Power-Delay Analysis of Consensus Algorithms on Wireless Networks with Interference

S. Vanka, V. Gupta and M. Haenggi[*]

November 7, 2008

Abstract

We study the convergence of the average consensus algorithm in wireless networks in the presence of interference. For regular lattices with periodic boundary conditions, we characterize the convergence properties of an optimal MAC protocol that maximizes the speed of convergence on these networks. We extend this analysis to hierarchical networks with a backbone of communication nodes supporting randomly placed sensor nodes. We provide analytical upper and lower bounds for the convergence rate for these networks. Our results show that in an interference-limited scenario the fastest converging interconnection topology for the consensus algorithm crucially depends on the geometry of node placement. In particular, we prove that asymptotically in the number of nodes, increasing the transmit power to allow long range interconnections improves the convergence rate in one-dimensional tori, while it has the opposite effect in higher dimensions.

Index Terms—Consensus algorithms, interference, MAC protocol, wireless networks.

1 Introduction

Consensus in general, and average consensus in particular, has become an area of increasing research focus in recent years (e.g. see [1, 11, 12, 14] and the references therein). Many applications, including distributed estimation [2, 16,

17], motion coordination [13] and load balancing of multiple processes [4], have been analyzed in this framework.

Given n nodes each with a scalar value and a possibly time-varying interconnection graph defined on these nodes, a consensus algorithm specifies the updating rule that every node should follow. The updated value of each node at every time step depends on the value held by itself and its neighbors at the previous time step. Initial results about such algorithms showed that the values held by the nodes converge to a common value provided that the interconnection graphs satisfy some connectivity constraints. Lately, the focus has shifted to analyzing the convergence properties in the face of communication constraints imposed by the channels between the nodes. Thus, effects such as quantization [10], packet erasures [2, 6], additive channel noise [7, 8], and delays [9] have begun to gain attention.

Such works typically assume that the communication channels between each pair of nodes are uncoupled. However, consensus algorithms are often employed over wireless networks, where models with independent channel realizations are not suitable. Wireless channels are inherently coupled due to their broadcast nature and the presence of interference. Moreover, in a wireless network, any two nodes can communicate by spending enough energy. Long range interconnections lead to a smaller graph diameter, but also to decreased spatial re-use. The effect of long-range interconnections on the rate of convergence of the consensus algorithm is thus, not clear. The communication topology in wireless networks thus depends on the network protocols and is, in fact, a design parameter. In this work, we take the first steps towards analyzing the effect of such communication constraints on consensus algorithms and designing the communication parameters for the consensus problem. In particular, we consider the rate of convergence of the average consensus algorithm while explicitly accounting for interference. We analyze the performance of scheduling algorithms that are optimal with respect to the rate of convergence. We also provide an analytical understanding of the impact of transmission power on the rate of convergence.

The paper is organized as follows. We begin by formulating the problem and introducing our notation. We concentrate on networks of nodes that are physically placed on a grid with periodic boundary conditions (Section 3). Section 4 extends these results to hierarchical networks with randomly placed sensor nodes and a regular communication backbone. Some avenues for future work are presented in Section 5.

2 Problem Formulation

2.1 Average Consensus Algorithm

In this paper, we will concentrate exclusively on the average consensus algorithm. Consider n nodes that aim to reach consensus with the final value being the average of their initial scalar values. Denote the value held by the i-th node at time k as $x_i(k)$. Also denote by $x(k)$ the n-dimensional vector obtained by stacking the values of all the nodes in a column vector.

We describe in brief the average consensus algorithm defined with a given interconnection topology among the nodes at every iteration step. The topology can be described by a *consensus graph*, with an edge present between two nodes if and only if they can exchange information. Denote the neighbor set of node i at time k by $\mathcal{N}_i(k)$, where the argument k is included to model dynamic interconnection topologies. An iteration consists of every node i exchanging its state variable $x_i(k)$ with all nodes in $\mathcal{N}_i(k)$. Assuming that all nodes are allowed to simultaneously broadcast their states to their neighbors, this exchange happens in a single packet transmission interval (also referred to as a *time slot* and normalized to 1). Thus the state of the system evolves as

$$x_i(k+1) = x_i(k) - h \sum_{j \in \mathcal{N}_i(k)} (x_i(k) - x_j(k)), \tag{1}$$

where h is a scalar constant designed to ensure convergence of the algorithm. In this case, from the perspective of the consensus algorithm, we say that the *iteration time* is one time slot.

Denote the interconnection graph at time k by $\mathcal{G}(k)$. The system thus evolves according to the discrete time equation

$$x(k+1) = (I - hL(k))\, x(k), \qquad x(0) = x_0, \tag{2}$$

where $L(k)$ denotes the Laplacian matrix of the graph $\mathcal{G}(k)$. It can be shown (see, e.g., [14]) that under proper connectivity assumptions, if the parameter h is small enough, consensus is achieved with each node assuming the average value $x_{av} = \frac{1}{n} \sum_i x_i(0)$. Throughout our presentation, we will assume that h is fixed and has a value $h < \frac{1}{2d_{\max}}$ where $d_{\max}$ is the maximum degree corresponding to any node in the consensus graph over all time. In other words, $d_{\max} = \max_{i,k} |\mathcal{N}_i(k)|$. To ensure that the nodes converge to the average of x_0, it is also essential that

the graph at every time step be balanced. The protocols we consider below will ensure that the graph is symmetric, which satisfies this condition.

The rate of convergence of the value of the nodes is a function of the graph topology. In the case of a static graph topology (i.e., $\mathcal{G}(k) = \mathcal{G}$ for all time k), it can be shown (see, e.g., [5, 14, 15]) that the convergence of the consensus protocol is geometric, with the rate being governed by the second largest eigenvalue modulus (SLEM) of the matrix $I - hL$. In general, a consensus algorithm on a graph with smaller SLEM converges more quickly. If the matrix L is symmetric, its SLEM can be written as its norm restricted to the subspace orthogonal to $\mathbf{1}_n \triangleq [1 \; 1 \; ...1]^*_{1 \times n}$.

However, in practice, a number of transmissions, each occupying a single time slot, may be necessary for this information exchange among these nodes to occur. This scenario is common in wireless networks, where concurrent transmissions in the same frequency band can interfere at a node, and hence may not be decodable. Therefore if each node can receive data from say at most one neighbor at any given time, the exchange of information necessary for iteration k will require at least $1 + \max_i |\mathcal{N}_i(k)|$. This idea is developed further in this paper.

2.2 Communication Protocols

Usual treatments of the average consensus algorithm assume the consensus graph as given. This presumes the existence of a communication channel between any two nodes connected by an edge. In typical applications, nodes communicate over wireless channels. In such situations, any two nodes can potentially communicate by expending enough power or by lowering the transmission rate. Moreover, the wireless medium is inherently multicast. At a receiving node, the effect of such constraints manifests as interference from unintended transmitters. This effect on the average consensus algorithm has not been studied previously.

We consider a situation in which the physical locations of the nodes are given. Every node then decides on the power with which it transmits. This power determines the communication radius of the node according to the relation

$$P = P_0 r_{\mathrm{c}}^{\alpha}, \tag{3}$$

where P_0 is a normalization constant, α is the path-loss exponent (typically

$2 \leq \alpha \leq 5$), P is the transmission power and r_c is the communication radius. All nodes at a distance smaller than r_c from the transmitter can receive the transmitted message.

Similar to the communication radius, we can also define an interference radius r_i. A node at position x can receive a message successfully from a node at position y only if $\|y - x\| < r_c$, and there is no node at position z that is simultaneously transmitting, such that $\|z - x\| < r_i$ (interference constraint). In this paper, for simplicity, we assume $r_c = r_i$. The results can be generalized to other cases.

Given the above condition for successful transmission, we require a medium access control (MAC) protocol for the nodes. We focus on Time Division Multiple Access (TDMA)-based MAC protocols in this paper rather than random access protocols. These protocols assure successful communication by scheduling transmissions in time such that messages do not interfere. They demonstrate better throughput than collision-based MAC protocols, at the expense of greater synchronization and co-ordination requirements among the nodes [18, 19].

2.3 Problem Formulation

The operation of the average consensus protocol can be divided into two phases that are repeated at every update of the node values. In the first phase, the nodes exchange their values through possibly multiple transmissions. We consider each transmission to consume one time slot. The effective communication graph at each update is composed of edges (i, j) such that node j has received the value of node i during the previous communication phase. In the second phase, the nodes update their values according to (1). As in the standard model, this step is assumed to be instantaneous. Therefore, due to multiple transmissions to set up the consensus graph, in our model, the state update does not occur at every time slot. In fact, assuming that each communication phase is completed in T time slots, the kth update can be expressed as

$$x(kT + T) = (I - hL)x(kT). \tag{4}$$

Therefore the effect of finite communication time, possibly due to interference, is to slow down the convergence rate.

We are interested in the following problem: Given a set of nodes at known locations, what is the effect of increasing transmit power on the convergence

rate of the consensus algorithm when the channel-access mechanism accounts for interference? In this context, we characterize the convergence of the consensus algorithm for the optimal MAC protocol that minimizes the number of time slots needed for communication in order to form a desired consensus graph $\mathcal{G}$ (thus maximizing the update rate). We analyze this problem for two physical distributions of the nodes on a torus:

1. A regular grid of sensor nodes.

2. A regular grid of nodes that form a communication network for sensor nodes that are distributed as a binomial point process.

The transmit power at each node determines its neighbors in the consensus algorithm. The periodic boundary condition is chosen for analytical tractability, and our analysis becomes accurate as the number of nodes becomes large. We assume the following:

- We assume equal transmission power for all nodes.

- We limit the transmission policy to be time-invariant.

- At the time of an update of the values of the nodes, we require that the effective communication graph be undirected, i.e., for any two nodes i, j in the network, $j \in \mathcal{N}_i \Leftrightarrow i \in \mathcal{N}_j$. Note that this is slightly stronger than the necessary and sufficient condition for convergence of the average consensus algorithm that the graph be balanced [14].

- We do not assume explicit routing of values through nodes since the consensus algorithm itself incorporates implicit routing and in-network computation.

- We assume half-duplex operation, and further assume that packets that suffer collisions cannot be decoded.

Under these assumptions, we are able to show the following results:

- We characterize the rate of convergence for the optimal MAC scheduling protocol for the average consensus algorithm for tori in n dimensions.

- We show that network geometry plays a key role in identifying the optimum power allocation that maximizes the speed of convergence. In particular, while the convergence rate increases with the transmission power in 1-dimensional tori, the opposite is true in higher dimensions.

- In hierarchical networks, we show that a positive fraction of nodes can always achieve consensus for certain scalings of backbone node density.

In the next section, we begin by studying the convergence properties of MAC protocols that maximize the speed of convergence for a given consensus graph $\mathcal{G}$. We begin by considering nodes placed on a regular grid with periodic boundary conditions.

3 Analysis of a Ring and a 2D Torus

3.1 The 1-D Case: Nodes on a Ring

Consider n nodes numbered $\{0, 1, \ldots n-1\}$ placed uniformly on a circle of radius r centered at the origin, as shown in Figure 1. Suppose that the transmission power is such that every node can transmit information to m of its nearest neighbors on either side. As an example, in Figure 1, $m = 1$. Define P_m, $m \leq \lfloor \frac{n}{2} \rfloor$ as the transmit power that provides a communication radius $r_{\mathrm{c}} = 2r \sin\left(\frac{m\pi}{n}\right)$. Hence

$$P_m \propto \left(2r \sin\left(\frac{m\pi}{n}\right)\right)^{\alpha}, \tag{5}$$

where $\alpha \geq 2$ is the path-loss exponent. As stated above, for simplicity, we will assume that the interference radius $r_{\mathrm{i}} = r_{\mathrm{c}}$.

We note here that an alternative interpretation of this geometry of node placement is to consider the n nodes placed on a regular one-dimensional torus $[0, 1]$ (hereafter called a "1-torus" or $\mathcal{T}_1(n)$). This interpretation is useful in connecting these results with those for higher dimensional tori, that are discussed later. In this case, if choose the location of the first node as the origin, the position of the the kth node is given by $\frac{k}{n}$, $0 \leq k \leq n - 1$, with a periodic boundary condition. In this geometry, the expression for P_m would appear as $P_m \propto \left(\frac{m}{n}\right)^{\alpha}$.

If the wireless channel could support simultaneous transmissions by every node, the system would evolve according to (2), with $I - hL$ being an $n \times n$ circulant matrix with the first row given by

$$\begin{bmatrix} 1 - 2mh & \mathbf{1}_m^* & 0 & 0 & \cdots & 0 & \mathbf{1}_m^* \end{bmatrix}, \tag{6}$$

where

$$\mathbf{1}_m^* = \begin{bmatrix} 1 & 1 & \cdots & 1 \end{bmatrix}_{1 \times m}. \tag{7}$$
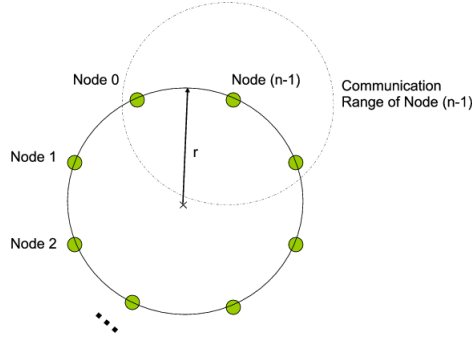
Figure 1: Schematic of nodes placed along a ring.

For future reference, denote by $\mathcal{G}_{1,m}$, $L_{1,m}$ and $F_{1,m}$ the consensus graph, the Laplacian and the update matrix respectively for such a situation. Given the nodes placed on a ring, $\mathcal{G}_{1,m}$ is the consensus graph with the highest connectivity that can be formed for a given P_m, and therefore will have the fastest convergence. The MAC protocol that we propose guarantees that the system evolves according to this matrix. However, the communication phase occurs over multiple steps.

3.1.1 Characterizing the Time for the Communication Phase

We begin by bounding the number of time slots required to form $\mathcal{G}_{1,m}$. Denote the smallest number of time slots used to form $\mathcal{G}_{1,m}$ by $T_1^*(m)$, or more compactly, as T_1^*. Also denote by $\mathcal{P}^*(m)$ the optimal TDMA protocol that forms the graph $\mathcal{G}_{1,m}$ in T_1^* number of steps. Observe that for all $m_1 \leq m_2$, $\mathcal{G}_{1,m_1} \subseteq \mathcal{G}_{1,m_2}$. Then $\mathcal{G}_{1,m}$ can always be formed in at most $T_1^*(m_1) \leq T_1^*(m_2)$ slots. This implies that $m_1 \leq m_2 \implies T_1^*(m_1) \leq T_1^*(m_2)$. In the following, we say a *link* is formed from node v to node u whenever the message from v is successfully decoded at u. Since $\mathcal{G}_{1,m}$ is undirected and balanced, an edge $e \in \mathcal{E}_{1,m}$ connecting v and u is formed iff both v and u form links with each other.

Lemma 1. *Consider the set-up described above, where the consensus graph $\mathcal{G}_{1,m}$ is to be formed in the smallest number of time slots. The optimal TDMA protocol forms $\mathcal{G}_{1,m}$ in the smallest possible number of time slots T_1^* where*

$$2m + 1 \leq T_1^*(m) \leq 4m + 1. \tag{8}$$

Proof. Suppose $\mathcal{G}_{1,m}$ is formed in T time slots. Each node is connected to m nearest neighbours on either side, and the node degree of $\mathcal{G}_{1,m}$ is $2m$. Suppose N_t links in $\mathcal{G}_{1,m}$ are formed in time slot t. Given that there are are $4mn$ links in the graph,

$$\sum_{t=1}^{T} N_t = 4mn. \tag{9}$$

Now suppose that $K(t)$ nodes $\{v_{i_1}, v_{i_2}, \ldots, v_{i_{K(t)}}\}$ transmit in time slot t. Let the respective power allocations be $\{P_{l_1}, P_{l_2}, \ldots, P_{l_{K(t)}}\}$. Clearly, the power allocation satisfies $P_1 \le P_{l_k} < P_{\lfloor \frac{n}{2} \rfloor}$ for $k = 1, 2, \ldots K(t)$.

Assume that the power allocated to node v_{i_k} allows it to broadcast its message to at most the l_k nearest neighbors on either side, or equivalently form at most $2l_k$ links. Since there are at most $2m$ of these links are part of $\mathcal{G}_{1,m}$,

$$N_t \le 2 \sum_{k=1}^{K(t)} \min\left(m, l_k\right) \le 2mK(t), \tag{10}$$

where the second inequality is obtained by choosing $\min(m, l_k) = m$ for all transmitting nodes. To minimize T in (9), we need to maximize N_t for all t.

Observe that any transmission to a neighbor beyond the m nearest neighbors does not contribute to the graph $\mathcal{G}_{1,m}$; this implies that choosing $l_k = m$ is sufficient. This argument holds for any of the $K(t)$ transmitters. Hence each on the transmitting nodes should form m links on either side, or equivalently transmit with power P_m.

Now note that the optimal protocol selects $\{v_{i_1}, v_{i_2}, \ldots, v_{i_{K(t)}}\}$ for all t while meeting this upper bound, and ensuring that T is minimized. As noted above, this minimum value of T has been called T_1^*.

Since the optimal protocol leads to each node transmitting to $m-$nearest neighbors whenever it communicates, for every $2m + 1$ adjacent nodes, only one node can transmit in any slot to avoid interference. As a result, a TDMA schedule cannot have fewer than $2m + 1$ slots. In other words, $T_1^* \ge 2m + 1$. Denote this interval of $2m + 1$ nodes as the transmission interval of the central node, with m nodes on each side that it forms a link with during that time slot.

An upper bound on T_1^* can be obtained by considering the length of a particular TDMA schedule. Consider a protocol in which each transmitter is allocated power P_m. Due to interference constraints, no two transmission intervals can touch overlap each other. Given that there are a total of n nodes uniformly placed on a ring, the maximum number of allowed transmitters in any time step

is

$$K_{\max} = \lfloor \frac{n}{2m+1} \rfloor. \tag{11}$$

The transmission schedule for this MAC protocol is as follows. Consider time slot 1. Suppose some node $v \in \mathcal{V}$ transmits at P_m. We require that all nodes $(2m+1)$ nodes apart should transmit as long as the half-duplex and interference constraints are satisfied. Since the maximum number of simultaneous transmissions possible is $\lfloor \frac{n}{2m+1} \rfloor$, in $2m+1$ time slots, $(2m+1)\lfloor \frac{n}{2m+1} \rfloor$ nodes can transmit. After $2m+1$ time slots,

$$n - (2m+1)\lfloor \frac{n}{2m+1} \rfloor = \operatorname{rem}(n, 2m+1) \leq 2m \tag{12}$$

nodes will not have transmitted with this protocol. So this schedule forms $\mathcal{G}_{1,m}$ in $T_u = (2m+1) + \operatorname{rem}(n, 2m+1) \leq 4m+1$. Hence we conclude $T_1^* \leq T_u \leq 4m+1$. $\square$

We have thus bounded the length of the shortest TDMA schedule that forms the consensus graph $\mathcal{G}_{1,m}$. In other words, we have bounded the smallest time T in the update equation (4).

3.1.2   Bounding the Rate of Convergence

To characterize the fastest convergence possible for a given $\mathcal{G}_{1,m}$ we need to use the above result in conjuction with the spectral properties of $\mathcal{G}_{1,m}$, which are presented below.

Theorem 2. *Consider the problem set-up described above. If the optimal TDMA protocol is used to construct $\mathcal{G}_{1,m}$ for each iteration, the error vector $\epsilon(k) = x(k) - \mathbf{1}_n x_{av}$ converges geometrically to zero with the rate of decay β bounded as*

$$\rho_1^{\frac{1}{2m+1}} \leq \beta \leq \rho_1^{\frac{1}{4m+1}} \tag{13}$$

where

$$\rho_1 = 1 - h(2m+1) + hS_1^{(m,n)} \tag{14}$$
$$S_p^{(m,n)} = \frac{\sin(\frac{(2m+1)\pi p}{n})}{\sin(\frac{\pi p}{n})}, \quad p = 0, 1, \dots n-1.$$

Proof. The consensus graph at each update step is balanced and connected.

10

Thus, the node values converge to the average of their initial values with the decay rate as the modulus of the second largest eigenvalue of $F_{1,m}$ [14]. Denote $e^{-j\frac{2\pi k}{n}}$ by $W_{k,n}$. Since $F_{1,m}$ is circulant, its kth eigenvalue ρ_k is

$$
\begin{aligned}
\rho_k &= 1 - 2mh + h\sum_{l=1}^{m}(W_{k,n} + W_{-k,n}) \\
&= 1 - (2m+1)h + 2h\sum_{l=0}^{m}\cos\left(\frac{2\pi kl}{n}\right) \\
&= 1 - (2m+1)h + hS_k^{(m,n)}, \ k = 0,\ldots,n-1.
\end{aligned}
\tag{15}
$$

It is easy to see $\rho_0 = 1$. The second largest eigenvalue is given by $\rho_1 = \rho_{n-1} < 1$, where ρ_1 was defined in (14). From Lemma 1, $2m+1 \leq T_1^* \leq 4m+1$. Thus the convergence rate is bounded as

$$
\rho_1^{\frac{1}{2m+1}} \leq \beta \leq \rho_1^{\frac{1}{4m+1}}.
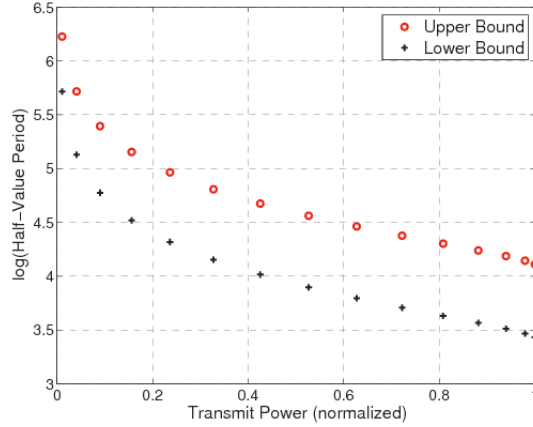\tag{16}
$$

$\square$

Remarks

Figure 2: Variation of the convergence rate with the transmission power for a ring of $n = 31$ nodes.

1. For any given transmission power P_m, we see that the MAC constraints reduce the rate by a factor of T where $2m+1 \leq T \leq 4m+1$.

2. The speed of convergence is an increasing function in m, and hence in P_m. An illustration of this fact is provided in Figure 2. For the purpose of the plot, we show the time taken for the error norm to become half, termed the "half-value period", as a function of transmission power for 31 nodes arranged regularly on a ring of radius 1 unit. We have assumed $\alpha = 2$, and the constant of proportionality in (5) to be unity. For each P_m, we chose $h \propto \frac{1}{2m+1}$. The results are somewhat counter-intuitive since the rate reduction due to a larger number of steps in the communication phase is always compensated by the increase in rate due to higher connectivity. That forming long range communication links would lead to faster convergence even in networks with interference was not evident *a priori*.

3. The effect of increasing the transmission power are the most prominent at small P_m. This can again be seen from Figure 2. If $\theta = p\pi/n$ and $p \ll n$,

$$\sin \theta \approx \theta - \theta^3/3. \tag{17}$$

We use (17) to express the spectral gap SG $\triangleq 1 - \rho_1^{\frac{1}{T}}$ when $m \ll n$ as

$$
\begin{aligned}
\text{SG} &= 1 - \left(1 - h(2m+1) + h\frac{\sin\left(\frac{(2m+1)\pi}{n}\right)}{\sin\left(\frac{\pi}{n}\right)} \right)^{\frac{1}{T}} \\
&\approx \eta m h(m+1)(2m+1)T^{-1}n^{-2}.
\end{aligned}
$$

where $\eta \triangleq \frac{4\pi^2}{3}$. Since $h \propto \frac{1}{2m+1}$ and $T = \Theta(m)$ for the optimal schedule, the spectral gap scales as $\frac{m}{n^2}$.

3.1.3 Nodes on a Two-Dimensional Torus

We now generalize our results to higher dimensional tori. Consider a set $\mathcal{T}_d(n)$ of $n = l^d$ regularly spaced points on a $d-$dimensional torus located at $[0,1]^d$. An example when $d = 2$ is shown in Figure 3. Choose a node as the origin, and label each node using its displacements along each of the d axes (referred to as the *d axial directions* of the torus in this paper). For example, in Figure 3, $n = 9$ and the node $(1,1)$ is located at $r_{11} \equiv r_{(1,1)} \triangleq \left(\frac{1}{3}, \frac{1}{3}\right)$. An alternative interpretation of a toroidal arrangement in the two-dimensional case is shown in Figure 4. Both these interpretations yield similar results in the limiting case of a large torus in which case local distances are not significantly affected by the curvature. We will focus on the former interpretation in this paper.
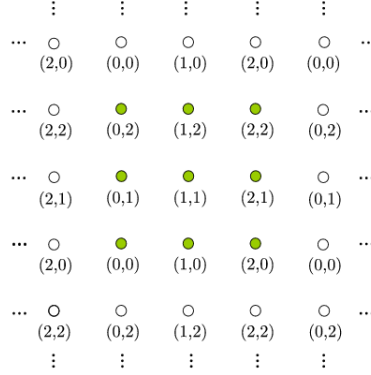
... ○ ○ ○ ○ ○ ...
(2,0) (0,0) (1,0) (2,0) (0,0)

... ○ ● ● ● ○ ...
(2,2) (0,2) (1,2) (2,2) (0,2)

... ○ ● ● ● ○ ...
(2,1) (0,1) (1,1) (2,1) (0,1)

... ○ ● ● ● ○ ...
(2,0) (0,0) (1,0) (2,0) (0,0)

... ○ ○ ○ ○ ○ ...
(2,2) (0,2) (1,2) (2,2) (0,2)

Figure 3: The toroidal lattice $\mathcal{T}_2(9)$. Colored nodes indicate those physically placed in $[0,1]^2$. Node $(0,0)$ represents the node at the origin, with each of the colored nodes (i,j) being placed at $(i/3, j/3)$. Nodes that left unfilled are the image nodes that arise due to the periodic boundary condition.

Suppose all nodes on a torus $\mathcal{T}_d(n)$ participate in an average consensus algorithm of the form (2) with a power allocation of P_m per node. As an extension to the 1-D case, we assume that with power P_m, the set of all reachable nodes will lie in a *communication sphere* of radius $\frac{m}{n^{1/d}}$ $(m < \lfloor \frac{l}{2} \rfloor)$ centered at the transmitter. We now have the desired consensus graph $\mathcal{G}_{d,m} = (\mathcal{V}, \mathcal{E}_{d,m})$ where the vertex set

$$\mathcal{V} = \{0, 1, \ldots, l-1\}^d, \tag{18}$$

is the set of all points in $\mathcal{T}_2(n)$. The edge set $\mathcal{E}_{d,m}$ is formed by connecting every node to all nodes on the torus that are within its communication sphere:

$$\mathcal{E}_{d,m} = \{\{v,u\} \in \mathcal{V}^2 : v \neq u, \ell_2(r_v, r_u) \leq \frac{m}{n^{1/d}}\}, \tag{19}$$

where $\ell_2(x,y)$ is the Euclidean norm between x and y.

In keeping with the notation developed for the one-dimensional case, we will denote the Laplacian and the update matrix for $\mathcal{G}_{d,m}$ by $L_{d,m}$ and $F_{d,m} \triangleq I - hL_{d,m}$ respectively. Assuming as before that each transmission occupies one time slot, we now study the convergence properties of the optimal MAC protocol that will form $\mathcal{G}_{d,m}$ in the smallest number of time slots. In this paper, we set $d = 2$; the results can be generalized to tori of higher dimensions.

Using similar arguments as in Section 3.1.1, we can show that the optimal TDMA protocol assigns a power P_m to each chosen transmitting node i in any

13

time slot. Observe that at most one node can transmit in a time slot inside the volume occupied by one communication sphere. Denote by $T_2^*(m)$, or more compactly by T_2^* the number of time slots required by an optimal schedule to construct $\mathcal{G}_{2,m}$. The optimal MAC schedule places the *maximum* number $N_2^*(t)$ of non-intersecting spheres on the torus in every time slot $t = 1, 2, \cdots, T_2^*$.

Note that increased network dimensionality plays an important role in finding N_2^* (and consequently, T_2^*). This makes the problem of analytically finding T_2^* non-trivial. This effect is illustrated in Figure 5 for a 2-torus of $n = 25$ nodes and $m = 1$. The transmitters are chosen from the entire two-dimensional lattice. With power P_1, each node can reach its 4 nearest neighbors as shown, with 5 nodes in each communication sphere. Since there can be at most one transmission within this sphere, forming the consensus graph $\mathcal{G}_{2,1}$ requires at least $4 + 1 = 5$ slots. Figure 5 shows the optimal transmitting set in the first time slot.

As before, we begin by characterizing the length T_2^* of the shortest TDMA schedule that constructs $\mathcal{G}_{2,m}$. Thereafter, we exploit the properties of the consensus algorithm along with optimality of the MAC protocol and constraints imposed by our problem to bound the convergence rates for a 2-torus.
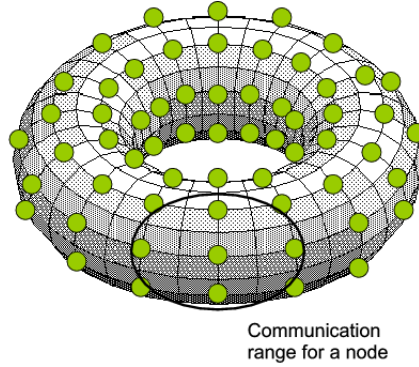

Communication
range for a node

Figure 4: Schematic of nodes placed along a 2-dimensional torus. The periodic square grid being considered can be considered a limiting case of a large torus, so that the effect of its curvature on small distances is not important.

As before, define P_m to be the transmit power that enables a node to form error-free links with m neighbors in the axial directions. Given that there are
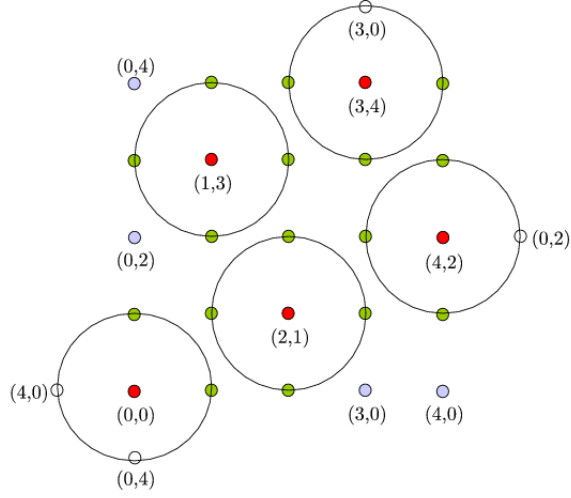
Figure 5: Effect of geometry in sphere packing, shown for $\mathcal{T}_2(25)$ and $m = 1$. The transmitters are shown in red (e.g., (0,0)). Note that the dimensionality is exploited to allow more concurrent transmissions. The nodes in light-blue (e.g., (4,0)) are covered by transmitters via their images (shown unfilled).

$\sqrt{n}$ nodes in either of the axial directions,

$$P_m \propto \left(\frac{m}{\sqrt{n}} \right)^\alpha, \tag{20}$$

where $\alpha \geq 2$ is the path-loss exponent. Define Π to be the elementary circulant matrix having as the first row $\begin{bmatrix} 0 & 1 & 0 \cdots & 0 & 0 \end{bmatrix}_{1 \times m}$. Then the update matrix $F_{2,m}$ can be written as

$$F_{2,m} = \sum_{ij} a_{ij} \Pi^i \otimes \Pi^j \tag{21}$$

where $\otimes$ denotes the Kronecker product and

$$a_{ij} = h \, \forall i, j \, \text{s.t.} \, i^2 + j^2 \leq m^2 \tag{22}$$

for $(i, j) \neq (0, 0)$ and $a_{00} = 1 - h|\{(i, j) \in \mathbb{Z}^2 : i^2 + j^2 \leq m^2\}|$. We now bound the number of time slots required to form $\mathcal{G}_{2,m}$.

Lemma 3. *If each node transmits at power P_m and the optimal schedule over the 2-torus constructs $\mathcal{G}_{2,m}$ in T_2^* time slots, for $1 < m < \lfloor l/2 \rfloor$, T_2^* always*

satisfies

$$T_l \leq T_2^* \leq T_u \tag{23}$$

where

$$T_l = 2m^2 + 2m + 1 \tag{24}$$

and

$$T_u = 16m^2 + 8m + 1. \tag{25}$$

Proof. Using similar arguments as in Lemma 1 it is easy to show that power allocation for any node should be at least P_m. Define a *feasible* TDMA schedule for a power allocation P_m per node as one that which constructs $\mathcal{G}_{2,m}$ while satisfying the half-duplex and interference constraints described in Section 2.

Without loss of generality, suppose node $(0,0)$ transmits in the first time slot with power P_m. Let $\mathcal{B}$ denote its communication disc, i.e., a circle of radius $\frac{m}{\sqrt{n}}$ centered at $(0,0)$. From the definition of a feasible schedule, other than $(0,0)$, no other node inside $\mathcal{B}$ can transmit at this time. This means a feasible schedule can allow at most one transmission inside $\mathcal{B}$. Therefore any feasible schedule needs at least as many time slots as the number of nodes inside a sphere of radius $\frac{m}{\sqrt{n}}$ centered at $(0,0)$, in order to form $\mathcal{G}_{2,m}$. In other words, a feasible schedule has *at least* $|\mathcal{B} \cap \mathcal{T}_2(n)|$ time slots. This involves counting the number of nodes of a square grid that fall inside the circle we are considering, which can be cumbersome. For our purposes, it will suffice to find a lower bound for this value by counting only those nodes that fall in a suitably chosen square $\mathcal{U}$ circumscribed by $\mathcal{B}$.

Consider a square $\mathcal{U}$ having as its vertices, the nodes $(m,0)$, $(0,m)$, $(-m,0)$ and $(0,-m)$. Clearly $\mathcal{U} \subseteq \mathcal{B}$. Each of its diagonals contains $2m+1$ nodes. The line segments joining nodes $(-m+|k|,k)$ and $(m-|k|,k)$ for $k = -m, -m+1, \ldots, m-1, m$ are all parallel to the segment joining $(-m,0)$ and $(m,0)$. Each such segment has exactly $2m+1-2|k|$ nodes. Counting all these nodes yields the number of nodes in $\mathcal{B}$:

$$
\begin{aligned}
|\mathcal{U} \cap \mathcal{T}_2(n)| &= \sum_{k=-m}^{m} (2m+1) - 2|k| \\
&= 2m+1 + 2\sum_{k=1}^{m} 2(m-k) \\
&= 2m^2 + 2m + 1.
\end{aligned}
$$

We can thus choose $T_l = 2m^2 + 2m + 1$ as the lower bound on the length of any feasible schedule that forms $\mathcal{G}_{2,m}$. In particular, T_2^* is the length of the shortest feasible schedule. Consequently, $T_2^* \geq T_l$.

If T_u is number of time slots taken by any feasible schedule to form $\mathcal{G}_{2,m}$, $T_2^* \leq T_u$. Consider the following schedule: in the first time slot, choose $(0,0)$ as a transmitter, and schedule nodes $(0 + p(2m+1), 0 + q(2m+1))$ for $p, q = 1, \cdots, \lfloor \frac{l}{2m+1} \rfloor$ to transmit. In other words, we attempt to tile the torus with squares of side $\frac{2m+1}{\sqrt{n}}$. Clearly this is feasible, since each node lies in at most one communication sphere. In each subsequent time slot, repeat this process by choosing some other node (i,j) inside the square of side $(2m+1)/\sqrt{n}$ centered at the origin and schedule nodes $(i+p(2m+1), j+q(2m+1))$ for transmission. Repeat this process until all the $(2m+1)^2$ nodes in this square have been chosen once. Using arguments similar to those used in Lemma 1, a maximum of $\lfloor \frac{l}{(2m+1)} \rfloor^2$ simultaneous transmissions can be scheduled per time slot. After $(2m+1)^2$ time slots,

$$
n - \lfloor \frac{l}{(2m+1)} \rfloor^2 (2m+1)^2
$$
$$
= 2\lfloor \frac{l}{2m+1} \rfloor (2m+1) \operatorname{rem}(l, (2m+1))
$$
$$
+ (\operatorname{rem}(l, 2m+1))^2
$$

nodes are yet to transmit. In the first term, we can schedule $\lfloor l/(2m+1) \rfloor$ nodes in each of the $2\operatorname{rem}(l, 2m+1) \leq 4m$ "rows", that require at most $4m(2m+1)$ additional time slots. Scheduling one node per time slot, all the remaining $(\operatorname{rem}(l, 2m+1))^2$ nodes can transmit in at most $4m^2$ time slots. Therefore the schedule constructs $\mathcal{G}_{2,m}$ in $(2m+1)^2 + 4m(2m+1) + 4m^2 = 16m^2 + 8m + 1$ time slots. Thus we conclude that $T_2^* \leq 16m^2 + 8m + 1$. □

As compared to the 1-torus, the optimal schedule for a 2-torus is bounded by two quadratic terms. This arises due to imposing interference constraints to the given geometry of node placement. As we shall see, this quadratic - rather than linear - dependence on m is the key to understanding the effect of transmit power on convergence behavior.

3.1.4   Bounding the Rate of Convergence

Finding the eigenvalues of $F_{2,m}$ also seems intractable. This difficultly arises from the fact that the nodes that can receive data from a particular node are

specified through circular disks (or spheres). While in a one-dimension such disks can cover the entire ring, such coverage is not possible in higher dimensions. For our purpose, we lower and upper bound such discs by squares of suitable side length that cover the entire region. The idea is illustrated in Figure 6. To this end, we begin with the following preliminary result.

Lemma 4. *Let $\hat{\mathcal{G}}_{2,m}$ be the consensus graph formed over $\mathcal{T}_2(n)$ by placing edges between each node (i,j) with all other nodes $(k,l) \neq (i,j)$ satisfying*

$$\ell_\infty(r_{ij}, r_{kl}) \leq \frac{m}{\sqrt{n}}, \qquad 1 \leq m < \lfloor \frac{\sqrt{n}}{2} \rfloor. \tag{26}$$

where ℓ_∞ denotes the $\infty-$norm. Also denote the Laplacian of $\hat{\mathcal{G}}_{2,m}$ by $\hat{L}_{2,m}$ and its maximum degree by d_{max} and the update matrix $\hat{F}_{2,m} = I - h\hat{L}_{2,m}$ for some $0 \leq h \leq \frac{1}{2d_{max}}$. Then the eigenvalues of the $\hat{F}_{2,m}$ are

$$\lambda_{a,b} = 1 - h(2m+1)^2 + hS_a^{(m,l)}S_b^{(m,l)}, \tag{27}$$

where, as defined above $S_a^{(m,l)} = \frac{\sin\left(\frac{(2m+1)\pi a}{l}\right)}{\sin\left(\frac{\pi a}{l}\right)}$, $a = 0, 1, \ldots l - 1$.

Proof. To begin, note that the Laplacian matrix $\hat{L}_{2,m}$ of $\hat{\mathcal{G}}_{2,m}$ is an $n \times n$ block circulant matrix with each of its l block rows being

$$\begin{bmatrix} A_0 & B_m & 0 & 0 & \cdots & 0 & B_m \end{bmatrix}_{l \times n} \tag{28}$$

where $B_m \triangleq \mathbf{1}_m^* \otimes A_1$. Here the $l \times l$ matrices A_0 and A_1 are also circulant, their first row being an $l \times 1$ matrix $\begin{bmatrix} d_k & -\mathbf{1}_m^* & 0 & 0 & \cdots & 0 & -\mathbf{1}_m^* \end{bmatrix}$, with $d_0 = (2m+1)^2 - 1$ and $d_1 = -1$. With this structure, we can use the properties of block circulant matrices and the fact that A_k's are themselves circulant (and consequently share the same eigenvectors) to compute the eigenvalues of $\hat{L}_{2,m}$ as

$$\mu_{r,s} = \sum_{t=0}^{l-1} \eta_{r,t} e^{-j\frac{2\pi st}{l}} \tag{29}$$

where $\eta_{r,t}$ is the rth eigenvalue of A_t $\forall r, s = 0, 1, \ldots, l - 1$. Using the 1-torus result from equation (15) for $\eta_{r,t}$ and simplifying, the eigenvalues of $\hat{F}_{2,m} = I - h\hat{L}_{2,m}$ are

$$\lambda_{a,b} = 1 - h(2m+1)^2 + hS_a^{(m,\sqrt{n})}S_b^{(m,\sqrt{n})} \tag{30}$$

which is the desired result. $\qquad\square$

18

We are now in a position to bound the rate of decay for the case of the torus. This is presented in the form of a theorem below.

Theorem 5. *Consider a consensus algorithm of the form (2) on $\mathcal{G}_{2,m}$. If each node transmits at P_m for $1 \leq m < \lfloor \frac{\sqrt{n}}{2} \rfloor$, the rate of convergence β of an optimal MAC schedule on $\mathcal{G}_{2,m}$ that drives $\delta(k) = x(k) - \mathbf{1}_n x_{av}$ to zero is bounded as*

$$\lambda_1^{\frac{1}{2m(m+1)}} < \beta < \tilde{\lambda}_1^{\frac{1}{16m^2+8m+1}} \tag{31}$$

where $\tilde{m} \triangleq \lfloor \frac{m}{\sqrt{2}} \rfloor$ and

$$\begin{aligned}
\lambda_1 &= \left(1 - h(2m+1)^2 + h(2m+1)S_1^{(m,\sqrt{n})} \right) \\
\tilde{\lambda}_1 &= \left(1 - h(2\tilde{m}+1)^2 + h(2\tilde{m}+1)S_1^{(\tilde{m},\sqrt{n})} \right)
\end{aligned}$$

Proof. Consider undirected graphs $\mathcal{G}_{\text{sub}}, \mathcal{G}$ and $\mathcal{G}_{\text{sup}}$ with a common vertex set $\mathcal{V}$ and edge sets $\mathcal{E}_{\text{sub}} \subseteq \mathcal{E} \subseteq \mathcal{E}_{\text{sup}}$, with the same edge weights. Call $\mathcal{G}$ the nominal graph and $\mathcal{G}_{\text{sub}}$ and $\mathcal{G}_{\text{sup}}$ as sub- and super-graphs of $\mathcal{G}$. Suppose all the graphs satisfy conditions to reach average consensus. The proof rests on two simple but important facts:

Fact 1: If iterations are done at the same rate and all the edge weights positive, the consensus algorithm on a graph cannot be slower than that on its sub-graphs. This follows from the fact that the second largest eigenvalue always decreases if edges are added to a graph (*cf.* [20, Thm. 3.2]).

Fact 2: Suppose the iterations on these graphs occur every T_u, T_2^* and T_l time slots respectively, with $T_u \geq T_2^* \geq T_l$. Then the graphs on which the consensus algorithm converges the fastest and slowest are still $\mathcal{G}_{\text{sup}}$ and $\mathcal{G}_{\text{sub}}$ respectively. In what follows, we will first define the nominal consensus graph $\mathcal{G}$ and choose appropriate super- and sub-graphs $\mathcal{G}_{\text{sup}}$ and $\mathcal{G}_{\text{sub}}$. We will then use the results from Lemma 3 to obtain the values of T_l, T_2^* and T_u.

Consider the graph $\hat{\mathcal{G}}_{2,m}$ as defined in Lemma 4. Note that

$$\ell_2(r_{ij}, r_{kl}) \leq c \implies \ell_\infty(r_{ij}, r_{kl}) \leq c. \tag{32}$$

Choosing $c = \frac{m}{\sqrt{n}}$, it is easy to see that every edge in $\mathcal{E}_{2,m}$ is present in $\hat{\mathcal{E}}_{2,m}$. It follows that $\hat{\mathcal{G}}_{2,m} \supseteq \mathcal{G}_{2,m}$. We now have a super-graph of $\mathcal{G}_{2,m}$. This is illustrated in Figure 6 for $n = 25, m = 2$.
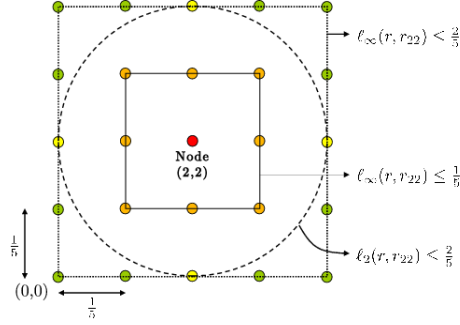
19

Figure 6: Edges to $(2,2)$ in $\mathcal{T}_2(25)$ in $\hat{\mathcal{G}}_{2,\tilde{m}}$, $\mathcal{G}_{2,m}$ and $\hat{\mathcal{G}}_{2,m}$ for $m = 2$. For nodes shown in orange (e.g. (2,3)), edges between (2,2) and these nodes exist in all these graphs. For nodes colored yellow (e.g. (4,2)), such edges exist only in $\mathcal{G}_{2,m}$ and $\hat{\mathcal{G}}_{2,m}$. For the green nodes (e.g. node (0,0)), such edges exist only in $\hat{\mathcal{G}}_{2,m}$.

Define $\tilde{m} = \lfloor \frac{m}{\sqrt{2}} \rfloor$ and form a graph $\hat{\mathcal{G}}_{2,\tilde{m}} = (\mathcal{V}, \hat{\mathcal{E}}_{2,\tilde{m}})$. It is easy to see that

$$
\begin{aligned}
\ell_\infty(r_{ij}, r_{kl}) &\leq \lfloor \frac{\tilde{m}}{\sqrt{n}} \rfloor \ell_2(r_{ij}, r_{kl}) \\
\implies \ell_2(r_{ij}, r_{kl}) &\leq \sqrt{2} \lfloor \frac{\tilde{m}}{\sqrt{n}} \rfloor \\
&\leq \lfloor \frac{\sqrt{2}\tilde{m}}{\sqrt{n}} \rfloor \leq \frac{m}{\sqrt{n}}.
\end{aligned}
$$

Therefore $\hat{\mathcal{G}}_{2,\tilde{m}} \subseteq \mathcal{G}_{2,m}$.

The differences in connnectivity in the graphs $\hat{\mathcal{G}}_{2,\tilde{m}}$, $\mathcal{G}_{2,m}$ and $\hat{\mathcal{G}}_{2,m}$ are illustrated in Figure 6, for a torus of $n = 25$ nodes and $m = 2$. Here $\tilde{m} = \lfloor 2/\sqrt{2} \rfloor = 1$. Consider now the nodes connected to node $(2,2)$. For $\hat{\mathcal{G}}_{2,1}$ over this torus, edges are placed between nodes $(2,2)$ and (k,l) iff $\ell_\infty(r_{22}, r) \leq \frac{1}{5}$. In Figure 6, these are the nodes on or inside the inner-most square centered at $(2,2)$. In $\mathcal{G}_{2,m}$ such edges are placed between nodes that are within a communication disk of radius $\frac{2}{5}$, i.e., when $\ell_2(r_{22}, r) \leq \frac{2}{5}$. These nodes comprise of the nodes shown on or inside the circle. In the case of $\hat{\mathcal{G}}_{2,2}$, all nodes on the torus satisfy $\ell_\infty(r_{22}, r) \leq \frac{2}{5}$. This comprises of all nodes inside the outer square.

The consensus graphs $\hat{\mathcal{G}}_{2,m}$ and $\mathcal{G}_{2,m}$ correspond to update matrices $\hat{F}_{2,m}$ and $F_{2,m}$ respectively for all m. We now use the results for eigenvalues of the Laplacian $\hat{L}_{2,m}$, derived in Lemma 4. The rate of convergence on $\hat{\mathcal{G}}_{2,m}$ is a function of the second largest eigenvalue modulus, that is obtained setting

$a = 0$ and $b = 1$ in (30):

$$\lambda_{0,1} = 1 - h(2m+1)^2 + h(2m+1)S_1^{(m,l)} \triangleq \lambda_1. \tag{33}$$

Making the substitution $m \leftarrow \tilde{m}$ gives

$$\tilde{\lambda}_1 \triangleq 1 - h(2\tilde{m}+1)^2 + h(2\tilde{m}+1)S_1^{(\tilde{m},l)}. \tag{34}$$

which is the corresponding value for $\hat{\mathcal{G}}_{2,\tilde{m}}$.

Consider now a consensus algorithm on $\mathcal{G}_{\text{sub}} = \hat{\mathcal{G}}_{2,\tilde{m}}$, $\mathcal{G} = \mathcal{G}_{2,m}$ and $\mathcal{G}_{\text{sup}} = \hat{\mathcal{G}}_{2,m}$. All the graphs are given equal edge weights h, determined by the maximum node degree in $\hat{\mathcal{G}}_{2,m}$. Suppose the optimal MAC schedule on the torus forms $\mathcal{G}_{2,m}$ in T_2^* time slots. In other words, an iteration on $\mathcal{G}_{2,m}$ occurs every T_2^* time slots. Furthermore, let each iteration on $\hat{\mathcal{G}}_{2,m}$ and $\hat{\mathcal{G}}_{2,\tilde{m}}$ occur be in T_l and T_u time slots respectively, where $T_l = 2m^2 + 2m + 1$ and $T_u = 16m^2 + 8m + 1$. From Lemma 3 we know that $T_u \geq T_2^* \geq T_l$. Since we now know the second largest eigenvalue modulus of these graphs $\hat{\mathcal{G}}_{2,m}$ and $\hat{\mathcal{G}}_{2,\tilde{m}}$, the convergence rates of these graphs are, respectively, λ_1^{1/T_l} and $\tilde{\lambda}_1^{1/T_u}$. The bound on the convergence rate β

$$\lambda_1^{1/T_l} < \beta < \tilde{\lambda}_1^{1/T_u} \tag{35}$$

of the $\mathcal{G}_{2,m}$ that is formed using the optimal schedule now follows from Fact 2. $\square$

To understand the effect of higher transmit power on the convergence rate in 2-tori, first simplify simplifying the expressions for λ_1 and $\tilde{\lambda}_1$ in Theorem 5 using $h = \gamma/(2m+1)^2$, $0 < \gamma < 1$:

$$\lambda_1 = 1 - \gamma + \gamma \frac{\sin\left((2m+1)\pi/\sqrt{n}\right)}{(2m+1)\sin\left(\pi/\sqrt{n}\right)}$$

$$\tilde{\lambda}_1 = 1 - \gamma \frac{(2\tilde{m}+1)^2}{(2m+1)^2}$$

$$+ \gamma \frac{(2\tilde{m}+1)}{(2m+1)^2} \frac{\sin\left((2\tilde{m}+1)\pi/\sqrt{n}\right)}{\sin\left(\pi/\sqrt{n}\right)}.$$

Comparing this to the 1-torus case with $h = \gamma/(2m+1)$,

$$\rho_1 = 1 - \gamma + \gamma \frac{\sin\left((2m+1)\pi/n\right)}{(2m+1)\sin\left(\pi/n\right)}. \tag{36}$$
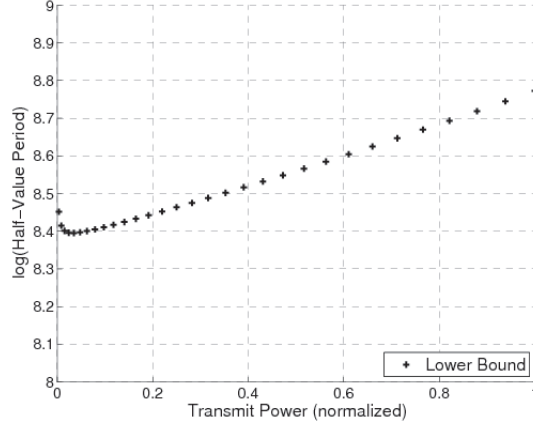
21

Figure 7: Variation of convergence rate with transmission power on a 2-torus.

Clearly λ_1 is of the same form as ρ_1 in Theorem 2, except for the $\sqrt{n}$. For large m, the behaviour of $\tilde{\lambda}_1$ will also be similar to λ_1.

However, there is a significant difference between the two cases in the length of the optimal schedule. This length was shown to be $\Theta(m)$ for the 1-torus, and $\Theta(m^2)$ for the 2-torus. This means that the effect of interference depends on the geometry of node placement. High transmit powers reduce available network throughput by causing more interference. In TDMA-scheduling MAC protocols, this effect is reflected in longer schedules. For a 1-torus, this is still offset by the resultant long-range connections. However, this is no longer true for higher dimensions. In Figure 7, we plot the bounds for β obtained from Theorem 5 for a 2-torus with $n = 4096$ nodes arranged as a 64×64 toroidal lattice. Observe that the convergence rate *worsens* with increasing transmit power. This is surprising when compared to the 1-torus, where the result is the opposite.

3.1.5   Tori in Arbitary Dimensions

The results in Lemma 3 can be extended to higher dimensional grids with toroidal boundary conditions. Indeed using similar arguments for power P_m the optimal schedule for a $d-$dimensional torus cannot be shorter than $|\mathcal{U} \cap \mathcal{T}_d(n)|$ where $\mathcal{U}$ is a polyhedron that can be circumscribed by the sphere of radius $m/\left(n^{1/d}\right)$. Similarly, to find the upper bound one can generalize the schedule described in Lemma 1 that was used to find an upper bound. It can be shown that the optimal schedule will be $\Theta(m^d)$.

The results in Lemma 4 can also be generalized to $d-$dimensions as $\lambda_1 = 1 - h(2m+1)^d + h(2m+1)^{d-1}S_1^{(m,l)}$ for the lower bound and choosing $\tilde{m} = \lfloor m/\sqrt{d} \rfloor$ to find $\tilde{\lambda}_1$ for an upper bound on the convergence rate. Thus the convergence rate increases with transmit power in geometries having dimension 2 or more.

4   Hierarchical networks

Since calculation of the rates of convergence for average consensus for arbitrary graphs is not possible even without the interference constraints, we do not expect to be able to extend our results for arbitrary graphs. In this section, we consider another useful class of graphs that allow us to state analytical results. We consider a variation on the random geometric graphs by adding a backbone of dedicated (long-distance) communication nodes. Thus we consider a hierarchical network with N sensing nodes uniformly randomly placed on a torus in $[0,1]^d$, and K^d identical regularly spaced *backbone* nodes on the torus, as shown in Figure 8 for $d = 1$.
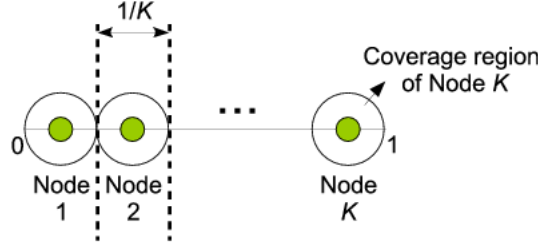

Figure 8: Schematic of backbone nodes placed along a 1-dimensional torus. The ith backbone node is placed at $r_i = \left(i - \frac{1}{2}\right)\frac{1}{K}$ from the origin, for $i = 1, 2, \ldots, K$.

We assume that the backbone nodes do not participate in sensing, and only communicate with each other in the network. Each backbone node has a fixed exclusive region of coverage, i.e., it (alone) collects data from all the sensing nodes within a sphere of radius $r = 1/2K$. Initially, each backbone node collects and averages the data from all the sensing nodes in its region of coverage. All the backbone nodes now run an average consensus algorithm among themselves with their respective averaged data as initial values. It is possible to pass on this (global) average back to all the nodes within their coverage regions in $O(1)$ steps. Therefore for analyzing the rate of convergence, it is sufficient to analyze

the time taken for collecting data by the backbone nodes and the time taken to reach consensus among these nodes.

We begin by characterizing the number of sensors reporting to each backbone node.

Lemma 6. *For large N and if the number of backbone nodes scales as $o(\frac{N}{\log N})$ w.h.p. the number of sensors per coverage area is $n = \frac{N\pi^{d/2}}{\Gamma(1+d/2)(2K)^d}$, where K is the number of nodes per dimension.*

Proof. This can be proved by a a variation of the argument used in the theory of random geometric graphs to show regularity. Consider a sphere $\mathcal{S}$ of radius r centered at point P on the torus. Marking the points $1, 2, \ldots, N$, we can associate with each point k, a random variable X_k $(1 \le k \le N)$ be defined as:

$$X_k = \mathbf{1}_{\mathcal{S}}(k) \tag{37}$$

where $\mathbf{1}(.)$ is the indicator function. That is, $\mathbf{1}_{\mathcal{S}}(k) = 1$ if $k \in \mathcal{S}$, and 0 otherwise. Since the sensing nodes are placed on the torus uniformly and independently of each other, $\{X_k\}$ are iid Bernoulli with success probability

$$p = \frac{\text{Vol. of sphere}}{\text{Vol. of torus}} = \frac{\pi^{d/2}}{\Gamma(1+d/2)(2K)^d}. \tag{38}$$

The number of sensors inside the sphere is thus a binomial random variable

$$\sum_{k=1}^{N} X_k \triangleq Y, \tag{39}$$

with $\mu \triangleq \mathbb{E}Y = Np$. We can now use the Chernoff bound:

$$\mathbb{P}(|Y - \mu| > \mu\delta) \le 2\exp\left(-\frac{\mu\delta^2}{2}\right). \tag{40}$$

For $0 < \delta < 1$, since the number of nodes $K^d = o(\frac{N}{\log N})$ we can always choose

$$\delta = \sqrt{\frac{4\log N}{N} \cdot \gamma K^d} = o(1), \ N \to \infty \tag{41}$$

for some $\gamma > 0$. Plugging in this value into the bound, we see that

$$\mathbb{P}(Y \notin [\mu(1-\delta), \mu(1+\delta)]) \le \frac{1}{N^{2\eta\gamma}} \le \frac{1}{N^3}, \tag{42}$$

24

where $\eta = \frac{\pi^{d/2}}{2^d \Gamma(1+d/2)}$ and γ is chosen such that $\eta\gamma > 2$. So w.h.p.,

$$Y = Np(1 \pm o(1)). \tag{43}$$

Thus, the probability that any coverage area does not have $n(1\pm o(1))$ nodes is

$$\mathbb{P}(\cup_k \{Y_k \notin [\mu(1-\delta), \mu(1+\delta)]\})$$
$$\leq \cup_k \mathbb{P}(Y_k \notin [\mu(1-\delta), \mu(1+\delta)])$$
$$= K^d \frac{1}{N^3} < N \frac{1}{N^3} = \frac{1}{N^2}$$

where we use the union bound. The result now follows readily. $\square$

Since each of the K^d backbone nodes has n sensors in its coverage region w.h.p. when N is large, a total of nK^d sensing nodes will be covered by the backbone network. Therefore for large N, it is always possible to achieve consensus over a positive fraction

$$\kappa = \frac{nK^d}{N} = \frac{\pi^{d/2}}{2^d \Gamma(1+d/2)} \tag{44}$$

of the sensing nodes, independent of N and K.

We will now study specific cases for $d = 1$ and 2. To begin, we note that for $d = 1$,

$$\kappa_1 = \frac{\pi^{1/2}}{2\Gamma(3/2)} = 1, \tag{45}$$

and for $d = 2$,

$$\kappa_2 = \frac{\pi}{4\Gamma(2)} = \frac{\pi}{4} \approx 78.5\%. \tag{46}$$

As stated above, we assume that each backbone node collects data by polling all the sensing nodes in its coverage region. This is done in parallel over all backbone nodes, by a suitable choice of transmit power. Assuming each node transmits in one time slot, we need n time slots to initialize the consensus algorithm, where w.h.p.

$$n = \frac{\alpha N}{K^d}. \tag{47}$$

During the consensus phase, the network topology is identical to the (regular) ring and torus topologies discussed previously. As before, let each node transmit

at fixed power P_m to reach $m \le \lfloor K/2 \rfloor$ neighbors per direction per dimension. Using the results in Theorems 2 and 5, we obtain that for $K^d = o\left(\frac{N}{\log N}\right)$, w.h.p.,

$$\left(1 - h(2m+1) + hS_1^{(m,K)}\right)^{\frac{1}{4m+1}} \ge \beta$$
$$\ge \left(1 - h(2m+1) + hS_1^{(m,K)}\right)^{\frac{1}{2m+1}}$$

for the ring or 1D torus and

$$\left(1 - h(2\tilde{m}+1)^2 + h(2\tilde{m}+1)S_1^{(\tilde{m},K)}\right)^{\frac{1}{16m^2+8m+1}}$$
$$> \beta > \left(1 - h(2m+1)^2 + h(2m+1)S_1^{(m,K)}\right)^{\frac{1}{2m(m+1)}}$$

for the 2D torus, where $\tilde{m} = \lfloor \frac{m}{\sqrt{2}} \rfloor$ as defined in Theorem 5.

5 Conclusions

We introduced a framework that considers the effects of realistic communication constraints on average consensus algorithms. In particular, we analytically characterize the performance of the medium access control algorithm that maximizes the speed of convergence. We study the effect of transmit power on convergence in the presence of interference. In inteference-limited wireless networks, the geometry of node placement plays a key role in deciding the fastest converging consensus graph. While forming long-range links (using more power) always improves the convergence on ring topologies, it is not so for higher-dimensional tori.

This work could be extended to other classes of graphs, like Cayley graphs and expander graphs that have good convergence properties [3]. Another issue is the effect of stochastic data loss through effects due to fading and interference, using a different framework as compared to [2, 6], to explicitly account for interference.

References

[1] V. Blondel, J. Hendrickx, A.Olshevsky, and J. Tsitsiklis. Convergence in multiagent coordination, consensus and flocking. *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 2996–3000, 2005.

[2] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE Transactions on Information Theory*, 52(6):2508–2530, 2005.

[3] R. Carli, F. Fagnani, A. Speranzon and S. Zampieri, "Communication constraints in the average consensus problem," *Automatica*, 44(3):671-684, March 2008.

[4] G. Cybenko. Dynamic load balancing for distributed memory multiprocessors. *Journal on parallel and distributed computing*, (7):279–301, 1989.

[5] M.P. Desai, and V.B. Rao. A new eigenvalue bound for reversible Markov chains with applications to the temperature-asymptotics of simulated annealing. *Proceedings of the IEEE International Symposium on Circuits and Systems (2):1211–1214*, 1990.

[6] P. Hovareshti, V. Gupta and J. S. Bars, "Average Consensus over Small World Networks: A Probabilistic Framework," IEEE Conference on Decision and Control (CDC 08) 2008, Submitted.

[7] M. Huang and J.H. Manton, "Stochastic double array analysis and convergence of consensus algorithms with noisy measurements," Proc. American Control Conference, New York, pp. 705-710, July 2007.

[8] M. Huang and J.H. Manton. "Stochastic Lyapunov analysis for consensus algorithms with noisy measurements," Proc. American Control Conference, New York, pp. 1419-1424, July 2007.

[9] A. Nedich and A. Ozdaglar, "Convergence Rate for Consensus with Delays," LIDS Technical Report 2774, MIT, Lab. for Information and Decision Systems.

[10] A. Nedich, A. Olshevsky, A. Ozdaglar, and J. N. Tsitsiklis, "On Distributed Averaging Algorithms and Quantization Effects," LIDS Technical Report 2778, MIT, Lab. for Information and Decision Systems.

[11] L. Fang and P. Antsaklis. On communication requirements for multi-agents consensus seeking. *Proceedings of Workshop NESC05: University of Notre Dame, Lecture Notes in Control and Information Sciences (LNCIS), Springer*, 331:53–68, 2006.

[12] A. Jadbabaie, J. Lin, and A. S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48(6):988–1001, 2003.

[13] W. Ren, R. W. Beard, and T. W. McLain. Coordination variables and consensus building in multiple vehicle systems. *"Proceedings of the Block Island Workshop on Cooperative Control, Lecture Notes in Control and Information Sciences series,Springer-Verlag"*, 309:171–188, 2004.

[14] R. Olfati Saber and R.M Murray. Consensus problems in networks of agents with switcing topology and time-delays. *IEEE Transactions on Automatic Control*, 49(9):1520–1533, 2004.

[15] E. Seneta. *Nonnegative Matrices and Markov Chains*. Springer, 2nd edition, 1981.

[16] L. Xiao and S. Boyd. Fast linear iterations for distributed averaging. *Proceedings of 42th IEEE Conference on Decision and Control*, pages 4997–5002, 2003.

[17] L. Xiao, S. Boyd, and S. Lall. A scheme for robust distributed sensor fusion based on average consensus. *Proceedings of International Conference on Information Processing in Sensor Networks*, pages 63–70, 2005.

[18] X. Liu and M. Haenggi. Throughput Analysis of Fading Sensor Networks with Regular and Random Topologies, *EURASIP Journal on Wireless Communications and Networking*, vol. 4, pp. 554-564, Aug. 2005. Special Issue on Wireless Sensor Networks.

[19] M. Xie and M. Haenggi. Delay Performance of Different MAC Schemes for Multihop Wireless Networks. *IEEE Global Communications Conference (GLOBECOM'05)*, (St. Louis, MO), Nov. 2005.

[20] B. Mohar. The Laplacian Spectrum of Graphs. "Graph Theory, Combinatorics, and Applications", Vol. 2, Ed. Y. Alavi, G. Chartrand, O. R. Oellermann, A. J. Schwenk, Wiley, 1991, pp. 871–898.