

# Near optimal signal detection for finite state Markov signals with application to magnetic resonance force microscopy

Michael Ting\*, *Student Member, IEEE*, Alfred O. Hero, *Fellow, IEEE*, Daniel Rugar†, *Member, IEEE*, Chun-yu Yip, *Student Member, IEEE*, and Jeffrey A. Fessler, *Senior Member, IEEE*

**Abstract**—Detection of a finite state Markov signal in additive white Gaussian noise (AWGN) can be done in an intuitive manner by applying an appropriate filter and using an energy detector. One might not expect this heuristic approach to signal detection to be optimal. However, in this paper, we show that for a certain type of finite state Markov signal, namely the discrete-time (DT) random telegraph, this filtered energy detector is approximately optimal under the following conditions of: symmetric transition probabilities, low signal-to-noise ratio (SNR), long observation time, and small probability of transition between two consecutive time instances. When these last three conditions hold, but the transition probabilities are not symmetric, we show that a variant of the filtered energy detector is approximately optimal. It is also shown, under low SNR conditions, that the likelihood ratio test (LRT) for a finite state DT Markov signal in AWGN reduces to the matched filter statistic with the minimum mean squared error (MMSE) predictor signal values used in place of the known signal values. Using this result, we propose a general methodology for obtaining an approximation to the LRT of a finite state DT Markov signal in AWGN. Specifically, instead of the conditional mean (also MMSE) estimators, affine estimators with lowest mean squared error (MSE) are used. This work is relevant to magnetic resonance force microscopy, an emerging technology that uses ultrasensitive force sensing to detect magnetic resonance. Sensitivity down to the single spin level was demonstrated in a recent experiment.

## I. INTRODUCTION

Detection of a finite state DT Markov signal in AWGN is widespread in many different fields. Detection of a random telegraph signal is used in the study of particle tunnelling [1], and in the study of low-frequency noise characteristics of light-emitting diodes [2]. Markov chains are used in [3] for the purpose of statistical network anomaly detection, and in [4], for the purpose of land mine detection. The focus application of this paper is magnetic resonance force microscopy (MRFM), which is a promising technique for three-dimensional imaging on the nanometer scale. Recent experiments at IBM have shown that MRFM is capable of detecting

and localizing individual electron spins associated with sub-surface atomic defects in silicon dioxide [5]. This single-spin detection milestone represents a factor of  $10^7$  improvement over conventional electron spin resonance detection and was achieved using energy detection methods similar to those described in this paper. Other recent MRFM experiments have demonstrated the ability to detect and manipulate naturally occurring statistical fluctuations in small spin ensembles [6]. With further development, single-spin MRFM may eventually lead to atomic-resolution magnetic resonance imaging and find application in quantum computing experiments [7].

The recursive structure of the LRT for a finite state DT Markov signal is given in [8]. In this paper, we specialize the noise to AWGN, and derive a new interpretation of the optimal LRT for a finite state DT Markov signal under low SNR conditions. It is shown that, under low SNR, the LRT reduces to the matched filter statistic with the MMSE predictor values used in place of the known signal values. Current single spin experiments operate under conditions of very low SNR; consequently, we are interested in the performance of detectors in the regime of low SNR and long observation time. Our results are applicable to [1], [2], [3], [4] under conditions of low SNR. When applied to the LRT of the DT random telegraph, the result is an estimator-correlator detector. This estimator-correlator structure appears in the LRT of problems whose probability density functions have an exponential form [9]. In particular, it applies when detecting a Gaussian signal in AWGN. This first result has a continuous-time (CT) analog: in CT, the LRT for detecting a random signal in AWGN has the form of the matched filter statistic with the MMSE predictor used in place of the known signal values. It is, however, exact under all SNR conditions [10], [11], [12]. There is another difference: the square of the conditional expectation of the random process is used instead of the conditional expectation of the squared value of the random process.

Secondly, when used to detect the DT random telegraph in AWGN, the filtered energy (FE) detector is approximately optimal under the following four conditions: symmetric transition probabilities, low SNR, long observation time, and a small probability of transition between two consecutive instances. The FE detector is no longer approximately optimal when the transition probabilities are asymmetric. We extend the FE detector to a hybrid second-order detector which combines the filtered energy, amplitude, and energy statistics. It is shown that the hybrid detector is approximately optimal for the DT

This work was partially supported by the DARPA Mosaic program under ARO contract DAAD19-02-C-0055.

† D. Rugar is with the IBM Research Division, Almaden Research Center, San Jose, CA 95120, USA (email: rugar@almaden.ibm.com, ph: 408 927 2027, fax: 408 927 2510).

The other authors are with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109-2108, USA. \*M. Ting (email: mting@umich.edu, ph: 734 764 5216), A. O. Hero (email: hero@eecs.umich.edu, ph: 734 763 0564), J. A. Fessler (email: fessler@eecs.umich.edu, ph: 734 763 1434), C.-Y. Yip (email: chun-yuy@umich.edu, phone: 734 730 2659). Fax: 734 763 8041

random telegraph model under only the last three conditions. This is an intuitively pleasing result, as the idea of performing detection of a finite state Markov signal by filtering the noisy observations and applying an energy detector is one that comes naturally.

Thirdly, the first result is used to obtain an approximation to the LRT of a general finite state DT Markov signal in AWGN. Suboptimal affine estimators of the random process and the squared value of the random process are used instead of the optimal conditional mean (also MMSE) estimators. When this general methodology is applied to the LRT of the DT random telegraph in AWGN, an approximation is obtained that closely matches the result that was obtained using a straightforward analysis. It has been noted in [13] that linear MMSE estimates of the CT symmetric random telegraph process are as efficient as the nonlinear MMSE estimates as the  $\text{SNR} \rightarrow 0$ . We also present simulations that suggest the optimality of a similar approximation for the DT random walk process. It would be interesting to investigate whether affine MMSE estimators are as efficient as nonlinear MMSE estimators for a general finite state DT Markov signal in AWGN as the  $\text{SNR} \rightarrow 0$ .

The outline of this paper is as follows. In Section II, we briefly review the basic principles of MRFM. This is followed by a discussion in Section III of two finite state DT Markov signal models: the random telegraph and random walk models. In Section IV, we describe existing detectors that are commonly used, namely the amplitude and filtered energy detectors, and compare them to the optimal detectors. We derive a new interpretation of the LRT under low SNR conditions. As well, the FE detector is extended to a hybrid version, and a general methodology to obtain an approximation to the LRT of a finite state DT Markov process in AWGN is presented. Simulation results are presented in Section V.

## II. BASIC PRINCIPLES OF MRFM SPIN DETECTION

MRFM experiments, in general, involve the measurement of magnetic force between a submicron-size magnetic tip and spins in a sample. The details of spin manipulation and signal detection depend on the exact MRFM protocol used. One particularly successful protocol is called OSCAR, which stands for Oscillating Cantilever-driven Adiabatic Reversal [14], [15]. A variation of this protocol, “interrupted OSCAR” (iOSCAR), was used in recent single spin experiments [5].

A schematic diagram of an OSCAR-type MRFM experiment is shown in Fig. 1. As shown in the figure, a submicron ferromagnet is placed on the tip of a cantilever and positioned close to an unpaired electron spin contained within the sample. An applied radio-frequency (rf) field serves to induce magnetic resonance of the spin when the condition  $B_0 = \omega_{\text{rf}}/\gamma$  is met. Here,  $B_0$  is the magnitude of the magnetic field from the tip, plus any externally applied static field that may be present. The constant  $\gamma = 5.6\pi \times 10^{10} \text{ T}^{-1}\text{s}^{-1}$  is the gyromagnetic ratio, and  $\omega_{\text{rf}}$  is the (single) frequency of the applied rf field. Because the magnetic field emanating from the tip is highly inhomogeneous, magnetic resonance is confined spatially to a thin bowl-shaped region called the “resonant slice”.

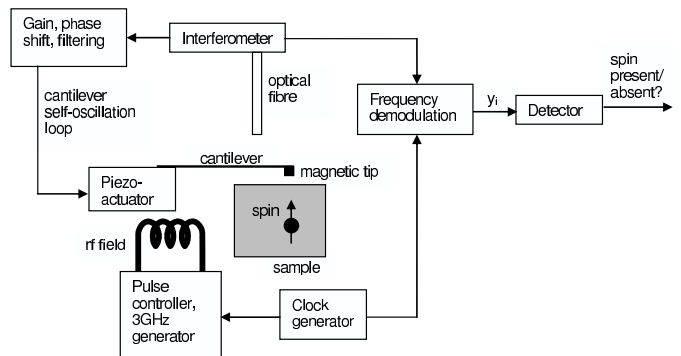


Fig. 1. Schematic of an OSCAR-type MRFM experiment.

In an OSCAR experiment, a gain-controlled positive feedback loop is used to oscillate the cantilever with a preset amplitude (typically 10-20 nm). The cantilever oscillation frequency is determined by the cantilever itself (specifically, by the fundamental flexural mode eigenfrequency), as well as by tip-sample interactions. As the tip of the cantilever vibrates, the resonant slice passes back and forth through the spin and, as a result, the spin direction is cyclically inverted due to an effect called adiabatic rapid passage [14], [16], [17]. The cyclic inversion is synchronous with the cantilever motion and affects the cantilever dynamics by slightly shifting the cantilever resonant frequency. The frequency shift depends on the angle  $\theta$  of the spin with respect to a vector called the “effective field in the rotating frame”. See [15], [16] for further details. The frequency shift can be written as

$$\Delta\omega = \Delta\omega_{\text{max}} \cos\theta, \quad (1)$$

where  $\Delta\omega_{\text{max}} = 2\omega_0 G\mu/\pi k x_{\text{pk}}$  [15]. Here  $\omega_0$  is the unperturbed cantilever frequency,  $G$  is the gradient of the magnetic field from the tip as measured at the spin location,  $\mu$  is the magnetic moment of the spin,  $k$  is the cantilever spring constant and  $x_{\text{pk}}$  is the peak amplitude of the cantilever vibration. The factor  $\cos\theta$  represents the normalized projection of the spin in the direction of the effective field.

There are several impediments to single spin detection. Firstly, because the force from a single spin is so tiny (a few attonewtons), the maximum cantilever frequency shift is only about one part per million for typical experimental parameters. This small frequency shift must be detected in the presence of the cantilever phase (or frequency) noise that originates from cantilever thermal vibrations and sample-induced force fluctuations. The resulting low SNR necessitates long integration times for signal detection. Secondly, the detection is complicated by environmental disturbance to the spin (i.e., relaxation effects) that can randomly flip the spin orientation and reverse the signal polarity during the signal integration time. A low operating temperature, on the order of 1K, can help reduce both the cantilever thermal excitations and the random spin flip rate. Nevertheless, the signal processing methodology must take these effects into account.

We consider two models for the spin behaviour in the presence of environmental disturbances. In a quantum mechanical measurement model, the spin is always found to be either

aligned or anti-aligned with the effective field, so that  $\cos \theta = \pm 1$ . Thus the frequency shift signal has only two levels:  $\Delta\omega = \pm\Delta\omega_{\max}$ , and the time sequence of the frequency shift is a random telegraph signal with a transition rate that depends on the spin relaxation rate. In a classical measurement model (which we consider for the sake of completeness),  $\cos \theta$  can take arbitrary values between +1 and -1. As a result of environmental disturbance,  $\Delta\omega$  will be uniformly distributed between  $-\Delta\omega_{\max}$  and  $+\Delta\omega_{\max}$ . For this case, a bounded random walk model is appropriate. Recent results strongly favour the random telegraph model [18]. Other publications of interest include [19], [20].

### III. MRFM SIGNAL MODELS

#### A. Model 1: Discrete-time Random Telegraph model

In the quantum measurement model, the frequency shift is characterized by random transitions between two discrete levels. The transition times are taken to be Poisson distributed [18]. Denote the DT random telegraph signal by  $\zeta_i$ , where  $t_i = iT_s$  are the sampling times, and  $T_s$  is the sampling time interval. In this paper, a Markovian process with a finite number of states will have a state space denoted by  $\Psi = \{\psi_1, \dots, \psi_d\}$ , where  $d$  is the number of states. Let the state space of the DT random telegraph be  $\Psi_{\pi}$ ; it has  $d = 2$  states and we shall take  $\psi_1 = -A$ ,  $\psi_2 = A$ , where  $A$  is the amplitude of the random telegraph ( $A$  corresponds to  $\Delta\omega_{\max}$  for the case of a MRFM signal). As an initial condition,  $\zeta_0$  is equally likely to be either  $\pm A$ . Then, a probability transition matrix  $\mathbf{P}_{\pi}$  can be associated with  $\zeta_i$  such that the  $(j, k)$ -th value of  $\mathbf{P}_{\pi}$  equals  $P(\zeta_i = \psi_k | \zeta_{i-1} = \psi_j)$  for  $1 \leq j, k \leq 2$  and  $i \geq 1$ . Assume that  $\mathbf{P}_{\pi}$  has the form:

$$\mathbf{P}_{\pi} = \begin{pmatrix} q & 1-q \\ 1-p & p \end{pmatrix}, \quad (2)$$

where  $0 < p, q < 1$ . If  $p = q$ , we say that the transition probabilities are *symmetric*, whereas if  $p \neq q$ , we shall say that they are *asymmetric*. Define the signal vector  $\underline{\zeta} = [\zeta_0, \dots, \zeta_{N-1}]^T$ , the noise vector  $\underline{w} = [w_0, \dots, w_{N-1}]^T$ , and the observation vector  $\underline{y} = [y_0, \dots, y_{N-1}]^T$ , where the superscript  $(\cdot)^T$  denotes the transpose operator, and  $N$  is the number of observations. The  $w_i$ 's are modelled as independent and identically distributed (i.i.d.) Gaussian random variables (r.v.s) with zero mean and variance  $\sigma^2$ . The detection problem is then to decide between:

$$\begin{aligned} H_0 \text{ (spin absent)} &: \underline{y} = \underline{w} \\ H_1 \text{ (spin present)} &: \underline{y} = \underline{\zeta} + \underline{w} \end{aligned} \quad (3)$$

Before proceeding further, let us define the SNR of a finite state DT Markov process  $\zeta_i$ . Let  $f_i$  be the density of  $y$  induced under hypothesis  $H_i$  for  $i = 0, 1$ . Similarly, let  $E_i[\cdot]$  denote the expectation under hypothesis  $H_i$  for  $i = 0, 1$ . We shall use [21, (3)], which is motivated by the error exponent of the optimum detector in a binary hypothesis test, with  $\alpha = 1/2$ . The result [21, (3)] assumes that  $f_0$  and  $f_1$  are Gaussian distributed. While  $f_0$  in this paper does have a Gaussian distribution, the same is not true for density  $f_1$ . Nevertheless, we shall approximate  $f_1$  with a Gaussian density. Let  $\underline{p}_{ss} \in \mathbb{R}^d$  denote

the vector of steady-state probabilities of the random process  $\zeta_i$ , i.e. it satisfies  $\underline{p}_{ss}^T = \underline{p}_{ss}^T \mathbf{P}$ . For a vector  $\underline{x} \in \mathbb{R}^d$ ,  $\text{diag}(\underline{x})$  is defined to be the  $d$ -by- $d$  matrix with  $\underline{x}$  along its main diagonal. Another approximation that we shall make is that the covariance matrix of  $\underline{y}$  under hypothesis  $H_1$  is approximately  $\sigma_1^2 \mathbf{I}$ , where  $\sigma_1^2 = \sigma^2 + \underline{p}_{ss}^T \mathbf{D}_{\zeta}^2 \underline{1}$ , and  $\mathbf{D}_{\zeta} \triangleq \text{diag}(\psi_1, \dots, \psi_d)$ . Lastly, we shall make the approximation that  $E_1[\zeta_i] \approx \delta\zeta \triangleq \underline{p}_{ss}^T \mathbf{D}_{\zeta} \underline{1}$  for  $0 \leq i < N$ . This is reasonable if  $N$  is large. See [22] for more details. Then, the SNR can be expressed as

$$\text{SNR} = N \left\{ -\frac{1}{2} \log \frac{2\sigma\sigma_1}{\sigma^2 + \sigma_1^2} + \frac{1}{4(\sigma^2 + \sigma_1^2)} (\delta\zeta)^2 \right\} \quad (4)$$

The SNR in dB is defined in the usual way as  $\text{SNR}_{\text{dB}} \triangleq 10 \log_{10} \text{SNR}$ .

The paper [5] uses another definition of SNR. Let  $\bar{\zeta}_i \triangleq \zeta_i - E_1[\zeta_i]$ ,  $i \geq 0$  be the mean-corrected version of  $\zeta_i$ . The -3dB bandwidth of the random process  $\bar{\zeta}_i$  is

$$W = \arccos \left( \frac{4r - 1 - r^2}{2r} \right) \quad (5)$$

where  $r \triangleq p + q - 1$ . Note that  $|r| < 1$ . In the symmetric case when  $p = q$ , and  $p \approx 1$ , (5) is approximately  $2\pi(1 - p)$ . The -3dB bandwidth is then proportional to the mean number of transitions per second, which is  $(1 - p)/T_s$ . The definition of SNR used in [5] is

$$\text{SNR}' = \frac{\text{Power of random process } \bar{\zeta}_i \text{ in } [-W, W]}{\text{Power of noise in } [-W, W]} \quad (6)$$

Under definition (6), the SNR in the single electron spin experiment was reported to be -6.7 dB [5]. In this paper, we shall use the SNR definition of (4). The condition of low SNR for the DT random telegraph will be taken to mean  $|A/\sigma| \ll 1$ .

Examples of noiseless and noisy random telegraph signals are given in Fig. 2.

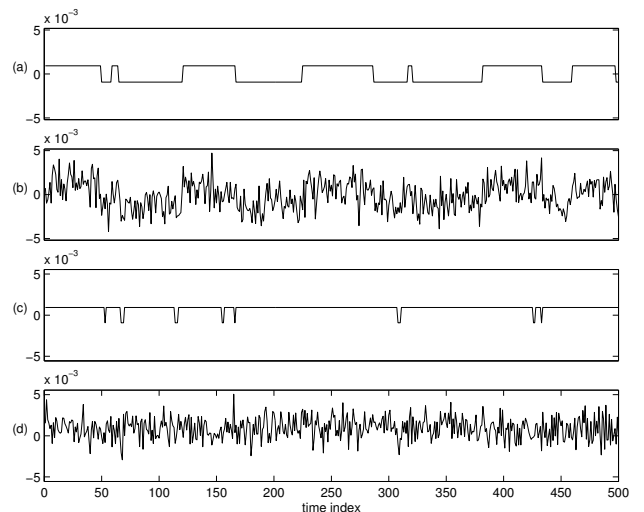


Fig. 2. (a) Noiseless random telegraph signal with symmetric transition probabilities  $p = q = 0.98$ . (b) Noisy version of (a) at SNR = 7.09 dB. (c) Noiseless random telegraph signal with asymmetric transition probabilities  $p = 0.98, q = 0.6$ . (d) Noisy version of (c) at SNR = 14.1 dB.

### B. Model 2: Discrete-time Random Walk model

In the classical spin detection model, the frequency shift signal is well approximated by a one dimensional random walk confined to the interval  $I = [-A, A]$ , where  $A = \Delta\omega_{\max}$  for the case of a MRFM signal. We discretize  $I$  into  $(2M + 1)$  states using a step size of  $s$ , where  $M \in \mathbb{N}$  and  $s > 0$  and define  $\zeta_i$  to be the random walk restricted to the discretized  $I$ ; we shall refer to this model as the DT random walk model. The state space  $\Psi_{\text{rw}}$  of the DT random walk will then have  $d = 2M + 1$  states, where  $\psi_j = (j - M - 1)s$  for  $j = 1, \dots, (2M + 1)$ . Associate with  $\zeta_i$  the probability transition matrix  $\mathbf{P}_{\text{rw}}$ , so that, as before, the  $(j, k)$ -th element of  $\mathbf{P}_{\text{rw}}$  is  $P(\zeta_i = \psi_k | \zeta_{i-1} = \psi_j)$  for  $1 \leq j, k \leq (2M + 1)$  and  $i \geq 1$ .  $\mathbf{P}_{\text{rw}}$  is defined such that, at each time step,  $\zeta_i$  changes by either  $\pm s$ . This implies that  $\mathbf{P}_{\text{rw}}$  is a tridiagonal matrix. We assume reflecting boundary conditions, and  $\zeta_0$  is equally likely to be either  $\pm s$ . The initial condition on  $\zeta_0$  was arbitrarily chosen. The regime of interest that we will focus on is large  $N$ , and so the effect of the initial condition will not be significant.

The detection problem is now to test (3) when  $\underline{\zeta}$  is modelled by a random walk. Note that the DT random walk model can almost be regarded as a multi-state generalization of the DT random telegraph model. There are, however, important differences. The DT random walk process cannot remain in the same state for two consecutive time instances. In contrast, it is possible for the DT random telegraph process to do so. Additionally, the DT random walk process has reflecting boundary conditions. The DT random telegraph process does not have this. In the limit as  $s \rightarrow 0, M \rightarrow \infty$ , the random walk converges to Brownian motion over the interval  $I$  [23].

Analogous to the DT random telegraph, the condition of low SNR for the DT random walk will be taken to mean  $|\psi_i/\sigma| \ll 1$  for  $i = 1, \dots, d$ . Examples of noiseless and noisy random walk signals are given in Figs. 3 and 4, where, at each state, a change of  $\pm s$  is equally likely.

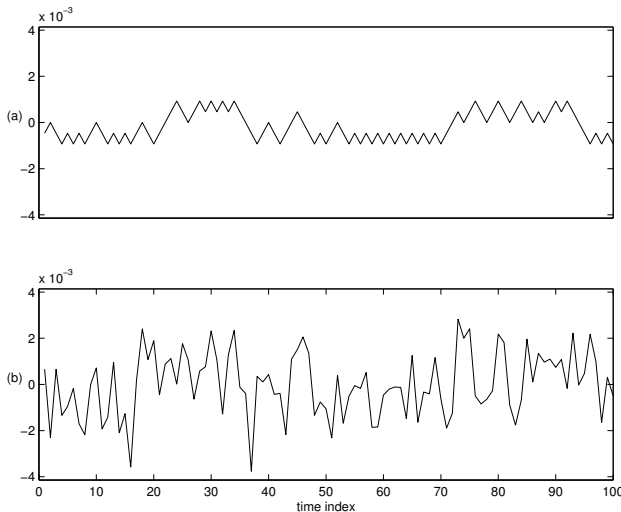


Fig. 3. (a) Noiseless random walk signal with 5 levels. (b) Noisy version of (a) at SNR = -7.33 dB.

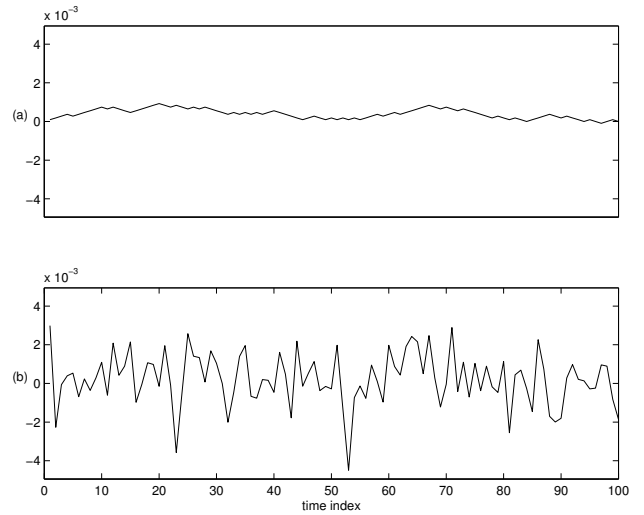


Fig. 4. (a) Noiseless random walk signal with 21 levels. (b) Noisy version of (a) at SNR = -8.23 dB.

### IV. DETECTION STRATEGIES

The detectors considered here can be placed into three categories: versions of existing detectors that are currently in use for MRFM; LRTs; and approximations to the LRT. The LRT is a most powerful test that satisfies the Neyman-Pearson criterion: it maximizes the probability of detection ( $P_D$ ) subject to a constraint on the probability of false alarm ( $P_F$ ) [24], which is set by the user. Consequently, it can be used as a benchmark with which to compare the other detectors. When the initial state value, the random transition times, and all subsequent state values are known, the optimal LRT is the matched filter, called the omniscient matched filter (MF) in this paper. Although unimplementable, the MF detector provides an absolute upper bound when comparing the various detectors' Receiver Operating Characteristic (ROC) curves.

#### A. Amplitude, energy, filtered energy detectors

The DT amplitude detector is

$$\left| \frac{1}{N} \sum_{i=0}^{N-1} y_i \right| \underset{H_0}{\overset{H_1}{\gtrless}} \eta \quad (7)$$

where  $\eta$  is set to satisfy the constraint on  $P_F$ . This is the optimal test under the assumption that  $y_i$  is the sum of an unknown constant and AWGN. This assumption would be true if there were no random spin flips. However, as the number of random transitions in  $y_i$  increases, the performance of the amplitude detector degrades. An intuitive explanation can be obtained by considering the detection of the DT random telegraph process and omitting the absolute value bars in the amplitude detector. Under  $H_1$ ,  $\sum_i y_i = \sum_i \zeta_i + \sum_i w_i$ . If  $\zeta_i = A$  for  $i = 0, \dots, N - 1$ ,  $E_1[\sum_i y_i] = NA$ . In contrast,  $E_0[\sum_i y_i] = 0$ . However, if  $\zeta_i$  is equally likely to be  $\pm A$ , then  $E_1[\sum_i y_i] = 0$ , which is the same value as  $E_0[\sum_i y_i]$ . In this case, it would be harder to distinguish between the two hypotheses, as the test statistic has the same expected value under both hypotheses.

An alternative test statistic is the DT signal energy, i.e. the sum of the squares of the  $y_i$  instead of the magnitude of the sum in (7). As the signal and noise are assumed to be independent, under hypothesis  $H_1$ , one would expect  $\underline{y}$  to have a higher energy on average than under hypothesis  $H_0$ . This can be reliably detected under a sufficiently high SNR. A natural improvement to the energy detector is to reject out-of-band noise by prefiltering  $\underline{y}$  over the signal passband. If the signal  $\zeta_i$  is baseband, which is the case for the DT random telegraph and random walk process, a lowpass filter (LPF) is appropriate. In particular, one might use a simple first-order, single-pole filter given by

$$H_{\text{LP}}(z) = \frac{1 - \alpha_{\text{LP}}}{2} \frac{1 + z^{-1}}{1 - \alpha_{\text{LP}}z^{-1}} \quad (8)$$

where we require  $|\alpha_{\text{LP}}| < 1$  for stability [25]. The time constant  $\alpha_{\text{LP}}$  dictates the bandwidth of the LPF. If  $\omega_c$  is the desired -3dB bandwidth of the filter, one should set

$$\alpha_{\text{LP}} = \frac{1 - \sin \omega_c}{\cos \omega_c} \quad (9)$$

The -3dB bandwidth used depends on the bandwidth of the random process  $\zeta_i$ . For example, the -3dB bandwidth of the DT random telegraph is given by (5). Suppose we have symmetric transition probabilities. The CT random telegraph model is typically characterized by the rate parameter  $\lambda$ , which corresponds to the mean number of transitions per second. One can equate the mean number of transitions per second in both DT and CT models to obtain

$$\lambda = \frac{1 - p}{T_s} \quad (10)$$

Since  $p > 0$ , we require  $T_s < \lambda^{-1}$  in order to use (10). In practice,  $p$  (or equivalently  $\lambda$ ) is only approximately known to the experimenter. As a result, a bank of LPFs with different  $\alpha_{\text{LP}}$ 's are used to perform detection [5].

Let “ $*$ ” be the convolution operator, so that  $s = y * h$  is defined to be  $s_i \triangleq \sum_k y_k h_{i-k}$ . The energy and filtered energy detector can be expressed as

$$\sum_{i=0}^{N-1} (y * h)_i^2 \underset{H_0}{\underset{H_1}{\geq}} \eta \quad (11)$$

where  $(y * h)_i$  is taken to be  $i$ -th value of  $y * h$ . For the energy detector,  $h$  is taken to be the unit impulse function  $\delta[i]$ , while for the filtered energy detector,  $h = h_{\text{LP}}$ , the impulse response of  $H_{\text{LP}}(z)$  in (8).

Note that the computational complexity for the amplitude, filtered energy, and energy detectors is  $\mathcal{O}(N)$ .

### B. Recursive equations for the optimal LRT detector of a general finite state DT Markov signal in AWGN

In this section, we shall consider the detection of a general finite state DT Markov process in AWGN and derive the LRT. The formulas that provide an initial starting point are given in [8]. We shall use the notation in [8]: while it is slightly different, the differences are superficial.

The hypothesis test that we consider is (3). Let the state space of  $\zeta_i$  be denoted by  $\Psi = \{\psi_1, \dots, \psi_d\}$ , where there are

$d$  possible states. Let  $\mathbf{P}^{(i)}, i \geq 1$  be the probability transition matrix associated with the process  $\zeta_i$  at the  $i$ -th time step, so that  $\mathbf{P}_{jk}^{(i)} = P(X_i = \psi_k | X_{i-1} = \psi_j)$ . The noise is denoted by  $w_i$ , and are independent Gaussian r.v.s with mean zero and variance  $R_i \in \mathbb{R}, i = 0, \dots, N - 1$ .

Let  $\mathbf{y}^i \triangleq [y_0, \dots, y_i]^T$  for  $i \geq 0$ . We shall define  $\underline{p}_i, \underline{r}_i \in \mathbb{R}^d, i \geq 0$  and  $\underline{q}_i \in \mathbb{R}^d, i \geq 1$  as

$$\begin{aligned} \underline{p}_i &\triangleq [P(\zeta_i = \psi_1), \dots, P(\zeta_i = \psi_d)]^T \\ \underline{r}_i &\triangleq [P(\zeta_i = \psi_1 | \mathbf{y}^i), \dots, P(\zeta_i = \psi_d | \mathbf{y}^i)]^T \\ \underline{q}_i &\triangleq [P(\zeta_i = \psi_1 | \mathbf{y}^{i-1}), \dots, P(\zeta_i = \psi_d | \mathbf{y}^{i-1})]^T \end{aligned} \quad (12)$$

Define  $\mathbf{\Omega}^{(i)} \in \mathbb{R}^{d \times d}, i \geq 1$  as

$$\mathbf{\Omega}_{jk}^{(i)} \triangleq \begin{cases} \frac{f_1(y_i | \zeta_i = \psi_k, \mathbf{y}^{i-1})}{f_1(y_i | \mathbf{y}^{i-1})} & j = k \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

*Proposition 1:*  $\underline{p}_i^T = \underline{p}_{i-1}^T \mathbf{P}^{(i)}, i \geq 1$   
Examine the  $j$ -th element of  $\underline{p}_i$ :

$$\begin{aligned} P(\zeta_i = \psi_j) &= \sum_{n=1}^d P(\zeta_i = \psi_j | \zeta_{i-1} = \psi_n) P(\zeta_{i-1} = \psi_n) \\ &= \underline{p}_{i-1}^T (\mathbf{j}\text{-th column of } \mathbf{P}^{(i)}) \quad \blacksquare \end{aligned}$$

*Proposition 2:*  $\underline{q}_i^T = \underline{r}_{i-1}^T \mathbf{P}^{(i)}, i \geq 1$   
Examine the  $j$ -th element of  $\underline{q}_i$ . By the Markov assumption,

$$\begin{aligned} P(\zeta_i = \psi_j | \mathbf{y}^{i-1}) &= \sum_{n=1}^d P(\zeta_i = \psi_j | \zeta_{i-1} = \psi_n, \mathbf{y}^{i-1}) \cdot \\ &\quad P(\zeta_{i-1} = \psi_n | \mathbf{y}^{i-1}) \\ &= \underline{r}_{i-1}^T \begin{pmatrix} P(\zeta_i = \psi_j | \zeta_{i-1} = \psi_1) \\ \vdots \\ P(\zeta_i = \psi_j | \zeta_{i-1} = \psi_d) \end{pmatrix} \\ &= \underline{r}_{i-1}^T (\mathbf{j}\text{-th column of } \mathbf{P}^{(i)}) \quad \blacksquare \end{aligned}$$

*Proposition 3:*  $\underline{r}_i^T = \underline{q}_i^T \mathbf{\Omega}^{(i)}, i \geq 0$   
Using [8, (39)]:

$$P(\zeta_i = \psi_j | \mathbf{y}^i) = \frac{f_1(y_i | \zeta_i = \psi_j, \mathbf{y}^{i-1}) P(\zeta_i = \psi_j | \mathbf{y}^{i-1})}{f_1(y_i | \mathbf{y}^{i-1})} \quad \blacksquare$$

Propositions 1 and 3 are derived in [8]. From Propositions 2 and 3,

$$\underline{q}_i^T = \underline{q}_{i-1}^T \mathbf{\Omega}^{(i-1)} \mathbf{P}^{(i)} \quad (14)$$

This result is incorrectly given in [8, (38)].

Define  $\varphi(x; \mu, \sigma^2) \triangleq \exp[-(x - \mu)^2 / 2\sigma^2] / \sqrt{2\pi}\sigma$ , i.e., a Gaussian r.v. with mean  $\mu$  and variance  $\sigma^2$ . Let  $\pi_k, k = 1, \dots, d$  denote the initial probability density of  $\zeta_0$ , and  $\underline{\pi} \triangleq [\pi_1, \dots, \pi_d]^T$ . Define

$$\underline{n}_i = [\varphi(y_i; \psi_1, R_i), \dots, \varphi(y_i; \psi_d, R_i)]^T \quad (15)$$

Since  $\mathbf{\Omega}^{(i)} = \text{diag}(\underline{n}_i) / (\underline{q}_{i-1}^T \underline{n}_i)$ , (14) can be written as

$$\underline{q}_i^T = \underline{q}_{i-1}^T \frac{\text{diag}(\underline{n}_{i-1})}{\underline{q}_{i-1}^T \underline{n}_{i-1}} \mathbf{P}^{(i)} \quad (16)$$

If we define  $\underline{q}_0 \triangleq \underline{\pi}$ , the log LRT for (3) is given by

$$\log \Lambda(\underline{y}) = \sum_{i=0}^{N-1} \log \frac{\underline{q}_i^T \underline{n}_i}{\varphi(\underline{y}_i; 0, R_i)} \quad (17)$$

Equations (16) and (17) are a recursive way to compute the LRT for a general finite state DT Markov signal in AWGN. See [22] for more details; note that this result does not appear in [8]. We see that the running time of (16) and (17) is of order  $\mathcal{O}(Nd^2)$  for general matrices  $\mathbf{P}^{(i)}$ . If each  $\mathbf{P}^{(i)}$  were tridiagonal, for example, the running time would be  $\mathcal{O}(Nd)$ .

The detection test would then be

$$\Lambda(\underline{y}) \underset{H_0}{\overset{H_1}{\geq}} \eta. \quad (18)$$

One can also take the log of both sides of (18). As the log function is strictly monotone increasing, the threshold would be altered, but the performance of the detector would be unaffected. The threshold  $\eta$  can be determined via simulation, if all of the parameters of the signal and noise models were known. The ROC curve could be generated, and the value of  $\eta$  that corresponded to the desired  $P_F$  could be obtained. Alternatively, the ROC curve could be generated via experimentation, and the threshold  $\eta$  selected in a similar fashion.

Define the transition likelihood ratio  $l(y_i|y^{i-1})$ ,  $i \geq 1$  as  $l(y_i|y^{i-1}) \triangleq f_1(y_i|y^{i-1})/f_0(y_i|y^{i-1})$ . Note that

$$\log \Lambda(\underline{y}) = \sum_{i=1}^{N-1} \log l(y_i|y^{i-1}) + \log \left( \frac{f_1(y_0)}{f_0(y_0)} \right) \quad (19)$$

In [9],  $\log l(y_i|y^{i-1})$ ,  $i \geq 1$  is given when  $P(y_i|y^{i-1}, \zeta^i)$  belongs to a class of exponential functions of the form  $K(\zeta^i) \exp[\zeta_i g(y^i) + B(y^i)]$  for functions  $K(\cdot)$ ,  $g(\cdot)$ ,  $B(\cdot)$  which give rise to a valid density. The expression for  $\log l(y_i|y^{i-1})$ , involves the conditional mean estimate

$$\hat{\zeta}_i(y^i) = \int_{\Psi} \zeta_i dP(\zeta_i|y^i) \quad (20)$$

and the function

$$G(y^i) = \int \hat{\zeta}_i(\xi^i) d\xi_i \Big|_{\xi^i=y^i} \quad (21)$$

For the hypothesis testing problem considered here

$$\hat{\zeta}_i(y^i) = \underline{r}_i^T \mathbf{D}_\zeta \underline{1} \quad (22)$$

Recall that  $\mathbf{D}_\zeta = \text{diag}(\psi_1, \dots, \psi_d)$ . It is possible to apply Propositions 2 and 3 to obtain a recursive equation for  $\underline{r}_i$  and then solve for a closed-form expression. It would be difficult, however, to evaluate (21).

### C. Approximation of the LRT under low SNR

Let us consider the log LRT under low SNR. Each transition likelihood ratio can be simplified as follows:

$$\begin{aligned} l(y_i|y^{i-1}) &= \sum_{n=1}^d P(\zeta_i = \psi_n|y^{i-1}) \\ &\quad \exp \left[ -\frac{1}{2R_i} (-2y_i\psi_n + \psi_n^2) \right] \\ &\approx \sum_{n=1}^d P(\zeta_i = \psi_n|y^{i-1}) \left( 1 + \frac{1}{R_i} y_i\psi_n - \frac{1}{2R_i} \psi_n^2 \right) \\ &= 1 + \frac{1}{R_i} y_i E_1[\zeta_i|y^{i-1}] - \frac{1}{2R_i} E_1[\zeta_i^2|y^{i-1}] \quad (23) \end{aligned}$$

where the approximation  $e^\delta \approx 1 + \delta$  for small  $\delta$  was used. Next, we shall use the approximation  $\log(1 + \delta) \approx \delta$  for small  $\delta$ . This is justified if the SNR is low so that  $|\psi_n/\sqrt{R_i}| \ll 1$  for all  $n = 1, \dots, d$ . So

$$\log l(y_i|y^{i-1}) \approx \frac{1}{R_i} y_i E_1[\zeta_i|y^{i-1}] - \frac{1}{2R_i} E_1[\zeta_i^2|y^{i-1}] \quad (24)$$

As well, the same approximation can be applied to  $\log[f_1(y_0)/f_0(y_0)]$ , so that

$$\log \left( \frac{f_1(y_0)}{f_0(y_0)} \right) \approx \frac{1}{R_0} y_0 E_1[\zeta_0] - \frac{1}{2R_0} E_1[\zeta_0^2]. \quad (25)$$

Define the conditional mean (also MMSE) estimator of  $\zeta_i$  under  $H_1$  as follows:  $\hat{\zeta}_i = E_1[\zeta_i|y^{i-1}]$  for  $i \geq 1$  and  $\hat{\zeta}_0 = E_1[\zeta_0]$ . Use a similar notation for  $\zeta_i^2$ , so that  $\hat{\zeta}_0^2 = E_1[\zeta_0^2]$  and  $\hat{\zeta}_i^2 = E_1[\zeta_i^2|y^{i-1}]$  for  $i \geq 1$ . Using (19), (24), and (25), the log LRT can be approximately written under low SNR as

$$\log \Lambda(\underline{y}) \approx \sum_{i=0}^{N-1} \frac{1}{R_i} y_i \hat{\zeta}_i - \frac{1}{2} \sum_{i=0}^{N-1} \frac{1}{R_i} \hat{\zeta}_i^2 \quad (26)$$

The right hand side of (26) is similar to the matched filter statistic, but with the MMSE estimates of  $\zeta_i$  and  $\zeta_i^2$  used instead of the known values. Note that the conditional mean estimator of  $\zeta_i$  used in (26) is different from that defined in (20). Schwartz's version in (20) includes the observation  $y_i$ , whereas ours does not.

In [8], it is shown that in detecting a finite-state DT Markov signal, the LRT is in general not expressible as the known form LRT with an estimator of  $\zeta_i$  used, i.e., the RHS of (26) but with  $(E_1[\zeta_i|y^{i-1}])^2$  used instead of  $E_1[\zeta_i^2|y^{i-1}]$  for  $i \geq 1$ .

### D. Comparison to the CT analog

Consider the CT analog of the hypothesis test problem (3) when the noise variances  $R_i$  are all equal. One has to decide between the following two hypotheses:

$$\begin{aligned} H_0 : \quad & \dot{y}(t) = \dot{w}(t), \quad t \in [0, T] \\ H_1 : \quad & \dot{y}(t) = \zeta(t) + \dot{w}(t), \quad t \in [0, T]. \quad (27) \end{aligned}$$

Define  $I \triangleq [0, T]$ . Here,  $\zeta(t)$  is a random process (not necessarily finite state) such that  $\int_I E[\zeta^2(t)] dt < \infty$  and  $\dot{w}(t)$  is AWGN with

$$E[\dot{w}(t)] = 0, \quad E[\dot{w}(t)\dot{w}(s)] = \sigma^2 \delta(t-s) \quad (28)$$

The LRT is given by [10], [11], [12]:

$$\exp\left(\frac{1}{\sigma^2} \int_I \hat{\zeta}_1(t)y(t)dt - \frac{1}{2\sigma^2} \int_I \hat{\zeta}_1^2(t)dt\right) \quad (29)$$

where  $\hat{\zeta}_1(t) = E_1[\zeta(t)|\{y(\xi), \xi < t\}]$  is the conditional mean estimate under hypothesis  $H_1$  given the previous observations, and the first integral in (29) is an Itô stochastic integral.

There are three noteworthy differences between (26) and (29). Firstly, result (26) approximately holds only under low SNR, whereas (29) is exact under all SNR conditions. Secondly,  $\zeta(t)$  is not constrained to be a Markov process, whereas its DT counterpart  $\zeta_i$  is Markovian. Lastly, in the second term of (26), the expected value of  $\zeta_i^2$  conditioned on the past observations is used. On the other hand, the square of the expected value of  $\zeta(t)$  conditioned on the past observations is used in the CT version. In general, for  $i \geq 1$ ,  $E_1[\zeta_i^2|y^{i-1}] \neq (E_1[\zeta_i|y^{i-1}])^2$ . Indeed, for a r.v.  $X$ ,

$$E[X^2] = (E[X])^2 \text{ iff } \text{var}(X) = 0, \quad (30)$$

By the Chebyshev inequality, for  $\delta > 0$ ,  $P[|X - E[X]| \geq \delta] \leq \text{var}(X)/\delta^2 = 0$ . So (30) holds iff  $X$  is some value  $c \in \mathbb{R}$  w.p. 1. As a result  $E_1[\zeta_i^2|y^{i-1}] = (E_1[\zeta_i|y^{i-1}])^2$  iff  $\zeta_i$  is a function of  $y^{i-1}$  w.p. 1.

#### E. Application to the detection of the DT random telegraph process

Under the regime of low SNR, long observation time ( $N \gg 1$ ), and  $r = p+q-1 \approx 1$  (the probability of transition between consecutive samples is small), the second-order expansion of  $\log \Lambda_\pi(\underline{y})$  is approximately equal to the hybrid detector with the test statistic

$$\sum_i (y * h_{\text{LP}})_i^2 + K_a \sum_i y_i + K_e \sum_i y_i^2, \quad (31)$$

where the constants  $K_a = K_a(p, q, A, \sigma)$  and  $K_e = K_e(p, q)$  are given in (63) of the appendix, and  $\alpha_{\text{LP}} = p + q - 1$ . Therefore, in the aforementioned regime, one expects the hybrid detector to have performance similar to the optimal LRT. When  $p = q$ , the second-order expansion of the LRT is approximately equal to the filtered energy detector, which is given by (11) with  $h = h_{\text{LP}}$  and  $\alpha_{\text{LP}} = 2p - 1$ . See the appendix for more details.

The previous result did not make use of the approximation under low SNR obtained in Subsection IV-C. Equation (26), however, provides a general approach to deriving an approximation to the LRT for a finite state DT Markov signal in AWGN. We shall illustrate by deriving an approximation to the LRT for the DT random telegraph using this general approach.

Let us specialize the result of Subsection IV-C to the DT random telegraph processes, so that  $R_i = \sigma^2$ ,  $i = 0, \dots, N-1$  and  $\mathbf{P}^{(i)} = \mathbf{P}_\pi$ ,  $i = 1, \dots, N-1$ . Since  $E_1[\zeta_0^2] = A^2$  and  $E_1[\zeta_i^2|y^{i-1}] = A^2$  for  $i \geq 1$ , the second term of (26) will be a constant and can be omitted. Then,

$$\log \Lambda_\pi(\underline{y}) \approx \frac{1}{\sigma^2} \sum_{i=0}^{N-1} y_i \hat{\zeta}_i \quad (32)$$

under low SNR conditions. We see that the log LRT for the random telegraph is an estimator-correlator detector. The estimator-correlator structure is known to be optimal for detecting Gaussian signals in AWGN [9].

Now, the conditional mean estimator  $\hat{\zeta}_i = E_1[\zeta_i|y^{i-1}]$ ,  $i \geq 1$  is a function of  $y_0, \dots, y_{i-1}$ . Its exact form will be dictated by the conditional probability mass function (p.m.f.)  $P(\zeta_i|y^{i-1})$ , and will be nonlinear in general. Suppose that we would like to find the estimator of  $\zeta_i$  with smallest MSE that lies in the linear span of  $\{1, y_0, \dots, y_{i-1}\}$ . Let  $\hat{\zeta}_i^A = c + \sum_{n=0}^{i-1} \gamma_n y_n$  be this estimator. It is also known as the best affine estimator of  $\zeta_i$  given  $y_0, \dots, y_{i-1}$ . One can apply the Projection Theorem [26] to obtain the well-known result that

$$\hat{\zeta}_i^A = E[\zeta_i] + \text{cov}_1(y^{i-1}, \zeta_i) \mathbf{\Gamma}_{i-1}^{-1} (y^{i-1} - E_1[y^{i-1}]) \quad (33)$$

where  $\text{cov}_1(a, b) = E_1[ab^T] - E_1[a]E_1[b]^T$  is the covariance of the random vectors  $a, b$  assuming hypothesis  $H_1$  and  $\mathbf{\Gamma}_{i-1} \triangleq \text{cov}_1(y^{i-1}, y^{i-1})$  is the covariance matrix of  $y^{i-1}$  under hypothesis  $H_1$ . The MSE achieved with  $\hat{\zeta}_i^A$  cannot be smaller than the MSE achieved with  $\hat{\zeta}_i = E_1[\zeta_i|y^{i-1}]$ , as the conditional mean achieves the smallest MSE out of all possible estimators.

For  $0 \leq n < i$ ,

$$\begin{aligned} \text{cov}_1(y_n, \zeta_i) &= E[(\zeta_n + w_n)\zeta_i] - E[\zeta_n + w_n]E[\zeta_i] \\ &= \text{cov}(\zeta_n, \zeta_i) \end{aligned} \quad (34)$$

since  $w_n$  is independent of  $\zeta_i$ . As well,  $E_1[y^{i-1}] = E[\zeta^{i-1} + w^{i-1}] = E[\zeta^{i-1}]$ , and

$$\mathbf{\Gamma}_{i-1} = \text{cov}_1(y^{i-1}, y^{i-1}) = \sigma^2 \mathbf{I} + \text{cov}(\zeta^{i-1}, \zeta^{i-1}) \quad (35)$$

Substituting these into (33) results in

$$\hat{\zeta}_i^A = E[\zeta_i] + \text{cov}(\zeta^{i-1}, \zeta_i) [\sigma^2 \mathbf{I} + \text{cov}(\zeta^{i-1}, \zeta^{i-1})]^{-1} \cdot (y^{i-1} - E[\zeta^{i-1}]) \quad (36)$$

It remains to compute the various quantities in (36). Recalling that  $r = p + q - 1$  and the definition of  $C_m$  in (56) of the appendix,

$$\begin{aligned} E[\zeta_i] &= AC_m(1 - r^i) \\ \text{cov}(\zeta_i, \zeta_j) &= 4A^2 \frac{(1-p)(1-q)}{(1-r)^2} r^{|i-j|} + \\ &\quad A^2 C_m^2 (2r^{\max(i,j)} - r^{i+j}) \end{aligned} \quad (38)$$

Using (37) and (38) in (36), and plugging the resulting expression of  $\hat{\zeta}_i^A$  into (32), one obtains

$$\log \Lambda_\pi(\underline{y}) \approx \frac{1}{\sigma^2} \sum_{i=0}^{N-1} y_i \hat{\zeta}_i^A \quad (39)$$

Unfortunately, (39) is not in a form that can be compared to the approximation derived previously without the benefit of (26). In order to do this, we have to explicitly evaluate (36). As the inversion of a matrix without any special structure is required, this will be difficult. Make the approximation that

$$\text{cov}(\zeta_i, \zeta_j) \approx 4A^2 \frac{(1-p)(1-q)}{(1-r)^2} r^{|i-j|} \quad (40)$$

This will be a good approximation when  $\max(i, j)$  is large. Moreover, since we are assuming a low SNR condition,  $|A/\sigma| \ll 1$ , so

$$\sigma^2 \mathbf{I} + \text{cov}(\zeta^{i-1}, \zeta^{i-1}) \approx \sigma^2 \mathbf{I} \quad (41)$$

With (40) and (41),

$$\hat{\zeta}_i^A \approx AC_m(1 - r^i) + C_f(r^i, \dots, r^1) \cdot (y^{i-1} - E[\zeta^{i-1}]) \quad (42)$$

$$\approx AC_m[1 - r^i - C_f \sum_{n=0}^{i-1} (r^{i-n} - r^i)] + C_f \sum_{n=0}^{i-1} r^{i-n} y_n \quad (43)$$

$$\approx AC_m \left[ 1 - r^i - C_f \left( \frac{r(1 - r^{i+1})}{1 - r} - ir^i \right) \right] + C_f \sum_{n=0}^{i-1} r^{i-n} y_n \quad (44)$$

where  $C_f \triangleq 4(1-p)(1-q)A^2/[\sigma^2(1-r)^2]$ .

As  $i \rightarrow \infty$ ,  $ir^i \rightarrow 0$  as  $|r| < 1$ , and  $(1 - r^{i+1})/(1 - r) \rightarrow 1/(1 - r)$ . Apply these to (44); then using the subsequent result in (39),

$$\log \Lambda_{\pi}(\underline{y}) \approx L_1 + L_2 \quad (45)$$

$$\text{where } L_1 = \frac{A}{\sigma^2} C_m \sum_{i=0}^{N-1} \left( 1 - r^i - C_f \frac{r}{1 - r} \right) y_i \quad (46)$$

$$L_2 = \frac{1}{\sigma^2} C_f \sum_{i=0}^{N-1} \sum_{n=0}^{i-1} r^{i-n} y_i y_n \quad (47)$$

The expressions (45)-(47) are similar to those in (52)-(54).

*F. General methodology for obtaining an approximation to the LRT of a finite state DT Markov signal in AWGN under low SNR*

In the previous subsection, the affine estimator of  $\zeta_i$ ,  $i \geq 1$  with the smallest MSE was used in place of the conditional mean  $E_1[\zeta_i | y^{i-1}]$ ,  $i \geq 1$  in order to obtain an approximation to the LRT of the DT random telegraph. This was obtained by finding the estimator of  $\zeta_i$  with smallest MSE that lied in the linear span of  $\{1, y_0, \dots, y_{i-1}\}$ . The best affine estimator of  $\zeta_i$  is necessarily suboptimal, as the conditional mean achieves the lowest MSE out of all estimators.

For the general case when the second term of (26) is present, one could also find a suboptimal estimator of  $\zeta_i^2$ ,  $i \geq 1$  in terms of the previous observations  $y_0, \dots, y_{i-1}$ . Following the idea with regard to  $\hat{\zeta}_i$ , a suboptimal estimator can be obtained by finding the estimator of  $\zeta_i^2$  with smallest MSE that lies in the linear span of  $\{1, y_0^2, \dots, y_{i-1}^2\}$ .

The suboptimal estimators of  $\zeta_i$  and  $\zeta_i^2$  can then be used in (26).

## V. SIMULATION RESULTS

The objective in this section is to compare the detection methods discussed in the previous section. The class of LRT detectors is optimal for their respective signal models, and

provides a good comparison benchmark. Comparison of the various detectors is done using: (1) ROC curves, each of which is a plot of  $P_D$  vs.  $P_F$ , and (2) power curves, each of which is a plot of  $P_D$  vs. SNR at a fixed  $P_F$ . Recall that  $P_D$  is the probability of detection and  $P_F$  is the probability of false alarm. To generate each ROC curve, 20 simulations were generated. Then, the average and the error bars of one standard deviation were plotted. In a similar fashion, the data that is plotted in each power curve is the average over the 20 simulations at each SNR value, along with error bars of one standard deviation.

Some of the parameters used in the simulation of the DT random telegraph and random walk models are as follows:  $k = 10^{-3} \text{ N m}^{-1}$ ,  $\omega_0 = 2\pi \times 10^4 \text{ rad s}^{-1}$ ,  $G = 2 \times 10^6 \text{ T m}^{-1}$ . The sampling period was  $T_s = 1 \text{ ms}$ , and signal durations of  $T = 60 \text{ s}$  and  $T = 150 \text{ s}$  were used. The amplitude of the rf field was  $B_1 = 0.2 \text{ mT}$ . The performance of the detectors varies as a function of  $T$ ; in general, a larger  $T$  results in better performance. Values of  $T$  used in iOSCAR MRFM experiments are on the order of tens of hours [5]. Nevertheless, the comparative results obtained from using the two values of  $T$  above are representative of larger values. Indeed, our approximations to the optimal detectors improve with increased  $T$ .

### A. Discrete-time random telegraph model

First, consider the DT random telegraph. Fig. 5 depicts the simulated ROC curves at SNR = -34.3 dB,  $\lambda = 0.5 \text{ s}^{-1}$ , and with symmetric transition probabilities ( $p = q$ ). With  $T_s = 1 \text{ ms}$ , this results in  $p = q = 0.9995$ . We examine the matched filter, DT random telegraph LRT (RT-LRT), filtered energy, hybrid (given by (31)), amplitude, and unfiltered energy detectors. The RT-LRT, filtered energy, and hybrid detector curves are virtually identical, which is consistent with our analysis. The unfiltered energy and amplitude detectors have performance that is poorer than the RT-LRT, as it should be since the RT-LRT is the optimal detector. The unfiltered energy detector has the worst performance out of the five detector methods considered. Lastly, the omniscient MF detector has the best performance.

A power curve was generated over a range of SNRs under the same conditions as before with a fixed  $P_F = 0.1$ ; it is illustrated in Fig. 6. For spin detection, an acceptable range for  $P_F$  is on the order of 0.05 to 0.1. The RT-LRT, filtered energy, and hybrid detector have similar performance from -25 dB to -55 dB. With this particular value of  $P_F$  and  $\lambda$ , the RT-LRT, filtered energy, and hybrid detector perform from 10 dB to 20 dB worse than the MF detector. Although the amplitude detector has worse performance than the RT-LRT and filtered energy detector, all three have comparable performance at -55 dB.

Fig. 7 shows the power curves generated using the bigger value of  $T = 150 \text{ s}$ . The RT-LRT, filtered energy, and hybrid detectors have the same performance from -20 dB to -50 dB. Note that the definition of SNR that we use scales with  $N = T/T_s$ . As a result, a larger  $T$  increases the SNR. It is intuitively pleasing that the same SNR results in the same  $P_D$  values in Figs. 6 and 7.



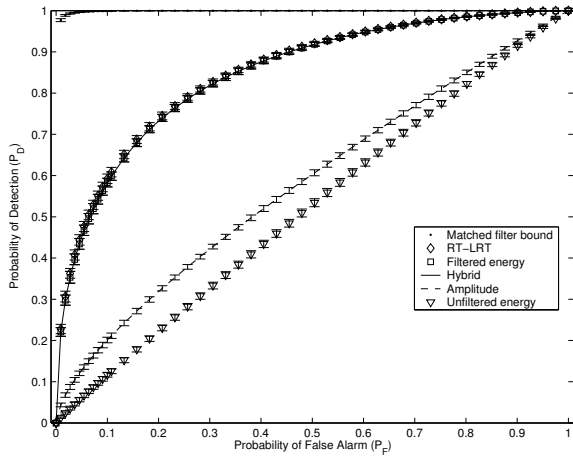


Fig. 5. Simulated ROC curves ( $P_D$  vs.  $P_F$ ) for the DT random telegraph model with symmetric transition probabilities at SNR = -34.3 dB,  $T = 60$  s, and  $\lambda = 0.5 \text{ s}^{-1}$  for the omniscient matched filter, DT random telegraph LRT (RT-LRT), filtered energy, hybrid, amplitude, and unfiltered energy detectors. The RT-LRT is theoretically optimal.

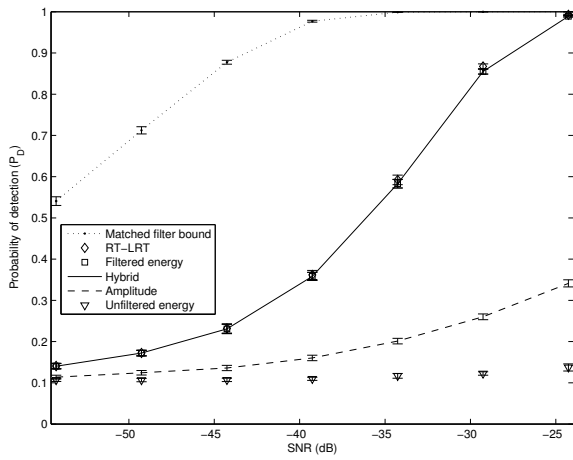


Fig. 6. Simulated power curves ( $P_D$  vs. SNR) for the DT random telegraph model with  $P_F$  fixed at 0.1 and  $\lambda = 0.5 \text{ s}^{-1}$ ,  $T = 60$  s. The RT-LRT is theoretically optimal.

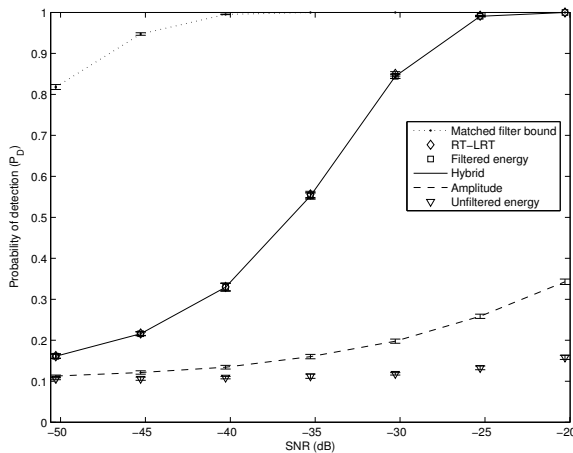


Fig. 7. Simulated power curves for the DT random telegraph model with  $P_F$  fixed at 0.1 and  $\lambda = 0.5 \text{ s}^{-1}$ ,  $T = 150$  s. The RT-LRT is theoretically optimal.

The ROC and power curve simulations were repeated with different values of  $\lambda$ , and the same relative performance was observed. In the interest of space, however, they will not be shown. Note that performance degrades as  $\lambda T_s$  increases. From (10), the probability of transition between consecutive time samples is  $1-p = \lambda T_s$ . A higher value of  $\lambda T_s$  results in a higher probability of transition, which decreases performance.

In the second set of simulations, we investigate the case in which the transition probabilities are asymmetric, i.e.  $p \neq q$ . Consider the scenario where  $p = 0.9998$ ,  $q = 0.9992$ , and all of the other parameters values are unchanged. The ROC curves for these parameter values are presented in Fig. 8. There are noticeable differences between the curves of the RT-LRT and filtered energy detectors. The hybrid detector's curve is slightly below that of the LRT, and it is better than that of the filtered energy detector. In fact, the filtered energy detector has worse performance than the amplitude detector. An asymmetry in  $p, q$  leads to a non-zero mean signal, which is why the amplitude detector's performance improves. Indeed, for the DT random telegraph model,  $\lim_{i \rightarrow \infty} E[\zeta_i] = A \frac{p-q}{1-r} = 0.6A$  for the values of  $p$  and  $q$  used here. Asymmetric transition probabilities can arise in some situations, e.g. the feedback-cooling-of-spins MRFM protocol proposed by Budakian [6].

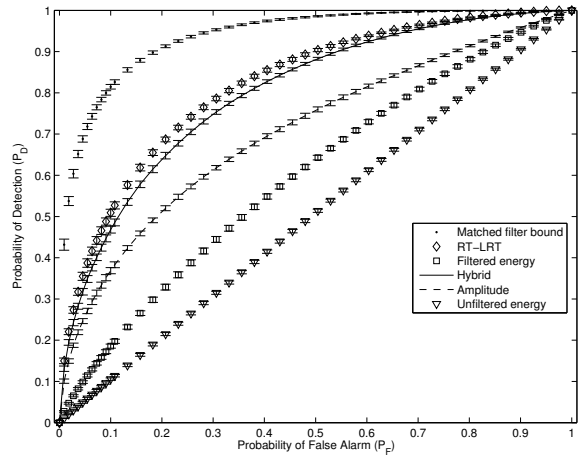


Fig. 8. Simulated ROC curves for the DT random telegraph model with asymmetric transition probabilities ( $p = 0.9998$ ,  $q = 0.9992$ ) at SNR = -6.71 dB,  $T = 150$  s. The RT-LRT is theoretically optimal.

Power curves from SNR = 3.3 dB to -16.7 dB were generated for the asymmetric case in Fig. 9. We used  $T = 150$  s for simulations of the asymmetric random telegraph model. A larger value of  $T$  is required when  $p \neq q$  for the hybrid detector to be a good approximation to the optimal LRT. It is important to recall that the hybrid detector given by (61) was derived for large  $N$ . An estimation of the error between (61) and the LRT of the DT random telegraph has not been conducted. It is likely, however, that when  $p \neq q$ , (61) is not as accurate as approximation (57). The hybrid detector has better performance than the amplitude and filtered energy detectors. It has performance that is comparable to the RT-LRT for lower SNR values.

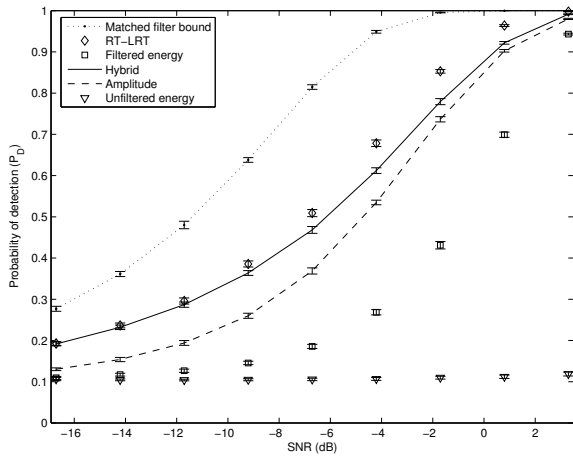


Fig. 9. Simulated power curves for the DT random telegraph model with  $P_F$  fixed at 0.1,  $p = 0.9998$ ,  $q = 0.9992$ , and  $T = 150$  s. The RT-LRT is theoretically optimal.

### B. Discrete-time random walk model

Recall that for the DT random walk model,  $\mathbf{P}_{\text{rw}}$  is tridiagonal. For the simulations, a particular subset of tridiagonal matrices was studied. Suppose for the moment that  $M$  is even. Recall that the random walk  $\zeta_i$  is confined to the interval  $[-Ms, Ms]$ . Define the lower-quartile transition probabilities as  $p_{l,1}, p_{l,2}$  and the upper-quartile transition probabilities as  $p_{u,1}, p_{u,2}$ . Let  $\mathbf{P}_{\text{rw}}^{(j,k)}$  be the  $(j, k)$ -th element of  $\mathbf{P}_{\text{rw}}$ . Here, we examine the performance of the detectors assuming the following reflecting boundary conditions:  $\mathbf{P}_{\text{rw}}^{(1,2)} = 1$ ,  $\mathbf{P}_{\text{rw}}^{(1,i)} = 0$  for  $i \neq 2$  and  $\mathbf{P}_{\text{rw}}^{(2M+1,2M)} = 1$ ,  $\mathbf{P}_{\text{rw}}^{(2M+1,i)} = 0$  for  $i \neq 2M$ . The rest of  $\mathbf{P}_{\text{rw}}$  is

$$\mathbf{P}_{\text{rw}}^{(j,k)} = \begin{cases} p_{l,1} & 2 \leq j < M/2 + 1, k = j - 1 \\ p_{l,2} & 2 \leq j < M/2 + 1, k = j + 1 \\ 0.5 & M/2 + 1 \leq j \leq 3M/2 + 1, k = j \pm 1 \\ p_{u,1} & 3M/2 + 1 < j \leq 2M, k = j - 1 \\ p_{u,2} & 3M/2 + 1 < j \leq 2M, k = j + 1 \end{cases} \quad (48)$$

Let  $\mathbf{A}_n(p_1, p_2)$  be a  $n \times (n + 2)$  matrix that looks like:

$$\mathbf{A}_n(p_1, p_2) = \begin{pmatrix} p_1 & 0 & p_2 & & & \\ & p_1 & 0 & p_2 & & \\ & & \ddots & \ddots & \ddots & \\ & & & p_1 & 0 & p_2 \end{pmatrix}$$

where the unspecified parts of the matrix are taken to be all zeros. In this section, the following subset of transition matrices for the DT random walk was studied:

$$\mathbf{P}_{\text{rw}} = \begin{pmatrix} 0 & 1 & & & \\ \mathbf{A}_{\frac{M}{2}-1}(p_{l,1}, p_{l,2}) & \mathbf{F} & & & \\ & & \mathbf{A}_{\frac{M}{2}-1}(p_{u,1}, p_{u,2}) & & \\ & & & 1 & 0 \end{pmatrix},$$

where  $\mathbf{F} = \mathbf{A}_{M+1}(0.5, 0.5)$ . Note that since each row of a probability transition matrix must sum to 1, one has  $p_{l,1} + p_{l,2} = 1$  and  $p_{u,1} + p_{u,2} = 1$ .

In the case of  $M$  odd, the ranges for the indices  $j, k$  would change in an obvious way. When  $p_{l,1} = p_{u,2}$  (or equivalently

$p_{l,2} = p_{u,1}$ ), we say that the transition probabilities are *symmetric*, and if not, that they are *asymmetric*. The matched filter, DT random walk LRT (RW-LRT), RT-LRT, filtered energy, amplitude, and unfiltered energy detectors are compared. In order to run the RT-LRT in the case of the symmetric DT random walk, an average autocorrelation function of the random walk was empirically generated; then  $p$  was selected (and choosing  $q = p$ ) so that the autocorrelation function of the symmetric DT random telegraph matched the empirical result. From this, the optimal  $\alpha_{\text{LP}}$  for the LPF of the filtered energy detector was also obtained.

The ROC curves for two symmetric cases are illustrated in Figs. 10 and 11. In the former,  $p_{l,1} = p_{l,2} = p_{u,1} = p_{u,2} = 0.5$ , while in the latter,  $p_{l,1} = p_{u,2} = 0.52$  and  $p_{l,2} = p_{u,1} = 0.48$ . In both cases, the performance of the RW-LRT, RT-LRT, and filtered energy detector are all approximately the same, i.e. the latter two detectors are nearly optimal. When the transition probabilities of the DT random walk are asymmetric however, as in the case of Fig. 12, the DT random walk LRT is noticeably better than the filtered energy detector.

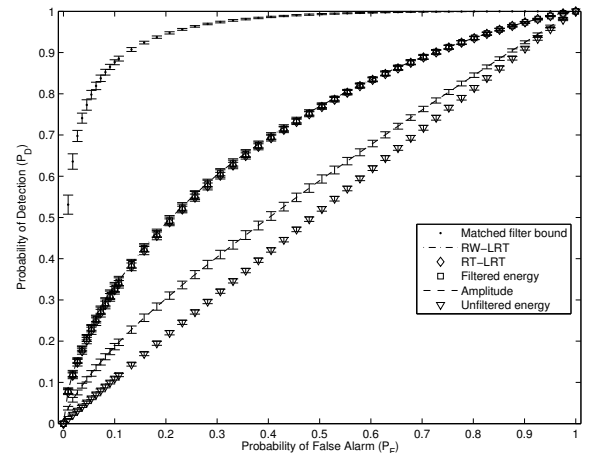


Fig. 10. Simulated ROC curves for the DT symmetric random walk  $p_{l,1} = p_{l,2} = p_{u,1} = p_{u,2} = 0.5$  at SNR = -44.0 dB,  $T = 60$  s for the matched filter, RW-LRT, RT-LRT, filtered energy, amplitude, and unfiltered energy detector. The RW-LRT is theoretically optimal.

## VI. CONCLUSION AND DISCUSSION

We have developed and compared optimal and non-optimal detectors under two single spin MRFM signal models. Recent experiments using the approximately optimal filtered energy detector have resulted in the successful detection of a single electron spin. This is strong evidence that the random telegraph signal model accurately describes the cantilever-single spin interaction.

The results of this paper lend strong theoretical and practical support to the use of the simple filtered energy detector for the current MRFM single spin research community. It has been shown that the existing baseband filtered energy detector that is in current use is approximately optimal in the case of the symmetric DT random telegraph model under the regime of low SNR, long observation time, and  $p$  close to 1. The last

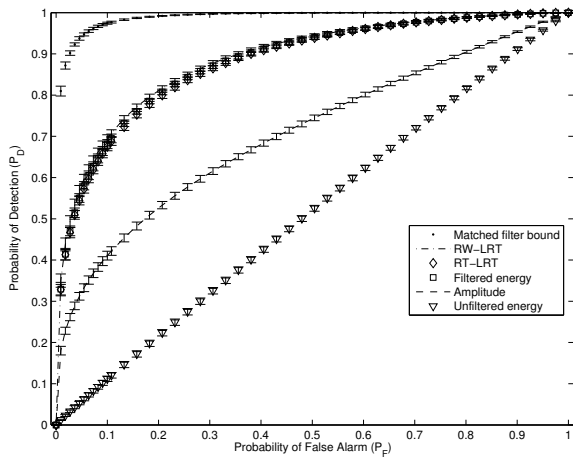


Fig. 11. Simulated ROC curves for the DT symmetric random walk  $p_{l,1} = p_{u,2} = 0.52$ ,  $p_{l,2} = p_{u,1} = 0.48$  at SNR = -39.1 dB,  $T = 60$  s. The RW-LRT is theoretically optimal.

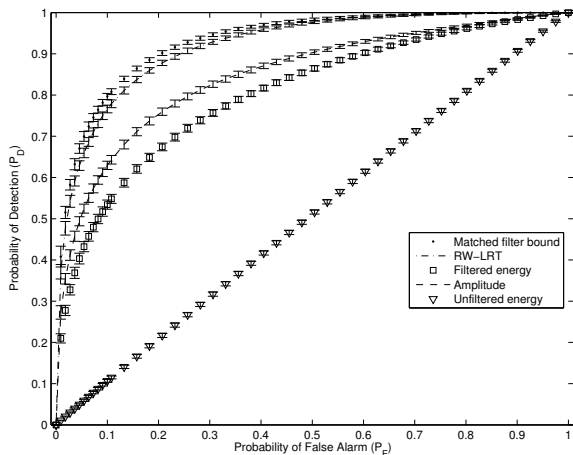


Fig. 12. Simulated ROC curves for the DT asymmetric random walk  $p_{l,1} = p_{u,1} = 0.45$ ,  $p_{l,2} = p_{u,2} = 0.55$  at SNR = -2.50 dB,  $T = 60$  s. The RW-LRT is theoretically optimal.

condition can be achieved by sampling at a sufficiently fast rate as compared to the rate of random transitions. This result has been extended to the case of the asymmetric DT random telegraph by using a hybrid filtered energy/amplitude/energy detector. Simulations were presented showing that the near optimality of the baseband filtered energy detector extends to the case of the symmetric DT random walk model. In the case of the asymmetric DT random walk, the filtered energy detector does not perform as well as the optimal LRT. We expect that a hybrid detector along the lines of that formulated for the DT random telegraph will perform close to the optimal for the asymmetric DT random walk. Mathematical analysis of the DT random walk model will be presented in a future paper.

A new interpretation of the LRT for a finite state DT Markov signal in AWGN under low SNR conditions was presented. Specifically, the LRT is approximately the matched filter statistic with the MMSE predictor signal values used in place of the known signal values.

The previous result can be used to obtain an approxi-

mation to the LRT for a general finite state DT Markov signal in AWGN under low SNR conditions. For ease of computation, we have proposed the use of affine estimators as suboptimal versions of the conditional mean estimators in (26). In particular, the suboptimal estimators of  $\zeta_i$  and  $\zeta_i^2$  for  $i \geq 1$  can be computed as the affine estimators with lowest MSE that lie in the linear span of  $\{1, y_0, \dots, y_{i-1}\}$  and  $\{1, y_0^2, \dots, y_{i-1}^2\}$  respectively. We apply this methodology to compute an approximation to the LRT for the DT random telegraph process in AWGN. The approximation matches the result previously obtained using a straightforward analysis.

## REFERENCES

- [1] K. Chun and N. O. Birge, "Dissipative quantum tunneling of a single defect in a disordered metal," *Physical Review B*, vol. 54, no. 7, pp. 4629–4637, 1996.
- [2] L. J. Vandamme, A. V. Belyakov, M. Y. Perov, and A. V. Yakimov, "Difference in dependence of  $1/f$  and RTS noise on current in quantum dots light emitting diodes," in *Noise in Devices and Circuits*, ser. Proceedings of the SPIE, vol. 5113, 2003, pp. 368–378.
- [3] S. Jha, K. Tan, and R. A. Maxion, "Markov chains, classifiers, and intrusion detection," in *Proceedings of the 14th IEEE Computer Security Foundations Workshop*, 2001, pp. 206–219.
- [4] P. D. Gader, M. Mystkowski, and Y. Zhao, "Landmine detection with ground penetrating radar using Hidden Markov Models," *IEEE Trans. Geosci. Remote Sensing*, vol. 39, no. 6, pp. 1231–1244, 2001.
- [5] D. Rugar, R. Budakian, H. J. Mamin, and B. W. Chui, "Single spin detection by magnetic resonance force microscopy," *Nature*, vol. 430, no. 6997, pp. 329–332, 2004.
- [6] R. Budakian, H. J. Mamin, B. W. Chui, and D. Rugar, "Creating order from random fluctuations in small spin ensembles," *Science*, vol. 307, no. 5708, pp. 408–411, 2005.
- [7] G. P. Berman, G. D. Doolen, P. C. Hammel, and V. I. Tsifrinovich, "Solid-state nuclear-spin quantum computer based on magnetic resonance force microscopy," *Physical Review B*, vol. 61, no. 21, pp. 14 694–14 699, 2000.
- [8] L. L. Scharf and L. W. Nolte, "Likelihood ratios for sequential hypothesis testing on Markov sequences," *IEEE Trans. Inform. Theory*, vol. IT-23, no. 1, pp. 101–109, 1977.
- [9] S. C. Schwartz, "The estimator-correlator for discrete-time problems," *IEEE Trans. Inform. Theory*, vol. IT-23, no. 1, pp. 93–100, 1977.
- [10] T. Kailath, "A general likelihood-ratio formula for random signals in Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-15, no. 3, pp. 350–361, 1969.
- [11] —, "A further note on a general likelihood formula for random signals in Gaussian noise," *IEEE Trans. Inform. Theory*, vol. IT-16, no. 4, pp. 393–396, 1970.
- [12] T. Kailath and M. Zakai, "Absolute continuity and Radon-Nikodym derivatives for certain measures relative to Wiener measure," *The Annals of Mathematical Statistics*, vol. 42, no. 1, pp. 130–140, 1971.
- [13] Y.-C. Yao, "Estimation of noisy telegraph processes: nonlinear filtering versus nonlinear smoothing," *IEEE Trans. Inform. Theory*, vol. IT-31, no. 3, pp. 444–446, 1985.
- [14] B. C. Stipe, H. J. Mamin, C. S. Yannoni, T. D. Stowe, T. W. Kenny, and D. Rugar, "Electron spin relaxation near a micron-size ferromagnet," *Physical Review Letters*, vol. 87, no. 27, pp. 277 602/1–4, 2001.
- [15] H. J. Mamin, R. Budakian, B. W. Chui, and D. Rugar, "Detection and manipulation of statistical polarization in small spin ensembles," *Physical Review Letters*, vol. 91, no. 20, pp. 207 604/1–4, 2003.
- [16] C. P. Slichter, *Principles of Magnetic Resonance*, 3rd ed. Springer-Verlag, Berlin, 1989.
- [17] K. Wago, D. Botkin, C. S. Yannoni, and D. Rugar, "Force-detected electron-spin resonance: adiabatic inversion, nutation, and spin echo," *Physical Review B*, vol. 57, no. 2, pp. 1108–1114, 1998.
- [18] G. P. Berman, F. Borgonovi, and V. I. Tsifrinovich, "A model for quantum jumps in magnetic resonance force microscopy," *Physics Letters A*, vol. 337, no. 3, pp. 161–165, 2005.
- [19] —, "Wave function collapses in a single spin magnetic resonance force microscopy," *Physics Letters A*, vol. 331, no. 3-4, pp. 187–192, 2004.

- [20] G. P. Berman, F. Borgonovi, V. N. Gorshkov, and V. I. Tsifrinovich, "Modeling and simulations of a single-spin measurement using MRFM," *IEEE Trans. Nanotechnol.*, vol. 4, no. 1, pp. 14–20, 2005.
- [21] A. O. Hero, B. Ma, O. J. J. Michel, and J. Gorman, "Applications of entropic spanning graphs," *IEEE Signal Processing Mag.*, vol. 19, no. 5, pp. 85–95, 2002.
- [22] M. Ting, "Signal processing for magnetic resonance force microscopy," Ph.D. thesis proposal report, University of Michigan, 2005.
- [23] H. Stark and J. W. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*, 2nd ed. Prentice-Hall, New Jersey, 1994.
- [24] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*. Wiley, New York, 1968, vol. 1.
- [25] S. K. Mitra, *Digital Signal Processing: A Computer-Based Approach*, 2nd ed. McGraw-Hill, New York, 2001.
- [26] D. G. Luenberger, *Optimization by vector space methods*. New York: John Wiley, 1969.

## APPENDIX

Define  $q_i(u)$ ,  $u \in \Psi_\pi$  to be the element of  $\underline{q}_i$  that corresponds to the state  $u$ . From (16) and (17), one can obtain the LRT for the DT random telegraph as

$$\Lambda_\pi(\underline{y}) = \prod_{i=0}^{N-1} \left[ q_i(A) e^{\frac{A}{\sigma^2} y_i} + q_i(-A) e^{-\frac{A}{\sigma^2} y_i} \right] \quad (49)$$

Let  $f(y_0, \dots, y_{N-1})$  denote the log LRT function of the DT random telegraph, i.e. the log of (49). Let  $g(y_0, \dots, y_{N-1})$  be the filtered energy detector function in (11). Let us analyze the two functions  $f$  and  $g$  under the regime of low SNR ( $|A/\sigma| \ll 1$ ) and long observation time ( $N \gg 1$ ).

We want to obtain the approximate Taylor series expansion of  $f$  about  $\underline{y} = \underline{0}$  and compare that with  $g$ . Define:

$$\theta_i \triangleq \frac{q_i(A) e^{\frac{A}{\sigma^2} y_i}}{q_i(A) e^{\frac{A}{\sigma^2} y_i} + q_i(-A) e^{-\frac{A}{\sigma^2} y_i}}$$

for  $i \geq 0$ . A recursive equation for  $\theta_i$  can be derived based on (16). Its approximate solution is

$$\theta_i \approx \beta_i + \frac{qA}{\sigma^2} \sum_{j=0}^i \xi_{ij} y_j, \quad i \geq 0$$

where  $\beta_i = \frac{1-q}{1-r} + \left( \frac{1}{2} - \frac{1-q}{1-r} \right) r^i$ ,  $i \geq 0$

$$\xi_{ij} = \frac{2(1-q)r^{i-j} + (2q-r-1)r^i}{1-r}, \quad 0 \leq j \leq i-1$$

$$\xi_{ii} = \frac{2(1-q)}{1-r} + \frac{r^i(2q-r-1)}{1-r} = 2\beta_i, \quad i \geq 0$$

(50)

and  $r = p + q - 1$ . Recall that  $p, q \in (0, 1)$  implies  $|r| < 1$ . Define  $s_i \triangleq \frac{A}{\sigma^2} y_i$ . Then,

$$f \approx \sum_i \left\{ \left[ s_i(2q_i(A) - 1) + \frac{1}{2} s_i^2 \right] - \frac{1}{2} \left[ s_i(2q_i(A) - 1) + \frac{1}{2} s_i^2 \right]^2 \right\} \quad (51)$$

By solving for  $q_i(A)$  in terms of  $\theta_i$  and using (50), one obtains

the approximate Taylor series expansion of  $f$  as

$$f \approx L_1 + L_{2a} + L_{2b} + \text{h.o.t.} \quad (52)$$

$$\text{with } L_1 = \frac{A}{\sigma^2} C_m \sum_i (1-r^i) y_i \quad (53)$$

$$L_{2a} = 2q \left( \frac{A}{\sigma^2} \right)^2 \sum_i \sum_{j=0}^{i-1} \left[ \frac{2(1-q)}{1-r} r^{i-j} - r^i C_m \right] y_i y_j \quad (54)$$

$$L_{2b} = \left( \frac{A}{\sigma^2} \right)^2 \sum_i \left\{ 4r \left( \frac{1-q}{1-r} \right)^2 + 2 \frac{(q-r)(1-q)}{(1-r)^2} - C_m(2q + C_m)r^i + \frac{1}{2} C_m^2 r^{2i} \right\} y_i^2 \quad (55)$$

$$C_m = \frac{p-q}{2-p-q} \quad (56)$$

In (52), "h.o.t." denotes the higher-order terms; specifically, terms of degree three or higher. The parameter  $C_m$  indicates the mismatch between the transition probabilities  $p$  and  $q$ . In the symmetric case when  $p = q$ , this results in  $C_m = 0$ , and one obtains a simpler expression for  $f$ . Let  $f_{\text{sym}}$  be the function  $f$  under symmetric transition probabilities, i.e.  $p = q$ . Then,

$$f_{\text{sym}} \approx 2p \left( \frac{A}{\sigma^2} \right)^2 \left\{ \sum_{i=1}^{N-1} \sum_{j=0}^{i-1} (2p-1)^{i-j} y_i y_j + \sum_{i=0}^{N-1} \left( 1 - \frac{1}{4p} \right) y_i^2 \right\} \quad (57)$$

For sufficiently large  $N$ , it can be shown that

$$g \approx D \left\{ \sum_{i=1}^{N-1} \sum_{j=0}^{i-1} \alpha_{\text{LP}}^{i-j} y_i y_j + \frac{\alpha_{\text{LP}}}{1 + \alpha_{\text{LP}}} \sum_{i=0}^{N-1} y_i^2 \right\} \quad (58)$$

where  $D = (1 - \alpha_{\text{LP}}^2)/2\alpha_{\text{LP}}$  is a constant. Note that  $D$  plays no role in the performance of the test statistic. Indeed, the detection test

$$\begin{array}{l} H_1 \\ g \geq \eta \\ H_0 \end{array}$$

has the same performance as

$$\begin{array}{l} H_1 \\ \gamma g \geq \gamma \eta \\ H_0 \end{array}$$

for a constant  $\gamma \in \mathbb{R}_+$  not dependent on the observations.

In order to compare the performance of  $f_{\text{sym}}$  and  $g$ , we consider their normalized versions  $\frac{1}{2p} \left( \frac{A}{\sigma^2} \right)^{-2} f_{\text{sym}}$  and  $\frac{1}{D} g$ . Denote these statistics by  $\tilde{f}_{\text{sym}}$  and  $\tilde{g}$  respectively. Both  $\tilde{f}_{\text{sym}}$  and  $\tilde{g}$  are a weighted sum of two terms: an energy term of the form  $\sum_i y_i^2$  and a second-order term of the form  $\sum_{j < i} \gamma^{i-j} y_i y_j$ , where  $\gamma = 2p - 1$  for  $\tilde{f}_{\text{sym}}$  and  $\gamma = \alpha_{\text{LP}}$  for  $\tilde{g}$ . If  $\alpha_{\text{LP}} = 2p - 1$ , then

$$|\tilde{f}_{\text{sym}} - \tilde{g}| \approx \frac{1}{4p} \sum_i y_i^2 \quad (59)$$

Now,  $E_1[\sum_{i=0}^{N-1} y_i^2] - E_0[\sum_{i=0}^{N-1} y_i^2] = A^2N$ . On the other hand, for large  $N$ ,

$$E_1 \left[ \sum_{i=1}^{N-1} \sum_{j=0}^{i-1} \alpha_{\text{LP}}^{i-j} y_i y_j \right] - E_0 \left[ \sum_{i=1}^{N-1} \sum_{j=0}^{i-1} \alpha_{\text{LP}}^{i-j} y_i y_j \right] \approx GA^2N \quad (60)$$

where  $G = \alpha_{\text{LP}}(2p-1)/[1 - \alpha_{\text{LP}}(2p-1)]$ . When  $\alpha_{\text{LP}} = 2p-1$ ,  $G = \frac{(2p-1)^2}{1-(2p-1)^2} = \frac{1}{4(1-p)} + \frac{1}{4p} - 1$ . For  $p$  close to 1,  $G \gg \frac{1}{4p}$ , and  $GA^2N \gg \frac{1}{4p}A^2N$ . So to the first moment, the difference of  $\frac{1}{4p} \sum_i y_i^2$  between  $\tilde{f}_{\text{sym}}$  and  $\tilde{g}$  does not represent a significant difference when  $p \approx 1$ . Under these conditions, we expect that the performance of the filtered energy detector and the DT random telegraph LRT to be similar.

It is possible to obtain an approximation to the DT random telegraph LRT that holds when we make no assumption about  $p$  being equal to  $q$ . When  $p \neq q$ , we have  $C_m \neq 0$ , and there are terms of the form  $r^i C_m$  and  $r^{2i} C_m^2$  in (53)-(55). Since  $|r| < 1$ ,  $r^i \rightarrow 0$  in the limit as  $i \rightarrow \infty$ . So drop these terms to get:

$$f \approx C \left\{ \frac{(p-q)\sigma^2}{4q(1-r)A} \sum_i y_i + \sum_i \sum_{j < i} r^{i-j} y_i y_j + \left[ \frac{1}{2} + \frac{r(1-q)}{2q(1-r)} \right] \sum_i y_i^2 \right\} \quad (61)$$

where  $C = 4q \frac{1-q}{1-r} \left( \frac{A}{\sigma^2} \right)^2$  is a constant. Define  $C_a \triangleq \frac{(p-q)\sigma^2}{4q(1-r)A}$  and  $C_e \triangleq \frac{r(1-q)}{2q(1-r)}$ . In order to equate the coefficients of the cross-terms  $y_i y_j$  between (61) and (58), we require  $\alpha_{\text{LP}} = r = p+q-1$ . In  $g$ , the ratio of the energy terms to the cross-terms is  $\alpha_{\text{LP}}/(1+\alpha_{\text{LP}})$ . For  $r = \alpha_{\text{LP}} \approx 1 \Rightarrow \alpha_{\text{LP}}/(1+\alpha_{\text{LP}}) \approx 1/2$ . The idea is to add the energy and amplitude statistics to  $g$  so that all three statistics are in the same ratio as in (61). Let  $g_{\text{hyb}}$  be the ‘‘extended’’ version of  $g$ , which we shall call the hybrid filtered energy/amplitude/energy detector:

$$g_{\text{hyb}} \triangleq g + \frac{1 - \alpha_{\text{LP}}^2}{2\alpha_{\text{LP}}} \left[ C_a \sum_i y_i + C_e \sum_i y_i^2 \right] = g + \frac{1 - \alpha_{\text{LP}}^2}{2\alpha_{\text{LP}}} C_a \sum_i y_i + \frac{1 - \alpha_{\text{LP}}^2}{2\alpha_{\text{LP}}} C_e \sum_i y_i^2 \quad (62)$$

We expect  $g_{\text{hyb}}$  to have performance that is similar to  $f$  under the conditions of large  $N$ , low SNR, and  $r \approx 1$ . The constants in (62) can be further simplified. Let  $K_a \triangleq C_a(1 - \alpha_{\text{LP}}^2)/2\alpha_{\text{LP}}$  and  $K_b \triangleq C_e(1 - \alpha_{\text{LP}}^2)/2\alpha_{\text{LP}}$ . As  $\alpha_{\text{LP}} = p+q-1$ , one obtains after some algebra

$$K_a = \frac{p^2 - q^2}{8q(p+q-1)} \left( \frac{A}{\sigma^2} \right)^{-1} \quad \text{and} \quad K_e = \frac{(p+q)(1-q)}{4q} \quad (63)$$