

Detection of Gauss-Markov Random Fields under Routing Energy Constraint

Animashree Anandkumar
ECE Department
Cornell University
Ithaca, NY 14853
aa332@cornell.edu

Lang Tong
ECE Department
Cornell University
Ithaca, NY 14853
ltong@ece.cornell.edu

Ananthram Swami
Army Research Laboratory
Adelphi
MD 20783
a.swami@ieee.org

Abstract—The problem of maximizing detection performance subject to an energy constraint is analyzed in an asymptotic setting. The correlation between the sensor measurements is incorporated through the Gauss-Markov random field with Euclidean nearest-neighbor dependency graph. An average energy constraint is imposed on a routing scheme with an approximation factor of two, and the resulting Neyman-Pearson error exponent is optimized with respect to the density of the deployed sensors. It is shown that the behavior of this optimal density crucially depends on the ratio between the measurement variances under the two hypotheses and displays a threshold behavior. Below the threshold value of the variance ratio, the optimal density tends towards infinity for any feasible energy constraint. On the other hand, when the variance ratio is above the threshold, the optimal density is characterized by the energy constraint.

Index Terms—detection, Gauss-Markov random fields, Routing, error exponent.

I. INTRODUCTION

The deployment of sensors is the first step in establishing a network. It influences the performance of the network, including the energy consumed in routing data. However, designing optimal deployment strategies requires finding optimal locations for all nodes and is not feasible for a large network. In contrast, deploying nodes randomly but according to an optimal density may be more tractable. In this paper, we assume that the nodes are placed IID according to the uniform or Poisson distribution with a fixed node density (Fig.1a).

In a network of sensors measuring a correlated signal field, the node density influences the extent of correlation among the measurements and thereby detection performance. Moreover, the energy required to route data typically depends on the inter-node distance, and in turn, the node density. Hence, we can design an optimal density that achieves the best tradeoff between the energy consumed in routing data and the resulting detection performance. Of course, the optimal density depends on the routing protocols employed.

This work was supported in part through the collaborative participation in the Communications and Networks Consortium sponsored by the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-2-0011 and by the Army Research Office under Grant ARO-W911NF-06-1-0346. The third author was partially supported by the DARPA IT-MANET program. The U. S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon.

Part of this work is presented in [1], [2].

In the classical approach, a layered architecture separates the design of routing from the application. For energy-constrained networks, application-specific routing may offer better tradeoffs, and we employ one such routing scheme.

To characterize the detection performance, we consider the Neyman-Pearson (NP) error exponent. Our objective is to find an optimal node density λ_* that maximizes the detection error exponent D , under a constraint \bar{E} on the average energy consumption \bar{C} of routing per node,

$$\lambda_* \triangleq \arg \max_{\lambda > 0} D_\lambda \quad \text{subject to } \bar{C} \leq \bar{E}. \quad (1)$$

Note that both the energy constraint and detection performance are asymptotic in the number of nodes.

We address the following questions: does an optimal density exist? And if so, what is its value? Is it one of the extremes, viz., zero or infinity? This is an important question, since if the optimal node density is either zero or infinity, then we can simply place the nodes in as small/large an observation area as possible. On the other hand, if this is not the case, then we need to deploy the nodes, based on the optimal density.

II. RELATED WORK

Detection of Gauss-Markov random field (GMRF) is related to the detection of Gauss-Markov random processes (GMRP), which is a classical problem [3]. There is an extensive literature on the large-deviations approach to the analysis of detection of GMRP, e.g., [4], [5]. However, the use of GMRP is not easily amenable to random placement of nodes in planar and higher dimensional spaces. Hence, we employ the model of a GMRF for randomly placed sensors on the Euclidean plane.

An overview of routing in wireless networks can be found in [6]. We are interested in routing spatially-correlated data for optimal detection and hence, we can undertake aggregation along the path. Such in-network processing algorithms for computing certain aggregate functions are dealt in the surveys [7], [8]. These data-aggregation schemes, however, are applicable for optimal detection, only for conditionally independent data.

The literature considering the tradeoff between detection performance and routing energy is less extensive. In [9]–[11],

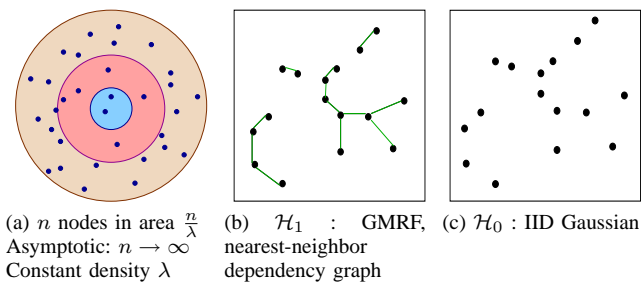


Fig. 1: Detection problem and random node deployment.

energy-efficient detection is considered under various formulations. However, they all assume that the measurements are conditionally independent. Detection of one-dimensional GMRP is considered in [12], [13]. A specific link cost for detection, based on the Chernoff information, is proposed in [12]. In [13], a power-density constraint is imposed, i.e., the power per unit distance is fixed, so that the power per node decreases linearly with the node density and a finite sensor density is shown to be optimal. We instead consider a per-node average energy constraint. Also, these works assume the one-dimensional model, not applicable when the nodes are on a plane. In [14], local vote decision fusion is proposed for non-parametric target detection; binary decisions are fused locally by a majority rule before transmission to a fusion center. However, no guarantee is provided on the energy savings obtained. In [15], a decision-theoretic approach to inference of a Markov random field with single-bit communication is considered and the network topology is predefined by a directed acyclic graph.

III. SUMMARY OF RESULTS

As a motivation to our results, we first consider the special case when the measurements are IID, conditioned under either hypothesis. In this case, the log-likelihood ratio (LLR) is a sum function, with each term depending only on one node measurement. Since the signal fields are uniform under both the hypotheses (same measurement variance at all the nodes), the detection performance is unaffected by the nature of node deployments. Hence, we only need to consider the effect of node density on routing.

The minimum-energy routing in this case is given by the directed minimum spanning tree (DMST), with the directions toward the fusion center. See Fig.2. The sum function can be calculated hierarchically along DMST, starting at the leaves and ending at the fusion center. It is intuitively clear that increasing the node density decreases the (random) node distances on average. Since the energy consumption typically follows a distance-attenuation law, node density has an inverse relationship with energy consumption. As the node density goes to infinity, the energy constraint is always satisfied and hence, infinite density is always a solution to the optimal tradeoff problem. However, incorporating correlation introduces new tradeoffs in the problem.

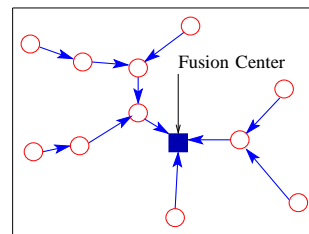


Fig. 2: For IID data, LLR is a sum function. Minimum energy routing is hierarchical aggregation along the minimum spanning tree, directed towards the fusion center.

We assume the presence of correlation under the alternative hypothesis \mathcal{H}_1 through the Gauss-Markov random field (GMRF) model. See (Fig.1b, 1c). Again, assuming an uniform signal field model (same measurement variance at every node) and a correlation function decaying with distance, we study the effect of node density on the error exponent. To this end, we unify the results from our previous works [1], [2], which independently characterize the detection performance and routing for a GMRF.

Intuitively, when both the hypotheses have the same measurement variance, they can be distinguished only by the presence of correlation under \mathcal{H}_1 . Correlation is maximized when all the nodes are clustered close to one another, since correlation decays with distance. Hence, ignoring the energy constraint, the optimal density should be infinite. We prove that this is indeed true.

For the general case, when the measurement variances are different, the behavior is decided by the ratio of the variances under \mathcal{H}_1 and \mathcal{H}_0 . We show that the optimal density is infinite when the variance ratio is below a threshold value (Fig.3). Moreover, imposing any feasible routing-energy constraint does not change the optimal density. We provide the expression for the threshold on the variance ratio.

We then show that the presence of correlation degrades the error exponent when the variance ratio is above the threshold. The optimal density is zero in this regime, if there is no energy constraint (Fig.3a), i.e., the nodes should be placed as far apart as possible. We then show that imposing a finite energy constraint leads to a strictly positive optimal density, for variance ratios above another threshold (Fig.3b).

Finding a minimum energy routing scheme that delivers the likelihood ratio of a Markov random field has been shown to be NP-hard [2] and hence, finding the optimal density in this regime is also NP-hard. We provide a bound on the optimal density, by analyzing the energy consumption of a 2-approximate routing scheme, proposed in [2]. Numerical investigation shows that the bound holds even in the gap region between the two thresholds.

Additionally, we analyze the extent to which the error exponent is influenced by the presence of correlation and in turn, by the value of node density. This enables us to gauge the usefulness of finding the optimal density. As expected, the maximum contribution from correlation to the

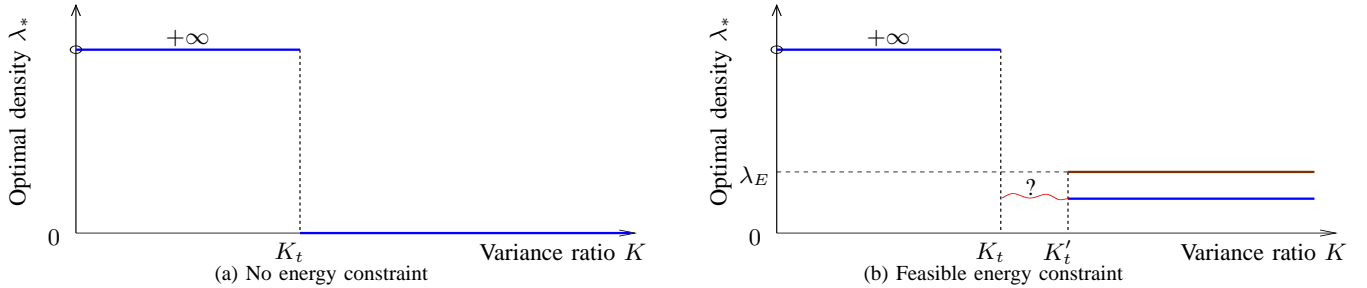


Fig. 3: Optimal node density vs. ratio of variances under \mathcal{H}_1 and \mathcal{H}_0 . See Theorem 5. λ_E is a bound on λ_* for $K > K'_t$.

error exponent is when the variances are equal under the two hypotheses. On the other hand, we show that at large variance ratio, the contribution from correlation decays to zero, implying that the error exponent is mostly insensitive to changes in node density.

IV. SYSTEM MODEL

An undirected graph \mathcal{G} is a tuple $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where \mathcal{V} is the vertex set and $\mathcal{E} = \{(i, j), i, j \in \mathcal{V}, i < j\}$ is the edge set. Let R_{ij} be the random variable denoting Euclidean edge-length of (i, j) . The neighborhood function of a node i is the set of all other nodes having an edge with it. For a directed graph, we denote the edges by $\mathcal{E} = \{< i, j >, i, j \in \mathcal{V}, i \neq j\}$, where the direction of the edge is from i to j .

For the matrix $\mathbf{A} = [A(i, j)]$, $A(i, j)$ denotes the element in the i^{th} row and j^{th} column and $|\mathbf{A}|$ its determinant. For two sets A and B , let $A \setminus B = \{i : i \in A, i \notin B\}$.

A. Gauss-Markov Random Field

A GMRF, in addition to being a Gaussian random field, satisfies special conditional independence properties. A simple example is the first-order auto-regressive process, where the conditional independence of the observations is based on causality. However, a spatial random field has a far richer set of conditional independencies, requiring a more general definition [16, p. 21].

Definition 1 (GMRF): Given a point set $\mathcal{V} = \{1, \dots, n\}$, $\mathbf{Y}_{\mathcal{V}} = \{Y_i : i \in \mathcal{V}\}$ is a GMRF with an (undirected) dependency graph $\mathcal{G}_d(\mathcal{V}, \mathcal{E})$ if $\mathbf{Y}_{\mathcal{V}}$ is a Gaussian random field, and $\forall i, j \in \mathcal{V}$, Y_i and Y_j are conditionally independent given observations at all other nodes if i and j are not neighbors,

$$Y_i \perp Y_j | \mathbf{Y}_{-ij} \iff (i, j) \notin \mathcal{E}_d, \forall i, j \in \mathcal{V}, i \neq j, \quad (2)$$

where \perp denotes conditional independence and $\mathbf{Y}_{-ij} \triangleq (Y_k : k \in \mathcal{V}, k \neq i, j)$.

B. Network and Energy Model

We assume that the energy consumed by a node i can be represented by the sum of a constant processing energy $C_p >$

0 and the transmission energy¹ [17]. For details, see [2]. The energy required for the transmission of a real number from i to j is denoted by $C_{i,j}$. We assume that

$$C_{i,j} = C_t |\text{dist}(i, j)|^\nu, 2 \leq \nu \leq 4, \quad (3)$$

where C_t is a constant, $\text{dist}(i, j)$ is the inter-node distance and ν is the channel path-loss exponent. Further, we assume that the processing energy C_p at every node is constant and independent of the number of transmissions from the node.

The set of transmissions in a node set \mathcal{V} can be represented as a digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Note that this transmission digraph is different from the dependency graph of the signal field. Let $C(\mathcal{G})$ denote the total transmission energy given by

$$C(\mathcal{G}) \triangleq C_t \sum_{e \in \mathcal{E}} R_e^\nu, \quad (4)$$

where r_e is the Euclidean edge length. We further require that $\mathcal{G} \subset \text{UDG}(\mathcal{V})$, the unit-disk graph, defined as the set of edges between any two nodes within an unit Euclidean distance of each other.

V. HYPOTHESIS-TESTING PROBLEM

Let $\mathcal{V} = \{1, \dots, n\}$ be a set of n nodes on the plane and let \mathbf{Y}_n be the random vector of observation samples $Y_i, i \in \mathcal{V}$,

$$\mathbf{Y}_n \triangleq [Y_1, \dots, Y_n]^T. \quad (5)$$

The hypothesis-testing problem is (Fig.1b,1c),

$$\mathcal{H}_0 : \mathbf{Y}_n \sim \mathcal{N}(\mathbf{0}, \sigma_0^2 \mathbf{I}) \quad \text{vs.} \quad \mathcal{H}_1 : \mathbf{Y}_n \sim \mathcal{N}(\mathbf{0}, \Sigma_{1,\mathcal{V}}), \quad (6)$$

where the covariance matrix $\Sigma_{1,\mathcal{V}}$ of a GMRF depends on the configuration of nodes in \mathcal{V} .

The optimal decision-rule under the Neyman-Pearson formulation is a threshold test based on the log-likelihood ratio (LLR). Let $p[\mathbf{Y}_n | \mathcal{V}; \mathcal{H}_j]$ be the conditional PDF of the observations given the point set \mathcal{V} under hypothesis j .

$$\text{LLR}(\mathbf{Y}_n, \mathcal{G}_d) \triangleq \log \frac{p[\mathbf{Y}_n, \mathcal{V}; \mathcal{H}_0]}{p[\mathbf{Y}_n, \mathcal{V}; \mathcal{H}_1]}. \quad (7)$$

¹We ignore the energy consumed at the receiver.

A. Covariance matrix of GMRF

We make additional assumption on the structure of the covariance matrix $\Sigma_{1,\mathcal{V}}$ of the GMRF under \mathcal{H}_1 viz., that the nodes have the same measurement variance for any node configuration \mathcal{V} , i.e.,

$$\Sigma_{1,\mathcal{V}}(i,i) \triangleq \sigma_1^2, \quad i = 1, \dots, n. \quad (8)$$

We denote the ratio between the variances under the alternative and the null hypothesis at each node by

$$K \triangleq \frac{\sigma_1^2}{\sigma_0^2}. \quad (9)$$

We also assume that under \mathcal{H}_1 , the amount of correlation between the neighbors i, j of the dependency graph is specified by an arbitrary function g , which has the Euclidean edge length R_{ij} as its argument. From (8), we have

$$g(R_{ij}) \triangleq \frac{\Sigma_{1,\mathcal{V}}(i,j)}{\sigma_1^2}, \quad \forall (i,j) \in \mathcal{E}. \quad (10)$$

The correlation function g is required to satisfy some regularity conditions. We assume that

$$g(0) \triangleq M < 1, \quad g(\infty) = 0, \quad (11)$$

and that g is a monotonically decreasing convex function of the edge length. This is a reasonable assumption, since the amount of correlation usually decays as nodes become farther apart. Moreover, partial correlation or $g(0) = M < 1$ arises due to signal imperfections. Note that this has the same effect on correlation as imposing an exclusion region on how near two nodes can be placed. Some examples of the correlation functions, satisfying our assumptions, are

$$g(r) = Me^{-ar}, \quad g(r) = \frac{M}{1+r^a}, \quad 0 \leq a, M < 1.$$

Note that we do not assume the presence of additive Gaussian measurement noise in the signal model, as considered in the literature before. This is because with the additive noise, the resulting GMRF no longer has a sparse dependency graph. As an alternative formulation, we model imperfect measurements by assuming partially-correlated nodes.

B. Expression for log-likelihood ratio

We incorporate the assumptions (8-10) in the theorem below to obtain the LLR for detection.

Theorem 1 (Log-likelihood ratio): Under the assumptions (8-10), the log-likelihood ratio in (7) for the hypothesis-testing problem in (6), given point set $\mathcal{V} = \{1, \dots, n\}$, is

$$\begin{aligned} \text{LLR}(\mathbf{Y}_n, \mathcal{G}_d) &= n \log \frac{\sigma_1}{\sigma_0} + \frac{1}{2} \left[\sum_{i \in \mathcal{V}} \left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_0^2} \right) Y_i^2 \right. \\ &+ \sum_{(i,j) \in \mathcal{E}} \left\{ \log[1 - g^2(R_{ij})] \right. \\ &+ \frac{g^2(R_{ij})}{1 - g^2(R_{ij})} \frac{Y_i^2 + Y_j^2}{\sigma_1^2} \\ &\left. \left. - \frac{2g(R_{ij})}{1 - g^2(R_{ij})} \frac{Y_i Y_j}{\sigma_1^2} \right\} \right], \quad (12) \end{aligned}$$

where R_{ij} is the Euclidean edge length of $(i, j) \in \mathcal{E}$.

Theorem 1 gives a closed-form expression for the log-likelihood ratio, in terms of the edges of the nearest-neighbor dependency graph of the GMRF. Note in (12), the cross-terms are only between the neighbors of the dependency graph, which is exploited to yield explicit data-fusion and routing schemes [2], summarized in section VI. Note that (12) contains two kinds of terms. The *edge potentials*, denoted by $\phi_{i,j}(Y_i, Y_j)$, that depend on a pair of observations, given by

$$\begin{aligned} \phi_{i,j}(i,j) &\triangleq \frac{1}{2} \log[1 - g^2(R_{ij})] - \frac{g(R_{ij})}{1 - g^2(R_{ij})} \frac{Y_i Y_j}{\sigma_1^2} \\ &+ \frac{g^2(R_{ij})}{1 - g^2(R_{ij})} \frac{Y_i^2 + Y_j^2}{2\sigma_1^2}, \quad (13) \end{aligned}$$

and the *node potentials* $\phi_i(Y_i)$, depending on a single observation, given by

$$\phi_i(Y_i) \triangleq \log \frac{\sigma_1}{\sigma_0} + \frac{1}{2} \left(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_0^2} \right) Y_i^2. \quad (14)$$

Hence, the LLR can be written in a compact form

$$\text{LLR}(\mathbf{Y}_n; \mathcal{G}_d) = \sum_{i \in \mathcal{V}} \phi_i(Y_i) + \sum_{(i,j) \in \mathcal{E}_d} \phi_{i,j}(Y_i, Y_j). \quad (15)$$

VI. MINIMUM ENERGY ROUTING

The aim of optimal routing for detection is to ensure the delivery of the LLR, given by (15), to the designated fusion center v_0 , while minimizing the total energy consumption. However, this is NP-hard and we instead employ a 2-approximation routing algorithm for data-fusion of a Markov random field (DFMRF). See Fig. 4 and [2]. We define two types of transmissions for routing.

Definition 2: The *data-transmission graph* $\text{DTG}(\mathcal{V})$ is defined as the transmission subgraph that consists of transmissions of the raw sensor data. The *likelihood-aggregation graph* $\text{AG}(\mathcal{V})$ is defined as the transmission subgraph that consists of transmissions of the aggregates of the log-likelihood ratio. The nodes that process data from other nodes i.e., those in the likelihood-aggregation graph, are known as the set of *aggregators*, denoted by \mathcal{V}_{AG} .

The data-transmission graph $\text{DTG}(\mathcal{V})$ needs to compute all the edge potentials of the LLR in (15) of the form

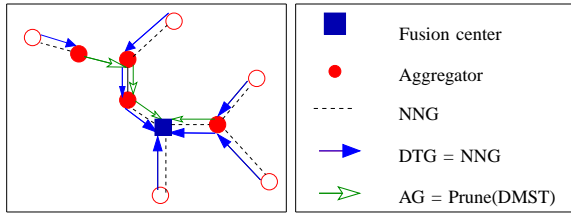


Fig. 4: DFMRF routing delivering LLR to fusion center.

$\phi_{i,j}(Y_i, Y_j)$. These edge potentials exist only between the neighbors in the dependency graph \mathcal{G}_d . Therefore, we consider $\text{DTG}(\mathcal{V})$ as \mathcal{G}_d itself, with arbitrarily assigned directions. The data-fusion mechanism has two phases. During the data-transmission phase, transmissions of raw data occurs along the $\text{DTG}(\mathcal{V})$. At the end of this phase, every internal node i computes

$$m(i) = \phi_i(Y_i) + \sum_{\langle j,i \rangle \in \text{DTG}(\mathcal{V})} \phi_{i,j}(Y_i, Y_j) + \sum_{\langle j,i \rangle \in \text{DTG}(\mathcal{V}), j \notin \mathcal{V}_{\text{AG}}} \phi_i(Y_j). \quad (16)$$

This local contribution is then aggregated and delivered to the fusion center, along the aggregation graph $\text{AG}(\mathcal{V})$, given by $\text{DMST}(\mathcal{V})$, the minimum spanning tree, directed towards the fusion center. On receiving the aggregates from all its direct predecessors in $\text{AG}(\mathcal{V})$, each node $i \neq v_0$ combines them with its local contribution and transmits $l(i)$ to its direct successor,

$$l(i) = \sum_{\langle j,i \rangle \in \text{AG}(\mathcal{V})} l(j) + m(i). \quad (17)$$

At the end of the aggregation process, the LLR is available at the fusion center and is given by $l(v_0)$.

Theorem 2 (Approx. for Arbitrary Placement): DFMRF is a 2-approximation algorithm for any arbitrary placement.

$$\frac{C(\text{DFMRF}(\mathcal{V}))}{C(\mathcal{G}_*(\mathcal{V}))} \leq 2. \quad (18)$$

VII. ASYMPTOTIC ANALYSIS

It is intractable to evaluate the routing energy consumption or the detection performance for an arbitrary point set. Therefore, we assume that the nodes are placed randomly, according to a point process defined on expanding regions. Let $(\mathcal{B}_n)_{n \geq 1}$ denote a sequence of “regular” Borel sets (such as squares or circles) of area $\frac{n}{\lambda}$, centered at the origin, for any positive constant λ . Let $\mathcal{U}_{n,\lambda}$ be a binomial point process on \mathcal{B}_n with intensity λ , i.e., n nodes distributed i.i.d uniform on the region \mathcal{B}_n with node density λ . Let $\mathcal{P}_{n,\lambda}$ be the homogeneous Poisson process on \mathcal{B}_n with constant intensity λ . The number of nodes n goes to infinity, with fixed node density λ (Fig.1a).

A. Average Energy for DFMRF

We now analyze the average energy consumption for the DFMRF scheme. The average energy for DFMRF under the binomial or Poisson process can be quantified using the law of large numbers (LLN) for graph functionals [18].

Theorem 3 (Average energy for DFMRF): The asymptotic average energy consumption of the optimal data-transmission graph is given by

$$\bar{C} \triangleq \lim_{n \rightarrow \infty} \frac{C(\text{DFMRF})}{n} \triangleq C_t \lim_{n \rightarrow \infty} \sum_{(i,j) \in \text{DFMRF}} \frac{R_{ij}^\nu}{n} + C_p = C_t \lambda^{-\frac{\nu}{2}} c_e(\nu) + C_p, \quad (19)$$

where $c_e(\nu)$ is a constant, independent of λ .

Proof: The power-weighted edges of the NNG, MST and the leaves of MST are all stabilizing graph functionals [18]. Therefore, the LLN for graph functionals holds. \square

The constant $c_e(\nu)$ is the same for both the binomial and the Poisson process, and can be found through simulations. For a 200 node network, we obtained $c_e(2) = 0.7134$, $c_e(3) = 0.6462$ and $c_e(4) = 0.6528$.

B. Detection Error Exponent

In this section, we characterize the detection error exponent. Under the Neyman-Pearson detection, for a fixed Type-I error bound, the exponent of the Type-II error is independent of the type-I error bound [5], and given by

$$D \triangleq \lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{p[\mathbf{Y}_n; \mathcal{H}_0]}{p[\mathbf{Y}_n | \mathcal{V}; \mathcal{H}_1]}, \quad \text{under } \mathcal{H}_0. \quad (20)$$

In the theorem below, we give a closed form for the error exponent, derived² in [1], in terms of the variables and functions defined below.

$$f_1(x) \triangleq \log[1 - x^2], f_2(x) \triangleq \frac{2x^2}{[1 - x^2]}, \quad (21)$$

$$h_i(x; g) \triangleq f_i(g(x)) - \frac{\pi}{2\omega} f_i(g(\sqrt{\frac{\pi}{\omega}}x)), \quad (22)$$

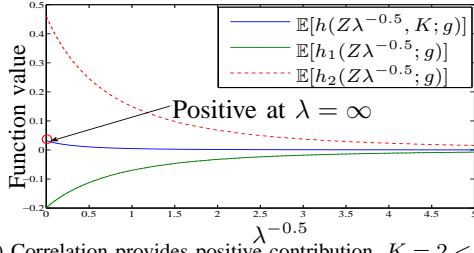
$$h(x, K; g) \triangleq h_1(x; g) + \frac{1}{K} h_2(x; g). \quad (23)$$

Let Z denote the Rayleigh random variable with variance $(2\pi)^{-1}$. Let ω be the area of the union of two unit- radii circles with centers unit distant apart, given by

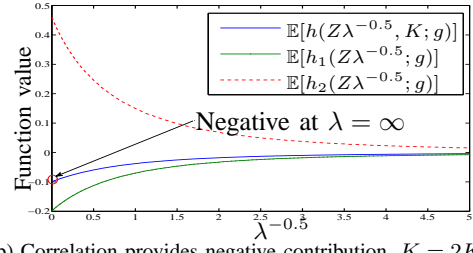
$$\omega = \frac{4\pi}{3} + \frac{\sqrt{3}}{2} \approx 5.06. \quad (24)$$

Theorem 4 (Expression for D): For a GMRF on NNG with correlation function g , with the nodes drawn from the binomial or the Poisson process with node density λ and

²The expression given in [1] is in a different form, but reduces to (25).



(a) Correlation provides positive contribution. $K = 2 < K_t$.



(b) Correlation provides negative contribution. $K = 2K_t$.

Fig. 5: Behavior of functions with exponential correlation $g(r) = M \exp[-ar]$, $M = 0.5, a = 1$. Functions h_1 and h_2 are independent of K . However, K affects their scaling in h . See (23).

region area $\frac{\eta}{\lambda}$, the error exponent D for Neyman-Pearson detection is

$$D(\lambda, K; g) = \frac{1}{2} \left[\mathbb{E}_\lambda h(Z\lambda^{-0.5}, K; g) + \log K + \frac{1}{K} - 1 \right], \quad (25)$$

where \mathbb{E}_λ is the expectation over the random variable Z .

Proof : Due to lack of space, we omit the derivation which employs a special law of large numbers for graph functionals [18] and can be found in [1]. \square

Note that in (25), the expectation term captures the correlation structure of the GMRF and the remaining terms represent the detection error exponent for two IID Gaussian processes with variance ratio K , denoted by $D_{IID}(K)$, i.e.,

$$D(\lambda, K; g) = \frac{1}{2} \mathbb{E}_\lambda h(Z\lambda^{-0.5}, K; g) + D_{IID}(K). \quad (26)$$

Hence, the effect of correlation on the error exponent is quantified in a compact form. It can be easily verified that the expectation term is zero, when $M = 0$ (no correlation). Note that the factor $\lambda^{-0.5}$ appears in the expectation term, since on an average, the edge-lengths scale by that amount. It is easy to see that $D_{IID}(K)$ is independent of node density.

C. Infinite Node Density

We first analyze the error exponent D , when the node density λ goes to infinity. This will provide insights to finding the optimal density. As $\lambda \rightarrow \infty$, we have

$$D(\infty, K; g) = D_{IID}(K) + \frac{1}{2} h(0, K, M),$$

where $h(0, K; g)$ depends on g only through $g(0) = M$, and

$$h(0, K, M) = \left(1 - \frac{\pi}{2\omega}\right) \left(\log[1 - M^2] + \frac{2M^2}{K[1 - M^2]} \right). \quad (27)$$

In the theorem below, we prove that the presence of correlation can either improve or degrade the error exponent, depending on the variance ratio K . We establish a threshold on K that determines the transition.

Lemma 1 (Behavior at Infinite Density ($\lambda = \infty$)): At $\lambda = \infty$, the correlation term $h(0, K, M)$ in (27) is positive, if the variance ratio K is below a threshold value $K_t(M)$,

$$h(0, K, M) \begin{cases} \geq 0, & \text{for } K < K_t(M), \\ < 0, & \text{for } K > K_t(M). \end{cases} \quad (28a)$$

For a fixed $g(0) = M < 1$, the threshold $K_t(M)$ is

$$K_t(M) = -\frac{1}{\log(1 - M^2)} \frac{2M^2}{1 - M^2}, \quad (29)$$

and $2 < K_t(M) < \frac{2}{1 - M^2}$.

Proof: From (27) and $(1 - \frac{\pi}{2\omega}) > 0$. \square

VIII. OPTIMAL NODE DENSITY

Our objective is to find an optimal node density λ_* maximizing the detection error exponent under a constraint on the average energy consumption of routing. Since, the minimum energy scheme is NP-hard and finding its energy consumption is intractable, we instead check whether the DFMRF routing scheme is feasible under the given energy constraint. From (19), the energy constraint \bar{E} in (1) translates to a constraint on λ under the DFMRF routing scheme,

$$\bar{C} \leq \bar{E} \Rightarrow \lambda \geq \lambda_E \triangleq \left(\frac{(\bar{E} - C_p)^+}{C_t c_e(\nu)} \right)^{\frac{2}{\nu}}. \quad (30)$$

$\bar{E} \geq C_p$ is needed for a feasible solution and in this case, when the density $\lambda \rightarrow \infty$, the constraint is always satisfied.

We now consider the modified density optimization problem in (1), restricting to the DFMRF scheme,

$$\bar{\lambda}_*(K, \bar{E}) = \arg \max_{\lambda \geq \lambda_E} D(\lambda, K; g), \quad (31)$$

where K is the variance ratio and g is the correlation function. For fixed K and \bar{E} , $\bar{\lambda}_* \geq \lambda_*$. This is because energy is inversely to node density, and a scheme more efficient than DFMRF can deliver LLR at a lower energy.

A. Threshold Behavior of Optimal Density

In this section, we provide the results for optimal density. In the theorem below, we show that the nature of optimal density is determined by the thresholds K_t and K'_t on K .

Theorem 5 (Result on $\lambda_(K, \bar{E})$):* The optimal density in (31) that maximizes the error exponent, for correlation function with $g(0) = M < 1$, feasible constraint \bar{E} , is

$$\lambda_*(K, \bar{E}) = \infty, \quad \forall K < K_t(M), \quad (32)$$

where threshold K_t is given by (29), and

$$\lambda_*(K, \bar{E}) \leq \bar{\lambda}_*(K, \bar{E}) = \lambda_E, \quad \forall K > K'_t(M), \quad (33)$$

where $\bar{\lambda}_*$ is the optimal density under the DFMRF routing scheme, λ_E is given by (30) and

$$K'_t(M) = \frac{2}{1 - M^2} > K_t(M). \quad (34)$$

Also, when energy constraint is infinite $\bar{E} \rightarrow \infty$, we have $\lambda_E = 0$, and the result is improved to

$$\lambda_*(K, \infty) = 0, \quad \forall K > K_t(M). \quad (35)$$

Proof: See Appendix A. \square

The above theorem states that when the variance ratio K is below the threshold K_t , for any feasible energy constraint, optimality is attained at infinite density. On the other hand, above another threshold K'_t , the constraint λ_E attains the optimal value. In the special case, when there is infinite energy ($\lambda_E = 0$), the constraint λ_E attains the optimal value even in the gap region between the two thresholds K_t and K'_t . On the other hand, for a general constraint $\lambda_E > 0$, we are not able to analytically derive the optimal density in the gap region. However, numerical investigation suggests that λ_E is optimal even in this region.

B. Sensitivity Analysis

In the previous section, we proved the result on optimal density λ_* . In this section, we analyze the extent to which the error exponent is dependent on λ . This enables us to gauge the usefulness of finding λ_* . To this end, we define

$$C(\lambda, K; g) \triangleq \frac{\mathbb{E}[Z\lambda^{-0.5}, K; g]}{2D(\lambda, K; g)}, \quad (36)$$

which is the fraction of contribution coming from correlation to the error exponent in (26), and hence, the part influenced by λ . Note, $0 \leq C(\lambda, K; g) \leq 1$ and $C(0, K; g) = 0$.

Lemma 2 (Sensitivity of D_λ to λ): At $K = 1$, the fraction of contribution from correlation to the error exponent $C(\lambda, K; g)$ is maximum,

$$C(\lambda, 1; g) = 1, \quad \forall \lambda > 0. \quad (37)$$

Also, in the large- K regime,

$$C(\lambda, K; g) \rightarrow 0, \quad \text{as } K \rightarrow \infty, \forall \lambda \in \mathfrak{R}^+. \quad (38)$$

Hence, node density greatly influences detection performance at $K = 1$. In Fig.6, we plot $C(\lambda, K; g)$ as a function of K . We see that it decays to zero as $K \rightarrow \infty$. Hence, the error exponent is insensitive to changes in density at high K .

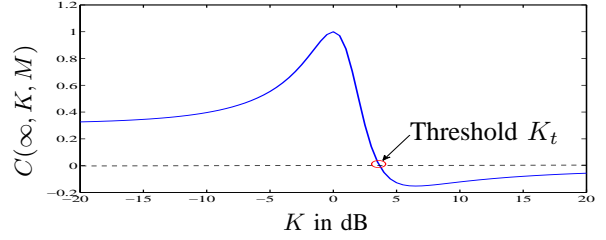


Fig. 6: The contribution from correlation to error exponent vs. variance ratio K at $\lambda = \infty$ and $M = 0.5$. See (36).

IX. NUMERICAL ANALYSIS

In this section, we consider a specific correlation function namely the exponential-correlation function,

$$g(r) = Me^{-ar}, \quad a > 0, 0 < M < 1. \quad (39)$$

Using Theorem 4, we numerically evaluate D through Monte-Carlo runs. We fix the correlation coefficient $a = 1$ in (39).

In Fig.5, we plot the expectations of functions h, h_1 and h_2 against $\lambda^{-0.5}$. In Fig.5a and Fig.5b, the value of K is below and above the threshold K_t . We observe that the behavior at $\lambda = \infty$ is different in the two plots. Note that the functions h_1 and h_2 are independent of K , but K affects their scaling in h .

In Fig.7a and Fig.7b, we numerically evaluate the optimal $\lambda_*(K, \lambda_E)$ for different values of M and variance ratio K . It is convenient to plot the results in terms of $\lambda^{-0.5}$, since the λ_* could be infinite. We observe the threshold behavior at K_t , as predicted. When $K < K_t$, $\lambda_* = \infty$. For $K > K_t$, in Fig.7a and Fig.7b, optimality is mostly attained at the other extreme point λ_E . This is consistent with Theorem 5.

X. CONCLUSIONS

The tradeoff between the energy consumption in routing data and the resulting detection performance at the fusion center, is an important problem in the context of sensor networks. In this paper, we aimed to gain a perspective on this problem and additionally incorporated correlation between the measurements through the Gauss-Markov random field model. We characterized the density of node deployment that maximizes the detection error exponent subject to a constraint on the average energy consumption. The measurement variance is crucial in determining whether the optimal node density is limited by the routing energy constraint and displays a threshold behavior. We have derived the threshold analytically and verified it with simulations.

The results on the optimal density are possible due to exploitation of the Markovian structure, in deriving the error exponent and the routing schemes. While acknowledging the limitations of its validity, we have made an attempt to characterize detection-energy tradeoffs for correlated measurements in two (and higher) dimensions, which to the best of our knowledge, has not been dealt before. Alternative formulations, not dealt in this paper, include selection of nodes

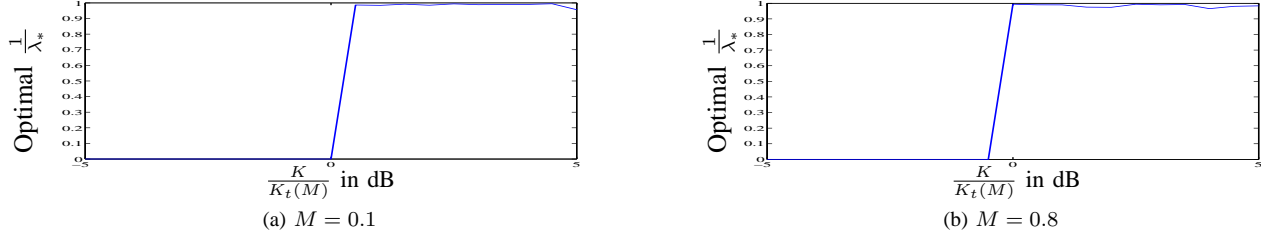


Fig. 7: The optimal density λ_* is decided by threshold $K_t(M)$ on variance ratio K . $\lambda_E = 1$. See (31), Theorem 5.

with “useful” data, and incorporating node and link failures. We have also not considered the issue of quantization of measurements at various nodes. Another scenario is mobile nodes having regions of coverage, instead of static nodes, as assumed here, with observations at single points. Extensions to such scenarios would be of interest.

APPENDIX

A. Proof of theorem 5

Define $f(x, K) \triangleq f_1(x) + \frac{1}{K} f_2(x)$,

$$f(x, K) = \log(1 - x^2) + \frac{2}{K} \frac{x^2}{1 - x^2}.$$

Since $g(0) = M$ and $g(\infty) = 0$, we have $x \in [0, M]$.

$$\frac{\partial f}{\partial x} = \frac{2x}{1 - x^2} \left(-1 + \frac{2}{K(1 - x^2)} \right).$$

Therefore, f has only one critical point in $(0, M]$. For $K < 2$, $\frac{\partial f}{\partial x} > 0$ and for $K > \frac{2}{1 - M^2}$, $\frac{\partial f}{\partial x} < 0$, $\forall x \in [0, M]$. There are no critical points. For $2 < K < \frac{2}{1 - M^2}$, the critical point is a minimum. Hence, maximum is attained at one of the boundary points $\{0, M\}$. For $K < K_t$, it is at $x = M = g(0)$ and hence,

$$f(g(R_e), K) \leq f(g(0), K; g), \quad \forall R_e \geq 0.$$

When the point sets are drawn from binomial or Poisson processes, and R_e are the edge-lengths of NNG $\mathcal{G}_d = (\mathcal{V}, \mathcal{E}_d)$,

$$\frac{1}{n} \sum_{e \in \mathcal{E}_d} f(g(R_e), K; g) \leq \frac{1}{n} \sum_{e \in \mathcal{E}_d} f(g(0), K; g).$$

Letting $n \rightarrow \infty$ on both sides,

$$\mathbb{E}[h(Z\lambda^{-0.5}, K; g)] \leq h(0, K, M), \quad \forall x \in \mathfrak{R}^+,$$

where the limits have been derived in [1]. Hence,

$$\lambda_*(K, \lambda_E) = \arg \max_{\lambda \geq \lambda_E} \mathbb{E}[h(Z\lambda^{-0.5}, K; g)] = \infty.$$

Similarly, for $K > K_t$, we have

$$f(R_e, K; g) \leq f(\infty, K; g) = 0, \quad \forall R_e \geq 0.$$

If $\bar{E} = \infty$, $\lambda_E = 0$ and $\lambda_*(K, 0) = 0$. Now, for $K > \frac{2}{1 - M^2}$, $f(g(R_e), K)$ is increasing in R_e . $R_e = \lambda^{-0.5} R'_e$, where R'_e is the edge-length in unit area. Hence, $f(g(R_e), K)$ is decreasing in λ . Therefore, the limit $\mathbb{E}[h(Z\lambda^{-0.5}, K; g)]$ is decreasing in λ . Hence, $\lambda_*(K, \lambda_E) = \lambda_E$. \square

REFERENCES

- [1] A. Anandkumar, L. Tong, and A. Swami, “Detection of Gauss-Markov random fields with nearest-neighbor dependency,” Tech. Rep. ACSP TR 01-07-01, Jan. 2007, submitted to IEEE Trans. Information Theory, available at <http://arxiv.org/abs/0706.1588v1>.
- [2] A. Anandkumar, L. Tong, and A. Swami, “Energy Efficient Routing for Statistical Inference of Markov Random Fields,” in *Proc. of CISS '07*, Baltimore, USA, March 2007.
- [3] T. Kailath and H.V. Poor, “Detection of stochastic processes,” *IEEE Trans. on Information Theory*, vol. 44, no. 6, pp. 2230–2259, Oct. 1998.
- [4] M. D. Donsker and S. R. S. Varadhan, “Large deviations for stationary Gaussian process,” *Commun. Math. Phys.*, vol. 97, pp. 187–210, 1985.
- [5] Po-Ning Chen, “General formulas for the Neyman-Pearson type-II error exponent subject to fixed and exponential type-I error bounds,” *IEEE Trans. on Information Theory*, vol. 42, no. 1, pp. 316–323, Jan. 1996.
- [6] Akkaya K and Younis M, “A Survey of Routing Protocols in Wireless Sensor Networks,” *Elsevier Ad Hoc Network Journal*, vol. 3/3, pp. 325–349, 2005.
- [7] A. Giridhar and P.R. Kumar, “Toward a theory of in-network computation in wireless sensor networks,” *Communications Magazine, IEEE*, vol. 44, no. 4, pp. 98–107, 2006.
- [8] R. Rajagopalan and P.K. Varshney, “Data aggregation techniques in sensor networks: A survey,” *Communications Surveys & Tutorials, IEEE*, vol. 8, no. 4, pp. 48–63, 2006.
- [9] Y. Yang and R.S. Blum, “Energy-efficient routing for signal detection under the Neyman-Pearson criterion in wireless sensor networks,” in *Proceedings of the 6th international conference on Information processing in sensor networks*, 2007, pp. 303–312.
- [10] S. Appadwedula, VV Veeravalli, and DL Jones, “Energy-Efficient Detection in Sensor Networks,” *Selected Areas in Communications, IEEE Journal on*, vol. 23, no. 4, pp. 693–702, 2005.
- [11] L. Yu, L. Yuan, G. Qu, and A. Ephremides, “Energy-driven detection scheme with guaranteed accuracy,” in *Proceedings of the fifth international conference on Information processing in sensor networks*, 2006, pp. 284–291.
- [12] Y. Sung, S. Misra, L. Tong, and A. Ephremides, “Cooperative Routing for Signal Detection in Large Sensor Networks,” *submitted to IEEE J. Select. Area Comm.*, Jan. 2006.
- [13] J.F. Chamberland and VV Veeravalli, “How Dense Should a Sensor Network Be for Detection With Correlated Observations?,” *Information Theory, IEEE Transactions on*, vol. 52, no. 11, pp. 5099–5106, 2006.
- [14] N. Katenka, E. Levina, and G. Michailidis, “Local vote decision fusion for target detection in wireless sensor networks,” in *Joint Research Conference on Statistics in Quality Industry and Technology*, Knoxville, TN, USA, June 2006.
- [15] O.P. Kreidl and A.S. Willsky, “Inference with Minimal Communication: a Decision-Theoretic Variational Approach,” in *Advances in Neural Information Processing Systems*, 2006.
- [16] H. Rue and L. Held, *Gaussian Markov Random Fields: Theory and Applications*, Chapman and Hall, London, 2005, vol. 104 of *Monographs on Statistics and Applied Probability*.
- [17] A. Ephremides, “Energy concerns in wireless networks,” *IEEE Wireless Communications*, no. 4, pp. 48–59, August 2002.
- [18] M.D. Penrose and J.E. Yukich, “Limit Theory for Random Sequential Packing and Deposition,” *Annals of Applied probability*, vol. 12, no. 1, pp. 272–301, 2002.