Decoupling Coupled Constraints Through Utility Design

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. In many systems this control design is further complicated by coupled constraints on the agents' behavior, e.g., cooperative robotics or distributed power control. This paper seeks to address the design of such algorithms using the field of game theory. In particular, is it possible to design local agent utility functions for an *unconstrained* game such that all resulting pure Nash equilibria (i) optimize the given system level objective and (ii) satisfy the given coupled constraint. Such developments would greatly simplify the control design by eliminating the need to explicitly consider the constraint. Unfortunately, we illustrate that the standard game theoretic framework is not suitable for this design objective. However, we demonstrate that by adding an additional state variable in the game environment, i.e., moving towards state based games, we can satisfy these performance criteria by utility design. The key to this realization is incorporating classical optimization approaches, in particular the lagrangian penalty and barrier function methods, into the design of agent utility functions.

I. Introduction

Many engineering systems can be characterized as a large scale collection of interacting subsystems where each subsystem makes local independent decisions in response to local information. The central challenge in these multiagent systems is to derive desirable collective behaviors through the design of individual agent control algorithms [1]–[7]. The potential benefits of distributed decision architectures include the opportunity for real-time adaptation and robustness to dynamic uncertainties. These benefits come with significant challenges, such as

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, Boulder, CO 80309, USA jason.marden@colorado.edu. Corresponding author.

the complexity associated with a potentially large number of interacting agents and the analytical difficulties of dealing with overlapping and partial information.

In many systems the desirable collective behavior must satisfy a given coupled constraint on the agents' actions [2]–[7]. One example is the problem of TCP control where the users' sending rates need to satisfy link capacity constraints [4], [8]. An alternative example is the problem of power control in MIMO interference systems where the users' sending rates need to satisfy quality of service constraints [5], [6]. Regardless of the specific application domain, these coupled constraints bring additional complexity to the control algorithm design.

There are two main research directions aimed at designing distributed control algorithms to satisfy performance criteria that involve coupled constraints. The first direction seeks to design algorithms where the coupled constraint is always satisfied, e.g., the well-studied consensus algorithm [1], [3], [9]. While theoretically appealing, such algorithms lack a robustness to environmental uncertainties, noisy measurements, and inconsistent clock rates amongst the agents. The second direction seeks to design algorithms to ensure that *only* the limiting behavior satisfies the coupled constraints, e.g., [4], [7], [10]. While there are efficient algorithms to achieve this, such as dual decomposition [10]–[12], these algorithms often require a two-time scale solution approach by introducing additional pricing terms. Depending on the application domain, this two-time scale approach may be prohibitive either by the informational dependence on the pricing terms or the rigidity of the update algorithm.

Recently, game theory has emerged as a viable design paradigm for distributed control in multiagent systems [2], [5], [7], [13]–[15]. A game theoretic approach to cooperative control involves the following two steps:

(i) ***Game design***: Define the agents as self-interested decision makers in a game theoretic environment.

(ii) ***Agent decision rules***: Define a distributed learning algorithm, i.e., a distributed control law, for each agent to coordinate group behavior.

The goal is to do both steps in conjunction with one another to ensure that the emergent global behavior is desirable with respect to a given system level objective [16], [17].

One of the central appeals of game theory for the design and control of multiagent systems is that it provides a hierarchical decomposition between the distribution of the optimization problem (game design) and the specific distributed control laws (agent decision rules) [15]. For example, if the game is designed as a *potential game* then a system designer can appeal to a wide class of distributed learning algorithms which guarantee convergence to a (pure) Nash equilibrium under a variety of informational dependencies [18]–[23]. Accordingly, 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 [17], [24], [25]. However, these methodologies typically provide no guarantees on the *locality* of the agent utility functions or the efficiency of the resulting Nash equilibria. Furthermore, the theoretical limits of what such approaches can achieve are poorly understood.

In this work we focus on multiagent systems where the desired global behavior possesses an inherent coupled constraint on the actions of the agents. We address the following question: Is it possible to design a game such that (i) all utility functions depend only on local information, (ii) all resulting Nash equilibria satisfy the desired performance criterions, and (iii) the resulting game has an underlying structure that can be exploited by distributed algorithms? In this paper we demonstrate that it is impossible to satisfy these objectives using the framework of strategic form games. However, we demonstrate that it is possible to achieve these objectives using the framework of *state based games* when the system level objective is convex and the constraints are linear in the decision variables. Before going into the details of our contribution we first introduce our model to make our contributions clear.

A. *Problem setup*

We consider a multiagent system consisting of n agents denoted by the set $N \triangleq \{1, \cdots, n\}$. Each agent $i \in N$ is endowed with a set of possible decisions (or values) denoted by $\mathcal{V}_i$ which we assume is a convex subset of $\mathbb{R}^{p_i}$ for some $p_i > 1$, i.e., $\mathcal{V}_i \subset \mathbb{R}^{p_i}$. For ease of exposition we let $p_i = 1$ for all $i \in N$; however, the forthcoming results also hold for cases where $p_i \geq 1$. We denote a joint decision by the tuple $v \triangleq (v_1, ..., v_n) \in \mathcal{V} \triangleq \prod_{i \in N} \mathcal{V}_i$ where $\mathcal{V}$ is referred to as the set of joint decisions. There is a global objective of the form $\phi : \mathcal{V} \to \mathbb{R}$ that a system designer seeks to minimize. We assume throughout that ϕ is differentiable and convex unless otherwise noted. Lastly, we focus on multiagent systems with an inherent coupled constraint represented by a set of linear inequalities $\{g_k(v) \triangleq \sum_i A_i^k v_i - C_k \leq 0\}_{k \in M \triangleq \{1,...,m\}}$. More formally, the optimization problem takes on the form

$$\begin{array}{lll} \min_{v_i \in \mathcal{V}_i, i \in N} & \phi(v_1, v_2, \dots, v_n) & \\ s.t. & g_k(v_1, v_2, \dots, v_n) \leq 0, & k \in M \end{array} \tag{1}$$

We do not explicitly highlight the equality case, $h(v) = 0$, since this can be handled by two inequalities of the form $g_1(v) = h(v) \leq 0$ and $g_2(v) = -h(v) \leq 0$.

The focus of this paper is different from that of direct optimization as our goal is to establish an interaction framework where each decision maker $i \in N$ makes a decision independently in response to "local" information. These informational dependencies are modeled by an undirected and connected communication graph $\mathcal{G} = (N, \mathcal{E})$ with nodes N and edges $\mathcal{E} \subseteq N \times N$ where the neighbor set of agent i is $N_i \triangleq \{j \in N : (i,j) \in \mathcal{E}\}$. By convention, we let $(i,i) \in \mathcal{E}$ for all $i \in N$. Hence, $i \in N_i$ for each player $i \in N$. The interaction framework produces a sequence of decisions $v(0), v(1), v(2), ...$, where at each iteration $t \in \{0, 1, ...\}$ the decision of each agent i is chosen *independently* according to a local control law of the form

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

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 $\{F_i(\cdot)\}_{i \in N}$ within the desired informational constraints such that the collective behavior converges to a joint decision v^* that solves the optimization problem in (1). Note that these informational constraints often limit the availability of established distributed optimization approaches.

Example 1 (Consensus, [1]). *Consider a set of agents $N = \{1, \cdots, n\}$ where each agent $i \in N$ has an initial value $v_i(0) \in \mathbb{R}$. This value could represent a measurement, a physical location, or a belief about a particular event. The goal of the consensus problem is to establish a set of local control laws $\{F_i(\cdot)\}_{i \in N}$ such that each agent's value converges to a consistent value, i.e., $\lim_{t \to \infty} v_i(t) = v^* = \frac{1}{n} \sum_{i \in N} v_i(0)$. Given that the communication graph $\mathcal{G}$ is connected, this can be formalized as the following optimization problem:*

$$\begin{array}{ll} \min_v & \sum_{i \in N, j \in N_i} ||v_i - v_j||_2^2 \\ s.t. & \sum_{i \in N} v_i = \sum_{i \in N} v_i(0). \end{array} \tag{3}$$

Example 2 (Power control, [11]). *Consider a set $N = \{1, \cdots, n\}$ of channels that are transmitted across a given optical fiber by wavelength-multiplexing. Here, each channel is capable* *of choosing it's transmission power level v_i which influences the quality of the transmission of all channels. The goal of the power control problem is to establish a set of local control laws $\{F_i(\cdot)\}_{i \in N}$ such that the power levels converge to levels $v^* = (v_1^*, ..., v_n^*)$ which minimizes the total transmission power subject to constraints on the optical signal to noise ratio (OSNR). See Section VI-B for a formal description.*

B. Our contributions

We focus on game theory as a tool for obtaining distributed solutions to the optimization problem (1). Accordingly, we formulate the optimization problem as a game where the player set is N, the action sets of each player is $\mathcal{V}_i$, and each agent is assigned a cost function of the form $J_i : \mathcal{V} \to \mathbb{R}$.[1] The goal of this paper is to establish a methodology for designing an unconstrained game G, represented by the tuple $G = \{N, \{\mathcal{V}_i\}, \{J_i\}\}$, that achieves the following three objectives:

(O-1): Each agent's cost function relies solely on "local" information and is of the form $J_i : \prod_{j \in N_i} \mathcal{V}_j \to \mathbb{R}$. Alternatively, any player $j \notin N_i$ has no impact on the objective function of agent i. This constraint on agent cost functions directly facilitates the design of distributed learning algorithms of the form (2).

(O-2): All Nash equilibria of the resulting game G solve the optimization problem in (1).

(O-3): The design methodologies lead to a hierarchical decomposition that permits the utilization of a wide class of distributed learning algorithms that guarantees convergence to a Nash equilibrium, e.g., potential games.

There are two main results in this paper. The first result in this paper is an impossibility result. That is, in general it is impossible to satisfy Objectives (O-1)–(O-3) using the framework of strategic form games. We demonstrate this impossibility result on the well studied consensus problem which fits into the class of optimization problems in (1).

Our second result establishes a systematic methodology for game design that satisfies Objectives (O-1)–(O-3) by considering a broader class of games than that of strategic form games. Specifically, we consider the class *state based games* which introduces an underlying state space

[1] We use the terms players and agents interchangeably. Furthermore, we use the term cost functions instead of utility functions as this is the convention for cost minimization systems.

to the game theoretic environment [26]. Now, each player's cost function takes on the local form

$$J_i : \prod_{j \in N_i} (X_j \times \mathcal{V}_j) \to \mathbb{R}. \tag{4}$$

where X_j is a set of local state variables for agent j. These local state variables are used as a mechanism for each agent to estimate constraint value $\{g_k(\cdot)\}$ using only available information. The novelty of our approach stems from integrating classical optimization techniques, in particular exterior penalty methods and barrier function methods, into the design of the cost functions in (4) as shown in Sections IV and V respectively. Both methodologies ensure that all three objectives are satisfied (Theorems 5 and 9). The core difference between the two approaches is that barrier functions can also be used to ensure that the constraint is satisfied dynamically in addition to asymptotically.

Lastly, we illustrate the developments in this paper on the problems of consensus and power control. While the designed game is not a potential game, it is a state based version of potential games which has a similar underlying game structure [26]. Accordingly, we appeal to the algorithm gradient play for which guarantees convergence to a class of equilibria termed recurrent state equilibria in state based potential games [27]. Since our game design methodologies ensure that all recurrent state equilibria solve the optimization problem in (1), our two step design approach achieves our desired objectives. One of the nice properties of state based potential game (or potential games) is that virtually any learning algorithm is guaranteed to converge to a recurrent state equilibrium [26]. Hence, there is an inherent robustness to variations in decision rules caused by delays in information, inconsistent clock rates, or inaccuracies in information.

The focus of this paper is on decoupling (or localizing) *coupled constraints* in multiagent systems. Accordingly, we assume throughout that Objectives (O-1)–(O-3) are achievable when considering the class of optimization problems in (1) without the coupled constraints. That is, we focus on objective functions where each agent can evaluate gradients (or marginal contributions) using only local information. Examples of systems with such objective functions include the aforementioned consensus and power control problems as well as many others, e.g., network coding [14], sensor coverage [28], and network routing [4], [29].

C. Related work

There are many approaches in the existing literature in distributed optimization for decoupling coupled constraints. Ultimately, the ability to implement these algorithms in a local fashion is contingent on the structure of the optimization problem in addition to uncertainty regarding the control environment. For example, consider the consensus problem defined in Example 1. One of the most prevalent distributed optimization algorithms, known as the "consensus algorithm" [1], [3], takes on the form

$$v_i(t) = v_i(t-1) + \epsilon \sum_{j \in N_i} (v_j(t-1) - v_i(t-1)),$$

where $\epsilon > 0$ is the step-size. This algorithm imposes the constraint that for all times $t \geq 0$ we have $\sum_{i \in N} v_i(t) = \sum_{i \in N} v_i(0)$, i.e., the average in invariant. Hence, if the agents reach consensus on a common value, this value must represent the average of the initial values. Notice that keeping this average invariant places a strong condition on the control environment which requires that all agents update synchronously and have perfect information regarding the current values of neighboring agents. In many large scale distributed systems these demands are not practical. Alternative algorithms, such as gossiping algorithms [30], can be used to overcome these challenges; however, the key to successfully implementing such algorithms requires characterizing and accounting for the uncertainty in the control environment.

There are several alternative approaches to distributed optimization ranging from dual decomposition [10]–[12], [31] to subgradient methods [31], [32]. These optimization approaches prescribe the decision making architecture rather than defining a suitable architecture within a given design space. Furthermore, the general theme for solving optimization problems with coupled constraints involves incorporating a two-time scale process where constraints are translated to costs by analyzing the associated dual problem. For example, in [33] the authors introduce a two time scale algorithm for a network coding situation where prices are fixed, players reach an equilibrium, and then prices are updated. This type of algorithm is difficult to implement as convergence rates for reaching an equilibrium are uncharacterized; therefore, when to adapt prices is hard to determine. Alternatively, in [11] the authors analyze a power control problem and specify a precise two step process where players update one period and prices are updated the next period. While the process achieves the desired behavior, there is no inherent robustness to variations on this sequential update process. Furthermore, these papers focus on specific problems where constraints and the associated pricing terms can be computed using only local information. This locality need not be the case in alternative problem settings making these approaches inapplicable. Recently, [32] and [31] seek to add locality as a design feature in the optimization problems focusing on a specific class of optimization problems that differ from the optimization problems considered in this paper. In spirit, these approaches are very similar to the approach presented in this paper as player's respond to estimates of coupled constraint violations which are updated using only local information. However, our approach is less algorithmic and more focused on exploiting the decomposition that game theory provides.

Lastly, games with coupled constraints have been extensively studied in the game theoretic literature for the past fifty years [7], [10], [11], [34]–[37]. A fundamental difference between the research direction in this body of work and our contribution is highlighted by the role of cost functions. In the existing literature both the agent cost functions and the coupled constraints are *inherited* while in our setting the coupled constraints are inherited but the cost functions are *designed*. Accordingly, in these coupled constraint games research focuses on establishing conditions that ensure the existence of a Nash equilibrium [11], [36], [37]. To a lesser extent there has been a degree of work centered on identifying distributed algorithms for finding such equilibria [7], [34], [35]; however, most of these algorithms typically are accompanied by strong implementation restrictions.

II. AN IMPOSSIBILITY RESULT FOR GAME DESIGN

In this section we review the framework of strategic form games. We demonstrate that it is impossible to design local cost functions within this framework that ensures all resulting equilibria satisfy both our system level objective and our desired coupled constraint. We focus on the problem of consensus to illustrate this limitation.

A. Background: 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. Consensus: A game theoretic formulation

We consider a game theoretic model for the consensus problem where the agents are modeled as self-interested decision makers each with the action set $\mathcal{V}_i$. Here, we focus on the design of cost functions $\{J_i\}_{i \in N}$ for the agents which results in desirable equilibria. We focus on the design of cost functions of the form

$$J_i(v_i, v_{-i}) \triangleq F\left(\{v_j, v_j(0)\}_{j \in N_i}\right) \tag{5}$$

meaning that an agent's cost function depends only on information regarding neighboring agents. The function $F(\cdot)$ defines each agent's cost and is invariant to specific indices assigned to agents. Hence, it gives rise to a degree of scalability as the design of such an $F(\cdot)$ leads to a well defined game irrespective of the player set N or the structure of the communication graph, i.e., $\{N_i\}_{i \in N}$. Is it possible to construct a single $F(\cdot)$ such that for any player set N, connected communication graph, and initial values $v(0) \in \mathcal{V}$ the following two conditions are satisfied:

(i) The set of Nash equilibria is nonempty and

(ii) Any Nash equilibrium v^* minimizes $\phi(v^*) = \sum_{i \in N} \sum_{j \in N_i} ||v_i^* - v_j^*||_2^2$ subject to the constraint that $\sum_i v_i^* = \sum_{i \in N} v_i(0)$.

Note that the solution to this optimization problem in (ii) is $v_i^* = (1/n) \sum_j v_j(0)$ for each player $i \in N$ which solves the consensus problem. The following example demonstrates that it is not possible to design cost functions of the form (5) that guarantee our two desired conditions.

Example 3. *Figure 1 highlights the communication graph and the initial values (either A or B) for two different setups for the average consensus problems. For example, in problem (a) we have the following: $N = \{1, 2, 3, 4\}$, $\mathcal{E} = \{\{1, 2\}, \{1, 3\}, \{3, 4\}\}$, $\{v_1(0), v_2(0), v_3(0), v_4(0)\} = \{A, B, A, A\}$. Define a cost function for each agent of the form (5) and use this same $F(\cdot)$ for both situations (a) and (b). Because of the structure of $F(\cdot)$, it is straightforward to show that if $v^{(a)} = (v^*, v^*, v^*, v^*)$ is a Nash equilibrium of (a) then $v^{(b)} = (v^*, v^*, v^*, v^*, v^*)$ is also a Nash equilibrium of (b). The impossibility comes from the fact that v^* cannot represent the average for both situation (a) and (b).*

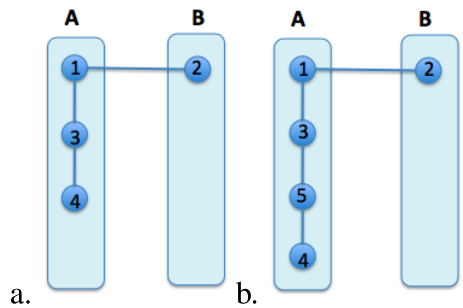

Fig. 1. Communication Graph for Consensus Problem in Example 1

III. STATE BASED GAMES

The previous section demonstrates that the framework of strategic form games is not suitable for handling coupled constraints through the design of agent cost functions. Accordingly, we focus on an extension to the framework of strategic form games, termed *state based games* [26], which introduces an underlying state space to the game theoretic framework.[2] 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 in a strategic form game 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 continuous 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 selects an action $a_i(t) \in \mathcal{A}_i$ according to some specified decision rule which depends on the current state $x(t)$. For example, if a player use a myopic Cournot adjustment

[2]State based games can be viewed as a simplification of Markov games [38]. We avoid formally defining state based games within the context of Markov games as the inherent complexity of Markov games is unwarranted in our research directions.

process then $a_i(t) \in \arg\min_{a_i \in \mathcal{A}_i(x)} J_i(x(t), a_i, a_{-i}(t-1))$. 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. Thus we can mimic the concepts for strategic form games and Markov processes to extend the definition of pure Nash equilibria to state based games.

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 [26].

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 $\Phi : X \times \mathcal{A} \rightarrow \mathbb{R}$ 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 [26], [27]. As to single/recurrent state equilibria in state based potential games, we have the following proposition:

Proposition 1. *Let G be a state based potential game with potential function Φ. If a state action pair $[x^*, a^*]$ satisfies $a^* = \mathrm{argmin}_{a \in \mathcal{A}(x)} \Phi(x^*, a)$, 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.*

IV. A GAME DESIGN USING EXTERIOR PENALTY FUNCTIONS

In this section we demonstrate that it is possible to achieve Objectives (O-1)–(O-3) listed in Section I-B by conditioning each agent's cost function on additional local state variables. Our design utilizes these local state variable as estimates of the constraint functions $\{g_k(\cdot)\}_{k \in M}$. Our first design integrates *exterior penalty functions* into the agents cost functions. Exterior penalty functions are widely used in optimization for transforming a constrained optimization problem into an equivalent unconstrained optimization problem where the constraints are translated to penalties that are directly integrated into the objective functions. More specifically, the constrained optimization problem (1) is transformed to the unconstrained optimization problem[3]

$$\min_{v \in \mathcal{V}} \phi(v) + \mu\alpha(v) \tag{6}$$

where $\alpha(v)$ is a penalty function associated with the constraints $\{g_k(\cdot)\}$ and $\mu > 0$ is a tradeoff parameter. For concreteness we focus on the case when

$$\alpha(v) \triangleq \sum_{k \in M} \left[\max\left(0, \sum_{i \in N} A_i^k v_i - C_k\right)\right]^2. \tag{7}$$

The value of this transformation centers on the fact that solving this unconstrained optimization problem is easier than solving the constrained optimization problem in (1). For each $\mu > 0$ the solution to the unconstrained optimization problem in (6) is actually an infeasible solution to the constrained optimization problem in (1); however, as $\mu \to \infty$, the limit of the solutions to the unconstrained optimization problem in (6) is an optimal solution to the constrained optimization problem in (1). See Theorem 9.2.2 in [39] for more details.

A. *Specifications of 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 $e = \{e_i^k\}_{k \in M, i \in N}$ is the profile of estimation terms where e_i^k is agent i's estimate for the k-th constraint, i.e., $e_i^k \sim \sum_{i=1}^n A_i^k v_i - C_k$.

Actions: Each agent i is assigned an action set $\mathcal{A}_i$ that permits agents to change their value and change their constraint estimate through communication with their neighboring agents. Specifically, an action $a_i \in \mathcal{A}_i$ is defined as a tuple $a_i = (\hat{v}_i, \{\hat{e}_i^1, ..., \hat{e}_i^m\})$ where $\hat{v}_i$ indicates a change in the agent's value and $\hat{e}_i^k$ indicates a change in the agent's estimate of the k-th constraint. Here, the estimation change is a tuple $\hat{e}_i^k = \{\hat{e}_{i \to j}^k\}_{j \in N_i}$ and indicates how the player exchanges estimation values with the neighboring set of agents. The term $\hat{e}_{i \to j}^k$ indicates the estimation value that player i exchanges (or passes) to player j with regards to the k-th constraint.

State Dynamics: For notational simplicity let $\hat{e}_{i \leftarrow \text{in}}^k \triangleq \sum_{j \in N_i} \hat{e}_{j \to i}^k$ and $\hat{e}_{i \to \text{out}}^k \triangleq \sum_{j \in N_i} \hat{e}_{i \to j}^k$ denote the total estimation passed to and from agent i regarding the k-th constraint respectively.

[3]We refer to this problem as unconstrained to differentiate it from the original constrained optimization problem in (6). However, there are uncoupled constraints $v_i \in \mathcal{V}_i \subset \mathbb{R}$ in this problem.

The state transition function $f(x,a)$ takes on the following form: for any state $x=(v,e)$ and an action $a=(\hat{v},\hat{e})$ the ensuing state $\tilde{x}=(\tilde{v},\tilde{e})=f(x,a)$ is

$$\tilde{v}_i = v_i + \hat{v}_i \tag{8}$$

$$\tilde{e}_i = \left\{e_i^k + A_i^k \hat{v}_i + \hat{e}^k_{i\leftarrow \mathrm{in}} - \hat{e}^k_{i\rightarrow \mathrm{out}}\right\}_{k\in M} \tag{9}$$

The specified state transition functions give rise to the following propositions:

Proposition 2. *For any initial state $x(0)=(v(0),e(0))$ and any action trajectory $a(0)$,$a(1)$,..., the resulting state trajectory $x(t)=(v(t),e(t))=f(x(t-1),a(t-1))$ satisfies the following equality for all $t\geq 1$ and constraints $k\in M$: $\sum_{i\in N} A_i^k v_i(t) - \sum_{i\in N} e_i^k(t) = \sum_{i\in N} A_i^k v_i(0) - \sum_{i\in N} e_i^k(0)$.*

Proposition 3. *If the initial estimation $e(0)$ is designed such that $\sum_{i\in N} A_i^k v_i(0) - \sum_{i\in N} e_i^k(0) = C_k$ for all $k\in M$. Then for all $t\geq 0$ and constraints $k\in M$ we have*

$$\sum_{i\in N} e_i^k(t) = \sum_{i\in N} A_i^k v_i(t) - C_k. \tag{10}$$

Therefore we have that for any constraint $k\in M$

$$\sum_{i\in N} e_i^k(t) \leq 0 \Leftrightarrow \sum_{i\in N} A_i^k v_i(t) - C_k \leq 0 \tag{11}$$

Propositions 2 and 3 demonstrate that the estimation terms give direct information to the individual agents as to whether the constraints are violated. We require that the initial estimation terms $e(0)$ satisfies the condition in Proposition 3. Note that there are several ways to define the initial estimation terms to satisfy this equality.

Admissible Action Sets: Since the optimization problem in (1) requires that $v_i \in \mathcal{V}_i$, we define the set of actions available to agent i given the state $x=(v,e)$ as $\mathcal{A}_i(x) \triangleq \{(\hat{v},\hat{e}) : v_i + \hat{v}_i \in \mathcal{V}_i\}$. Since the constraint is purely on the values, we will also use the notation $\mathcal{A}_i^v(x) \triangleq \{\hat{v} : v_i + \hat{v}_i \in \mathcal{V}_i\}$. Given our state dynamics, the null action $\mathbf{0}$ takes on the form

$$\hat{v}_i = 0,\ \forall i \in N$$

$$\hat{e}^k_{i\rightarrow j} = 0,\ \forall i\in N, j\in N_i, k\in M \tag{12}$$

It is straightforward to check that $x = f(x,\mathbf{0})$ for all $x\in X$.

Agent Cost Functions: Before specifying the agent's cost function, we define an agent's marginal contribution to the objective for any value profile $v \in \mathcal{V}$ as

$$M_i(v) \triangleq \phi(v) - \phi(v_i^0, v_{-i}). \tag{13}$$

where $v_i^0 \in \mathcal{V}_i$ is any fixed baseline value for player i. As stated previously, we assume throughout that the marginal contribution $M_i(\cdot)$ can be evaluated using only local information, i.e., for each agent i this marginal contribution is of the functional form $M_i : \prod_{j \in N_i} \mathcal{V}_j \to \mathbb{R}$.[4] We define the state based cost function for each agent $i \in N$ as a summation of two terms relating to the objective ϕ and the constraints $\{g_k\}$ respectively as

$$J_i(x,a) = J_i^{\phi}(x,a) + \mu J_i^{g}(x,a) \tag{14}$$

where $\mu \geq 0$ is a tradeoff parameter. For any state $x \in X$ and admissible action profile $a \in \prod_{i \in N} \mathcal{A}_i(x)$ the components take on the form

$$J_i^{\phi}(x,a) \;=\; M_i(\tilde{v}) \tag{15}$$

$$J_i^{g}(x,a) \;=\; \sum_{j \in N_i} \sum_{k=1}^{m} \left[\max(0, \tilde{e}_j^k)\right]^2 \tag{16}$$

where $\tilde{x} = (\tilde{v}, \tilde{e}) = f(x,a)$ represents the ensuing state. The cost functions satisfy

$$J_i(x,a) = J_i(\tilde{x}, \mathbf{0}) = J_i^{\phi}(\tilde{x}, \mathbf{0}) + \mu J_i^{g}(\tilde{x}, \mathbf{0}) \tag{17}$$

where $\mathbf{0}$ is the null action defined in (12). We utilize this property in the forthcoming proofs.

B. Analytical properties of designed game

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

Theorem 4. *Model the constrained optimization problem in (1) as a state based game G with a fixed trade-off parameter $\mu > 0$ as depicted in Section IV-A. The state based game is a state based potential game with potential function*

$$\Phi(x,a) = \phi(\tilde{v}) + \mu \sum_{i \in N} \sum_{k=1}^{m} \left[\max(0, \tilde{e}_i^k)\right]^2 \tag{18}$$

[4] This marginal contribution is in fact localized in a variety of applications including sensor coverage [28], network coding [14], network routing [4], [29], consensus and rendezvous [1] among many others.

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

Proof: Consider any two state action pairs of the form $[x, a]$ and $[x, a']$ where $a' = (a'_i, a_{-i})$ for some agent $i \in N$ with action $a'_i \neq a_i$. Denote $\tilde{x} = (\tilde{v}, \tilde{e}) = f(x, a)$ and $\bar{x} = (\bar{v}, \bar{e}) = f(x, a')$. Focusing on $J_i^\phi(\cdot)$ it is straightforward to verify

$$
\begin{aligned}
J_i^\phi(x, a'_i, a_{-i}) - J_i^\phi(x, a) &= M_i(\bar{v}) - M_i(\tilde{v}) \\
&= \phi(\bar{v}) - \phi(\tilde{v}).
\end{aligned}
$$

Focusing on $J_i^g(\cdot)$ we have

$$
\begin{aligned}
J_i^g(x, a'_i, a_{-i}) - J_i^g(x, a) &= \mu \sum_{j \in N_i} \sum_{k=1}^{m} [\max(0, \bar{e}_j^k)]^2 - \mu \sum_{j \in N_i} \sum_{k=1}^{m} [\max(0, \tilde{e}_j^k)]^2 \\
&= \mu \sum_{i \in N} \sum_{k=1}^{m} [\max(0, \bar{e}_i^k)]^2 - \mu \sum_{i \in N} \sum_{k=1}^{m} [\max(0, \tilde{e}_i^k)]^2
\end{aligned}
$$

where the second equality comes from the fact that any agent $j \notin N_i$ is not impacted by the change from a_i to a'_i. Hence, the game satisfies the first condition of Definition 3 with the potential function as in (18). The second condition of Definition 3 is satisfied directly from the definition of the Φ which ensures $\Phi(x, a) = \Phi(\tilde{x}, \mathbf{0})$. This completes the proof. ■

The next theorem demonstrates the equivalence between the solutions of the unconstrained optimization problem in (6) and the recurrent state equilibria of the designed game. Therefore, as $\mu \to \infty$ all equilibria of our designed game are solutions to the constrained optimization problem in (1). Note that this equivalence is surprising given the fact that we impose a restriction on the underlying decision making architecture, i.e., distributed instead of centralized. Recall that our only requirement on the communication graph $\mathcal{G}$ is that $\mathcal{G}$ is undirected and connected.

Theorem 5. *Model the constrained optimization problem in (1) as a state based game G with a fixed trade-off parameter $\mu > 0$ as depicted in Section IV-A. 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 four conditions are satisfied:*

(i) The estimation profile e satisfies $\max\left(0, e_i^k\right) = \frac{1}{n} \max\left(0, \sum_{i \in N} A_i^k v_i - C_k\right)$ for all players $i \in N$ and constraints $k \in M$.

(ii) The value profile v is an optimal point of the unconstrained optimization problem $\min_{v \in \mathcal{V}} \phi(v) + \frac{\mu}{n} \alpha(v)$ where $\alpha(v)$ is defined in (7).

(iii) The change in value profile satisfies $\hat{v}_i = 0$ for all players $i \in N$.

(iv) The change in estimation profile satisfies $\hat{e}^k_{i\leftarrow \text{in}} = \hat{e}^k_{i\to \text{out}}$ for all players $i \in N$ and constraints $k \in M$.

Proof: In the following proof we express the ensuing state for a state action pair $[x, a] = [(v, e), (\hat{v}, \hat{e})]$ as $\tilde{x} = (\tilde{v}, \tilde{e})$, i.e. $\tilde{x} \triangleq f(x, a)$.

($\Leftarrow$) We start by proving that if a state action pair $[x, a]$ satisfies conditions (i)–(iv) then $[x, a]$ is a recurrent state equilibrium. The proof essentially follows Proposition 1 which tells us that if $a \in \text{argmin}_{\check{a}\in\mathcal{A}(x)}\Phi(x, \check{a})$ and $x = f(x, a)$, then $[x, a]$ is a recurrent state equilibrium. It is straightforward to show that if a satisfies condition (iii)–(iv) then $x = f(x, a)$. Hence, the proof requires showing that if a state action pair $[x, a]$ satisfies condition (i)–(iv) then $a \in \text{argmin}_{\check{a}\in\mathcal{A}(x)}\Phi(x, \check{a})$. Let $\check{a} \triangleq (\check{v}, \check{e}) \in \mathcal{A}(x)$. Since $\Phi(x, \check{a}) = \Phi(x, (\check{v}, \check{e}))$ is convex over $(\check{v}, \check{e})$, the necessary and sufficient conditions for $a = (\hat{v}, \hat{e})$ to be an optimal solution of the optimization problem $\min_{\check{a}\in\mathcal{A}(x)} \Phi(x, \check{a})$ are:

$$
\begin{aligned}
\left.\frac{\partial\Phi(x,\check{a})}{\partial \check{e}^k_{i\to j}}\right|_a &= 0, \qquad \forall i \in N,\ j\in N_i,\ k\in M\\
\left.\frac{\partial\Phi(x,\check{a})}{\partial \check{v}_i}\right|_a \cdot (\hat{v}'_i - \hat{v}_i) &\geq 0, \qquad \forall i\in N,\ \hat{v}'_i \in \mathcal{A}^v_i(x)
\end{aligned}
$$

Since $\Phi(x, a) = \Phi(\tilde{x}, \mathbf{0})$ these conditions simplify to

$$
\max(0, \tilde{e}^k_j) - \max(0, \tilde{e}^k_i) = 0, \qquad \forall i\in N,\ j\in N_i,\ k\in M, \tag{19}
$$

$$
\left[\left.\frac{\partial\phi}{\partial v_i}\right|_{\tilde{v}} + 2\mu\sum_{k\in M} A^k_i \max(0, \tilde{e}^k_i)\right]\cdot(\tilde{v}'_i - \tilde{v}_i) \geq 0, \qquad \forall i\in N,\ \tilde{v}'_i\in\mathcal{V}_i. \tag{20}
$$

To complete the proof of this direction we actually prove the following stronger statement: if the ensuing state $\tilde{x}$ of a state action pair $[x, a]$ satisfies conditions (i)–(ii), then $a \in \text{argmin}_{\check{a}\in\mathcal{A}(x)}\Phi(x, \check{a})$. This statement also implies that $[x, a]$ is a single state equilibrium by Proposition 1. If the ensuing state $\tilde{x}$ of a state action pair $[x, a]$ satisfies conditions (i)-(ii), it is straightforward to show that $\tilde{x}$ satisfies the following conditions:

$$
\max(0, \tilde{e}^k_i) = \max(0, \tilde{e}^k_j) = \tfrac{1}{n}\max\left(0, \textstyle\sum_{i=1}^n A^k_i\tilde{v}_i - C_k\right), \qquad \forall i, j\in N,\ k\in M \tag{21}
$$

$$
\left[\left.\tfrac{\partial\phi}{\partial v_i}\right|_{\tilde{v}} + \tfrac{2\mu}{n}\textstyle\sum_{k\in M} A^k_i \max\left(0, \sum_{i=1}^n A^k_i\tilde{v}_i - C_k\right)\right]\cdot(\tilde{v}'_i - \tilde{v}_i) \geq 0, \quad \forall i\in N,\ \tilde{v}'_i\in\mathcal{V}_i, \tag{22}
$$

Equation (21) is from Condition (i) and Proposition 3. Equation (22) is the optimal condition of the optimization problem $\min_{v\in\mathcal{V}} \phi(v) + \frac{\mu}{n}\alpha(v)$. Substituting (21) into (22) proves that $\tilde{x}$ satisfies the two optimality conditions in (19) and (20). Hence, $a \in \text{argmin}_{\check{a}\in\mathcal{A}(x)}\Phi(x,\check{a})$. Therefore, we can conclude that such $[x,a]$ is a recurrent state equilibrium.

($\Rightarrow$) Now we prove the other direction of Theorem 5. First notice that if $[x,a]$ is a recurrent state equilibrium then the action profile $a=(\hat{v},\hat{e})$ must satisfy conditions (iii)–(iv). Otherwise, $x=(v,e)\notin \bar{X}(x,a;f)$ which violates the second condition of Definition 2. Moreover, Conditions (iii)–(iv) give us that $x=\tilde{x}=f(x,a)$. To complete the proof of this direction we actually prove the following stronger statement: if a state action pair $[x,a]$ is a single state equilibrium then the ensuing state $\tilde{x}=(\tilde{v},\tilde{e})$ satisfies condition (i)–(ii) listed in the theorem. Recall that the set of recurrent state equilibria is a subset of the set of single state equilibria (Definition 2).

If a state action pair $[x,a]$ is a single state equilibrium, then for each $i\in N$, we have $J_i(x,a_i,a_{-i}) = \min_{\check{a}_i\in\mathcal{A}_i(x)} J_i(x,\check{a}_i,a_{-i})$. Since $J_i(x,\check{a}_i,a_{-i})$ is a convex function on $\check{a}_i \triangleq (\check{v}_i,\check{e}_i)\in\mathcal{A}_i(x)$, we know that

$$\begin{aligned}\left.\frac{\partial J_i(x,\check{a}_i,a_{-i})}{\partial \check{e}_i}\right|_a &= 0, \quad \forall i\in N, k\in M &(23)\\ \left[\left.\frac{\partial J_i(x,\check{a}_i,a_{-i})}{\partial \check{v}_i}\right|_a\right]\cdot(\hat{v}_i'-\hat{v}_i) &\geq 0, \quad \forall i\in N, \hat{v}_i'\in\mathcal{A}_i^v(x) &(24)\end{aligned}$$

which is equivalent to

$$\begin{aligned}2\mu\left(\max\left(0,\tilde{e}_l^k\right)-\max\left(0,\tilde{e}_i^k\right)\right) &= 0, \quad \forall i\in N, j\in N_i, k\in M &(25)\\ \left[\left.\frac{\partial M_i}{\partial v_i}\right|_{\tilde{v}} + 2\mu\sum_{k\in M}A_i^k\max(0,\tilde{e}_i^k)\right]\cdot(\tilde{v}_i'-\tilde{v}_i) &\geq 0, \quad \forall i\in N,\ \tilde{v}_i'\in\mathcal{V}_i. &(26)\end{aligned}$$

Equation (25) implies that $\max\left(0,\tilde{e}_i^k\right)=\max\left(0,\tilde{e}_j^k\right)$ for all agents $i,j\in N$ and constraints $k\in M$ since the communication graph $\mathcal{G}$ is connected. Applying Proposition 3, we have that for all players $i\in N$ and constraints $k\in M$, $\max\left(0,\tilde{e}_i^k\right)=\frac{1}{n}\max\left(0,\sum_{i=1}^n A_i^k\tilde{v}_i - C_k\right)$. Substituting this equality into equation (26), and applying the equality $\left.\frac{\partial M_i}{\partial v_i}\right|_{\tilde{v}} = \left.\frac{\partial \phi}{\partial v_i}\right|_{\tilde{v}}$ gives us

$$\left(\left.\frac{\partial \phi}{\partial v_i}\right|_{\tilde{v}} + \frac{2\mu}{n}\sum_{k\in M}A_i^k\max\left(0,\sum_{i=1}^n A_i^k\tilde{v}_i - C_k\right)\right)\cdot(\tilde{v}_i'-\tilde{v}_i)\geq 0 \tag{27}$$

for all $\tilde{v}_i'\in\mathcal{V}_i$. Hence, $\tilde{v}$ the optimal solution to $\phi(v)+\frac{\mu}{n}\alpha(v)$. This completes the proof. ■

C. A special case involving pairwise decomposable objectives

In the previous section we identified an equivalence between the recurrent state equilibria of our designed games and the solutions to an unconstrained (or simple) optimization problem parameterized by a tradeoff parameter μ. Ensuring that the recurrent state equilibria were optimal points of the original constrained optimization problem in (1) requires $\mu \to \infty$. In this section we discuss special classes of optimization problems where achieving optimality of the original constrained optimization problem in (1) does not require limiting cases of μ.

For a differentiable convex objective function ϕ the necessary and sufficient conditions for $v^* \in \mathcal{V}$ to be an optimal solution to the constrained optimization problem in (1) are the KKT conditions. That is, there exists a unique $\lambda \in \mathbb{R}^m$, termed the Lagrange multipliers, such that

$$\begin{array}{lccr} \text{(KKT-1):} & \left(\frac{\partial \phi}{\partial v_i}\Big|_{v^*} + \sum_{k=1}^m \lambda_k A_i^k \right)(v_i - v_i^*) \geq 0, & \forall i \in N,\ v_i \in \mathcal{V}_i, & \\ \text{(KKT-2):} & \lambda_k \geq 0, & \forall k \in M, & \\ \text{(KKT-3):} & \lambda_k \left(\sum_{i=1}^n A_i^k v_i^* - C_k \right) = 0, & \forall k \in M, & \\ \text{(KKT-4):} & \sum_{i=1}^n A_i^k v_i^* \leq C_k, & \forall k \in M. & \end{array} \tag{28}$$

In some cases Conditions (KKT-1)–(KKT-2) can determine a unique zero λ. This implies that the optimal value profile satisfies $\left(\frac{\partial \phi}{\partial v_i}\Big|_{v^*} \right)(v_i - v_i^*) \geq 0$ for all $v_i \in \mathcal{V}_i$. Hence, the optimal profile v^* is also a solution to the unconstrained optimization problem $\min_{v \in \mathcal{V}} \phi(v)$. The following proposition identifies one such class of optimization problems.

Proposition 6. *Suppose the optimization problem in (1) satisfies the following conditions:*

(a) The objective function ϕ is convex, differentiable, and is of the form

$$\phi(v) = \sum_{i \in N} \sum_{j \in N_i} d_{ij}(v_j - v_i) \tag{29}$$

where $d_{ij}(v_j - v_i) = d_{ji}(v_i - v_j)$ for any two agent $i, j \in N$,

(b) The value profiles for each agent $i \in N$ is $\mathcal{V}_i = \mathbb{R}$,

(c) The constraints coefficients $\{\sum_{i \in N} A_i^k\}_{k \in M}$ are either all positive or all negative.

Then Conditions (KKT-1)–(KKT-2) determine a unique zero Lagrange multiplier λ.

Proof: By KKT conditions, there exists a Lagrange multiplier $\lambda_k \geq 0$ for each constraint $k \in M$ such that $\frac{\partial \phi}{\partial v_i} + \sum_{k=1}^m \lambda_k A_i^k = 0$ for each $i \in N$. Summing up these n equations gives us

$$\sum_{i=1}^n \frac{\partial \phi}{\partial v_i} + \sum_{k=1}^m \lambda_k \left(\sum_{i=1}^n A_i^k \right) = 0.$$

Focusing on the first set of terms in the above summation gives us

$$\begin{aligned}
\sum_{i=1}^{n} \frac{\partial \phi}{\partial v_i} &= \sum_{i=1}^{n} \sum_{j \in N_i} \frac{\partial d_{ij}(v_j - v_i)}{\partial v_i} + \sum_{i=1}^{n} \sum_{j \in N_i} \frac{\partial d_{ji}(v_i - v_j)}{\partial v_i} \\
&= \sum_{i=1}^{n} \sum_{j \in N_i} \frac{\partial d_{ij}(v_j - v_i)}{\partial v_i} + \sum_{j=1}^{n} \sum_{i \in N_j} \frac{\partial d_{ji}(v_i - v_j)}{\partial v_i} \\
&= \sum_{i=1}^{n} \sum_{j \in N_i} \frac{\partial d_{ij}(v_j - v_i)}{\partial v_i} + \sum_{i=1}^{n} \sum_{j \in N_i} \frac{\partial d_{ij}(v_j - v_i)}{\partial v_j} \\
&= \sum_{i=1}^{n} \sum_{j \in N_i} \left(- \left.\frac{\partial d_{ij}(\delta)}{\partial \delta}\right|_{\delta = v_j - v_i} + \left.\frac{\partial d_{ij}(\delta)}{\partial \delta}\right|_{\delta = v_j - v_i} \right) \\
&= 0
\end{aligned}$$

The second equality stems from the fact that the communication graph is undirected. The third equality comes from interchanging the subscripts i and j. Therefore, for any decomposable ϕ in the form of (29), we have $\sum_{k=1}^{m} \lambda_k (\sum_{i=1}^{n} A_i^k) = 0$. This directly implies that $\lambda_k = 0$ for all $k \in M$, since $\sum_{i=1}^{n} A_i^k$ are either all positive or all negative and $\lambda_k \geq 0$ for all $k \in M$. ■

The above proposition identifies a broad class of pairwise decomposable objective functions where there is an equivalence between the solutions of the constrained and unconstrained optimization problems. For such objective functions we can provide stronger analytical guarantees on the recurrent state equilibria as depicted in the following theorem.

Theorem 7. *Model the constrained optimization problem in (1) as a state based game G with a fixed trade-off parameter $\mu > 0$ as depicted in Section IV-A. Suppose Conditions (KKT-1)-(KKT-2) result in a zero Lagrange multiplier vector λ, i.e. $\lambda_k = 0$ for all $k \in M$. 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 four conditions are satisfied:*

(i) The estimation profile e satisfies $e_i^k \leq 0$ for all players $i \in N$ and constraints $k \in M$.

(ii) The value profile v is an optimal point of the constrained optimization problem (1).

(iii) The change in value profile satisfies $\hat{v}_i = 0$ for all players $i \in N$.

(iv) The change in estimation profile satisfies the following for all players $i \in N$ and constraints $k \in M$, $\hat{e}_{i \leftarrow \text{in}}^k = \hat{e}_{i \rightarrow \text{out}}^k$.

Proof: For notational simplicity we denote the ensuing state of a state action pair as $[x, a]$ as $\tilde{x} = (\tilde{v}, \tilde{e}) \triangleq f(x, a)$.

($\Leftarrow$) We start by proving that if a state action pair satisfies Conditions (i)–(iv), then it is a recurrent state equilibrium. In the proof of Theorem 5, we prove that $a = \mathrm{argmin}_{\check{a}\in\mathcal{A}(x)}\Phi(x,\check{a})$ if and only if equations (19, 20) are satisfied. If $\tilde{v}$ satisfies Condition (ii) of the theorem, then it satisfies the four KKT conditions with a zero Lagrange multiplier λ. This implies that $\tilde{v}$ satisfies

$$\textstyle\sum_i^n A_i^k \tilde{v}_i \le C_k, \qquad \forall k \in M \tag{30}$$

$$\left[\frac{\partial\phi}{\partial v_i}\bigg|_{\tilde{v}}\right](\tilde{v}_i' - \tilde{v}_i) \ge 0, \quad \forall i \in N, \tilde{v}_i' \in \mathcal{V}_i \tag{31}$$

If $\tilde{e}$ satisfies Condition (i) of the theorem, combining with equation (30) we know that for any agent $i \in N$ and constraint $k \in M$

$$\max\left(0, \tilde{e}_i^k\right) = \frac{1}{n}\max\left(0, \sum_i^n A_i^k \tilde{v}_i - C_k\right) = 0. \tag{32}$$

Hence, (19) satisfied. Substituting (32) into (20), we know that (20) is also satisfied based on equation (31). Thus we prove that such $[x, a]$ satisfies $a = \mathrm{argmin}_{\check{a}\in\mathcal{A}(x)}\Phi(x,\check{a})$. The rest of the proof follows directly after the proof of Theorem 5.

($\Rightarrow$) We next show that if $[x, a]$ is a recurrent state equilibrium then Conditions (i)–(iv) of the theorem are satisfied. As shown in the proof for Theorem 5, if $[x, a]$ is a single state equilibrium then (27) is satisfied. Define $\eta_k \triangleq \frac{2\mu}{n}\max\left(0, \sum_{i=1}^n A_i^k \tilde{v}_i - C_k\right)$. We have:

$$\begin{array}{lccc} (i) & \left(\frac{\partial\phi}{\partial v_i}\Big|_{\tilde{v}} + \sum_{k=1}^m \eta_k A_i^k\right)(\tilde{v}_i' - \tilde{v}_i) \ge 0 & \forall i \in N,\ \tilde{v}_i' \in \mathcal{V}_i, \\ (ii) & \eta_k \ge 0 & \forall k \in M, \end{array} \tag{33}$$

These are precisely the first two KKT conditions. Hence, we know that for each constraint $k \in M$ we have $\eta_k = 0$ which implies that $\sum_{i=1}^n A_i^k \tilde{v}_i \le C_k$. This is equivalent to $\tilde{e}_i^k \le 0$. Hence, Condition (i) of the theorem is satisfied. Condition (ii) of the theorem is also satisfied since $\tilde{v}$ and η satisfy the KKT conditions (28). Conditions (iii)–(iv) are satisfied directly by Theorem 5. This also gives us that the ensuing state $\tilde{x} = (\tilde{v}, \tilde{e}) = f(x, a)$ satisfies $\tilde{v} = v$ and $\tilde{e} = e$. This completes the proof. ■

When the constraint in the optimization problem (1) is only an equality constraint $\sum_i A_i v_i = c$, it seems that the result in the previous proposition does not hold since if we write the constraint in terms of two inequality constraints $\sum_i A_i v_i \le c$ and $\sum_i (-A_i) v_i \le -c$, Condition (c) in the Proposition 6 is not satisfied. However, if we follow the same argument, we can get a similar result, which is stated in the following proposition.

Proposition 8. *Model the constrained optimization problem in (1) as a state based game G with a fixed trade-off parameter $\mu > 0$ as depicted in Section IV-A. Suppose that the objective function satisfies Condtion (a) in Proposition 6 and the constraints $\{g_k(\cdot)\}_{k=1,2}$ represent a single equality constraint, i.e., $\sum_{i\in N} A_i v_i = c$, where $\sum_{i\in N} A_i \neq 0$. 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 four conditions are satisfied:*

(i) The estimation profile e satisfies $e_i^k = 0$ for all players $i \in N$ and constraints $k = 1, 2$.

(ii) The value profile v is an optimal point of the constrained optimization problem (1).

(iii) The change in value profile satisfies $\hat{v}_i = 0$ for all players $i \in N$..

(iv) The change in estimation profile satisfies the following for all players $i \in N$ and constraints $k = 1, 2$, $\hat{e}^k_{i\leftarrow \text{in}} = \hat{e}^k_{i\to\text{out}}$.

Proof: The proof is similar to the proof of Theorem 7. Similarly we denote the ensuing state of a state action pair as $[x,a]$ as $\tilde{x} = (\tilde{v},\tilde{e}) \triangleq f(x,a)$.

($\Leftarrow$) We start by proving that if Conditions (i)–(iv) of the Proposition are satisfied then the state action pair $[x,a]$ is a recurrent state equilibrium. In the proof of Theorem 5, we prove that $a = \text{argmin}_{\check{a}\in\mathcal{A}(x)}\Phi(x,\check{a})$ if and only if equations (19, 20) are satisfied. Since here we have $\mathcal{V}_i = \mathbb{R}$, equation (20) is equivalent to

$$\left.\frac{\partial \phi}{\partial v_i}\right|_{\tilde{v}} + \frac{2\mu}{n}\sum_{k\in M} A_i^k \max(0, \sum_{i=1}^{n} A_i^k \tilde{v}_i - C_k) = 0, \forall i \in N. \tag{34}$$

If $\tilde{v}$ satisfies Condition (ii) of the proposition, then it satisfies the four KKT conditions. This implies that $\tilde{v}$ satisfies:

$$\begin{aligned} \textstyle\sum_i^n A_i\tilde{v}_i &= c \\ \left.\tfrac{\partial \phi}{\partial v_i}\right|_{\tilde{v}} + \lambda_1 A_i - \lambda_2 A_i &= 0, \forall i \in N. \end{aligned} \tag{35}$$

where λ_1 and λ_2 are the Lagrange multipliers corresponding to the two constraints $\sum_{i\in N} A_i\tilde{v}_i \le c$ and $\sum_{i\in N}(-A_i)\tilde{v}_i \le -c$ respectively. Since the objective function is pairwise decomposable, we know from the proof of Proposition 6 that $\sum_{i\in N} \left.\frac{\partial\phi}{\partial v_i}\right|_{\tilde{v}} = 0$. Thus we have $(\lambda_1 - \lambda_2)\left(\sum_{i\in N} A_i\right) = 0$, which tells us that $\lambda_1 = \lambda_2$ since $\sum_{i\in N} A_i \neq 0$. As a result, we have

$$\left.\frac{\partial \phi}{\partial v_i}\right|_{\tilde{v}} = 0, \forall i \in N. \tag{36}$$

If $\tilde{e}$ satisfies Condition (i) of the proposition, combining with the first equation in (35), we have

$$\max(0, \tilde{e}_i^k) = \frac{1}{n} \max\left(0, \sum_i^n A_i \tilde{v}_i - c\right) = 0, k = 1, 2. \tag{37}$$

Therefore, equation (19) is satisfied. Combining with second equation of (35), we have equation (34) satisfied based on equation (36). The rest of the proof follows directly after the proof of Theorem 5.

($\Rightarrow$) We next show that if $[x, a]$ is a recurrent state equilibrium then Conditions (i)–(iv) of the proposition are satisfied. As shown in the proof of Theorem 7, if $[x, a]$ is a recurrent state equilibrium then (27) must be satisfied. Therefore, representing the equality constraint by two inequality constraints gives us

$$\left.\frac{\partial \phi}{\partial v_i}\right|_{\tilde{v}} + A_i \left(\frac{2\mu}{n} \max(0, \sum_{i \in N} A_i \tilde{v}_i - c)\right) - A_i \left(\frac{2\mu}{n} \max(0, \sum_{i \in N} (-A_i \tilde{v}_i) + c)\right) = 0 \tag{38}$$

for all agents $i \in N$. Define $\lambda_1 \triangleq \frac{2\mu}{n} \max\left(0, \sum_{i \in N} A_i \tilde{v}_i - c\right)$ and $\lambda_2 \triangleq \frac{2\mu}{n} \max\left(0, \sum_{i \in N} (-A_i \tilde{v}_i) + c\right)$ and substituting into (38) we have for all agents $i \in N$

$$\left.\frac{\partial \phi}{\partial v_i}\right|_{\tilde{v}} + (\lambda_1 - \lambda_2) A_i = 0.$$

As we shown before, we know we have $\lambda_1 = \lambda_2$ which gives us that $\max\left(0, \sum_{i \in N} A_i \tilde{v}_i - c\right) = \max\left(0, -\left(\sum_{i \in N} A_i \tilde{v}_i - c\right)\right)$. Thus the constraint is satisfied, i.e., Condition (i) of the proposition is satisfied, as $\sum_{i \in N} A_i \tilde{v}_i = c$. Substituting this equality into (38) gives us that for any agent $i \in N$, we have $\left.\frac{\partial \phi}{\partial v_i}\right|_{\tilde{v}} = 0$. Therefore, $\tilde{v}$ is an optimal solution of the constrained optimization problem (1), i.e., Condition (ii) of the proposition is satisfied. The rest of the proof is exactly the same as the proof of Theorem 7. ■

D. Gradient play for penalty function methods

In this section we discuss the learning algorithm *gradient play* which is guaranteed to converge to a recurrent state equilibrium in state based potential games where the potential function is continuous, differentiable, and convex [27]. The state based game design using penalty functions depicted in Section IV-A is a state potential game with a potential function that meets the above specifications (18). Accordingly, we can implement the algorithm gradient play to reach a recurrent state equilibrium provided that $\mathcal{V}_i$ is a closed convex set for all $i \in N$. The algorithm gradient play can be described as follows: At each time $t \geq 0$ each agent i selects an action $a_i(t) \triangleq (\hat{v}_i(t), \hat{e}_i(t))$ given the state $x(t) = (v(t), e(t))$ according to:

$$
\begin{aligned}
\hat{v}_i(t) &= \left[-\epsilon \cdot \frac{\partial J_i(x(t), a)}{\partial \hat{v}_i} \bigg|_{a=\mathbf{0}} \right]^+ \\
&= \left[-\epsilon \cdot \left(\frac{\partial \phi(v)}{\partial v_i} \bigg|_{v=v(t)} + 2\mu \sum_{k \in M} A_i^k \max\left(0, e_i^k(t)\right) \right) \right]^+ \\
\hat{e}_{i \to j}^k(t) &= -\epsilon \cdot \frac{\partial J_i(x(t), a)}{\partial \hat{e}_{i \to j}^k} \bigg|_{a=\mathbf{0}} \\
&= \epsilon \cdot 2\mu \left(\max\left(0, e_i^k(t)\right) - \max\left(0, e_j^k(t)\right) \right)
\end{aligned} \tag{39}
$$

where $[\cdot]^+$ represents the projection onto the closed convex set $\mathcal{A}_i^{\hat{v}}(x(t)) \triangleq \{\hat{v}_i : v_i(t) + \hat{v}_i \in \mathcal{V}_i\}$ and ϵ is the step size which is a positive constant. If the step size is sufficiently small then the state action pair $[x(t), a(t)]$ will asymptotically converge to a recurrent state equilibrium $[x(t), \mathbf{0}]$.

V. A GAME DESIGN USING BARRIER FUNCTIONS

In this section we introduce our second design which integrates *barrier functions* as opposed to exterior penalty functions into the agent cost functions. Both design approaches focus on solving the constrained optimization problem through a sequence of subproblems which generates a trajectory of value profiles $v(1)$, $v(2)$, ..., that converges to the optimal solution. The key difference between the two approaches lies in the feasibility of the intermediate solutions as barrier functions can be used to ensure that the intermediate solutions are in fact feasible. More specifically, the constrained optimization problem (1) is transformed to the following problem:

$$
\begin{array}{ll}
\min_{v \in \mathcal{V}} & \phi(v) + \mu B(v) \\
s.t. & g(v) < 0
\end{array} \tag{40}
$$

where $B(v)$ is a barrier function associated with the constraints $g(\cdot)$ and $\mu > 0$ is a given constant or tradeoff parameter. For concreteness, we focus on the case when

$$
B(v) = -\sum_{k \in M} \log \left(C_k - \sum_{i \in N} A_i^k v_i \right). \tag{41}
$$

The barrier function is added to the objective function and ultimately prohibits the generated points from leaving the feasible region. Hence, this method generates a feasible point for each μ. When $\mu \to 0$, the limit of those feasible points is an optimal solution to the original problem. See Theorem 9.4.3 in [39] for more details.

A. Specifications of designed game

State Space, Actions, State Dynamics: These three parts are identical as in Section IV-A.
Admissible Action Sets: Let $x = (v, e)$ represent any state where the value profile v is in the interior of the feasible set, i.e., $g_k(v) < 0$ for all $k \in M$; and the estimation profile e satisfies $e_i^k < 0$ for all $i \in N$ and $k \in M$. Define the admissible action set for each agent $i \in N$ as

$$\mathcal{A}_i(x) \triangleq \left\{(\hat{v}_i, \hat{e}_i) : v_i + \hat{v}_i \in \mathcal{V}_i,\ e_i^k + A_i^k \hat{v}_i - \hat{e}_{i\to\text{out}}^k < 0,\ \hat{e}_{i\to j}^k < 0, \forall k \in M\right\}. \tag{42}$$

It is straightforward to check that this admissible action set ensures that the ensuing state profile $[\tilde{v}, \tilde{e}] = f(x, a)$ also satisfies that $g_k(v) < 0$ and $e_i^k < 0$ for all $i \in N$ and $k \in M$ by using Proposition 3. If the initial value $v(0)$ is in the interior of the feasible set and initial estimation $e(0)$ is negative then all subsequent value profiles $v(1)$, $v(2), \ldots$ will also be in the interior of the feasible set and the estimation profile $e(1), e(2), \ldots$ will also be negative. Furthermore, each agent i can evaluate the admissible action set $\mathcal{A}_i(x)$ using only local information.
Agent Cost Functions: Each agent $i \in N$ is assigned a cost function of the form (14) where the sole difference is the penalty function on the estimation term which now takes the form

$$J_i^g(x, a) = -\sum_{j\in N_i}\sum_{k=1}^{m} \log\left(-\tilde{e}_j^k\right). \tag{43}$$

where $\tilde{x} = (\tilde{v}, \tilde{e}) = f(x, a)$ represents the ensuing state. Here, the cost function $J_i^g(\cdot)$ utilizes a barrier function which penalizes values close to the constraint boundary.

B. Analytical properties of designed game

The barrier function state based game designed in this section provides the same guarantees as the penalty function state based game depicted in Section IV-A.

Theorem 9. *Model the constrained optimization problem in (1) as a state based game G with a fixed trade-off parameter $\mu > 0$ as depicted in Section V-A. The state based game is a state based potential game with potential function*

$$\Phi(x, a) = \phi(\tilde{v}) - \mu \sum_{i\in N}\sum_{k=1}^{m} \log\left(-\tilde{e}_i^k\right)$$

where $\tilde{x} = (\tilde{x}, \tilde{e}) = f(x, a)$ represents the ensuing state. Furthermore, the state action pair $[x, a] = [(v, e), (\hat{v}, \hat{e})]$ is a recurrent state equilibrium in the game G if and only if the following conditions are satisfied:

(i) The estimation profile e *satisfies that* $e_i^k = \frac{1}{n}\left(\sum_{i\in N} A_i^k v_i - C_k\right)$ *for all players* $i \in N$ *and constraints* $i \in N$.

(ii) The value profile v *is an optimal point of the constrained problem* $\min_{\{v\in\mathcal{V}:g(v)<0\}} \phi(v) + n\mu B(v)$ *where* $B(v)$ *is defined in equation (41).*

(iii) The change in value profile satisfies $\hat{v}_i = 0$ *for all players* $i \in N$.

(iv) The change in estimation profile satisfies the following for all players $i \in N$ *and constraints* $k \in M$, $\hat{e}^k_{i\leftarrow \text{in}} = \hat{e}^k_{i\rightarrow \text{out}}$.

As $\mu \to 0$ all state recurrent equilibria of our designed game are solutions to the constrained optimization problem in (1). We omit the proof for this Theorem 9 as it is virtually identical to the proof of Theorems 4-5.

C. Gradient play for barrier function methods

The state based game design using barrier functions depicted in Section IV-A is a state potential game with a potential function that is continuous, differentiable, and convex. Accordingly, we can implement the algorithm gradient play to reach a recurrent state equilibrium provided that $\mathcal{V}_i$ is a closed convex set for all $i \in N$ [27]. The algorithm gradient play can be described as follows: At each time $t \geq 0$ each agent i selects an action $a_i(t) \triangleq (\hat{v}_i(t), \hat{e}_i(t))$ given the state $x(t) = (v(t), e(t))$ according to:

$$\hat{v}_i = -\beta(t)\cdot\epsilon\cdot \left.\frac{\partial J_i(x(t),a)}{\partial \hat{v}_i}\right|_{a=\mathbf{0}} \tag{44}$$

$$\hat{e}^k_{i\rightarrow j} = -\beta(t)\cdot\epsilon\cdot \left.\frac{\partial J_i(x(t),a)}{\partial \hat{e}^k_{i\rightarrow j}}\right|_{a=\mathbf{0}} \tag{45}$$

where ϵ is the step size which are positive constants and $\beta(t) = \left(\frac{1}{2}\right)^{k(t)}$ where $k(t)$ is the smallest nonnegative integer k for which

$$\left(-\left(\frac{1}{2}\right)^k \epsilon\cdot \left.\frac{\partial J_i(x(t),a)}{\partial \hat{v}_i}\right|_{a=\mathbf{0}}, -\left(\frac{1}{2}\right)^k \epsilon\cdot \left.\frac{\partial J_i(x(t),a)}{\partial \hat{e}^k_{i\rightarrow j}}\right|_{a=\mathbf{0}}\right) \in \mathcal{A}_i(x(t))$$

VI. Illustrations

In this section we illustrate the results contained in this paper on two separate cooperative control problems where the desired global behavior must satisfy a given coupled constraint.

A. Consensus

Consider the consensus problem described in Example 1. Note that the system level objective function is pairwise decomposable and there is a single equality constraint. Therefore, this optimization problem fits into the class of problems described in Proposition 8. Accordingly, the exterior penalty method design in Section IV-A guarantees that we can achieve the optimal solution of the consensus problem (3) for any fixed constant $\mu > 0$.

We simulated a (weighted) average consensus problem with 61 agents and the results are presented in Figure 2. The figure on the left shows the initial conditions, i.e., the agents, the communication graph, the initial values, and the (weighted) average of the initial values. The figure on the right illustrates the convergence of the dynamics by applying the state based game design proposed in Section IV-A and the gradient play dynamics (39). Figure 3 demonstrates the evolution of $\Phi(x(t), a(t))$ (defined by equation (18)) during the gradient play learning process, which shows that $\Phi(x(t), a(t))$ converges to 0 very fast. Moreover, the evolution of the individual components of the potential function $\Phi(\cdot)$ which pertain to the value and estimation terms respectively, $\Phi^v(x(t), a(t)) = \phi(\tilde{v})$ and $\Phi^e(x(t), a(t)) = \mu \sum_{k=1}^{m} \left[\max(0, \tilde{e}_i^k)\right]^2$, are plotted in Figure 3. Since the agents are initially dispersed, the magnitude of $\Phi(x(t), a(t))$ is primarily caused by the value component $\Phi^v(x(t), a(t))$. Gradually the agents values move closer and estimation cost $\Phi^e(x(t), a(t))$ is the main contributor to $\Phi(\cdot)$. This plot demonstrates that agents are usually responding to inaccurate estimates of the value profile. However, this inaccuracies do not impact the limiting behavior of the dynamics.

B. Power control

Consider an optical channel problem where there is a set of channels $N = \{1, \ldots, n\}$ that are transmitted across an optical fiber by wavelength-multiplexing [11]. The system level objective is to minimize the total power while ensuring that the optical signal-to-noise ratio (OSNR) is larger than a specified threshold. For any channel $i \in N$ let v_i is the optical power at the input, p_i is the optical power at the output, and n_i^{in} and n_i^{out} is the noise at the input and output. The OSNR of channel $i \in N$ at the receiver is defined as

$$OSNR_i = \frac{p_i}{n_i^{\text{out}}} = \frac{\Gamma_{ii} v_i}{n_i^{\text{in}} + \sum_{j \in N} \Gamma_{i,j}\ v_j}$$

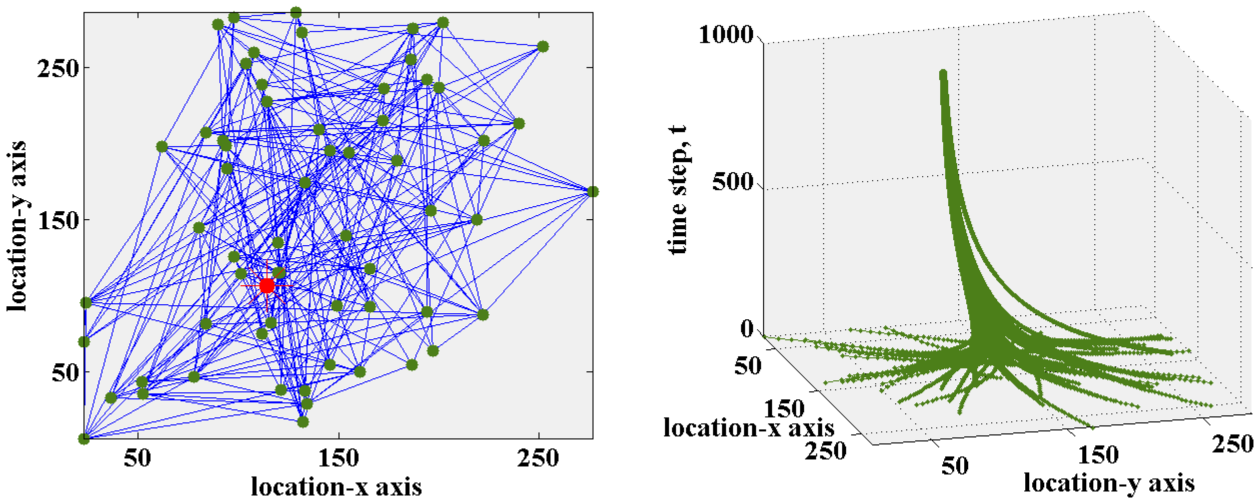

Fig. 2. The left figure shows the communication graph: Green nodes represent agents and their positions represent their initial values $v_i(0)$, which are two dimensional. Blue lines represent the communication links and the red node is the (weighted) average. The right figure demonstrates that the dynamics converge. XY plane is for the value information v and Z axis shows the time step. Each curve is a trajectory of one agent value v_i.

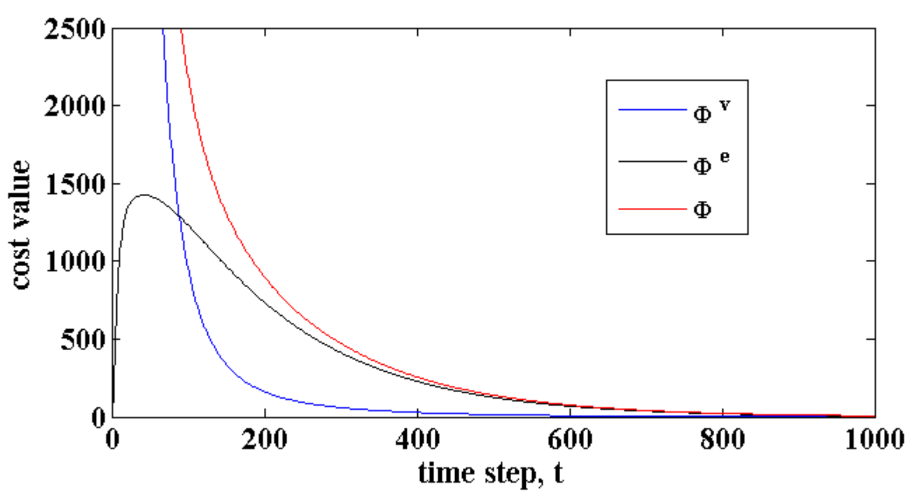

Fig. 3. Dynamics of $\Phi(x(t), a(t))$, $\Phi^{\phi}(x(t), a(t))$, and $\Phi^{e}(x(t), a(t))$

where $\Gamma = [\Gamma_{i,j}]$ is a system matrix. We consider a power control problem for this channel, which is defined as:

$$\begin{aligned} \min_{v_i \geq 0} \quad & \sum_{i \in N} v_i \\ s.t. \quad & \frac{\Gamma_{i,i} v_i}{n_i^{\text{in}} + \sum_{j \neq i} \Gamma_{i,j} v_j} \geq r_i, \forall i \in N \end{aligned} \tag{46}$$

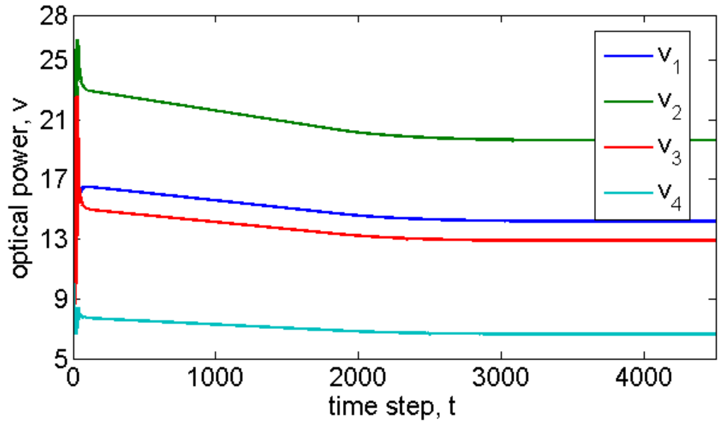

Fig. 4. Power control game design: v_i's dynamic during gradient play learning process by using exterior penalty function method design.

where r_i is the threshold for channel i's OSNR. It is easy to see that the constraints can be transformed to the following linear constraints:

$$r_i n_i^{\text{in}} + r_i \sum_{j \in N} \Gamma_{i,j} v_j - \Gamma_{i,i} v_i \leq 0, \forall i \in N \tag{47}$$

Moreover, the marginal cost for each channel is $M_i(v) = v_i$ which is a function based on its own action. Therefore, we can use the state based game designs depicted in Section IV-A and Section V-A to solve this channel problem. We consider the following setup with $N = 4$, $n_i^{\text{in}} = 0.5$, OSNR thresholds $\mathbf{r} = (4, 5, 3, 2)$ and system matrix

$$\Gamma = \begin{bmatrix} 4 & 0.5 & 0.25 & 0.1 \\ 0.5 & 4 & 0.5 & 0.25 \\ 0.25 & 0.5 & 4 & 0.5 \\ 0.1 & 0.25 & 0.5 & 4 \end{bmatrix}$$

The communication graph is shown as $1 \leftrightarrow 2 \leftrightarrow 3 \leftrightarrow 4$. By using linear programming we can solve for the optimal values which is $(v_1, v_2, v_3, v_4) = (14.1863, 19.6096, 12.8744, 6.6291)$.

Figure 4 shows the simulation of the power control problem using the state based game design incorporating exterior penalty functions and the learning algorithm gradient play. We set $\mu = 50$ and $\epsilon = 0.0015$. Figure 4 illustrates that the value profile v converges to $(v_1, \ldots, v_4) = (14.1974, 19.6243, 12.8838, 6.6322)$ which is close to the optimal results.

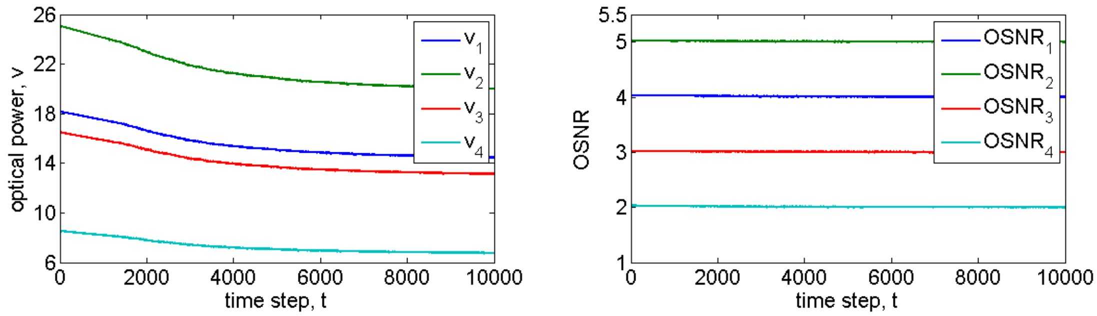

Fig. 5. Power control game design: The left figure shows v_i's dynamic during gradient play learning process by using barrier function method game design. The right figure shows $OSNR_i$'s dynamic during the learning process.

Figure 5 shows the simulation of the power control problem using the state based game design incorporating barrier functions and the learning algorithm gradient play. We set $\mu = 0.05$ and $\epsilon = 0.01$. Figure 5 illustrates that the value profile v converges to $(v_1, \ldots, v_4) = (14.30, 19.77, 12.98, 6.68)$ which is close to the optimal result. Furthermore, note that the constraints are always satisfied when using the algorithm gradient play.

VII. Conclusion

In this paper we investigate game design for multiagent systems where the desired global behavior involves coupled constraints. First, we demonstrate that the standard game theoretic framework is not suitable for dealing with such systems. Then we introduce a novel game design using the framework of state based games that ensures all resulting recurrent state equilibria satisfy the desired performance criterion. In essence, we introduce an underlying state variable for equilibrium refinement. The designed game fits into a class of games termed state based potential games which extends potential games to the realm of state based games. One of the nice properties of potential game (or state based potential games) is that virtually any learning algorithm is guaranteed to converge to a Nash equilibrium. Hence, there is an inherent robustness to variations in decision rules caused by delays in information, inconsistent clock rates, or inaccuracies in information. A significant future direction involves developing alternative methodologies for game design to meet the inherent challenges in multiagent systems, e.g., locality of objectives, efficiency of resulting equilibria, or more general coupled constraints.

REFERENCES

[1] J. N. Tsitsiklis, D. P. Bertsekas, and M. Athans. Distributed asynchronous deterministic and stochastic gradient optimization algorithms. *IEEE Transactions on Automatic Control*, **35**(9):803–812, 1986.

[2] E. Campos-Naòez, A. Garcia, and C. Li. A game-theoretic approach to efficient power management in sensor networks. *Operation Research*, 56(3):552–561, 2008.

[3] 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.

[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, October 2007.

[5] G. Arslan, M. F. Demirkol, and S. Yuksel. A game theoretic approach to power optimization in mimo interference systems.

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

[7] 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.

[8] D. X. Wei, C. Jin, S. H. Low, and S. Hegde. Fast tcp: motivation, architecture, algorithms, performance. *IEEE/ACM Transactions on Networking*, 14(6):1246–1259, December 2006.

[9] R. Olfati-saber and R. M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, 49(9):1520–1533, 2004.

[10] A. Rantzer. Dynamic dual decomposition for distributed control. In *Proceedings of American Control Conference*, St. Louis, June 2009.

[11] L. Pavel. An extension of duality to a game-theoretic framework. *Automatica*, 43:226–237, 2007.

[12] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. Technical report, working paper on line, Stanford, Univ, 2010.

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

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

[15] R. Gopalakrishnan, J. R. Marden, and A. Wierman. An architectural view of game theoretic control. In *Proceedings of ACM Hotmetrics*, 2010.

[16] J. R. Marden, G. Arslan, and J. S. Shamma. Cooperative control and potential games. *IEEE Transactions on Systems, Man and Cybernetics. Part B: Cybernetics*, 39, 2009.

[17] J. Marden and A. Wierman. Distributed welfare games.

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

[19] 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.

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

[21] J. R. Marden, G. Arslan, and J. S. Shamma. Regret based dynamics: Convergence in weakly acyclic games. In *Proceedings of the 2007 International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, Honolulu, Hawaii, May 2007.

[22] H.P. Young. *Strategic learning and its limits*. Oxford University Press, USA, 2004.

[23] D. Fudenberg and D.K. Levine. *The theory of learning in games*, volume 2. The MIT press, 1998.

[24] G. arslan, J. R. Marden, and J. S. Shamma. Autonomous vehicle-target assignment: a game theoretical formulation. *ASME Journal of Dynamic Systems, Measurement and Control*, 129:584–596, September 2007.

[25] D. Wolpert and K. Tumor. An overview of collective intelligence. In J. M. Bradshaw, editor, *Handbook of Agent Technology*. AAAI Press/MIT Press, 1999.

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

[27] N. Li and J. R. Marden. Designing games for distributed optimization. under submission, 2011.

[28] Y. Zou and K. Chakrabarty. Uncertainty-award and coverage-oriented deployment for sensor networks. *Journal of Parallel and Distributing Computing*, 64(7):788–798, 2004.

[29] A. Garcia, D. Reaume, and R. Smith. Fictitious play for finding system optimal routings in dynamic traffic networks. *Transportation Research B, Methods*, 34(2):147–156, January 2004.

[30] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Gossip algorithms: Design, analysis and applications. In *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, volume 3, pages 1653–1664. IEEE, 2005.

[31] M. Zhu and S. Martinez. An approximate dual subgradient algorithm for distributed cooperative non-convex constrained optimization. *IEEE Transactions on Automatic Control*, submitted, 2010.

[32] I. Lobel, A. Ozdaglar, and D. Feijer. Distributed multi-agent optimization with state-dependent communication. 2010. under submission.

[33] 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.

[34] S. Uryasev and R. Y. Rubinstein. On relaxation algorithms in computation of non-cooperative equilibria. *IEEE Transactions on Automatic Control*, 39:1263–1267, June 1994.

[35] J. Krawczyk. Numerical solutions to coupled-constraint (or generalised nash) equilibrium problems. *Computational Management Science*, 4(2):183–204, April 2007.

[36] H. Nikaido and K. Isoda. Note on non-cooperative convex games. *Pacific J. Math*, 5:807–815, 1955. Suppl. 1.

[37] J. B. Rosen. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica*, 33(3):520–534, 1965.

[38] 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.

[39] M. Bazaraa, H. Sherali, and C. Shetty. *Nonlinear Programming: Theory and Algorithms, 3rd edition*. WILEY-Interscience, 2006.