DETECTION OF GAUSS-MARKOV RANDOM FIELD ON NEAREST-NEIGHBOR GRAPH

Animashree Anandkumar, Lang Tong *

Electrical and Computer Engineering,
Cornell University, Ithaca, NY, 14853 USA.

Ananthram Swami†

Army Research Laboratory,
Adelphi, MD 20783 USA.

ABSTRACT

The problem of hypothesis testing against independence for a Gauss-Markov random field (GMRF) with nearest-neighbor dependency graph is analyzed. The sensors measuring samples from the signal field are placed IID according to the uniform distribution. The asymptotic performance of Neyman-Pearson detection is characterized through the large-deviation theory. An expression for the error exponent is derived using a special law of large numbers for graph functionals. The exponent is analyzed for different values of the variance ratio and correlation. It is found that a more correlated GMRF has a higher exponent (improved detection performance) at low values of the variance ratio, whereas the opposite is true at high values of the ratio.

***Index Terms*— Signal detection, Gaussian processes, Markov processes, Error analysis, Graph theory.**

1. INTRODUCTION

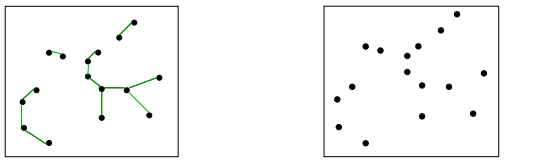
For distributed detection, the so-called conditionally IID assumption is mathematically convenient and is widely assumed in the literature. The assumption states that conditioned on a particular hypothesis, the observations at sensors are independent and identically distributed. In practice, however, sensors observe correlated data, since natural spatial signals have stochastic dependence. Moreover, spatial random signals are typically acausal in contrast to temporal signals. In the literature, the two are usually distinguished by referring to acausal signals as random fields (RF) and causal signals as random processes (RP).

In this paper, we consider the detection problem of a Gauss-Markov random field. See Fig.1. We consider the Neyman-Pearson (NP) formulation, where the detector is optimal at a fixed false-alarm probability. For any positive fixed level of false alarm, when the miss-detection probability $P_M(n)$ of the NP detector decays exponentially with the sample size n, we have the error exponent

$$D \triangleq -\lim_{n\to\infty} \frac{1}{n} \log P_M(n). \tag{1}$$

The error exponent is an important performance measure since a large exponent implies faster decay of error probability with increasing sample size.

In this paper, we evaluate the error exponent for NP detection of GMRF, restricting to a class of GMRFs with nearest-neighbor dependency graph. This assumption has been employed in areas of applied science, including the social sciences, geography and ecology, where proximity data is often important.



(a) $\mathcal{H}_1$: Gauss-Markov random field on nearest-neighbor graph.

(b) $\mathcal{H}_0$: Independent observations.

Figure 1: Illustration of the hypothesis-testing problem.

We assume that the sensors observing the signal field are placed IID according to the uniform distribution. This results in a non-stationary GMRF (for the definition of stationary GMRF, see [1, p. 57]). We analyze the detection performance assuming access to all the observations, when the number of sensors goes to infinity, by way of the coverage area of the nodes going to infinity, while keeping the node density fixed.

1.1. Related work and contributions

The kind of hypothesis testing we consider is called testing against independence. In [2], problems of this kind are considered, with rate constraints on the channels and for only two sources, with large number of samples at each source. In this paper, we assume that there are no constraints on the channel and that the observations have a specific correlation structure of the GMRF. However, our formulation is different since there is a single observation at every sensor, and the number of sensors goes to infinity.

The detection of Gauss-Markov random processes (GMRP) in Gaussian noise is a classical problem [3]. There is an extensive literature on the large-deviations approach to the analysis of detection of GMRP, but closed-form expressions have been derived only for some special cases, *e.g.,* [4]. An approach to characterizing the GMRP via inversion algorithms for block banded matrices has been presented in [5]. However, these approaches are not amenable to the extension of the problem to planar and higher dimensional spaces, since they deal with random processes rather than random fields, or to the random placement of nodes.

To our knowledge, the asymptotic performance of detection of acausal non-stationary GMRF has not been analyzed in the past. We pursue a graph-theoretic approach and exploit recent advances in computational geometry [6]. We provide an expression for the log-likelihood ratio of detection by exploiting the properties of the nearest-neighbor graph. By casting the error exponent as the limit of a graph functional, we are able to apply the law of large numbers for functionals on graphs derived in [6]. We then numerically evaluate the exponent for different values of the variance ratio and correla-

*Email:{aa332@,ltong@ece.}cornell.edu. 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 National Science Foundation under Contract CNS-0435190.

†E-mail: a.swami@ieee.org. The third author was partially supported by the DARPA ITMANET program.

tion, assuming exponential-correlation function. We conclude for a fixed node density that a more correlated GMRF has a higher exponent at low values of variance ratio whereas the situation is reversed at high values.

1.2. Notation and organization

Vectors and matrices are written in boldface. Random variables are in capital letters, random processes and random fields in boldface capitals and sets in calligraphic font. 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.

An undirected graph $\mathcal{G}$ is a tuple $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where $\mathcal{V}$ is the vertex set and $\mathcal{E} = \{(i,j),\, i,j \in \mathcal{V}, i \neq j\}$ is the edge set. When i and j have an edge between them, i and j are neighbors denoted by $i \sim j$ (otherwise it is $i \nsim j$). The neighborhood function of a node i is the set of all other nodes having an edge with it, *i.e.,*

$$\mathcal{N}_e(i) = \{j \in \mathcal{V} : j \neq i, (i,j) \in \mathcal{E}\}. \tag{2}$$

The number of neighbors of a node i is called its degree, $\text{Deg}(i)$. Let r_{ij} denote the Euclidean edge length of (i,j). A node with a single edge *i.e.,* its degree is 1 is known as a leaf and the corresponding edge as a leaf edge. In this paper, we consider the terms node, vertex and sensor interchangeable.

Our paper is organized as follows. We provide the description of the GMRF in section 2 and describe the nearest-neighbor graph and the correlation function in sections 2.1 and 3.1. We provide the problem statement in section 3. In section 4, we derive a closed-form expression for the log-likelihood ratio. We evaluate the error exponent in section 5 and provide numerical results for the exponent in section 5.1. Section 6 concludes the paper.

2. GAUSS-MARKOV RANDOM FIELD (GMRF)

Gauss-Markov random fields, in addition to being Gaussian random fields, satisfy conditional-independence properties. A simple example for GMRF is the first order auto-regressive process. In this case, the conditional independence of the observations is based on causality. However, for spatial observations we need a more general definition [1, p. 21].

Definition 1 (GMRF) *Given a point set $\mathcal{V} = \{1, \ldots, n\}$, $\mathbf{Y}(\mathcal{V}) = \{Y_i : i \in \mathcal{V}\}$ is a GMRF with an (undirected) dependency graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ 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 iff i and j are not neighbors, i.e.,*

$$Y_i \perp Y_j | \mathbf{Y}_{-ij} \iff i \nsim j,\ \forall i,j \in \mathcal{V}, i \neq j, \tag{3}$$

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

This implies that the conditional distribution at a node given the observations at its neighbors is independent of the rest of the network. A common approach to formulating a GMRF is to specify the graph through a neighborhood rule, and then to specify the correlation function between these neighbors. In general the neighbors of a node are those which are in its 'proximity', usually in terms of the Euclidean distance [1, 7]. With a regular lattice structure (*e.g.,* in image processing), a fixed set of neighbors can be specified in a straight-forward manner [7]. However, the situation is more complicated for arbitrarily placed nodes. We assume the dependency graph to be the nearest-neighbor graph (NNG), described below.

2.1. Nearest-neighbor graph

The nearest-neighbor function of a node $i \in \mathcal{V}$, is defined as,

$$\text{nn}(i) \triangleq \arg \min_{j \in \mathcal{V}, j \neq i} \text{dist}(i,j), \tag{4}$$

where $\text{dist}(\cdot,\cdot)$ is the Euclidean distance. For the random point sets considered here, the inter-point distances are unique with probability 1. Therefore, (4) is a well-defined function for every node, almost surely. The nearest-neighbor (undirected) graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is given by

$$(i,j) \in \mathcal{E} \iff i = \text{nn}(j) \text{ or } j = \text{nn}(i). \tag{5}$$

NNG is acyclic with a maximum node degree of 5 almost surely [8].

3. PROBLEM STATEMENT

Let $(\mathcal{B}_n)_{n \geq 1}$ denote a sequence of 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. We are interested in the detection performance when the number of nodes goes to infinity, with fixed node density, *i.e.,* $n \to \infty$ with λ fixed.

Let $\mathbf{Y}_n$ be the random vector of observation samples Y_i,

$$\mathbf{Y}_n \triangleq [Y_1, \ldots, Y_n]^T. \tag{6}$$

Given a set of nodes $\mathcal{V}$ drawn from $\mathcal{U}_{n,\lambda}$, the hypothesis-testing problem is as follows,

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

where $\mathbf{\Sigma}_{1,\mathcal{V}}$ is the covariance matrix of a GMRF, which depends on the configuration of nodes in $\mathcal{V}$, described in detail in section 3.1.

We assume that the point process $\mathcal{U}_{n,\lambda}$ is the same under both hypotheses and also that the node locations are known. Therefore, the optimal decision rule is a threshold test based on the conditional log-likelihood ratio. Let $p[\mathbf{Y}_n | \mathcal{V}; \mathcal{H}_j]$ be the conditional PDF of the observations given $\mathcal{V} \sim \mathcal{U}_{n,\lambda}$ under hypothesis j. The Kullback-Leibler rate with respect to the sample size n is given by

$$\frac{1}{n} \log \frac{p[\mathbf{Y}_n | \mathcal{V}; \mathcal{H}_0]}{p[\mathbf{Y}_n | \mathcal{V}; \mathcal{H}_1]}.$$

The error exponent under the Neyman-Pearson detection is given by the almost-sure limit of the Kullback-Leibler rate under the null hypothesis, assuming that the limit exists [9]. Therefore, the error exponent D of NP detection in (1) is given by

$$\begin{aligned} D &= \lim_{n \to \infty} \frac{1}{n} \log \frac{p[\mathbf{Y}_n | \mathcal{V}; \mathcal{H}_0]}{p[\mathbf{Y}_n | \mathcal{V}; \mathcal{H}_1]},\ a.s.\,[\mathcal{H}_0] \\ &= \lim_{n \to \infty} \frac{1}{2n} \Big(\log \frac{|\mathbf{\Sigma}_{1,\mathcal{V}}|}{|\sigma_0^2 \mathbf{I}|} + \mathbf{Y}_n^T \{ \mathbf{\Sigma}_{1,\mathcal{V}}^{-1} \\ &\quad - (\sigma_0^2 \mathbf{I})^{-1} \} \mathbf{Y}_n \Big)\ a.s.\,[\mathcal{H}_0], \end{aligned} \tag{7}$$

where $a.s.\,[\mathcal{H}_0]$ denotes the almost-sure limit under null hypothesis.

3.1. Correlation function

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

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

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

$$K \triangleq \frac{\sigma_1^2}{\sigma_0^2}. \tag{9}$$

We assume that the correlation function between the neighbors of the nearest-neighbor graph is specified by an arbitrary function g, which has the Euclidean edge length r_{ij} as its argument *i.e.,*

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

In general, g is a monotonically non-increasing function of the edge length, and $g(0) = M < 1$ (defined as the nugget), in geo-statistics [10]. We assume that $g(r) < 1, r \in \Re^+$, *e.g.,*

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

4. EXPRESSION FOR LOG-LIKELIHOOD RATIO

The inverse of the covariance matrix of a GMRF is known as the potential matrix or precision matrix or information matrix. The non-zero elements of the potential matrix $\mathbf{A}$ of a GMRF are in one to one correspondence with the edges of its graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ [1, Theorem 2.2] in the sense that

$$i \nsim j \iff A(i,j) = 0, \forall i, j \in \mathcal{V}, i \neq j. \tag{11}$$

The simple relationship between the conditional independence of the GMRF and the zero structure of its potential matrix is not evident in covariance matrix, which is a completely dense matrix. Therefore, it is easier to evaluate the log-likelihood ratio in (7) through the potential matrix. We now provide expressions for the joint distribution of the GMRF through the potential matrix and its determinant. Note, the joint distribution can also be derived in terms of the marginal probability of the nodes and the joint probability at the edges of the acyclic dependency graph [11].

Theorem 1 *[Coefficients and determinant of potential matrix] Under assumptions (8-10), given the correlation function g and edge lengths $r_{ij} > 0$, the coefficients of potential matrix $\mathbf{A} \triangleq \mathbf{\Sigma}_{1,\mathcal{V}}^{-1}$ are*

$$A(i,i) = \frac{1}{\sigma_1^2} + \sum_{j \in \mathcal{N}_e(i)} \frac{g^2(r_{ij})}{\sigma_1^2(1-g^2(r_{ij}))}, \tag{12}$$

$$A(i,j) = \begin{cases} \frac{-g(r_{ij})}{\sigma_1^2(1-g^2(r_{ij}))} & \text{if } i \sim j, \\ 0 & \text{o.w.} \end{cases} \tag{13}$$

The determinant of the potential matrix of $\mathbf{A}$ is given by

$$|\mathbf{A}| = \frac{1}{\sigma_1^{2n}} \prod_{\substack{(i,j) \in \mathcal{E} \\ i<j}} \frac{1}{1-g^2(r_{ij})}. \tag{14}$$

Proof : Direct inverse, by exploiting the acyclicity of NNG. For the determinant, using (12) and (13), we recursively derive the determinant of a component, starting at the leaf. An acyclic graph with atleast an edge has a leaf. See [12]. □

Using the expressions for the coefficients and determinant, we have a closed-form expression for (7). Furthermore, (12-14) lead to explicit data fusion and routing schemes [13]. To derive the error exponent, we need to evaluate the Kullback-Leibler rate under the null hypothesis, where the observations Y_i's are i.i.d and independent of the point process $\mathcal{U}_{n,\lambda}$. To do this, we cast the error exponent as the limit of a graph functional, with the nodes drawn from a marked binomial point process with marking variable Y_i.

Lemma 1 (D as a graph functional) *Under the assumptions (8-10), the error exponent D for Neyman-Pearson detection of GMRF with nearest-neighbor dependency graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, expressed as the sum of edge and node functionals of a marked point set $\mathcal{V} \sim \mathcal{U}_{n,\lambda}$ with marking variable Y_i, is*

$$\begin{aligned} D \ = \ & \log\frac{\sigma_1}{\sigma_0} + \lim_{n\to\infty} \frac{1}{2n}\Bigg[\sum_{i\in\mathcal{V}}\Big(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_0^2}\Big)Y_i^2 \\ & + \sum_{\substack{(i,j)\in\mathcal{E} \\ i<j}} \Big\{\log[1-g^2(R_{ij})] + \frac{g^2(R_{ij})}{1-g^2(R_{ij})}\frac{Y_i^2+Y_j^2}{\sigma_1^2} \\ & - \frac{2g(R_{ij})}{1-g^2(R_{ij})}\frac{Y_iY_j}{\sigma_1^2}\Big\}\Bigg], \ Y_i \overset{i.i.d}{\sim} \mathcal{N}(0,\sigma_0^2), \end{aligned} \tag{15}$$

where R_{ij} denotes the (random) Euclidean edge length of $(i,j) \in \mathcal{E}$, that depends on the underlying binomial point process $\mathcal{U}_{n,\lambda}$. The condition $i < j$ ensures that every edge is counted only once.
Proof : Substitute (12-14) in (7). □

5. ERROR EXPONENT

We showed in lemma 1 that the error exponent reduces to the limit of a graph functional. In this section, we apply the law of large numbers for graph functionals to evaluate the limit.

Theorem 2 (An expression for D) *The error exponent D for Neyman-Pearson detection with node density λ and correlation function g is*

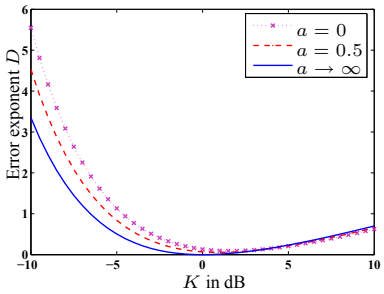
$$D = \frac{1}{2}\big[\mathbb{E}f(g(Z_1)) - \frac{\pi}{2\omega}\mathbb{E}f(g(Z_2)) + \log K + \frac{1}{K} - 1\big], \tag{16}$$

where f is defined by,

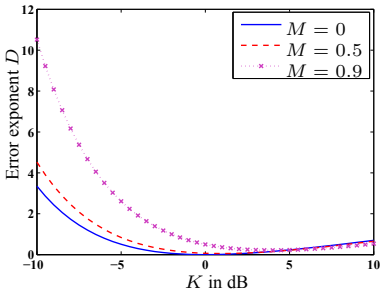
$$f(x) \triangleq \log[1-x^2] + \frac{2x^2}{K[1-x^2]}, \tag{17}$$

Z_1 and Z_2 are Rayleigh distributed with variances $(2\pi\lambda)^{-1}$ and $(2\omega\lambda)^{-1}$, and ω is given by

$$\omega = \frac{4\pi}{3} + \frac{\sqrt{3}}{2} \approx 5.06, \tag{18}$$



(a) Different values of correlation coefficient a, nugget $g(0) = M = 0.5$.



(b) Different values of nugget $g(0) = M$, correlation coefficient $a = 0.5$.

Figure 2: Error exponent D vs. ratio of variances K, node density $\lambda = 1$. See (16-19).

and is the area of the union of two unit-radii circles with centers unit distant apart.

Proof : By law of large numbers for graph functionals derived in [6, p. 287]. See [12] for details. □

In (16), except for the first two f-terms which capture the correlation structure of the GMRF, the remaining terms represent the detection-error exponent for two IID Gaussian processes.

5.1. Numerical results

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

$$g(r) = M\mathrm{e}^{-ar},\ a > 0, 0 < M < 1. \tag{19}$$

Using theorem 2, we numerically evaluate D through Monte-Carlo runs. For fixed values of K and M, we have

$$D(K, M, \lambda, a) = D(K, M, 1, \frac{a}{\sqrt{\lambda}}), \tag{20}$$

which we obtain by changing the integration variable in the expectation term in (16). Therefore, in terms of the error exponent, increasing the node density λ is equivalent to a lower correlation coefficient at unit density. Here, we plot the effects of correlation coefficient a and nugget M on D.

In Fig.2(a), we plot the error exponent at $\lambda = 1$ and $M = 0.5$, for different values of correlation coefficient a. Note, the cases $a = 0$ and $a \to \infty$ correspond to constant correlation (of value M) and the independent case. We notice that a more correlated GMRF or the one with smaller a has a higher exponent at low value of K, whereas the situation is reversed at high K. Equivalently, increasing the node density λ improves the exponent at low value of K, but not at high K. Also, when the variance ratio K is large enough, D appears to increase linearly with K (in dB), and the correlation coefficient a and nugget M appear to have little effect, as expected from theorem 2. In Fig.2(b), we plot the exponent at constant correlation coefficient $a = 0.5$, for different values of the nugget M. We notice a similar behavior as the correlation coefficient. A higher value of M results in a higher exponent at low K, but not at high K.

6. CONCLUSION

In this paper, we derived a closed-form expression for the likelihood function of a Gauss-Markov random field and then employed the law of large numbers for graph functionals to derive its Neyman-Pearson detection-error exponent. Energy-efficient data fusion and routing schemes to achieve this exponent in a distributed way are currently under investigation [13]. Although, we have assumed identical variance at every sensor, a spatially-varying variance model can be incorporated into our results. We have focused on the GMRF with nearest-neighbor dependency graph, which is a simplifying assumption. Although, the law of large numbers is valid for a number of proximity graphs, which have edges between "nearby" points, the actual evaluation of the log-likelihood ratio and the exponent are intractable for most of these graphs. Moreover, GMRF with small neighborhood structure has been demonstrated to approximate the hidden GMRF [14] as well as the Gaussian field with long correlation lengths [15], reasonably well.

7. REFERENCES

[1] 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*.

[2] Te Sun Han and S. Amari, "Statistical inference under multiterminal data compression," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2300–2324, Oct. 1998.

[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] 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. Information Theory*, vol. 52, no. 4, pp. 1354–1365, April 2006.

[5] A. Asif and J.M.F. Moura, "Block Matrices with L-Block-Banded Inverse: Inversion Algorithms," *IEEE Trans. SP*, vol. SP-53, no. 2, pp. 630–642, Feb. 2005.

[6] 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.

[7] J.M.F. Moura and N. Balram, "Recursive Structure of Noncausal Gauss-Markov Random Fields," *IEEE Trans. IT*, vol. IT-38, no. 2, pp. 334–354, March 1992.

[8] D. Eppstein, M. S. Paterson, and F. F. Yao, "On Nearest-Neighbor Graphs," *Discrete and Computational Geometry*, vol. 17, pp. 263–282, 1997.

[9] R. R. Bahadur, S. L. Zabell, and J. C. Gupta, "Large deviations, tests, and estimates," *Asymptotic Theory of Statistical Tests and Estimation:In Honor of Wassily Hoeffding*, 1980, Academic Press, New York.

[10] M.B. Jaksa, P.I. Brooker, and W.S. Kaggwa, "Inaccuracies Associated with Estimating Random Measurement Errors," *Journal of Geotechnical and Geoenvironmental Engineering*, vol. 123, no. 5, pp. 393–401, 2006.

[11] R.G. Cowell, *Probabilistic Networks and Expert Systems*, Springer, 1999.

[12] 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, Also submitted to IEEE Trans. Information Theory, `http://acsp.ece.cornell.edu/pubR.html`.

[13] A. Anandkumar, L. Tong, and A. Swami, "Energy-efficient data fusion in Gauss-Markov random field," Tech. Rep. ACSP TR 01-07-02, Jan. 2007, `http://acsp.ece.cornell.edu/pubR.html`.

[14] H. Rue, I. Steinsland, and S. Erland, "Approximating hidden Gaussian Markov random fields," *Journal of the Royal Statistical Society Series B(Statistical Methodology)*, vol. 66, no. 4, pp. 877–892, 2004.

[15] H. Rue and H. Tjelmeland, "Fitting Gaussian Markov Random Fields to Gaussian Fields," *Scandinavian Journal of Statistics*, vol. 29, no. 1, pp. 31–49, 2002.