

# Consensus Problem under Diffusion-based Molecular Communication

Arash Einolghozati, Mohsen Sardari, Ahmad Beirami, Faramarz Fekri

School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332

Email: {einolghozati, mohsen.sardari, beirami, fekri}@ece.gatech.edu

**Abstract**—We investigate the consensus problem in a network where nodes communicate via diffusion-based molecular communication (DbMC). In DbMC, messages are conveyed via the variation in the concentration of molecules in the medium. Every node acquires sensory information about the environment. Communication enables the nodes to reach the best estimate for that measurement, e.g., the average of the initial estimates by all nodes. We consider an iterative method for communication among nodes that enables information spreading and averaging in the network. We show that the consensus can be attained after a finite number of iterations and variance of estimates of nodes can be made arbitrarily small via communication.

**Index Terms**—Molecular Communication, Diffusion, Consensus, Information networks, Distributed averaging

## I. INTRODUCTION

There has been numerous evidence of the existence of a form of communication using molecules in nature. At the microorganism scale, molecular signals are used for communication and control among cells in living tissues. Communication enables single cells to process sensory information about their environment (in a way similar to neural networks) and evaluate and react to chemical stimuli. The use of chemical signaling by living cells has been under extensive study. For example, it is known that some bacteria use a process named “Quorum Sensing” to estimate the density of their kind in the environment [1]–[6]. Quorum Sensing is a decentralized coordination process which allows bacteria to estimate the density of their population and regulate their behavior according to the estimated density. To estimate the local population density, bacteria release specific signaling molecules. These molecules are subject to diffusion process that would make the molecules drift away instead of accumulating in the bacteria vicinity. Therefore, the concentration near the bacteria tends to the average concentration in the medium. As the local density of bacteria increases, so will the density of molecules in the medium. Bacteria have molecules receptors that can estimate the molecules density and thus the bacteria population density. Bacteria use quorum sensing to coordinate actions (mostly energy expensive) that cannot be carried out by a single bacterium. This phenomenon, captures most of the important components of a communication system in micro-scale. Arguably, the most dominating form of communication at the scale of microorganisms is Diffusion based Molecular Communication (DbMC), i.e., embedding the information in the alteration of the concentration of the molecules and rely on diffusion to transfer the information.

Despite its importance, study of DbMC is still in infancy. Very recently, new theoretical frameworks are being developed for molecular communication [7]–[9]. Broadly speaking, the purpose of understanding DbMC is three-fold. First, in the field of system biology. For example, the new field of system biology enables us to engineer microorganisms for certain objectives [10]–[12]. Our study can help to model and manipulate biological networks comprised of such engineered microorganisms. Secondly, recently, there is a new trend of designing micro-scale networks of nano-scale devices to perform tasks similar to their biological counterparts [13]. There is a large number of applications that such networks could apply to. One may envision molecular based networks built using these nano-scale devices that can be deployed over or inside the human body to monitor glucose, sodium, and cholesterol levels, to detect the presence of different infectious agents, or to identify specific types of cancer. Such networks will also enable new smart drug administrative systems to release specific drugs inside the body with great accuracy and in a timely manner. These networks are to operate in the environments similar to those of bacteria and other living organisms. Hence, the same principles hold for these networks and DbMC is the most promising form of communication in these bio-inspired networks. There is no need for complex computation in this type of communication which has been performed by bacteria for millions of years. Therefore, it can be done by primitive nano-scale devices that have limited computation capabilities. Third, understanding DbMC can help to find out whether there is any optimality in natural complex bio-systems.

In this work, we study the consensus problem in a network governed by DbMC. In a general consensus problem, nodes in a network communicate with each other to obtain the best estimate given their initial estimates. Average value of these initial estimates is considered to be an important measure that can be considered as a goal in a network. In this work, we capture the situation where a network of nodes (agents) must achieve a consistent opinion through information exchanges via molecular signaling with their neighbors.

The key element of this type of the consensus problem is diffusion. Diffusion describes the spread of particles through random motion from regions of higher concentration to regions of lower concentration. Every node has the capability of sensing the concentration of molecules in the environment and producing new molecules with desired rates. Molecules diffuse from the transmitter to the entire media. It is important to

note that this type of molecular communication is different from other communication schemes where Brownian motion of single molecules is studied or molecules are directed toward specific directions.

The rest of the paper is organized as follows: in Sec. II, we capture the effect of diffusion by studying the Fick's second law of diffusion and show as to how the production of molecules by each node affects the concentration level sensed by other nodes in the network. In Sec. III, the consensus problem is presented and formulated in a matrix form. Then, in Sec. IV, consensus problem is studied for compact and uniform networks and an iterative algorithm is proposed for the uniform type. In Sec. V, the convergence of the proposed algorithm is verified and the rate of the convergence is discussed. Finally, in Sec. VI, the results in the previous sections is simulated and verified for a specific network.

## II. BACKGROUND

The scheme for a typical molecular communication network is depicted in Fig. 1. The communication between the shaded nodes in the network is modeled in Fig. 2. Each agent has the capability of sensing the concentration of molecules in the medium and release molecules at a specific rate back into the medium. In other words, each agent in such networks is a transceiver. Channel is the medium that molecules are injected into and carried depending on the diffusion coefficient of the medium. In molecular communication, information is encoded into the variations of molecule concentration, e.g., it can be in the form of different concentration levels corresponding to Amplitude Modulation.

Two main features that distinguish the consensus problem under DbMC from that of the traditional electromagnetic communication are the broadcast nature of DbMC and lingering of the molecules in the shared medium. The molecules that are produced by nodes stay in the medium and change the concentration sensed by all the other nodes; there is no need for specialized routing which makes it fit for *in-network* processing. Hence, the first step is to characterize the temporal and spatial variations of molecules in the channel which follows the general diffusion equations. According to Fick's second law of diffusion, the concentration of molecules  $c(x, t)$  at position  $x$  at time  $t$  in an  $m$ -dimensional space (i.e. medium) is computed using the molecule production rate  $r(x, t)$  at the source, as follows:

$$\frac{\partial c(x, t)}{\partial t} = D\nabla^2 c(x, t) + r(x, t). \quad (1)$$

Here,  $x$  is the distance of any point in the environment from the source (assuming the source is at the zero position) and  $D$  is the diffusion coefficient of the medium. The impulse response of (1), corresponding to  $r(x, t) = \delta(x)\delta(t)$ , is the Green's function  $g_d(x, t)$  whose expression is as follows:

$$g_d(x, t) = \frac{1}{(4\pi Dt)^{\frac{m}{2}}} \exp\left(-\frac{|x|^2}{4Dt}\right).$$

This impulse response is given for an  $m$ -dimensional medium. Note that diffusion in an  $m$ -dimensional space is equivalent to

$m$  simultaneous 1-D diffusions. Since the diffusion equation is a linear equation, the solution to (1) for an arbitrary input  $r(x, t)$ , denoted by  $c^*(x, t)$ , can be calculated using

$$c^*(x, t) = g_d(x, t) \otimes r(x, t),$$

where  $\otimes$  denotes multi-dimensional convolution operation on  $x$  and  $t$ .

In our setup, the only molecule producers are nodes. Therefore, we have  $r(x, t) = F(t)\delta(x)$ , where  $F(t)$  is the molecule production rate. Hence, we will have

$$c^*(x, t) = \int_0^\infty F(\tau) \frac{1}{(4\pi D(t-\tau))^{\frac{m}{2}}} \exp\left(-\frac{x^2}{4D(t-\tau)}\right) d\tau. \quad (2)$$

This response is valid for open free medium where the only boundary conditions are at the transmitter. Note that in this model, we do not consider the delay due to the travel time of molecules between the nodes. That is we assume molecules reach the nodes instantly. This does not affect our analysis of consensus, because it only shifts the time that molecules are received at the nodes.

## III. PROBLEM STATEMENT

We consider collaborative networks and refer to agents as network nodes. The goal is to spread the information about an event or any other variation through the network with minimum latency. The problem setup is such that each node in an  $N$ -node network initially has a measurement value. These initial measurement values are assumed to be formed from estimating a specific parameter in the environment. These estimates are assumed to derive from a random variable. All the nodes try to obtain the best estimate for this random variable through the communication in the network. Node  $i$  maps its estimate of the environment parameter into the level of molecule concentration  $\rho_i$  corresponding to this estimate. Network nodes exchange their estimates via producing proper molecular concentrations in the medium, to arrive at consensus. We assume these estimated concentrations are drawn from the same distribution  $N(\mu, \sigma_0^2)$ , i.e., a Gaussian distribution with expected value of  $\mu$  and variance of  $\sigma_0^2$ .

The best unbiased estimate for  $\mu$  is the average of initial estimates of nodes, i.e.  $\hat{\mu} = \rho_{av} = \frac{1}{N} \sum_{i=1}^N \rho_i$ . This estimate has variance of  $\frac{\sigma_0^2}{N}$  which can be arbitrarily small when  $N$  is large enough. Consensus is reached when estimate of each node approaches this value and the variance of their estimate approaches  $\frac{\sigma_0^2}{N}$ . We propose an algorithm that ensures all nodes to arrive at an estimate for the average consensus in finite amount of time. After evaluating the average of these concentrations, nodes can map their molecule production rate back into the average of the parameter that they measured in the environment.

In this network, nodes communicate with each other through diffusing molecules into the environment with different rates. These rates of molecule production change the concentration of molecules at vicinity of other nodes and convey the information from one node to another. Hence, each node needs

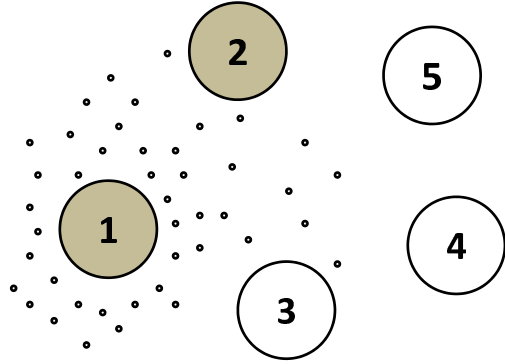


Fig. 1: Node 1 Diffusing Molecules into a Network of 5 Nodes.

to choose the best rate based on its estimate and convey this estimate to other nodes. Based on received molecule concentration and their previous estimate, nodes improve their estimates iteratively until they reach a consensus.

Assume node  $i$  has an initial estimate  $\rho_i$ . The goal is to obtain the average  $\rho_{av} = \frac{1}{N} \sum_{i=1}^N \rho_i$  by every node. From (1), the concentration of molecules  $c_i(t)$  in the vicinity of node  $i$  is given by

$$c_i(t) = \sum_{j=1}^N c_j^*(x_{ij}, t) \quad \text{for } i = 1, 2, \dots, N.$$

where  $c_j^*(x_{ij}, t)$  is the response at the position of node  $i$  due to production by node  $j$ . Furthermore,  $x_{ij}$  denotes the distance between nodes  $i$  and  $j$ . In its classical communication analogue, we have a network of  $N$  transceivers. Transmitter  $i$  sends the signal  $F_i(t)$  to  $N$  receivers through  $N$  different channels with the impulse response  $g_d(x_{ij}, t)$ ,  $j \in \{1, 2, \dots, N\}$ . At each receiver, the superposition of outputs from  $N$  different channels is observed.

The  $N$  equations above should be optimized with respect to  $F_i(t)$  to produce the average of the initial estimates in the vicinity of each node at a specific time  $T_0$ . We assume node  $i$  transmits molecules with a constant rate  $F_i$  in the interval  $[0, T_0]$ . Therefore, at time  $T_0$ , the concentration of molecules at node  $i$  is given by

$$c_i(t) = \sum_{j=1}^N F_j X(i, j), \quad (3)$$

where

$$X(i, j) = \int_0^{T_0} \frac{1}{(4\pi D(T_0 - \tau))^{\frac{m}{2}}} \exp\left(-\frac{|x_{ij}|^2}{4D(T_0 - \tau)}\right) d\tau. \quad (4)$$

Note that  $X(i, j)$  only depends on the geometry of the network.

Let  $F$  be a vector consisting of the production rate of all the nodes. Likewise, let  $\mathbf{c}$  be the vector of concentration of molecules at vicinity of the network nodes. Then, we have

$$\mathbf{c} = \mathbf{X}F. \quad (5)$$

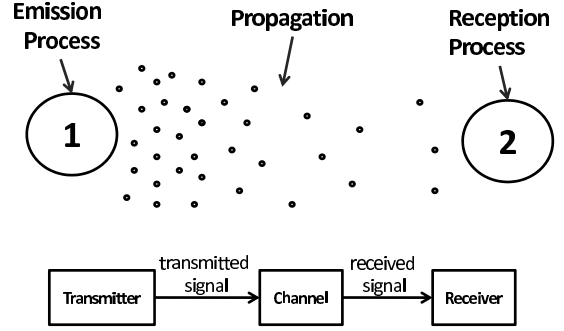


Fig. 2: Model of Molecular Communication from node 1 to node 2.

In (5),  $\mathbf{X}$  is a symmetric matrix whose entries  $X(i, j)$  are calculated according to (4). If we set every entry of  $\mathbf{c}$  as  $\rho_{av}$ , we will get a linear system of equations that can be solved to obtain  $F_i$ , the  $i^{\text{th}}$  element of  $F$ . If  $\mathbf{X}$  is invertible, there would be a unique solution for  $F$ . The important issue here is that each agent chooses  $F_i$  based solely on its estimate whereas the estimates of other nodes are unknown for the agent. Hence,  $F_i$  in (5) can depend only on  $\rho_i$  but the general solution for  $F$  does not necessarily satisfy this condition. Therefore, we propose an iterative solution for reaching the consensus while imposing this dependency constraint.

#### IV. PROPOSED ALGORITHM

We consider a discrete model of time epochs of length  $T_0$  for analyzing the network. The length of an epoch depends on the topology of the network. In particular, it depends on the effective radius  $R$  of each node (which is explained later) and the diffusion coefficient of the medium  $D$ . We can assume  $T_0 = k \frac{R^2}{D}$  where  $k$  is a constant. We arbitrarily choose  $k$  to be equal to one. Although our analysis is valid for other values of  $k$ . In each time epoch, node  $i$  emits molecules with a constant rate  $F_i$ . At the beginning of each time epoch, each node optimizes  $F_i$  based on its previous estimate and measurement of the molecular concentration in the environment, such that the estimate of all nodes becomes closer to  $\rho_{av}$ . Hence, at the end of each interval ( $t = T_0$ ), the concentration at node  $i$  can be obtained by (5).

Note that after each interval, nodes need to wait for a specific amount of time, which will be again in the form of  $k \frac{R^2}{D}$ , before releasing the molecules for the next epoch. This waiting interval is needed to allow the molecules in the environment diffuse away or equivalently to reset the channel. For the rest of the paper, an iteration includes both the injection of molecules and the waiting intervals.

**Case I:** First, we study the special case of compact networks. In such networks, nodes are in vicinity of each other and approximately sense the same concentration of molecules within the radius of the network, i.e.  $c_i = c_j, \forall i, j$ . Molecules produced by each node would have the same share in making

up the concentration in the network radius. Therefore, the matrix equation in (5) results in a single equation for the common concentration of nodes  $c(T_0)$  at the center of the network at time  $T_0$ :

$$\begin{aligned} c(T_0) &= \sum_{j=1}^N \left[ \int_0^{T_0} \frac{F_j}{(4\pi D(T_0 - \tau))^{\frac{m}{2}}} \right. \\ &\quad \left. \exp\left(-\frac{|x_j|^2}{4D(T_0 - \tau)}\right) d\tau \right] \\ &= \sum_{j=1}^N F_j X_j. \end{aligned} \quad (6)$$

Here,  $x_j$  is the distance between the node  $j$  and the center of the network (note that we choose the center for the notation simplification). Furthermore,  $X_j$  is the result of integration in (6) which depends only on  $x_j$  and the network constants.

Equation (6) implies that by setting  $F_j = \frac{\rho_j}{X_j N}$ , we can equate the concentration  $c(T_0)$  to  $\rho_{av} = \frac{1}{N} \sum_{j=1}^N \rho_j$ . We observe that the previously mentioned dependency constraint is taken into account. In this scenario, all the nodes are able to observe each other. Hence, by setting appropriate rates, they can reach the consensus in one step without any further iterations. In the following, we study the more challenging case that the network is extended and the concentration of molecules differs at the vicinity of each node. In addition, nodes may not be able to interact with all the other nodes. In this case, the information is conveyed from one side of the network to another by nodes acting as relays.

**Case II:** We consider a network with nodes that are uniformly deployed in an  $m$ -dimensional medium. In a sufficiently large network, each node observes the same relative distances to the other nodes in the network (except for the nodes on the boundary of the network whose effects are negligible when the number of nodes is large enough). Since the network is extended, each node has effective communications with only a specific number of nodes. To obtain this number, we consider a circle of radius  $r$  around each node which includes the nodes that it can have effective communications. The number of nodes at distance  $r$  from node  $i$  increases linearly with  $r$ . However, based on (2), the effect they have on node  $i$  decreases as  $e^{-r^2}$ . Hence, the collective effect of all the nodes at distance  $r$  varies as a factor of  $re^{-r^2}$  which approaches zero when  $r$  becomes large. By setting a threshold  $\epsilon$  for this factor, we can find  $R$ , the effective radius around each node. Therefore, the number of nodes  $N'$  that each node is able to interact with, is equal to  $\pi R^2 d$  in a 2- $D$  medium, where  $d$  is density of nodes in the medium.

Based on the above discussion, the matrix  $\mathbf{X}$ , will be an  $(N \times N)$  symmetric matrix where the columns are the permutation of each other. In addition, at each column (or equivalently at each row), there are  $N'$  nonzero elements whereas the rest of  $N - N'$  elements are zero. These zero elements correspond to the nodes that the node cannot interact directly. In this model, the effect of nodes on the boundary

is disregarded. We note that in a general case,  $\mathbf{X}$  would be a general symmetric matrix and the analysis will be more difficult. However, in the uniformly distributed case, we can propose an iterative method that ensures the reduction of variance in each iteration and convergence to the average value.

Let  $\boldsymbol{\rho}(n)$  be the estimate vector of nodes at epoch  $n$  and  $\rho_i(n)$  be the  $i^{\text{th}}$  element of the vector. In addition, let  $\mathbf{c}(n)$  denotes the vector of molecule concentrations at the vicinity of nodes at epoch  $n$ . Assume each node has an initial estimate  $\rho_i(0)$ . After each epoch, nodes update their estimate based on the molecules received from other nodes. Suppose  $S_i$  to be the sum of entries in the  $i^{\text{th}}$  column of  $\mathbf{X}$ . Because of the isotropic setting,  $S_i$  is the same for all columns, denoted by  $S$ . The nodes set their molecule production rate as  $F(n) = \frac{\rho(n)}{S}$  (we assume that each node is aware of the location of other nodes at its effective radius and hence can compute  $S$ ). Hence, based on (5), after communication among nodes, we will have

$$\mathbf{c}(n) = \mathbf{X} \frac{\boldsymbol{\rho}(n)}{S} = \frac{\mathbf{X}}{S} \boldsymbol{\rho}(n). \quad (7)$$

We observe that sum of each row or column in the matrix  $\tilde{\mathbf{X}} = \frac{\mathbf{X}}{S}$  is equal to one. Hence  $\tilde{\mathbf{X}}$  is a doubly stochastic matrix. We assume that each node  $i$  updates its estimate by putting  $\rho_i(n+1) = c_i(n)$ . Thus, we have an iterative equation for estimate of nodes.

$$\boldsymbol{\rho}(n) = \tilde{\mathbf{X}} \boldsymbol{\rho}(n-1). \quad (8)$$

In the following, we examine the convergence of (8).

## V. CONVERGENCE ANALYSIS

First, we verify that the iterative algorithm proposed in (8) results in an unbiased estimate for  $\mu = \mathbf{E}[\rho_i(0)]$  where  $\mu$  is the expected value of the Gaussian distribution that the initial estimate is from. We denote by  $\mathbf{E}[\boldsymbol{\rho}(n)]$  the vector of expected values of the estimates at iteration  $n$ . We know that  $\mathbf{E}[\boldsymbol{\rho}(0)] = \mu \mathbf{1}$ , where  $\mathbf{1}$  is the  $(N \times 1)$  vector whose elements are all one. Since the matrix  $\tilde{\mathbf{X}}$  is assumed to be constant during the iterations,  $\mathbf{E}[\boldsymbol{\rho}(n)]$  is given by

$$\mathbf{E}[\boldsymbol{\rho}(n)] = \mathbf{E}[\tilde{\mathbf{X}} \boldsymbol{\rho}(n-1)] = \tilde{\mathbf{X}} \mathbf{E}[\boldsymbol{\rho}(n-1)]. \quad (9)$$

Since the sum of each row of  $\tilde{\mathbf{X}}$  is equal to one,  $\tilde{\mathbf{X}} \mathbf{E}[\boldsymbol{\rho}(0)] = \mathbf{E}[\boldsymbol{\rho}(0)]$ . Hence, by continuing the chain in (9), we conclude that  $\mathbf{E}[\boldsymbol{\rho}(n)] = \mathbf{E}[\boldsymbol{\rho}(0)] = \mu \mathbf{1}$ . This implies that the estimate of nodes at each step  $n$  is an unbiased estimate for the initial parameter.

In order to study the variance of the estimates in each iteration, we need to elaborate on some characteristics of the matrix  $\tilde{\mathbf{X}}$ . This doubly stochastic matrix resembles the transition matrix in a Markov chain. We observe that this transition matrix is aperiodic because the entries on the main diagonal, which shows the effect of each node on its own, are nonzero. It is also irreducible because the graph representation of this matrix is connected and we can reach other nodes from each node. Hence, we use the Perron-Frobenius Theorem [14], regarding the eigenvalues ( $\lambda$ ) of  $\tilde{\mathbf{X}}$ :

- 1)  $\lambda_1 = 1$  and it is a singular root,
- 2)  $|\lambda_N| \leq |\lambda_{N-1}| \cdots \leq |\lambda_2| < 1$ .

Since the sum of each row is one, the eigenvector corresponding to  $\lambda_1$  is a uniform vector which is the consensus vector. The normalized form of this vector is  $v_1 = \frac{1}{\sqrt{N}}\mathbf{1}$ . The eigenvalue decomposition of  $\tilde{\mathbf{X}}$  is given by

$$\tilde{\mathbf{X}} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T. \quad (10)$$

Here,  $\mathbf{Q}$  is a matrix that whose columns are the eigenvectors of  $\tilde{\mathbf{X}}$  and  $\mathbf{\Lambda}$  is a diagonal matrix containing the eigenvalues of  $\tilde{\mathbf{X}}$ . The result in (10) comes from the fact that  $\tilde{\mathbf{X}}$  is a symmetric matrix. Hence, eigenvectors are orthogonal to each other and we have  $\tilde{\mathbf{X}}^{-1} = \tilde{\mathbf{X}}^T$ . Based on (10), we have  $\tilde{\mathbf{X}}^k = \mathbf{Q}\mathbf{\Lambda}^k\mathbf{Q}^T$ . Therefore, the estimate of nodes at iteration  $n$  is given by

$$\begin{aligned} \boldsymbol{\rho}(n) &= \tilde{\mathbf{X}}\boldsymbol{\rho}(n-1) \\ &= \tilde{\mathbf{X}}^n\boldsymbol{\rho}(0) \\ &= \mathbf{Q}\mathbf{\Lambda}^n\mathbf{Q}^T\boldsymbol{\rho}(0). \end{aligned} \quad (11)$$

Since all the eigenvalues except for  $\lambda_1 = 1$  are smaller than one, all the entries in the main diagonal of matrix  $\mathbf{\Lambda}$  approaches zero except for  $\Lambda_{11} = 1$ . Hence, we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \boldsymbol{\rho}(n) &= v_1 v_1^T \boldsymbol{\rho}(0) \\ &= \frac{1}{N} \left( \sum_{i=1}^N \rho_i(0) \right) \mathbf{1} \\ &= \rho_{av} \mathbf{1}. \end{aligned} \quad (12)$$

As we see in (12), after sufficient number of iterations, the estimate of nodes approaches to the average of initial beliefs. In order to quantify the rate of convergence, we look into the variance of estimates of nodes. As mentioned before, the smallest variance that can be achieved is  $\frac{\sigma_0^2}{N}$  which corresponds to the average value. The covariance matrix  $Cov(n)$  of the estimates at iteration  $n$  is given by

$$\begin{aligned} Cov(n) &= \mathbf{E} [(\boldsymbol{\rho}(n) - \mathbf{E}[\boldsymbol{\rho}(n)])(\boldsymbol{\rho}(n) - \mathbf{E}[\boldsymbol{\rho}(n)])^T] \\ &= \tilde{\mathbf{X}}^n \mathbf{E} [\boldsymbol{\rho}(0) - \mathbf{E}[\boldsymbol{\rho}(0)]] \\ &\quad (\boldsymbol{\rho}(0) - \mathbf{E}[\boldsymbol{\rho}(0)])^T (\tilde{\mathbf{X}}^T)^n \\ &= \tilde{\mathbf{X}}^n Cov(0) \tilde{\mathbf{X}}^n. \end{aligned} \quad (13)$$

where the fact that all powers of a symmetric matrix are also symmetric is used. Since the initial estimates are considered to be independent of each other,  $Cov(0) = \sigma_0^2 \mathbf{I}_{N \times N}$  where  $\mathbf{I}$  is the identity matrix. Hence, from (13),  $Cov(n) = \sigma_0^2 \tilde{\mathbf{X}}^{2n}$ . We denote the diagonal elements of the matrix  $Cov(n)$  by  $Cov_{ii}(n)$  which gives the variance of the estimate of each node at iteration  $n$ . For large  $n$ , we consider only the effect of  $\lambda_2$ , the second largest eigenvalue of  $\tilde{\mathbf{X}}$ . Thus, by using the decomposition in (10), we have

$$Cov_{ii}(n) = \sigma_0^2 \left( \frac{1}{N} + v_{2i}^2 \lambda_2^{2n} \right), \quad (14)$$

where  $v_{2i}$  denotes the  $i^{th}$  element of eigenvector  $v_2$  corresponding to  $\lambda_2$ . Since the norm of eigenvector  $v_2$  is one, then

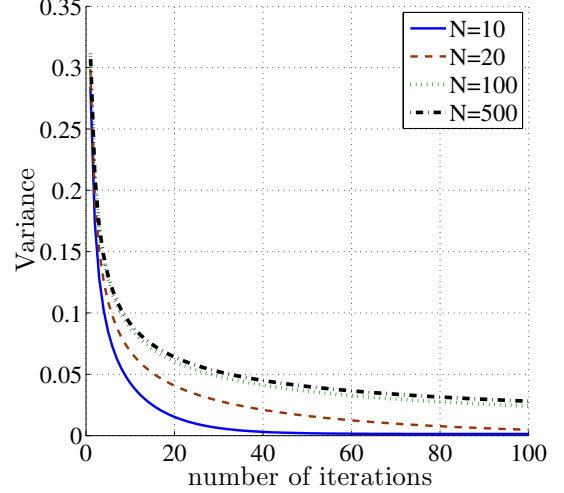


Fig. 3: Convergence of the iterative algorithm versus the number of nodes in the network.

$v_{2i} \leq 1$ ,  $i \in 1, 2, \dots, N$ . Hence, the variance of the estimate of node  $i$  at iteration  $n$  is obtained as

$$\sigma_i^2(n) \leq \sigma_0^2 \left( \frac{1}{N} + \lambda_2^{2n} \right) \quad i \in 1, 2, \dots, N. \quad (15)$$

By deploying more iterations, the upper-bound in (15) can become arbitrarily close to  $\frac{\sigma_0^2}{N}$  which approaches to zero in a network with large number of nodes. It is obvious that smaller  $\lambda_2$  will result in a faster convergence. It can be proved that when matrix  $\tilde{\mathbf{X}}$  becomes more sparse, i.e. the columns and the rows contain more zeros,  $\lambda_2$  becomes closer to one. This can be explained by the fact that a more sparse  $\tilde{\mathbf{X}}$  means a more extended network. Hence, more number of relays are needed to spread the information in the network and hence, reaching consensus will be more time consuming. In particular when  $\tilde{\mathbf{X}}$  becomes a diagonal matrix,  $\lambda_2$  is equal to one and  $\sigma_i^2$  does not converge to  $\frac{1}{N}\sigma_0^2$ . This case is equivalent to the scenario in which none of the nodes are able to communicate with each other and their initial beliefs cannot be improved.

## VI. CASE STUDY

In this section, we present an example to verify our results in the previous section. We assume a network of  $N$  nodes that are placed on a line where the distance between two successive nodes is a constant  $a$ . This network would satisfy the uniform network condition in Sec. IV as long as this line extends to infinity in both ends. For a realistic situation, we assume  $N$  to be finite. We assume nodes communicate with each other through 2-D DbMC in the medium and the size of nodes is small such that it does not interfere with diffusion of molecules.

We compute the elements of matrix  $\mathbf{X}$  based on (4) and by normalizing the columns, we obtain  $\tilde{\mathbf{X}}$ . Since, the boundary nodes are taken into account, the symmetric matrix of  $\tilde{\mathbf{X}}$  would not be doubly stochastic. We compute the variance of columns of powers of  $\tilde{\mathbf{X}}$  and verify the convergence of the matrix that governs the convergence of the estimates. The results for

variant  $N$  and fixed effective radius, i.e. fixed  $N'$ , is shown in Fig. 3. We assumed  $T_0$  to be  $\frac{q^2}{D}$  and  $N'$  equal to 5. In this plot, sum of the variance of columns is depicted versus the number of iterations. As we observe in the graph, as  $N$  becomes larger,  $\tilde{\mathbf{X}}$  becomes more sparse and the rate of convergence decreases since the second largest eigenvalue becomes larger.

## VII. CONCLUSION

In this paper, we studied the consensus problem in a network that employs DbMC for communication between nodes. We first considered the simple case of compact networks and then extended the concept to more general distributed networks. In particular, we considered a network with uniformly distributed nodes in the medium. Then, we developed a protocol as to how each node should behave to arrive at an unbiased estimate. The resulting protocol is an iterative scheme which the variance of the estimates can be reduced arbitrarily. Future works include the general case of networks and the case that nodes do not have complete information about the other nodes in their vicinity.

## REFERENCES

- [1] B. L. Bassler, "How bacteria talk to each other: regulation of gene expression by quorum sensing." *Curr Opin Microbiol*, vol. 2, no. 6, pp. 582–587, Dec 1999.
- [2] —, "Small talk. cell-to-cell communication in bacteria." *Cell*, vol. 109, no. 4, pp. 421–424, May 2002.
- [3] B. K. Hammer and B. L. Bassler, "Quorum sensing controls biofilm formation in vibrio cholerae." *Mol Microbiol*, vol. 50, no. 1, pp. 101–104, Oct 2003.
- [4] J. M. Henke and B. L. Bassler, "Bacterial social engagements." *Trends Cell Biol*, vol. 14, no. 11, pp. 648–656, Nov 2004. [Online]. Available: <http://dx.doi.org/10.1016/j.tcb.2004.09.012>
- [5] P. Mehta, S. Goyal, T. Long, B. L. Bassler, and N. S. Wingreen, "Information processing and signal integration in bacterial quorum sensing." *Mol Syst Biol*, vol. 5, p. 325, 2009. [Online]. Available: <http://dx.doi.org/10.1038/msb.2009.79>
- [6] W.-L. Ng and B. L. Bassler, "Bacterial quorum-sensing network architectures." *Annu Rev Genet*, vol. 43, pp. 197–222, 2009. [Online]. Available: <http://dx.doi.org/10.1146/annurev-genet-102108-134304>
- [7] K. V. Srinivas, R. S. Adve, and A. W. Eckford, "Molecular communication in fluid media: The additive inverse gaussian noise channel," arXiv:1012.0081v2.
- [8] A. W. E. Sachin Kadloor, Raviraj S. Adve, "Molecular communication using brownian motion with drift," arXiv:1006.3959v1.
- [9] M. Pierobon and I. F. Akyildiz, "A physical end-to-end model for molecular communication in nanonetworks," *IEEE Journal of Selected Areas in Communications*, vol. 28, pp. 602–611, 2010.
- [10] T. Danino, O. Mondragón-Palomino, L. Tsimring, and J. Hasty, "A synchronized quorum of genetic clocks." *Nature*, vol. 463, no. 7279, pp. 326–330, Jan 2010.
- [11] S. Choudhary and C. Schmidt-Dannert, "Applications of quorum sensing in biotechnology." *Appl Microbiol Biotechnol*, vol. 86, no. 5, pp. 1267–1279, May 2010. [Online]. Available: <http://dx.doi.org/10.1007/s00253-010-2521-7>
- [12] F. Dressle and O. B. Akan, "A survey on bio-inspired networking," *Computer Networks*, vol. 56, pp. 881–900, 2010.
- [13] I. Akyildiz and J. M. Jornet, "Nano communication networks," *NanoComNet*, vol. 1, pp. 3–19, 2010.
- [14] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*. Oxford University Press, 2001.