

Consensus in opinion dynamics as a repeated game [★]

Dario Bauso ^{a,b}, Mark Cannon ^c,

^a*Department of Automatic Control and Systems Engineering, The University of Sheffield, Mappin Street Sheffield, S1 3JD, United Kingdom*

^b*Dipartimento di Ingegneria Chimica, Gestionale, Informatica, Meccanica, Università di Palermo, V.le delle Scienze, 90128 Palermo, Italy*

^c*Department of Engineering Science, University of Oxford, Parks Road, Oxford, OX1 3PJ, UK*

Abstract

We study an n -agent averaging process with dynamics subject to controls and adversarial disturbances. The model arises in multi-population opinion dynamics with macroscopic and microscopic intertwined dynamics. The averaging process describes the influence from neighbouring populations, whereas the input term indicates how the distribution of opinions in the population changes as a result of dynamical evolutions at a microscopic level (individuals' changing opinions). The input term is obtained as the vector payoff of a two player repeated game. We study conditions under which the agents achieve robust consensus to some predefined target set. Such conditions build upon the approachability principle in repeated games with vector payoffs.

Key words: Game theory; networks; allocations; robust receding horizon control.

1 Introduction

We consider an n -agent averaging process in which each agent is described by a dynamic system with controlled and uncontrolled inputs, the latter being adversarial disturbances.

We specialize the model to multi-population opinion dynamics. The averaging process describes the influence from neighbouring populations, whereas the input term indicates how the distribution of opinions in the population changes as a result of dynamical evolutions at a microscopic level (individuals' changing opinions). The input term is obtained as the vector payoff of a two player repeated game [6, 8, 17]. Motivations for the dynamics can be found in coalitional games with Transferable Utilities (TU games) [26], bargaining [7, 19], consensus [18, 20, 23, 24], opinion dynamics [1, 2, 3, 4, 9, 10, 12, 13, 16, 22, 25] and in multi-population games with macroscopic and microscopic dynamics.

The main contribution of this paper is to introduce a

distributed multi-stage receding horizon control strategy that ensures the existence of invariant and contractive sets for the collective dynamics, and which can be used to enforce convergence of consensus to a specified set.

This paper improves [5] as it links the model to opinion dynamics and multi-population games, it includes exponential stability and identifies regions of attraction that are dependent and independent of the horizon, and it provides new numerical results. An alternative way to deal with the problem is to include a deterministic adversarial disturbance in the spirit of set inclusion theory [14, 15].

The paper is organized as follows. In Section 2 we formulate the problem. In Section 3 we discuss motivations. Section 4 gives the main control theoretic results. Numerical illustrations are presented in Section 5, and concluding remarks are provided in Section 6.

Notation. We denote the Euclidean norm of a vector x as $\|x\|$, and we use a_j^i or $[A]_{ij}$ to denote the ij th entry of a matrix A . We say that $A \in \mathbb{R}^{n \times n}$ is row-stochastic if $a_j^i \geq 0$ for all $i, j \in \{1, \dots, n\}$ and $\sum_{j=1}^n a_j^i = 1$ for all $i \in \{1, \dots, n\}$. Matrix A is doubly stochastic if both A and its transpose A^\top are row-stochastic. We use $|S|$ for the cardinality of a given finite set S . We write $P_X[x]$ to

[★] A preliminary conference version of this work was presented in the 2014 IFAC World Congress [5]. Corresponding author D. Bauso

Email addresses: `d.bauso@sheffield.ac.uk` (Dario Bauso), `mark.cannon@eng.ox.ac.uk` (Mark Cannon).

denote the projection of a vector x on a set X , and we write $|x|_X$ for the distance from x to X , i.e., $P_X[x] = \arg \min_{y \in X} \|x - y\|$ and $|x|_X = \|x - P_X[x]\|$, respectively.

2 Model and problem set-up

For each i in a set $N = \{1, \dots, n\}$, agent i is characterized by a state $x_i(t) \in \mathbb{R}^{\tilde{n}}$. At every time t this state evolves according to a distributed averaging process representing the interaction of the agent with its neighbours, and under the influence of an input variable $u_i(t)$.

Formally, the state $x_i(t)$ of agent i evolves as follows

$$x_i(t+1) = \sum_{j=1}^n a_j^i(t) x_j(t) + u_i(t), \quad t = 0, 1, \dots \quad (1)$$

where $a^i = (a_1^i, \dots, a_n^i) \in \mathbb{R}^n$ is a vector of nonnegative scalar weights relating to the *communication graph* $\mathcal{G}(t) = (N, \mathcal{E}(t))$. A link $(j, i) \in \mathcal{E}(t)$ exists (and hence $a_j^i(t) \neq 0$) if agent j is a neighbour of agent i at time t .

For each agent $i \in N$, the input $u_i(\cdot)$ is the payoff of a repeated two-player game between player i (Player A) and an (external) adversary (Player B). Let S_A and S_B be the finite sets of actions of players A and B respectively and let us denote the set of mixed action pairs by $\Delta(S_A) \times \Delta(S_B)$ (set of probability distributions on S_A and S_B). For any pair of mixed strategies $(p(t), q(t)) \in \Delta(S_A) \times \Delta(S_B)$ for player A and B at time t , the expected payoff is

$$\begin{cases} u_i(t) = \sum_{j \in S_A, k \in S_B} p_j^i(t) \phi(j, k) q_k^i(t), \\ \sum_{j \in S_A} p_j^i(t) = 1, \sum_{k \in S_B} q_k^i(t) = 1, p_j^i, q_k^i \geq 0. \end{cases} \quad (2)$$

Essentially, in the above game $\phi(j, k) \in \mathbb{R}^{\tilde{n}}$ is the vector payoff when players A and B play pure strategies $j \in S_A$ and $k \in S_B$ respectively. Figure 1 illustrates the continuous action sets for the two players, for the case that $S_A = \{1, 2, 3\}$ and $S_B = \{1, 2, 3\}$.

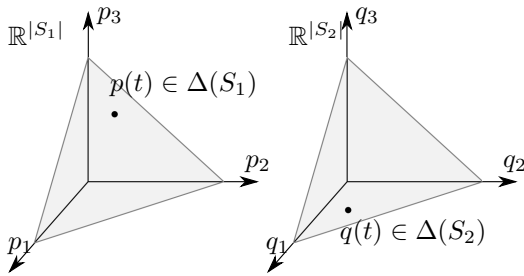


Fig. 1. Spaces of mixed strategies for the two players.

Let $X \subset \mathbb{R}^{\tilde{n}}$ be a closed convex target set, and assume that player A seeks to drive the state $x_i(t)$ to X , while player B tries to push the state far from it. The resulting

strategy can be formulated as the solution of a robust optimization problem, with one player minimizing and the other maximizing the distance of the state from X .

In compact form the problem with finite horizon $[0, T]$ to be solved by agent i takes the form:

$$\begin{aligned} & \min_{p^i(0)} \max_{q^i(0)} \cdots \min_{p^i(T-1)} \max_{q^i(T-1)} \sum_{t=0}^T |x_i(t)|_X^2 \\ & \left. \begin{aligned} & p^i(t) \in \Delta(S_A), \quad q^i(t) \in \Delta(S_B), \\ & x_i(t+1) = y_i(t) + u_i(t), \\ & u_i(t) = \sum_{j \in S_A, k \in S_B} p_j^i(t) \phi(j, k) q_k^i(t) \end{aligned} \right\} t = 0, \dots, T-1 \end{aligned} \quad (3)$$

where $y_i(t)$ is the *space average* defined as

$$y_i(t) = \sum_{j=1}^n a_j^i(t) x_j(t). \quad (4)$$

Through the above problem we can study contractivity and invariance of sets for the collective dynamics (1)-(2). In the following we simplify notation and drop the dependence on i of p and q .

3 Multi-population opinion dynamics

A simple model of opinion dynamics is derived from a classical model of consensus dynamics that also arises in the Kuramoto oscillator model [22]. In this perspective, the dynamic model (1) appears as a discrete-time model of a consensus problem [21], in which the coupling term accounts for emulation (an individual's opinion is influenced by those of its neighbours), and which includes an additional input term (the natural opinion changing rate). In addition, the target set X can be used to enforce consensus. For instance we can set $X := \{x\} \subset \mathbb{R}^{\tilde{n}}$, in which case $\lim_{t \rightarrow \infty} x_i(t) = x$ for all $i \in N$ also implies that $\lim_{t \rightarrow \infty} x_i(t) - x_j(t) = 0$ for all $i, j \in N$. Note that this notion of consensus may be in general different from the consensus studied in distributed algorithms [21].

In the following we consider n distinct populations of agents interacting according to a predefined topology. Let the collective state be $\xi(t) = (x_1(t), \dots, x_n(t))$, which we now see as a collection of n macro-states. For each population, and at every time $t \in [0, T]$, a probability distribution function $x_i(t)$, $i \in N$, describes the probability distribution of agents over a discrete set of micro-states. In other words, consider a finite discrete space of micro-states $\{1, \dots, \tilde{n}\}$, and let a probability distribution function be given, $m_i : \{1, \dots, \tilde{n}\} \times [0, +\infty) \rightarrow [0, 1]$, $(j, t) \mapsto m_i(j, t)$, which satisfies $\sum_{j \in \{1, \dots, \tilde{n}\}} m_i(j, t) = 1$ for every t . Now, let us collect all distribution values

$m_i(j, t)$, $j \in \{1, \dots, \tilde{n}\}$ in the macro-state vector of population i , namely:

$$x_i(t) := (m_i(1, t), m_i(2, t), \dots, m_i(\tilde{n}, t)) \in [0, 1]^{\tilde{n}}.$$

Thus, the averaging term in (1) describes the influence from neighbour populations.

As for the input term, consider, from a microscopic perspective, the case that the political opinions in a single population are distributed between two states, vote *left* and vote *right*, and such a distribution is subject to transitions from one state to the other. This is represented by the network depicted in Fig. 2 where nodes 1 and 2 correspond to the two states. Two persuaders, one of which is the controller (player A), the other the disturbance (player B), can influence the transitions described by the controlled flows \hat{v}_j , $j = 1, \dots, 4$ and disturbance parameters \hat{w}_k , $k = 3, 4$. In particular, player A can influence all the transitions, while player B has influence only on the transitions from node 2.

More generally, the terms \hat{v}_j and \hat{w}_k determine the transition rates between state 1 (vote *left*) and state 2 (vote *right*). In other words, a political campaign can make voters change their political opinion, and the controlled transition rates \hat{v}_j , $j = 2, 4$ represent the rates of change from one state to the other as a consequence of such a deliberate action. The parameters \hat{w}_k modulate these flows and are representative of unpredicted or uncontrolled events that can influence voters' opinions.

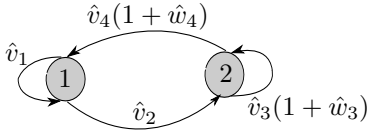


Fig. 2. Two opinion states (vote *left*, vote *right*) and corresponding transition functions.

In this case $\tilde{n} = 2$ and the evolution of the distribution is given by

$$x_i(t+1) = \left(I + B\tilde{B}^\top(\hat{v}_2(t), \hat{v}_4(t)) + D\tilde{D}^\top(\hat{v}_4(t), \hat{w}_4(t)) \right) x_i(t) \quad (5)$$

where

$$B = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}, \quad \tilde{B}(\hat{v}_2, \hat{v}_4) = \begin{bmatrix} \hat{v}_2 & 0 \\ 0 & \hat{v}_4 \end{bmatrix}, \quad (6)$$

$$D = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}, \quad \tilde{D}(\hat{v}_4, \hat{w}_4) = \begin{bmatrix} 0 & 0 \\ 0 & \hat{v}_4 \hat{w}_4 \end{bmatrix}.$$

System (5) corresponds to the set of difference equations

$$\begin{aligned} m_i(1, t+1) &= (1 - \hat{v}_2(t))m_i(1, t) \\ &\quad + \hat{v}_4(t)(1 + \hat{w}_4(t))m_i(2, t) \\ m_i(2, t+1) &= (1 - \hat{v}_4(t)(1 + \hat{w}_4(t)))m_i(2, t) \\ &\quad + \hat{v}_2(t)m_i(1, t) \end{aligned} \quad (7)$$

where $\hat{v}_2(t) \in [0, 1]$, $\hat{v}_4(t)(1 + \hat{w}_4(t)) \in [0, 1]$ and

$$\begin{aligned} \hat{v}_1(t) + \hat{v}_2(t) &= 1, \\ \hat{v}_3(t)(1 + \hat{w}_3(t)) + \hat{v}_4(t)(1 + \hat{w}_4(t)) &= 1. \end{aligned}$$

From (5) or (7), the input term is equal to the variation of distribution, and is given by

$$\begin{aligned} u_i(t) &:= x_i(t+1) - x_i(t) \\ &= \left(B\tilde{B}^\top(\hat{v}_2(t), \hat{v}_4(t)) + D\tilde{D}^\top(\hat{v}_4(t), \hat{w}_4(t)) \right) x_i(t) \\ &= \begin{bmatrix} -\hat{v}_2(t)m_i(1, t) + \hat{v}_4(t)(1 + \hat{w}_4(t))m_i(2, t) \\ -\hat{v}_4(t)(1 + \hat{w}_4(t))m_i(2, t) + \hat{v}_2(t)m_i(1, t) \end{bmatrix}. \end{aligned} \quad (8)$$

Let us assume that player A can control transitions by selecting (\hat{v}_2, \hat{v}_4) as one of the following transition configurations:

$$\begin{aligned} (v_2^{(1)}, v_4^{(1)}) &= (0, 0), & (v_2^{(2)}, v_4^{(2)}) &= (0, 0.2), \\ (v_2^{(3)}, v_4^{(3)}) &= (0.2, 0), & (v_2^{(4)}, v_4^{(4)}) &= (0.2, 0.2). \end{aligned}$$

For player B , we consider the following actions for \hat{w}_4 :

$$w_4^{(1)} = -0.5, \quad w_4^{(2)} = 0.5.$$

Consider a game where players select $p \in \Delta(S_A)$ and $q \in \Delta(S_B)$, where p_j and q_k are the probabilities assigned to the controlled flows $(v_2^{(j)}, v_4^{(j)})$ and to the modulating disturbance $w_4^{(k)}$, respectively. Thus $\Delta(S_A) = \{p \in \mathbb{R}^4 : p_j \geq 0, p_1 + \dots + p_4 = 1\}$ and $\Delta(S_B) = \{q \in \mathbb{R}^2 : q_1, q_2 \geq 0, q_1 + q_2 = 1\}$. We define the cumulative payoff at time t corresponding to the mixed strategies $(p(\tau), q(\tau)) \in \Delta(S_A) \times \Delta(S_B)$, $\tau \in [0, t]$, as the function

$$\begin{aligned} x_i(t+1) &= x_i(0) + \sum_{\tau=0}^t \left[B\tilde{B} \left(\sum_{j \in S_A} p_j(\tau) (v_2^{(j)}, v_4^{(j)}) \right) \right. \\ &\quad \left. + D\tilde{D} \left(\sum_{j \in S_A, k \in S_B} p_j(\tau) v_4^{(j)} q_k(\tau) w_4^{(k)} \right) \right] x_i(\tau) \\ &= x_i(0) + \sum_{\tau=0}^t \sum_{j \in S_A, k \in S_B} p_j(\tau) q_k(\tau) \phi(j, k). \end{aligned}$$

In the expression above

$$\begin{aligned}\phi(j, k) &= (B\tilde{B}^\top(v_2^{(j)}, v_4^{(j)}) + D\tilde{D}^\top(w_4^{(k)}))x_i(t) \\ &= \begin{bmatrix} -v_2^{(j)}m_i(1, t) + v_4^{(j)}(1 + w_4^{(k)})m_i(2, t) \\ -v_4^{(j)}(1 + w_4^{(k)})m_i(2, t) + v_2^{(j)}m_i(1, t) \end{bmatrix}.\end{aligned}$$

The complete matrix of vector payoffs is then obtained from Table 1, where each entry represents a possible vector payoff $\phi(j, k)$, and where we use $m(1) = m_i(1, t)$, $m(2) = m_i(2, t)$ to simplify notation.

3.1 Main assumptions

Following [20] (see also [19]) we make the following assumptions on the information structure of the model (1). Let $A(t)$ be the weight matrix with (i, j) th element $a_j^i(t)$.

Assumption 1 *The matrix $A(t)$ is doubly stochastic with positive diagonal. Furthermore, there exists a scalar $\alpha > 0$ such that $a_j^i(t) \geq \alpha$ whenever $a_j^i(t) > 0$ for all t .*

The instantaneous graph $\mathcal{G}(t)$ need not be connected at any given time t , however the union of the graphs $\mathcal{G}(t)$ over a period of time is assumed to be connected.

Assumption 2 *There exists an integer $Q \geq 1$ such that the graph $(N, \bigcup_{\tau=tQ}^{(t+1)Q-1} \mathcal{E}(\tau))$ is strongly connected for every non-negative integer t .*

For simplicity the one-shot vector-payoff game (S_A, S_B, x_i) is denoted by G . Let $\lambda \in \mathbb{R}^{\tilde{n}}$ and denote by $\langle \lambda, G \rangle$ the zero-sum one-shot game whose set of players and their action sets are as in the game G , and for which the payoff that player B pays to player A is $\lambda^\top \phi(j, k)$ for every $(j, k) \in S_A \times S_B$. We refer to $\langle \lambda, G \rangle$ as the *projected game*.

The projected game $\langle \lambda, G \rangle$ is described by the matrix

$$\Phi_\lambda = [\lambda^\top \phi(j, k)]_{j \in S_A, k \in S_B},$$

and as a zero-sum one-shot game it has a value v_λ , where

$$v_\lambda := \min_{p \in \Delta(S_A)} \max_{q \in \Delta(S_B)} p^\top \Phi_\lambda q = \max_{q \in \Delta(S_B)} \min_{p \in \Delta(S_A)} p^\top \Phi_\lambda q.$$

Following [15] (see also [14], Corollary 5.1), we make the following approachability assumption. In Section 4 we show that this ensures that any closed convex target set X can be made exponentially stable by a receding horizon strategy based on the repeated game (3).

Assumption 3 *For all $\lambda \in \mathbb{R}^{\tilde{n}}$ we have*

$$\min_{p \in \Delta(S_A)} \max_{q \in \Delta(S_B)} \left(2p^\top \Phi_\lambda q + \left\| \sum_{j \in S_A, k \in S_B} p_j \phi(j, k) q_k \right\|^2 \right) \leq 0.$$

The minimax inequality in Assumption 3 is a condition that the projected game $\langle \lambda, G \rangle$ must satisfy. This assumption also recalls Blackwell's Approachability Principle in [8].

We now strengthen Assumption 3 to include a condition which ensures that a receding horizon strategy based on (3) exponentially stabilizes the target set X with a region of attraction that is independent of the horizon length, T .

Assumption 4 *For given $R > 0$ there exists a scalar $\gamma \in (0, 1]$ such that*

$$\begin{aligned} \min_{p \in \Delta(S_A)} \max_{q \in \Delta(S_B)} \left(2p^\top \Phi_\lambda q + \left\| \sum_{j \in S_A, k \in S_B} p_j \phi(j, k) q_k \right\|^2 \right) \\ \leq -\gamma \|\lambda\|^2 \end{aligned}$$

for all $\lambda \in \mathbb{R}^{\tilde{n}}$ such that $\|\lambda\| \leq R$.

The condition of Assumption 4 is among the foundations of approachability theory since it requires that the value of the projected game satisfies $v_\lambda < 0$ whenever $\lambda \neq 0$. This is sufficient to guarantee that the average vector payoff of a two-player repeated game is locally almost surely convergent to the target set X (see e.g. [8], and also [11], chapter 7).

In order to apply the above assumptions to the multi-population example, one needs to take into account that the state trajectory lies in the positive quadrant (states represent distributions) and λ is in the 2nd and 4th quadrants ($m(1)$ increases if $m(2)$ decreases and vice versa).

The preceding assumptions allow an exponentially stable receding horizon control strategy to be constructed from (3) as follows. Let $\xi(t) = (x_1(t), \dots, x_n(t))$ denote the collective state of all agents in N at time t . We introduce a value function $V_{i,\tau}(\xi(t), t)$ representing the optimal cost over τ steps starting at $x_i(t)$, where $\tau = T - t$ for $t \in [0, T]$. Using dynamic programming and the Bellman principle, we require that the value function satisfies the following recursion

$$\begin{aligned} V_{i,\tau}(\xi(t), t) &= |x_i(t)|_X^2 \\ &+ \min_{p(t) \in \Delta(S_A)} \max_{\substack{q(t) \in \Delta(S_B) \\ u_j(t), j \neq i, j \in N}} V_{i,\tau-1}(\xi(t+1), t+1) \\ &\text{subject to } |x_i(t+1)|_X \leq |y_i(t)|_X, \end{aligned} \quad (9)$$

$10\phi(j, k)$	$k = 1$	$k = 2$
$j = 1$	$(0, 0)$	$(0, 0)$
$j = 2$	$(m(2), -m(2))$	$(3m(2), -3m(2))$
$j = 3$	$(-2m(1), 2m(1))$	$(-2m(1), 2m(1))$
$j = 4$	$(-2m(1) + m(2), 2m(1) - m(2))$	$(-2m(1) + 3m(2), 2m(1) - 3m(2))$

Table 1

The possible vector payoffs $\phi(j, k)$ scaled by a factor of 10.

with final value $V_{i,0}(\xi(T), T) = |x_i(T)|_X^2$. The minimizing solution for $p(t)$ in (9) with $\tau = T$ is then the optimal solution for $p^i(t)$ in the T -stage min-max problem (3) under the worst case allocations to all other agents $j \neq i$ and subject to the additional constraint that $|x_i(t+1)|_X \leq |y_i(t)|_X$. This additional constraint is included in the optimal control formulation to ensure stability of the resulting receding horizon control law. Since it involves only $u_i(t)$ and $y_i(t)$, this constraint does not require knowledge of the allocations of the other agents, which may not be known to any degree of accuracy. In Section 4 we show that it nonetheless results in a global contractivity property.

4 Main result

The main results of this paper establish contractivity and invariance for the collective dynamics (1) under the multi-stage receding horizon strategy defined in (9). Before stating the result we introduce four lemmas. The first of these establishes that the space averaging process in (1) reduces the total squared distance (i.e. the sum of squared distances) of the states x_i , $i \in N$, from the set X . The reader is referred to the Appendix for proof of each of the results given in this section.

Lemma 1 *Let Assumption 1 hold. Then the total squared distance from X decreases when states $x_i(t)$ are replaced by their space averages $y_i(t)$, i.e.,*

$$\sum_{i=1}^n |y_i(t)|_X^2 \leq \sum_{i=1}^n |x_i(t)|_X^2.$$

As a preliminary step to the next result, observe that, from the definition of $|\cdot|_X$ and from (1) and (3), we can write

$$\begin{aligned} |x_i(t+1)|_X^2 &= \|x_i(t+1) - P_X[x_i(t+1)]\|^2 \\ &\leq \|x_i(t+1) - P_X[y_i(t)]\|^2 \\ &= \|y_i(t) - P_X[y_i(t)]\|^2 + \|u_i(t)\|^2 \\ &\quad + 2(y_i(t) - P_X[y_i(t)])^\top u_i(t). \end{aligned} \quad (10)$$

The following lemma states that, under the approachability assumption, there necessarily exists an input $u_i(t)$ given by (2) that places the successor state $x_i(t+1)$ closer to X than the space average $y_i(t)$.

Lemma 2 *If Assumptions 1-3 hold, then, for all $\xi(t) = (x_1(t), \dots, x_n(t)) \in \mathbb{R}^{\tilde{n}} \times \dots \times \mathbb{R}^{\tilde{n}}$, there exists $u_i(t)$ satisfying (2) and*

$$|x_i(t+1)|_X^2 \leq |y_i(t)|_X^2 \quad (11)$$

for each $i \in N$. For $r > 0$ let $\Psi(r)$ denote the set

$$\Psi(r) = \left\{ (x_1, \dots, x_n) \in \mathbb{R}^{\tilde{n}} \times \dots \times \mathbb{R}^{\tilde{n}} \mid \sum_{i=1}^n |x_i|_X^2 \leq r^2 \right\}.$$

If Assumptions 1-4 hold, then, for all $\xi(t) \in \Psi(R)$ and each $i \in N$, there exists $u_i(t)$ satisfying (2) and

$$|x_i(t+1)|_X^2 \leq (1 - \gamma)|y_i(t)|_X^2. \quad (12)$$

for some $\gamma \in (0, 1]$.

As a consequence of Lemma 2, the constraint incorporated in the receding horizon strategy (9) is feasible for all collective states $\xi = (x_1, \dots, x_n) \in \mathbb{R}^{\tilde{n}} \times \dots \times \mathbb{R}^{\tilde{n}}$. The following lemma uses this property to show that $\Psi(r)$ is invariant for all $r > 0$.

Lemma 3 *If Assumptions 1-3 hold, then, for any $r > 0$, $\Psi(r)$ is invariant for (1) under the receding horizon strategy defined by (9) for all $i \in N$.*

We next give upper and lower bounds on the collective value function $\sum_{i=1}^n V_{i,T}(\xi, t)$ in terms of the sum of squared distances of individual agents' states from X .

Lemma 4 *Under Assumptions 1-3, the value functions $V_{i,T}(\xi, \cdot)$, $i \in N$ satisfy, for all $\xi \in \mathbb{R}^{\tilde{n}} \times \dots \times \mathbb{R}^{\tilde{n}}$,*

$$\sum_{i=1}^n |x_i|_X^2 \leq \sum_{i=1}^n V_{i,T}(\xi, \cdot) \leq (T+1) \sum_{i=1}^n |x_i|_X^2. \quad (13)$$

If Assumptions 1-4 hold, then the following bounds apply for all $\xi \in \Psi(R)$,

$$\sum_{i=1}^n |x_i|_X^2 \leq \sum_{i=1}^n V_{i,T}(\xi, \cdot) \leq \frac{1 - (1 - \gamma)^{T+1}}{\gamma} \sum_{i=1}^n |x_i|_X^2. \quad (14)$$

Let $\Psi(r_T)$ define a set of initial conditions $\xi(0)$ such that the state $x_i(T)$ of (1) is steered into X for all $i \in N$

by the optimal strategy for (9) with fixed terminal time $t = T$. Accordingly we define r_T by

$$r_T = \max \left\{ r \mid \sum_{i=1}^n |\hat{x}_i(T)|_X^2 = 0 \text{ for all } \xi(0) \in \Psi(r) \right\}$$

where $\hat{x}_i(t)$ for $t = 0, \dots, T$ denotes the evolution of (1) under the min-max strategy with optimal value function $V_{i,T-t}((\hat{x}_1(t), \dots, \hat{x}_n(t)), t)$ for all $i \in N$, with $\hat{x}_i(0) = x_i(0)$. Since Lemma 3 implies that X is invariant under any control law that satisfies the constraints of (9), it follows that r_T is monotonically non-decreasing in T , and hence $\Psi(r_T) \subseteq \Psi(r_{T+1})$ for each $T = 0, 1, \dots$

We are now ready to state the main results concerning the stabilizing properties of the receding horizon control law defined by (9). The proofs of the following two theorems are given in the Appendix.

The first result establishes exponential stability of X , with region of attraction dependent on T .

Theorem 1 *Let Assumptions 1-3 hold. For the system (1) with the receding horizon strategy with optimal cost $V_{i,T}(\xi(t), t)$ for all $i \in N$, the set X is exponentially stable with a region of attraction that contains $\Psi(r_T)$, i.e. for all $\xi(0) \in \Psi(r_T)$ and each $t = 0, 1, \dots$, we have*

$$\sum_{i=1}^n |x_i(t)|_X^2 \leq \left(\frac{T}{T+1} \right)^t \sum_{i=1}^n V_{i,T}(\xi(0), 0). \quad (15)$$

The next result establishes exponential stability of X , with region of attraction independent of T .

Theorem 2 *Let Assumptions 1-4 hold. Then for (1) under the receding horizon strategy with optimal cost $V_{i,T}(\xi(t), t)$ for all $i \in N$, the set X is exponentially stable and has a region of attraction that contains $\Psi(R)$, i.e. for all $\xi(0) \in \Psi(R)$ and each $t = 0, 1, \dots$, we have*

$$\sum_{i=1}^n |x_i(t)|_X^2 \leq \left(\frac{1-\gamma^2 - (1-\gamma)^{T+1}}{1 - (1-\gamma)^{T+1}} \right)^t \sum_{i=1}^n V_{i,T}(\xi(0), 0). \quad (16)$$

Exponential stability of the target set X implies that the state of (1) converges to X faster than an exponentially decaying function of time. Note that $(\frac{T}{T+1})$ is less than one and hence the term $(\frac{T}{T+1})^t$ in (15) decreases exponentially with time t . A similar comment applies to the term $(\frac{1-\gamma^2 - (1-\gamma)^{T+1}}{1 - (1-\gamma)^{T+1}})^t$ in (16) since $\gamma \in (0, 1]$. From (15) and (16) it also follows that, under the conditions of Theorems 1 and 2, if $x_i(t_0) \in X$ for all $i \in N$, then $x_i(t)$ must remain in X for all $i \in N$, for all $t > t_0$. Thus whenever the state of (1) lies in the target set X

for all $i \in N$ we have robust consensus, namely all states of the model have converged to a unique value with a given tolerance that depends on the size of X .

5 Numerical analysis

We consider $n = 5$ populations, each consisting of 200 agents, and a horizon length of $T = 25$. We assign to the j th agent in the i th population a state $\zeta_{ij}(t) \in [0, 1]$ and a target state $\zeta_{ij}^{\text{ref}}(t) \in \{0, 1\}$, where $\zeta_{ij}^{\text{ref}}(t) = 0$ or $\zeta_{ij}^{\text{ref}}(t) = 1$ indicates that the agent's opinion is changing to state 1 (vote *left*) or to state 2 (vote *right*), respectively, (see Fig. 2). We consider smooth trajectories that capture the inertia with which the agents change opinions. Thus the evolution of the opinion of an agent (microscopic dynamics) is in accordance with the dynamics:

$$\zeta_{ij}(t+1) = \text{sat} \left(\zeta_{ij}(t) + \beta (\zeta_{ij}^{\text{ref}}(t) - \zeta_{ij}(t)) + W_{ij}(t) \right) \quad (17)$$

where $\beta \in (0, 1)$, $\{W_{ij}(0), \dots, W_{ij}(T-1)\}$ is a Gaussian white noise sequence, and sat is the saturation function: $\text{sat}(\zeta) = \min\{\max\{0, \zeta\}, 1\}$.

For each population $i \in N$, a probability distribution function $x_i(t) := (m_i(1, t), \dots, m_i(\tilde{n}, t)) \in [0, 1]^{\tilde{n}}$ tracks the distribution of agents within the population. We assume transitions of opinions within each population occur according to the scheme in Fig. 2. Hence $\tilde{n} = 2$ and the distribution function over the space of target states $\{0, 1\}$ is $x_i(t) := (m_i(1, t), m_i(2, t)) \in [0, 1]^2$, where $m_i(1, t)$ and $m_i(2, t)$ are the fractions of the population in state 1 (vote *left*) and 2 (vote *right*), respectively. The evolution of $x_i(t)$ is determined for each $i \in N$ by

$$x_i(t+1) = y_i(t) + u_i(t),$$

where $u_i(t)$ is specified by (3) and $y_i(t)$ accounts for the transfer of opinions between populations via a communication graph, as in (4).

The target state $\zeta_{ij}^{\text{ref}}(t)$ for each agent in the population is updated at each time t based on the value of $x_i(t)$. Specifically, for $j < 200m_i(1, t)$ the target state $\zeta_{ij}^{\text{ref}}(t)$ is set to 0, whereas the target state of every other agent in population i is set to 1. For example, if $m_i(1, t) = 0.7$, then we set $\zeta_{ij}^{\text{ref}}(t) = 0$ for 70% of agents (their opinions are changing to *vote left*) and set $\zeta_{ij}^{\text{ref}}(t) = 1$ for the remaining 30% of agents (their opinions are changing to *vote right*). The set of initial states $\{\zeta_{ij}(0), j = 1, \dots, 200\}$ is uniformly distributed with mean equal to $m_i(2, 0)$.

The vector payoffs $\phi(j, k)$ in (3) are defined by Table 1, and these satisfy the minimax condition in Assumption 3. To see this, note that the payoffs $\phi(j, k)$ consist of pairs of opposite sign, and hence G is a zero-sum game where the payoff entries for player A are inverse to those

for player B . Therefore we can assume without loss of generality that the payoff matrix Φ_λ for the projected game (λ, G) is defined in terms of $\lambda = (\lambda_1, -\lambda_1)$, for $\lambda_1 \in \mathbb{R}$, so that $p^\top \Phi_\lambda q = 2\lambda_1 p^\top M q$ with

$$M = 10^{-1} \begin{bmatrix} 0 & 0 \\ m(2) & 3m(2) \\ -2m(1) & -2m(1) \\ m(2) - 2m(1) & 3m(2) - 2m(1) \end{bmatrix}. \quad (18)$$

In order to satisfy Assumption 3 for all $\lambda_1 \in \mathbb{R}$, we require that, for all $q \in \Delta(S_B)$, there exists $p \in \Delta(S_A)$ such that $p^\top M q \in [-2\lambda_1, 0]$ if $\lambda_1 \geq 0$, and such that $p^\top M q \in [0, -2\lambda_1]$ if $\lambda_1 < 0$. This is clearly satisfied by the matrix in (18) since, for each column $k = 1, 2$, there exists a convex combination of elements of that column which can be made arbitrarily small and either positive or negative. More formally, for all $\lambda_1 \in \mathbb{R}$ we have

$$\min_{p \in \Delta(S_A)} \max_{q \in \Delta(S_B)} (4\lambda_1 p^\top M q + 2(p^\top M q)^2) \leq 0, \quad (19)$$

from which it follows that Assumption 3 holds.

Case I The first set of simulations investigates the speed of the microscopic dynamics and the role of the communication graph topology on the macroscopic dynamics for the uncontrolled case (i.e. $u_i(t) = 0$ for all t). Consider first the connected communication graph depicted in Fig. 3 (left), with weights $a_j^i = 0.8$ for $i = j$, and $a_j^i = 0.05$ for $i \neq j$, for all $i, j \in N$. Figure 4 (left) shows the effect of varying β on the microscopic evolution of each agent's state $\xi_{ij}(t)$, and compares (right) the macro-state components $m_1(1, t), \dots, m_5(1, t)$ (solid lines) with the population averages of $1 - \xi_{ij}(t)$ (dashed lines). Clearly a larger coefficient β implies faster microscopic dynamics and hence faster convergence of agents to their target states. The average of $m_1(1, t), \dots, m_5(1, t)$ is independent of t since A is doubly stochastic, and the connectedness of the network topology therefore implies consensus on average.

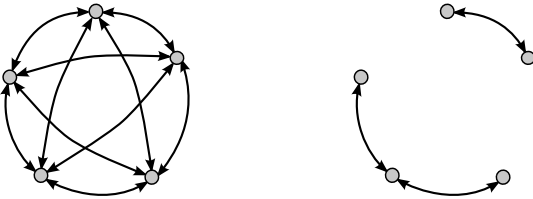


Fig. 3. Two communication graph topologies. Left: fully connected graph. Right: union of two disjoint subgraphs.

Figure 5 compares the responses of the two communication graphs shown in Fig. 3, with $u_i(t) = 0$ and $\beta = 0.5$ in each case. The lower plots indicate asymptotic polarization of opinion; this is due to the presence of two connected components in the graph on the right of Fig. 3.

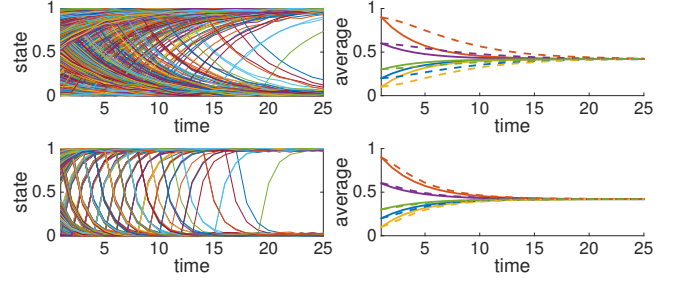


Fig. 4. Microscopic states $\xi_{ij}(t)$ (left) and macroscopic states $m_i(1, t)$ (right) for $\beta = 0.2$ (top) and $\beta = 0.6$ (bottom).

The oscillatory transient responses macroscopic states in the lower plot result from large off-diagonal weights (0.7) in the connected component consisting of two nodes.

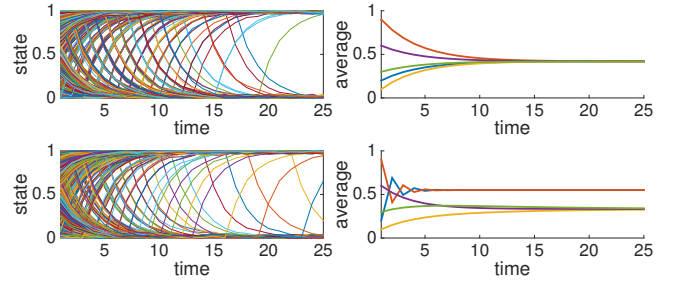


Fig. 5. Microscopic states $\xi_{ij}(t)$ (left) and macroscopic states $m_i(1, t)$ (right) for communication graphs with one (upper plots) and two (lower plots) connected components.

Case II The second set of simulations investigates the influence of the input $u_i(t)$ in (1). Suppose we wish the distribution of opinions in each population to converge to a given proportion, ρ , of agents in state 1, corresponding to a target set $X = \{(\rho, 1 - \rho)\}$. For this simple example the optimal strategy for the min-max problem (3) with the payoffs of Table 1 is given for any horizon $T \geq 1$ by

$$\begin{aligned} p^i(t) &= (0, 1, 0, 0) & \text{if } y_{i1}(t) - \rho < -\frac{1}{5}m_i(2, t) \\ p^i(t) &= (1 - \theta(t), \theta(t), 0, 0) & \text{if } y_{i1}(t) - \rho \in [-\frac{1}{5}m_i(2, t), 0) \\ p^i(t) &= (1 - \psi(t), 0, \psi(t), 0) & \text{if } y_{i1}(t) - \rho \in [0, \frac{1}{5}m_i(1, t)) \\ p^i(t) &= (0, 0, 1, 0) & \text{if } y_{i1}(t) - \rho \geq \frac{1}{5}m_i(1, t) \end{aligned}$$

where $y_{i1}(t)$ is the first element of $y_i(t)$ and

$$\theta(t) = \frac{\rho - y_{i1}(t)}{\frac{1}{5}m_i(2, t)}, \quad \psi(t) = \frac{y_{i1}(t) - \rho}{\frac{1}{5}m_i(1, t)}.$$

For the two communication graphs in Fig. 3 the evolution of this optimal control law with $\rho = 0.8$ is shown in the upper and middle plots of Figure 6. The disturbance is chosen to be worst-case (in the sense of maximizing (3)) at each simulation time-step. Although the macroscopic states in the middle plot are initially clustered due to the two connected components of the network topology, both clusters converge to the target point and thus consensus is achieved. The bottom plots in Fig. 6

are included here to illustrate a situation in which Assumption 3 does not hold. In this case one population is given a different target, $\rho' = 0.5$, and the effect of this population on the other populations, each of which has target $\rho = 0.8$, is equivalent to a disturbance (recall that the analysis in Section 4 relies on all populations having the same target set). However, unlike the disturbances considered in Section 3, this disturbance cannot be completely neutralised through the action of the controller. Consequently neither micro- nor macro-states of the system converge asymptotically to the target point in this case.

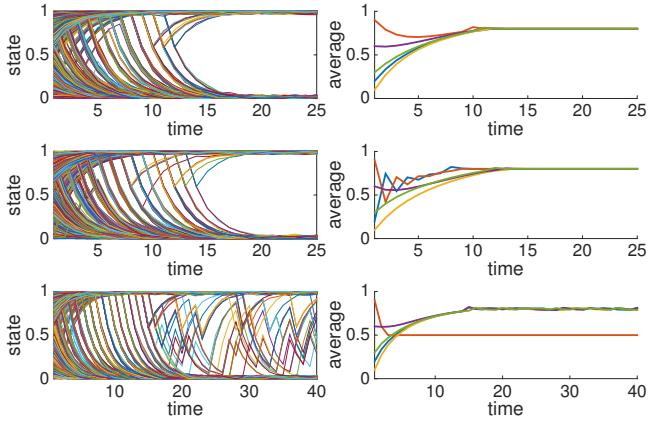


Fig. 6. Microscopic states $\xi_{ij}(t)$ (left) and macroscopic states $m_i(1, t)$ (right) for $\beta = 0.2$ (top) and 0.6 (bottom).

6 Conclusions

We have studied a multi-population opinion dynamics with macroscopic and microscopic intertwined dynamics. An averaging term is used to capture emulation, while an input term accounts for the changes in the population opinion distribution. Such a term is modeled as the vector payoff of a two player repeated game. We study conditions under which the agents achieve robust consensus to some predefined target set. Such conditions build upon the approachability principle in repeated games with vector payoffs.

References

- [1] D. Acemoğlu, G. Como, F. Fagnani, and A. Ozdaglar. Opinion fluctuations and disagreement in social networks. *Math. of Operation Research*, 38(1):1–27, 2013.
- [2] D. Acemoğlu and A. Ozdaglar. Opinion dynamics and learning in social networks. *International Review of Economics*, 1(1):3–49, 2011.
- [3] D. Aeyels and F. De Smet. A mathematical model for the dynamics of clustering. *Physica D: Nonlinear Phenomena*, 237(19):2517–2530, 2008.
- [4] AV. Banerjee. A simple model of herd behavior. *Quarterly Journal of Economics*, 107(3):797–817, 1992.
- [5] D. Bauso, M. Cannon, and J. Fleming. Robust consensus in social networks and coalitional games. In *Proceedings of 2014 IFAC World Congress*, pages 1537–1542, 2014.
- [6] D. Bauso, E. Lehrer, E. Solan, and X. Venel. Attainability in repeated games with vector payoffs. *INFORMS Mathematics of Operations Research*, 2014. accepted.
- [7] D. Bauso and G. Notarstefano. Distributed n -player approachability via time and space average consensus. In *3rd IFAC Workshop on Distributed Estimation and Control in Networked Systems*, pages 198–203, Santa Barbara, CA, USA, Sept 2012.
- [8] D. Blackwell. An analog of the minmax theorem for vector payoffs. *Pacific J. Math.*, 6:1–8, 1956.
- [9] V. D. Blondel, J. M. Hendrickx, and J. N. Tsitsiklis. Continuous-time average-preserving opinion dynamics with opinion-dependent communications. *SIAM J. Control and Optimization*, 48(8):5214–5240, 2010.
- [10] C. Castellano, S. Fortunato, and V. Loreto. Statistical physics of social dynamics. *Rev. Mod. Phys.*, 81:591–646, 2009.
- [11] N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.
- [12] G. Como and F. Fagnani. Scaling limits for continuous opinion dynamics systems. *The Annals of Applied Probability*, 21(4):1537–1567, 2011.
- [13] R. Hegselmann and U. Krause. Opinion dynamics and bounded confidence models, analysis, and simulations. *Journal of Artificial Societies and Social Simulation*, 5(3), 2002.
- [14] J. Hofbauer, M. Benaïm, and S. Sorin. Stochastic approximations and differential inclusions. *SIAM Journal on Control and Optimization*, 44(1):328–348, 2005.
- [15] J. Hofbauer, M. Benaïm, and S. Sorin. Stochastic approximations and differential inclusions, part ii: Applications. *INFORMS Mathematics of Operations Research*, 31(4):673–695, 2006.
- [16] U. Krause. A discrete nonlinear and non-autonomous model of consensus formation. In *Communications in Difference Equations*, S. Elaydi, G. Ladas, J. Popena, and J. Rakowski editors, Gordon and Breach, Amsterdam, pages 227–236, 2000.
- [17] E. Lehrer. Allocation processes in cooperative games. *International J. of Game Theory*, 31:341–351, 2002.
- [18] H. Liu, G. Xie, and L. Wang. Necessary and sufficient conditions for containment control of networked multi-agent systems. *Automatica*, 48(7):1415–1422, 2012.
- [19] A. Nedić and D. Bauso. Dynamic coalitional tu games: Distributed bargaining among players’ neighbors. *IEEE Trans. Autom. Control*, 58(6):1363–1376, 2013.
- [20] A. Nedić, A. Ozdaglar, and P.A. Parrilo. Constrained consensus and optimization in multi-agent

networks. *IEEE Trans. Autom. Control*, 55(4):922–938, 2010.

- [21] R. Olfati-Saber, J.A. Fax, and R.M. Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1):215–233, 2007.
- [22] A. Pluchino, V. Latora, and A. Rapisarda. Compromise and synchronization in opinion dynamics. *The European Physical Journal B - Condensed Matter and Complex Systems*, 50(1-2):169–176, 2006.
- [23] S. Sundhar Ram, A. Nedić, and V.V. Veeravalli. Incremental stochastic subgradient algorithms for convex optimization. *SIAM Journal on Optimization*, 20(2):691–717, 2009.
- [24] G. Shi and Y. Hong. Global target aggregation and state agreement of nonlinear multi-agent systems with switching topologies. *Automatica*, 45(5):1165–1175, 2009.
- [25] A.-S. Sznitman. Topics in propagation of chaos. *Springer Lecture Notes in Mathematics*, 1464:165–251, 1991.
- [26] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton Univ. Press, 1944.

Appendix

Proof of Lemma 1. By convexity of the distance function $|\cdot|_X$ we have $|y_i(t)|_X \leq \sum_{j=1}^n a_j^i(t) |x_j(t)|_X$. Hence convexity of $(\cdot)^2$ implies

$$|y_i(t)|_X^2 \leq \sum_{j=1}^n a_j^i(t) |x_j(t)|_X^2.$$

Summing both sides over $i = 1, \dots, n$ we obtain

$$\begin{aligned} \sum_{i=1}^n |y_i(t)|_X^2 &\leq \sum_{i=1}^n \sum_{j=1}^n a_j^i(t) |x_j(t)|_X^2 \\ &= \sum_{j=1}^n \left(\sum_{i=1}^n a_j^i(t) \right) |x_j(t)|_X^2 = \sum_{j=1}^n |x_j(t)|_X^2. \end{aligned}$$

Here the first inequality and last equality follow from Assumption 1 that $A(t)$ is doubly-stochastic. \square

Proof of Lemma 2. Rearranging the inequality in (10) we obtain

$$|x_i(t+1)|_X^2 - |y_i(t)|_X^2 \leq \|u_i(t)\|^2 + 2(y_i(t) - P_X[y_i(t)])^\top u_i(t). \quad (20)$$

With $\lambda = y_i(t) - P_X[y_i(t)]$, Assumption 3 implies that there exists a mixed strategy $p(t) \in \Delta(S_A)$ for player A such that, for any mixed strategy $q(t) \in \Delta(S_B)$ of player

B, $u_i(t) = \sum_{j \in S_A} \sum_{k \in S_B} p_j(t) \phi(j, k) q_k(t)$ satisfies

$$\|u_i(t)\|^2 + 2(y_i(t) - P_X[y_i(t)])^\top u_i(t) \leq 0$$

for all $y_i(t) \in \mathbb{R}^{\tilde{n}}$. Therefore the bound (11) follows from (20).

If $\xi(t) \in \Psi(R)$, then $\lambda = y_i(t) - P_X[y_i(t)]$ satisfies $\|\lambda\| \leq R$ since Lemma 1 implies that $|y_i(t)|_X^2 \leq R^2$ for all $i \in N$. Therefore the bound (12) follows from (20) and from Assumption 4. \square

Proof of Lemma 3. From the constraint in (9) (which, by Lemma 2, is necessarily feasible) and Lemma 1, we obtain $\sum_{i=1}^n |x_i(t+1)|_X^2 \leq \sum_{i=1}^n |y_i(t)|_X^2 \leq \sum_{i=1}^n |x_i(t)|_X^2$. Hence $\xi(t+1) \in \Psi(r)$ if $\xi(t) \in \Psi(r)$. \square

Proof of Lemma 4. The lower bounds in (13) and (14) follow directly from (9) and the fact that $V_{i,T-1}(\xi, \cdot) \geq 0$ for any horizon $T \geq 1$ and all $\xi \in \mathbb{R}^{\tilde{n}} \times \dots \times \mathbb{R}^{\tilde{n}}$. To prove the upper bound in (14), first consider first the case of $T = 1$. Since $\xi(t) \in \Psi(R)$ by assumption, condition (12) and the definition of $V_{i,0}(\xi, \cdot) = |x_i(\cdot)|_X^2$ yield

$$\begin{aligned} V_{i,1}(\xi(t), t) &= |x_i(t)|_X^2 + \min_{p \in \Delta(S_A)} \max_{q \in \Delta(S_B)} |x_i(t+1)|_X^2 \\ &\leq |x_i(t)|_X^2 + (1 - \gamma) |y_i(t)|_X^2. \end{aligned}$$

for all $i \in N$. Summing over $i \in N$ and using Lemma 1, we obtain

$$\sum_{i=1}^n V_{i,1}(\xi(t), t) \leq (2 - \gamma) \sum_{i=1}^n |x_i(t)|_X^2. \quad (21)$$

Consider next the case of $T > 1$, and assume that the upper bound in (14) holds for a horizon of $T - 1$. Bounding the RHS of (9) using (12) and Lemma 1 then yields

$$\begin{aligned} \sum_{i=1}^n V_{i,T}(\xi(t), t) &\leq \sum_{i=1}^n |x_i(t)|_X^2 \\ &\quad + \sum_{i=1}^n \min_{p(t) \in \Delta(S_A)} \max_{q(t) \in \Delta(S_B)} \frac{1 - (1 - \gamma)^T}{\gamma} |x_i(t+1)|_X^2 \\ &\leq \left(1 + (1 - \gamma) \frac{1 - (1 - \gamma)^T}{\gamma} \right) \sum_{i=1}^n |x_i(t)|_X^2 \\ &= \frac{1 - (1 - \gamma)^{T+1}}{\gamma} \sum_{i=1}^n |x_i(t)|_X^2. \end{aligned} \quad (22)$$

The upper bound in (14) therefore follows for all $T = 1, 2, \dots$ by induction using (21) and (22) and the fact that $\Psi(R)$ is invariant. Finally we note that the upper bound in (13) coincides with the upper bound in (14) in

the limit as $\gamma \rightarrow 0$, and that this bound must hold for all ξ under Assumption 3. \square

Proof of Theorem 1. The bound in (15) follows from the definition of r_T and the positive invariance of $\Psi(r_T)$. Thus, if $\xi(0) \in \Psi(r_T)$, then for all $i \in N$, the terminal state of (1) satisfies $|\hat{x}_i(T)|_X = 0$ under the min-max strategy with optimal value function $V_{i,T-t}((\hat{x}_1(t), \dots, \hat{x}_n(t)), t)$ for $t = 0, \dots, T$ and $\hat{x}_i(0) = x_i(0)$. Therefore

$$V_{i,T}(\xi, \cdot) = V_{i,T-1}(\xi, \cdot) \quad \forall \xi \in \Psi(r_T).$$

Furthermore $\xi(0) \in \Psi(r_T)$ implies $\xi(t) \in \Psi(r_T)$ for all $t = 0, 1, \dots$, and hence

$$\begin{aligned} V_{i,T}(\xi(t), t) &= \min_{p(t) \in \Delta(S_A)} \max_{\substack{q(t) \in \Delta(S_B) \\ u_j(t), j \neq i, j \in N}} V_{i,T-1}(\xi(t+1), t+1) \\ &\quad + |x_i(t)|_X^2 \\ &\geq V_{i,T}(\xi(t+1), t+1) + |x_i(0)|_X^2 \end{aligned}$$

for all $i \in N$. Summing this inequality over $i \in N$ and using the upper bound of (13) gives

$$\begin{aligned} \sum_{i=1}^n [V_{i,T}(\xi(t+1), t+1) - V_{i,T}(\xi(t), t)] &\leq - \sum_{i=1}^n |x_i(t)|_X^2 \\ &\leq - \frac{1}{T+1} \sum_{i=1}^n V_{i,T}(\xi(t), t). \end{aligned}$$

Hence $\sum_{i=1}^n V_{i,T}(\xi(t), t) \leq \left(\frac{T}{T+1}\right)^t \sum_{i=1}^n V_{i,T}(\xi(0), 0)$, and the lower bound of (13) yields (15). \square

Proof of Theorem 2. By optimality of (9) and the bounds (14) we have, for all $\xi(1) \in \mathbb{R}^{\bar{n}} \times \dots \times \mathbb{R}^{\bar{n}}$,

$$\sum_{i=1}^n V_{i,T}(\xi(1), 1) \leq \sum_{i=1}^n V_{i,T-1}(\xi(1), 1) + (1-\gamma) \sum_{i=1}^n |\hat{x}_i(T)|_X^2,$$

where $\hat{x}_i(t)$ for $t = 1, \dots, T$ is the state of (1) under the min-max optimal strategy for (9) with value function $V_{i,T-t}((\hat{x}_1(t), \dots, \hat{x}_n(t)), t)$ for all $i \in N$, and $\hat{x}_i(1) = x_i(1)$. For all $\xi(0) \in \Psi(R)$, it follows that

$$\begin{aligned} \sum_{i=1}^n V_{i,T}(\xi(0), 0) &\geq \sum_{i=1}^n |x_i(0)|_X^2 + \sum_{i=1}^n V_{i,T}(\xi(1), 1) \\ &\quad - (1-\gamma) \sum_{i=1}^n |\hat{x}_i(T)|_X^2 \\ &\geq \gamma \sum_{i=1}^n |x_i(0)|_X^2 + \sum_{i=1}^n V_{i,T}(\xi(1), 1), \end{aligned}$$

where the inequality $\sum_{i=1}^n |\hat{x}_i(t)|_X^2 \leq \sum_{i=1}^n |\hat{x}_i(t-1)|_X^2$ (which follows from (9) and Lemma 1) has been used for $T \geq t \geq 1$. For all $\xi(t) \in \Psi(R)$, the upper bound of (14) therefore yields

$$\begin{aligned} \sum_{i=1}^n [V_{i,T}(\xi(t+1), t+1) - V_{i,T}(\xi(t), t)] \\ \leq -\gamma \sum_{i=1}^n |x_i(t)|_X^2 \leq -\frac{\gamma^2}{1-(1-\gamma)^{T+1}} \sum_{i=1}^n V_{i,T}(\xi(t), t). \end{aligned}$$

Hence $\sum_{i=1}^n V_{i,T}(\xi(t), t)$ is bounded from above by $\left(\frac{1-\gamma^2-(1-\gamma)^T}{1-(1-\gamma)^{T+1}}\right)^t \sum_{i=1}^n V_{i,T}(\xi(0), 0)$, and the lower bound of (14) implies (15). \square