}{2}\mathfrak{m}_A^2\mathfrak{m}_C^2 + \frac{3}{2}\mathfrak{m}_A\mathfrak{m}_C - \frac{1}{2}\mathfrak{m}_A - \frac{1}{2}\mathfrak{m}_A^2) = N_{\Sigma_1}$ (say) flops. Similarly $(B_{d,i}Q_{d,i}B_{d,i}^T)$ requires $(\frac{3}{2}\mathfrak{m}_A^2\mathfrak{n}_B^2 + \frac{3}{2}\mathfrak{m}_A\mathfrak{n}_B - \frac{1}{2}\mathfrak{m}_A - \frac{1}{2}\mathfrak{m}_A^2) = N_{\Sigma_2}$ flops. Now, flop counts needed for calculating $(A_{d,i} - L_{d,i}C_{d,i})$ is $\mathfrak{m}_A^2 + \mathfrak{m}_A^2(2\mathfrak{m}_C - 1) = 2\mathfrak{m}_C\mathfrak{m}_A^2$. Once $(A_{d,i} - L_{d,i}C_{d,i})$ and $\Sigma_{m/m-1,i}$ are available, flop counts for $((A_{d,i} - L_{d,i}C_{d,i})\Sigma_{m/m-1,i}(A_{d,i} - L_{d,i}C_{d,i})^T)$ can be calculated similar to described above which is $(\frac{3}{2}\mathfrak{m}_A^4 + \mathfrak{m}_A^2 - \frac{1}{2}\mathfrak{m}_A) = N_{\Sigma_3}$. So, computation of $((A_{d,i} - L_{d,i}C_{d,i})\Sigma_{m/m-1,i}(A_{d,i} - L_{d,i}C_{d,i})^T)$ needs $N_{\Sigma_3} + (2\mathfrak{m}_C\mathfrak{m}_A^2)$ flops. Let N_Σ denote flops needed for the r.h.s. of (19). Flops needed for adding $((A_{d,i} - L_{d,i}C_{d,i})\Sigma_{m/m-1,i}(A_{d,i} - L_{d,i}C_{d,i})^T)$, $(B_{d,i}Q_{d,i}B_{d,i}^T)$ and $(L_{d,i}R_{d,i}L_{d,i}^T)$ is $\mathfrak{m}_A(\mathfrak{m}_A+1)$. $(A_{d,i} - L_{d,i}C_{d,i})$, $(B_{d,i}Q_{d,i}B_{d,i}^T)$ and $(L_{d,i}R_{d,i}L_{d,i}^T)$ are time invariant and need to be computed only once for the first count of the recursive update. Clearly, for the first recursive update $N_\Sigma = N_{\Sigma_1} + N_{\Sigma_2} + N_{\Sigma_3} + (2\mathfrak{m}_C\mathfrak{m}_A^2) + \mathfrak{m}_A(\mathfrak{m}_A + 1)$. From the second recursive stage onwards, the matrices $(B_{d,i}Q_{d,i}B_{d,i}^T + L_{d,i}R_{d,i}L_{d,i}^T)$ and $(A_{d,i} - L_{d,i}C_{d,i})$ are already available implying $N_\Sigma = N_{\Sigma_3} + \mathfrak{m}_A(\mathfrak{m}_A + 1)/2$. Thus

$$N_\Sigma = \begin{cases} \left(\frac{3}{2}\mathfrak{m}_A^4 + \mathfrak{m}_A^2\left(\frac{3}{2}(\mathfrak{m}_C^2 + \mathfrak{n}_B^2) + 2\mathfrak{m}_C + 1\right) \\ \quad + \mathfrak{m}_A\left(\frac{3}{2}(\mathfrak{m}_C + \mathfrak{n}_B) - \frac{1}{2}\right)\right), & \text{for first update} \\ \frac{3}{2}\left(\mathfrak{m}_A^4 + \mathfrak{m}_A^2\right), & \text{for subsequent updates.} \end{cases} \tag{A.6}$$

Let N_Ω denote the flops per update for (20) which can be easily shown to be $N_\Omega = \frac{3}{2}(\mathfrak{m}_C^2\mathfrak{m}_A^2 + \mathfrak{m}_C\mathfrak{m}_A)$. Let N_p denote flops needed for update of (18). $\Omega_{m/i}$ is positive definite so its determinant and inverse can be calculated by Cholesky decomposition which is $\Omega_{m/i} = L_\Omega L_\Omega^T$. Standard algorithm for Cholesky factorization requires $(\frac{1}{3}\mathfrak{m}_C^3 + \frac{1}{2}\mathfrak{m}_C^2 + \frac{1}{6}\mathfrak{m}_C)$ flops (see Hunger, 2005). Then, $|\Omega_{m/i}|^{-\frac{1}{2}} = |L_\Omega|^{-1}$. L_Ω being lower triangular, $|L_\Omega|$ can be computed by multiplying its diagonal elements needing $(\mathfrak{m}_C - 1)$ multiplications. So, when the matrix L_Ω is available then $|\Omega_{m/i}|^{-\frac{1}{2}}$

needs $(\mathfrak{m}_C - 1 + 1) = \mathfrak{m}_C$ flops. Now, $\Omega_{m/i}^{-1} = L_\Omega^{-T} L_\Omega^{-1}$. Inversion of L_Ω needs $(\frac{1}{3}\mathfrak{m}_C^3 + \frac{2}{3}\mathfrak{m}_C)$ flops and computing $L_\Omega^{-T} L_\Omega^{-1}$ needs $(\frac{1}{3}\mathfrak{m}_C^3 + \frac{1}{2}\mathfrak{m}_C^2 + \frac{1}{6}\mathfrak{m}_C)$ flops (Hunger, 2005). So, $|\Omega_{m/i}|^{-\frac{1}{2}}$ and $\Omega_{m/i}^{-1}$ can be computed using $(\frac{1}{3}\mathfrak{m}_C^3 + \frac{1}{2}\mathfrak{m}_C^2 + \frac{1}{6}\mathfrak{m}_C) + (\mathfrak{m}_C) + (\frac{1}{3}\mathfrak{m}_C^3 + \frac{2}{3}\mathfrak{m}_C) + (\frac{1}{3}\mathfrak{m}_C^3 + \frac{1}{2}\mathfrak{m}_C^2 + \frac{1}{6}\mathfrak{m}_C) = (\mathfrak{m}_C^3 + \mathfrak{m}_C^2 + 2\mathfrak{m}_C)$ flops. Evaluation of $(-\frac{1}{2}\tilde{y}_{m,i}^T \Omega_{m/i}^{-1} \tilde{y}_{m,i})$ needs $1 + (\frac{3}{2}\mathfrak{m}_C^2 + \frac{3}{2}\mathfrak{m}_C - 1) = \frac{3}{2}(\mathfrak{m}_C^2 + \mathfrak{m}_C)$ flops when $\Omega_{m/i}^{-1}$ and $\tilde{y}_{m,i}$ are available. As a result, total number of flops required for computation of $p(\theta_i|\mathcal{Y}_m)$ is given by $N_p = (\mathfrak{m}_C^3 + \mathfrak{m}_C^2 + 2\mathfrak{m}_C) + \frac{3}{2}(\mathfrak{m}_C^2 + \mathfrak{m}_C) + 3 + N_{exp} = \mathfrak{m}_C^3 + \frac{5}{2}\mathfrak{m}_C^2 + \frac{7}{2}\mathfrak{m}_C + 3 + N_{exp}$. Here, N_{exp} is the flop counts required to compute $\exp(.)$ operation. N_{exp} may vary depending on the algorithm used to implement $\exp(.)$. For simplicity assume that $N_{exp} = 1$. So, $N_p = \mathfrak{m}_C^3 + \frac{5}{2}\mathfrak{m}_C^2 + \frac{7}{2}\mathfrak{m}_C + 4$. Since $i \in 1, 2$, flops needed per recursive update is $2(N_\Sigma + N_\Omega + N_p)$.

References

Agarwal, P. (2010). *Model mismatch paradigm for probe based nanoscale imaging.* (Ph.D. thesis).

Agarwal, P., & Salapaka, M. (2009). Real time estimation of equivalent cantilever parameters in tapping mode atomic force microscopy. *Applied Physics Letters*, 95(8), 083113–083113.

Aggarwal, T., Materassi, D., Davison, R., Hays, T., & Salapaka, M. (2012). Detection of steps in single molecule data. *Cellular and molecular bioengineering*, 5(1), 14–31.

Anderson, B. D., & Moore, J. B. (2012). *Optimal filtering.* Courier Dover Publications.

Bhaskaran, H., Gotsmann, B., Sebastian, A., Drechsler, U., Lantz, M. A., Despont, M., Jaroenapibal, P., Carpick, R. W., Chen, Y., & Sridharan, K. (2010). Ultralow nanoscale wear through atom-by-atom attrition in silicon-containing diamond-like carbon. *Nature nanotechnology*, 5(3), 181–185.

Bhaskaran, H., Sebastian, A., Drechsler, U., & Despont, M. (2009). Encapsulated tips for reliable nanoscale conduction in scanning probe technologies. *Nanotechnology*, 20(10), 105701.

Binnig, G., Quate, C., & Gerber, C. (1986). Atomic force microscope. *Physical Review Letters*, 56(9), 930–933.

Bray, M. T., Cohen, S. H., & Lightbody, M. L. (2013). *Atomic force microscopy/scanning tunneling microscopy.* Springer Science & Business Media.

Cai, K., Qin, Z., Zhang, S., Ng, Y., Chai, K., & Radhakrishnan, R. (2010). Modeling, detection, and ldpc codes for bit-patterned media recording. In *GLOBECOM workshops, GC Wkshps, 2010 IEEE* (pp. 1910–1914). IEEE.

Campardo, G., Tiziani, F., & Iaculo, M. (2011). *Memory mass storage.* Springer Science & Business Media.

Cannara, R. J., Gotsmann, B., Knoll, A., & Dürig, U. (2008). Thermo-mechanical probe storage at mbps single-probe data rates and tbit in- 2 densities. *Nanotechnology*, 19(39), 395305.

Carosino, M., Chen, Y., Belzer, B. J., Sivakumar, K., Murray, J., & Wettin, P. (2013). Iterative detection and decoding for the four-rectangular-grain tdmr model. In *Communication, control, and computing, Allerton, 2013 51st annual allerton conference on* (pp. 653–659). IEEE.

Chen, Y., & Srinivasa, S. G. (2013). Joint self-iterating equalization and detection for two-dimensional intersymbol-interference channels. *Communications, IEEE Transactions on*, 61(8), 3219–3230.

Cho, Y., Hashimoto, S., Odagawa, N., Tanaka, K., & Hiranaga, Y. (2006). Nanodomain manipulation for ultrahigh density ferroelectric data storage. *Nanotechnology*, 17(7), S137.

Datta, A. (2012). *Adaptive internal model control.* Springer Science & Business Media.

Feng, G., JianQiang, W., & Keqiang, L. (2006). Hierarchical switching control of longitudinal acceleration with large uncertainties. In *Vehicular electronics and safety, 2006. ICVES 2006. IEEE International conference on* (pp. 270–275). IEEE.

Forney, G., Jr (1972). Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference. *Information Theory, IEEE Transactions on*, 18(3), 363–378.

Gao, F., Li, S. E., Kum, D., & Zhang, H. (2015). Synthesis of multiple model switching controllers using $h\infty$ theory for systems with large uncertainties. *Neurocomputing*, 157, 118–124.

Gao, F., Li, X., & Li, Y. (2014). Control of system with large parametric uncertainty using multiple robust controllers and switching. In *Intelligent control and automation, WCICA, 2014 11th World Congress on* (pp. 4408–4413). IEEE.

Ghosal, S., Saraswat, G., Ramamoorthy, A., & Salapaka, M. (2013). Topography detection using innovations mismatch method for high speed and high density dynamic mode afm. In *American control conference, ACC, 2013* (pp. 5500–5505). IEEE.

Gotsmann, B., Knoll, A., Pratt, R., Frommer, J., Hedrick, J., & Duerig, U. (2010). Designing polymers to enable nanoscale thermomechanical data storage. *Advanced Functional Materials*, 20(8), 1276–1284.

Hamann, H. F., O'Boyle, M., Martin, Y. C., Rooks, M., & Wickramasinghe, H. K. (2006). Ultra-high-density phase-change storage and memory. *Nature materials*, 5(5), 383–387.

Han, Z., & Narendra, K. S. (2012). New concepts in adaptive control using multiple models. *IEEE Transactions on Automatic Control*, 57(1), 78–89.

Huang, M., Wang, X., & Wang, Z. (2015). Multiple model adaptive control for a class of linear-bounded nonlinear systems. *IEEE Transactions on Automatic Control*, 60(1), 271–276.

Hunger, R. (2005). *Floating point operations in matrix–vector calculus.* Munich University of Technology, Inst. for Circuit Theory and Signal Processing.

Kailath, T., Sayed, A. H., & Hassibi, B. (2000). *Linear estimation: vol. 1.* NJ: Prentice Hall Upper Saddle River.

Kappenberger, P., Luo, F., Heyderman, L., Solak, H., Padeste, C., Brombacher, C., Makarov, D., Ashworth, T., Philippe, L., Hug, H., et al. (2009). Template-directed self-assembled magnetic nanostructures for probe recording. *Applied Physics Letters*, 95(2), 023116.

Kavcic, A., & Moura, J. (2000). The viterbi algorithm and markov noise memory. *Information Theory, IEEE Transactions on*, 46(1), 291–301.

Khatami, S. M., & Vasic, B. (2013). Detection for two-dimensional magnetic recording systems. In *Computing, networking and communications, ICNC, 2013 international conference on* (pp. 535–539). IEEE.

Knoll, A., Bächtold, P., Bonan, J., Cherubini, G., Despont, M., Drechsler, U., Dürig, U., Gotsmann, B., Häberle, W., Hagleitner, C., et al. (2006). Integrating nanotechnology into a working storage device. *Microelectronic Engineering*, 83(4), 1692–1697.

Koelmans, W.W., Engelen, J.B., & Abelmann, L. Probe-based data storage, arXiv preprint arXiv:1511.08755, 2015.

Krishnan, A. R., Radhakrishnan, R., Vasic, B., Kavcic, A., Ryan, W., & Erden, F. (2009). 2-d magnetic recording: Read channel modeling and detection. *Magnetics, IEEE Transactions on*, 45(10), 3830–3836.

Kryder, M. H., Gage, E. C., McDaniel, T. W., Challener, W. A., Rottmayer, R. E., Ju, G., Hsia, Y.-T., & Erden, M. F. (2008). Heat assisted magnetic recording. *Proceedings of the IEEE*, 96(11), 1810–1835.

Kumar, N., Agarwal, P., Ramamoorthy, A., & Salapaka, M. (2010). Maximum-likelihood sequence detector for dynamic mode high density probe storage. *Communications, IEEE Transactions on*, 58(6), 1686–1694.

Lai, S. K. (2008). Flash memories: Successes and challenges. *IBM Journal of Research and Development*, 52(4.5), 529–535.

Lantz, M. A., Gotsmann, B., Jaroenapibal, P., Jacobs, T. D., O'Connor, S. D., Sridharan, K., & Carpick, R. W. (2012). Wear-resistant nanoscale silicon carbide tips for scanning probe applications. *Advanced Functional Materials*, 22(8), 1639–1645.

Lantz, M. A., Wiesmann, D., & Gotsmann, B. (2009). Dynamic superlubricity and the elimination of wear on the nanoscale. *Nature nanotechnology*, 4(9), 586–591.

Mamin, H., Ried, R., Terris, B., & Rugar, D. (1999). High-density data storage based on the atomic force microscope. *Proceedings of the IEEE*, 87(6), 1014–1027.

Moon, J., & Carley, L. R. (2012). *Sequence detection for high-density storage channels: vol. 187.* Springer Science & Business Media.

Morse, A. (1993). Supervisory control of families of linear set-point controllers. In *Decision and control, 1993. Proceedings of the 32nd IEEE conference on* (pp. 1055–1060). IEEE.

Morse, A. S. (1996). Supervisory control of families of linear set-point controllers part i. exact matching. *IEEE Transactions on Automatic Control*, 41(10), 1413–1431.

Morse, A. S. (1997). Supervisory control of families of linear set-point controllers. 2. robustness. *IEEE Transactions on Automatic Control*, 42(11), 1500–1515.

Narendra, K. S., & Balakrishnan, J. (1994). Improving transient response of adaptive control systems using multiple models and switching. *IEEE Transactions on Automatic Control*, 39(9), 1861–1866.

Narendra, K. S., & Balakrishnan, J. (1997). Adaptive control using multiple models. *IEEE Transactions on Automatic Control*, 42(2), 171–187.

Richter, H. (2009). Density limits imposed by the microstructure of magnetic recording media. *Journal of Magnetism and Magnetic Materials*, 321(6), 467–476.

Sahoo, D. R., De Murti, T., & Salapaka, V. (2005). Observer based imaging methods for atomic force microscopy. In *Proceedings of the 44th IEEE conference on decision and control* (pp. 1185–1190). IEEE.

Sanders, J., & Verhulst, F. (1985). *Averaging methods in nonlinear dynamical systems: vol. 59.* Springer.

San Paulo, A., & Garcia, R. (2002). Unifying theory of tapping-mode atomic-force microscopy. *Physical Review B*, 66(4), 041406.

Saraswat, G., Agarwal, P., Haugstad, G., & Salapaka, M. (2013). Real-time probe based quantitative determination of material properties at the nanoscale. *Nanotechnology*, 24(26), 265706.

Sebastian, A., Gannepalli, A., & Salapaka, M. (2007). A review of the systems approach to the analysis of dynamic-mode atomic force microscopy. *IEEE Transactions on Control Systems Technology*, 15(5), 952–959.

Sebastian, A., Gannepalli, A., & Salapaka, M. (2004). The amplitude phase dynamics and fixed points in tapping-mode atomic force microscopy. In *American control conference, 2004. Proceedings of the 2004: vol. 3* (pp. 2499–2504). IEEE.

Sebastian, A., Sahoo, D. R., & Salapaka, M. V. (2003). An observer based sample detection scheme for atomic force microscopy. In *Decision and control, 2003. Proceedings. 42nd IEEE Conference on: vol. 3* (pp. 2132–2137). IEEE.

Sebastian, A., Salapaka, M., Chen, D., & Cleveland, J. (2001). Harmonic and power balance tools for tapping-mode atomic force microscope. *Journal of Applied Physics*, 89, 6473.

Sun, Z., & Ge, S. S. (2005). Analysis and synthesis of switched linear control systems. *Automatica*, 41(2), 181–195.

Tanaka, K., & Cho, Y. (2010). Actual information storage with a recording density of 4 tbit/ in. 2 in a ferroelectric recording medium. *Applied physics letters*, 97(9).

Vettiger, P., Cross, G., Despont, M., Drechsler, U., Durig, U., Gotsmann, B., Haberle, W., Lantz, M., Rothuizen, H., Stutz, R., et al. (2002). The millipede-nanotechnology entering data storage. *Nanotechnology, IEEE Transactions on*, *1*(1), 39–55.

Wiesmann, D., Rawlings, C., Vecchione, R., Porro, F., Gotsmann, B., Knoll, A., Pires, D., & Duerig, U. (2009). Multi tbit/in2 storage densities with thermomechanical probes. *Nano letters*, *9*(9), 3171–3176.

Woodbury, M. A. (1950). *Inverting modified matrices, Memorandum report*: vol. 42 (p. 106).

Wright, C. D., Aziz, M. M., Shah, P., & Wang, L. (2011). Scanning probe memories–technology and applications. *Current Applied Physics*, *11*(2), e104–e109.

Govind Saraswat received his Bachelors of Technology degree in Electrical Engineering (EE) in 2007 from Indian Institute of Technology (IIT), Delhi, India. He received his Ph.D. from the Department of Electrical and Computer Engineering at University of Minnesota (Twin Cities) in January 2014. His current research interest is in applying system and control concepts for the faster and better (higher resolution) imaging using atomic force microscopes.

Sayan Ghosal received his B.E. degree in Electronics and Telecommunication Engineering in 2007 from Jadavpur University, Kolkata, India. Subsequently, he acquired M.Tech. degree in Communication Systems from Indian Institute of Technology, Madras, India in 2010. From September 2010 to present, he has been pursuing his Ph.D. at the Department of Electrical and Computer Engineering, University of Minnesota (Twin Cities), USA. His research interests include advanced control systems, system modeling and identification, estimation and detection, machine learning and atomic force microscopy.

Murti Salapaka obtained his Bachelors of Technology degree in Mechanical Engineering from Indian Institute of Technology, Madras in 1991. He obtained his Masters and Ph.D. degrees from University of California, Santa Barbara in the years 1993 and 1997 respectively. He was at Electrical Engineering Department at Iowa State University from 1997–2007. He is currently a faculty in the Electrical and Computer Engineering Department at University of Minnesota (Twin Cities), USA. His research interests span control theory and its applications to nano-interrogation and bio-manipulation at the molecular scale using laser tweezers and atomic force microscopes.