DETECTING RANDOM WALKS HIDDEN IN NOISE: PHASE TRANSITION ON LARGE GRAPHS

Ameya Agaskar[1,2] *and Yue M. Lu*[1]

[1]Harvard University, Cambridge, MA 02138, USA
[2]MIT Lincoln Laboratory, Lexington, MA 02420, USA
Email: {aagaskar, yuelu}@seas.harvard.edu

ABSTRACT

We consider the problem of distinguishing between two hypotheses: that a sequence of signals on a large graph consists entirely of noise, or that it contains a realization of a random walk buried in the noise. The problem of computing the error exponent of the optimal detector is simple to formulate, but exhibits deep connections to problems known to be difficult, such as computing Lyapunov exponents of products of random matrices and the free entropy density of statistical mechanical systems. We describe these connections, and define an algorithm that efficiently computes the error exponent of the Neyman-Pearson detector. We also derive a closed-form formula, derived from a statistical mechanics-based approximation, for the error exponent on an arbitrary graph of large size. The derivation of this formula is not entirely rigorous, but it closely matches the empirical results in all our experiments. This formula explains a phase transition phenomenon in the error exponent: below a threshold SNR, the error exponent is nearly constant and near zero, indicating poor performance; above the threshold, there is rapid improvement in performance as the SNR increases. The location of the phase transition depends on the entropy rate of the random walk.

***Index Terms*—** Neyman-Pearson detection, random walks, hidden Markov processes, Lyapunov exponent, phase transitions

1. INTRODUCTION

Detecting a hidden Markov process is a fundamental problem in statistical signal processing, with broad applications (see, e.g., [1]–[5]). The central issues are to derive the optimal detectors and to understand their performance within different signal-to-noise ratio (SNR) regimes. We consider this problem in the setting that the hidden Markov process is a signal generated by a random walk on a *large* graph and corrupted with additive white Gaussian noise. Our goal is to distinguish between this hypothesis and the null hypothesis that the observations made on the graph consist entirely of noise.

A scenario in which this problem arises is the detection of an intruder via a sensor network; the motion of a potential intruder might be modeled as a random walk on a graph representing the network, and one is tasked with testing the hypothesis that an intruder is currently present based on noisy measurements from each sensor.

We will describe the optimal detector for this hypothesis testing problem and analyze its asymptotic performance via the (type II) error exponent [6], which completely characterizes the asymptotic rate of decay of the probability of miss as more data is collected, with the probability of false alarm held fixed. Using tools from statistical mechanics [7, 8], we derive a closed-form approximation for the error exponent, which reveals an important phase transition phenomenon: below a threshold SNR which depends on the entropy of the random walk, the optimal detector performs very poorly and does not improve with marginal improvements to the SNR; but above the threshold the error exponent improves almost linearly with SNR.

1.1. Related and Prior Work

Detecting a continuous Gauss-Markov process in Gaussian noise is a classical signal processing problem that has been extensively studied (see, e.g., [9, 10].) Hypothesis testing that tries to distinguish between two different *finite-state* Markov chains based on *noiseless* realizations is also well-understood [11]–[13]. In this work, we focus on the related problem of detecting random walks on finite graphs (which are finite state Markov chains) based on observations that are perturbed by additive Gaussian noise. These observations do not satisfy the Markov property and are not jointly Gaussian, making the problem a difficult one.

The structure of the optimal detector for a finite-state Markov chain in noise was addressed in [1, 14]. We are interested in going further and characterizing the asymptotic performance of the optimal detector by computing the error exponent. For the Gauss-Markov case, a closed-form expression for the error exponent was derived by Sung *et al.* [2] using a state space representation. Our problem turns out to be more challenging. Leong *et al.* [4] described a numerical technique to find the error exponent for detecting a two-state Markov chain in noise by approximating the solution to a certain integral equation. Little is known about computing the error exponents for detecting general Markov chains with more than two states; this is the focus of our work.

Our problem of finding the detection error exponent is also related to computing the entropy rate of hidden Markov processes (HMPs), a long-standing problem in information theory. Jacquet *et al.* [15] studied a two-state Markov chain observed through a binary symmetric channel and showed that computing the entropy rate of the underlying HMP is equivalent to finding the top Lyapunov exponent of the product of an infinite sequence of random matrices [16], a problem known to be difficult [17]. They, and others, focused on local asymptotic approximations to the entropy rate by deriving its partial derivatives with respect to various parameters of the system (e.g. [15, 18, 19]).

In this paper, we study the detection error exponent when the underlying graph (i.e., Markov chain) is large, a context that, to our knowledge, has not been addressed before. The high-dimensional setting allows us to use tools from large sample theory [20] and statistical mechanics [7, 8] to reach a closed-form approximation for the error exponent, and to uncover a phase transition phenomenon in the

This work was sponsored in part by the Department of the Air Force under Contract FA8721-05-C-0002. Opinions, interpretations, conclusions and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

detector performance.

1.2. Paper Outline and Summary of Contributions

In Section 2 we first mathematically formulate the hypothesis-testing problem, and then describe the optimal detector and define its error exponent. We provide in Section 3 several alternative expressions for the likelihood ratio, and give a simple Monte Carlo algorithm for numerically computing the error exponent. Our main contributions are presented in Section 4 and Section 5. Using the random energy model (REM) from statistical mechanics [8], we derive in Section 4 an *exact* asymptotic expression for the error exponent when the random walk takes place on a large unweighted and complete graph. This result is then extended in Section 5 to an analytic expression *approximating* the error exponent for arbitrary graphs. Based on a form of the asymptotic equipartition theorem for Markov chains [19] and again on the REM model, our derivations in reaching the general formula for arbitrary graphs are not entirely rigorous, but simulation results demonstrate the approximation's accuracy. They also illustrate the phase transition effect predicted by the analytic formula. We conclude the paper in Section 6.

2. PROBLEM STATEMENT

Given a sequence $(\boldsymbol{y}_i)_{i=1}^N$, $\boldsymbol{y}_i \in \mathbb{R}^K$, we consider the problem of testing the following hypotheses:

$$\begin{aligned} &\mathcal{H}_0 : \boldsymbol{y}_i = \boldsymbol{z}_i; \\ &\mathcal{H}_1 : \boldsymbol{y}_i = \boldsymbol{f}(s_i) + \boldsymbol{z}_i, \qquad i = 1, \ldots, N, \end{aligned} \tag{1}$$

where $\boldsymbol{z}_i \overset{\text{i.i.d}}{\sim} \mathcal{N}(0, \Sigma)$ is the noise process and $(s_i)_{i=1}^N \in \{1, \ldots, M\}^N$ is the N-step trajectory-sequence of a homogeneous random walk on a connected graph with M vertices. This random walk can be described by a Markov chain with an $M \times M$ one-step transition matrix $\boldsymbol{P}$. We let $\boldsymbol{f} : \{1, \ldots, M\} \to \mathbb{R}^K$ be a function that assigns a vector to each state, so that $\boldsymbol{f}(s)$ is the observable "signature" of the state s.

In this paper, we consider the special case with independent and identically-distributed (i.i.d.) noise (so $\Sigma = \sigma^2 I_K$), $K = M$, and the state map $\boldsymbol{f}(m) = \boldsymbol{e}_m$, where $\{\boldsymbol{e}_m\}_{m=1}^M$ is the standard basis in $\mathbb{R}^M$. In this case, we can think of each realization $\boldsymbol{y}_i$ as a dynamic process on a weighted graph, having at each time step some value at each vertex in the graph. Under $\mathcal{H}_0$, the process is just noise, i.i.d. in time and space. Under $\mathcal{H}_1$, a "particle" undergoes a random walk on the graph, and at each time step the data consist of an impulse at the current location of the particle corrupted with i.i.d. Gaussian noise. The more general case, as in (1), will be studied further and presented in a follow-up work.

To simplify the notation, we will concatenate the column vectors $\boldsymbol{y}_i$ to form a matrix $\boldsymbol{Y}^N$, whose entry $y_{k,l}$ is the kth entry of $\boldsymbol{y}_l$. By the Neyman-Pearson Lemma [6], the most powerful test of a given size compares the likelihood ratio to a threshold and rejects $\mathcal{H}_0$ if the ratio surpasses a threshold:

$$\delta^N(\boldsymbol{Y}^N) = \begin{cases} 1 & \text{if } L_N(\boldsymbol{Y}^N) > \tau_N \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

where $L_N(\boldsymbol{Y}^N) \overset{\text{def}}{=} \frac{\Pr(\boldsymbol{Y}^N|\mathcal{H}_1)}{\Pr(\boldsymbol{Y}^N|\mathcal{H}_0)}$. Several expressions for the likelihood ratio will be given in the sequel.

We consider the (type-II) error exponent η of the Neyman-Pearson detector, which is the exponential decay rate of the miss probability under a fixed false alarm probability as the number of observations increases. To be precise [2],

$$\eta \overset{\text{def}}{=} -\lim_{N\to\infty} \frac{1}{N} \log P_{\text{miss}}\left(\delta^N\right). \tag{3}$$

The error exponent can be computed as [2]

$$\eta = -\lim_{N\to\infty} \frac{1}{N} \log L_N(\boldsymbol{Y}^N) \tag{4}$$

where the limit requires almost sure convergence under $\mathcal{H}_0$.

The natural question to ask is how η changes with the parameters of the problem: the SNR, given by $1/\sigma^2$, and the graph on which the random walk takes place. We address this problem by providing a close-form expression approximating the error exponent in the asymptotic limit of large graphs.

3. COMPUTING THE ERROR EXPONENT

To begin, we compute the likelihood ratio given data $\{\boldsymbol{y}_i\}_{i=1}^N$. We suppose that the initial state is chosen according to the probability vector $\boldsymbol{\pi}_0$, where $(\boldsymbol{\pi}_0)_k = \Pr(s_1 = k) > 0$ for $k = 1, \ldots, M$, and that the elements of the transition matrix P are $p_{ij} = \Pr(s_k = j|s_{k-1} = i)$.

Then

$$\begin{aligned} L_N(\boldsymbol{Y}^N) &= \sum_{S=(s_1,s_2,\ldots,s_N)} \Pr(S) \exp\left(\sum_{i=1}^N \frac{y_{s_i,i} - 1/2}{\sigma^2}\right) \\ &= e^{-N/(2\sigma^2)}\, \boldsymbol{\pi}_0^T D_1 P D_2 P \cdots D_{N-1} P D_N \mathbf{1} \end{aligned} \tag{5}$$

where the second equality was shown in [14], with $\mathbf{1}$ denoting a vector of all ones and $D_i \overset{\text{def}}{=} \operatorname{diag}\left[\exp(\frac{y_{1,i}}{\sigma^2}), \ldots, \exp(\frac{y_{M,i}}{\sigma^2})\right]$. Here, $y_{k,l}$ are i.i.d. zero-mean Gaussian with variance σ^2, as under $\mathcal{H}_0$.

It then follows from (4) that the error exponent is given by the almost sure limit

$$\begin{aligned} \eta &= \frac{1}{2\sigma^2} - \lim_{N\to\infty} \frac{1}{N} \log \boldsymbol{\pi}_0^T D_1 P D_2 P \cdots D_{N-1} P D_N \mathbf{1} && (6) \\ &= \frac{1}{2\sigma^2} - \underbrace{\lim_{N\to\infty} \frac{1}{N} \log ||D_1 P D_2 P \cdots D_N P||_\star}_{\gamma}, && (7) \end{aligned}$$

where $||\cdot||_\star$ is any matrix norm (e.g., 2-norm or 1-norm). The equivalence between (6) and (7) can be established by noting that the quantities inside the logarithms of both formulae are equal to within a finite, constant multiplicative factor, which vanishes in the limit due to the logarithm and division by N. Due to space constraint, we omit the proof of this claim, which can be verified by using the Cauchy-Schwarz inequality and the equivalence of finite-dimensional norms.

We note that the second term, γ, in (7) denotes the exponential (growth or decay) rate of a product of i.i.d. random matrices—namely, the matrices $D_i P$. Such product of random matrices has a long history in mathematics and statistical physics, and the quantity γ is often referred to as the (top) Lyapunov exponent in the literature. In a classical paper [16], Furstenberg and Kesten showed that the almost sure limit in (7) is well defined, and that it is equal to its expectation. Incidentally, this result also guarantees that our error exponent given in (4) is well defined.

Computing the Lyapunov exponent analytically is known to be hard [17]. It generally requires solving an integral equation to obtain the invariant measure of a continuous diffusion process on a real projective space [21]. In low dimensions this can be done with numerical quadrature [4], but this is not tractable for our high dimensional problem. In Algorithm 1, we present a simple Monte Carlo approach to directly approximate the Lyapunov exponent.

Algorithm 1 `Estimate-Lyapunov-Exponent`

$\boldsymbol{v}_0 \leftarrow \frac{1}{\sqrt{M}}\mathbf{1}$
for $i = 1 \rightarrow N$ **do**
  $\boldsymbol{z}_i \sim \mathcal{N}(\mathbf{0}, \frac{1}{\sigma^2} I_M); D_{N+1-i} \leftarrow \mathrm{diag}(e^{\boldsymbol{z}_i})$
  $x_i \leftarrow ||D_{N+1-i} P\, \boldsymbol{v}_{i-1}||$ {*Any norm will do.*}
  $\boldsymbol{v}_i \leftarrow D_{N+1-i} P\, \boldsymbol{v}_{i-1}/x_i$
end for
$\widehat{\gamma}_N \leftarrow \frac{1}{N}\sum_{i=1}^{N} \log x_i$

Proposition 1. *The output $\widehat{\gamma}_N$ of Algorithm 1 converges, as N tends to infinity, almost surely to the Lyapunov exponent γ.*

Proof. We use the fact that each $\boldsymbol{v}_i$ is unit norm and recursively expand $\boldsymbol{v}_N$:

$$\begin{aligned} 1 = ||\boldsymbol{v}_N|| &= \left|\left|\frac{D_1 P \boldsymbol{v}_{N-1}}{x_N}\right|\right| \\ &= \left|\left|\frac{D_1 P D_2 P \boldsymbol{v}_{N-2}}{x_N x_{N-1}}\right|\right| = \ \ldots \quad = \left|\left|\frac{D_1 P \cdots D_N P \boldsymbol{v}_0}{x_N x_{N-1} \cdots x_1}\right|\right|. \end{aligned}$$

So we obtain

$$\begin{aligned} \lim_{N\to\infty} \widehat{\gamma}_N &= \lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \log x_i = \lim_{N\to\infty} \frac{1}{N} \log(x_1 x_2 \cdots x_N) \\ &= \lim_{N\to\infty} \frac{1}{N} \log ||D_1 P \cdots D_N P\, \boldsymbol{v}_0|| = \gamma, \end{aligned}$$

where the fact that $\boldsymbol{v}_0$ has no zero entries as well as the equivalence of finite-dimensional norms ensures the last equality. □

In the sequel, we will describe an analytic approximation for η, and we will use Algorithm 1 to compute the error exponents of a given Markov chain $\boldsymbol{P}$ at various SNRs and compare these numerical results to the analytic approximation.

4. RANDOM WALKS ON COMPLETE GRAPHS

We present in this section a special case of the problem that can be solved exactly in the limit of large graph size M. Suppose that the random walk takes place on a complete graph with loops: every vertex is connected to every other vertex and itself, with equal transition probabilities. The random walk is thus just i.i.d. draws from $\{1, \ldots, M\}$, and the transition matrix $P = \frac{1}{M}\mathbf{1}\mathbf{1}^T$. To compute the error exponent, we return to the likelihood expression (5). Using the fact that the almost sure limit is equal to the limit in expectation [16] and the fact that the state sequence is i.i.d., we have

$$\begin{aligned} \eta &= \frac{1}{2\sigma^2} - \lim_{N\to\infty} \frac{1}{N}\mathbb{E}\log \sum_{s_1=1}^{M} \cdots \sum_{s_N=1}^{M} M^{-N} \exp\left(\sum_{i=1}^{N} \frac{y_{s_i,i}}{\sigma^2}\right) \\ &= \log(M) + \frac{1}{2\sigma^2} \\ &\quad - \lim_{N\to\infty} \frac{1}{N}\mathbb{E}\log\left[\left(\sum_{s_1=1}^{M} e^{\frac{y_{s_1,1}}{\sigma^2}}\right)\cdots\left(\sum_{s_N=1}^{M} e^{\frac{y_{s_N,N}}{\sigma^2}}\right)\right] \\ &= \log(M) + \frac{1}{2\sigma^2} - \mathbb{E}\log\sum_{s=1}^{M} \exp\left(\frac{y_s}{\sigma^2}\right), \end{aligned} \tag{8}$$

where $y_s \overset{\text{i.i.d.}}{\sim} \mathcal{N}\left(0, \sigma^2\right)$.

We know of no simple way to compute the last term in (8). However, it becomes tractable if we take an appropriate limit as the graph size M grows very large. In fact, in this limiting case, it is equivalent to the solution of a well-known problem in statistical mechanics: the free entropy density of the *random energy model* (REM) [8].

In statistical mechanics, a system's macroscopic behavior can be determined by computing thermodynamic potentials such as the free entropy; these potentials depend on the energy levels of each of the system's "microstates". The REM is a simplified system in which there are 2^n microstates, each of which has independent, Gaussian-distributed energy with zero mean and variance $n/2$. This model has been fully characterized in the asymptotic regime [22], and by making the appropriate substitutions, we can use it to obtain an expression for the error exponent.

To that end, we let the system size $n = \log_2(M)$, $\beta = \sqrt{\frac{2\log 2}{\sigma^2 \log M}}$ and the energy levels $E_s = -y_s\sqrt{\frac{\log M}{2\sigma^2 \log 2}} \sim \mathcal{N}(0, n/2)$ for $1 \le s \le 2^n = M$. In the literature, the parameter β is referred to as the inverse temperature. It follows from (8) that

$$\eta = \log(M) + \frac{1}{2\sigma^2} - n \cdot \underbrace{\frac{1}{n}\mathbb{E}\log\sum_{s=1}^{2^n} \exp\left(-\beta E_s\right)}_{\Phi^{(n)}(\beta)}. \tag{9}$$

As the system size n grows, $\Phi^{(n)}(\beta)$ converges to $\Phi(\beta)$, the free entropy density of the random energy model. This is the limiting case we are interested in.

To find $\Phi(\beta)$, it is assumed that $n \to \infty$, and the density of states with energies in a neighborhood of nE is found to leading exponential order in n. This density is then used to integrate $\exp(-\beta E)$—with the integral approximating the desired sum. The resulting expression is [8]

$$\Phi(\beta) = \begin{cases} \log 2 + \frac{\beta^2}{4}, & \text{if } \beta < 2\sqrt{\log 2} \\ \sqrt{\log 2}\,\beta, & \text{otherwise.} \end{cases} \tag{10}$$

Substituting in our values for the REM parameters, and adding in a correction term to account for finite n (which for practical reasons will be necessary, as n is only logarithmic in M), we obtain

$$\eta \doteq \begin{cases} 0, & \text{if } \frac{1}{\sigma^2} < 2\log M \\ \frac{1}{2\sigma^2} + \log M - \sqrt{2\log M \frac{1}{\sigma^2}} + \frac{\log\log M}{2\sqrt{2\sigma^2 \log M}}, & \text{otherwise,} \end{cases} \tag{11}$$

where $\doteq$ indicates exact asymptotic equality for very large M. Due to space constraints, we will leave the detailed derivation of this expression to a follow-up work.

The expression (11) indicates an interesting phase transition: when the SNR (i.e., $1/\sigma^2$) is below a threshold, the error exponent for large M is nearly zero, indicating poor performance; above the threshold, there is rapid improvement as the SNR increases. It is also interesting to note that the threshold in (11), $2\log M$, is exactly equal to twice the entropy of a uniform distribution on $\{1, 2, \ldots, M\}$. In the next section, we will extend this result to arbitrary graphs.

5. GENERAL GRAPHS

Returning to (5) we can rewrite the error exponent as

$$\eta = \frac{1}{2\sigma^2} - \lim_{N\to\infty}\frac{1}{N}\log\sum_{S\in\text{paths}(P,N)} \Pr(S)\exp\left(\sum_{i=1}^{N}\frac{y_i^{(s_i)}}{\sigma^2}\right), \tag{12}$$

where paths(P, N) is the set of all length-N sample paths of the Markov chain. Using a form of the asymptotic equipartition theorem for Markov chains [19], for large N there is a typical set

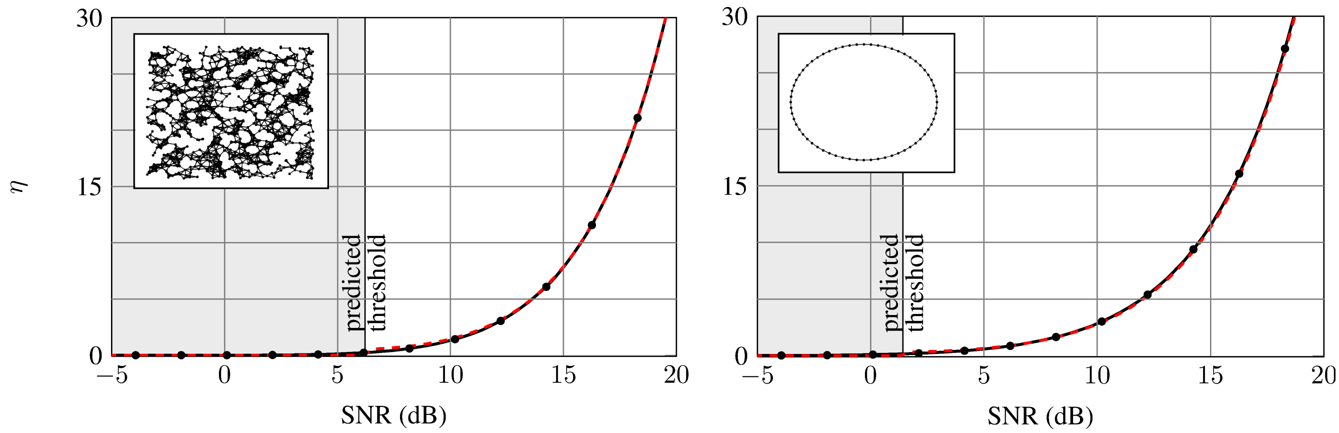

Fig. 1. Error exponent curves for two graphs. The solid black curve is the numerically computed error exponent. The dashed red curve is the proposed analytic approximation. The shaded area represents the predicted sub-threshold regime where the performance is poor. The curves for two graphs (shown in insets) are plotted. Left: a random geometric graph, M = 1000, H = 2.1 nats. Right: a cycle graph, M = 5000, H = 0.693 nats. At the same SNR level, the higher the entropy rate of the Markov chain, the worse the detector performance. The theoretical curve deviates slightly from experiments only in a small region around the phase transition point.

typ(P, N) of roughly $\exp(NH(P))$ paths S each having probability $\Pr(S) \approx \exp(-NH(P))$, where $H(P)$ is the entropy rate of the Markov chain. If the Markov chain is irreducible and aperiodic, it has a unique stationary distribution $\boldsymbol{\pi}$, and the entropy rate is given by

$$H(P) = -\sum_i \pi_i \sum_j P_{ij} \log P_{ij}, \tag{13}$$

where P_{ij} are the transition probabilities. The entropy rate gives the expected value—under the equilibrium distribution—of the conditional entropy of the next state given the current state. For example, for a complete, unweighted graph with loops, $H(P) = \log M$, the largest possible value for Markov chains of size M. Meanwhile, a cycle graph can be shown to have $H(P) = \log 2$, even though the random walk does not converge to a steady-state distribution.

By omitting atypical paths from the sum in (12), the error exponent can be approximated as

$$\begin{aligned}\eta &\approx \frac{1}{2\sigma^2} + H(P) - \lim_{N\to\infty} \frac{1}{N} \log \sum_{S\in \text{typ}(P,N)} \exp\left(\sum_{i=1}^{N} \frac{y_i^{(s_i)}}{\sigma^2}\right) \\ &= \frac{1}{2\sigma^2} + H(P) - \underbrace{\lim_{N\to\infty} \frac{1}{N} \log \sum_{S=1}^{e^{NH(P)}} \exp(-\beta E_S)}_{\Phi(\beta)\frac{H(P)}{\log 2}}, \end{aligned} \tag{14}$$

where we now set $\beta = \sqrt{\frac{2\log(2)}{H(P)\sigma^2}}$ and $E_S = \sqrt{\frac{H(P)}{2\log(2)\sigma^2}} \sum_{i=1}^{N} y_i^{(s_i)} \sim \mathcal{N}\left(0, \frac{NH(P)}{2\log(2)}\right)$. We will approximate $\Phi(\beta)$ by the free entropy of a system with $\exp(NH(P)) = 2^{NH(P)/\log(2)}$ random states. This approximation is not exact because the state energy levels are not independent. However, we will proceed as if they were independent and check the resulting expression numerically. Substituting the REM solution [scaled by $H(P)/\log(2)$], and again adding back the correction term from the complete graph model, gives

$$\eta \approx \begin{cases} 0, & \text{if } \frac{1}{\sigma^2} < 2H(P) \\ \frac{1}{2\sigma^2} + H(P) - \sqrt{2H(P)\frac{1}{\sigma^2}} + \frac{\log\log M}{\sqrt{2\sigma^2 \log M}}, & \text{otherwise.} \end{cases} \tag{15}$$

The phase transition in (15) indicates that below a certain threshold SNR, the performance of the optimal detector is poor, whereas above the threshold SNR, the error exponent grows almost linearly with the SNR. Note that the phase transition point increases with the entropy rate of the Markov chain. This is to be expected—knowledge of the Markov chain parameters is far less informative when the entropy rate is high (e.g., for a complete graph) than when it is low (e.g., for a cycle graph), so our detector performance is better in the latter case than the former.

Two main approximations were made to obtain (15). First, we used an asymptotic equipartition theorem to simplify the expression for the probability of any given state sequence in (14). Second, to apply the random energy model, we assumed that the state energy levels were independent, which is not true. Despite these approximations, numerical results indicate that the expression is valid. Figure 1 shows a comparison between the error exponents computed using Algorithm 1 and the closed form expression (15) for random walks on two different graphs. Curves are plotted illustrating the effect of SNR on the error exponent. The empirical curve follows the analytical approximation very closely. The phase transition is also quite evident—below a certain threshold, performance is very poor, then rises rapidly as the SNR is increased. We have done extensive numerical tests on various other graphs, and have found the match between the theoretical and empirical curves to be quite consistent.

6. CONCLUSIONS

We considered the problem of detecting a random walk on a graph hidden in additive Gaussian noise. We described an algorithm for numerically estimating the error exponent of the optimal Neyman-Pearson detector, and derived a closed-form expression approximating the error exponent. Numerical simulations show that the closed-form expression closely matches the empirical results. The error exponent exhibits phase transition behavior, indicating that performance of the optimal detector is poor below a threshold SNR (which depends on the entropy rate of the Markov chain) but then improves when the SNR is increased past the threshold. Future work will entail more rigorously justifying the closed-form approximation, and generalizing to arbitrary HMPs.

7. REFERENCES

[1] M. Ting, A. Hero, D. Rugar, C.-Y. Yip, and J. Fessler, “Near-optimal signal detection for finite-state Markov signals with application to magnetic resonance force microscopy,” *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2049–2062, Jun. 2006.

[2] Y. Sung, L. Tong, and H. V. Poor, “Neyman-Pearson detection of Gauss-Markov signals in noise: Closed-form error exponent and properties,” *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1354–1365, Apr. 2006.

[3] P. Runkle, L. Nguyen, J. McClellan, and L. Carin, “Multi-aspect target detection for SAR imagery using hidden Markov models,” *IEEE Trans. on Geoscience and Remote Sensing*, vol. 39, no. 1, pp. 46–55, Jan. 2001.

[4] A. Leong, S. Dey, and J. Evans, “Error exponents for Neyman-Pearson detection of Markov chains in noise,” *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 5097–5103, Oct. 2007.

[5] O. Missaoui, H. Frigui, and P. Gader, “Land-mine detection with ground-penetrating radar using multistream discrete hidden Markov models,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 6, pp. 2080–2099, Jun. 2011.

[6] H. V. Poor, *An Introduction to Signal Detection and Estimation*. Springer, Mar. 1994.

[7] R. Ellis, *Entropy, Large Deviations, and Statistical Mechanics*. Springer, Jun. 1985.

[8] M. Mézard and A. Montanari, *Information, Physics, and Computation*. Oxford University Press, USA, Mar. 2009.

[9] T. Kailath, “The innovations approach to detection and estimation theory,” *Proc. IEEE*, vol. 58, no. 5, pp. 680–695, May 1970.

[10] L. Scharf and L. Nolte, “Likelihood ratios for sequential hypothesis testing on Markov sequences,” *IEEE Trans. Inf. Theory*, vol. 23, no. 1, pp. 101–109, Jan. 1977.

[11] D. Kazakos, “The Bhattacharyya distance and detection between Markov chains,” *IEEE Trans. Inf. Theory*, vol. 24, no. 6, pp. 747–754, Nov. 1978.

[12] K. Nakagawa and F. Kanaya, “On the converse theorem in statistical hypothesis testing for Markov chains,” *IEEE Trans. Inf. Theory*, vol. 39, no. 2, p. 629–633, 1993.

[13] B. C. Levy, *Principles of Signal Detection and Parameter Estimation*. Springer, Jul. 2008.

[14] M. Ting and A. Hero, “Detection of a random walk signal in the regime of low signal to noise ratio and long observation time,” in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc.*, vol. 3, May 2006, pp. 293–296.

[15] P. Jacquet, G. Seroussi, and W. Szpankowski, “On the entropy of a hidden Markov process,” *Theoretical Computer Science*, vol. 395, no. 2–3, pp. 203–219, May 2008.

[16] H. Furstenberg and H. Kesten, “Products of random matrices,” *The Annals of Mathematical Statistics*, vol. 31, no. 2, pp. 457–469, Jun. 1960.

[17] J. N. Tsitsiklis and V. D. Blondel, “The Lyapunov exponent and joint spectral radius of pairs of matrices are hard—when not impossible—to compute and to approximate,” *Mathematics of Control, Signals, and Systems (MCSS)*, vol. 10, no. 1, pp. 31–40, 1997.

[18] E. Ordentlich and T. Weissman, “New bounds on the entropy rate of hidden Markov processes,” in *Information Theory Workshop, 2004. IEEE*, 2004, p. 117–122.

[19] R. Gharavi and V. Anantharam, “An upper bound for the largest Lyapunov exponent of a Markovian product of nonnegative matrices,” *Theoretical computer science*, vol. 332, no. 1, p. 543–557, 2005.

[20] T. S. Ferguson, *A Course in Large Sample Theory: Texts in Statistical Science*. CRC Press, Jul. 1996.

[21] A. Comtet, C. Texier, and Y. Tourigny, “Products of random matrices and generalised quantum point scatterers,” *Journal of Statistical Physics*, vol. 140, no. 3, p. 427–466, 2010.

[22] B. Derrida, “Random-energy model: An exactly solvable model of disordered systems,” *Physical Review B*, vol. 24, no. 5, p. 2613, 1981.