Designing Games for Distributed Optimization

Na Li and Jason R. Marden

***Abstract*—The central goal in multiagent systems is to design *local* control laws for the individual agents to ensure that the emergent global behavior is desirable with respect to a given system level objective. Ideally, a system designer seeks to satisfy this goal while conditioning each agent's control law on the least amount of information possible. This paper focuses on achieving this goal using the field of game theory. In particular, we derive a systematic methodology for designing local agent objective functions that guarantees (i) an equivalence between the resulting Nash equilibria and the optimizers of the system level objective and (ii) that the resulting game possesses an inherent structure that can be exploited in distributed learning, e.g., potential games. The control design can then be completed utilizing any distributed learning algorithm which guarantees convergence to a Nash equilibrium for the attained game structure. Furthermore, in many settings the resulting controllers will be inherently robust to a host of uncertainties including asynchronous clock rates, delays in information, and component failures.**

I. INTRODUCTION

The central goal in multiagent systems is to design *local* control laws for the individual agents to ensure that the emergent global behavior is desirable with respect to a given system level objective, e.g., [2]–[7]. These control laws provide the groundwork for a decision making architecture that possesses several desirable attributes including real-time adaptation and robustness to dynamic uncertainties. However, realizing these benefits requires addressing the underlying complexity associated with a potentially large number of interacting agents and the analytical difficulties of dealing with overlapping and partial information. Furthermore, the design of such control laws is further complicated by restrictions placed on the set of admissible controllers which limit informational and computational capabilities.

Game theory is beginning to emerge as a powerful tool for the design and control of multiagent systems [6]–[10]. Utilizing game theory for this purpose requires two steps. The first step is to model the agent as self-interested decision makers in a game theoretic environment. This step involves defining a set of choices and a local objective function for each decision maker. The second step involves specifying a distributed learning algorithm that enables the agents to reach a desirable operating point, e.g., a Nash equilibrium of the designed game.

One of the core advantages of game theory is that it provides a hierarchical decomposition between the distribution of the optimization problem (*game design*) and the specific local decision rules (*distributed learning algorithms*) [11]. For example, if the game is designed as a potential game [12] then there is an inherent robustness to decision making rules as a wide class of distributed learning algorithms can achieve convergence to a pure Nash equilibrium under a variety of informational dependencies [13]–[16]. Several recent papers focus on utilizing this decomposition in distributed control by developing methodologies for designing games, in particular agent utility functions, that adhere to this potential game structure [6], [9], [11], [17]. However, these methodologies typically provide no guarantees on the locality of the agent utility functions or the efficiency of the resulting pure Nash equilibria. Furthermore, the theoretical limits of what such approaches can achieve are poorly understood.

The goal of this paper is to establish a methodology for the design of *local* agent objective functions that leads to desirable system-wide behavior. We define the locality of an objective function by the underlying interdependence, i.e., the set of agents that impact this objective function. For convention, we refer to this set of agents as the neighbor set. Accordingly, an objective function (A) is more local than an objective function (B) if the neighbor set of (A) is strictly smaller than the neighbor set of (B). The existing utility design methodologies, i.e., the wonderful life utility [6], [9] and the Shapley value utility [18], [19], prescribe procedures for deriving agent objective functions from a given system level objective function. While both procedures guarantee that the resulting game is a potential game, the degree of locality in the agent objective functions is an artifact of the methodology and the underlying structure of the system level objective. Hence, these methodologies do not necessarily yield agent objective functions with the desired locality.

The main contribution of this paper is the development of a systematic methodology for the design of local agent objective functions that guarantees the efficiency of the resulting equilibria. In particular, in Theorem 3 we prove that our proposed methodology ensures (i) that there is an equivalence between the equilibria of the resulting game and the optimizers of the system level objective and (ii) that the resulting game is a *state based potential game* as introduced in [20].[1] A state based potential game is an extension of a potential game where there is an underlying state space introduced into the

This research was supported by AFOSR grants #FA9550-09-1-0538 and #FA9550-12-1-0359 and ONR grant #N00014-12-1-0643. The conference version of this work appeared in [1].

Na Li is a graduate student of Control and Dynamical Systems, California Institute of Technology, Pasadena, CA 91125, USA nali@caltech.edu

Jason R. Marden is with the Department of Electrical, Computer & Energy Engineering, University of Colorado at Boulder, Boulder, CO 80309, USA jason.marden@colorado.edu

[1]It is important to highlight that [20] focuses predominantly on learning in state based potential games with finite action sets. The design of agent utility functions to ensure the efficiency of the resulting equilibria, which is the focus of this manuscript, is not addressed in [20].

game structure. Our design utilizes these state variables as a coordinating entity to decouple the system level objective into agent specific objectives of the desired interdependence.

Our second result focuses on learning in state based potential games with continuous action sets. Much like potential games, state based potential games possess an underlying structure that can be exploited in distributed learning. Accordingly, in this paper we prove that the learning algorithm gradient play, introduced in [21], [22] in the context of strategic form games, converges to an equilibrium in any state based potential game (see Theorem 4). Moreover, we provide a characterization of the convergence rate of gradient play for state based potential games (see Theorem 5). This work is complimentary to the results in [20] which provide similar results for state based potential games with finite action sets.

The design of multiagent systems parallels the theme of distributed optimization which can be thought of as a concatenation between a designed game and a distributed learning algorithm. One core difference between these two domains is the fact that multiagent systems frequently place restrictions on the set of admissible controllers. In terms of distributed optimization, this places a restriction on the set of admissible distributed algorithms. Accordingly, the applicability of some common approaches to distributed optimization, e.g, subgradient methods [23]–[28], consensus based methods [2], [3], [29], [30], or two-step consensus based approaches [10], [31], [32], may be limited by the structure of the system level objective.

There are also a family of distributed algorithms that are similar in spirit to the algorithms presented in this paper. In particular, the algorithms presented in [27] and [32] introduce a communication protocol between the agents with the purpose of providing the agents with sufficient degrees of information so that the agents can estimate their gradient to the system level objective. While the proposed algorithms provide the desired asymptotic guarantees, the robustness to variations in clock rates, delays in information, and component failures is currently uncharacterized. Furthermore, the complexity regarding the analysis of these algorithms could make providing such a characterization challenging. In contrast to [27] and [32], our focus is on a methodological decomposition of the system level objective into local agent objective functions. Through this decomposition, we can take advantage of existing results in the field of learning in games to derive distributed algorithms that are robust to delays in information and heterogeneous clock rates. This follows directly from [14] and [20] which prove that any *reasonable* distributed learning algorithm will converge to a pure Nash equilibrium in any (finite) potential game or (finite) state based potential game.

This paper focuses on establishes a systematic approach for distributed optimization. Accordingly, we focus predominantly on a general class of optimization problem with the realization that many problem instantiations relevant to multiagent systems can be represented within this problem formulation. Examples include collaborative sensing in a distributed PTZ camera network and the design of local control strategies for mobile sensor networks [33], [34]. For concreteness, in Section V-B we formally describe a distributed routing problem and illustrate how the proposed methodology can lead to desirable system behavior even when the agents possess incomplete information regarding the network behavior.

II. Problem Setup and Preliminaries

We consider a multiagent system consisting of n agents denoted by the set $N = \{1, \cdots, n\}$. Each agent $i \in N$ is endowed with a set of decisions (or values) denoted by $\mathcal{V}_i$ which is a nonempty convex subset of $\mathbb{R}$. We denote a joint value by the tuple $v = (v_1, \cdots, v_n) \in \mathcal{V} = \prod_i \mathcal{V}_i$ where $\mathcal{V}$ is referred to as the set of joint values. Last, there is a global cost function of the form $\phi : \mathbb{R}^N \to \mathbb{R}$ that a system designer seeks to minimize. More formally, the optimization problem takes on the form:[2]

$$\begin{array}{ll} \min_v & \phi(v_1, v_2, \ldots, v_n) \\ \text{s.t.} & v_i \in \mathcal{V}_i, \forall i \in N. \end{array} \tag{1}$$

We assume throughout that ϕ is differentiable convex and a solution of this optimization problem is guaranteed to exist.[3]

The focus of this paper is to establish an interaction framework where each agent $i \in N$ chooses its value independently in response to local information. The information available to each agent is represented by an undirected and connected communication graph $\mathcal{G} = \{N, \mathcal{E}\}$ with nodes N and edges $\mathcal{E}$. Define the neighbors of agent i as $N_i = \{j \in N : (i,j) \in \mathcal{E}\}$ and we adopt the convention that $i \in N_i$ for each i. This interaction framework produces a sequence of values $v(0), v(1), v(2), \ldots$, where at each iteration $t \in \{0, 1, \ldots\}$ each agent i chooses a value independently according to a local control law of the form:

$$v_i(t) = F_i\left(\{\text{Information about agent } j\}_{j \in N_i}\right) \tag{2}$$

which designates how each agent processes available information to formulate a decision at each iteration. The goal in this setting is to design the local controllers $\{F_i(\cdot)\}_{i \in N}$ such that the collective behavior converges to a joint value v^* that solves the optimization problem in (1).

A. An illustrative example

We begin by presenting a simple example to motivate the theoretical developments in this paper. Consider the following instance of (1) where

$$\phi(v_1, v_2, v_3) = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}^T \begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 4 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} + [1\ 1\ 1] \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \tag{3}$$

and $\mathcal{V}_i = \mathbb{R}$ for all agents $N = \{1, 2, 3\}$. Here, the goal is to derive local agent control laws of the form (2) that converge to the minimizer of the cost function in (3) while adhering to the communication graph $1 \leftrightarrow 2 \leftrightarrow 3$. Note that this communication graph implies that the control policy of agent 1 is not able to depend on the true value of agent 3.

[2]For ease of exposition we let $\mathcal{V}_i \subseteq \mathbb{R}$, which is just one dimension. In general, $\mathcal{V}_i$ can be any convex subset of $\mathbb{R}^{d_i}$ for any dimension $d_i \geq 1$. The results in this paper immediately extends to the cases where $d_i > 1$ and $d_i \neq d_j$ for $i \neq j$. Furthermore, this work focuses on problems with decoupled constraints on agents' actions, i.e., $v_i \in \mathcal{V}_i$. The forthcoming methodologies can also incorporate coupled constraints using the approach demonstrated in [35].

[3]There are many sufficient conditions to guarantee the existence of the optimal solution, e.g., compactness of $\mathcal{V}$ or coercivity of ϕ.

1) Gradient methods: Gradient methods represent a popular algorithm for solving nonlinear optimization problem [36]. A gradient method for the optimization problem in (3) takes on the form

$$v_i(t+1) = v_i(t) - \epsilon \frac{\partial \phi}{\partial v_i}, \tag{4}$$

where

$$\begin{aligned} \frac{\partial \phi}{\partial v_1} &= 4v_1 + 2v_2 + 2v_3 + 1 \\ \frac{\partial \phi}{\partial v_2} &= 2v_1 + 6v_2 + 2v_3 + 1 \\ \frac{\partial \phi}{\partial v_3} &= 2v_1 + 2v_2 + 8v_3 + 1. \end{aligned}$$

and ϵ is a positive step size. Note that both agent 1 and 3 require global information to calculate their gradients which is not admissible in our setting.

2) A game theoretic approach: Since ϕ in (3) does not possess a locally decomposable structure, the resulting gradient descent algorithms were not of the desired locality. A game theoretic approach introduces an intermediate step to the control design where each agent is assigned an objective function of the form $J_i : \prod_{j \in N_i} \mathcal{V}_j \to \mathbb{R}$. Here the goal is to embed the information admissibility constraints directly into the agents' objective function. For example, if we design agent objective functions of the form:

$$\begin{aligned} &J_1 : \mathcal{V}_1 \times \mathcal{V}_2 \to \mathbb{R} \\ &J_2 : \mathcal{V}_1 \times \mathcal{V}_2 \times \mathcal{V}_3 \to \mathbb{R} \\ &J_3 : \mathcal{V}_2 \times \mathcal{V}_3 \to \mathbb{R} \end{aligned}$$

and each agent follows a gradient-based approach to their local objectives, i.e., for any agent $i \in N$,

$$v_i(t+1) = v_i(t) - \epsilon \frac{\partial J_i}{\partial v_i},$$

then the resulting agents' control policies will satisfy our locality constraints. However, the convergence properties of such an algorithm are not as straightforward as the gradient algorithm given in (4), which leads to the work of this paper.

B. Preliminaries: Potential games

A strategic form game is characterized by a set of agents $N = \{1, \ldots, n\}$ where each agent $i \in N$ has an action set $\mathcal{A}_i$ and a cost function $J_i : \mathcal{A} \to \mathbb{R}$ where $\mathcal{A} = \prod_{i \in N} \mathcal{A}_i$ denotes the set of joint actions. For an action profile $a = (a_1, ..., a_n)$, let a_{-i} denote the action profile of agents other than agent i, i.e., $a_{-i} = (a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$.

One class of games that plays a prominent role in engineering multiagent systems is that of potential games [12].

Definition 1. *(Potential Games) A game $\{N, \{\mathcal{A}_i\}, \{J_i\}\}$ is called an (exact) potential game if there exists a global function $\Phi : \mathcal{A} \to \mathbb{R}$ such that for every agent $i \in N$, $a_{-i} \in \mathcal{A}_{-i}$ and $a_i', a_i'' \in \mathcal{A}_i$,*

$$J_i(a_i', a_{-i}) - J_i(a_i'', a_{-i}) = \Phi(a_i', a_{-i}) - \Phi(a_i'', a_{-i}).$$

There are three main properties regarding potential games which makes them an attractive paradigm for distributed engineering systems. First, in a potential game a pure Nash equilibrium, i.e., an action profile $a^* \in \mathcal{A}$ such that

$$J_i(a_i^*, a_{-i}^*) = \min_{a_i \in \mathcal{A}_i} J_i(a_i, a_{-i}^*), \forall i \in N,$$

is guaranteed to exist. Second, there are several available distributed learning algorithms with proven asymptotic guarantees that could be utilized for the control design [12]–[16]. Lastly, learning pure Nash equilibria in potential games is inherently robust [14]. That is, any "reasonable" learning algorithm where players seek to optimize their individual objective function will converge to a pure Nash equilibrium in potential games [14]. Hence, issues such as heterogeneous clock rates and informational delays are not problematic to learning pure Nash equilibria in such games.

C. Preliminaries: state based potential games

State based games, a simplification of stochastic games [37], represent an extension to strategic form games where an underlying state space is introduced to the game theoretic environment [20]. The class of state based games considered in this paper consists of the following elements:

(i) an agent set N,
(ii) a state space X,
(iii) a state dependent action set, $\mathcal{A}_i(x)$, for each agent $i \in N$ and state $x \in X$,
(iv) a state dependent cost function of the form $J_i(x, a) \in \mathbb{R}$, for each agent $i \in N$, $x \in X$, and $a \in \mathcal{A}(x) = \prod_{i \in N} \mathcal{A}_i(x)$, and
(v) a deterministic state transition function $f(x, a) \in X$ for $x \in X$ and $a \in \mathcal{A}(x)$.

Furthermore, we focus on state based games where for any $x \in X$ there exists a null action $\mathbf{0} \in \mathcal{A}(x)$ such that $x = f(x, \mathbf{0})$. This implies that the state will remain unchanged if all of the agents take the null action. We will frequently denote a state based game by $G = \{N, X, \mathcal{A}, J, f\}$, where $\mathcal{A} = \bigcup_{x \in X} \mathcal{A}(x)$.

Repeated play of a state based game produces a sequence of action profiles $a(0), a(1), \cdots$, and a sequence of states $x(0), x(1), \ldots$, where $a(t) \in \mathcal{A}$ is referred to as the action profile at time t and $x(t) \in X$ is referred to as the state at time t. At any time $t \geq 0$, each agent $i \in N$ selects an action $a_i(t) \in \mathcal{A}_i(x(t))$ according to some specified decision rule which depends on the current state $x(t)$. The state $x(t)$ and the joint action profile $a(t) = (a_1(t), \ldots, a_n(t)) \in \mathcal{A}(x(t))$ determine each agent's one stage cost $J_i(x(t), a(t))$ at time t. After all agents select their respective action, the ensuring state $x(t+1)$ is chosen according to the deterministic state transition function $x(t+1) = f(x(t), a(t))$ and the process is repeated.

In this paper we focus on the class of games termed state based potential games which represents an extension of potential games to the framework of state based games.

Definition 2. *(State Based Potential Game) A (deterministic) state based game G with a null action $\mathbf{0}$ is a (deterministic) state based potential game if there exists a potential function $\Phi : X \times \mathcal{A} \to \mathbb{R}$ satisfying the following two properties for every state $x \in X$:*

(D-1): For every agent $i \in N$, action profile $a \in \mathcal{A}(x)$ and action $a_i' \in \mathcal{A}_i(x)$

$$J_i(x, a_i', a_{-i}) - J_i(x, a) = \Phi(x, a_i', a_{-i}) - \Phi(x, a)$$

(D-2): For every action profile $a \in \mathcal{A}(x)$ and the ensuing state $\tilde{x} = f(x, a)$, the potential function satisfies

$$\Phi(x, a) = \Phi(\tilde{x}, \mathbf{0}).$$

The first condition states that each agent's cost function is aligned with the potential function in the same fashion as in potential games (Definition 1). The second condition relates to the evolution on the potential function along the state trajectory.[4] We focus on the class of state based potential games as dynamics can be derived that converge to the following class of equilibria (see Theorem 4).

Definition 3. *(Stationary State Nash Equilibrium) A state action pair $[x^*, a^*]$ is a stationary state Nash equilibrium if*

(D-1): For any agent $i \in N$,

$$a_i^* \in \arg \min_{a_i \in \mathcal{A}_i(x^*)} J_i(x^*, a_i, a_{-i}^*).$$

(D-2): The state x^* is a fixed point of the state transition function, i.e., $x^* = f(x^*, a^*)$.

Note that in the case of a single state, i.e., $X = 1$, the definition of Stationary State Nash Equilibrium is precisely that of a Nash equilibrium since Condition (D-2) is satisfied trivially. The following proposition proves the existence of a stationary state Nash equilibrium in any state based potential game.

Proposition 1. *Let G be a state based potential game with potential function Φ and a null action $\mathbf{0}$. If $x^* \in \text{argmin}_{x \in X} \Phi(x, \mathbf{0})$, then $[x^*, \mathbf{0}]$ is a stationary state Nash equilibrium. Moreover, for any $a \in \mathcal{A}(x^*)$ such that $x^* = f(x^*, a)$, $[x^*, a]$ is also a stationary state Nash equilibrium.*

Proof: In order to prove that $[x^*, \mathbf{0}]$ is a stationary state Nash equilibrium we only need to show that $\mathbf{0} \in \text{argmin}_{a \in \mathcal{A}(x^*)} \Phi(x^*, a)$ because $x = f(x, \mathbf{0})$ for any $x \in X$ and Φ is a potential function of the game G. Let $a^* \in \text{argmin}_{a \in \mathcal{A}(x^*)} \Phi(x^*, a)$. Thus $\Phi(x^*, \mathbf{0}) \geq \Phi(x^*, a^*)$. However since $x^* \in \text{argmin}_{x \in X} \Phi(x, \mathbf{0})$, we have that $\Phi(x^*, a^*) = \Phi(\tilde{x}^*, \mathbf{0}) \geq \Phi(x^*, \mathbf{0})$ where $\tilde{x}^* = f(x^*, a^*)$. Therefore we have $\Phi(x^*, \mathbf{0}) = \Phi(x^*, a^*) = \min_{a \in \mathcal{A}(x^*)} \Phi(x^*, a)$. Hence $[x^*, \mathbf{0}]$ is a stationary state Nash equilibrium. For any a such that $x^* = f(x^*, a)$, we have $\Phi(x^*, a) = \Phi(x^*, \mathbf{0}) = \min_{a \in \mathcal{A}(x^*)} \Phi(x^*, a)$ implying that $[x^*, a]$ is also a stationary state Nash equilibrium. ∎

III. State Based Game design

In this section we introduce a state based game design for the distributed optimization problem in (1). The goal of our design is to establish a state based game formulation that satisfies the following four properties:

(i) The state represents a compilation of local state variables, i.e., the state x can be represented as $x = (x_1, \ldots, x_n)$ where each x_i represents the state of agent i. Furthermore, the state transition f should also rely only on local information.

(ii) The objective function for each agent i is local and of the form $J_i(\{x_j, a_j\}_{j \in N_i}) \in \mathbb{R}$.

(iii) The resulting game is a state based potential game.

(iv) The stationary state Nash equilibria are optimal in the sense that they represent solutions to the optimization problem in (1), i.e., $v_i = v_i^*$.[5]

A. A state based game design for distributed optimization

We now introduce the specifics of our designed game.

State Space: The starting point of our design is an underlying state space X where each state $x \in X$ is defined as a tuple $x = (v, e)$, where

- $v = (v_1, \ldots, v_n) \in \mathbb{R}^n$ is the profile of values and
- $e = (e_1, \ldots, e_n)$ is the profile of estimation terms where $e_i = (e_i^1, \cdots, e_i^n) \in \mathbb{R}^n$ is agent i's estimation for the joint action profile $v = (v_1, \ldots, v_n)$. The term e_i^k captures agent i's estimate of agent k's actual value v_k.

The estimation terms are introduced as a means to relax the degree of information available to each agent. More specifically, each agent is aware of its own estimation as opposed to the true value profile which may in fact be different, i.e., e_i^k need not equal v_k.

Action Sets: Each agent i is assigned an action set $\mathcal{A}_i$ that permits agents to change their value and change their estimation through communication with neighboring agents. Specifically, an action for agent i is defined as a tuple $a_i = (\hat{v}_i, \hat{e}_i)$ where $\hat{v}_i \in \mathbb{R}$ indicates a change in the agent's value v_i and $\hat{e}_i = (\hat{e}_i^1, \cdots, \hat{e}_i^n)$ indicates a change in the agent's estimation terms e_i. We represent each of the estimation terms $\hat{e}_i^k$ by the tuple $\hat{e}_i^k = \{\hat{e}_{i \to j}^k\}_{j \in N_i \setminus \{i\}}$ where $\hat{e}_{i \to j}^k \in \mathbb{R}$ represents the estimation value that agent i passes to agent j regarding to the value of agent k.

State Dynamics: Define $\hat{e}_{i \leftarrow \text{in}}^k = \sum_{j \in N_i \setminus \{i\}} \hat{e}_{j \to i}^k$ and $\hat{e}_{i \to \text{out}}^k = \sum_{j \in N_i \setminus \{i\}} \hat{e}_{i \to j}^k$ denote the total estimation passed to and from agent i regarding the value of the k-th agent respectively. We represent the state transition function $f(x, a)$ by a set of local state transition functions $\{f_i^v(x, a)\}_{i \in N}$ and $\left\{f_{i,k}^e(x, a)\right\}_{i,k \in N}$. For a state $x = (v, e)$ and an action

[4] The definition of state based games differs slightly from [20] as we focus on state dependent actions sets and games where there exists null actions.

[5] There is a significant body of work in the field of algorithmic game theory that focuses on analyzing the inefficiency of Nash equilibria [38]. A common measure for this inefficiency, termed price of anarchy, is the worst case ratio between the system level performance of a Nash equilibrium and the optimal systems level performance. The vast literature in this area is predominantly analytical where the price of anarchy is characterized for situations where both the system level objective function and the agent cost functions are given. This work, on the other hand, focuses on the the counterpart of this analytical direction. In particular, is it possible to design local agent cost functions such that the price of anarchy is 1 for given a system level objective function? For the class of optimization problems considered in this manuscript, we provide a systematic methodology for accomplishing this task.

$a = (\hat{v}, \hat{e})$, the ensuing state $\tilde{x} = (\tilde{v}, \tilde{e})$ is given by

$$\begin{array}{rcl} \tilde{v}_i = f_i^v(x,a) & = & v_i + \hat{v}_i \\ \tilde{e}_i^k = f_{i,k}^e(x,a) & = & e_i^k + n\delta_i^k \hat{v}_i + \hat{e}_{i\leftarrow \text{in}}^k - \hat{e}_{i\rightarrow \text{out}}^k \end{array} \quad (5)$$

where δ_i^k is an indicator function, i.e., $\delta_i^i = 1$ and $\delta_i^k = 0$ for all $k \neq i$. Since the optimization problem in (1) imposes the requirement that $v_i \in \mathcal{V}_i$, we condition agents' available actions on the current state. That is, the available action set for agent i given state $x = (v, e)$ is defined as

$$\mathcal{A}_i(x) = \{(\hat{v}_i, \hat{e}_i) : v_i + \hat{v}_i \in \mathcal{V}_i\} \quad (6)$$

Invariance associated with state dynamics: Let $v(0) = (v_1(0), ..., v_n(0))$ be the initial values of the agents. Define the initial estimation terms $e(0)$ to satisfy $\sum_{i\in N} e_i^k(0) = n \cdot v_k(0)$, for each agent $k \in N$; hence, the initial estimation values are contingent on the initial values. Note that satisfying this condition is trivial as we can set $e_i^i(0) = n \cdot v_i(0)$ and $e_i^j(0) = 0$ for all agents $i, j \in N$ where $i \neq j$. Define the initial state as $x(0) = [v(0), e(0)]$. It is straightforward to show that for any action trajectory $a(0), a(1), \cdots$, the resulting state trajectory $x(t) = (v(t), e(t)) = f(x(t-1), a(t-1))$ satisfies the following equalities for all times $t \geq 1$ and agents $k \in N$

$$\sum_{i=1}^{n} e_i^k(t) = n \cdot v_k(t) \quad (7)$$

Agent Cost Functions: The cost functions possess two distinct components and takes on the form

$$J_i(x,a) = J_i^{\phi}(x,a) + \alpha \cdot J_i^e(x,a) \quad (8)$$

where $J_i^{\phi}(\cdot)$ represents the component centered on the objective function ϕ, $J_i^e(\cdot)$ represents the component centered on the disagreement of estimation based terms e, and α is a positive constant representing the tradeoff between the two components.[6] We define each of these components as

$$\begin{array}{rcl} J_i^{\phi}(x,a) & = & \sum_{j\in N_i} \phi(\tilde{e}_j^1, \tilde{e}_j^2, ..., \tilde{e}_j^n) \\ J_i^e(x,a) & = & \sum_{j\in N_i} \sum_{k\in N} \left[\tilde{e}_i^k - \tilde{e}_j^k\right]^2 \end{array} \quad (9)$$

where $\tilde{x} = (\tilde{v}, \tilde{e}) = f(x, a)$ represents the ensuing state. The null action $\mathbf{0}$ is characterized by

$$\hat{v}_i = 0, \quad \hat{e}_{i\rightarrow j}^k = 0, \forall i, j, k \in N.$$

Since $x = f(x, \mathbf{0})$, the agents' cost functions satisfy $J_i(x, a) = J_i(\tilde{x}, \mathbf{0})$.

B. Analytical properties of designed game

In this section we derive two analytical properties of the designed state based game. The first property establishes that the designed game is a state based potential game.

Theorem 2. *The state based game depicted in Section III-A is a state based potential game with potential function*

$$\Phi(x,a) = \Phi^{\phi}(x,a) + \alpha \cdot \Phi^e(x,a) \quad (10)$$

where

$$\begin{array}{rcl} \Phi^{\phi}(x,a) & = & \sum_{i\in N} \phi(\tilde{e}_i^1, \tilde{e}_i^2, ..., \tilde{e}_i^n) \\ \Phi^e(x,a) & = & \frac{1}{2} \sum_{i\in N} \sum_{j\in N_i} \sum_{k\in N} \left[\tilde{e}_i^k - \tilde{e}_j^k\right]^2 \end{array} \quad (11)$$

and $\tilde{x} = (\tilde{v}, \tilde{e}) = f(x, a)$ represents the ensuing state.

Proof: It is straightforward to verify that Conditions (D-1)-(D-2) of state based potential games in Definition 2 are satisfied using the state based potential function defined in (10). ■

The following theorem demonstrates that *all* equilibria of our designed game are solutions to the optimization problem in (1).

Theorem 3. *Let G be the state based game depicted in Section III-A. Suppose that ϕ is a differentiable convex function, the communication graph $\mathcal{G}$ is connected and undirected, and at least* ***one*** *of the following conditions is satisfied:*

(i) The communication graph $\mathcal{G}$ is non-bipartite.[7]

(ii) The communication graph $\mathcal{G}$ contains an odd number of nodes, i.e., the number of agents is odd;

(iii) The communication graph $\mathcal{G}$ contains at least two agents which have a different number of neighbors, i.e., $|N_i| \neq |N_j|$ for some agents $i, j \in N$;

(iv) For each agent $i \in N$ the actions set $\mathcal{V}_i$ is open.

Then the state action pair $[x, a] = [(v, e), (\hat{v}, \hat{e})]$ is a stationary state Nash equilibrium if and only if the following conditions are satisfied:

(a) The estimation profile e satisfies that $e_i^k = v_k$, $\forall i, k \in N$;

(b) The value profile v is an optimal solution for problem (1);

(c) The change in value profile satisfies $\hat{v} = \mathbf{0}$;

(d) The change in estimation profile satisfies that for all agents $i, k \in N$, $\hat{e}_{i\leftarrow \text{in}}^k = \hat{e}_{i\rightarrow \text{out}}^k$.

The above theorem proves that the resulting equilibria of our state based game coincide with the optimal solutions to the optimization problem in (1) under relatively minor conditions on the communication graph. Hence, our design provides a systematic methodology for distributing an optimization problem under virtually any desired degree of locality in the agents' objective functions. A natural question arises as to whether the results in Theorem 2 and 3 could have been attained using the framework of strategic form games. In Appendix-A we prove that it is impossible to accomplish such a task.

IV. Gradient Play

In this section we prove that the learning algorithm *gradient play*, studied previously in [21] and [22] for strategic form games, converges to a stationary state Nash equilibrium in state based potential games. Since the designed game depicted in Section III-A is a state based potential game, the algorithm gradient play can be utilized to design control laws of the form (2) that guarantee convergence to the optimal solution of (1).

[6] We will show that for any positive α, the results demonstrated in this paper holds. However, choosing the right α is important for the learning algorithm implementation, e.g., the convergence rate of the learning algorithm.

[7] A bipartite graph is a graph that does not contain any odd-length cycles.

A. Gradient play for state based potential games

Given a state based potential game $G = \{N, \mathcal{A}, X, J, f\}$, suppose that $\mathcal{A}_i(x)$ is a closed convex set for all $i \in N$ and $x \in X$. Let $x(t)$ represent the state at time t. Accordingly to the learning algorithm gradient play, each agent $i \in N$ selects an action $a_i(t) \in \mathcal{A}_i(x(t))$ according to

$$a_i(t) = \left[-\epsilon_i \cdot \frac{\partial J_i(x(t), a)}{\partial a_i} \bigg|_{a=\mathbf{0}} \right]^+ \tag{12}$$

where $[\cdot]^+$ represents the projection onto the closed convex set $\mathcal{A}_i(x(t))$ and ϵ_i is the step size which is a positive constant. Note that the agents' step sizes can be heterogeneous.

Before establishing the convergence results, we make the following assumptions for the state based potential game G:

A-1: $\Phi(x, \mathbf{0})$ is continuously differentiable and bounded below on x and $\Phi(x, a)$ is convex and differentiable on variable a.

A-2: $\nabla_a \Phi(x, a)$ is a Lipschitz function on variable a, i.e., there exist a constant L such that for any $x \in X$ and for any $a, a' \in \mathcal{A}(x)$, $||\nabla_a \Phi(x, a) - \nabla_a \Phi(x, a')||_2 \leq L||a - a'||_2$ where $\nabla_a \Phi(x, a) = (\frac{\partial \Phi}{\partial a_1}, \ldots, \frac{\partial \Phi}{\partial a_n})$.

Theorem 4. *Let G be a state based potential game with a potential function $\Phi(x, a)$ that satisfies Assumption (A-1,2). If the step size ϵ_i is smaller than $2/L$ for all $i \in N$, then the state action pair $[x(t), a(t)]$ of the gradient play process in (12) asymptotically converges to a stationary state Nash equilibrium of the form $[x, \mathbf{0}]$.*

Proof: From the definition of the state based potential game, we have $\Phi(x(t+1), \mathbf{0}) = \Phi(x(t), a(t))$ where $x(t+1) = f(x(t), a(t))$. We will first prove that $\Phi(x(t+1), \mathbf{0})$ is monotonically decreasing during the gradient play process provided that the step size is sufficiently small. The gradient play process in (12) can be expressed using the state based potential function as

$$\begin{aligned} a_i(t) &= \left[-\epsilon \cdot \frac{\partial J_i(x(t), a)}{\partial a_i} \bigg|_{a=\mathbf{0}} \right]^+ \\ &= \left[-\epsilon \cdot \frac{\partial \Phi(x(t), a)}{\partial a_i} \bigg|_{a=\mathbf{0}} \right]^+ \end{aligned} \tag{13}$$

Therefore, we have

$$\begin{aligned} \Phi(x(t+1), \mathbf{0}) - \Phi(x(t), \mathbf{0}) &= \Phi(x(t), a(t)) - \Phi(x(t), \mathbf{0}) \\ &\leq a(t)^T \left. \frac{\partial \Phi(x(t), a)}{\partial a} \right|_{a=\mathbf{0}}^T + \frac{L}{2} \|a(t)\|_2^2 \end{aligned}$$

where the second inequality is based on Proposition A.24 in [36]. By the Projection Theorem (Proposition 2.1.3 in [36]), we know that

$$\left(-\epsilon_i \cdot \frac{\partial \Phi(x(t), a)}{\partial a_i} \bigg|_{a=\mathbf{0}} - a_i(t) \right)^T \cdot (-a_i(t)) \leq 0$$

which is equivalent to

$$a_i(t)^T \cdot \frac{\partial \Phi(x(t), a)}{\partial a_i} \bigg|_{a=\mathbf{0}} \leq -\frac{1}{\epsilon_i} a_i(t)^T a_i(t).$$

If ϵ_i is smaller than $\frac{2}{L}$ for all $i \in N$, we have that

$$\Phi(x(t+1), \mathbf{0}) - \Phi(x(t), \mathbf{0}) \leq \sum_i \left(\frac{L}{2} - \frac{1}{\epsilon_i} \right) \|a_i(t)\|_2^2 \leq 0$$

and the equality holds in the second inequality if and only if $a(t) = \mathbf{0}$. Therefore, $\Phi(x(t), \mathbf{0})$ is monotonically decreasing along the trajectory $x(t)$. Since $\Phi(x(t), \mathbf{0})$ is bounded below, $\Phi(x(t), \mathbf{0})$ keeps decreasing until it reaches a fixed point, which means $a(t) = \mathbf{0}$. By Lemma 12 in Appendix-C, we know that such a fixed point is a stationary state Nash equlibrium. Hence $[x(t), a(t)]$ converges to a stationary state Nash equilibrium in the form of $[x, \mathbf{0}]$. ■

First note that the asymptotic guarantees given in Theorem 4 hold for heterogenous step sizes. This implies that the agents can take actions synchronously or asynchronously without altering the asymptotic guarantees. Second, the rate of convergence of gradient play depends on the structure of the potential function Φ, the state transition function f, and the stepsize ϵ_i. Larger step sizes ϵ_i generally leads to faster convergence but can also lead to instability. The bound on the stepsize ϵ_i in Theorem 4 is conservative as larger stepsize can usually be used without losing stability. Moreover, the stepsizes can vary with time as long as some additional mild conditions are satisfied.[8]

The following theorem establishes the convergence rate of the gradient play algorithm for state based potential games. For ease of exposition, we let $\epsilon_i = \epsilon_j = \epsilon$ for all the agents $i, j \in N$ and $\mathcal{A}_i(x) = \mathbb{R}^{d_x}$ for some dimension d_x, which means that the gradient play algorithm in (12) takes on the form: $a_i(t) = -\epsilon \cdot \frac{\partial J_i(x(t), a)}{\partial a_i} \Big|_{a=\mathbf{0}}$. Additionally, we make the following assumptions.

A-3 : The state transition rule is linear, namely that $\tilde{x} = f(x, a) = x + Ba$. Thus $\Phi(x, a) = \Phi(x + Ba, \mathbf{0})$ for all $a \in \mathcal{A}(x)$.

A-4 : There exit constants $M, m > 0$ such that for any $[x, a] \in X \times \mathcal{A}$,

$$\frac{m}{2}||a||^2 \leq \Phi(x, a) - \Phi(x, 0) - a^T \cdot \nabla_a \Phi|_{(x,0)} \leq \frac{M}{2}||a||^2.$$

Note that if $\Phi(x, a)$ is a strict convex function on variable a, one choice for M, m is that

$$\begin{aligned} M &= \max_{[x,a] \in X \times \mathcal{A}} \left(\sigma_{\max} \nabla_a^2 \Phi(x, a) \right); \\ m &= \min_{[x,a] \in X \times \mathcal{A}} \left(\sigma_{\min} \nabla_a^2 \Phi(x, a) \right). \end{aligned}$$

Here $\nabla_a^2 \Phi(x, a)$ denotes the Hessian matrix of Φ on variable a and σ denotes the singular values of this matrix.

Theorem 5. *Let G be a state based potential game that satisfies Assumptions (A-1,3,4). If the step size ϵ is smaller than $2/M$, then the state action pair $[x(t), a(t)]$ of the gradient play process asymptotically converges to a stationary state Nash equilibrium of the form $[x^*, \mathbf{0}]$. Moreover, $\Phi(x(t), a(t))$ is monotonically non-increasing and for all $t > 1$,*

$$\Phi(x(t), a(t)) - \Phi(x^*, \mathbf{0}) \leq \theta \cdot \left(\Phi(x(t-1), a(t-1) - \Phi(x^*, \mathbf{0}) \right)$$

[8] This is similar with the gradient methods in optimization literature [39].

where $\theta = \left(1 - 2m(\epsilon - \frac{M}{2}\epsilon^2)\right)$.

Proof: Please see Appendix. ■

B. Gradient play for our designed game

Suppose that $\mathcal{V}_i$ is a closed convex set for all $i \in N$. The gradient play algorithm applied to the game depicted in Section III-A takes on the following form. At each time $t \geq 0$, given the state $x(t) = (v(t), e(t))$, each agent i selects an action $a_i = (\hat{v}_i, \hat{e}_i)$ according to

$$\begin{aligned} \hat{v}_i(t) &= \left[-\epsilon_i^v \cdot \frac{\partial J_i\left(x(t), a\right)}{\partial \hat{v}_i}\bigg|_{a=0} \right]^+ \quad (14) \\ &= \left[-\epsilon_i^v (n\ \phi_i|_{e_i(t)} + 2n\alpha \sum_{j\in N_i} (e_i^i(t) - e_j^i(t))) \right]^+ \\ \hat{e}_{i\to j}^k(t) &= -\epsilon_{i\to j}^{k,e} \cdot \frac{\partial J_i\left(x(t), a\right)}{\partial \hat{e}_{i\to j}^k}\bigg|_{a=0} \\ &= \epsilon_{i\to j}^{k,e} \cdot \big(\phi_k|_{e_i(t)} - \phi_k|_{e_j(t)} + 2\alpha \left(e_i^k(t) - e_j^k(t)\right) \\ &\quad + 2\alpha \sum_{l\in N_i} \cdot \left(e_i^k(t) - e_l^k(t)\right) \big) \quad (15) \end{aligned}$$

where $[\cdot]^+$ represents the projection onto the closed convex set $\mathcal{A}_i^{\hat{v}}(x) = \{\hat{v}_i : v_i + \hat{v}_i \in \mathcal{V}_i\}$; and ϵ_i^v and $\left\{\epsilon_{i\to j}^{k,e}\right\}_{j\in N_i}$ are the stepsizes which are positive constants.

If $\phi(v)$ in (1) is a bounded differentiable convex function, it is straightforward to verify that the designed state based potential game satisfies Assumptions (A-1,2). Therefore, if the step sizes are sufficiently small, Theorem 4 ensures that the gradient play algorithm (14,15) will converge to a stationary state Nash equilibrium in the form of $[(v, e), \mathbf{0}]$, where v is the optimal solution of (1). Moreover, notice that the station transition rule given in (5) is linear; hence Theorem 5 guarantees a linear convergence rate.

V. Illustrations

In this section we illustrate the theoretical developments in this paper on two independent problems. The first problem rigorously explores our state based game design on the motivational example given in Section II-A. The second problem focuses on distributed routing with information constraints.

A. A simple example

Following the state based game design rule given in Section III-A, each agent $i \in \{1, 2, 3\}$ in the example in Section II-A is assigned a local state variable of the form $x_i = (v_i, e_i^1, e_i^2, e_i^3)$ where e_i^k is agent i's estimate of agent k's value v_k. Agent i's action a_i is of the form $a_i = (\hat{v}_i, \hat{e}_i^1, \hat{e}_i^2, \hat{e}_i^3)$ where $\hat{e}_i^k = \left\{\hat{e}_{i\to j}^k\right\}_{j\in N_i}$ for $k = 1, 2, 3$. The state transition rule and local cost function are defined in (5) and (8) respectively.

For concreteness, consider agent 1 as an example.

- A state associated with agent 1 is of the form $x_1 = (v_1, e_1^1, e_1^2, e_1^2)$.
- An action associated with agent 1 is of the form $a_1 = (\hat{v}_1, \hat{e}_{1\to 2}^1, \hat{e}_{1\to 2}^2, \hat{e}_{1\to 2}^3)$.
- The state transition rule is of the form $[\tilde{v}, \tilde{e}] = f\left([v, e], [\hat{v}, \hat{e}]\right)$ where

$$\begin{aligned} \tilde{v}_1 &= v_1 + \hat{v}_1, \\ \tilde{e}_1^1 &= e_1^1 + \hat{v}_1 - \hat{e}_{1\to 2}^1 + \hat{e}_{2\to 1}^1 \\ \tilde{e}_1^2 &= e_1^2 - \hat{e}_{1\to 2}^2 + \hat{e}_{2\to 1}^2 \\ \tilde{e}_1^3 &= e_1^3 - \hat{e}_{1\to 2}^3 + \hat{e}_{2\to 1}^3. \end{aligned}$$

- The local cost function of agent 1 is of the form

$$J_1\left([v, e], [\hat{v}, \hat{e}]\right) = \phi(\tilde{e}_1^1, \tilde{e}_1^2, \tilde{e}_1^3) + \frac{\alpha}{2} \sum_{k=1,2,3} \left(\tilde{e}_1^k - \tilde{e}_2^k\right)^2$$

Figure 1 shows simulation results associated with this example. The top figure includes the following: (i) the red curve shows the dynamics of ϕ using a centralized gradient method, (ii) the blue curve shows the dynamics of ϕ using our proposed state based game design with gradient play where agents take actions synchronously with a homogeneous step size $\epsilon = 0.02$, and (iii) the black curve shows the dynamics of ϕ using our proposed state based game design with gradient play where agents take actions asynchronously with heterogeneous step sizes, $\epsilon_1 = 0.01$, $\epsilon_2 = 0.02$, and $\epsilon_3 = 0.015$. In the asynchronous simulation, each agent took an action with probability 0.9 or took the null action $\mathbf{0}$ with probability 0.1. Lastly, we set $\alpha = 1$ for the above simulation. These simulations demonstrate that our state based game design can efficiency solve the optimization problem under the presented informational constraints. Furthermore, the agents achieve the correct estimate of the true value v as highlighted in the bottom figure. Note that the bottom figure only highlights the estimation errors for agent 1 as the plots for agents 2 and 3 are similar.

B. Distributed routing problem

In this section we focus on a simple distributed routing problem with a single source, a single destination, and a disjoint set of routes $\mathcal{R} = \{r_1, ..., r_m\}$. There exists a set of agents $N = \{1, ..., n\}$ each seeking to send an amount traffic, represented by $Q_i \geq 0$, from the source to the destination. The action set $\mathcal{V}_i$ for each agent is defined as

$$\left\{ v_i = (v_i^{r_1}, , v_i^{r_m}) : 0 \leq v_i^r \leq 1, \forall r \in \mathcal{R}; \sum_{r\in\mathcal{R}} v_i^r = 1 \right\} \quad (16)$$

where v_i^r represents that percentage of traffic that agent i designates to route r. Alternatively, the amount of traffic that agent i designates to route r is $v_i^r Q_i$. Lastly, for each route $r \in \mathcal{R}$, there is an associated "congestion function" of the form: $c_r : [0, +\infty) \to \mathbb{R}$ that reflects the cost of using the route as a function of the amount of traffic on that route.[9] For a given routing decision $v \in \mathcal{V}$, the total congestion in the network takes the form

$$\phi(v) = \sum_{r\in\mathcal{R}} f_r \cdot c_r(f_r)$$

[9]This type of congestion function is referred to an anonymous in the sense that all agents contribute equally to traffic. Non-anonymous congestion functions could also be used for this example.

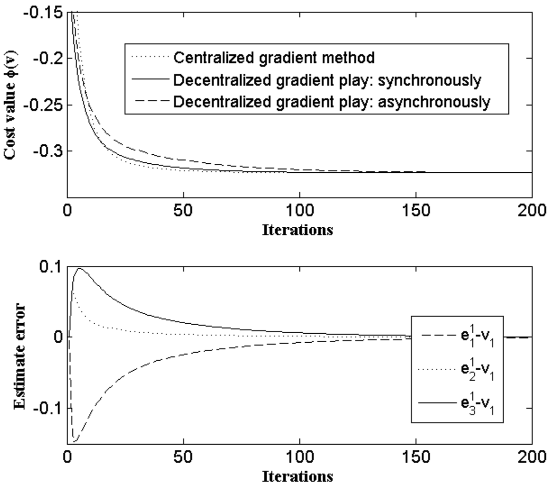

Fig. 1. Simulation results for the optimization problem in (II-A). The top figure shows the evolution of the system cost $\phi(v)$ using (i) centralized gradient algorithm, (ii) our proposed state based game design with gradient play, homogeneous step sizes, and synchronously updates (blue), and (iii) our proposed state based game design with gradient play, heterogeneous step sizes, and asynchronously updates (black). The bottom figure shows the evolution of agent 1's estimation errors, i.e., $e_1^1 - v_1$, $e_1^2 - v_2$, and $e_1^3 - v_3$, during the gradient play algorithm with homogeneous step sizes and synchronous updates.

where $f_r = \sum_{i \in N} v_i^r Q_i$. The goal is to establish a local control law for each agent that converges to the allocation which minimizes the total congestion, i.e., $v^* \in \arg\min_{v \in \mathcal{V}} \phi(v)$. One possibility for a distributed algorithm is to utilize a gradient decent algorithm where each agent adjust traffic flows according to

$$\frac{\partial \phi}{\partial v_i^r} = Q_i \cdot \left(c_r' \left(\sum_{i \in N} Q_i v_i^r \right) + c_r \left(\sum_{i \in N} Q_i v_i^r \right) \right)$$

where $c_r'(\cdot)$ represents the gradient of the congestion function. Note that implementing this algorithm requires each agent to have complete information regarding the decision of all other agents. In the case of non-anonymous congestion functions this informational restriction would be even more pronounced.

Using the methodology developed in this paper, we can localize the information available to each agent by allowing them only to have estimates of other agents' flow patterns. Consider the above routing problem with 10 agents and the following communication graph

$$1 \leftrightarrow 2 \leftrightarrow 3 \leftrightarrow \cdots \leftrightarrow 10.$$

Now, each agent is only aware of the traffic patterns for at most two of the other agents and maintaining and responding to estimates of the other agents' traffic patterns. Suppose we have 5 routes where each route $r \in \mathcal{R}$ has a quadratic congestion function of the form $c_r(k) = a_r k^2 - b_r k + c_r$ where $k \geq 0$ is the amount of traffic, and a_r, b_r, and c_r are positive and randomly chosen coefficients. Set the tradeoff parameter α to be 900. Figure 2 illustrates the results of the gradient play algorithm presented in Section IV coupled with our game design in Section III. Note that our algorithm does not perform as well as the centralized gradient descent algorithm in transient. This

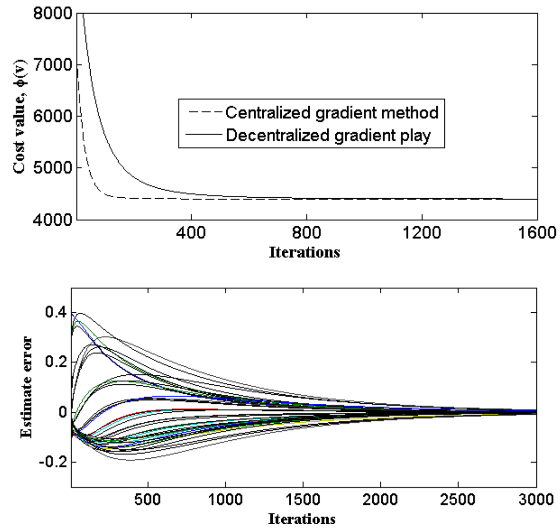

Fig. 2. Simulation results: The upper figure shows the evolution of the system cost ϕ using the centralized gradient decent algorithm (red) and our proposed algorithm (black). The bottom figure shows the evolution of agent 1's estimation error, i.e., $e_1^{k,r} - v_k^r$ for each route $r \in \mathcal{R}$ and each agent $k \in N$.

is expected since the informational availability to the agents is much lower. However, the convergence time is comparable which is surprising.

VI. Conclusion

This work presents an approach to distributed optimization using the framework of state based potential games. In particular, we provide a systematic methodology for localizing the agents' objective function while ensuring that the resulting equilibria are optimal with regards to the system level objective function. Furthermore, we proved that the learning algorithm gradient play guarantees convergence to a stationary state Nash equilibria in any state based potential game. By considering a game theoretic approach to distributed optimization, as opposed to the more traditional algorithmic approaches, we were able to attain immediate robustness to variation in clock rates and step sizes as highlighted in Sections III and IV. There are several open and interesting questions that this paper promotes. One in particular is regarding the communication requirements on the agents. In our design, each agent possessed n additional state variables as estimates for the n components of the value profile v. Could similar guarantees be attained with less variables? What happens if we transition from a fixed to time varying communication topology? Lastly, how does this approach extend to alternative classes of system level objective functions?

References

[1] N. Li and J. R. Marden. Design games for distributed optimization. In *Proceedings of 50th IEEE Conference on Decision and Control*, 2011.

[2] J. Tsitsiklis and M. Athans. Convergence and asymptotic agreement in distributed decision problems. *IEEE Transactions on Automatic Control*, 29(1):42–50, 1984.

[3] 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.

[4] A. Tang, J. Wang, S. H. Low, and M. Chiang. Equilibrium of heterogeneous congestion control: Existence and uniqueness. *IEEE/ACM Transactions on Networking*, 15(4):824–837, 2007.

[5] V. Mhatre, K. Papagiannaki, and F. Baccelli. Interference mitigation through power control in high density 802.11. In *Proceedings of IEEE INFOCOM*, 2007.

[6] E. Campos-Nanez, A. Garcia, and C. Li. A game-theoretic approach to efficient power management in sensor networks. *Operations Research*, 56(3):552, 2008.

[7] G. Scutari, D. P. Palomar, J. Pang, and F. Facchinei. Flexible design of cognitive radio wireless systems. *IEEE Signal Processing Magazine*, 26(5):107–123, 2009.

[8] L. Chen, S. H. Low, and J. C. Doyle. Random access game and medium access control design. *IEEE/ACM Transactions on Networking*, (4):1303–1316, 2010.

[9] J. R. Marden and M. Effros. The price of selfiness in network coding. *IEEE Transactions on Information Theory*, 58(4):2349–2361, 2012.

[10] V. Reddy, S. Shakkottai, A. Sprintson, and N. Gautam. Multipath wireless network coding: a population game perspective. In *Proceedings of IEEE INFOCOM*, 2010.

[11] R. Gopalakrishnan, J. R. Marden, and A. Wierman. An architectural view of game theoretic control. *ACM SIGMETRICS Performance Evaluation Review*, 38(3):31–36, 2011.

[12] D. Monderer and L.S. Shapley. Potential games. *Games and Economic Behavior*, 14:124–143, 1996.

[13] D. Fudenberg and D. K. Levine, editors. *The Theory of Learning in Games*. MIT Press, Cambridge, MA, 1998.

[14] H. P. Young. *Strategic Learning and Its Limits*. Oxford University Press, Oxford, UK, 2004.

[15] J.R. Marden, G. Arslan, and J.S. Shamma. Joint strategy fictitious play with inertia for potential games. *IEEE Transactions on Automatic Control*, 54(2):208–220, 2009.

[16] J. R. Marden, H. P. Young, G. Arslan, and J. S. Shamma. Payoff based dynamics for multi-player weakly acyclic games. *SIAM Journal on Control and Optimization*, 48(1):373–396, 2009.

[17] J. R. Marden and A. Wierman. Overcoming limitations of game-theoretic distributed control. In *48th IEEE Conference on Decision and Control*, 2009.

[18] LS Shapley. A value for n-person games. *Contributions to the theory of games*, 2:307–317, 1953.

[19] E. Anshelevich, A. Dasgupta, J. Kleinberg, E. Tardos, T. Wexler, and T. Roughgarden. The price of stability for network design with fair cost allocation. *SIAM Journal on Computing*, 38(4):1602–1623, 2008.

[20] J. R. Marden. State based potential games. *Automatica*, 48(12):3075–3088, 2012.

[21] S. D. Flam. Equilibrium, evolutionary stability and gradient dynamics. *International Game Theory Review*, 4(4):357–370, 2002.

[22] J. Shamma and G. Arslan. Dynamic fictitious play, dynamic gradient play, and distributed convergence to nash equilibria. *IEEE Transactions on Automatic Control*, 50(3):312–327, 2005.

[23] M. V. Solodov. Incremental gradient algorithms with stepsizes bounded away from zero. *Computational Optimization and Applications*, 11(1):23–35, 1998.

[24] D. Blatt, A.O. Hero, and H. Gauchman. A convergent incremental gradient method with a constant step size. *SIAM Journal on Optimization*, 18(1):29–51, 2008.

[25] J. Tsitsiklis, D. Bertsekas, and M. Athans. Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *IEEE Transactions on Automatic Control*, 31(9):803–812, 1986.

[26] A. Nedic, A. Olshevsky, A. Ozdaglar, and J.N. Tsitsiklis. On distributed averaging algorithms and quantization effects. *IEEE Transactions on Automatic Control*, 54(11):2506–2517, 2009.

[27] I. Lobel and A. Ozdaglar. Distributed subgradient methods for convex optimization over random networks. *IEEE Transactions on Automatic Control*, (99):1–1, 2010.

[28] M. Zhu and S. Martinez. On distributed convex optimization under inequality and equality constraints via primal-dual subgradient methods. *Arxiv preprint arXiv:1001.2612*, 2010.

[29] A. Jadbabaie, J. Lin, and A.S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48(6):988–1001, 2003.

[30] Y. Hatano and M. Mesbahi. Agreement over random networks. *IEEE Transactions on Automatic Control*, 50(11):1867–1872, 2005.

[31] IP Androulakis and GV Reklaitis. Approaches to asynchronous decentralized decision making. *Computers and Chemical Engineering*, 23(3):339–354, 1999.

[32] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Foundation and Trends in Machine Learning*, 3(1):1–122, 2010.

[33] C. Ding, B. Song, A. Morye, J.A. Farrell, and A.K. Roy-Chowdhury. Collaborative sensing in a distributed ptz camera network. *IEEE Transactions on Image Processing*, 21(7):3282–3295, 2012.

[34] M.S. Stankovic, K.H. Johansson, and D.M. Stipanovic. Distributed seeking of nash equilibria with applications to mobile sensor networks. *IEEE Transactions on Automatic Control*, 57(4):904–919, 2012.

[35] N. Li and J. R. Marden. Decouple coupled constraints through utility design. *Under Submission*, 2011. Available on http://ecee.colorado.edu/marden/publications.html.

[36] D.P. Bertsekas. *Nonlinear programming, 2nd edition*. Athena Scientific Belmont, MA, 2008.

[37] L. S. Shapley. Stochastic games. *Proceedings of the National Academy of Science of the United States of America*, 39(10):1095–1100, 1953.

[38] N. Nisan, T. Roughgarden, E. Tardos, and V.V. Vazirani. *Algorithmic Game Theory*. Cambridge University Press, 2007.

[39] M.S. Bazaraa, H.D. Sherali, and C.M. Shetty. *Nonlinear programming: theory and algorithms, 3rd edition*. John Wiley & Sons, Inc., NJ, 2006.

[40] N. Li and J. R. Marden. Designing games to handle coupled constraints. In *Proceedings of 49th IEEE Conference on Decision and Control*, 2010.

Appendix

A. An impossibility result for game design

This section addresses the question as to whether the results in Theorem 2 and 3 could have been attained using the framework of strategic form games. More specifically, is it possible to design agent objective functions that achieves the following four objectives:

- Each agent's cost function relies solely on local information as defined by the communication graph. Moreover, agents' cost functions should possess a degree of scalability with regards to the size of the system and the topology of the communication graph.
- All Nash equilibria of the resulting game represent solutions to the optimization problem (1);
- The resulting game possesses an underlying structure that can be exploited by distributed learning algorithm, e.g., potential games.

Accomplishing these objectives would ensure that the agents' control policies resulting from the designed game plus a suitable learning algorithm would be of the local form in (2).

In the following we demonstrate that achieving these objectives using the framework of strategic form games is impossible in general. To show this we focus on the following optimization problem

$$\begin{array}{ll} \min_v & \left(\sum_{i\in N} v_i\right)^2 \\ \text{s.t.} & v_i \in [c_i, d_i] \subset \mathbb{R}. \end{array} \tag{17}$$

To make the control laws $\{F_i(\cdot)\}_{i\in N}$ scalable as to the agent set and the communication graph $\mathcal{G}$, we require that the underlying control design must be invariant to the agents' indices. This implies that if two agents (i,j) have the same number of neighbors, i.e., $|N_i| = |N_j|$, and for each agent k in N_i there is an agent h in N_j such that $v_k = v_h$ and $[c_k, d_k] = [c_h, d_h]$, and vice versa, then the control policies of agent i, j should be the same, i.e., $F_i\left(\{v_k, c_k, d_k\}_{k\in N_i}\right) = F_j\left(\{v_k, c_k, d_k\}_{k\in N_j}\right)$.

Accordingly, we formulate the optimization problem as a game where the agent set is N, the action set of each agent is the set $\mathcal{A}_i = [c_i, d_i]$, and each agent is assigned a cost function of the form $J_i : \prod_{j\in N_i} \mathcal{V}_j \to \mathbb{R}$. To facilitate the design of scalable agent control policies, we focus on the design of agent cost functions of the form:

$$J_i(v) = J\left(\{v_j, c_j, d_j\}_{j\in N_i}\right) \quad (18)$$

where the function $J(\cdot)$ is invariant to specific indices assigned to agents. Notice that this design of $J(\cdot)$ leads to a well defined game irrespective of the agent set N, constraint sets $[c_i, d_i]$ or the structure of the communication graph $\{N_i\}_{i\in N}$. The following proposition demonstrates that it is impossible to design $J(\cdot)$ such that for any game induced by a constraint profile $[c, d]$ and communication graph $\mathcal{G}$ all resulting Nash equilibria solve the optimization problem in (3).

Proposition 6. *There does not exist a single $J(\cdot)$ such that for any game induced by a connected communication graph $\mathcal{G}$, a constraint profile $[c, d]$, and agents' cost functions of the form (18), the Nash equilibria of the induced game represent solutions to the optimization problem in (17).*

Proof: Suppose that there exists a single $J(\cdot)$ that satisfies the proposition. We will now construct a counterexample to show that this is impossible. Consider two optimization problems of the form (17) with a single communication graph given by

$$1 \leftrightarrow 2 \leftrightarrow 3 \leftrightarrow 4 \leftrightarrow 5 \leftrightarrow 6.$$

Here, we have $N = \{1, 2, 3, 4, 5, 6\}$ and $\mathcal{E} = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 5\}, \{5, 6\}\}$. In the first optimization problem the constraint profile is: $[c_1, d_1] = [c_6, d_6] = [-1, -\frac{21}{22}]$, $[c_2, d_2] = [c_3, d_3] = [c_4, d_4] = [\frac{6}{11}, \frac{7}{11}]$, and $[c_5, d_5] = [0, 0]$. In the second optimization problem, the constraint profile is:$[c_1, d_1] = [c_6, d_6] = [-1, -\frac{21}{22}]$ and $[c_2, d_2] = [c_3, d_3] = [c_4, d_4] = [c_5, d_5] = [\frac{6}{11}, \frac{7}{11}]$. We call the settings for the two optimization problems as setting (a) and (b) respectively. Under those constraints, the optimal solution for setting (a) is $v^a = (v_1^a, v_2^a, v_3^a, v_4^a, v_5^a, v_6^a) = (-\frac{21}{22}, \frac{7}{11}, \frac{7}{11}, \frac{7}{11}, 0, -\frac{21}{22})$ and the optimal solutions for setting (b) is $v^b = (v_1^b, v_2^b, v_3^b, v_4^b, v_5^b, v_6^b) = (-1, \frac{6}{11}, \frac{6}{11}, \frac{6}{11}, \frac{6}{11}, -1)$.

We start by defining agent cost functions of the form (18) which ensures that v^a is a Nash equilibrium for setting (a). This implies that for any agent $i \in N$, we have

$$J\left(\{v_i^a, c_i, d_i\}, \{v_j^a, c_j, d_j\}_{j\in N_i\backslash\{i\}}\right) \leq J\left(\{v_i, c_i, d_i\}, \{v_j^a, c_j, d_j\}_{j\in N_i\backslash\{i\}}\right) \quad (19)$$

for any $v_i \in \mathcal{V}_i$. By writing down the Nash equilibrium condition in (19) for setting (b), it is straightforward to see that agents $1, 2, 3, 4, 5, 6$ in setting (b) have the same structure form of the cost function as agents $1, 2, 3, 3, 2, 1$ in setting (a) respectively. For example, agent 4 in setting (a) has an identical cost function to agent 3 in setting (b). Since v^a represents a Nash equilibrium for setting (a) then no agent $i \in \{1, \ldots, 6\}$ has a unilateral incentive to deviate from v^a. As agents $1, 2, 3, 4, 5, 6$ in (b) can be mapped precisely to agents $1, 2, 3, 3, 2, 1$ in (a), $v^* = (v_1^*, v_2^*, v_3^*, v_4^*, v_5^*, v_6^*) = (-\frac{21}{22}, \frac{7}{11}, \frac{7}{11}, \frac{7}{11}, \frac{7}{11}, -\frac{21}{22})$ is a Nash equilibrium of setting (b) since no agent $i \in \{1, \ldots, 6\}$ has a unilateral incentive to deviate from v^*. The impossibility comes from the fact that v^* is not an optimal solution to setting (b). ■

B. Proof of Theorem 3

Since the designed state based game is a state based potential game, we can apply Proposition 1 to prove the sufficient condition of the theorem. The proof involves two steps: (i) If x^* satisfy Condition (a)-(b) listed in the theorem, then $x^* \in \operatorname{argmin}_{x\in X} \phi(x, \mathbf{0})$; (ii) if a^* satisfy Condition (c)-(d) in the theorem, then $x = f(x, a)$ for all $x \in X$. Therefore it is straightforward to prove that if a state action pair $[x, a]$ satisfies Conditions (a)-(d) listed in the theorem, then $[x, a]$ is a stationary state Nash equilibrium.

Let us prove the necessary condition of Theorem 3. Suppose $[x, a]$ is a stationary state Nash equilibrium. First notice that to ensure $[x, a]$ satisfies Condition (D-2) of Definition 3, i.e. $x = f(x, a)$, the action profile $a = (\hat{v}, \hat{e})$ should satisfy Condition (c)-(d) of this theorem. To prove Condition (a)-(b), we will use a series of lemmas to prove that under one of Cases (i)-(iv) of this theorem, if a station action pair $[x, a]$ satisfies Condition (D-1) of Definition 3, i.e. $a_i \in \operatorname{argmin}_{\check{a}} J_i(x, \check{a}_i, a_{-i})$ for all $i \in N$, then the ensuing state $\tilde{x} = f(x, a)$ satisfies the following conditions:

1) *Estimation alignment:* The ensuing estimation terms are aligned with the ensuing value profile, i.e., for all agents $i, k \in N$ we have $\tilde{e}_i^k = \tilde{v}_k$ where $(\tilde{v}, \tilde{e}) = f(x, a)$. (Lemma 7 for Case (i)–(ii), Lemma 8 for Case (iii) and Lemma 10 for Case (iv).)
2) *Optimality alignment:* The ensuing value profile $\tilde{v}$ is an optimal solution to (1). (Lemma 9 for Case (i)–(iii) and Lemma 10 for Case (iv).)

Combining with the fact that $\tilde{x} = f(x, a) = x$, we can conclude that under one of Cases (i)-(iv) of this theorem if $[x, a]$ is a state based Nash equilibrium, then Condition (a)-(d) must be satisfied.

In the subsequent lemmas we consistently express the ensuing state for a state action pair $[x, a] = [(v, e), (\hat{v}, \hat{e})]$ as $(\tilde{v}, \tilde{e}) = f(x, a)$.

Lemma 7. *Suppose the communication graph $\mathcal{G}$ satisfies either Condition (i) or (ii) of Theorem 3. If $[x, a]$ satisfies $a_i \in \operatorname{argmin}_{\check{a}\in\mathcal{A}_i(x)} J_i(x, \check{a}, a_{-i})$ for all $i \in N$, then all agent have correct estimates of the value profile. That is, for all agents $i, k \in N$ we have $\tilde{e}_i^k = \tilde{v}_k$.*

Proof: If $a_i \in \operatorname{argmin}_{\check{a}_i=(\check{v}_i,\check{e}_i)\in\mathcal{A}_i(x)} J_i(x, \check{a}_i, a_{-i})$ for all $i \in N$, then

$$\left.\frac{\partial J_i(x, \check{a}_i, a_{-i})}{\check{e}_{i,l}^k}\right|_{a_i} = 0, \forall i, k \in N, l \in N_i\backslash\{i\}$$

which is equivalent to

$$\phi_k|_{\tilde{e}_i} + 2\alpha \sum_{j\in N_i} \left(\tilde{e}_i^k - \tilde{e}_j^k\right) = \phi_k|_{\tilde{e}_l} - 2\alpha \left(\tilde{e}_i^k - \tilde{e}_l^k\right) \quad (20)$$

where $\phi_k|_{\tilde{e}_i}$ represents the derivative of ϕ relative to $\tilde{e}_i^k$ for the profile $\tilde{e}_i$, i.e., $\phi_k|_{\tilde{e}_i} = \frac{\partial\phi(\tilde{e}_i)}{\partial\tilde{e}_i^k}$. Consider any two connected agents $i, j \in N$, i.e., $(i,j) \in \mathcal{E}$. The equality in (20) translates to

$$\begin{aligned} \phi_k|_{\tilde{e}_i} + 2\alpha \textstyle\sum_{l \in N_i} \left(\tilde{e}_i^k - \tilde{e}_l^k\right) &= \phi_k|_{\tilde{e}_j} - 2\alpha \left(\tilde{e}_i^k - \tilde{e}_j^k\right) \\ \phi_k|_{\tilde{e}_j} + 2\alpha \textstyle\sum_{l \in N_j} \left(\tilde{e}_j^k - \tilde{e}_l^k\right) &= \phi_k|_{\tilde{e}_i} - 2\alpha \left(\tilde{e}_j^k - \tilde{e}_i^k\right). \end{aligned}$$

Adding these two equalities gives us

$$\textstyle\sum_{l \in N_i} (\tilde{e}_i^k - \tilde{e}_l^k) = - \sum_{l \in N_j} (\tilde{e}_j^k - \tilde{e}_l^k) \tag{21}$$

for all agents $i, j, k \in N$ such that $(i,j) \in N_j$. Since our communication graph is connected, the equality condition in (21) tells us that the possible values for the summation terms $\sum_{l \in N_i} (\tilde{e}_i^k - \tilde{e}_l^k)$ for each agent $i \in N$ can be at most one of two possible values that differ purely with respect to sign, i.e., for any agent $i \in N$ we have

$$\textstyle\sum_{l \in N_i} (\tilde{e}_i^k - \tilde{e}_l^k) \in \left\{ e_{\text{diff}}^k, -e_{\text{diff}}^k \right\} \tag{22}$$

where $e_{\text{diff}}^k \in \mathbb{R}$ is a constant. We can utilize the underlying topology of the communication graph coupled with (22) to demonstrate that $e_{\text{diff}}^k = 0$.

1) If there exists a cycle in the communication graph with an odd number of nodes, applying equality (21), we can get that $e_{\text{diff}}^k = -e_{\text{diff}}^k$, which tells us that $e_{\text{diff}}^k = 0$.
2) Since the communication graph is undirected we know that $\sum_{i \in N} \sum_{l \in N_i} (\tilde{e}_i^k - \tilde{e}_l^k) = 0$. If the number of agents n is odd, condition (22) tells that $\sum_{i \in N} \sum_{l \in N_i} (\tilde{e}_i^k - \tilde{e}_l^k) = h \cdot e_{\text{diff}}^k$ where h is a nonzero integer. Hence $e_{\text{diff}}^k = 0$.

In summary, if the total number of agents is odd or there exists a cycle in the communication graph with odd number of nodes we have that for all $i, k \in N$, $\sum_{l \in N_i} (\tilde{e}_i^k - \tilde{e}_l^k) = 0$. Since the communication graph is connected and undirected, it is straightforward to show that for all agents $i, j \in N$, $\tilde{e}_i^k = \tilde{e}_j^k, \forall k \in N$ where the proof is the same as the proof of Theorem 1 in [40].[10] Combining with the equality (7), we get that for all agents $i, k \in N$, $\tilde{e}_i^k = v_k$. ■

Remark 1. *It is important to note that alternative graph structures may very well provide the same guarantees.*

Lemma 8. *Suppose the objective function ϕ and communication graph $\mathcal{G}$ satisfy Condition (iii) of Theorem 3. If $[x,a]$ satisfies $a_i \in \operatorname{argmin}_{\check{a}_i \in \mathcal{A}_i(x)} J_i(x, \check{a}_i, a_{-i})$ for all $i \in N$, then all agent have correct estimates of the value profile. That is, for all agents $i, k \in N$ we have $\tilde{e}_i^k = \tilde{v}_k$.*

Proof: In the proof of last lemma, we have proved that if $a_i \in \operatorname{argmin}_{\check{a}_i} J_i(x, \check{a}_i, a_{-i})$, then equation (20) should satisfy. Consider any agent $i \in N$, and any pair of agents $j_1, j_2 \in N_i$, equation (20) tells us that

$$\begin{aligned} \phi_k|_{\tilde{e}_i} + 2\alpha \textstyle\sum_{j \in N_i} \left(\tilde{e}_i^k - \tilde{e}_j^k\right) &= \phi_k|_{\tilde{e}_{j_1}} - 2\alpha \left(\tilde{e}_i^k - \tilde{e}_{j_1}^k\right) \\ \phi_k|_{\tilde{e}_i} + 2\alpha \textstyle\sum_{j \in N_i} \left(\tilde{e}_i^k - \tilde{e}_j^k\right) &= \phi_k|_{\tilde{e}_{j_2}} - 2\alpha \left(\tilde{e}_i^k - \tilde{e}_{j_2}^k\right). \end{aligned} \tag{23}$$

Combining the two equations, we have the following equality

$$\phi_k|_{\tilde{e}_{j_1}} - \phi_k|_{\tilde{e}_{j_2}} - 2\alpha \left(\tilde{e}_{j_2}^k - \tilde{e}_{j_1}^k\right) = 0.$$

[10]The main idea of this proof is to write $\sum_{l \in N_i} (\tilde{e}_i^k - \tilde{e}_l^k) = 0, \forall i \in N$ in a matrix form for each $k \in N$. The rank of this matrix is $n-1$ resulting from the fact that the communication graph is *connected* and *undirected* hence proving the result.

Note that agents j_1 and j_2 are not necessarily connected but are rather siblings as both agents are connected to agent i. Therefore, the above analysis can be repeated to show that for any siblings $j_1, j_2 \in N$ that are siblings we have the equality

$$\phi_k|_{\tilde{e}_{j_1}} - \phi_k|_{\tilde{e}_{j_2}} = 2\alpha \left(\tilde{e}_{j_2}^k - \tilde{e}_{j_1}^k\right). \tag{24}$$

for all agents $k \in N$. Applying Lemma 11 in the appendix, Condition (24) coupled with the fact that ϕ is a convex function implies that for any siblings $j_1, j_2 \in N$,

$$\tilde{e}_{j_1} = \tilde{e}_{j_2}. \tag{25}$$

Since the communication graph is connected and undirected, Equality (25) guarantees that there exist at most two different estimation values which we denote by $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$, i.e.,

$$\tilde{e}_i \in \{x, y\}, \forall i \in N. \tag{26}$$

Now applying equality (22), for each $i \in N$, we have that either $e_{\text{diff}}^k = 2n_i(x_k - y_k)$ or $e_{\text{diff}}^k = -2n_i(x_k - y_k)$, where $n_i = |N_i| - 1 > 0$. If there exist two agents having different number of neighbors, we can derive that $x = y$, i.e. $\tilde{e}_i = \tilde{e}_j, \forall i, j \in N$. Following the same arguments as the previous proof, we have that $\tilde{e}_i^k = v_k, \forall i, k \in N$. ■

Lemma 9. *Suppose that at least one of conditions (i)–(iii) of Theorem 3 is statisfied. If $[x,a]$ satisfies $a_i \in \operatorname{argmin}_{\check{a}_i \in \mathcal{A}_i(x)} J_i(x, \check{a}_i, a_{-i})$ for all $i \in N$, then $\tilde{v}$ is an optimal solution to (1).*

Proof: If $a_k \in \operatorname{argmin}_{\check{a}_i} J_k(x, \check{a}_k, a_{-k})$, where $\check{a}_k = (\check{v}_k, \check{e}_k)$,we have

$$\left.\frac{\partial J_k(x, \check{a}_k, a_{-k})}{\check{v}_k}\right|_{a_k} \cdot (\hat{v}_k' - \hat{v}_k) \geq 0, \forall \hat{v}_k' \in \mathcal{A}_i^{\hat{v}}(x)$$

which is equivalent to

$$\left[n \, \phi_k|_{\tilde{e}} + 2n_k \sum_{j \in N_k} (\tilde{e}_k^k - \tilde{e}_j^k) \right] \cdot (\hat{v}_k' - \hat{v}_k) \geq 0 \tag{27}$$

We have shown in Lemma 7 and Lemma 8 that if $[x,a] = [(v,e),(\hat{v},\hat{e})]$ satisfies $a_i \in \operatorname{argmin}_{\check{a}_i} J_i(x, \check{a}_i, a_{-i})$, then $\tilde{e}_i^k = v_k, \forall i, k \in N$. Therefore, equation (27) tells that

$$\phi_k|_{(\tilde{v})} \cdot (\tilde{v}_k' - \tilde{v}_k) \geq 0, \forall \tilde{v}_k' \in \mathcal{V}_k. \tag{28}$$

This implies that $\tilde{v}$ is an optimal profile for the optimization problem (1) given that ϕ is convex over $\mathcal{V}$. ■

Lemma 10. *Suppose Condition (iv) of Theorem 3 is satisfied. If $[x,a]$ satisfies $a_i \in \operatorname{argmin}_{\check{a}_i \in \mathcal{A}_i(x)} J_i(x, \check{a}_i, a_{-i})$ for all $i \in N$, then $\tilde{e}_i^k = \tilde{v}_k$ for all $i, j \in N$, and $\tilde{v}$ is an optimal profile for the optimization problem (1).*

Proof: In the proof of Lemma 7 and Lemma 8, we have shown that if $[x,a]$ satisfies $a_i \in \operatorname{argmin}_{\check{a}_i} J_i(x, \check{a}_i, a_{-i})$, equations (20) and (27) should satisfy. Since $\mathcal{V}_k$ is open, equation (27) is equivalent to

$$\phi_k|_{\tilde{e}_k} + 2 \sum_{j \in N_k} \left(\tilde{e}_k^k - \tilde{e}_j^k\right) = 0, \ \forall k \in N \tag{29}$$

Substituting this equation into equation (20), we have

$$\phi_k|_{\tilde{e}_l} + 2\tilde{e}_l^k = 2\tilde{e}_k^k,\ \forall l \in \mathcal{N}_k, k \in \mathcal{N} \tag{30}$$

Since ϕ is a convex function, we already have equality (26) as shown in the proof of Lemma 8. We will show that $x = y$. Suppose $x \neq y$. For each $i \in N$, either $\tilde{e}_i = x$ or $\tilde{e}_i = y$. Suppose $\tilde{e}_i = x$. Then for all $j \in N_i$, $\tilde{e}_j = y$; otherwise if $\tilde{e}_j = x$ for some $j \in N_i$, equation (25) implies that $\tilde{e}_j = x, \forall j \in N$, i.e. $x = y$. Equation (29) tells us that

$$\phi_k|_x = 2n_k(y_k - x_k)$$

where $n_k = |N_k| - 1$. While equation (30) tells us that

$$\phi_k|_y = 2(x_k - y_k)$$

If $\tilde{e}_k = y$, similarly we will have:

$$\begin{array}{rcl} \phi_k|_y &=& 2n_k(x_k - y_k) \\ \phi_k|_x &=& 2(y_k - x_k) \end{array}$$

In both cases, we have $\phi_k|_x - \phi_k|_y = 2(n_k+1)(y_k - x_k)$. Applying Lemma 11, we know $x = y$. Now we can conclude that $\tilde{e}_i = \tilde{e}_j$ and hence $\tilde{e}_i^k = v_k, \forall i, k \in N$. Substituting those equalities into equation (29), we have:

$$\phi_k|_{(\tilde{v}_1,\ldots,\tilde{v}_n)} = 0, \forall k \in N$$

which implies that $\tilde{v}$ is an optimal point of the optimization problem (1) given that ϕ is an convex function and $\mathcal{V}$ is open. ■

Lemma 11. *Given a continuously differentiable convex function $\phi(x_1, x_2, \ldots, x_n)$ and two vectors $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$, if for all $k = 1, 2, \ldots n$, we have $\phi_k|_x - \phi_k|_y = \alpha_k(y_k - x_k)$ where $\alpha_k > 0$, then $x = y$.*

Proof: Since ϕ is a convex function, we have

$$\begin{array}{rcl} \phi(x) &\geq& \phi(y) + (x-y)^T \nabla\phi|_y \\ \phi(y) &\geq& \phi(x) + (y-x)^T \nabla\phi|_x \end{array}$$

Adding up the two inequalities, we have

$$0 \geq (x-y)^T(\nabla\phi|_y - \nabla\phi|_x)$$

Since $\phi_k|_x - \phi_k|_y = \alpha(y_k - x_k)$ for all k, we have

$$0 \geq \sum_k \alpha_k(x_k - y_k)^2 \geq 0$$

Therefore $x = y$. ■

C. A Lemma for Gradient Play

Lemma 12. *Let G be a state based potential game and the potential function $\Phi(x, a)$ is a differentiable convex function on variable a. Suppose all agents are using the gradient play algorithm and the state at time t is $x(t) = [v(t), e(t)]$. The action profile at time t is the null action, i.e., $a(t) = \mathbf{0}$, if and only if the state action pair $[x(t), \mathbf{0}]$ is a stationary state Nash equilibrium of the state based game G.*

Proof: Since $f(x(t), \mathbf{0}) = x(t)$, by Definition 3 we know that $[x(t), \mathbf{0}]$ is a stationary state Nash equilibrium if and only if $0 \in \mathrm{argmin}_{a_i \in \mathcal{A}_i(x(t))} J_i(x(t), a_i, 0)$ for all $i \in N$. This is equivalent to

$$\left(\frac{\partial J_i(x(t), a)}{\partial a_i}\bigg|_{a=\mathbf{0}} \right) \cdot a_i \geq 0$$

for all $i \in N$ and $a_i \in \mathcal{A}_i(x(t))$. By Projection Theorem, this inequality is equivalent to the fact that the projection of $-\epsilon_i \frac{\partial J_i(x(t),a)}{\partial a_i}\Big|_{a=\mathbf{0}}$ onto $\mathcal{A}_i(x(t))$ is 0, i.e.

$$a_i(t) = \left[-\epsilon \cdot \frac{\partial J_i(x(t), a)}{\partial a_i}\bigg|_{a=\mathbf{0}} \right]^+ = 0, \forall i \in N.$$

■

D. Proof of Theorem 5

1) From Assumption (A-4), we have

$$\begin{aligned} \Phi(x(t+1), \mathbf{0}) &= \Phi(x(t), a(t)) \\ &\leq \Phi(x(t), \mathbf{0}) + a(t)^T \nabla_a \Phi(x(t), \mathbf{0}) \\ &\quad + \tfrac{M}{2}||a(t)||_2^2 \\ &\leq \Phi(x(t), \mathbf{0}) - (\epsilon - \tfrac{M}{2}\epsilon^2)||\nabla_a \Phi(x(t), \mathbf{0})||_2^2 \end{aligned} \tag{31}$$

Therefore if $\epsilon < \frac{2}{M}$, $\Phi(x(t+1), \mathbf{0}) \leq \Phi(x(t), \mathbf{0})$.

2) Assumption (A-4) also implies the following inequality:

$$\begin{aligned} \Phi(x + Ba, \mathbf{0}) &= \Phi(x, a) \\ &\geq \Phi(x, \mathbf{0}) + a^T \cdot \nabla_a \Phi(x, 0) + \tfrac{m}{2}||a||^2 \\ &\geq \min_a \left(\Phi(x, \mathbf{0}) + a^T \cdot \nabla_a \Phi(x, 0) + \tfrac{m}{2}||a||^2 \right) \\ &= \Phi(x, \mathbf{0}) - \tfrac{1}{2m}||\nabla_a \Phi(x, \mathbf{0})||_2^2 \end{aligned}$$

Since the state transition rule is $x(t+1) = f(x(t), a(t)) = x(t) + Ba(t)$, we have:

$$\begin{array}{rcl} \Phi(x(T), \mathbf{0}) &=& \Phi(x(T-1) + Ba(T-1), \mathbf{0}) \\ &=& \ldots \\ &=& \Phi(x(t) + B\sum_{\tau=t}^{T-1} a(t), \mathbf{0}) \\ &\geq& \Phi(x(t), \mathbf{0}) - \frac{1}{2m}||\nabla_a \Phi(x(t), \mathbf{0})||_2^2 \end{array} \tag{32}$$

for any $T > t \geq 0$. If we pick $t = 0$, we know that $\{\Phi(x(T), \mathbf{0})\}_{T \geq 0}$ is bounded below. As we showed in the proof for Theorem 4, we know that $(x(t), a(t))$ will asymptotically converge to a stationary state Nash equilibrium $[x^*, \mathbf{0}]$ and $\Phi(x(t), \mathbf{0}) \geq \Phi(x^*, \mathbf{0})$ for any $t \geq 0$.

3) Since $x(T) = x(t) + B\sum_{\tau=t}^{T-1} a(\tau)$ and $\lim_{T\to\infty} x(T) = x^*$, we know that $\lim_{T\to\infty} x(t) + B\sum_{\tau=t}^{T-1} a(\tau) = x^*$. Combining with Inequality (32), we have:

$$\Phi(x^*, \mathbf{0}) - \Phi(x(t), \mathbf{0}) \geq -\frac{1}{2m}||\nabla_a \Phi(x(t), \mathbf{0})||_2^2 \tag{33}$$

for any $t \geq 0$. Substituting this into Inequality (31), we have

$$\begin{aligned} \Phi(x(t+1), \mathbf{0}) \leq{}& \Phi(x(t), \mathbf{0}) \\ &- 2m(\epsilon - \frac{M}{2}\epsilon^2)\left(\Phi(x(t), \mathbf{0}) - \Phi(x^*, \mathbf{0})\right) \end{aligned}$$

which gives the following inequality:

$$\Phi(x(t+1), \mathbf{0}) - \Phi(x^*, \mathbf{0}) \leq \theta\left(\Phi(x(t), \mathbf{0}) - \Phi(x^*, \mathbf{0})\right) \tag{34}$$

where $\theta = \left(1 - 2m(\epsilon - \frac{M}{2}\epsilon^2)\right)$. Therefore we can conclude the statement in this theorem.

□