

Distributed Averaging in Dense Wireless Networks

S. Vanka, M. Haenggi, and V. Gupta

Department of Electrical Engineering

University of Notre Dame, Notre Dame, IN 46556, USA

E-mail: {svanka,mhaenggi,vgupta2}@nd.edu

Abstract—We consider the effect of network throughput on the convergence of a specific class of distributed averaging algorithms, called consensus algorithms. These algorithms rely on iterative computation of the desired average by message passing among the nodes. It is thus assumed that the rate of convergence should benefit from greater network connectivity. However, one must also account for the additional network resources that establishing such a connectivity would entail. In this paper, we study this problem in the context of randomly-placed consensus-seeking nodes that are connected through a dense wireless network, i.e., whose capacity is interference-limited. By analyzing the outage of each communication link along with results from mixing times of Markov chains, we obtain scaling laws for the mixing times of fastest-converging consensus topologies over such networks.

Keywords—Distributed Estimation, Consensus, Wireless Networks, Scaling Laws, MAC Protocols.

I. INTRODUCTION

The advent of wireless sensor and ad hoc networks has motivated the need for information processing algorithms that are inherently distributed, thereby allowing each node to run the algorithm by utilizing only local information. An archetypal instance of such an algorithm is the distributed averaging algorithm that computes the global average of sensor observations based on purely local computations at each sensing node. Distributed averaging algorithms provide a useful building block to solve more complex problems such as distributed estimation or detection, optimization or compression. The reason lies in the ease with which they can be generalized to compute a linear projection of measurements.

A well-studied distributed averaging algorithm is the average consensus algorithm, where a set of nodes in a (possibly time-varying) interconnection topology iteratively compute the global average of their initial states, see e.g., [1]–[5] and references therein. The connectivity properties of the graph that ensure convergence have been well-studied (eg. [6], [7]). Of late, the focus has shifted to studying convergence in the face of communication constraints, like quantization [8], [9], packet drops [10] and noise [11]. A closely associated algorithm is the gossip algorithm, studied for instance in [12] and [2].

Given the broadcast nature of the wireless channel, the geometry of the network plays a important role in network performance. As noted in [13], wireless networks may be broadly classified as either *dense* or *extended* networks. While

the former are an abstraction for large networks of spatially localized radios, the latter are a useful model to study constant-density deployments of radios over a large area. Wireless links in extended wireless networks can be either power- or interference- limited, whereas those in dense networks are only interference-limited.

In this paper, our objective is to derive scaling laws that quantify the convergence behavior of a large wireless network of consensus-seeking nodes where throughput is limited by interference alone. This problem is formulated as a distributed averaging problem in a spatially localized deployment of sensing nodes – hence we assume a dense network model. Furthermore, practical networks are packetized, i.e., a received message is meaningless if the packet that encapsulates the message cannot be decoded correctly. In other words, a message from a source to a destination is considered valid iff the link is *not* in outage. Thus we believe that the model above provides a key ingredient to a more realistic study of this problem.

The remainder of this paper is organized as follows. In Section II we provide a precise statement of the problem. Using the connectivity model described therein, we derive the scaling laws for the achievable link throughput in Section III. This is followed by an interpretation of these results to analyze the convergence behavior of average consensus in Section IV. Section V provides some directions for further work. The asymptotic notation used in this paper is as follows. For two functions f and g of a variable n , as $n \rightarrow \infty$, we write

- $g = \mathcal{O}(f)$ if the ratio g/f is asymptotically finite. $g = o(f)$ if this limit is zero.
- $g = \Omega(f)$ if $f = \mathcal{O}(g)$. $g = \omega(f) \iff f = o(g)$.
- $g = \Theta(f)$ if $g = \mathcal{O}(f)$ and $g = \Omega(f)$.

II. PROBLEM FORMULATION

A. Consensus and Random Walks

Consider a graph G specified by a set of vertices $\mathcal{V} = \{0, 1, \dots, n-1\}$ and a set of (undirected) edges E . Associate with each vertex i a state x_i that is updated in a synchronous manner to x_i^+ as

$$x_i^+ = x_i + h \sum_{j \in \mathcal{N}_i} (x_j - x_i) \quad (1)$$

where \mathcal{N}_i denotes the set of all neighboring nodes $j \neq i \in \mathcal{V}$ such that the edge $(i, j) \in E$, h is a constant chosen so that $h < \frac{1}{2d_{\max}}$ where d_{\max} is the maximum node degree of G . If

The partial support of NSF (grant CNS 04-47869) and DTRA (grant N00164-07-8510) is gratefully acknowledged. The work for the third author was supported partially by the NSF award 0846631.

one stacks the individual states x_i to form the column vector x , the state update (1) can be expressed as

$$x^+ = Wx, \quad (2)$$

where $W \triangleq I - hL$ and L is the Laplacian matrix of the graph G . Observe that this algorithm is distributed, in that each node only needs the knowledge of the states of its neighbors to implement (1). For an arbitrary vector of initial states $x(0) \triangleq [x_i(0)]_{n \times 1}$, if G is connected and balanced, it can be shown [6] that under this algorithm the state of each vertex asymptotically reaches $\frac{1}{n} \sum_i x_i(0) = \mathbf{1}^* \frac{x(0)}{n}$, where $\mathbf{1}$ denotes an n -dimensional all-ones vector, and A^* denotes the transpose of the matrix A . Thus, each state of each vertex converges to the average of the initial state $x(0)$.

The construction of the update matrix W guarantees that it is always stochastic. Moreover, if G is balanced, all the column sums are also unity, resulting in a doubly stochastic matrix W . The stochasticity of W allows us to formulate the convergence of (2) as follows. Consider a random walk on a graph G , where the probability mass function of the position of the particle is given by x and the transition probability matrix by W . The equilibrium probability mass function of the position of the particle now corresponds to the asymptotic state (if it exists) x for the consensus problem. It is well-known [14] that that the constraints imposed on G can be understood in terms of the recurrence, aperiodicity and the time reversibility of the Markov chain \mathcal{M} associated with this walk.

The *rate* of convergence determines how fast a Markov chain converges to its stationary distribution (if it exists). One approach to measure this rate is to consider the Total Variational (TV) distance¹ (cf. [15, Chap. 4]) between the stationary distribution $\pi \triangleq \mathbf{1}/n$ and the distribution $p(k)$ after k iterations. Since W is a contraction operator on \mathbb{R}^n whose fixed point is the stationary distribution π , such a distance is always monotonically decreasing. The idea, therefore, is to quantify the number of iterations it takes to reach a pre-determined “distance” from stationarity. Using this definition, the (ϵ -)mixing time of the Markov chain \mathcal{M} can be informally understood as the smallest number of iterations τ_ϵ beyond which the total variational distance of the distribution $p(k; p(0))$ after k iterations from the stationary distribution π is guaranteed to be smaller than ϵ , for any initial state $p(0)$. Formally,

$$\tau_\epsilon(W) = \sup_{p(0)} \inf_k \{k : \|p(k; p(0)) - \pi\|_{TV} \leq \epsilon\}. \quad (3)$$

Without loss of generality, let $x_i(0) > 0$, and define $y_i(0) \triangleq x_i(0) / \sum_i x_i(0)$ as the normalized initial state. Note that the state y is a valid probability distribution. Applying (3) to express the mixing time and expressing y in terms of the original state x we get

$$\tau_\epsilon(W) = \sup_{x(0)} \inf_k \{k : \|x(k) - n^{-1} \mathbf{1} s_0\|_{TV} \leq \epsilon s_0\} \quad (4)$$

¹The TV distance between two distributions μ and ν over \mathcal{S} is defined as $\|\mu - \nu\|_{TV} \triangleq \frac{1}{2} \sum_{i \in \mathcal{S}} |\mu_i - \nu_i|$ (essentially the ℓ_1 norm).

where $s_0 \triangleq \sum_i x_i(0)$. In this paper, we are interested in consensus over random geometric graphs based on the so-called disc model [16]. Below we define the graph in question and cite work related to the mixing times of random walks over these graphs.

Let $G_d(n, s)$ denote a d -dimensional random geometric graph formed by n vertices that are placed uniformly randomly on a torus $[0, 1]^d$, independent of each other. An edge (i, j) exists between vertices i and j iff $\|r_i - r_j\| \leq s$. The set of neighbors j for a vertex i is given by

$$\mathcal{N}_i(s) \triangleq \{j : \|r_j - r_i\| \leq s, j \neq i\}$$

This is the so-called disc graph model. Unless stated otherwise, we assume $s = \omega((\frac{\log n}{n})^{1/d})$. This scaling of s guarantees that $G_d(n, s)$ is connected and regular asymptotically almost surely (a.a.s.) [17]. For brevity, define $s_c(n) \triangleq (\frac{\log n}{n})^{\frac{1}{d}}$.

Consider a random walk on $G_d(n, s)$. Given a vertex $v \in G_d(n, s)$. With a given probability $q > 0$ jump to any neighbor of $u \in \mathcal{N}(v)$ chosen uniformly at random. With probability $1 > 1 - |\mathcal{N}(v)|q > 0$ remain at v . Such a random walk can be described by a Markov chain with a transition probability matrix W given by

$$u_{ij} = \begin{cases} q, & \|r_j - r_i\| \leq s, i \neq j \\ 1 - |\mathcal{N}(i)|q, & i = j \end{cases} \quad (5)$$

for $i, j = 0, 1, \dots, n-1$. Note that if $s = \omega(s_c)$, a.a.s. $|\mathcal{N}(i)| = n(\pi s^2) + o(1)$. The mixing time of such a random walk on $G_d(n, s)$, for an error $\epsilon = 1/n^\delta$, $\delta > 0$, is known [2]:

$$\tau(W) = \Theta(s^{-2} \log n). \quad (6)$$

It is that important to note that s scales with n as mentioned above. We will refer to s as the *interaction radius* in the rest of the paper. Consider an average consensus algorithm over $G_d(n, s)$. From (4) it is clear that $\tau_\epsilon(W) = \Theta(s^{-2} \log n)$. Indeed this correspondence has been exploited before [2], [18]. In particular, the scaling behavior of mixing time of fast-mixing random walks on geometric random graphs was analyzed in [2].

B. Network Model

The state update (1) requires each vertex i in G to have access to the state values $\{x_j : j \in \mathcal{N}_i\}$ of its neighbors, or equivalently, that each i can communicate with all its neighbors \mathcal{N}_i without errors. If all nodes share a common communication medium (e.g. space, as in a wireless network), this requirement limits the network resources available to each node. In such a shared medium, the transmission rate available to any transmitter-receiver pair can be much lower than what would be achievable with independent point-to-point channels.

In this paper, we study this problem of implementing (2) in the context of a wireless network, where interference from concurrent transmissions limits the throughput between two nodes. Thus the spatial locations of the vertices of G become important. We make the following assumptions:

- Nodes are placed in a finite volume. For concreteness, we assume they are placed on a d -dimensional torus $[0, 1]^d$.
- Node locations $\{r_i\}$ are iid and uniform on $[0, 1]^d$. Such a model maximizes the uncertainty in node locations.
- Messages $\{x_j : j \in \mathcal{N}_i\}$ for all i have no quantization errors. A node uses exactly one packet to encode its message.
- A time slot is defined as one packet transmission time and normalized to 1.
- There is no fading.
- The network throughput is limited by interference alone. Furthermore, a packet from node i can be received at j iff the Signal-to-Interference-Ratio (SIR) at node j , SIR_{ij} , is greater than a known constant $\beta > 0$. Therefore for any sender i and receiver j , the link $i \rightarrow j$ will be in outage iff

$$\frac{\|r_j - r_i\|^{-\alpha}}{\sum_{k \in \mathcal{S} \setminus \{i\}} \|r_j - r_k\|^{-\alpha}} < \beta.$$

where \mathcal{S} is the set of all senders that transmit in the same slot as i .

- All messages are transmitted at the same rate $R = R(\beta)$ and at the same power. Transmission rate is close to channel capacity.
- Receivers do not perform interference cancellation.
- The medium-access scheme is TDMA with spatial re-use.
- The path-loss exponent α is greater than the dimension d of the network, such that the interference remains finite a.s. as the number of nodes grows.
- The (desired) neighbor set $\mathcal{N}_i(s)$ for each node i consists of all nodes located at a distance at most s , i.e., $\mathcal{N}_i(s) \triangleq \{j : \|r_j - r_i\| \leq s\}$.
- All edges in the consensus graph correspond to wireless links.

C. Problems Studied in this Paper

We study iterative distributed averaging algorithms of the form (2) on a geometric random graph as defined in Section II-A. Our key contribution lies in explicitly quantifying the effect of interference on the convergence speed of a distributed averaging algorithm in a dense wireless network. Consider a set of n nodes on a torus $[0, 1]^d$ that exchange messages at a rate $R(\beta)$ over an interaction radius s . We study the following problems:

- How does the achievable throughput between a transmitter and a receiver separated by a distance s scale with the number of nodes n and the transmission rate R , given that the transmitter must be able to communicate with *all* receivers at a distance less than s from itself?
- How does the mixing time scale with the interaction radius s and the transmission rate R ?

D. Asymptotic Behavior

From Sections II-A and II-B we notice that the problem involves three parameters, namely:

- The size of the network n .

- The interaction radius s . Unless stated otherwise, this scales as $\omega(s_c)$ to ensure asymptotic almost sure (a.a.s.) connectivity of the network.
- The transmission rate R which is a function of the SIR threshold β . Clearly this can be chosen independent of either the network size n or the interaction radius s .

Of interest is the scaling behavior as $n \rightarrow \infty$. The functional forms of the dependence of throughput on the interaction radius s and the SIR threshold β have been derived for this regime.

III. INTERFERENCE-LIMITED THROUGHPUT

We now prove two results that follow from the assumptions made in Section II-B.

Proposition 1. *Consider a d -dimensional geometric random graph $G(n, s)$ on $[0, 1]^d$, with $s = \omega(s_c)$ and the assumptions stated in Section II-B. The length of the shortest TDMA schedule that ensures that each node i is connected to all nodes $\{j : \|r_j - r_i\| \leq s\}$ cannot have fewer than $C_1 n s^d (\beta^{1/\alpha} - 1)^d$ time slots a.a.s., for some positive constant C_1 .*

Proof: Let \mathcal{S} be the set of concurrent transmitters at any given time. Suppose node j is an intended receiver of a transmitter $i \in \mathcal{S}$. Then i 's message is decoded correctly iff

$$\frac{\|r_j - r_i\|^{-\alpha}}{\sum_{k \in \mathcal{S} \setminus \{i\}} \|r_j - r_k\|^{-\alpha}} \geq \beta.$$

This condition is satisfied only if, for all $k \in \mathcal{S} \setminus \{i\}$,

$$\frac{\|r_j - r_i\|^{-\alpha}}{\|r_j - r_k\|^{-\alpha}} \geq \beta. \quad (7)$$

Using the triangle inequality and (7), we can write

$$\begin{aligned} \|r_k - r_i\| &= \|r_k - r_j - (r_i - r_j)\| \\ &\geq \|r_k - r_j\| - \|r_j - r_i\| \\ &\geq \|r_j - r_i\| (\beta^{\frac{1}{\alpha}} - 1). \end{aligned}$$

Since this must hold for every intended receiver $j \in \mathcal{N}_i$,

$$\begin{aligned} \|r_k - r_i\| &\geq \sup_j \|r_j - r_i\| (\beta^{\frac{1}{\alpha}} - 1) \\ &= s (\beta^{\frac{1}{\alpha}} - 1). \end{aligned} \quad (8)$$

Thus in any given slot, there cannot be more than one transmitter in a ball of radius $(\beta^{\frac{1}{\alpha}} - 1)s$. From our assumptions it is clear that:

- All nodes in the network should transmit at least once to form the required message passing graph.
- Since there cannot be more than one transmitter in a ball of radius $(\beta^{\frac{1}{\alpha}} - 1)s$, any TDMA protocol that constructs the message passing graph requires at least $n|b(0, s(\beta^{\frac{1}{\alpha}} - 1))|$ time slots.
- For $s = \omega(s_c)$, each such ball of radius s has $n|b(0, s)|$ nodes a.a.s.

Hence any feasible TDMA protocol has $n|b(0, s(\beta^{1/\alpha} - 1))| = |b(0, 1)| n s^d (\beta^{1/\alpha} - 1)^d$ slots a.a.s., where $|b(0, 1)|$ is the volume of a d -dimensional unit ball. Choose $C_1 = 0.5|b(0, 1)|$.

In particular, the length of the shortest feasible schedule is also lower bounded by $|b(0,1)|ns^d(\beta^{1/\alpha}-1)^d$, a.a.s. We note that the lower bound is proportional to $(\beta^{\frac{1}{\alpha}}-1)^d$. ■

From the above result, increasing the SIR threshold β - or equivalently, the transmission rate - reduces the spatial density of concurrent transmitters, and hence cannot reduce the mixing time. The effect also becomes more pronounced in higher dimensions. However, higher transmission rates result in smaller packet transmission times. In Section IV we include both these effects in analyzing mixing times.

We now show that the shortest TDMA schedule requires less than $C_2ns^d\beta^{\frac{d}{\alpha}}$ transmissions, for some constant $C_2 > 0$, to construct the desired message passing graph.

Proposition 2. *Consider n nodes distributed uniformly randomly in a d -dimensional torus $\mathcal{T}_d(n)$ lying on $[0,1]^d$. Suppose each node is to send its message to all nodes that are at most $s = \omega(s_c)$ away. Assume a physical model for connectivity, with SIR threshold β . Then for some positive constant C_2 , shortest TDMA schedule that guarantees such connectivity for each node cannot have more than $C_2ns^d\beta^{d/\alpha}$ time slots a.a.s.*

Proof: The proof involves construction of a feasible TDMA schedule whose length is $C_2ns^d\beta^{d/\alpha}$. We prove the result for $d = 2$. The proof can be easily generalized to higher dimensions.

Let $x \triangleq \gamma s$ for some fixed $\gamma > 1$. Consider the lattice \mathbb{L} that consists of points on the scaled integer lattice $x\mathbb{Z}^2$ that also lie on the torus. In other words, $\mathbb{L} = x\mathbb{Z}^2 \cap \mathcal{T}_2(n)$. Partition \mathbb{L} into sublattices as follows:

- $\mathbb{L}_{00} \triangleq \{(ix, jx) \in \mathbb{L} : i \text{ and } j \text{ are even}\}$
- $\mathbb{L}_{01} \triangleq \{(ix, jx) \in \mathbb{L} : i \text{ even, } j \text{ odd}\}$
- $\mathbb{L}_{10} \triangleq \{(ix, jx) \in \mathbb{L} : i \text{ odd, } j \text{ even}\}$
- $\mathbb{L}_{11} \triangleq \{(ix, jx) \in \mathbb{L} : i \text{ and } j \text{ are odd}\}$

With each lattice site $p \in \mathbb{L}$ one can associate the tile $\tau_p = p + [0, x]^2$ that lies within the torus $\mathcal{T}_2(n)$. Denote by \mathbb{T}_{ij} the set of such tiles associated with each of the points in \mathbb{L}_{ij} , $i, j = 0, 1$. For example, $\mathbb{T}_{00} \triangleq \{\tau_p : p \in \mathbb{L}_{00}\}$. Thus $\{\mathbb{T}_{ij}\}$ partition the torus $\mathcal{T}_2(n)$.

The idea behind such a partition is to enable spatial re-use. Consider the following four-phase MAC protocol consisting of phases 00, 01, 10, 11. In phase ij at most one node from each tile in \mathbb{T}_{ij} is allowed to transmit. The protocol ensures that each node transmits exactly once.

The next step is to show that this protocol provides the desired connectivity to each node every $C_2ns^2\beta^{2/\alpha}$ time slots for some positive C_2 . To this end, we first show that the interference at each intended receiver is bounded from above, and can be made smaller than any arbitrary $\beta > 0$ by a suitable choice of γ .

Consider one such transmission in phase 00. Let $\mathcal{S} \in \mathbb{T}_{00}$ be the set of all transmitters. Consider a transmitting node i in tile τ_p where $p = (0, 0)$, i.e., a tile at the origin. To remain

feasible, the protocol must ensure

$$\frac{\|r_j - r_i\|^{-\alpha}}{\sum_{k \in \mathcal{S} \setminus \{i\}} \|r_k - r_j\|^{-\alpha}} \geq \beta. \quad (9)$$

For any i, j, k , it is clear that

$$\begin{aligned} \|r_k - r_j\| &= \|r_k - r_i - (r_j - r_i)\| \\ &\geq \|r_k - r_i\| - \|r_j - r_i\| \\ &\geq \|r_k - r_i\| - \sup_j \|r_j - r_i\| \\ &= \|r_k - r_i\| - s \\ \implies \|r_k - r_j\|^{-\alpha} &\leq (\|r_k - r_i\| - s)^{-\alpha}. \end{aligned}$$

Therefore for a transmitter at r_i , the interference power at any intended receiver j can be upper bounded as

$$\sum_{k \in \mathcal{S} \setminus \{i\}} \|r_k - r_j\|^{-\alpha} \leq \sum_{k \in \mathcal{S} \setminus \{i\}} (\|r_k - r_i\| - s)^{-\alpha}, \quad (10)$$

where the right hand side is independent of j . By the design of the protocol, an interferer k for any intended receiver of the message from i must lie in a tile distinct from $\tau_{(0,0)}$. Moreover, such a tile should lie within \mathbb{T}_{00} ; thus the protocol imposes a lower bound on the minimum distance between any two concurrent transmitters. Using geometrical arguments (see Figure 1), the right hand side of (10) is upper bounded as

$$\begin{aligned} &\sum_{k \in \mathcal{S} \setminus \{i\}} (\|r_k - r_i\| - s)^{-\alpha} \\ &\leq \sum_{l=1}^{\infty} 8l((2l-1)\gamma s - s)^{-\alpha} \\ &= 8s^{-\alpha} \sum_{l=1}^{\infty} l((2l-1)\gamma - 1)^{-\alpha} \\ &\leq 8s^{-\alpha} \left((\gamma-1)^{-\alpha} + \sum_{l=2}^{\infty} l((2l-1)\gamma - \gamma)^{-\alpha} \right) \\ &= 8s^{-\alpha} \left((\gamma-1)^{-\alpha} + 2^{-\alpha}\gamma^{-\alpha} \sum_{l=2}^{\infty} l(l-1)^{-\alpha} \right) \\ &\leq \xi s^{-\alpha} (\gamma-1)^{-\alpha}, \end{aligned} \quad (11)$$

for some fixed $\xi > 0$. The SIR condition (9) is guaranteed to be satisfied at every intended j , if γ is chosen such that

$$\begin{aligned} \frac{s^{-\alpha}}{\xi s^{-\alpha} (\gamma-1)^{-\alpha}} &\geq \beta \\ \implies \gamma &\geq 1 + (\xi\beta)^{\frac{1}{\alpha}}. \end{aligned}$$

For a suitable choice of ξ , we can set $\gamma = 10(\xi\beta)^{\frac{1}{\alpha}}$.

For $s = \omega(\rho_c)$, the number of nodes in each tile is nx^2 a.a.s. Hence as $n \rightarrow \infty$, the protocol constructed requires $4 \times nx^2 = 4n\gamma^2 s^2 = C_2ns^2\beta^{2/\alpha}$ transmissions almost surely to establish the necessary connectivity to each node in the network, where $C_2 = 400\xi^{2/\alpha}$. By optimality, the number of slots T^* in the shortest TDMA schedule cannot exceed this number. Note that the upper bound is proportional to $\beta^{d/\alpha}$. ■

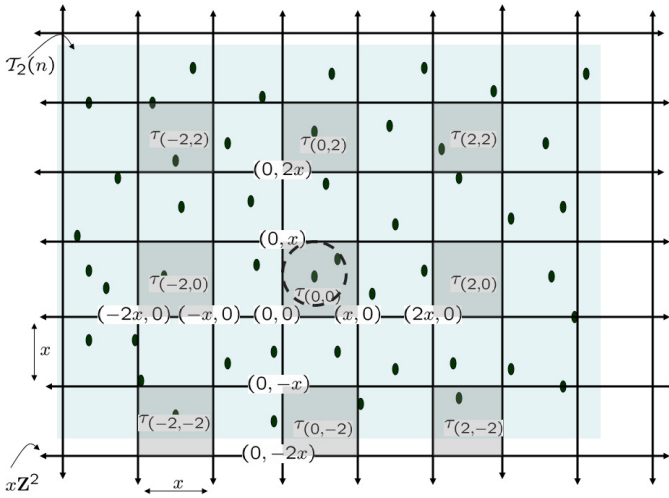


Figure 1: Geometric reasoning underlying the proof of Proposition 2. The location of a typical transmitter in $\tau_{0,0}$ and one of its intended receivers is shown. The nearest interferers and their respective tiles are $\tau_{2,0}, \tau_{2,2}, \tau_{0,2}, \tau_{-2,2}, \tau_{-2,0}, \tau_{-2,-2}, \tau_{0,-2}, \tau_{2,-2}$. The signal power from any of one these interferers at the intended receiver is no larger than that received from the closest interferer allowed by the protocol. The protocol ensures that this nearest distance is no smaller than $x = \gamma s$.

The results from Propositions 1 and 2 lead to the following corollary.

Corollary 3. *If $T^* = T^*(n, s, \beta)$ denotes the length of the shortest TDMA schedule, then a.a.s.:*

- 1) For fixed β , $T^* = \Theta(ns^d)$.
- 2) As $\beta \rightarrow \infty$, $T^* = \Theta(ns^d)\Theta(\beta^{\frac{d}{\alpha}})$.

Proof: Claim 1 is evident from the results of Propositions 1 and 2.

For some constants C_1 and C_2 , we have from Propositions 1 and 2, a.a.s. for large n ,

$$C_1 ns^d (\beta^{1/\alpha} - 1)^d \leq T^* \leq C_2 ns^d \beta^{d/\alpha}.$$

Clearly as $\beta \rightarrow \infty$, $(\beta^{1/\alpha} - 1)^d \sim \beta^{d/\alpha}$. Then the T^* will scale as $T^* = \Theta(ns^d)\Theta(\beta^{d/\alpha})$. ■

If all nodes had independent point-to-point channels between one another, the throughput would have scaled as $\Theta(1)$. For a wireless channel, however, Corollary 3 bounds that the maximum achievable throughput by two decreasing functions of n . It is therefore not surprising that this loss in throughput directly affects the time needed to establish the desired exchange of state variables $\{x_i\}$ which will be discussed in the following section.

IV. CONVERGENCE IN DENSE NETWORKS

A. Mixing Time

We now analyze the asymptotic convergence behavior of the distributed averaging algorithm (2) in a dense network at high transmission rate (large β). To begin with, (4) only counts the *number of iterations* necessary to reach an ϵ -ball around the equilibrium distribution. However, from Corollary 3 even

the optimal TDMA protocol requires $\Theta(ns^d)$ time slots to establish the desired connectivity. Assume for simplicity that the time needed to compute a state update from the received packets is zero. The time taken for each iteration is thus determined only by the time necessary to facilitate the desired exchange of information among the nodes.

Hence the ϵ -mixing time if such an algorithm were to be implemented over a dense wireless network scales as

$$T_{\text{mix}}(W) = \Theta(s^{d-2}n \log n) \quad (12)$$

time slots. From Section II-B, we assume that all the nodes transmit close to channel capacity. Since the channels are all assumed to be Gaussian, the transmission rate $R(\beta)$ can be written as

$$R(\beta) = B \log(1 + \beta) - \delta, \quad (13)$$

where B is the message bandwidth and $\delta > 0$ is an arbitrarily small constant. For fixed packet sizes K , the duration of each time slot is $\Theta(1/R)$. In the light of (12),

$$T_{\text{mix}}(W) \sim s^{d-2} \frac{\beta^{\frac{d}{\alpha}}}{R} n \log n. \quad (14)$$

B. Choice of Interaction Radius

The mixing time in (14) scales polynomially in s for $d > 1$. Interestingly, for $d = 1$, the mixing time scales as the inverse of s . This suggests that increasing the interaction radius can improve the rate of convergence. For $d = 2$, however, the mixing time scales *independently* of s , suggesting that these two effects exactly cancel each other, a rather non-intuitive result. For higher dimensions, the scaling law has a positive exponent in s - implying that the mixing time cannot benefit from increasing s .

This dependence on network dimension can be understood as follows. If the network is one-dimensional, although a transmitter is an isotropic radiator, its effect on the network is seen only along the line $[0, 1]$. Although the throughput provided by the optimal TDMA protocol only scales as $\Theta(n^{-1}s^{-1})$ for a given β from Corollary (3), the mixing time scales as $\Theta(s^{-2} \log n)$, thereby offsetting this loss. In d -dimensions, however, while the throughput scales as $\Theta(n^{-1}s^{-d})$, the mixing time only scales as $\Theta(s^{-2} \log n)$. As a result, improving spatial re-use becomes more important than increasing connectivity.

C. Choice of Transmission Rate

The transmission rate affects the mixing time in two ways. On the one hand, a higher transmission rate reduces the packet transmission time. This is captured by the term R^{-1} in (14). However, a higher transmission rate reduces the spatial re-use possible, which is modeled by the term $(\beta^{1/\alpha} - 1)$ in the lower bound. Using the scaling law for mixing time (14) and the relationship between β and R (13), it is clear that the benefit of smaller packet transmission times is *always* outweighed by reduced spatial re-use for large β .

V. CONCLUSIONS

The convergence behavior of consensus algorithms in dense wireless networks was analyzed using their connection with the mixing time of a Markov chain. We bounded the maximum achievable throughput when the desired connectivity is realized using single-hop transmissions. These two results were used to quantify the scaling behavior of the mixing time with transmission rate, interaction radius, and the SIR threshold. In the scaling limit, higher transmission rates cannot improve convergence speed, since reduced spatial re-use always outweighs benefits of faster transmissions. However, the benefits of an increased interaction radius depend on the network dimension. While increasing the interaction radius can benefit convergence in one-dimensional dense networks, it does not for higher dimensional networks. For two-dimensional dense networks the scaling behavior does not depend on interaction radius.

REFERENCES

- [1] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," in *Proc. IEEE Conf. on Decision and Control*, Dec. 2003.
- [2] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Trans. Info. Theory*, vol. 52 (6), pp. 2506–2530, 2006.
- [3] C. C. Moallemi and B. van Roy, "Consensus propagation," *IEEE Trans. Info. Theory*, vol. 52(11), pp. 4753–4766, Nov. 2006.
- [4] V. Blondel, J. Hendrickx, A. Olshevsky, and J. Tsitsiklis, "Convergence in multiagent coordination, consensus and flocking," in *Proceedings of the 44th IEEE Conference on Decision and Control*, pp 2996–3000, 2005.
- [5] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Proceedings of International Conference on Information Processing in Sensor Networks*, pp 63–70, 2005.
- [6] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. on Automatic Control*, vol. 49(9), pp. 1520–1533, Sep. 2004.
- [7] W. Ren and R. W. Beard, "Consensus seeking in multi-agent systems using dynamically changing interaction topologies," *IEEE Trans. Auto. Control*, vol. 50(5), pp. 655–61, May 2005.
- [8] A. Nedic, A. Olshevsky, A. Ozdaglar, and J. Tsitsiklis, "On distributed averaging algorithms and quantization effects," *IEEE Trans. on Auto. Control (to appear)*, 2009.
- [9] S. Kirti, A. Scaglione, and R. Thomas, "A scalable wireless communication architecture for average consensus," in *46th IEEE Conference on Decision and Control (CDC), New Orleans, LA, USA*, Dec 12-14, 2007.
- [10] F. Fagnani and S. Zampieri, "Average consensus with packet drop communication," *SIAM J. Control Optim.*, vol. 48(1), pp. 102–33, 2009.
- [11] R. Rajagopal and M. J. Wainwright, "Network-based consensus averaging with general noisy channels," tech. rep., UC Berkeley, Dept. of Statistics, Tech. Report 751, May 2008.
- [12] T. C. Aysal, M. E. Yildiz, and A. Scaglione, "Broadcast gossip algorithms," in *Proc. IEEE Information Theory Workshop (ITW)*, pp. 343–347, May 5-9, 2008.
- [13] A. Ozgur, O. Leveque, and D. Tse, "Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks," *IEEE Transactions on Information Theory*, vol. 53(10), pp. 3549 – 3572, Oct. 2007.
- [14] M. H. deGroot, "Reaching a consensus," *Journal of the American Statistical Association*, vol. 69, pp. 118–121, March 1974.
- [15] D. A. Levin, Y. Peres, and E. L. Wilmer, *Markov Chains and Mixing Times*. American Mathematical Society, 2008.
- [16] M. Penrose, *Random Geometric Graphs*. Oxford University Press, 2003.
- [17] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Info. Theory*, vol. 46(2), pp. 388–404, Mar. 2000.
- [18] W. Li, Y. Zhang, and H. Dai, "Location aided fast distributed consensus," *IEEE Trans. Info. Theory*, submitted, 2007.