Detection Error Exponent for Spatially Dependent Samples in Random Networks

Animashree Anandkumar*†, Lang Tong* and Alan Willsky†

*ECE Dept., Cornell University, Ithaca, NY 14853, Email: {aa332@,ltong@ece}.cornell.edu
‡EECS Dept., MIT, Cambridge, MA 02139 Email: {animakum,willsky}@mit.edu

***Abstract*—The problem of binary hypothesis testing is considered when the measurements are drawn from a Markov random field (MRF) under each hypothesis. Spatial dependence of the measurements is incorporated by explicitly modeling the influence of sensor node locations on the clique potential functions of each MRF hypothesis. The nodes are placed i.i.d. in expanding areas with increasing sample size. Asymptotic performance of hypothesis testing is analyzed through the Neyman-Pearson type-II error exponent. The error exponent is expressed as the limit of a functional over dependency edges of the MRF hypotheses for acyclic graphs. Using the law of large numbers for graph functionals, the error exponent is derived.**

***Index Terms*—Error exponent, Markov random field, random graphs, law of large numbers for graph functionals.**

I. Introduction

The assumption that the observations are i.i.d. under each hypothesis is often used in the literature [1]. While the i.i.d. assumption leads to elegant results, it is often violated in practice. In this paper, we focus on the case when under each hypothesis, the observation samples are correlated according to a Markov random field (MRF) model which depends on the spatial locations from where the samples are collected.

For hypothesis testing, the probability of making an error is a key performance measure. It is desired that this error decay exponentially with increasing sample size. The rate of exponential decay of error probability is known as the *detection error exponent*, which serves as a performance measure for large-scale networks. It is not always tractable to find the error exponent in closed form. Although there are established results for the error exponent for general hypotheses [2], further simplifications are possible only for special cases, such as for i.i.d. or stationary samples [3].

It is challenging to find the error exponent for general hypotheses, especially for spatially-dependent samples collected from irregular locations. We consider random distribution for spatial location of the measurement samples thereby introducing additional randomness into the hypothesis-testing problem. As a result, the error-exponent analysis is influenced by the node-location distribution, and we study it here.

This work was supported in part by collaborative participation in Communications and Networks Consortium sponsored by U. S. Army Research Laboratory under Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011 and by Army Research Office under Grant ARO-W911NF-06-1-0346. The first author is supported by the IBM Ph.D Fellowship 2008-09 and is currently a visiting student at MIT, Cambridge, MA 02139. The third author is supported in part by ARO Grant W911NF-06-1-0076, in part by AFOSR through Grant FA9550-08-1-1080, in part under a MURI through AFOSR Grant FA9550-06-1-0324 and in part by Shell International Exploration and Production, Inc. The U. S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon.

A. Related Work and Contributions

The large-deviation analysis for the test of simple hypotheses with general distributions exists [2], [4], but closed-form expressions are possible only for certain cases. Such an analysis for homogeneous Gauss-Markov random fields on lattices have been considered in [5], [6]. However, their techniques are not easily generalized to irregular spatially-dependent distributions, considered here. In [7], an expression for the Kullback-Leibler (KL) divergence rate is derived when the two distributions are Markov chains of arbitrary order, which is a special case of the formulation here. In [8]–[10], approximations to the KL divergence rates for hidden Markov chains are derived.

In [11], we considered a special case of the problem here, viz., of testing a Gauss-Markov random field (GMRF) with nearest-neighbor dependency graph against independence when the nodes are uniformly distributed. In this paper, we extend the results to more general distributions, and node location distributions.

In this paper, we derive the error exponent in closed form when the MRF hypotheses have acyclic stabilizing dependency graphs, such as the nearest-neighbor graph. The key issue we address is the influence of node placement on the resulting detection error exponent. This provides guidelines for efficient node placements to maximize detection performance. We express the error exponent as a functional over the edges of a stabilizing graph and then use the law of large numbers (LLN) to evaluate it, recently proposed by Penrose and Yukich [12].

II. System Model

A. Stochastic model of sensor locations

We assume that n sensor nodes are placed randomly with sensor i located at $V_i \in \mathbb{R}^2$ and $\mathbf{V}_n{:=}[V_1, \ldots, V_n]$. We consider a sequence of sensor populations on expanding square regions $Q_{\frac{n}{\lambda}}$ of area $\frac{n}{\lambda}$ centered at the origin, where we fix λ as the overall sensor density and let the number of sensors $n \to \infty$. To generate sensor locations V_i, first let $Q_1 := [-\frac{1}{2}, \frac{1}{2}]^2$ be the unit area square, and $X_i \overset{i.i.d.}{\sim} \tau, 1 \leq i \leq n$, be a set of n independent and identically distributed (i.i.d.) random variables distributed on Q_1 according to τ. We next generate V_i by scaling X_i accordingly: $V_i = \sqrt{\frac{n}{\lambda}} X_i \in Q_{\frac{n}{\lambda}}$. Let $\mathcal{P}_\lambda$ be the homogeneous Poisson distribution on $\mathbb{R}^2$ with density λ.

B. Graphical inference model

The inference problem we consider is the simple binary hypothesis testing $\mathcal{H}_0$ vs. $\mathcal{H}_1$ on a pair of Markov random fields (MRF). A MRF is defined by its (undirected) dependency graph $\mathcal{G}$ and an associated pdf $f(\cdot \mid \mathcal{G})$. Under each hypothesis $\mathcal{H}_m$, let $\mathcal{G}_m(\mathbf{V}_n)$ be the dependency graph of the MRF, where $\mathbf{V}_n = \{V_1, \cdots, V_n\}$ is the set of random node, described in Sec II-A. We denote the (random) measurements from all the sensors in a set $\mathbf{V}$ by $\mathbf{Y}_\mathbf{V}$. The inference problem can be stated as:

$$\begin{aligned}\mathcal{H}_0 :&[\mathbf{Y}_{\mathbf{V}_n}, \mathbf{V}_n] \sim f(\mathbf{y}_{\mathbf{v}_n} \mid \mathcal{G}_0(\mathbf{v}_n), \mathcal{H}_0) \prod_{i=1}^{n} \tau(\sqrt{\tfrac{\lambda}{n}} v_i),\\ \mathcal{H}_1 :&[\mathbf{Y}_{\mathbf{V}_n}, \mathbf{V}_n] \sim f(\mathbf{y}_{\mathbf{v}_n} \mid \mathcal{G}_1(\mathbf{v}_n), \mathcal{H}_1) \prod_{i=1}^{n} \tau(\sqrt{\tfrac{\lambda}{n}} v_i). \quad (1)\end{aligned}$$

The conditional pdfs $f(\mathbf{y}_{\mathbf{v}_n} \mid \mathcal{G}_m(\mathbf{v}_n), \mathcal{H}_m)$ are defined on the Lesbegue measure, and $f(\mathbf{y}_{\mathbf{v}_n} \mid \mathcal{G}_0(\mathbf{v}_n), \mathcal{H}_0)$ is *absolutely continuous*[1] with respect to $f(\mathbf{y}_{\mathbf{v}_n} \mid \mathcal{G}_1(\mathbf{v}_n), \mathcal{H}_1)$ [13]. Note that sensor locations have the same distribution under either hypothesis. Therefore, only the conditional distribution under each hypothesis is relevant for inference.

The celebrated Hammersley-Clifford theorem states that, under the positivity condition [14], the likelihood function is

$$f(\mathbf{y}_{\mathbf{v}_n} \,|\, \mathcal{G}_m(\mathbf{v}_n), \mathcal{H}_m) = \frac{1}{Z_m(\mathbf{v}_n)} \exp[-\sum_{c \in \mathcal{C}_m} \psi_{m,c}(\mathbf{y}_c)], \quad (2)$$

where $\mathcal{C}_m$ is a collection of (maximal) cliques in $\mathcal{G}_m(\mathbf{v}_n)$, the functions $\psi_{m,c}$, known as *clique potentials*, are real valued, and $Z_m(\mathbf{v}_n) > 0$ is the normalization constant, also known as the *partition function*. In general, it is NP-hard to evaluate the partition function for given potential functions, although for the Gaussian distribution, it can be evaluated in polynomial time since it reduces to the evaluation of a determinant.

C. Spatial Modeling: Dependency Graph and Potentials

A key modeling feature in this paper is to incorporate the spatial dependence of sensor measurements. This is achieved by explicitly specifying the influence of (random) node locations on the MRF dependency graph and the conditional distributions of the measurements given the node locations.

We restrict our attention to proximity-based local dependency graphs such as the (undirected) (k-NNG) or the disk graph (also known as continuum percolation). An important localization property of these graphs is *stabilization* facilitating asymptotic scaling analysis.

We assume that a set of clique potentials $\psi_{m,c} > 0$ under either hypothesis can be parameterized locally by the sensor locations of the clique members and their l-hop neighbors, for some finite l, in a translation-invariant manner, i.e.,

$$\begin{aligned}\psi_{m,c}(\mathbf{y}_c; \mathbf{v}_n) &= \psi_{m,c}(\mathbf{y}_c; \mathbf{v}_n + v), \quad \forall c \in \mathcal{C}_m, v \in \mathbb{R}, \quad (3)\\ \psi_{m,c}(\mathbf{y}_c; \mathbf{v}_n) &= \psi_{m,c}(\mathbf{y}_c; \{v_i : \mathcal{N}^l(i) \in c\}), \ \forall c \in \mathcal{C}_m, \quad (4)\end{aligned}$$

[1] We use the convention that $0 \log \frac{0}{q} = 0$ and $p \log \frac{p}{0} = \infty$.

where $\mathcal{N}^l$ is the set of all 0 to l-hop neighbors. Further conditions are imposed for acyclic graphs in Section III-A.

D. Error Exponent

We consider the Neyman-Pearson (NP) formulation, where the detector is optimal at a fixed false-alarm probability. We focus on the large-network scenario, where the number of observations goes to infinity. Under Neyman-Pearson formulation, for any positive level of the false alarm or the type-I error probability, when the mis-detection or the type-II error probability $P_M(n)$ of the NP detector decays exponentially with the sample size n, we have the error exponent

$$\mathcal{D} := - \lim_{n \to \infty} \frac{1}{n} \log P_M(n). \quad (5)$$

In this paper, we are interested in evaluating the error exponent in (5) for random networks under MRF hypotheses.

Given the node locations $\mathbf{V}_n = \mathbf{v}_n$, let $D_{\mathbf{v}_n}$ denote the Kullback-Leibler divergence between the conditional pdfs $f(\mathbf{y}\mathbf{v}_n \mid \mathcal{G}_0(\mathbf{v}_n), \mathcal{H}_0)$ and $f(\mathbf{y}_{\mathbf{v}_n} \mid \mathcal{G}_1(\mathbf{v}_n), \mathcal{H}_1)$,

$$D_{\mathbf{v}_n} := \int_{\mathbf{y}_{\mathbf{v}_n}} \log \frac{f(\mathbf{y}_{\mathbf{v}_n} | \mathcal{G}_0(\mathbf{v}_n), \mathcal{H}_0)}{f(\mathbf{y}_{\mathbf{v}_n} | \mathcal{G}_1(\mathbf{v}_n), \mathcal{H}_1)} f(\mathbf{y}_{\mathbf{v}_n} | \mathcal{G}_0(\mathbf{v}_n), \mathcal{H}_0) \mathbf{d}\mathbf{y}_{\mathbf{v}_n}. \quad (6)$$

In Section IV, we relate the error exponent $\mathcal{D}$ in (5) to the KL-divergence in (6).

III. Error Exponent as a Graph Functional

The binary hypothesis-testing problem defined in (1) involves two different graphical models, each with its own dependency graph and an associated likelihood function. The optimal detection test is based on the log-likelihood ratio (LLR). With the substitution of (2), it is given by

$$\begin{aligned}L(\mathbf{y}_\mathbf{v}) \quad &:= \quad \log \frac{f(\mathbf{y}_\mathbf{v} \mid \mathcal{G}_0(\mathbf{v}), \mathcal{H}_0)}{f(\mathbf{y}_\mathbf{v} \mid \mathcal{G}_1(\mathbf{v}), \mathcal{H}_1)} \quad (7)\\ &= \quad \sum_{a \in \mathcal{C}_1} \psi_{1,a}(\mathbf{y}_a) - \sum_{b \in \mathcal{C}_0} \psi_{0,b}(\mathbf{y}_b) + \log \frac{Z_1}{Z_0}.\end{aligned}$$

Hence, the LLR is a functional on the two dependency graphs $\mathcal{G}_0$ and $\mathcal{G}_1$.

The *spectrum* of the LLR [2], [4] is defined as the distribution of the normalized LLR evaluated under the null hypothesis

$$\frac{L(\mathbf{Y}_{\mathbf{V}_n})}{n}, \quad [\mathbf{Y}_{\mathbf{V}_n}, \mathbf{V}_n] \text{ under } \mathcal{H}_0.$$

In [2], [4] it is proven that for Neyman-Pearson detection under a fixed type-I error bound, the LLR spectrum can fully characterize the type-II error exponent of the hypothesis-testing system, and is independent of the type-I bound.

When LLR spectrum converges in probability to a constant $\mathcal{D}$, the error exponent $\mathcal{D}$ of NP detection in (5) is [4]

$$\mathcal{D} = \mathrm{p}\lim_{n \to \infty} \frac{1}{n} L(\mathbf{Y}_{\mathbf{V}_n}), \quad [\mathbf{Y}_{\mathbf{V}_n}, \mathbf{V}_n] \text{ under } \mathcal{H}_0, \quad (8)$$

where $\mathrm{p}\lim$ denotes the limit in probability, assuming it exists.

When $\mathbf{Y}_{\mathbf{V}_n}$ are i.i.d. conditioned under both $\mathcal{H}_0$ and $\mathcal{H}_1$, the result in (8) reduces to Stein's lemma [3, Theorem 12.8.1] and the limit in (8) is the node Kullback-Leibler (KL) divergence, i.e., when $Y_{V_i} \overset{i.i.d.}{\sim} g_k$ under $\mathcal{H}_k$,

$$\mathcal{D} = D_{V_1} := \int_y \log \frac{g_0(y)}{g_1(y)} g_0(y) dy. \tag{9}$$

In Section IV, we evaluate the error exponent for MRF hypotheses through the limit in (8). Due to random node placement and spatial dependence of the MRF hypotheses, the error exponent in (8) is the limit of a random-graph functional, and we can appeal to the LLN for graph functionals [12].

A. Acyclic Dependency Graphs

We consider the case when the dependency graphs under either MRF hypothesis $\mathcal{G}_0$ and $\mathcal{G}_1$ are *acyclic* and also stabilizing, such as the Euclidean nearest-neighbor graph.

Given a fixed set of points $\mathbf{v}_n$, the joint pdf of MRF for an acyclic dependency graph $\mathcal{G}(\mathbf{v}_n)$ admits a factorization [14]

$$f(\mathbf{y}_{\mathbf{v}_n}) = \prod_{i \in \mathbf{v}_n} f_i(y_i) \prod_{\substack{(i,j) \in \mathcal{G}(\mathbf{v}_n) \\ i<j}} \frac{f_{i,j}(y_i, y_j)}{f_i(y_i) f_i(y_j)}, \tag{10}$$

where f_i are the node marginal pdfs and $f_{i,j}$ are the pairwise pdfs on the edges. Recall that instead of fixed node locations, we have random locations $\mathbf{V}_n$ here, and hence, we consider the conditional pdf $f(\mathbf{y}_{\mathbf{V}_n} | \mathcal{H}_m, \mathcal{G}_m(\mathbf{V}_n))$ under each hypothesis $\mathcal{H}_m$. From (10), for an acyclic dependency graph $\mathcal{G}_m(\mathbf{V}_n)$, we can specify the conditional pdf $f(\mathbf{y}_{\mathbf{V}_n} | \mathcal{H}_m, \mathcal{G}_m(\mathbf{V}_n))$ through the conditional node pdfs $f_i(y_i | \mathcal{H}_m, \mathcal{G}_m)$ and the conditional pairwise edge pdfs $f_{i,j}(y_i, y_j | \mathcal{H}_m, \mathcal{G}_m)$.

We consider here a special form of spatial dependence in (4) by having identical node marginal pdfs for all node locations and edge marginal pdfs which are dependent only on the respective edge lengths. Under hypothesis $\mathcal{H}_m$, for $m = 0, 1$,

$$f_i(y_i \mid \mathcal{G}_m, \mathcal{H}_m) = g_m(y_i),\ i \in \mathbf{V}_n, \tag{11}$$
$$f_{i,j}(y_i, y_j \mid \mathcal{G}_m, \mathcal{H}_m) = h_m(y_i, y_j \mid R_{ij}),\ (i,j) \in \mathcal{G}_m, \tag{12}$$

where g_m is the node pdf and h_m is the pairwise pdf at the edges conditioned on R_{ij}, the Euclidean length of edge (i, j).

By using (10), (11) and (12), we simplify (8) as

$$\mathcal{D} = \mathrm{p}\lim_{n \to \infty} \frac{1}{n} \Big[\sum_{i \in \mathbf{V}_n} \log \frac{g_0(Y_i)}{g_1(Y_i)} + \sum_{\substack{(i,j) \in \mathcal{G}_0 \\ i<j}} \log \frac{h_0(Y_i, Y_j \mid R_{ij})}{g_0(Y_i) g_0(Y_j)}$$
$$- \sum_{\substack{(i,j) \in \mathcal{G}_1 \\ i<j}} \log \frac{h_1(Y_i, Y_j \mid R_{ij})}{g_1(Y_i) g_1(Y_j)} \Big],\ [\mathbf{Y}_{\mathbf{V}_n}, \mathbf{V}_n] \text{ under } \mathcal{H}_0, \tag{13}$$

Note that the above expression is a graph functional, based on the edge lengths of random graphs $\mathcal{G}_0$ and $\mathcal{G}_1$ with additional randomness from the conditional distribution of the sensor measurements given the edge lengths.

IV. Detection Error Exponent

In this section, we derive the error exponent for general MRF hypotheses.

A. Testing Against Independence

We first provide the closed-form error exponent for the special case when the null hypothesis has i.i.d. measurements with no spatial dependence, $f(\mathbf{y}_{\mathbf{v}_n} | \mathcal{G}_0(\mathbf{v}_n), \mathcal{H}_0) = \prod_{i \in \mathbf{v}_n} g_0(y_i)$. Here, the dependency graph is trivial, $\mathcal{G}_0 = \emptyset$, and the error exponent in (13) simplifies as

$$\mathcal{D} = \mathrm{p}\lim_{n \to \infty} \frac{1}{n} \Big[- \sum_{\substack{(i,j) \in \mathcal{G}_1 \\ i<j}} \log \frac{h_1(Y_i, Y_j \mid R_{ij})}{g_1(Y_i) g_1(Y_j)}$$
$$+ \sum_{i \in \mathbf{V}_n} \log \frac{g_0(Y_i)}{g_1(Y_i)} \Big], \quad Y_{V_i} \overset{i.i.d.}{\sim} g_0, \sqrt{\tfrac{\lambda}{n}} V_i \overset{i.i.d.}{\sim} \tau. \tag{14}$$

The above expression is a graph functional defined over a marked point process, where the marks are the sensor measurements Y_{V_i} drawn i.i.d from the pdf g_0.

We can now appeal directly to the LLN for marked point processes [12, Thm. 2.1] to simplify (14). Define a functional on the edge lengths

$$\xi_1(r_{ij}) := \mathbb{E}\Big[-\log \frac{h_1(Y_i, Y_j)}{g_1(Y_i) g_1(Y_j)} \Big| R_{ij} = r_{ij}, \mathcal{H}_0 \Big], \tag{15}$$
$$= - \int_{y_i} \int_{y_j} \log \frac{h_1(y_i, y_j \mid R_{ij} = r_{ij})}{g_1(y_i) g_1(y_j)} g_0(y_i) g_0(y_j) dy_j dy_i,$$

where the expectation is over the measurements conditioned on the node locations.

ξ_1 is said to satisfy moments condition of order $p > 0$ if

$$\sup_{n \in \mathbb{N}} \mathbb{E}\,[\sum_{j \in \mathcal{N}(\mathbf{0}), j \in \mathbf{V}_n} \xi_1(R_{\mathbf{0}j})^p] < \infty, \tag{16}$$

where $\mathcal{N}(\mathbf{0})$ denotes the neighbors of the origin in $\mathcal{G}_1$ and the expectation is over the node locations. We require that $p = 1$ or 2. In Section V, we prove that ξ_1 satisfies the moment condition for the Gaussian distribution under some simple constraints on the covariance matrix.

Recall that $\mathcal{P}_\lambda$ is the homogeneous Poisson distribution on $\mathbb{R}^2$ with density λ. We now provide the result below.

Lemma 1 (Testing Acyclicity Against Independence): When ξ_1 satisfies the moments condition in (16), the error exponent for testing against independence has the form

$$\mathcal{D} = D_{V_1} + \frac{1}{2} \int_{Q_1} \mathbb{E}\Big[\sum_{j:(\mathbf{0},j) \in \mathcal{G}_1(\mathcal{P}_{\lambda\tau(x)} \cup \{\mathbf{0}\})} \xi_1(R_{\mathbf{0}j}) \Big] \tau(x) dx, \tag{17}$$

where D_{V_1} is the node KL-divergence given by (9).

Proof: The first term follows from LLN for i.i.d variables. For the second term, ξ_1 is a stabilizing functional since it is a functional of edges of a stabilizing graph $\mathcal{G}_1$ and bounded-moments condition in (16) holds. Hence, the LLN in [12] guarantees L^2 convergence to the above constant, which in turn implies convergence in probability. $\square$

Remark 1: When the node locations are uniform ($\tau(x) \equiv 1$), the error exponent in (17) simplifies as

$$\mathcal{D} = D_{V_1} + \frac{1}{2}\mathbb{E}\Big[\sum_{j:(\mathbf{0},j)\in\mathcal{G}_1(\mathcal{P}_\lambda\cup\{\mathbf{0}\})} \xi_1(R_{\mathbf{0}j})\Big]. \tag{18}$$

B. General Hypothesis Testing

In this section, we extend the results to any general distribution under the null hypothesis. For such cases, we cannot directly use the LLN for marked point process to evaluate (13), since the marks are required to be i.i.d. for the LLN to hold.

We now additionally assume *uniform integrability* [13, (16.21)] to convert the functional on a marked point process in (8) to a functional on an unmarked process. In Section V, we show that the Gaussian distribution satisfies uniform integrability.

Proposition 1 (Uniform Integrability): When the normalized spectrum, given by the sequence $\{\frac{1}{n}L(\mathbf{Y}_{\mathbf{V}_n})\}_{n\geq 1}$ is uniformly integrable and converges in probability under $\mathcal{H}_0$, the error exponent in (8) is the KL-divergence rate,

$$\begin{aligned}\mathcal{D} &= \lim_{n\to\infty}\frac{D_{\mathbf{V}_n}}{n}, &(19)\\ &= \mathrm{p}\lim_{n\to\infty}\frac{1}{n}\Big[\sum_{a\in\mathcal{C}_1}\mathbb{E}(\psi_{1,a}(\mathbf{Y}_a)\mid\mathbf{V}_n,\mathcal{H}_0)\\ &\quad -\sum_{b\in\mathcal{C}_0}\mathbb{E}(\psi_{0,b}(\mathbf{Y}_b)\mid\mathbf{V}_n,\mathcal{H}_0)+\log\frac{Z_1(\mathbf{V}_n)}{Z_0(\mathbf{V}_n)}\Big], &(20)\end{aligned}$$

where $D_{\mathbf{V}_n}$ is the KL-divergence in (6), $\psi_{i,c}$ is potential of clique $c\in\mathcal{C}_i$ of the MRF under hypothesis $\mathcal{H}_i$ in (7).

Proof:
$$\begin{aligned}\mathcal{D} &= \lim_{n\to\infty}\frac{D_{\mathbf{V}_n}}{n} = \lim_{n\to\infty}\frac{1}{n}\mathbb{E}[L(\mathbf{Y}_{\mathbf{V}_n})\mid\mathcal{H}_0], &(21)\\ &= \mathrm{p}\lim_{n\to\infty}\frac{1}{n}\mathbb{E}[L(\mathbf{Y}_{\mathbf{V}_n})\mid\mathbf{V}_n,\mathcal{H}_0]. &(22)\end{aligned}$$

Now evaluating the conditional expectation using the form of LLR for a MRF in (7), we have the result. $\square$

Hence, we have (20), which is a functional on an unmarked process. Since the clique potential functions in (20) are parameterized by the node locations, (20) is a functional over a random graph. Note that we do not need the dependency graphs to be acyclic for the above result. We now specialize the above result for acyclic dependency graphs.

Lemma 2 (Acyclic Graphs): For acyclic graphs $\mathcal{G}_0$ and $\mathcal{G}_1$,

$$\begin{aligned}\mathcal{D} =& D_{V_1} + \mathrm{p}\lim_{n\to\infty}\frac{1}{n}\Big[\sum_{\substack{(i,j)\in\mathcal{G}_1\setminus\mathcal{G}_0\\ i<j}}\xi_1(R_{ij})\\ &+\sum_{\substack{(i,j)\in\mathcal{G}_0\cap\mathcal{G}_1\\ i<j}}\xi_2(R_{ij})\sum_{\substack{(i,j)\in\mathcal{G}_0\setminus\mathcal{G}_1\\ i<j}}\xi_3(R_{ij})\Big], &(23)\end{aligned}$$

where D_{V_1} and ξ_1 are given by (9) and (15), and the edge functionals ξ_2 and ξ_3 are defined as

$$\begin{aligned}&\xi_2(r_{ij}){:=}\mathbb{E}\Big[\log\frac{h_0(Y_i,Y_j|R_{ij}{=}r_{ij})}{h_1(Y_i,Y_j|R_{ij}{=}r_{ij})}\Big|R_{ij}{=}r_{ij},\mathcal{H}_0\Big]{-}2D_{V_1} &(24)\\ &\xi_3(r_{ij}){:=}I(Y_i;Y_j\mid R_{ij}{=}\,r_{ij},\mathcal{H}_0), &(25)\end{aligned}$$

where $I(X;Y)$ is mutual information between X and Y and $I(X;Y\mid Z=z)$ is mutual information conditioned on $Z=z$.

Proof: From (13) and Proposition 1. $\square$

We now provide the error exponent for MRF hypotheses.

Theorem 1 (Exponent For Stabilizing Acyclic Graphs): When ξ_i for $i=1,2,3$ satisfy the bounded-moments condition in (16), the error exponent for stabilizing acyclic dependency graphs is given by

$$\mathcal{D} = D_{V_1} + \frac{1}{2}\sum_{i=1}^{3}\int_{Q_1}\mathbb{E}\left[\sum_{j:(\mathbf{0},j)\in E_{i,\tau(x)}}\xi_i(R_{\mathbf{0}j})\right]\tau(x)dx, \tag{26}$$

where $E_{1,\tau(x)}{:=}\mathcal{G}_1\setminus\mathcal{G}_0(\mathcal{P}_{\lambda\tau(x)}\cup\{\mathbf{0}\})$, $E_{2,\tau(x)}{:=}\mathcal{G}_0\cap\mathcal{G}_1(\mathcal{P}_{\lambda\tau(x)}\cup\{\mathbf{0}\})$, and $E_{3,\tau(x)}{:=}\mathcal{G}_0\setminus\mathcal{G}_1(\mathcal{P}_{\lambda\tau(x)}\cup\{\mathbf{0}\})$.

Proof: Since $\mathcal{G}_0$ and $\mathcal{G}_1$ are stabilizing, its subgraphs with edges $E_{i,\tau(x)}$ for $i=1,2,3$ can be shown to be stabilizing. The moments condition in (16) holds. Hence, the LLN follows. $\square$

Remark 2: When the node locations are uniform ($\tau(x)\equiv 1$), the error exponent in (26) simplifies as

$$\mathcal{D} = D_{V_1} + \frac{1}{2}\sum_{i=1}^{3}\mathbb{E}\sum_{j:(\mathbf{0},j)\in E_{i,1}}\xi_i(R_{\mathbf{0}j}). \tag{27}$$

V. Gaussian Distribution on Acyclic Graphs

In this section, we simplify the results of the previous section on acyclic graphs when the distribution under each hypothesis $\mathcal{H}_m$ is Gaussian $\mathcal{N}(\boldsymbol{\mu}_m,\mathbf{\Sigma}_{m,\mathbf{v}_n})$, given the node locations $\mathbf{V}_n=\mathbf{v}_n$. In this case, the MRF factorization in (2) leads to a special relationship between the coefficients of the covariance matrix and its inverse, called the *potential matrix* . Specifically, there is a one-to-one correspondence between the non-zero elements of the potential matrix $\mathbf{\Sigma}^{-1}_{m,\mathbf{v}_n}$ and the dependency graph edges $\mathcal{G}_m(\mathbf{v}_n)$. Moreover, for acyclic graphs $\mathcal{G}_m(\mathbf{v}_n)$, further simplifications are possible [11, Thm. 1].

The additional constraints of spatial dependence for acyclic graphs in (11) and (12) imply that under each hypothesis, the mean and the variance at all the nodes are equal and that the correlation coefficient between any two neighboring nodes is only dependent on the inter-node distance, i.e., under hypothesis $\mathcal{H}_m$, for $m=0,1$, we have $\boldsymbol{\mu}_m=\mu_m\mathbf{I}$, $\Sigma_{m,\mathbf{v}_n}(i,i,)=\sigma_m^2$, and for $(i,j)\in\mathcal{G}_m(\mathbf{v}_n)$, we have $\Sigma_{m,\mathbf{v}_n}(i,j)=\rho_m(R_{ij})\sigma_m^2$. Here, the correlation function $\rho_m(\cdot)<1$ is positive and monotonically decreasing in the edge length, for each $m=0,1$.

With the above assumptions, the covariance matrix under hypothesis $\mathcal{H}_m$ is given by

$$\Sigma_{m,\mathbf{v}_n}(i,j) = \begin{cases} \sigma_m^2 > 0, & i = j, \quad (28a) \\ \sigma_m^2 \prod\limits_{(a,b)\in \text{Path}(i,j;\mathcal{G}_m(\mathbf{v}_n))} \rho_m(R_{a,b}), & \text{o.w.} \quad (28b) \end{cases}$$

where $\text{Path}(i,j;\mathcal{G}_m(\mathbf{v}_n))$ is the set of edges of the acyclic graph $\mathcal{G}_m(\mathbf{v}_n)$ belonging to the *unique* path[2] connecting the nodes i and j. It can be shown that $\Sigma_{m,\mathbf{v}_n}$ in (28) is positive definite for any node configuration $\mathbf{v}_n$ when $\rho_m(\cdot) < 1$.

Under the above assumptions, we now provide closed-form expression for the Gaussian error exponent. Recall that $\text{Path}(\mathbf{0},j;\mathcal{G}_0)$ denotes the set of edges in $\mathcal{G}_0$ connecting the origin $\mathbf{0}$ with some node j. Let $\Delta\mu{:=}\mu_1 - \mu_0, \quad K{:=}\frac{\sigma_1^2}{\sigma_0^2}$.

Theorem 2 (Gaussian Error Exponent): For Gaussian distribution under each hypothesis, the error exponent is given by (26), with the terms simplifying as

$$D_{V_1} = \frac{1}{2}\Big(\log(K) + \frac{1}{K} - 1 + \frac{\Delta\mu^2}{\sigma_1^2}\Big), \tag{29}$$

$$\xi_1(R_{\mathbf{0}j}) = \frac{\rho_1(R_{\mathbf{0}j})[\rho_1(R_{\mathbf{0}j}) - \prod\limits_{(k,l)\in\text{Path}(\mathbf{0},j;\mathcal{G}_0(\mathcal{P}_{\lambda\tau(x)}\cup\mathbf{0}))} \rho_0(R_{kl})]}{[1-\rho_1^2(R_{\mathbf{0}j})]K} + \frac{\log[1-\rho_1^2(R_{\mathbf{0}j})]}{2}, \tag{30}$$

$$\xi_2(R_{\mathbf{0}j}) = \frac{\rho_1(R_{\mathbf{0}j})[\rho_1(R_{\mathbf{0}j}) - \rho_0(R_{\mathbf{0}j})]}{[1-\rho_1^2(R_{\mathbf{0}j})]K} + \frac{1}{2}\log\frac{1-\rho_1^2(R_{\mathbf{0}j})}{1-\rho_0^2(R_{\mathbf{0}j})} - \frac{\Delta\mu^2\rho_1(R_{\mathbf{0}j})}{\sigma_1^2(1+\rho_1(R_{\mathbf{0}j}))}, \tag{31}$$

$$\xi_3(R_{\mathbf{0}j}) = -\frac{\log[1-\rho_0^2(R_{\mathbf{0}j})]}{2}. \tag{32}$$

Proof: From [11, Thm. 1], we have the expressions for determinant and potential matrix coefficients for acyclic graphs, and we use them to simplify terms in the error exponent.

The moments condition in (16) holds for ξ_m for $m = 1,2,3$ since the terms are bounded for correlation functions $\rho_k(R_{ij})$ which are decreasing in edge lengths and $\rho_k(0) < 1$. For uniform integrability [13, (16.21)] of normalized spectrum, it is sufficient to show that for any $n > 0$

$$\lim_{\alpha\to\infty} \int\limits_{|\mathbf{y}^T(\mathbf{\Sigma}_0^{-1}-\mathbf{\Sigma}_1^{-1})\mathbf{y}|\geq n\alpha} \frac{1}{n}|\mathbf{y}^T(\mathbf{\Sigma}_0^{-1}-\mathbf{\Sigma}_1^{-1})\mathbf{y}| \exp[-\frac{\mathbf{y}^T\mathbf{\Sigma}_0^{-1}\mathbf{y}}{2}]\mathbf{dy} = 0$$

From positive definiteness, this reduces to showing

$$\lim_{\alpha\to\infty} \int\limits_{|\mathbf{y}^T(\mathbf{\Sigma}_0^{-1}+\mathbf{\Sigma}_1^{-1})\mathbf{y}|\geq \alpha} |\mathbf{y}^T(\mathbf{\Sigma}_0^{-1}+\mathbf{\Sigma}_1^{-1})\mathbf{y}| \exp[-\frac{\mathbf{y}^T\mathbf{\Sigma}_0^{-1}\mathbf{y}}{2}]\mathbf{dy} = 0,$$

which is true. □

[2] $\Sigma_{m,\mathbf{v}_n}(i,j) = 0$ if no path exists between i and j in $\mathcal{G}_m$.

VI. Conclusion

In this paper, we considered hypothesis testing of spatially-dependent Markov random field observations collected from randomly located sensors. We derived the error exponent by appealing to the law of large number results for random graph functionals. This allows us to study the influence on the node placement distribution on the error exponent. In [15], [16], we addressed the issue of energy consumption for routing measurements for testing of MRF hypotheses, and proposed energy-efficient fusion schemes. In future, we plan to build on these works to investigate efficient node placement strategies that maximize the error exponent as well as meet energy constraints.

Acknowledgement

The authors thank J.E. Yukich, A. Swami and D. Shah for extensive comments and discussions.

References

[1] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1994.

[2] T. Han, "Hypothesis Testing with the General Source," *IEEE Tran. Inform. Theory*, vol. 46, no. 7, pp. 2415–2427, Nov. 2000.

[3] T. Cover and J. Thomas, *Elements of Information Theory*. John Wiley & Sons, Inc., 1991.

[4] P.-N. Chen, "General formulas for the Neyman-Pearson type-II error exponent subject to fixed and exponential type-I error bounds," *IEEE Tran. on Information Theory*, vol. 42, no. 1, pp. 316–323, Jan. 1996.

[5] H. Künsch, "Thermodynamics and Statistical Analysis of Gaussian Random Fields," *Probability Theory and Related Fields*, vol. 58, no. 3, pp. 407–421, 1981.

[6] Y. Sung, H. Yu, and H. V. Poor, "Information, Energy and Density for Ad-hoc Sensor Networks over Correlated Random Fields: Large-deviation Analysis," in *IEEE ISIT*, July 2008, pp. 1592–1596.

[7] Z. Rached, F. Alajaji, and L. Campbell, "The Kullback-Leibler divergence rate between Markov sources," *Information Theory, IEEE Transactions on*, vol. 50, no. 5, pp. 917–921, 2004.

[8] L. Xie, V. Ugrinovskii, and I. Petersen, "Probabilistic distances between finite-state finite-alphabet hidden Markov models," *Automatic Control, IEEE Transactions on*, vol. 50, no. 4, pp. 505–511, 2005.

[9] H. Liang, R. Anderson-Sprecher, R. Kubichek, and G. Talwar, "A Novel Approach to Approximate Kullback-Leibler Distance Rate for Hidden Markov Models," in *Asilomar Conf. on Signals, Systems and Computers*, 2005, pp. 869–873.

[10] M. Vidyasagar, "Bounds on the kullback-leibler divergence rate between hidden markov models," in *Conf. on Decision and Control*, Dec. 2007, pp. 6160–6165.

[11] A. Anandkumar, L. Tong, and A. Swami, "Detection of Gauss-Markov Random Fields with Nearest-neighbor Dependency," *IEEE Tran. Information Theory*, vol. 55, no. 2, pp. 816–827, Feb. 2009.

[12] M. Penrose and J. Yukich, "Weak Laws Of Large Numbers In Geometric Probability," *Annals of Applied Probability*, vol. 13, no. 1, pp. 277–303, 2003.

[13] P.Billingsley, *Probability and Measure*. New York, NY: Wiley Inter-Science, 1995.

[14] P. Brémaud, *Markov Chains: Gibbs fields, Monte Carlo simulation, and queues*. Springer, 1999.

[15] A. Anandkumar, L. Tong, and A. Swami, "Optimal Node Density for Detection in Energy Constrained Random Networks," *IEEE Tran. Signal Proc.*, vol. 56, no. 10, pp. 5232–5245, Oct. 2008.

[16] A. Anandkumar, J. Yukich, L. Tong, and A. Swami, "Energy Scaling Laws for Distributed Inference in Random Networks," *Accepted to IEEE J. Selec. Area Comm., available on Arxiv*, Dec. 2008.