FAST DISTRIBUTED CONSENSUS ALGORITHMS BASED ON ADVECTION-DIFFUSION PROCESSES

S. Sardellitti 1, M. Giona 2, S. Barbarossa 1

INFOCOM Dept. 1, Chemical Engineering Dept. 2, Univ. of Rome “La Sapienza”, Via Eudossiana 18, 00184, Rome, Italy
E-mail: {s.sardellitti@unicas.it, max@giona.ing.uniroma1.it, sergio@infocom.uniroma1.it}

ABSTRACT
Distributed consensus algorithms have recently gained a large interest in sensor networks as a way to achieve globally optimal decisions in a totally decentralized way, that is without the need of sending all the data collected by the sensors to a fusion center. The goal of this work is to show that modeling a consensus algorithm as the homogenization process of a fluid through an advection-diffusion process provides a fundamental clue to design innovative consensus algorithms whose convergence rate can be increased by acting on the (equivalent) advection mechanism, without increasing the coverage radius of any sensor. In particular, we show the increase of convergence rate resulting from a proper interplay between advection and diffusion mechanisms.

1. INTRODUCTION

Consensus mechanisms have been recently shown to represent useful distributed approaches to achieve globally optimal decisions, either detection or estimations, in a sensor network without the need of sending all the measurements to a fusion center or sink node [1], [2]. Decentralized processing algorithms reduce the congestion around sink nodes and increase the robustness of the network against node failures or unpredictable switches to sleeping mode. The bottleneck of these decision mechanisms is that they are inherently iterative. This aspect represents a very critical issue for sensor networks where energy consumption is a fundamental concern. As pointed out in [3], what really matters in a sensor network is the energy necessary to reach a decision that is proportional to the sum of transmit powers at each node and the convergence time necessary to achieve a consensus, with a prescribed accuracy. For undirected connected graphs, the convergence rate is proportional to the graph algebraic connectivity, given by the second smallest eigenvalue of the graph Laplacian 1.

For this reason, some recent works addressed the problem of speeding up the consensus algorithms by increasing the algebraic connectivity of the graph modeling the network. Two basic, alternative, approaches have been proposed: i) decide how to establish links among nodes in order to model the network topology as a small world graph, whose algebraic connectivity is known to be larger than a random geometric graph [5]; or, alternatively, ii) given a graph topology, supposed to be undirected, choose the weights on each link in order to maximize the algebraic connectivity of the graph, as the solution of a convex optimization problem [6]. In the first case, substantial increases of algebraic connectivity are achievable, but at the price of increasing the transmit power necessary to establish links between, possibly, far nodes. Therefore, the energy spent to achieve consensus may not be the minimum. In the second case, the result is optimal, but it assumes that the topology is given, the graph is undirected, and the weights associated to each link do not vary with time.

In this work, to minimize the power consumption, we assume that each node is connected only with its first neighbors. But, besides this assumption, we allow the graph to be directed and the weights to be, possibly, time-varying, if convenient from the point of view of the convergence rate. In general, looking for the optimal time-varying law to associate to each link would be a very demanding task. For this reason we restrict our search to interaction mechanisms that can be derived by drawing a similitude between our consensus problem and transport phenomena in fluid flows. The achievement of a consensus can in fact be seen as the perfect mixing of different fluids placed in a common environment. Building on this analogy, we can exploit the well known knowledge coming from fluid dynamics and engineering practice, that mixing (homogenization) processes in a fluid can be speeded up considerably (orders of magnitude) by using a proper stirring mechanism, i.e., letting an external velocity field act on the fluid mixture and drive distant fluid elements into close contact with each others, thus enhancing the smoothening action of diffusion [7, 8].

The purpose of this paper is then how to translate the stirring mechanism into a consensus algorithm, with exchange of data only among immediate neighbors. Our goal is the

1The analysis of directed graphs may be done as in [4].
choice of the advection law and the proper balancing between advection and diffusion mechanisms, in order to get a substantial speed increase with respect to conventional consensus algorithms working over undirected, time-invariant graphs.

2. CONSENSUS AS AN ADVECTION/DIFFUSION PROCESS

In this section, we formulate the consensus algorithm as a space-time discretization of an advection-diffusion process. To avoid border effects, we consider a two-dimensional toroidal domain obtained by folding the square \( \{ (x, y) \in \mathbb{R}^2 | 0 \leq x, y \leq 1 \} \). Let us denote with \( \phi(x, y; t) \) the continuous space time function corresponding to the difference of the concentration between two reacting species. This function will be regarded as the state function of a sensor network. The advection-diffusion equation (ADE) is expressed by the following differential equation

\[
\frac{\partial \phi(x, y; t)}{\partial t} = -\mathbf{v}(x, y; t) \cdot \nabla \phi(x, y; t) + \eta \nabla^2 \phi(x, y; t),
\]

where \( \mathbf{v}(x, y; t) \) is the forced velocity field, generally a space-time function, while \( \nabla (\cdot) \) and \( \nabla^2 (\cdot) \) represent, respectively, the gradient and the Laplacian operators; \( \eta^{-1} = Pe \) is an adimensional positive parameter, known in fluid dynamics as the Péclet number.\(^2\) We can note that, \(-\mathbf{v}(x, y; t) \cdot \nabla \phi(x, y; t)\) accounts for the advection contribution associated with an external velocity field, and \( \eta \nabla^2 \phi(x, y; t)\) expresses the action of molecular diffusion that, as \( Pe \to 0 \), becomes the dominant process. In (1), it has been assumed a solenoidal velocity vector, i.e. \( \nabla \cdot \mathbf{v} = 0 \). This physically corresponds to an incompressible (liquid) mixture, and in the network formulation this implies that the sum of the network (discrete) velocities entering each node is identically equal to zero. Hence, we can write \( \mathbf{v}(x, y; t) = (v_x, v_y) \) with \( v_x = v_x(y, t) \) and \( v_y = v_y(x, t) \). Now, assuming that \( \phi(x, y; t) \) is a sufficiently smoothed function, we can discretize (1) using space-time discretization techniques, to derive an equivalent system of linear equations. For simplicity, in this work we assume an uniform spatial sampling over a square grid, and we denote by \( \phi(i, j, n) \) the \( n \)-th time sample of the state of the sensor located in the position \( (x_i = i\delta_1, y_j = j\delta_1) \) for \( i, j = 1, \ldots, N \), where \( \delta_1 = 1/N \) is the spatial sampling, supposed to be the same over all axes.\(^3\) We use the so-called upwind finite difference approximation for the first spatial derivatives with respect to \( x \) (and analogously to \( y \), i.e., we assume the forward difference approximation \( \frac{\partial \phi(x, y; t)}{\partial x} \approx \frac{\phi(x + \Delta x, y; t) - \phi(x, y; t)}{\Delta x} \) when \( v_x < 0 \) and the backward difference approximation \( \frac{\partial \phi(x, y; t)}{\partial x} \approx \frac{\phi(x, y; t) - \phi(x - \Delta x, y; t)}{\Delta x} \) if \( v_x > 0 \). The versus of the derivative is thus switched so that only the points located in the area from which the flow is coming are used to update a given point. For the second derivative, instead, we assume the central finite difference approximation, while for the derivative with respect to \( t \) the forward one with time increment \( \epsilon \). Replacing all these expressions in (1), after some simple algebraic manipulations we can write (1) as

\[
\phi(n + 1) = W(n)\phi(n),
\]

where the components \( \phi(i, j, n) \) have been stacked in the \( N^2 \) dimensional vector \( \phi(n) \) according to the above described order of the nodes on the grid. The \( N^2 \times N^2 \) matrix \( W(n) \) is given by \( W(n) = I - \epsilon L(n) \) with \( L(n) = D_f + C(n) \), where the matrices \( D_f \) and \( C(n) \) represent the diffusion and advection processes, respectively. We have assumed that the advection process might be time-varying, to have a further degree of freedom to speed up the consensus mechanism. In particular, for the diffusive contribution, we have

\[
D_f = \text{bcirc}(A_1, -I, O, \ldots, O, -I) \cdot \beta
\]

wherein \( D_f \) is an \( N^2 \times N^2 \) block circulant matrix and

\[
\beta = \eta/\delta_1^2, \quad A_1 = \text{circ}(4, -1, 0, \ldots, 0, -1).
\]

As far as the advection matrix is concerned, we can write \( C(n) = C_x(n) + C_y(n) \) and, depending on the sign of each velocity vector component, we have to follow the forward or the backward approximation. Here, we assume that the components \( v_x \) and \( v_y \) are of constant sign at each time, thus

\[
C_x(n) = \begin{cases} \text{bcirc}(A_2(n), O, \ldots, O, -A_2(n)) & \text{if } v_x(x, n) > 0 \\ \text{bcirc}(-A_2(n), A_2(n), \ldots, O, O) & \text{if } v_x(x, n) < 0 \end{cases}
\]

with \( A_2(n) = \text{diag}(\alpha_2(1, n), \alpha_2(2, n), \ldots, \alpha_2(N, n)) \) an \( N \times N \) diagonal matrix and

\[
C_y(n)^5 = \begin{cases} F(n) \otimes A_3 & \text{if } v_y(y, n) > 0 \\ F(n) \otimes A_4 & \text{if } v_y(y, n) < 0 \end{cases}
\]

with \( F(n) = \text{diag}(\alpha_x(1, n), \alpha_x(2, n), \ldots, \alpha_x(N, n), A_3 = \text{circ}(1, 0, 0, \ldots, -1), A_4 = \text{circ}(-1, 1, 0, \ldots, 0) \) and \( \alpha_x(j, n) = v_x(y_j = j\delta_1) = \alpha_y(i, n) = v_y(x_i = i\delta_1) \).

\(^2\)The Péclet number represents the ratio between the characteristic diffusion and advection times.

\(^3\)The \( N^2 \) nodes on the grid are ordered from left to right along the rows starting from the bottom side.
for \( i, j = 1, \ldots, N \). We can now exploit some properties of the matrix \( W(n) = I - \epsilon L(n) \). In particular, using the upwind scheme and, under the assumption that the coefficients \( \alpha_x(j, n) \) (or, \( \alpha_y(i, n) \)) have all the same sign, we observe that each matrix \( L(n) \) is diagonally dominant with positive elements on the main diagonal and non positive off-diagonal elements. Thus, it is possible to write \( L(n) = D(n) - A(n) \) with \( D(n) \) the degree diagonal matrix and \( A(n) \) the adjacency matrix. Observing also that \( L(n)1 = 0 \) and \( 1^T L(n) = 0^T \), it can be deduced that the matrix \( L(n) \) is the Laplacian of a balanced digraph [1].

Since between each couple of nodes on the torus there is a directed path, the graph is strongly connected and then \( L(n) \) has a simple zero eigenvalue [1]. To ensure that all the eigenvalues of \( W(n) \) are in modulus less than 1 (necessary condition to prove in the following the convergence to the average consensus of the iterative system in (2)), we assume that \( \epsilon \in (0, 1/d_{max}(n)) \) with \( d_{max}(n) \) the maximum diagonal entry of \( D(n) \). Thus \( W(n) \) falls in the class of the non negative matrices and, according to the property that the associated digraph is strongly connected, \( W(n) \) is an irreducible non negative matrix. Furthermore, for each matrix \( W(n) \), we have positive main diagonal entries and also \( W(n)1 = 1, 1^T W(n) = 1^T \), i.e. it is a doubly stochastic primitive matrix for each \( n \). Let us assume now that the matrix \( W(n) \) is chosen, for each \( n \), from a finite set of doubly stochastic primitive matrices \( \mathcal{M} = \{ W_1, \ldots, W_K \} \). Since the class of primitive and doubly stochastic matrices is closed with respect to the product [9], we can apply the Wolfowitz theorem [10] asserting that for each infinite sequence \( W_{i_1}, W_{i_2}, \ldots \) of primitive row stochastic matrices, there exists a vector \( c \) such that

\[
\lim_{n \to \infty} W_{i_n} W_{i_{n-1}} \cdots W_{i_1} = c e^T .
\]

(7)

In our context, denoting with \( W_{i_n} \in \mathcal{M} \) the chosen matrix at the time \( n \) and observing that our sequences are doubly stochastic, the Wolfowitz theorem is still valid on the transpose of each sequence thus

\[
\lim_{n \to \infty} (W_{i_n} W_{i_{n-1}} \cdots W_{i_1})^T = c_1 e^T = c e^T .
\]

(8)

and so

\[
\lim_{n \to \infty} W_{i_n} W_{i_{n-1}} \cdots W_{i_1} = \frac{1}{N^2} .
\]

(9)

This proves the convergence to the average consensus of (2) for \( \epsilon \in (0, \min_{i=1,\ldots,K} 1/d_{max}(i)) \).

3. FAST CONSENSUS WITH TIME-PERIODIC ANOSOV FLOW

Having established the link between consensus algorithms and fluid dynamics, the first question to pose is: Can we find suitable velocity fields such that the convergence can be sped up by the advection process, with respect to the only diffusive case? To give an answer to this question, we exploit again the analogy with the fluid dynamics. Specifically, in two-dimensional systems, efficient fluid mixing can be achieved if the velocity field gives rise to chaotic kinematics, this means that \( v(x, y, t) \) should be at least time-periodic. Thus the use of time-periodic Anosov map has appeared to be a possible solution [8], time switching the velocity field between two orthogonal vectors. In particular, the Anosov flow assumes

\[
v(x, y, t) = \begin{cases}
  v_1 = (v_x(y), 0) & \text{for } 2p \leq t < 2p + 1 \\
v_2 = (0, v_y(x)) & \text{for } 2p + 1 \leq t < 2(p+1)
\end{cases}
\]

(10)

where \( v(\xi) = \xi \), with \( \xi \in [0, 1] \) and \( p \in \mathbb{N} \cup \{0\} \). In words, the vector field periodically switches between two linear spatial distributions, once with a velocity along the \( x \) axis only and then along the \( y \) axis only. Each velocity field is applied for a period \( T = N_p \alpha = 1^6 \) and, observing the system in (2) at \( t = nT_p \) with \( T_p = 2 \), we can write

\[
\phi(m) = (W_f)^m \phi(0)
\]

(11)

with \( W_f = W_{N_x}^N W_{N_y}^N \), where \( W_i \), for \( i = 1, 2 \), are the matrices associated to the velocities \( v_1 \) and \( v_2 \), respectively. In particular \( W_i = I - \epsilon (D_i + C_i) \) and we use to calculate \( C_i \) the expressions in (5) and (6) with positive velocity fields and spatial discretization \( v_x(y_j = j \delta_t) = v_y(x_i = i \delta_t) \) for \( i, j = 1, \ldots, N \). We can again note that \( W_f \) is a doubly stochastic non negative irreducible matrix thus simply applying the Perron-Frobenius Theorem we have

\[
\lim_{m \to \infty} W_f^m = \frac{1}{N^2} I
\]

(12)

proving in this way the average consensus convergence of the system. In [11] it has been shown that for balanced networks the average \( \alpha = \sum_{i=1}^{N^2} \phi_i / N^2 \) is an invariant quantity for each iteration, thus, we can use for the state vector the following decomposition \( \phi = \alpha 1 + \delta \) with \( \delta \) the group disagreement vector orthogonal to \( 1 \) and evolving according to the following dynamic

\[
\delta(m) = W_f \delta(m-1) .
\]

(13)

Then, defining \( W_f^\alpha = W_f - \frac{1}{N^2} \), we have that for a consistent norm

\[
\| \delta(m) \| \leq \| (W_f^\alpha)^m \delta(0) \| \leq \| (W_f^\alpha)^m \| \| \delta(0) \| .
\]

(14)

and since the spectral radius of \( W_f^\alpha \) is defined as \( \rho(W_f^\alpha) = \lim_{m \to \infty} \| (W_f^\alpha)^m \|^{1/m} \), for \( m \) large, we can approximate

\[
N^2 = 1/\epsilon \]

has been rounded towards nearest integer.
\[ \|(W_f^-)^m\| \approx \rho(W_f^-)^m \] and then write
\[ \|\delta(m)\| \leq \rho(W_f^-)^m \|\delta(0)\|, \] (15)

thus proving that for large \( m \) both the dynamic systems in (13) and in (11) converge with asymptotic rate \( \rho(W_f^-) \).

In order to confirm the effectiveness of the above analysis we now discuss some numerical results using the time-periodic Anosov flow. For the simulation we have considered \( \beta = 1 \), \( L_i = \gamma C_i/a_c + (1-\gamma)D_f/a_d \) for \( i = 1, 2 \) with \( 0 \leq \gamma < 1 \) a coefficient that allows a trade off between the advection and the diffusion contributions. Furthermore it has been ensured, through the scaling coefficients \( a_d \) and \( a_c \), that the norms \( \|D_f/a_d\| = \sum_{i,j} |D_f(i,j)/a_d| = \|C_i/a_c\| = \|L_i\| \) for \( i = 1, 2 \). It has to be remarked that also a scaling on the period of the Anosov map has to be introduced, i.e. each velocity field is now applied for a period \( T_c = a_c/\gamma = N_c \epsilon \). In Figure 1, we report an example of the state vector evolution, for \( N^2 = 625 \). It is evident that, for the time periodic Anosov flow, the convergence to the average consensus is faster than in the only diffusive case. Furthermore, in Figure 2 we have plotted the numerical values of the second largest (in modulus) eigenvalue of \( W(n) \), \( \mu_2(W(n)) \), versus time, for \( \gamma = 0.9 \) and \( \gamma = 0.4 \) and the corresponding \( \rho(W_f^-)^m \) asymptotic rate (observing the system at \( t = m2T_c \)). It can be noted the good agreement between them and, as expected, a faster convergence rate when the advection dominates (i.e., \( \gamma = 0.9 \)). The only diffusive case is also reported, as a comparison term, showing \( \mu_2(W_d^n) \) for both the simulation results and the corresponding analytical values \((1-\epsilon4n^2/(a_dN^2))^n\), that in this case can be easily computed since \( D_f \) is a circulant matrix. Finally, in Figure 3 we plot the norm of the disagreement vector as a function of \( \gamma \), at different time values. From Figure 3, we can observe that this norm decreases with \( \gamma \) and the decrease is faster and faster as time increases.

![Fig. 2. Rate of convergence versus the time with \( \epsilon = 0.9 \).](image)

## 4. CONCLUSIONS

In this paper, we have proposed a strategy to increase the convergence rate of distributed consensus algorithms exploiting a conceptual link between consensus mechanisms and fluid homogenization through advection-diffusion processes. In this initial work, we have assumed a two-dimensional toroidal domain and a simple linear distribution of the velocity field. Interesting extensions are under study to investigate the case of random sensor locations and more complicated velocity fields.

## 5. REFERENCES


Fig. 3. Norm of the disagreement vector versus $\gamma$.


