

Collective optimization over average quantities

Luca Scardovi and Rodolphe Sepulchre

Abstract—This paper addresses the design of algorithms for the collective optimization of a cost function defined over average quantities in the presence of limited communication. We argue that several meaningful collective optimization problems can be formulated in this way. As an application of the proposed approach, we propose a novel algorithm that achieves synchronization or balancing in phase models of coupled oscillators under mild connectedness assumptions on the (possibly time-varying and unidirectional) communication graphs.

I. INTRODUCTION

Collective design problems have received considerable attention in the recent years. Applications include formation control of autonomous vehicles [1], [2] and sensor networks [3], [4]. In those applications, the collective design can be formalized as the design of decentralized algorithms for the collective minimization of a suitable cost function characterizing a common objective. The natural –e.g. gradient-based– optimization algorithms require all-to-all communication because the cost function depends on the entire state. In the present paper we call such algorithms *global information algorithms*. However the communication constraints restrict the information available to a given agent at a given instant of time. In the present paper we call the algorithms that fulfill the communication constraints *local information algorithms*. The optimization based design of local information algorithms either requires to constrain the cost function in accordance with the communication constraints or to approximate the global information algorithm with a local one. The first solution –adapting the cost function– is systematic but challenging when the communication constraints are uncertain and might change over time, which is the typical situation encountered in practice. The present paper focuses on the second solution, which consists in approximating the global information algorithm.

More specifically, we assume in this paper that the cost function is defined over *average quantities*. This situation is frequent in applications and we provide two meaningful examples: a synchronization problem, where the average quantity is a measure of synchrony between the agents, and an optimal experiment design, where the average quantity is the Fisher Information Matrix, which plays a fundamental role in information measures.

This paper presents research results of the Belgian Programme on Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Office. The scientific responsibility rests with its authors.

L. Scardovi and R. Sepulchre are with the Department of Electrical Engineering and Computer Science, University of Liège, Belgium. l.scardovi@ulg.ac.be, r.sepulchre@ulg.ac.be

When the cost function is defined over average quantities, the evaluation of its partial derivatives (e.g. its gradient) requires a same average quantity as the sole global information. This suggests a simple (local) approximation of the (global) partial derivatives, in which the *global* average quantities are replaced by *local* estimates. The vast literature on consensus algorithms provide a good source of *local* estimators of *global* averages. The basic idea in this paper is thus very simple: starting from a global information optimization algorithm, one constructs a local information algorithm which replaces the global average information by local estimates provided by consensus algorithms.

In this approximation process, the dynamics of the optimization algorithm become coupled with the dynamics of the estimation (consensus) algorithm. As a consequence, the convergence of the approximated algorithm cannot be guaranteed a priori and must be studied in specific situations. In this paper, we successfully apply the general approximation idea to a specific problem where the convergence properties of the global information algorithm are recovered with a local one. We propose a local information algorithms that minimizes or maximizes the synchronization in phase models of coupled oscillators. The proposed algorithm recovers the *global* properties of the global one from which it is derived. This is in contrast with recent results in the literature which have focused on local information algorithms with *local* convergence properties for phase models of oscillators [5], [6].

The paper is organized as follows. In the next section, we define the problem of *collective optimization over average quantities* and we illustrate two meaningful problems that fit the proposed framework. In Section III we review the problem of estimating an average quantity in a limited communication setting. Subsequently, in Section IV, our local information optimization scheme is presented. Section V and Section VI, present a nontrivial example in which a global convergence analysis of the proposed local information algorithm can be established. Finally, in Section VII, we conclude with some technical observations about the proposed scheme.

II. COLLECTIVE PROBLEMS INVOLVING AVERAGES

A *collective optimization* problem is a problem where a group of N autonomous agents act together to optimize a common objective. We assume that each agent has a state $x_k \in X$, $k = 1, \dots, N$, where X is a *smooth manifold*, and that the cost function is a continuously differentiable and lower bounded function $V(x)$ with $x \triangleq \text{col}(x_1, \dots, x_N) \in X \times \dots \times X$.

The distinct feature of a *collective* optimization problem with respect to a standard optimization problem is that the available information to agent k is not the entire state x but rather the state x_l of a few ‘neighbors’. The optimization algorithm must take into account the limited communication between agents, which could moreover change over time.

In this paper, we assume that the cost function is a cost over average quantities, that is,

$$V(x) = \tilde{V}(\bar{z}(x)), \quad (1)$$

where

$$\bar{z}(x) = \frac{1}{N} \sum_k z(x_k),$$

and where z is the (arithmetic) mean over all the agents of the continuously differentiable function from X to \mathbb{R}^q . In general, descent algorithms that minimize the cost function V will depend on the entire state x . For instance, a gradient-based algorithm will require for agent k the partial derivative

$$\frac{\partial V}{\partial x_k} = \frac{\partial \tilde{V}}{\partial \bar{z}}(\bar{z}(x)) \cdot \frac{\partial \bar{z}}{\partial x_k}(x_k) \quad (2)$$

which depends on the entire state x through the average quantity $\bar{z}(x)$. A local information optimization algorithm must replace this global information by local information. This suggests to equip each agent with an estimator of the average quantity $\bar{z}(x)$ and to update this estimator with the locally available information. Before proceeding with this idea, we illustrate two collective optimization problems that fit the proposed framework.

A. Synchronization in phase models of oscillators

The problem of phase synchronization in large ensembles of oscillators has been extensively studied in the literature [7], [8], [9]. Phase models are popular in this context: each oscillator is modeled by a phase variable $\theta_k \in S^1$ that, in the absence of coupling, obeys the trivial dynamics $\dot{\theta}_k = \omega_k$ where ω_k is the natural frequency of oscillator k . In the present paper, we consider a population of identical oscillators, that is, $\omega_k = \omega$ for all k . The centroid of N oscillators is defined as

$$p_\theta = \frac{1}{N} \sum_{k=1}^N e^{i\theta_k} = |p_\theta| e^{i\psi}. \quad (3)$$

The parameter $|p_\theta|$ is a measure of synchrony of the phase variables θ . It is maximal when all phases are synchronized (identical). It is minimal when the phases balance to result in a vanishing centroid. Hence synchronization and balancing correspond to maximizing or minimizing the cost function

$$V(\theta) = \frac{N}{2} |p_\theta|^2. \quad (4)$$

This is an example of a cost function defined over averages, with $x_k = \theta_k \in S^1$ and $z(x_k) = e^{i\theta_k} \in \mathbb{C}$. Its gradient is computed as

$$\frac{\partial V}{\partial \theta_k} = \langle p_\theta, i e^{i\theta_k} \rangle = \frac{1}{N} \sum_{j=1}^N \sin(\theta_j - \theta_k), \quad (5)$$

where the inner product $\langle \cdot, \cdot \rangle$ is defined by $\langle z_1, z_2 \rangle = \text{Re}\{\bar{z}_1 z_2\}$ for $z_1, z_2 \in \mathbb{C} \approx \mathbb{R}^2$. A continuous-time gradient algorithm associated to the cost function (4) is

$$\dot{\theta}_k = \omega - \frac{K}{N} \sum_{j=1}^N \sin(\theta_j - \theta_k) = \omega - K \langle p_\theta, i e^{i\theta_k} \rangle, \quad (6)$$

for $k = 1, \dots, N$, where the sign of the parameter K determines a descent or ascent algorithm for the cost (4). This all-to-all sinusoidal coupling is the most frequently studied coupling in the literature of coupled oscillators [7], [9]. Its application in the context of collective stabilization of steered particles in the plane is discussed in [10].

Recently, phase synchronization or phase balancing in the context of local information algorithms has been an active research area [2], [6], [5]. In those algorithms, one needs to replace the global information p_θ that appears in the gradient (5) by a local information that might change over time. The approach proposed in the present paper is to augment the state of each oscillator with an estimator of the centroid p_θ .

B. Collective Optimal Experiment Design

Consider a collective Optimal Experiment Design Problem where N Decision Makers can acquire information about an unknown quantity by means of a noisy measurement process. A typical setting for this problem is as follows: one considers the measurement equation

$$q_k = H(x_k) w + \eta_k, \quad k = 1, \dots, N, \quad (7)$$

where w is the vector of parameters to estimate, q_k is the measurement vector, η_k is an i.i.d. Gaussian noise vector with zero mean and diagonal covariance matrix with all equal elements σ^2 , and $H(\cdot)$ is a matrix that is a function of the state of the Decision Maker k . The objective is to optimize the measurements intake such that the information about the unknown parameters w is maximized. By rewriting (7) in a vector form we obtain

$$q = \tilde{H}(x) w + \eta, \quad k = 1, \dots, N, \quad (8)$$

where

$$\tilde{H}(x) = [H^T(x_1), \dots, H^T(x_N)]^T.$$

Standard information measures (both in the bayesian setting and in the classical one) are scalar functions of the Fisher Information Matrix [11]:

$$M(x) \triangleq \frac{1}{\sigma^2} \tilde{H}^T(x) \tilde{H}(x) = \frac{1}{\sigma^2} \sum_k H^T(x_k) H(x_k). \quad (9)$$

For instance, the so-called D-optimal criterion aims at minimizing the cost

$$V_1(x) = \log \det M^{-1}(x),$$

while the so-called A-optimal experiment design aims at minimizing the cost

$$V_2(x) = \text{Trace}(M^{-1}(x)).$$

Because (9) is an average quantity, the two cost functions V_1 and V_2 , as well as any other scalar function of the Fisher

information matrix, fit the framework of the present paper. When the communication among the experimenters is limited and is constrained by some communication topology, optimization algorithms must rely on local information, which requires to replace the Fisher Information Matrix by a local estimate. In sensor networks the problem of moving sensors to acquire information about a dynamic or static process is particularly in vogue [4], [12], [3]. Applications in underwater robotics and environment exploration are particularly interesting [3]. It is not hard to imagine that in such situations the communication topology can be time variant and not all-to-all. The technique proposed in this paper opens new directions to address these problems under mild assumptions on the communication topology.

III. DECENTRALIZED ESTIMATION OF AVERAGES

In this section we review the problem of estimating an average quantity with local information. This problem, known as the *Consensus Problem*, has received considerable attention in the recent years, see for instance [13], [5], [14].

Let $G = (\mathcal{V}, \mathcal{E}, A)$ be a weighted digraph (directed graph) where $\mathcal{V} = \{v_1, \dots, v_N\}$ is the set of nodes, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges, and A is a weighted adjacency matrix with nonnegative elements a_{kj} . The node indices belong to the set of positive integers $\mathcal{I} \triangleq \{1, \dots, N\}$. Assume that there are no self-cycles i.e. $a_{kk} = 0, \forall k \in \mathcal{I}$. The graph Laplacian L associated to the graph G is defined as

$$L_{kj} = \begin{cases} \sum_i a_{ki}, & j = k \\ -a_{kj}, & j \neq k. \end{cases}$$

The k -th row of L is defined by L_k . The in-degree (respectively out-degree) of node v_k is defined as $d_k^{in} = \sum_{j=1}^N a_{kj}$ (respectively $d_k^{out} = \sum_{j=1}^N a_{jk}$). The digraph G is said to be *balanced* if the in-degree and the out-degree of each node are equal, that is,

$$\sum_j a_{kj} = \sum_j a_{jk}, \quad \forall k \in \mathcal{I}.$$

It is both of theoretical and practical interest to consider time-varying communication topologies. For example, in a network of moving agents, some of the existing links can fail and new links can appear when other agents enter an effective range of detection. In the following we assume that the communication topology is described by a time-varying graph $G(t) = (\mathcal{V}, \mathcal{E}(t), A(t))$, where $A(t)$ is piece-wise continuous and bounded and $a_{kj}(t) \in \{0\} \cup [\beta, \gamma], \forall k, j$, for some finite scalars $0 < \beta \leq \gamma$ and for all $t \geq 0$. The set of neighbors of node v_k at time t is denoted by $\mathcal{N}_k(t) \triangleq \{v_j \in \mathcal{V} : a_{kj}(t) \geq \beta\}$. We recall two definitions that characterize the concept of uniform connectivity for time-varying graphs.

Definition 1: Consider a graph $G(t) = (\mathcal{V}, \mathcal{E}(t), A(t))$. A node v_k is said to be connected to node v_j ($v_j \neq v_i$) in the interval $I = [t_a, t_b]$ if there is a path from v_k to v_j which respects the orientation of the edges for the directed graph $(\mathcal{N}, \cup_{t \in I} \mathcal{E}(t), \int_I A(\tau) d\tau)$.

Definition 2: $G(t)$ is said to be uniformly connected if there exists $T > 0$ such that for all t there is one node connected with all the other nodes across $[t, t + T]$.

Consider a group of N agents with state $y_k \in Y$, where Y is an Euclidean space. The communication between the N -agents is defined by the graph G : each agent can sense only the neighboring agents, i.e. agent j receives information from agent i iff $i \in \mathcal{N}_j(t)$. We use the notation $k \sim j$ to indicate the presence of a communication link from agent j to agent k , i.e. $k \sim j$ iff $v_j \in \mathcal{N}_k$.

The following proposition follows directly from the results of [13],[15] and [5]:

Proposition 1: Let Y be a finite-dimensional Euclidean space. Let $G(t)$ be a uniformly connected digraph and $L(t)$ the corresponding Laplacian matrix bounded and piecewise continuous in time. Then the solutions of

$$\dot{y} = -L(t)(y), \quad (10)$$

asymptotically converge to a consensus value $\alpha \mathbf{1}$ for some $\alpha \in Y$. Furthermore if $G(t)$ is balanced for all t , then $\alpha = \frac{1}{N} \sum_{i \in \mathcal{I}} y_i(0)$. \square

A general proof for Theorem 1 is based on the property that the convex hull of vectors $y_k \in X$ is non expanding along the solutions. For this reason, the assumption that Y is an Euclidean space is essential (see e.g. [5]). Under the additional balancing assumption on $G(t)$, the norm $y^T y$ is non increasing. Moreover, the balancing assumption implies $\mathbf{1}^T L(t) = 0$, which implies that the average $\frac{1}{N} \sum_{j \in \mathcal{I}} y_j$ is an invariant quantity along the solutions.

Remark 1: For the sake of simplicity, we have considered scalar valued y_k . The extension to vector or matrix valued quantities is straightforward and can be done by substituting L with $\hat{L} \triangleq L \otimes I_b$ where $b = \dim(y_k)$ (in the case of matrix valued y_k , it is sufficient to vectorize the matrix and to proceed in the same way). The results presented in the paper can be easily extended to these cases by relying on the properties of the Laplacian matrix and the Kronecker product.

In the following Section we will see that this filter plays a central role in collective optimization problems.

IV. COLLECTIVE OPTIMIZATION

We now return to the problem of designing local information algorithms to optimize the cost function $V(\bar{z}(x))$. The main idea is to estimate the average quantity $\bar{z}(x)$ with a (local information) consensus filter and to substitute the local estimate in the expression of the gradient (2). In other words, using a certainty equivalence principle, one can use y_k as a local estimate of the average quantity \bar{z} in order to transform a global information algorithm into a local one. For instance, one can replace the (global information) continuous-time

gradient algorithm

$$\dot{x}_k = \frac{\partial \tilde{V}}{\partial \bar{z}}(\bar{z}(x)) \cdot \frac{\partial \bar{z}}{\partial x_k}(x_k) =: F(\bar{z}) \cdot \frac{\partial \bar{z}}{\partial x_k}(x_k), \quad (11)$$

for $k = 1, \dots, N$, with the (local information) algorithm

$$\dot{x}_k = F(y_k) \cdot \frac{\partial \bar{z}}{\partial x_k}(x_k), \quad \forall k \in \mathcal{I} \quad (12)$$

Where y_k is a local estimate provided by the consensus filter. The algorithm (12) is ‘‘local’’ because each agent uses its own estimate y_k of the average \bar{z} . The limited communication between agents is used to asymptotically reach a consensus between the local estimators.

There is of course no guarantee that the local information algorithm (12) will asymptotically behave as the global information algorithm (11). This is because the dynamic coupling between the gradient-like dynamics and the estimation dynamics may destroy the convergence properties of both subsystems. Simulations indicate that this will indeed occur in general. A natural way to recover the convergence properties of the consensus filter and of the gradient system is to enforce a time-scale separation between the estimation process and the minimization process. This is nevertheless of limited interest in practice because the time-scale of the estimation process is graph-dependent and therefore uncertain, leading to conservative bounds on the allowable time-scale for the minimization process. The next section presents a nontrivial situation in which convergence of the local information algorithm can be established without assuming time-scale separation between the estimation process and the optimization process.

V. GLOBAL SYNCHRONIZATION IN PHASE MODELS

As a first illustration of the proposed approach, we apply the local information algorithm (12) to the global synchronization of the phase model discussed in Section II-A. By expressing the dynamics in a rotating frame, without loss of generality, we can set the natural frequencies ω equal to zero. Synchronized states coincide with the global maxima of the cost function $V = \frac{N}{2} |p_\theta|^2$. We seek to replace the (global information) gradient algorithm

$$\dot{\theta}_k = \frac{1}{N} \sum_{j=1}^N \sin(\theta_j - \theta_k) = \langle p_\theta, i e^{i\theta_k} \rangle \quad (13)$$

by the local information algorithm

$$\begin{aligned} \dot{\theta}_k &= \langle y_k, i e^{i\theta_k} \rangle, \quad \forall k \in \mathcal{I}, \\ \dot{y}_k &= -L_k(t)(y), \quad y(0) = e^{i\theta(0)}. \end{aligned} \quad (14)$$

In this particular situation, convergence of the local information algorithm can be guaranteed because the estimation (consensus) algorithm is decoupled from the optimization algorithm.

Theorem 1: Suppose that the communication graph $G(t)$ is uniformly connected and that $L(t)$ is bounded and piecewise continuous. Then all the solutions of the algorithm (14) asymptotically converge to an equilibrium. Moreover,

the only stable equilibrium in the shape space T^N/S^1 is the synchronized state characterized by N identical phases. Furthermore, if $G(t)$ is balanced for all t , then the asymptotic consensus value for $e^{i\theta_k}$ is $\alpha = (\frac{1}{N} \sum_{i \in \mathcal{I}} e^{i\theta_i(0)})$, that is the centroid $p_\theta(0)$ of the initial condition. \square

Proof: Proposition 1 guarantees that each local estimate y_k converges to a consensus value $\alpha = |\alpha| e^{i\phi}$. As a consequence, the optimization algorithm asymptotically converges to

$$\dot{\theta}_k = \frac{1}{N} \langle \alpha, i e^{i\theta_k} \rangle = |\alpha| \sin(\theta_k - \phi), \quad (15)$$

for $k = 1, \dots, N$.

Since the consensus dynamics for $y(t)$ are invariant with respect to translations in the plane, for any particular graph sequence, α has an equal probability to take any value in the complex plane if the initial conditions $e^{i\theta(0)}$ are randomly chosen (in the complex plane). This is sufficient to conclude that $\alpha \neq 0$ with probability 1.

Solutions of the complete system (14) are known to converge to a chain recurrent set of the limiting (autonomous) system (15) [17]. The limiting system is decoupled into N identical scalar systems whose only chain recurrent sets are the two equilibria of (15) (one stable node and one unstable node). Then the only limit sets of the local information algorithm (14) are equilibria that satisfy $\theta_k = \phi \bmod \pi$ for all k . The synchronized equilibrium $\theta = \mathbf{1}\phi$ is exponentially stable while all other equilibria are exponentially unstable. If $G(t)$ is balanced, it follows from Proposition 1 that $\alpha = \frac{1}{N} \sum_{i \in \mathcal{I}} e^{i\theta_i(0)} = p_\theta(0)$. \square

We conclude that synchronization on the circle can be achieved with a local information algorithm whose exchanged information is not the relative phase but rather the (relative) estimate of a vector that serves as a consensus reference direction. The global convergence analysis obtained in this way is in contrast with the local one proposed in [5], [6], for the (local information) algorithm

$$\dot{\theta}_k = \frac{1}{N} \langle L_k(t) e^{i\theta}, i e^{i\theta_k} \rangle = \frac{1}{N} \sum_{k \sim j} a_{kl} \sin(\theta_l - \theta_k), \quad (16)$$

for $k = 1, \dots, N$. The local information algorithm (16) can be considered as an approximation of the global information algorithm (13), in which the average p_θ is estimated by a local average $\sum_{k \sim l} e^{i\theta_l}$.

Convergence to the synchronized state in the system (16) can be guaranteed only if the maximal initial difference between any two phases does not exceed π , that is, if all phases initially lie on the same half-circle. The numerical simulation in Fig.1 illustrates a situation where the (dynamic) algorithm (14) achieves synchronization while the (static) algorithm (16) converges to a balanced state where $p_\theta = 0$. In this example, the communication is a fixed ring topology (see Fig. 3) and the initial phase distribution spreads over more than half a circle (see Fig. 2).

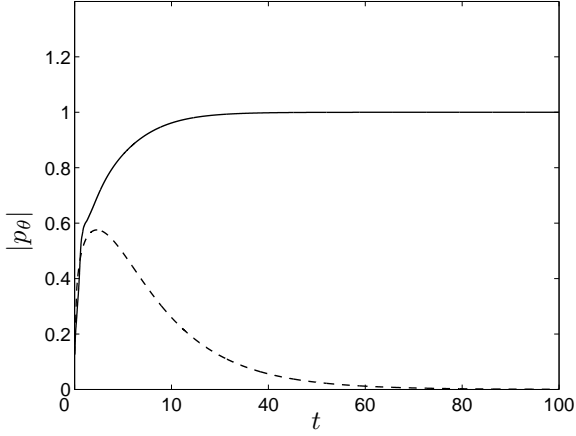


Fig. 1. Comparison of the behavior of the synchronization parameter $|p_\theta|$ for two different local information algorithms: the (dynamic) algorithm (14) with the initial conditions randomly selected (solid line); the (static) algorithm (16) (dash line). Only the first algorithm achieves synchronization.

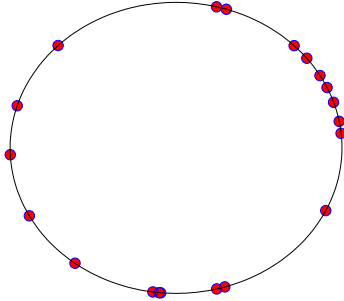


Fig. 2. Initial positions of the particles on the circle for the simulation in Fig 1.

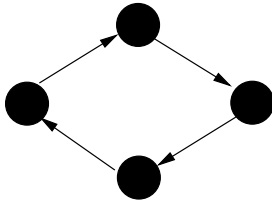


Fig. 3. Ring-directed communication topology used in the simulation in Fig 1.

Remark 2: The local information algorithm (14) extends in a straightforward way the global convergence properties of the global information algorithm (13). This extension is not restricted to the phase model discussed in the present paper. For instance, the paper [10] proposes a gradient-based approach to stabilize various collective motions in the plane. All collectives in that paper are specified as extrema of cost functions over averages. The resulting global information algorithms can be extended to the local ones using the approach of the present paper.

VI. GLOBAL BALANCING IN PHASE MODELS

Balanced states of the phase model discussed in Section II-A coincide with the global minima of the cost function $V = \frac{N}{2} |p_\theta|^2$. Similarly to the synchronization algorithm discussed in the previous section, we replace the (local information) gradient algorithm

$$\dot{\theta}_k = -\frac{1}{N} \sum_{j=1}^N \sin(\theta_j - \theta_k) = -\langle p_\theta, i e^{i\theta_k} \rangle, \quad (17)$$

for $k = 1, \dots, N$, by the local information algorithm

$$\begin{aligned} \dot{\theta}_k &= -\langle y_k, i e^{i\theta_k} \rangle, \quad \forall k \in \mathcal{I}, \\ \dot{y}_k &= -L_k(t)y + \frac{d}{dt} e^{i\theta_k}, \quad y(0) = e^{i\theta(0)}. \end{aligned} \quad (18)$$

In contrast to the synchronization algorithm (14) the gradient dynamics are here decoupled with the consensus dynamics. However the next result shows that the consensus filter asymptotically estimates the centroid and provides a globally convergent algorithm without any time-scale separation between the estimation algorithm and the optimization algorithm.

Theorem 2: Suppose that the communication graph $G(t)$ is uniformly connected and balanced for all $t \geq 0$ and that $L(t)$ is bounded and piecewise continuous. Then all the solutions of the local information algorithm (18) asymptotically converge to an equilibrium. Moreover, the only stable limit set is the set of balanced states characterized by $p_\theta = 0$. \square

Proof: The convergence analysis uses the Lyapunov function

$$W(y) = \frac{1}{2} \langle y, y \rangle.$$

Because the graph is balanced, $L(t)$ is a positive semi-definite matrix [16]. As consequence the time derivative along the solutions of (18) is non increasing:

$$\begin{aligned} \dot{W} &= -\langle L(t)y, y \rangle - \sum_{k=1}^N \langle y_k, i e^{i\theta_k} \rangle^2 \\ &= -\langle L(t)y, y \rangle - \sum_{k=1}^N \dot{\theta}_k^2 \leq 0. \end{aligned} \quad (19)$$

We deduce from (19) that $\dot{\theta}$ is a function in $L_2(0, \infty)$ and that y is uniformly bounded. To prove that $\dot{\theta}$ asymptotically converges to zero observe that

$$\dot{\theta}_k = \langle L_k(t)y, i e^{i\theta_k} \rangle + (\langle y_k, e^{i\theta_k} \rangle - 1) \dot{\theta}_k$$

is uniformly bounded, which implies that $\dot{\theta}$ is Lipschitz continuous. We conclude that $\dot{\theta}$ is uniformly continuous.

Then $\dot{\theta}$ is a uniformly continuous function in $L_2(0, \infty)$ and from Barbalat's Lemma we obtain that $\dot{\theta} \rightarrow 0$ as $t \rightarrow \infty$ [18].

Thanks to the balancing assumption on the graph, $\mathbf{1}$ is a left eigenvector of $L(t)$, and we obtain from (18) that

$$\frac{1}{N} \langle \mathbf{1}, \dot{x} \rangle = \frac{1}{N} \langle \mathbf{1}, \frac{d}{dt} e^{i\theta} \rangle. \quad (20)$$

Integrating both sides of (20), and using the fact that $y_k(0) = e^{i\theta_k(0)}$, one concludes that $\frac{1}{N} \sum_i y_i(t) = p_\theta$ for all $t \geq 0$. Because $y(t)$ converges to a consensus equilibrium, each component y_k must asymptotically converge to p_θ . As a consequence, the dynamics $\dot{\theta}_k = - \langle y_k, e^{i\theta_k} \rangle$ asymptotically converge to the time-invariant dynamics

$$\dot{\theta}_k = - \langle p_\theta, i e^{i\theta_k} \rangle, \quad \forall k \in \mathcal{I}. \quad (21)$$

Since $\dot{\theta}$ is asymptotically convergent to zero, the solutions asymptotically converge to a set of equilibria of (21). We conclude that $\theta(t)$ asymptotically converges to the critical set of V and that only the set of balanced states is asymptotically stable. \square

VII. FORMULATION IN SHAPE COORDINATES

The closed loop vector field associated to the feedback control proposed in the previous section, is not invariant under the action of the symmetry group S^1 . One of the consequences is that the agents must rely on a common coordinate frame (global information is needed). To recover the rotational invariance we must express our decentralized scheme in *shape coordinates*. Defining

$$r_k = (y_k) e^{-i\theta_k}, \quad \forall k \in \mathcal{I}.$$

We rewrite (14) as

$$\begin{aligned} \dot{\theta}_k &= \langle r_k, i \rangle, \quad \forall k \in \mathcal{I}, \\ \dot{r}_k &= -i r_k \dot{\theta}_k - \sum_{j=1}^N L_{kj} r_j e^{i(\theta_j - \theta_k)}, \end{aligned} \quad (22)$$

where $r_k(0), \forall k \in \mathcal{I}$ is randomly chosen in \mathbb{C} . Likewise we rewrite (18) as

$$\begin{aligned} \dot{\theta}_k &= - \langle r_k, i \rangle, \quad \forall k \in \mathcal{I}, \\ \dot{r}_k &= i(1 - r_k) \dot{\theta}_k - \sum_{j=1}^N L_{kj} r_j e^{i(\theta_j - \theta_k)}, \end{aligned} \quad (23)$$

where $r_k(0) = 1, \forall k \in \mathcal{I}$.

Now the vector field is invariant to the action of the symmetry group S^1 and the closed-loop dynamics evolve on a reduced quotient manifold. This $(N-1)$ -dimensional manifold is called the shape space and corresponds to the space of all relative orientations. Since only a change of variable has been introduced, the convergence analysis performed in the previous sections is still valid (see [19] for more details).

VIII. CONCLUSION

In this paper, we have argued that several collective optimization problems can be formulated as the design of local information algorithms for the minimization of a cost function defined over average quantities. Classical – e.g. gradient-based – algorithms associated to such cost functions require all-to-all communication because the gradient-based update of each agent uses average quantities. This

observation suggests to define a local information algorithm by substituting local estimates of the average quantities in the global information algorithm. The vast literature on consensus algorithms provides the local estimators of average quantities. The remaining challenge is to obtain convergence proofs in local information algorithms that couple the estimation algorithm with the optimization algorithm. Such a global convergence analysis is provided in the present paper for the meaningful application of maximizing or minimizing the synchronization of a population of oscillators described by phase models.

REFERENCES

- [1] A. Jadbabaie, J. Lin, and S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. on Automatic Control*, vol. 48, pp. 988–1001, 2003.
- [2] R. Sepulchre, D. Paley, and N. Leonard, "Group coordination and cooperative control of steered particles in the plane," in *Group Coordination and Cooperative Control*, (K. Y. Pettersen, J. Gravedahl, H. Nijmeijer (Eds.)), Lecture Notes in Control and Information Sciences, vol. 336, Springer, pp. 217–232, 2006.
- [3] N. Leonard, D. Paley, F. Lekien, D. Frantoni, R. Sepulchre, and R. Davis, "Collective motion, sensor networks and ocean sampling," to appear in *Proceedings of the IEEE*, 2006.
- [4] W. Li and C. G. Cassandras, "Distributed cooperative coverage control of sensor networks," in *Proceedings of the 44th IEEE Conference on Decision and Control and European Control Conference*, Seville, Spain, 2005, pp. 2542–2547.
- [5] L. Moreau, "Stability of multi-agent systems with time-dependent communication links," *IEEE Trans. on Automatic Control*, vol. 50, pp. 169–182, 2005.
- [6] A. Jadbabaie, N. Motee, and M. Barahona, "On the stability of the kuramoto model of coupled nonlinear oscillators," in *Proceedings of the 42nd IEEE American Control Conference*, Boston, Ma, 2004, pp. 4296–4301.
- [7] Y. Kuramoto, *Chemical oscillations, waves and turbulence*. Springer-Verlag, 1984.
- [8] S. H. Strogatz, *Sync: The Emerging Science of Spontaneous Order*. New York: Hyperion Press, 2003.
- [9] —, "From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators," *Physica D*, vol. 143, pp. 1–20, 2000.
- [10] R. Sepulchre, D. Paley, and N. Leonard, "Stabilization of planar collective motion with all-to-all communication," to appear in *IEEE Trans. on Automatic Control*.
- [11] V. V. Fedorov, *Theory of optimal experiments*. Academic Press, 1972.
- [12] D. Ucinski and Y. Q. Chen, "Time-optimal path planning of moving sensors for parameter estimation of distributed systems," in *Proceedings of the 44th IEEE Conference on Decision and Control and European Control Conference*, Seville, Spain, 2005, pp. 5257–5262.
- [13] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. on Automatic Control*, vol. 49, pp. 1520–1533, 2004.
- [14] V. D. Blondel, J. M. Hendrickx, A. Olshevsky, and J. N. Tsitsiklis, "Convergence in multiagent coordination, consensus, and flocking," in *Proceedings of the 44th IEEE Conference on Decision and Control and European Control Conference*, Seville, Spain, 2005.
- [15] D. P. Spanos, R. Olfati-Saber, and R. M. Murray, "Dynamic consensus for mobile networks," in *Proceedings of the 16th IFAC World Congress*, Prague, Czech Republic, 2005.
- [16] J. C. Willems, "Lyapunov functions for diagonally dominant systems," *Automatica J. IFAC*, vol. 12, pp. 519–523, 1976.
- [17] K. Mischaikow, H. Smith, and H. Thieme, "Asymptotically autonomous semiflows; chain recurrence and Lyapunov functions," *Trans. of the American Mathematical Society*, vol. 347, pp. 1669–1685, 1995.
- [18] H. K. Khalil, *Nonlinear Systems*. Prentice Hall, 2001.
- [19] L. Scardovi, A. Sarlette, and R. Sepulchre, "Synchronization and balancing on the N -torus," *Systems and control Letters*, 2006, submitted.