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. Unfortunately, there are no existing methodologies for addressing this design challenge. The goal of this paper is to address this challenge using the field of game theory. Utilizing game theory for the design and control of multiagent systems requires two steps: (i) defining a local objective function for each decision maker and (ii) specifying a distributed learning algorithm to reach a desirable operating point. One of the core advantages of this game theoretic approach is that this two step process can be decoupled by utilizing specific classes of games. For example, if the designed objective functions result in a *potential game* then the system designer can utilize distributed learning algorithms for potential games to complete step (ii) of the design process. Unfortunately, designing agent objective functions to meet objectives such as locality of information and efficiency of resulting equilibria within the framework of potential games is fundamentally challenging and in many case impossible. In this paper we develop a systematic methodology for meeting these objectives using a broader framework of games termed *state based potential games*. State based potential games is an extension of potential games where an additional state variable is introduced into the game environment hence permitting more flexibility in our design space. Furthermore, state based potential games possess an underlying structure that can be exploited by distributed learning algorithms in a similar fashion to potential games hence providing a new baseline for our decomposition.**

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., [1]–[6]. These control laws provide the groundwork for a decision making architecture that possess 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 begining to emerge as a powerful tool for the design and control of multiagent systems [5]–[9]. 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*) [10]. For example, if the game is designed as a potential game [11] 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 [12]–[15], e.g., gradient play, fictitious play, and joint strategy fictitious play. Several recent papers focus on utilizing this decomposition in distributed control by developing methodologies for designing games, or more specifically agent utility functions, that adhere to this potential game structure [5], [8], [10], [16]. 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. 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 [5], [8] and the Shapley value utility [17], [18], 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 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 agent objective functions that satisfy virtually any degree of locality while ensuring the desirability of the resulting Nash equilibria. The key enabler for this result is the addition of local state variables to the game environment, i.e., moving towards *state based games* [16], [19]. Our design utilizes these state variables as a coordinating entity to decouple the system level objective into agent specific objectives of the desired interdependence. This work is complimentary to our previous work in [20] where we utilized a similar state based formulation

This research was supported by AFOSR grant #FA9550-09-1-0538

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

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

to localize coupled constraints on agents' available actions. However, in [20] we restricted attention to a special class of system level objectives that naturally decouples while this work considers more general system level objective functions. Both approaches to game design guarantee that the resulting game is a state based potential game. State based potential games possess an underlying structure that can be exploited by distributed learning algorithms much like potential games [16], [19].

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 of the core differences 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 of the common approaches to distributed optimization, e.g, subgradient methods [21]–[26], consensus based methods [1], [2], [27], or two-step consensus based approaches [9], [28], is predicated on the structure of the system level objective. There are similarities between our contributions and the algorithmic structure of existing distributed algorithms [25], [28] where an underlying state space is introduced to estimate parameters relevant to the gradients. However, a core difference is that our focus is on the decomposition as opposed to a particular algorithm. Exploiting this decomposition could lead to *rich* set of tools for both game design and learning design that permits a broad class of distributed learning algorithms within an admissible set. For example, if the designed game is of the desired interdependence then an admissible distributed algorithm can be realized by using gradient play on this game. Furthermore, if the designed game is a potential game then this algorithm also guarantees convergence to a Nash equilibrium.

II. Problem Setup and Background

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 possible decisions (or values) denoted by $\mathcal{V}_i$ which is a nonempty convex subset of $\mathbb{R}^{p_i}$, i.e. $\mathcal{V}_i \subseteq \mathbb{R}^{p_i}$.[1] We denote a joint decision by the tuple $(v_1, \cdots, v_n) \in \mathcal{V} := \prod_i \mathcal{V}_i$ where $\mathcal{V}$ is referred to as the set of joint decisions. 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 the form:[2]

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

Throughout the paper we assume that ϕ is continuously differentiable convex and thus a solution is guaranteed to exist.

The focus of this paper is to establish an interaction framework where each decision maker $i \in N$ makes its decision independently in response to local information. The information available to each agent is represented by an undirected and connected communication (or interaction) graph $\mathcal{G} = \{N, \mathcal{E}\}$ with nodes N and edges $\mathcal{E}$.[3] Define the neighbors of agent i as $N_i := \{j \in N : (i,j) \in \mathcal{E}\}$. This interaction framework produces a sequence of decision $v(0), v(1), v(2), \ldots$ where at each iteration $t \in \{0, 1, \ldots\}$ each agent i makes a decision 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 decision v^* that solves the optimization problem in (1). We focus on game theory as a tool for obtaining distributed solutions to the optimization problem (1).

A. Strategic Form Games

A strategic form game consists of a set of players $N \triangleq \{1, 2, \cdots, n\}$ where each player $i \in N$ has an action set $\mathcal{A}_i$ and a cost function $J_i : \mathcal{A} \to \mathbb{R}$ where $\mathcal{A} \triangleq \mathcal{A}_1 \times \ldots \times \mathcal{A}_n$ is referred to as the set of joint action profiles. For an action profile $a = (a_1, \ldots, a_n)$, let a_{-i} denotes the action profile of players other than player i, i.e., $a_{-i} = (a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$. An action profile $a^* \in \mathcal{A}$ is called a *pure Nash equilibrium* if for all $i \in N$, $J_i(a_i^*, a_{-i}^*) = \min_{a_i \in \mathcal{A}_i} J_i(a_i, a_{-i}^*)$.

B. State Based Games

In this paper we consider an extension to strategic form games, termed *state based games* [8], [16], which introduces an underlying state space to the game theoretic framework.[4] In the proposed state based games we focus on myopic players and "static" equilibrium concepts similar to that of pure Nash equilibrium. The state is introduced as a coordinating entity used to improve system level behavior and can take on a variety of interpretations ranging from dynamics for equilibrium selection to the addition of "dummy" players that are preprogrammed to behave in a set fashion.

A state based games consists of a player set N and an underlying state space X. At each state x, each agent $i \in N$ has a state dependent action set $\mathcal{A}_i(x)$ and a state dependent cost function $J_i : X \times \mathcal{A} \to \mathbb{R}$, where $\mathcal{A} \triangleq \prod_i \mathcal{A}_i$ and $\mathcal{A}_i \triangleq \prod_x \mathcal{A}_i(x)$. Lastly, there is a deterministic state transition function $f : X \times \mathcal{A} \to X$. In this paper, we consider continuous state based games in which $\mathcal{A}_i(x)$ is a convex subset of $\mathbb{R}^{p_a}$ for some dimension p_a, X is a subset of R^{p_x} for some dimension p_x, and both $J_i(\cdot)$ and $f(\cdot)$ are continuously differentiable functions.

Repeated play of a state based game produces a sequence of action profiles $a(0)$, $a(1)$, ..., and a sequence of states $x(0)$, $x(1)$, ..., 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 player $i \in N$ myopically

[1] For ease of exposition we let $p_i = 1$ for all $i \in N$. The results in this paper also hold for cases where $p_i > 1$.

[2] Due to the space considerations we focus on optimization problems with decoupled constraints, i.e., $v_i \in \mathcal{V}_i$. The forthcoming methodologies can also incorporate coupled constraints using the approach demonstrated in [20].

[3] By convention, we let $(i,i) \in \mathcal{E}$ for all $i \in N$

[4] State based games can be interpreted as a simplification of Markov games [29]. We avoid formally defining the framework of state based games within the context of Markov games as the inherent complexity of Markov games is unwarranted in our proposed research directions.

selects an action $a_i(t) \in \mathcal{A}_i$ according to some specified decision rule which depends on the current state $x(t)$. The state $x(t)$ and the action profile $a(t) \triangleq (a_1(t), \ldots, a_n(t))$ together determine each player's one-stage cost $J_i(x(t), a(t))$ at time t. After all players select their respective action, the ensuing 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.

We focus on myopic players and static equilibrium concepts similar to that of Nash equilibria. Before defining our notion of equilibria for state based games, we introduce the notion of reachable states. The set of reachable states by an action invariant state trajectory starting from the state action pair $[x^0, a^0]$ is defined as

$$\bar{X}(x^0, a^0; f) \triangleq \{x^0, x^1, x^2, ...\}$$

where $x^{k+1} = f(x^k, a^0)$ for all $k \in \{0, 1, ...\}$. Notice that a fixed action choice a^0 actually defines a state trajectory.

Definition 1. *(Single state equilibrium) A state action pair $[x^*, a^*]$ is called a single state equilibrium if for every agent $i \in N$, we have $J_i(x^*, a^*) = \min_{a_i \in \mathcal{A}_i(x)} J_i(x^*, a_i, a^*_{-i})$.*

Definition 2. *(Recurrent state equilibrium) A state action pair $[x^*, a^*]$ is a recurrent state equilibrium if*

(D-1): $[x, a^*]$ is a single state equilibrium for each state $x \in \bar{X}(x^*, a^*; f)$ and

(D-2): $x^* \in \bar{X}(x^*, a^*; f)$.

Recurrent state equilibria represent fixed points of the Cournot adjustment process for state based games. That is, if a state action pair at time t, i.e., $[x(t), a(t)]$, is a recurrent state equilibrium, then $a(\tau) = a(t)$ for all times $\tau \geq t$ if all players adhere to the Cournot adjustment process. In this paper we focus on state based games where there exists a null action $\mathbf{0} \in \prod_i \mathcal{A}(x)$ for every state $x \in X$ that leaves the state unchanged, i.e., for any state $x \in X$ we have $x = f(x, \mathbf{0})$. The motivation for this structure stems from a control theoretic perspective where an action choice (or control) influences the state of the system. Accordingly, if a state action pair $[x(t), a(t)] = [x^*, \mathbf{0}]$ is a recurrent state equilibrium, then $x(\tau) = x^*$ and $a(\tau) = \mathbf{0}$ for all times $\tau \geq t$ if all players adhere to the Cournot adjustment process.

Given any state based game, a state recurrent equilibrium does not necessarily exist. We now introduce the class of state based potential games for which such an equilibrium is guaranteed to exist [19].

Definition 3. *(State Based Potential Game) A (deterministic) state based game G is a (deterministic) state based potential game if there exists a potential function $\Phi : X \times \mathcal{A} \to \mathbb{R}$ that satisfies the following two properties for every state action pair $[x, a] \in X \times \mathcal{A}$:*

(D-1): For any player $i \in N$ 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): The potential function satisfies $\Phi(x, a) = \Phi(\tilde{x}, \mathbf{0})$ where $\tilde{x} = f(x, a)$.

The first condition states that each agent's cost function is aligned with the potential function in the same fashion as in potential games. The second condition relates to the evolution on the potential function along the state trajectory. As in potential games, a recurrent state equilibrium is guaranteed to exist and there are distributed learning algorithms that converge to recurrent state equilibria in state base potential games [19], [20].

Proposition 1. *Let G be a state based potential game with potential function Φ. If a state action pair $[x^*, a^*]$ satisfies $a^* = \operatorname{argmin}_{a \in \mathcal{A}(x)} \Phi(x^*, a)$, then $[x^*, a^*]$ is a single state equilibrium. Furthermore, if $[x^*, a^*]$ also satisfies $x^* = f(x^*, a^*)$, then $[x^*, a^*]$ is a recurrent state equilibrium.*

III. State Based Game design

In this section we introduce a state based game design for the distributing 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 transitions also rely only on local information.

(ii) The objective function of each agent i is local and of the form $J_i : \prod_{j \in N_i} (X_j \times \mathcal{A}_j) \to \mathbb{R}$

(iii) The resulting game is a state based potential game. The significance of this is the availability of distributed learning algorithm which guarantees convergence to a recurrent state equilibrium.

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

A. A state based game design for distributed optimization

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 player i's estimation for the joint action profile v. The term e_i^k captures player i's estimate of player k's actual value 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 $\hat{e}_i^k := \{\hat{e}_{i\to j}^k\}_{j \in N_i}$ where $\hat{e}_{i\to j}^k \in \mathbb{R}$ represents the estimation value that player i passes to player j regarding to the value of player k.

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 $x(0) = [v(0), e(0)]$. Before specifying the state dynamics we introduce the following notation. Define

$\hat{e}_{i\leftarrow \text{in}}^k := \sum_{j\in N_i} \hat{e}_{j\to i}^k$ and $\hat{e}_{i\to \text{out}}^k := \sum_{j\in N_i} \hat{e}_{i\to j}^k$ 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 $a=(\hat{v},\hat{e})$ we have

$$\begin{aligned} f_i^v(x,a) &= v_i + \hat{v}_i \\ f_{i,k}^e(x,a) &= e_i^k + n\delta_i^k \hat{v}_i + \hat{e}_{i\leftarrow \text{in}}^k - \hat{e}_{i\to \text{out}}^k \end{aligned} \tag{3}$$

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 the available actions to an agent on the current state. That is, the available action set for agent i given state $x=(v,e)$ is defined as

$$\mathcal{A}_i(x) := \{(\hat{v},\hat{e}) : v_i + \hat{v}_i \in \mathcal{V}_i\} \tag{4}$$

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 $k \in N$:

$$\sum_{i=1}^{n} e_i^k(t) = n \cdot v_k(t) \tag{5}$$

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

$$J_i(x,a) = J_i^{\phi}(x,a) + \alpha \cdot J_i^e(x,a) \tag{6}$$

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.[5] We define each of these components as follows:

$$\begin{aligned} J_i^{\phi}(x,a) &= \textstyle\sum_{j\in N_i} \phi(\tilde{e}_j^1, \tilde{e}_j^2, ..., \tilde{e}_j^n) \\ J_i^e(x,a) &= \textstyle\sum_{j\in N_i} \sum_{k\in N} \left[\tilde{e}_i^k - \tilde{e}_j^k\right]^2 \end{aligned} \tag{7}$$

where $\tilde{x} = (\tilde{v}, \tilde{e}) = f(x,a)$ represents the ensuing state. Let $\mathbf{0}$ represent the null action, that is where $\hat{v}_i = 0$ and $\hat{e}_{i\to j}^k = 0$ for all agents $i,j,k \in N$. Given our state dynamics we know that $x = f(x, \mathbf{0})$. Accordingly, our designed cost functions possess the following simplifications: $J_i(x,a) = J_i(\tilde{x}, \mathbf{0})$

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 possesses an underlying structure that guarantees the existence of an equilibrium while facilitating the use of distributed algorithms to reach such equilibria.

Theorem 2. *Model the optimization problem in (1) as a state based game G as depicted in Section III-A with any positive constant α. The state based games is a state based potential game with potential function*

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

where

$$\begin{aligned} \Phi^{\phi}(x,a) &= \textstyle\sum_{i\in N} \phi(\tilde{e}_i^1, \tilde{e}_i^2, ..., \tilde{e}_i^n) \\ \Phi^e(x,a) &= \textstyle\frac{1}{2} \sum_{i\in N} \sum_{j\in N_i} \sum_{k\in N} \left[\tilde{e}_i^k - \tilde{e}_j^k\right]^2 \end{aligned} \tag{9}$$

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

Proof: It is straightforward to verify that the properties of state based potential games in Definition 3 are satisfied using the state based potential function in (8).

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

Theorem 3. *Model the optimization problem in (1) as a state based game G as depicted in Section III-A with any positive constant α. Suppose the undirected and connected communication graph $\mathcal{G} = \{N, \mathcal{E}\}$ satisfies at least **one** of the following conditions*

(i) The communication graph $\mathcal{G}$ is non-bipartite.[6]
(ii) The communication graph $\mathcal{G}$ contains an odd number of nodes, i.e., the number of players is odd;
(iii) The communication graph $\mathcal{G}$ contains at least two players which have a different number of neighbors, i.e., $|N_i| \neq |N_j|$ for some players $i,j \in N$;

Then the state action pair $[x,a] = [(v,e),(\hat{v},\hat{e})]$ is a recurrent state equilibrium in game G 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 the following for all agents $i,k \in N$, $\hat{e}_{i\leftarrow \text{in}}^k = \hat{e}_{i\to \text{out}}^k$.

The above theorem demonstrates 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 agent objective functions.

C. Proof of Theorem 3

It is straightforward to prove the sufficient condition of the theorem by utilizing the fact that the state based game we designed is a state based potential game with potential function defined in (8). Applying Proposition 1, we can conclude that if a state action pair $[x,a]$ satisfies Conditions (a)-(d) listed in the theorem, then $[x,a]$ is a recurrent state equilibrium.

We prove the necessary condition of Theorem 3 by a series of lemmas. Notice that a recurrent state equilibrium is a single state equilibrium by Definition 2. The main part of the proof is to establish necessary conditions for a *single state equilibrium* firstly. We demonstrate that a single state equilibrium should satisfy the following conditions:

1) *Estimation alignment:* An equilibrium must exhibit an alignment between the estimation terms and the value

[5] We will show that as long as α is positive, all 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.

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

profile, i.e., for all agents $i, k \in N$ we have $\tilde{e}_i^k = \tilde{v}_k$ where $(\tilde{v}, \tilde{e})$ is the ensuing state. (Lemma 4 for case (i)–(ii) and Lemma 5 for case (iii).)
2) *Optimality alignment:* An equilibrium must be optimal. That is, the ensuing value profile $\tilde{v}$ is an optimal solution to (1). (Lemma 6 for cases (i)–(iii))

Conclusion the proof completes the proof by establishing more thorough conditions on the resulting recurrent state equilibria.

In the subsequent claims we 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 4. *If $[x, a] = [(v, e), (\hat{v}, \hat{e})]$ is a single state equilibrium and the communication graph $\mathcal{G} = \{N, \mathcal{E}\}$ satisfies either condition (i) or (ii) of Theorem 3, 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 $[x, a]$ is a single state equilibrium then $a_i \in \operatorname{argmin}_{\check{a}_i \triangleq (\check{v}_i, \check{e}_i) \in \mathcal{A}_i(x)} J_i(x, \check{a}_i, a_{-i})$ for all $i \in N$. The necessary condition for the optimality of a_i is that: $\left.\frac{\partial J_i(x, \check{a}_i, a_{-i})}{\check{e}_{i,l}^k}\right|_{a_i} = 0, \forall i, k \in N$, which is equivalent to

$$\phi_k|_{\tilde{e}_i} + 2\alpha \textstyle\sum_{j \in N_i} \left(\tilde{e}_i^k - \tilde{e}_j^k\right) = \phi_k|_{\tilde{e}_l} - 2\alpha \left(\tilde{e}_i^k - \tilde{e}_l^k\right) \tag{10}$$

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 players $i, j \in N$, i.e., $j \in N_i$ and $i \in N_j$. The equality in (10) 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 equality constraints gives us

$$\textstyle\sum_{l \in N_i} (\tilde{e}_i^k - \tilde{e}_l^k) = - \sum_{l \in N_j} (\tilde{e}_j^k - \tilde{e}_l^k) \tag{11}$$

for all agents $i, j, k \in N$ such that $j \in N_i$ and $i \in N_j$. Since our communication graph is connected, the equality condition in (11) 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 player $i \in N$ can be at most one of two possible values that differ purely with respect to sign, i.e., for any player $i \in N$ we have

$$\textstyle\sum_{l \in N_i} (\tilde{e}_i^k - \tilde{e}_l^k) \in \{e_{\text{diff}}^k, -e_{\text{diff}}^k\} \tag{12}$$

where $e_{\text{diff}}^k \in \mathbb{R}$ is a constant. We can utilize the underlying topology of the communication graph coupled with (12) 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 (11), 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 (12) 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 [30].[7] Combining with the equality (5), we get that for all agents $i, k \in N$, $\tilde{e}_i^k = v_k$.

Remark 1. *While we identify two graph structures that lead to our result this is by no means exhaustive as there are alternative graph structures that provide the same guarantees.*

Lemma 5. *Suppose the objective function and communication graph satisfies condition (iii) of Theorem 3. If $[x, a] = [(v, e), (\hat{v}, \hat{e})]$ is a single state equilibrium, 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 $[x, a]$ is a single state equilibrium, then equation (10) should satisfy. Consider any player $i \in N$, any player $j \in N$, and any pair of agents $j_1, j_2 \in N_i$, equation (10) tells us that

$$\begin{aligned} \phi_k|_{\tilde{e}_i} + 2\alpha \textstyle\sum_{j \in N_i} \left(\tilde{e}_i^k - \tilde{e}_j^k\right) &= \phi_k|_{\tilde{e}_{j_1}} - 2\alpha \left(\tilde{e}_i^k - \tilde{e}_{j_1}^k\right) \\ \phi_k|_{\tilde{e}_i} + 2\alpha \textstyle\sum_{j \in N_i} \left(\tilde{e}_i^k - \tilde{e}_j^k\right) &= \phi_k|_{\tilde{e}_{j_2}} - 2\alpha \left(\tilde{e}_i^k - \tilde{e}_{j_2}^k\right). \end{aligned} \tag{13}$$

Combining the two equations, we have the following equality

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

Note that players j_1 and j_2 are not necessarily connected but are rather siblings as both players are connected to player 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

$$\phi_k|_{\tilde{e}_{j_1}} - \phi_k|_{\tilde{e}_{j_2}} = 2\alpha \left(\tilde{e}_{j_2}^k - \tilde{e}_{j_1}^k\right). \tag{14}$$

for all players $k \in N$. Applying Lemma 8 in the appendix, condition (14) coupled with the fact that ϕ is a convex function implies that for any siblings $j_1, j_2 \in N$, $\tilde{e}_{j_1} = \tilde{e}_{j_2}$. This property 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., $\tilde{e}_i \in \{x, y\}, \forall i \in N$, since the communication graph is connected and undirected, Now applying equality (12), 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 players 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 argument as previous proof, we have that $\tilde{e}_i^k = v_k, \forall i, k \in N$.

Lemma 6. *If $[x, a] = [(v, e), (\hat{v}, \hat{e})]$ is a single state equilibrium and the communication graph satisfies any of conditions (i)–(iii) of Theorem 3, then $\tilde{v}$ is an optimal solution to (1).*

Proof: If $[x, a]$ is a single state equilibrium, then we know that a_k is an optimal solution of $\min_{\check{a}_k \in \mathcal{A}_k} J_k(x, \check{a}_k, a_{-k})$, where $\check{a}_k \triangleq (\check{v}_k, \check{e}_k)$. Accordingly, we have $\left.\frac{\partial J_k(x, \check{a}_k, a_{-k})}{\check{v}_k}\right|_{a_k} \cdot (\hat{v}_k' - \hat{v}_k) \geq 0$. which is equivalent to

$$\left[n\, \phi_k|_{\tilde{e}} + 2n_k \sum_{j \in N_k} \left(\tilde{e}_k^k - \tilde{e}_j^k\right) \right] \cdot (\hat{v}_k' - \hat{v}_k) \geq 0 \tag{15}$$

[7]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 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.

We have shown in Lemma 4 and Lemma 5 that if $[x, a] = [(v, e), (\hat{v}, \hat{e})]$ is a single state equilibrium, then $\tilde{e}_i^k = v_k, \forall i, k \in N$. Therefore, equation (15) tells that

$$\phi_k|_{(\tilde{v})} \cdot (\tilde{v}'_k - \tilde{v}_k) \geq 0, \forall \tilde{v}'_k \in \mathcal{V}_k. \tag{16}$$

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

Conclusion the proof Lemma 4-6 has demonstrated that if $[x, a]$ is a single state equilibrium, then the ensuing state $\tilde{x} = (\tilde{v}, \tilde{e}) = f(x, a)$ has accurate estimation $\tilde{e}$ and optimal value $\tilde{v}$. Since a recurrent state equilibrium $[x, a]$ is a single state equilibrium, the ensuing state $\tilde{x} = (\tilde{v}, \tilde{e})$ satisfies the same conditions. Moreover, the action profile a of a recurrent state equilibrium $[x, a]$ should satisfy that $\hat{v} = 0$ and $\hat{e}_{i\leftarrow \text{in}} = \hat{e}_{i\rightarrow \text{out}}$ for all $i \in N$. Otherwise, we can check that $x = (v, e) \notin \bar{X}(x, a; f)$, which violates Condition (2) of Definition 2. Combining those facts we are able to complete the proof of Theorem 3.

IV. Gradient Play

We will develop a distributed learning algorithm for the state based game depicted in section III. The proposed gradient play algorithm extends the convergence results for the algorithm gradient play [7], [31], [32] to state based potential games. In this section, we assume that $\mathcal{V}_i$ is a closed convex set for all $i \in N$. Consider the following algorithm: at each time $t \geq 0$, given the state $x(t) = (v(t), e(t))$, each agent i selects an action $a_i \triangleq (\hat{v}_i, \hat{e}_i)$ according to:

$$\begin{aligned}
\hat{v}_i(t) &= \left[-\epsilon \cdot \left. \frac{\partial J_i\left(x(t), a\right)}{\partial \hat{v}_i} \right|_{a=0} \right]^+ \qquad (17)\\
&= \left[-\epsilon (n\, \phi_i|_{e_i(t)} + 2n\alpha \sum_{j \in N_i} (e_i^i(t) - e_j^i(t))) \right]^+ \\
\hat{e}_{i\rightarrow j}^k(t) &= -\epsilon \cdot \left. \frac{\partial J_i\left(x(t), a\right)}{\partial \hat{e}_{i\rightarrow j}^k} \right|_{a=0} \\
&= \epsilon \cdot \big(\phi_k|_{e_i(t)} - \phi_k|_{e_j(t)} + 2\alpha \left(e_i^k(t) - e_j^k(t)\right) \\
&\quad + 2\alpha \sum_{l \in N_i} \cdot \left(e_i^k(t) - e_l^k(t)\right) \big) \qquad (18)
\end{aligned}$$

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 is the stepsize which is a positive constant. The following theorem establishes the convergences of the gradient play.

Theorem 7. *Suppose each agent selects an action according to the gradient play algorithm in (17,18) at each time $t \geq 0$. If the stepsizes are sufficiently small, and the sequence $x(1), x(2), \cdots$ produced by the algorithm is contained in a compact subset of $\mathbb{R}^{2n}$, then $[x(t), a(t)] := [((v(t), e(t)), a(t))]$ asymptotically converges to the recurrent state equilibrium $[(v^*, \mathbf{v}^*), \mathbf{0}]$.*

Proof: The main idea is to explore the properties of the state based potential function $\Phi(x, a) = \Phi(\tilde{x}, \mathbf{0})$ and show that the potential function keeps decreasing during the gradient play process as long as the stepsize is small enough. Because of space consideration, we omit the detailed proof.

V. Illustrations

For illustration 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:

$$\left\{ v_i \triangleq (v_i^{r_1}, ..., v_i^{r_m}) : 0 \leq v_i^r \leq 1, \forall r \in \mathcal{R}; \sum_{r \in \mathcal{R}} v_i^r = 1 \right\} \tag{19}$$

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: $c_r : [0, +\infty) \rightarrow \mathbb{R}$ that reflects the cost of using the route as a function of the amount of traffic on that route.[8] For a given routing decision $v \in \mathcal{V}$, the total congestion in the network takes on the form $\phi(v) = \sum_{r \in \mathcal{R}} f_r \cdot c_r(f_r)$ 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.

Using the theory 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 players 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 1 illustrates the results of the algorithm proposed in Section IV coupled with our game design in Section III. Note that our algorithm does not perform as well in transient as the true gradient descent algorithm. This is expected since the informational availability to the agents is much lower. However, the convergence time is comparable which is surprising.

VI. Conclusion

We utilize the framework of state based potential games to develop a systematic methodology for the design of local agent objective functions that satisfy virtually any degree of locality while ensuring the optimality of the resulting Nash equilibria. This work, along with previous work, demonstrates the framework of state based potential games leads to a value hierarchical decomposition that can be an extremely

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

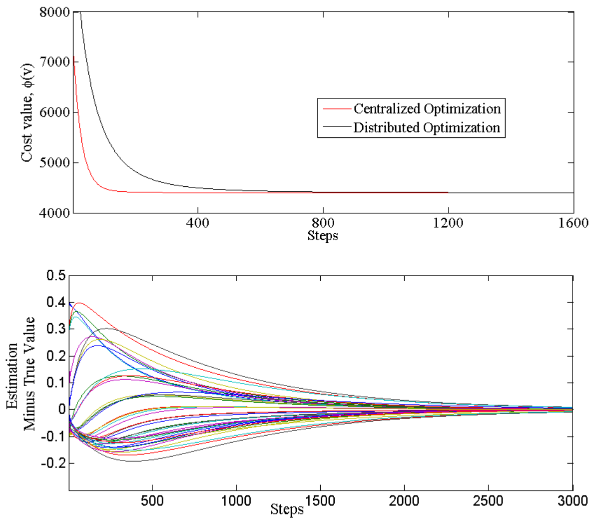

Fig. 1. Simulation results: The upper figure shows the evolution of the system cost using the true gradient decent algorithm (red) and our proposed algorithm (black). The bottom figure shows the evolution of one agent's estimation error, i.e., $e_i^{k,r} - v_k^r$ for each route $r \in \mathcal{R}$ and each agent $k \in N$. Note that the error converges to 0 illustrating that the agent's estimate converge to the right values as proved in Lemmas 4 and 5.

powerful for the design and control of multiagent systems. An important future direction is to enrich the tool set for both game design and learning design in state based potential games. Examples include (i) developing alternative learning algorithms to gradient play and characterizing their convergence rates and (ii) extend the analysis of the approach in this paper to a dynamical changing communication topology.

REFERENCES

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

[2] R. Olfati-saber, J. A. Fax, and R. M. Murray. Consensus and cooperation in networked multi-agent systems. In *Proceedings of the IEEE*, volume 95, January 2007.

[3] 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, October 2007.

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

[5] 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.

[6] G. Scutari, D. P. Palomar, and J. Pang. Flexible design of cognitive radio wireless systems: from game theory to variational inequality theory. *IEEE Signal Processing Magazine*, 26(5):107–123, September 2009.

[7] L. Chen, S. H. Low, and J. C. Doyle. Random access game and medium access control design. *IEEE/ACM Transactions on Networking*, 2010.

[8] J. R. Marden and M. Effros. The price of selfiness in network coding. In *Workshop on Network Coding, Theory, and Applications*, June 2009.

[9] V. Reddy, S. Shakkottai, A. Sprintson, and N. Gautam. Multipath wireless network coding: a population game perspective. In *INFOCOM, 2010 Proceedings IEEE*, pages 1–9. IEEE, 2010.

[10] 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.

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

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

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

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

[15] 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*, 48:373–396, February 2009.

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

[17] L. S. Shapley. A VALUE FOR n-PERSON GAMES1. *Classics in game theory*, page 69, 1997.

[18] E. Anshelevich, A. Dasgupta, J. Kleinberg, E. Tardos, T. Wexler, and T. Roughgarden. The price of stability for network design with fair cost allocation. 2004.

[19] J. R. Marden. State based potential games. *under submission*, 2011.

[20] N. Li and J. R. Marden. Decouple coupled constraints through game design. *Under Submission*, 2011.

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

[22] 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.

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

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

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

[26] 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.

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

[28] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *working paper, Stanford Univ*, 2010.

[29] L. S. Shapley. Stochastic games. In *Proceedings of the National Academy of Science of the United States of America*, volume 39, pages 1095–1100, 1953.

[30] N. Li and J. R. Marden. Designing games to handle coupled constraints. In *49th IEEE Conference on Decision and Control*, pages 250–255, 2010.

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

[32] 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, March 2005.

APPENDIX

Lemma 8. *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

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

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

$$\begin{aligned} 0 &\geq \sum_k \alpha_k(x_k - y_k)^2 \\ &\geq 0 \end{aligned}$$

Therefore $x = y$.