Statistical Mechanics of Optimal Networked Source Coding

Nima Sarshar and Xiaolin Wu
sarshan@mcmaster.ca, xwu@ece.mcmaster.ca
Department of Electrical and Computer Engineering, McMaster University
1280 Main St. West, Hamilton, Ontario, Canada

***Abstract*— We consider a large scale network of autonomous sensors each equipped with a local encoder and a local decoder. The nodes collaborate to reproduce a distributed source using a multiple description code (MDC). Each local encoder is a side encoder of MDC, producing one of the many descriptions and broadcasts its encoding to all its neighbors via an unreliable network. Each local decoder reconstructs the source using descriptions received from its neighbors along with the locally encoded description. Since duplicate descriptions arrived at a node will benefit it less, side MDC encoders should be assigned to nodes in such a way that, in an average sense, nodes receive the most number of distinct descriptions. This optimization problem is particularly challenging if the assignment has to be determined locally without central knowledge of the network. For simple network topologies, we consider the global orchestration of the nodes using a family of stochastic optimization algorithms, namely simulated annealing. Analogies to statistical mechanical models such as the antiferromagnetic Ising models are drawn which allow us to analytically derive the average distortion at steady state and characterize the ordering (and lack of it) in the assignments as a function of protocol parameters.**

I. Introduction

Diversity source coding schemes, like multiple-description codes (MDC), improve robustness of unreliable networks [5]. The idea is to add redundancy across different source code packets to combat packet loss. In a collaborative source coding application, such as a sensor network, where a large amount of correlation exists between the source measurements at neighboring nodes, the task of diversity coding can be distributed among multiple nodes via an MDC technique. Neighboring nodes collaborate by each producing one of the many MDC descriptions, i.e., each being a side encoder of MDC, and sharing it with the neighbors. In this setting the interplay between the network topology and the placement of MDC side encoders becomes an important issue that affects the overall system performance. The question is how to orchestrate the nodes of the network to perform MDC encoding and decoding under a given fidelity criterion and network constraints. The problem becomes more challenging in a large scale network of antonymous agents where any adaptation protocol should act locally without central control.

There are many applications for the scenario being studied. An example is a network of sensors of limited computational and communication capabilities deployed in an unreliable distributed environment. The network aims at gathering as precise an information as possible about a physical phenomenon distributed in the environment, such as the temperature or the concentration of a chemical in the atmosphere. Each node measures and encodes a lossy version of the source in its proximity and then communicates its encoding to all its neighbors. In consideration of the lossy nature of the network communications however, an MDC scheme is adopted. Constrained by low cost, low bandwidth, and low power consumption, each sensor has limited precision and each node is only able to encode one of the multiple descriptions locally and relies on its neighbors to supply the rest of the descriptions. The correlation of the source at neighboring nodes will help a node to improve its measurement upon receiving the coded observations of its neighbors. Networked MDC coding based on low-rate multiple description quantization [4] can be used to provide a robust solution to the problem. In this case we want to assign side encoders to different nodes with respect to a given network topology and channel statistics to achieve minimum expected estimation error overall all nodes.

Optimally designing the multiple description coding strategies and collaboration protocols for a large network is a challenging task. A number of code design algorithms for similar problems are considered in [1], [4]. One needs to consider the correlation structure of the source, the topology of the network and the specifics of the lossy communication medium.

In this paper however, we solely consider the collaboration aspects of the problem in large scale networks. Local coordination protocols remain the main challenge in almost any networked application [2]. We raise the following question: For optimal utilization of the network resources, how should the side encoders of MDC be strategically placed in the network so that the sum of the expected distortion of all the nodes is minimized? Algorithmically, can an optimal placement of the MDC side encoders be computed by local decisions made at each node? In a design process, for a given protocol, will an iterative greedy assignment of MDC descriptions to the nodes converge to a steady state, if yes, how fast is the convergence? How well will this greedy assignment algorithm perform in terms of approaching the globally minimum of the expected distortion?

The answers to these questions will shed light on large scale cooperative source coding strategies. In answering them, we restrict ourselves to a special case, where there are only two descriptions to choose from and the network topology is a simple $N \times N$ mesh, for some $N \gg 1$. These assumptions will enable us to make proper connections to and borrow tools from some elegant statistical mechanics models such as the Ising model on square lattices. As it turns out, even this simple scenario reveals the inherent structure of the problem arising from considering dynamics of a cooperative system.

The problem is formulated in Section II. Connections to statistical mechanics models are made in Section III by introducing a general Markovian dynamics which also resembles the simulated annealing process. This facilitates a detailed study in Section IV of the statistics of the emerging assignment of the MDC descriptions to nodes for the case of very unreliable channels. Different phases of the steady state assignments are explicitly characterized. The steady states for greedy algorithms and general forms of the simulated annealing are also considered. The average distortion in the steady state is found in terms of the protocol parameters. Phase transitions, where the characteristics of the emergent assignments drastically change with a slight change of the protocol parameters are fully characterized. Simulations and numerical observations for more general models that are not analytically solvable are reported in Section V followed by concluding remarks and a sketch of the future work.

II. Problem Formulation

A. General Formalism

The problem is abstracted as follows: Consider a network represented by the graph $G = \langle V, E \rangle$ and a multiple description code of m descriptions: $\mathcal{C} = \{C_1, C_2, ..., C_m\}$. Each node $v \in V$ is assumed to be able to locally encode its observation of a source into one of these m descriptions, denoted by $C(v) \in \mathcal{C}$. We denote the assignment of MDC descriptions to nodes by the vector $\mathbf{\Lambda}$.

Each node broadcasts its encoding to all its neighbors. Let $n(v) \subset V$ be the set of the neighbors of v. Each neighbor $l \in n(v)$ receives the encoding $C(v)$ with probability p.

Any node has a decoder that can reconstruct the source from any subset of $\mathcal{C}$ available to it. Let $S(v)$ be the set of MDC descriptions available to node v. When $p = 1$ and all communications are successful $S(v) = \{C(v)\} \cup \{\bigcup_{v' \in n(v)} C(v')\}$, whereas $S(v) = \{C(v)\}$ for $p = 0$. Define a function $d : 2^{\mathcal{C}} \rightarrow \mathcal{R}^+$ that represents the average reconstruction error as a function of the subset of the MDC descriptions available for decoding. The average distortion of all the nodes in the network is thus:

$$\mathcal{D}(\mathbf{\Lambda}) = |V|^{-1} \sum_{v \in V} E\{d(S(v))\}$$

where the expectation is taken, given v, over all possible sets of MDC descriptions $S(v)$. Note that $\mathcal{D}(\mathbf{\Lambda})$ is solely a function of the assignment vector $\mathbf{\Lambda}$.

B. Optimization of the Assignment Vector $\mathbf{\Lambda_t}$

An optimum assignment of the MDC descriptions to nodes is defined by

$$\mathbf{\Lambda}^* = \arg \min_{\mathbf{\Lambda} \in \mathcal{C}^{|V|}} \mathcal{D}(\mathbf{\Lambda})$$

Computing $\mathbf{\Lambda}^*$ for a general topology is most likely a hard combinatorial optimization problem. Computation of the optimal assignment vector $\mathbf{\Lambda}^*$ through local distributed decisions, as is the subject of this paper, proves to be a challenging problem even for simple network topologies. Let us consider an iterative process to obtain a locally optimal solution. We start with some (perhaps random) assignment $\mathbf{\Lambda}_0$. The assignment vector at the t^{th} iteration is called $\mathbf{\Lambda}_t$. In the t^{th} iteration a randomly selected node v is to locally update the choice of its description. The decision should be made solely on the basis of the t^{th} MDC assignment, namely, $C_t(v)$ and $C_t(u)$ for all $u \in n(v)$. The update changes the MDC assignment vector $\mathbf{\Lambda}_t$ to $\mathbf{\Lambda}_{t+1}$. By this construction, the two vectors $\mathbf{\Lambda}_t$ and $\mathbf{\Lambda}_{t+1}$ can differ in at most one position.

The above iterative optimization algorithm is said to converge if the $\lim_{t\rightarrow\infty} \mathcal{D}(\mathbf{\Lambda}_t)$ exists. If so the value of the limit $\mathcal{D}_\infty = \mathcal{D}(\mathbf{\Lambda}_\infty)$ is called the steady state distortion. The dynamics is said to be terminating if there exists a finite time t' such that $\mathbf{\Lambda}_t = \mathbf{\Lambda}_{t'}$ for any $t \geq t'$.

Central to the performance of the above iterative algorithm is the update rule that is a (perhaps stochastic) function of $C_t(v)$ and $C_t(u)$ for all $u \in n(v)$. A greedy update rule, denoted by $G : \mathbf{\Lambda}_{t+1} = G(\mathbf{\Lambda}_t)$, is the one that chooses $C_{t+1}(v)$ for node v solely to minimize the distortion d_v. While a greedy decision at a given node v is the best move for v alone, it might increase the overall distortion $\mathcal{D}(\mathbf{\Lambda}_{t+1})$ due to the interference with the neighbors of v. Also the simple greedy MDC assignment algorithm might neither be convergent nor terminating. Even when converging, the convergence might be very slow, as we will see later on.

To overcome these drawbacks of the greedy algorithm, simulated annealing algorithms will allow a node v to increase its distortion with some (usually small) probability. This probability depends on some scalar called the temperature T and on the amount the new selection will affect the distortion at the node v. We will examine a class of simulated annealing algorithms that are guaranteed to converge in finite time. Most of this paper is devoted to characterizing the steady state MDC assignment resulting from these algorithms.

These notions are made precise in the following section where we choose a simple, yet practically interesting network topology and restrict ourselves to simple two-description MDC codes.

III. Two-Descriptions Codes on a Mesh

A. The Problem Setup

Here after we assume that the network topology is an $N \times N$ planar mesh for a very large N. The MDC code has only two descriptions: $\mathcal{C} = \{C_1, C_2\}$. The distortion given C_1 and C_2 only is δ_1 and δ_2 respectively, while the distortion given both C_1, C_2 is δ_0.

Each node not on the boundary of the mesh has exactly four neighbors. Define the following auxiliary vector, called the state vector: $\mathbf{z} = (z_1, z_2, ..., z_M)$, $M = N^2$, where $z_v = 1$ if the node v chooses the code 1 (i.e., $C(v) = C_1$) and $z_v = -1$ otherwise (i.e., if $C(v) = C_2$). With this assumption, the average distortion at a node v is the following:

$$d_v(z_v) = (1 - q_v)[I(z_v = 1)d_1 + I(z_v = -1)d_2] + q_v d_0 \quad (1)$$

where I is the indicator function and

$$q_v = 1 - (1-p)^{|n_1(v)|I(z_v=1)+|n_2(v)|I(z_v=-1)},$$

in which $n_1(v), n_2(v)$ are defined as the set of the neighbors of v which code C_1, C_2 respectively.

B. Simulated Annealing: The Metropolis-Hastings Algorithm

We describe a general Markovian assignment algorithm. Any node updates its choice of MDC description independently from other nodes. The update rule is as follows. If $d_v(-z_v) < d_v(z_v)$, i.e., node v benefits from swapping between the two MDC descriptions, then the node will change its MDC description (and therefore $z_v \to -z_v$). If $d_v(-z_v) \geq d_v(z_v)$, on the other hand, the node v may still flip its MDC description but does so with probability $q = e^{-\frac{d_v(-z_v) - d_v(z_v)}{T}}$ for a constant T called the temperature. For $T \to 0^+$, this algorithm reduces to the greedy algorithm, that is, only the swaps that locally decrease the distortion at a node v are accepted. This update algorithm, a special form of simulated annealing [7], is called the Metropolis-Hastings [8] algorithm and is used to simulate more general Markovian processes.

C. Convergence of the Update Algorithm

The Metropolis-Hastings update algorithm defines a Markov chain with states uniquely determined by $\mathbf{z}(t)$, the state vector of iteration t. For any $T > 0$, this Markov chain is non-null recurrent, i.e, the transition from any state to any other state is possible with finite probability. As such, the distribution on the state of the chain $\mathbf{z}(t)$ will converge to a steady state [3]. We can easily show that the steady state probability of the system having the assignment vector $\mathbf{z}$ is

$$Pr\{\mathbf{z}\} = e^{-\frac{\mathcal{D}(\mathbf{z})}{T}} U^{-1} \quad (2)$$

where U is a normalizing factor not depending on $\mathbf{z}$.

This can be done through the following consistency check: Consider in steady state, two state vectors $\mathbf{z}$ and $\mathbf{z}'$ that differ only in the state of a single node v. The following identity would therefore be true:

$$Pr\{\mathbf{z}\}Pr\{\mathbf{z} \to \mathbf{z}'\} = Pr\{\mathbf{z}'\}Pr\{\mathbf{z}' \to \mathbf{z}\}$$

Where $Pr\{\mathbf{z}\}, Pr\{\mathbf{z}'\}$ are given by (2). Thus for $Pr\{\mathbf{z}\}$, given by (2), the *probability flow* from $\mathbf{z}$ to $\mathbf{z}'$ is equal to the flow from $\mathbf{z}'$ to $\mathbf{z}$. This is sufficient (though not necessary) to prove the convergence of the Markov chain to a steady state distribution given by (2) [8].

Why not Greedy: As $T \to 0^+$, the probability of finding the system in any other state except the one that minimizes the overall distortion would become small. This however is only true if the steady state distribution in (2) can be achieved. As was mentioned before, and will be confirmed through the simulations, this steady state might not be achievable for $T = 0$. A simulation result comparing the greedy algorithm with a case of gradually decreasing T is given in later sections (see Fig. 6). While the steady state is achieved for any $T > 0$, the speed of convergence might be very slow for very small T. The convergence time can be shown to scale as $O(\log 1/T)$ [7] for small T. Characterization of the overall distortion and the state of the network at a small positive T will thus become important.

The quantities of interest, including the average distortion at steady state as well as its rate of change as a function of T are calculated in the following section.

D. Unreliable Networks and Aniferromagnetic Ising Model

We start by confining our attention to small values of the probability of successful transmission p. The case for more general p is treated in later sections. Assuming $p \ll 1$, the probability of successful communication of the second description can be approximated as:

$$q_v \approx p(|n_1(v)|I(z_v = 1) + |n_2(v)|I(z_v = -1))$$

The average distortion $d_v(z_v)$ is approximately:

$$\begin{aligned} d_v(z_v) \quad &\approx \quad (1 - |n_1(v)|I(z_v = 1)p)\delta_1 \\ &+ \quad (1 - |n_2(v)|I(z_v = -1)p)\delta_2 \\ &+ \quad p[I(z_v = 1)|n_1(v)| + I(z_v = -1)|n_2(v)|]\delta_0 \end{aligned} \quad (3)$$

It is easy to verify that $I(z_v = 1)|n_1(v)| + I(z_v = -1)|n_2(v)|$, the number of neighbors of v that have a code different from v, is: $(4 - \sum_{n' \in n(v)} z_v z_{v'})/2$. Also, define the following constants: $\alpha \equiv (\delta_1 + \delta_2)/2$ and $\beta \equiv (\delta_1 - \delta_2)/$, $\gamma \equiv (d_1 + d_2 - 2d_0)/2$.

Inserting these observations into (3) results in:

$$\begin{aligned} d_v(z_v) \quad &\approx \quad \alpha + \beta z_v - \frac{p}{2}(4 - \sum_{v' \in n(v)} z_v z_{v'})(\gamma + \beta z_v) \\ &= \quad \alpha - 2p\gamma + \beta z_v + p\gamma \sum_{v' \in n(v)} z_v z_{v'} \\ &- \quad 2p\beta z_v + (p\beta/2) \sum_{v' \in n(v)} z_{v'} \end{aligned} \quad (4)$$

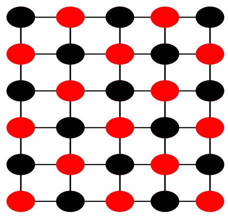

Fig. 1. The assignment of codes result in the minimum distortion. The nodes with code C_1 are colored red, while those coding C_2 are colored black. Another minimum distortion assignment results from swapping the red and black nodes.

where to get the last term we have used the fact that $z_v^2 = 1$. The overall average distortion is found by averaging over all v:

$$\begin{aligned}\mathcal{D}(\mathbf{z}) &\approx |V|^{-1} \sum_{v \in V} d_v(z_v) \qquad (5)\\ &= |V|^{-1} \sum_{v \in V} \left((\alpha - 2p\gamma) + \beta z_v + p\gamma \sum_{v' \in n(v)} z_{v'} z_v \right)\end{aligned}$$

To get this, note that neglecting boundaries, each node v has exactly 4 neighbors. Thus when summed over all the nodes, the last two terms in d_v will cancel each other.

The three terms of $\mathcal{D}(\mathbf{z})$ are as follows: The first term is a constant (independent of the choice of z_v's). The second term indicates the preference in choosing the better description. If both descriptions have the same importance ($d_1 = d_2$) then $\beta = 0$, and there is no difference between the choice of the descriptions. The third term, interestingly, describes the preference of nodes in having neighbors that encode different descriptions. When $\beta = 0$, the choice of minimum distortion is a checker board binary assignment as depicted in Fig. 1.

E. Physical Interpretation

The distortion in (5) can represent the *energy* of a statistical mechanical Ising model [6] on a square lattice. The assignments z_v's are analogously the *spin* directions. The term $\sum_{v' \in n(v)} z_{v'} z_v$ is the nearest neighbor interaction. The constant $\gamma' = p\gamma$ represents the magnitude of this interaction. The constant β on the other hand can be thought of as an external magnetic filed. Since $\gamma' > 0$, this corresponds to the so called *antiferromagnetic* Ising model; when the external field (β) is zero, the minimum energy of the system corresponds to the checker board assignment of the spins as in Fig. 1.

IV. Steady State Behaviors

The Ising model has been long studied and well understood. The two dimensional Ising model on the square lattice, as considered here, has been exactly solved for the case of zero external filed, corresponding to balanced descriptions in our case. In this section, we will borrow these analytical results to gain insight into the steady state MDC assignments in our iterative design process.

A. Ordered and Unordered MDC Assignment

Consider first the case of balanced descriptions $\beta = 0$. For large T, the MDC assignment will be random. Thus, the probability that any neighbor of a given node has description C_1 (or C_2) will converge to $1/2$. On the other hand, as $T \to 0^+$, the neighboring nodes of v are more likely to have a different description from v.

To quantify this bias, we define two sub-meshes as follows: The odd sub-mesh V_o is the set of all nodes on the lattice with coordinates (i, j) such that $i + j$ is odd. Similarly, V_e is the set of all nodes whose sum of their lattice coordinates are even. Let $V_{e,1}, V_{e,2}$ be the subset of all nodes in V_e that are assigned C_1, C_2 respectively, and $V_{o,1}, V_{o,2}$ be the corresponding subsets of V_o. Now define:

$$\begin{aligned}\eta_o &= \frac{|V_{o,1}| - |V_{o,2}|}{|V_o|}\\ \eta_e &= \frac{|V_{e,1}| - |V_{e,2}|}{|V_e|} \qquad (6)\end{aligned}$$

Since $\beta = 0$, the minimum distortion is achieved when all the nodes of a sub-lattice encode one description while the nodes of the other sub-lattice encode the other description. Thus, η_o and η_e measure how close the assignments are to the optimal assignment. The checker board assignment in Fig. 1 is the one for which $\eta_e = -\eta_o$ and $|\eta_e| = |\eta_o| = 1$. The analogy to the Ising model allows us to analytically calculate η_e and η_o as a function of the temperature T, as reported in Section IV-B and depicted in Fig. 5. As the temperature raises from 0^+, the values of $|\eta_o|, |\eta_e|$ decrease from 1, but are greater than zero as far as $T < T_c$ for some critical temperature T_c. For any $T > T_c$, $|\eta_o| = |\eta_e| = 0$, indicating that the ordering of the sub-lattices is lost. A visualization of this phase transition (from an ordered assignment to one without order) is given in Fig. 2.

B. Average Distortion

For $\beta = 0$, the average distortion as well as the value of $|\eta_o|$ or $|\eta_e|$, and the value of the critical temperature can be exactly found (see [6] pp. 242-245).

$$\begin{aligned}|\eta_o| &= |\eta_e| = \frac{(s^4 - 1)^{1/8}}{s^{1/2}}\\ s &= sinh\,(2p\gamma/T) = sinh\left(2\frac{p(d - d_0)}{T}\right) \qquad (7)\end{aligned}$$

The phase transition happens when $T = T_c$ such that $s = 1$, from which:

$$\frac{2p(d - d_0)}{T} = \ln(1 + \sqrt{2})/2$$

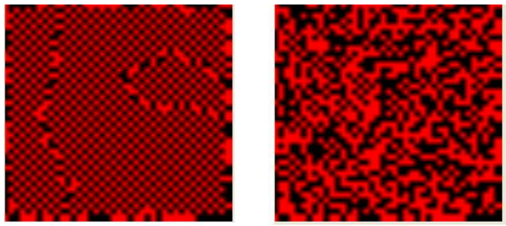

Fig. 2. The snapshot of code assignments at steady state for two values of temperature $T/(p(d-d_0)) = 1$ (left) and $T/(p(d-d_0)) = 4$ (right) when $p = 0.1$. The left image corresponds to a regime where the assignment of the codes is *ordered* ($T < T_c$). The right snapshot however, corresponds to a regime where there is no ordering in the code assignments ($T > T_c$).

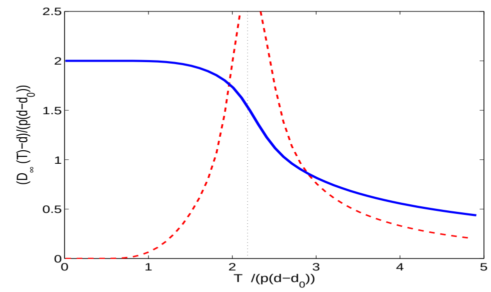

Fig. 3. Properly normalized distortion at the steady state as a function of the temperature. Also included, (dashed curve) is the derivative of this distortion. The phase transition point is also marked, where the derivative of the distortion reaches its maximum. The distortion is almost constant as far as $T \lesssim p(d-d_0)$

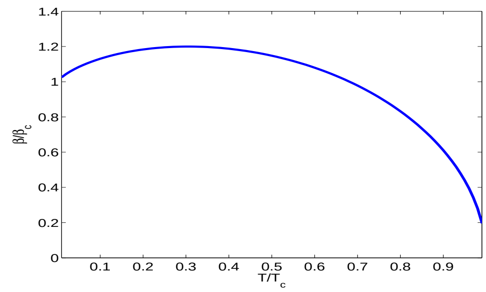

Fig. 4. The critical imbalance between the distortion as a function of the temperature. For β greater than this critical value, all the nodes will end up choosing the better description.

or

$$T_c = \frac{4p(d-d_0)}{\ln(1+\sqrt{2})} \approx 2.13p(d-d_0)$$

The overall average distortion in the steady state is

$$\mathcal{D}_\infty(T) = d - p(d-d_0)(4-u)$$

where u is given by

$$\begin{aligned} u &= coth(2K)[1 + (2/\pi)\{2tanh^2(2K) - 1\}\kappa(a)] \\ K &= \frac{p(d-d_0)}{T} \\ a &= 2sinh(2K)sech^2(2K) \\ \kappa(x) &= \frac{\pi}{2}[1 + (1/2)^2 x^2 + \left(\frac{1.3}{2.4}\right)^2 x^4 + ...] \end{aligned}$$

As $T \to \infty$, $u \to 2$, which corresponds to a totally random MDC assignment, we have

$$\mathcal{D}_\infty(T) = (1-2p)d + 2pd_0.$$

This is because on average, each node has two other neighbors with a different description. So on average, around $2p$ of all nodes (recall that $p \ll 1$) will have both descriptions available for decoding, hence the above result.

As $T \to 0^+$ on the other hand, all four neighbors of any node will take the description different from that node. Thus the average distortion at steady state becomes $(1-4p)d+4pd_0$. The maximum reduction in distortion compared to random MDC assignment is therefore $\Delta = D_\infty(\infty) - D_\infty(0) = 2p(d-d_0)$. The average distortion is depicted in the full range of T in Fig. 3.

C. Unbalanced Descriptions

When $d_1 = d_2 = d$ and thus $\beta = 0$, there is no difference between the two descriptions. The problem is significantly different for unbalanced descriptions where $\beta \neq 0$. The optimization process will have a tendency to choose the better description more frequently. On the other hand, by choosing the better description, a node might lose the opportunity of receiving distinct descriptions from its neighbors. Thus, a tradeoff would exist between choosing the better description and maximizing the diversity. For small values of $|\beta|$, the need for diversity dominates, making the MDC assignment close to the assignment in Fig. 1. If $|\beta|$ is large, there can come a point where the minimum distortion is achieved by assigning the better description to all the nodes. This critical point $\beta_c(T)$ would depend on T, but can be easily calculated only for $T \to 0^+$. From (5), the distortion for the uniform assignment is $H - |\beta| + 4p\gamma$, whereas the average distortion of the checker board assignment in Fig. 1 is $H - 4p\gamma$. Thus the transition (for $T = 0$) happens when

$$|\beta| \geq 8p\gamma = \beta_c(0)$$

or

$$|d_1 - d_2| \geq 4p(d_1 + d_2 - 2d_0)$$

For $T > 0$ on the other hand, the situation is far more involved. There is no exact solution for the critical value $\beta_c(T)$ such that for $\beta > \beta_c(T)$ all the nodes end up choosing the same description. Closed form approximations however do exist for this critical imbalance as a function of T. One such approximation (called the mean-filed approximation [6]) is depicted in Fig. 4.

D. The Case of General p

Up to now the update rule in iterative MDC optimization was based on the specific expression of the distortion $d_v(z_v)$

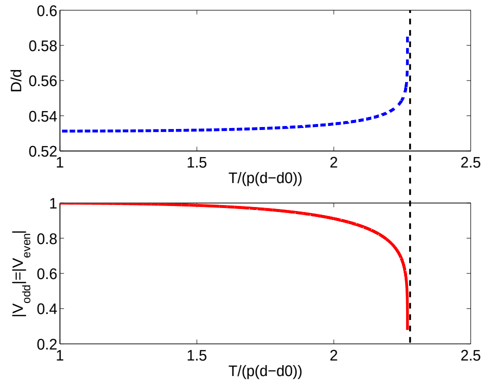

Fig. 5. The average distortion (top) and the value of η as a function of the temperature for T_c. The position of the phase transition is marked with a dashed line. As $T \to T_C^-$ the average distortion quickly climbs towards its value indicated by a random code assignment. It is assumed that $p = 0.5$ and $d = 2d_0$.

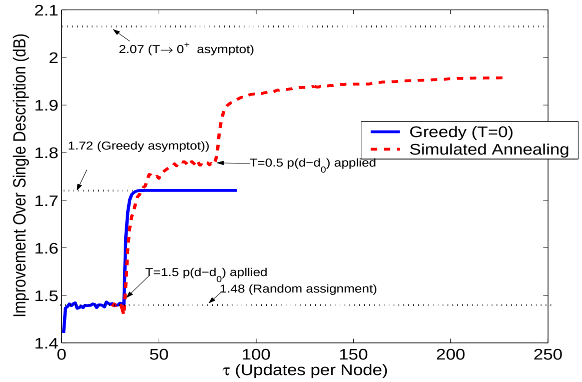

Fig. 6. Simulations of a system for $N = 100$. The average distortion as a function of τ, the number of updates per node, are depicted. Two identical systems start from a random ($T \gg 1$) configuration at time $\tau = 0$. At time $\tau = 30$, the temperature is set to exactly $T = 0$ (corresponding to greedy updates) for one system. The average distortion of this system as a function of τ is depicted with the solid curve. The distortion quickly saturates and the system is trapped in a local minima. For the second system however the temperature is decreased in two steps: At time $\tau = 30$ it is turned to $T = p(d - d_0) < T_c$ and then at time $\tau = 80$ it is further reduced to $T = 0.5p(d - d_0)$. The dashed curve shows the superiority of this algorithm to the greedy algorithm.

given by (4), which in turn was accurate only for small p. For general p however, one can still apply the Metropolis-Hastings algorithm pretending that the distortion form (5) still holds. In fact since the exact value of p might not be known, the update algorithm can simply assume $p = 1$ in (5). The actual distortion measure however should be taken from the exact distortion formula $d_v(z_v)$ in (1).

The average distortion at steady state can be found for a general value of p through the so-called mean-filed approximation [6]. The mean-filed approximation assumes that for any two *neighboring* nodes v, v', and for any real function f, all the expectations $Ef(z_v z_{v'})$ can be replaced with $f(E\{z_v\}E\{z_{v'}\}) = f(E\{z_v\}^2)$ (Note that the expectations are taken over the steady state distribution of the state vector **z**.

Using the mean-filed approximation, one gets $E\{q_v\} = 1 - (1-p)^{2(1+|\eta|^2)}$ with $\eta = |\eta_o| = |\eta_e|$, where $|\eta_o|, |\eta_e|$ are given by (7), with q_v being the probability of having both descriptions at node v. The distortion $d_v(z_v)$ in (1) (assuming $d_1 = d_2 = d$) can be written as:

$$\mathcal{D}_{mf} = E\{d_v\} = d(1 - E\{q_v\}) + d_0 E\{q_v\}$$

The distortion as a function of T is depicted in Fig. 5. A simulation with gradually decreasing temperature and its comparison with greedy algorithm ($T = 0$) is given in Fig. 6 for the case of $p = 0.5$.

V. Concluding Remarks

This paper considered the problem of local optimization of the diversity in a network of encoding/decoding sensors. This will help gather more accurate data at each source in an unreliable environment. The goal was to assign MDC descriptions to network nodes in such a way that the expected distortion over all nodes is minimized. We studied a simple iterative MDC assignment algorithm, and modelled it by a general Markovian dynamics. The convergence of this algorithm was guaranteed. Thus the nodes deployed to measure a distributed source can quickly adopt their coding strategies. The greedy algorithm, where each node makes its choice of MDC description only to benefit itself, was shown to be unable to reach the optimal MDC assignment in a finite time. Simulated annealing optimization strategies on the other hand where shown to converge in finite time with small penalties compared to the globally optimal assignments. The future work consists of generalizing the problem setup to more than two descriptions and to more complicated distortion measures and network topologies.

References

[1] S. D. Servetto, ”Lattice Quantization with Side Information”, Proceedings of the IEEE Data Compression Conference, Snowbird, UT, March 2000.

[2] D. Estrin, R. Govindan, J. Heidemann and S. Kumar, ”Next Century Challenges: Scalable Coordination in Sensor Networks.”, MobiCOM ’99, August 1999, Seattle, Washington.

[3] Cinlar, E., ”Introduction to Stochastic Processes”, Engelwood Cliffs, N.J.: Prentice Hall, 1975.

[4] R. Cristescu, B. Beferull-Lozano, and M. Vetterli, ”On network correlated data gathering,” in INFOCOM 2004.

[5] Fernandez Diaz,Dick Epema, Jan de Jongh, ” Multiple Description Source Coding and Diversity Routing: A Joint Source Channel Coding Approach to Real-Time Services over Dense Networks”, 13th International Packet Video Workshop, Nantes, 28-29 April 2003.

[6] D. A. Lavis, G. M. Bell, Statistical Mechanics of Lattice Systems Statistical Mechanics of Lattice Models, Springer Verlag, February 1999.

[7] Kirkpatrick, S.; Gelatt, C. D.; and Vecchi, M. P. ”Optimization by Simulated Annealing.” Science 220, 671-680, 1983.

[8] Chib, Siddhartha and Edward Greenberg: ”Understanding the MetropolisHastings Algorithm”. American Statistician, 49, 327335, 1995

[9] J. M. Kahn, R. H. Katz and K. S. J. Pister, ”Mobile Networking for Smart Dust”, ACM/IEEE Intl. Conf. on Mobile Computing and Networking (MobiCom 99), Seattle, WA, August 17-19, 1999