Consensus for Agents with General Dynamics Using Optimistic Optimization

Lucian Buşoniu and Irinel-Constantin Morărescu

Abstract— An important challenge in multiagent systems is consensus, in which the agents must agree on certain controlled variables of interest. So far, most consensus algorithms for agents with nonlinear dynamics exploit the specific form of the nonlinearity. Here, we propose an approach that only requires a black-box simulation model of the dynamics, and is therefore applicable to a wide class of nonlinearities. This approach works for agents communicating on a fixed, connected network. It designs a reference behavior with a classical consensus protocol, and then finds control actions that drive the nonlinear agents towards the reference states, using a recent optimistic optimization algorithm. By exploiting the guarantees of optimistic optimization, we prove that the agents achieve practical consensus. A representative example is further analyzed, and simulation results on nonlinear robotic arms are provided.

I. Introduction

We consider the problem of *consensus* in multi-agent systems, in which the agents must reach agreement on controlled variables of interest, in a decentralized fashion [1], [2]. This problem has important applications in the control of networks, such as transport, communication, and sensor networks. Classical consensus algorithms are designed for simple linear agents, e.g. [1]–[8]. In this linear setting, the behavior of the agents is well understood and useful properties can be ensured. Consensus approaches have also been proposed for nonlinear agent dynamics, such as second-order systems with nonlinear acceleration dynamics [9], [10], nonholonomic robots [11], and Euler-Lagrange dynamics [12]. The Lyapunov design techniques used in these works use the explicit forms of the agents' dynamics, in order to derive tailored control laws.

In this paper, we exploit a well-understood linear consensus technique to design reference next states for the nonlinear agents. Then, a predictive method is used to find sequences of control actions (inputs) that drive each of the nonlinear agents close to its reference state. New such states are designed, and the whole cycle is repeated in closed loop. Unlike the techniques reviewed above, our approach does not exploit the particular structure of the dynamics.

Instead, it only requires a black-box simulation model of these dynamics, and it is therefore applicable to a wide class of nonlinear agents.

The main ingredient lending this approach its generality is the global optimization algorithm used at the control stage, called *sequential optimistic optimization* [13]. This algorithm is selected since it guarantees closeness to the reference states for very general dynamics, and crucially, it only needs to sample the black-box model instead of using the explicit form of the dynamics. The distance to the optimal solution diminishes with increasing computation, at a rate modulated by a so-called near-optimality dimension [13]; few global optimization algorithms have such guarantees. Combining these properties with the nice convergence of the classical reference behavior, our main result proves a practical form of consensus.

The properties of the novel approach, including the near-optimality dimension, are analyzed for a simple type of nonlinear agent dynamics. Simulation experiments are performed for the consensus of robotic arms. It should be noted that our approach is also related to distributed model-predictive control, which has been applied for the consensus of linear agents in e.g. [14], [15].

Next, in Section II we state the consensus problem and, in Section III, introduce optimistic optimization. Section IV describes the proposed nonlinear consensus approach and gives the main result. Section V provides an analytical example, and Section VI presents a numerical experiment. Finally, Section VII concludes.

II. Problem Statement

We consider a set of M agents with decoupled nonlinear dynamics $x_{i,k+1} = f(x_{i,k}, u_{i,k})$, $i = 1, \ldots, M$, with $x_i \in \mathbb{R}^n$ and u_i belonging to a compact set U. Define also the collective state $\boldsymbol{x} = [x_1^\top, \cdots, x_M^\top]^\top \in \mathbb{R}^{Mn}$, containing the states of all the agents. The ideal goal is to achieve full-state consensus:

$$\lim_{k \to \infty} \|x_{i,k} - x_{j,k}\| = 0 \quad \forall i, j = 1, \ldots, M$$

where $\| \cdot \|$ is the Euclidean 2-norm, here as well as in the sequel. Note that our approach is not difficult to extend to the partial-state consensus case.

To achieve consensus, the agents must exchange information among themselves. An agent can receive information only from its neighbors on an interconnection graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. The set of nodes $\mathcal{V} = \{1, \ldots, M\}$ represents the agents, and the edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ are the communication

links. Denote by $\mathcal{N}_i = \{j \mid (i,j) \in \mathcal{E}\}$ the set of neighbors of node i. A path through the graph is a sequence of nodes $i_1, \ldots, i_L$ so that $(i_l, i_{l+1}) \in \mathcal{E}, 1 \leq l < L$. A graph is connected if there is a path between any pair of nodes i, j. In this paper, the communication graph is taken time-invariant and connected.

We impose the following assumptions on f.

Let $X_0 \subset \mathbb{R}^n$ be a user-defined set of interesting initial agent states, from which consensus might be required. Let $\mathbf{B}(x_c, R)$ be the smallest Euclidean ball containing X_0, which has center x_c and radius R. Then, define an enlarged set $\mathcal{X}$ as a concentric ball that is larger by a constant amount: $\mathcal{X} = \mathbf{B}(x_c, 3R + \Delta^+)$, where $\Delta^+ > 0$ is an arbitrarily small constant. Finally, denote by $\tilde{u}_k = [u_k^\top, \ldots, u_{k+K-1}^\top]^\top$ a sequence of K actions (inputs) starting at step k, and by $\tilde{f}(x_k, \tilde{u}_k)$ the resulting state after applying this sequence (K steps later). Note that we state the assumptions on f in generic terms, so no agent indices are used.

Assumption 1 (Controllability): There exists a finite K such that, for any $x, x^* \in \mathcal{X}$, there exists a sequence $\tilde{u} \in U^K$ so that:
$$\tilde{f}(x, \tilde{u}) = x^*$$
Define an *invertibility* property of the extended dynamics $\tilde{f}$, which requires that for every pair $(x, x^*) \in \mathcal{X}^2$ the sequence $\tilde{u}$ achieving $\tilde{f}(x, \tilde{u}) = x^*$ is unique. The *inverse* $\tilde{f}^{-1} : \mathcal{X}^2 \to U^K$ is defined as the mapping between state pairs and their corresponding action sequences.

Assumption 2 (Continuity and invertibility): The dynamics f are Lipschitz-continuous in $\mathcal{X} \times U$. Furthermore, the K-step dynamics $\tilde{f}$ has a unique inverse $\tilde{f}^{-1}$, which is also Lipschitz continuous in its domain $\mathcal{X}^2$.

Remarks: Controllability properties similar to those of Assumption 1 are studied by [16], where Lie-algebraic controllability conditions are provided and the size of reachable sets is characterized. In our setting, the controllable set is enlarged to $\mathcal{X}$ to deal with the errors made by the algorithm in unfavorable initial conditions, such as when most agents are close together at one edge of X_0 and one agent is diametrally opposite.

Since $\mathcal{X}$ and U are compact, the Lipschitz conditions of Assumption 2 are not overly restrictive. Invertibility is also an important topic in nonlinear system theory, and several types of invertibility conditions are provided by [17]–[19]. □

A practical requirement of our method is that K *must be known and small*. We will show that small values are appropriate in some interesting examples.

III. PRELIMINARIES: SIMULTANEOUS OPTIMISTIC OPTIMIZATION

Our method will use a predictive approach that finds the control actions with a recent branch-and-bound, sample-based optimization method called *simultaneous optimistic optimization* (SOO). The main appeal of SOO is that it allows a quantitative analysis of the convergence rate to the optimum as computation increases, for very general functions and without knowing the smoothness of the function [13]. This

has practical benefits in our consensus problem: the agents can approach arbitrarily close to the reference states, while their nonlinear dynamics can be treated as a black box.

The goal is to maximize the objective function $r : U \to \mathbb{R}$. The optimization proceeds by hierarchically partitioning the domain U. This partitioning is represented by a tree structure $\mathcal{T}$ in which each node (d, j) is labeled by a point $u_{d,j}$ and represents a subset of U denoted $U_{d,j} \ni u_{d,j}$. Here, $d \geq 0$ is the depth in the three and j is the node index at a given depth. The root of the tree represents the entire domain U, and the tree is defined so that the children of a node form a partition of the set represented by their parent. The collection of leaves of the currently explored tree is denoted by $\mathcal{L}$, and the leaves at depth d by $\mathcal{L}_d$.

An essential element in SOO is a semimetric $\ell : U \times U \to [0, \infty)$ defined over the domain U (a semimetric satisfies all properties of a metric except possibly the triangle inequality). We require the objective function to be Lipschitz w.r.t. ℓ around an optimum, and the partitioning procedure to ensure that the sets shrink with depth in the tree and are well-shaped, again in the semimetric ℓ. Formally, define a (pseudo-)diameter of each $U_{d,j}$: $\delta_{d,j} := \sup_{u \in U_{d,j}} \ell(u_{d,j}, u)$.

Assumption 3 (Optimization): The objective function and the partitioning satisfy the following conditions:

3.i There exists an optimum u^* so that:
$$r(u^*) - r(u) \leq \ell(u, u^*) \quad \forall u \in U \tag{1}$$

3.ii There exist $c > 0$ and $\gamma \in (0, 1)$ such that for any d, $\delta_{d,j} \leq c\gamma^d$ for all nodes j at depth d.

3.iii There exists a constant ν such that any subset $U_{d,j}$ contains a ball with center $u_{d,j}$ and radius $\nu c\gamma^d$ in the semimetric ℓ.

Remark: In our application to consensus the objective function will already be Lipschitz (due to Lipschitz dynamics from Assumption 2), so Assumption 3.i reduces to the rather mild existence of an optimum. Note also that for convenience, ℓ incorporates the Lipschitz constant.

Denote $\delta(d) = c\gamma^d$, the maximal diameter at depth d. The guarantees can be generalized to the case where $\delta(d)$ is any decreasing sequence [13], but only exponential partitions will be used in the sequel. □

Figure 1 exemplifies a partitioning of a scalar U.

If the diameters $\delta(d)$ were known, an upper bound on the values of all points in a set $U_{d,j}$ could be defined: $b(U_{d,j}) = r(u_{d,j}) + \delta(d)$. An optimistic optimization procedure could then be implemented by partitioning at every iteration a leaf set that has the largest upper bound, and therefore the best chance of containing the optimum. In fact, such an approach exists and is called deterministic optimistic optimization [20], but it assumes knowledge of ℓ by using $\delta(d)$. This knowledge is difficult to obtain in general, and in particular in our consensus setting it would require some information about the structure of the agent dynamics f_i.

We will therefore use the SOO algorithm, which does not require to know ℓ. Instead, at each round, SOO *simultaneously expands all potentially optimal leaf sets*: those for which the upper bound could be largest under *any* semimetric
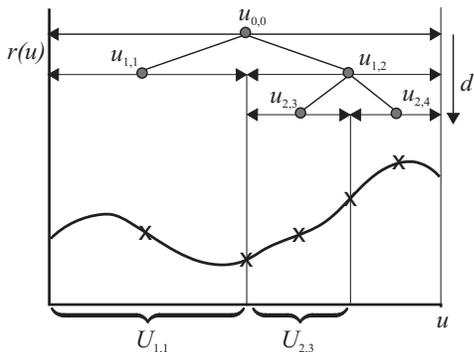
Fig. 1. Illustration of the tree structure that is used by optimistic optimization. In this example, U is an interval and binary partitions are used. Assuming a Euclidean metric, since samples are at the centers of intervals, the pseudodiameter is half the length of the interval.

ℓ. With a little thought, a set can only contain a largest upper bound if its sample value is at least as good as the values of all sets with diameters larger than its own; we say that the set is not dominated by larger sets. Since further, $\delta(d)$ decreases with d, we only have to compare with leaves higher up the tree. At each iteration t, the algorithm expands at most one leaf set at each depth. If we define $\mathcal{L}_{\leq d}$ as the set of leaf nodes having depths $d' \leq d$, then a leaf (d, j) is only expanded if $r(u_{d,j}) = \max_{(d',j') \in \mathcal{L}_{\leq d}} r(u_{d',j'})$; if there are several such leaves at d, one is chosen arbitrarily. SOO additionally limits the tree depth at each iteration t with a function $d_{\max}(t)$, a parameter of the algorithm that controls the tradeoff between deeper or more uniform exploration. A typical schedule is $d_{\max}(t) = t^a$ with $a \in (0, 1)$. Algorithm 1 summarizes SOO.

Algorithm 1 Simultaneous Optimistic Optimization

Input: r, $d_{\max}$, N, partitioning of U
1: initialize $\mathcal{T} \leftarrow U_{0,0}$; $t \leftarrow 1$
2: **loop**
3:     $r_{\max} \leftarrow -\infty$
4:     **for** $d = 0$ to $\min(\text{depth}(\mathcal{T}), d_{\max}(t))$ **do**
5:         $j^\dagger \leftarrow \arg\max_{j \in \mathcal{L}_d} r(u_{d,j})$
6:         **if** $r(u_{d,j^\dagger}) \geq r_{\max}$ **then**
7:             expand $(d, j^\dagger)$, add children to $\mathcal{T}$
8:             $r_{\max} \leftarrow r(u_{d,j^\dagger})$
9:             $t \leftarrow t + 1$
10:            **if** $t > N$ **then** stop the algorithm
11:            **end if**
12:        **end if**
13:    **end for**
14: **end loop**
Output: $\widehat{u} = \arg\max_{(d,j) \in \mathcal{T}} r(u_{d,j})$

The convergence of SOO depends on how smooth the function is in the semimetric ℓ. This is formalized as the near-optimality dimension β. First, define the near-optimal sets $U_\varepsilon = \{u \in U \mid r(u^*) - r(u) \leq \varepsilon\}$. For any ε, the packing number of U_ε is defined as the maximal number of disjoint ℓ-balls with centers in U_ε and equal radii $\nu\varepsilon$ (recall ν from

Assumption 3.iii). Then, the near-optimality dimension is the smallest β so that there exists a positive constant C, such that the packing number is smaller than $C\varepsilon^{-\beta}$ for any $\varepsilon > 0$.

Under Assumption 3, the following has been shown in [13].

Proposition 1 (SOO near-optimality): SOO returns a solution $\widehat{u}$ satisfying:[1]

$$r(u^*) - r(\widehat{u}) = \begin{cases} O(N^{-(1-a)/\beta}) & \text{if } \beta > 0 \\ O(\gamma^{\sqrt{N}/C'}) & \text{if } \beta = 0 \end{cases}$$

where $C' \geq 1$ is a problem-specific constant, and we choose $d_{\max}(t) = t^a$ when $\beta > 0$ and $d_{\max}(t) = \sqrt{t}$ when $\beta = 0$.

The optimization problem is easier to solve when the semimetric ℓ captures more closely the smoothness of the objective function r around u^*, in which case fewer balls are needed to pack U_ε – or, more precisely, their number grows more slowly with decreasing ε, meaning β is smaller. In particular, the ideal case is when U_ε shrinks at the same rate as an ℓ-ball of radius ε, which means that β is 0, i.e. that ℓ and r have the same smoothness; in this case near-optimality decreases exponentially in $\sqrt{N}$. Note also a non-obvious fact: since the guarantees hold for any semimetric satisfying the assumptions, SOO converges at the fastest rate allowed by any such semimetric.

IV. NONLINEAR CONSENSUS ALGORITHM AND ANALYSIS

Our nonlinear consensus approach works over iterations consisting of the following two stages. Once every K steps, each agent i computes a reference (desired) next state, which is a weighted average of the current states of its neighbors $\mathcal{N}_i$. Then, the agent uses SOO to optimize a sequence of actions so that after K steps, it approaches the reference state as closely as possible. All agents apply their sequences in open loop and, once the K steps have elapsed, close the loop and repeat the process.

Formally, at step k – a multiple of K – each agent i computes a reference state with:

$$x^*_{i,k+K} = \sum_{j=1}^{M} p_{i,j} x_{j,k} \tag{2}$$

where:

$$p_{i,j} \in [0, 1], p_{i,j} = p_{j,i} \quad \forall i, j$$
$$p_{i,j} \neq 0 \text{ iff } i = j \text{ or } (i, j) \in \mathcal{E}$$
$$\sum_{j=1}^{M} p_{i,j} = 1 \quad \forall i \tag{3}$$

Note that (2) can easily be implemented in a decentralized way, since each agent's reference state only uses the state of its neighbors. Define also a matrix $P \in \mathbb{R}^{M \times M}$ whose components are $p_{i,j}$ – a symmetric stochastic matrix in which any elements corresponding to disconnected agents are 0. This reference state is inspired from classical consensus algorithms [1], which are well studied and have useful properties for some classes of linear agents.

---

[1]The so-called 'big-O' notation $g = O(h)$ means that g asymptotically grows at most as fast as h.

Then, each agent uses SOO to solve the optimal control problem:

$$\tilde{u}_{i,k}^* = \arg\max_{\tilde{u}_{i,k}} -\|\tilde{f}(x_{i,k}, \tilde{u}_{i,k}) - x_{i,k+K}^*\|$$

(recall the definitions of the K-step actions and dynamics). A near-optimal extended action $\widehat{u}_{i,k}$ is found, which will get close to the reference state. Algorithm 2 summarizes the overall procedure, for the case where all the agents have the same optimization parameters.

Algorithm 2 Nonlinear consensus at agent i

Input: row P_i of P, budget N, partitioning of U, $d_{\max}$
1: $k \leftarrow 0$
2: **loop**
3:      measure $x_{i,k}$ and send it to $j \in \mathcal{N}_i$
4:      receive $x_{j,k}$ from $j \in \mathcal{N}_i$
5:      set reference state $x_{i,k+K}^* = \sum_{j \in \{\{i\} \cup \mathcal{N}_i\}} p_{i,j} x_{j,k}$
6:      $r_{i,k}(\tilde{u}_{i,k}) := -\|\tilde{f}(x_{i,k}, \tilde{u}_{i,k}) - x_{i,k+K}^*\|$
7:      $\widehat{u}_{i,k} \leftarrow \text{SOO}(r_{i,k}, d_{\max}, N, \text{partitioning of } U)$
8:      apply $\widehat{u}_{i,k}$ in open loop
9:      $k \leftarrow k + K$
10: **end loop**

Regarding parameter choices, P should be selected to impose a desired consensus behavior – of course, while still obeying the conditions explained above. A computationally simple strategy to select it is presented in [21]. In this method, the agents first agree on some value $\alpha \in (0, \frac{1}{\max_i |\mathcal{N}_i|})$ where $|\cdot|$ denotes set cardinality. Each agent i sets $p_{i,i} = 1 - \alpha|\mathcal{N}_i|$ and $p_{i,j} = \alpha, \forall j \in \mathcal{N}_i$. This strategy will be used in the sequel. Other methods to select P in a decentralized way are given by [22], [23].

The budget N should be chosen large enough – by default, as large as the computational resources of the agents allow. The partitioning of U will typically be defined as a collection of hyperboxes in a Km-dimensional space, where m denotes the number of action variables. A large hyperbox contains the whole action space, and each node expansion corresponds to splitting the parent hyperbox into 2^{Km} subboxes, in half along all dimensions. It should be clear now why K must be small: with this type of exponentially shrinking partitioning, the complexity of the algorithm is exponential in K.

A default choice for the maximum depth function $d_{\max}(t)$ is $\sqrt{t}$, expecting that a semimetric that yields a near-optimality dimension of 0 exists (like in the upcoming example of Section V).

A final assumption is made on the finiteness of the near-optimality dimension β (defined in Section III). The optimization problem of an agent is parameterized by its index i and the current collective state x. While the optimization problem also depends on the reference state, this state is computed from x using the fixed P, so it does not have to be included in the dependence. Denote by $\beta(i, x)$ the optimization-problem dependent near-optimality dimension.

Assumption 4: For any $i \in \{1, \dots, M\}$ and $x \in \mathcal{X}^M$, a finite $\beta(i, x)$ exists.

This assumption is technical, and we expect most problems to satisfy it in practice.

In a straightforward fashion, the guarantees of SOO imply the following result, which we will not prove here.

Lemma 2 (Uniformly small error): Let $d_{\max}(t) = \sqrt{t}$. Under Assumptions 1–4, for any desired error $\varepsilon^* > 0$, there exists a finite N using which all agents approach their reference states closer than ε^*, for any collective state $x \in \mathcal{X}^M$.

Our main result follows.

Theorem 3: Under Assumptions 1–4, for sufficiently large N, Algorithm 2 achieves practical consensus. This means there exist k_0 and an inter-agent disagreement Δ^{a} so that for all $k \geq k_0$, the distance $\|x_{i,k} - x_{j,k}\| \leq \Delta^{\mathrm{a}}$ for all i, j.

Proof: The proof will be developed in three parts. In part (i), we will show that the agents achieve practical *average* consensus at steps multiple of K, under the assumption that they remain controllable at every such step. This notion of consensus means that there exist l_0 and Δ so that for all $l \geq l_0$, $\|x_{i,lK} - \bar{x}\| \leq \Delta$ for all i, where $\bar{x} \in \mathbb{R}^n$ is the average of the initial states of the agents, $\bar{x} = \frac{1}{M}\sum_{i=1}^{M} x_{i,0}$. In part (ii), the agents are shown to actually remain controllable. Part (iii) proves practical consensus at steps in-between multiples of K.

(i) Matrix P is a symmetric stochastic matrix associated to a connected graph. Therefore, by iterating the ideal protocol $x_{i,(l+1)K}^* = \sum_{j=1}^{M} p_{i,j} x_{j,lK}^*$ starting from $x_{i,0}$, the states would converge exponentially to $\bar{x}$ – achieving average consensus. While our algorithm in fact implements an approximate version of this protocol, since the approximation error can be made arbitrarily small, practical average consensus is still achieved.

Specifically, let λ be the largest eigenvalue of P that is different from 1 (due to its structure, P will always have an eigenvalue of 1). Then, $0 < \lambda < 1$ and:

$$\|x_{(l+1)K}^* - \bar{x}\| = \lambda\|x_{lK} - \bar{x}\| \tag{4}$$

Further, Lemma 2 ensures that each agent reaches its reference state with at most ε^* error, leading to:

$$\|x_{lK} - x_{lK}^*\| \leq \sqrt{M}\varepsilon^*, \forall l \tag{5}$$

Choose a large enough budget N to have $\varepsilon^* \leq \frac{(\rho - \lambda)\Delta}{\sqrt{M}}$, for some $\Delta > 0$ and $\rho \in (\lambda, 1)$ (we can choose any values satisfying these conditions). Combining (5) and (4), Δ-practical consensus with a convergence rate ρ will be ensured.

Define $\bar{x} = [\bar{x}, \dots, \bar{x}]^\top$, a collective average state. Assume first that the initial disagreement is larger than Δ. Then, for all x_{lK} such that $\|x_{lK} - \bar{x}\| > \Delta$ one has:

$$\begin{aligned}
\|x_{(l+1)K} - \bar{x}\| &\leq \|x_{(l+1)K} - x_{(l+1)K}^*\| + \|x_{(l+1)K}^* - \bar{x}\| \\
&\leq \sqrt{M}\varepsilon^* + \lambda\|x_{lK} - \bar{x}\| \\
&\leq (\rho - \lambda)\Delta + \lambda\|x_{lK} - \bar{x}\| \\
&< \rho\|x_{lK} - \bar{x}\|
\end{aligned}$$

This exponential decay ensures that after a finite $l_0 = \left\lceil \log_\rho \frac{\Delta}{\|x_0 - \bar{x}\|} \right\rceil$, the distance to $\bar{x}$ will drop below Δ.

Once this is true, i.e., if $\|\boldsymbol{x}_{lK} - \bar{\boldsymbol{x}}\| < \Delta$, then:

$$
\begin{aligned}
\|\boldsymbol{x}_{(l+1)K} - \bar{\boldsymbol{x}}\| &\leq \|\boldsymbol{x}_{(l+1)K} - \boldsymbol{x}^*_{(l+1)K}\| + \|\boldsymbol{x}^*_{(l+1)K} - \bar{\boldsymbol{x}}\| \\
&\leq \sqrt{M}\varepsilon^* + \lambda\|\boldsymbol{x}_{lK} - \bar{\boldsymbol{x}}\| \\
&\leq (\rho - \lambda)\Delta + \lambda\Delta = \rho\Delta < \Delta
\end{aligned}
$$

so the state will remain at distance Δ from $\bar{x}$. If the initial disagreement is already below Δ, then $l_0 = 0$ and the derivation of the exponential decay is no longer needed.

(ii) To ensure controllability, the states of all the agents must remain inside the controllable set $\mathcal{X}$ at each step lK. Recall first the definitions of $X_0 \subseteq \mathbf{B}(x_c, R)$ and $\mathcal{X} = \mathbf{B}(x_c, 3R + \Delta^+)$. Define $\Delta_0 = \|\boldsymbol{x}_0 - \bar{\boldsymbol{x}}\|$, the *collective* initial disagreement, and $\Delta_0^{\mathrm{a}} = \max_i \|x_{i,0} - \bar{x}\|$, the *per-agent* initial disagreement. Take a desired collective disagreement of $\Delta \leq 2R$; part (i) allows imposing any $\Delta > 0$, and since $2R$ upper-bounds the diameter of X_0, a larger value makes little sense.

If $\Delta \geq \Delta_0$, then already, for any $l \geq 0$, $\|x_{i,lK} - \bar{x}\| \leq \|\boldsymbol{x}_{lK} - \bar{\boldsymbol{x}}\| \leq \Delta$. Thus the agents remain in $\mathbf{B}(\bar{x}, \Delta)$ and since $\bar{x} \in \mathbf{B}(x_c, R)$, they remain in $\mathbf{B}(x_c, 3R) \subset \mathcal{X}$.

Now if $\Delta < \Delta_0$, then for any l and i we have:

$$
\begin{aligned}
\|x_{i,(l+1)K} - \bar{x}\| &\leq \|x_{i,(l+1)K} - x^*_{i,(l+1)K}\| + \|x^*_{i,(l+1)K} - \bar{x}\| \\
&\leq \varepsilon^* + \max_j \|x_{j,lK} - \bar{x}\| \\
&\leq (l+1)\varepsilon^* + \Delta_0^{\mathrm{a}} \qquad (6)
\end{aligned}
$$

where the second inequality implies the third by induction. Furthermore, by part (i), as long as the collective disagreement is larger than Δ it decreases exponentially:

$$
\|x_{i,lK} - \bar{x}\| \leq \|\boldsymbol{x}_{lK} - \bar{\boldsymbol{x}}\| \leq \rho^l \Delta_0
$$

This allows taking a finite l_1 so that $\rho^{l_1}\Delta_0 \leq \max\{\Delta, \Delta_0^{\mathrm{a}}\}$, from which $\|x_{i,lK} - \bar{x}\| \leq \max\{\Delta, \Delta_0^{\mathrm{a}}\}\ \forall i, l \geq l_1$. Combining this with (6), we have $\|x_{i,lK} - \bar{x}\| \leq \max\{\Delta, l_1\varepsilon^* + \Delta_0^{\mathrm{a}}\}, \forall i, l \geq 0$. Therefore, finally imposing $\varepsilon^* \leq \Delta^+/l_1$ and noticing that $\Delta_0^{\mathrm{a}} \leq 2R$, the states are guaranteed to remain in $\mathcal{X}$.

(iii) Take $l \geq l_0$. Then $\forall i, j$, $\|x_{i,lK} - x_{j,lK}\| \leq 2\Delta$, due to $\|\boldsymbol{x}_{lK} - \bar{\boldsymbol{x}}\| \leq \Delta$ and the triangle inequality. Due to Assumption 2, we have $\tilde{u}_{i,lK} = \tilde{f}^{-1}(x_{i,lK}, x_{i,(l+1)K})$, $\tilde{u}_{j,lK} = \tilde{f}^{-1}(x_{j,lK}, x_{j,(l+1)K})$ and denoting the Lipschitz constant of $\tilde{f}^{-1}$ by L_{-1}:

$$
\begin{aligned}
&\|\tilde{u}_{i,lK} - \tilde{u}_{j,lK}\| \\
&\leq L_{-1}(\|x_{i,lK} - x_{j,lK}\| + \|x_{i,(l+1)K} - x_{j,(l+1)K}\|) \\
&\leq 4L_{-1}\Delta \qquad (7)
\end{aligned}
$$

Then, at steps $lK + k$ with $k = 1, \ldots, K - 1$, we have:

$$
\begin{aligned}
\|u_{i,lK+k} - u_{j,lK+k}\| &\leq \sqrt{m}\,\|\tilde{u}_{i,lK} - \tilde{u}_{j,lK}\|_\infty \\
&\leq 4\sqrt{m}L_{-1}\Delta
\end{aligned}
$$

Denote $L' = 4\sqrt{m}L_{-1}$, and L the Lipschitz constant of f. Finally, by a straightforward derivation we get:

$$
\|x_{i,lK+k} - x_{j,lK+k}\| \leq \left(L^k + L'\sum_{k'=1}^{k} L^{k'}\right)\Delta
$$

Taking $\Delta^{\mathrm{a}} = \max\{2, L^k + L'\sum_{k'=1}^{k} L^{k'}\}\Delta$, the desired result has been obtained. ∎

Remarks: From the proof, the agents exponentially approach and then remain inside a region of size Δ around the average state $\bar{x}$ – but only at steps multiple of K. In-between these steps and above $l_0 K$, the agents are only close to each other but not necessarily to $\bar{x}$. Indeed, the region containing the states can travel in the state space, with the constraint that it must always return around $\bar{x}$ at each multiple of K. Note that the region may grow somewhat in-between the multiples of K, but its size remains proportional to Δ. Figure 2 illustrates these guarantees. □
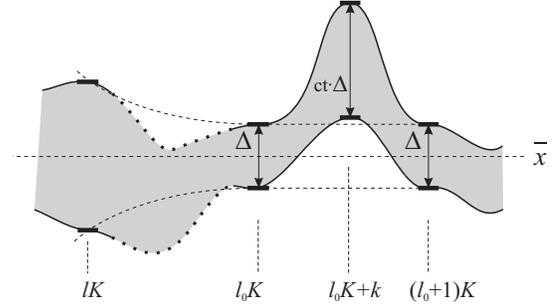
Fig. 2. Symbolic illustration of consensus guarantees. The agent states will stay within the shaded area. The dotted contour indicates regions where the analysis does not constrain the states.

V. ANALYSIS FOR A SIMPLE TYPE OF NONLINEAR AGENTS

To gain more insight into the approach, we clarify the meaning of the assumptions and find the near-optimality dimension for a simple type of nonlinear agents. It must be emphasized that this example is analyzed here in detail for illustrative purposes, using the structure of the agent dynamics. Indeed, the whole point of our algorithm is to avoid the need to use this structure. In practice, the user does not need to consider all these details.

Consider that the agents have the following dynamics:

$$
\begin{aligned}
y_{k+1} &= y_k + z_k \\
z_{k+1} &= z_k^2 + z_k + u_k
\end{aligned}
$$

where again we do not include the agent index. Here the state is $x = [y, z]^\top$. The set X_0 of interesting initial states considered is the ball $\mathbf{B}(0, R)$. It follows that $\mathcal{X} = \mathbf{B}(0, 3R + \Delta^+)$ for an arbitrary $\Delta^+ > 0$. Denote $R_{\mathcal{X}} = 3R + \Delta^+$.

Taking $K = 2$, the 2-step dynamics are:

$$
\begin{aligned}
y_{k+2} &= y_k + z_k^2 + 2z_k + u_k \\
z_{k+2} &= (z_k^2 + z_k + u_k)^2 + (z_k^2 + z_k + u_k) + u_{k+1}
\end{aligned}
$$

Denote u_k by v and u_{k+1} by w; these are the optimized variables. Actions to reach a desired state $x^*_{k+2} = [y^*, z^*]^\top$ can be explicitly written:

$$
\begin{aligned}
v^* = u_k^* &= y^* - y_k - z_k^2 - 2z_k \\
w^* = u_{k+1}^* &= z^* + z_k + y_k - y^* + (y^* - y_k - z_k)^2
\end{aligned}
\qquad (8)
$$

Since in $\mathcal{X}$, $|y| \leq R_\mathcal{X}$ and $|z| \leq R_\mathcal{X}$, to ensure Assumption 1 with $K = 2$ it is sufficient to take the set of actions $\tilde{U}_k = \left\{ \tilde{u}_k = [v, w]^\top \mid |v| \leq R_\mathcal{X}^2 + 4R_\mathcal{X}, |w| \leq 9R_\mathcal{X}^2 + 4R_\mathcal{X} \right\}$. For simplicity, we use instead a larger set $\tilde{U}_k = [-R_U, R_U]^2$ with $R_U \geq 9R_\mathcal{X}^2 + 4R_\mathcal{X}$.

Note that in fact (8) gives the inverse of the 2-step dynamics, which are clearly Lipschitz in their compact domain, as are the direct dynamics; so Assumption 2 also holds.

The optimal control problem requires maximizing:

$$r(\tilde{u}_k) = -\|x_{k+2} - x_{k+2}^*\|$$

To find a semimetric ℓ satisfying Assumption 3.i, note that:

$$\begin{aligned}
r(\tilde{u}_k^*) - r(\tilde{u}_k) &= \|x_{k+2} - x_{k+2}^*\| \leq \|x_{k+2} - x_{k+2}^*\|_1 \\
&= |v - v^*| + |v - v^* + w - w^* \\
&\quad + (z_k^2 + z_k + v)^2 - (z_k^2 + z_k + v^*)^2| \\
&\leq |v - v^*| + |v - v^* + w - w^* \\
&\quad + (2(z_k^2 + z_k) + v + v^*)(v - v^*)| \\
&\leq [2(R_\mathcal{X}^2 + R_\mathcal{X} + R_U) + 2]|v - v^*| + |w - w^*| \\
&\leq [2(R_\mathcal{X}^2 + R_\mathcal{X} + R_U) + 2] \|\tilde{u}_k - \tilde{u}_k^*\|_1
\end{aligned} \tag{9}$$

where we used the fact that $r(\tilde{u}_k^*) = 0$ and rewrote x_{k+2}^* using the optimal actions from above. So, an appropriate semimetric (in fact, a metric) is $\ell(\tilde{u}_k, \tilde{u}_k^*) = \overline{\gamma} \|\tilde{u}_k - \tilde{u}_k^*\|_1$ with some 'Lipschitz constant' $\overline{\gamma}$ greater than the multiplying factor above.

If we create a natural, exponential partitioning of the two-dimensional, square space $\tilde{U}_k$, by recursively splitting it in half along each of the two actions (resulting in 4 new, smaller squares at each expansion), then Assumptions 3.ii and 3.iii are satisfied with metric ℓ.

To find the near-optimality dimension β, the following sets must be packed with ℓ-balls:

$$\tilde{U}_\varepsilon = \{\tilde{u} \mid r(\tilde{u}^*) - r(\tilde{u}) \leq \varepsilon\}$$

The left hand side of the inequality is lower-bounded as follows:

$$\begin{aligned}
r(\tilde{u}_k^*) - r(\tilde{u}_k) &= \|x_{k+2} - x_{k+2}^*\| \geq \frac{1}{\sqrt{2}} \|x_{k+2} - x_{k+2}^*\|_1 \\
&= \frac{1}{\sqrt{2}} \big(|v - v^*| + |2(z_k^2 + z_k)(v - v^*) \\
&\quad + w - w^* + v^2 - v^{*2}|\big) \\
&\geq \frac{1}{2\sqrt{2}(z_k^2 + z_k) + \sqrt{2}} (|v - v^*| + |w - w^* + v^2 - v^{*2}|)
\end{aligned}$$

It can be shown that for v and w sufficiently close to their optimal values, there exist $c_1 > 0, c_2 > 0$ so that:

$$|v - v^*| + |w - w^* + v^2 - v^{*2}| \geq c_1|v - v^*| + c_2|w - w^*|$$

Then $\tilde{U}_\varepsilon$ is included in:

$$\tilde{U}_\varepsilon' = \left\{\tilde{u} \,\middle|\, \frac{\min\{c_1, c_2\}}{2\sqrt{2}(z_k^2 + z_k) + \sqrt{2}} (|v - v^*| + |w - w^*|) \leq \varepsilon \right\}$$

Further, if v and w are *not* close to v^* and w^*, then ε is large, and any such nonasymptotic behavior which does not obey

the inequality can be taken care of by making the constant C larger in the definition of β. It follows that $\tilde{U}_\varepsilon'$ is packed by a constant number of ℓ-balls, which means the same is true of the smaller set $\tilde{U}_\varepsilon$. So, finally, the near-optimality dimension is 0, meaning Assumption 4 is satisfied, and also that the optimization problem is simple.

VI. Application to Nonlinear Robot Arms

As a representative problem involving nonlinear dynamics, the consensus of two-link robotic arms operating in a horizontal plane is considered. This type of consensus problem can appear in decentralized manipulation and teleoperation.

Three robots are connected on the graph shown at the top of Figure 3. The state variables of a robot are the angles and angular velocities of the two links, $x_i = [\theta_{i,1}, \dot{\theta}_{i,1}, \theta_{i,2}, \dot{\theta}_{i,2}]^\top$, and the actions are the two motor torques $u_i = [\tau_{i,1}, \tau_{i,2}]^\top$. The model is standard so we do not provide the details here, instead referring the reader to [24] where the equations and physical parameters are given. The sampling time is $T_s = 0.05\,\text{s}$, and time discretization is performed with the fourth-order Runge-Kutta method (10 steps per sampling time). Full-state consensus is required, from the initial states $x_{1,0} = [-\pi, 0, -\pi/2, 0]^\top, x_{2,0} = 0, x_{3,0} = [\pi/2, 0, \pi, 0]^\top$.

We choose $K = 2$ and the two-step action space $\tilde{U} = [-50, 50]\text{Nm} \times [-50, 50]\text{Nm}$. The neighbor weight $\alpha = 0.45$, close to its upper limit 0.5, and the budget $N = 300$. Figure 3, bottom shows the resulting trajectories of the agents.

The algorithm easily deals with these nonlinear dynamics. Since the desired agent state $x_{i,k+2}^*$ is designed using only the current state $x_{i,k}$, without exploiting the structure of the dynamics, the resulting trajectory is not smooth. Further, in this simulation the average consensus state has zero velocity and is thus an equilibrium, which allows the states at all steps to eventually reach a *constant* ball around it. If the consensus
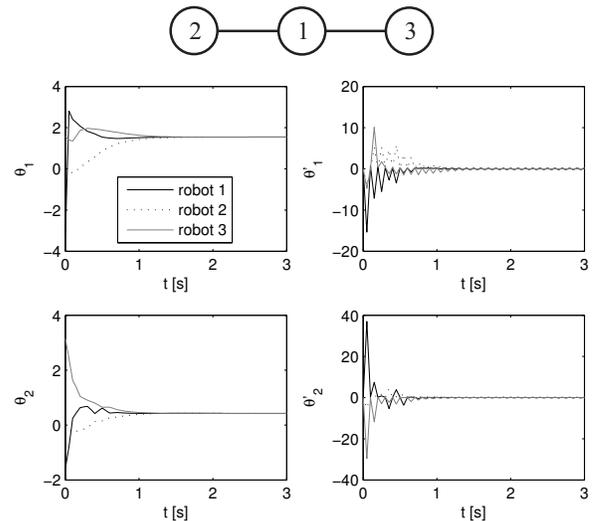
Fig. 3. Consensus of robotic arms. Top: communication graph. Bottom: state trajectories.

state were not an equilibrium, the ball would travel in the state space, returning to the average state at each multiple of $K = 2$, as illustrated in Figure 2.

VII. Conclusions

We presented a consensus method for general nonlinear agents, which guarantees practical consensus while treating the agent dynamics as a black box. The main component lending this method its generality is an optimistic optimization algorithm used to find the control actions. An illustrative example was analyzed, and the algorithm achieved practical consensus for nonlinear robot arms.

The main limitation of the method is that it works in increments of a small number K of steps, and a natural extension that avoids this limitation would be to apply long- or infinite-horizon optimistic predictive methods, called optimistic planning [20], [25]. The idea of approximating proven consensus methods with nonlinear optimization should extend to other open problems in nonlinear consensus, such as maintaining connectivity of time-varying communication graphs.

References

[1] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[2] W. Ren and R. W. Beard, *Distributed Consensus in Multi-Vehicle Cooperative Control: Theory and Applications*, ser. Communications and Control Engineering. Springer, 2008.

[3] W. Ren and R. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Transactions on Automatic Control*, pp. 655–661, 2005.

[4] L. Moreau, "Stability of multiagent systems with time-dependent communication links," *IEEE Transactions on Automatic Control*, vol. 50, pp. 169–182, 2005.

[5] J. Tsitsiklis, D. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE Transactions on Automatic Control*, vol. 31, pp. 803–812, 1986.

[6] L. Fang, P. Antsaklis, and A. Tzimas, "Asynchronous consensus protocols: Preliminary results, simulations and open questions," in *Proceedings of the Joint 44th IEEE Conference on Decision and Control, and the European Control Conference*, 2005, pp. 2194–2199.

[7] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, pp. 1520–1533, 2004.

[8] W. Michiels, I.-C. Morarescu, and S.-I. Niculescu, "Consensus problems with distributed delays, with application to traffic flow models," *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 77–101, 2009.

[9] H. Su, G. Chen, X. Wang, and Z. Lin, "Adaptive second-order consensus of networked mobile agents with nonlinear dynamics," *Automatica*, vol. 47, no. 2, pp. 368–375, 2011.

[10] J. Zhou, X. Wu, W. Yu, M. Small, and J. Lu, "Flocking of multi-agent dynamical systems based on pseudo-leader mechanism," *Systems & Control Letters*, vol. 61, no. 1, pp. 195–202, 2012.

[11] H. Tanner, A. Jadbabaie, and G. Pappas, "Flocking in teams of nonholonomic agents," in *Cooperative Control*, ser. Lecture Notes in Control and Information Sciences, V. Kumar, N. Leonard, and A. Morse, Eds. Springer, 2005, vol. 309, pp. 458–460.

[12] J. Mei, W. Ren, and G. Ma, "Distributed coordinated tracking with a dynamic leader for multiple euler-lagrange systems," *IEEE Transactions on Automatic Control*, vol. 56, no. 6, pp. 1415–1421, 2011.

[13] R. Munos, "Optimistic optimization of a deterministic function without the knowledge of its smoothness," in *Advances in Neural Information Processing Systems 24*, J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. C. N. Pereira, and K. Q. Weinberger, Eds., 2011, pp. 783–791.

[14] T. Keviczky and K. Johansson, "A study on distributed model predictive consensus," in *Proceedings 17th IFAC World Congress (IFAC-08)*, Seoul, Korea, 6–11 July 2008, pp. 1516–1521.

[15] G. Ferrari-Trecate, L. Galbusera, M. Marciandi, and R. Scattolini, "Model predictive control schemes for consensus in multi-agent systems with single- and double-integrator dynamics," *IEEE Transactions on Automatic Control*, vol. 54, no. 11, pp. 2560–2572, 2009.

[16] B. Jakubczyk and E. D. Sontag, "Controllability of nonlinear discrete-time systems: A lie-algebraic approach," *SIAM Journal of Control and Optimization*, vol. 28, pp. 1–33, 1990.

[17] J. Grizzle, "A linear algebraic framework for the analysis of discrete-time nonlinear systems," *SIAM Journal of Control and Optimization*, vol. 31, no. 4, pp. 1026–1044, 1993.

[18] Y. Zheng and R. J. Evans, "Minimal order discrete-time nonlinear system inversion," in *Proceedings 15th IFAC World Congress*, Barcelona, Spain, 21–26 July 2002, pp. 1119–1125.

[19] M. Fliess, "Reversible linear and nonlinear discrete-time dynamics," *IEEE Transactions on Automatic Control*, vol. 37, no. 8, pp. 1144–1153, 1992.

[20] R. Munos, "The optimistic principle applied to games, optimization and planning: Towards foundations of monte-carlo tree search," *Foundations and Trends in Machine Learning*, 2012, submitted.

[21] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transaction on Automatic Control*, vol. 49, pp. 1520–1533, 2004.

[22] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *System and Control Letters*, vol. 53, pp. 65–78, 2004.

[23] A. Olshevsky and J. Tsitsiklis, "Convergence speed in distributed consensus and averaging," *SIAM Journal of Control and Optimization*, vol. 48, no. 1, pp. 33–55, 2009.

[24] L. Buşoniu, D. Ernst, B. De Schutter, and R. Babuška, "Approximate dynamic programming with a fuzzy parameterization," *Automatica*, vol. 46, no. 5, pp. 804–814, 2010.

[25] L. Buşoniu, R. Munos, and R. Babuška, "A review of optimistic planning in Markov decision processes," in *Reinforcement Learning and Adaptive Dynamic Programming for Feedback Control*, F. Lewis and D. Liu, Eds. Wiley, 2012.