Nearly Optimal Distributed Configuration Management Using Probabilistic Graphical Models

Sung-eok Jeon and Chuanyi Ji

Abstract—
This work studies distributed configuration management of large wireless sensor networks, where management objectives are achieved by local cooperation of individual nodes. Specifically, we study *when* distributed configuration management is nearly optimal, and *how* to obtain a nearly-optimal configuration through decentralized adaptation.

We first derive a spatial network model that is determined by internal network characteristics and management requirements. We next show that a sufficient condition for distributed configuration management to be nearly-optimal is that the spatial network model belongs to a class of coupled Markov Random Fields also known as Random-Bond Ising model. Such graphs possess a cross-layer spatial Markov property. We specify the sufficient conditions for the nearly-optimality under different channels and density of nodes. We derive a nearly-optimal distributed algorithm using the probabilistic inference based on the derived network model. The algorithm is applied to spatial-reuse TDMA which configures a logical topology.

I. Introduction

Wireless infrastructureless networks provide a natural setting for dense and sparse sensor networks, wireless routers, mobile robots, and intelligent agents. Such a network can be large with hundreds of heterogeneous devices that communicate among themselves randomly. This poses a significant challenge to the management of such networks, where there is no centralized authority available and distributed management is a necessity.

Configuration management is an important task of wireless sensor networks. Optimal configurations can be obtained through global optimization of management objectives under multiple constraints. As centralized management schemes are infeasible for large wireless sensor networks, a question is whether configuration management can be accomplished in a distributed setting.

Numerous distributed algorithms and protocols have been developed for topology formation based on empirical investigations [14] [1]. These studies provide promising results on importance of configuration management. The performance of these algorithms though is usually tested through simulation. Simulations, however, would not provide quantifiable conditions on when and how distributed self-configuration can achieve a nearly-optimal configuration with a bounded error. Such conditions need to be derived based on trade-offs between the optimality of a configuration and the amount of information exchange in distributed configuration management [10]. In general, it has been shown to be a difficult problem to develop a distributed algorithm whose performance is within a tolerable degradation from that of the optimal scheme [18]. Hence, the open issues are

(a) *When* is it possible for distributed management to obtain a nearly-optimal configuration?

(b) *How* to obtain a nearly-optimal configuration using distributed algorithms?

We apply machine learning approaches to study these issues.

(a) **Global Model**: Our first step is to develop a network model which can represent the spatial dependence in a network configuration. This requires a bottom-up approach to incorporate both internal network characteristics and management constraints. The motivation on such a bottom-up model-based approach is that "when and how" should be determined and understood by intrinsic properties of networks.

We adopt an analogy between link activities and node positions of wireless sensor networks and interacting particles in statistical physics [5] [19]. This allows us to derive a so-called "configuration Hamiltonian" [11] which has been widely-used in statistical physics. When applied to a wireless sensor network, the configuration Hamiltonian corresponds to the system energy that combines a physical topology, link activities, and management constraints into one quantity. The configuration Hamiltonian is used to obtain a spatial probabilistic network model in the form of a Boltzmann distribution [19], which provides a global spatial model of node positions and link activities. An optimal configuration is the one that maximizes the global likelihood function.

(b) **Probabilistic Graphs and Local Model:** As the spatial dependence in a network configuration is complex, we resort to probabilistic graphical models in machine learning [5] [7]. A probabilistic graph provides a simple and explicit representation of the statistical spatial dependence among network nodes and node-pair communications. If a network configuration can be represented by a probabilistic graph with a nested spatial Markov dependence, the distributed configuration management is possible and optimal. A sufficient condition for the near-optimality is to show that the relative distance between a sub-optimal configuration achieved by distributed management and a global optimal configuration is within a given error bound.

(c) **Distributed Algorithm**: A nearly-optimal distributed algorithm can then be derived based on the network model and the corresponding cross-layer probabilistic graph. The range of information exchange is determined by an interference neighborhood which corresponds to the local connectivity on the dependency graph. The actual information exchanged includes relative positions of neighbors and activities of adjacent links.

S. Jeon and C. Ji are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA. E-mail: {sejeon,jic}@ece.gatech.edu

0-7803-9466-6/05/$20.00 ©2005 IEEE

We apply the distributed algorithm to spatial-reuse TDMA (STDMA) with a minimum cycle. In sensor networks, STDMA can be implemented for a fair and efficient link scheduling[4] [2].

II. Problem Formulation

Consider a wireless network with N nodes. Let X_i be the location of node i, and $\underline{X} = \{X_1, \cdots, X_N\}$ characterize a physical topology of the network. X_i is regarded as a random variable due to either noisy estimates of sensor locations [3] or random node movements. Assume that any pair of nodes within a communication range can communicate with an omni-directional antenna sharing a common channel. Assume that a MAC protocol (i.e., 802.11) provides a basis for the access of the shared wireless channel, upon which an advanced scheme for configuration management can be developed.

Let σ_{ij} be random activity of link (i, j), $\sigma_{ij} = 1$ if node i is transmitting to node j; and $\sigma_{ij} = -1$, otherwise. σ_{ij} is referred to as a "communication dipole" for $1 \leq i, j \leq N$, and $\underline{\sigma} = \{\sigma_{_{1,2}}, \cdots, \sigma_{_{N,N-1}}\}$ denotes a logical topology which consists of link activities of all node-pairs in the network. A network configuration is a combination of both a physical and a logical topology $(\underline{\sigma}, \underline{X})$.

Assume that wireless channel follows a path loss model with a power attenuation factor α where $\alpha \in [2 \sim 6]$. Assume also that transmission powers are deterministic, and shadowing or multi-path fading is not considered for simplicity. Let $P(\underline{\sigma}, \underline{X})$ be a global spatial probabilistic model of a network configuration. This work intends to investigate the following questions.

(a) How to obtain $P(\underline{\sigma}, \underline{X})$ based on the given assumptions and management requirements?

(b) How to represent the complex spatial dependence of a configuration?

(c) Under what conditions is it possible to obtain a nearly-optimal configuration through a distributed algorithm?

III. Global Model: A Cross-Layer Spatial Network Model

In this section, we develop a network model as a basis for studying the feasibility and algorithms of distributed configuration management. We introduce a quantity known as "configuration Hamiltonian" by drawing upon an analogy between link activities and node positions in wireless networks and interacting particles in statistical physics [5]. The configuration Hamiltonian results in a cross-layer spatial network model based on a Boltzmann distribution [11].

A. Logical Configuration: Random Link Activities

1) Configuration Hamiltonian

Given node positions $\underline{X}$, link activities $\underline{\sigma}$ depend on signal power attenuation (i.e., communication range), MAC channel contention, and link quality constraints. Based on the assumptions given in Section II, for a transmitter i and a receiver j, the received power $P_j = P_i l_{ij}^{-\alpha} \frac{\sigma_{ij}+1}{2}$, where P_i is the power at transmitter i and $l_{ij} = \|X_i - X_j\|$. If a dipole is active, $\sigma_{ij} = 1$, P_j=$P_i l_{ij}^{-\alpha}$; otherwise, $\sigma_{ij} = -1$, $P_j = 0$.

TABLE I
Correspondence between Dipole System and Lattice Gas

Dipole System	Lattice Gas [11]
active(+1) / inactive(-1)	occupied(+1) / empty(-1)
interference	repulsive interaction energy
communication potential	chemical potential
ordered configuration	liquid or gas

A network of dipoles $\underline{\sigma}$ can be compared with gas molecules in statistical physics [19] as shown in Table I, showing a strong resemblance to a particle system.

As the active dipoles interfere with one another, we denote N_{ij}^c as the range of channel contention for an active dipole $\sigma_{ij} = 1$. Due to strong interference, the neighboring dipoles within N_{ij}^c cannot be active simultaneously. This results in channel contention due to either MAC or SINR constraints [6]. Consider a circle, $C(X_i, r_m(X_i))$, centered at a transmitter positioned at X_i with a radius $r_m(X_i) = (\frac{P_{max}}{P_{th}})^{-\alpha}$, where P_{max} is the maximum transmission power and P_{th} is a certain power threshold. For an active dipole $\sigma_{ij} = 1$, the MAC constraint defines a contention region which is a union of $C(X_i, r_m(X_i))$ and $C(X_j, r_m(X_j))$.

Similarly, consider a circle centered at receiver X_j with radius $r_s(\sigma_{ij}) = min\{r|\ \text{SINR}_r \geq \text{SINR}_{th}\}$, where $\text{SINR}_r = \frac{P_i l_{ij}^{-\alpha}}{N_b + \sum_{m \in \{l_{mj} \geq r\}} P_m l_{mj}^{-\alpha} \frac{\sigma_{mn}+1}{2}}$ is the signal to interference ratio at receiver j due to transmitters at least r distance apart, N_b is the noise power, and SINR_{th} is a SINR threshold. For an active dipole $\sigma_{ij} = 1$, a SINR constraint defines a SINR-contention region $C(X_j, r_s(\sigma_{ij}))$. Thus for $\sigma_{mn} = -1$, $\forall\ m$ and n, $\|X_m - X_j\| < r_s(\sigma_{ij})$. We denote the minimum SINR contention range of active dipoles as r_c, where r_c=min$\{r_s(\sigma_{ij})\}$ for $\forall\ \sigma_{ij} = 1$.

This work considers the case of channel contention due to SINR constraint, which corresponds to the physical model in [6].

One dipole can be on within a contention range, and multiple dipoles can be active concurrently outside the contention region, resulting in interference. Relevant interference neighbors are those within a so-called interference range r_i. For an active dipole σ_{ij}, the set of active dipoles outside $r_s(\sigma_{ij})$ but within r_i is denoted as N_{ij}^I.

The effective system energy of a communication configuration can then be characterized by the total received communication power at each receiver summing over all transmitters in the network, $\sum_{ij} P_j \frac{\sigma_{ij}+1}{2}$. Following the definitions in statistical physics [19], the "configuration Hamiltonian" is the negative system energy,

$$H^{'}(\underline{\sigma}|\underline{X}) = -\sum_{ij} P_j \cdot \eta_{ij} + \beta \cdot \sum_{ij} (\text{SINR}_{ij} - \text{SINR}_{th})^2 \eta_{ij}, \qquad (1)$$

where $\eta_{ij} = \frac{\sigma_{ij}+1}{2}$, $\text{SINR}_{ij} = \frac{P_i l_{ij}^{-\alpha} \eta_{ij}}{\sum_{mn \neq ij} P_m l_{mj}^{-\alpha} \eta_{mn} + N_{b_{ij}}}$, and β is a large positive constant. For simplicity, we assume $N_{b_{ij}} = N_b$ for $1 \leq i, j \leq N$. $\beta \cdot (\text{SINR}_{ij} - \text{SINR}_{th})^2$ serves as a penalty term for the SINR constraint.

To show the interactions among dipoles, the resulting Hamiltonian can be rewritten as

$$H^{'}(\underline{\sigma}|\underline{X}) = R_1(\underline{\sigma},\underline{X}) + R_2(\underline{\sigma},\underline{X}) + R_3(\underline{\sigma},\underline{X}) + R_I(\underline{\sigma},\underline{X}), \quad (2)$$

where $R_1(\underline{\sigma},\underline{X}) = \sum_{ij}\alpha_{ij}\eta_{ij}$ consists of individual dipoles, $R_2(\underline{\sigma},\underline{X}) = \sum_{ij}\sum_{mn\in N_{ij}^I}\alpha_{ij,mn}\eta_{ij}\eta_{mn}$ is the second-order energy with products of two dipoles, $R_3(\underline{\sigma},\underline{X}) = \sum_{ij}\sum_{mn\in N_{ij}^I}\sum_{uv\in\{N_{ij}^I, N_{mn}^I\}}\alpha_{ij,mn,uv}\eta_{ij}\eta_{mn}\eta_{uv}$ is the third-order energy with products of three dipoles, $R_I(\underline{\sigma},\underline{X}) = \sum_{ij}R_{I_{ij}}(\underline{\sigma},\underline{X})$ is the total interference outside the interference range, and $R_{I_{ij}}(\underline{\sigma},\underline{X})$ is the interference outside the interference range experienced by an active dipole $\sigma_{ij}=1$.

In Equation (2), link activities $\underline{\sigma}$ are coupled by coefficients which depend on node positions,

$$\begin{aligned}
\alpha_{ij} &= -P_i l_{ij}^{-\alpha} + \beta\cdot(P_i l_{ij}^{-\alpha} - \text{SINR}_{th}N_b)^2, \quad (3)\\
\alpha_{ij,mn} &= 2\sqrt{P_iP_m}l_{ij}^{-\frac{\alpha}{2}}l_{mj}^{-\frac{\alpha}{2}} - P_m l_{mj}^{-\alpha} + \beta\text{SINR}_{th}^2P_m^2 l_{mj}^{-2\alpha}\\
&\quad -2\beta(P_i l_{ij}^{-\alpha} - \text{SINR}_{th}N_b)\cdot\text{SINR}_{th}P_m l_{mj}^{-\alpha},\\
\alpha_{ij,mn,uv} &= -2\sqrt{P_mP_u}l_{mj}^{-\frac{\alpha}{2}}l_{uj}^{-\frac{\alpha}{2}} + \beta\cdot(\text{SINR}_{th}^2P_mP_u l_{mj}^{-\alpha}l_{uj}^{-\alpha}),
\end{aligned}$$

Note that the Hamiltonian in Equation (2) is obtained exactly according to given channel conditions and SINR constraint.

B. Random Node Positions

Now consider random node positions $\underline{X}$. First, without any management purpose, nodes can move at random as "free movement". Such a behavior can be characterized by a two-dimensional random-walk around initial positions $\underline{X_0}$, and $P(\underline{X})$ is a multi-variate Gaussian distribution, a special case of which is Gaussian distribution with an exponent $H(\underline{X}) = \sum_i\frac{(X_i-X_i(0))^2}{2\sigma^2}$, where σ^2 is the variance of node movement. Such a distribution also applies when X_i is an estimated sensor location with an independent Gaussian measurement error.

With management constraints on topology, nodes move cooperatively. A simple and commonly-used constraint is on the 1-connectivity. The 1-connectivity can be achieved by constructing a Yao-like graph, where each node has an edge with the nearest neighbor for every θ ($\leq\frac{2\pi}{3}$) radian [20]. Such a constraint can be represented as a cost function,

$$C(X_i,X_j) = \begin{cases} 0 & ,\frac{|l_{ij}-l_{th}|}{l_{th}} \leq \epsilon_0\\ |l_{ij}-l_{th}| & ,\text{otherwise}\end{cases} \quad (4)$$

where ϵ_0 is a small constant, l_{th} is a threshold, l_{ij} is the distance between nodes i and j, and $j\in N_i^\theta$ with N_i^θ is the set of the nearest neighbors of node i for every angle θ (θ=90^o in this work). This cost function is referred to as the "Yao-like graph constraint."

The extended Hamiltonian for the physical topology is

$$H^{'}(\underline{X}) = H(\underline{X}) + \zeta\cdot\sum_i\sum_{j\in N_i^\theta}C(X_i,X_j), \quad (5)$$

where $H(\underline{X})$ is due to free movements, and $\zeta>0$ is a large weighting factor.

C. Joint Distribution of Link Activities and Node Positions

1) Overall Configuration Hamiltonian

Combining the Hamiltonians for physical and logical configurations results in an overall configuration Hamiltonian of the network,

$$H^{'}(\underline{\sigma},\underline{X}) = \varsigma_\sigma\cdot H^{'}(\underline{\sigma}|\underline{X}) + \varsigma_X\cdot H^{'}(\underline{X}), \quad (6)$$

where ς_σ and ς_X are scaling constants, $0\leq\varsigma_\sigma,\varsigma_X\leq 1$, which weight the relative importance of the two Hamiltonians [11] [1].

2) Boltzmann Distribution

The configuration Hamiltonian can be related to a probabilistic network model through Boltzmann distribution [11], widely-used in statistical physics and machine learning. In a particle system, the effective system energy, $H(\omega)$, is used to describe configuration ω of interacting particles of two states. $H(\omega)$ is known as the configuration Hamiltonian and has been widely used to describe the most probable energy distribution of a particle system [5]. The probability distribution of configuration Hamiltonian $H(\omega)$ obeys the Boltzmann distribution, $P(\omega) = Z_0^{-1}\cdot\exp\frac{-H(\omega)}{T}$, where Z_0 is a normalizing constant and T is the temperature [5].

Cross-layer network model:

$$P(\underline{\sigma},\underline{X}) = {Z_0}^{-1}\cdot\exp\frac{-H^{'}(\underline{\sigma},\underline{X})}{T}, \quad (7)$$

where $Z_0 = \sum_{(\underline{\sigma},\underline{X})}\exp\frac{-H^{'}(\underline{\sigma},\underline{X})}{T}$ is a normalization constant, and T corresponds to the temperature.

3) Minimum Hamiltonian and Optimal Configuration

An optimal configuration is the one that maximizes the likelihood function, i.e., minimizes the configuration Hamiltonian.

$$(\underline{\sigma}^*,\underline{X}^*) = \arg\max_{(\underline{\sigma},\underline{X})}P(\underline{\sigma},\underline{X}) = \arg\min_{(\underline{\sigma},\underline{X})}H(\underline{\sigma},\underline{X}). \quad (8)$$

Therefore, from an optimization standpoint, an optimal configuration should satisfy the management objectives, which, for example, correspond to a 1-connected physical topology and SINR constraint. When the constraints are satisfied, the penalty terms should be diminishing. An optimal configuration should also achieve the spatial channel reuse maximization, provided that active dipoles are sufficiently separated to satisfy the SINR constraint. Rigorous analysis will be provided in our future work to show this result formally.

IV. Local Model: Probabilistic Graphs

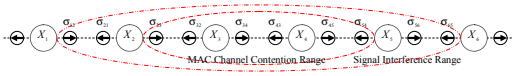
We now investigate when fully distributed algorithms are possible for configuration management. This suffices to examine whether there exists an approximation $P^l(\underline{\sigma},\underline{X})$ to the global model $P(\underline{\sigma},\underline{X})$, which possesses a certain spatial Markov property. The approach we use is probabilistic graphical models.

A. Graphical Representation

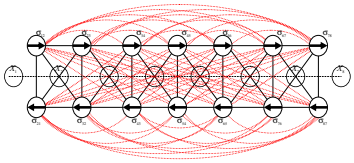
Graphical models relate a probability distribution of random variables with a corresponding dependency graph [5] [7] [8]. A node in the graph represents a random variable and a link between two nodes characterizes their statistical dependence. In particular, a set of random variables $\underline{v}$ forms Gibbs Random Field (GRF) if it obeys a Gibbs distribution [11]. A Gibbs distribution has the same form as Boltzmann distribution, which satisfies the positivity condition, meaning that all configurations have a positive probability. Hammersley-Clifford theorem shows that, if an additional spatial Markov condition is satisfied, a Gibbs distribution can be represented as a Markov Random Field (MRF), and vice versa.

Hammersley-Clifford Theorem [11]: A random field $\underline{v}$ is an MRF on S with respect to $\aleph$ if and only if $\underline{v}$ is a Gibbs Random Field (GRF) on S with respect to $\aleph$, where S is the set of nodes $S = \{1, \cdots, N\}$, $\underline{v}$ is a set of random variable of nodes $\underline{v} = \{v_1, \cdots, v_N\}$, and $\aleph$ is the neighborhood system $\aleph = \{N_i | i \in S\}$

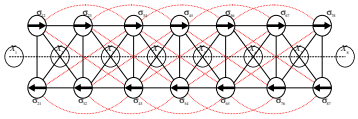
This theorem shows that Markov Random Fields correspond to an interesting type of graphical models where a random variable is conditionally independent of the other nodes given its neighbors. This conditional independence can be shown explicitly through local connections among nodes in a dependency graph.



(a) MAC and Interference Neighborhood on Linear Topology



(b) Dependence of $\underline{\sigma}$ with $R_I(\underline{\sigma}, \underline{X})$



(c) Dependence of $\underline{\sigma}$ while Ignoring $R_I(\underline{\sigma}, \underline{X})$

Fig. 1. Coordination (i.e., Dependence) Graph of $(\underline{\sigma}|\underline{X})$

Consider an example of a one-dimensional topology and the corresponding dependency graph for $(\underline{\sigma}|\underline{X})$ such as in Figure 1. Nodes in the graph represent binary random variable σ_{ij}'s, and the links represent their spatial dependence. For example, solid lines show the dependence due to channel contention, and the dashed lines indicate the dependence due to interference.

Since link activities are all dependent due to interference, the graph is fully connected as in Figure 1 (b). This shows an uninteresting case of MRF where the neighborhood of a node on the graph is the entire network. Therefore, obtaining an optimal configuration requires a centralized maximization of the global likelihood function, where each node exchanges information with all the rest in the network.

However, the interference outside the interference range, $R_I(\underline{\sigma}, \underline{X})$, can be relatively small compared to the first three terms of the configuration Hamiltonian in Equation (2). The third-order term, $R_3(\underline{\sigma}, \underline{X})$, also can be small compared to the second-order term. Hence, we may use the first two terms to approximate the configuration Hamiltonian, where

$$H^l(\underline{\sigma}|\underline{X}) = -\sum_{ij} \alpha_{ij}(\underline{X}) \cdot \eta_{ij} + \sum_{ij} \sum_{mn \in N_{ij}^I} \alpha_{ij,mn}(\underline{X}) \cdot \eta_{ij}\eta_{mn}, \quad (9)$$

and the corresponding Boltzmann distribution is

$$P^l(\underline{\sigma}|\underline{X}) = Z_\sigma^{-1} \cdot \exp\frac{-H^l(\underline{\sigma}|\underline{X})}{T}. \quad (10)$$

As the sum in Equation (9) only involves $(mn) \in N_{ij}^I$, the resulting dependency graph now has a small neighborhood as shown in Figure 1 (c). In fact this approximated Markov Random Field is the well-known second-order Ising model [11] where the Hamiltonian $H^l(\underline{\sigma}|\underline{X})$ consists of both the first-order and the second-order terms of σ_{ij}'s.

We now examine the spatial dependence of node positions $\underline{X}$. In general, the random field of node positions $\underline{X}$ may not be Markovian and thus correspond to a fully connected graph. However, several well-known management objectives result in physical topologies with Markovian dependence, e.g., a second-order MRF with the Yao-like graph constraint (Equation (4)). [1]

B. Cross-layer Markov Random Fields

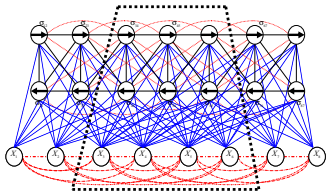


Fig. 2. Cross-Layer Coordination Graph and Clique of $(\underline{\sigma}, \underline{X})$

The probabilistic graph of an overall configuration can be obtained by combining the two graphs for the logical and physical configuration. For ease of illustration, we still consider the one-dimensional network in Figure 1 (a). A cross-layer graph is shown in Figure 2 as an approximation of the original overall configuration. The entire graph is thus locally-connected, exhibiting Markovian spatial dependence at both layers. The dashed trapezoid indicates a cross-layer clique, where each node depends on the adjacent nodes of dipoles within the interference range.

This cross-layer graph corresponds to a coupled MRF [5], where an Ising model and a second-order MRF are combined together. The graph is also known as a Random-Bond Markov

[1] Non-Markov constraints will be studied in future work.

Random Field [11], where dipoles are connected by random bonds which are determined by node positions. The cross-layer MRF ($\underline{\sigma}$, $\underline{X}$) can also be represented by a chain graph [16] of two MRF blocks, one for $\underline{X}$ and the other for $\underline{\sigma}$. The cross-layer probabilistic graph thus maps the complex spatial dependence of a wireless network to an explicit representation.

V. Optimality and Complexity of Local Model

In this section, we derive conditions for which a distributed algorithm is feasible/infeasible for finding a nearly-optimal configuration.

A. Optimality Conditions

As an optimal configuration can be obtained through minimizing a Hamiltonian, the performance can be measured by the difference between Hamiltonians.

Definition 2: *Performance: Near-optimality of a configuration*

Let $(\underline{\sigma}^*, \underline{X}^*)$ and $(\underline{\hat{\sigma}}, \underline{\hat{X}})$ be configurations minimizing $H()$ and $H^l()$ in Equations (1) and (9), respectively. The performance of a sub-optimal configuration $(\underline{\hat{\sigma}}, \underline{\hat{X}})$ is defined as the approximation error Δ, which is the relative difference between $(\underline{\sigma}^*, \underline{X}^*)$ and $(\underline{\hat{\sigma}}, \underline{\hat{X}})$ measured by the global Hamiltonian $H()$, $\Delta = |\frac{H(\underline{\sigma}^*, \underline{X}^*) - H(\underline{\hat{\sigma}}, \underline{\hat{X}})}{H(\underline{\sigma}^*, \underline{X}^*)}|$. Given a desired performance $\epsilon_{th} > 0$, if $\Delta \leq \epsilon_{th}$, $(\underline{\hat{\sigma}}, \underline{\hat{X}})$ is considered to be nearly-optimal.

We now derive conditions for the near-optimality. This suffices to investigate when the residual interference can/cannot be neglected in the global Hamiltonian. For feasibility of analysis, we assume a homogeneous network where N nodes are positioned uniformly, transmit at the same power level (P_0), and have the same desired SINR threshold (SINR_{th}) and interference range (r_i). To evaluate the performance, we begin with a dipole σ_{ij} and investigate whether and when the residual term $R_{I_{ij}}(\underline{\sigma}, \underline{X})$ is negligible for a given channel and transmission power.

Lemma 1: *Let I_b be an upper bound of the residual interference $R_{I_{ij}}(\underline{\sigma}, \underline{X})$ of an active dipole $\sigma_{ij} = 1$. Then $I_b = \sum_{k=1}^{k_0} P_{max}\, \pi\, \frac{(r_i + (k-1)r_c)^{1-\alpha}}{r_c}$, where k_0 is a constant that satisfies $\sum_{k=1}^{k_0 - 1} 2\pi\, \frac{(r_i + (k-1)r_c)}{r_c} < \lfloor \frac{N-2}{2} \rfloor \leq \sum_{k=1}^{k_0} 2\pi\, \frac{(r_i + (k-1)r_c)}{r_c}$. Then, $\frac{I_b}{2} \leq R_{I_{ij}}(\underline{\sigma}, \underline{X}) \leq I_b$.*

The proof is given in Appendix I. Lemma 1 provides a lower and an upper bound of the residual interference of a single active dipole, and can be used to derive conditions for the local model to either satisfy or violate the SINR constraint.

Lemma 2: *Consider a threshold $V_{th} = \frac{P_{max}}{\text{SINR}_{th}} - \sum_{k=1}^{\lfloor \frac{r_i}{r_c} \rfloor} \frac{\pi}{sin^{-1}(\frac{r_c}{kr_c})} P_{max}(kr_c)^{-\alpha} - N_b$. Consider also SINR penalty term $\beta \cdot U(\text{SINR}_{th} - \text{SINR}_{ij})$ for each active dipole σ_{ij}, where $U(x)$ is the unit step function, $U(x)$=1 if $x \geq 0$, and $U(x) = 0$, otherwise. If $I_b \leq V_{th}$, $\beta \cdot U(\text{SINR}_{th} - \text{SINR}_{ij})$ =0 for $\forall \sigma_{ij} = 1$. If $\frac{I_b}{2} \geq V_{th}$, $\beta \cdot U(\text{SINR}_{th} - \text{SINR}_{ij})$ =β for $\forall \sigma_{ij} = 1$.*

The proof is provided in Appendix II. This lemma shows the conditions when a local model can either satisfy or violate the SINR constraint, respectively. Assume now that a configuration satisfies the SINR constraint so that the penalty term is negligible. By aggregating $R_{I_{ij}}$ for all active dipoles, the upper bound in Lemma 1 can then be used to compare the Hamiltonians.

Theorem 1: *Consider the Hamiltonians, $H(\underline{\sigma}^*|\underline{X})$ and $H(\underline{\hat{\sigma}}|\underline{X})$, where $(\underline{\sigma}^*|\underline{X})$ and $(\underline{\hat{\sigma}}|\underline{X})$ are the configurations that minimize the Hamiltonian of the global and local models, respectively. Assume that these configurations satisfy the SINR constraint for a given physical topology $\underline{X}$. Then $\Delta = |\frac{H(\underline{\sigma}^*|\underline{X}) - H(\underline{\hat{\sigma}}|\underline{X})}{H(\underline{\sigma}^*|\underline{X})}| \leq \epsilon$, where $\epsilon = \frac{1}{\text{SINR}_{th} N_b} \cdot (\sum_{k=1}^{\infty} \frac{\pi}{sin^{-1}\frac{\frac{r_c}{2}}{r_i + (k-1)r_c}} P_{max} \cdot (r_i + (k-1)r_c)^{-\alpha} + \sum_{k_1=1}^{\lfloor \frac{r_i}{r_c} \rfloor} \sum_{k_2=1}^{\lfloor \frac{r_i}{r_c} \rfloor} 2P_{max} k_1^{\frac{-\alpha}{2}} k_2^{\frac{-\alpha}{2}} r_c^{-\alpha}) \cdot \frac{r_i^2}{r_c^2}$.*

The proof is given in Appendix III. The upper bound ϵ is on the performance of a sub-optimal configuration $(\underline{\hat{\sigma}}|\underline{\hat{X}})$. The bound takes a complex form but can be evaluated numerically.

Corollary 1: *Assume that the physical and logical configurations are obtained sequentially in two separate steps. Then $|\frac{H(\underline{\sigma}^*, X^*) - H(\underline{\hat{\sigma}}, \hat{X})}{H(\underline{\sigma}^*, \underline{X}^*)}| \leq \epsilon$.*

The proof is given in Appendix IV. Hence, for a desired performance ϵ_{th}, if $\epsilon \leq \epsilon_{th}$, $(\underline{\hat{\sigma}}, \underline{\hat{X}})$ is nearly-optimal.

B. Complexity

Definition 3: *Communication Complexity*

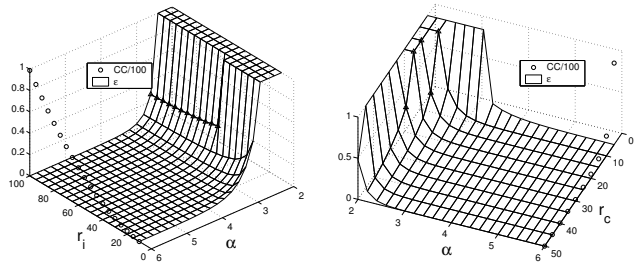
The communication complexity (CC) of a dipole σ_{ij} is defined as the total number of active dipoles within the interference range of the receiver j.

Lemma 4: *The communication complexity (CC) satisfies $CC \leq \lceil \frac{r_i^2 - r_c^2}{r_c^2} \rceil$, where $\lceil y \rceil$ is the smallest integer greater or equal to y.*

The proof can be obtained by packing the circular region between r_c and r_i, and is omitted.

The communication complexity can also be explained through probabilistic graphical representations. Consider the example given in Figure 1 (b). The neighborhood size of an active dipole σ_{ij} is $|N_{ij}^I|$, which corresponds to the communication complexity.

C. Optimality and Complexity Trade-Off



(a) varying r_i and α (b) varying r_c and α

Fig. 3. Trade-off between performance and complexity

Reducing the communication complexity results in a simpler model for a network configuration. The performance, however,

may degrade accordingly. Therefore, a trade-off between the optimality and the complexity needs to be explored.

Figure 3 plots the performance bound ϵ and CC as a function of α r_c and r_i, where CC is scaled by 100 for convenience of illustration. As the communication complexity is independent of the power attenuation factor α (see Lemma 1), we only depict CC for $\alpha = 6$.

Small Interference: First we examine the region where the power attenuation factor is moderate, i.e., $3 < \alpha \leq 6$. This corresponds to such channel environments as shadowed urban areas ($\alpha \in 3 \sim 5$), and obstructed regions in buildings ($\alpha \in 4 \sim 6$) [13]. The corresponding upper bound of the approximation error shown in Figure 3 is small. Moreover, in the range of $\alpha \geq 4$, as the interference range r_i increases, although the communication complexity increases, there is no significant performance gain, as the performance is already upper bounded by a small value (less than 1 %). Thus, these results advocate the approximated local model.

Large Interference: Next we examine the region where the power attenuation factor α is small, i.e., $2 \leq \alpha \leq 3$. This corresponds to such channel environments as free space $\alpha = 2$, obstructed areas in factories $\alpha = 2 \sim 3$, and urban areas $\alpha = 2.7 \sim 3.5$. Those α's correspond to the flat region in Figure 3 (a), where the performance upper bound is large due to a strong interference, and thus truncated to remain flat. Such a large upper bound, however, is not informative of the actual performance. Therefore, we resort to the lower bound of the interference, i.e., $\frac{I_{lb}}{2}$ given in Lemma 2.

The lower bound is evaluated for the flat region. There, the SINR constraint is violated, resulting in a large value for the Hamiltonian $H(\underline{\hat{\sigma}}, \underline{\hat{X}})$ due to a large penalty term (a large β). However, the minimum Hamiltonian $H(\underline{\sigma}^*, \underline{X}^*)$ corresponding to the globally optimal configuration is close to zero. Hence, the approximation error is large, and the local model is not a good approximation to the global network model.

This suggests that the interference from far away dipoles outside of the interference region cannot always be ignored, and the spatial model is not always Markovian with a small-size neighborhood. This is especially true for dense sensor networks with many closely-spaced nodes.

How does the density of nodes affect the performance? We can vary the size of contention region r_c to adjust the density of active dipoles. The range of r_c corresponds to the cases that channel contention is over $2 \sim 25$ hops, which covers most of the feasible scenarios on channel contention.

Figure 3 (b) shows that for a given interference range r_i, as r_c becomes larger, the active dipoles are more sparsely spaced, reducing the upper bound on the approximation error and the communication complexity. This suggests that a logical configuration can be managed so that a dense sensor network can be reduced to a sparse network where the active dipoles are sufficiently far apart to satisfy the SINR requirement.

Theorem 2: *Given α, P_0, l_{th}, and r_i, there exists a minimum r_c, $min(r_c)$, so that $R_{I_{ij}}(\underline{\sigma}, \underline{X})$ is bounded for a given SINR$_{th}$, where* $min(r_c) = \left(\frac{\sum_{i=1}^{\infty} \frac{\pi k^{1-\alpha}}{sin^{-1}(\frac{1}{2k})} P_0 \cdot SINR_{th}}{P_0 l_{th}^{-\alpha} (1+\epsilon_0)^{-\alpha} - Nb \cdot SINR_{th}} \right)^{\frac{1}{\alpha-1}}$.

The proof is given in Appendix V.

VI. Distributed Algorithms

Assume that the near optimality condition holds so that a cross-layer Markov Random Field is a good network model.

A. Distributed Algorithm

The nearly-optimal configuration maximizes the approximated likelihood function,

$$(\underline{\hat{\sigma}}, \underline{\hat{X}}) \quad = \quad \arg\max_{(\underline{\sigma}, \underline{X})} P^l(\underline{\sigma}, \underline{X}) = \arg\min_{(\underline{\sigma}, \underline{X})} H^l(\underline{\sigma}, \underline{X}). \tag{11}$$

Due to the spatial Markov property, maximizing the global likelihood function reduces to maximizing the local likelihood function at cliques, i.e., $P^l(X_i = x_i | X_{N_i})$ and $P^l(\sigma_{ij} = \sigma | X_{N_i}, \sigma_{N_{ij}})$.

Stochastic relaxation can be applied for each node to make local decisions on its new position and transmission activity (shown in Table II). Specifically, at time $t+1$ of kth slot, $\hat{X}_i^k(t+1)$ is the new position that node i should move to, and $\hat{\sigma}_{ij}^k(t+1)$ is the activity of link (i, j). N_i is a neighborhood of node i, N_{ij} is the set of neighboring active dipoles of σ_{ij}. The cooling scheduler is applied to the temperature $T(t) = \frac{T_0}{\log(1+t)}$ with T_0=3.0. This allows an almost-sure convergence to the minimum Hamiltonian [5].

TABLE II
Distributed Stochastic Algorithm

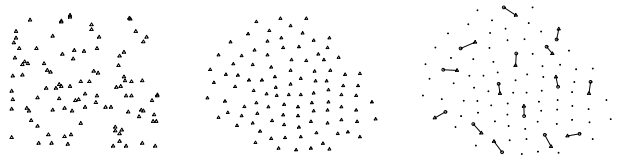
Initial Configuration $(X_i^1(0), \sigma_{ij}^1(0))$,	for $\forall\, i, j > 1$
For $\underline{X}$	1st time slot
$\hat{X}_i^1(t+1)$	$= x_i^1$ with $P^l(X_i^1(t+1) = x_i^1 \| X_{N_i}^1(t))$, which is $\frac{\exp(-\psi_i(x_i^1)/T(t+1))}{\sum_{\forall x_i^1} \exp(-\psi_i(x_i^1)/T(t+1))}$, where $\psi_i(x_i^1)$ $= \frac{(x_i^1 - X_i(0))^2}{2\sigma^2} + \zeta \sum_{j \in N_i^\theta} C(x_i^1, X_j^1(t))$
For $\underline{\sigma}$	kth time slot, for $k \geq 1$
$\sigma_{ij}^k(0)$	$= \sigma_{ij}^{k-1}(\infty)$
$\hat{\sigma}_{ij}^k(t+1)$	$= \sigma_{ij}^k$ with $P^l(\sigma_{ij}^k(t+1) = \sigma_{ij}^k \| X_{N_i}^k(t), \sigma_{N_{ij}}^k(t))$, which is $\frac{\exp(-\psi_{ij}(\sigma_{ij}^k)/T(t+1))}{\sum_{\forall \sigma_{ij}^k} \exp(-\psi_{ij}(\sigma_{ij}^k)/T(t+1))}$, where $\psi_{ij}(\sigma_{ij}^k)$ $= (\alpha_{ij} + \sum_{mn \in N_{ij}^I} \alpha_{ij,mn} \frac{\sigma_{mn}^k(t)+1}{2}) \cdot \frac{\sigma_{ij}^k+1}{2}$
$\sigma_{ij}^{k+1}(t)$	$= -1$ for $t > 0$ if $\sigma_{ij}^k(\infty) = 1$

B. Information Exchange

The clique structure of the cross-layer Markov Random Fields determines the range of information exchange. The type of information exchanged is due to the local likelihood function. For a node i, the update information is either (X_i, P_i, $\sigma_{ij} = 1$) or (X_i, P_i, $\emptyset$). Each node can also update the information of its neighbors according to the neighbors' broadcasting message. For node m that receives a message (X_i, $\sigma_{ij} = 1$), if $\|X_m - X_i\| < r_i$, it can relay this message to its neighbors; otherwise, the message can be dropped.

VII. Self-Configuration

Spatial reuse TDMA is used as an example of distributed self-configuration. STDMA is a scheme which schedules transmissions of nodes in time so that the maximum spatial channel-reuse can be achieved while maintaining fairness, 1-connectivity, and SINR constraints. In particular, STDMA is composed of multiple time slots which result in a so-called STDMA-cycle. For fairness, these active links cannot be active again until all their neighboring links access the shared channel.



(a) $\underline{X_0}$ (b) $\underline{X}$ (c) $\underline{\sigma}$ given $\underline{X}$

Fig. 4. Self-configuration with local optimal algorithm

Simulations: We choose a sensor network of 100 nodes with an initial configuration $(\underline{\sigma_0}, \underline{X_0})$ shown in Figure 4 (a). To implement a Yao-like graph, $l_{th} = 2$ and $\theta = \frac{\pi}{2}$ are chosen for the 1-connectivity constraint. Assume for now an ideal case where each node is equipped with a global positioning system (GPS) and broadcasts its position to the neighbors [2]. The iteration of local rules in Table II stops when the steady state is reached, i.e., $\frac{|\,\|X_i - X_j\| - l_{th}\,|}{l_{th}} < 0.01$ for $j \in N_i^\theta$. Figures 4 (b) and (c) show the resulting configuration at the $1st$ time slot of a STDMA cycle.

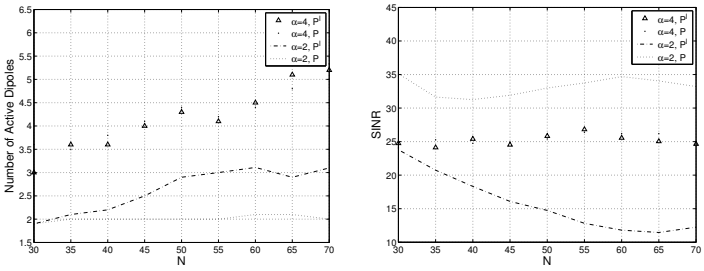
Figure 5 (a) shows the one-hop capacity with the change of number of nodes in the network. As the one-hop capacity of centralized schemes is $O(N)$ [6] [9], the distributed algorithm achieves the one-hop capacity of the same rate, which is linear in N. Figure 5 (b) shows that when the interference decays sufficiently fast e.g., $\alpha = 4$, the distributed algorithm based on local decisions can satisfy the SINR constraint (i.e., $\text{SINR}_{th} = 20$). But when the interference decays slowly, e.g., $\alpha = 2$, the local model communicates with too few neighbors by ignoring the interference from far-away nodes, and thus fail to satisfy the SINR constraint. Therefore, when the interference is strong, the distributed algorithm based on local information exchange is insufficient.

VIII. Other Related Work

Optimality: Optimality of network capacity has been investigated for wireless ad hoc and sensor networks [6] [9] [15]. The algorithmic optimality on "how to self-configure using a near-optimal distributed algorithm" needs to be further studied [10]. In this work, we show that if the configuration Hamiltonian can be well approximated into local potential functions with a bounded approximation error, a near-optimal configuration can be obtained using distributed algortihms.

Topology-formation: Topology-formation has been investigated in the context of emerging behavior for mobile nodes [1], intelligent agents [17], information management units, and mobile robots [14]. The angle to approach topology formation is usually to model the behavior of a system externally. In this work, we show that not only a physical topology, but also a logical topology and their combinations can be configured in a fully distributed fashion.

Link scheduling: Fair link-scheduling is investigated through distributed approaches at the MAC layer using the contention graph [12] [9]. STDMA link-scheduling is formulated by [2] using linear programming and an optimal centralized algorithm. This provides a benchmark for the performance of heuristic algorithms, but distributed algorithm is lacking. A distributed TDMA algorithm at the MAC layer is developed by [4] for sensor networks based on graph coloring. The performance of the approach is assessed through simulations.



(a) One-hop capacity (b) Achieved SINR

Fig. 5. Performance: One-hop capacity and SINR. P^l is the local model and P the global model.

IX. Conclusion

In this work, we have studied *when* distributed configuration management is nearly-optimal and *how* to obtain a nearly-optimal configuration of a wireless sensor network using a distributed algorithm. We have obtained the following findings.

(a) A global model can be obtained on a network configuration as the spatial distribution of node positions and link activities with management constraints.

(b) The complex spatial dependence in a configuration can be represented explicitly by a probabilistic graphical model. When the graph is a Markov Random Field, distributed configuration management is optimal.

(c) Optimality conditions on when an approximated model is valid are derived for distributed configuration management under various channel conditions and density of sensor nodes. Communication complexity is defined as the neighborhood size in an interference range and the corresponding cross-layer Random Fields. Trade-off between the optimality and the complexity is studied for various network conditions.

(d) Using probabilistic inference, distributed algorithms can be derived directly from the Markov Random Field model when the optimality condition holds. The distributed algorithms have been applied to STDMA slot allocation.

Different from top-down approaches in machine learning which usually assume a model, we have taken a bottom-up approach so that conditions and algorithms on "when and how" can be studied through internal network properties and management constraints. More complex network conditions will be studied in our future work.

[2]This requirement may be relaxed as only relative positions are needed.

Acknowledgement Support from NSF ECS 9908578 is gratefully acknowledged. The authors would like to thank Mo Li, John Barry, Guanglei Liu, Zesheng Chen, Sung Ho Hwang, Junbeum Kim for many useful discussions.

APPENDIX

I. PROOF OF LEMMA 1

Proof: For an active dipole σ_{ij}, the set of kth nearest neighboring active dipoles is denoted with G_k, which is depicted in Figure 6 (a). The cardinality of G_k, $|G_k|$, is upper bounded by $|G_k| \leq \frac{2\pi}{\theta} = \frac{\pi}{sin^{-1}\frac{\frac{r_c}{2}}{r_i+(k-1)r_c}} < \frac{2\pi(r_i+(k-1)r_c)}{r_c}$ (refer Figure 6 (a) for θ). We denote $max\{r_s(\sigma_{ij})\}$ for $\forall\ \sigma_{ij} = 1$ with r_{max}. Since we consider the configuration that maximizes the spatial channel reuse and the nodes are assumed to be uniformly and densely distributed, as the node density increases in the network, the probability P($r_{max} \leq 2r_c$) asymptotically reaches one. Thus, the cardinality of G_k is lower bounded by $|G_k| > \frac{\pi(r_i+(k-1)r_c)}{r_c}$.

We denote $\sum_{k=1}^{k_0} P_{max}\ \pi\ \frac{(r_i+(k-1)r_c)^{1-\alpha}}{r_c}$ with I_b, where k_0 is a constant that satisfies $\sum_{k=1}^{k_0-1} 2\pi\ \frac{(r_i+(k-1)r_c)}{r_c} < \lfloor\frac{N-2}{2}\rfloor \leq \sum_{k=1}^{k_0} 2\pi\ \frac{(r_i+(k-1)r_c)}{r_c}$. The value $N-2$ denotes the number of nodes except nodes i and j.

Thus, $\frac{I_b}{2} \leq R_{I_{ij}}(\underline{\sigma}, \underline{X}) \leq I_b$. ■

II. PROOF OF LEMMA 2

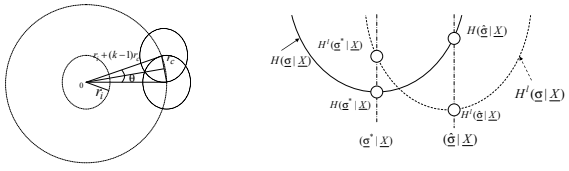
Proof: For an active dipole σ_{ij}, $\frac{P_{max}}{\sum_{k=1}^{\lfloor\frac{r_i}{r_c}\rfloor} \frac{\pi}{sin^{-1}(\frac{\frac{r_c}{2}}{kr_c})} P_{max}(kr_c)^{-\alpha} + R_{I_{ij}}(\underline{\sigma},\underline{X})) + N_b} \geq \text{SINR}_{th}$. Let V_{th} denote a certain threshold, $\frac{P_{max}}{\text{SINR}_{th}} - \sum_{k=1}^{\lfloor\frac{r_i}{r_c}\rfloor} \frac{\pi}{sin^{-1}(\frac{\frac{r_c}{2}}{kr_c})} P_{max}(kr_c)^{-\alpha} - N_b$.

If $I_b \leq V_{th}$, then the local model $P^l()$ always satisfies the SINR constraints, i.e., $\beta \cdot U(\text{SINR}_{th} - \text{SINR}_{ij}) \cdot \eta_{ij} = 0$. Similarly, if $\frac{I_b}{2} \geq V_{th}$, then the local model $P^l()$ always violates the SINR constraints, i.e., $\beta \cdot U(\text{SINR}_{th} - \text{SINR}_{ij}) \cdot \eta_{ij} = 1$. ■

III. PROOF OF THEOREM 1

Proof: $|\frac{H(\underline{\sigma}^*|\underline{X}) - H(\underline{\hat{\sigma}}|\underline{X})}{H(\underline{\sigma}^*|\underline{X})}| \leq \frac{|H(\underline{\sigma}^*|\underline{X}) - H(\underline{\hat{\sigma}}|\underline{X})|^u}{|H(\underline{\sigma}^*|\underline{X})|^l}$, where the super-scripts u and l denote the upper and lower bound of corresponding quantity.

Figure 6 (b) shows the relative position of related Hamiltonian values, i.e., $H(\underline{\sigma}^*|\underline{X})$, $H(\underline{\hat{\sigma}}|\underline{X})$, $H^l(\underline{\sigma}^*|\underline{X})$, and $H^l(\underline{\hat{\sigma}}|\underline{X})$. Note that $H(\underline{\sigma}^*|\underline{X}) \leq H(\underline{\hat{\sigma}}|\underline{X})$ and $H^l(\underline{\hat{\sigma}}|\underline{X}) \leq H^l(\underline{\sigma}^*|\underline{X})$. Note that $H(\underline{\sigma}|\underline{X}) = H^l(\underline{\sigma}|\underline{X}) + R_I(\underline{\sigma}, \underline{X})$, where $R_I(\underline{\sigma}, \underline{X}) > 0$ from definition in Equation (1). Thus $H^l(\underline{\hat{\sigma}}|\underline{X}) \leq H(\underline{\sigma}^*|\underline{X}) \leq H(\underline{\hat{\sigma}}|\underline{X})$. As a result, $|H(\underline{\sigma}^*|\underline{X}) - H(\underline{\hat{\sigma}}|\underline{X})| \leq |H(\underline{\hat{\sigma}}|\underline{X}) - H^l(\underline{\hat{\sigma}}|\underline{X})|$.



(a) Cardinality of kth Circle (b) Optimal and Suboptimal Hamiltonian

Fig. 6. Cardinality and Hamiltonians

For any configuration $(\underline{\sigma_a}|\underline{X_a})$, $|H(\underline{\sigma_a}|\underline{X_a}) - H^l(\underline{\sigma_a}|\underline{X_a})| = |R_I(\underline{\sigma_a}, \underline{X_a}) + R_3(\underline{\sigma_a}, \underline{X_a})| \leq |R_I(\underline{\sigma_a}, \underline{X_a})| + |R_3(\underline{\sigma_a}, \underline{X_a})|$. Thus, $|H(\underline{\sigma}^*|\underline{X}) - H(\underline{\hat{\sigma}}|\underline{X})| \leq |H(\underline{\hat{\sigma}}|\underline{X}) - H^l(\underline{\hat{\sigma}}|\underline{X})| = |R_I(\underline{\hat{\sigma}}, \underline{X}) + R_3(\underline{\hat{\sigma}}, \underline{X})| \leq |R_I(\underline{\hat{\sigma}}, \underline{X})| + |R_3(\underline{\hat{\sigma}}, \underline{X})| \leq (\sum_{k=1}^{\infty} \frac{\pi}{sin^{-1}\frac{\frac{r_c}{2}}{r_i+(k-1)r_c}} P_{max}\ (r_i + (k-1)r_c)^{-\alpha} + \sum_{k_1=1}^{\lfloor\frac{r_i}{r_c}\rfloor} \sum_{k_2=1}^{\lfloor\frac{r_i}{r_c}\rfloor} 2P_{max}\ (k_1 r_c)^{\frac{-\alpha}{2}}\ (k_2 r_c)^{\frac{-\alpha}{2}}\ ) \cdot \hat{N_\sigma}$, where $\hat{N_\sigma}$ is total number of active dipoles in $(\underline{\hat{\sigma}},\underline{X})$. Thus, $|H(\underline{\sigma}^*|\underline{X}) - H(\underline{\hat{\sigma}}|\underline{X})|^u = (\sum_{k=1}^{\infty} \frac{\pi}{sin^{-1}\frac{\frac{r_c}{2}}{r_i+(k-1)r_c}} P_{max}(r_i + (k-1)r_c)^{-\alpha} + \sum_{k_1=1}^{\lfloor\frac{r_i}{r_c}\rfloor} \sum_{k_2=1}^{\lfloor\frac{r_i}{r_c}\rfloor} 2P_{max} k_1^{\frac{-\alpha}{2}} k_2^{\frac{-\alpha}{2}} r_c^{-\alpha}) \cdot \hat{N}_\sigma$.

$|H(\underline{\sigma}^*|\underline{X})| \geq min\{(S_{ij} - I_{ij}) \cdot \frac{\sigma_{ij}+1}{2}\} \cdot N_\sigma^* \geq \text{SINR}_{th} N_b \cdot N_\sigma^*$, where S_{ij} and I_{ij} are the signal and the interference of an active dipole σ_{ij}, N_b is power of background noise, and N_σ^* is total number of active dipoles in $(\underline{\sigma}^*|\underline{X})$. Thus, $|H(\underline{\sigma}^*|\underline{X})|^l = \text{SINR}_{th} N_b \cdot N_\sigma^*$. $\hat{N_\sigma} = \frac{\pi r_i^2}{\pi(r_c')^2}$ and $N_\sigma^* = \frac{\pi r_i^2}{\pi(r_c^*)^2}$, thus $\frac{\hat{N_\sigma}}{N_\sigma^*} = \frac{(r_c^*)^2}{(r_c')^2} \leq \frac{r_i^2}{r_c^2}$. As a result, $|\frac{H(\underline{\sigma}^*|\underline{X}) - H(\underline{\hat{\sigma}}|\underline{X})}{H(\underline{\sigma}^*|\underline{X})}| \leq \frac{1}{\text{SINR}_{th} N_b} \cdot (\sum_{k=1}^{\infty} \frac{\pi}{sin^{-1}\frac{\frac{r_c}{2}}{r_i+(k-1)r_c}} P_{max}(r_i + (k-1)r_c)^{-\alpha} + \sum_{k_1=1}^{\lfloor\frac{r_i}{r_c}\rfloor} \sum_{k_2=1}^{\lfloor\frac{r_i}{r_c}\rfloor} 2P_{max} k_1^{\frac{-\alpha}{2}} k_2^{\frac{-\alpha}{2}} r_c^{-\alpha}) \cdot \frac{r_i^2}{r_c^2}$. ■

IV. PROOF OF COROLLARY 1

Proof: $|\frac{H(\underline{\sigma}^*,\underline{X}^*) - H(\hat{\underline{\sigma}},\hat{X})}{H(\underline{\sigma}^*,\underline{X}^*)}| \leq \varsigma_\sigma\ |\frac{H(\underline{\sigma}^*|\underline{X}^*) - H(\hat{\underline{\sigma}}|\hat{X})}{H(\underline{\sigma}^*,\underline{X}^*)}| + \varsigma_X\ |\frac{H(\underline{X}^*) - H(\hat{X})}{H(\underline{\sigma}^*,\underline{X}^*)}|$.
In case the management decisions on $(\underline{\sigma}, \underline{X})$ are done sequentially from $\underline{X}$ to $\underline{\sigma}$, $\underline{X}^*$=$\hat{\underline{X}}$. Thus, $|\frac{H(\underline{\sigma}^*,\underline{X}^*) - H(\hat{\underline{\sigma}},\hat{X})}{H(\underline{\sigma}^*,\underline{X}^*)}| \leq \varsigma_\sigma\ |\frac{H(\underline{\sigma}^*|\underline{X}^*) - H(\hat{\underline{\sigma}}|\underline{X}^*)}{H(\underline{\sigma}^*,\underline{X}^*)}| \leq \varsigma_\sigma \cdot \epsilon \leq \epsilon$. ■

V. PROOF OF THEOREM 2

Proof: We assume $P_i = P_0$ for $\forall i$. For an active dipole σ_{ij}, $\text{SINR}_{ij} \geq \frac{P_0 l_{th}^{-\alpha(1+\varepsilon_0)^{-\alpha}}}{N_b + max(I_{ij})}$, where I_{ij} is the total interference. The sufficient condition for $\text{SINR}_{ij} \geq \text{SINR}_{th}$ is to show that $\frac{P_0 l_{th}^{-\alpha(1+\varepsilon_0)^{-\alpha}}}{N_b + max(I_{ij})} \geq \text{SINR}_{th}$. Fromr Figure 6 (a), $max(I_{ij}) \leq \sum_{k=1}^{\infty} \frac{2\pi k r_c}{2sin^{-1}(\frac{1}{2k})} P_0 (kr_c)^{-\alpha}$. As a result, from $\text{SINR}_{th} \leq \frac{P_0 l_{th}^{-\alpha(1+\varepsilon_0)^{-\alpha}}}{N_b + max(I_{ij})}$, $min(r_c) = \left(\frac{\sum_{i=1}^{\infty} \frac{\pi k^{1-\alpha}}{sin^{-1}(\frac{1}{2k})} P_0 \cdot \text{SINR}_{th}}{P_0 l_{th}^{-\alpha(1+\epsilon_0)^{-\alpha}} - N_b \cdot \text{SINR}_{th}} \right)^{\frac{1}{\alpha-1}}$. ■

REFERENCES

[1] J. Baras, and X. Tan, *"Control of Autonomous Swarms Using Gibbs Sampling," IEEE CDC*, 2004.

[2] P. Bjorklund, P. Varbrand, and D. Yuan, *"Resource Optimization of Spatial TDMA in Ad Hoc Radio Networks: A Column Generation Approach," IEEE Infocom*, 2003.

[3] N. Bulusu, J. Heidemann, and D. Estrin, *"GPS-less Low Cost Outdoor Localization for very small devices," IEEE Personal Communications*, 2000.

[4] S. Gandham, M. Dawande, and R. Prakash *"Link Scheduling in Sensor Networks: Distributed Edge Coloring Revisited," IEEE Infocom*, 2005.

[5] S. Geman, and D. Geman, *"Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," IEEE Trans. PAMI*, 1984.

[6] P. Gupta, and P.R. Kumar, *"The Capacity of Wireless Networks," IEEE Trans. Information Theory*, 2000.

[7] M. Jordan, and Y. Weiss, *"Graphical Models: Probabilistic Inference," Handbook of Neural Networks and Brain Theory*, 2002.

[8] F. Kschischang, B. Frey, H. Loeliger *"Factor graphs and the sum-product algorithm," IEEE Trans. on Information Theory,* 2001.

[9] J. Li, C. Blake, D. De Couto, H. Lee, and R. Morris, *"Capacity of Ad Hoc Wireless Networks,", ACM Mobicom*, 2001.

[10] P. Kyasanur, and N. Vaidya *"Detection and Handling of MAC Layer Misbehavior in Wireless Networks," DCC at the DSN*, 2003

[11] S. Z. Li, *"Markov Random Field Modeling in Computer Vision," Springer-Verlag*.

[12] H. Luo, S. Lu, V. Bharghavan, J. Cheng, and G. Zhong, *"A Packet Scheduling Approach to QoS Support in Multihop Wireless Networks," ACM MONET*, 2004.

[13] T. Rappaport, *"Wireless Communications: Principles and Practice," Prentice Hall*.

[14] Z. Butler, K. Kotay, D. Rus, and K. Tomita *"Generic Decentralized Control for a Class of Self-Reconfigurable Robots," ICRA*, 2002.

[15] S. Shakkottai, R. Srikant, and N. Shroff, *"Unreliable Sensor Grids: Coverage, Connectivity and Diameter," IEEE Infocom*, 2003.

[16] P. Smyth, D.Heckerman, and M. Jordan, *"Probabilistic independence networks for hidden Markov probability Models," Neural computation*, 1997.

[17] M. Steenstrup, *"Emergent Structure Among Self-Organizing Devices," NSF-RPI Workshop on PCN*, 2004.

[18] J. Tsitsiklis and M. Athens, *" On the Complexity of Decentralized Decision Making and Detection Problem," IEEE Trans. Automatic Control* , 1985.

[19] A. Wagner and V. Anantharam, *" Wireless sensor network design via interacting particles," Allerton*, 2002.

[20] R. Wattenhofer, L. Li, P. Bahl, and Y. Wang, *"Distributed topology control for power efficient operation in multihop wireless networks," IEEE Infocom*, 2001.