Randomized and Distributed Self-Configuration of Wireless Networks: Two-Layer Markov Random Fields and Near-Optimality

Sung-eok Jeon and Chuanyi Ji, *Senior Member, IEEE*

***Abstract*—This work studies the near-optimality versus the complexity of distributed configuration management for wireless networks. We first develop a global probabilistic graphical model for a network configuration which characterizes jointly the statistical spatial dependence of a physical- and a logical-configuration. The global model is a Gibbs distribution that results from the internal network properties on node positions, wireless channel and interference; and the external management constraints on physical connectivity and signal quality. A local model is a two-layer Markov Random Field (i.e., a random bond model) that approximates the global model with the local spatial dependence of neighbors. The complexity of the local model is defined through the communication range among nodes which corresponds to the number of neighbors in the two-layer Markov Random Field. The local model is near-optimal when the approximation error to the global model is within a given bound. We analyze the tradeoff between approximation error and complexity. We then derive sufficient conditions on the near-optimality of the local model. For a fast decaying wireless channel with power attenuation factor $\alpha > 4$, a node only needs to communicate with $O(1)$ neighbors for a local model to be near optimal. For a slowly decaying channel with a power attenuation factor $2 \leq \alpha < 4$, a node may have to communicate with more than $O(N^{(4-\alpha)/4})$ neighbors to result in a bounded approximation error. If the communication range is kept to be $O(1)$, a bounded approximation error can also be achieved by reducing the density of active links to $O(N^{(\alpha-4)/(\alpha+4)})$ for $\alpha < 4$ and $O(1)$ for $\alpha > 4$. The two-layer Markov Random Fields enable a class of randomized distributed algorithms such as the stochastic relaxation that allows a node to self-configure based on information from neighbors. We validate the model, the analysis and the randomized distributed algorithms through simulation.**

***Index Terms*— Near-optimality, randomized and distributed management, self-configuration, two-layer Markov Random Field.**

I. Introduction

WIRELESS infrastructureless networks such as sensor- and actor-networks, wireless mesh networks, wireless cognitive networks, and mobile agent networks [2], [6] can be characterized by a physical configuration (node positions) and a logical configuration (node-to-node communications). Both a physical and a logical configuration can vary due to channel conditions, node or link failures, and environmental changes. Self-configuration is for network nodes to independently adapt a physical and a logical configuration to support end-to-end communications between nodes. Locality, randomness and spatial dependency are three important characteristics of configuration management of infrastructureless wireless networks. For example, nodes can either fail locally or move away from a geographical proximity. Local adaptation is thus desirable for preventing incessant network-wide reconfiguration. Furthermore, local adaptation is often inevitable for infrastructureless networks where nodes have information only from their neighbors. Local information available at a node includes locations of neighboring nodes and communication activities of their adjacent links. Such local information exhibits randomness due to uncertainty of user behavior and incomplete information. The corresponding network variables are spatially dependent due to either constraints on physical and logical configurations or information exchange among neighbors. Distributed self-configuration should take locality, randomness and dependency into consideration and allow nodes to self-adapt local configurations, resulting in a network-wide optimal configuration collectively.

The range of information exchange determines the communication complexity. When the information-exchange is among close neighbors, the complexity is moderate; and the resulting distributed self-configuration scales gracefully with network size. An open issue is whether distributed self-configuration would result in a near-optimal configuration with a moderate complexity. This work studies the open issue by developing an analytical model for distributed self-configuration and quantifying near-optimality versus complexity using the model.

Prior Work: Numerous distributed algorithms and protocols have been developed for topology formation [28], self-organizing sensor networks [6], and p2p self-stabilizing networks (see [17] and references therein). These distributed algorithms are deterministic and do not involve random variables. Srinivasan *et al.* [25] develops a randomized and distributed algorithm for edge coloring. There, Markov dependence is assumed among variables using a bounded nodal degree on bipartite graphs. Modiano *et al.* [24] develops gossiping algorithms to show that the throughput of wireless networks can be maximized in a fully distributed fashion. Baras *et al.* [1] develops a probabilistic model with spatial Markov assumptions to characterize randomness in node positions. Doyle *et al.* [3] develops a framework based on Markov Random Fields for routing in ad hoc networks. Meshkova *et al.* applied simulated annealing to learn parameters of graphical models

Manuscript received November 09, 2009; accepted May 01, 2010. Date of publication June 07, 2010; date of current version August 11, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Huaiyu Dai. The authors gratefully acknowledge the support from NSF ECS 0300605 and ECS 990857.

S. Jeon is with the Microsoft, Redmond, WA 98052 USA (e-mail: sujeon@microsoft.com).

C. Ji is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: jic@ece.gatech.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2010.2051806

for cross-layer optimization [23]. These randomized algorithms assume Markovian spatial dependence that warrants the optimality of the distributed algorithms. Elbatt *et al.* [5] and Madan *et al.* [22] show that the general SINR-based interference models are non-Markovian. There exists spatial dependence with far interferers when aggregated interference from far interferers is not negligible [5], [22]. Kauffman *et al.* [15] develops a Gibbs distribution for measurement-based self-organization for 802.11 wireless access networks, where each node is aware of all the other nodes' decisions via base stations. Different from access networks, individual nodes in infrastructureless wireless networks often do not have complete information of the entire network.

Two research issues arise. One is how to characterize a large number of dependent random variables in configuration management such as inaccurate node positions, wireless channel conditions, interference, and locally interacting wireless nodes. This issue relates to an open question, i.e., how to systematically develop a Gibbs distribution [8]. The other is how to derive conditions on near-optimal distributed configuration when spatial Markovian assumption does not hold. This issue is related to a challenging problem, i.e., how to develop a distributed algorithm with a predictable performance [26]. This work focuses on the combination of these two issues, and attempts to answer the following questions: i) How to characterize a large number of dependent random variables in a wireless network configuration? ii) What measures the optimality of centralized configuration and the near-optimality of distributed-configuration? and iii) When is distributed configuration near-optimal? This work develops a framework for studying these questions through modeling, analysis, and algorithms.

Global Model and Optimality: Modeling is a key step to characterize dependencies among a large number of random variables. Thus this work intends to derive a model through the first principle, i.e., to model the spatial dependencies in a network configuration based on simple yet commonly used characteristics. In particular, we consider three factors in a model of wireless network configuration management: Randomness resulting from a wireless network internally, management constraints imposed externally, and distributed decisions made by individual nodes with limited neighborhoods. Randomness in wireless networks is challenging to model [12]. This work begins with simple scenarios where the randomness results from perturbed node positions, interference, and node-to-node communication. Wireless channel fading is not considered for simplicity. Management constraints include requirements on the connectivity of a physical topology, SINR, and reconfiguration cost. Nodes make asynchronous and randomized decisions independently, adjusting a configuration in a distributed fashion. The model is developed from bottom-up by mapping these three factors onto a probability distribution.

To obtain an analytical form of the probabilistic distribution of a wireless configuration management, we adopt an analogy between link activities and node positions of a wireless network and interacting particles in statistical physics [11], [27]. Such an analogy allows the use of "configuration Hamiltonian" [18] for quantifying a network configuration. Configuration Hamiltonian combines a physical topology, a logical topology, and management constraints into a single quantity. Configuration Hamiltonian is then used to obtain a probabilistic model as a Gibbs distribution [11]. Such a Gibbs distribution is for an entire network-configuration and thus corresponds to a global probabilistic model of a wireless network. The optimality is defined as a maximum likelihood solution [8], [20], where complete network information is assumed to be available.

Local Model and Probabilistic Graph: A local model is an approximation of the global model. The approximation is based on the spatial dependence among nodes, which can be viewed as a one-to-one mapping between a Gibbs distribution and a probabilistic graphical model in machine learning [8], [14]. Probabilistic graphs provide an explicit graphical representation of statistical spatial dependence in a network configuration. We show that a probabilistic graph of the global model belongs to a two-layer random-field. One layer is for a physical configuration, and the other layer is for a logical configuration. The graph is fully connected, i.e., non-Markovian, where the long-range spatial dependence results from the interference among far-away nodes and the signal-to-interference-plus-noise ratio (SINR) constraint. When the long-range interference can be neglected, the global model can be approximated by a two-layer coupled Markov Random Field which is also called a random bond model [8], [18]. The corresponding dependency graph exhibits a nested spatial Markov dependence for both a physical and logical configuration.

We define an approximation error to measure the difference between the local and global model. When the approximation error is within a given bound, a local model is near-optimal. We derive bounds for the approximation error. We derive sufficient conditions for the local model to be near-optimal, i.e., when the total interference from far-away nodes decreases faster than the size of a network.

Distributed Algorithm: A local model results in a class of distributed algorithms where nodes self-configure through information exchange with neighbors. Node decisions are made by applying randomized distributed algorithms using the graphical model of a wireless network. We apply the distributed algorithm to the self-configuration of wireless network.

The contributions of this work lie in a) developing a two-layer probabilistic graphical model as a mapping of an infrastructureless wireless network configuration under given conditions and b) deriving the approximation error and the bounds for the near-optimality of distributed configuration management. The distributed algorithm naturally results from the graphical model (i.e., Gibbs distribution).

II. Problem Formulation

A. Assumptions

Physical Layer: All nodes share a common frequency channel. A pair of nodes within a communication range can communicate directly with an omni-directional antenna. The wireless channel follows a path-loss model with power attenuation factor α, for $2 \leq \alpha \leq 6$. Shadowing and multi-path fading are not considered in this work for simplicity. Node i transmits with power P_i, where $0 \leq P_i \leq P_{\max}, 1 \leq i \leq N$, $P_{\max}$ is the maximum transmission power, and N is the number of nodes in the network. Power control is not considered in this work for simplicity.

MAC Layer: Let $\mathrm{SINR}_{\mathrm{th}}$ be a given threshold for the SINR requirement. Node i can successfully transmit to node j when

the SINR requirement is satisfied, i.e., $\mathrm{SINR}_{ij} = (P_i l_{ij}^{-\alpha}/(N_b + \sum_{(m,n)\neq(i,j)} P_m l_{mj}^{-\alpha})) \geq \mathrm{SINR}_{\mathrm{th}}$, where l_{ij} is the distance between nodes i and j, and N_b is the noise power.

Configuration Management: Let X_{i0} and X_i be a desired and an actual location of node i, for $1 \leq i \leq N$. $\boldsymbol{X} = \{X_1, \cdots, X_N\}$ is a set of random positions of nodes where the randomness results from perturbed positions, random movements and measurement errors (e.g., GPS measurement error). Let σ_{ij} denote channel-access of link (i,j), where $\sigma_{ij} = 1$ if node i is transmitting to node j; and $\sigma_{ij} = -1$, otherwise. σ_{ij} is referred to as a "communication dipole", for $1 \leq i, j \leq N$, $i \neq j$, and $\boldsymbol{\sigma} = \{\sigma_{1,2}, \cdots, \sigma_{N,N-1}\}$ denotes a set of link activities which forms a logical configuration. Link activities are assumed to be random as they are triggered by network-layer random traffic demands. A network configuration is a vector $(\boldsymbol{\sigma}, \boldsymbol{X})$.

Distributed self-configuration is for each node to iteratively determine its position and the activity of its adjacent links given the positions and the link activities of its neighbors. The objectives of distributed configuration management are a) to form a desired 1-connected physical topology, b) to schedule resource utilization at a given time to maximize the spatial channel-reuse with a desired SINR requirement, and c) to reconfigure upon failures by minimizing reconfiguration cost.

B. Formulation

Let $\mathrm{P}(\boldsymbol{\sigma}, \boldsymbol{X})$ be a probabilistic global model of a network configuration based on the above assumptions.

Definition 1: Optimal Configuration: $(\boldsymbol{\sigma}^*, \boldsymbol{X}^*)$ is an optimal configuration if it maximizes the global likelihood [8], [20],

$$(\boldsymbol{\sigma}^*, \boldsymbol{X}^*) = \arg\max_{(\boldsymbol{\sigma}, \boldsymbol{X})} P(\boldsymbol{\sigma}, \boldsymbol{X}). \tag{1}$$

Let $P^l(\boldsymbol{\sigma}, \boldsymbol{X})$ be a local model that approximates the global model $\mathrm{P}(\boldsymbol{\sigma}, \boldsymbol{X})$.

Definition 2: Approximated Configuration: $(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{X}})$ is an approximated configuration if it maximizes the local likelihood,

$$(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{X}}) = \arg\max_{(\boldsymbol{\sigma}, \boldsymbol{X})} P^l(\boldsymbol{\sigma}, \boldsymbol{X}). \tag{2}$$

Definition 3: Approximation Error and Near-Optimal Configuration: An approximation error is defined as the average relative difference between the log likelihood at the globally optimal and the locally optimal configuration, i.e.,

$$E[\Delta] = E\left[\left|\frac{\log P(\boldsymbol{\sigma}^*, \boldsymbol{X}^*) - \log P(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{X}})}{\log P(\boldsymbol{\sigma}^*, \boldsymbol{X}^*)}\right|\right] \tag{3}$$

where the expectation is over $\hat{X}, X^*, \sigma^*$, and $\hat{\sigma}$; and the randomness results from possibly multiple solutions for $\hat{X}, X^*, \sigma^*$, and $\hat{\sigma}$ [13]. For a given $\epsilon > 0$, if $E[\Delta] \leq \epsilon$, the local model $P^l(\boldsymbol{\sigma}, \boldsymbol{X})$ and the corresponding realization $(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{X}})$ are near-optimal.

If $P^l(\boldsymbol{\sigma}, \boldsymbol{X})$ is factorizable, i.e., $P^l(\boldsymbol{\sigma}, \boldsymbol{X}) = \prod_{ij} g_{ij}(\boldsymbol{\sigma}, \boldsymbol{X})$, where $g_{ij}(\boldsymbol{\sigma}, \boldsymbol{X})$ is a localized probability density function that depends on decision variables in neighboring nodes i, j and link (i,j), for $1 \leq i, j \leq N$, the maximization from (2) reduces to a set of coupled local maximizations, e.g., $(\hat{\sigma}_{ij}, \hat{X}_i) = \arg\max_{(\sigma_{ij}, X_i)} g_{ij}(\boldsymbol{\sigma}, \boldsymbol{X})$ for $1 \leq i, j \leq N$. Note that when variables exhibit spatial Markovian dependence, distributed configuration management is naturally optimal [8]. But when variables are non-Markovian, the near-optimality becomes important. $E[\Delta]$ provides a performance measure for distributed configuration management and needs to be less than a given error bound.

TABLE I
NOTATIONS

σ_{ij}	Activity of link (i,j), $\sigma_{ij} \in \{-1, 1\}$, $\eta_{ij} = \frac{\sigma_{ij}+1}{2}$
SINR_{ij}	$\frac{P_i l_{ij}^{-\alpha} \eta_{ij}}{N_b + \sum_{(m,n)\neq(i,j)} P_m l_{mj}^{-\alpha} \eta_{mn}}$
r_f	Interference range from a receiver
N_{ij}^I	Set of neighboring links of link (i,j)
X_i	Position of node i
$E(\Delta)$	Approximation error of the local model
ϵ_Δ	An upper-bound of approximation error $E(\Delta)$

TABLE II
CORRESPONDENCE BETWEEN WIRELESS (DIPOLE) NETWORK AND PARTICLE SYSTEMS

Wireless (Dipole) Network	Particle Systems (Lattice Gas [18])
active (+1) / inactive (-1)	occupied (+1) / empty (-1)
interference	interaction energy
system potential energy	chemical potential
logical configuration	system state (e.g., liquid or gas)

The rest of the paper describes how to a) obtain a global model $\mathrm{P}(\boldsymbol{\sigma}, \boldsymbol{X})$ from the given assumptions; b) characterize the spatial dependence of decision variables $(\boldsymbol{\sigma}, \boldsymbol{X})$ using a graphical representation of $\mathrm{P}(\boldsymbol{\sigma}, \boldsymbol{X})$, and obtain a simplified graph and a mathematical representation for $P^l(\boldsymbol{\sigma}, \boldsymbol{X})$; c) obtain sufficient conditions for $P^l(\boldsymbol{\sigma}, \boldsymbol{X})$ to result in a near-optimal configuration; d) derive a distribution algorithm, and apply the algorithm to self-configuration of a wireless network. Table I summarizes key notations used in the paper.

III. GLOBAL MODEL

We begin by developing a global model that characterizes the probabilistic spatial dependence in a wireless network. Our approach is bottom-up so that the probabilistic model can be obtained from the ground truth based on the given assumptions and management constraints.

A. Logical Configuration

We first develop a probabilistic model for logical configuration $\boldsymbol{\sigma}$ given a set of node positions $\boldsymbol{X}$. A logical configuration is considered as random since link activities are triggered by random user demands (see [13] for an example). The basic principle for model development is drawn from particle systems in statistical physics [18]. There, a particle corresponds to a dipole with binary states, active or inactive; and a particle system consists of statistically distinguishable dipoles [18]. Using an analogy between particle systems and wireless communication networks, we regard link activity σ_{ij} as a communication dipole, and logical configuration $\boldsymbol{\sigma}$ as a set of communication dipoles. Table II summarizes similarities between particle systems and wireless infrastructureless networks.

Configuration Hamiltonian of a particle system describes the state of a set of particles under the following conditions: a) active particles are statistically distinguishable, and b) interactions

between particles are weak. Similar conditions hold for the wireless networks, where active communication dipoles are statistically distinguishable; and interactions among communication dipoles are weak due to decaying interference. Hence, we extend the notion of configuration Hamiltonian to the wireless networks. We define a system energy of a logical configuration as the summation of the received power at individual receivers in the network,

$$\sum_{ij} P_j \eta_{ij} \tag{4}$$

where $\eta_{ij} = ((\sigma_{ij}+1)/2)$; for a link (i,j), node i is a transmitter and node j is a receiver. A dipole is inactive, i.e., $\sigma_{ij} = -1$ and $P_j = 0$, if node i does not transmit to node j. For an active link (i,j), P_j denotes the net received-power at the receiver j, which includes the signal and the interference.

Following the definition in statistical physics [11], the configuration Hamiltonian of a dipole system is the negative system energy [1], [8], [12], [20], where the negative sign ensures that a more likely system configuration has a larger probability,

$$H(\boldsymbol{\sigma}|\boldsymbol{X}) = -\sum_{ij} P_j \eta_{ij} + \beta \sum_{ij} (\mathrm{SINR}_{ij} - \mathrm{SINR}_{\mathrm{th}})^2 \eta_{ij} \tag{5}$$

where SINR_{ij} is the SINR of dipole σ_{ij}, $\mathrm{SINR}_{\mathrm{th}}$ is a given SINR threshold, and $\beta > 0$ is a weighting factor. $\beta(\mathrm{SINR}_{ij} - \mathrm{SINR}_{\mathrm{th}})^2$ serves as a penalty term for the SINR constraint, and an equivalence of which is used for simplicity,[1] i.e., $\beta(P_i l_{ij}^{-\alpha}\eta_{ij} - \mathrm{SINR}_{\mathrm{th}}(\sum_{mn\neq ij} P_m l_{mj}^{-\alpha}\eta_{mn} + N_b))^2$.

For an active dipole $\sigma_{ij} = 1$, interfering nodes within a certain neighborhood of the receiver j are considered as significant interferers; and this neighborhood is denoted by N_{ij}^I as the interference range of node j. That is, if the interference range of node j is r_f, $N_{ij}^I = \{(m,n); |X_m - X_j| \le r_f, m \neq i, m \neq j, n \neq i, n \neq j\}$. Note that interference range is usually referred to a transmitter. Here we use the notion of interference range at a receiver in order to characterize the performance of the receiver due to the accumulated residual interference. The interference range r_f is a distance that the cooperating nodes can accommodate the complexity of information exchange and of distributed decisions. For simplicity, we assume the interference range of all receivers is the same value. The Hamiltonian can be rewritten as

$$H(\boldsymbol{\sigma}|\boldsymbol{X}) = R_1(\boldsymbol{\sigma},\boldsymbol{X}) + R_2(\boldsymbol{\sigma},\boldsymbol{X}) + R_3(\boldsymbol{\sigma},\boldsymbol{X}) + R_I(\boldsymbol{\sigma},\boldsymbol{X}) \tag{6}$$

where $R_1(\boldsymbol{\sigma},\boldsymbol{X}) = \sum_{ij}\alpha_{ij}\eta_{ij}$ is the first-order energy of individual dipoles, $R_2(\boldsymbol{\sigma},\boldsymbol{X}) = \sum_{ij}\sum_{mn\in N_{ij}^I}\alpha_{ij,mn}\eta_{ij}\eta_{mn}$ is the second-order energy with products of two dipoles within the interference range, $R_3(\boldsymbol{\sigma},\boldsymbol{X}) = \sum_{ij}\sum_{mn\in N_{ij}^I}\sum_{uv\in\{N_{ij}^I, N_{mn}^I\}}\alpha_{ij,mn,uv}\eta_{ij}\eta_{mn}\eta_{uv}$ is the third-order energy with products of three dipoles within the interference range, $R_I(\boldsymbol{\sigma},\boldsymbol{X}) = \sum_{ij} R_{I_{ij}}(\boldsymbol{\sigma},\boldsymbol{X})$ is the total interference outside the interference range where $R_{I_{ij}}(\boldsymbol{\sigma},\boldsymbol{X})$ is the residual interference outside the interference range of an active dipole $\sigma_{ij} = 1$. The coefficients of the link activities $\boldsymbol{\sigma}$ are

$$\begin{aligned}
\alpha_{ij} &= -P_i l_{ij}^{-\alpha} + \beta\cdot\left(P_i l_{ij}^{-\alpha} - \mathrm{SINR}_{\mathrm{th}} N_b\right)^2,\\
\alpha_{ij,mn} &= 2\sqrt{P_i P_m} l_{ij}^{-\frac{\alpha}{2}} l_{mj}^{-\frac{\alpha}{2}} - P_m l_{mj}^{-\alpha} + \beta \mathrm{SINR}_{\mathrm{th}}^2 P_m^2 l_{mj}^{-2\alpha}\\
&\quad - 2\beta\left(P_i l_{ij}^{-\alpha} - \mathrm{SINR}_{\mathrm{th}} N_b\right)\cdot \mathrm{SINR}_{\mathrm{th}} P_m l_{mj}^{-\alpha},\\
\alpha_{ij,mn,uv} &= -2\sqrt{P_m P_u} l_{mj}^{-\frac{\alpha}{2}} l_{uj}^{-\frac{\alpha}{2}}\\
&\quad + \beta\left(\mathrm{SINR}_{\mathrm{th}}^2 P_m P_u l_{mj}^{-\alpha} l_{uj}^{-\alpha}\right).
\end{aligned} \tag{7}$$

Intuitively, α_{ij} corresponds to the increased power when dipole σ_{ij} becomes active, $\alpha_{ij,mn}$ relates to the interference experienced by σ_{ij} resulting from a neighboring active dipole σ_{mn}, and $\alpha_{ij,mn,uv}$ relates to the interference experienced by σ_{ij} from two different dipoles σ_{mn} and σ_{uv}.

B. Physical Configuration

Node position $\boldsymbol{X}$ is assumed to be random where the randomness results from random movements and GPS measurement error. For example, desired positions can be pre-determined by a management constraint to form a regular topology. However, the actual node positions may deviate from their desired positions, resulting in an irregular topology.

Management constraints can be imposed on physical connectivity. 1-connectivity is an example that there exists at least one connected path between any source-destination pair. Yao-graph [28] provides a sufficient condition of 1-connected physical topology, where a Yao-graph constraint of any two adjacent nodes i and j can be represented as

$$h(X_i, X_j) = \begin{cases} 0, & \left|\frac{l_{ij}-l}{l}\right| \le \epsilon_0 \\ |l_{ij} - L_{ij}|, & \text{otherwise} \end{cases} \tag{8}$$

where $j \in N_i^\theta$, and N_i^θ is a set of nearest neighbors of node i for every θ radian; ϵ_0 is a small positive constant; l is a positive constant and can be $l = L_{ij}$ for simplicity; $L_{ij} = |X_{i0} - X_{j0}|$ is a desired distance, and X_{i0} is a desired position of node i. The resulting Hamiltonian for the physical topology is

$$H(\boldsymbol{X}) = \sum_i \frac{(X_i - X_{i0})^2}{2\sigma^2} + \zeta \sum_i \sum_{j\in N_i^\theta} h(X_i, X_j) \tag{9}$$

where σ is the variance that is assumed to be the same for all nodes for simplicity, and ζ is a positive weighting constant.

C. Network Configuration

We now consider a network configuration which consists of both a physical and a logical configuration. Network configuration Hamiltonian combines the Hamiltonians from the physical and the logical configuration,

$$H(\boldsymbol{\sigma},\boldsymbol{X}) = \alpha_m H(\boldsymbol{\sigma}|\boldsymbol{X}) + \alpha_n H(\boldsymbol{X}) \tag{10}$$

where α_m and α_n are weighting factors between 0 and 1, and for simplicity, we assume $\alpha_m = \alpha_n = 1$.

[1]The quadratic function is used for ease of derivation. Such a function approximates the hard constraint, i.e., $\beta U(\mathrm{SINR}_{\mathrm{th}} - \mathrm{SINR}_{ij})$ where $U(x) = 1$ for $x > 0$; 0, otherwise.

Gibbs distribution relates a configuration Hamiltonian with a probabilistic model [11], [18], and applies to network configuration as follows.

Model of Logical Configuration: For a logical configuration $\boldsymbol{\sigma}$ given node positions $\boldsymbol{X}$, a Gibbs distribution $P(\boldsymbol{\sigma}|\boldsymbol{X})$ is obtained using configuration Hamiltonian $H(\boldsymbol{\sigma}|\boldsymbol{X})$ [1], [8], [20],

$$P(\boldsymbol{\sigma}|\boldsymbol{X}) = Z_{\sigma}^{-1} \cdot \exp\left(\frac{-H(\boldsymbol{\sigma}|\boldsymbol{X})}{T}\right) \tag{11}$$

where $Z_{\sigma} = \sum_{\boldsymbol{\sigma}} \exp((-H(\boldsymbol{\sigma}|\boldsymbol{X}))/T)$ is a normalizing constant and also called the partition function [8]. The variable $T > 0$, called as the temperature [8], is used as a computational variable to obtain the most probable configuration (see Section VI in [8] for details).

Model of Physical Configuration: Similarly, the Gibbs distribution of node positions is obtained as

$$P(\boldsymbol{X}) = Z_{X}^{-1} \cdot \exp\left(\frac{-H(\boldsymbol{X})}{T}\right) \tag{12}$$

where $Z_X = \sum_{\boldsymbol{X}} \exp(-H(\boldsymbol{X})/T)$ is a normalizing constant.

Two-Layer Network Model: The Gibbs distribution of an entire network configuration is obtained using the network configuration Hamiltonian,

$$P(\boldsymbol{\sigma}, \boldsymbol{X}) = Z_{0}^{-1} \cdot \exp\left(\frac{-H(\boldsymbol{\sigma}, \boldsymbol{X})}{T}\right) \tag{13}$$

where $Z_0 = \sum_{(\boldsymbol{\sigma},\boldsymbol{X})} \exp((-H(\boldsymbol{\sigma}, \boldsymbol{X}))/T)$ is a normalization constant.

Optimal Configuration: An optimal configuration maximizes the likelihood function,

$$(\boldsymbol{\sigma}^*, \boldsymbol{X}^*) = \arg\max_{(\boldsymbol{\sigma},\boldsymbol{X})} P(\boldsymbol{\sigma}, \boldsymbol{X}) = \arg\min_{(\boldsymbol{\sigma},\boldsymbol{X})} H(\boldsymbol{\sigma}, \boldsymbol{X}) \tag{14}$$

where $H(\boldsymbol{\sigma}, \boldsymbol{X}) = -\log(P(\boldsymbol{\sigma}, \boldsymbol{X}))/T - \log(Z_0)$. Note that the system energy $H(\boldsymbol{\sigma}, \boldsymbol{X})$ incorporates both the randomness and the external management requirements. When the constraints are satisfied, the penalty terms can be ignored. Therefore, an optimal configuration should satisfy the management objectives and the constraints.

IV. Local Model

We now seek a local model $P^l(\boldsymbol{\sigma}, \boldsymbol{X})$ that is a good approximation to the global model $P(\boldsymbol{\sigma}, \boldsymbol{X})$. We resort to the probabilistic graphical models.

A. Graphical Representation

Probabilistic graphical models relate a probability distribution of random variables with a spatial dependency graph [8], [14], [16]. A node in the dependency graph represents a random variable, and a link between two nodes characterizes their dependence. In particular, a set of random variables forms a Gibbs random field (GRF) if it obeys a Gibbs distribution [18]. Hammersley-Clifford theorem shows an equivalence between a probabilistic dependency graph and a Gibbs distribution [18], where a random field is a Markov Random Field *if and only if* the probability distribution of the random field follows a Gibbs distribution. This theorem shows an interesting type of probabilistic graphical model where a random variable is conditionally independent of the others given its neighbors. In particular, when the neighborhood is much smaller than the size of a network, the conditional independence implies an interesting type of spatial Markov dependence, i.e., a node depends on its far neighbors through neighbors' neighbors. Such a nested dependence can be shown explicitly through local connections among nodes in a dependency graph, and the probability distribution is thus factorizable to local probability distributions.

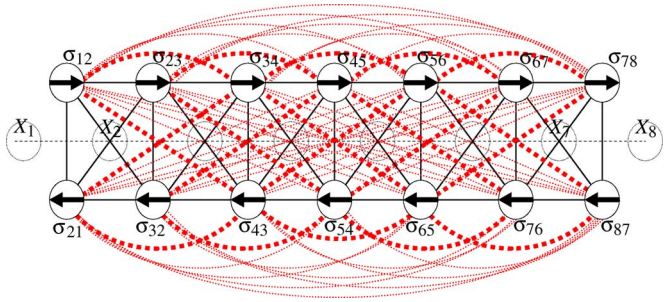

Fig. 1. Spatial dependency graph of link activities $\boldsymbol{\sigma}$ given a set of node positions $\boldsymbol{X}$.

For example, consider a one-dimensional physical topology and the corresponding dependency graph for link activities $\boldsymbol{\sigma}$ given $\boldsymbol{X}$, as shown in Fig. 1. Nodes in the graph are binary random variables $\{\sigma_{ij}\}$, and the links represent their spatial dependence. The nodes on the first and second rows are the communication dipoles of directional links. The black real lines show that any two links cannot be active at the same time. The red dark dashed links show that two nodes are within the interference range, and the red light dashed links show that two nodes are outside the interference range. All communication dipoles are fully connected and thus exhibit non-Markovian spatial dependence. The fully connected dependency graph shows an uninteresting case of a random field where the neighborhood of a node is the entire network. This implies that obtaining an optimal configuration requires each node to exchange information with all the other nodes in the network.

B. Approximation

Due to power decay, the interference outside the interference range, $R_I(\boldsymbol{\sigma}, \boldsymbol{X})$, can be relatively small compared to the first three terms of the configuration Hamiltonian in (6). The third-order term, $R_3(\boldsymbol{\sigma}, \boldsymbol{X})$, can be small also compared to the second-order term. For example, shown in (6), a third term includes $l_{ij}^{-\alpha/2} l_{mj}^{-\alpha/2} l_{uj}^{-\alpha/2}$ where $l_{mj} \gg l_{ij}$ and $l_{uj} \gg l_{ij}$. This term is usually much smaller than a second term $l_{ij}^{-\alpha/2} l_{mj}^{-\alpha/2}$. Hence, if the first two terms are used to approximate the configuration Hamiltonian, we have

$$H^l(\boldsymbol{\sigma}|\boldsymbol{X}) = \sum_{ij} \alpha_{ij}(\boldsymbol{X}) \eta_{ij} + \sum_{ij} \sum_{mn \in N_{ij}^I} \alpha_{ij,mn}(\boldsymbol{X}) \eta_{ij} \eta_{mn}, \tag{15}$$

and the corresponding Gibbs distribution is

$$P^l(\boldsymbol{\sigma}|\boldsymbol{X}) = Z_l^{-1} \cdot \exp\left(\frac{-H^l(\boldsymbol{\sigma}|\boldsymbol{X})}{T}\right) \tag{16}$$

where Z_l is a normalization constant.

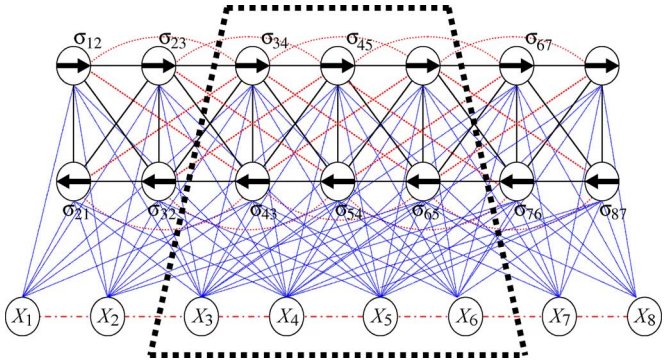

Fig. 2. Two-layer graph of $(\boldsymbol{\sigma}, \boldsymbol{X})$. Dashed-line box: clique.

As the sum in (15) only involves neighboring dipoles which are within the interference range, the resulting dependency graph of dipoles has a small neighborhood as shown in Fig. 2. In fact, this approximated Markov Random Field is the well-known second-order Ising model [18] where the Hamiltonian $H^l(\boldsymbol{\sigma}|\boldsymbol{X})$ consists of both the first- and second-order terms of $\{\sigma_{ij}\}$. Such an approximation can also be obtained directly from the probabilistic dependency graph of the global model by removing all edges outside the interference range of each node.

C. Spatial Dependence in Physical Topology

In general, spatial dependence in physical topology $\boldsymbol{X}$ is not always Markovian since a management objective can induce a long-range spatial dependence. However, for many practical management objectives, a node only needs to interact with close neighbors. Thus we consider a physical topology that exhibits the second-order Markov dependence in this work. The 1-connectivity constraint shown in (8) is such an example, where a random field of node positions $\boldsymbol{X}$ corresponds to a second-order Markov Random Field, and the interactions are only with the first-order neighbors.

D. Two-Layer Markov Random Fields

The probabilistic graph of a network configuration can be obtained by combining the graph of the logical and the physical configuration. For the above example network, a two-layer graph is shown in Fig. 2 as an approximation of the original network configuration. The upper layer graph is for the logical configuration that assumes the local interference among neighboring active dipoles. The lower-layer is for the physical configuration that shows Markov dependence of node positions due to the 1-connectivity constraint in (8). The entire graph thus exhibits spatial Markov dependence at both layers. The red dash lines between dipoles are due to interference, the blue real lines are the dependence between dipoles and node positions, and the red dash-dot lines are the spatial dependence of node positions.

This two-layer graph corresponds to a coupled MRF [8], where an Ising model and a second-order MRF are combined together. The graph is also known as a Random-Bond model [18]. A random bond connects two dipoles on the graph and decays fast with respect to the relative distance between two dipoles. The two-layer MRF $(\boldsymbol{\sigma}, \boldsymbol{X})$ can also be represented by a chain graph [20] of two MRF layers, one for $\boldsymbol{X}$ and the other for $\boldsymbol{\sigma}$. The two-layer probabilistic graph thus maps the complex spatial dependence of a multi-hop wireless network to an explicit graphical representation, which provides a solid ground for the statistical and distributed management algorithm.

The corresponding likelihood function can be represented as

$$P^l(\boldsymbol{\sigma}, \boldsymbol{X}) \propto \prod_{i,j} g_{ij}(\boldsymbol{\sigma}, \boldsymbol{X}) \tag{17}$$

where $g_{ij}(\boldsymbol{\sigma}, \boldsymbol{X})$ is a local probability density function and can be represented as a function of the sum of clique potentials [8], [20]:

$$g_{ij}(\boldsymbol{\sigma}, \boldsymbol{X}) = \exp\left(\frac{-\sum_{c \in C_{ij}} \psi_c(\boldsymbol{\sigma}, \boldsymbol{X})}{T}\right) \tag{18}$$

where C_{ij} is the set of all cliques including node i, node j, and link (i, j); and $\psi_c(\boldsymbol{\sigma}, \boldsymbol{X})$ is a clique potential function of a clique c. As an example, consider a clique for nodes 3 and 4 as well as link (3,4) shown in Fig. 2. The corresponding potential is a collection of related clique functions, i.e., $\sum_{c \in C_{34}} \psi_c(\boldsymbol{\sigma}, \boldsymbol{X}) = \alpha_{34} + \sum_{mn \in \sigma_{N_{34}^I}} (\alpha_{34,mn} + \alpha_{mn,34})((\sigma_{mn} + 1)/2) + (((X_3 - X_3(0))^2/2\sigma^2) + \zeta \sum_{j \in \{2,4\}} h(X_3, X_j)) + (((X_4 - X_4(0))^2/2\sigma^2) + \zeta \sum_{j \in \{3,5\}} h(X_4, X_j))$, where α values from (7), and $\sigma_{N_{34}^I} = \{\sigma_{12}, \sigma_{21}, \sigma_{23}, \sigma_{32}, \sigma_{43}, \sigma_{45}, \sigma_{54}, \sigma_{56}, \sigma_{65}\}$ corresponds to a set of neighboring dipoles of σ_{34} within the interference range. In general, a clique is determined by the chosen interference range.

V. Analysis: Near-Optimality and Complexity

We now derive sufficient conditions for a local model to be a near optimal approximation of the global model.

A. Communication Complexity of Cooperating Nodes

The neighborhood size in a Markov Random Field corresponds to the range of information-exchange of a node with its neighbors and thus characterizes communication complexity of cooperating neighbors. Specifically, the communication complexity[2] of an active dipole can be regarded as the maximum number of active dipoles within its interference range. As such a maximum number is random and varies with respect to dipoles, we use a deterministic bound for the number of active dipoles within the interference range.

Assume that an active dipole σ_{ij} satisfies the SINR requirement. Consider a circle centered at the receiver j of the active dipole σ_{ij} within which there cannot exist any active dipoles for the SINR requirement to hold. The radius of the circle is the contention range for node j. Let r_c be the minimum contention range for all active dipoles. Consider the interference range outside the contention range where multiple dipoles can be active concurrently, resulting in interference. We bound the interference region using a circular region of radius r_f although the actual interference range of a receiver may not be symmetrical in all directions. The region that includes active dipoles outside r_c but within r_f as shown in Fig. 3 is considered as the relevant interference neighborhood of a node. By packing the circular region with small circles of radius r_c, we can obtain an upper

[2]Here, the number of neighboring nodes is expected to be proportional to the complexity and thus signifies the information bits transmitted among cooperating nodes.

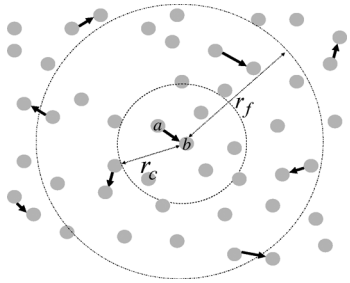

Fig. 3. Contention range r_c and interference range r_f of an active dipole.

bound of the maximum number of active dipoles in the interference neighborhood.

Definition 3: Communication Complexity $\mathcal{C}$: Communication complexity of a dipole σ_{ij} is defined as an upper bound of the maximum number of active dipoles within the interference range, i.e., $\mathcal{C} = (r_f/r_c)^2$, for $1 \le i, j \le N$, $i \ne j$.

B. Near-Optimality Conditions

We now derive sufficient conditions for a local model to be a good approximation of the global model. We consider homogeneous networks for the convenience of analysis.

Theorem 1: Consider a network where N nodes are uniformly distributed and satisfy the 1-connectivity constraint in (8). The distance between any node and its neighbors for every θ radian is between $l(1-\epsilon_0)$ and $l(1+\epsilon_0)$, l is the distance between two nodes if there were no random perturbation, and $0 < l(1+\epsilon_0) < r_c$. Assume that N nodes transmit at the same power level ($P_t > 0$), and that they have the same desired SINR threshold $\mathrm{SINR}_{\mathrm{th}}$ and the same circular interference range r_f. Let $\alpha \ge 2$ be the power attenuation factor of the channel. Let $f = (l/r_c)$ that characterizes the density of active dipoles on the regular topology. The average approximation error can be bounded as $E[\Delta] \le \epsilon_\Delta$, where

$$\epsilon_\Delta = \begin{cases} \mathcal{I}\left(\frac{(2+\ln \mathcal{C})^2}{4\pi} f^2 + f(1+\sqrt{N/f})\right), & \alpha=2 \\ \mathcal{I}\left(\frac{\left(2-\frac{1}{\mathcal{C}}\right)^2}{2\pi l^2} f^4 \right. \\ \quad \left. + \frac{f^2}{l^2}\left(\frac{1}{\sqrt{\mathcal{C}}} + \ln\left(1+\sqrt{\frac{Nf}{\mathcal{C}}}\right)\right)\right), & \alpha=4 \\ \mathcal{I}\left(\frac{\left(\alpha-2\mathcal{C}^{\frac{2-\alpha}{2}}\right)^2 f^\alpha}{(\alpha-2)^2 2\pi l^{\alpha-2}} + \frac{f^{\frac{\alpha}{2}}}{l^{\alpha-2}} \right. \\ \quad \times \left[\mathcal{C}^{\frac{2-\alpha}{4}} + \frac{2}{4-\alpha} \right. \\ \quad \left. \left. \times \left(-\mathcal{C}^{\frac{4-\alpha}{4}} + (\sqrt{\mathcal{C}} + \sqrt{Nf})^{\frac{4-\alpha}{2}}\right)\right]\right), & \text{else} \end{cases}$$

with $\mathcal{I} = 8\pi/(l(1 - 1/\mathrm{SINR}_{\mathrm{th}} - N_b l^\alpha/P_t))$.

The proof is given in Appendix I. The bound $\mathrm{E}[\Delta]$ shows that a local model results in two components of the approximation error. The first is due to using a model of a lower order of $\{\sigma_{ij}\}$ within the interference range, and bounded by the first terms in the above expressions. For a given communication complexity $\mathcal{C}$, the first type of error can be made small by reducing the density f of active dipoles. This results in sparsely spaced (i.e., a large r_c) active dipoles and thus reduces the spatial dependencies. The second type of error is due to neglecting the aggregated interference outside the interference range, and upper bounded by the remaining terms in the above expressions. The second type of error can be made small if the aggregated interference outside the interference range is reduced. Consider a slow decaying channel $\alpha = 2$, the second type of error is independent of $\mathcal{C}$, and thus can only be reduced by sufficiently sparse active dipoles. In fact the fraction of active dipoles would have to decrease with respect to $O(N)$ for the error bound to be $O(1)$. For a fast decaying channel $\alpha > 4$, the error can be bounded even when complexity $\mathcal{C} = O(1)$ is a constant. Hence whether and how fast $\mathcal{C}$ grows with N depends on the power decay of the channel, since the aggregated interference outside the interference range depends on α.

Channel, Contention and Network Size: Using the upper bound of the approximation error ϵ_Δ, a sufficient condition can be obtained on the density f of active dipoles for a large network so that the local model be near-optimal.

Corollary 1: Let ϵ be a desired performance-bound. Assume $N \gg 1$, and $\mathcal{C}$ does not grow with respect to N. We have $\epsilon_\Delta \le \epsilon$ if $0 < f$ and

$$f \le \begin{cases} O\left(N^{\frac{\alpha-4}{4+\alpha}}\right), & 2 < \alpha < 4 \\ O\left(\frac{1}{\sqrt{\ln N}}\right), & \alpha = 4 \\ O(1), & \alpha > 4 \end{cases}$$

where $O()$ represents the order for large N.

The proof can be obtained through simple algebraic manipulations from Theorem 1, and is thus omitted. Consider a large interference channel where $2 < \alpha < 4$, then $Nf = O(N^{2\alpha/(4+\alpha)})$, showing a sparsely activated network. This implies that to satisfy the same error bound, if the network is densely activated, the complexity may have to increase with N. Hence if a Markov Random Field with a small neighborhood may not be an efficient approximation to the global model.

Now consider a small interference channel where $4 < \alpha \le 6$, then $Nf = O(N)$. This implies that the network has densely activated dipoles; that is, the number of active nodes can be a fraction of N. The corresponding Markov Random Field with a small neighborhood is an efficient approximation to the global model.

Performance-Complexity Tradeoff: How does $\mathcal{C}$ vary with respect to α for large N? Shown by Theorem 1, the larger the communication complexity $\mathcal{C}$, the larger the first type of error, since more dependency is neglected by the local model within the inference range. As shown by Corollary 1, the first type of error can be made small for choosing a sufficiently small f. Hence, we discuss the role of $\mathcal{C}$ focusing on the second type of error. The second type of error is expected to be smaller for a larger $\mathcal{C}$ since the dependence is considered by including more active dipoles. This can be quantified from the second term in Theorem 1.

Corollary 2: Let ϵ be a desired performance-bound. Assume $N \gg 1$, and f is given to be $O(1)$. We have the second terms in $\epsilon_\Delta \le \epsilon$ if

$$C > \begin{cases} \Omega(Nf), & \alpha = 4 \\ \Omega(1), & \alpha > 4 \end{cases}$$

where $\Omega()$ represents the order for large N used in a lower bound.

The proof can be obtained through simple algebraic manipulations, and thus omitted. For $2 \le \alpha \le 4$, the bound for the

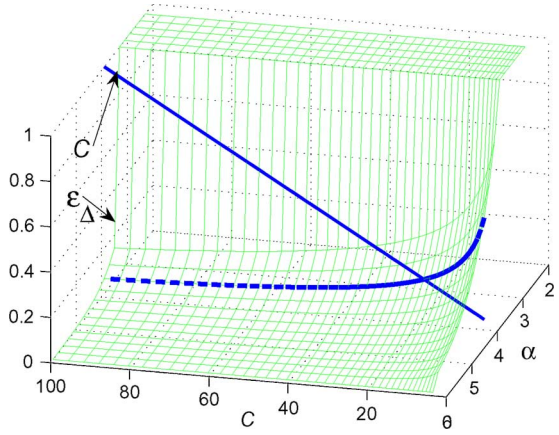

Fig. 4. Tradeoff between performance and complexity. Dashed line: approximation error for $\alpha = 4$.

second type of error is nearly independent of $\mathcal{C}$ for large network size N. This suggests that the accumulated interference outside interference range is not negligible but dominates for a slow decaying channel, and a Markov Random Field may not efficiently approximate the global model. For $\alpha > 4$, the spatial dependence within rather than outside an interference range dominates for a rapidly decaying channel. Therefore, for $\alpha > 4$, Corollary 2 shows that the second error term can be made arbitrarily small if $\mathcal{C}$ grows at a slow rate with respect to N. Markov Random Fields are thus an efficient approximation of the global model.

Fig. 4 shows an example of the tradeoff between the approximation error and the communication complexity for $\alpha = 4$. Other parameters are $\mathrm{SINR}_{\mathrm{th}} = 20$, $N_b = 0.1$, $r_c = 10$, $r_f \in \{20 \sim 100\}$, $N = 1000$, and $l = 2$ m.[3] The intersection $\mathcal{C}$ and ϵ_Δ corresponds to an optimal neighborhood $\mathcal{C} = 16$. This corresponds to $r_f = 40$ and $r_c = 10$. In general, for a given r_f, r_c can be adjusted to vary $\mathcal{C}$ so that a proper tradeoff can be obtained.

VI. Distributed Algorithm

The distributed algorithm obtains a near-optimal configuration by maximizing the approximated likelihood function $P^l(\boldsymbol{\sigma}, \boldsymbol{X})$ as in (2). Due to the spatial Markovian property, $P^l(\boldsymbol{\sigma}, \boldsymbol{X})$ is factorizable over cliques. Hence maximizing the likelihood function reduces to maximizing the local probability density functions (i.e., clique potentials),

$$(\hat{\sigma_{ij}}, \hat{X}_i) = \arg \max_{(\sigma_{ij}, X_i)} P^l\left(\sigma_{ij}, X_i | X_{N_i}, \sigma_{N_{ij}}\right) \qquad (19)$$

where $P^l(\sigma_{ij}, X_i | X_{N_i}, \sigma_{N_{ij}}) = g_{ij}(\boldsymbol{\sigma}, \boldsymbol{X})$, $X_{N_i} = \{X_j, j \in N_i\}$ and $\sigma_{N_{ij}} = \{\sigma_{mn}, (m,n) \in N_{ij}^I\}$ are in the neighborhood of node i and dipole σ_{ij}, respectively. These local probability density functions are composed of the neighboring nodes and dipoles, and the configuration can thus be updated locally.

This work uses stochastic relaxation which is a randomized algorithm and converges to a maximum asymptotically with probability one [8]. The main steps of the algorithm are summarized below. Let $\hat{X}_i(t+1)$ and $\hat{\sigma}_{ij}(t+1)$ denote the new position that node i and the activity of link (i,j) at time $t+1$,

[3]This corresponds to sensor networks for habitat monitoring, battlefield surveillance, and mechanical measurement and monitoring.

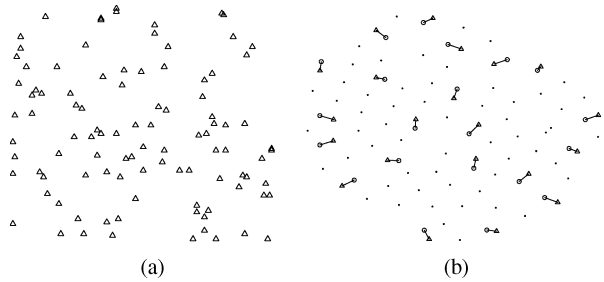

Fig. 5. Self-configuration with localized algorithm. (a) $\boldsymbol{X}_0$; (b) $(\boldsymbol{\sigma}, \boldsymbol{X})$ given $\boldsymbol{X}_0$

respectively. Node i determines these values locally as follows: for $t \geq 1$ and $1 \leq i \leq N$,

i) Given X_{N_i}, $\hat{X}_i(t+1) = x_i$, with probability $P^l(X_i(t+1) = x_i | X_{N_i}(t)) = (\exp(-\psi_i(x_i)/T(t+1)) / \sum_{\forall x_i} \exp(-\psi_i(x_i)/T(t+1))$, where $\psi_i(x_i) = (((x_i - X_{i0})^2)/2\sigma^2) + \zeta \sum_{j \in N_i^\theta} h(x_i, X_j(t))$.

ii) Given link activities $\sigma_{N_{ij}}$ and positions X_{N_i}, determine whether node i transmits to node j for $\forall j \in N_i$, i.e., $\hat{\sigma}_{ij}(t+1) = \sigma_{ij}$ for $\sigma_{ij} \in \{-1, 1\}$ with probability $P^l(\sigma_{ij}(t+1) = \sigma_{ij} | X_{N_i}(t), \sigma_{N_{ij}}(t)) = (\exp(-\psi_{ij}(\sigma_{ij})/T(t+1)) / \sum_{\forall \sigma_{ij}} \exp(-\psi_{ij}(\sigma_{ij})/T(t+1)))$, where $\psi_{ij}(\sigma_{ij}) = (\alpha_{ij} + \sum_{mn \in N_{ij}^I} (\alpha_{ij,mn} + \alpha_{mn,ij})((\sigma_{mn}(t)+1)/2))((\sigma_{ij}+1)/2)$. Temperature $T(t)$ is a cooling constant in the algorithm [8], where $T(t) = T_0 / \log(1+t)$ varies with time t and $T_0 = 3$ [1], [18]. The result of the distributed algorithm is the network configuration $(\hat{\mathbf{X}}, \hat{\sigma})$.

Information Exchange: At time t, node i broadcasts its position and adjacent links activities, $(X_i(t), \sigma_{ij}(t) = 1)$, to the neighboring nodes. At time $t+1$, node i makes decisions using the information received from neighbors, i.e., $(X_m(t), \sigma_{mn}(t) = 1)$ for $m \in N_i$. Such information exchange bears a similar spirit to that of Bellman-Ford routing algorithm.

VII. Self-Configuration: Example and Validation

The distributed algorithm is now applied to the examples of self-configuration of wireless networks.

A. Example of Self-Configuration

We simulate the distributed algorithm for self-configuration using the following parameters: Network size $N = 100$, the threshold of the inter-node distance $l = 2$ and $\theta = (\pi/2)$ for the 1-connectivity of the physical topology, $\alpha = 4$ for the channel, and $\mathrm{SINR}_{\mathrm{th}} = 20$. Fig. 5(a) shows an example of a random initial configuration $(\boldsymbol{\sigma_0} = \mathbf{0}, \boldsymbol{X_0})$. Each node then updates jointly its position and link activities based on the distributed algorithm in Section VI. The iteration stops when a steady state is reached. Fig. 5(b) shows a resulting network configuration.

Now, consider that certain nodes fail in a network configuration. Upon failures, the closest neighbors of failed nodes first adjust their positions and select other nodes for transmission. This may cause adjustments to the neighbors and the neighbors' neighbors, resulting in cascading changes across the entire network. Thus localizing failure events is important. Additional penalty terms can be introduced to the Hamiltonian as reconfiguration costs to penalize cascading changes: $\xi \cdot |(\boldsymbol{\sigma}, \boldsymbol{X}) - (\boldsymbol{\sigma}_s, \boldsymbol{X}_s)|$, where $(\boldsymbol{\sigma}_s, \boldsymbol{X}_s)$ denotes the steady-state

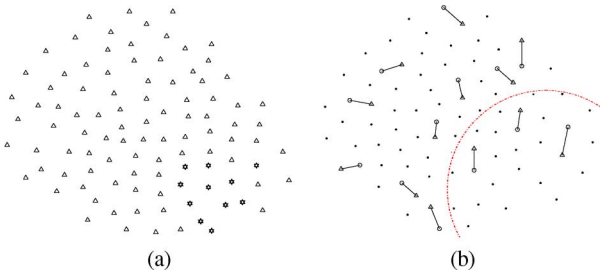

Fig. 6. Localized recovery from random node failures. (a) Random failure of nodes; (b) localized recovery of $\boldsymbol{\sigma}$.

network configuration, and ξ is a positive weighting constant that characterizes the cost of change in node positions and/or link activities. Such a constraint would localize the change. Fig. 6(a) shows failures of wireless nodes that are marked as stars. Localized recovery of the configuration is shown in Fig. 6(b). The nodes outside the arc are not affected by failures. The resulting configuration, however, is no longer globally optimal, as a tradeoff with the reconfiguration cost.

B. Model Validation

We now validate the performance of the local model, the communication complexity and the tradeoffs.

Performance and Comparison: We first examine through simulation the performance of the local model compared with that of the global model and commonly-used protocol model in the prior works [7], [9], [21].

A wireless network is generated with 50 randomly-positioned nodes, $\alpha = 4$, and $\mathrm{SINR}_{\mathrm{th}} = 20$. The results of ten simulations conducted over random initial topologies and link activities are averaged. The interference range is the entire network for the global model, and $r_f = 4$ for the local model. The protocol model assumes a contention range to be r_s without using an interference range. To avoid any confusion with r_c, we call the contention range of the protocol model as the separation range and denote with r_s.

Fig. 7 shows the one-hop capacity achieved by the global, local and protocol model. For a large r_s, the protocol model fails to maximize the spatial channel-reuse because the contention constraint is too stringent. For a small r_s, the protocol model over-utilizes the channel resource, resulting in a violation of SINR constraint and a reduced spatial-reuse. Even at the optimal separation range $(r_s = 3)$, the spatial-reuse of the protocol model is less than that of the local model by 23%. The inefficient spatial-reuse is because the protocol model does not consider interference range. Compared to the global model, the local model has 8% less spatial reuse on average, showing performance degradation due to modeling error.

Now consider the tightness of the bound on the approximation error. Specifically, the approximation error is measured from the simulation and compared with an upper bound in Theorem 1. A linear topology is selected, where 100 nodes are randomly placed and a node communicates with two neighbors. Other parameters used in the simulations are $\alpha = 4$, $\mathrm{SINR}_{\mathrm{th}} = 10$, $N = 100$, $r_c = 10$, and a varying r_f to obtain various values for $\mathcal{C}$. As shown in Fig. 8, the measured approximation error decreases sharply as $\mathcal{C}$ increases and is indeed bounded by $E(\Delta)$. The bound follows the same trend as the actual approximation

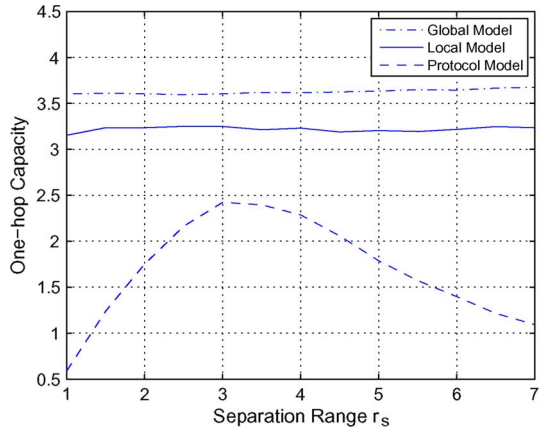

Fig. 7. Comparison of one-hop capacity: global, local, and protocol model.

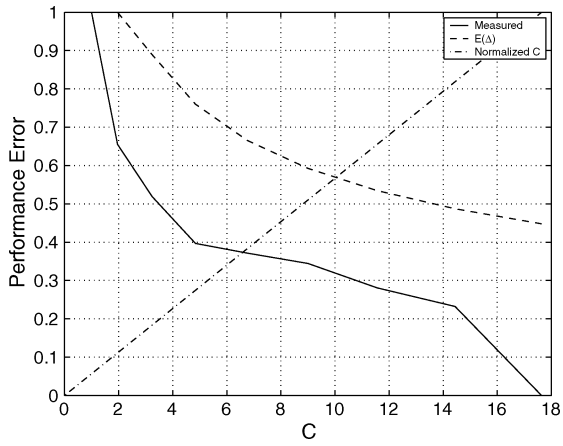

Fig. 8. Approximation error and upper bound $E(\Delta)$.

error. The difference between the measured approximation error and the bound may result from the linear topology whose density of active links is much sparser than the assumptions in Theorem 1. Intersections between $\mathcal{C}$ and $E[\Delta]$ illustrate the performance-complexity tradeoff.

Complexity: The actual communication complexity is obtained by counting the number of the active links within the interference range of each active dipole, and averaged over all active dipoles and 10 simulation runs, using the same parameters as in the above simulation. Fig. 9 shows the communication complexity for both the global and the local model as a function of network size N. The communication complexity of the centralized global optimization increases linearly with N since each node uses the information from all other nodes in the network. The complexity increases slowly with N for the distributed algorithm.

Extension: We now consider an extension to multiple time slots to show the promise of the distributed algorithm to scheduling in a more realistic setting. This is motivated by the prior work [24] that considers packet-level scheduling, queue size and packet arrivals, but uses a simplified interference model. Our model does not include packet arrivals and queue occupancy but uses a general interference model and adopts physical topology. To compare these two works at a common ground, we assume that a physical topology is given, there is one queue for each communication link and the queue is always busy. Our local

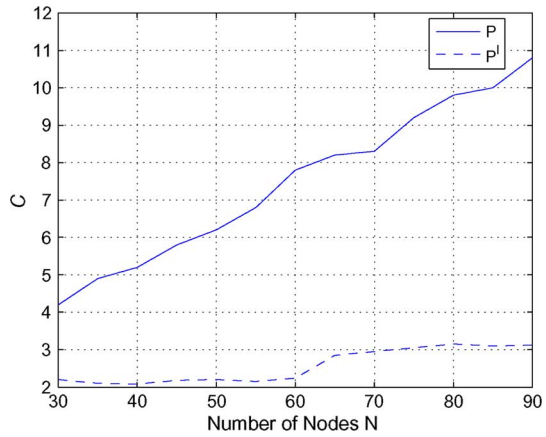

Fig. 9. Communication complexity $\mathcal{C}$ vs network size N. P and P^l: global and local model, respectively.

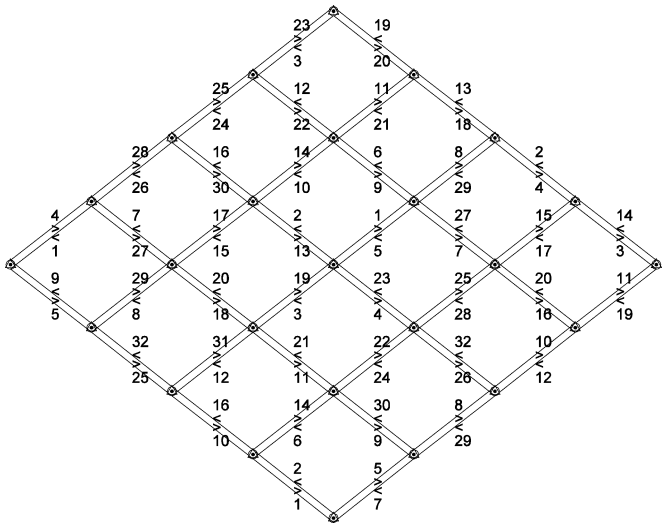

Fig. 10. Time slot allocation of spatial time division multiple access.

probabilistic model is extended directly to include multiple time slots in the coefficients, and details are given in Appendix II. We conduct a simulation for scheduling using the following parameters: $\alpha = 4, N = 25, l = 2$, and $r_f = 6$ for regular topology in Fig. 10. There are 80 communication links and 32 assigned time slots. The figure shows that the optimal allocations are obtained by the local model and the distributed algorithm while satisfying the SINR constraints. Hence, compared with [24] when a physical topology is given, an advantage of the local model is its ability to approximate the general (non-Markovian) SINR model. A disadvantage of our local model is a lack of specifications of network traffic.

VIII. Conclusion

This work has developed an analytical framework in which approximation, near-optimality, and randomized distributed algorithms have been studied for self-configuration. Our findings are summarized as follows.

a) A global model of the network configuration is first developed that characterizes the ground truth in regard to network assumptions and management constraints. The resulting model is a Gibbs distribution where the exponent can be regarded as a cost function. This relates modeling with optimization so that both randomness in a network and objectives can be combined naturally. This approach differs from the pure "emerging behavior" where nodal decisions are not governed by an optimal performance. This approach also differs from the "black-box" method where models are learned externally. A shortcoming of our approach is the simple assumptions in the current model. Fading has not been included in our model. Traffic demands at network-layer are yet to be included.

b) The mathematical representation of the global model quantifies the statistical spatial dependence of node positions and node-node communications. The complex spatial dependence is represented explicitly by a probabilistic spatial dependency graph. The graph is hierarchical which we did not expect. The hierarchy differs from the existing models and results naturally from the ground truth. The local model is a two-layer Markov Random Field or a random bond model. The complexity of the local model is the maximum size of cliques in the Markov Random Field. The local model is obtained by removing long links in the graph, and a more sophisticated approach may be explored to obtain a better approximation.

c) We have derived sufficient conditions on the near-optimality of the local model in terms of different channel conditions, density of nodes, network size and complexity. The conditions show a tradeoff between the near-optimality and the complexity. For example, a local model is near-optimal with a moderate complexity if a channel has a fast power decay of at least 4; and may not be near-optimal, otherwise. The near-optimality conditions thus complement the empirical choices in protocol designs (e.g. 802.11), and can be utilized for studying jointly the effect of channel attenuation, densities of nodes, and node-node communications.

d) Near-optimal local models render a class of randomized distributed algorithms for self-configuration. Local self-configuration collectively achieves a near-optimal global configuration at a bounded communication complexity. Hence the algorithmic advantage is the strength of the local model. We have shown examples of stochastic scheduling and reconfiguration upon failures, and compared the performance and complexity with existing protocol models. One disadvantage of stochastic relaxation is the slow convergence that has been explored in the prior work [10]. Other simple randomized algorithms have been discussed in [19]. It would be beneficial to study simpler distributed algorithms for self-configuration.

Appendix I

Proof of Theorem 1: We first bound the error, i.e., $|((H(\boldsymbol{\sigma}^*|\boldsymbol{X}) - H(\hat{\boldsymbol{\sigma}}|\boldsymbol{X}))/H(\boldsymbol{\sigma}^*|\boldsymbol{X}))| \leq (|H(\boldsymbol{\sigma}^*|\boldsymbol{X}) - H(\hat{\boldsymbol{\sigma}}|\boldsymbol{X})|^u / |H(\boldsymbol{\sigma}^*|\boldsymbol{X})|^l)$, where the super-scripts u and l denote an upper and a lower bound of the corresponding quantity.

To find an upper bound of the numerator, we have $|H(\boldsymbol{\sigma}^*|\boldsymbol{X}) - H(\hat{\boldsymbol{\sigma}}|\boldsymbol{X})| = |(H(\boldsymbol{\sigma}^*|\boldsymbol{X}) - H^l(\boldsymbol{\sigma}^*|\boldsymbol{X})) + H^l(\hat{\boldsymbol{\sigma}}|\boldsymbol{X}) - H(\hat{\boldsymbol{\sigma}}|\boldsymbol{X}) + (H^l(\boldsymbol{\sigma}^*|\boldsymbol{X}) - H^l(\hat{\boldsymbol{\sigma}}|\boldsymbol{X}))| \leq |H(\boldsymbol{\sigma}^*|\boldsymbol{X}) - H^l(\boldsymbol{\sigma}^*|\boldsymbol{X})| + |H(\hat{\boldsymbol{\sigma}}|\boldsymbol{X}) - H^l(\hat{\boldsymbol{\sigma}}|\boldsymbol{X})|$,

where $H(\boldsymbol{\sigma}^*|\boldsymbol{X}) \leq H(\hat{\boldsymbol{\sigma}}|\boldsymbol{X})$, and $H^l(\hat{\boldsymbol{\sigma}}|\boldsymbol{X}) \leq H^l(\boldsymbol{\sigma}^*|\boldsymbol{X})$ by definition.

For any configuration $(\boldsymbol{\sigma}|\boldsymbol{X}), |H(\boldsymbol{\sigma}|\boldsymbol{X}) - H^l(\boldsymbol{\sigma}|\boldsymbol{X})| \leq |R_I(\boldsymbol{\sigma}|\boldsymbol{X})| + |R_3(\boldsymbol{\sigma}|\boldsymbol{X})|$. Furthermore,

$$\begin{aligned}|R_3(\boldsymbol{\sigma}|\boldsymbol{X})| &\leq \sum_{k_1=1}^{\frac{r_f}{r_c}} \sum_{k_2=1}^{\frac{r_f}{r_c}} 2P_t k_1^{\frac{-\alpha}{2}} k_2^{\frac{-\alpha}{2}} r_c^{-\alpha} \\ &\leq 2P_t \left(1 + \left(\int_{k=1}^{\frac{r_f}{r_c}} k^{\frac{-\alpha}{2}} dk\right)\right)^2 \cdot r_c^{-\alpha} \\ &= \begin{cases} 2P_t r_c^{-\alpha}(1+0.5\ln \mathcal{C})^2, & \alpha = 2 \\ \frac{2P_t r_c^{-\alpha}}{(\alpha-2)^2}\left(\alpha - 2(\mathcal{C})^{\frac{2-\alpha}{2}}\right)^2, & \alpha > 2 \end{cases} \\ &= I_3.\end{aligned}$$

The penalty term in the configuration Hamiltonian is not included as I_3 is an upper bound. To find a bound for $R_I(\boldsymbol{\sigma}|\boldsymbol{X})$, we extend the idea of packing applied to analyzing ALOHA protocol to this work [4]. Consider an active dipole σ_{ij}, and let G_k be the set of neighboring active dipoles that are $r_f + kr_c$ apart from the receiver. The cardinality of G_k is upper bounded, i.e., $|G_k| \leq (2\pi/\theta) = \pi/\sin^{-1}((r_c/2)/(r_f + (k-1)r_c)) < 2\pi(r_f + (k-1)r_c)/r_c$. For an active dipole σ_{ij},

$$\begin{aligned}R_{I_{ij}}(\boldsymbol{\sigma}|\boldsymbol{X}) &\leq \sum_{k=1}^{k_U} 2P_t 2\pi \frac{(r_f + (k-1)r_c)^{\frac{2-\alpha}{2}}}{r_c} (l(1+\epsilon_0))^{\frac{-\alpha}{2}} \\ &= I_{k_U},\end{aligned}$$

where k_U is an integer that satisfies the inequality

$$\frac{(N-2)(1+\epsilon_0)}{r_c} \leq \sum_{k=1}^{k_U} 2\pi \frac{r_f + (k-1)r_c}{r_c}. \qquad (20)$$

The value $(N-2)l(1+\epsilon_0)/r_c$ denotes an upper bound of the maximum number of available active dipoles except σ_{ij} (with two nodes) in a network with total N nodes. k_U can be solved from the above inequality as $k_U \leq r_c + \sqrt{Nlr_c}$. Using this bound for k_U and replacing r_f by $\sqrt{\mathcal{C}}r_c$, we have

$$\begin{aligned}I_{k_U} &\leq \frac{2P_t 2\pi}{r_c (l(1+\epsilon_0))^{\frac{\alpha}{2}}} \left(r_f^{\frac{2-\alpha}{2}} + \int_1^{k_U^u} (r_f + (k-1)r_c)^{\frac{2-\alpha}{2}} dk\right) \\ &\leq \begin{cases} \frac{P_t 4\pi}{r_c^2 (l(1+\epsilon_0))^{\frac{\alpha}{2}}} \\ \quad \times \left(\frac{1}{\sqrt{\mathcal{C}}} + \ln(1 + \sqrt{Nl/(\mathcal{C}r_c)})\right), & \alpha = 4 \\ \frac{P_t 4\pi}{r_c^2 (l(1+\epsilon_0))^{\frac{\alpha}{2}}} \\ \quad \times \left(\mathcal{C}^{\frac{2-\alpha}{4}} r_c^{\frac{4-\alpha}{2}} + \frac{2}{4-\alpha} \cdot \left(-(\sqrt{\mathcal{C}}r_c)^{\frac{4-\alpha}{2}} \right.\right. \\ \quad \left.\left. + (\sqrt{\mathcal{C}}r_c + \sqrt{Nlr_c})^{\frac{4-\alpha}{2}}\right)\right), & \alpha \neq 4 \end{cases} \\ &= I_R.\end{aligned}$$

Thus, $|H(\boldsymbol{\sigma}^*|\boldsymbol{X}) - H(\hat{\boldsymbol{\sigma}}|\boldsymbol{X})| \leq |H(\hat{\boldsymbol{\sigma}}|\boldsymbol{X}) - H^l(\hat{\boldsymbol{\sigma}}|\boldsymbol{X})| + |H(\boldsymbol{\sigma}^*|\boldsymbol{X}) - H^l(\boldsymbol{\sigma}^*|\boldsymbol{X})| \leq 2(I_R + I_3)N_\sigma^*$, where N_σ^* is total number of active dipoles in $\boldsymbol{\sigma}^*$ given $\boldsymbol{X}$.

To obtain a lower bound for $|H(\boldsymbol{\sigma}^*|\boldsymbol{X})|$, we take only the net energy term in $H(\boldsymbol{\sigma}^*|\boldsymbol{X})$ as the penalty term is usually small, i.e., close to zero when the constraint is satisfied. Then $|H(\boldsymbol{\sigma}^*|\boldsymbol{X})| \geq \min\{P_t l_{ij}^{-\alpha} - \sum_{mn \neq ij} P_t l_{mj}^{-\alpha/2} l_{ij}^{-\alpha/2}\} N_\sigma^*$ for $\forall \sigma_{ij} = 1$. For $\sigma_{ij} = 1$, to satisfy a given $\mathrm{SINR}_{\mathrm{th}}$, a sufficient condition is $(P_t(\sum_{mn\neq ij} l_{mj}^{-\alpha/2})^2 + N_b)/(P_t(l_{ij}^{-\alpha/2})^2) \leq 1/\mathrm{SINR}_{\mathrm{th}}$, thus $(\sum_{mn\neq ij} l_{mj}^{-\alpha/2})^2 \leq ((l_{ij}^{-\alpha}/\mathrm{SINR}_{\mathrm{th}}) - (N_b/P_t))$. Therefore, $|H(\boldsymbol{\sigma}^*|\boldsymbol{X})| \geq (P_t l^{-\alpha}(l+\epsilon_0)^{-\alpha} - P_t l^{-\alpha/2}(l-\epsilon_0)^{-\alpha/2}\sqrt{l^{-\alpha/2}(1-\epsilon_0)^{-\alpha/2}\mathrm{SINR}_{\mathrm{th}}^{-1} - (N_b/P_t)})N_\sigma^*$.

Hence, $|((H(\boldsymbol{\sigma}^*|\boldsymbol{X}) - H(\hat{\boldsymbol{\sigma}}|\boldsymbol{X}))/H(\boldsymbol{\sigma}^*|\boldsymbol{X}))| \leq \epsilon_\Delta$, where $\epsilon_\Delta = 2(I_R + I_3)/I_D$, $I_D = P_t l^{-\alpha}(l+\epsilon_0)^{-\alpha} - P_t l^{-\alpha/2}(l-\epsilon_0)^{-\alpha/2}\sqrt{l^{-\alpha/2}(1-\epsilon_0)^{-\alpha/2}/\mathrm{SINR}_{\mathrm{th}} - (N_b/P_t)}$. Assume $\epsilon_0 = 0$ for a simple representation,

$$\epsilon_\Delta = \begin{cases} \mathcal{I}\left(\frac{(2+\ln\mathcal{C})^2}{4\pi} f^2 + f(1+\sqrt{N/f})\right), & \alpha = 2 \\ \mathcal{I}\left(\frac{\left(2-\frac{1}{\mathcal{C}}\right)^2}{2\pi l^2} f^4 \right. \\ \quad \left. + \frac{f^2}{l^2}\left(\frac{1}{\sqrt{\mathcal{C}}} + \ln\left(1 + \sqrt{\frac{Nf}{\mathcal{C}}}\right)\right)\right), & \alpha = 4 \\ \mathcal{I}\left(\frac{\left(\alpha - 2\mathcal{C}^{\frac{2-\alpha}{2}}\right)^2 f^\alpha}{(\alpha-2)^2 2\pi l^{\alpha-2}} + \frac{f^{\frac{\alpha}{2}}}{l^{\alpha-2}}\left[\mathcal{C}^{\frac{2-\alpha}{4}} + \frac{2}{4-\alpha}\right.\right. \\ \quad \left.\left. \times \left(-\mathcal{C}^{\frac{4-\alpha}{4}} + (\sqrt{\mathcal{C}} + \sqrt{Nf})^{\frac{4-\alpha}{2}}\right)\right]\right), & \text{else} \end{cases}$$

where $\mathcal{I} = 8\pi/(l(1 - 1/\mathrm{SINR}_{\mathrm{th}} - N_b l^\alpha/P_t))$, and $f = (l/r_c)$. Moreover, since the bounds derived above hold for any (σ, X), they also hold for the expected value, i.e., $E[|((H(\boldsymbol{\sigma}^*, \boldsymbol{X}^*) - H(\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{X}}))/H(\boldsymbol{\sigma}^*, \boldsymbol{X}^*))|] \leq \epsilon_\Delta$.

APPENDIX II

Model Extension to Multiple Time Slots: We assign a time slot S_{ij} to a dipole (i,j), for $1 \leq S_{ij} \leq S_{\max}$, where $S_{\max}$ is the TDMA cycle length. Since $S_{\max}$ is unknown, we then generalize the coefficients in the Hamiltonian in (6) as functions of time slots, $\alpha'_{ij} = \alpha_{ij} \cdot \omega_s(S_{ij})$, $\alpha'_{ij,mn} = \alpha_{ij,mn}\delta(S_{ij}, S_{mn}) \cdot \omega_s(S_{ij})$, $\alpha'_{ij,mn,uv} = \alpha_{ij,mn,uv}\delta(S_{ij}, S_{mn})\delta(S_{ij}, S_{uv}) \cdot \omega_s(S_{ij})$, where $\delta(S_{ij}, S_{mn}) = 1$ if $S_{ij} = S_{mn}$; $\delta(S_{ij}, S_{mn}) = 0$, otherwise. $\omega_s(S_{ij}) = 1/S_{ij}$ in this work, which makes the first time slot gets the highest weight and utilized first and so on. As a result, the length of TDMA cycle $S_{\max}$ is minimized.

The resulting MRF model is now defined over multiple orthogonal time slots, and a communication dipole σ_{ij} now takes an extended form of $\sigma_{ij} \cdot \delta(\boldsymbol{S}_{ij} - s)$, where $\delta(\boldsymbol{S}_{ij} - s) = 1$ if $\boldsymbol{S}_{ij} = s$; 0, otherwise, for $1 \leq s \leq S_{\max}$.

ACKNOWLEDGMENT

The authors would like to thank anonymous reviewers for helpful comments. C. Ji would like to thank C. S. Ji and J. Modestino for helpful discussions on Markov Random Fields.

REFERENCES

[1] J. Baras and X. Tan, "Control of autonomous swarms using Gibbs sampling," in *Proc. IEEE Conf. Decision and Control (CDC)*, Atlantis, Bahamas, Dec. 2004, vol. 5, pp. 4752–4757.

[2] Z. Butler, K. Kotay, D. Rus, and K. Tomita, "Generic decentralized control for a class of self-reconfigurable robots," in *Proc. Int. Conf. Robotics and Automation (ICRA)*, Washington, DC, May 2002, pp. 809–815.

[3] L. E. Doyle, A. C. Kokaram, S. J. Doyle, and T. K. Forde, "Ad hoc networking, Markov Random Fields, decision-making frameworks," *IEEE Signal Process. Mag., Special Issue on Signal Processing for Wireless Ad Hoc Communication Networks*, vol. 23, no. 5, pp. 63–73, Sep. 2006.

[4] M. Durvy and P. Thiran, "A packing approach to compare slotted and non-slotted medium access control," in *Proc. IEEE INFOCOM*, Barcelona, Spain, Apr. 2006, pp. 1–12.

[5] T. Elbatt and A. Ephremides, "Joint scheduling and power control for wireless ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 74–85, Jan. 2004.

[6] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar, "Next century challenges: Scalable coordination in sensor networks," in *Proc. ACM MOBICOM*, Seattle, WA, Aug. 1999, pp. 263–270.

[7] S. Gandham, M. Dawande, and R. Prakash, "Link scheduling in sensor networks: Distributed edge coloring revisited," in *Proc. IEEE INFOCOM*, Miami, FL, Mar. 2005, vol. 4, pp. 2492–2501.

[8] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, pp. 721–741, Jun. 1984.

[9] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[10] B. Hajek and G. Sasaki, "Link scheduling in polynomial time," *IEEE Trans. Inf. Theory*, vol. 34, no. 4, pp. 910–917, Sep. 1988.

[11] K. Huang, *Statistical Mechanics*. New York: Wiley, 1987.

[12] S. Jeon and C. Ji, "Nearly optimal distributed configuration management using probabilistic graphical models," in *Proc. IEEE Resource Provisioning and Management in Sensor Network (RPMSN) Workshop*, Washington, DC, Nov. 2005, pp. 1–8.

[13] S. Jeon, "Near-optimality of distributed network management with a machine learning approach" Ph.D. dissertation, Georgia Inst. of Technology, Atlanta, Jul. 2007 [Online]. Available: http://smartech.gatech.edu/dspace/handle/1853/16136

[14] M. Jordan and Y. Weiss, "Graphical models: Probabilistic inference," in *Handbook of Neural Networks and Brain Theory*. Cambridge, MA: MIT Press, 2002.

[15] B. Kauffmann, F. Baccelli, A. Chaintreau, V. Mhatre, D. Papagiannaki, and C. Diot, "Measurement-based self organization of interfering 802.11 wireless access networks," in *Proc. IEEE INFOCOM*, Anchorage, AL, 2007.

[16] F. Kschischang, B. Frey, and H. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

[17] B. Ko and D. Rubenstein, "A distributed, self-stabilizing protocol for placement of replicated resources in emerging networks," in *Proc. ACM Int. Conf. Network Protocols (ICNP)*, Atlanta, GA, Nov. 2003, pp. 6–15.

[18] S. Z. Li, *Markov Random Field Modeling in Computer Vision*. New York: Springer-Verlag, 1995.

[19] X. Lin, N. Shroff, and R. Srikant, "A tutorial on cross-layer optimization in wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1452–1463, Jun. 2006.

[20] G. Liu and C. Ji, "Cross-layer graphical models for resilience of all-optical networks under crosstalk attacks," *IEEE J. Sel. Areas Commun. Optical Commun. Netw. Series*, vol. 25, pp. 2–17, 2007.

[21] H. Luo, S. Lu, V. Bharghavan, J. Cheng, and G. Zhong, "A packet scheduling approach to QoS support in multihop wireless networks," *ACM Mobile Netw. Appl. (MONET)*, vol. 9, no. 3, pp. 193–206, Jun. 2004.

[22] R. Madan, S. Cui, S. Lall, and A. Goldsmith, "Cross-layer design for lifetime maximization in interference-limited wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 11, pp. 3142–3152, Nov. 2006.

[23] E. Meshkova, J. Riihijarvi, and P. Mahonen, "Exploring simulated annealing and graphical models for optimization in cognitive wireless networks," presented at the IEEE GLOBECOM, Honolulu, HI, Nov. 2009.

[24] E. Modiano, D. Shah, and G. Zussman, "Maximizing throughput in wireless networks via gossiping," in *Proc. ACM SIGMETRICS/IFIP Performance*, Saint-Malo, France, Jun. 2006, pp. 27–38.

[25] A. Srinivasan and A. Panconesi, "Fast randomized algorithms for distributed edge coloring," in *Proc. ACM Symp. Principles of Distributed Computing*, Vancouver, Canada, Aug. 1992, pp. 251–262.

[26] J. Tsitsiklis and M. Athens, "On the complexity of decentralized decision making and detection problem," *IEEE Trans. Autom. Control*, vol. 30, no. 5, pp. 440–446, May 1985.

[27] A. Wagner and V. Anantharam, "Wireless sensor network design via interacting particles," in *Proc. Allerton Conf.*, Oct. 2002.

[28] R. Wattenhofer, L. Li, P. Bahl, and Y. Wang, "Distributed topology control for power efficient operation in multihop wireless networks," in *Proc. IEEE INFOCOM*, Anchorage, AL, Apr. 2001, vol. 3, pp. 1388–1397.

Sung-eok Jeon received the B.S. degree from Yonsei University, Seoul, Korea, in 1996, the M.S. degree from KAIST, Daejeon, Korea, in 1999, the Ph.D. degree from the Georgia Institute of Technology, in 2007, all in electrical engineering.

He is now in Microsoft, Redmond, WA. His research interest lies in probabilistic system modeling, statistical near-optimal distributed management of large networks, probabilistic graphical models, information abstraction, and data mining.

Chuanyi Ji (S'85–M'91–SM'03) received the B.S. (Hons.) degree from Tsinghua University, Beijing, China, in 1983, the M.S. degree from the University of Pennsylvania, Philadelphia, in 1986, and the Ph.D. degree from the California Institute of Technology, Pasadena, in 1992, all in electrical engineering.

She was on the faculty at Rensselaer Polytechnic Institute, Troy, NY, from 1991 to 2001. She is currently an Associate Professor in the Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta. Her research interests are in understanding and managing complex networks, applications of machine learning to network management and security, large-scale measurements, learning theory and algorithms, and information theory. Her research lies in the areas of networking, machine learning, and their interfaces.

Dr. Ji received an NSF Career Award in 1995, and an Early Career Award from Rensselaer Polytechnic Institute in 2000.